Aging predisposes B cells to malignancy by activating c-Myc and perturbing the genome and epigenome

Anastasia V. Shindyapina[1]*, José P. Castro[1*$], Alessandro Barbieri[2], Olga S. Strelkova[3,4], João A. Paulo[5], Csaba Kerepesi[1], Anna P. Petrashen[6], Marco Mariotti[1], Margarita Meer[1], Yan Hu[1], Grigoriy Losyev[1], Artur A. Indzhykulian[3,4], Steven P. Gygi[5], John M. Sedivy[6], John P. Manis[2], Vadim N. Gladyshev[1#]

[1] Brigham and Women's Hospital, Harvard Medical School, Boston, 02115, USA, [2] Boston Children's Hospital, Harvard Medical School, Boston, 02115, USA, [3] Eaton-Peabody Laboratories, Massachusetts Eye and Ear Infirmary, Boston 02114 USA, [4] Department of Otolaryngology – Head and Neck Surgery, Harvard Medical School, Boston, 02114, USA, [5] Department of Cell Biology, Harvard Medical School, Boston, 02115, USA, [6] Department of Molecular Biology, Cell Biology and Biochemistry, Brown University, Providence, RI 02912, USA

* These authors contributed equally

[$] *Current address: i3S, Instituto de Investigação e Inovação em Saúde, Universidade do Porto, 4200-135, Porto, Portugal and Aging and Aneuploidy Laboratory, IBMC, Instituto de Biologia Molecular e Celular, Universidade do Porto, 4200-135, Porto, Portugal,*

[#] Corresponding author: vgladyshev@rics.bwh.harvard.edu

Abstract

Age is the single major risk factor for human cancer; however, naturally occurring cancers are rarely studied in aged animal models. Laboratory mouse strains spontaneously develop cancer with age and some predominantly die from B-cell lymphoma. Here, we uncover how B-cell lymphoma develops as a consequence of the aging immune system. We found that aged B cells undergo clonal expansions driven by genetic and epigenetic changes and established cell and spleen size as early markers of malignant transformation. High-throughput and omics assays of aged B cells and the use of mouse models revealed that c-Myc is a master regulator of B cell size and clonal expansion. A single-cell RNA-seq analysis suggested that clonal B cells originate from age-associated B cells, memory B cells that accumulate during aging. Further studies showed that c-Myc becomes activated in B cells in response to the aging microenvironment. Thus, c-Myc, aging environment, somatic mutations and the epigenome cooperate to give rise to clonal age-accelerated B cells, which we named Myc+ cells. We further show the relevance of this model to aged human B cells in blood and spleen. This study characterized a first mouse model that captures a natural transition of B cells to a prevalent type of cancer during aging.

Introduction

Aging involves a progressive accumulation of damage, whose rate is influenced by genetic, environmental and stochastic factors (Gladyshev, 2016), leading to a loss of function and ultimately to the increased risk of death. Aging is also the single major risk factor for several chronic diseases (Niccoli and Partridge, 2012), thereby limiting both lifespan and healthspan. Understanding the molecular causes of aging may help develop tools to treat and prevent age-related diseases, such as cancer. The risk of developing invasive cancer grows exponentially with age starting from midlife (Niccoli and Partridge, 2012). Although cancer can occasionally develop at a young age, the median age of its diagnosis in the US is 66 years old (Howlader et al., 2020). Mortality rate for cancer and aging follows the same trend consistent with the idea that cancer is a disease of aging (Zenin et al., 2019). In recent years, aging tissues were shown to

accumulate somatic mutations (Martincorena et al., 2018; Yizhak et al., 2019; Yokoyama et al., 2019). Clonal hematopoiesis of indeterminate potential (CHIP) was established as a link between aging and cancer (Genovese et al., 2014; Jaiswal et al., 2014). Clonal mutations in the aged blood fall onto genes involved in DNA methylation, which underlies a crucial role of epigenetic changes in clonal selection of immune cells. However, little is known about the molecular mechanisms that give advantage to one clone over another, and the role of the aging environment in the selection process.

Aberrant DNA methylation is known to drive cancer, especially blood malignancies. Mouse knockout models of DNA methylation regulators, *Tet1*, *Tet2*, *Tet3* and *Dnmt3a,* develop myeloid leukemias or lymphomas (An et al., 2015; Mayle et al., 2015; Zhao et al., 2015). Age-related clonal expansion of blood cells in humans is predominantly driven by mutations in *TET2*, *DNMT3A*, and *ASXL1* (Genovese et al., 2014; Jaiswal et al., 2014). Human cancer cells often exhibit a remodeled DNA methylome, in particular hypermethylated promoters (Esteller et al., 2001; Herman and Baylin, 2003; Saghafinia et al., 2018). At the same time, recurrent age-related blood DNA methylation changes provide the foundation for the development of epigenetic aging clocks (Hannum et al., 2013; Horvath, 2013; Petkovich et al., 2017). It is possible that some aging cells acquire fitness advantage through an altered DNA methylome, allowing them to outcompete other cells and ultimately progress to cancer. Interestingly, *Dnmt3a* and *Tet2* knockout mice develop tumors with one to two-year latency indicating that the DNA methylation change alone is insufficient for malignant transformation of immune cells. Epigenetic changes may cooperate with the aging environment and somatic mutations in clonal selection and subsequent malignancy of blood cells.

c-Myc is among most potent oncogenes, but alone it can hardly initiate neoplastic transformation (Gabay et al., 2014). High c-Myc levels stimulate cell proliferation or transition to apoptosis as a possible defense mechanism against cancer (Gabay et al., 2014). Transgenic c-Myc mice develop cancer (Adams et al., 1985; Harris et al., 1988), while c-Myc haploinsufficient mice are protected from it and are longer lived (Hofmann et al., 2015). Interestingly, c-Myc also drives cell competition, a hallmark of cancer, in mouse heart and embryos (Clavería et al., 2013; Muñoz-Martín et al., 2019) as well as in healthy Drosophila tissues (de la Cova et al., 2004; Di Gregorio et al., 2016; Simpson and Morata, 1981). Therefore, c-Myc is a great candidate for providing growth advantage to age-related clones, constituting an early step in the transition to cancer.

Aging microenvironment is thought to be a major driver of cancer transformation (Fane and Weeraratna, 2020), e.g. manifesting through chronic inflammation (Coussens and Werb, 2002; Greten and Grivennikov, 2019). Age-associated B cells (ABC) are a prominent example of the contribution of the pro-inflammatory environment. ABCs are a group of cells negative for CD21 and CD23, that express T-box transcription factor *Tbx21*, and present in patients with autoimmune diseases (Rubtsov et al., 2011; Rubtsova et al., 2013). ABCs replace healthy B cells in mouse spleen and blood with age and express pro-inflammatory cytokines such as IFNg, IL-6, and IL-4 (Hao et al., 2011). It is an attractive possibility that they could provide an environment that supports malignant transformation of aging cells. Myeloid bias (age-related generation of myeloid over lymphoid cells) is another prominent feature of the aging immune system with the potential to create a pro-inflammatory environment (Elias et al., 2017). However, it remains unclear whether ABCs or myeloid bias play any role in age-related clonal selection and cancer.

