Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/6/16/eaaz0356/DC1

Supplementary Materials for

PHF20L1 as a H3K27me2 reader coordinates with transcriptional repressors to promote breast tumorigenesis

Yongqiang Hou, Wei Liu, Xianfu Yi, Yang Yang, Dongxue Su, Wei Huang, Hefen Yu, Xu Teng, Ying Yang, Wei Feng, Tao Zhang, Jie Gao, Kai Zhang, Rongfang Qiu, Yan Wang*

*Corresponding author. Email: yanwang@tmu.edu.cn

Published 15 April 2020, *Sci. Adv.* **6**, eaaz0356 (2020)
DOI: 10.1126/sciadv.aaz0356

This PDF file includes:

Supplementary Materials and Methods
Tables S1 to S4
Figs. S1 to S6

Supplementary Materials and Methods

EdU assays

Briefly, breast cancer cells were incubated with EdU in medium (50 mM) for 2 h. Then, the cells were fixed in 4% paraformaldehyde for 10 min, permeabilized with 0.5% triton-X-100, and stained with Alexa Fluor® 488, according to the manufacturer's instructions of the Click-iT® EdU Alexa Fluor® 488 imaging Kit (Life Technologies). Then, cells were stained with DAPI (Sigma-Aldrich). Photographs of the cells were scanned independently using a multi-tracking mode with an OLYMPUS confocal microscope. EdU^+ cells were identified by ImagePro Plus software.

.

Cell flow cytometry

MDA-MB-231 cells transfected with siRNA against PHF20L1 were synchronized using thymidine blockade for 12 h and then released for 12 h. After that, cells were blocked for another 12 h and released with 10% FBS for an appropriate period. Cells were then trypsinized, washed with PBS, and fixed in 70% ethanol at 4°C overnight. After washing with PBS, cells were incubated with RNase A (Sigma) in PBS for 30 min at 37°C and then stained with 50 mg/ml propidium iodide (PI). Cell cycle data were collected using FACS Calibur (Becton Dickinson) and analyzed with ModFit LT 3.0 (Verity Software House Inc., Topsham, ME).

FPLC chromatography

HEK293T cells or HEK293T cells stably expressing FLAG-tagged PHF20L1 were washed twice with cold PBS, scraped, and collected by centrifugation at 1500 × *g* for 5 min. Nuclear and cytoplasmic proteins were extracted using a nuclear-cytosol extraction kit (Applygen Technologies Inc, Beijing, China). The supernatant (nuclear fraction) was collected by centrifugation at 13,000 × *g* for 30 min at 4°C. The protein concentration was determined using a BCA Pierce protein assay kit (Thermo Scientific), and 6 mg of nuclear proteins or FLAG-tagged PHF20L1-coupled proteins were used for the FPLC assay. The proteins were concentrated to 1 ml using a Millipore Ultra free centrifugal filter apparatus (10 kDa nominal molecular mass limit), and then applied to an 850 x 20 mm Superose size 6 exclusion column (Amersham Biosciences) that had been equilibrated with buffer D containing 1 mM dithiothreitol and calibrated with protein standards (blue dextran, 2,000 kDa; thyroglobulin, 669 kDa; ferritin, 440 kDa; catalase, 232 kDa; bovine serum albumin, 67 kDa; and RNase A, 13.7 kDa, all from Amersham Biosciences). The column was eluted at a flow rate of 0.5 ml/min and fractions were collected for further western blotting analysis.

Supplementary Table S1. Mass spectrometry results of PHF20L1-containing protein complex in HEK293T cell.

Description	Score	Coverage	Proteins	Unique Peptides	Peptides	PSMs	AAs
PHF20L1	1229.50	38.64	7	3	39	337	1017
Mi-2β	905.69	25.88	8	47	55	334	1905
MTA2	645.63	52.99	2	28	33	226	668
GATAD2A	447.86	57.66	13	26	28	139	633
GATAD2B	369.55	53.79	1	28	30	115	593
MTA1	338.35	55.30	13	25	31	120	698
RBBP4	320.74	31.95	17	6	13	100	410
RBBP7	301.26	37.18	7	1	14	90	425
HDAC2	259.74	28.82	17	6	11	96	458
Mi-2α	255.28	13.15	6	17	25	114	2000
HDAC1	228.57	29.05	8	8	13	86	482
MTA3	195.26	37.16	8	11	19	70	514
MBD3	173.79	50.52	5	18	18	65	291
MTA2	44.66	11.23	2	5	7	21	668
EED	6.72	8.00	4	3	3	3	400
SUZ12	4.97	3.07	2	2	2	2	716
EZH2	2.12	1.46	5	1	1	1	707

Supplemental Table S2. Mass spectrometry results of PHF20L1-containing protein complex in MDA-MB-231 cell.

Description	Score	Coverage	Proteins	Unique Peptides	Peptides	PSMs	AAs
PHF20L1	441.45	52.21	7	3	55	211	1017
Mi-2β	79.38	25.20	5	24	31	42	1905
MTA2	68.74	32.49	2	14	18	44	668
GATAD2A	60.62	41.45	13	13	15	33	633
GATAD2B	96.41	53.63	1	17	19	46	593
MTA1	66.92	33.52	10	14	20	35	698
RBBP4	115.03	54.10	12	8	11	33	410
RBBP7	105.94	56.00	7	9	12	32	425
HDAC2	124.52	48.57	15	10	16	53	458
Mi-2α	16.16	9.30	6	7	12	14	2000
HDAC1	114.12	36.31	6	7	13	50	482
MTA3	31.68	27.24	8	6	12	22	514
MBD3	24.21	41.58	3	10	11	26	291
MTA2	68.74	32.49	2	14	18	44	668
EED	5.80	12.67	6	4	4	5	400
SUZ12	6.00	4.55	2	3	3	4	716
EZH2	6.09	2.31	10	2	2	3	707

Supplemental Table S3. siRNA and shRNA sequences.

