Molecular Cell, Volume 42

Supplemental Information

L3MBTL2 Protein Acts in Concert with PcG Protein-Mediated Monoubiquitination of H2A to Establish a Repressive Chromatin Structure

Patrick Trojer, Alina R. Cao, Zhonghua Gao, Yan Li, Jin Zhang, Xiaoqin Xu, Guohong Li, Regine Losson, Hediye Erdjument-Bromage, Paul Tempst, Peggy J. Farnham, and Danny Reinberg

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Cell lines, transfections

Full-length L3MBTL2 cDNA was inserted into pCMV-Tag4A (Stratagene) and this constructs or empty vector transfected in 293F cells using the FuGENE HD (Roche) transfection reagent. Selection with G418 for two weeks led to the identification of single clones that stably expressed an L3MBTL2-FLAG fusion protein. One clone was expanded and adapted to growth in suspension culture by changing the culture medium to a fully synthetic FreeStyle™ 293 expression medium (Invitrogen).

Biochemical purification of L3MBTL2-F and associated polypeptides

Extracts were subjected to a 15 ml DE52 (Whatman) cation exchange resin using gravity flow. Loading and washes were carried out at 100 mM KCl, elution of bound proteins was performed at 350 mM KCl in buffer B (10 mM Tris, pH 7.9, 0.1 mM EDTA, 10% glycerol, 5 mM 2-mercaptoethanol). The elution fraction was incubated in batch with 500 µl of FLAG(M2) agarose (Sigma) for 4 h at 4°C under constant rotation. The resin was washed with 20 column

volumes (cv) of buffer B containing 350 mM KCl and 0.05% NP-40. Bound proteins were eluted with 250 μg/ml FLAG peptide (Sigma) in buffer B containing 100 mM KCl. Eluted proteins were loaded by gravity flow onto an anti-HP1γ immuno-affinity column (250 μl protein G agarose (Roche) and 2 ml of anti-HP1γ ascites was cross-linked with dimethyl pimelimidate (Pierce) according to the manufacturer's instructions). The resin was washed with 2 cv of buffer B containing 500 mM KCl and 0.05% NP-40 and 10 cv of buffer B containing 200 mM KCl and 0.05% NP-40. Elution was carried out using 0.1 M glycine, pH2.7 and eluted proteins were dialyzed into buffer B containing 100 mM KCl. Affinity-purified L3MBTL2-F fractions were resolved by SDS-PAGE and analyzed by silver staining, western blotting and tandem mass spectrometry as described previously (Trojer et al., 2007). Briefly, gel-resolved proteins were digested with trypsin (Sebastiaan Winkler et al., 2002), the mixtures fractionated on a Poros 50 R2 RP micro-tip, and resulting peptide pools analyzed by matrix-assisted laser-desorption/ionization reflectron time-of-flight (MALDI-reTOF) MS using a BRUKER UltraFlex TOF/TOF instrument (Bruker; Bremen, Germany), as described (Erdjument-Bromage et al., 1998). Selected experimental masses (m/z) were taken to search the human segment of a non-redundant protein database ('NR'; ~223,695 entries; National Center for Biotechnology Information; Bethesda, MD), utilizing the Mascot Peptide Mass Fingerprint (PMF) program, version 2.2.04 for Windows (www.matrixscience.com), with a mass accuracy restriction better than 35 ppm, and maximum two missed cleavage site allowed per peptide. To confirm observed peptides with calculated mono-isotopic fragments matching to the experimental values , mass spectrometric sequencing of selected peptides were done by MALDI-TOF/TOF (MS/MS) analysis on the same prepared samples, using the UltraFlex instrument in 'LIFT' mode. Fragment ion spectra were taken to search NR using the MASCOT MS/MS Ion Search program

(Matrix Science). Any tentative result thus obtained was verified by comparing the computer-generated fragment ion series of the predicted tryptic peptide with the experimental MS/MS data.

Antibodies, immunoprecipitations (IPs)

For the production of L3MBTL2 and MBLR antibodies we expressed L3MBTL2 fragments (the L3MBTL2 N-terminal fragment comprised residues 1-178 and L3MBTL2 C-terminal fragment residues 606-705) or full-length mouse MBLR in E. Coli as hexahistidine (HIS) tagged fusion proteins. Recombinant proteins were purified using NiNTA Agarose (Qiagen), dialyzed into phosphate buffered saline (PBS) and injected subcutaneously into rabbits as a mixture with Freund's Adjuvant. Boosts were performed and rabbit sera obtained 2 weeks afterwards. The IgG fraction was purified from the serum proteins using Protein A Sepharose (Roche) according to standard protocols. Purified IgGs were used for all experiments shown in this study.

IPs were performed using 1mg of HeLa nuclear extract and 5 µg of antibodies in a final volume of 300 µl buffer B. After incubation for 4 h at 4°C under constant rotation 25 µl of protein A beads were added and the mixture incubated for one more hour. Beads were precipitated by centrifugation (2 min at 500 x g) and washed three times with 400 µl of IP buffer (20 mM Tris, pH7.9, 200 mM NaCl, 0.1 mM EDTA, 0.05% NP-40, 5 mM 2-mercaptoethanol). Washed beads were mixed with Laemmli sample buffer and analyzed by SDS-PAGE with subsequent western blotting.

GST and peptide pulldown experiments:

GST fusion proteins or biotinylated histone peptides (2-5 µg) were incubated with proteins of interest (5 µg) in 300 µl of IP buffer (20 mM Tris, pH7.9, 200 mM NaCl, 0.1 mM EDTA, 0.05%

NP-40, 5 mM 2-mercaptoethanol) either for 4 h or overnight at 4°C on a rotating wheel. Equilibrated glutathione sepharose 4FF beads (GE Healthcare) or streptavidin agarose beads (Millipore) were added and the reactions further incubated for 1 h. Beads were precipitated by centrifugation (30 sec at 500 x g) and washed three times with 400 µl of IP buffer Washed beads were mixed with Laemmli sample buffer and analyzed by SDS-PAGE with subsequent western blotting. Histone octamers were pulled down as described above except that HBS-P buffer (Biacore) was used for incubation and washes.

Lentivirus shRNAs, transfection siRNAs:

Five different L3MBTL2 short hairpin (sh) RNA constructs were obtained from the MISSION™ TRC-Hs 1.0 (Human) shRNA library (Sigma) and used as described previously (Zufferey et al., 1998). For the production of Lentivirus particles that harbour L3MBTL2 specific shRNA sequences we used the Virapower™ Lentiviral packaging mix (Invitrogen) according to the manufacturer's instructions. The packaging mix was transfected together with shRNA constructs into 293T cells using Lipofectamine 2000 (Invitrogen), and the supernatant obtained 72 h post transfection (containing a high titer of virus particles) was subsequently used for Lentivirus transduction experiments. Selection of shRNA containing 293F cells was initiated 48 h post transduction with 1.5 µg/ml puromycin.

ChIP-chip

ChIP-chip assays (1 x 10^7 cells/assay) were performed following the protocol provided at http://genomics.ucdavis.edu/farnham and http://genomecenter.ucdavis.edu/expression_analysis. For tthe primary antibodies used in this study see Supplementary Table 1. For PCR analysis of the ChIP samples prior to amplicon generation, QIAquick-purified immunoprecipitates were

dissolved in 50 µl of water. Standard PCR reactions using 2 µl of the immunoprecipitated DNA were performed, and primer sequences are available upon request. Amplicons were prepared by adapting the standard protocol for whole genome amplification using the Sigma GenomePlex WGA kit as described in O'Geen et al. (O'Geen et al., 2006). Briefly, the initial random fragmentation step was eliminated and DNA from the entire ChIP sample was used for amplification. A detailed protocol for the WGA method is provided at http://genomics.ucdavis.edu/farnham and http://genomecenter.ucdavis.edu/expression_analysis.