Until now, there was no appropriate animal model to test the role of c-Myc, DNA methylation and their interaction with the aging environment in cancer and clonal selection. As commonly used C57BL/6 and BALB/c mice spontaneously develop and die from B-cell lymphoma (BCL) at old age (Brayton et al., 2012; Turturro et al., 2002),

we hypothesized that aged mice develop clonal expansions of B cells that ultimately progress to lymphoma, and can serve as a convenient model to study mechanisms by which cancer develops with age. Indeed, we demonstrate that c-Myc, somatic mutations and recurrent DNA methylation alterations support clonal expansion of aged B cells marked by a conspicuous B cell size increase. Human B cells also clonally expand with age in blood and spleen driven by somatic mutations, demonstrating relevance of the aging mouse model.

Results

B cell size increases with age

Larger B cells is a common feature of human age-related B-cell lymphoma (BCL), such as follicular and diffuse large BCL (Ott et al., 2013), that also carries poorer prognosis. Another prominent feature of BCL is splenomegaly (i.e., enlarged spleen). Since aging rodents die predominantly from BCL (Brayton et al., 2012; Turturro et al., 2002), we first examined changes in B cell size and spleen weight with age in young (6 months old) and old (27 months old) mice. We found that aged mice developed both splenomegaly and increased cellularity (Fig.1A,B). Aged spleen was characterized by the decreased proportion of B cells and increased proportion of myeloid cells (Fig.1C, Fig.S1A,B), in agreement with previous reports on myeloid bias (Elias et al., 2017). We further found a striking increase of B cell size with age in both spleen and blood using forward scatter (FSC) in flow cytometry as a measure of cell size (Fig.1D,E). T cells, but not myeloid cells, also increased in size, although much less prominently than B cells (Fig.S1C).

As splenic B cells mature through multiple stages and change their size during these transitions, we analyzed various B cell populations to determine which population accounted for the average B cell size increase. B cells were stained for surface markers of follicular, marginal zone, germinal center, and plasma cells, and examined for their cell size and relative abundance by flow cytometry (Fig.S1D-F). The proportions of follicular and marginal zone B cells significantly decreased with age, while $CD21^{-}CD23^{-}$ B cells populated most of the B-cell compartment (Fig.1F,H,Fig.S1F) (Hao et al., 2011). Of all populations, splenic and blood $CD21^{-}CD23^{-}$ B cells increased in size most dramatically (Fig.1G,I), and follicular and marginal zone B cells slightly increased in the spleen. The size of other B cells remained unchanged with age (Fig.S1G). We thus conclude that accumulation and cell size increase of $CD21^{-}CD23^{-}$ B cells are drivers of average cell size enlargement that we observed in bulk aged B cells.

To examine earlier stages of B cell maturation, we analyzed bone marrow (Fig.S1H) and observed an age-related increase in the ratio of Pro- and IgM+ immature B cells to $B220^{+}$ cells and a drop in Pre-B cells (Fig.S1I), in agreement with previous reports (Labrie et al., 2004). However, there was no change in B cell size (Fig.S1J). In addition to C57BL/6 mice, we examined old BALB/c mice and found that they exhibit most of the phenotypes of C57BL/6 mice: spleen enlargement, an increase in average B cell size and cell size of $CD21^{-}CD23^{-}$ B cells, and an increase in splenic cell number with age (Fig.S2). Thus, increased cell size, splenomegaly and accumulation of $CD21^{-}CD23^{-}$ B cells are universal in aging mice. As those phenotypes take place in the spleen and blood, but not in the bone marrow, we further focused on these tissues.

Because aged mice mostly die of lymphoma, and human clonal hematopoiesis precedes blood cancers, we tested whether B cells undergo clonal expansion with age. We further hypothesized that enlarged B cells characteristic of aged spleens are premalignant. We thus sorted B cells according to their size and sequenced CDR3 regions of Ig heavy chains (Immunoseq) to calculate Ig diversity and clone size (Fig.1J-

P). We defined enlarged B cells as cells exceeding the size of young B cells based on FSC (Fig.1J). We analyzed three groups : B cells ($CD45^{+}CD3^{-}CD11b^{-}CD19^{+}CD80^{-}$) from young animals, regular sized B cells from old animals and enlarged B cells from old animals (FSC^{hi}). Clone size was calculated as percentage of reads supporting the clone divided by the total number of reads in the sample, with the larger proportion representing larger B cell clones (Fig.1K). Strikingly, aged B cells contained large expanded clones, with single clones occupying on average 38% of a sample (Fig.1L). Larger clones were formed by enlarged B cells and originated from larger spleens (Fig.1M). Accordingly, Ig diversity (Fig.1N) decreased with age, and the difference was more pronounced in enlarged B cells (Fig.1O). Interestingly, enlarged and regular sized B cells belonged to the same Ig clones (Fig.1P), suggesting that enlarged B cells originate in the spleen from regular sized B cells. Moreover, top clones did not carry somatic mutations within CDR3 regions (Fig.S3A), indicating that aged clonal B cells have not undergone somatic hypermutation in germinal centers. We further found that the CDR3 regions of 5 top clones of old enlarged B cells were significantly shorter than those of young B cells (Fig.S3B). Together, these findings demonstrate that aged mouse B cells are clonally expanded, accompanied by an increase in B cell size.

Mouse and human B cells undergo clonal expansions with age

As B cells may clonally expand for many reasons, we sequenced the exomes of clonal aged B cells and examined their somatic mutations. For each animal, we performed whole exome sequencing (WES) on three types of samples: regular sized B cells, enlarged B cells (same samples for which we performed Immunoseq), and liver as a healthy tissue. We found that aged B cells carry hundreds of somatic mutations (300-2,000 per sample, depending on filter threshold for coverage), but only 5-10% of them were shared between regular and enlarged B cells, suggesting an independent path for clonal selection of enlarged B cells. Enlarged B cells formed larger clones as measured by variant allele fraction (VAF) (Fig.2A). Top clones of mouse B cells carried somatic mutations in genes commonly mutated in human lymphoma, e.g. *Trp53*, *Pim1* and *Myh11* (Fig.2B), suggesting common genetic mechanisms of clonal selection between human and mouse B cells. We further estimated top clone size using VAF and examined them against Ig clone size for the same samples. The largest Ig clones and genetic clones were almost identical in size in all of the samples (Fig.2C), suggesting that mutated B cells expand after VDJ rearrangement, most likely in the spleen. This finding also suggests that Ig clones are dependent on the identified somatic mutations. B cell size significantly correlated with genomic clone size (Fig.2D), again suggesting that B cell size is a convenient marker of premalignant B cells, especially when surface markers are unknown.

To distinguish malignant and premalignant cases in our analysis, we performed necropsy of mice analyzed by WES and Immunoseq and found only 3 cases of lymphoma out of 12 (Fig.S7). We conclude that clonal expansions, splenomegaly and B cell size can serve as markers of both malignant and premalignant B cells, so they can be conveniently used to detect cancer at early stages.

