Supplementary Information

Supplementary Figure Legends

Figure S1. CARM1 is required for estrogen-induced gene transcriptional activation.

(A) Box plot showing the expression of CARM1 (FPKM) in a cohort of clinical breast cancer (n=1,102) and normal (n=113) samples from TCGA (The Cancer Genome Atlas).

(B, C) Kaplan Meier survival analyses for DMFS (distal metastasis free survival) (n=1,747) and OS (overall survival) (n=1,402) of breast cancer patients using CARM1 as input.

(D, E) RNA-seq as described in Figure 1A for specific genes were shown as indicated.

(F) Correlation of CARM1’s effects on whole transcriptome in response to estrogen between two RNA-seq biological repeats (n=11,816).

(G) Correlation of CARM1 effects on whole transcriptome in response to estrogen based on RNA-seq between two independent siRNAs targeting CARM1 (siCARM1 (1) and siCARM1 (2)) (n=12,234).

(H, I) MCF7 cells as described in (G) were subjected to immunoblotting (IB) (H) and RT-qPCR (I) analysis to examine the levels of indicated protein and mRNA, respectively. Actin was served as a loading control.

(J) MCF7 cells as described in Figure 1E were subjected to RT-qPCR analysis to examine the expression of genes as indicated.

(K) MCF7 cells as described in Figure 1E were subjected to immunoblotting (IB) analysis to examine the protein levels of CARM1. Actin was served as a loading control.

(L) Genomic DNA was extracted from CARM1 knockout cells, followed by PCR using specific primer sets spanning gRNA targeting region (boxed in blue). The resultant PCR products were subjected to Sanger sequencing as shown. Nucleotide being inserted was highlighted in red.
MCF7 cells as described in Figure 1F were subjected to immunoblotting (IB) analysis to examine the protein levels of CARM1 by using anti-specific CARM1 antibodies from two independent vendors as indicated. Actin was served as a loading control.

Pol II ChIP-seq signals across specific genes as indicated. Units are mean tags per bin for 16 bins across the transcribed gene region with 2 kb upstream (4 bins of 500 bp each) and 5 kb downstream flanking regions (10 bins of 500 bp each).

Pol II ChIP-seq tag density distribution centered on cognate enhancers of specific genes as indicated (± 4,000 bp).

Heat map (R) and box plots (S) representation of the expression levels (FPKM) for genes induced by estrogen and dependent on CARM1 in the cohort of clinical samples, both normal and tumor, as described in (A).

**Figure S2. CARM1 is recruited onto ERα-bound active enhancers upon estrogen stimulation.**

MCF7 cells treated with or without estrogen (E2, 10^{-7} M, 1 hr) were subjected to ChIP with anti-CARM1 antibody followed by qPCR analysis with primers specifically targeting enhancer (e) regions as indicated. ChIP signals were presented as percentage of inputs (± s.e.m.).

ChIP-seq tag distribution, including ERα, H3K4me1, H3K4me2, H3K4me3, H3K27Ac, P300, MED1, MED12, H3K9me3 and H3K27me3, centered on all of its own sites (left panels) or estrogen-induced CARM1 sites (right panels) (± 4,000 bp).

Genome browser views of CARM1, ERα, H3K4me1, H3K4me2, H3K4me3, H3K27Ac, P300, MED1, MED12, H3K9me3 and H3K27me3 ChIP-seq in the presence and absence of estrogen on selected active enhancer regions, such as the ones in the vicinity of estrogen-induced coding gene NRIP1 (left panel on the top), PGR (right panel on the top), SIAH2 (left panel on the bottom) and
APOAI (right panel on the bottom), were shown. Boxed regions indicated active enhancers. ChIP-seq views, except for CARM1, on SIAH2 have been shown in our previous study [1].

**Figure S3. Global mapping of CARM1 substrates.**

(A) Experimental workflow for global mapping of CARM1 substrates. FASP: filter aided sample preparation.

(B) The distribution of proteins with different number of arginine methylation sites. All the proteins with more than ten methylation sites were shown in oval.

