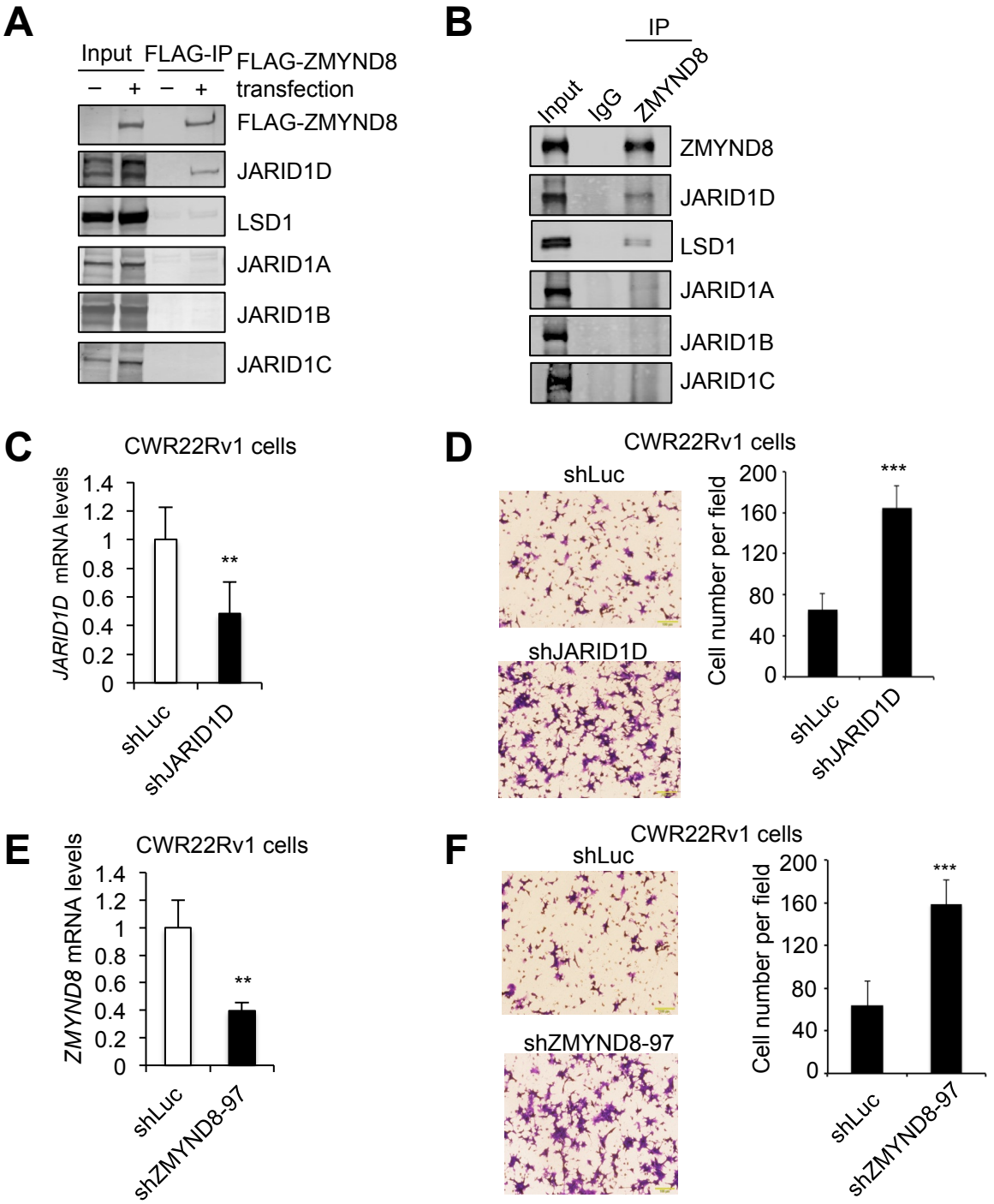
Figure S1



<u>Figure S1</u>. ZMYND8 weakly interacts with LSD1, and knockdown of JARID1D or ZMYND8 increases the invasiveness of the androgen-independent cell line CWR22Rv1 (Related to Figures 1 and 3).

(**<u>A</u>**) FLAG-immunoprecipitation (IP) showed that FLAG-ZMYND8 interacted with JARID1D and to a lesser extent LSD1 but not with JARID1A-1C. DU145 cells were transfected with FLAG-ZMYND8 expression plasmid. The whole cell extracts were used for FLAG-IP.

(**<u>B</u>**) Endogenous IP using nuclear extracts of DU145 cells showed that ZMYND8 interacted with JARID1D and to a lesser extent LSD1 but not with JARID1A-1C.

(**<u>C & E</u>**) *JARID1D* (**C**) and *ZMYND8* (**E**) mRNA levels were quantitatively measured using CWR22Rv1 cells that were treated with shLuciferase (shLuc) and shJARID1D, or shZMYND8-97 lentiviruses.

(**<u>D & F</u>**) Cell invasion assays showed that JARID1D (**D**) and ZMYND8 (**F**) knockdown increased the invasive abilities of CWR22Rv1 cells.

