

## Supplementary Information

### A prelude to the proximity interaction mapping of CXXC5

Gamze Ayaz<sup>1,4,6\*</sup>, Gizem Turan<sup>1\*</sup>, Çağla Ece Olgun<sup>1\*</sup>, Gizem Kars<sup>1</sup>, Burcu Karakaya<sup>1</sup>, Kerim Yavuz<sup>1</sup>, Öykü Deniz Demiralay<sup>1</sup>, Tolga Can<sup>3</sup>, Mesut Muyan<sup>1,2,6</sup>, and Pelin Yaşar<sup>1,5\*</sup>

<sup>1</sup>Department of Biological Sciences, Middle East Technical University, Ankara 06800, Turkey; <sup>2</sup>Cansyl Laboratories, Middle East Technical University, Ankara 06800, Turkey; <sup>3</sup>Department of Computer Engineering Middle East Technical University, Ankara 06800, Turkey

<sup>4</sup>Current Address: Cancer and Stem Cell Epigenetics Section, Laboratory of Cancer Biology and Genetics, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, 20892, USA

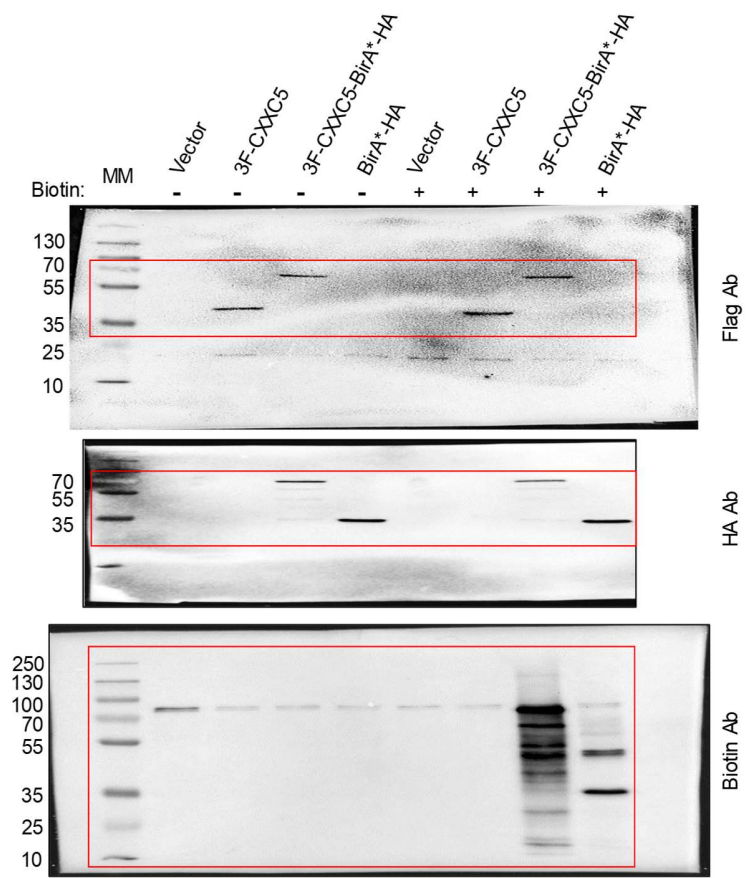
<sup>5</sup>Current Address: Epigenetics and Stem Cell Biology Laboratory, Single Cell Dynamics Group, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA

<sup>6</sup>Corresponding Authors: [gamze.ayazsen@nih.gov](mailto:gamze.ayazsen@nih.gov) and [mmuyan@metu.edu.tr](mailto:mmuyan@metu.edu.tr)

\*Equal contributions; should be considered as first authors

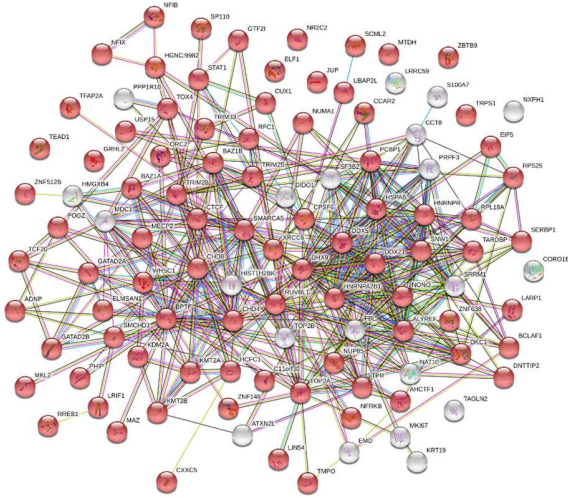
**Short title:** The proximity interaction partners of CXXC5

**Keywords:** CXXC5, BioID, Proximity interactions, EMD, MAZ, MeCP2

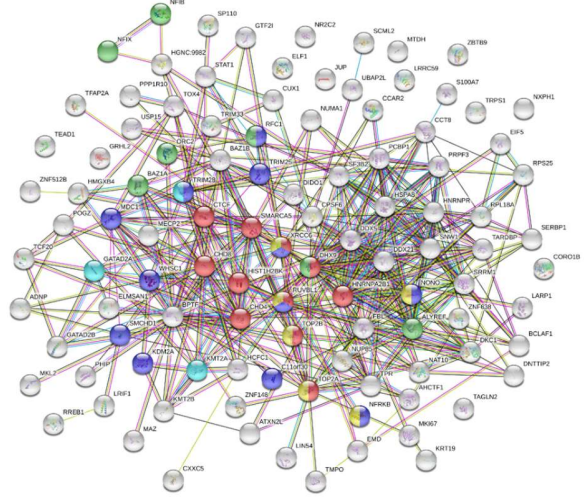


**Figure S1. Biotinylation of endogenous proteins in MCF7 cells.** MCF7 cells were transiently transfected for 24h with an expression vector (pcDNA3.1) bearing the cDNA for 3F-CXXC5, 3F-CXXC5-BirA\*-HA, or BirA\*-HA. Cells were then treated without (-) or with (+) biotin (50  $\mu$ M) and ATP (1 mM) for 16h. Total protein extracts of MCF7 cells were subjected to SDS-10%PAGE followed by WB using the Flag, the HA, or the Biotin antibody followed by an HRP-conjugated goat-anti mouse secondary antibody for the Flag (Advansta R-05071-500) or goat-anti-rabbit secondary antibody for the HA or the Biotin antibody (Advansta R-05072-500). Molecular masses (MM) in kDa are indicated. Cropped images for Fig. 1 are indicated with red boxes.