(C) Schematic representation of domain architecture of representative proteins with clustered arginine methylation sites in the “CARM1 methylome”. Arginine methylation sites were shown by matchsticks. CID: CTD-interacting domain; CTD: Spt5 C-terminal domain; G-patch: glycine rich nucleic binding domain; DRBM: double strand RNA-binding domain; PHD: the plant homeodomain zinc finger; TFS2M: transcription factor S-II (TFIIS) domain in the central regions of transcription elongation factor S-II (and elsewhere); RING: ring finger; BBC: B-Box C-terminal domain; BBOX: B-Box-type zinc finger; BROMO: bromo domain; TFS2N: transcription factor S-II (TFIIS) domain in the N-terminus of transcription elongation factor S-II (and elsewhere); ZnF_C3H1: (CCCH) type Zinc finger (Znf) domain; KH: K homology RNA-binding domain; IF rod: Intermediate filament (IF) rod domain; WW: domain with 2 conserved Trp (W) residues; FF: a novel motif that often accompanies WW domains and contains two conserved Phe (F) residues.

(D, E) Rates of somatic mutations at (D) and in the vicinity of (E) CARM1 methylation sites in human cancers. Rates of somatic mutations at and in the vicinity of (±5 nucleotides) CARM1-dependent methylation sites (n=589) were significantly higher than those of arginines in the proteome (P = 1.62e-11 and 7.42e-45, respectively).

**Figure S4. CARM1 hypermethylates MED12.**
(A) Human MED12 C-terminus (amino acid (aa) 1616-2177). Arginine residues with mono-methylation only, di-methylation only and both mono- and di-methylation were highlighted in red, green and yellow, respectively.

**Figure S5. CARM-mediated methylation is involved in estrogen-induced gene transcriptional activation.**

(A) MCF7 cells as described in Figure 5F were subjected to RT-qPCR analysis to examine the expression of CARM1.

(B) MCF7 cells as described in Figure 5F were subjected to immunoblotting analysis to examine the expression of MED12 and CARM1. Actin was served as a loading control.

(C) Correlation of MED12’s effects on transcriptome in response to estrogen based on RNA-seq between two independent siRNAs targeting MED12 (siMED12 (1) and siMED12 (2)).

(D) Correlation of CARM1’s effects on MED12 binding based on ChIP-seq between two independent siRNAs targeting CARM1 (siCARM1 (1) and siCARM1 (2)).

(E, F) MCF7 cells were transfected with control siRNA (siCTL) or two independent siRNAs specific against CARM1 (siCARM1 (1) and siCARM1 (2)) in stripping medium for three days, and treated with or without estrogen (E$_2$, $10^{-7}$ M, 6 hrs), followed by RT-qPCR (E) and immunoblotting (IB) (F) analysis to examine the mRNA and protein levels of MED12, respectively. Actin was served as a loading control.

(G) MCF7 cells as described in Figure 5J were subjected to RT-qPCR analysis to examine the fold induction by estrogen for selected genes as indicated (± s.e.m., **P<0.01, ***P<0.001).

(H, I) MCF7 cells as described in Figure 5J were subjected to RT-qPCR analysis to examine the expression of CARM1 (H) and MED12 (I) (± s.e.m., ***P<0.001).

**Figure S6. TDRD3 is required for estrogen-induced gene transcriptional activation.**
(A-C) MCF7 cells as described in Figure 6G were subjected to qPCR analysis to examine the expression of selected estrogen-induced genes as indicated.

(D-F) MCF7 cells as described in Figure 6I were subjected to qPCR analysis to examine the expression of selected estrogen-induced genes as indicated.

**Figure S7. CARM1 is required for estrogen-induced breast cancer cell growth.**

(A) WT and CARM1 KO MCF7 cells were subjected to cell proliferation assay (± s.e.m., ***P<0.001).

(B) MCF7 cells as described in (A) were subjected to FACS analysis.

(C, D) T47D (C) and BT474 (D) cells transfected with siCTL or siCARM1 were subjected to cell proliferation assay (± s.e.m., *P<0.01, **P<0.01, ***P<0.001).

(E, F) T47D (E) and BT474 (F) cells transfected with siCTL or siCARM1 were subjected to colony formation assay.

**Supplementary Table Legend**

**Table S1. Global mapping of CARM1 substrates.** The total number of arginine methylation sites (the sum of mono- and asymmetrical di-methylation sites after removing duplicates), and the number of mono- or asymmetrical di-methyl arginine methylation sites for proteins which had arginine methylation sites identified (sheet 1), had arginine methylation sites on which methylation signals decreased at least two-fold (sheet 2) or completely abolished (sheet 3), and had all its methylation sites abolished (sheet 4) were shown.

