

Published in final edited form as:

Nat Cell Biol. 2017 September ; 19(9): 1093–1104. doi:10.1038/ncb3597.

Early loss of *Crebbp* confers malignant stem cell properties on lymphoid progenitors

Sarah J. Horton^{1,2,3}, George Giotopoulos^{1,2,3}, Haiyang Yun^{1,2,3}, Shabana Vohra^{1,2,3}, Olivia Sheppard^{1,2,3}, Rachael Bashford-Rogers³, Mamunur Rashid⁴, Alexandra Clipson⁸, Wai-In Chan^{1,2,3}, Daniel Sasca^{1,2,3}, Loukia Yiangou¹, Hikari Osaki^{1,2,3}, Faisal Basheer^{1,2,3}, Paolo Gallipoli^{1,2,3}, Natalie Burrows³, Aysegul Erdem^{1,2,3}, Anastasiya Sybirna¹, Sarah Foerster¹, Wanfeng Zhao⁵, Tonci Sustic^{1,2}, Anna Petrunkina Harrison⁶, Elisa Laurenti^{1,2}, Jessica Okosun⁷, Daniel Hodson^{1,2}, Penny Wright⁵, Ken G. Smith³, Patrick Maxwell³, Jude Fitzgibbon⁷, Ming Q. Du⁸, David J. Adams⁴, and Brian J.P Huntly^{1,2,3,9,10}

¹Wellcome Trust-MRC Cambridge Stem Cell Institute, Cambridge, UK

²Department of Haematology, University of Cambridge, Cambridge, UK

³Cambridge Institute for Medical research, Cambridge Biomedical Campus, Hills Road Cambridge, CB2 0XY, UK

⁴Experimental Cancer Genetics, Wellcome Trust Sanger Institute, Hinxton, UK

⁵Dept of Pathology, Cambridge University Hospitals, Hills Road Cambridge, CB2 0QQ, UK

⁶Dept of Medicine, University of Cambridge, Hills Road Cambridge, CB2 0QQ, UK

⁷Barts Cancer Institute, Charterhouse Square, London EC1M 6BQ, UK

⁸Dept of Pathology, University of Cambridge, Hills Road Cambridge, CB2 0QQ, UK

⁹Dept of Haematology, Cambridge University Hospitals, Hills Road Cambridge, CB2 0QQ, UK

Abstract

Loss-of-function mutations of cyclic-AMP response element binding protein, binding protein (CREBBP) are prevalent in lymphoid malignancies. However, the tumour suppressor functions of CREBBP remain unclear. We demonstrate that loss of *Crebbp* in murine haematopoietic stem and

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

¹⁰Correspondence should be addressed to B.J.P.H.

Author Contributions

S.J.H designed and performed experiments, analysed data and wrote the paper. G.G performed FACS, serial replating and transplantation experiments. H.Y performed ChIP-Seq and Q-RT-PCR. S.V performed bioinformatics analyses. O.S maintained the mouse lines and performed experiments. R.B-R performed BCR amplification, sequencing and analysis. M.R performed exome sequencing and bioinformatics analyses. A.C designed and performed the allele specific PCR. D.S performed Western blotting. L.Y performed flow cytometry. H.O, F.B, P.G, A.E, S.F and T.S provided technical assistance. N.B and P.M provided the CD19-Cre mice. A.S optimized immunofluorescence experiments. W.Z performed immunohistochemistry. A.P-H and E.L helped with design and sorting of HSPC from patients. J.O and J.F provided patient samples. D.H and K. S designed experiments, P.W analysed histology and immunohistochemistry. A.C and M.D designed and analysed allele specific PCR. D.A analysed exome sequencing data. B.J.P.H designed and analysed experiments and wrote the paper.

Competing financial interests

The authors declare no competing financial interests.

progenitor cells (HSPC) leads to increased development of B-cell lymphomas. This is preceded by accumulation of hyperproliferative lymphoid progenitors with a defective DNA damage response (DDR) due to a failure to acetylate p53. We identify a pre-malignant lymphoma stem cell population with decreased H3K27ac, which undergoes transcriptional and genetic evolution due to the altered DDR, resulting in lymphomagenesis. Importantly, when *Crebbp* is lost later in lymphopoiesis, cellular abnormalities are lost and tumour generation attenuated. We also document that *CREBBP* mutations may occur in HSPC from patients with *CREBBP*-mutated lymphoma. These data suggest that earlier loss of *Crebbp* is advantageous for lymphoid transformation and inform the cellular origins and subsequent evolution of lymphoid malignancies.

Introduction

The Cyclic-AMP response element binding protein, binding protein CREBBP (also known as CBP and KAT3A) is a general transcriptional co-activator. CREBBP and its paralog EP300 (also known as p300 and KAT3B) control processes during development and homeostasis through binding to multiple protein partners and acetylating lysine residues of histone (including H3K27Ac, H3K18Ac and H3K56Ac) and non-histone substrates¹. Germline loss-of-function mutations of *CREBBP* in the Rubinstein-Taybii cancer predisposition syndrome suggested a tumour suppressor role for CREBBP2, and early mouse models subsequently confirmed this³. More recently, somatic mutations, predominantly hemizygous and affecting the acetyltransferase domain, or deletions of *CREBBP*, have been described across a large number of sporadic tumours including myeloid haematological malignancies^{4,5}. However, *CREBBP* mutations are particularly frequent in both primitive and mature lymphoid malignancies, occurring in around 20% of relapsed B-cell Acute Lymphoblastic Leukaemias (ALL)⁶, 40% of diffuse large B-cell lymphomas (DLBCL)⁷, 60% of follicular lymphomas (FL)⁸ and also in T-ALL⁹ and cutaneous T-cell lymphomas¹⁰. The mutations may occur throughout disease development, with variant allele frequency analysis demonstrating their very early acquisition in FL but longitudinal studies documenting their enrichment in relapsed ALL^{6, 11,12}. Despite this, the full extent of CREBBPs function as a tumour suppressor and the reason for its predilection for the lymphoid lineage remain unanswered questions.

Many cancers are dependent upon a population of stem or initiating cells for their continued growth and relapse, identifying a critical target population for therapeutic eradication¹². However, for mature lymphoproliferative disorders (LPD), such as lymphomas, although malignant stem cell populations are predicted, direct evidence for their existence is controversial¹³ and if present, their identity and provenance are mysterious. Moreover, the molecular and cellular perturbations that direct the evolution of this transformed cell towards a fully blown lymphoma remain entirely unknown. The target cell for transformation in mature lymphoid malignancies had previously been presumed to be a cell with inherent self-renewal and capable of antigenic memory^{14,15}. However, the demonstration of clonal human lymphopoietic reconstitution in murine xenotransplant recipients of haematopoietic stem and progenitor cell (HSPC) populations from chronic lymphocytic leukaemia (CLL) patients¹⁶ and the existence of “driver” mutations, such as *BRAF* and *SF3B1* mutations, in

HSPC from hairy cell leukaemia (HCL) and CLL patients^{17,18} has challenged this hypothesis¹⁹.

In this study, we investigate the tumour suppressor functions of *Crebbp* in isolation and describe murine models with conditional inactivation of *Crebbp* at different stages of lymphopoiesis. Mice with early loss of *Crebbp* within the HSPC compartment demonstrate alterations of transcription, epigenetic regulation and DNA damage response (DDR) and an increased frequency of an aggressive LPD/lymphoma. This lymphoma is preceded by a distinct pre-malignant phase, allowing the interrogation of transcriptional, epigenetic and genetic events occurring during lymphoma evolution. In contrast, loss of *Crebbp* in committed lymphoid cells significantly abrogates the cellular phenotype and markedly reduces tumour development. Finally, we demonstrate the relevance of this for human disease, detecting a *CREBBP* mutation in the HSPC compartment of a patient whose lymphoma carried the same mutation. Taken together, these data have profound implications for the potential cellular origins and subsequent evolution of lymphoid malignancies.

Results

***Crebbp* loss predisposes to an aggressive B-cell malignancy**

Initially, we aged a cohort of mice where excision of *Crebbp* occurs within the HSPC compartment, following pIpC-mediated Mx1-Cre recombinase expression (hereafter Mx-*Crebbp*^{-/-} mice)^{20,21}. Compared with wild-type littermate controls, Mx-*Crebbp*^{-/-} mice displayed a significantly shorter survival (Figure 1a, $p < 0.0001$), with the incidence of haematological malignancies, particularly B-cell malignancies, more than doubled (accounting for 29% of all deaths, Figure 1b and Table S1-3). These mice developed a long-latency aggressive LPD/Lymphoma with predominantly blood and splenic involvement (Figure 1c-d), although ~10% also demonstrated lymphadenopathy. Tumour analysis revealed low and high-grade morphology (Figure 1d), a mature surface phenotype (B220⁺, CD19⁺, sIgM⁺), with aberrant expression of the B1-like markers Mac1 and CD5 (Figure 1e), and reflected aspects of human lymphoid malignancies associated with high frequencies of *CREBBP* mutations. The LPD also efficiently transferred disease to secondary recipients (Figures 1f)

***Crebbp* loss alters transcriptional and epigenetic regulation to change lymphoid progenitor function**

We have previously shown that loss of *Crebbp* in HSPC leads to stem and progenitor alterations, with a significant reduction of mature B and T-cells apparent at later time-points^{21,22}. At earlier time-points (8-10 weeks post-pIpC), Mx-*Crebbp*^{-/-} mice demonstrated a significant increase in lineage negative cells (Lin⁻) within bone marrow (BM) and peripheral blood (PB), with an increase of IL-7receptor α positive (IL-7Ra⁺) progenitors in BM when compared with wild type (WT) controls (Figure 2a-b). Despite this, the frequency of common lymphoid progenitors (CLPs) and all downstream B cell compartments including mature B and T cell populations were not altered at this time point (Figure S1a-d). This IL-7Ra⁺ compartment demonstrated lymphoid and limited myeloid potential in methylcellulose assays (Figures 2c and S1e). *Crebbp*^{-/-} IL-7Ra⁺ cells showed

increased proliferation and clonogenic potential in lymphoid colony assays and enhanced self-renewal when compared with WT controls (Figures 2c and S1e). RNA-Seq analysis of this early IL-7Ra⁺ progenitor population revealed subtle differential gene expression between Mx-*Crebbp*^{-/-} and WT mice (Table S4), with 97 genes including oncogenes *Pbx3*, *Ets2*, *Yes1* and *Yap1* upregulated and 101 genes including the lymphoid transcription factors and tumour suppressor genes *Tcf7* and *Bcl11b*, and the Notch inhibitor *Dtx123* downregulated (Figure 2d). Geneset enrichment analysis (GSEA) demonstrated significant enrichment for genes differentially regulated in *CREBBP* mutant lymphomas²⁴ and also core ES cell, HSC and proliferation signatures (Figures 2e and S1f). Genomewide H3K27ac binding by ChIP-Seq revealed a larger number of differentially bound peaks; 416 peaks were downregulated from WT to *Crebbp*^{-/-} progenitors, although only a small number of these overlapped with the downregulated genes (9/101, 10%) at this early time point (Figure 2f-g). These data link *Crebbp* loss in the HSPC compartment with altered transcriptional and epigenetic regulation, disordered lymphoid differentiation and expansion of a hyperproliferative progenitor population with *in vitro* self-renewal potential.

***Crebbp*^{-/-} progenitors acquire *in vivo* self-renewal and premalignant potential**

Serial analysis of PB in Mx-*Crebbp*^{-/-} mice identified a population of aberrant lymphocytes, with a B220⁺/Mac1^{Int} phenotype, similar to that observed in the overt LPD. This population was detected in the blood of every animal that succumbed to the disease, appearing as early as 10 weeks after pIpC administration and was often detectable many months prior to the development of disease (Figure 3a-b). We hypothesised that this population was a pre-malignant ancestral population from which the overt Lymphoma evolved and tested its stem cell properties *in vivo*. “Pre-malignant” (PM) Mx-*Crebbp*^{-/-} mice were identified by the presence of a B220⁺ Mac1^{Int} population and a normal white cell count (WCC) on PB testing, absence of clinical symptoms and a normal spleen weight upon necropsy. The LSK population was isolated from the BM and the B220⁺/Mac1^{Int} putative PM stem cell population was isolated from the spleen and were transplanted independently across a congenic background (Figure 3c). Remarkably, highly efficient reconstitution (25% of PB) by the B220⁺/Mac1^{Int} population was demonstrated in every recipient animal by 4 weeks (Figure 3d left panel). In addition, this population expanded *in vivo*, and over a highly variable time course recipient animals developed an overt LPD indistinguishable from the disease in primary animals (Figure 3d right panel). Secondary transplantations performed using tumour cells from two recipient animals demonstrated a highly significant reduction in disease latency (median 48 vs 132 days, p=0.0002) in keeping with evolution of the B220⁺/Mac1^{Int} population to frankly malignant cells (Figure 3e).

