Supplementary Information for

Inactivating mutations of acetyltransferase genes in B-cell lymphoma

Laura Pasqualucci[1,2*], David Dominguez-Sola[1], Annalisa Chiarenza[1], Giulia Fabbri[1], Adina Grunn[1], Vladimir Trifonov[3], Lawryn H. Kasper[4], Stephanie Lerach[4], Honjyan Tang[1], Jing Ma[5], Davide Rossi[6], Vundavalli V. Murty[1,2], Charles G. Mullighan[7], Gianluca Gaidano[6], Raul Rabadan[3], Paul K. Brindle[4] and Riccardo Dalla-Favera[1,2,8*]

[1]Institute for Cancer Genetics and the Herbert Irving Comprehensive Cancer Center, [2]Department of Pathology & Cell Biology, [3]Department of Biomedical Informatics and the Center for Computational Biology and Bioinformatics, and [8]Department of Genetics & Development, Columbia University, New York, NY 10032, USA.

[4]Department of Biochemistry, [5]The Hartwell Center for Bioinformatics and Biotechnology, and *[7]Department of Pathology, St Jude Children's Research Hospital, Memphis, TN 38105, USA.*

[6]Division of Hematology, Department of Clinical and Experimental Medicine and IRCAD, Amedeo Avogadro University of Eastern Piedmont, Novara, Italy.

[*]To whom correspondence should be addressed. E-mail: lp171@columbia.edu; rd10@columbia.edu

SUPPLEMENTARY METHODS

Cell lines. The following DLBCL cell lines were used in the study: OCI-Ly3, OCI-Ly10, HBL1, SUDHL-2, U2932, RIVA, RC-K8, OCI-Ly1, OCI-Ly4, OCI-Ly7, OCI-Ly8, OCI-Ly18, SUDHL-4, SUDHL-5, SUDHL-6, SUDHL-7, SUDHL-8, SUDHL-10, DB, BJAB, FARAGE, VAL, and WSU. All cell lines were maintained in Iscove's Modified Dulbecco Medium (IMDM) supplemented with 10% FBS, 100 μg/mL penicillin, 100 μg/mL streptomycin and 2 mM L-glutamine, except for OCI-Ly10 and OCI-Ly4, which were cultured in IMDM with 20% heparinized human plasma and 55 μM β-mercaptoethanol. HEK293T, H1299 and HeLa cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% FBS, 100 μg/mL penicillin, and 100 μg/mL streptomycin. The following BL cell lines were included in the mutation analysis of *CREBBP* and *EP300*: Ramos, P3HRI, BL113, DAUDI, RAJI, ODHI1, MUTU1.

Primary tumor samples. Primary biopsies from 111 newly diagnosed, previously untreated DLBCL patients were obtained as paraffin-embedded and/or frozen material from the archives of the Departments of Pathology at Columbia University, Weill Cornell Medical College, and the Department of Clinical and Experimental Medicine at the University of Novara, after approval by the respective Institutional Review Boards. The fraction of tumor cells, assessed by southern blot analysis of the rearranged immunoglobulin heavy chain locus and/or by histologic analysis of frozen sections isolated before and after obtaining tissue for molecular studies, corresponded to >80% in most of the cases and to >50% in all cases. The DLBCL cohort (cell lines and primary biopsies) comprised 65 GCB-DLBCL, 54 ABC/non-GC-DLBCL, and 5 unclassified (NC) cases, as determined previously[1] based on gene expression profile analysis[2] (available for 113

samples) or immunohistochemical stains[3] (n=11). None of the patients included in the study had a history of RTS. High molecular weight genomic DNA from 46 FL, 53 CLL, 23 BL and 11 MZL patients were provided by the Department of Clinical and Experimental Medicine, University of Novara, and by the Department of Pathology at Columbia University.

Copy number determination of *CREBBP* and *EP300* in DLBCL by high-density SNP array analysis. Genome-wide DNA profiles were obtained from high molecular weight genomic DNA of DLBCL patients using the Affymetrix Genome-Wide Human SNP Array 6.0 (Affymetrix, Santa Clara, CA, USA) following the manufacturer's instructions. Image data analysis and quality control for the hybridized samples were performed using the Affymetrix Genotyping Console 3.0.1 software, and only samples passing the Affymetrix recommended contrast QC and SNP call rates threshold (in the Birdseed v2.0 algorithm) were considered for analysis. Affymetrix CEL files and corresponding SNP genotype call files generated by the Affymetrix Genotyping Console tool were then analyzed using the dCHIP software[4] according to a previously published workflow[5,6]. Model-based expression was performed using the perfect-match (PM) model to summarize signal intensities for each probe set. Probe intensity data for each array were normalized using a diploid reference set of 3 normal (non tumor) DNA samples that had been processed and hybridized in the same experiment as the tumor samples. The standard invariant-set normalization approach in dCHIP was implemented by a karyotype guided normalization method as described in detail in Ref. 5 and 6. Candidate genomic regions of amplification and deletion were identified by applying the circular binary segmentation (CBS) algorithm

to the SNP array data as described[5,7], with the following criteria: i) mean log2 ratios of ≥ 0.2 or ≤ -0.2 (≥ 0.15 or ≤ -0.15 in two cases where the % of tumor cells in the biopsy was estimated to be ~60%; ii) ≥8 SNP markers (SNP and/or CNV) within a segment. The results of the CBS algorithm were then compared to those of dChip. To exclude calls of genomic gains or loss arising from inherited genomic CNV, the dChipSNP algorithm was also applied to 130 normal DNAs from an independent study[5] as well as to 230 normal DNAs from the HapMap project; alterations identified in the pool of reference samples were assumed to be inherited and therefore excluded. In addition, CNV were excluded if present in the Database of Genomic Variants (http://projects.tcag.ca/cgi-bin/variation/gbrowse/hg18/).

cDNA synthesis and RT-PCR analysis. Total RNA was treated with DNAse prior to cDNA synthesis, according to the manufacturer's instructions. The cDNA was then used as a template in PCR amplification reactions. For detection of the normal and mutant *CREBBP* and *EP300* alleles in affected samples, primers were designed based on the transcript sequences surrounding the validated mutations, and the amplified PCR products were then analyzed by direct sequencing or by cloning and sequencing.

Gene expression data. Gene expression profile analysis of normal B cells and primary DLBCL cases was performed using Affymetrix HG-U133Plus 2.0 arrays as part of an independent study (GEO database GSE12195)[1]. The probe sets used in Fig. S12 are 228177_at (CREBBP) and 202221_s_at (EP300).

Construction of mammalian expression vectors encoding for mutated HA-tagged CREBBP proteins. All CREBBP missense mutants where generated by a site-directed mutagenesis protocol using a previously described mammalian expression vector encoding the mouse Crebbp protein with a C-terminal HA tag as a template (pCIN4-CREBBP-HA, a kind gift of Dr W. Gu). The presence of the desired mutations and the integrity of the open reading frame (ORF) where confirmed in all constructs by enzymatic digestion and full-length sequencing. In western blot analyses, the R1360X construct, which contains a stop codon N-terminal to the HA tag, can only be detected by the anti-CREBBP specific N-22 antibody.

Modeling of CREBBP HAT domain missense mutations. The structural view of the CREBBP HAT domain was generated in the PyMOL v0.99 software (available at http://www.pymol.org) using the coordinates of the crystal structure of the EP300 HAT domain (85% identity with CREBBP) in complex with the Lys-CoA (MMDB ID 62423).

Immunofluorescence analysis of exogenous CREBBP proteins. For analysis of mutant CREBBP subcellular localization, HeLa cells were transiently transfected with plasmids encoding wild-type or mutant CREBBP-HA cDNAs, harvested by trypsinization 30 hr after transfection, and used to prepare cytospins according to standard methods. After fixation in 10% buffered formalin for 15 minutes at room temperature, cells were permeabilized for 10 minutes in 0.2% Triton X-100 in PBS, blocked for 30 min in PBS-Tween with 3% bovine serum albumin (BSA), and incubated 90 min at room temperature in blocking buffer containing a 1:200 dilution of a FITC-conjugated anti-HA antibody

(clone 11.1, Covance). Images were captured using the Nikon NES software coupled to a Nikon ES400 fluorescence microscope with a B/W camera, and artificially colored using Adobe Photoshop.

Isolation of recombinant GST-p53 and CREBBP-HA proteins. Rosetta *E.coli* bacteria containing a GST-p53 expression plasmid were kindly provided by W.Gu and the procedure used for purification of the recombinant protein is described in detail elsewhere[8]. Briefly, GST-p53 expression was induced with 1mM IPTG overnight at 23-25°C and the recombinant protein was obtained from bacteria cell lysates in BC500 using gluthatione-coated agarose beads, upon 90 minutes incubation at 4°C. After several washing steps, GST-p53 recombinant protein was eluted using BC100 buffer containing 20 mM gluthatione. The concentration of GST-p53 in the eluates was assessed by commassie staining upon SDS-PAGE (Figure S8b). Recombinant CREBBP-HA proteins (WT and selected mutants) were isolated from transiently transfected HEK293T cells. Briefly, 48 hr after transfection the cells were harvested in cold PBS and lysed in BC500 (20mM Tris pH 7.5, 500 mM NaCl, 10% Glycerol, 0.2 mM EDTA, 1% Triton X-100, 1mM PMSF, 0.1 mM sodium orthovanadate and protease inhibitors). Cleared lysates were supplemented with 5M NaCl to a final concentration to 650 mM, and CREBBP-HA was immunoprecipitated overnight at 4°C using agarose-conjugated HA beads (Sigma). After several washes in BC500 and BC100, proteins were eluted in BC100 containing 500 µg/mL HA peptide (Sigma) twice, at 4°C. The two fractions were pooled and resolved by SDS-PAGE, followed by commassie blue staining to assess protein purity and concentration (see Fig. S8c).

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1. The *CREBBP* locus is targeted by focal deletions in DLBCL. **a,** Distribution of SNP markers covering the *CREBBP* locus in the Affymetrix Genome-Wide Human SNP array 6.0, as obtained from the University of California Santa Cruz genome browser (http://genome.ucsc.edu/cgi-bin/hgGateway) using the NCBI36.1/hg18 annotation. **b,** Graphic representation of segmentation data from the 9 DLBCL samples carrying CREBBP deletions, visualized using the Integrative Genomics Viewer software (http://www.broadinstitute.org/igv). Each track represents one sample, where white denotes a normal (diploid) copy number, blue indicates a region of copy number loss, and red corresponds to a region of copy number gain (data range for minimum, baseline and maximum value: -1.5, 0, 1.5). Note that, due to the presence of non-tumor cells infiltrating the biopsy, the inferred copy number, and the corresponding color intensity, may vary across samples (see actual values in Table S2). Individual genes in the region are aligned in the bottom panel, and the boxed area (defined by the red bar at the top) corresponds to the *CREBBP* locus. Numbers indicate the genomic coordinates of the displayed region on chromosome 16p. **c,** Copy number plots of the four DLBCL samples showing focal deletions encompassing the *CREBBP* locus (boxed area), as compared to a normal diploid DNA. In each plot, a red line denotes a baseline level of 2.

Supplementary Figure 2. *CREBBP* is mutated in a large fraction of FL cases. a, Distribution of FL-associated point mutations along the *CREBBP* gene and protein, graphically represented as described in Fig. 1a. **b,** Prevalence of *CREBBP* mutated cases in

various B-NHL types. **c,** Overall frequency of *CREBBP* mutations, according to mutation type.