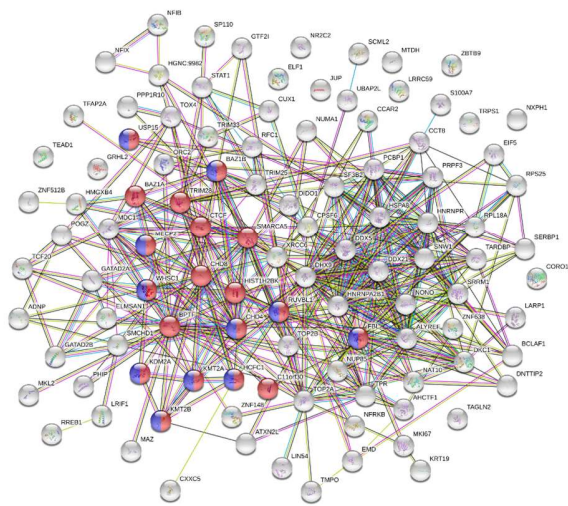
### Gene Expression



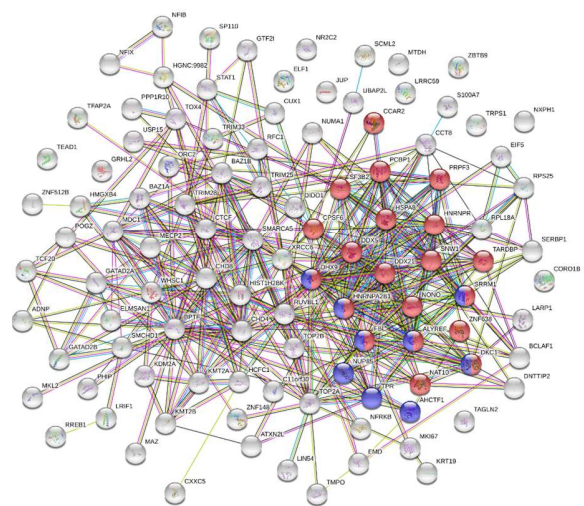
### DNA



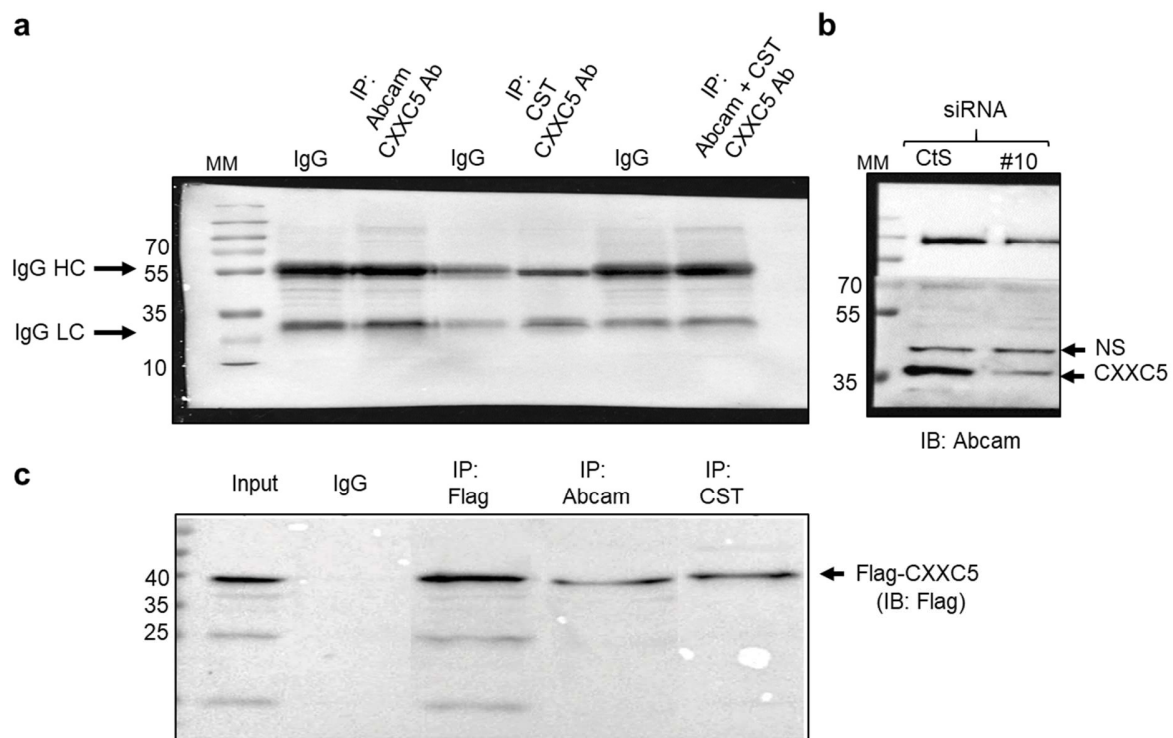
### Chromatin/Histone



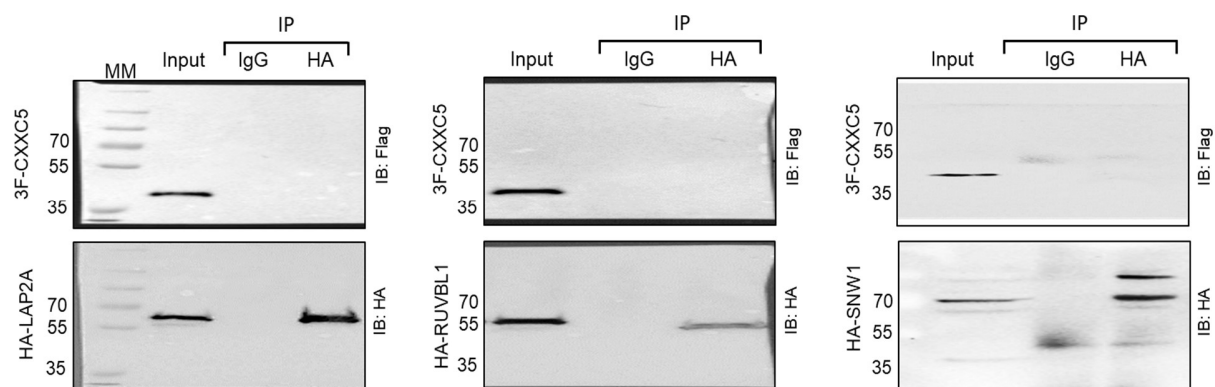
### RNA



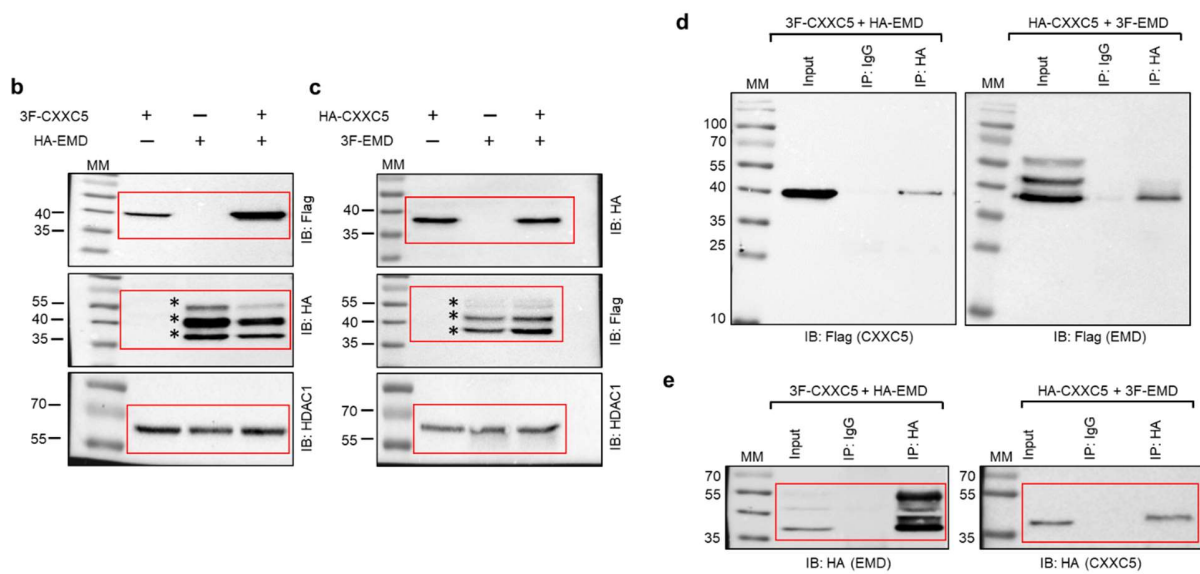
**Figure S2. The proximity interaction mapping of CXXC5 with GO term enrichment analyses by the use of the STRING interaction database.** Images generated by STRING are shown. The majority of proteins with the proximity interactions are involved in the Regulation of Gene Expressions [(GO:0010468, (red))]. GO terms can further be functionally assigned to DNA [GO:0071103, DNA conformation change (red); GO:0006281, DNA repair (dark blue); GO:0006260, DNA replication (green); GO:0006310, DNA recombination (yellow); GO:0006306, DNA methylation (light blue)]; Chromatin and Histone [(GO:0006325, Chromatin Organization (red); GO:0016570, Histone modifications (dark blue)]; and, RNA [GO:0006396, RNA processing (red); GO:0031123, GO:0006403, RNA localization (dark blue)] processes.



**Figure S3. Assessing antibody efficiencies in immunoprecipitation.** (a) The Abcam antibody (106533) is generated against a 15 aa-long peptide from near the amino terminus of the human CXXC5 protein (<https://www.abcam.com/cxxc5-antibody-ab106533.html>); whereas, the CST D1O4P antibody is generated against a peptide corresponding to residues surrounding Leu172 of human CXXC5 protein (<https://www.cellsignal.com/products/primary-antibodies/cxxc5-d1o4p-rabbit-mab/84546>). Nuclear extracts of MCF7 cells (500  $\mu$ g) were subjected to immunoprecipitation by using 5  $\mu$ g normal rabbit IgG (IgG), Abcam, or CST (Cell Signaling Technology, D1O4P) antibody. Since the different regions on CXXC5 are used to generate CXXC5 antibodies, the efficiency of immunoprecipitation could be increased by subjecting the nuclear extracts of MCF7 cells (500  $\mu$ g) to the combination of 2.5  $\mu$ g Abcam and 2.5  $\mu$ g CST antibodies. Immunoprecipitants were subjected to SDS-10% PAGE followed by immunoblotting with the Abcam antibody for visualization. (b) While both antibodies effectively detect CXXC5 in WB analysis of nuclear extracts of MCF7 cell transiently transfected with the control siRNA (CtS) or siRNA#10, shown with the Abcam antibody, the antibodies were ineffective in precipitating the endogenous CXXC5. Molecular masses (MM) are in kDa. IgG heavy (55 kDa) and IgG light (26 kDa) chains are indicated. (c) The efficiencies of antibodies to immunoprecipitate CXXC5 were also tested in transiently transfected HEK293 cells with an expression vector bearing sequences encoding a Flag-tag at the amino-terminus of CXXC5. Nuclear extracts (500  $\mu$ g) were immunoprecipitated by using 5  $\mu$ g of normal mouse IgG (IgG), the Flag-M2 (Sigma-Aldrich, F1804), Abcam, or CST antibody. Precipitates were then subjected to SDS-10% PAGE followed by immunoblotting with the Flag-M2 antibody. Input indicates 10% of the precipitate. Although the efficiency of the Abcam or the CST antibody to immunoprecipitate CXXC5 is lower than that of the Flag-M2 antibody, higher levels of the protein in cells allow an operational immunoprecipitation of CXXC5 by both antibodies.

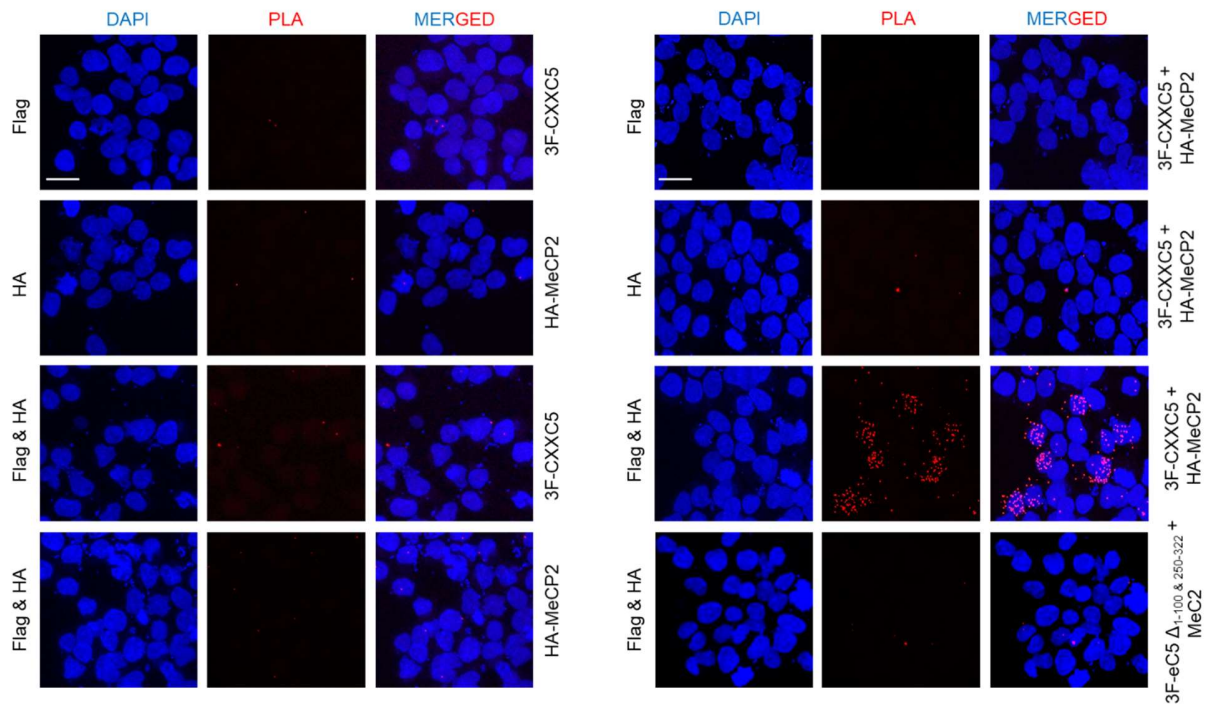


**Figure S4. Initial assessment of CXXC5 interaction with LAP2A, RUVBL1, or SNW1.** The nuclear extracts of HEK293 cells transiently co-transfected with the expression vectors bearing 3Flag-CXXC5 and HA-LAP2A, HA-RUVBL1, or HA-SNW1 were subjected to Co-IP using the HA antibody or the isotype-matched IgG followed by immunoblotting using the Flag or the HA antibody.