**Table S2. MS2 spectrum of arginine-methylated peptides in MED12.** MS2 spectrum of methylated arginine residues in MED12 C-terminus as described in Figure 4D were shown as indicated.
Table S3. Sequence information for all qPCR primers used in the current study. Sequence information of qPCR primers designed to detect the expression of GREB1, NRIP1, PGR, SIAH2, TSKU, APOA1, ERa, PRSS23, PDCD2L, PSAT1, CAD, CARM1, MED12, TDRD3, SMN, SND1 and SPF30 mRNA, and GREB1, NRIP1, PGR, SIAH2, TSKU and APOA1 enhancer RNA (eRNA), and to detect the binding of CARM1 on TFF1 enhancer (e), GREB1 (e), FOXC1 (e), PGR (e), SIAH2 (e), SMAD7 (e), TSKU (e) and APOA1 (e) regions were shown. F: forward; R: reverse.

Supplementary Reference

1. Gao WW, Xiao RQ, Zhang WJ, Hu YR, Peng BL, Li WJ, et al. JMJD6 Licenses ERalpha-Dependent Enhancer and Coding Gene Activation by Modulating the Recruitment of the CARM1/MED12 Co-activator Complex. Mol Cell. 2018; 70: 340-57 e8.
**Figure S1**

**A**

FPKM (log2)

Normal Tumor

**B**

p=0.0062

CARM1 low expression

CARM1 high expression

**C**

p=0.025

CARM1 low expression

CARM1 high expression

**D**

1kb

chr11:76,493,000-76,511,000

**E**

0.5kb

chr11:116,706,000-116,709,000

**F**

Pearson Correlation Coefficient = 0.75

**G**

Pearson Correlation Coefficient = 0.75

**H**

siCTL (CTL)

siCARM1 (CTL)

siCARM1 (E)

E

IB1

CARM1

IB

ERα

**I**

Normalized Values

**J**

Normalized Values

**K**

shCTL + + - -

shCARM1 - - + +

E2 + - + +

IB: CARM1

IB: Actin

**L**

KO:

WT:

CARM1 (CTST)

CARM1 (Bethyl)

IB: CARM1

IB: Actin

**M**

WT KO

WT KO

**N**

Pol II ChIP-seq tag density

Tag Density

2Kb TSS TTS 5Kb

**O**

Pol II ChIP-seq tag density

Tag Density

2Kb TSS TTS 5Kb

**P**

Pol II ChIP-seq tag density

Tag density

-4000-2000 0 2000 4000

Distance to the center of TSKU enhancer

**Q**

Pol II ChIP-seq tag density

Tag density

-4000-2000 0 2000 4000

Distance to the center of GREB1 enhancer

**R**

Normalized Values

Genes induced by E2

Genes dependent on CARM1

(n=461)

**S**

FPKM (log2)

Normal Tumor

**Coefficient = 0.75**

**Pearson Correlation**

**$p=0.025$**

**$p=0.0062$**

**$2.42E-05$**
Centered on all of its own sites or E\textsubscript{2}-induced CARM1 sites.

Figure S2
Figure S3
Figure S4

A

MED12 (NP_005111.2) aa1616-2177
LAKKLQKELGERQSDSLEKVRULLPLPKQTRDVITCEPQGSLIDTKGN
KIAGFDSIFKKEGLQVSTKQKISPWDLFEGLKPSAPLSWGWGTFVVD
RRVARGEQQRLLLHYTHLRPRAYLEPLPLPDEDDEPAPTTLEP
EKKAKEPPKTDKPGAAPSTEERKKKSTKGGKRSQPATKTEDYGMG
PGRSGPYGVTVPDDLHHNPNGSITHLNYRQGSIGLYTQTNQPLPAGGP
RDVYPVFLPMQKLPPTPTYPQVLPTTMTGVMLEGPSYKTSSYQ
QQPAPVQGQRLQQLQQSGMLGQSSVHOMTPSYYGLQTSQGYT
PYVSHQVLQNTGPGATMVPSYSSQPYQSTHPSTNPTLVDPTRHLQ
QRPSGYVHQAPTYGHGLTSTQRFSHOTLQQTMISTMPSAOGQVQ
AGVRSTAILPEQQQQQQQQQQQQQQQQQQQQQQQQYHIRQQ
QQQLRQQQQQQQQQQQQQQQQQQQQQQQQQQQQHQQQQQQQQQAAP
PQPQPQSQPFQRQGLQQTQQQQQTAALVRQLQQLSNTQPQPST
NIFGTY
Figure S6
Figure S7

A

B

C

D

E

F