Supplementary Figure 3. The *EP300* gene is mutated in a small fraction of DLBCL and FL. a, Schematic representation of the EP300 protein, with its key functional domains. Color-coded symbols indicate distinct types of mutational events, and the corresponding number is given in brackets. **b,** Overall frequency of *EP300* mutated samples in various B-NHL, including DLBCL cell lines, DLBCL primary cases, FL, BL, and MZL. In the DLBCL bar graph, red and blue colors are used to indicate the phenotypic subtype of the affected cases. **c,** Graphic representation of segmentation data from 6 DLBCL primary biopsies carrying *EP300* deletions, visualized as described in Fig. S1b. Individual genes in the region are aligned in the bottom panel, and numbers indicate the genomic coordinates of the displayed region on chromosome 22q13.2. Two additional cell lines carrying aberrant *EP300* alleles do not appear in the plot since the loss of *EP300* was documented by different methods. These include SUDHL2, where deletion of one allele was documented by FISH analysis, and RC-K8, which has been reported in the literature[9] (see also Table S4). **d,** Venn diagram showing the overlap between cases carrying genomic lesions of *CREBBP* and *EP300* in DLBCL (top) and FL (bottom). Note that four of the six DLBCL samples harboring alterations in both genes are represented by cell lines. Diagrams were generated by using the interactive tool VENNY at *http://bioinfogp.cnb.csic.es/tools/venny/index.html.*

Supplementary Figure 4. The residual wild-type *CREBBP* allele is expressed, suggesting haploinsufficiency. Chromatograms of PCR products amplified from genomic DNA (gDNA) and cDNA (bottom) of representative DLBCL cell lines harboring *CREBBP* mutations. The presence of a double peak in the cDNA sequence (arrows) documents expression of both wild-type and mutated allele.

Supplementary Figure 5. Most CREBBP missense mutations cluster in the HAT domain, proximal to the lysine acetylation reaction center. a, Crystal structure of the EP300 HAT domain (85% identity with CREBBP) in complex with the Lys-CoA (MMDB ID 62423). Residues targeted by somatic point mutations in DLBCL and FL are highlighted in blue (or red if recurrently mutated). Note their clustering around the pocket that accommodates the Lys-CoA bisubstrate, mimicking the lysine side chain from native protein substrates, in close proximity to Ac-CoA[10]. **b,** Ribbon diagram of the HAT domain (front view). Lys-CoA is shown in aquamarine stick figure representation, while the mutated residues are color-coded in pink. **c,** Ribbon diagram of the HAT domain (back view), with the Lys-CoA represented in aquamarine, and the mutated residues color-coded in pink.

Supplementary Figure 6. CREBBP acetylates and inactivates the BCL6 transcriptional repressor. a, Co-immunoprecipitation of CREBBP-HA and BCL6 in transfected HEK293T cells confirms their physical interaction, as previously shown for EP300[11]. **b,** Dose dependent response of a 5X-BCL6 luciferase reporter construct to BCL6, in the absence or presence of increasing CREBBP-HA amounts. Results are shown as

relative activity compared to the basal reporter activity, set as 1 (white bar) (mean +/- standard deviation, as obtained from one of two independent experiments, performed in duplicate). Western blot analysis using anti-HA and anti-BCL6 antibodies monitors for the relative expression levels of exogenous CREBBP and BCL6 in the same lysates (bottom panels).

Supplementary Figure 7. Failure of CREBBP missense mutants to acetylate BCL6 and p53 is not due to loss of physical interaction. a, Subcellular localization of wild-type and mutant CREBBP proteins in HeLa cells, transiently transfected with expression plasmids encoding for the corresponding HA-tagged cDNAs and analyzed by immuofluorescence using specific antibodies directed against the HA tag. Nuclei are counterstained with DAPI, and the signal corresponding to HA is artificially colored in red. **b,** Western blot analysis of whole cell extracts from HEK293T cells co-transfected with Flag-p53 and the indicated CREBBP-HA derivatives, before (inputs) and after immunoprecipitation with Flag/M2 antibodies. **c,** HEK293T cells were transiently transfected with plasmids expressing Flag-tagged human BCL6 and HA-tagged mouse CREBBP proteins (wild-type vs the indicated DLBCL-derived point mutants), and the presence of exogenous CREBBP in the complex was measured by western blot analysis using anti-HA antibodies after immunoprecipitation with Flag/M2 beads (top panel). Western blot analysis with anti-BCL6 antibodies in the total protein extracts before immunoprecipitation (inputs) controls for the expression levels of exogenous BCL6, and β-actin is used as protein loading control. The acetylation status of BCL6 was assessed in the same Flag immunoprecipitates by using anti-acetyl-lysine specific antibodies. Note that the

amounts of pCMV-Flag-BCL6 and pCIN4-CREBBP-HA DNAs were adjusted in the transfection in order to obtain comparable expression levels for each of the proteins across the samples. Numbering of the mutated residues is based on the human CREBBP amino acid sequence.

Supplementary Figure 8. Coomassie blue staining of purified recombinant CREBBP-HA (a) and GST-p53 (b) proteins used in the *in vitro* acetylation assays. In (b), each lane corresponds to increasing dilutions of the original eluate stock.

Supplementary Figure 9. DLBCL-associated CREBBP mutations have variable effects on cAMP-dependent gene expression and are incapable of rescuing the growth phenotype of conditional *Crebbp*/*Ep300* double null MEFs. a, Immunofluorescence analysis of CREBBP protein expression in wild-type and *Crebbp/Ep300* null (dKO) MEFs transduced with retroviral vectors expressing various HA-tagged CREBBP proteins. DAPI identifies the nuclei. **b,** Heat map of hierarchically clustered qRT-PCR gene expression data in *Crebbp/Ep300* dKO MEFs reconstituted with retroviruses expressing wild-type CREBBP (+CREBBP) or various CREBBP mutants (as indicated), and treated with forskolin + IBMX (FI) for 90 min (n=2-5). Data are expressed as log ratio of the wild-type MEF signal. All four DLBCL-derived CREBBP mutants tested were generally deficient for cAMP-responsive transcription, as was the W1502A/Y1503S HAT-dead mutation[12] (used as negative control), although individual CREB target genes displayed different levels of dependence on HAT activity. **c,** Proliferative capacity of conditional *Crebbp/Ep300* dKO MEFs reconstituted with retroviruses expressing wild-type (WT)

CREBBP or the indicated CREBBP mutant alleles. From each reconstituted population, equivalent numbers of YFP^+ (dKO) MEFs were seeded at day 1, and the total number of YFP^+ cells was measured on Day 11 by flow cytometry, as described in the methods section. Data are expressed as a percentage of the total cell number (n=1-3; mean+/- S.E.M.).

Supplementary Figure 10. Quantitative RT-PCR analysis of CREB target gene expression. a-j, qRT-PCR gene expression data from wild-type MEFs, untransduced *Crebbp*/*Ep300* dKO MEFs, and dKO MEFs reconstituted with retroviruses expressing wild-type (+CREBBP) or the indicated mutant HA-tagged CREBBP, and treated for 90 min with forskolin+IBMX (FI) or Ethanol vehicle (EtOH). The W1502A/Y1503S HAT-dead mutation is included as control[12]. Gene expression of individual genes was normalized to expression of Pgk1, and is shown as mean values ± SEM, as obtained from 2-5 independent experiments.

Supplementary Figure 11. DLBCL-associated mutations in *CREBBP* affect H3K18 acetylation. a, Immunofluorescence analysis of histone H3 lysine 18 acetylation (H3K18Ac) in wild type and dKO MEFs, non-transduced or transduced with retrovirus expressing HA-tagged CREBBP proteins (wild type and the indicated mutants). White arrows point to cells that did not delete endogenous *Crebbp* and *Ep300*, and do not express the HA-tagged Crebbp construct, thus serving as a reference for endogenous levels of H3K18Ac. **b,** Signal intensity ratio of H3K18Ac versus exogenous CREBBP (HA) for individual nuclei from the same populations shown in (a) (mean +/- SEM; n=34-65).

Mutated CREBBP proteins showed significantly reduced ability to induce H3K18Ac, independent of retrovirus transduction efficiency (Tukey's post-hoc test of one-way ANOVA for the pairs indicated).

Supplementary Figure 12. Expression of CREBBP and EP300 mRNA in normal and transformed B cells. a, b, CREBBP and EP300 mRNA expression levels in purified normal B cell subpopulations and DLBCL samples (cell lines and primary biopsies), as measured by Affymetrix U133p2 gene expression profile analysis. Data are expressed as absolute probe values after MAS5 normalization. Each dot corresponds to an individual sample, color-coded based on the presence or absence of distinct genetic lesions. CB/CC, GC-derived centroblasts and centrocytes; Na/Me, non-GC naïve and memory B cells (each purified from 5 different individuals); ABC, activated B cell type; GCB, germinal center B cell type. Arrow points to the homozygously deleted DLBCL case 2147. The SUDHL2 cell line was negative for EP300 expression by northern blot analysis (see Fig. 3d).

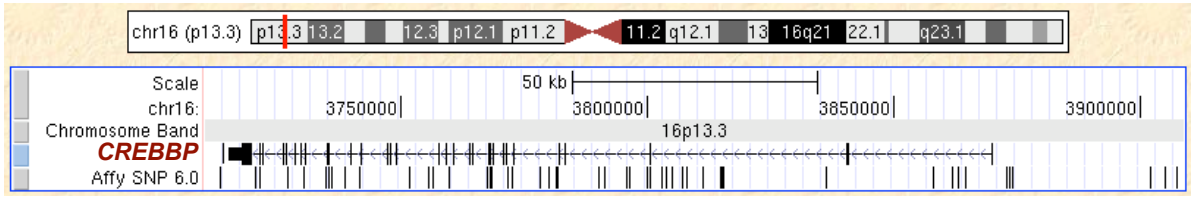
References

1. Compagno, M. et al. Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. *Nature* **459**, 717-21 (2009).
2. Wright, G. et al. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A* **100**, 9991-6 (2003).
3. Hans, C.P. et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood* **103**, 275-82 (2004).
4. Lin, M. et al. dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. *Bioinformatics* **20**, 1233-40 (2004).
5. Mullighan, C.G. et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* **446**, 758-64 (2007).
6. Pounds, S. et al. Reference alignment of SNP microarray signals for copy number analysis of tumors. *Bioinformatics* **25**, 315-21 (2009).

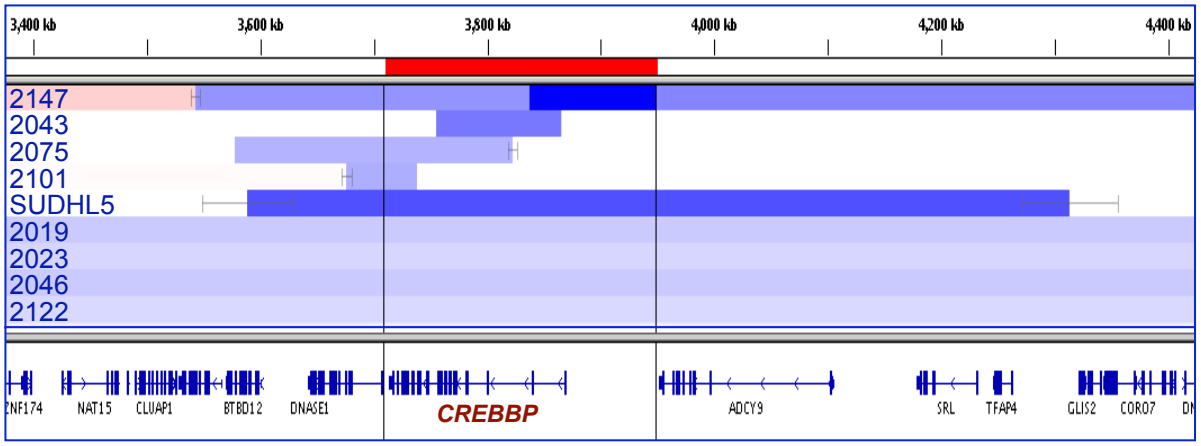
7. Venkatraman, E.S. & Olshen, A.B. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics* **23**, 657-63 (2007).
8. Gu, W., Shi, X.L. & Roeder, R.G. Synergistic activation of transcription by CBP and p53. *Nature* **387**, 819-23 (1997).
9. Garbati, M.R., Alco, G. & Gilmore, T.D. Histone acetyltransferase p300 is a coactivator for transcription factor REL and is C-terminally truncated in the human diffuse large B-cell lymphoma cell line RC-K8. *Cancer Lett* **291**, 237-45 (2010).
10. Liu, X. et al. The structural basis of protein acetylation by the p300/CBP transcriptional coactivator. *Nature* **451**, 846-50 (2008).
11. Bereshchenko, O.R., Gu, W. & Dalla-Favera, R. Acetylation inactivates the transcriptional repressor BCL6. *Nat Genet* **32**, 606-13 (2002).
12. Bordoli, L. et al. Functional analysis of the p300 acetyltransferase domain: the PHD finger of p300 but not of CBP is dispensable for enzymatic activity. *Nucleic Acids Res* **29**, 4462-71 (2001).