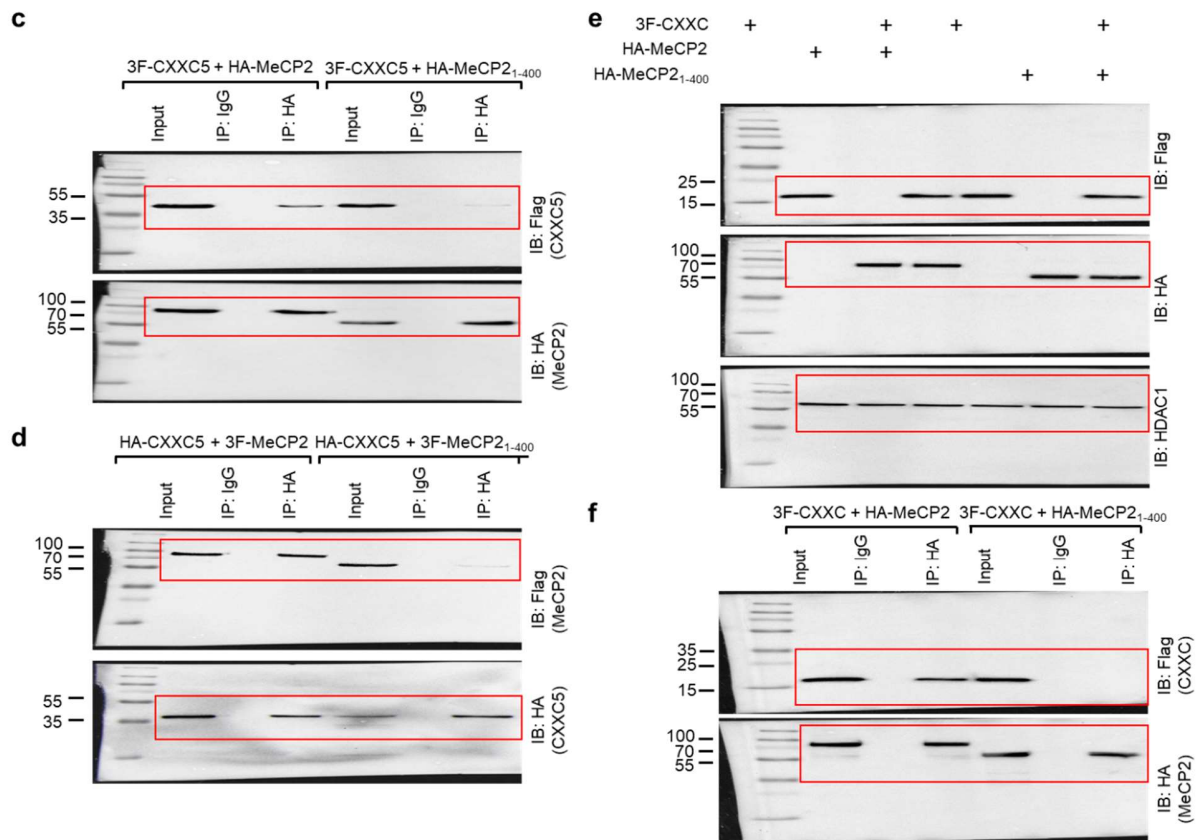


**Figure S5. Interaction of EMD and CXXC5.** (b & c) To examine the protein synthesis, HEK293 cells were transfected with the expression vector bearing (b) 3F-CXXC5 and/or HA-EMD cDNA; or (c) HA-CXXC5 and/or 3F-EMD cDNA for 48h. The synthesis of proteins was assessed by WB using the HA or the Flag antibody. HDAC1 used as a loading control was probed with the HDAC1 antibody. Star denotes distinct EMD species. (d & e) The nuclear extracts (500  $\mu$ g) of transiently co-transfected HEK293 cells were subjected to Co-IP with the HA (d) or the isotype-matched IgG. 50  $\mu$ g of nuclear lysates was used as input control. The precipitates were subjected to SDS-10%PAGE followed with WB using analyzed using the Flag (d) or the HA (e) antibody. Molecular masses (MM) in kDa are indicated. Cropped images for Fig. 2 are indicated with red boxes.



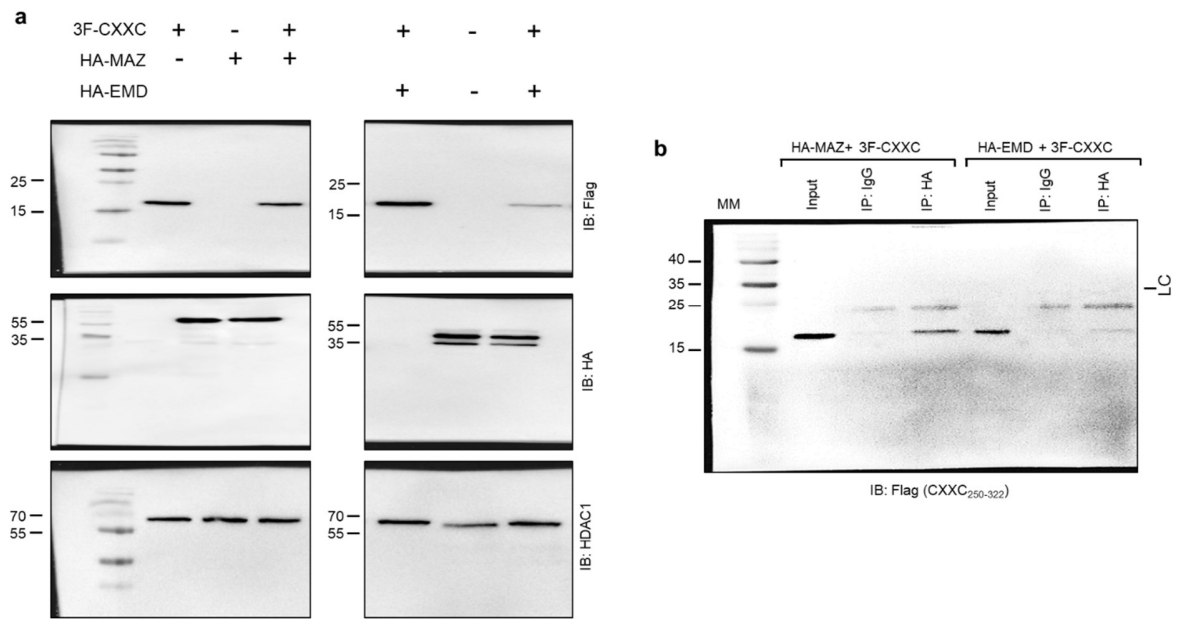


**Figure S6. *In cellula* interaction of CXXC5 and MeCP2.** Proximity ligation assay (PLA). To assess the *in cellula* interaction of CXXC5 and MeCP2, HEK293 cells grown in coverslips were transiently transfected the expression vector bearing the 3F-CXXC5 and/or HA-MeCP2 cDNA. Cells were also transiently co-transfected the expression vector bearing the 3F-eCXXC5 $\Delta$ <sub>1-100 & 250-322 (3F-C5 $\Delta$ <sub>1-100 & 250-322</sub>) and HA-MeCP2 cDNA. Cells were fixed, permeabilized, blocked, and probed with the Flag and/or the HA antibody. Cells were then subjected to fluorescent probes for circular DNA amplification. DAPI was used for nuclear staining. The scale bar is 25  $\mu$ m.</sub>

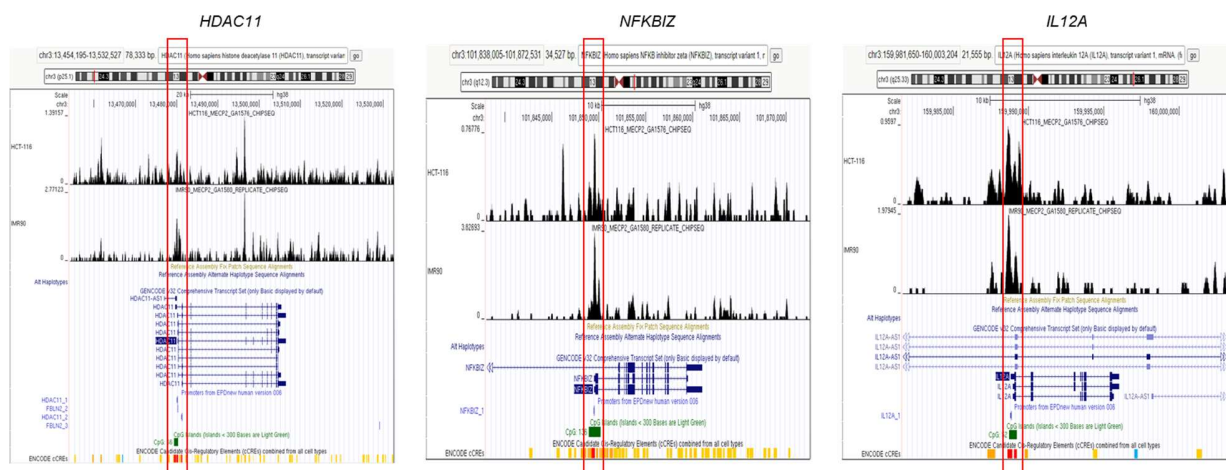


**Figure S7. Identification of a sub-region of MeCP2 critical for interacting with CXXC5.** (c & d) HEK293 cells were transiently co-transfected with the expression vector bearing cDNA for 3F-CXXC5 and HA-MeCP2 or HA-MeCP2<sub>1-400</sub>. Nuclear extracts were subjected to Co-IP with the HA or the isotype-matched IgG. The precipitates were subjected to WB using the Flag antibody. The membrane was re-probed with the HA antibody. 10% of nuclear extracts was used as input control. Molecular masses (MM) in kDa are indicated. (e) Nuclear extracts of HEK293 cells transiently co-transfected with an expression vector bearing cDNA for the 3F-CXXC domain (3F-CXXC), HA-MeCP2, or HA-MeCP2<sub>1-400</sub>, were subjected to WB using the Flag, HA or HDAC1 antibody. Molecular masses (MM) in kDa are indicated. (f) Nuclear extracts, 500 µg, co-synthesizing CXXC, and HA-MeCP2, or HA-MeCP2<sub>1-400</sub>, were subjected to Co-IP with the HA antibody or the isotype-matched IgG. The precipitates were subjected to WB using the Flag antibody. The membrane was also re-probed with the HA antibody. 10% of nuclear extracts was used as input control. Molecular masses (MM) in kDa are indicated. Cropped images for Fig. 7c-f are indicated with red boxes.