Figure S2

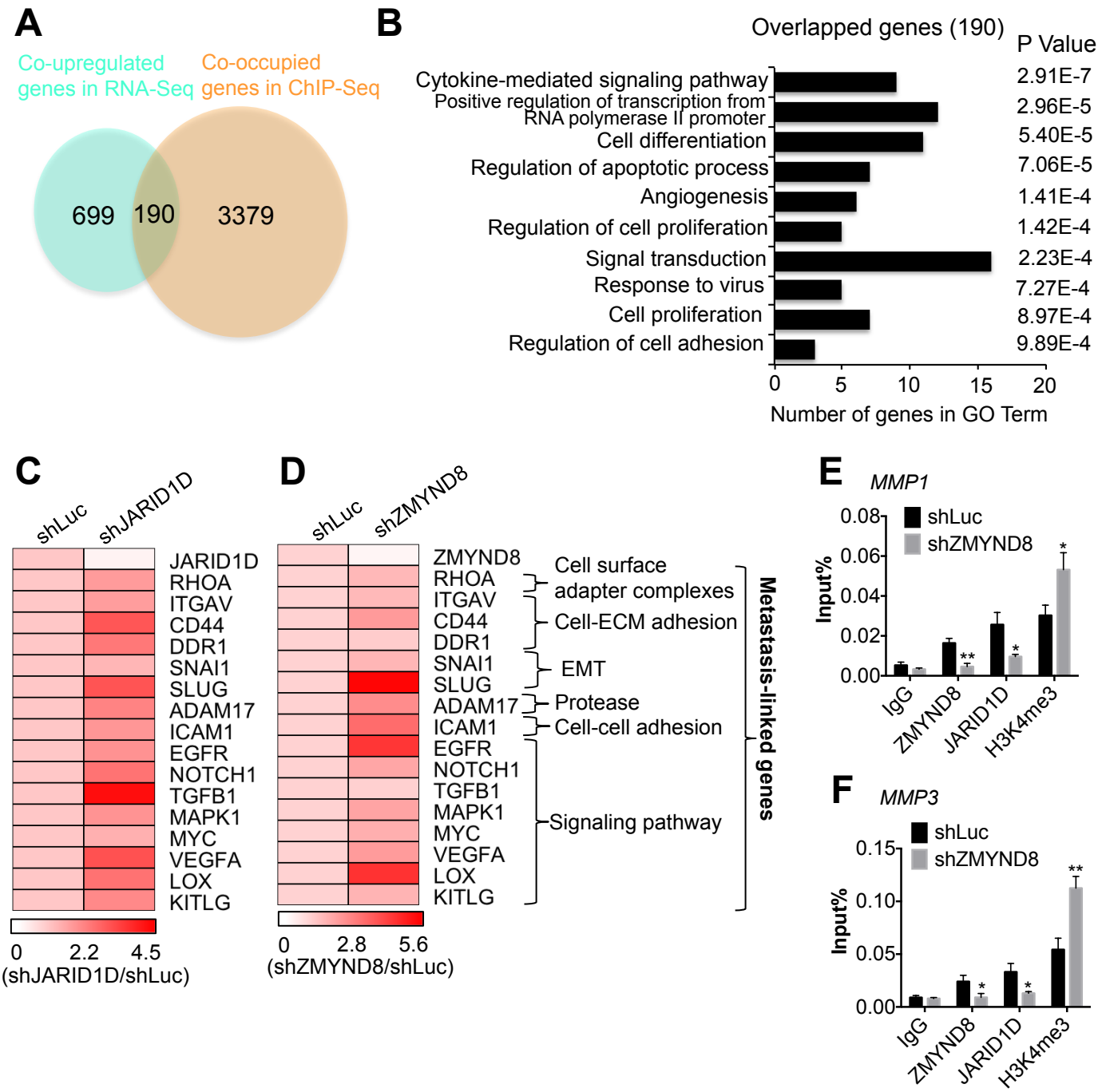


Figure S2. Knockdown of ZMYND8 or JARID1D increases expression of multiple metastasis-linked genes (Related to Figure 4).

(**A and B**) Venn diagrams (**A**) and gene ontology analysis (**B**) are shown for genes that were co-upregulated by JARID1D and ZMYND8 knockdown and were co-occupied by JARID1D and ZMYND8. RNA-Seq was performed for the whole transcriptome analysis.

(**C and D**) Quantitative RT-PCR results showed that multiple metastasis-linked genes were up-regulated by JARID1D (**C**) or ZMYND8 (**D**) knockdown. These genes were co-occupied by JARID1D and ZMYND8 according to ChIP-Seq data (See also **Figure S5**). Metastasis-linked genes are grouped by their functions. ECM, Extracellular matrix; EMT, epithelial-mesenchymal transition.

(**E and F**) ZMYND8 knockdown decreased JARID1D and ZMYND8 occupancy and increased H3K4me3 levels at the promoter regions of *MMP1* (**E**) and *MMP3* (**F**) genes.

