Genomic characterisation of Eμ-Myc mouse lymphomas identifies Bcor as a Myc co-operative tumour-suppressor gene

Marcus Lefebure1,2,* , Richard W. Tothill1,2,3,* , Elizabeth Kruse1, Edwin D. Hawkins4, Jake Shortt1,5, Geoffrey M. Matthews6, Gareth P. Gregory1, Benjamin P. Martin1, Madison J. Kelly1, Izabela Todorovski1, Maria A. Doyle1, Richard Lupat1, Jason Li1, Jan Schroeder4, Meaghan Wall7,8, Stuart Craig1, Gretchen Poortinga1, Don Cameron1, Megan Bywater1, Lev Kats1,2, Micah D. Gearhart9, Vivian J. Bardwell9, Ross A. Dickens10, Ross D. Hannan1,11, Anthony T. Papenfuss1,2,4 & Ricky W. Johnstone1,2

The Eμ-Myc mouse is an extensively used model of MYC driven malignancy; however to date there has only been partial characterization of MYC co-operative mutations leading to spontaneous lymphomagenesis. Here we sequence spontaneously arising Eμ-Myc lymphomas to define transgene architecture, somatic mutations, and structural alterations. We identify frequent disruptive mutations in the PRC1-like component and BCL6-corepressor gene Bcor. Moreover, we find unexpected concomitant multigenic lesions involving Cdkn2a loss and other cancer genes including Nras, Kras and Bcor. These findings challenge the assumed two-hit model of Eμ-Myc lymphoma and demonstrate a functional in vivo role for Bcor in suppressing tumorigenesis.

* These authors contributed equally to this work. Correspondence and requests for materials should be addressed to R.W.T. (email: Richard.Tothill@petermac.org) or to R.W.J. (email: Ricky.Johnstone@petermac.org).
Getically engineered mouse models (GEMMs) have been extensively used to validate oncogene or tumour suppressor function and serve as tractable models for pre-clinical evaluation of new therapeutics. Reminiscent of human malignancies, considerable disease heterogeneity can be observed in GEMMs despite the mostly inbred homozygous genotype and their controlled housing environment. Similar to humans, inter-tumour heterogeneity in GEMMs is likely caused by additional somatic lesions that occur spontaneously and are required for tumorigenesis. GEMMs in conjunction with high-throughput genomics can therefore be used to ‘discover’ resources to identify and validate spontaneous co-operating mutations and thus providing important information on the genetic interactions that underpin tumour onset and progression.

The Eμ-Myc mouse model employs a transgene mimicking the t(8:14) translocation of cMyc and IgH enhancer elements characteristic of human Burkitt lymphoma. Eμ-Myc mice develop B-cell leukaemia/lymphoma like malignancy with 100% penetrance. Disease in these mice typically manifests as a period of premalignant polyclonal precursor-B cell expansion followed by self-limiting remission and onset of clonal lymphoid neoplasia ranging from a pre-B to naïve B phenotype. Numerous articles have been published using this model to interrogate the efficacy and mechanisms of action of diverse anti-cancer agents and to study putative functional interactions between Myc and candidate cancer causing genes. For example, a landmark study by Strasser and colleagues using the Eμ-Myc model provided definitive evidence that MYC and BCL2 can functionally cooperate to accelerate lymphoma development. In addition, the role of the p19ARF-MDM2-p53 axis in regulating MYC-mediated apoptosis and lymphomagenesis has been comprehensively dissected at the genetic, biochemical and biological level using the Eμ-Myc model. Oncogenic Ras can also have an important co-operative role in MYC transformation and spontaneous activating Nras mutations were identified in Eμ-Myc lymphomas. Taken together, these studies highlight the utility of the model as a cancer gene-discovery resource and its extensive application for accurately studying the biology of the Myc oncogene.

Despite Eμ-Myc lymphomas being initiated by a single oncogene on a defined genetic background, there is a remarkable degree of inter-tumour heterogeneity, reflected by gene expression profiling and the latency of disease onset. This is consistent with studies indicating that Eμ-Myc lymphomas arise through acquisition of secondary or tertiary mutations that de-regulate tumour suppressor pathways such as those mediated by p19ARF and p53 that normally counteract the oncogenic effects of MYC. The intrinsic apoptotic pathway is clearly implicated in Eμ-Myc lymphomagenesis. Overexpression of Bcl2, forced expression of other pro-survival BCL2 family members (for example, Bclxl, Mcl1) or deletion of pro-apoptotic BH3-only genes (for example, Bim, Bmf, Bad) can accelerate lymphoma development in the Eμ-Myc mouse. However, whether mutational dysregulation of the BCL2 family actually occurs in spontaneously arising tumours is not yet defined. Disruption of non-apoptotic tumour suppressive mechanisms are also implicated in Eμ-Myc lymphomagenesis, in particular oncogene induced senescence pathways and immune surveillance. Despite clear evidence that naturally occurring mutations in Trp53, Cdkn2a and Nras co-operate with MYC in the Eμ-Myc model, other co-operating de novo genetic lesions remain unknown in up to half of all cases. In particular, it is important to determine if disruption of genes controlling disparate apoptotic and senescence pathways that can co-operate with MYC, when synthetically de-regulated are actually affected in spontaneously arising Eμ-Myc lymphomas. We therefore used massively-parallel sequencing to catalogue genetic lesions arising in Eμ-Myc lymphomas. We found that in addition to the expected mutations or deletions in Trp53, Cdkn2a and Nras, deleterious mutations in Bcl6-corepressor gene Bcor frequently occurred in Eμ-Myc lymphomas. Subsequent functional studies confirmed the tumour suppressor role of Bcor, and Eμ-Myc lymphomas with experimental depletion or deletion, or spontaneous mutation of Bcor presented with a unique gene expression signature indicating that TGFβ signalling was abnormally expressed in these lymphomas. Finally, we discovered that Eμ-Myc lymphomas co-occur with loss of Cdkn2a and either activating mutations in Ras or deleterious mutations in Bcor providing evidence that loss of Cdkn2a alone may not be sufficient to cooperate with overexpressed Myc to drive lymphomagenesis.