Transcriptional, epigenetic and genetic alterations in *Crebbp*^{-/-} tumours mimic changes in human lymphomas

RNA-seq of WT B220⁺ lymphocytes, and lymphomas from Mx-*Crebbp*^{-/-} and WT mice revealed marked alterations of gene expression (Tables S5-6), with 3546 and 3071 genes differentially expressed between *Crebbp*^{-/-} lymphomas and WT B220⁺ cells (Figure 4a-b). Genes upregulated in frank LPD included the lymphoma oncogenes and self-renewal genes *Ezh225*, *Myc*, *Myb* and *Gfi-1b26*, with downregulated genes including the lymphoma tumour suppressors *Pten27*, *Id328* and *Irf829*, master lymphoid transcription factors *Ebf18*

and *Pax530* and class II MHC genes *H2-Ab1*, *H2-Ob* and *H2-Eb2*, previously demonstrated to be downregulated in *CREBBP* mutated lymphomas⁴¹. Genotype specific changes between *Crebbp*^{-/-} and WT lymphomas were also evident, including downregulation of *Pten* and upregulation of *Ccnd2*. Tumours were demonstrated to be clonal by IgH analysis (Figure S2). GSEA analysis showed enrichment for the *CREBBP* mutant signature (Figure 4c). *CREBBP* is known to acetylate multiple lysine residues on histones, with H3K27, H3K18 and H3K56 acetylation linked to transcription and DNA damage. We detected global loss of H3K27Ac in *Crebbp*^{-/-} LPD compared to WT B220⁺ B-cells (Figure S3a). We next performed genome-scale analysis of these modifications between *Crebbp*^{-/-} LPD and WT B-Cells, as well as *Crebbp* binding in WT B-cells, by ChIP-Seq. Although some loss of acetylation was demonstrated for all marks, this was most obvious for H3K27Ac, where 3360 peaks showed decreased acetylation in *Crebbp*^{-/-} lymphomas. Furthermore, H3K27Ac correlated best with *Crebbp* binding (Figure 4d-e) and with genes downregulated, where 1308 of the differential H3K27Ac peaks was linked to a downregulated gene (Figures 4e-f, S3b and Tables S7-12). The majority of differential peaks were in enhancers rather than promoters, with both super-enhancers and normal enhancers dysregulated (Figure 4g-h). *De novo* motif analysis of differentially bound peaks suggested altered regulation of transcriptional networks, including those regulated by *Ets*, *Irf*, *Ebf* and *Mef2b* transcription factors (Figure 4i).

We performed exome-sequencing in 14 *Crebbp*^{-/-} and all (n=3) WT LPD (Table S13-20). Single nucleotide variants and insertion/deletion events included recurrent alterations in the *Card11* and *Gna13* genes and alteration of the *Hist1h1d* gene in a single case, all genes mutated in human lymphoma^{8, 29} (Figures 5a and S4a). Copy number analysis demonstrated recurrent losses and gains, including loss of the *Cdkn2a* and *2b* locus, a common alteration in human lymphomas³¹. We also noted mutations in Ras- and Jak-Stat intermediates, pathways known to be activated in human lymphomas^{7,8,32,33}, and could demonstrate activation of these pathways in our *Crebbp*^{-/-} and WT LPD (Figure S4b). Analysis of mutation patterns revealed predominantly C > T transitions and C > A transversions. Cosine analysis correlated this single signature best with human cancer signature 134 (Figures 5b and S4c-d), which reflect the effects of ageing. Taken together, our data demonstrate transcriptional, epigenetic and genetic alterations within the *Crebbp*^{-/-} lymphomas that overlap with human disease.

***Crebbp* modulates DNA damage and protects genomic integrity in lymphoid progenitors**

Crebbp has been associated with DNA repair and the cellular response to DNA damage in a number of cell types³⁵. This was also suggested by GSEA analysis of differential gene expression in *Crebbp*^{-/-} progenitors (Figure 5c) and we hypothesised that this might contribute to tumour evolution within the *Crebbp*^{-/-} LPD. Therefore, we measured H2AX phosphorylation (H2AX γ foci), a surrogate for DNA double stranded breaks, and the functional consequences of the DNA damage response (DDR) in IL-7Ra⁺ progenitors. At baseline we found no difference in H2AX γ foci between Mx-*Crebbp*^{-/-} and WT mice. However, when dynamically induced with hydroxyurea (HU) treatment, DNA damage was significantly increased in *Crebbp*^{-/-} mice (Figure 5d-e, p=0.047). However, surprisingly, given the degree of DNA damage, there was no expected increase in apoptosis in *Crebbp*^{-/-}

progenitors (Figure 5f). Similarly, in colony forming assays clonogenic potential and proliferation were almost fully preserved (Figure 5g). The DDR is orchestrated by p53, which in turn requires acetylation by p300 or Crebbp for its full transcriptional output^{36,37}. While WT mice demonstrated a brisk acetylation of p53 following DNA damage induction, none was evident for mice lacking *Crebbp* (Figure 5h). In addition, reduced induction of p53 target genes following DNA damage further demonstrated defective p53 activity (Figure 5i). Collectively, these data suggest that loss of *Crebbp* abrogates optimal cellular response to DNA damage, with these subtle changes creating a permissive state for the retention and accumulation of subsequent mutations, facilitating transformation.

Evolution at the clonal idiotype, transcriptional and genetic levels in *Crebbp*^{-/-} lymphomas

The presence and accessibility of the B220⁺/Mac1^{Int} premalignant stem cell population provided us with a unique opportunity to study the evolution of a mature lymphoid malignancy, as did comparisons between the premalignant donor cells and their lymphomatous progeny harvested from the transplant recipients in Figure 3d. Using a next generation sequencing (NGS) method we have developed³⁸, we could demonstrate an early reduction in IgH diversity following *Crebbp* loss (Figure 6a). A further contraction of the IgH repertoire was evident upon longitudinal analysis as the LPD/lymphoma developed in individual primary mice *in vivo*, as well as in transplant recipients of premalignant cells (Figure 6a-c). Of note, in a small number of cases two IgH rearrangements were evident, and sequencing of RNA could demonstrate expression of both rearrangements in keeping with the transformation of multiple independent clones (Figure 6b).

Comparing transcriptional changes across WT B cells, B220⁺/Mac1^{Int} PM stem cells and *Crebbp*^{-/-} lymphoma cells demonstrated significant transcriptional evolution (Figures 6d-e and S5a-b). Only 469 and 587 genes were differentially regulated up and down respectively at this earlier time point between normal and PM cells (table S21), compared to 3546 and 3071 genes respectively in the lymphomas (table S22). These already included upregulation of *Meis3* and multiple *Hoxa* and *b* paralogues as well as downregulation of multiple linker histone genes that are mutated in FL, including *Hist1h1b* and *Hist1h1d*. In addition, differential expression of 205/469 (44%) of upregulated genes and 88/587 (15%) of downregulated genes was also evident in the later lymphoma samples (Fig 6e).

Exome sequencing performed across lymphoma evolution in 2 *Crebbp*^{-/-} LSK HSPC and 8 premalignant B220⁺/Mac1^{Int} stem cells, also demonstrated evidence of genetic evolution when compared to both *Crebbp*^{-/-} and WT lymphomas (Figure 6f). This demonstrated a sequential increase in the average number of total mutations within these populations ($p=0.001$ between the PM and the LPD), but no difference between *Crebbp*^{-/-} and WT LPD. Exome data was available for the premalignant population transplanted, as well as 6 of the 7 tumours that developed in recipient mice in Figure 3e. Comparisons of SNV in the PM and subsequent tumours demonstrated significant divergent evolution; the transplanted PM cells only contained a single SNV (*Cbx2*), however recipient tumours demonstrated variable numbers of SNV (25-47, Table S13), with only 3 SNV common across all cases (*Cbx2*, *Celf1* and *Ceacam11*, Figure 6g).

Loss of *Crebbp* in committed lymphoid cells attenuates disease generation and *CREBBP* mutations are present in patient HSPC

The cellular phenotype observed upon *Crebbp* deletion predominantly reads out in earlier lymphoid compartments. To assess whether the tumour suppressor functions of *Crebbp* were necessary at specific stages of lymphoid development and to identify the initial cell of transformation in our model, we generated a second mouse strain where *Crebbp* is deleted following lymphoid commitment, using a Cd19-Cre39 (hereafter Cd19-*Crebbp*^{-/-}). Although there was no difference in the efficiency of recombination, we could find no differences in the size, composition or function of the HSPC compartment in Cd19-*Crebbp*^{-/-} mice, in marked contrast to early loss of *Crebbp* (Figures 7a and S6a-c). Similarly, there was no difference in basal or induced DNA damage following HU treatment in the early compartment (Figure 7b). Only the pro-B population was decreased when later lymphoid populations were examined (Figure S6d-e). Although the B220⁺ Mac1^{Int} population could be detected in Cd19-*Crebbp*^{-/-} mice and occasional mice did develop B-cell LPD, no difference in survival was seen between Cd19-*Crebbp*^{-/-} and WT mice and the number of B-Cell LPD, and other malignancies, were markedly reduced (Figures 7c-e, S6e and Table S23-25). To determine the relevance of these findings for human disease, we screened the HSPC compartment of 3 patients with *CREBBP*-mutated lymphoma (1 DLBCL and 2 FL) who lacked BM lymphoma involvement. HSPC were flow sorted and grown as colonies in methycellulose (MC) under conditions that only permitted myeloid cell outgrowth and DNA harvested (Figure 7f). Sensitive (~1 mutant in 10⁵ wild-type cells) bespoke forward and reverse allele-specific PCR assays were developed to screen for sentinel *CREBBP* mutations (Figure S7a-b). Although no evidence for a *CREBBP* mutation could be found in 2 patients, critically, in the remaining patient, the pathogenetic mutation (a 3 base pair deletion leading to pS1608delS in the HAT-domain)⁷ identified in the tumour was also detected in the CD19⁻/CD34⁺/CD38⁺ HSPC compartment and confirmed by Sanger sequencing (Figure 7g and S7c-d). These data demonstrate that early loss of *Crebbp* is advantageous for tumour formation in our model and that *CREBBP* mutations may occur in the HSPC compartment of patients with lymphoma.

Discussion

CREBBP is mutated or deleted across a wide range of human cancers, identifying it as a generic tumour suppressor. In this report, we describe an increased frequency of a number of malignancies in Mx-*Crebbp*^{-/-} mice, and further address the most common phenotype, that of mature B-cell malignancies. We mechanistically demonstrate specific loss of H3K27Ac that occurs early in lymphoid differentiation and progresses during the evolution of the lymphomas. Although, these changes correlate poorly with downregulated genes in early pre-malignant progenitors, they and many more enhancers that lose H3K27Ac significantly correlate with later downregulated gene expression programmes in overt lymphomas (Figure S5b). Of interest, although the acetylation and transcription alterations we report in our *Crebbp*^{-/-} lymphomas mirror recent reports of murine models of follicular lymphomas where *Crebbp* loss or knockdown is combined with *Bcl2*-overexpression^{40,41}, differences are also apparent. In particular, we did not see any enrichment for *Bcl6* targets by GSEA or *de novo* motif analysis at *Crebbp* binding sites or sites with differential H3K27Ac, raising

the possibility that loss of *Crebbp* collaborates with Bcl6 only in the context of *Bcl2* overexpression.

We demonstrate that in the *Mx1-Crebbp*^{-/-} mice lymphoid progenitor populations accumulate, are hyperproliferative and demonstrate altered cellular responses to DNA damage upon *Crebbp* deletion, which we link to the loss of post-translational activation of p53. This constellation of alterations cumulatively creates a “perfect storm” for the acquisition (hyperproliferation) and retention (inefficient DDR) of mutations and we postulate that the “primed” epigenetic state may also favour subsequent transformation by these mutations (See models in Figure 8). CREBBP is a global transcriptional coactivator⁴², that interacts with multiple master regulator transcription factors, therefore it is likely that aspects of the tumour suppressor phenotype we describe in the context of lymphomagenesis also occur across a range of other tumours. However, lymphocytes are highly proliferative and undergo additional, significant endogenous physiological DNA damage, through antigen receptor rearrangement and somatic hypermutation¹⁴. We suggest that our data defining a defective DDR provide a plausible explanation for the particularly high rates of *CREBBP* mutations in lymphoid malignancies.

Our model fits with the premise that loss of *Crebbp* predisposes to the subsequent development of mature lymphoid malignancies. This hypothesis would be in keeping with the observation that mutation or loss of *CREBBP* occurs across multiple lymphoma phenotypes. This would also fit with our description of *Crebbp* loss leading to the generation of a premalignant stem and progenitor population, with our model being, to our knowledge, the only one to demonstrate the existence of such a population for a mature lymphoid malignancy. Our study also sheds mechanistic light on the target of initial transformation in mature lymphoid malignancies. Early loss of *Crebbp* in the HSPC compartment resulted in a marked cellular phenotype and the development of an aggressive LPD. However, loss in a later, fully lymphoid restricted compartment, almost completely abrogated these abnormalities significantly attenuating tumour formation. Moreover, we demonstrate that *CREBBP* mutations may occur in rigorously isolated HSPC from patients with *CREBBP*-mutated lymphomas. A requirement for early loss of CREBBP function would be in keeping with phylogenetic sequencing studies in human follicular lymphoma, which show that *CREBBP* mutation occurs in the earliest inferable lymphoid progenitor^{11,12}, with the description of identical *EP300* mutations in bone-marrow transplant donor-recipient pairs that both went on to develop lymphomas⁴³ and also with the identification that *CREBBP* mutations can rarely occur in HSPC to cause age-related clonal haematopoiesis⁴⁴. Furthermore, mutations of other clonal haematopoiesis genes *TET2* and *DNMT3A* initially occurring in the HSPC compartment have been shown to generate predominantly T-cell lymphomas in some patients^{45,46}. Evidence from other mouse models would also be in keeping with events in earlier progenitor compartments contributing to the development of mature B-cell malignancies^{47–49}. Moreover, it is increasingly apparent that for some human LPD such as CLL and HCL, driver mutations such as *BRAF*^{V600E} occur within phenotypic HSPC^{16–18}. Taken together, our murine and human data are therefore compatible with *CREBBP* mutations occurring throughout HSPC and later lymphoid ontogeny. Although the clinical relevance of this is currently unclear, it merits further investigation and it will be of interest to determine whether the first cell type in which the *CREBBP* mutation occurs as an

initiating event has a bearing on the phenotype or clinical outcome of the lymphoma. Furthermore, our study models *Crebbp* loss, and it remains to be seen if altered rather than reduced function, perhaps related to specific HAT domain mutants, contribute to lymphoma formation. Finally, our findings also provide a mechanistic explanation as to why acquisition of the same mutation earlier rather than later in differentiation would be advantageous for lymphoma development.