Figure S3

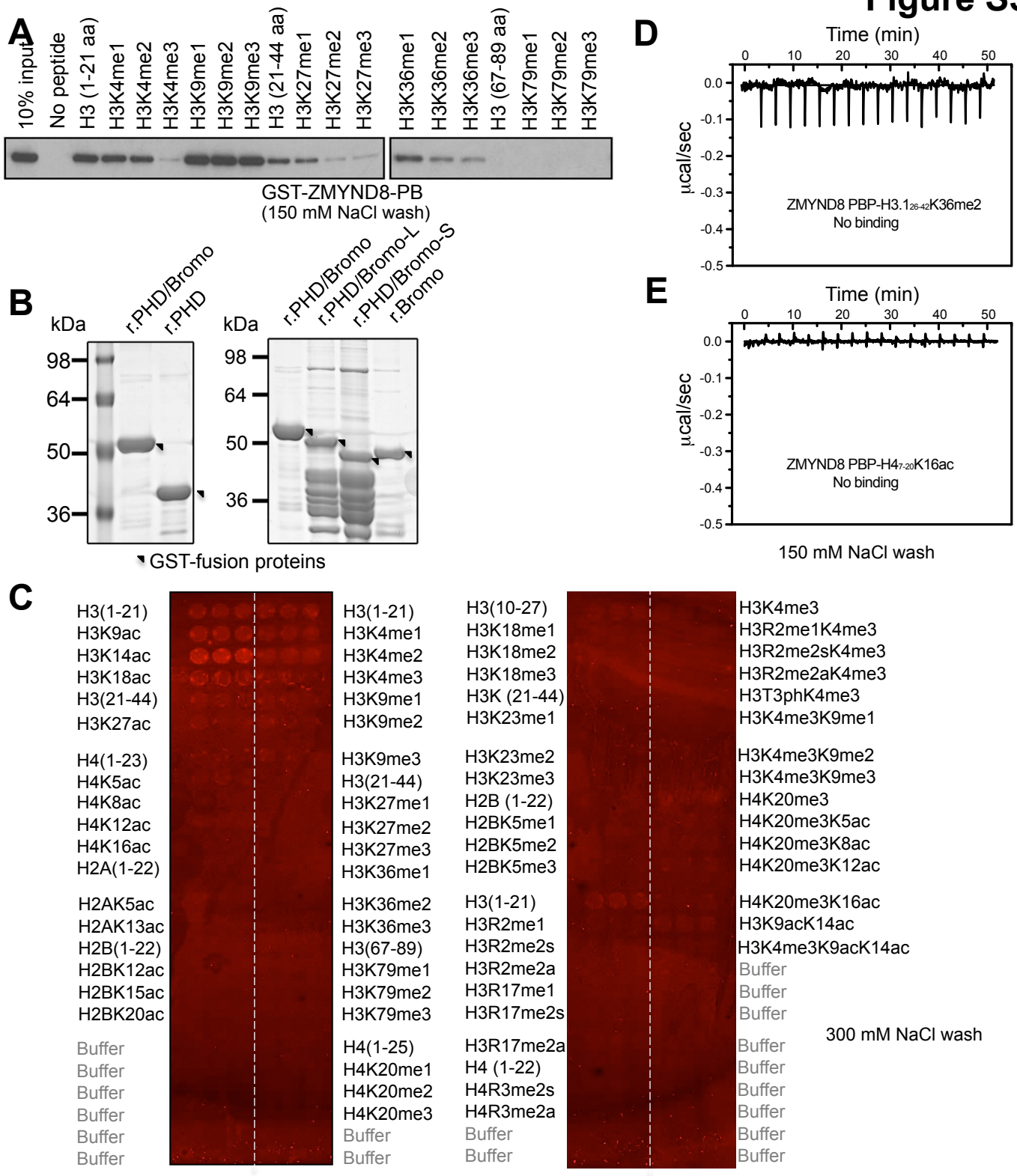


Figure S3. ZMYND8 PBP binds to H3K4me0, H3K4me1, and H3K14ac peptides (Related to Figure 5).

(**A**) Peptide pull-down assay in a low salt wash (150 mM NaCl) using recombinant GST-ZMYND8-PB.

(**B**) Analysis of recombinant GST-tagged fusion proteins (PHD/Bromo or its deletion mutants) by Coomassie blue staining.

(**C**) Peptide array containing peptides with indicated modifications were incubated using recombinant GST-ZMYND8-PBP. Three replicates were used for each binding assay.

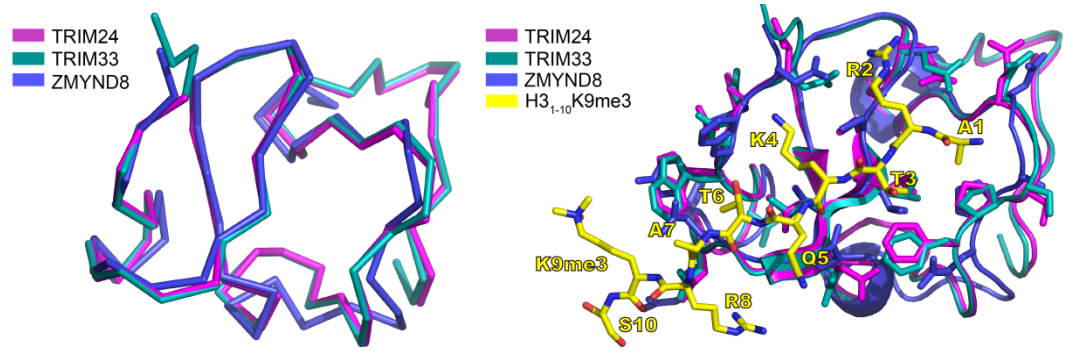
(**D and E**) Isothermal titration calorimetry (ITC) assays showed that ZMYND8-PBP did not have any significant binding activity to H3K36me2 (**D**) and H4K16ac (**E**) .

Figure S4

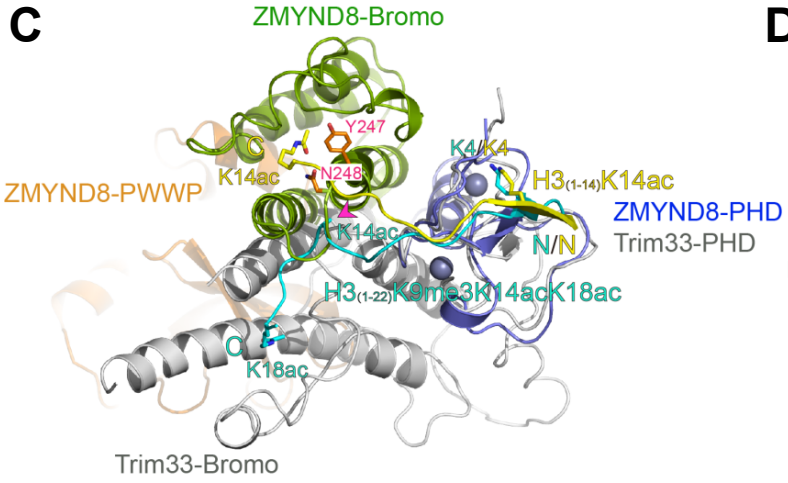
A PHD domain

```
ZMYND8-HUMAN      93  LTTDPVDVVPQDGRNDFYCWVCHREGQVLCCELCPRVYHAKCLR--LTSEPEGDWFCPECEKITV 155
ZMYND8-Mus        97  LTTDPVDVVPQDGRNDFYCWVCHREGQVLCCELCPRVYHAKCLR--LTSEPEGDWFCPECEKITV 159
ZMYND8-Rattus     73  LTTDPVDVVPQDGRNDFYCWVCHREGQVLCCELCPRVYHAKCLR--LTSEPEGDWFCPECEKITV 135
ZMYND8-Xenopus   113  LTTDPVDVVPQDGRNDFYCWVCHREGQVLCCELCPRVYHAKCLK--LTAEPEGDWFCPECEKITV 175
TRIM24-PHD       812  PLHVGETRKEDDP-NEDWCAVCQNGGELLCCEKCPKVFHLSCHVPTLTNFPSGEWICTFCRDLSK 875
TRIM33-PHD       873  RIGGDGNNKDDDP-NEDWCAVCQNGGDLLCCEKCPKVFHLTCHVPTLLSFPSGDWICTFCRDIGK 936
BHC80-PHD        474  VSLPSPTSTDGDI-HEDFCSVCRKSGQLLMCDTCSRVYHLDCLDPPLKTIPKGMWICPRCQDQML 537
```