To examine the origin of somatic mutations in B cells, we deconvoluted the observed mutation patterns into mutational signatures (Alexandrov et al., 2020). Mutations were enriched in ID5 and DBS7 signatures (Fig.2E,G). Burden of mutations in the ID5 signature has a clock-like behavior and correlates with the age of cancer diagnosis; this signature is viewed as an aging signature. The DBS7 signature is associated with defective DNA mismatch repair, suggesting that B cells accumulate mutations in part due to impaired DNA repair mechanisms. Single base substitutions were enriched in SBS29 (73.8% of mutations), SBS42 and SBS46 signatures, with small contribution from SBS5 (Fig.S3C). The suggested etiology of SBS29 is tobacco chewing; however, in a recent analysis of human cancers this signature was detected in kidney, bladder

and liver cancers (Alexandrov et al., 2020). As the analyzed mice were not exposed to exogenous mutagens, SBS29 may have an unknown endogenous origin.

It was reported that Ig repertoire and clone size can be reliably reconstructed from RNA-seq data using mixcr software (Bolotin et al., 2017). Indeed, the clone sizes for our B-cell samples estimated by Immunoseq and RNA-seq/mixcr strongly correlated (Fig.S3D). Thus, we used mixcr to analyze diversity of Ig repertoire in human aging spleen and blood from The Genotype-Tissue Expression (GTEx) project (GTEx Consortium, 2013). Similar to mice, diversity of Ig repertoire decreased with age in both blood and spleen (Fig.S3E,F). Using GTEx whole genome sequences of blood, we predicted cases of CHIP and found that they have an even more pronounced decrease of Ig repertoire diversity (Fig.S3G). Moreover, clone size of mutated B cells (based on VAF) significantly correlated with top Ig clone size (Fig.S3H), indicating that a) CHIP somatic mutations are crucial for B cell clonal expansion with age, and b) B cells clonally expand after rearranging VDJs, most likely in spleen. Overall, the data suggest that B cells clonally expand with age both in mice and humans and do this in a similar manner.

Aged B cells mimic the transcriptional and epigenetic programs of malignant cells

To get insights into molecular mechanisms that drive clonal expansions with age, we examined gene expression, protein levels, protein phosphorylation, and DNA methylation of the same B cells analyzed for the Ig repertoire. PCA analyses of gene expression (Fig.3A), gene promoter methylation (Fig.3B) and protein expression (Fig.3C) revealed significant perturbations in clonal and enlarged B cells. Aged B cell samples containing the largest B cell clones and larger cells clustered apart from young B cells and unaffected aged B cell samples (Fig.3A-C). We further carried out reduced representation bisulfute sequencing (RRBS) of the samples and subjected them to the analyses of biological age using aging clocks. The estimated age of clonal B cells measured by two DNA methylation clocks significantly correlated with clone size and B cell size (Fig.3D,E, Fig.S4C). Thus, clonally enlarged B cells are epigenetically older than regular sized B cells within the same animal.

To further investigate the molecular mechanisms driving clonal expansion, we correlated gene expression, protein expression and average DNA methylation of promoters and gene bodies for each gene with the top clone size in a sample. These analyses revealed that top B cell clones upregulate targets of transcriptional factor c-Myc, metabolic pathways, MTORC1 and cell cycle genes (Fig.3F). Clonal B cells also up-regulated ribosome biogenesis pathways (Fig.S4A,B). Up-regulation of c-Myc targets was further confirmed at the protein level based on proteomic analyses by MS/MS (Fig.3F). Most DNA methylation changes do not seem to be involved in the regulation of these pathways, except for oxidative phosphorylation, which was up-regulated in clonal samples and lost promoter methylation at the same time (Fig.3F). Targets of c-Myc were significantly up-regulated in enlarged B cells (Fig.3G), and c-Myc tended to up-regulate genes with lower promoter methylation (Fig.3H). These findings suggest that c-Myc drives clonal expansion and cell size increase with age.

DNA methylation changes in clonal cells mostly suppressed gene expression (Fig.3I). In line with that and with what is known about human cancer, we observed an increase in global methylation of gene promoters in clonal cells (Fig.3J). We also analyzed promoter methylation of tumor suppressor genes. Negative correlation was found between promoter methylation and gene expression of *Id3* and *Mitf* in clonal B cells (Fig.3K). *Id3* is a human tumor suppressor gene recurrently mutated in Burkitt lymphoma (Greenough and Dave, 2014; Richter et al., 2012), while *Id3* knockout resulted in T cell lymphomas (Li et al., 2010), accelerated B-cell lymphoma in Myc

transgenic mice (Love et al., 2015), impaired B cell development, and enhanced B cell proliferation (Chen et al., 2016). We further found that oncogene *Ptpn1* was up-regulated in clonal B cells, while its promoter lost methylation (Fig.S4D). The data suggest that DNA methylation plays an important role in clonal selection by silencing tumor suppressor genes and activating oncogenes, as well as by global silencing of gene expression through increased methylation of promoters.

We also carried out phosphoproteome profiling of clonal and enlarged B cells, which revealed activation of MAPK1 and p38a kinases (Supplementary Table 1). Both kinases are known to be activated by extracellular signals of growth hormones and pro-inflammatory cytokines (Zhang and Liu, 2002), and c-Myc is a well described downstream target. The data suggest that MAPK1 and p38a activate c-Myc expression in B cells in response to extracellular signals, possibly from the aging microenvironment.

Single-cell RNA-seq reveals a novel B-cell population that is clonal and may originate from age-associated B cells

To shed light on the origin of clonal B cells, we analyzed a single-cell RNA-seq dataset of aged mouse spleen (Kimmel et al., 2019). We first established which B cells are age-related and clonal. There were two major clusters of B cells specific for aged spleens (Kimmel et al., 2019). We found that both of them were negative for CD21 (*Cr2*) and CD23 (*Fcer2a*) (Fig.S5A), and therefore represented the $CD21^{-}CD23^{-}$ population in our flow cytometry analysis (Fig.1H,I). One of these populations included ABCs , based on the expression of canonical ABC markers: T cell transcription factor T-bet (*Tbx21*) and CD11c (*Itgax*) (Rubtsova et al., 2013) (Fig.S5B, Fig.4A). The other cluster showed elevated expression of c-Myc targets (Fig. S5D), and we named it the Myc+ cluster (Fig.4A). To test which aged B-cell population was clonal, we calculated a clonality signature score for each cell as the average expression of 50 top genes minus average expression of 50 bottom genes from the clonality signature of bulk RNA-seq (Supplementary Table 2). With this metric, B cells with higher scores should be clonally expanded, and we found that Myc+ B cells had the highest scores among aged B cells (Fig.4E). We further reconstructed CDR3 regions of Ig kappa chains in individual cells and found that most B cells from the largest clone belonged to the Myc+ cluster (Fig.4B). Myc+ B cells harbor cells from larger CDR3 clones than either ABC or regular B cells (which are mostly follicular (FO) and marginal zone (MZ) B cells) (Fig.4F). Note that the clonality signature score significantly correlated with Igk clone size (Fig.S5C), demonstrating predictive power of the signature. We further compared the transcriptomes of B cells from the largest clone to the smallest clones and found significant enrichment for c-Myc targets in clonal B cells (Fig.4D).