**Results**

**Transgene architecture in Eμ-Myc lymphoma.** To comprehensively characterize the genetic architecture of the Eμ-Myc transgenic mouse, we applied whole-genome sequencing (WGS) to spontaneous Eμ-Myc lymphomas (#88) and its matching germline DNA extracted from tail tissue of a hemizygous transgenic animal. Augmenting the mouse reference genome with the pUC12 vector sequence of the Eμ-Myc transgene enabled mapping of break points within the transgene and enumeration of transgene copy number in germline and lymphoma (Fig. 1a). We identified all the expected elements of the transgene cassette including the cloned translocation between chr12 (IgH enhancer) and chr15 (Myc) in addition to the flanking cloning sites within the pUC12 vector DNA (ref. 7). A single breakpoint was detected between chr19 and pUC12 vector sequence 5’ of the Eμ-Myc transgene. FISH analysis using fluorescently labelled BAC probes confirmed juxtaposition of Myc and jak2 (located on chr19) proximal to transgene insertion on chr19, which is consistent with observations from a previous study (Supplementary Fig. 1). Copy-number analysis of WGS data revealed that five copies of Myc were present in the hemizygous Eμ-Myc germline, corresponding to three transgene copies and two endogenous copies of Myc. We also observed a copy-number gain of a ~3 Mb segment of chr19 proximal to site of transgene insertion (Fig. 1b), as previously reported. The chr19 segment plus one extra copy of the Eμ-Myc transgene underwent further somatic gain (+1) in the #88 Eμ-Myc lymphoma, which has not been previously reported. Additional mate-pair WGS indicated that the transgene copies are likely arranged as concatenated repeats (Fig. 1b).

We confirmed germline amplification of the chr19 segment in Eμ-Myc transgenic mice bred at two additional institutions (Walter and Eliza Hall Institute (WEHI) Melbourne Australia, Cold Spring Harbor Laboratories USA) (Supplementary Fig. 2). This indicates that these structural arrangements are almost certainly founding events that occurred following the pronuclear injection of the transgenic vector. The amplicon on murine chr19 is syntenic to the human 9p24.1 region, which is frequently amplified in Hodgkin’s Lymphoma (HL) and primary mediastinal B-cell lymphoma (PMBCL) and contains the tumour-promoting genes JAK2 and CD274 (PD-L1). Fusello et al. previously showed that pre-malignant and transformed B-cells from Eμ-Myc mice exhibited higher PD-L1 expression compared with wild type B-cells, concluding that this was causally related to the chr19 amplification. However, we found that PD-L1 protein expression did not correlate with gene dosage inferred from additional somatic gain of the chr19 amplicon (Supplementary Fig. 3). We also found that Eμ-Myc lymphomas did not show appreciable activation of JAK2 signalling as assessed by phosphorylation of the downstream target STAT5 and furthermore, JAK2 inhibition using FDA-approved small molecule inhibitor ruxolitinib had no effect on survival of...
**Figure 1 | Transgene architecture in Eμ-Myc lymphoma.** (a) Circos plot showing the Eμ-Myc transgenic cassette (pUC12, phiX, chr12 (Eμ), chr15 (Myc), identified breakpoints and estimated copy number. Outer coloured bars depict chromosome and transgene segments. Grey bars and links represent germline copies with two copies per increment. Red bars represent additional somatic copy-number gain and breakpoints. (b) Schematic showing the arrangement of the Eμ-Myc transgene (upper), the three repeats (TG1-3) of the transgene in the Eμ-Myc germline (middle) and the extra gain of the transgene and the segment of chr19 in lymphoma #88 (lower). This figure was adapted from Adams et al. (ref. 7). R in red band indicates repetitive DNA elements.