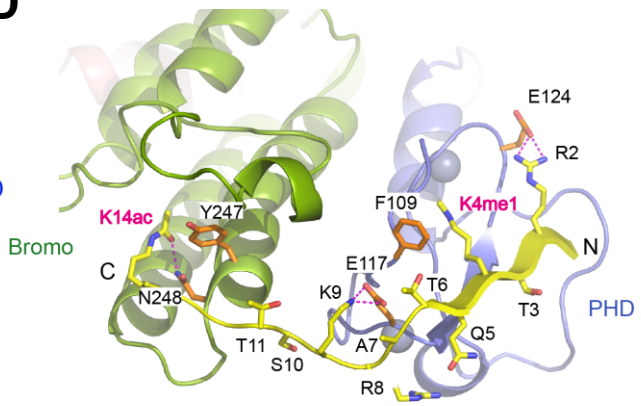
B



C



D



E

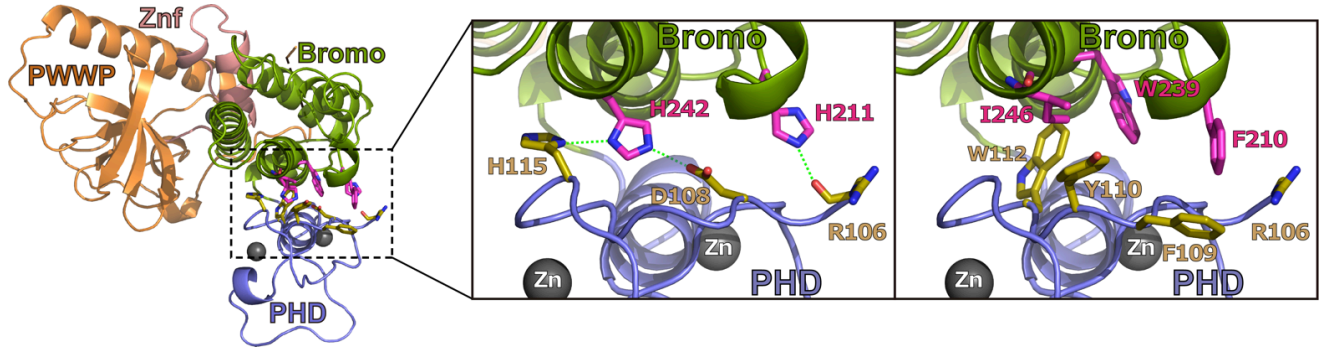


Figure S4. Structural modelling for the interaction between ZMYND8-PBP and H3K4me0K14ac (1-14 aa) or between ZMYND8-PBP and H3K4me1K14ac (1-14 aa) (Related to Figure 6).

(**A**) **Sequence alignment of PHD domains in ZMYND8, ZMYND8's three homologues, TRIM24, TRIM33, and BHC80.** The amino acids in red color indicate the C4HC3 motifs of the PHD domains.

(**B**) **Structure-based superimposition of TRIM24 (PDB code: 3O37) and TRIM33 (PDB code: 3U5N) with ZMYND8 PHD fingers**. Left, alignment at Cα trajectory level. Right, superimposition of all three PHD fingers with key H3K4me0-binding residues depicted as sticks. An H3K9me3 (1-10) peptide from TRIM33 complex (PDB code: 3U5N) was shown as yellow sticks.

(**C**) **Structural superimposition of free ZMYND8-PHD/Bromo with TRIM33-PHD/Bromo bound to $H3_{(1-22)}$K9me3K14acK18ac peptide** (PDB code: 3U5O). The sidechains of two residues Y247 and N248 are shown in orange sticks with blue for nitrogen and red for oxygen atom. TRIM33-PHD/Bromo is shown in grey cartoon. Histone $H3_{(1-22)}$K9me3K14acK18ac peptide is depicted as cyan cartoon with sidechains of K4, K14ac and K18ac shown in sticks (blue for nitrogen and red for oxygen atoms). Modelled $H3_{(1-14)}$K4me0K14ac peptide is shown in yellow cartoon with sidechains of K4 and K14ac shown in sticks (blue for nitrogen and red for oxygen atoms). Zinc atoms are shown in blue spheres.

(D) Structural modelling of ZMYND8-PBP bound to H3K4me1K14ac (1-14 aa) peptides. Blue spheres, Zinc atoms; Magenta dashes, hydrogen bonds. The modelled H3 peptide is depicted as sticks with yellow for carbon, blue for nitrogen and red for oxygen atoms.

(**E**) **Interaction analysis in the interface between ZMYND8-PHD and ZMYND8-Bromo**. Left close-up view, hydrogen bond network at the PHD/Bromo interface; right close-up view, hydrophobic clusters at the PHD/Bromo interface