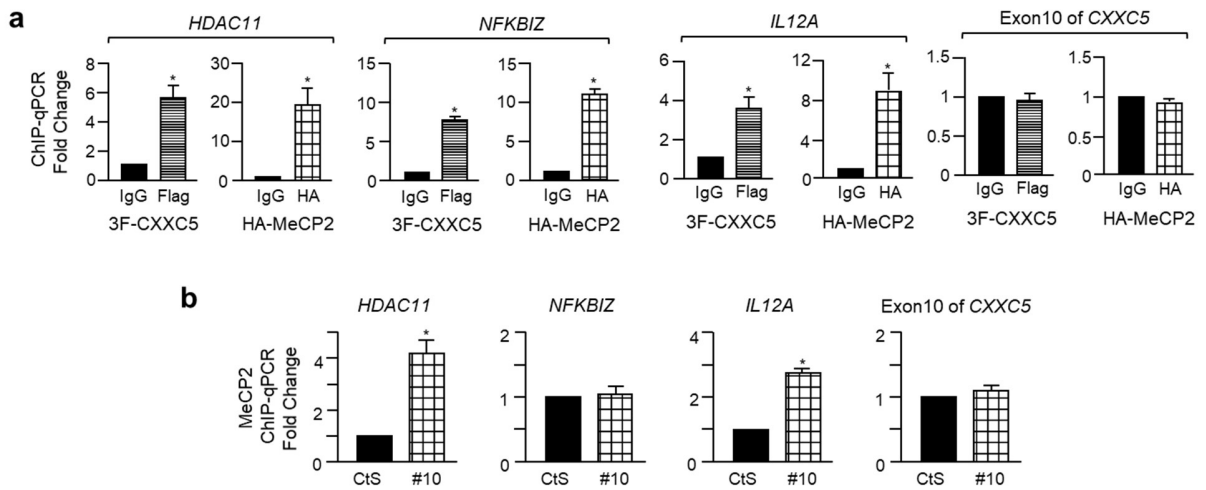




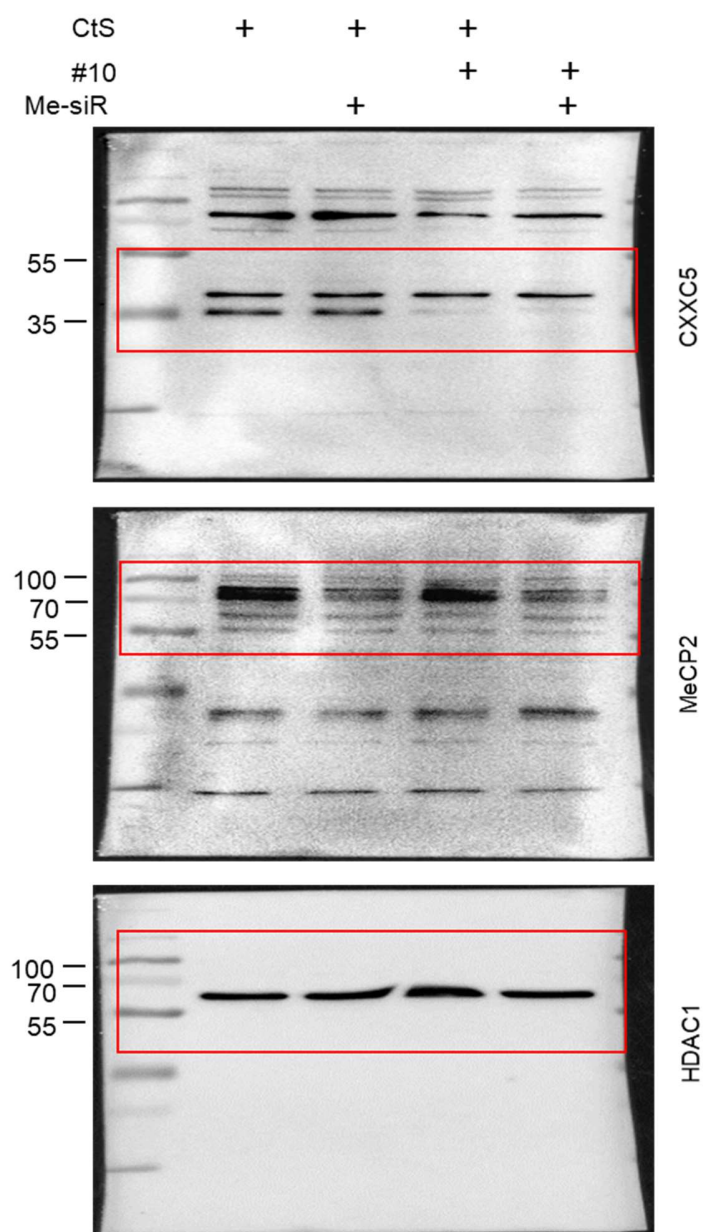
**Figure S8. Interaction of the CXXC domain of CXXC5 with EMD or MAZ.** To assess the interaction between the 3-Flag-CXXC domain (3F-CXXC) of CXXC5 and HA-MAZ or HA-EMD, HEK293 cells were transfected with the expression vector bearing the 3F-CXXC and/or the HA-MAZ or the HA-EMD cDNA for 48h. **(a)** Nuclear extracts (100  $\mu$ g) were subjected to SDS-PAGE followed with WB using the HA or the Flag antibody. HDAC1 used as a loading control was probed with the HDAC1 antibody. **(b)** The nuclear extracts (500  $\mu$ g) of transiently co-transfected HEK293 cells were subjected to Co-IP with the HA or the isotype-matched IgG. 50  $\mu$ g of nuclear lysates were used as the input control. The precipitates were subjected to SDS-PAGE followed with WB using the Flag antibody. Molecular masses (MM) in kDa and the light chain of IgG (LC) are indicated.



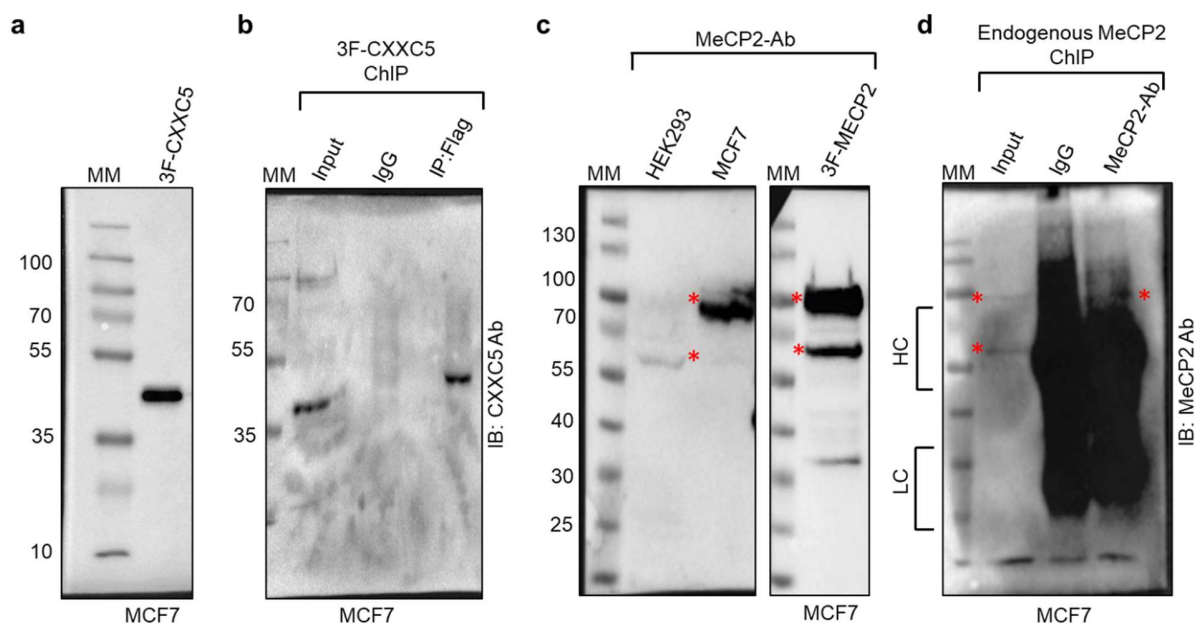
**Figure S9. Binding profile of MeCP2 to the promoter, located within a CGI, of *HDAC11*, *NFKBIZ*, or *IL12A* gene.** Snapshots of the binding profile of MeCP2 (Red rectangle) generated with Cistrome at gene loci using the UCSC genome browser with CGI (green), promoters (light blue arrows), and cis-regulatory elements (red boxes). To assess whether MeCP2 is enriched at the promoter region of *HDAC11*, *NFKBIZ*, or *IL12A*, datasets of IMR-90, a human lung fibroblast cell line, and HCT-166 cells derived from human colon carcinoma at the Cistrome Data Browser were visualized with the UCSC genome browser. Snapshots of the gene loci are shown. The red rectangle indicates the promoter region.



**Figure S10. Assessing the interplay between CXXC5 and MeCP2 in CXXC5 target gene expressions. (a)** To assess whether CXXC5 or MeCP2 is associated with target gene promoters, MCF7 cells were transfected with a vector bearing the 3F-CXXC5 or HA-MeCP2 cDNA for 48h. Cells were processed for ChIP using IgG, the Flag, or the HA antibody. Precipitates were subjected to qPCR using primer sets specific for target gene promoters. Results depict fold changes compared to IgG, which was set to 1. **(b)** To assess whether alterations in CXXC5 levels affect the MeCP2 loading on target promoters, MCF7 cells transfected with CtS or siRNA#10 for 48h were subjected to ChIP using IgG or a MeCP2 antibody. Recovered DNAs were subjected to qPCR using primer sets for target gene promoters. Results, normalized to IgG, depict fold changes compared to CtS, which was set to 1.