Eμ-Myc lymphomas cultured in vitro (Supplementary Fig. 4). This suggests germline or somatic amplification of Jak2 is unlikely to contribute to initiation and maintenance of the Eμ-Myc lymphomas. The role of other genes within the amplified chr19 region (for example, Il33, Pdcd11g2) (Supplementary Data 1) remains to be clarified and is the subject of ongoing experiments.

**WES identifies new gene drivers of Eμ-Myc lymphoma.** We next applied whole-exome sequencing (WES) to explore the number, type and frequency of somatic mutations in 23 spontaneous Eμ-Myc lymphomas (Supplementary Data 2). Sixteen cases were derived prospectively and seven cases were taken from retrospectively archived lymphomas. WES of normal tail DNA from Eμ-Myc animal #88 and a littermate control from the prospective series (ML62) was used to filter germline polymorphisms from the data. Furthermore, dbSNP annotated variants or any single nucleotide variants (SNV) or insertion/deletion (InDel) recurring two or more times in closely related animals were removed from analysis as these variants were most likely polymorphisms or recurrent sequencing artifacts. A subset of SNVs and InDels were further validated using targeted amplicon-based sequencing confirming the high specificity of 92.6% (75/81) of our variant calling pipeline. Finally, to confirm the somatic driver mutations (described below) we sequenced matching isolated circulating B-cells from individual mice at 4-weeks of age showing undetectable or very low variant allele frequency (VAF) compared with VAF measured in blood at disease end point (Supplementary Fig. 5).

The lymphomas from the prospective cohort each had 3–31 SNVs or InDels predicted to truncate or alter the translated amino acid sequence of proteins encoded by annotated genes (Fig. 2a). As expected we detected Trp53 (n = 4) and Nras (n = 2) mutations. In addition we identified Kras (n = 4) mutations that have not been previously described in Eμ-Myc lymphomas (Fig. 2b). The most frequently mutated gene was the BCL6-copressor (Bcor), recurrently mutated in seven lymphomas (32%) with either frameshift InDels or nonsense mutations predicted to cause premature protein truncation. A splice-site variant in the PRC2 complex subunit Ezh2, a gene previously shown to function as a tumour-suppressor in the Eμ-Myc model24, was observed in one case (ML33). Lymphoma ML39 harboured a mutation in the gene Mtor affecting a residue conserved between mice and humans and predicted to be damaging using SIFT algorithm. Deregression of the AKT-MTOR-EIF4E pathway has been shown to co-operate with MYC in Eμ-Myc lymphomagenesis25 and treatment of Eμ-Myc transgenic mice with mTORC1 inhibition by everolimus during the pre-malignant stage of disease significantly delayed lymphoma onset26. Activating MTOR mutations have also previously been described in several human cancer types including diffuse large B-cell lymphoma27,28. Lymphoma #88 harboured a mutation in mir142 that has been reported in ~20% of diffuse large B-cell lymphomas28. Lymphomas ML352 and #299 had mutations in the ribosomal protein Rpl10, a gene found recurrently mutated in T-ALL (ref. 29). Three lymphomas (#218, #219, ML20) did not harbour SNVs or InDels in genes that have any obvious link to cancer.

We analysed WES (n = 9) and low to medium coverage WGS data (n = 13) (unpaired) to detect somatic copy-number alterations (SCNAs) in lymphoma samples. Focal deletion events of Cdkn2a were detected in five cases, validated by quantitative PCR (Supplementary Fig. 6). Somatic amplification of the Eμ-Myc transgene and the proximal chr19 amplified region as shown for lymphoma #88 (Fig. 1b), was evident in 8/22 tumours. We could also detect recurrent amplification and deletion events involving known human cancer genes30 (Supplementary Data 3 and Supplementary Fig. 7). Most of these SCNAs were characterized by large segmental amplifications or whole chromosome duplications involving chr3 (including Nras), chr6 (including Kras), chr10, chr12, chr15 (including Myc) and chr18 and monosomy or deletion of part of chr13. Focal SCNAs included low-level gains on chr1 (Cxcr7 and chr7 (Fgfr2), and hemizygous loss on chr4 (Psip1), chr9 (Nckipsd) and chr11 (Msi2). Taken together these studies highlight the genetic diversity of individual Eμ-Myc lymphomas and identify Bcor as the most frequently mutated gene in this model of Myc-driven malignancy.

**Tertiary driver mutations occur in Eμ-Myc lymphomas.** A long-held paradigm established from previous studies using the Eμ-Myc model was that a single co-operating lesion in the
prospective homozygous deletion along with an activating lymphomas #4242 (oncogenic lesions in the context of lymphomas was the identification of additional putative lesions in the same cell (Supplementary Data 4). Collectively, our data suggest that while heterogeneous loss of Cdkn2a creates a selective pressure resulting in loss of the wild type Cdkn2a allele, it does not ameliorate selective pressure for gain- or loss-of-function of other genes outside the p53-axis. These data are consistent with a previous report indicating that p19ARF inactivation is itself insufficient to trigger lymphomagenesis in Eμ-Myc transgenic mice and indeed the authors of that study posited that 'cryptic mutations other than p19ARF loss accompany the conversion of premalignant Eμ-Myc B cells’ in the Eμ-Myc;Cdkn2a+/− setting. Herein we identify activating mutations in Bcor and loss of function mutations in Bcor as two such events.