DNA Microarrays: Amplicons were applied either to ENCODE arrays or to 1.5 kb promoter arrays (see Supplementary Table S2 and www.nimblegen.com for details). For ChIP-chip analysis, the labeling and the hybridization of DNA samples were performed by NimbleGen Systems, Inc. Briefly, each DNA sample (1 µg) was denatured in the presence of 5'-Cy3- or Cy5-labeled random nonamers (TriLink Biotechnologies) and incubated with 100 units (exo-) Klenow fragment (NEB) and dNTP mix (6 mM each in TE buffer (10 mM Tris/1 mM EDTA, pH 7.4; Invitrogen)) for 2 h at 37°C. Reactions were terminated by addition of 0.5 M EDTA (pH 8.0), precipitated with isopropanol, and resuspended in water. Then, 13 µg of the Cy5-labeled ChIP sample and 13 µg of the Cy3-labeled total sample were mixed together, dried down, and resuspended in 40 µl of NimbleGen Hybridization Buffer (NimbleGen Systems) plus 1.5 ug of human COT1 DNA. After denaturation, hybridization was carried out in a MAUI Hybridization System (BioMicro Systems) for 18 h at 42°C. The arrays were washed using NimbleGen Wash Buffer System (NimbleGen Systems), dried by centrifugation, and scanned at 5-µm resolution using the GenePix 4000B scanner (Axon Instruments). Fluorescence intensity raw data were obtained from scanned images of the arrays using NIMBLESCAN 2.0 extraction software (NimbleGen Systems). For each spot on the array, log2-ratios of the Cy5-labeled test sample

versus the Cy3-labeled reference sample were calculated. Then, the biweight mean of this log2 ratio was subtracted from each point; this procedure is approximately equivalent to mean-normalization of each channel.

Data analysis: Binding sites on the ENCODE arrays were identified using the highest stringency level, L1 (six consecutive probes above the 98th percentile threshold, $P<0.0001$, with S=50 meaning that a 100 nt gap was allowed between peaks), of the Tamalpais peak calling algorithm previously described (Bieda et al., 2006); see also http://genomics.ucdavis.edu/farnham. For 1.5kb promoter array analysis, the overall enrichment value for each promoter was calculated by NimbleGen software based on the median value of the top 11 of the 15 probes for each promoter region. Functional annotations were performed using the program Database for Annotation, Visualization, and Integrated Discovery (DAVID) 2.1 (http://david.abcc.ncifcrf.gov), as previously described (Squazzo et al., 2006).

ChIP sequencing

For ChIP sequencing (ChIP-seq), L3MBTL2 ChIP samples were prepared from K562 cells as follows: cultures of 1×10^8 K562 cells were harvested at a density of 10^6 cells/ml and cross-linked with 1% formaldehyde for 10 min at room temperature. Cross-linking was stopped by the addition of glycine to 125 mM final concentration and cells were washed twice with 1 × PBS. The cell pellet was then resuspended in 2 ml ChIP lysis buffer (50 mM Tris-Cl pH 8.0, 5 mM EDTA, 1% SDS, 1 complete protease inhibitor tablet (Roche)) and incubated on ice for 30 min. Samples were sonicated for 30 min, with 30 sec pulses, 1 min resting, using the Bioruptor sonicator (Diagenode) to produce chromatin fragments of 0.5 kb on average. After clarification by centrifugation, sonicated extracts were precleared

with 20 µl/10^7 StaphA cells blocked with 10 mg/ml BSA. The precleared extracts were diluted 1:10 with ChIP dilution buffer (1% triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris-Cl pH 8.0, 5 mM PMSF). Anti-L3MBTL2 antibody was added at a concentration of 2 µg/ 10^7 cells and incubated for 12 h at 4°C. Complexes were recovered with StaphA cells for 15 min at room temperature and washed 5 times with ice cold RIPA buffer. Precipitates were resuspended in 100 µl ChIP elution buffer (1% SDS, 0.1 M $NaHCO_3$), incubated at 65°C for 12 h and treated with 10 µg RNaseA for 20 min at 37°C. After pooling, the DNA was recovered from the eluate using the QIAquick PCR Purification kit (QIAGEN) according to the manufacturer's instructions. To analyze the modified histones, 5×10^6 K562 cells were used per sample and protein G magnetic beads were used instead of Staph A cells to collect the immunoprecipitates; specific antibodies (see Table S1) were added to the ChIP reactions. 30 µl of protein G magnetic beads (CST cat#9006) were added to each ChIP reaction and incubated at 4°C on a rotating platform for 2 h and washed with cold RIPA buffer and IP wash buffer (100 mM Tris-Cl pH 9, 500 mM LiCl, 1% Igepal, 1% deoxycholic acid) for a total of five washes. Precipitates were purified in the same manner as L3MBTL2 ChIP samples.

ChIP libraries were created as described previously (Robertson et al., 2007), using 15 cycles of amplification. Libraries were run on a 2% agarose gel and the 200-400 and 400-600 fractions of the library was extracted and purified. To estimate the yield of library and its relative amplification value, library DNA was quantitated using a Nanodrop and serial dilutions of 1.25 nM library were compared to a reference library by real-time PCR using primers complementary to the library adapters. Before sequencing, libraries were analyzed using positive and negative control regions by quantitative real-time PCR (qPCR) was performed on a Bio-Rad DNA Engine Opticon Real-Time PCR System using SYBR® Green Master PCR Mix according to the manufacturer's instructions (Invitrogen). The fold enrichment of each target site was calculated

as 2 to the power of the cycle threshold (cT) difference between input chromatin and ChIP samples. In general, we have found that the 400-600 size libraries contain higher enrichments for heterochromatic regions and for factors that bind to H3K9me3. The amplification value relative to the reference library was used to estimate the flowcell loading concentration. The ChIP-seq libraries were run on an Illumina GA2 by the DNA Technologies Core Faciity at the University of California-Davis (http://genomecenter.ucdavis.edu/dna_technologies/). Peaks were called from the ChIP-seq data using Sole-search (Blahnik et al., 2010).

SUPPLEMENTAL REFERENCES

Bieda, M., Xu, X., Singer, M.A., Green, R., and Farnham, P.J. (2006). Unbiased location analysis of E2F1-binding sites suggests a widespread role for E2F1 in the human genome. Genome Res 16, 595-605.

Blahnik, K.R., Dou, L., O'Geen, H., McPhillips, T., Xu, X., Cao, A.R., Iyengar, S., Nicolet, C.M., Ludascher, B., Korf, I., and Farnham, P.J. (2010). Sole-Search: an integrated analysis program for peak detection and functional annotation using ChIP-seq data. Nucleic Acids Res 38, e13.

Erdjument-Bromage, H., Lui, M., Lacomis, L., Grewal, A., Annan, R.S., McNulty, D.E., Carr, S.A., and Tempst, P. (1998). Examination of micro-tip reversed-phase liquid chromatographic extraction of peptide pools for mass spectrometric analysis. J Chromatogr A 826, 167-181.

O'Geen, H., Nicolet, C.M., Blahnik, K., Green, R., and Farnham, P.J. (2006). Comparison of sample preparation methods for ChIP-chip assays. Biotechniques 41, 577-580.