siRNA sequences	
PHF20L1#1	CUAUUACCCUGCCAAGAUU
PHF20L1#2	UUGGACAGACUGUCGCUAU
PHF20L1#3	GCAAGAAUCUUCAGUACCA
SETDB1#1	CCCGAGGCUUUGCUCUUAA
SETDB1#2	GCUGAGACACCAAACGUCA
TDRKH#1	UUAGAGGUUCCCUGAUAUA
TDRKH#2	GUACACAAAGGAUACGCAA
FXR1#1	CGAGCUGAGUGAUUGGUCA
FXR1#2	GAAACGGAAUCUGAGCGUA
LBR#1	UGAUUGGAUGGGUGGUUAU
LBR#2	AGCGUGUGCCCUACCGUAU
KDM4C#1	GCUUGAAUCUCCCAAGAUA
KDM4C#2	CAAAGUAUCUUGGAUCAAA
SND1#1	GGUACCAUCCUUCAUCCAA
SND1#2	GGACAAGGCCGGCAACUUU
FMR1#1	GUUUGGAGAGAUUACAAAU
FMR1#2	GCGUUUGGAGAGAUUACAA
UHRF2#1	CGUCUCUUCUUCCAUUACA
UHRF2#2	GUAAUGUGGCUUAUCAUAU
PHF19#1	CUCGUGACUUUCGAAGAUA
PHF19#2	CCUCGUGACUUUCGAAGAU
MSL3#1	CAGUGAAAUAUUCGUGGAA
MSL3#2	CCAGUCCGCCUUUGUUGAA
KDM4A#1	GUUGAGGAUGGUCUUACCU

siRNA sequences	
KDM4A#2	GGACUUAGCUUCAUAACUA
SMNDC1#1	GCCAGGUAAAGAGGAGUAU
SMNDC1#2	GAAAGUGAAAUGGCAACAA
UHRF1#1	GGUCAAUGAGUACGUCGAU
UHRF1#2	CUUUGAUUCGUUCCUUCUU
TDRD6#1	GUUUAGGCAGCUACCAAUA
TDRD6#2	AGUGCACCCUGGCUGAUAU
ARID4B#1	AGAUAGAGGUACACCUAUU
ARID4B#2	ACCUUAGGGUGGCUUUAAU
MYC#1	ACUCGGUGCAGCCGUAUUU
MYC#2	CCUGAGACAGAUCAGCAAC
MYC#3	CAGGAACUAUGACCUCGAC
MYC#4	CCCAAGGUAGUUAUCCUUA
MTA1#1	UCACGGACAUUCAGCAAGA
MTA1#2	GGACCAAACCGCAGUAACA
EZH2#1	CGGCUCCUCUAACCAUGUU
EZH2#2	CCCAACAUAGAUGGACCAA
shRNA sequences	
shPHF20L1	CTATTACCCTGCCAAGATT
shEZH2	GAGCTCCTCTGAAGCAAAT
shMTA1	TCACGGACATTCAGCAAGA

Supplemental Table S4. The primers used in qRT-PCR and qChIP.

qRT-PCR primers			
Species	**Genes**	**Forward Primer Sequences**	**Reverse Primer Sequences**
human	PHF20L1	GGGCAAGATGTTGGTCCATTT	TCTCAAGGGGTCGCAATCTAT
human	ACTB	CATGTACGTTGCTATCCAGGC	CCTTATCAAGATGCGAACTCACA
human	SETDB1	AGGAACTTCGGCATTTCATCG	TGTCCCGGTATTGTAGTCCCA
human	TDRKH	TCTACTGAACGGACTTCTTGGA	CTGCTTTCCCTATACCTGCGG
human	FXR1	GAGAAGACGGTATGGTTCCATTT	AGGCGTTCCATTCTTAGCTGT
human	LBR	CGAGGGAGTCGATCAAGGTCA	CTTCAGAATCAGCGGAGTCAAT
human	KDM4C	CATGGAGTCTAAAGGAGCCCA	TGTACTGAGTGAACAGTCCTGA
human	SND1	CCTGAGCGGCAGATCAACC	AGGTAGATCATGCCATACTCTCG
human	FMR1	TATGCAGCATGTGATGCAACT	TTGTGGCAGGTTTGTTGGGAT
human	UHRF2	ATTGAGGACGTGTCTCGCAAA	GGTCTGGGCGAACTAGCAG
human	PHF19	ACTCGGGACTCCTATGGTGC	CCTCCGTCAGTTTGGACATCA
human	MSL3	AACAGGAGGAAACGGTTAGTGA	TGTGGCATAACGTGATGGTGA
human	KDM4A	ATCCCAGTGCTAGGATAATGACC	ACTCTTTTGGAGGAACAACCTTG
human	SMNDC1	AGCGGAGATTGAGGAGATAGATG	GGTTCAACAGTGGAGTCACTTC
human	UHRF1	GCCATACCCTCTTCGACTACG	GCCCCAATTCCGTCTCATCC
human	MTF2	GTCCACCTGGCCCATATACAA	CCGTGAAATCCACATCTGAGG
human	TDRD6	GACGTGCATCCCGATGTGAT	AGCAGGAAGACACGGCTCT
human	BRCA1	GAAACCGTGCCAAAAGACTTC	CCAAGGTTAGAGAGTTGGACAC
human	RASSF1	AGGACGGTTCTTACACAGGCT	TGGGCAGGTAAAAGGAAGTGC
human	FBXW7	GGCCAAAATGATTCCCAGCAA	ACTGGAGTTCGTGACACTGTTA
human	KISS1	AGCAGCTAGAATCCCTGGG	AGGCCGAAGGAGTTCCAGT
human	HIC1	GTCGTGCGACAAGAGCTACAA	CGTTGCTGTGCGAACTTGC
human	LDHA	ATGGCAACTCTAAAGGATCAGC	CCAACCCCAACAACTGTAATCT
human	PGK1	TGGACGTTAAAGGGAAGCGG	GCTCATAAGGACTACCGACTTGG
human	SIRT1	TAGCCTTGTCAGATAAGGAAGGA	ACAGCTTCACAGTCAACTTTGT
human	ENO1	AAAGCTGGTGCCGTTGAGAA	GGTTGTGGTAAACCTCTGCTC
human	PKM	ATGTCGAAGCCCCATAGTGAA	TGGGTGGTGAATCAATGTCCA
human	MYC	GGCTCCTGGCAAAAGGTCA	CTGCGTAGTTGTGCTGATGT
human	MTA1	ACGCAACCCTGTCAGTCTG	GGGCAGGTCCACCATTTCC