Pasqualucci_Figure S1

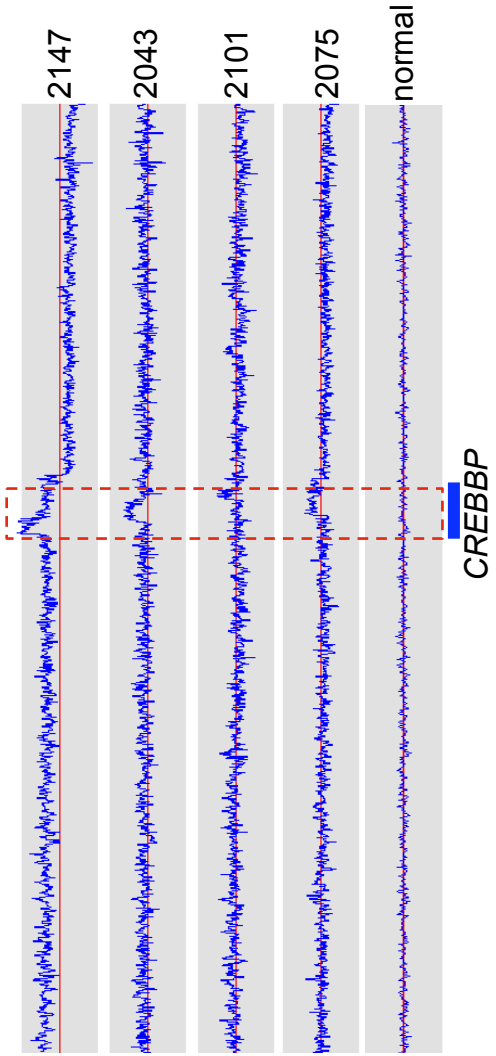
a



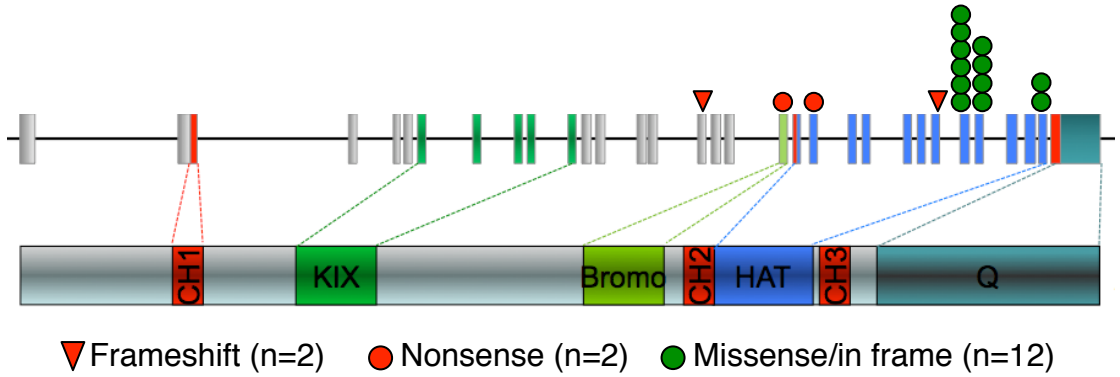
b



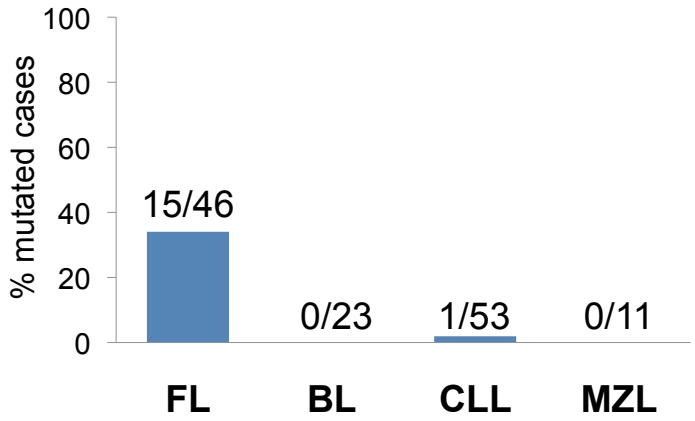
c



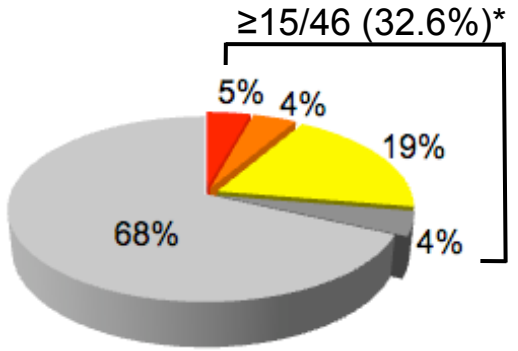
a


CH1
KIX
Bromo
CH2
HAT
CH3
Q
Frameshift (n=2)
Nonsense (n=2)
Missense/in frame (n=12)

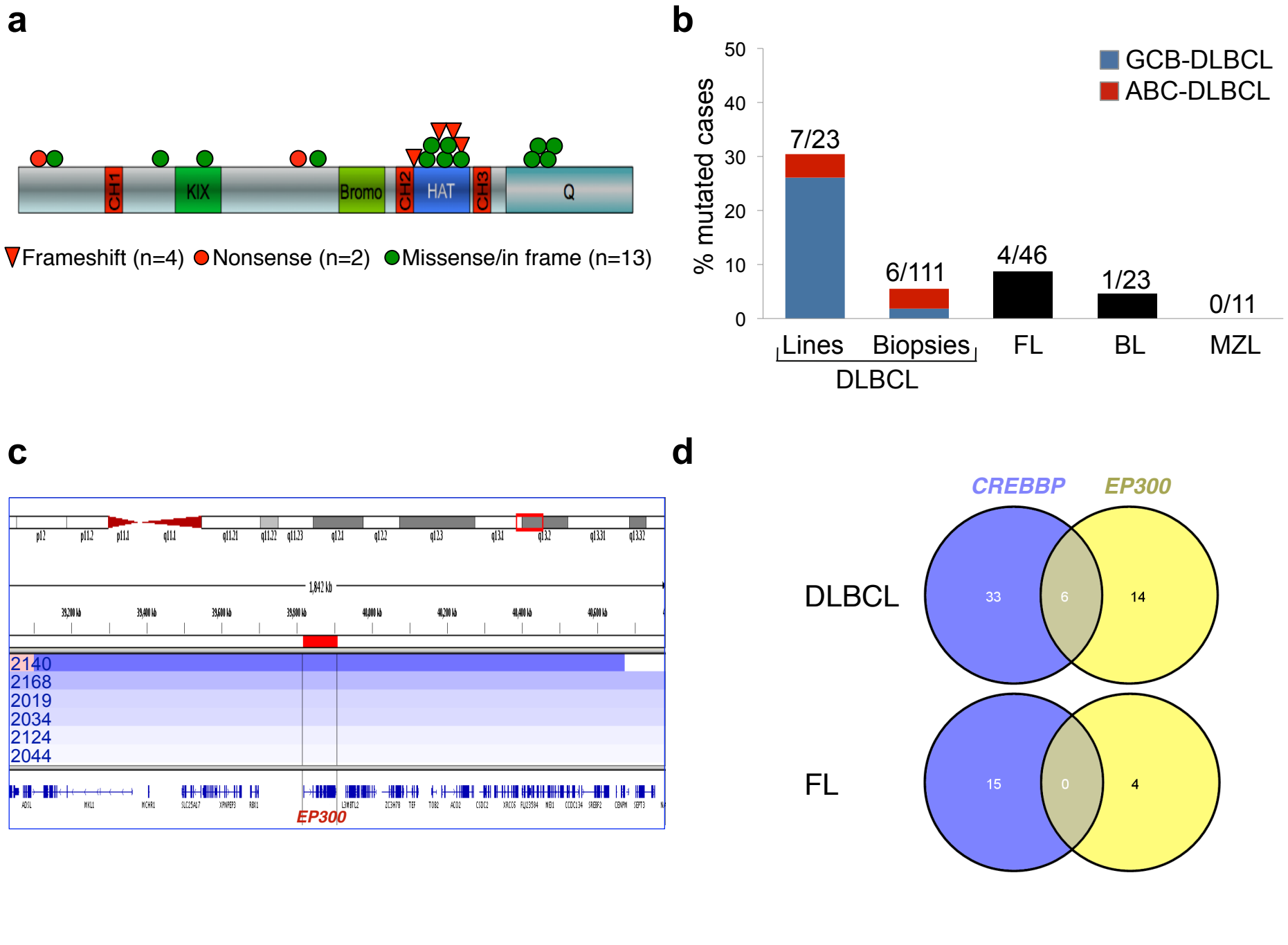
b


% mutated cases
100
80
60
40
20
0
15/46
0/23
1/53
0/11
FL
BL
CLL
MZL

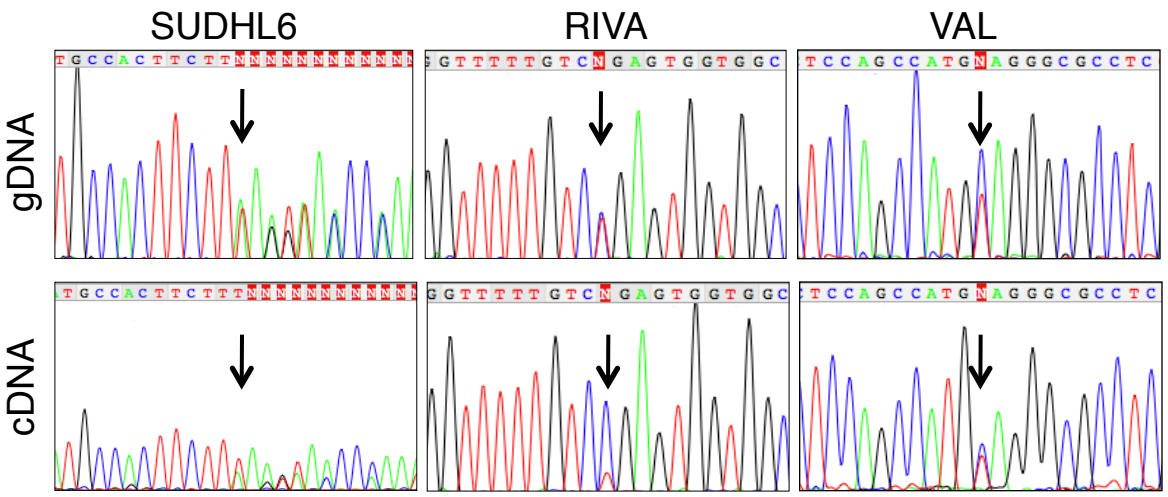
c


≥15/46 (32.6%)*
5%
4%
19%
4%
68%
Frameshift mutation
Nonsense mutation
Missense mutation, HAT domain
Missense mutation, outside HAT
Unmutated
*Deletion data not available