Since both aged B-cell populations were $CD21^{-}CD23^{-}$, Myc+ cells appear to originate from ABCs. We found that both populations shared many deregulated genes compared to the rest of splenic B cells that were mostly represented by follicular and marginal zone B cells (FO/MZ) (Fig.4C). Myc+ and ABCs also shared CDR3 Igk sequences (Supplementary Table 3). Thus, premalignant Myc+ B cells may arise from ABCs, providing a further link between aging and cancer transformation.

We further applied scRNA-seq to identify surface markers for Myc+ and ABC cells. We established CD29 as a surface marker for both populations and CD24a as a surface marker for Myc+ cells (Fig.4G). Flow cytometry of young and old spleens confirmed that Myc+ cells ($CD19^{+}CD21^{-}CD23^{-}CD29^{+}CD24^{+}$) accumulate with age (Fig.4H, Fig.S5E) and revealed that they are larger than follicular B cells (Fig.4I), which we further confirmed by immunofluorescence microscopy (Fig.4J,K). Moreover, Myc+ cells had higher levels of c-Myc fluorescence intensity suggesting higher levels of protein (Fig.4L), in agreement with single-cell RNA-seq analysis results. We conclude that

aged clonal B cells are represented by a novel population of Myc+ B cells, rather than ABC or FO/MZ B cells, and that c-Myc is a potential driver of B cell clonal expansion and cell size increases with age. We also established a panel of antibodies to differentiate premalignant Myc+ cells from ABCs in downstream analysis.

We analyzed human blood RNA-seq from GTEx for changes in gene expression of markers of Myc+ and ABC cells with clonal expansions, similar to the mouse model. We found that expression of ABC markers changed with the increase of CDR3 clones similarly to old mouse B cells: CD21 (*CR2*), CD23 (*FCER2*) and *CD19* decreased, while CD29 (*ITGB1*) increased (Fig.S5F). This analysis suggests that ABC cells may accumulate in human blood with age, and that their accumulation is associated with B cell clonal expansion.

Myc as a driver of cell size increase and clonality

In addition to RNA-seq enrichment results, qRT-PCR confirmed higher c-Myc expression in enlarged B cells (Fig.5A). It also corresponded to the elevated RNA/DNA ratio (Fig.5B), rRNA abundance (Fig.5C) and the larger size of the nucleus (Fig.5D), pointing to an increased ribosome biogenesis and cell mass, the traits controlled by c-Myc (Dang, 2012). To test a causal role of c-Myc in age-related clonal expansions, we analyzed c-Myc heterozygous mice (*Myc*$^{+/-}$) (Hofmann et al., 2015) and mice transgenic for c-Myc specifically in B cells (Emu-Myc) (Harris et al., 1988). *Myc*$^{+/-}$ mice are longer-lived and have a reduced lymphoma burden, whereas the lifespan of Emu-Myc mice is greatly reduced and they die almost exclusively of BCL. We found that the loss of one copy of c-Myc protected mice from splenomegaly (Fig.5E,F), B cell size increase (Fig.5G), and RNA accumulation in B cells (Fig.5H), compared to age-matched wild type (Wt) siblings. Moreover, Ig diversity was better preserved in *Myc*$^{+/-}$ mice with age, than in Wt controls (Fig.5I). At the same time, B cells from Emu-Myc mice had higher levels of RNA (Fig.S6F), were larger (Fig.S6G), and exhibited a reduced diversity of the Ig repertoire than non-carriers (Fig.S6H). Together, these data strongly suggest that c-Myc regulates cell size and fosters clonal expansion of B cells.

Myc$^{+/-}$ mice also accumulated fewer CD21^{-}CD23^{-} B cells in proportion to splenic B cells (Fig.5J-K). In addition, CD21^{-}CD23^{-} B cells were smaller in *Myc*$^{+/-}$ mice (Fig.5L), and this effect was mostly driven by the decrease in the number and size of clonal Myc+ cells, but not of ABCs (Fig.5M-O). The fact that ABCs did not change in response to c-Myc deficiency suggests that the ABC origin is independent of c-Myc levels, and that c-Myc activation is a later event in malignant transformation.

We set to compare how the transcriptome of B cells changes in response to c-Myc activation, c-Myc deficiency and clonal expansion with age. To this end, we calculated gene expression changes specific to clonal B cells in aged and Emu-Myc B cells by correlating log-transformed normalized transcript levels with the top clone size per sample. We also calculated differentially-expressed genes in *Myc*$^{+/-}$ B cells versus aged-matched controls, and performed pathway enrichment downstream of both these analyses. Interestingly, old enlarged B cells and Emu-Myc B cells showed changes in the same pathways and in the same direction in the largest clones, whereas c-Myc deficient B cells showed the opposite direction (Fig.5P). Genes that belong to Myc targets, metabolic pathways and inflammation pathways cooperatively respond to changes in c-Myc level, and thus the majority of age-related changes in the clonal B cell transcriptome can be attributed to c-Myc activation. Collectively, our data demonstrate the direct role of c-Myc in both B cell clonal expansion and cell size increase with age.

Longitudinal study reveals enlarged and CD21^{-}CD23^{-} B cells to be predictors of lifespan, and co-culture experiments point to their origin.

We analyzed the blood at seven time points starting at 20 months old (n=25) (Fig.6A) and as control at 6 months old (n=5). Enlarged B cells accumulated in the mouse blood starting at the age of 28 months (Fig.6B). Interestingly, myeloid switch started earlier, at the age of 25 months (Fig.6B), while T cells and $CD11b^{+}CD19^{+}$ cells (as in Fig. 1C) did not change over time. The trajectory of changes became even more apparent when the mice were aligned by the last blood collection before death (Fig.6C), wherein enlarged B cells and myeloid cells started to dominate in their niches 3 to 4 months prior to death. We additionally analyzed changes in $CD21^{-}CD23^{-}$ cells as they accumulate with age at the expense of follicular B cells and include clonal B cells (Fig.1H and Fig.4B,E-F). We found no significant longitudinal changes in their proportion over the analyzed time frame (Fig.6D, left plot). However, when the mice were aligned by last blood collection before death, the $CD21^{-}CD23^{-}$ cells were consistently accumulated, doubling their numbers by the time of death (Fig.6D, right plot). Exactly the same pattern, but in the opposite direction was observed for follicular B cells (Fig.6D, right plot). This suggests a prominent role of the $CD21^{-}CD23^{-}$ B cells, enlarged B cells and myeloid bias in lifespan control.

We calculated custom scores (Methods) for capturing changes in B cell size and abundance of $CD21^{-}CD23^{-}$ B cells, which we designated i) the FSC Score, and ii) the CD21-CD23- Score. We also introduced the Myeloid Score, because myeloid switch precedes an increase in B cell size (Fig.6B). The CD21-CD23- score correlated with the myeloid score at early ages, and proceeded to correlate with the FSC score at later ages (Fig.6E). This suggests that myeloid cells predate the formation of $CD21^{-}CD23^{-}$ B cells, and later $CD21^{-}CD23^{-}$ B cells predate the enlargement of B cells. We further found that mice with the higher sum of CD21-CD23- and FSC scores had a 69 day shorter maximum lifespan and a 62 day shorter median lifespan (Fig.6F). Thus, the accumulation of $CD21^{-}CD23^{-}$ enlarged B cells increases the risk of death from lymphoma or/and accelerates aging.