qRT-PCR primers			
Species	**Genes**	**Forward Primer Sequences**	**Reverse Primer Sequences**
human	EZH2	AATCAGAGTACATGCGACTGAGA	GCTGTATCCTTCGCTGTTTCC
human	TUBB	TGGACTCTGTTCGCTCAGGT	TGCCTCCTTCCGTACCACAT
mouse	Phf20l1	CGTCCTGGAATCACTTTTGAGA	TTGCTGTCCCAGTAAATCCAC
mouse	Ldha	CAAAGACTACTGTGTAACTGCGA	TGGACTGTACTTGACAATGTTGG
mouse	Pgk1	ATGTCGCTTTCCAACAAGCTG	GCTCCATTGTCCAAGCAGAAT
mouse	Gapdh	AGGTCGGTGTGAACGGATTTG	GGGGTCGTTGATGGCAACA
mouse	Eno1	TGCGTCCACTGGCATCTAC	CAGAGCAGGCGCAATAGTTTTA
mouse	Hk2	ATGATCGCCTGCTTATTCACG	CGCCTAGAAATCTCCAGAAGGG
mouse	Sirt1	TGATTGGCACCGATCCTCG	CCACAGCGTCATATCATCCAG
mouse	Pkm	CGCCTGGACATTGACTCTG	GAAATTCAGCCGAGCCACATT
qChIP primers			
Species	**Genes**	**Forward Primer Sequences**	**Reverse Primer Sequences**
human	PHF20L1#1	CAAGTGTCATTATTGGAAGACC	CAGTATTCAAGTTTCTGATGGT
human	PHF20L1#2	GAGGCTAGAACTTTAGCTGCAA	CTATGGAGATTTGGGTGCTGGA
human	PHF20L1#3	ACCTGGATGGTAGAGCAG	TAAAATGTAATTAGAAGTGAGA
human	PHF20L1#4	TGCTCAGTCTAGCTCAGTAGCT	TTTTCTCTTGGTTTCCACTGTA
human	PHF20L1#5	GTTTGGGAGGGTAGGTAT	TATTACTAAGATGTGTGTACGC
human	PHF20L1#6	AGTTTGCCAAGGGTCCTG	GGTTACTGGAGGAAAGAACG
human	BRCA1	GGGGATTGGGACCTCTTC	TCTACCTGAGTTTGCCATAAAG
human	GATA2	CAATTACCGACTGTCAATCCCG	TCCTCCAGCCCTCTTCCCT
human	GSTM2	TGAATCCCTGAGCACCAA	CCGCTGCACTCTAGGTTAG
human	HIC1	GGGCAGGACTTTTCCGAACT	CCTACCCGCAGGTTCGCC
human	KISS1	ACATTTCACACTGACACTTCT	GGACTTTCTCCTTCTTACCG
human	STMN3	CAGACCTGGACCCTGTTGG	AGATGAGCCCCCATTCCCCA
human	VILL	AAGTGCCAAGACATAACAA	GTTCAAGTGATCTTCCCGCT
human	ZNF512B	GGATGCGGCGGAATCAGGT	CGCTGGCCTGTGGTTCCC
human	ACTB	CCAGCACCCCAAGGCG	GGCTTCGCCGCACAGT

fig. S1

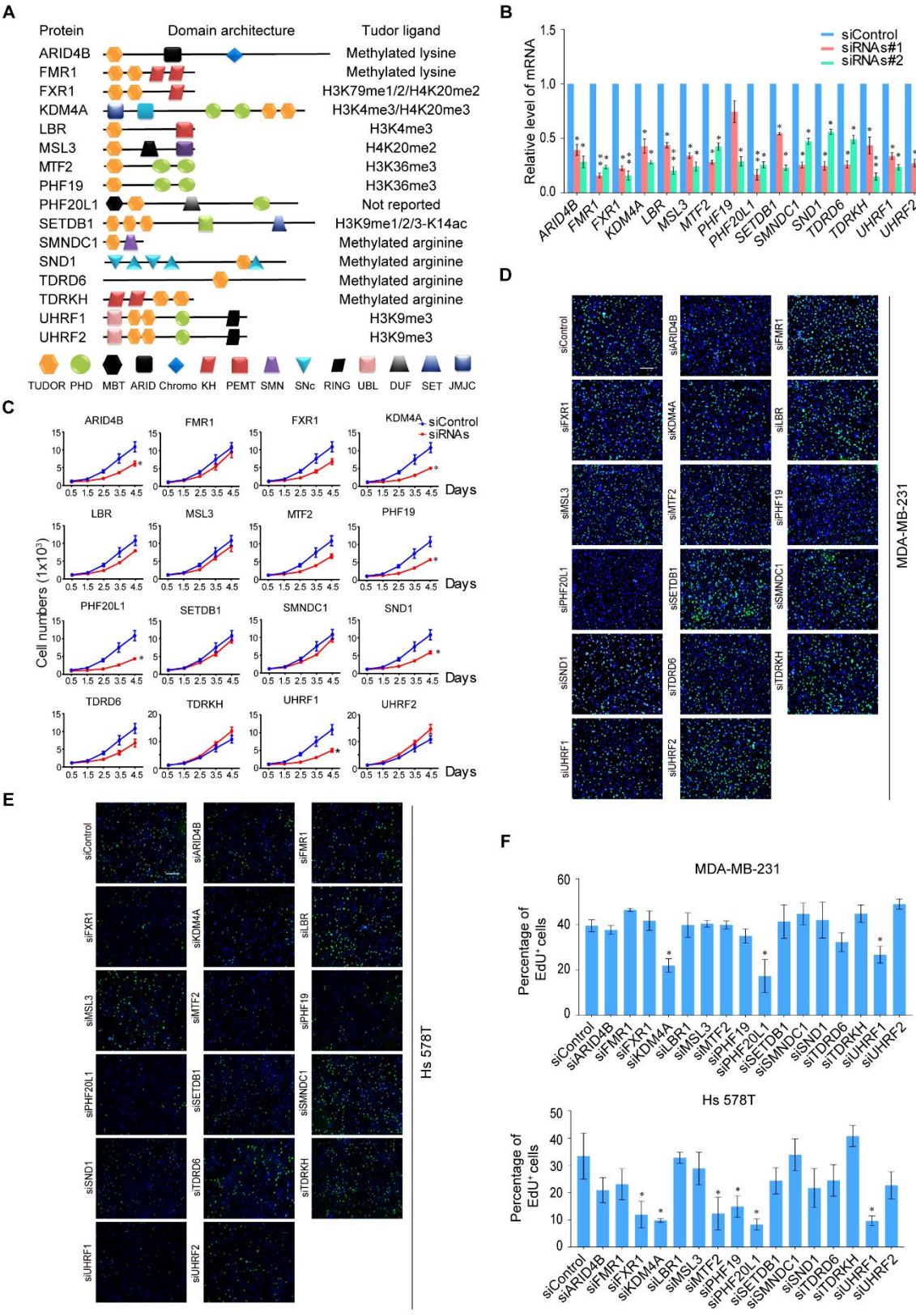
A
Protein
Domain architecture
Tudor ligand
ARID4B
Methylated lysine
FMR1
Methylated lysine
FXR1
H3K79me1/2/H4K20me2
KDM4A
H3K4me3/H4K20me3
LBR
H3K4me3
MSL3
H4K20me2
MTF2
H3K36me3
PHF19
H3K36me3
PHF20L1
Not reported
SETDB1
H3K9me1/2/3-K14ac
SMNDC1
Methylated arginine
SND1
Methylated arginine
TDRD6
Methylated arginine
TDRKH
Methylated arginine
UHRF1
H3K9me3
UHRF2
H3K9me3
TUDOR PHD MBT ARID Chromo KH PEMT SMN SNc RING UBL DUF SET JMJC
B
Relative level of mRNA
siControl
siRNAs#1
siRNAs#2
ARID4B FMR1 FXR1 KDM4A LBR MSL3 MTF2 PHF19 PHF20L1 SETDB1 SMNDC1 SND1 TDRD6 TDRKH UHRF1 UHRF2
C
Cell numbers (1x10³)
siControl
siRNAs
Days
D
siControl siARID4B siFMR1 siFXR1 siKDM4A siLBR siMSL3 siMTF2 siPHF19 siPHF20L1 siSETDB1 siSMNDC1 siSND1 siTDRD6 siTDRKH siUHRF1 siUHRF2
MDA-MB-231
E
Hs 578T
F
MDA-MB-231
Percentage of EdU⁺ cells
Hs 578T