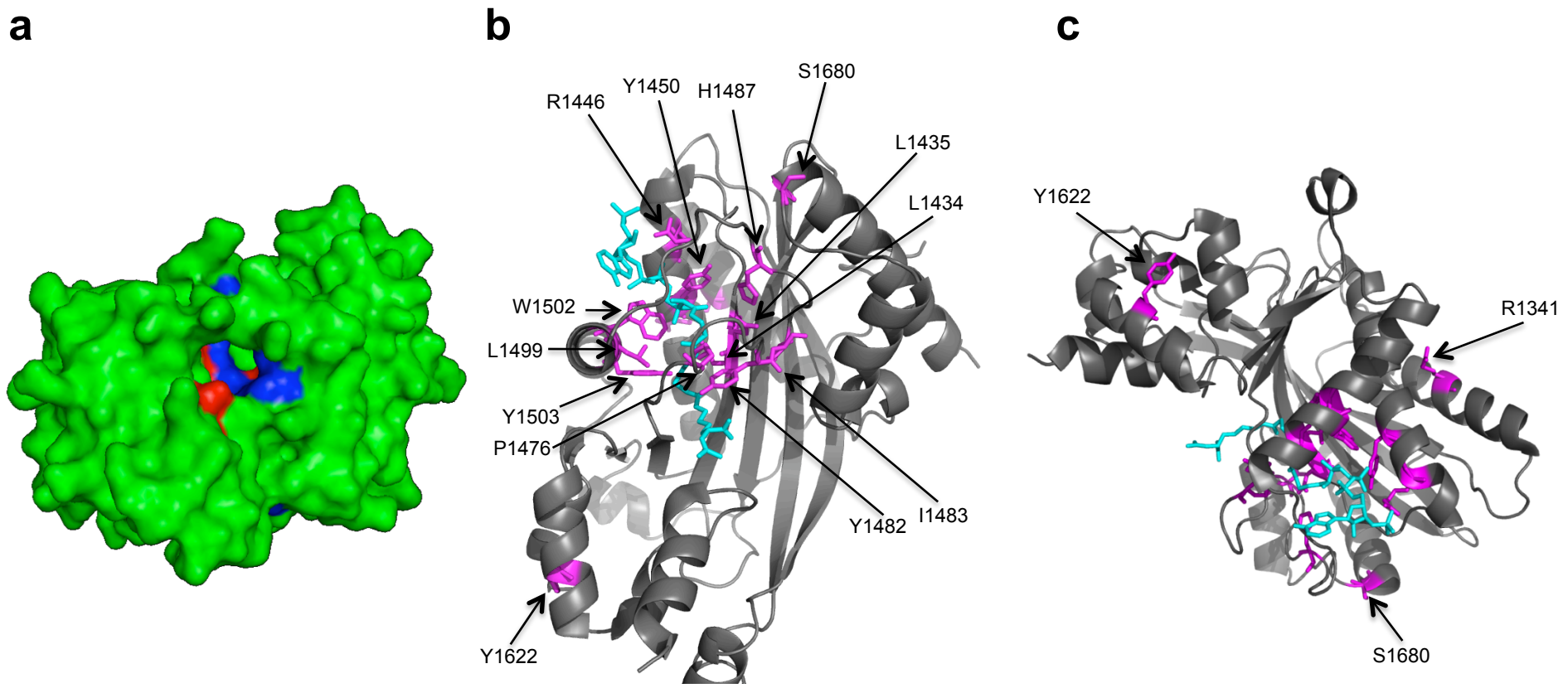
Pasqualucci_Figure S3



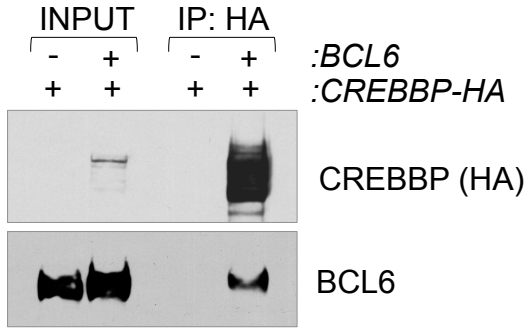
Pasqualucci_Figure S4



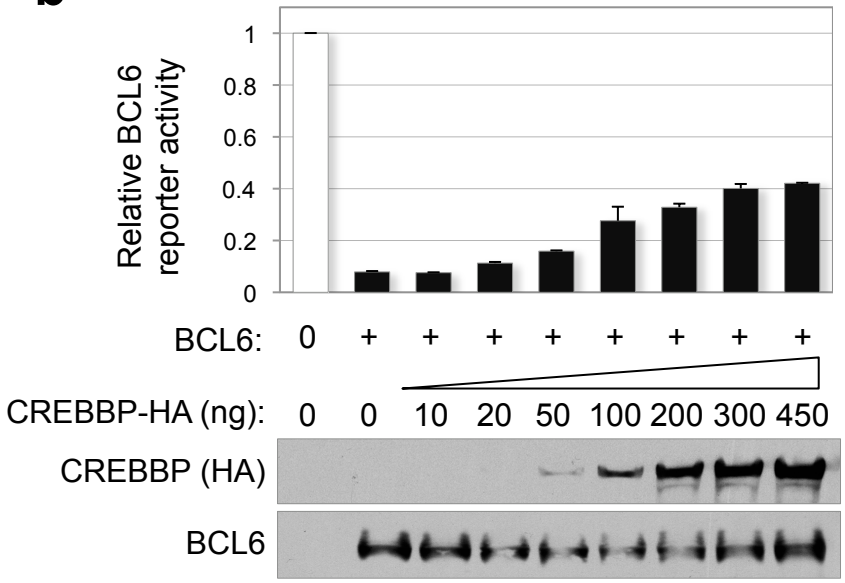
Pasqualucci_Figure S5



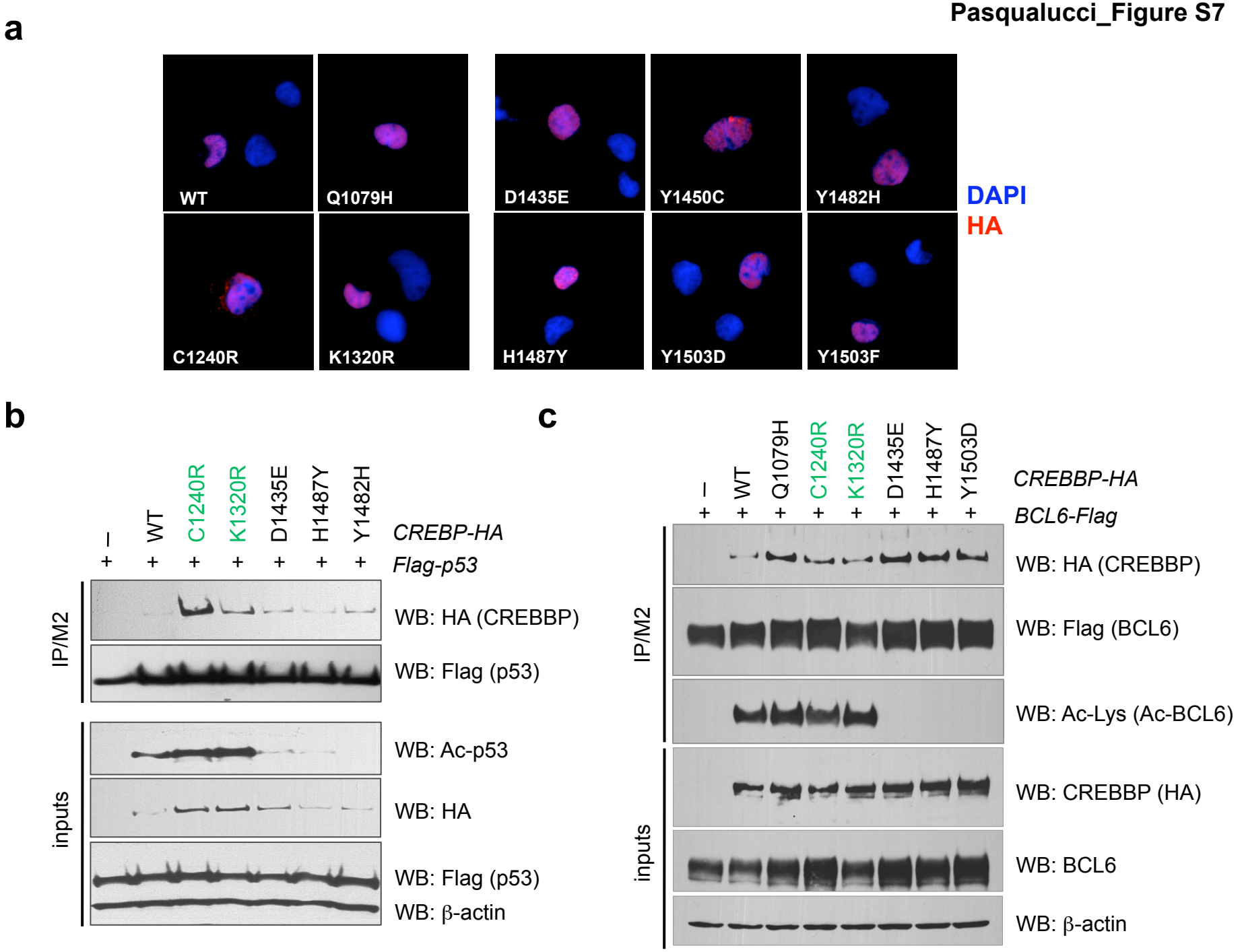
a


INPUT
IP: HA
- + - +
+ + + +
:BCL6
:CREBBP-HA
CREBBP (HA)
BCL6

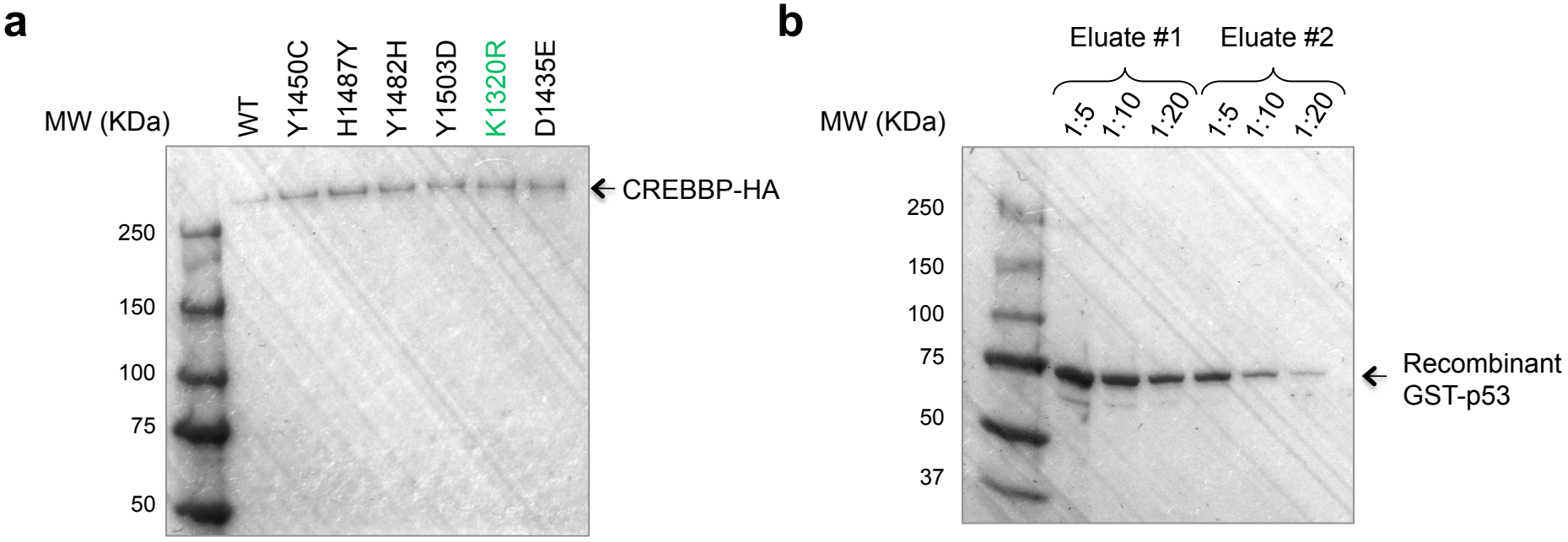
b


Relative BCL6 reporter activity
1
0.8
0.6
0.4
0.2
0
BCL6: 0 + + + + + + + +
CREBBP-HA (ng): 0 0 10 20 50 100 200 300 450
CREBBP (HA)
BCL6