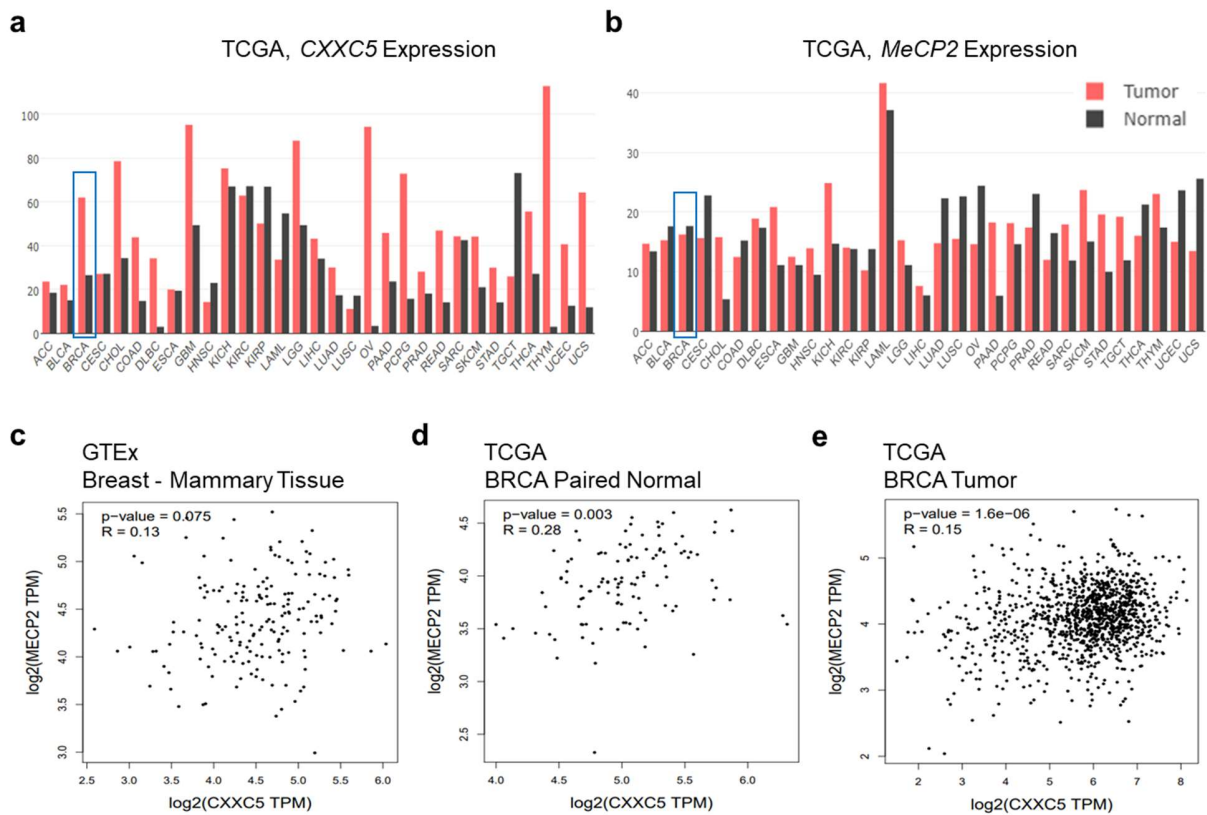
**Figure S11. The effects of siRNA on protein levels of CXXC5 and/or MeCP2 in MCF7 cells.** MCF7 cells were untransfected (UT), or were transiently transfected with CtS, siRNA specific for CXXC5 (#10), and/or siRNA pool for MeCP2 (Me-SiR) for 48h. To equalize the total amount of siRNA (20 nM) used in co-transfection experiments, 10 nM gene-specific siRNA was used together with 10 nM CtS. Nuclear extracts of transfected cells were subjected to WB using the CXXC5 or MeCP2 antibody. HDAC1 was probed with an HDAC1-specific antibody. MMs in kDa are indicated. (c) To assess the effect of reduction in CXXC5 and/or MeCP2 levels on gene expressions, MCF7 cells were transfected with CtS, #10, and/or Me-SiR as indicated for 48h. Cropped images for Fig. 8b are indicated with red boxes.



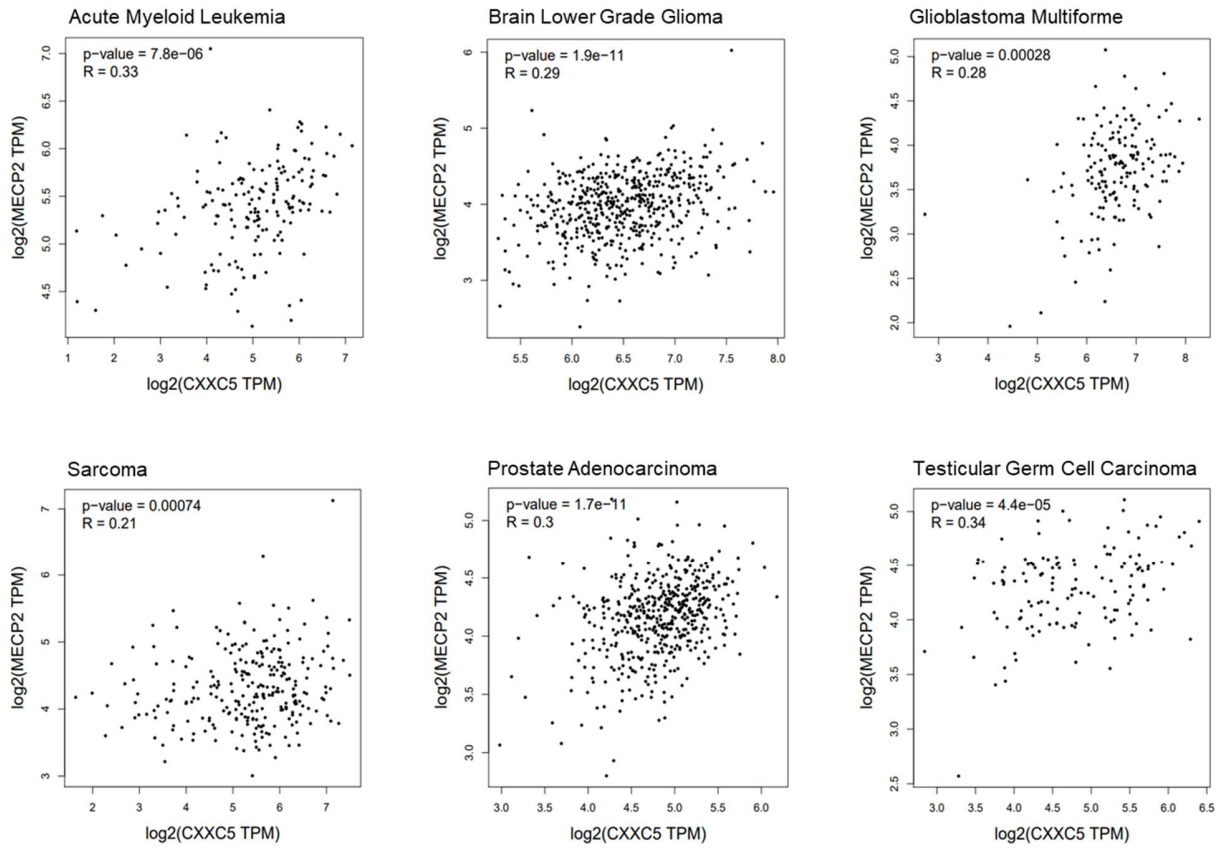
**Figure S12. Immunoprecipitation of 3F-CXXC5 and ChIP of MeCP2.** (a & b) To assess the efficiency of the Flag antibody to precipitate 3F-CXXC5 protein in ChIP assay, (a) nuclear extracts of MCF7 cells transiently transfected with an expression vector bearing the 3F-CXXC5 cDNA for 48h were subjected to SDS-10%PAGE followed with WB using an antibody specific to CXXC5. (b) Transfected MCF7 cells were also subjected to ChIP using the Flag antibody or the isotype-matched IgG followed by immunoblotting using the CXXC5-specific antibody. Molecular masses (MM) in kDa are indicated. (c) To assess the efficiency of the MeCP2 antibody to detect the endogenous MeCP2 protein in WB, the synthesis of the endogenous MeCP2 protein in HEK293 or MCF7 cells in comparison with that of 3F-MeCP2 in transiently transfected MCF7 was assessed in WB with an antibody specific to MeCP2. Molecular masses (MM) in kDa are shown. Asterisks (red) indicate MeCP2 isoforms with molecular masses of approximately 60 and 80 kDa, the latter which is likely a post-translationally modified form. (d) To assess the efficiency of the MeCP2 antibody to precipitate the endogenous MeCP2 protein in ChIP assay, MCF7 cells were also subjected to ChIP using the MeCP2-specific antibody or the isotype-matched IgG followed by immunoblotting using the MeCP2-specific antibody. HC and LC indicate the heavy and light chain of IgG. 5% of ChIP was used as input control.