Methods

Mice

C57Bl6 Mx-*Crebbp*^{-/-} (*Crebbp* FL/FL; Mx-Cre) conditional knockout mice and the induction of Cre-mediated recombination in these animals have been described previously²¹. To induce Cre-mediated recombination, 6- to 10-week-old Mx-*Crebbp*^{-/-} and WT (*Crebbp* FL/FL; WT) mice were administered 5 doses of poly(I)-poly(C) (pIpC) (300 µg per dose; Sigma) by intraperitoneal injection every other day over 10 days. All experiments were performed using littermates that were not selected for gender but age-matched in terms of date of birth and time of pIpC administration. CD19-*Crebbp*^{-/-} (*Crebbp* FL/FL; CD19-Cre) conditional knockout mice were generated by breeding *Crebbp* FL/FL mice with CD19^{Cre/+} transgenic mice, both on a C57BL/6 background. Deletion efficiency was determined in both cohorts by PCR of genomic DNA extracted from blood or spleen. Peripheral blood was collected and counted as described previously²¹. All mice were housed in an SPF animal facility and experiments were conducted under UK Home Office regulations. This research has been regulated under the Animals (Scientific Procedures) Act 1986 Amendment Regulations 2012 following ethical review by the University of Cambridge Animal Welfare and Ethical Review Body (AWERB). Power calculations were used to generate sample sizes. Randomisation was genetic and the investigators were not blinded to group allocations. However, the animal technicians who delivered care and decided upon which animals needed to be sacrificed were blinded.

Flow cytometry analysis

Single cell suspensions of bone marrow (BM) cells and splenocytes were prepared as described previously⁵⁰. All staining was performed in Dulbecco's phosphate-buffered saline (DPBS; Gibco Life Technologies) supplemented with 2% fetal bovine serum (FBS; Sigma) on ice. Cells were blocked using FcR blocking reagent (Miltenyi Biotech, 1:20) and stained with combinations of the antibodies listed in Table S27. Lymphoid progenitors were analysed using APC-mouse lineage antibody cocktail (558074, BD Biosciences), FITC anti-CD127 (IL7-Ra), PE-Cy7 anti-c-Kit and PB anti- Sca1 (Table S27). Dead cells were excluded by gating on 7AAD (420403, Biolegend) negative cells. Fluorescence-minus-one controls were used to set appropriate gates. Flow cytometry analysis were performed on a CyAn ADP flow cytometer (Dako) or an LSR Fortessa cell analyser (BD Biosciences) and data were analysed with FlowJo software v 10 (Tree Star).

Histopathology and immunohistochemistry

Tissues were fixed, embedded and sectioned as described previously⁵⁰. Immunohistochemistry was performed using an autoimmunostainer with cover tile

technology (Bond-III system, Leica Biosystems). The heat induced antigen retrieval was performed using Bond Epitope Retrieval solution 2 (Leca, Biosystem) for 30 minutes. Staining was performed using anti-Ki-67 (abcam ab16667, clone SP6, 1:50) followed by biotin-SP-conjugated donkey anti-Rabbit IgG (711-066-152; Jackson ImmunoResearch Laboratories, Inc) and Vectastain elite ABC kit (Vector labs, Pk-6100). The Bond™ Polymer Refine Detection kit (Leica, Biosystems) was used for visualisation.

Isolation of murine cell populations by flow sorting

$Lin^- CD45^+ IL7Ra^+$ populations were isolated from the BM of *Mx-Crebbp^{-/-}* and WT mice 8-10 weeks after poly(I) · poly (C) treatment or from *CD19-Crebbp^{-/-}* and WT mice 8-10 weeks after birth. Whole skeletons (except skull and ribs) were pooled and crushed from 3 animals of each genotype and resuspended in DPBS / 2% FCS. Cells were filtered through a 70- μ m cell strainer (BD Biosciences) and red blood cells (RBC) were lysed using RBC lysis buffer (5 PRIME). Lineage-positive cells were depleted prior to sorting using the Easysep mouse hematopoietic progenitor cell enrichment kit (19756, Stem Cell Technologies). Lineage-depleted cells were stained with V500-streptavidin, CD45.2-FITC and CD127-PE-Cy7. CD45.2 was used to gate hematopoietic cells and 7AAD was used to exclude dead cells. The $Lin^- c-Kit^+ Sca1^+$ (LSK) and $B220^+ Mac1^{int}$ populations from the BM and spleens respectively of (pre)-malignant mice were isolated as follows. Lineage-depleted cells were prepared from the whole skeleton of individual pre-malignant mice as previously described. Cells were stained with V500-streptavidin, CD45.2-FITC, c-Kit-PE-Cy7 and Sca1-PB (Table S27). 7AAD was used to exclude dead cells. Total splenocytes were blocked and stained with B220-APC and CD11b-FITC. DAPI was used to exclude dead cells. All sorting was performed using an Influx cell sorter (BD Biosciences).

Isolation of human cell populations by flow sorting

Bone marrow samples from 3 patients with documented *CREBBP* mutations were thawed in DPBS/ 2% FCS and stained with CD34-APC, CD38-PE-Cy7 and CD19-FITC (Table S27). Dead cells were excluded by gating on Zombie Aqua (423101, Biolegend) negative cells. Fluorescence-minus-one controls were used to set appropriate gates. $CD34^+ CD38^- CD19^-$ and $CD34^+ CD38^+ CD19^-$ cells were sorted using a FACS Aria III (BD Biosciences) and data analysed with FlowJo software (Tree Star). Cells were cultured in Methocult H4531 methylcellulose medium (StemCell Technologies) for 7 days and DNA extracted from the colonies. Informed consent was obtained from all patients and approved (UK IRAS, REC reference: 10/H0704/65, protocol number: 006650QM).

Two allele-specific PCRs were performed. Reaction 1; allele-specific forward primer 5'-CACTGGGAGTTCTCCTTGCG-3' and reverse primer 5'-ACAGGATGCTTCGTCAGACC-3', reaction 2; forward primer 5'-GACTAAAGCCCCCTCCTCTCC-3' and allele-specific reverse primer 5'-GAGCGGCGCAAGGAGAAC-3'. PCR reactions contained 2.5 ng DNA, 400 nM of each primer and AmpliTaq Gold® 360 DNA Polymerase. PCR conditions were as follows: a pre-denaturation of 95 °C for 10 min, 35 cycles of amplification for reaction 1 and 38 cycles for reaction 2 (95 °C for 20 s, 66 °C for 30 s and 72 °C for 30 s) and a final extension at 72 °C

for 5 min. PCR products were separated on 10 % polyacrylamide gel by electrophoresis and confirmed by Sanger sequencing analysis.

Transplantation assays

One million unfractionated (or B220⁺ Mac1^{int} sorted) splenocytes isolated from individual malignant mice were injected intravenously into sub-lethally (550 rads) or lethally irradiated (550 rads; 2 doses) C57BL/6 recipients. In other experiments 1 x 10⁶ B220⁺ Mac1^{int} sorted splenocytes isolated from individual pre-malignant mice or 8 x 10³ – 2 x 10⁴ LSK cells were injected intravenously into lethally irradiated C57BL/6 congenic recipients (CD45.1). When recipients were lethally irradiated 2 x 10⁵- 1 x 10⁶ unfractionated CD45.1⁺ helper BM cells were mixed with the donor cells prior to injection. PB chimerism was monitored monthly by flow cytometry analysis.

Serial replating assays

Lin⁻, CD45⁺, IL7Ra⁺ cells were plated in MethoCult M3630 methylcellulose medium (StemCell technologies) supplemented with 100ng/ml stem cell factor (SCF) and 10ng/ml Flt3Ligand (Flt3L; Peprotech) to assess lymphoid potential or Methocult M3434 methylcellulose medium (StemCell technologies) to assess myeloid potential. Cells were plated in duplicate and after 7 days colonies were scored, pooled and identical numbers of cells were re-plated under the same conditions.

RNA-Seq

Total RNA was extracted from Lin⁻ CD45⁺ IL7Ra⁺ cells isolated from Mx-*Crebbp*^{-/-} and WT mice 8-10 weeks after poly(I) · poly (C) treatment using an RNeasy Plus Micro kit (Qiagen). In other experiments RNA was extracted from unfractionated splenocytes of Mx-*Crebbp*^{-/-} malignant mice (in which >80% of the cells were B220⁺) or sorted B220⁺ cells isolated from the splenocytes of WT mice using an RNeasy Plus mini kit (Qiagen). RNA quantity and quality was analysed by qubit and bioanalyzer. Ribosomal RNA (rRNA) was depleted using a ribozero kit (Illumina) and purified with Qiagen RNA columns. For the WT vs malignant LPD comparison approximately 40-95ng of depleted rRNA was used for library preparation using a Rapid Directional RNA-seq Kit (5138-08 NEXTflex). Library quality was checked and 20nm of each barcoded library was pooled together and sequenced at the CRUK Cambridge Institute genomics core (<http://www.cruk.cam.ac.uk/core-facilities/genomics-core>).

For the lymphoid progenitor comparison and pre-malignant (B220⁺ Mac1⁺) versus WT B220⁺ comparison, RNA libraries were generated from sorted cells using the Ovation Mouse RNA-Seq System (0348-32, NuGEN). Ribosomal RNA was then depleted with Insert Dependent Adapter Cleavage (InDA-C technology; NuGEN). Library quality was checked and then libraries were pooled in equimolar amounts prior to sequencing. All experiments were performed in duplicate on biologically independent samples.

RNA-Seq Data Analysis

RNA-Seq reads were aligned to mouse genome (UCSC mm10) using RNA-seq aligner STAR (version 2.4.0)51 and the number of uniquely mapping reads were estimated using

HTSeq-Count52 and normalized as counts per million. Gene expression read counts were analysed to identify differentially expressed genes using DESeq2 (version 1.12.4)53 after filtering genes with low read count. Features with adjusted p value <0.1 and absolute 1 logarithmic base 2 fold change >1 were considered as having significant altered expression. Differentially expressed genes were analysed for gene set enrichment using GSEA against a collection of gene sets and published information54.

ChIP-seq

Lin⁻ CD45⁺ IL7Ra⁺ cells (lymphoid progenitors) were isolated from *Mx-Crebbp*^{-/-} and WT mice 8-10 weeks after pIpC treatment. In other experiments unfractionated splenocytes of *Mx-Crebbp*^{-/-} malignant mice (in which >80% of the cells were B220⁺) or sorted B220⁺ cells isolated from the splenocytes of WT mice were used. Chromatin immunoprecipitation was performed with iDeal ChIP-seq kits for Histones and Transcription Factors (Diagenode). In brief, 10⁶ or 5x10⁶ cells were crosslinked with 1% formaldehyde (Thermo Scientific) for 5 minutes (H3K27ac) or 15 minutes (Crebbp). Fixation was stopped using Glycine (Fisher Scientific) and cells were washed twice with ice-cold PBS. Lysis buffer iL1 (or iL1b) and iL2 were used to prepare the nuclei pellets. Chromatin was sheared in Shearing buffer iS1 using the Bioruptor Plus (Diagenode) for 2-2.5 runs of 10 cycles (each cycle; 30 seconds “ON” and 45 seconds “OFF”) at high power setting. Immunoprecipitation mix was prepared by the addition of ChIP buffer iC1 (or iC1b), BSA, protease inhibitor cocktails, pre-washed protein A-coated magnetic beads, antibody (1 µg for anti-H3K27ac antibody, ab4729, Abcam; 1 µg for anti-H3K56ac antibody, ab76307, Abcam; 1 µg for anti-H3K18ac antibody, ab1191, Abcam; 4 µg for anti-Crebbp antibody, sc-369, Santa Cruz) to the sheared chromatin and incubated at 4 °C overnight with rotation. Anti-Crebbp antibody was pre-incubated in the mix without sheared chromatin at 4 °C for 3-3.5 hours prior to overnight incubation. 10% of the sheared chromatins were kept aside as input samples. The next day, sequential washes of the chromatin-antibody-beads were performed with ice-cold Wash buffer iW1, iW2, iW3 and iW4 and the beads were captured on a magnetic rack. Elution buffer iE1 and iE2 were added to the beads and incubated at 65 °C for 4 hours to elute and decrosslink the chromatin DNA. Finally, DNA was purified using QIAquick PCR Purification Kit (Qiagen) and eluted in 30 µl EB buffer. ChIP-seq library preparation of ChIP DNA or input DNA was performed using TruSeq ChIP Sample Prep Kit (Illumina). Library DNA was quantified using KAPA Library Quantification Kit (Complete Kit (Universal) for Illumina platforms) (Kapa Biosystems) and library average size was determined by Agilent DNA 1000 Kit (Agilent Technologies) run on a 2100 Bioanalyser System (Agilent Technologies). Libraries were pooled for Single-Read 50bp sequencing on an Illumina HiSeq 2500 or 4000 platforms. Experiments were performed in duplicate on biologically independent samples apart from Crebbp in WT cells and H3K27Ac in IL7Ra⁺ cells which were performed once.