Validating the tumour-suppressor function of Bcor. Protein truncating BCOR mutations have previously been reported in both solid and liquid human neoplasms. The recurrence of inactivating BCOR mutations in human tumours, a recent in vitro study showing that BCOR regulates myeloid cell proliferation and differentiation and our observations in the B-cell lineage of function mutations in Bcor and loss of Bcor and loss of function mutations in Bcor as two such events.

p19ARF-MDM2-p53 pathway or the RAS pathway would be sufficient to drive tumour development. However, our data challenges this hypothesis, at least with regard to the role of Cdkn2a, as more than one putative 'driver lesion' could be detected in some lymphomas. Examples include: (A) clone #6066 that harboured a Cdkn2a homozygous deletion and an activating NrasQ61K mutation; (B) clone ML353 that carried a Cdkn2a homozygous deletion along with an activating KrasQ61H point mutation and (C) clone #4242 that had a Cdkn2a deletion and Bcor frameshift deletion. These co-occurring lesions had similar VAF in the individual samples that were validated by deep sequencing and qPCR CNV analysis of individual clones confirmed a single cells and expanded each clone in vitro. DNA sequencing and qPCR CNV analysis of individual clones confirmed a single barcode integration and coalescing lesions in Cdkn2a and Bcor in lymphoma #4242, and Cdkn2a and Nras lesions in lymphoma #6066, confirming that both genetic lesions detected in the two lymphomas must be present in the same cell (Supplementary Fig. 8).

A surprising finding from our sequencing of Eμ-Myc lymphomas was the identification of additional putative oncogenic lesions in the context of Cdkn2a loss, namely for lymphomas #4242 (Bcor frameshift deletion), ML353 (KrasQ61H) and #6066 (NrasQ61K). To further determine if Eμ-Myc lymphomas with knockout of one Cdkn2a allele not only lost the second allele but also acquired additional mutations we crossed Eμ-Myc to Cdkn2a knockout mice11 and sequenced spontaneous Eμ-Myc;Cdkn2a+/− lymphomas. As expected Eμ-Myc;Cdkn2a+/− mice demonstrated accelerated lymphomagenesis compared with Eμ-Myc transgenic animals (Supplementary Fig. 9). Further WES analysis of six Eμ-Myc; Cdkn2a+/− lymphomas revealed that all had lost the wild type Cdkn2a allele and one lymphoma harboured a heterozygous pathogenic KrasQ61H mutation (VAF = 67%), indicating selective outgrowth of an Eμ-Myc lymphoma with loss of INK4A/ARF and expression of oncogenic Kras (Supplementary Fig. 10 and Supplementary Data 4). Collectively, our data suggest that while heterogeneous loss of Cdkn2a creates a selective pressure resulting in loss of the wild type Cdkn2a allele, it does not ameliorate selective pressure for gain- or loss-of-function of other genes outside the p53-axis. These data are consistent with a previous report indicating that p19ARF inactivation is itself insufficient to trigger lymphomagenesis in Eμ-Myc transgenic mice and indeed the authors of that study posited that 'cryptic mutations other than p19ARF loss accompany the conversion of premalignant Eμ-Myc B cells’ in the Eμ-Myc;Cdkn2a+/− setting. Herein we identify activating mutations in Bcor and loss of function mutations in Bcor as two such events.

Figure 2 | Somatic mutations detected by exome-sequencing of Eμ-Myc lymphomas. (a) Mutation frequency across 23 Eμ-Myc lymphomas from a prospective (n = 16) and retrospective series (n = 7). (b) Mutations identified in known cancer genes in Eμ-Myc lymphomas.
was significant enrichment of IgM/MYC to drive oncogenesis remains unclear. We found that there is cell proliferation or survival, so how BCOR loss co-operates with BCOR has not previously been demonstrated to control tumour range of human malignancies. 



Figure 3 | Variant allele frequency determined by targeted amplicon sequencing. Targeted amplicon sequencing (TAM-seq) was performed on the prospective Eμ-Myc tumour cohort to validate mutations from WES screen. Variant allele frequency is shown on the left y-axis for the mutations that were validated, with genes present in the cancer gene census list highlighted. The right y-axis displays relative copy number of Cdkn2a in the samples where gene deletion is suspected based on WES read depth and qPCR analysis.