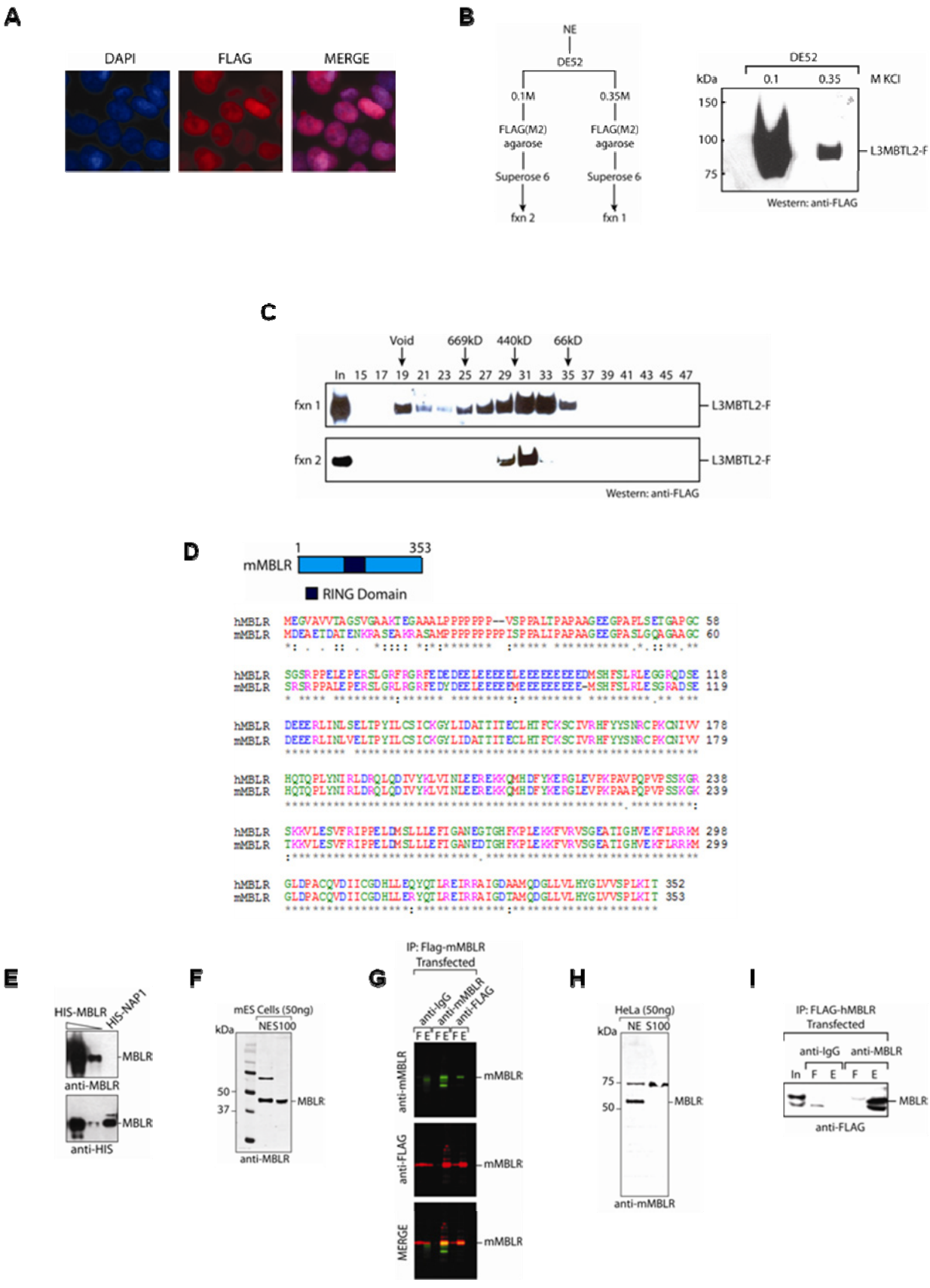
Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y., Zeng, T., Euskirchen, G., Bernier, B., Varhol, R., Delaney, A., et al. (2007). Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. Nature methods 4, 651-657.

Sebastiaan Winkler, G., Lacomis, L., Philip, J., Erdjument-Bromage, H., Svejstrup, J.Q., and Tempst, P. (2002). Isolation and mass spectrometry of transcription factor complexes. Methods 26, 260-269.

Squazzo, S.L., O'Geen, H., Komashko, V.M., Krig, S.R., Jin, V.X., Jang, S.W., Margueron, R., Reinberg, D., Green, R., and Farnham, P.J. (2006). Suz12 binds to silenced regions of the genome in a cell-type-specific manner. Genome Res 16, 890-900.

Trojer, P., Li, G., Sims, R.J., 3rd, Vaquero, A., Kalakonda, N., Boccuni, P., Lee, D., Erdjument-Bromage, H., Tempst, P., Nimer, S.D., et al. (2007). L3MBTL1, a histone-methylation-dependent chromatin lock. Cell 129, 915-928.

Zufferey, R., Dull, T., Mandel, R.J., Bukovsky, A., Quiroz, D., Naldini, L., and Trono, D. (1998). Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. J Virol 72, 9873-9880.


A
DAPI
FLAG
MERGE
B
NE
DE52
0.1M
0.35M
FLAG(M2) agarose
FLAG(M2) agarose
Superose 6
Superose 6
fxn 2
fxn 1
DE52
kDa
0.1
0.35
M KCl
150
100
75
L3MBTL2-F
Western: anti-FLAG
C
Void
669kD
440kD
66kD
In 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47
fxn 1
fxn 2
L3MBTL2-F
L3MBTL2-F
Western: anti-FLAG
D
mMBLR
1
353
RING Domain
hMBLR MEGVAVVIAGSVGAAKIEGAAALPPPPPPP--VSPPALIPAPAAGEEGPAPLSETGAPGC 58
mMBLR MDEAETDATENKPASEAKRASAMPPPPPPPPISPPALIPAPAAGEEGPASLGQAGAAGC 60
hMBLR SGSRPPELEPERSLGRFRGRFEDEDEELEEEELEEEEEEEEDMSHFSLRLEGGRQDSE 118
mMBLR SRSRPPALEPERSLGRLRGRFEDYDEELEEEEEMEEEEEEEEE-MSHFSLRLESGRADSE 119
hMBLR DEEERLINLSELTPYILCSICKGYLIDATTITECLHTFCKSCIVRHFYYSNRCPKCNIVV 178
mMBLR DEEERLINLVELTPYILCSICKGYLIDATTITECLHTFCKSCIVRHFYYSNRCPKCNIVV 179
hMBLR HQTQPLYNIRLDRQLQDIVYKLVINLEEREKKQMHDFYKERGLEVPKPAVPQPVPSSKGR 238
mMBLR HQTQPLYNIRLDRQLQDIVYKLVINLEEREKKQMHDFYKERGLEVPKPAAPQPVPSSKGK 239
hMBLR SKKVLESVFRIPPELDMSLLLEFIGANEGTGHFKPLEKKFVRVSGEATIGHVEKFLRRKM 298
mMBLR TKKVLESVFRIPPELDMSLLLEFIGANEDTGHFKPLEKKFVRVSGEATIGHVEKFLRRKM 299
hMBLR GLDPACQVDIICGDHLLEQYQTLREIRRAIGDAAMQDGLLVLHYGLVVSPLKIT 352
mMBLR GLDPACQVDIICGDHLLERYQTLREIRRAIGDTAMQDGLLVLHYGLVVSPLKIT 353
E
HIS-MBLR
HIS-NAP1
MBLR
anti-MBLR
MBLR
anti-HIS
F
mES Cells (50ng)
NES100
kDa
50
37
MBLR
anti-MBLR
G
IP: Flag-mMBLR
Transfected
anti-IgG
anti-mMBLR
anti-FLAG
F E F E F E
anti-mMBLR
mMBLR
anti-FLAG
mMBLR
MERGE
mMBLR
H
HeLa (50ng)
NE S100
kDa
75
50
MBLR
anti-mMBLR
I
IP: FLAG-hMBLR
Transfected
anti-IgG
anti-MBLR
In F E F E
MBLR
anti-FLAG