ChIP-Seq Data Analysis

Sequences were aligned to the mouse reference genome (assembly mm10, Genome Reference Consortium GRCm38) using Bowtie software (version 2.1.0)55 allowing a maximum mismatch of 2. Duplicates were removed using PICARD tools. The identification of ChIP-seq enriched regions (peaks) was performed using MACS (version 2.0.10)56 with a

p-value cut-off of 0.005 and the read counts were normalized against input sample. Differential binding analysis was performed using the DiffBind package (version 2.0.10)⁵⁷.

Sample preparation for Exome Sequencing

LSK cells isolated from the BM of pre-malignant mice were expanded in MethoCult M3434 methylcellulose medium (StemCell technologies). After 7 days colonies were pooled and genomic DNA was extracted using a DNeasy blood and tissue kit (Qiagen). DNA was extracted from B220⁺ Mac1^{int} cells sorted from the spleens of pre-malignant mice using the same kit. Genomic DNA was isolated from unfractionated BM or spleen cells from malignant mice and tail tissue using a Puregene Core kit A (Qiagen).

Sequencing Metric and Somatic Variant Calling Section

Whole exome sequencing was performed using the Agilent whole exome capture kit (SureSelect Mouse All Exon, http://www.genomics.agilent.com/article.jsp?pageId=3102&_requestid=112619). Captured material was indexed and sequenced on the Illumina platform at the Wellcome Trust Sanger Institute. Raw pair end sequencing reads were aligned with BWA-mem to the GRCm38 mouse reference genome.

Aligned DNA sequence data from paired tumour/normal DNA samples was presented to the Cake pipeline, using the somatic variant callers Bambino, Mutect, SAMTools Mpileup, Somaticsniper and VarScan2. As described previously⁵⁸ we used a somatic caller merging approach to identify somatic variants selecting only those detected by four or more of these algorithms for further analysis. Five (5) of 27 tumours had matched germline DNA, and for the remaining tumours we used DNA from mice with the same background. Data was further filtered using a pool of germline variants generated by combining the calls from the 5 available matched normal samples. These calls were further refined by removing germline variants identified inhouse by the Wellcome Trust Sanger Institute mouse genome Project.

Mutation Signature Analysis

The mutational catalogue of somatic variants was interpreted using EMu, a probabilistic algorithm that infers the number of mutational processes operative in a tumour, and their individual signatures⁵⁹. All somatic point mutations (including silent mutations) were mapped to the 96 possible trinucleotide combinations taking into account the opportunity of each mutation occurring in the context of each trinucleotide type within the mouse exome. We identified one statistically significant mutation signature; xC:G>T:Ax.

We also compared our catalog of mouse mutation signatures against human cancer signatures³⁴, after accounting for mutational opportunity. A cosine similarity analysis reveals that the the xC:G>T:Ax has strongest correlation with human cancer signature 1.

BCR amplification and sequencing

Genomic DNA was extracted from RBC lysed peripheral blood samples using a KAPA Express Extract kit (KAPA Biosystems). DNA was extracted from malignant tissues as described previously. PCR amplification of ~10ng of DNA samples was performed with the JH reverse primers and the forward primer set pools (group 1 or group 2, 0.25 µM each),

using 0.5µl Phusion[®] High-Fidelity DNA Polymerase (Finnzymes), 1 µl dNTPs (0.25 mM), 1 µl DTT (0.25 mM), per 50 µl reaction. The following PCR program was used: 3 minutes at 94°C, 40 cycles of 30 seconds at 94°C, 15 seconds at 60°C and 30 seconds minute at 72 °C, with a final extension cycle of 7 minutes at 72 °C on an MJ Thermocycler. Group 1 and group 2 primed products were pooled together for each sample. DNA was purified using QIAquick PCR Purification Kit (Qiagen) and eluted in 30 µl EB buffer. For each sample 1 ug DNA was used to prepare the sequencing library using TruSeq DNA PCR-Free Library Prep kit (Illumina) (starting from end-repairing onwards). Library DNA was quantified and multiplexed for Paired-End 300bp sequencing on Illumina MiSeq platform. MiSeq reads were filtered for base quality (median >32) using QUASR60, and paired-end reads merged if they contained identical overlapping regions of >65bp, or otherwise discarded. Non-immunoglobulin sequences were removed and only reads with significant similarity to reference IgHV genes in the IMGT database⁶¹ by BLAST62 were retained (<1x10⁻¹⁰ E-value). Primer sequences were trimmed from reads, and sequences retained for analysis only if both forward and reverse primer sequences were identified and sequence lengths were greater than 240bp for MiSeq. The network generation algorithm and network properties were calculated as previously described³⁸: each vertex represents a unique sequence, where relative vertex size is proportional to the number of identical reads. Edges join vertices that differ by single nucleotide non-indel differences and clusters are collections of related, connected vertices. Mutational analysis was performed using IMGT/V-QUEST⁶³.

Cell culture

No cell lines were used in this study. Lin⁻ CD45⁺ IL7Ra⁺ cells isolated from Mx-*Crebbp*^{-/-} and WT mice 8-10 weeks after pIpC treatment or from Cd19-Cre and WT mice 8-10 weeks after birth were cultured overnight in Iscove's Modified Dulbecco's Medium supplemented with 5% FBS, 1% penicillin/streptomycin, 0.1 mM non-essential amino acids, 1 mM sodium pyruvate, 2 mM glutamine and 50 µM 2-mercaptoethanol (all from Sigma) and 100 ng/ml SCF, 20 ng/ml Interleukin-7 (IL-7) and 10 ng/ml Flt3L (peprotech). In some experiments cells were cultured overnight with or without a low dose (100 µM) of Hydroxyurea (HU; Sigma). For immunoblotting experiments cells were counted the following morning and equal cell densities were cultured with 5 µM Trichostatin A (TSA; Sigma) with or without 4 mM HU for 1 hour. For Q-RT-PCR experiments cells were cultured with 4 mM HU for 4 hours.

Immunofluorescence

A maximum of 5 x10⁴ cells were adhered to poly-L-lysine (Sigma) coated slides. Soluble and cytoskeletal proteins were removed using a modified pre-extraction protocol⁶⁴. Slides were immersed in cytoskeleton buffer (10 mM PIPES (pH 6.8), 100 mM NaCl, 300 mM sucrose, 3 mM MgCl₂, 1mM EGTA and 0.5% Triton X-100) for 5 min on ice and then cytoskeleton stripping buffer (10 mM Tris-HCL (pH 7.4), 10 mM NaCl, 3 mM MgCl₂, 1% Tween-40 and 0.5% sodium deoxycholate) for 5 min on ice. After washing with PBS on ice for 5 min, cells were fixed with 4% paraformaldehyde (PFA) for 10 min at room temperature (RT), permeabilised in 0.15% Triton X-100 / PBS for 5 min at RT and blocked in 1% Bovine Serum Albumin (BSA) / PBS overnight at 4°C. Cells were stained with anti-phospho-Histone H2AX (Ser139, Millipore, 05-636, clone JBM301, 1:500) for 1 hour at 37°C and

then washed with PBS 3 x 5min at RT. Cells were stained with Alexa Fluor 488 goat anti-mouse (ThermoFisher Scientific, R37120, 1:500) for 1 hour at 37°C, washed with PBS 3 x 5min at RT and mounted with ProLong Gold mounting medium with DAPI (Invitrogen). A minimum of 100 cells were scored using a Zeiss AxioSkop 2 microscope and the number of phospho-H2AX foci per cell were recorded from 3 independent experiments. Images were acquired using a Zeiss AxioImager Z2 and processed in Zeiss ZEN software.

Apoptosis

A maximum of 5×10^4 cells were adhered to poly-L-lysine (Sigma) coated slides and then apoptotic cells were detected using an In Situ Cell Death Detection kit, Fluorescein (Roche) according to the manufacturer's instructions. A minimum of 100 cells per condition were scored using a Zeiss AxioSkop 2 microscope.

Immunoblotting

Whole cell lysates were prepared from 7×10^5 cells in 2 x SDS-DTT sample buffer. Lysates were run on SDS-PAGE gels and transferred to PVDF membranes (Millipore) using standard protocols. Membranes were probed with the following primary antibodies at 1:1000 dilution unless otherwise stated: Acetyl-p53 Lys379 (1:500) (2570; Cell Signalling), total p53 (1C12; Cell Signalling), Phospho-p44/42 MAPK (Erk1/2) Thr202/Tyr204 (9101; Cell Signalling), p44/42 MAPK (Erk1/2) (9102; Cell Signalling), Phospho-Stat3 (Tyr705) (9131; Cell Signalling), Stat3 (79D7) (4904; Cell Signalling), Histone H3 (acetyl K27) (ab4729, Abcam), Histone H3 (ab1791, Abcam), Beta Actin (ab8227, Abcam). Membranes were probed with secondary antibodies conjugated to horseradish peroxidase (HRP) (GE Healthcare; 1:5000) and proteins detected using SuperSignal West Femto kit (Thermo Scientific).

q-RT-PCR

Total RNA was isolated using an RNeasy Mini Kit (Qiagen). cDNA was prepared from 0.1 - 1 µg RNA using the SuperScript III First-Strand Synthesis System (Invitrogen) Q- RT-PCR was performed on diluted cDNA (1:10 in water), using Brilliant III Ultra-Fast QPCR Master Mix (Agilent) and gene specific primers (Sigma-Aldrich) on an Mx3000p qPCR system (Agilent) and standard cycling setup. Primer sequences were retrieved from Primer3 (<http://primer3.ut.ee>) and are listed below (F, forward; R, reverse; all 5'-3'): Noxa_F, ATAAGTGTGGTTCTGGCGCA, and Noxa_R, TCCTTCAAGTCTGCTGGCAC; Puma_F, GCGGCGGAGACAAGAAGA, and Puma_R, AGTCCCATGAAGAGATTGTACATGAC; Bax_F, TGGAGCTGCAGAGGATGATTG, and Bax_R, AGCTGCCACCCGGAAGA; Cdkn1a_F, TTGCACTCTGGTGTCTGAGC, and Cdkn1a_R, TGATAGAAATCTGTCAGGCTGGT; Mdm2_F, GGTCCTGTCCCTTTGATCCG, and Mdm2_R, CAAGATCCTGATGCGAGGGC; Bcl2_F, GAGCGTCAACAGGGAGATGT, and Bcl2_R, GCCGGTTCAGGTACTCAGTC; Gapdh_F, TGACGTGCCGCTGGAGAAA, and Gapdh_R, AGTGTAGCCCAAGATGCCCTTCAG. Gene expression levels were determined by 2^{-CT} method following normalization to *Gapdh*.

Code Availability

All codes have been referenced in the appropriate methods section and are also available upon request.

Statistics and Reproducibility

Unless otherwise stated, all statistical analyses used unpaired Student's *t*-test on raw data. All error bars represent the standard error of the mean (SEM). P-values ≤ 0.05 were considered statistically significant. All box plots show the median and interquartile range and whiskers represent minimum and maximum values. Survival curves were constructed using the Kaplan-Meier method and statistical significance was determined using log-rank analysis. Hypergeometric tests were performed using *phyper* function in R stats package. The number of independent experiments used to generate statistically significant data are defined in the relevant figure legends. The histopathology and immunohistochemistry images shown in Figure 1d are representative of 7 mice. The flow cytometry images in Figure 1e are representative of 7 mice. Immunofluorescence data shown in Figure 5d and 5e were generated from 3 independent experiments. The western blot shown in Figure 5h is representative of 3 independent experiments. The western blots shown in Figures S3a and Figure S4b are representative of 2 independent experiments.

Data Availability

All RNA-Seq and Chip-Seq data have been deposited in the GEO database (GSE923344). All data that support the conclusions are available from the authors on reasonable request and source data from all graphs are present in Supplementary Table 26.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The Huntly lab is funded by the ERC (grant 647685 COMAL), KKLf, MRC, Bloodwise, the Wellcome Trust (WT) and the Cambridge NIHR BRC. We acknowledge the WT/MRC centre grant (097922/Z/11/Z) and support from WT strategic award 100140. D.S is a Postdoctoral Fellow of the Mildred-Scheel Organisation, German Cancer Aid. We thank Chiara Cossetti, Michal Maj and Reiner Schulte at the CIMR Flow Cytometry Core for their help with cell sorting. This research was supported by the Cambridge NIHR BRC Cell Phenotyping Hub. EL is funded by a Sir Henry Dale fellowship from the WT/Royal Society.