described herein would strongly support a tumour-suppressive function of this gene; however this has not been formally demonstrated through functional in vivo studies. To experimentally validate the tumour-suppressor function of Bcor in Eμ-Myc lymphoma we knocked down or deleted Bcor using shRNA or CRISPR-Cas9 methods, respectively, in Eμ-Myc haematopoietic progenitor cells and transplanted them into lethally irradiated syngeneic mice (Fig. 4a). We knocked down or deleted Trp53 as a reference for expected accelerated lymphomagenesis as well as including a negative control (scrambled) shRNA and CRISPR sgRNA sequences. As expected, Eμ-Myc lymphomagenesis was significantly accelerated when p53 was knocked down or deleted (Fig. 4b,c). Strikingly, depletion or deletion of Bcor using similar techniques significantly accelerated lymphomagenesis demonstrating that Bcor restrains Myc-induced lymphomagenesis (Fig. 4b,c). Western blot analysis of Bcor depleted/deleted lymphomas confirmed the loss of Bcor protein expression (Fig. 4d,e). Furthermore, RNA-seq analysis of Eμ-Myc lymphomas with CRISPR-Cas9-mediated targeted deletion of Bcor showed deletions or insertions proximal to the expected CRISPR-Cas9 trigger sequence in exon 4 that would be expected to cause Bcor loss of function (Supplementary Fig. 11). These experiments demonstrate, for the first time, that Bcor can function in vivo as a tumour suppressor gene, and plays an important role in Myc-driven lymphomagenesis. To complement these studies, we rescued loss of Bcor in the Bcor-mutant #4242 Eμ-Myc lymphoma. We generated a traceable system introducing wild type BCOR co-expressed with GFP by retroviral transduction. Longitudinal proliferation studies in vitro demonstrated that #4242 Eμ-Myc lymphomas ectopically expressing wild type Bcor were negatively selected in a competitive proliferation assay while expression of GFP alone had no effect (Fig. 5a,b). This finding was confirmed in a similar experiment using the human Namalwa Burkitt lymphoma cell line that expressed very low endogenous levels of BCOR (Fig. 5c,d). Our novel functional data using Bcor depletion/deletion unequivocally demonstrated that mutation of Bcor, seen at a higher frequency in in Eμ-Myc lymphomas than lesions in the known cancer genes Trp53, Cdkn2a and Ras, is a ‘driver’ oncogenic event and provides mechanistic context to observations by others that mutations in Bcor occur in a range of human malignancies.

Unique gene expression signature in BcorMut lymphomas. BCOR has not previously been demonstrated to control tumour cell proliferation or survival, so how BCOR loss co-operates with MYC to drive oncogenesis remains unclear. We found that there was significant enrichment of IgM+/IgD- BcorMut lymphomas in our tumour cohort (Supplementary Fig. 12). Nine lymphomas with BCOR loss-of-function were IgM+/IgD-, while only one was IgM+/IgD-. Eight BcorWT tumours were IgM-/IgD-, while 10 were IgM+/IgD-. We therefore conclude that Bcor mutations are associated with an IgM-/IgD- profile (Chi-squared test, P<0.05). To begin to elucidate the underlying biology of the Bcor-mutant lymphomas we applied RNA-seq analysis of Eμ-Myc lymphomas with shRNA-mediated knockdown of Bcor (Eμ-Myc;shBcor, n = 6) or Trp53 (Eμ-Myc;shp53, n = 6); or Eμ-Myc lymphomas overexpressing NrasQ61K (Eμ-Myc;NrasQ61K, n = 5). Gene expression profiling identified 393 significantly differentially expressed genes in Eμ-Myc;shBcor lymphomas compared with those with knockdown of p53 or expressing oncogenic Nras (Supplementary Data 5). To determine if Eμ-Myc lymphomas with CRISPR/CAS9-mediated knockout of Bcor (Eμ-Myc;sgBcor) or those harbouring spontaneous mutations in Bcor carried similar gene expression patterns, we clustered gene expression data using the Eμ-Myc;shBcor signature gene set (Fig. 6). All sporadic Bcor mutant lymphomas and Eμ-Myc;sgBcor lymphomas clustered with the Eμ-Myc;shBcor lymphomas. This demonstrates that BCOR loss-of-function, either engineered or through spontaneous mutation, drives a unique and reproducible transcriptional pattern in Eμ-Myc lymphomas.

To determine if the BCOR loss-of-function gene expression signature provides any molecular insight into the functional interaction between MYC and BCOR we performed pathway analysis using PANTHER. This identified upregulation of TGFβ signalling as the most affected molecular pathway in Eμ-Myc;shBcor lymphomas (Bonferroni corrected P = 0.0058), with enhanced expression of TGFβ pathway members (Cited1, Bambi, Acrv2b, Smad3, Mapk12, Tgfb2). The importance of TGFβ signalling in Eμ-Myc lymphomagenesis has been previously reported, where this cytokine derived from macrophages in the tumour microenvironment augmented oncogene-induced senescence19. There it was shown that knockout of the senescence-related histone methyltransferase Suv39h1 accelerated Myc-driven lymphomagenesis with those lymphomas possessing an enhanced TGFβ signature reminiscent of the signature observed in Eμ-Myc lymphomas with loss of Bcor. It was proposed that TGFβ drives Myc-induced senescence in a Suv39h1-dependent manner raising the intriguing possibility that BCOR also regulates this unique tumour suppressive mechanism.