Fig. S1. PHF20L1 is critical for breast cancer cell proliferation. (**A**) Schematic representation of TDRDs and the TUDOR ligand involved. (**B**) The knockdown efficiency are measured by qRT-PCR in MDA-MB-231 cells. (**C**) Growth curve assays were performed in MDA-MB-231 cells transfected with siRNAs against the indicated TDRDs. The cell number was measured by Operetta CLS™ for biological triplicate experiments. Data represent the mean ± SD. Two-tailed unpaired t test, $*p < 0.05$, $**p < 0.01$ at the final day. (**D** and **E**) EdU assays were performed in TDRDs-deficient MDA-MB-231 or Hs 578T cells. Representative images are shown. Scale bar, 50 μm. (**F**) Statistical analysis of EdU assays in TDRDs-deficient MDA-MB-231 cells (upper panel) or Hs 578T cells (lower panel) are shown. Data represent the mean ± SD. Two-tailed unpaired t test, $*p < 0.05$, $**p < 0.01$ (B and F).

fig. S2

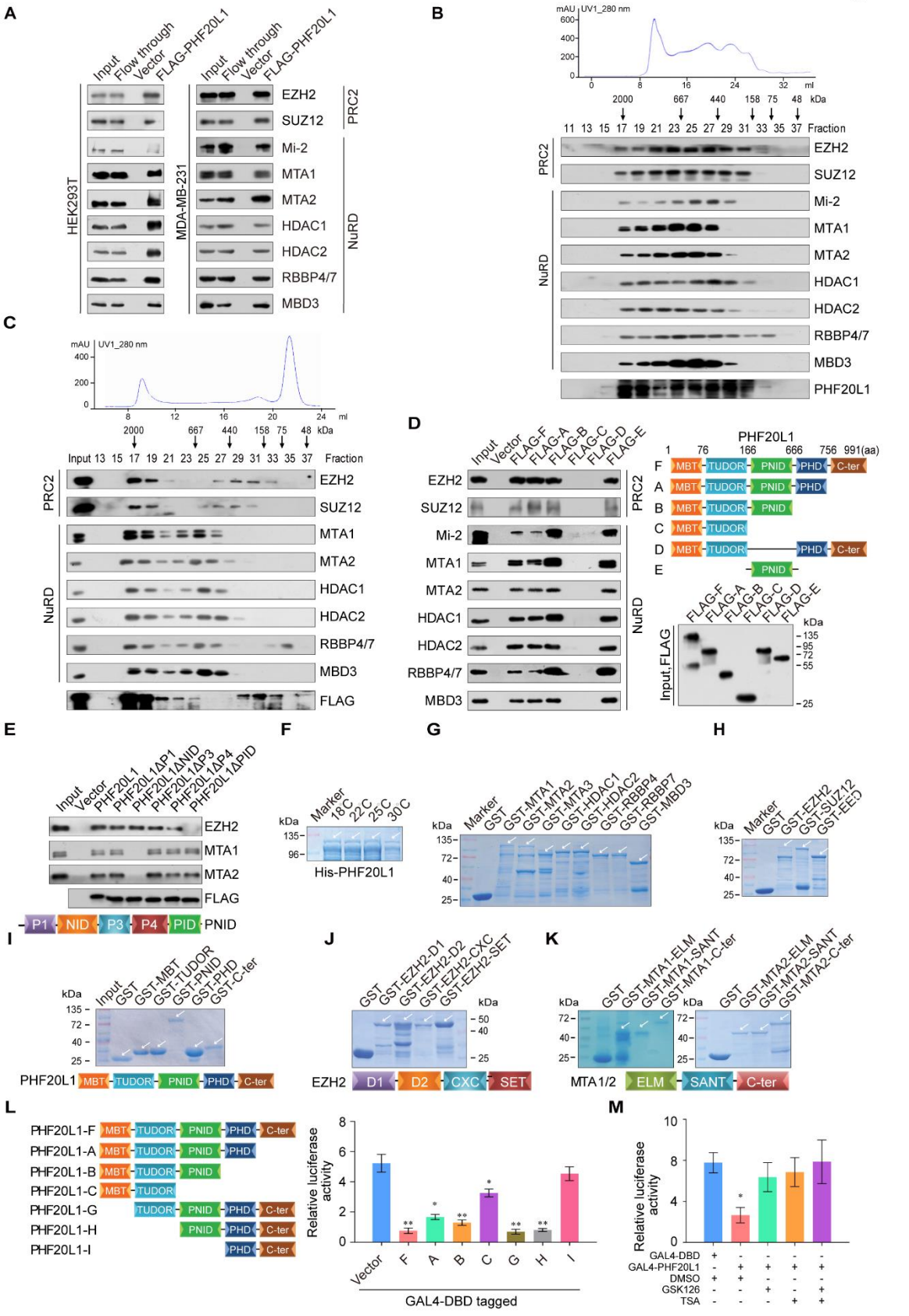
A
Input
Flow through
Vector
FLAG-PHF20L1
HEK293T
MDA-MB-231
EZH2
SUZ12
Mi-2
MTA1
MTA2
HDAC1
HDAC2
RBBP4/7
MBD3
PRC2
NuRD
B
mAU UV1_280 nm
Fraction
C
FLAG
D
PHF20L1
MBT
TUDOR
PNID
PHD
C-ter
E
Input
Vector
PHF20L1
PHF20L1ΔP1
PHF20L1ΔNID
PHF20L1ΔP3
PHF20L1ΔP4
PHF20L1ΔPID
P1
NID
P3
P4
PID
PNID
F
Marker
18°C
22°C
25°C
30°C
His-PHF20L1
G
GST
GST-MTA1
GST-MTA2
GST-MTA3
GST-HDAC1
GST-HDAC2
GST-RBBP4
GST-RBBP7
GST-MBD3
H
GST-EZH2
GST-SUZ12
GST-EED
I
GST-MBT
GST-TUDOR
GST-PNID
GST-PHD
GST-C-ter
J
GST-EZH2-D1
GST-EZH2-D2
GST-EZH2-CXC
GST-EZH2-SET
D1
D2
CXC
SET
K
GST-MTA1-ELM
GST-MTA1-SANT
GST-MTA1-C-ter
GST-MTA2-ELM
GST-MTA2-SANT
GST-MTA2-C-ter
ELM
SANT
C-ter
L
PHF20L1-F
PHF20L1-A
PHF20L1-B
PHF20L1-C
PHF20L1-G
PHF20L1-H
PHF20L1-I
Relative luciferase activity
Vector
GAL4-DBD tagged
M
GAL4-DBD
GAL4-PHF20L1
DMSO
GSK126
TSA