Figure S5

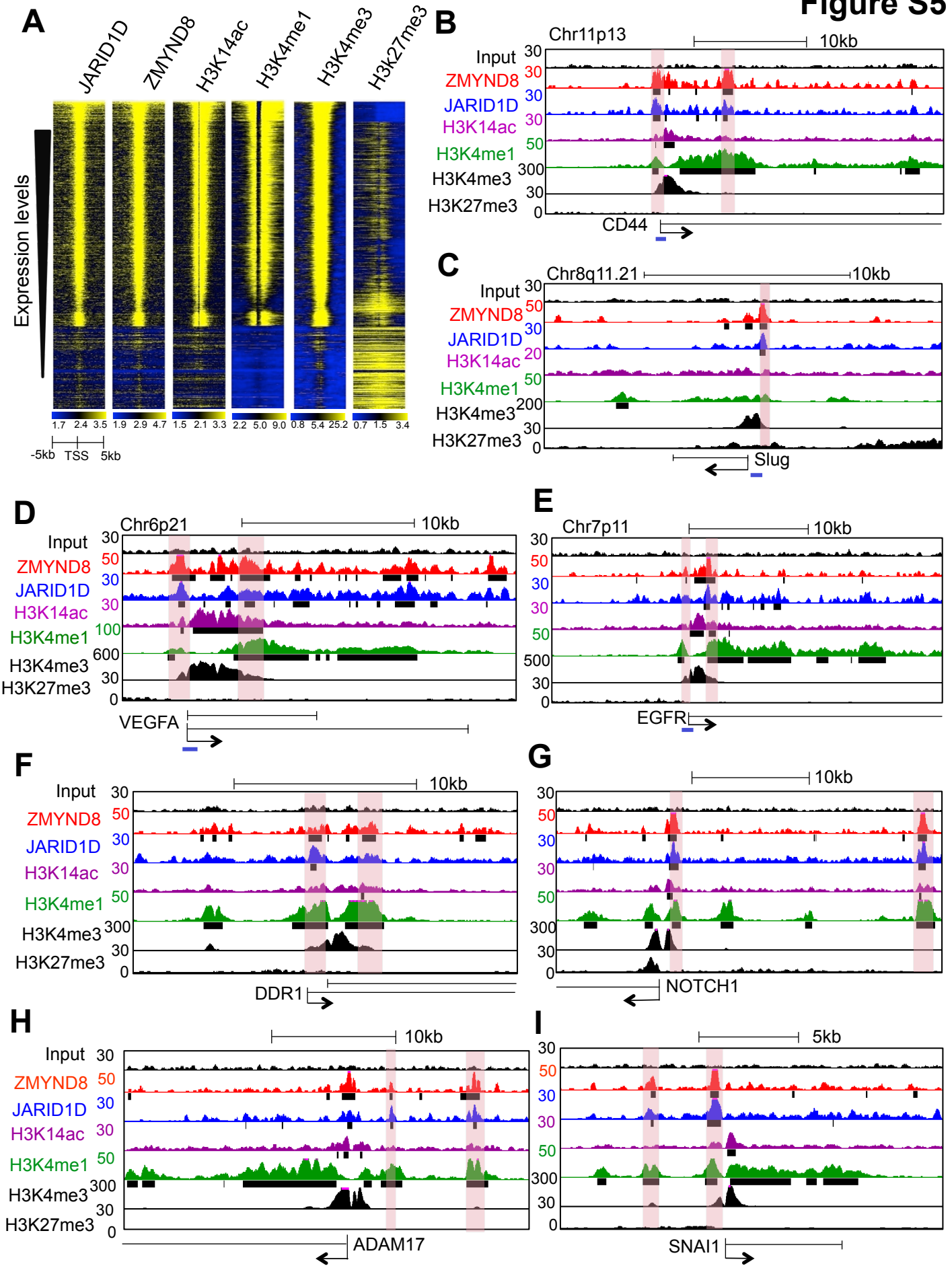
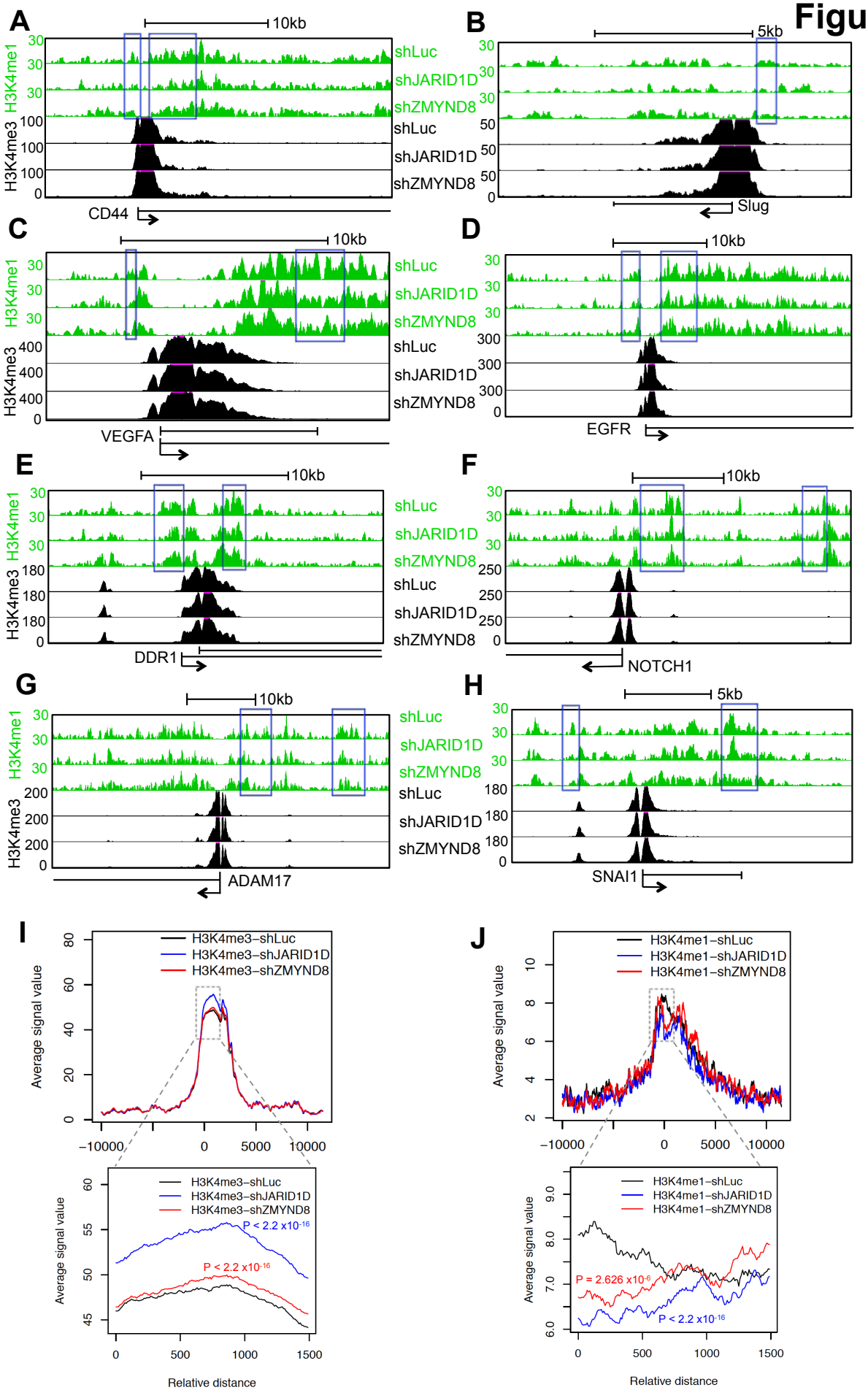

A
JARID1D
ZMYND8
H3K14ac
H3K4me1
H3K4me3
H3K27me3
Expression levels
1.7 2.4 3.5
1.9 2.9 4.7
1.5 2.1 3.3
2.2 5.0 9.0
0.8 5.4 25.2
0.7 1.5 3.4
-5kb TSS 5kb
B
Chr11p13
10kb
Input
ZMYND8
JARID1D
H3K14ac
H3K4me1
H3K4me3
H3K27me3
CD44
C
Chr8q11.21
10kb
Slug
D
Chr6p21
10kb
VEGFA
E
Chr7p11
10kb
EGFR
F
10kb
DDR1
G
10kb
NOTCH1
H
10kb
ADAM17
I
5kb
SNAI1

Figure S5. ZMYND8 and JARID1D peaks overlap H3K4me1 and H3K14ac peaks at metastasis-linked genes (Related to Figure 7).