Discussion

In summary, we have performed the first comprehensive genome-wide sequencing analysis of lymphoid malignancies arising in the Eμ-Myc transgenic mouse model. Despite extensive use of the model for cancer gene discovery through forward genetic
of-function may subvert TGFβ signalling to drive lymphoma development. Collectively our studies provide important information on the genomic architecture of one of the most utilized GEMMs, identifies Bcor as a bona fide tumour suppressor gene and provides important information regarding the functional interaction between Myc and p19ARF in the context of Myc-Myb lymphomagenesis.

Methods

**Eμ-Myc Lymphomas.** All animal work was performed with approval from the Peter MacCallum Cancer Centre Animal Experimentation Ethics Committee. A prospective series of sixteen heterozygote Eμ-Myc mice originally obtained from the Walter and Eliza Hall Institute (Melbourne, Australia) on a C57BL/6 background were bred for discovery analysis and a retrospective series of seven lymphomas were collated from archived laboratory resources. The age and gender of mice used in the study are detailed in Supplementary Table 1. Matched normal tail tissue was available for lymphoma #88 and a littermate control ML62. We collected peripheral blood for isolation of B-cells from 4-weeks of age and at 2-week intervals until the mice were sacrificed with advanced disease.

**Nucleic acid extraction.** DNA was extracted using the DNeasy blood and tissue kit (Qiagen, Hilden, Germany). RNA was extracted using the miRNeasy mini kit (Qiagen) and nucleic acids were quantified by spectrophotometer. RNA integrity was assessed using the Agilent Bioanalyzer 2100 (Agilent, CA, USA).
Whole-genome sequencing. Approximately 1 μg of gDNA was prepared for fragment-sequence libraries and then processed according to standard protocols using TruSeq chemistry (Illumina, San Diego, CA, USA). Mate-pair libraries were prepared using Nextera Mate Pair Sample Prep Kit according to manufacturer instructions (Illumina, San Diego, CA, USA). Paired-end sequencing (2 × 100 bp) on HiSeq2000 (Illumina) was applied to achieve 1-30-fold mapped read coverage across the entire genome depending on required depth (Supplementary Data 7). Reads were aligned using Bowtie2 v2.1.0 (Langmead and Salzberg 2012) to mouse reference genome mm10 (Dec. 2011 GRCm38/mm10). Duplicates were identified and removed from the aligned data using Picard’s Mark Duplicates v1.89 (Broad Institute, Boston, MA, USA). BreakDancer v1.3.5 (ref. 37), CREST v0.0.1 (ref. 38) and Socrates v0.9.5 (ref. 39) were employed to identify structural variants. Whole genome de novo assembly was undertaken using Gossamer to compare contigs with the genomic sequence relevant to the transgene. Control-FREQ (ref. 41) was used to call somatic copy-number alterations using taI88 as the germline control. A window size of 50,000 bp was used.

Exome sequencing. Approximately 1 μg of gDNA prepared as above was processed for whole exome-capture sequencing according to standard protocols (Agilent Technologies, Santa Clara, CA, USA) on a HiSeq 2000 achieving a mean sequence depth of 110-fold and at least 10-fold coverage in 98% of targeted bases (Supplementary Data 8). Variant calling methods were based on Genome Analysis Tool Kit (GATK)42,43 and Strelka44. An intersection of variant callers was used to optimize specificity and sensitivity by including only SNVs called by at least two methods. Variants were annotated using the ENSEMBL database45. Somatic copy-number calls from exome data were generated using ADTEx (ref. 46) against three germline control. A window size of 50,000 bp was used.

RNA sequencing and data analysis. Approximately 1 μg of RNA was used to generate polyA enriched cDNA libraries using TruSeq sample preparation kit (Illumina, San Diego, CA, USA) and paired-end RNA-seqencing (2 × 50 bp) was performed on the HiSeq 2000, generating 4 × 10^7 reads per sample. Reads were quality checked using FastQC and trimmed if necessary for low base quality or adaptor using Cutadapt v1.6 (ref. 47). RNA-seqencing reads were aligned using the short read aligner software – TopHat48. The raw RNA-seqencing data was converted to feature counts using HTSeq software package allowing generation of expression matrices collating read counts per gene based on ENSEMBL annotation45. Edger v3.0 (Biorad, Fredd Hutchinson Cancer Research Center, Seattle, WA, USA) statistical software was used in the statistical programming tool R (http://www.r-project.org/) to normalize gene expression and perform statistical analysis by linear regression49. Gene Cluster 3.0 was used to cluster samples and genes by average linkage centred clustering. Cluster diagrams and heat maps were visualized and then exported from Treeview v1.6.2. Pathway mapping of significant genes (P value < 0.05) was undertaken using PANTHER pathway analysis software50.