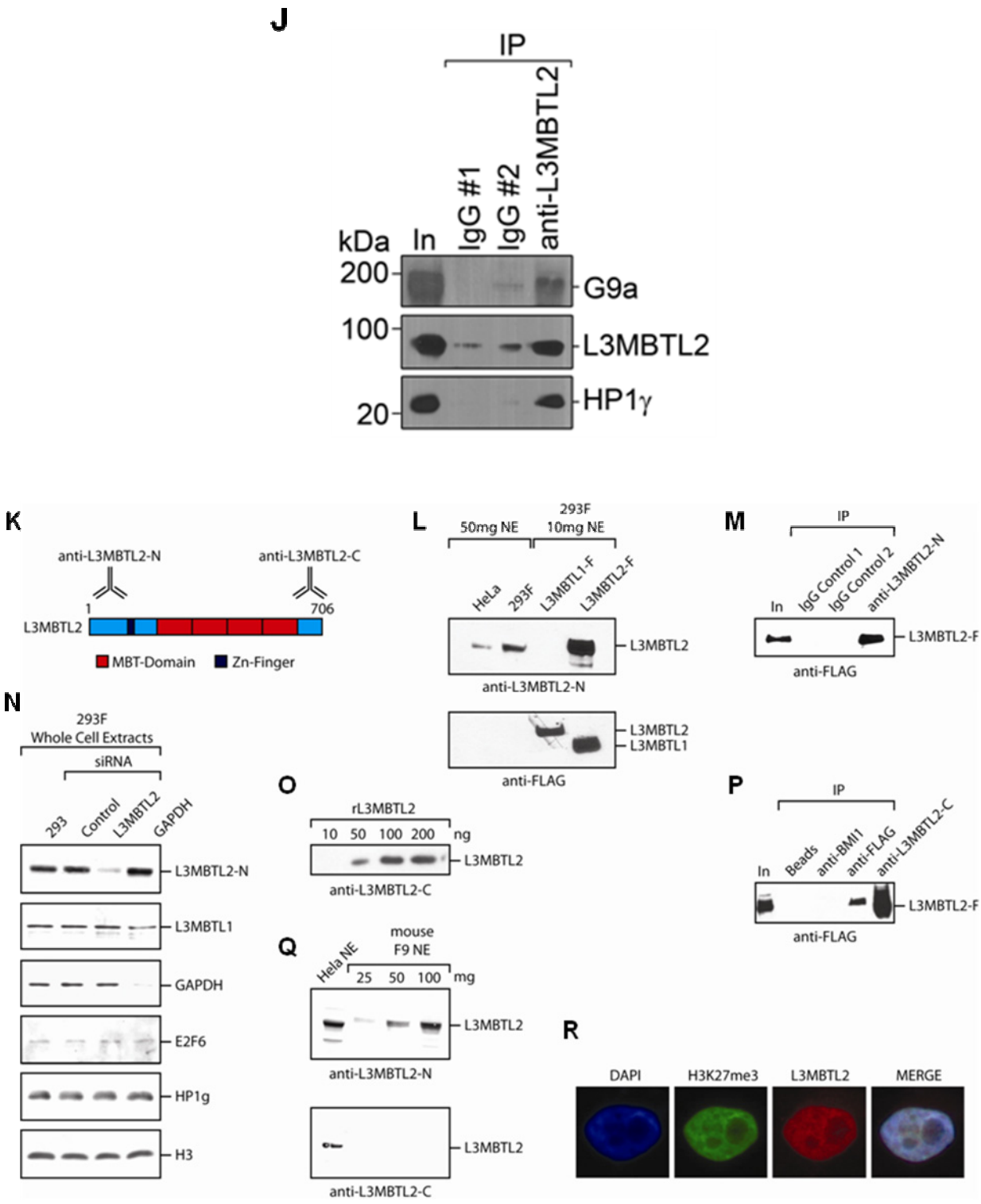

J
IP
anti-L3MBTL2
IgG #2
IgG #1
In
kDa
200
100
20
G9a
L3MBTL2
HP1γ
K
anti-L3MBTL2-N
anti-L3MBTL2-C
1
706
L3MBTL2
MBT-Domain
Zn-Finger
L
293F
50mg NE
10mg NE
HeLa
293F
L3MBTL1-F
L3MBTL2-F
L3MBTL2
anti-L3MBTL2-N
L3MBTL2
L3MBTL1
anti-FLAG
M
IP
In
IgG Control 1
IgG Control 2
anti-L3MBTL2-N
L3MBTL2-F
anti-FLAG
N
293F
Whole Cell Extracts
siRNA
293
Control
L3MBTL2
GAPDH
L3MBTL2-N
L3MBTL1
GAPDH
E2F6
HP1g
H3
O
rL3MBTL2
10 50 100 200 ng
L3MBTL2
anti-L3MBTL2-C
P
IP
In
Beads
anti-BMI1
anti-FLAG
anti-L3MBTL2-C
L3MBTL2-F
anti-FLAG
Q
Hela NE
mouse
F9 NE
25 50 100 mg
L3MBTL2
anti-L3MBTL2-N
L3MBTL2
anti-L3MBTL2-C
R
DAPI
H3K27me3
L3MBTL2
MERGE

Figure S1. Generation of an L3MBTL2-F expressing 293F cell line and purification of L3MBTL2-associated polypeptides, G9a IPs, MBLR and L3MBTL2 antibody characterization

(A) Ectopically expressed FLAG-tagged L3MBTL2 is found exclusively in the nucleus of 293F cells as visualized by indirect immunofluorescence (IF). DNA was stained with DAPI (left), the presence of L3MBTL2 analyzed by anti-FLAG (M2) antibodies (center) and localization to the nucleus determined by merging DAPI and anti-FLAG images (MERGE; right). (B) Purification scheme for L3MBTL2-F protein complexes. L3MBTL2-F was found in both the flow-through (0.1M KCl) and the bound (0.35M KCl) fractions of the DE52 anion exchange chromatography as determined by western blot (right panel). (C) Bound (fxn 1) and flow-through (fxn 2) fractions of the DE52 were separately purified using anti-FLAG immuno-affinity resin and subsequently subjected to a Superose 6 gel filtration column. The L3MBTL2-F elution profiles were determined by western blot using anti-FLAG(M2) antibodies. Fraction numbers and migration of MW markers are indicated on top. (D) Schematic representation of the MBLR domain organization with the RING domain indicated in dark blue (top panel). CLUSTALW2 amino acid alignment of human (h) and mouse (m) MBLR sequences is shown (bottom panel) with asterisks indicating amino acid identity, double dots strong similarity and single dots weak similarity between the two sequences. Notably, mouse and human MBLR are 88% identical and more than 95% similar over their entire amino acid sequences. The mMBLR sequence was expressed as a recombinant protein in *E. coli* and used as an immunogen in rabbits. (E) Anti-MBLR antibodies recognized recombinant hexahistidine (HIS)-tagged mMBLR protein but not HIS-tagged NAP-1 protein. The upper panel shows a western blot using anti-MBLR, the lower panel the same western blot membrane probed with anti-HIS antibodies. (F) Anti-MBLR recognized a specific band at approximately 43 kD in nuclear extract (NE) and cytoplasmic extracts (S100) from mouse embryonic stem (ES) cells. This corresponds well with the mMBLR calculated MW of 39 kDa. The approximately 73 kD species in NE is considered a cross-reacting protein. (G) Full-length FLAG-tagged mMBLR was ectopically expressed in 293F cells and whole cell extracts were used for immunoprecipitation (IP) experiments using anti-MBLR, anti-FLAG or an IgG control. Both, anti-MBLR and anti-FLAG, but not the IgG control efficiently precipitated FLAG-mMBLR. The top panel shows a western blot with immunoprecipitates analyzed by anti-MBLR antibodies (using an Infrared 800 dye conjugated anti-rabbit secondary

antibody which results in green signals), the center panel shows the same western blot analyzed by anti-FLAG antibodies (using an Infrared 700 dye conjugated anti- mouse secondary antibody which results in red signals) and the bottom panel shows an overlay of both signals. (H) The anti-MBLR antibodies detect a specific band at 53 kD in HeLa NE. A band at 75 kD detected in NE and S100 is considered a cross-reactive species. (I) Full-length FLAG-tagged hMBLR was ectopically expressed in 293F cells and whole cell extracts were used for IP experiments using anti-MBLR or an IgG control. Analysis of supernatants (F) and precipitates (E) by western blot showed that anti-MBLR but not the IgG control precipitated FLAG-hMBLR, confirming that the MBLR antibodies recognize both mouse and human MBLR. (J) HeLa nuclear extracts (1mg each) were incubated with two IgG control (IgG #1, #2) or with anti-L3MBTL2 antibodies and precipitated using protein A sepharose. Precipitates were analyzed by western blot. Shown are 50 μg of nuclear extract (In) and the precipitates from IgG#1, IgG#2 and anti L3MBTL2 antibodies. Immunodetection was carried out with the antibodies indicated on the right. Size markers are indicated on the left. (K) Schematic representation of the L3MBTL2 domain organization with the MBT domains indicated in red and the Zn-finger in dark blue. The immunogen for the antibody directed against the L3MBTL2 N-terminus (α-L3MBTL2-N) comprised residues 1-178 and for the antibody against the L3MBTL2 C-terminus (α-L3MBTL2-C) residues 606-705. (L) The L3MBTL2-N antibodies recognized ectopically expressed FLAG-tagged L3MBTL2 but not FLAG-L3MBTL1 in 293F nuclear extracts (lanes 3 and 4) and endogenous L3MBTL2 in HeLa and 293F nuclear extracts (lanes 1 and 2). The upper panel shows a western blot using anti-L3MBTL2-N, the lower panel the same membrane probed with anti-FLAG(M2) antibodies. (M) Full-length FLAG-tagged L3MBTL2 was ectopically expressed in 293F cells and whole cell extracts were used for IP experiments using anti-L3MBTL2-N or two different IgG controls. Anti-L3MBTL2-N but not the IgG controls precipitated FLAG-L3MBTL2. (N) RNAi experiment to confirm the identity of the protein recognized by anti-L3MBTL2-N. 293F cells were transfected with siRNAs ([5 nM] final in medium) against L3MBTL2 or GAPDH or with nonspecific siRNAs (Control) and whole cell extracts were prepared 72 hours post-transfection. Anti-L3MBTL2-N recognized a band at the calculated MW of L3MBTL2 that was significantly reduced upon L3MBTL2 siRNA treatment, but remained unchanged in the presence of other siRNA reagents. GAPDH was only reduced in GAPDH siRNA treated cells. L3MBLT1, H3, E2F6 and HP1γ served as loading controls. (O) The antibody directed against the L3MBTL2 C-