Fig. S2. PHF20L1 is a transcriptional repressor that interacts with PRC2 and the NuRD complex. (**A**) Column-bound proteins were analyzed using western blotting with antibodies against the indicated proteins, related to Fig. 3A. (**B**) Co-fractionation of PHF20L1, PRC2, and the NuRD complex by fast protein liquid chromatography. Nuclear extracts of HEK293T cells were fractionated on Superose 6 size exclusion columns. Chromatographic elution profiles and western blotting analysis of the chromatographic fractions are shown. (**C**) Co-fractionation of PHF20L1, PRC2, and the NuRD complex by fast protein liquid chromatography. FLAG-PHF20L1 multiprotein complexes were fractionated on Superose 6 size exclusion columns. Chromatographic elution profiles and western blotting analysis of the chromatographic fractions are shown. (**D**) Co-IP analysis of the molecular interaction between PRC2, NuRD and PHF20L1 with cellular lysates from HEK293T cells expressing FLAG-tagged full-length or deletions of PHF20L1. The conserved domains of PHF20L1 were divided into the MBT, TUDOR, PNID, PHD, and C-terminal domain. (**E**) Co-IP analysis of the molecular interaction between PNID domain and PRC2 or NuRD with cellular lysates from HEK293T cells expressing FLAG-tagged deletions of PHF20L1. The PNID domain of PHF20L1 were divided into P1, P2 (NID), P3, P4, and P5 (PID). (**F**-**K**) GST/His fused proteins purified from BL21 *Escherichia coli*. used in Fig. 3. (**L**) The PNID domain is responsible for the transcriptional repression activity of PHF20L1. HEK293T cells were transfected with the indicated Gal4-deletions and the Gal4 luciferase reporter activity was measured. (**M**) Effect of GSK126 and TSA treatment on PHF20L1 repressive activity. HEK293T cells were transfected as indicated constructs along with treatment of inhibitor against EZH2 or MTA1, and Gal4 luciferase reporter activity was measured. Data shown are mean ± SD. Two-tailed unpaired t test. $*p < 0.05$, $**p < 0.01$ (L and M).

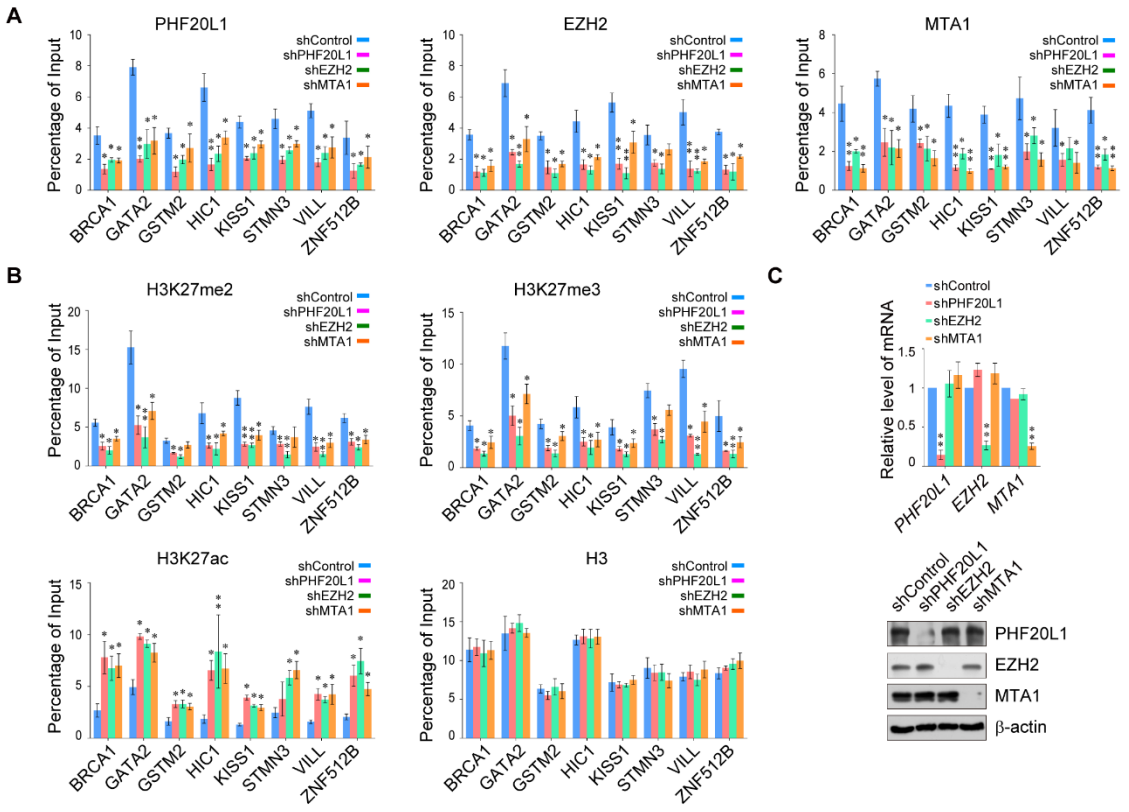

Fig. S3. PHF20L1 loss-of-function impairs the deposition of PRC2 and the NuRD complex. (**A** and **B**) qChIP analysis using anti-PHF20L1, anti-EZH2, anti-MTA1, anti-H3K27me2, anti-H3K27me3, anti-H3K27ac, and anti-H3 antibodies in control, PHF20L1 KD, EZH2 KD and MTA1 KD MDA-MB-231 cells. (**C**) The knockdown of PHF20L1 could not affect the expression of EZH2 and MTA1 in MDA-MB-231 cells. Data shown are mean ± SD. Two-tailed unpaired t test, *$p < 0.05$, **$p < 0.01$ (A-C).