Variant validation by amplicon-based gene sequencing. Targeted amplicon-based massively-parallel sequencing was used to validate mutations detected by exome-sequencing and for analysis of serial blood samples. PCR Primers were designed flanking mutated bases incorporating a 5-prime extended universal sequence for secondary PCR enabling addition of full-length illumina adaptors (Supplementary Data 9). Primary PCR reaction mixture included 10 ng genomic DNA template, 10 μl of 2 × Phusion high-fidelity PCR master-mix (New England Biolabs, Ipswich, UK), 2.5 μl of 4 μm forward and reverse oligonucleotide primers and made up to contain a total reaction volume of 20 μl with PCR-grade H2O. The PCR reaction mixture was cycled at 95 °C for 2 min, (95 °C for 30 s, 68 °C for 1 min) for 15 cycles and (95 °C for 30 s, 68 °C for 1 min) for 32 cycles and 72 °C for 5 min. Primary PCR products were then pooled and combined with AMPure XP magnetic beads (Beckman Coulter, Pasadena, CA, USA) in a ratio of 1:9.9 and then purified as per manufacturer protocol. Purified pooled amplicons were diluted 1:100 in PCR-grade H2O and prepared in PCR reaction mix of 1 μl template DNA, 10 μl of 2 × Phusion high-fidelity PCR master-mix, 2.5 μl of 4 μm forward and reverse oligonucleotide primers and made up to contain a total reaction volume of 20 μl with PCR-grade H2O. The PCR reaction mix was then subject to an additional PCR cycle using indexed Fluidigm CS1 and CS2 primers (Fluidigm, CA, USA) at conditions of 95 °C for 5 min, (95 °C for 15 s, 60 °C for 30 s, 72 °C for 1 min) for 15 cycles and 72 °C for 3 min. Secondary PCR products were then purified using AMpure XP.
All CrispR.BcorT2 and sporadic bar represents median normalized log2-fold gene-expression.

Nras pLMS.sh.

Figure 6 | Gene-expression profiling of Eμ-Myc lymphomas identified a reproducible signature of Bcor mutation or knockdown. RNA-seq analysis was first used to identify 393 significantly differentially expressed genes (FDR-corrected < 0.05 and log, fold-change ≥ 2) between pLMS.sh.Bcor

and combined pLMS.sh.Bcor

and overexpressing NrasQ61K fetal liver-derived Eμ-Myc lymphomas. RNA-seq data for shRNA fetal liver-derived lymphomas plus CRISPR-Cas9 (CRISPR.BcorT2) and sporadic Eμ-Myc lymphomas was then clustered using the 393 gene set. All CRISPR.BcorT2 and sporadic Bcor mutant lymphomas (red arrows) cluster with the pLMS.sh.Bcor overexpressing fetal liver-derived lymphomas. The sporadic NrasQ61K mutant cell line #6066 (blue arrow) also clustered correctly with NrasQ61K overexpressing fetal liver-derived lymphomas. Heatmap and scale bar represents median normalized log2-fold gene-expression.
2. Andrechek, E. R. & Nevins, J. R. Mouse models of cancers: opportunities to address heterogeneity of human cancer and evaluate therapeutic strategies. *J. Med. Med.* 88, 1095–1100 (2010).

3. Kim, M. et al. Comparative oncogenomics identifies NEDD9 as a melanoma metastasis gene. *Cell* 125, 1269–1281 (2006).

4. McFadden, D. G. et al. Genetic and clonal dissection of murine small cell lung carcinoma progression by genome sequencing. *Cell* 156, 1298–1311 (2014).

5. Masers, R. S. et al. Chromosomally unstable mouse tumours have genomic alterations similar to diverse human cancers. *Nature* 447, 966–971 (2007).

6. Wartman, L. D. et al. Sequencing a mouse acute promyelocytic leukaemia genome reveals genetic events relevant for disease progression. *J. Clin. Invest.* 121, 1445–1455 (2011).

7. Adams, J. M. et al. Integrative analysis reveals selective 9p24.1 amplification, altered PD-1 ligand expression, and further induction via JAK2 in nodular sclerosing Hodgkin lymphoma and primary mediastinal large B-cell lymphoma. *Blood* 122, 3169–3177 (2013).

8. Grossmann, V. et al. Whole-exome sequencing identifies mutations of BCR in acute myeloid leukaemia with normal karyotype. *Blood* 118, 5163–5166 (2011).

9. Chen, K. et al. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat. Methods* 6, 677–681 (2009).

10. Wang, J. et al. CREST maps somatic structural variation in cancer genomes with base-pair resolution. *Nat. Methods* 8, 652–654 (2011).

11. Mori, S. et al. Socrates: identification of genomic rearrangements in tumour genomes by re-aligning soft clipped reads. *Bioinformatics* 30, 1064–1072 (2014).

12. Conway, T., Wazny, J., Bromage, A., Zobel, J. & Beresford-Smith, B. Gossamer—a resource-efficient de novo assembler. *Bioinformatics* 28, 1913–1920 (2012).

13. Schroder, J. et al. MicroRNA-142 is mutated in about 20% of diffuse large B-cell lymphomas. *Blood* 118, 1269–1281 (2011).

14. De Keersmaecker, K. et al. Precision weights unlock linear model analysis tools for RNA-seq read counts. *Gene* 510, 732 (2014).

15. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105–1111 (2009).