terminus (anti-L3MBTL2-C) detected recombinant full-length L3MBTL2 from Sf9 cells. (P) Full-length FLAG-tagged L3MBTL2 was ectopically expressed in 293F cells and whole cell extracts were used for IP experiments using anti-L3MBTL2-C, anti-FLAG or anti-BMI1 (as control). As well, extracts were incubated with protein A agarose without antibody addition (Beads). Anti-L3MBTL2-C and anti-FLAG but not anti-BMI1 or beads precipitated FLAG-L3MBTL2. (Q) To determine if these antibodies raised against human L3MBTL2 also recognize mouse L3MBTL2, nuclear extracts were prepared from mouse F9 cells. Only anti-L3MBTL2-N (upper panel) but not L3MBTL2-C (lower panel) was able to recognize mouse L3MBTL2. (R) Endogenous L3MBTL2 is found exclusively in the nucleus of HeLa cells as visualized by IF. DNA was stained with DAPI (left), chromatin was visualized by immunodetection using H3K27me3 specific antibodies (second from the left), L3MBTL2 was immunodetected using anti-L3MBTL2-N (second from the right) and co-localization of all signals to the nucleus determined by merging DAPI and immunostained images (MERGE; right).

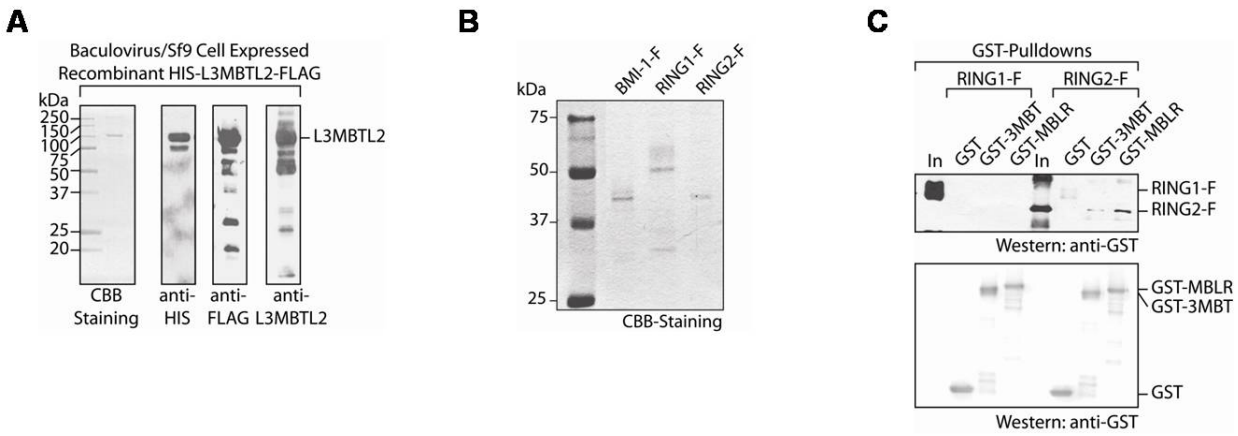


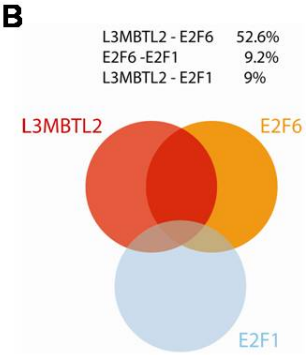
Figure S2. Recombinant PRC1L4 components were used for *in vitro* interaction studies

(A) HIS- and FLAG-tagged, full-length L3MBTL2 was purified from Sf9 cells using a NiNTA agarose and FLAG (M2) agarose affinity purification strategy. Purified L3MBTL2 was resolved by SDS-PAGE and visualized by Coomassie Brillant Blue (CBB) staining. Marker sizes are indicated on the left. The identity of the purified protein was confirmed by immunoblotting using anti-HIS, anti-FLAG and anti-L3MBTL2 specific antibodies. (B) Sf9 cell derived, FLAG-tagged, full-length BMI1, RING1 and RING2 were purified by FLAG (M2) agarose affinity purification and purified proteins were resolved by SDS-PAGE and visualized by CBB staining. Marker sizes are indicated on the left. (C) GST pull-down experiment using GST, GST-3MBT (comprising the three MBT domains of human L3MBTL1) and GST-MBLR recombinant proteins which were incubated with recombinant, full-length RING1-FLAG (RING1-F; lanes 2-4) and RING2-F (lanes 6-8) and precipitated with glutathione sepharose. Shown are western blots of inputs (lanes 1 and 5) and precipitates (lanes 2-4 and 6-8) using anti-FLAG (upper panel) and anti-GST (lower panel) antibodies.

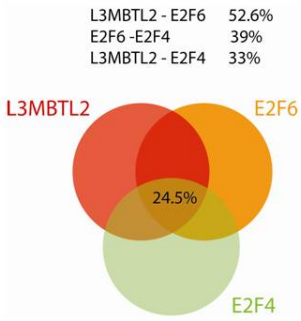
A

Comparison	Top	Overlap #:
E2F1 vs E2F4	1000	98
E2F1 vs E2F6	1000	92
E2F4 vs E2F6	1000	394
E2F1 vs L3MBTL2	1000	90
E2F4 vs L3MBTL2	1000	328
E2F6 vs L3MBTL2	1000	526
E2F6 vs L3MBTL1	1000	156
E2F1 vs L3MBTL1	1000	60
L3MBTL1 vs L3MBTL2	1000	183