Pasqualucci_Figure S7

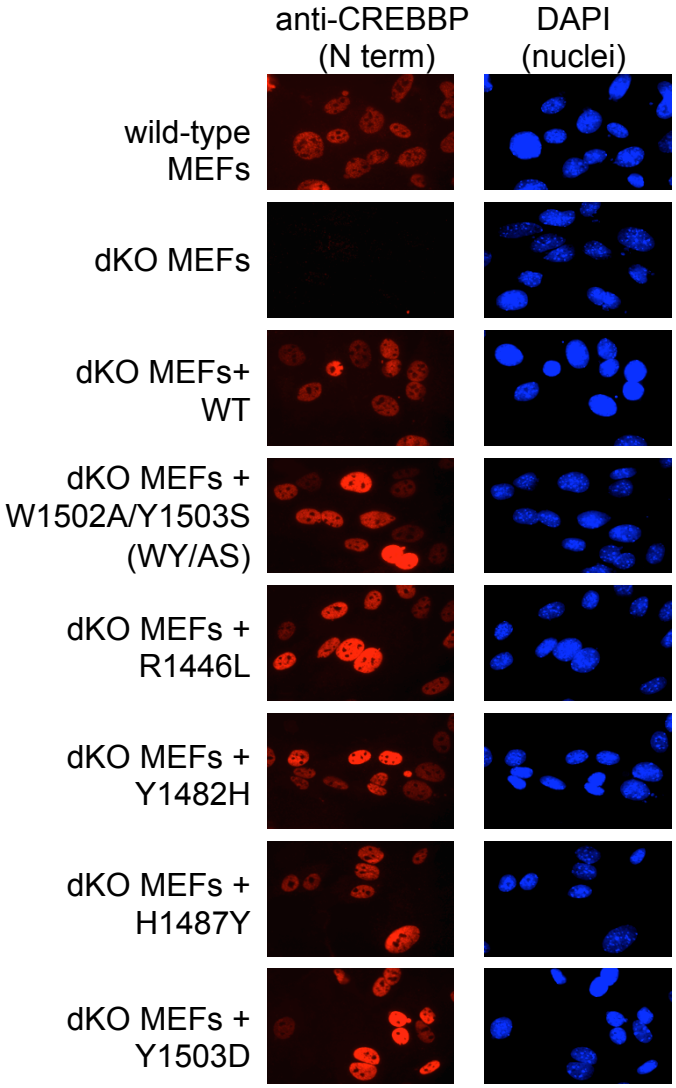


Pasqualucci_Figure S8

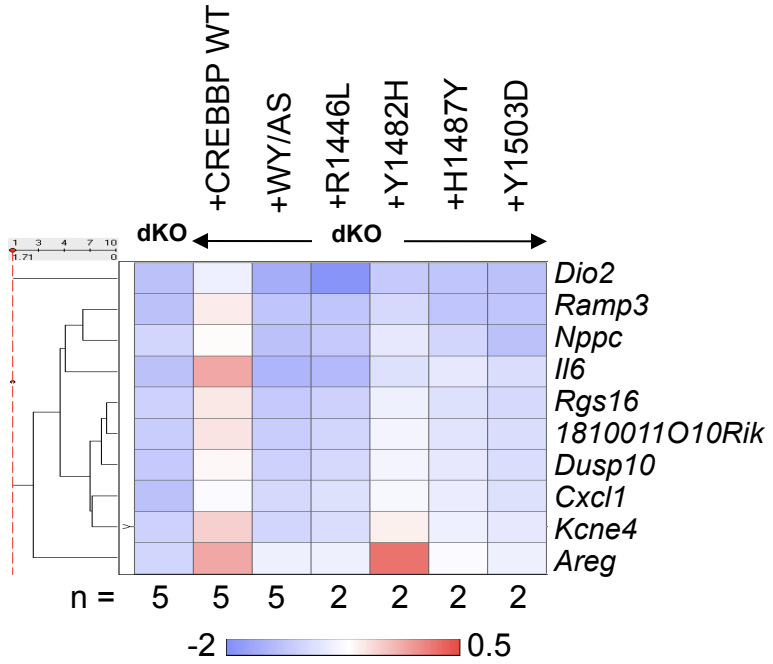


Pasqualucci_Figure S9

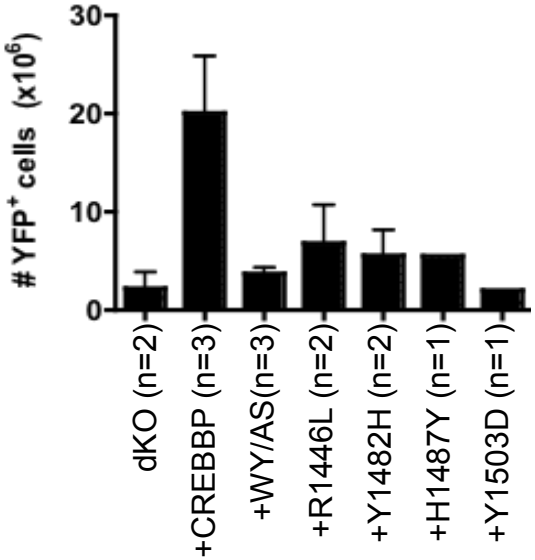
a



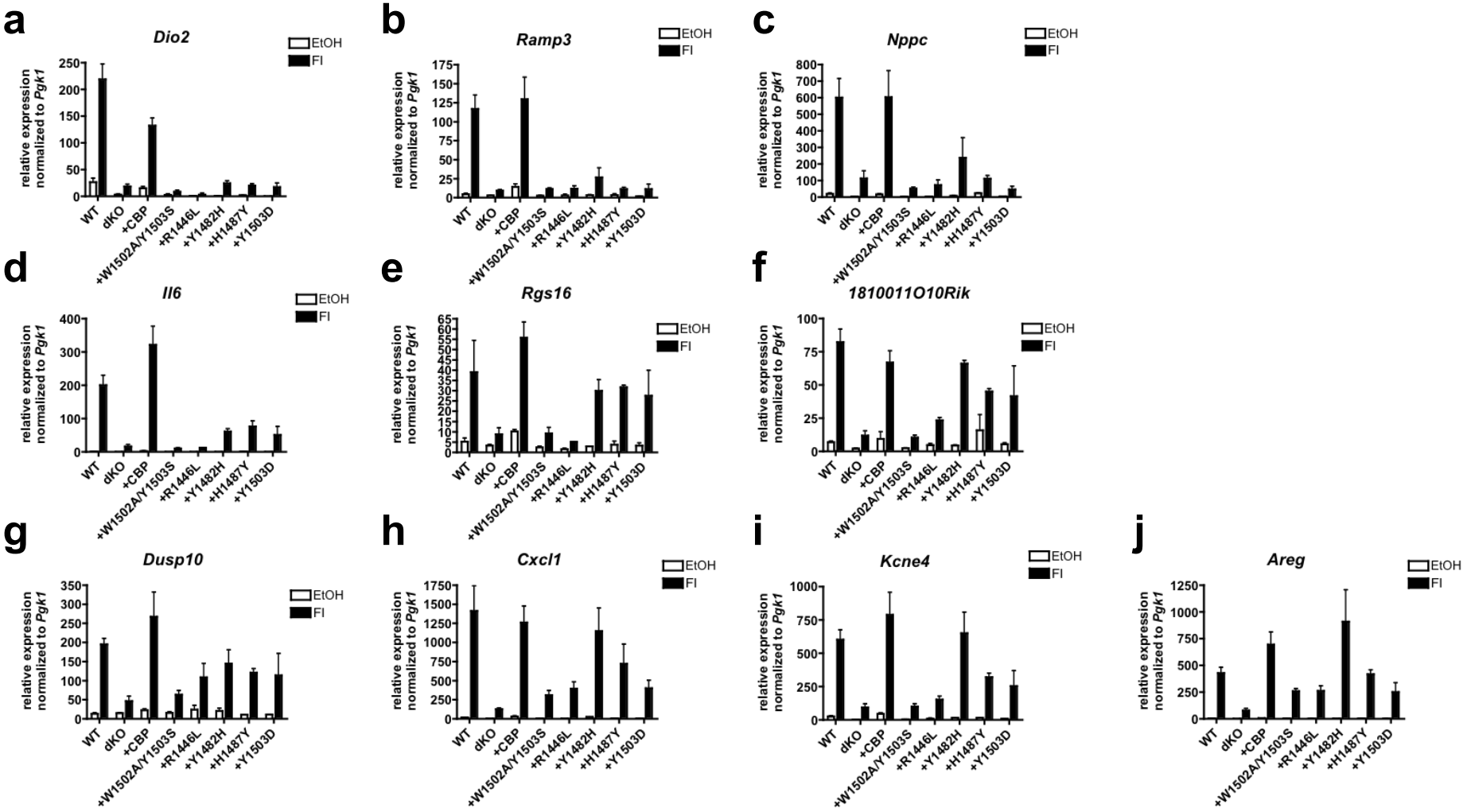
b



c

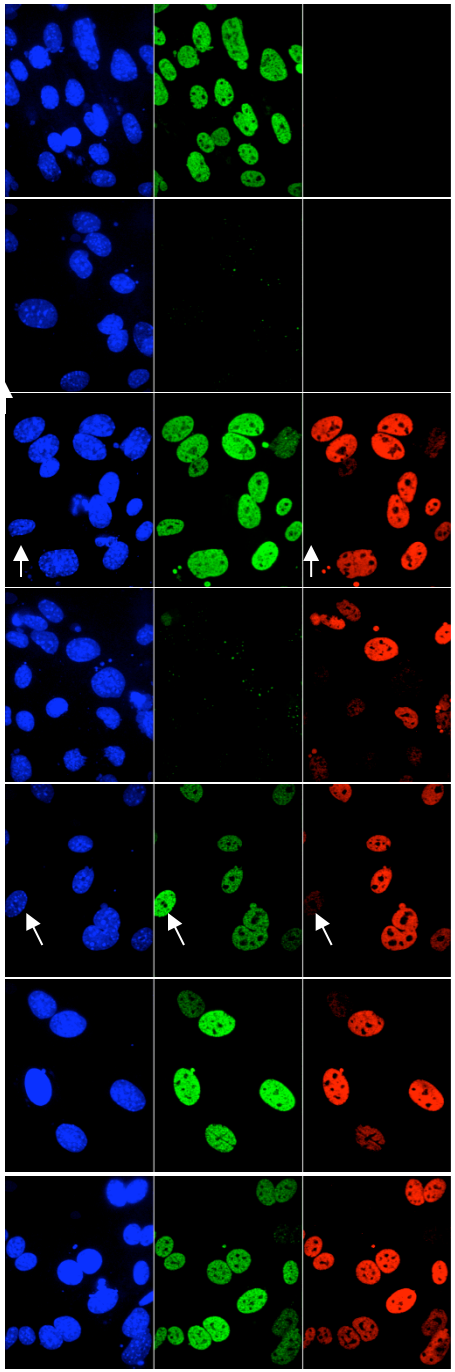


Pasqualucci_Figure S10

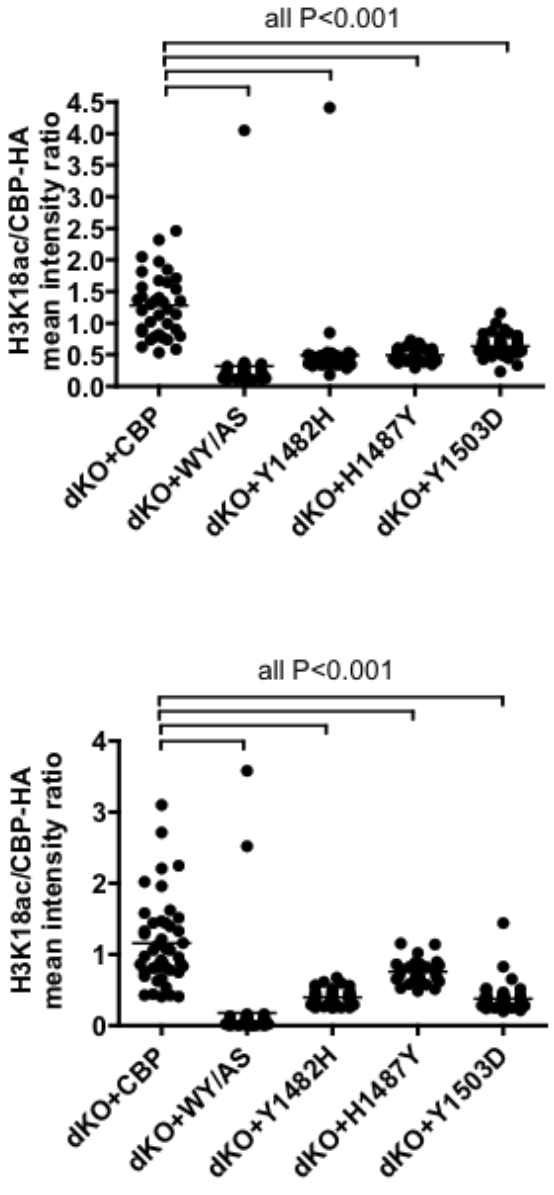


Pasqualucci_Figure S11

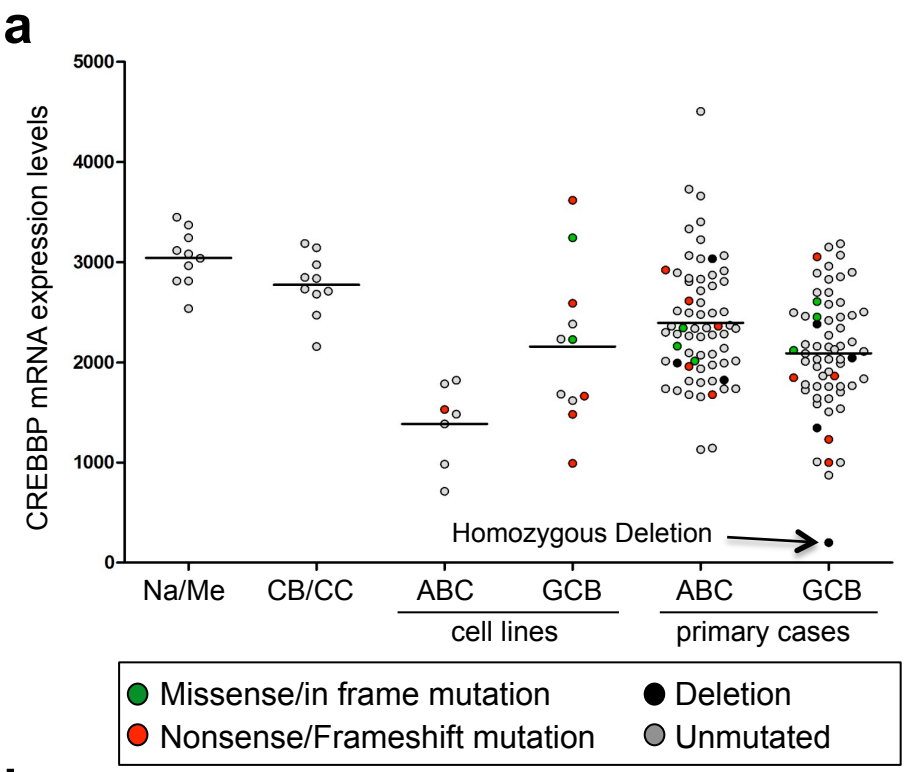
a



b



Pasqualucci_Figure S12



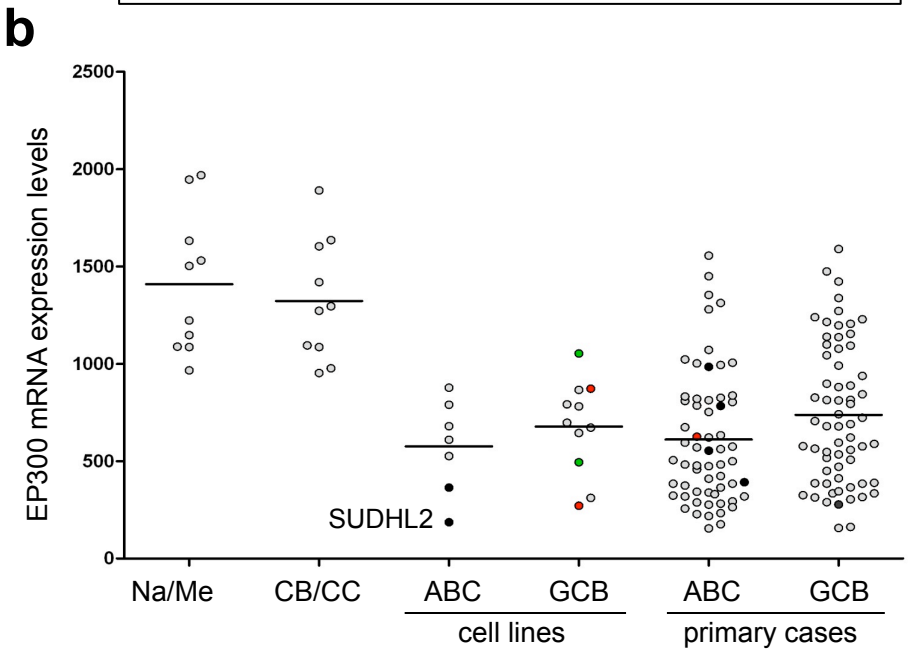


Table S1. Mutations of the *CREBBP* gene in B-NHL

Sample ID	Diagnosis	Exon	Nucleotide change*	AA change*	Affected Domain	HAT activity^^
Missense/In-frame Mutations						
2108	ABC-DLBCL	E16	C3362T	P1053L	interdomain region	retained
FARAGE	GCB-DLBCL	E16	G3441T	Q1079H	interdomain region	retained
Ly8	GCB-DLBCL	E20	T3922C	C1240R	PHD domain	retained
2040^	GCB-DLBCL	E23	A4163G	K1320R	HAT domain	retained
2026	ABC-DLBCL	E24	G4226A	R1341Q	HAT domain	retained
2062	GCB-DLBCL	E26	T4509A	D1435E	HAT domain	lost
2126^	GCB-DLBCL	E26	G4541T (+/+)	R1446L	HAT domain	lost
SUDHL7	GCB-DLBCL	E26	A4553G	Y1450C	HAT domain	lost
34t	GC-DLBCL	E27	C4631G	P1476R	HAT domain	nd
2134	GCB-DLBCL	E27	T4648C	Y1482H	HAT domain	lost
WSU	GCB-DLBCL	E27	C4663T	H1487Y	HAT domain	lost
2015^	ABC-DLBCL	E27	T4711G	Y1503D	HAT domain	lost
33t^	GC-DLBCL	E27	T4711C (+/+)	Y1503H	HAT domain	nd
2109^	GCB-DLBCL	E27	A4712T	Y1503F	HAT domain	lost
SUDHL7	GCB-DLBCL	E30	ΔTCC (5239-5241)	ΔS1680#	HAT domain	lost
2109^	GCB-DLBCL	E30	ΔTCC (5239-5241)	ΔS1680#	HAT domain	lost
2191	GC-DLBCL	E30	ΔTCC (5239-5241)	ΔS1680#	HAT domain	lost
FL22^	FL	E26	A4508G	D1435G	HAT domain	nd
BOR	FL	E26	G4541A	R1446H	HAT domain	lost
FL15^	FL	E26	G4541T	R1446H	HAT domain	lost
FL28^	FL	E26	C4540T	R1446C	HAT domain	nd
FL6^	FL	E26	C4540T	R1446C	HAT domain	nd