Because mice with larger B cell size had a higher CD21-CD23- switch score (Fig.6E), we hypothesized that the former may cause the latter through c-Myc activation. In fact, inflammation has been linked to c-Myc expression through STAT3 in immune cells, and $CD21^{-}CD23^{-}$ B cells are known to express proinflammatory cytokines (Hao et al., 2011). For this reason, we tested how the aging environment affects follicular B cell size and c-Myc expression *in vitro* (Fig.6G). Follicular B cells co-cultured with ABC cells increased in size and increased c-Myc expression (Fig.6H-I). Interestingly, the young spleen environment failed to support an increase in c-Myc expression (Fig.6H-I). Thus, we show that the aged environment of the spleen triggers an increase of cell size and of c-Myc expression in B cells.

Finally, we found that follicular B cells convert into $CD21^{-}CD23^{-}$ B cells in the presence of an old spleen environment, but not a young one (Fig.6J). Altogether, our data suggest that old follicular B cells turn into ABC cells under the influence of aging microenvironment, and ABCs later activate c-Myc expression, contributing to clonal selection and malignant transformation of B cells.

Discussion

To understand how cancer develops in the context of aging we studied commonly used mouse strains, which spontaneously develop B-cell lymphoma at old ages in 50-60% of cases (Brayton et al., 2012; Turturro et al., 2002). We discovered a novel population of B cells that clonally expands with age. These cells are larger than their healthy counterparts and represent most of the aged B cells. We found a strong correlation between B cell size and top clone size, suggesting that cell size increase goes hand in hand with clonal expansion. Using whole exome sequencing we confirmed that top B

cell clones carry somatic mutations in oncogenes and tumor suppressors, and exhibit mutational signatures characteristic of aging and cancer (Alexandrov et al., 2020). It would be interesting to determine whether additional mouse tissues also exhibit clonal expansions with age, which is likely since mice die mostly from cancer (Brayton et al., 2012; Turturro et al., 2002), and clonal selection is a necessary step for it.

Clonal expansions are characteristic of aging human tissues. Age-related clonal hematopoiesis increases the risk of cancer and cardiovascular diseases (Genovese et al., 2014; Jaiswal et al., 2014). However, the molecular mechanisms of clonal expansion are poorly understood. Here, we carried out multiple omics analyses of mouse clonal B cells to shed light on these mechanisms. At the transcript and protein levels, we revealed activation of targets of transcription factor c-Myc as well as c-Myc activation itself. c-Myc is known to be crucial for i) cancer development, ii) cell growth, and iii) cell competition (Dang, 2012), suggesting that it is a strong candidate to drive clonal expansions with age. We further analyzed DNA methylation and found that clonal B cells exhibited a remodeled DNA methylome, in particular featuring hypermethylated promoters that silenced gene expression. Some of the recurrently silenced genes were tumor suppressors, while others were genes regulating the response to inflammation and cell fate. We further found that c-Myc preferably up-regulated targets with lower promoter methylation, which is in line with *in vitro* studies showing that c-Myc prefers un-methylated binding sites (Tate and Bird, 1993; Yin et al., 2017). Together, these results suggested that the altered DNA methylome and activated c-Myc provide a selective advantage to B cell clones, which therefore expand with age. Further research may test whether other mouse and human cell types develop clones driven by c-Myc or epigenetic changes with age.

We tested the causal role of c-Myc in clonal expansions with age by using mouse models of c-Myc activation and deficiency. In 2015, it was shown that c-Myc deficiency extends the lifespan of mice and reduces the spread of lymphoma (Hofmann et al., 2015). Moreover, similar findings were observed in Drosophila pointing to a conserved mechanism of action (Greer et al., 2013). In our study, c-Myc deficient mice were protected from B cell size increase and preserved Ig diversity with age, while c-Myc transgenic mice accumulated larger B cells and clonal expansions. Thus, the causal role of c-Myc in age-related clonal expansions is strongly supported by our work.

To understand the origin of clonal B cells, we carried out a scRNA-seq analysis of aged mouse spleen (Kimmel et al., 2019). Two major age-related populations of B cells were found, both of which were CD21-CD23-, a characteristic of age-associated B cells (ABCs), which have been described as atypical memory B cells (Hao et al., 2011). However, only one population expressed transcription factor *Tbx21* (Tbet) and receptor *Itgax* (Cd11c) and thus represented canonical ABCs (Rubtsova et al., 2013). The reconstructed Ig sequences revealed that the second population is clonal and shares transcriptional changes with B cell clones from our bulk analysis. We designated this novel B-cell population as Myc+ B cells. ABCs shared the majority of deregulated genes and several Ig clones with Myc+ B cells. The data suggest that these clonal, premalignant B cells originate from ABCs or with their support. In humans, ABCs accumulate in HIV patients that are also susceptible to Burkitt lymphoma (Atallah-Yunes et al., 2020; Knox et al., 2017). Furthermore, patients with autoimmune diseases accumulate ABCs (Rubtsov et al., 2011; Wu et al., 2019) and are susceptible to diffuse-large B-cell lymphomas (Zintzaras et al., 2005). Further studies are needed to test if ABCs give rise to cancerous B cells and facilitate their development in humans similarly to aging mice.

We tested if the aging environment was responsible for c-Myc activation, since c-Myc has been linked to inflammation and cancer (Greten and Grivennikov, 2019;

Grivennikov et al., 2010). Indeed, we exposed follicular B cells to an aged environment (including ABC cells) *in vitro* and this was sufficient to trigger an increase of c-Myc and loss of CD23 and CD21 surface markers. Moreover, the young environment failed to increase c-Myc expression indicating an age-dependent effect. In future work, it would be interesting to establish which age-related signals trigger c-Myc expression in aged B cells, and whether targeting these signals can delay clonal expansion and age-related cancer. Another important question is whether low-grade inflammation facilitates clonal expansions through activation of c-Myc in other tissues.

Altogether, we here describe for the first time how aged B cells transform into cancer cells through several steps that comprise c-Myc activation stimulated by an aging environment, DNA methylation and transcriptome remodeling, and somatic mutations accumulation. We also establish aged mice as a model to study how aging favors naturally occurring cancer.

Materials and methods

Animals

Balb/c and C57BL6/JNia mice were obtained from the National Institute on Aging Aged Rodent Colony. Only female mice were used in the present study. 6-wk-old female Emu-Myc mice and control non-carriers were purchased from The Jackson Laboratory (cat. #002728). Animals were euthanized with CO_2. Spleens were harvested and placed in cold PBS until analysis, and liver samples were immediately frozen in liquid nitrogen and stored at -80°C. For bone marrow analyses, bones were stripped of the muscle in cold PBS, cut from both ends, and cells were aspirated with 5 ml of cold PBS. The aspirates were filtered through 40 µm Falcon Cell Strainers (Corning). Spleens from 24-27 months old Myc haploinsufficient (Myc +/-) mice and their wild type siblings were collected at Brown University and transferred on cold PBS to Harvard Medical School for further analysis. All experiments using mice were performed in accordance with institutional guidelines for the use of laboratory animals and were approved by the Brigham and Women's Hospital and Harvard Medical School Institutional Animal Care and Use Committees.