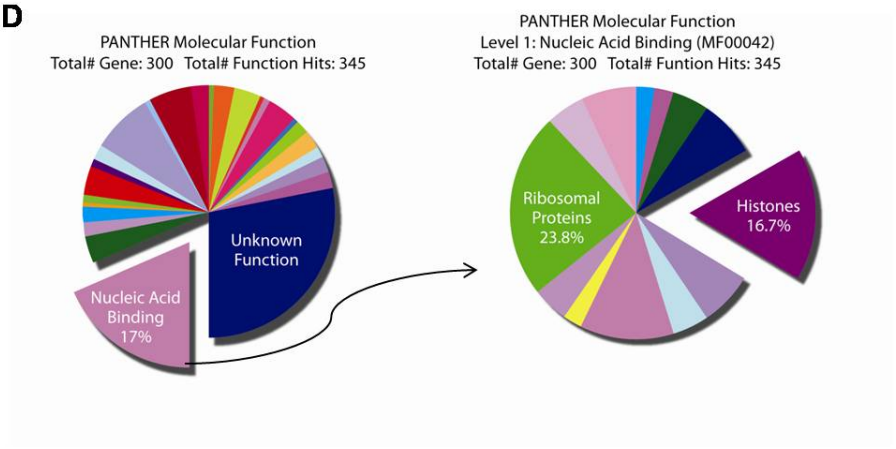
B



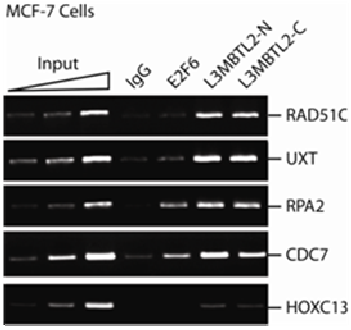
C



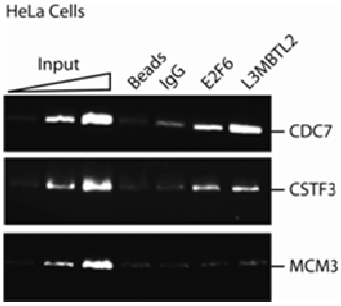
D



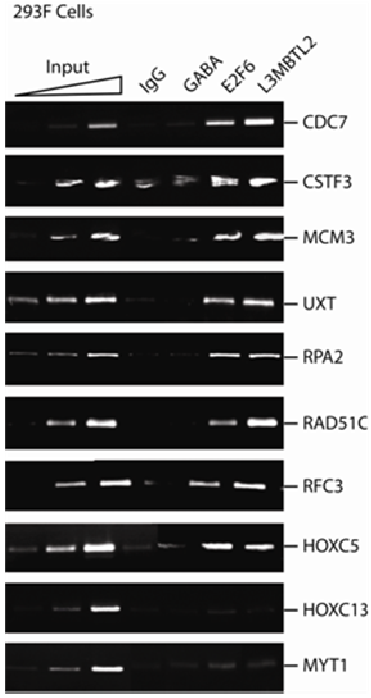
E



F



G



H

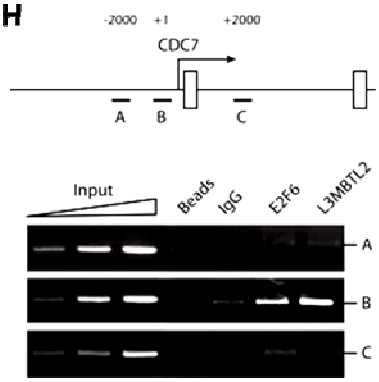
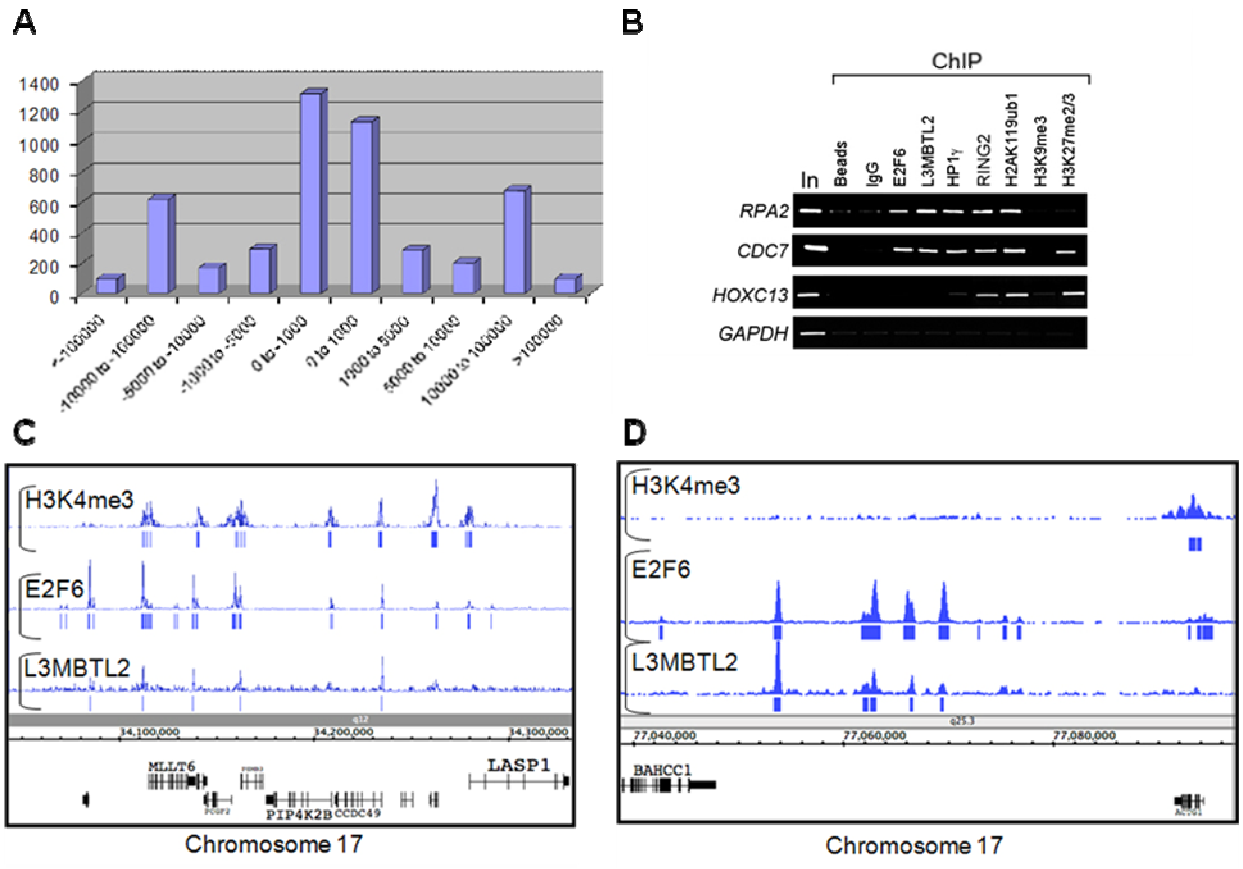


Figure S3. Various aspects of the genome-wide position analysis for L3MBTL2 and E2F transcription factors and manual ChIP assays to confirm L3MBTL2 target genes in various cell lines

(A) Table summarizing the overlap in co-occupancy of the top 1000 binding sites of MBT proteins L3MBTL1 and L3MBTL2 and E2F transcription factors E2F1, E2F4 and E2F6 as determined by a 1.5 kb promoter array ChIP-chip campaign from MCF7 cells. (B) Venn diagrams illustrate the overlap between L3MBTL2, E2F6 and E2F1 target genes in MCF7 cells. While 526 genes were occupied by L3MBTL2 and E2F6, there was only below random co-occupancy of E2F1 and L3MBTL2 and of E2F6 and E2F1 across all genomic promoters. (C) Venn diagrams illustrate the overlap between L3MBTL2, E2F6 and E2F4 target genes in MCF7 cells. The overlap of promoters that are occupied by all three factors is 24.5% which is in the range of random occupancy. (D) Functional classification of L3MBTL2 and E2F6 co-occupied genes based on biological processes as annotated in PANTHER. (E) ChIP assays from asynchronously growing MCF-7 cells were performed to determine occupancy of PRC1L4 complex components on the *RAD51C*, *UXT*, *RPA2*, *CDC7* and *HOXC13* promoters. (F) ChIP assays from asynchronously growing HeLa cells were performed to determine occupancy of PRC1L4 complex components on the *CDC7*, *CSTF3*, and *MCM3* promoters. (G) ChIP assays from asynchronously growing 293F cells were performed to determine occupancy of PRC1L4 complex components on the gene promoters indicated on the right. (H) Schematic of the *CDC7* genomic locus with the TSS indicated with an arrow and exons indicated with boxes. The locations of the primer sets are at -2000 bp, 0 bp and +2000 bp from the TSS and are indicated as A, B and C, respectively. ChIP assays from asynchronously growing 293F cells were performed to determine occupancy of PRC1L4 complex components on various locations along the *CDC7* gene. The antibodies used for ChIP are indicated above. ChIPs using protein G agarose in the absence of antibodies (Beads), IgG control antibodies or anti-GABA antibodies served as negative controls. The inspected gene is indicated on the right of each panel.