(**A**) Heat maps of ChIP-Seq enrichment of JARID1D, ZMYND8, H3K14ac, H3K4me1, H3K4me3 and H3K27me3 in the regions near TSS in DU145 cells. Each column represents a 10-kb window (- 5 kb to +5 kb). All the genes were sorted by their mRNA expression levels in DU145 cells.

(**B-I**) Genome browser views of ZMYND8, JARID1D, H3K14ac, H3K4me1, H3K4me3 and H3K27me3 ChIP-Seq peaks at at *CD44* (**B**), *Slug* (**C**), *VEGFA* (**D**), *EGFR* (**E),** *DDR1* (**F**), *NOTCH1* (**G**), *ADAM17* (**H**), and *SNAI1* (**I**) genes in DU145 cells. ChIP-Seq signals of JARID1D, ZMYND8, H3K4me1, H3K14ac, H3K4me3, and H3K27me3 in the tracks represent their remainder signals after input signals were subtracted from their original signals. Peak tracks in the BigBed format are shown as black lines along the bottom of ZMYND8, JARID1D, H3K14ac, and H3K4me1 ChIP-Seq signals. The peak tracks were determined using a statistical cutoff with Poisson test P value (1.00×10^{-15} for peak calling and 1.00×10^{-10} for peak extending). ChIP-Seq signals in Figures **7A–7D** are re-shown in **B–E**.

Figure S6


A
10kb
H3K4me1
H3K4me3
shLuc
shJARID1D
shZMYND8
CD44
B
5kb
Slug
C
VEGFA
D
EGFR
E
DDR1
F
NOTCH1
G
ADAM17
H
SNAI1
I
Average signal value
H3K4me3–shLuc
H3K4me3–shJARID1D
H3K4me3–shZMYND8
P < 2.2 x10-16
Relative distance
J
H3K4me1–shLuc
H3K4me1–shJARID1D
H3K4me1–shZMYND8
P = 2.626 x10-6
P < 2.2 x10-16

<u>Figure S6</u>. Knockdown of JARID1D or ZMYND8 moderately increases H3K4me3 levels and decreases H3K4me1 levels (Related to Figure 7).

(**<u>A-H</u>**) **Knockdown of JARID1D or ZMYND8 moderately augmented H3K4me3 levels and reduced H3K4me1 levels** at *CD44* (**A**), *Slug* (**B**), *VEGFA* (**C**), *EGFR* (**D),** *DDR1* (**E**), *NOTCH1* (**F**), *ADAM17* (**G**), and *SNAI1* (**H)** genes**.** Genome browser views of H3K4me1, H3K4me3 ChIP-Seq are shown for DU145 cells treated with shLuc, shJARID1D or shZMYND8 lentiviruses. Average signal of the spike-in normalized ChIP-Seq signals from two biological replicates are presented for each track.

(<u>I</u>) JARID1D knockdown upregulated average H3K4me3 level, and ZMYND8 knockdown slightly increased average H3K4me3 level. Average H3K4me3 ChIP-Seq signals of 190 genes in shLuc, shJARID1D or shZMYND8 DU145 cells are shown. These genes were co-upregulated by JARID1D and ZMYND8 knockdown and were co-occupied by JARID1D and ZMYND8. In the enlarged figure, Wilcoxon Signed-rank test was performed for statistic analysis between shLuc and shJARID1D (P value < 2.2×10^{-16}) and between shLuc and shZMYND8 (P value < 2.2×10^{-16}).

(<u>J</u>) Knockdown of JARID1D or ZMYND8 moderately decreased average H3K4me1 levels. Average H3K4me1 ChIP-Seq signals of 190 genes in shLuc, shJARID1D or shZMYND8 DU145 cells are shown. These genes were co-upregulated by JARID1D and ZMYND8 knockdown and were co-occupied by JARID1D and ZMYND8. In the enlarged figure, Wilcoxon Signed-rank test was used for statistic analysis between shLuc and shJARID1D (P value < 2.2×10^{-16}) and between shLuc and shZMYND8 (P value = 2.626×10^{-06}). For **I** and **J**, the center of the plot represents the regions overlapped by JARID1D and ZMYND8 ChIP-Seq peaks that are located between - 5kb and the transcription termination sites.

Supplemental Experimental Procedures

Purification of the JARID1D-containing complex

Nuclear extracts were obtained from a H1299 cell line that stably expressed FLAG-tagged JARID1D. Anti-FLAG M2 affinity resin (Sigma) was used to purify the complex. JARID1D-associated proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and identified by liquid chromatography-tandem mass spectroscopy.

Immunoprecipitation assay

To test the endogenous interaction between JARID1D and ZMYND8, nuclear extracts were isolated from DU145 cells. Nuclear extracts were then pre-cleared by rProtein A beads for 3h at 4°C and immunoprecipitated with anti-JARID1D or anti-ZMYND8 antibody for 16-18h at 4°C. rProtein A beads were added and incubated for 1–2 h. The beads were washed extensively and eluted using 1 x SDS-PAGE sample buffer. The eluates were analyzed by SDS-PAGE, followed by Western blotting using the indicated antibodies. FLAG-IP was also performed to determine the interaction between JARID1D and exogenously expressed FLAG-ZMYND8 in DU145 cells.