Flow cytometry and sorting

mAbs used for staining included: anti-CD19 [6D5], anti-CD43 [S11], anti-IgM [RMM-1], anti-CD11b [M1/70], anti-B220, anti-CD80 [16-10A1], anti-CD3 [17A2], anti-Fas [SA367H8], anti-CD138 [281-2], anti-CD23 [B3B4], anti-CD21 [7E9], anti-CD24 [M1/69], anti-CD29 [HMβ1-1], anti-CD93, and anti-CD45 [30-F11] (all from Biolegend) with fluorophores as in Supplementary table 4. Dead cells were excluded by DAPI staining. Data was collected on a Cytek DXP11 and analyzed by FlowJo software (BD). Cells were sorted on BD Aria Fusion. Spleens were gently pressed between microscopy slides to get single-cell suspensions. One ml of cell suspensions were 1) incubated with 14 ml of red blood lysis buffer for 10 min on ice, 2) centrifuged at 4°C, 250 g for 10 min, 3) washed once with 1 ml of FACS buffer (1% FBS in PBS), 4) stained in 100 ul of AB solution (2 ng/ul of each AB) at 4°C for 20 min protected from light, 5) washed again, 6) filtered into tubes with cell strainer snap cap (Corning), and 7) analyzed with flow cytometry or FACS-sorted into 1 ml of FACS buffer. To establish surface markers of ABCs and Myc+ populations we selected genes that: 1) differentially up-regulated in both populations in scRNA-seq, 2) produce surface proteins based on GO terms, 3) differentially up-regulated in clonal B cells in our bulk RNA-seq data. It yielded a few candidates, of which we chose CD29 (*Itgb1*) as a positive marker for both age-related populations, and CD24 to distinguish Myc+ cells from ABCs.

RNA sequencing and qRT-PCR

Total RNA, DNA and proteins were extracted from fresh or snap-frozen FACS-sorted 2-5 million splenic B cells using AllPrep DNA/RNA/Protein Mini Kit (Qiagen) following the manufacturer's instructions. RNA was eluted with 42 ul of RNAse-free water. RNA concentration was measured with Qubit using RNA HS Assay kit. Libraries were prepared with TruSeq Stranded mRNA LT Sample Prep Kit according to TruSeq Stranded mRNA Sample Preparation Guide, Part # 15031047 Rev. E. After the libraries had been quantified using the Bioanalyzer (Agilent), samples were sequenced with Illumina NovaSeq6000 S4 (2x150bp) (reads trimmed to 2x100bp) to get 20M read depth coverage per sample. The BCL (base calls) binary were converted into FASTQ using the Illumina package bcl2fastq. Fastq files were mapped to mm10 (GRCm38.p6) mouse genome. and gene counts were obtained with STAR v2.7.2b (Dobin et al., 2013). For quantitative RT-PCR, RNA samples were normalized by DNA concentration in the same sample, then 2 ul of diluted RNA were mixed with iTaq Universal SYBR (Bio-Rad) and primers for c-Myc (F: TTCCTTTGGGCGTTGGAAAC, R: GCTGTACGGAGTCGTAGTCG) or Gapdh (F: AGGTCGGTGTGAACGGATTTG, R: TGTAGACCATGTAGTTGAGGTCA) for the final volume of 10 ul, and loaded into Multiplate 96-Well PCR Plates (Bio-rad). Data was acquired for 40 cycles on Bio-Rad C1000, CFX96 Thermal Cycler. Each sample was loaded in duplicate.

Whole exome sequencing

Total DNA was extracted with AllPrep DNA/RNA/Protein Mini Kit (Qiagen) following the manufacturer's instructions. DNA was eluted with 100 ul of EB. DNA concentration was measured with Qubit using DNA BR Assay kit. Libraries were prepared with Agilent SureSelect XT Mouse All Exon according to SureSelectXT Target Enrichment System for Illumina Version B.2, April 2015. After the libraries had been quantified using the Bioanalyzer (Agilent), samples were sequenced with Illumina NovaSeq6000 S4 (2x150bp) (reads trimmed to 2x100bp) to get 100X throughput depth (roughly 50X on-target) coverage per sample. The BCL (base calls) binary were converted into FASTQ using the Illumina package bcl2fastq. Genomic reads were mapped to the GRCm38.p2 mouse genome assembly using BWA-MEM v0.7.15-r1140 (Li, 2013) and sorted using Samtools v1.6 (Li et al., 2009). Somatic mutations were called with Mutect2 from the GATK package v4.1.8.0 using liver as matched control, high quality variants were selected with GATK FilterMutectCalls (Van der Auwera et al., 2013). Variants were annotated with the Ensembl Variant Effect Predictor v100.2 (with option --everything) (McLaren et al., 2016). Only the variants that 1) were supported by 5 reads or more (DP.ALT>4), and 2) were in positions covered by 30 reads or more (DP.ALT+DP.REF>29), were taken for VAF analysis. Signatures of somatic mutations were extracted and deconvoluted using Sigproextractor v.1.0.20 with 100 nmf replicates.

Proteomics

Tandem mass tag (TMTpro) isobaric reagents were from ThermoFisher Scientific (Waltham, MA). Trypsin was purchased from Pierce Biotechnology (Rockford, IL) and LysC from Wako Chemicals (Richmond, VA). Samples were prepared as described previously (Li et al., 2020; Navarrete-Perea et al., 2018). Briefly, cell pellets were syringe-lysed in 8M urea complemented with protease and phosphatase inhibitors. Samples were reduced using 5mM TCEP for 30 min and alkylated with 10 mM iodoacetamide for 30 min. The excess of iodoacetamide was quenched with 10 mM DTT for 15 min. Protein was quantified using the BCA protein assay. Approximately 50 µg of protein were chloroform-methanol precipitated and reconstituted in 100 µL of 200 mM EPPS (pH 8.5). Protein was digested using Lys-C overnight at room temperature followed by trypsin for 6h at 37°C, both at a 100:1 protein:protease ratio. After digestion, the samples were labeled using the TMTpro16 reagents for 90 mins, the reactions were quenched using hydroxylamine (final concentration of 0.3% v/v). The samples were combined equally and subsequently desalted.

We enriched phosphopeptides from the pooled TMT-labeled mixtures using the Pierce High-Select Fe-NTA Phosphopeptide Enrichment kit ("mini-phos") (Navarrete-Perea et al., 2018; Paulo et al., 2018) following manufacturer's instructions. The unbound fraction was retained and fractionated using basic pH reversed-phase (BPRP) HPLC. Ninety-six fractions were collected and then consolidated into 12 which were analyzed by LC-MS3 (Paulo et al., 2016a).