FL8	FL	E26	T4505C	L1434P	HAT domain	nd
FL25	FL	E27	T4652G	I1483S	HAT domain	nd
1518^	FL	E27	T4700C	L1499P	HAT domain	nd
BOR	FL	E27	G4710T	W1502C	HAT domain	nd
1647	FL	E27	A4712T	Y1503F	HAT domain	lost
1514	FL	E30	ΔTCC (5239-5241)	ΔS1680#	HAT domain	lost
FL26^	FL	E30	ΔTCC (5239-5241)	ΔS1680#	HAT domain	lost
CLL44^	B-CLL	E29	A5069G	Y1622C	HAT domain	nd
Nonsense Mutations						
2209^	non-GC-DLBCL	E2	C427T	R75X	KIX, BD, HAT, IBiD domain	lost
2152^	ABC-DLBCL	E2	C490T	Q96X	KIX, BD, HAT, IBiD domain	lost
2161	GCB-DLBCL	E2	C670T	Q156X	KIX, BD, HAT, IBiD domain	lost
VAL	GCB-DLBCL	E2	C898T	Q232X	KIX, BD, HAT, IBiD domain	lost
2179	GCB-DLBCL	E12	C2458T	Q752X#	KIX, BD, HAT, IBiD domain	lost
2106	GCB-DLBCL	E18	C3721T	R1173X#	BD, HAT, IBiD domain	lost
2040^	GCB-DLBCL	E23	A4171T	K1323X	HAT, IBiD domain	lost
2020	GCB-DLBCL	E24	C4225T	R1341X	HAT, IBiD domain	lost
RIVA	ABC-DLBCL	E24	C4282T	R1360X	HAT, IBiD domain	lost
FL7	FL	E18	T3579A	Y1125X#	BD, HAT, IBiD domain	lost
FL17	FL	E20	T3906G	Y1234X	BD, HAT, IBiD domain	lost
Frameshift/Splice Site Mutations						
2183	ABC-DLBCL	E6	dupl CCAT (1535-1538)	L446fs	KIX, BD, HAT, IBiD domain	lost
SUDHL8	GCB-DLBCL	I6	T(+2)C	L525fs	KIX, BD, HAT, IBiD domain	lost
SUDHL6	GCB-DLBCL	E6	ΔT (1611)	L470X	KIX, BD, HAT, IBiD domain	lost
Ly1	GCB-DLBCL	E8	ΔC (1922)	T573fs	KIX, BD, HAT, IBiD domain	lost
Ly8	GCB-DLBCL	E17	ΔCT (3505-3506)	L1101fs	BD, HAT, IBiD domain	lost
2170	ABC-DLBCL	E17	dupl 40bp (3465-3505)	L1101fs	BD, HAT, IBiD domain	lost
2069^	ABC-DLBCL	E28	dupl GCACT (4907-4911)	S1572fs	HAT domain	lost
SUDHL10	GCB-DLBCL	E31	ΔC (6035)	A1944fs	IBiD domain	nd
1546	FL	E15	A3263T (splice site)	E1020fs	BD, HAT, IBiD domain	lost
1536^	FL	E25	ΔC (4428)	C1408fs	HAT domain	lost

*Numbering according to GenBank accession No. NM_004380.2 (mRNA) and NP_004371.2 (protein) respectively.