fig. S4

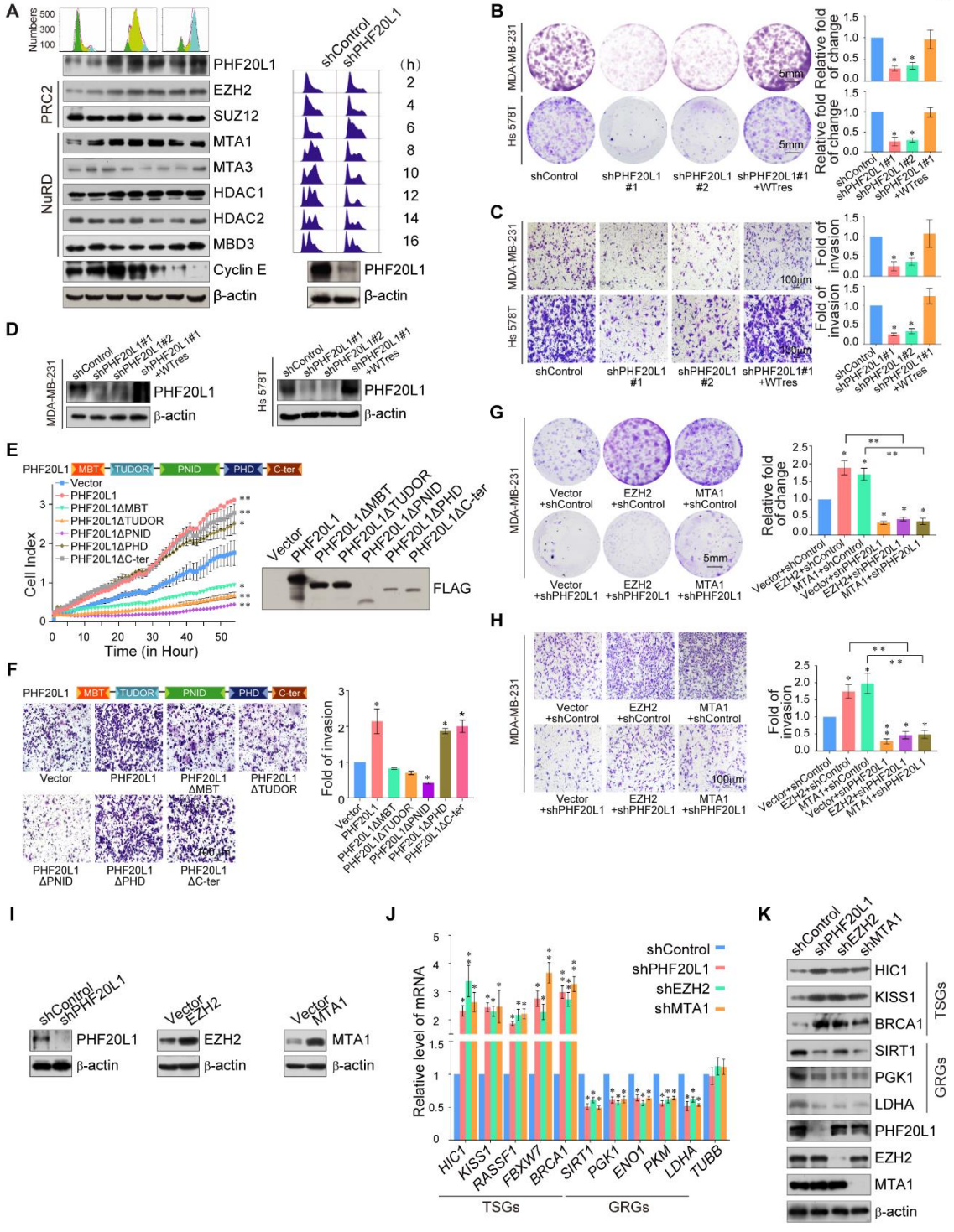
A
Numbers
PHF20L1
PRC2
EZH2
SUZ12
NuRD
MTA1
MTA3
HDAC1
HDAC2
MBD3
Cyclin E
β-actin
shControl
shPHF20L1
(h)
B
MDA-MB-231
Hs 578T
shControl
shPHF20L1 #1
shPHF20L1 #2
shPHF20L1#1 +WTres
Relative fold of change
C
Fold of invasion
100μm
D
E
Cell Index
Time (in Hour)
Vector
PHF20L1
PHF20L1ΔMBT
PHF20L1ΔTUDOR
PHF20L1ΔPNID
PHF20L1ΔPHD
PHF20L1ΔC-ter
FLAG
F
G
Vector +shControl
EZH2 +shControl
MTA1 +shControl
Vector +shPHF20L1
EZH2 +shPHF20L1
MTA1 +shPHF20L1
H
I
J
Relative level of mRNA
shControl
shPHF20L1
shEZH2
shMTA1
HIC1
KISS1
RASSF1
FBXW7
BRCA1
SIRT1
PGK1
ENO1
PKM
LDHA
TUBB
TSGs
GRGs
K
HIC1
KISS1
BRCA1
SIRT1
PGK1
LDHA
PHF20L1
EZH2
MTA1
β-actin

Fig. S4. PHF20L1 acts in concert with its associated corepressor complexes to promote breast cancer carcinogenesis. (**A**) Western blotting analysis of the expression of PHF20L1, PRC2, and NuRD subunits at different cell cycle stages (left panel). Propidium iodide FACS analysis of the cell cycle following PHF20L1 depletion in MDA-MB-231 cells are shown (right panel). (**B**) Colony formation assays were performed in control, PHF20L1 KD and PHF20L1 KD MDA-MB-231 or Hs 578T cells stably expressing shRNA-resistant PHF20L1 (WTres). (**C**) Transwell invasion assays were performed in cells as in (B). (**D**) Western blotting analysis of PHF20L1 in cells as in (B and C). (**E**) The growth curves were measured in MDA-MB-231 cells stably expressing vector, full-length PHF20L1 or deletion mutations of the MBT, TUDOR, PNID, PHD, and C-terminal domains. Data shown are mean ± SD. Two-tailed unpaired t test, $*p < 0.05$, $**p < 0.01$ at the final minute. (**F**) Transwell invasion assays were performed in MDA-MB-231 cells as in (E). (**G**) MDA-MB-231 cells infected with lentiviruses carrying shControl, shPHF20L1, shControl + EZH2, shControl + MTA1, shPHF20L1 + EZH2, shPHF20L1 + MTA1, then colony formation assays were performed. (**H**) Transwell invasion assays were performed in cells as in (G). (**I**) Western blotting analysis of PHF20L1, EZH2 and MTA1 in cells as in (G and H). (**J** and **K**) The expression of the indicated tumor suppressor genes and glycolysis related genes were measured by qRT-PCR (J) and western blotting analysis (K) in control, PHF20L1 KD, EZH2 KD and MTA1 KD MDA-MB-231 cells. *TUBB* served as an irrelevant control gene. The mRNA levels were normalized to those of *ACTB*. β-actin served as loading control. All error bars represent mean ± SD. Two-tailed unpaired t test. $*p < 0.05$, $**p < 0.01$ (B, C, F, G, H, and J).