A
1400
1200
1000
800
600
400
200
0
B
ChIP
In
Beads
IgG
E2F6
L3MBTL2
HP1γ
RING2
H2AK119ub1
H3K9me3
H3K27me2/3
RPA2
CDC7
HOXC13
GAPDH
C
H3K4me3
E2F6
L3MBTL2
MLLT6
LASP1
PIP4K2B
CCDC49
Chromosome 17
D
H3K4me3
E2F6
L3MBTL2
BAHCC1
Chromosome 17

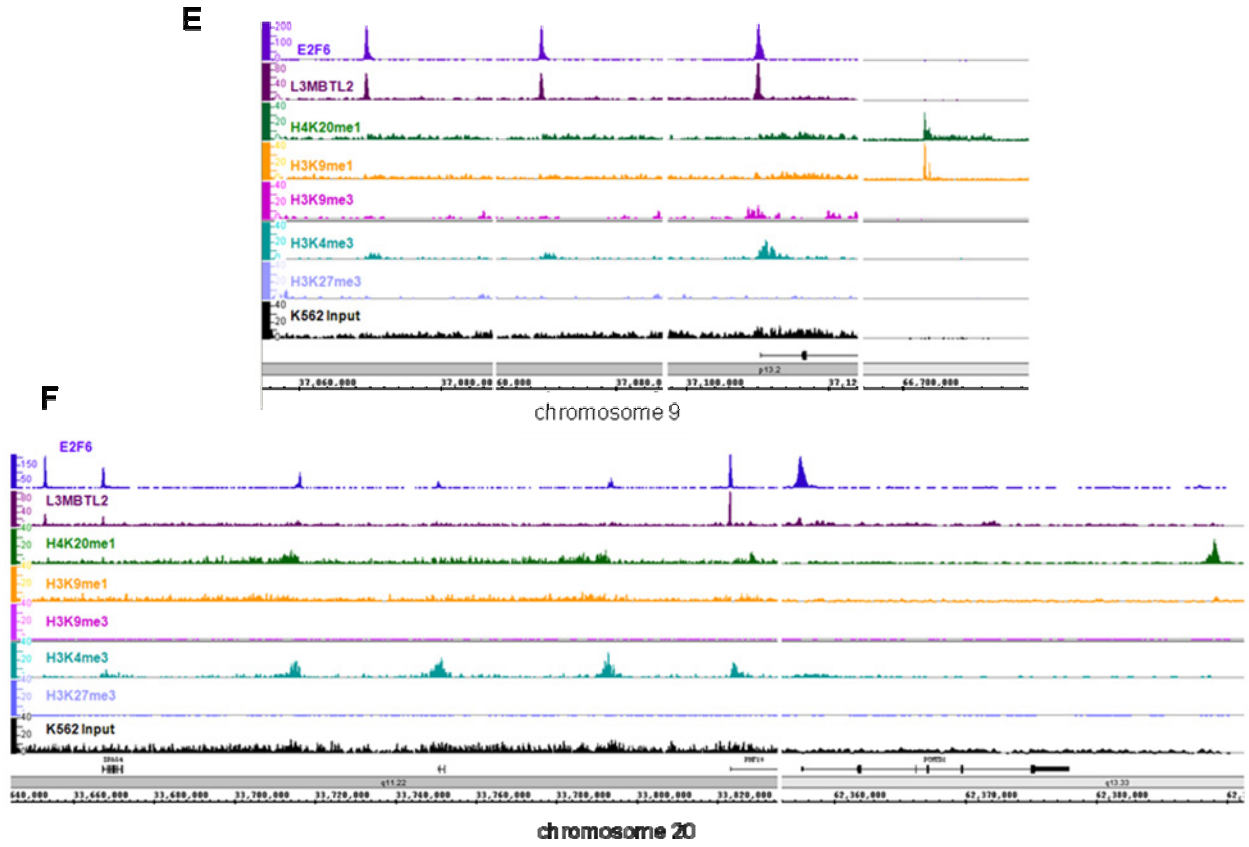


Figure S4. ChIP sequencing in K562 cells to determine L3MBTL2 and E2F6 genomic binding sites and examples illustrating the overlap (or the lack thereof) between L3MBTL2 and E2F6 genomic binding sites and histone methylation marks

(A) L3MBTL2 binding sites as determined by ChIP sequencing experiments were compared for their location relative to transcriptional start sites (TSS). The majority of L3MBTL2 bound within 2 kb of the TSS. (B) ChIP assays in 293F cells were performed to determine binding of PRC1L4 complex components and to query for the presence of several histone methylation marks on L3MBTL2 target genes. The antibodies used are indicated on the top. Binding to the proximal promoter regions of *RPA2* and *CDC7* (two genes identified as L3MBTL2 target genes by ChIP-chip in MCF7 cells) were compared to binding at *HOXC13* (previously identified as occupied by PcG proteins including RING2) and *GAPDH* (as a constitutively active housekeeping gene). (C) A cluster of L3MBTL2 and E2F6 genomic binding sites were found on chromosome 17 that are co-occupied by H3K4me3. (D) A cluster of L3MBTL2 and E2F6

genomic binding sites were found on chromosome 17 that are not co-occupied by H3K4me3. (E) A cluster of L3MBTL2 and E2F6 genomic binding sites were found on chromosome 9 that do not correlate with the presence of H3K27me3, H3K9me3, H3K9me1 and H4K20me1. A certain level of H3K4me3 was detected on all these binding sites. On the right, a different region of chromosome 9 is shown illustrating that there are well defined areas of H3K9me1 and H4K20me1 enrichment. (F) As in (E) but shown is a cluster of L3MBTL2 and E2F6 genomic binding sites on chromosome 20.

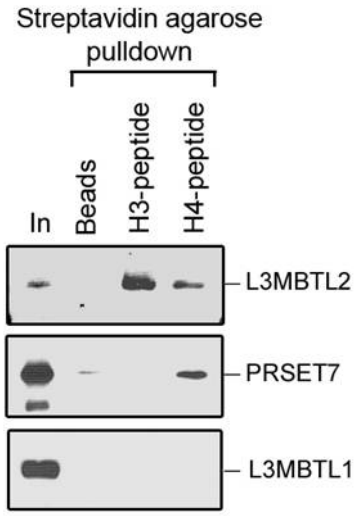


Figure S5. L3MBTL2 interacts with the unmodified histone H3 and H4 N-terminal peptide sequences

Biotinylated peptides corresponding to the N-terminal 20 histone H3 and 24 N-terminal histone H4 residues were incubated with full-length, recombinant L3MBTL2, PR-SET7 and L3MBTL1 proteins. Peptides were precipitated using streptavidin agarose and precipitated fractions were analyzed by western blot. L3MBTL2 interacted with both, the H3 and H4 peptides, while PR-SET7 only interacted with H4. L3MBTL1 did not interact with any of the tested peptides.