References

1. Blobel GA. CREB-binding protein and p300: molecular integrators of hematopoietic transcription. *Blood*. 2000; 95:745–755. [PubMed: 10648382]
2. Roelfsema JH, Peters DJ. Rubinstein-Taybi syndrome: clinical and molecular overview. *Expert reviews in molecular medicine*. 2007; 9:1–16.
3. Kung AL, et al. Gene dose-dependent control of hematopoiesis and hematologic tumor suppression by CBP. *Genes & development*. 2000; 14:272–277. [PubMed: 10673499]
4. Cancer Genome Atlas Research, N. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med*. 2013; 368:2059–2074. DOI: 10.1056/NEJMoa1301689 [PubMed: 23634996]

5. Papaemmanuil E, et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood*. 2013; 122:3616–3627. quiz 3699. DOI: 10.1182/blood-2013-08-518886 [PubMed: 24030381]
6. Mullighan CG, et al. CREBBP mutations in relapsed acute lymphoblastic leukaemia. *Nature*. 471:235–239.
7. Pasqualucci L, et al. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature*. 471:189–195.
8. Okosun J, et al. Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. *Nat Genet*. 2014; 46:176–181. DOI: 10.1038/ng.2856 [PubMed: 24362818]
9. Vicente C, et al. Targeted sequencing identifies associations between IL7R-JAK mutations and epigenetic modulators in T-cell acute lymphoblastic leukemia. *Haematologica*. 2015; 100:1301–1310. DOI: 10.3324/haematol.2015.130179 [PubMed: 26206799]
10. da Silva Almeida AC, et al. The mutational landscape of cutaneous T cell lymphoma and Sezary syndrome. *Nat Genet*. 2015; 47:1465–1470. DOI: 10.1038/ng.3442 [PubMed: 26551667]
11. Green MR, et al. Mutations in early follicular lymphoma progenitors are associated with suppressed antigen presentation. *Proc Natl Acad Sci U S A*. 2015; 112:E1116–1125. DOI: 10.1073/pnas.1501199112 [PubMed: 25713363]
12. Green MR, et al. Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma. *Blood*. 2013; 121:1604–1611. DOI: 10.1182/blood-2012-09-457283 [PubMed: 23297126]
13. Kelly PN, Dakic A, Adams JM, Nutt SL, Strasser A. Tumor growth need not be driven by rare cancer stem cells. *Science*. 2007; 317:337. doi: 10.1126/science.1142596 [PubMed: 17641192]
14. Kuppers R. Mechanisms of B-cell lymphoma pathogenesis. *Nat Rev Cancer*. 2005; 5:251–262. DOI: 10.1038/nrc1589 [PubMed: 15803153]
15. Huntly BJ, Gilliland DG. Leukaemia stem cells and the evolution of cancer-stem-cell research. *Nat Rev Cancer*. 2005; 5:311–321. DOI: 10.1038/nrc1592 [PubMed: 15803157]
16. Kikushige Y, et al. Self-renewing hematopoietic stem cell is the primary target in pathogenesis of human chronic lymphocytic leukemia. *Cancer cell*. 2011; 20:246–259. DOI: 10.1016/j.ccr.2011.06.029 [PubMed: 21840488]
17. Chung SS, et al. Hematopoietic stem cell origin of BRAFV600E mutations in hairy cell leukemia. *Sci Transl Med*. 2014; 6:238ra271. doi: 10.1126/scitranslmed.3008004
18. Damm F, et al. Acquired initiating mutations in early hematopoietic cells of CLL patients. *Cancer Discov*. 2014; 4:1088–1101. DOI: 10.1158/2159-8290.CD-14-0104 [PubMed: 24920063]
19. Weigert O, Weinstock DM. The evolving contribution of hematopoietic progenitor cells to lymphomagenesis. *Blood*. 120:2553–2561.
20. Kang-Decker N, et al. Loss of CBP causes T cell lymphomagenesis in synergy with p27Kip1 insufficiency. *Cancer cell*. 2004; 5:177–189. [PubMed: 14998493]
21. Chan WI, et al. The transcriptional coactivator Cbp regulates self-renewal and differentiation in adult hematopoietic stem cells. *Mol Cell Biol*. 2011; 31:5046–5060. MCB.05830-11 [pii]. DOI: 10.1128/MCB.05830-11 [PubMed: 22006020]
22. Xu W, et al. Global transcriptional coactivators CREB-binding protein and p300 are highly essential collectively but not individually in peripheral B cells. *Blood*. 2006; 107:4407–4416. DOI: 10.1182/blood-2005-08-3263 [PubMed: 16424387]
23. Wang HC, Perry SS, Sun XH. Id1 attenuates Notch signaling and impairs T-cell commitment by elevating Deltex1 expression. *Mol Cell Biol*. 2009; 29:4640–4652. DOI: 10.1128/MCB.00119-09 [PubMed: 19564409]
24. Jiang Y, et al. CREBBP Inactivation Promotes the Development of HDAC3 Dependent Lymphomas. *Cancer Discov*. 2016; doi: 10.1158/2159-8290.CD-16-0975
25. Morin RD, et al. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat Genet*. 2010; 42:181–185. DOI: 10.1038/ng.518 [PubMed: 20081860]

26. Baron BW, et al. GFI1B, EVI5, MYB--additional genes that cooperate with the human BCL6 gene to promote the development of lymphomas. *Blood Cells Mol Dis.* 2014; 52:68–75. DOI: 10.1016/j.bcmd.2013.07.003 [PubMed: 23910958]
27. Wang X, Huang H, Young KH. The PTEN tumor suppressor gene and its role in lymphoma pathogenesis. *Aging (Albany NY).* 2015; 7:1032–1049. [PubMed: 26655726]
28. Schmitz R, et al. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature.* 2012; 490:116–120. DOI: 10.1038/nature11378 [PubMed: 22885699]
29. Li H, et al. Mutations in linker histone genes HIST1H1 B, C, D, and E; OCT2 (POU2F2); IRF8; and ARID1A underlying the pathogenesis of follicular lymphoma. *Blood.* 2014; 123:1487–1498. DOI: 10.1182/blood-2013-05-500264 [PubMed: 24435047]
30. Prasad MA, et al. Ebf1 heterozygosity results in increased DNA damage in pro-B cells and their synergistic transformation by Pax5 haploinsufficiency. *Blood.* 2015; 125:4052–4059. DOI: 10.1182/blood-2014-12-617282 [PubMed: 25838350]
31. Oricchio E, et al. Frequent disruption of the RB pathway in indolent follicular lymphoma suggests a new combination therapy. *J Exp Med.* 2014; 211:1379–1391. DOI: 10.1084/jem.20132120 [PubMed: 24913233]
32. Lohr JG, et al. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc Natl Acad Sci U S A.* 2012; 109:3879–3884. DOI: 10.1073/pnas.1121343109 [PubMed: 22343534]
33. Morin RD, et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature.* 2011; 476:298–303. DOI: 10.1038/nature10351 [PubMed: 21796119]
34. Alexandrov LB, et al. Signatures of mutational processes in human cancer. *Nature.* 2013; 500:415–421. DOI: 10.1038/nature12477 [PubMed: 23945592]
35. Ogiwara H, Kohno T. CBP and p300 histone acetyltransferases contribute to homologous recombination by transcriptionally activating the BRCA1 and RAD51 genes. *PLoS one.* 7:e52810.
36. Gu W, Roeder RG. Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. *Cell.* 1997; 90:595–606. [PubMed: 9288740]
37. Tang Y, Zhao W, Chen Y, Zhao Y, Gu W. Acetylation is indispensable for p53 activation. *Cell.* 2008; 133:612–626. [PubMed: 18485870]
38. Bashford-Rogers RJ, et al. Network properties derived from deep sequencing of human B-cell receptor repertoires delineate B-cell populations. *Genome Res.* 2013; 23:1874–1884. DOI: 10.1101/gr.154815.113 [PubMed: 23742949]
39. Rickert RC, Roes J, Rajewsky K. B lymphocyte-specific, Cre-mediated mutagenesis in mice. *Nucleic Acids Res.* 1997; 25:1317–1318. [PubMed: 9092650]
40. Zhang J, et al. The CREBBP Acetyltransferase Is a Haploinsufficient Tumor Suppressor in B-cell Lymphoma. *Cancer Discov.* 2017; 7:322–337. DOI: 10.1158/2159-8290.CD-16-1417 [PubMed: 28069569]
41. Garcia-Ramirez I, et al. Crebbp loss cooperates with Bcl2 over-expression to promote lymphoma in mice. *Blood.* 2017; doi: 10.1182/blood-2016-08-733469
42. Goodman RH, Smolik S. CBP/p300 in cell growth, transformation, and development. *Genes & development.* 2000; 14:1553–1577. [PubMed: 10887150]
43. Weigert O, et al. Molecular ontogeny of donor-derived follicular lymphomas occurring after hematopoietic cell transplantation. *Cancer Discov.* 2012; 2:47–55. DOI: 10.1158/2159-8290.CD-11-0208 [PubMed: 22585168]
44. Xie M, et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med.* 2014; 20:1472–1478. DOI: 10.1038/nm.3733 [PubMed: 25326804]
45. Quivoron C, et al. TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. *Cancer cell.* 2011; 20:25–38. DOI: 10.1016/j.ccr.2011.06.003 [PubMed: 21723201]
46. Couronne L, Bastard C, Bernard OA. TET2 and DNMT3A mutations in human T-cell lymphoma. *N Engl J Med.* 2012; 366:95–96. DOI: 10.1056/NEJMc1111708 [PubMed: 22216861]

47. Green MR, et al. Transient expression of Bcl6 is sufficient for oncogenic function and induction of mature B-cell lymphoma. *Nat Commun.* 2014; 5:3904.doi: 10.1038/ncomms4904 [PubMed: 24887457]
48. Zhang J, et al. Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. *Nat Med.* 2015; 21:1190–1198. DOI: 10.1038/nm.3940 [PubMed: 26366712]
49. Ortega-Molina A, et al. The histone lysine methyltransferase KMT2D sustains a gene expression program that represses B cell lymphoma development. *Nat Med.* 2015; 21:1199–1208. DOI: 10.1038/nm.3943 [PubMed: 26366710]
50. Giotopoulos G, et al. A novel mouse model identifies cooperating mutations and therapeutic targets critical for chronic myeloid leukemia progression. *J Exp Med.* 2015; 212:1551–1569. DOI: 10.1084/jem.20141661 [PubMed: 26304963]
51. Dobin A, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013; 29:15–21. DOI: 10.1093/bioinformatics/bts635 [PubMed: 23104886]
52. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics.* 2015; 31:166–169. DOI: 10.1093/bioinformatics/btu638 [PubMed: 25260700]
53. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014; 15:550.doi: 10.1186/s13059-014-0550-8 [PubMed: 25516281]
54. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005; 102:15545–15550. DOI: 10.1073/pnas.0506580102 [PubMed: 16199517]
55. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012; 9:357–359. DOI: 10.1038/nmeth.1923 [PubMed: 22388286]
56. Zhang Y, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 2008; 9:R137.doi: 10.1186/gb-2008-9-9-r137 [PubMed: 18798982]
57. Stark R, Brown G. DiffBind: differential binding analysis of ChIP-Seq peak data. *Bioconductor.* 2011 <http://bioconductor.org/packages/release/bioc/html/DiffBind.html>.
58. Thomas RK, et al. High-throughput oncogene mutation profiling in human cancer. *Nat Genet.* 2007; 39:347–351. DOI: 10.1038/ng1975 [PubMed: 17293865]
59. Fischer A, Illingworth CJ, Campbell PJ, Mustonen V. EMu: probabilistic inference of mutational processes and their localization in the cancer genome. *Genome Biol.* 2013; 14:R39.doi: 10.1186/gb-2013-14-4-r39 [PubMed: 23628380]
60. Watson SJ, et al. Viral population analysis and minority-variant detection using short read next-generation sequencing. *Philosophical Transactions of the Royal Society B-Biological Sciences.* 2013; 368 doi:Artn 20120205. doi: 10.1098/Rstb.2012.0205
61. Lefranc MP, et al. IMGT, the international ImMunoGeneTics information system. *Nucleic acids research.* 2009; 37:D1006–1012. DOI: 10.1093/nar/gkn838 [PubMed: 18978023]
62. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *Journal of molecular biology.* 1990; 215:403–410. DOI: 10.1016/S0022-2836(05)80360-2 [PubMed: 2231712]
63. Lefranc MP. IMGT unique numbering for the variable (V), constant (C), and groove (G) domains of IG, TR, MH, IgSF, and MhSF. *Cold Spring Harb Protoc.* 2011; 2011:633–642. DOI: 10.1101/pdb.ip85 [PubMed: 21632789]
64. Mirzoeva OK, Petrini JH. DNA damage-dependent nuclear dynamics of the Mre11 complex. *Mol Cell Biol.* 2001; 21:281–288. DOI: 10.1128/MCB.21.1.281-288.2001 [PubMed: 11113202]