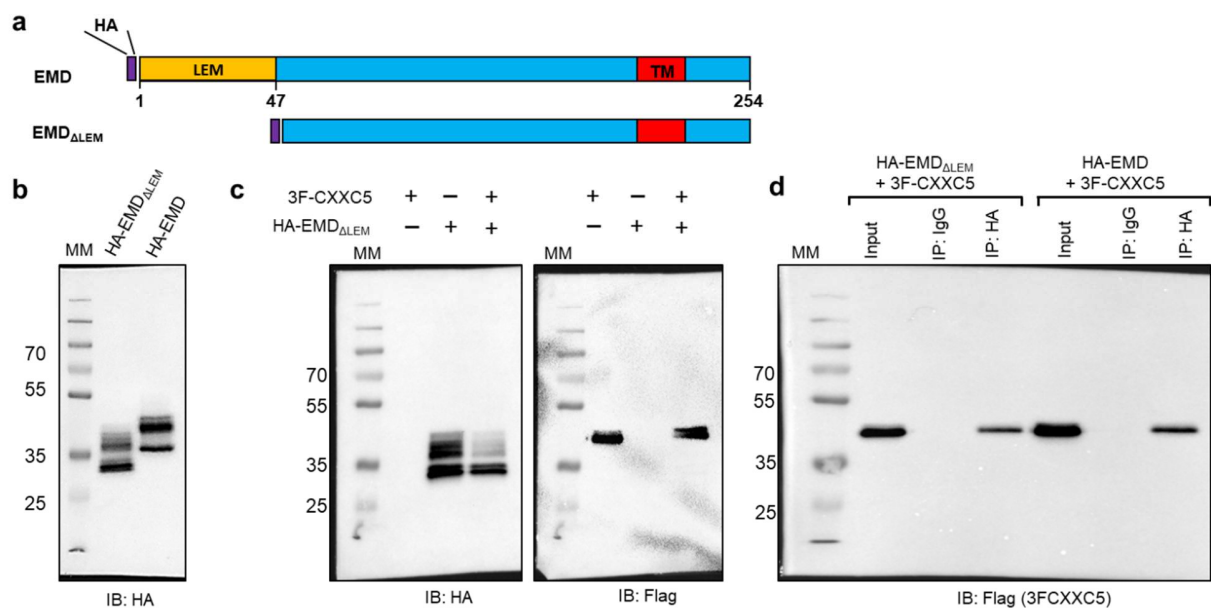




**Figure S14. Expression analyses of *CXXC5* and *MeCP2* across paired normal tissues and tumor samples and the correlation between mRNA expressions of *CXXC5* and *MeCP2* in normal and breast tumor samples. (a & b)** Snapshots of bar-plot analyses of the median *CXXC5* (a) or *MeCP2* (b) expression, as transcripts *per* million (TPM), in various tissue samples are shown. The red bar-plot indicates expressions in tumors and the black bar-plot denotes expressions in paired normal tissues. BRCA (blue rectangle) indicates breast invasive carcinoma. (c-e) Snapshots of dot-plot analyses with the use of the Pearson correlation coefficient indicate the correlation between expressions of *CXXC5* and *MeCP2* in GTEx normal breast tissue (c), TCGA paired normal (d), and TCGA breast tumor samples (e) as log<sub>2</sub> TPM.



**Figure S15. The correlation between mRNA expressions of *CXXC5* and *MeCP2* in tumor samples of various tissues.** Snapshots of dot-plot analyses with the use of the Pearson correlation coefficient indicate the correlation between expressions of *CXXC5* and *MeCP2* in TCGA tumor samples of various tissues as  $\log_2$  transcript per million (TPM) are shown.



**Figure S16. Interaction of EMD lacking LEM domain (EMD<sub>ΔLEM</sub>) with CXXC5.** (a) Schematics of EMD and EMD lacking LEM domain (HA-EMD<sub>ΔLEM</sub>). (b) To assess synthesis of HA-EMD and HA-EMD<sub>ΔLEM</sub>, HEK293 cells were transiently transfected for 48h followed by WB using the HA antibody. (c) HEK293 cells were also transiently co-transfected for 48h with an expression vector (pcDNA3.1) bearing the 3F-CXXC5 and the HA-EMD<sub>ΔLEM</sub> cDNAs. The synthesis of proteins (50 μg of lysate) was assessed by WB using the HA or the Flag antibody. (d) The nuclear extracts (500 μg) of HEK293 cells co-synthesizing 3x F-CXXC5 and HA-EMD<sub>ΔLEM</sub> or 3x F-CXXC5 and HA-EMD were subjected to Co-IP with the HA or the isotype-matched IgG. 10% of nuclear lysates was used as input control. The precipitates were subjected to SDS-10%PAGE followed with WB using analyzed using the Flag antibody. Molecular masses (MM) in kDa are indicated.

TABLE S1. CXXC5 Proximity Interactors, G AYAZ et al.