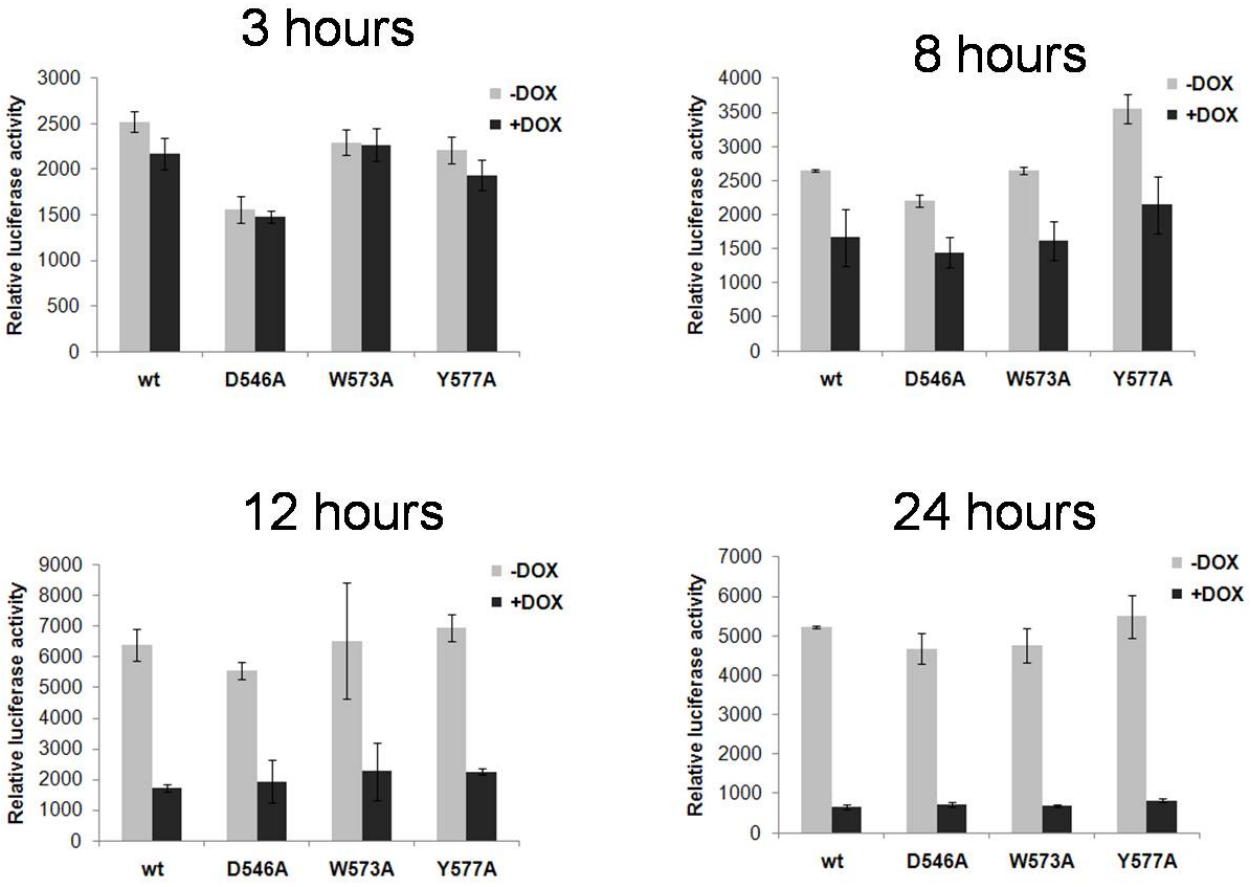


Figure S6. L3MBTL2 repressed a luciferase transgene irrespective of the methyl-lysine binding capability of the fourth MBT domain

GAL4-L3MBTL2 efficiently repressed a luciferase transgene 24 hours upon induction with doxycycline (+DOX). No repression was observed after 3 hours (upper left panel), repression became evident after 8 hours (upper right panel) and was further increasing after 12 hours (lower left panel) of doxycycline addition. GAL4-L3MBTL2 versions, either representing the wildtype sequence or a sequence with single mutated residues (D546A, W573A, Y577A) is indicated below the graph. Data are represented as the mean of three independent experiments ±SD.

Table S1. List of antibodies used in this study

Shown are the names and origins of all antibodies used in this study.

description	supplier	product ID#	used for ChIP-chip	used for ChIP-seq
anti-ART27 (UXT)	kind gift from Susan Logan			
anti-beta actin	AMBION	AM4302		
anti-BMI1	Millipore	05-637		
anti-CDC7	Abcam	ab17880		
anti-E2F1	Santa Cruz	sc-251		
anti-E2F1	Upstate	05-379	☑	
anti-E2F6	Abcam	ab11952		
anti-E2F6	Santa Cruz	sc-8175		
anti-E2F6	Santa Cruz	sc-22823X	☑	☑
anti-FLAG(M2)	SIGMA	F3165		
anti-G9a	Nakatani Lab			
anti-G9a	Bethyl Lab.	A300-933A		
anti-GABA (as negative control)	Santa Cruz	sc-81877		
anti-GAPDH	AMBION	AM4300		
anti-GST	GE Healthcare	21457701V		
anti-H2A	Upstate	07-146		
anti-H2AK119ub1	Upstate	05-678		
anti-H3K4me3	Cell Signaling	9751S		☑
anti-H3K27me2/3	Reinberg Lab			
anti-H3K27me3	Cell Signaling	9733S		☑
anti-H3K9me1	Millipore	17-680		
anti-H3K9me1	Abcam	ab8896		☑
anti-H3K9me2	Millipore	04-768		
anti-H3K9me3	Diagenode	pAb-056-050		☑
anti-H3K9me3	Millipore	07-442		
anti-H4	Cell Signaling	2935S		
anti-H4K20me1	Abcam	ab9051		☑
anti-hexahistidine(HIS)	Qiagen	34660		
anti-HP1alpha	Losson Lab			
anti-HP1beta	Losson Lab			
anti-HP1gamma	Losson Lab			
anti-L3MBTL1	Reinberg Lab		☑	
anti-L3MBTL2 (C-terminal)	Reinberg Lab			
anti-L3MBTL2 (N-terminal)	Reinberg Lab		☑	☑
anti-MBLR	Reinberg Lab			
anti-MYC	Cell Signaling	9402		
anti-Rb (Ab-5)	Calbiochem	OP#66		
anti-RING1	Santa Cruz	sc-28736		
anti-RING1B	Rockland	600-101-292		
anti-RPA2	Calbiochem	NA19L		
anti-SNF2H	Reinberg Lab			
non-specific rabbit IgG	Alpha Diagnostics	210-561-9515 or 20009-5	☑	
secondary rabbit anti-mouse IgG	MP Biomedicals	55436	☑	

Table S2. List of arrays used for ChIP-chip experiments

ARRAY #	CELLS	PLATFORM	DESIGN FILE	CHIP ID	FACTOR	positive control	primer 1	primer 2	negative control	primer 1	primer 2
1	MCF7	ENCODE, hg17	2005-04-25_HG17_ENCODE_50mer	99783	L3MBTL2	Mfap1	cccagccgtagactgaat	cggactcaggtggaaggaat	DHFR3' UTR	ctgatgtccaggaggagaaagg	agcccgacaatgtcaaggactg
2	MCF7	ENCODE, hg17	2005-04-25_HG17_ENCODE_50mer	99254	E2F6	Mfap1	cccagccgtagactgaat	cggactcaggtggaaggaat	DHFR3' UTR	ctgatgtccaggaggagaaagg	agcccgacaatgtcaaggactg
3	MCF7	ENCODE, hg17	2005-04-25_HG17_ENCODE_50mer	98593	L3MBTL1	Myc	GGCTTCTCAGAGGCTTGGCGGG	TCCAGCGTCTAAGCAGCTGCAA	DHFR3' UTR	ctgatgtccaggaggagaaagg	agcccgacaatgtcaaggactg
4	MCF7	1.5kb,hg17	2005-04-18_HGS17_min_promoter_set	87636	L3MBTL2	Mfap1	cccagccgtagactgaat	cggactcaggtggaaggaat	DHFR3' UTR	ctgatgtccaggaggagaaagg	agcccgacaatgtcaaggactg
5	MCF7	1.5kb,hg17	2005-04-18_HGS17_min_promoter_set	88253	E2F6	Mfap1	cccagccgtagactgaat	cggactcaggtggaaggaat	DHFR3' UTR	ctgatgtccaggaggagaaagg	agcccgacaatgtcaaggactg
6	MCF7	1.5kb,hg17	2005-04-18_HGS17_min_promoter_set	87953	L3MBTL1	Myc	GGCTTCTCAGAGGCTTGGCGGG	TCCAGCGTCTAAGCAGCTGCAA	DHFR3' UTR	ctgatgtccaggaggagaaagg	agcccgacaatgtcaaggactg

Table S3. Annotation of top 5% L3MBTL2 and E2F6 target genes as determined by a 1.5 kb promoter array ChIP-chip campaign from MCF cells

Shown are gene IDs and median enrichment values (log2 scale) for L3MBTL2 and E2F6 bound genomic fragments that were hybridized to a 1.5-kb promoter genome tiling array. An overlap of the top 5% of L3MBTL2+E2F6 genomic binding sites is shown.

Table S4. Annotation of L3MBTL2, E2F6 and E2F4 co-occupied genomic binding sites as determined by a 1.5 kb promoter array ChIP-chip campaign from MCF cells

Shown are gene IDs, median enrichment values (log2 scale) for L3MBTL2, E2F6 and E2F4, top 1000 promoters and gene descriptions for all co-occupied binding sites.

Table S5. Annotation of L3MBTL2 genomic binding sites as determined by ChIP sequencing from K562 cells

Shown are gene IDs for L3MBTL2 genomic binding sites that have been identified by a ChIPsequencing campaign from K562 cells. (Sheet 1) L3MBTL2 ChIP sequencing peak information. (Sheet 2) Peak values and genomic location for L3MBTL2 targeted regions.

Number of peaks:	8487
Average peak height:	21.16
Median peak height:	18
Highest peak:	181
Lowest peak:	12
Average peak width:	309.8
TotalReads:	69,898,270
MappedReads:	33,218,228
UniqueReads:	20,706,571