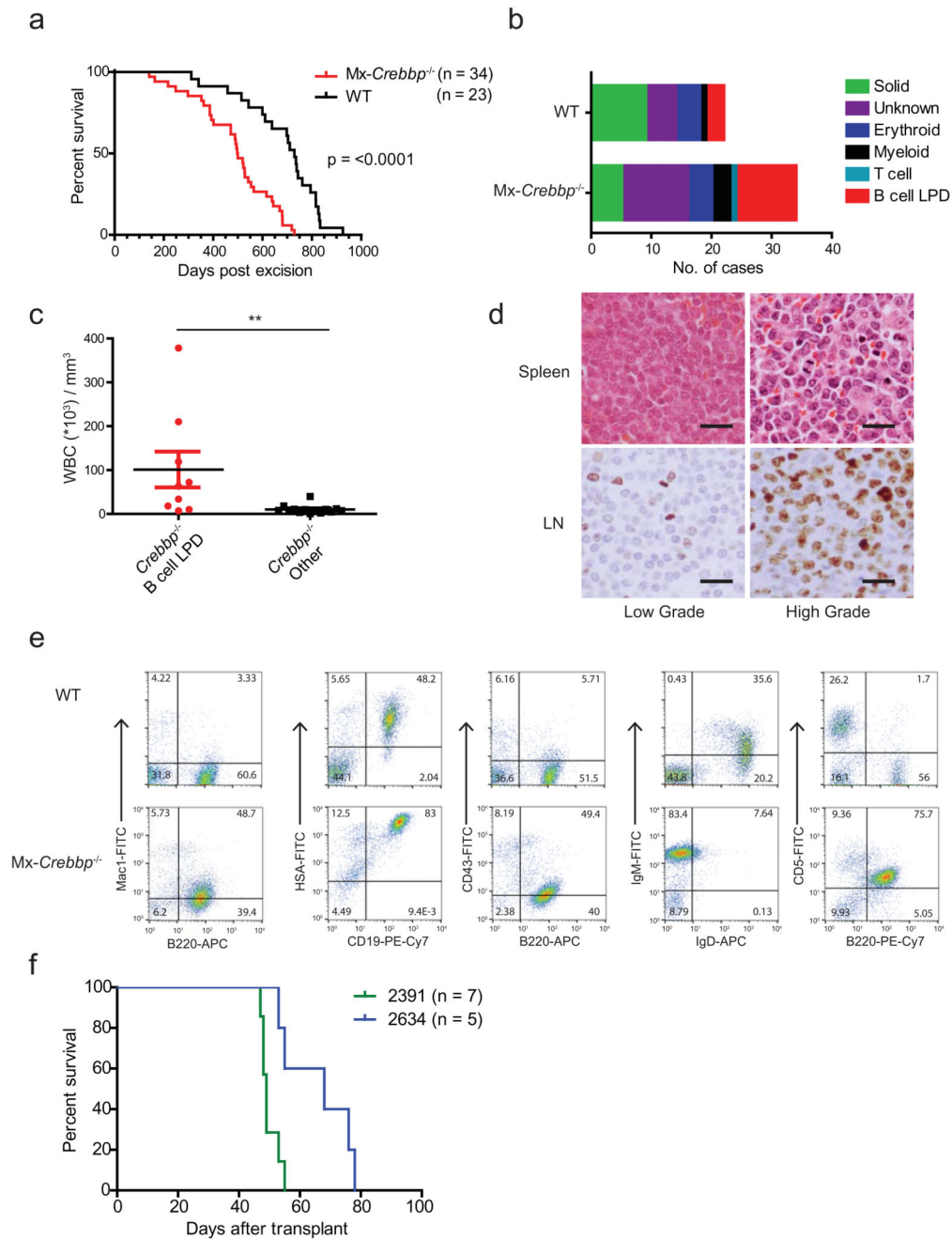


Figure 1. Loss of *Crebbp* generates an aggressive mature LPD/Lymphoma.

a. Kaplan-Meier (KM) graph demonstrating significantly reduced survival for $Mx-Crebbp^{-/-}$ mice (n = 34 animals) over WT littermates, (n = 23 animals, $p < 0.0001$) Log-rank (Mantel-Cox) test. **b.** Increased incidence of haematological malignancies of $Mx-Crebbp^{-/-}$ mice in comparison to WT mice (18 vs 7 cases), that is particularly marked for B-cell LPD (10 vs 3 cases). **c.** Mice with B-cell LPD (n = 9 animals) demonstrate a significantly raised WCC count in comparison to other $Mx-Crebbp^{-/-}$ mice (n = 14 animals) (** $p=0.0034$, Mann Whitney test). The mean \pm standard error of the mean (SEM) is shown. **d.** Spleen (upper

panels H + E) and Lymph node (lower, Ki-67 staining) histology from *Mx-Crebbp*^{-/-} mice with B-cell LPD/Lymphoma. Scale bar: 25µm. **e.** Immunophenotypic analysis confirming a mature B-cell phenotype (CD19⁺, B220⁺, sIgM⁺) with additional aberrant markers (Mac1^{Int}, CD5⁺). **f.** The disease is transplantable to secondary recipients. Two representative tumours 2391 and 2634 (7 and 5 recipient animals respectively) are shown.

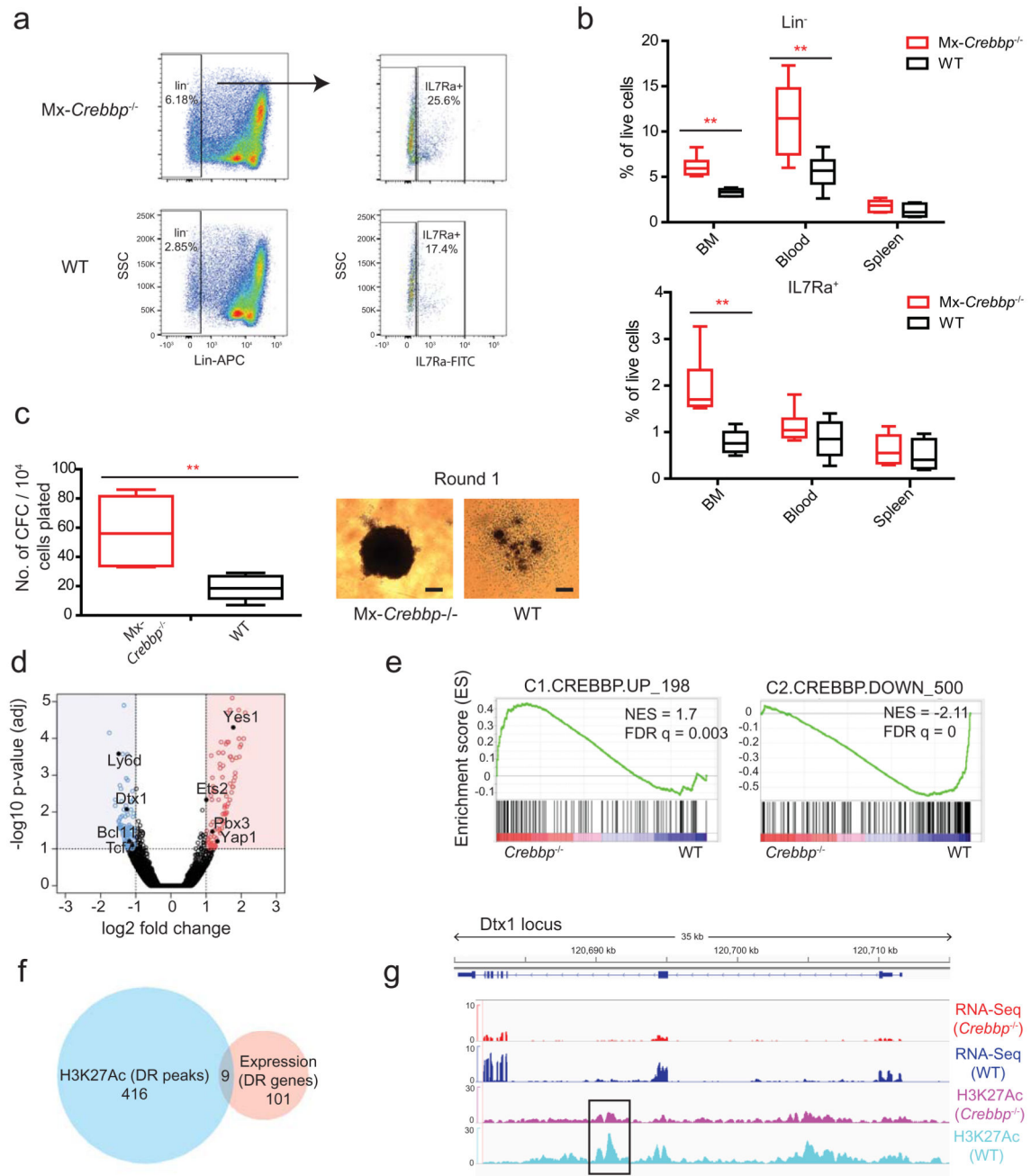


Figure 2. *Crebbp* loss alters stem and progenitor function.

a Representative FACS plots of bone marrow reveal **b**, a general increase in lin⁺ cells in BM ($p = 0.0012$ by 2-sided t-test with Welch's correction), and PB ($p = 0.01$ by 2-sided t-test) and in IL-7Ra⁺ progenitors in BM ($p = 0.0065$ by 2-sided t-test with Welch's correction) of Mx-*Crebbp*^{-/-} animals ($n = 6$ animals per genotype; 3 individual animals per genotype, quantified in 2 independent experiments). **c**. IL-7Ra⁺ progenitor population from Mx-*Crebbp*^{-/-} mice demonstrates an increased clonogenic (left panel) and proliferative (right photomicrographs, scale bar 50 μ m) potential in CFC assays (left, $n = 6$ animals per

genotype; 2 individual animals per genotype, quantified in 3 independent experiments, ** $p = 0.0024$ by 2-sided t-test). All graphs show the median with boxes representing the interquartile range and whiskers representing the minimum and maximum values. **d.** Volcano plot of differential gene expression between Mx-*Crebbp*^{-/-} and WT mice (n = 6 animals per genotype). P value generated using a negative binomial generalized linear model (DEseq2) **e.** GSEA analysis demonstrates enrichment for a published *CREBBP* mutant lymphoma signature. **f.** Integration of downregulated differentially bound H3K27Ac peaks and downregulated genes demonstrates only minimal overlap. **g.** Integrated plot showing RNA-Seq and H3K27Ac ChIP-Seq at the *Dtx1* locus demonstrates that downregulated gene expression is accompanied by a statistically significant loss of H3K27ac at an intragenic enhancer (boxed) in *Crebbp*^{-/-} progenitors. All mice were treated with pIpC between 6-10 weeks after birth (7.7 weeks in Fig a-b, 6.3 weeks in Fig c and 6.9 weeks in Fig d) and IL7Ra⁺ cells isolated from the BM 8-10 weeks (9.6 weeks in Fig a-b, 8.1 weeks in Fig c and 9 weeks in Fig d) after pIpC administration.

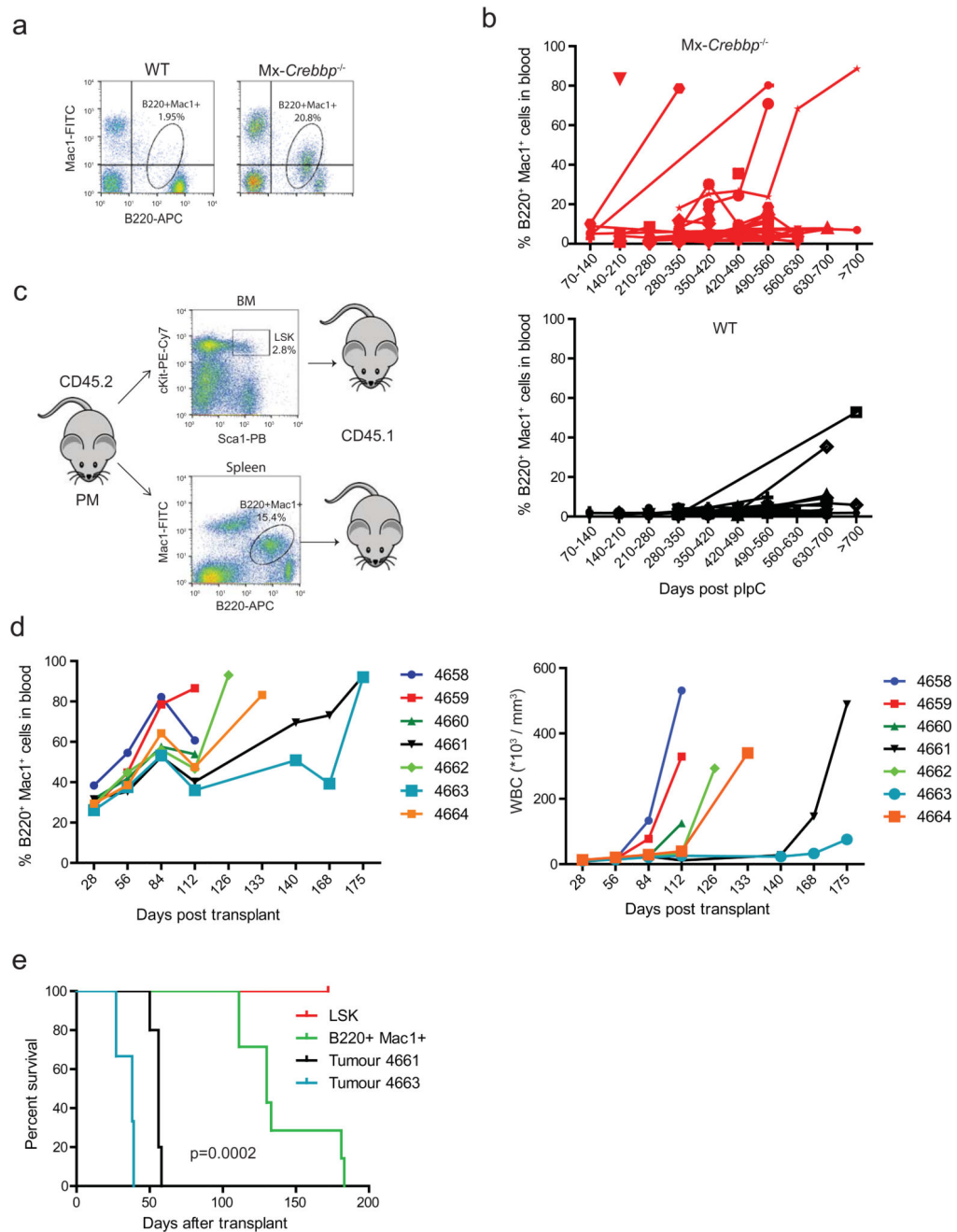


Figure 3. Development of a pre-malignant stem cell population upon *Crebbp* loss.