To assess a ZMYND8 domain that is accountable for the interaction with JARID1D, FLAG-JARID1D and either HA-ZMYND8 or its deletion mutants were transiently expressed in 293T cells. FLAG-IP was performed using whole cell extracts. To determine a JARID1D domain that is responsible for the interaction with ZMYND8, recombinant FLAG-JARID1D, its deletion mutants, and full-length ZMYND8 proteins were purified from baculovirus-infected insect cells as previously described (Lee et al., 2007). Full-length ZMYND8 and either recombinant FLAG-JARID1D or its deletion mutants were mixed, and IP was performed using anti-ZMYND8

antibody.

ChIP-Seq analysis

For ChIP-Seq analysis, we used Bowtie to map ChIP-Seq reads to the human reference genome version hg19, while achieving single best match for each read across the genome. We used the function Dregion in DANPOS2 to calculate reads density from the mapped reads and defined enriched peaks using a high statistical cutoff (Poisson test P value 1.00 $x10^{-50}$ for peak calling and 1.00 $x10^{-20}$ for peak extending) that enabled us to prioritize a gene list for subsequent genome-wide analysis (such as gene ontology analysis and Venn diagram overlaps). For each ChIP-Seq sample without spike-in control, we normalized the total reads number to 25 millions. For each ChIP-Seq sample with spiked-in *Drosophila melanogaster* chromatin as a minor fraction of the total chromatin, we normalized the total reads number to be the reads number of the sample containing the least number of *Drosophila melanogaster* reads within the same antibody group. The function Dregion in DANPOS2 also was used to calculate variations between biological replicates of the same ChIP experiment. We set the extending length to 200bp and bin size to 10bp in the calculation of reads density. We set the smooth width to 0bp to not use any smoothing step in the calculation. Input effect was subtracted from the ChIP-Seq data by DANPOS2. For reference gene set, we downloaded the KnownGene in the GENCODE version 22 provided at the Table Browser page of UCSC Genome Browser (http://genome.ucsc.edu/cgi-bin/hgTables). We used the function Profile in DANPOS2 to plot average ChIP-Seq reads density around all transcription start sites and to generate data matrix for heatmap of ChIP-Seq reads density around each transcription start site and used the tool MEV to plot the heatmap.

We used the function Selector in DANPOS2 to retrieve peaks that are on exon, on intron, within 3kb distance to transcription start site, or in intergenic region. The Selector function also retrieves the genes associated with those peaks. We used the function Overlap in

DANPOS2 to calculate number of overlap between any two sets of peaks. We used the function RetrieveDNA in DANPOS2 to retrieve the genomic DNA underlying top 1000 peaks of JARID1D or ZMYND8 ranked by their maximal binding density value within each peak. We then divided the sequences into three groups that were occupied by both JARID1D and ZMYND8, only JARID1D, or only by ZMYND8.

We used the sub-command *intersect* of the bedtools suite to identify co-binding regions of JARID1D and ZMYND8. Using the same sub-command, we also retrieved the subset of co-binding regions overlapping with either the 5kb upstream regions of TSS or transcript bodies of the 889 genes co-upregulated by JARID1D and ZMYND8 knockdown. As described above, we used the function Profile in DANPOS2 and customized R script to plot average reads density around all identified overlapping regions for H3K4me1 and H3K4me3 ChIP-Seq samples. Specifically, we normalized the length of each overlapping region to 1.5kb, with flanking region of both side being 10kb. For H3K4me1 and H3K4me3 ChIP-Seq, we performed statistical test for difference of signal strength between control and knockdown samples using wilcoxon Signed-rank test.

ChIP PCR

ChIP-enriched DNAs were subjected to quantitative PCR. PCR primers were as follows: CD44 (Forward: ATCTTGCTCCAGCCGGATTC; Reverse: GGACAGAGGATGACCGAACC), Slug (Forward: AGTGTGAGAGAATGTCCGGTGGT; Reverse: GCGGCCGCGTGCAAATTAAG), EGFR (Forward: ATGGGGCTCACAGCAAACTT; Reverse: GCCAGACTCGCTCATGTTCT), VEGFA (Forward: TAGCAAAGAGGGAACGGCTC; Reverse: AACTCTGTCCAGAGACACGC).

Knockdown and ectopic expression experiments

For the knockdown of ZMYND8 and JARID1D, lentiviruses were generated by co-

transfecting HEK293T cells with shRNA-expressing plasmids with puromycin resistance, a packing plasmid (deltaR8), and an envelope plasmid (VSV-G). Fourteen hours later, the medium was replaced with Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum. Media containing virus particles with shJARID1D, shZMYND8, shLuciferase or shScramble were collected and used to infect prostate cancer cells for three times. Following a 72-hour infection, knockdown efficiency was determined by qRT-PCR and Western blot analysis.

To ectopically express JARID1D, JARID1D mutants, ZMYND8, and ZMYND8 mutants, we transfected their cDNA constructs into cells using Lipofectamine 3000 (life technologies) or Continuum transfection reagent (GEMINI) according to the manufacturer's instructions. Following incubation for 24-48 hours, cells were harvested, and total RNA was isolated for further analysis.

Cell proliferation assay

Cells were infected with viruses containing shLuc or shZMYND8 and then seeded in triplicate at a density of 15,000 cells per well in 24-well plate, and cell numbers were determined at 0, 1, 2 and 4 days.

Migration and invasion assays

Migration assays were performed using inserts with membrane filters (8 μm of pore size) in 24-well plates. DU145 cells (1.0×10^5) in 500μl of serum-free DMEM media were placed on the membrane in each insert, and 500μl of DMEM media containing 10% fetal bovine serum was added in each well. Twenty-four hours later, cells that had migrated to the other side of the membrane were fixed with 4% paraformaldehyde and stained with crystal violet. For the cell invasion assay, DU145 (1.0×10^5) or CWR22Rv1 cells (2.0×10^5) cells were suspended in 500

μl of 0.1% bovine serum albumin in DMEM media and seeded on the Matrigel-coated membrane in each insert. Twenty four hours later, cells that had invaded the Matrigel and migrated to the other side of the membrane were stained and counted. Data are presented as the mean ± SEM (error bars).

RNA isolation, cDNA synthesis, qRT-PCR, and RNA-seq

Total RNAs were extracted using RNeasy kits (Qiagen). cDNAs were synthesized using the iScript cDNA synthesis kit (Bio-Rad) according to the manufacturer's instructions. Gene expression levels were normalized to GAPDH. The relative fold represents the fold change compared to the control. Data are presented as the mean ± SEM (error bars).