These mutations were also reported in RTS patients (see Roelfsema and Peters, 2007; Schorry et al, 2008; and Leiden Open Variation Database at http://chromium.liacs.nl/LOVD2/home.php).

^ For these samples, paired normal DNA was available and confirmed the somatic origin of the mutation.

^^ as assessed by in vivo acetylation assays (missense mutations) and/or by prediction (nonsense mutations)

Abbreviations: ABC, activated B cell type; GCB, germinal center B cell type; NC, unclassified; non-GC and GC, non germinal center and germinal center type, respectively (based on IHC); FL, follicular lymphoma; B-CLL, B-cell chronic lymphocytic leukemia; Δ, deletion; i, intron; fs, frameshift; dupl, duplication; +/+, hemizygous mutation; nd, not determined.

KIX, CREB-binding; BD, bromodomain; PHD, plant homeodomain; HAT, histone acetyltransferase;IBiD, IRF3-binding

Table S2. SNP array copy number analysis of the *CREBBP* gene in DLBCL

Sample ID	DLBCL Subtype	Chr	Cytoband	Start position*	End position*	Mean CN	Size (Kb)	Aberration	No.of Genes	Annotated genes in the region**
2101	non-GC	16	16p13.3	3681563	3738462	1.48	56.9	loss	2	CREBBP, TRAP1
2043	ABC	16	16p13.3	3756551	3864655	1.21	108.1	loss	1	CREBBP
2147	GCB	16	16p13.3	3838219	3949571	0.65	111.4	loss	1	CREBBP
2147	GCB	16	16p13.3	3547655	3836389	1.35	288.7	loss	5	BTBD12, CREBBP, DNASE1, NLRC3, TRAP1
2075	GCB	16	16p13.3	3580275	3818710	1.51	238.4	loss	4	BTBD12, CREBBP, DNASE1, TRAP1
SUDHL5^	GCB	16	16p13.3	3636366	4282743	1.05	646.4	loss	7	BTBD12, CREBBP, DNASE1, TRAP1, ADCY9, SRL, TFAP4
2023	GCB	16	16p13.3	765	5547879	1.80	5547.1	loss	206	CREBBP and >200 other genes
2046	GCB/NC	16	16p13.3	765	5547879	1.78	5547.1	loss	206	CREBBP and >200 other genes
2019	ABC	16	16p13.3 - 16q21	26113	62521051	1.67	62494.9	loss	525	CREBBP and >200 other genes
2122	GC	16	16p13.3 - 16q24.3	765	88815024	1.64	88814.3	loss (entire chromosome)	776	CREBBP and >200 other genes

* Numbering according to NCBI Build 36.1. Position 765 corresponds to the first marker on chromosome 16 in the Affymetrix SNP6.0 array.
**NCBI RefSeq database.

Abbreviations: ABC, activated B cell type; GCB, germinal center B cell type; NC, unclassified; non-GC and GC, non germinal center and germinal center type, respectively (based on immunohistochemistry).

^ Data for the SUDHL5 cell line were obtained from the Tumorscape database at http://www.broadinstitute.org/tumorscape, and were confirmed by FISH analysis using BAC clone RP11-292B10 and a control centromere 16 probe.

Table S3. Mutations of the *EP300* gene in B-NHL

Sample ID	Diagnosis	Exon	Nucleotide change*	AA change*	Affected Domain
Missense/In-frame Mutations					
2180	ABC-DLBCL	E2	A771G	M126V	none
2161	GCB-DLBCL	E6	A1914G	S507G	none
2141	ABC-DLBCL	E9	G2163C	A590P	KIX domain
2025^	ABC-DLBCL	E14	C3168A	P925T	none
FL1	FL	E26	T4584C	H1397Y	HAT domain
DB	GCB-DLBCL	E28	G4912T	G1506V	HAT domain
SUDHL8	GCB-DLBCL	E29	A5105T	K1570N	HAT domain
SUDHL10	GCB-DLBCL	E30	G5191A	R1599H	HAT domain
SUDHL6	GCB-DLBCL	E30	C5274T	R1627W	HAT domain
2041	NC-DLBCL	E31	A6106C	Q1904P	none
FL14	FL	E31	A6918C	M2175L	none
2135	ABC-DLBCL	E31	ΔCAGCAGCAACAG (6969-6980)	Δ4Q (2192-2195)	none
FL19	FL	E31	Δ12bp (7022-7033)	ΔQFQQ (2210-2213)	none
Nonsense Mutations					
RAMOS	BL	E2	C582T	Q63X	KIX, BD, HAT, IBiD domain
SUDHL2	ABC-DLBCL	E14	C2856T (+/+)	Q821X	KIX, BD, HAT, IBiD domain
Frameshift/Splice Site Mutations					
BJAB	GCB-DLBCL	E24/I24	G(i24+1)A	ΔE24 (R1292fs)**	HAT, IBiD domain
FARAGE	GCB-DLBCL	E27	ΔA4803	M1470fs	HAT, IBiD domain
SUDHL8	GCB-DLBCL	E29	ΔA5031	K1547fs	HAT, IBiD domain
1658	FL	E28/I28	G(i28+1)A	V1540fs	HAT, IBiD domain

*Numbering according to GenBank accession No NM_001429.3 (mRNA) and NP_ 001420.2 (protein) respectively.

Note that none of these mutations has been previously observed in RTS patients (see Roelfsema and Peters, 2007; Bartsch et al, 2009; and Leiden Open Variants Database at http://chromium.liacs.nl/LOVD2/home.php).

** as verified by cDNA amplification followed by both direct sequencing and sequencing of individual clones.

^ For this sample, paired normal DNA was available and confirmed the somatic origin of the mutation. One additional sequence variant identified in 4 patients (G1026A, corresponding to AA change G211S) most likely represents a germline polymorphism since, although no matched normal tissue was available for these patients, the same change was present in normal DNA in a separate study (Ref. 34).

Abbreviations: ABC, activated B cell type; GCB, germinal center B cell type; NC, unclassified; FL, follicular lymphoma; BL, Burkitt lymphoma; Δ, deletion; I, intron; fs, frameshift; +/+, hemyzigous mutation

Table S4. SNP array copy number analysis of the *EP300* gene in DLBCL

Sample ID^	DLBCL Subtype	Chr	Cytoband	Start position*	End position*	Mean CN	Size (Kb)	Aberration	No.of Genes	Annotated genes in the region**
2140	GC	22	22q13.1 - 22q13.2	39102833	40674243	1.10	1571.41	loss	28	EP300 and 27 other genes
2168	ABC	22	22q11.21 - 22q13.33	20175343	49535360	1.60	29360.02	loss	352	EP300 and >200 other genes
2019	ABC	22	22q11.1 - q13.33	14432516	49581309	1.64	35148.79	loss (entire chromosome)	416	EP300 and >200 other genes
2034	ABC	22	22q11.1 - 22q13.33	14805205	49518684	1.44	34713.48	loss (entire chromosome)	412	EP300 and >200 other genes
2124	GCB	22	22q11.1 - 22q13.33	14432516	49581309	1.81	35148.79	loss (entire chromosome)	416	EP300 and >200 other genes
2044	ABC	22	22q11.1 - 22q13.33	14857504	49581309	1.82	34723.81	loss (entire chromosome)	414	EP300 and >200 other genes

^ Two additional cell lines carrying aberrant *EP300* alleles include SUHDL2, which showed a monoallelic gene deletion by FISH analysis using a labeled probe for the *EP300* locus (BAC RP11-1078O11) and a control probe for chromosome 22, and RC-K8 (Garbati et al., Ref. 30 of the manuscript)

* Numbering according to NCBI Build 36.1.

**NCBI RefSeq database.

Abbreviations: ABC, activated B cell type; GCB, germinal center B cell type; GC, germinal center type (based on immunohistochemistry).

Table S5. Oligonucleotides used for genomic amplification of the *CREBBP* gene

Exon	Oligonucleotide name	Position^	Oligonucleotide Sequence (5'-3')
Exon 1	CREBBP-E1F1	13*	GTTGCTGTGGCTGAGATTTG
	CREBBP-E1R1	+463	GACAGTCCTGCGACGAACTT
Exon 2-I	CREBBP-E2F1	+781	ACTTTCCAACTGCCTGACCTG
	CREBBP-E2R1	743*	ACTTTAACCAGACCCACCCAG
Exon 2-II	CREBBP-E2F2	943*	AAACCTGCGTTAGGGTCTCAG
	CREBBP-E2R2	368*	ACAGTGGGAACCTTGTTCCAG
Exon 2-III	CREBBP-E2F3	779*	TTACTATTGAGGAGGCCTGGG
	CREBBP-E2R3	-73	TCATAGAAACGTGGCAGTTGG
Exon 3	CREBBP-E3F	+200	CCTCACCTATCTCTCCTGTTGC
	CREBBP-E3R	-80	AACTGTGTGAGCATTTCCCAG
Exon 4	CREBBP-E4F	+323	TGAACCAGAGAGCTGCTGTAAG
	CREBBP-E4R	-107	TGGTCGGTATTATCCATCAGC
Exon 5	CREBBP-E5F	+202	CTGCCCACTCCCTACCTACTC
	CREBBP-E5R	-240	GCCCTTTATCACCCATCCTTC
Exon 6	CREBBP-E6F	+316	TTGGGTTCCATCACTCCATTC
	CREBBP-E6R	-149	TTCTCCCTAGGCTTGTTTACTGC
Exon 7	CREBBP-E7F	+327	AAGTCAGGAGAGCCAAGTGATG
	CREBBP-E7R	-101	CACCAATTGTTTATATGGTGGC
Exon 8	CREBBP-E8F	+249	GCACGTGACTTGTATAGGCTCC
	CREBBP-E8R	-124	ACATCACTTGGCTCTCCTGAC
Exon 9	CREBBP-E9F	+263	GGAAGTCTCCTTGGTCAGTGG
	CREBBP-E9R	-161	GCCAGGCAGATCTCAAACTG
Exon 10	CREBBP-E10F	+298	ACATCAACAGCTTCTGCAGGG
	CREBBP-E10R	-145	TGGGCTTCTGGTGTATCAGG
Exon 11	CREBBP-E11F	+148	GAAAGGGTTAGAAAGAAATATACAGGG
	CREBBP-E11R	-191	GGACACTGAGTTTCTCTCTTGGG
Exon 12	CREBBP-E12F	+208	CAAGTGACATGAATTCTGCTGC
	CREBBP-E12R	-112	TTCTGTTGCCTGTGCGTTC
Exon 13	CREBBP-E13F	+300	AGACATGAAATGTGCATTCTGG
	CREBBP-E13R	-78	GGATCATTCTGGCTCACCTTG
Exon 14	CREBBP-E14F1	+558	CAAGTGATTCTCCCACCTTGG
	CREBBP-E14R1	-11	ATTTCTGGTAGGGACAGGTGCC
Exon 15	CREBBP-E15F	+311	CCGAACCCACATTCAAACAG
	CREBBP-E15R	-83	ATTGTAGGTTGCATGAGCAGC
Exon 16	CREBBP-E16F	+350	CCTGAGTGTTTCTGCAGGG
	CREBBP-E16R	-83	AGGCTTGTAAGAGTCTTCCCG
Exon 17	CREBBP-E17F	+231	AGAACAATCTTCAAGGCAGGG
	CREBBP-E17R	-93	ATGTCACAGCAGCCAGGATTC
Exon 18	CREBBP-E18F	+328	CCAGTATACAGGCGTGGTCTC
	CREBBP-E18R	-131	TTCTTACTGTTGGGAATGGAAGT
Exon 19	CREBBP-E19F1	-163	TCACATGCTATCCCAAAATGTC
	CREBBP-E19R1	+176	TTCAGGAAAGAAATAATGTACA
Exon 20	CREBBP-E20F	+176	GGCACCGGTACCTTCCTTATAG
	CREBBP-E20R	-149	TTTCATTGTCCTCACTGACCC
Exon 21	CREBBP-E21F	+139	ACCCACAACCCACTCCATAAG
	CREBBP-E21R	-259	GGGCAACAGAGCAAGACC
Exon 22	CREBBP-E22F	+238	AAGCGGACAAACGCTTAGAAC
	CREBBP-E22R	-83	GTGGACGCACACACAGACTTC
Exon 23	CREBBP-E23F	+165	GAACCATGTGTTGAGAGGAACC