a. An aberrant cell population with similar B220⁺/Mac1^{Int} surface phenotype to the tumours can be found in the PB **b.** Increase in this aberrant cell population over time following plpC administration (*Mx-Crebbp*^{-/-}; 67 mice, WT; 64 mice). **c.** Transplant of this pre-malignant stem-cell population into 7 congenic CD45.1 recipient mice generates **d.** high-level engraftment (>25% of peripheral blood cells at 4 weeks, left panel) with sequential *in vivo* amplification of this population seen over time (left panel) ultimately giving rise to elevated WBC counts and LPD (right panel). Each separately coloured number (4658, 4659, 4660,

4661, 4662, 4663 and 4664) in both panels represents the same individual transplant recipient. **e.** KM graph of survival of the recipient animals in (d) receiving pre-malignant B220⁺/Mac1^{Int} (median survival 132 days) in contrast to LSK cells (n = 7 mice per cell population) and secondary recipients of tumour cells from 4661 (n = 5 recipient mice) and 4663 (n = 3 recipient mice) (median survival 48 days, p = 0.0002, Log-rank (Mantel-Cox) test).

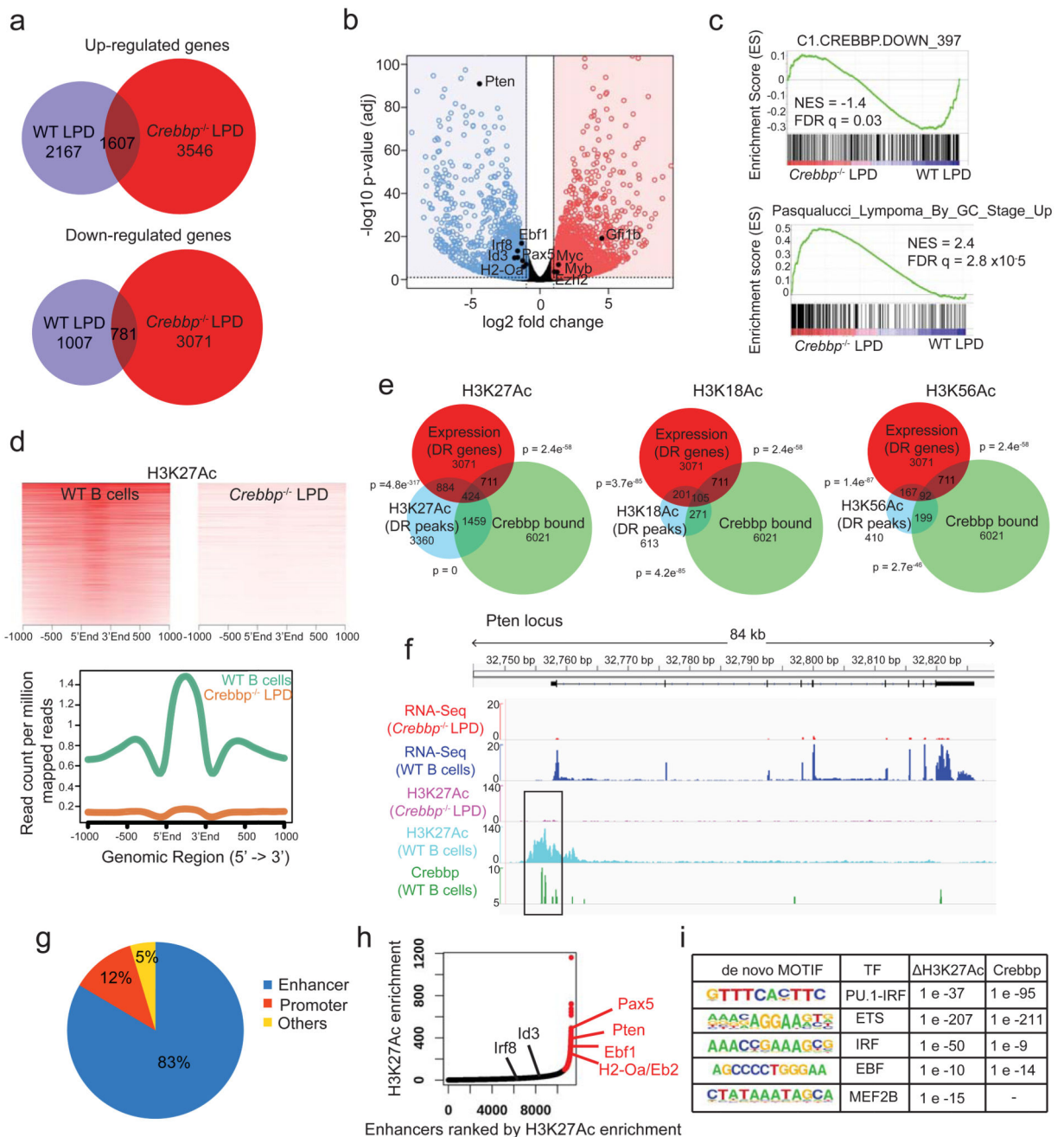


Figure 4. Transcriptional and epigenetic alterations in *Crebbp*^{-/-} deficient lymphomas.

a. Venn diagrams show overlaps in differentially expressed genes between WT and Mx-*Crebbp*^{-/-} lymphomas when compared to WT B220+ cells. Numbers in outer circles are inclusive of overlaps, overlapping gene numbers are shown. **b.** Volcano plot of differential gene expression between Mx-*Crebbp*^{-/-} lymphomas (from n = 2 animals) and WT B220 cells (from n = 6 animals; B220+ cells from 3 WT animals pooled per sequencing reaction). P value generated using a negative binomial generalized linear model (DEseq2). **c.** GSEA analysis of genes differentially expressed between Mx-*Crebbp*^{-/-} and WT lymphomas

demonstrates enrichment of a *CREBBP* mutant human signature²⁴ and a previously published dataset of murine lymphomas. **d.** H3K27Ac modification measured by specific density plots (upper panel) and averaged read counts (lower panel) for genes downregulated in *Mx-Crebbp*^{-/-} lymphomas. **e.** Significant overlap is seen for Crebbp binding, downregulated H3K27Ac peaks (left panel), H3K18Ac peaks (middle panel), H3K56Ac peaks (right panel) and genes downregulated in *Mx-Crebbp*^{-/-} lymphomas compared to WT B220+ cells (p values for hypergeometric comparisons are shown). **f.** Integrated RNA-Seq and ChIP-Seq analysis of the *Pten* locus, a known lymphoma tumour suppressor, in normal B220+ B-cells and the B-cell LPD/lymphoma demonstrates decreased H3K27Ac at the boxed Crebbp bound peak. **g.** Downregulated H3K27Ac peaks occur most frequently at enhancer regions and **h** appear to alter the function of genes regulated in WT B220+ cells by superenhancers (red) and normal enhancers (black). **i.** Table demonstrating selected motifs enriched in the downregulated H3K27Ac peaks in *Mx-Crebbp*^{-/-} LPD (H3K27Ac) and Crebbp bound peaks in WT B220 cells (Crebbp), as determined by *de novo* motif analysis.

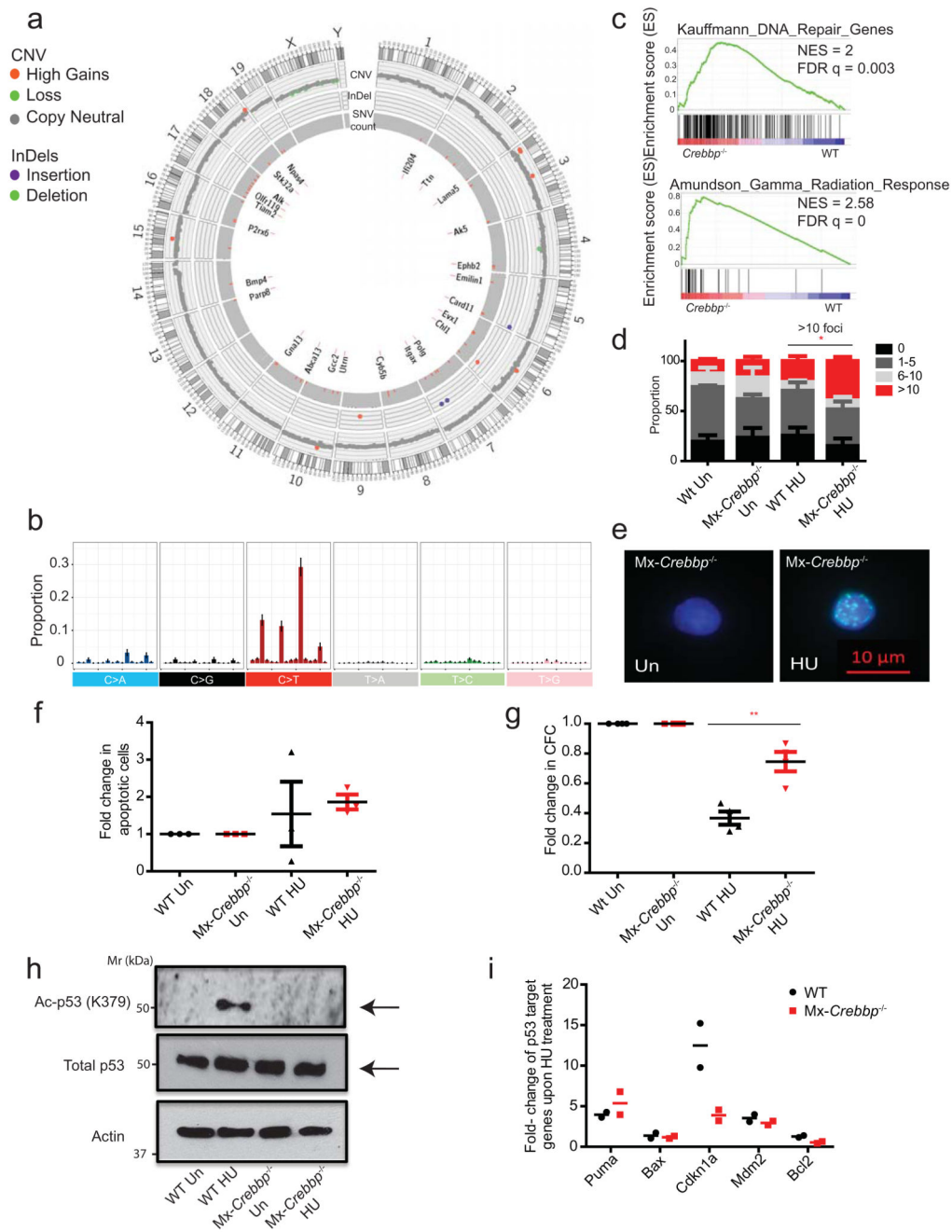


Figure 5. Genetic evolution in *Crebbp* deficient mice is linked to a defective DNA damage response.

a. Circos plot showing SNV and copy number analysis for a representative *Crebbp*^{-/-} lymphoma. **b.** Mutational signature analysis of exome sequencing results for *Mx-Crebbp*^{-/-} tumours (after Alexandrov³⁴), demonstrates an ageing signature. **c.** GSEA analysis in IL-7Ra⁺ progenitors demonstrates enrichment for DNA repair genes and genes associated with DNA damage. **d.** Bar chart showing the number of double stranded DNA breaks as evidenced by H2AX γ foci in IL-7Ra⁺ progenitors from WT and *Mx-Crebbp*^{-/-} mice,

demonstrating no basal difference, but a marked difference in cells with multiple foci in *Mx-Crebbp*^{-/-} progenitors upon the induction of DNA damage with HU. The mean \pm SEM is shown, $n = 3$ independent experiments ($*p = 0.047$ by 2-sided t-test). See Supplementary Table 26 for scatter plots of >10 foci. **e.** Representative image of H2AX γ foci in untreated and HU treated *Mx-Crebbp*^{-/-} cells. Scale bar 10 μ m. **f.** Despite the increase in DNA damage in *Mx-Crebbp*^{-/-} mice, no difference in the degree of progenitor apoptosis following HU treatment was seen between WT and *Mx-Crebbp*^{-/-} mice ($n = 3$ biologically independent experiments, 2-sided t-test). **g.** Preservation of clonogenic capacity in CFC assays for *Mx-Crebbp*^{-/-} when compared to WT lymphoid progenitors following HU. The mean \pm SEM is shown, $n = 4$ mice per genotype; 2 mice per genotype quantified in 2 independent experiments ($**p = 0.003$ by 2-sided t-test). **h.** Western blot analysis demonstrates that following HU, *Mx-Crebbp*^{-/-} lymphoid progenitors fail to properly acetylate p53 and induce an optimal DNA damage response (upper arrow, acetylated (K379) p53, lower arrow, total p53). Unprocessed scans of blots are shown in Figure S8. **i.** Transcriptional activation of targets genes suggests a loss of p53 activity. The graph shows data from two independent experiments with a line representing the mean.