16. Law, C. V. et al. Detecting somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* 31, 213–219 (2013).

17. Amasarisinghe, K. C. et al. Dissecting p53 tumor suppressor functions with RNA-Seq content using next-generation sequencing data. *Bioinformatics* 28, 123–125 (2012).

18. De Keersmaecker, K. et al. Integrative analysis reveals selective 9p24.1 amplification, altered PD-1 ligand expression, and further induction via JAK2 in nodular sclerosing Hodgkin lymphoma and primary mediastinal large B-cell lymphoma. *Blood* 122, 3169–3177 (2013).

19. Geisinger-Kern, P., Schmitt, C. A., Frenzel, A. et al. The mTORS1 inhibitor everolimus prevents and treats Emu-Myc lymphoma by restoring oncogene-induced senescence. *Cancer Discov.* 3, 82–95 (2013).

20. Bolitho, P. et al. Perforin-mediated suppression of B-cell lymphoma. *Proc. Natl Acad. Sci. USA* 106, 7273–7278 (2009).

21. Fusello, A. et al. Histone H2AX suppresses translocations in lymphomas of Emu-c-Myc transgenic mice that contain a germline amplicon of tumor-promoting genes. *Cell Cycle* 12, 2867–2875 (2013).

22. Green, M. R. et al. Integrative analysis reveals selective 9p24.1 amplification, increased PD-1 ligand expression, and further induction via JAK2 in nodular sclerosing Hodgkin lymphoma and primary mediastinal large B-cell lymphoma. *Blood* 116, 3268–3277 (2010).

23. Lee, S. C. et al. Polycomb repressive complex 2 (PRC2) suppresses Emu-myc lymphoma. *Blood* 122, 2654–2663 (2013).

24. Wendel, H.-G. et al. Survival signalling by Akt and eIF4E in oncogenesis and cancer therapy. *Nature* 428, 332–337 (2004).

25. Grabinger, B. C. et al. A diverse array of cancer-associated MTOR mutations are hyperactivating and can predict rapamycin sensitivity. *Cancer Discov.* 4, 554–561 (2014).

26. Zhang, J. et al. Genetic heterogeneity of diffuse large B-cell lymphoma. *Proc. Natl Acad. Sci. USA* 106, 1398–1403 (2013).

27. Kwanhian, W. et al. MicroRNA-142 is mutated in about 20% of diffuse large B-cell lymphoma. *Cancer Med.* 1, 141–155 (2012).

28. De Keersmaecker, K. et al. Exome sequencing identifies mutation in CNOT3 and ribosomal genes RPL5 and RPL10 in T-cell acute lymphoblastic leukemia. *Nat. Genet.* 45, 186–190 (2013).

30. Futreal, P. A. et al. A census of human cancer genes. *Nat. Rev. Cancer* 4, 171–183 (2004).

31. Schmitt, C. A. et al. Dissecting p53 tumor suppressor functions in vivo. *Cancer Cell* 1, 289–298 (2002).

32. Bertwistle, D. & Sherr, C. J. Regulation of the Arf tumor suppressor in Emiuro-Myc transgenic mice: longitudinal study of Myc-induced lymphomagenesis. *Blood* 109, 792–794 (2007).

33. Damm, F. et al. BCOIR and BCOIR1 mutations in myelodysplastic syndromes and related disorders. *Blood* 122, 3169–3177 (2013).

34. Grossmann, V. et al. Whole-exome sequencing identifies mutations of BCR in acute myeloid leukaemia with normal karyotype. *Blood* 118, 5163–5166 (2011).

35. Pugh, T. J. et al. Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. *Nature* 488, 106–110 (2012).

36. Cao, Q. et al. BCOIR regulates myeloid cell proliferation and differentiation. *Leukemia* 30, 1155–1165 (2016).

37. Schütz, A. et al. The tumor suppressor p53 is hyperactivating in most human cancers: implications for cancer therapy. *Cell Cycle* 10, 1445–1455 (2011).

38. Wall, M. et al. Somatic mutations in the B-cell lymphoma 22 gene are hyperactivating and can predict rapamycin sensitivity. *Cancer Discov.* 3, 82–95 (2013).

39. Boeva, V. et al. Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics* 28, 423–425 (2012).

40. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *Eur. J. Biochem.* 316, 1982–1988 (2012).

41. Boeva, V. et al. Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics* 28, 423–425 (2012).

42. DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43, 491–498 (2011).

43. Gerhard, D. M. et al. The Human Protein Reference Database—2009 update. *Nucleic Acids Res.* 37, D769–D774 (2009).

44. Boeva, V. et al. Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics* 28, 423–425 (2012).

45. Schütz, A. et al. The tumor suppressor p53 is hyperactivating in most human cancers: implications for cancer therapy. *Cell Cycle* 10, 1445–1455 (2011).
How to cite this article: Lefebure, M. et al. Genomic characterisation of Eμ-Myc mouse lymphomas identifies Bcor as a Myc co-operative tumour-suppressor gene. Nat. Commun. 8, 14581 doi: 10.1038/ncomms14581 (2017).

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.