fig. S5

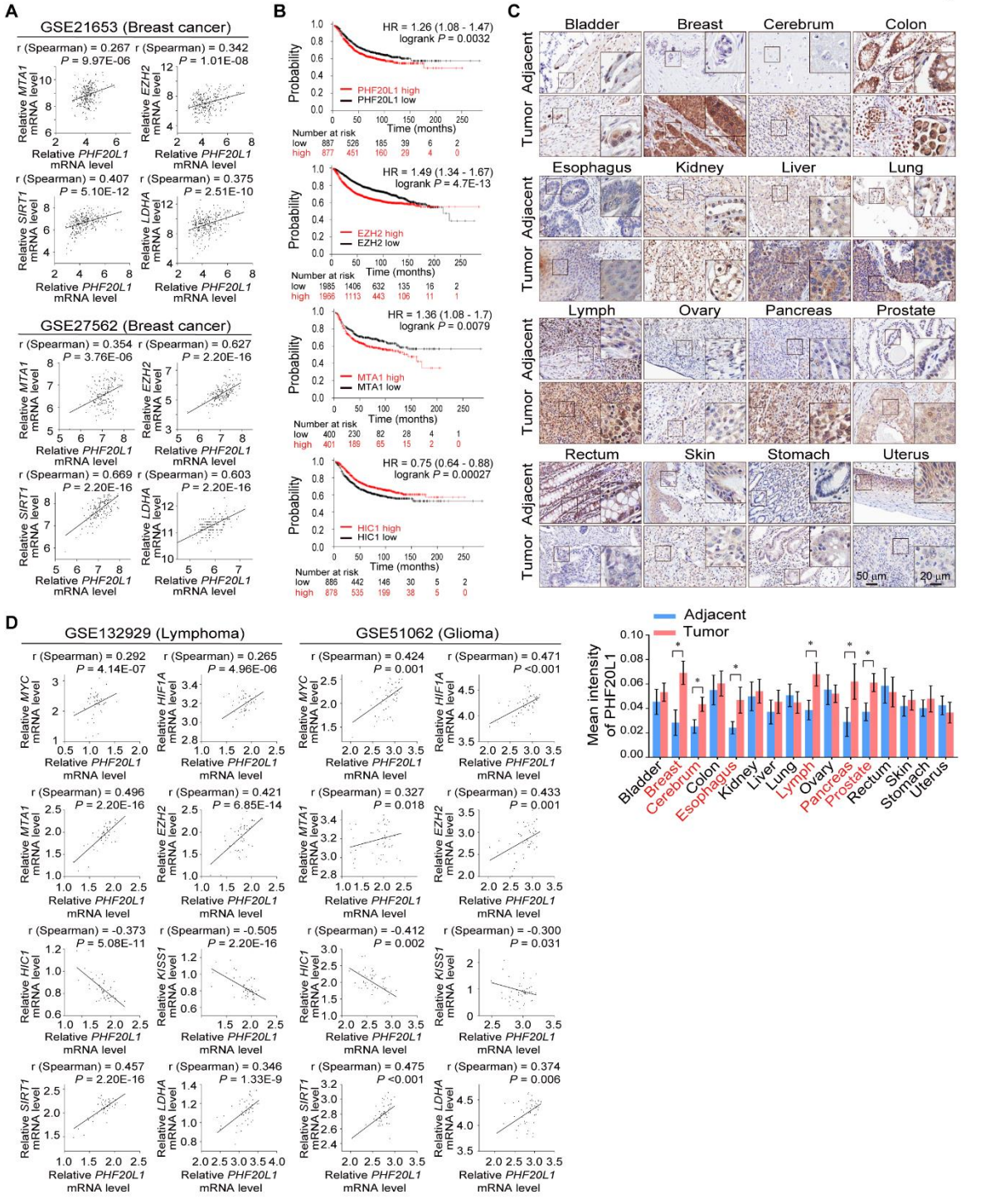
A
GSE21653 (Breast cancer)
r (Spearman) = 0.267
P = 9.97E-06
r (Spearman) = 0.342
P = 1.01E-08
r (Spearman) = 0.407
P = 5.10E-12
r (Spearman) = 0.375
P = 2.51E-10
Relative MTA1 mRNA level
Relative EZH2 mRNA level
Relative SIRT1 mRNA level
Relative LDHA mRNA level
Relative PHF20L1 mRNA level
GSE27562 (Breast cancer)
r (Spearman) = 0.354
P = 3.76E-06
r (Spearman) = 0.627
P = 2.20E-16
r (Spearman) = 0.669
P = 2.20E-16
r (Spearman) = 0.603
P = 2.20E-16
B
HR = 1.26 (1.08 - 1.47)
logrank P = 0.0032
PHF20L1 high
PHF20L1 low
HR = 1.49 (1.34 - 1.67)
logrank P = 4.7E-13
EZH2 high
EZH2 low
HR = 1.36 (1.08 - 1.7)
logrank P = 0.0079
MTA1 high
MTA1 low
HR = 0.75 (0.64 - 0.88)
logrank P = 0.00027
HIC1 high
HIC1 low
Probability
Time (months)
Number at risk
C
Bladder
Breast
Cerebrum
Colon
Esophagus
Kidney
Liver
Lung
Lymph
Ovary
Pancreas
Prostate
Rectum
Skin
Stomach
Uterus
Adjacent
Tumor
50 μm
20 μm
D
GSE132929 (Lymphoma)
GSE51062 (Glioma)
r (Spearman) = 0.292
P = 4.14E-07
r (Spearman) = 0.265
P = 4.96E-06
r (Spearman) = 0.424
P = 0.001
r (Spearman) = 0.471
P <0.001
r (Spearman) = 0.496
P = 2.20E-16
r (Spearman) = 0.421
P = 6.85E-14
r (Spearman) = 0.327
P = 0.018
r (Spearman) = 0.433
P = 0.001
r (Spearman) = -0.373
P = 5.08E-11
r (Spearman) = -0.505
P = 2.20E-16
r (Spearman) = -0.412
P = 0.002
r (Spearman) = -0.300
P = 0.031
r (Spearman) = 0.457
P = 2.20E-16
r (Spearman) = 0.346
P = 1.33E-9
r (Spearman) = 0.475
P <0.001
r (Spearman) = 0.374
P = 0.006
Relative MYC mRNA level
Relative HIF1A mRNA level
Relative MTA1 mRNA level
Relative EZH2 mRNA level
Relative HIC1 mRNA level
Relative KISS1 mRNA level
Relative SIRT1 mRNA level
Relative LDHA mRNA level
Relative PHF20L1 mRNA level
Adjacent
Tumor
Mean intensity of PHF20L1
Bladder
Breast
Cerebrum
Colon
Esophagus
Kidney
Liver
Lung
Lymph
Ovary
Pancreas
Prostate
Rectum
Skin
Stomach
Uterus