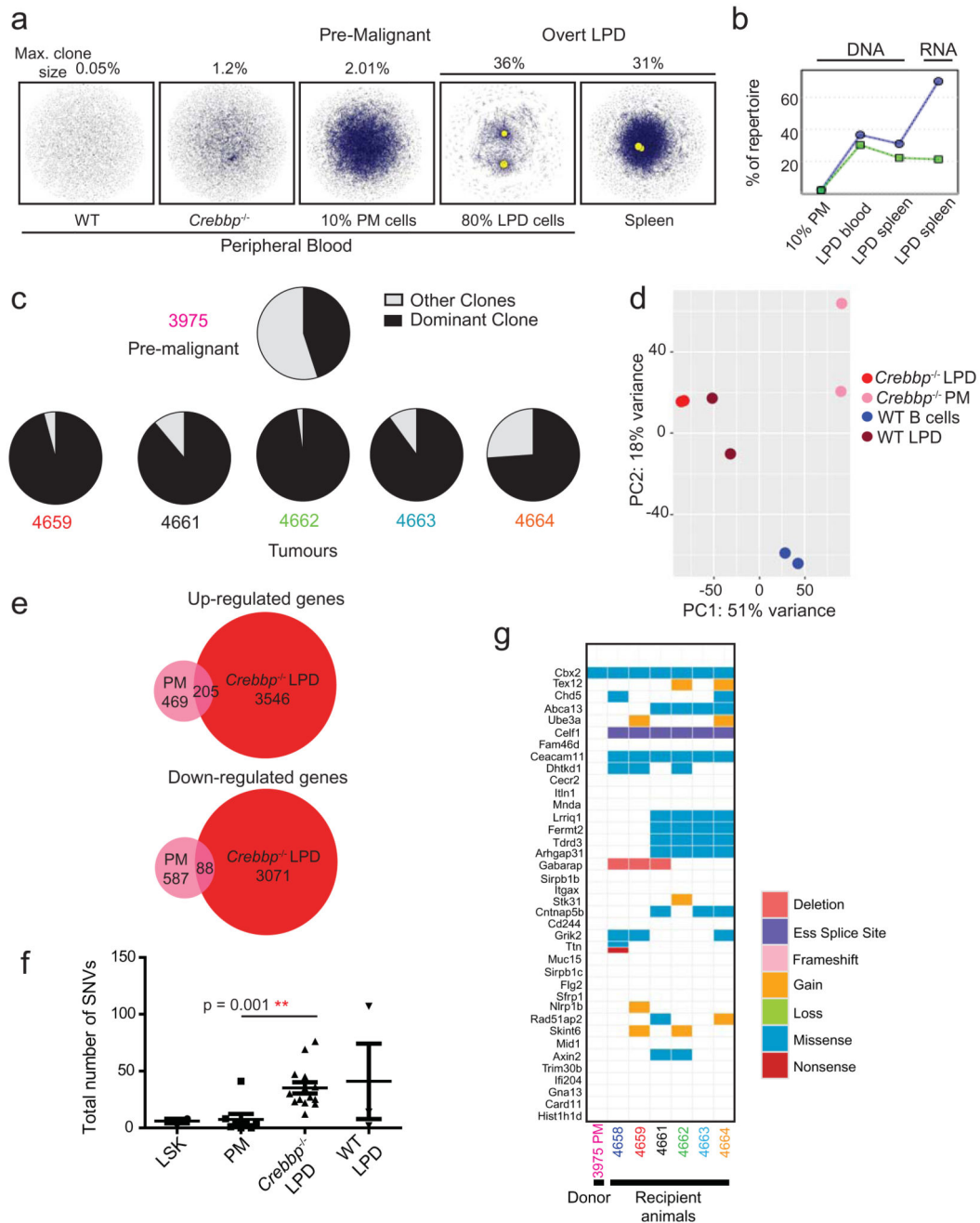


Figure 6. *Crebbp* deficient pre-malignant stem cells demonstrate clonal, transcriptional and genetic evolution during lymphomagenesis.

a. Left 2 panels, clonal analysis of IgH DNA rearrangements demonstrates a reduced IgH repertoire diversity (number of dots) and an increased maximum clone size (relative size of dots, maximum clone size given above each panel) for *Mx-Crebbp*^{-/-} mice when compared to WT littermates. This repertoire narrows further as the pre-malignant (PM) B220⁺/Mac1^{Int} stem cell population increases from 10% to 80% in the PB (2 middle panels) of the same animal during disease evolution. **b.** Similar information in bar chart form. RNA analysis

from the last time point demonstrates that both clones are expressed suggesting a cellularly biclonal tumour. **c.** Pie charts to figuratively illustrate the growth of the dominant clone donated from the premalignant animal 3975 in the transplant recipients. Colour scheme and animal identifiers are as per Figure 3. **d.** Principle component analysis (PCA) demonstrates clear transcriptional differences along the continuum of lymphoma development from WT B cells through the PM stage to the *Mx-Crebbp*^{-/-} lymphoma and between *Mx-Crebbp*^{-/-} tumours and WT lymphomas. **e.** Venn diagrams document that genes differentially expressed in PM cells remain differentially expressed in the same direction in *Mx-Crebbp*^{-/-} lymphomas. Numbers in outer circles are inclusive of overlaps, overlapping numbers are shown. **f.** Scatter plot showing that the total number of SNVs increases along lymphoma evolution. Mean \pm SEM is shown, LSK n = 2 mice, PM n = 8 mice, *Crebbp*^{-/-} LPD n = 14 mice, WT LPD n = 3 mice. P value calculated using 2-sided t-test. **g.** Mutations in the most recurrently altered genes. Results for the premalignant (PM) donor cells from animal 3975 are aligned beside the tumours that developed from these cells in animals 4659, 4661, 4662, 4663 and 4644, demonstrating obvious and divergent genetic evolution within the individual lymphomas.

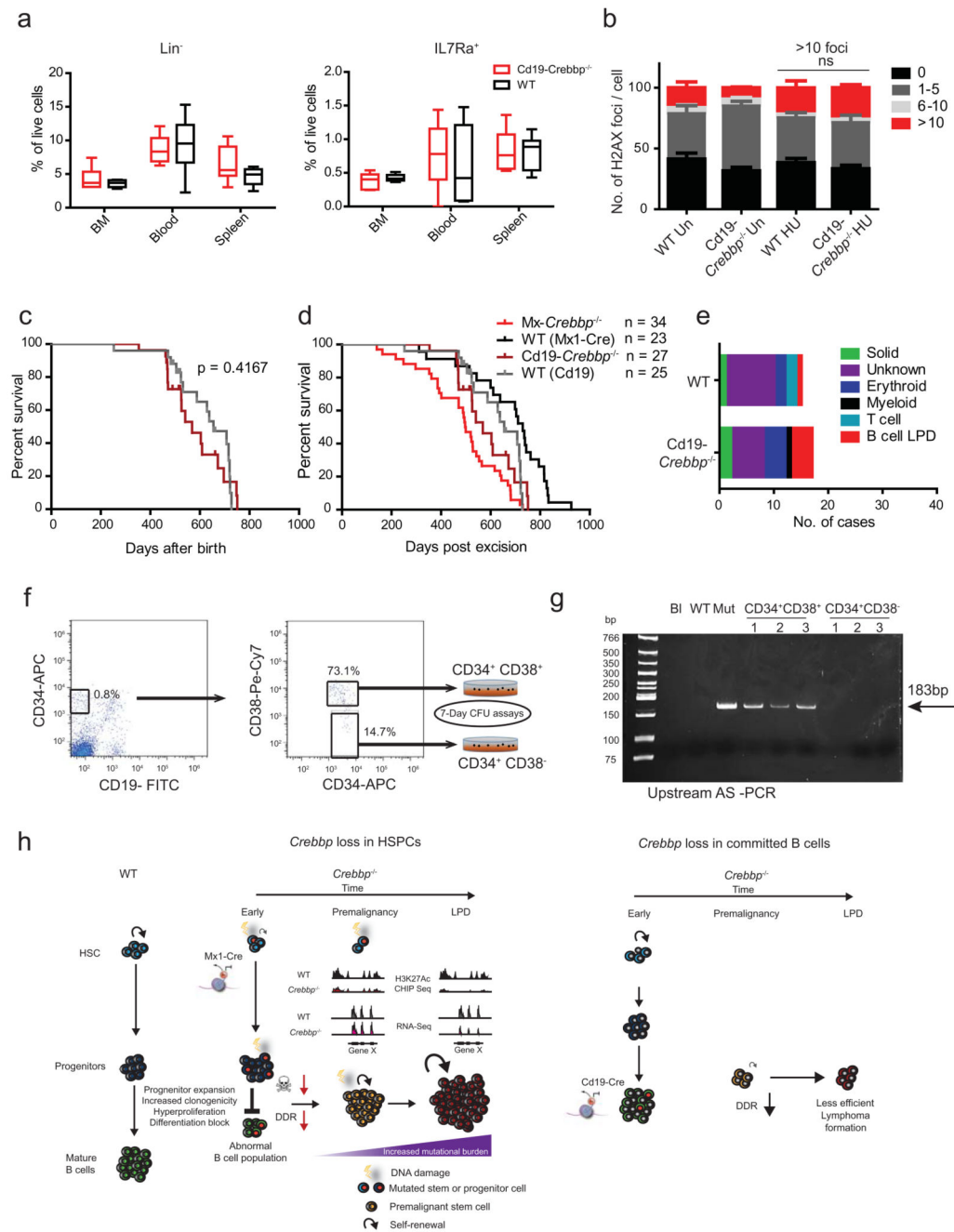


Figure 7. Loss of *Crebbp* in committed lymphoid cells attenuates disease generation and *CREBBP* mutations occur in HSPC from a lymphoma patient.

a. Analysis of the Lin⁻ and IL7Ra⁺ compartments in Cd19-*Crebbp*^{-/-} mice. Box plots show the median with interquartile range and whiskers represent the minimum and maximum values, n = 6 mice per genotype were quantified. Non-significance confirmed using 2-sided t-test. **b.** Comparison of H2AX γ foci in the IL7Ra⁺ progenitor populations from Cd19-*Crebbp*^{-/-} mice. Data represent the mean \pm SEM from n = 3 independent experiments, non-significance confirmed using 2-sided t-test. See Supplementary Table 26 for scatter plots of

>10 foci. **c.** KM analysis showing that no significant alterations in survival were seen for Cd19-*Crebbp*^{-/-} mice (n = 27 animals) vs WT mice (n = 25 animals). P value calculated using Log-rank (Mantel-Cox) test. **d.** in contrast to that seen in *Mx-Crebbp*^{-/-} mice. **e.** B-cell disease was very rare in this cohort (Cd19-*Crebbp*^{-/-} LPD 4 cases, WT LPD 1 case). **f.** Flow sorting strategy to functionally isolate specific HSPC populations from patients with *CREBBP* mutated lymphomas. **g.** Allele specific PCR demonstrates the presence of the same mutation (183 bp band) identified from the tumour (Mut lane) within the CD34⁺/CD38⁺ HSPC population. CFU-colony forming assay, Bl- blank, WT- wild type. **h.** Model of mechanisms and stage specific nature of *Crebbp* tumour suppression in lymphoid malignancies. Following early loss of *Crebbp* activity in the HSPC compartment, abnormal epigenetic regulation, post-translational modifications of substrate proteins and altered transcription lead to expansion of lymphoid progenitors, blocked differentiation, increased proliferation and clonogenicity within this expanded progenitor pool. In association with an alteration of the DNA damage response mediated through suboptimal *Crebbp* acetylation of p53, DNA double strand breaks accumulate and accompanied by a relative decrease in apoptosis, lead to the retention of mutations within B-cell progenitors. We illustrate the epigenetic priming of specific loci at earlier time points, where H3K27ac ChIP-Seq and RNA-Seq at the critical gene X locus read out in terms of downregulation of gene expression only at later timepoints during lymphoma evolution. We speculate that the molecular aberrations cumulatively acquired interact with this primed epigenetic state and lead to the restoration of self-renewal in this abnormal pre-malignant stem cell population and to the evolution of Lymphoma.