	CREBBP-E23R	-178	AAGTTCTAAGCGTTTGTCCGC
Exon 24	CREBBP-E24F1	-132	TTCCGTGTTTGGAGAACATT
	CREBBP-E24R1	+317	TAAAACCCCAGACAGGACAA
Exon 25	CREBBP-E25F	+221	GGACACTTAAGAGCCCTGGTC
	CREBBP-E25R	-225	CATTCACAGAGGTGCAGTTCC
Exon 26	CREBBP-E26F	+268	CACCTGGAAAGAGGAGCTTTG
	CREBBP-E26R	-78	CAGGGTGTTGTTTGTTGCTTG
Exon 27	CREBBP-E27F	+376	CTCCAACTGTGCTGCTCTCAG
	CREBBP-E27R	-96	TCCTGGCTTTAGTCCTTGCTC
Exon 28	CREBBP-E28F	+259	AGGACCTAACAGTCGACACGC
	CREBBP-E28R	-122	CACACATGCATGGGACTCTG
Exon 29	CREBBP-E29F	+268	ACTTCCCTCCCACCACAGAC
	CREBBP-E29R	-111	GTGACCTACTTTGGCCTGAGC
Exon 30	CREBBP-E30F	+389	CAGCCACCATCAGGTACAGAC
	CREBBP-E30R	-107	CTCAGCCACCTGCCTATTCTG
Exon 31-I	CREBBP-E31F12	5756*	CGGAGCTTGTGTTTGATGTTG
	CREBBP-E31R12	-198	GTCCATGATCCCATCTTGTCC
Exon 31-II	CREBBP-E31F11	5965*	AGCCAGCTGGTGACATGC
	CREBBP-E31R11	5581*	CCATCCTGCCAGAAGATGAAG
Exon 31-III	CREBBP-E31F10	6030*	CACCTGGCTGGTAGGCTTC
	CREBBP-E31R10	5731*	TGCCTCAACATCAAACACAAG
Exon 31-IV	CREBBP-E31F9	6230*	CTCACCTGGTTGGGTCGG
	CREBBP-E31R9	5842-*	CAGAGTCTGCCTTCTCCTACCTC
Exon 31-V	CREBBP-E31F8	6395*	GGTGAGATGCTCCTGGGTG
	CREBBP-E31R8	6006*	CACAGGGAAGCCTACCAGC
Exon 31-VI	CREBBP-E31F7	6509*	TTGATGAAAGCTGCCATTAGC
	CREBBP-E31R7	6113*	AGCACCTGTACCGGGTGAAC
Exon 31-VII	CREBBP-E31F6	6639*	CATGGCATTCAGGTTCTGC
	CREBBP-E31R6	6313*	CCTGTGATATCCATGCAGGC
Exon 31-VIII	CREBBP-E31F5	6792*	CTGCCTCCGTAACATTTCTCG
	CREBBP-E31R5	6400*	GCTCTGCAAGACCTGCTGC
Exon 31-IX	CREBBP-E31F4	6913*	GTCCTTGAGGCTGCTGGAAC
	CREBBP-E31R4	6521*	CCAAGTACGTGGCCAATCAG
Exon 31-X	CREBBP-E31F3	7234*	GCACCTGGTTACTAAGGGACG
	CREBBP-E31R3	6714*	CTTGAACATCATGAACCCAGG
Exon 31-XI	CREBBP-E31F2	7406*	CCCAAGTGTCCCTGATCTATG
	CREBBP-E31R2	7060*	AACATCCAGCAAGCCCTG
Exon 31-XII	CREBBP-E31F1	7655*	ATCCACCCTTCCATGGCTC
	CREBBP-E31R1	7269*	GTCCCAGCCTCCACATTCC

^ Numbers indicate the distance from the corresponding exon, except where indicated by an asterisk (primers annealing to cDNA sequences)

* Numbering according to GenBank accession No NM_004380.2

Table S6. Oligonucleotides used for genomic amplification of the *EP300* gene

Exon	Oligonucleotide name	Position^	Oligonucleotide Sequence (5'-3')
Exon 1	EP300-E1F	206*	TTTCTATCGAGTCCGCATCC
	EP300-E1R	+73	AATGGAAGAATAAAGGCGCAC
Exon 2	EP300-E2F	-111	CATGGAGTGAGGTTGGGAAAT
	EP300-E2R	+82	ACAATGTAAGGCAAACCCTCC
Exon 3	EP300-E3F	-151	TCACGTTGCCCAAGCTGTAG
	EP300-E3R	+111	GCTCCTAGTGGGTACAAATCCC
Exon 4	EP300-E4F	-148	TGGTTTATGCATTCCCTGTGT
	EP300-E4R	+130	CATACAGACACCATCACCACAA
Exon 5	EP300-E5F	-201	CTTTGTGCAAATTGCTTACCC
	EP300-E5R	+100	CAACACCACAGGTCCCTCAC
Exon 6	EP300-E6F	-135	TGGAAGACAAGATCCACATACTC
	EP300-E6R	+189	GGAAGATACAAACCAGTGGCA
Exon 7	EP300-E7F	-171	GGTGCTTCAAATCAGCCTTTAC
	EP300-E7R	+106	AAAGACATCCTCAAACCGAGG
Exon 8	EP300-E8F	-136	CCCTGCCTAGCTCCTTAATGC
	EP300-E8R	+192	ATGATGGTGGGAAAGGTTGAG
Exon 9	EP300-E9F	-202	AGAGTCATTTCTTATATTGTGAACGG
	EP300-E9R	+100	CCGTCTGAAGCATATGTACTACTTG
Exon 10	EP300-E10F	-125	TGGCACCAGTTCTTAATGCAG
	EP300-E10R	+203	GTTAAGTACCATGCCCAGACATC
Exon 11	EP300-E11F	-207	CAACAAATCCACTTGGAGGC
	EP300-E11R	+84	TAAAGCGGGTGTTCAGGTAGC
Exon 12	EP300-E12F	-190	AGAATGGTGTGAACCCAGGTG
	EP300-E12R	+200	CACTGACACTCCAGGGACAAG
Exon 13	EP300-E13F	-90	GCCTGACGTTAGGAGCATTTG
	EP300-E13R	+95	ACCCACTATTTGCTGCCACTC
Exon 14	EP300-E14F	-169	TGTCCAAAGATACATGCCCAG
	EP300-E14R	+81	AGAGAATGGAAATGGCCCAG
Exons 15-16	EP300-E15F	-100	GGTGGCTAATTCTGCTATCCTG
	EP300-E16R	+102	CAATAATGGCAACTTCTGAGGC
Exon 17	EP300-E17F	-181	AACAAAGCGGGGCTTAGAAT
	EP300-E17R	+88	TGGCTATACTGTTTGGAATGTGA
Exon 18	EP300-E18F	-124	GGGAATATAGACAGGCCAGAAAC
	EP300-E18R	+91	CAGAAGCAGGATATTCTCTTATCCC
Exon 19	EP300-E19F	-140	TCAGGAACTGAATTAGCCCATC
	EP300-E19R	+86	CATGACTGTATGTGTGCTGGC
Exon 20	EP300-E20F	-72	GCTTCGTTGCTTGGCTTG
	EP300-E20R	+300	CCCATTGCTGACATATTCCC
Exon 21	EP300-E21F	-232	TTCCTGGGTTCTCCATTTCTG
	EP300-E21R	+88	TCATCTGATTGGTCATGCAAAC
Exon 22	EP300-E22F1	-177	CCACTCCAGCCTGTACAACA
	EP300-E22R1	+132	GCCAAACCCAAAGAAAACAA
Exon 23	EP300-E23F	-204	TTTCTTCATGTCTGTTGCTTGA
	EP300-E23R	+299	GCCAATCATTTCCACACCTTA
Exons 24-25	EP300-E24F	-191	GATTAGCATGTTCCCTGCACTC
	EP300-E25R	+114	TTGCATTTCCAAACCAAACAC
Exon 26	EP300-E26F	-106	CAAAGAGCCTGGGAGAGTGAG
	EP300-E26R	+190	GGGCCAACATATCCATTTCTC
Exon 27	EP300-E27F	-94	GGTATCTATATCAACTCCAACTTGTGG
	EP300-E27R	+198	GAATGGCATGAACCCAGG
Exon 28	EP300-E28F	-134	GCTTAGGTATAAAGTCTCTGCCAGC
	EP300-E28R	+178	AAAGTTTAATACCACTGAAACAGAACC
Exon 29	EP300-E29F	-356	ACCAAGGTCTTCTGGCTCC
	EP300-E29R	+373	GCCCAGCGAACAGTCAGT

Exon 30	EP300-E30F	-128	CCATGGTGGGATAATTGCTTG
	EP300-E30R	+74	AAATACGTGGCTGCATGGC
Exon 31-I	EP300-E31F1	-126	GGGCAGAGCTGAAGAGGC
	EP300-E31R1	6069*	CTTGAGTCCTGGGCAAGTAGG
Exon 31-II	EP300-E31F2	5811*	TGCCTAAACATCAAGCAGAAGC
	EP300-E31R2	6546*	AGGCTTGTTGAGACACAGTGC
Exon 31-III	EP300-E31F3	6463*	AACCTTTGAACATGGCTCCAC
	EP300-E31R3	7202*	CATCTGTTGCTGAAGGAGTCG
Exon 31-IV	EP300-E31F4	6918*	ATGCCTTCACAATTCCGAGAC
	EP300-E31R4	7515*	AAAGCATTGAATTCTGGTCCG
Exon 31-V	EP300-E31F5	7272*	CACCTACAAGGCCAGCAGAT
	EP300-E31R5	7778*	CGGCTACTGCACAGTTCTTATG

^ Numbers indicate the distance from the corresponding exon, except where indicated by an asterisk (primers annealing to cDNA sequences)

* Numbering according to GenBank accession No. NM_001429.3