Fig. S5. Clinical relevance of the MYC/HIF1α-(PHF20L1-EZH2-MTA1)-HIC1/KISS1 axis in cancers. (**A**) Analysis of public datasets (GSE21653 and GSE27562) from breast cancer for the correlation of *MTA1*, *EZH2*, *SIRT1*, *LDHA*, and *PHF20L1*, related to Fig. 5K. (**B**) Kaplan-Meier survival analysis for the relationship between survival time and expression levels of PHF20L1, EZH2, MTA1, and HIC1 in breast cancer using the online tool (http://kmplot.com/analysis/). (**C**) Immunohistochemical staining of PHF20L1 in paired samples of multiple carcinomas versus adjacent normal tissues and values for the stainings were determined by ImagePro Plus software and are presented with statistical histogram. Representative images of 200-fold magnifications of each type of paired tumor section are presented. Data shown are mean ± SD. Two-tailed unpaired t test. $*p < 0.05$, $**p < 0.01$. (**D**) Analysis of two published clinical datasets (GSE132929 and GSE51062) from lymphoma and glioma for the correlation of MYC, HIF1A, MTA1, EZH2, HIC1, KISS1, SIRT1, LDHA, and PHF20L1.

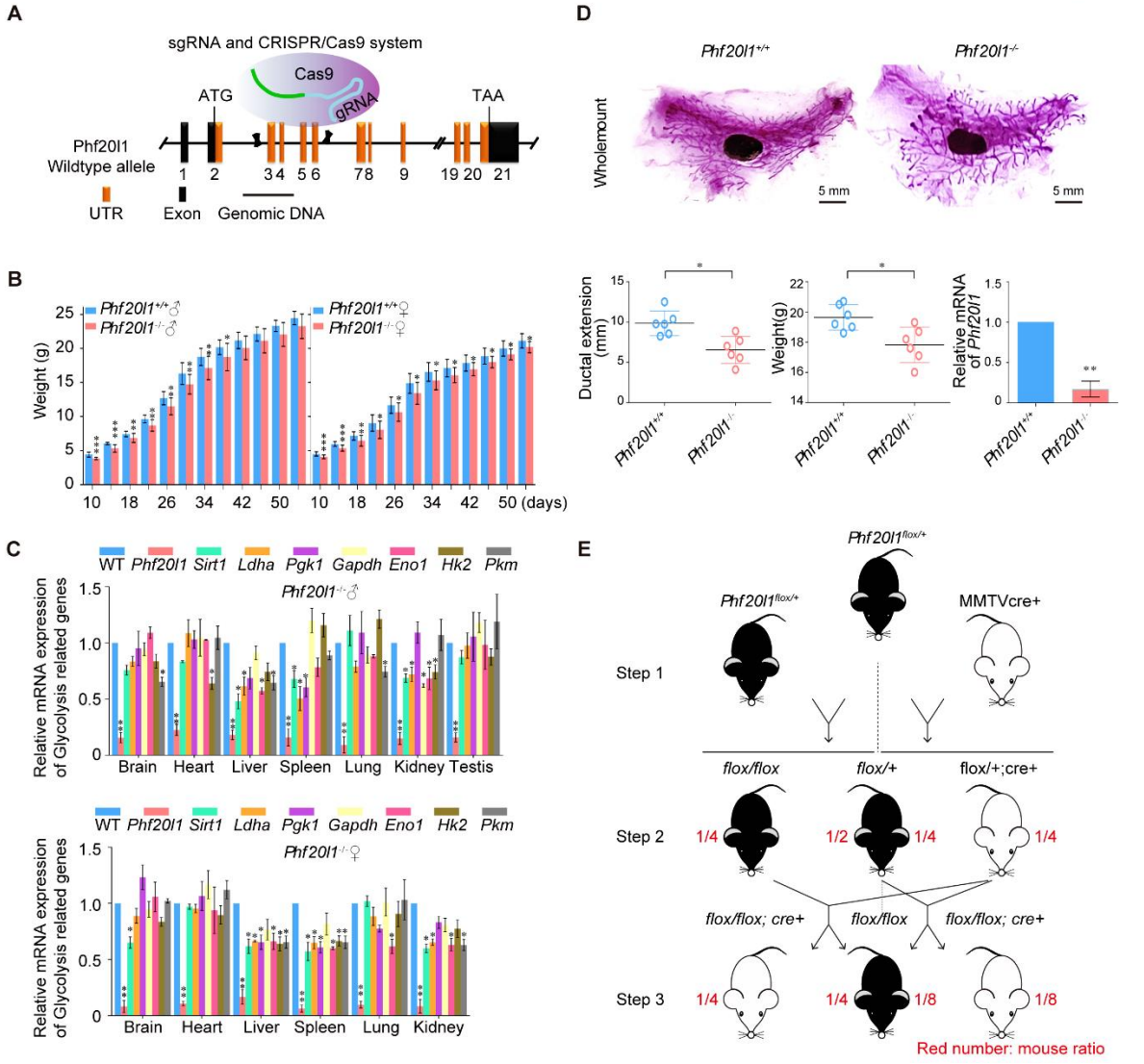

Fig. S6. *Phf20l1* deletion induces growth retardation and delay of mammary ductal outgrowth *in vivo*. (**A**) Schematic of the depletion of the *Phf20l1* gene using the CRISPR/Cas9 system. (**B**) Bodyweights of *Phf20l1* KO and wild-type mice. n=15. (**C**) The expression profiles of glycolysis-related genes of various major organs in 4-week-old *Phf20l1*-null mice or littermate wild-type mice. The expression levels of the indicated genes were measured by qRT-PCR. (**D**) Mammary ductal developmental defects in *Phf20l1* knockout mice at about 8-week-old. (**E**) Strategy for the generation of $Phf20l1^{flox/flox}$; MMTV-Cre mice. All error bars represent mean ± SD. Two-tailed unpaired t test, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$ (B-D).