| No | Accession #   | String Id                   | Preferred Name  | Description   | # AAs      | MW [kDa]     | calc. pI    | Biological Replicate I  |             |                    |           |                         |              |                    |           |                         |             |                    |           |                         |             |                    |              | Biological Replicate II |          |          |           |  |  |  |  |  |  |  |  |  |  |  |  |
|----|---------------|-----------------------------|-----------------|---|------------|--------------|-------------|-------------------------|-------------|--------------------|-----------|-------------------------|--------------|--------------------|-----------|-------------------------|-------------|--------------------|-----------|-------------------------|-------------|--------------------|--------------|-------------------------|----------|----------|-----------|--|--|--|--|--|--|--|--|--|--|--|--|
|    |               |                             |                 |   |            |              |             | 1st Technical Replicate |             |                    |           | 2nd Technical Replicate |              |                    |           | 1st Technical Replicate |             |                    |           | 2nd Technical Replicate |             |                    |              |                         |          |          |           |  |  |  |  |  |  |  |  |  |  |  |  |
|    |               |                             |                 |   |            |              |             | Coverage                | Σ# Proteins | Σ# Unique Peptides | Σ# PSMs   | Coverage                | Σ# Proteins  | Σ# Unique Peptides | Σ# PSMs   | Coverage                | Σ# Proteins | Σ# Unique Peptides | Σ# PSMs   | Coverage                | Σ# Proteins | Σ# Unique Peptides | Σ# PSMs      |                         |          |          |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 1  | Q9H2P0        | 9606.ENSP00000379346        | ADNP            | Activity-dependent neuroprotector homeobox protein        | 1102       | 123.49       | 7.34        | 15.06                   | 2           | 14                 | 14        | 25                      | 18.87        | 2                  | 15        | 15                      | 26          | 6.81               | 2         | 5                       | 5           | 10                 | 6.81         | 2                       | 5        | 5        | 10        |  |  |  |  |  |  |  |  |  |  |  |  |
| 2  | P05549        | 9606.ENSP00000368924        | TFAP2A          | Transcription factor AP-2-alpha                           | 437        | 48.03        | 8.02        | 24.49                   | 14          | 5                  | 14        | 33                      | 26.54        | 14                 | 8         | 16                      | 38          | 11.44              | 13        | 3                       | 4           | 7                  | 11.44        | 15                      | 4        | 4        | 10        |  |  |  |  |  |  |  |  |  |  |  |  |
| 3  | Q8WWM7        | 9606.ENSP00000378917        | ATXN2L          | Ataxin-2-like protein                                     | 1075       | 113.30       | 8.02        | 24.49                   | 9           | 2                  | 2         | 4                       | 2.14         | 8                  | 1         | 2                       | 4           | 2.14               | 8         | 1                       | 2           | 4                  | 2.14         | 8                       | 1        | 2        | 4         |  |  |  |  |  |  |  |  |  |  |  |  |
| 4  | Q9NRL2        | 9606.ENSP0000035458         | BAZ1A           | Bromodomain adjacent to zinc finger domain protein 1A     | 1556       | 178.59       | 6.60        | 1.99                    | 2           | 2                  | 2         | 5                       | 4.05         | 2                  | 3         | 8                       | 5.14        | 3                  | 6         | 12                      | 2.76        | 3                  | 4            | 4                       | 8        | 8        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 5  | Q9UIG0        | 9606.ENSP00000342434        | BAZ1B           | Tyrosine-protein kinase BAZ1B                             | 1483       | 170.80       | 8.48        | 18.81                   | 2           | 25                 | 25        | 49                      | 19.82        | 2                  | 26        | 26                      | 53          | 18.88              | 2         | 20                      | 42          | 16.93              | 2            | 17                      | 17       | 36       |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 6  | Q9NYF8        | 9606.ENSP00000435210        | BCLAF1          | Bcl-2-associated transcription factor 1                   | 920        | 106.06       | 9.98        | 8.15                    | 10          | 6                  | 6         | 10                      | 4.57         | 9                  | 3         | 4                       | 5.33        | 10                 | 3         | 3                       | 8           | 5.22               | 9            | 3                       | 3        | 8        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 7  | Q12830        | 9606.ENSP00000307208        | BPTF            | Nucleosome-remodeling factor subunit BPTF                 | 3046       | 338.05       | 6.54        | 3.05                    | 6           | 8                  | 13        | 3                       | 3.61         | 6                  | 9         | 17                      | 1.74        | 5                  | 4         | 8                       | 2.43        | 6                  | 5            | 5                       | 10       | 8        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 8  | Q8N163        | 9606.ENSP00000310670        | CCAR2           | Cell cycle and apoptosis regulator protein 2              | 923        | 102.84       | 5.22        | 1.52                    | 3           | 1                  | 1         | 2                       | 1.52         | 3                  | 1         | 2                       | 1.52        | 3                  | 1         | 2                       | 1.52        | 3                  | 1            | 1                       | 2        | 3        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 9  | Q14839        | 9606.ENSP00000349508        | CHD4            | Chromodomain-helicase-DNA-binding protein 4               | 1912       | 217.87       | 5.86        | 9.36                    | 3           | 11                 | 14        | 28                      | 12.45        | 3                  | 15        | 19                      | 39          | 10.41              | 6         | 12                      | 16          | 33                 | 10.41        | 4                       | 13       | 17       | 35        |  |  |  |  |  |  |  |  |  |  |  |  |
| 10 | Q9HCK8        | 9606.ENSP00000382863        | CHD8            | Chromodomain-helicase-DNA-binding protein 8               | 2581       | 290.34       | 6.47        | 7.25                    | 7           | 12                 | 13        | 24                      | 8.21         | 6                  | 12        | 14                      | 28          | 5.85               | 4         | 9                       | 18          | 4.49               | 4            | 7                       | 7        | 14       |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 11 | Q9BR76        | 9606.ENSP00000377471        | CORO1B          | Coronin-1B  | 489        | 54.20        | 5.88        | 15.13                   | 4           | 5                  | 5         | 10                      | 10.43        | 4                  | 3         | 6                       | 5.2         | 4                  | 3         | 6                       | 5.2         | 4                  | 2            | 2                       | 6        | 6        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 12 | Q16630        | 9606.ENSP00000266679        | CPSF6           | Cleavage and polyadenylation specificity factor subunit 6 | 551        | 59.17        | 7.15        | 2.72                    | 4           | 1                  | 1         | 2                       | 9.44         | 4                  | 3         | 6                       | 2.54        | 4                  | 1         | 2                       | 7.08        | 4                  | 2            | 2                       | 2        | 4        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 13 | P49711        | 9606.ENSP00000264010        | CTCF            | Transcriptional repressor CTCF                            | 727        | 82.73        | 6.96        | 8.39                    | 2           | 4                  | 4         | 9                       | 10.73        | 2                  | 5         | 9                       | 8.1         | 4                  | 4         | 9                       | 11.2        | 5                  | 5            | 5                       | 9        | 9        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 14 | P39880        | 9606.ENSP00000353401        | CUX1            | Homeobox protein cut-like 1                               | 1505       | 164.09       | 5.90        | 16.68                   | 11          | 18                 | 18        | 33                      | 14.68        | 11                 | 16        | 16                      | 32          | 5.05               | 11        | 6                       | 12          | 4.98               | 11           | 5                       | 5        | 10       |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 15 | <b>Q7FLH8</b> | <b>9606.ENSP00000302543</b> | <b>CXXC5</b>    | <b>CXXC5-type zinc finger protein 5</b>                   | <b>322</b> | <b>32.96</b> | <b>9.10</b> | <b>45.03</b>            | <b>12</b>   | <b>22</b>          | <b>22</b> | <b>108</b>              | <b>45.03</b> | <b>12</b>          | <b>20</b> | <b>20</b>               | <b>116</b>  | <b>31.37</b>       | <b>12</b> | <b>8</b>                | <b>8</b>    | <b>50</b>          | <b>31.37</b> | <b>12</b>               | <b>8</b> | <b>8</b> | <b>59</b> |  |  |  |  |  |  |  |  |  |  |  |  |
| 16 | Q9NR30        | 9606.ENSP00000346120        | DDX21           | Nucleolar RNA helicase 2                                  | 783        | 87.29        | 9.28        | 5.75                    | 2           | 4                  | 4         | 6                       | 5.49         | 2                  | 3         | 5                       | 2.81        | 2                  | 1         | 2                       | 2.81        | 2                  | 1            | 1                       | 2        | 2        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 17 | P17844        | 9606.ENSP00000225792        | DDX5            | Probable ATP-dependent RNA helicase DDX5                  | 614        | 69.10        | 8.92        | 16.29                   | 14          | 9                  | 9         | 17                      | 8.92         | 14                 | 11        | 11                      | 20          | 8.63               | 9         | 4                       | 4           | 8                  | 8.63         | 9                       | 4        | 4        | 8         |  |  |  |  |  |  |  |  |  |  |  |  |
| 18 | Q08211        | 9606.ENSP00000356520        | DHX9            | ATP-dependent RNA helicase A                              | 1270       | 140.87       | 6.84        | 5.75                    | 2           | 6                  | 6         | 10                      | 4.41         | 1                  | 6         | 9                       | 5.51        | 2                  | 5         | 10                      | 4.09        | 2                  | 4            | 4                       | 8        |          |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 19 | Q9BTC0        | 9606.ENSP00000266070        | DIDO1           | Death-inducible oligomer 1                                | 2240       | 243.72       | 7.88        | 1.38                    | 2           | 2                  | 2         | 4                       | 3.75         | 2                  | 4         | 8                       | 3.93        | 2                  | 4         | 7                       | 13          | 3.93               | 2            | 6                       | 6        | 14       |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 20 | Q0832         | 9606.ENSP00000358563        | DKC1            | H/ACA ribonucleoprotein complex subunit 4                 | 514        | 57.64        | 9.42        | 10.31                   | 6           | 3                  | 3         | 6                       | 6.61         | 6                  | 3         | 4                       | 2.53        | 5                  | 1         | 2                       | 2.53        | 5                  | 1            | 1                       | 2        | 2        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 21 | P32519        | 9606.ENSP00000239882        | ELF1            | ETS-related transcription factor E1F-1                    | 619        | 67.46        | 5.21        | 17.61                   | 11          | 7                  | 7         | 15                      | 9.69         | 6                  | 3         | 4                       | 8           | 5.82               | 2         | 2                       | 4           | 2.91               | 2            | 1                       | 1        | 2        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 22 | Q6PJG2        | 9606.ENSP00000286523        | ELMSAN1         | ELM2 and SANT domain-containing protein 1                 | 1045       | 114.92       | 9.19        | 5.17                    | 2           | 3                  | 3         | 6                       | 6.6          | 3                  | 4         | 8                       | 1.82        | 1                  | 2         | 2                       | 2           | 2                  | 4            | 2.87                    | 1        | 3        | 3         |  |  |  |  |  |  |  |  |  |  |  |  |
| 23 | Q8WYF5        | 9606.ENSP00000355465        | ELYS            | Protein ELYS  | 2266       | 252.34       | 6.60        | 3.18                    | 3           | 4                  | 4         | 7                       | 2.43         | 3                  | 3         | 6                       | 1.37        | 3                  | 2         | 3                       | 0.71        | 3                  | 1            | 1                       | 2        | 2        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 24 | P50402        | 9606.ENSP00000358857        | EMD             | Emerin  | 254        | 28.98        | 5.50        | 10.24                   | 2           | 2                  | 2         | 3                       | 10.24        | 2                  | 2         | 2                       | 3           | 12.24              | 3         | 2                       | 2           | 3                  | 14.1         | 4                       | 2        | 2        | 4         |  |  |  |  |  |  |  |  |  |  |  |  |
| 25 | Q7Z589        | 9606.ENSP00000433205        | C11orf30 (EMSY) | Protein EMSY  | 1322       | 141.38       | 9.33        | 6.43                    | 12          | 5                  | 5         | 11                      | 14.07        | 13                 | 11        | 11                      | 17          | 4.54               | 9         | 3                       | 3           | 5                  | 2.04         | 8                       | 2        | 2        | 4         |  |  |  |  |  |  |  |  |  |  |  |  |
| 26 | P22087        | 9606.ENSP00000221801        | FBL             | rRNA 2-methyltransferase fibrillarin                      | 321        | 33.76        | 10.18       | 18.38                   | 10          | 4                  | 4         | 8                       | 24.92        | 10                 | 6         | 6                       | 12          | 22.43              | 10        | 5                       | 10          | 12.77              | 10           | 3                       | 3        | 6        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 27 | Q6ISB3        | 9606.ENSP00000251808        | GRHL2           | Grainyhead-like protein 2 homolog                         | 625        | 71.06        | 6.44        | 19.04                   | 2           | 7                  | 7         | 16                      | 23.84        | 2                  | 9         | 9                       | 22          | 15.04              | 2         | 6                       | 11          | 17.6               | 2            | 7                       | 7        | 12       |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 28 | P78347        | 9606.ENSP00000460070        | GTIF2           | General transcription factor II-I                         | 998        | 112.35       | 6.39        | 2.51                    | 4           | 2                  | 2         | 4                       | 1.5          | 4                  | 1         | 1                       | 2           | 1.6                | 4         | 1                       | 3           | 2.91               | 5            | 2                       | 2        | 6        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 29 | Q08014        | 9606.ENSP00000349430        | HIST1H2BK       | Histone H2B type 1-K                                      | 126        | 13.88        | 10.32       | 69.84                   | 2           | 8                  | 13        | 5                       | 74.6         | 3                  | 1         | 16                      | 53          | 19.05              | 17        | 2                       | 14          | 19.05              | 17           | 2                       | 2        | 12       |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 30 | P51610        | 9606.ENSP00000309555        | HCFC1           | Host cell factor 1  | 2035       | 208.60       | 7.46        | 7.27                    | 5           | 12                 | 12        | 25                      | 7.52         | 5                  | 12        | 12                      | 22          | 5.45               | 5         | 8                       | 18          | 5.45               | 5            | 10                      | 10       | 20       |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 31 | Q9UGU5        | 9606.ENSP00000216106        | HMGXB4          | HMG domain-containing protein 4                           | 601        | 65.67        | 9.32        | 6.16                    | 3           | 2                  | 2         | 3                       | 6.16         | 3                  | 3         | 5                       | 2.83        | 3                  | 1         | 2                       | 2.83        | 3                  | 1            | 1                       | 2        | 2        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 32 | Q43390        | 9606.ENSP00000363745        | HNRNPR          | Heterogeneous nuclear ribonucleoprotein R                 | 633        | 70.90        | 8.13        | 7.9                     | 11          | 3                  | 3         | 6                       | 9.64         | 6                  | 4         | 5                       | 10          | 7.74               | 5         | 2                       | 5           | 9                  | 4.27         | 5                       | 1        | 3        | 6         |  |  |  |  |  |  |  |  |  |  |  |  |
| 33 | P08107        | 9606.ENSP00000364802        | HSP71 (HSPA1A)  | Heat shock 70 kDa protein 1A/1B                           | 641        | 70.01        | 5.66        | 8.74                    | 9           | 3                  | 3         | 4                       | 9.2          | 10                 | 3         | 4                       | 6           | 2.5                | 3         | 1                       | 4           | 6.55               | 4            | 3                       | 3        | 5        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 34 | P11142        | 9606.ENSP00000432083        | HSPA8           | Heat shock cognate 71 kDa protein                         | 646        | 70.85        | 5.52        | 8.82                    | 16          | 4                  | 5         | 10                      | 4.49         | 15                 | 2         | 3                       | 6           | 4.49               | 12        | 2                       | 4           | 4.18               | 13           | 2                       | 2        | 4        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 35 | P55010        | 9606.ENSP00000216554        | IF5 (EIF5)      | Eukaryotic translation initiation factor 5                | 431        | 49.19        | 5.58        | 12.06                   | 3           | 5                  | 5         | 8                       | 12.53        | 1                  | 6         | 6                       | 10          | 16.01              | 7         | 6                       | 11          | 6.96               | 7            | 3                       | 3        | 6        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 36 | P08727        | 9606.ENSP00000355124        | KRT19           | Keratin, type I cytoskeletal 19                           | 400        | 44.08        | 5.14        | 11.15                   | 15          | 2                  | 5         | 10                      | 19.75        | 17                 | 4         | 7                       | 13          | 25.75              | 9         | 3                       | 12          | 28                 | 21.5         | 8                       | 2        | 12       | 24        |  |  |  |  |  |  |  |  |  |  |  |  |
| 37 | Q9Y2K7        | 9606.ENSP00000432786        | KDM2A           | Lysine-specific demethylase 2A                            | 1162       | 132.71       | 7.58        | 1.46                    | 5           | 1                  | 2         | 2                       | 2.5          | 5                  | 2         | 2                       | 4           | 1.46               | 5         | 1                       | 2           | 1.46               | 5            | 1                       | 1        | 2        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 38 | P46013        | 9606.ENSP00000357643        | KI67 (MKI67)    | Antigen Ki-67   | 3256       | 358.47       | 9.45        | 13.85                   | 3           | 29                 | 29        | 55                      | 17.57        | 3                  | 35        | 35                      | 66          | 11.92              | 3         | 24                      | 50          | 12.07              | 3            | 26                      | 26       | 50       |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 39 | Q03164        | 9606.ENSP00000436786        | KMT2A           | Histone-lysine N-methyltransferase 2A                     | 3969       | 431.50       | 9.09        | 3.2                     | 4           | 9                  | 9         | 17                      | 3.63         | 4                  | 11        | 11                      | 21          | 6.35               | 5         | 16                      | 30          | 5.29               | 5            | 14                      | 14       | 27       |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 40 | Q9JMN6        | 9606.ENSP00000398837        | KMT2B           | Histone-lysine N-methyltransferase 2B                     | 2715       | 293.33       | 8.22        | 0.88                    | 2           | 2                  | 2         | 4                       | 0.99         | 2                  | 2         | 2                       | 4           | 1.18               | 1         | 2                       | 2           | 1.18               | 1            | 2                       | 2        | 4        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 41 | P42166        | 9606.ENSP00000266732        | LAP2A (TMPO)    | Lamina-associated polypeptide 2, isoform alpha            | 694        | 75.45        | 7.66        | 43.37                   | 2           | 18                 | 25        | 46                      | 41.07        | 2                  | 19        | 26                      | 53          | 29.68              | 2         | 15                      | 29          | 19.18              | 1            | 6                       | 11       | 21       |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 42 | P42167        | 9606.ENSP00000266732        | LAP2B (TMPO)    | Lamina-associated polypeptide 2, isoform beta/gamma       | 454        | 50.64        | 9.38        | 37.22                   | 5           | 9                  | 16        | 35                      | 46.92        | 5                  | 10        | 17                      | 34          | 26.65              | 5         | 3                       | 8           | 15                 | 23.79        | 5                       | 2        | 7        | 15        |  |  |  |  |  |  |  |  |  |  |  |  |
| 43 | Q6PKG0        | 9606.ENSP0000036721         | LARP1           | La-related protein 1                                      | 1096       | 123.43       | 8.82        | 2.19                    | 4           | 1                  | 1         | 2                       | 1.19         | 4                  | 1         | 2                       | 2.65        | 5                  | 2         | 4                       | 2           | 2.65               | 5            | 2                       | 2        | 3        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 44 | Q6MZP7        | 9606.ENSP00000341947        | LINS4           | Protein lin-54 homolog                                    | 749        | 79.44        | 9.01        | 8.54                    | 5           | 4                  | 4         | 8                       | 6.81         | 4                  | 3         | 3                       | 6           | 4.94               | 4         | 3                       | 6           | 4.94               | 4            | 3                       | 3        | 6        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 45 | Q6AG44        | 9606.ENSP00000225972        | LRRCS9          | Leucine-rich repeat-containing protein 59                 | 307        | 34.91        | 9.57        | 18.24                   | 1           | 4                  | 4         | 7                       | 13.36        | 1                  | 4         | 4                       | 7           | 5.21               | 1         | 1                       | 2           | 8.14               | 1            | 2                       | 2        | 4        |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 46 | Q5T3J3        | 9606.ENSP00000358778        | LRIF1           | Ligand-dependent nuclear receptor-interacting factor 1    | 769        | 84.52        | 9.72        | 13.26                   | 1           | 6                  | 6         | 10                      | 12.74        | 1                  | 5         | 7                       | 7.8         | 1                  | 4         | 8                       | 7.67        | 1                  | 4            | 4                       | 8        |          |           |  |  |  |  |  |  |  |  |  |  |  |  |
| 47 | Q8UJE4        | 9606.ENSP00000338235        | MTDH            | Pro   |            |              |             |                         |             |                    |           |                         |              |                    |           |                         |             |                    |           |                         |             |                    |              |                         |          |          |           |  |  |  |  |  |  |  |  |  |  |  |  |