All data were collected on an Orbitrap Fusion Lumos mass spectrometer coupled to a Proxeon NanoLC-1000 UHPLC. The peptides were separated using a 100 µm capillary column packed with ≈35 cm of Accucore 150 resin (2.6 µm, 150 Å; ThermoFisher Scientific). The mobile phase was 5% acetonitrile, 0.125% formic acid (A) and 95% acetonitrile, 0.125% formic acid (B). For BPRP fractions, the data were collected using a DDA-SPS-MS3 method with online real-time database searching (RTS) (Schweppe et al., 2020a) to reduce ion interference (Gygi et al., 2019; Paulo et al., 2016b). Each fraction was eluted using a 90 min method over a gradient from 6% to 30% B. Peptides were ionized with a spray voltage of 2,600 kV. The instrument method included Orbitrap MS1 scans (resolution of 120,000; mass range 400−1400 m/z; automatic gain control (AGC) target $2x10^5$, max injection time of 50 ms and ion trap MS2 scans (CID collision energy of 35%; AGC target $1x10^4$; rapid scan mode; max injection time of 120 ms). RTS was enabled and quantitative SPS-MS3 scans (resolution of 50,000; AGC target $2.5x10^5$; max injection time of 250 ms).

Phosphopeptides were analyzed with FAIMS/hrMS2 using our optimized workflow for multiplexed phosphorylation analysis (Pfammatter et al., 2019; Saba et al., 2009; Schweppe et al., 2020b). Briefly, the Thermo FAIMSpro device was operated with default parameters (inner and outer electrodes were set at 100°C, yielding a FWHM between 10 V to 15 V and dispersion voltage (DV) was set at -5000 V). Each "mini-phos" was analyzed twice by the mass spectrometer using a 2.5h method having a gradient of 6% to 30%.

Raw files were first converted to mzXML. Database searching included all mouse entries from UniProt (downloaded March 2020). The database was concatenated with one composed of all protein sequences in the reversed order. Sequences of common contaminant proteins were also included. Searches were performed using a 50ppm precursor ion tolerance and 0.9 Da (low-resolution MS2) or 0.03 Da (high-resolution MS2) product ion tolerance. TMTpro on lysine residues and peptide N termini (+304.2071 Da) and carbamidomethylation of cysteine residues (+57.0215 Da) were set as static modifications, and oxidation of methionine residues (+15.9949 Da) was set as a variable modification. For phosphopeptide analysis, +79.9663 Da was set as a variable modification on serine, threonine, and tyrosine residues.

PSMs (peptide spectrum matches) were adjusted to a 1% false discovery rate (FDR) (Elias and Gygi, 2007, 2010). PSM filtering was performed using linear discriminant analysis (LDA) as described previously (McAlister et al., 2012), while considering the following parameters: XCorr, ΔCn, missed cleavages, peptide length, charge state, and precursor mass accuracy. Protein-level FDR was subsequently estimated. Phosphorylation site localization was determined using the AScore algorithm (Beausoleil et al., 2006). A threshold of 13 corresponded to 95% confidence that a given phosphorylation site was localized.

For reporter ion quantification, a 0.003 Da window around the theoretical m/z of each reporter ion was scanned, and the most intense m/z was used. Peptides were filtered to include only those with a summed signal-to-noise ratio ≥100 across all channels. For each protein, the filtered signal-to-noise values were summed to generate protein

quantification values. To control for different total protein loading within an experiment, the summed protein quantities of each channel were adjusted to be equal in the experiment. For each protein in a TMTpro experiment, the signal-to-noise was scaled to sum to 100 to facilitate comparisons across experiments.

Spectral counts values were analyzed with R in Rstudio. Proteome and phosphoproteome data was normalized using the RLE method and log transformed using *edgeR* package. Values for phospho sites were normalized to corresponding protein level and differentially changed sites were calculated with *limma* package. Ranked phospho sites were then assessed for enrichment for targets of mouse kinases using PTM-SEA resources (Krug et al., 2019) and kinact software (Wirbel et al., 2018).

Single-cell RNA sequencing analysis

To identify genes differentially expressed in this newly identified B cell cluster and other B cells, we downloaded the single cell RNA seq data from Calico's murine aging cell atlas (https://mca.research.calicolabs.com/data, Spleen single cell count data, filtered) (Kimmel et al., 2019). Preprocessing of the downloaded data was performed using scanpy (Wolf et al., 2018). We first removed cells with a high (>0.05) percentage of mitochondrial reads. Cells not annotated as B cells were also removed. We then normalized the read counts by total reads number per cell and multiplied by a rescaling factor of 10000. Normalized reads were log transformed after adding a pseudo-count of 1. We scaled the log-transformed data to unit variance and zero mean and clipped maximum value to 10. After the above data preprocessing, we selected the cells corresponding to our newly identified B cell cluster, which we have shown to be C130026I21Rik^{+}Apoe^{+}Cr2^{-}Fcer2a^{-}. For each cell, we used a linear combination of the RNA level of these four marker genes to calculate a score (score = *C130026I21Rik* + *Apoe* - *Cr2* - *Fcer2a*), and within this cluster we selected the ABC cluster that is Tbx21+ and another one that is Myc+ using the same linear system. We then used a score threshold of 2.5 to select cells in the cluster of interest. Differential expression analysis was performed between the cell cluster of interest and all other remaining cells using the rank_genes_groups function in scanpy (Wilcoxon test). Details of the analysis can be found in our jupyter notebook B_cell_scRNAseq.ipynb. To reconstruct CDR3 regions of single cells, we demultiplexed bulk fastq files into single cell fastq files with *scruff* package in R (Wang et al., 2019), mapped it to GRCm38.p6 genome and obtained gene counts with STAR v2.7.2b (Dobin et al., 2013), and reconstructed CDR3 regions of Ig kappa chain using *mixcr* (Bolotin et al., 2017). Clone sizes were calculated as percent of templates supporting the current Ig kappa chain to the total number of reconstructed templates of Ig kappa chain. Cells with fewer than 200 gene counts were removed. Clonality signature score was calculated for each cell as transcript level of top 50 genes minus transcript level of bottom 50 genes taken from differentially expressed genes in top clones in our bulk data (Supplementary Table 2). Cells were marked as B cells if the sum of gene counts for genes *Cd19*, *Cd74*, *Cd79a*, *Cd79b*, *H2-Aa*, *H2-Eb1*, *H2-Ab1* was above 35. Cells were marked as ABC if 1) it was marked as B cell, 2) 30 top DEGs in the ABC cluster (defined above) had a sum of gene counts over 70. A cell was marked as Myc+ if 1) it was marked as a B cell, 2) 30 top DEGs in the Myc+ cluster (defined above) had a sum of gene counts over 25. tSNEs were calculated using the *M3C* package in R (John et al., 2019).

Reconstruction of Ig CDR3 regions

Genomic CDR3 regions of Ig heavy chains were analyzed with Immunoseq (Adaptive Biotechnologies). CDR3 regions were reconstructed from RNA sequencing raw data using mixcr software (Bolotin et al., 2017) with recommended settings for transcriptome data. Filtering of reconstructed regions and diversity analysis was done with VDJtools software (Shugay et al., 2015).