The total RNA of shLuciferase, shZMYND8 and shJARID1D DU145 cells from two biological replicates were used for RNA-seq experiments. The sequencing was performed by HiSeq 3000 (MD Anderson SMF). For RNA-seq data, we used TopHat to map reads to the hg19 human reference genome and the cuffdiff tool of cufflink suite to calculate read counts for each gene. To identify differential expressed genes based on read counts of different RNA-Seq samples, we used the edgeR package.

Peptide pull-down assays

GST-tagged bacterial recombinant proteins were purified as previously described (Dhar et al., 2012). For peptide pull-down assays, 1μg of biotinylated peptides were incubated with 2 μg of recombinant proteins in 300 μL of binding buffer (50 mM Tris-HCl [pH 7.5], 150mM or 300 mM NaCl, 0.05% [v/v] Nonidet P-40, 1 mM PMSF, 1 μg/mL aprotinin, 2.5 μg/mL leupeptin, and 1 μg/mL pepstatin) for 4 h or overnight at 4°C and then mixed with streptavidin beads (Amersham Biosciences) for 1 h. The beads were washed four times and eluted with 1x SDS-PAGE sample buffer (62.5 mM Tris, 5% β-mercaptoethanol, 2% SDS, 10% glycerol, and 0.01% Bromophenol blue [pH 6.8]).

Isothermal titration calorimetry

For ITC measurements, synthetic histone peptides (>97% purity) and recombinant proteins (ZMYND8-PB, ZMYND8-PBP, and their mutants) were all prepared under the same titration buffer (150 mM NaCl, 25 mM HEPES [pH 7.5], and 5% glycerol). Protein concentration was measured on the basis of its UV_{280nm} absorption. Aliquots of the sample were freeze-dried until further use. ITC titrations were performed using a MicroCal iTC200 system (GE Healthcare) at 15°C. Each titration consisted of 17 successive injections (the first at 0.4 μl and the remaining 16 at 2.41 μl). Usually, histone H3 peptides at a concentration of 1.0 – 1.5 mM were titrated into ZMYND8-PBP at a concentration of 0.06 − 0.1 mM in the sample cells. The resultant titration curves were processed using the Origin 7.0 software program (OriginLab) according to the “one set of sites” fitting model.

Protein purification, crystallization, data collection, and structure determination

Recombinant ZMYND8-PBP (93-426aa) was cloned into a modified pET28b vector and expressed with the N-terminal 10xHis-SUMO tag in the *E. coli* strain BL21(DE3) (Novagen) in the presence of 0.1 mM IPTG plus 50 μM $ZnCl_2$ at 12°C. Overnight induced cells were collected and suspended in lysis buffer: 150 mM NaCl, 25 mM HEPES [pH 7.5], 10 mM imidazole, 5% glycerol, and 1 mM phenylmethylsulfonyl fluoride. The cells were lysed using an Emulsiflex C3 (Avestin) high-pressure homogenizer. After centrifugation, the supernatant was applied to a HisTrap (GE Healthcare) nickel column and the protein was eluted with a linear imidazole gradient from 20 mM to 500 mM. The resultant protein was treated with 6xHis-tagged Ulp1 (homemade) digestion overnight and then exchanged to lysis buffer before being loaded onto the HisTrap (GE Healthcare) nickel column for 10xHis-SUMO tag and Ulp1 removal. The tag-free ZMYND8-PBP was further purified by a HiLoad 16/60 Superdex 200 (GE Healthcare) gel filtration column. The monomer peak in elution buffer (150 mM NaCl, 25 mM HEPES [pH 7.5], and 5% glycerol) was pooled, concentrated to about 12 mg/ml, and stored in a -80°C freezer. All

chromatographic steps were performed using AKTA Purifier 10 systems (GE Healthcare).

Before crystallization, ZMYND8-PBP was mixed with histone $H3_{(1-15)}$K4me0 peptide (>97% purity) at a molar ratio of 1:2. After incubation on ice for about 2 hours, crystallization was performed via the sitting drop vapor diffusion method at 18°C by mixing equal volumes (0.2-1.0 μl) of protein and reservoir solution. After 1 day, ZMYND8-PBP/$H3_{(1-15)}$K4me0 crystals appeared from a reservoir solution containing 0.1mM HEPES [pH 7.5] and 1.4 M sodium citrate tribasic dihydrate. The complex crystals were briefly soaked in a cryoprotectant drop containing the reservoir solution supplemented with 20% glycerol and then flash–frozen in liquid nitrogen for data collection.

Diffraction data sets for native and zinc absorption were collected at the beamline BL17U of the Shanghai Synchrotron Radiation Facility at wavelengths of 0.9790 Å and 1.2821 Å, respectively. All diffraction images were indexed, integrated, and merged using the HKL2000 program package (Otwinowski and Minor, 1997). The phase of ZMYND8-PBP bound to $H3_{(1-15)}$K4me0 was determined by the PHENIX program (Adams et al., 2010) on the basis of the zinc single-wavelength anomalous dispersion method. Model building and structural refinement was carried out using the programs COOT and PHENIX (Adams et al., 2010). Details of the data collection and refinement statistics are summarized in **Table 1**. Structural figures were created using PYMOL (http://www.pymol.org/). The model of ZMYND8-PBP in complex with histone $H3_{(1-6)}$K4me0/1 and $H3_{(1-14)}$K4me0/1-K14ac peptides was obtained by superimposing the free ZMYND8-PBP structure and Trim33-H3 complex structure (Protein Data Bank ID: 3U5N and 3U5O) and manually adjusting them in COOT to remove apparent steric clashes. The resultant model was then sent to PHENIX (http://www.phenix-online.org) for geometry minimization. Three macro cycles of refinement were performed with the energy term of the targets dropping from 7864 to 5599.

Supplemental References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr *66*, 213-221.

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 & development *26*, 2749-2762.

Lee, M.G., Norman, J., Shilatifard, A., and Shiekhattar, R. (2007). Physical and functional association of a trimethyl H3K4 demethylase and Ring6a/MBLR, a polycomb-like protein. Cell *128*, 877-887.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. Methods in Enzymology *276*, 19.