**TABLE S2. Primers, G AYAZ et al.**

**Cloning Primers**

CXXC5  
EMD  
EMD ΔLEMD  
FL-MAZ  
MAZΔN  
MeCP2  
MeCP21-400

CGCATATACTCGAGATTACCATGGAAGCTAGCATGTCGAGCCTCGGCGGTGGCTC  
CTGATCTCGAGCCATGGCTAGCGACAACACGAGATCTTTCC  
TTAGTGTAGCAGCTCGTCCGCCCTCCTCTTATAG  
CTGATGCTAGCTTCCCGGTGTTTCCCTGACGCTGCTG  
CGCATCTCGAGACGCGTAACCATGGCTAGCGTGCCTGAGCCTCCTGAGC  
CTGATGATAGCTAGCTGGGATGTTAGGGCTC  
CTGATGATAGCTAGCTGGGATGTTAGGGCTC

**Tag sequences (DNA and aminoacid)**

Flag  
3xFlag  
HA  
NLS (Nuclear Localization Signal)

GATTACAAGGATGACGACGATAAG  
GACTACAAGACCATGACGGTATTATAAAGATCATGACATCGACTACAAGACGATGACGACAAG  
TACCCATACGATGTTCCAGATTACGCT  
CCCAAGAAAAAGAGGAAGGTGGGCTACCCCAAGAAAAAGAGGAAGGTGGCGGCTCA

**CXXC5 Truncation and Deletion Primers**

3F-CXXC5  
3F-CXXC5-Δ101-149  
3F-CXXC-Δ150-199  
3F-CXXC-Δ200-249  
3F-CXXC5-Δ1-100  
3F-CXXC5-Δ250-322  
3F-CXXC5-Δ1-100 & 250-322  
3F-CXXC  
3F-CXXC-DBM  
3F-CXXC5-DBM

CGCATGCTAGCACCGGTTTCGAGCCTCGGCGGTGGCTCC  
GCAGCATGATGGCGGAGAGCGGGCCACGAGCTGGCAGC  
GCCTGCTGAGCAAGGCAGAGGAAGCCCTCAATGGCCAGTC  
TGGAGGCTGTGGCAGGTGCGCCCTCTGCCATCAGCTCCGG  
CGCATGCTAGCACCGGTTCTGCTGACAAGGCCACTGCG  
CGCATGCTAGCACCGGTTTCGAGCCTCGGCGGTGGCTCC  
CGCATGCTAGCACCGGTTCTGCTGACAAGGCCACTGCG  
CGCATGCTAGCACCGGTTGCTCTGCCATCAGCTCCGGC  
CGCATGCTAGCACCGGTTGCTCTGCCATCAGCTCCGGC  
CGCATGCTAGCACCGGTTTCGAGCCTCGGCGGTGGCTCC

**Generation of DBM (AAAA-188-191)**

1st round  
2nd round  
Overlapping PCR (using template 1st & 2nd rounds)

CGCATATACTCGAGATTACCATGGAAGCTAGCATGTCGAGCCTCGGCGGTGGCTC  
GAGCGAGGCGGGTGTCTGCCGACGGACATGG  
CGCATATACTCGAGATTACCATGGAAGCTAGCATGTCGAGCCTCGGCGGTGGCTC

**RT-qPCR Primers**

CXXC5  
HDAC11  
NFKBIZ  
IL12A

CACCAGGCATCTCTGTTGTGG  
CAGCTGTGGCCCTGATCCAAATC  
ACTCAGCTGCAAAGATAAAGTTAC  
TGCAAAGCTTCTGATGGATCC

**ChIP Primers**

HDAC11  
NFKBIZ  
IL12A  
Exon10 of CXXC5

CTGTGGGGCATTATTACG  
ACGTACGCACCTTAGCCATC  
CGCTTTCATTTTGGCCGAG  
TCAACCCAGGCCTTTCATT

**Bisulfite Primers**

HDAC11  
NFKBIZ  
IL12A

GTTAGTTAAGGTTATATAGTGAAGTTT  
GGGGTTTAGAGGTTAGGGATGGTT  
TATGTAAGTGGGAGGATTTTTTTTTTTTA

**Reverse Primer Sequence (5' to 3')**

CGCATGGGATCCTTTATTAGAATTCCTGAAACCACCGGAAGGCGG  
CTGATGGGATCCTTTATTAGTCGACGAAGGGGTTGCCTTCTCAGCCTGC  
CTGATGGGATCCTTTATTAGTCGACGAAGGGGTTGCCTTCTCAGCCTGC  
CTGATGGATCCTTTATTACCAGGTTGGGAGGGAAGTGGCTGAGAGCT  
CGCATGGGATCCTCTAGATTTATTAGAATTCGACAGGTGGGCTGTGGCTGGGG  
CGCATGGGATCCTTTATTAGAATTCGCTAACTCTCTCGGTACAGGG  
CGCATGGGATCCTTTATTAGAATTCGGTGGGGTCTCGGAGCTCTCGGG

**Amino Acid Sequence**

DYKDDDDK  
DYKDHGDYKDHIDYKDDDDK  
YPYDVFDYA  
PKKKRKYVSPKKRKYVSS

**Reverse Primer Sequence (5' to 3')**

CGCATGGGATCCTTTATTAGAATTCCTGAAACCACCGGAAGGCGG  
GCTGCCAGTCCGTGGCCGCTCTCCGCCATCATGCTGC  
GACTGGCATTGAGGGCTTCTCTGCTTCTCAGCAGGC  
CCGGAGCTGATGGCAGAGGGCCACCTGCCACAGCTCCA  
CGCATGGGATCCTTTATTAGAATTCCTGAAACCACCGGAAGGCGG  
CGCATGGGATCCTTTATTAGAATTCAGCTCTCCCTGCATGGGGTAC  
CGCATGGGATCCTTTATTAGAATTCAGCTCTCCCTGCATGGG  
CGCATGGGATCCTTTATTAGAATTCCTGAAACCACCGGAAGGCGG  
CGCATGGGATCCTTTATTAGAATTCCTGAAACCACCGGAAGGCGG  
CGCATGGGATCCTTTATTAGAATTCCTGAAACCACCGGAAGGCGG

CCATGTCGGCTGCGGCAGCACCCGCCTCGCTC  
CGCATGGGATCCTTTATTAGAATTCCTGAAACCACCGGAAGGCGG  
CGCATGGGATCCTTTATTAGAATTCCTGAAACCACCGGAAGGCGG

**Reverse Primer Sequence (5' to 3')**

TTGTCTGCTGCTCCTGCCTTT  
ATCTTGAACCTTCTCCCTGC  
ACATTATTTTCTTGGCGCTGG  
AAAATCCGGTTCTTCAAGGGA

**Reverse Primer Sequence (5' to 3')**

TGCTTAACTGGCTGCCTTT  
AGAGAGCGAGCGATCTCCTG  
ACATCAGCTTCTCGGTGACA  
TGGCCAGTCTTTCGATTCTC

**Reverse Primer Sequence (5' to 3')**

ACCAACAAATCTAAACCTTACCTAA  
CCATAAAAATACTCAAATTAAC  
CTACTACTATTCCAATCTCTCCTTAAACAA