Immunofluorescence microscopy
B cells were FACS-sorted at 500 thousand cells per well and incubated with poly l-lysine treated coverslips for 1 hour in 24 well plates. Cells were permeabilized with 0.1% Triton X-100 2 times for 30 seconds, fixed in 3.7% PFA in PBS for 10 minutes and washed three times with PBS, incubated with the blocking buffer until further analysis (1% BSA, 0.1% Triton X-100 in PBS). Samples were incubated with primary antibodies overnight at 4 ℃ (1:100, Abcam #ab32072), then washed with PBS five times and incubated overnight with Alexa Fluor 568 - conjugated anti-rabbit secondary antibodies (1:500, Biotium cat#20098) and DAPI dye. Cells were washed with PBS and mounted with ProLong Diamond antifade (Thermo Fisher Scientific). Samples were imaged using Leica SP8 confocal microscope. Images were analyzed with ImageJ (Schindelin et al., 2015). Raw Z-stacks were converted to the maximum intensity projection images. Nuclei and cell borders were detected using manual thresholding and “Analyze Particle” function. All crowded groups and not-round shaped cells were manually removed from the analysis.

Reduced representation bisulfite sequencing (RRBS)
Libraries were prepared and sequenced as in (Meer et al., 2018). Bisulfite sequence reads were trimmed by TrimGalore v0.4.1 and mapped to the mouse genome sequence (mm10/GRCm38.p6) with Bismark v0.15.0 (Krueger and Andrews, 2011). We kept CpG sites that were covered by five reads or more. Promoter regions were determined as the [-1500, +500] bp from the transcription start site (following the direction of the transcription) taken from Ensembl annotation file (Mus_musculus.GRCm38.100.chr.gtf). The start and end positions of gene bodies were taken from the Ensembl gene predictions (Mus_musculus.GRCm38.cds.all.fa). The mean methylation levels were calculated for regions that have at least 5 covered CpG sites with average methylation level above 1%. To determine ribosomal DNA methylation (rDNAm) age, we developed a blood rDNAm clock in a similar way as described in (Wang and Lemos, 2019). Briefly, we applied ElasticNet regression on ribosomal DNA (BK000964.3) CpG methylation levels of 153 control fed C57BL/6 blood samples with an age range from 0.67 to 35 months (GSE80672).

Longitudinal blood collection and analysis
Mice were anesthetized with isoflurane and then locally with topical anesthetic, restrained, and approximately 100 ul of blood was collected from mouse tails into EDTA-coated tubes (BD). Blood was incubated on ice until further analysis (2-3 hours), then mixed with 1 ml of blood cell lysis buffer and centrifuged at 250g for 10 minutes at 4°C. Pellets were washed once with a FACS buffer (PBS with 1% FBS), split equally into 2 tubes and incubated for 20 minutes at 4°C with antibodies against B cell, T cell and myeloid cell markers, or follicular, marginal zone, and plasma cell markers. Stained cells were washed again, resuspended in 200 ul of FACS buffer and analyzed with FACS, with 20,000 events being recorded. Dead cells were gated by DAPI staining. Cell size was measured with forward scatter.

Longitudinal blood scores
We defined the FSC score as the difference between the last measurement of mean B cell size prior to death (if mouse died) or B cell size at the given round (if mouse was alive) and the mean B cell size of young mice at the same round. CD21.score was calculated as delta(CD21-CD23-) - delta(Follicular). Delta(CD21-CD23-) was calculated as percentage of $CD21^{-}CD23^{-}$ of total B cells before death (if mouse died) or $CD21^{-}CD23^{-}$ percentage at the given round (if mouse was alive) minus percentage of $CD21^{-}CD23^{-}$ of B cells in young mice at the same round. delta(Follicular) was calculated the same way for the percentage of follicular B cells of B cells. The myeloid score was calculated as delta(Myeloid) - delta(B cells). delta(Myeloid) was calculated as percentage of myeloid cells of $CD45^{+}$ cells before death (if mouse died) or percentage

at the given round (if mouse was alive) minus the percentage of myeloid cells in young mice at the same round. delta(B cells) was calculated the same way for the percentage of B cells of $CD45^+$ cells. A higher CD21.score indicates a higher proportion of $CD21^-$ $CD23^-$ cells to total B cells and/or lower proportion of follicular B cells to total B cells. A higher myeloid score indicates a higher proportion of myeloid cells to $CD45^+$ cells and/or lower proportion of B cells to CD45+. A higher FSC score indicates a larger increase of B cell size.

Cell culture

Freshly FACS-sorted cells were plated into 96 wells at 400 thousand cells per well for each cell type and cultured for 24-48 hours in 200 ul of RPMI medium with 10% FCS (Gibco), 2 mM glutamine (ThermoFisher), 1% oxaloacetic acid (15 mg/ml), 5 mg/ml sodium pyruvate (ThermoFisher), 1% non-essential amino acids (ThermoFisher), and 50 µM 2-ME (Sigma). Where only follicular B cells were incubated, 800 thousand of cells were plated. After co-incubation, cells were centrifuged at 250 g for 10 minutes at 4°C with slow deceleration. The media were carefully removed leaving ~50 ul, cells were washed once in 100 ul of FACS buffer, then resuspended in 100 ul of AB solution (1:100) and incubated at 4°C for 20 minutes protected from light, washed again and resuspended in 200 ul of FACS buffer, and filtered through 40 µm Falcon Cell Strainers (Corning). Follicular B cells (20-100 thousand) were then gated as $CD19^+CD21^{int}CD23^+$ and FACS-sorted into 300 ul of Trizol. RNA was extracted with Direct-zol RNA Microprep (Zymo Research).

Necropsy analysis

Mice were euthanized with CO_2. The chest and abdomen were opened, and the body was immersed into formalin solution and stored at 4°C until further analysis. For necropsy analysis, all organs, including small endocrine organs, were dissected, trimmed at 5 mm thickness and embedded in paraffin blocks. Paraffin blocks were sectioned at 5 um and stained with hematoxylin and eosin. The slides were examined blindly by a pathologist. Lymphoma was diagnosed when any tissue had areas containing uniform sheets of atypical lymphocytes with large nuclei. Lymphocytic hyperplasia was diagnosed when the liver or lung had small infiltrates of lymphocytes.

Data analysis and availability

All data were analyzed and plotted with R in Rstudio using a standard set of packages until otherwise specified. RNA sequencing, DNA methylation and proteomics data were analyzed for differential changes and GSEA with *limma* (Ritchie et al., 2015), *edgeR* (McCarthy et al., 2012) and *clusterprofiler* (Yu et al., 2012) packages. PCA analysis was done with the *factoextra* package. Color schemes are from the *ggsci* package. Raw reads for RNA-sequencing, whole exome sequencing, and RRBS are available at SRA (PRJNA694093).

References

Adams, J.M., Harris, A.W., Pinkert, C.A., Corcoran, L.M., Alexander, W.S., Cory, S., Palmiter, R.D., and Brinster, R.L. (1985). The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. Nature *318*, 533–538.
Alexandrov, L.B., Kim, J., Haradhvala, N.J., Huang, M.N., Tian Ng, A.W., Wu, Y., Boot, A., Covington, K.R., Gordenin, D.A., Bergstrom, E.N., et al. (2020). The repertoire of mutational signatures in human cancer. Nature *578*, 94–101.
An, J., González-Avalos, E., Chawla, A., Jeong, M., López-Moyado, I.F., Li, W., Goodell, M.A., Chavez, L., Ko, M., and Rao, A. (2015). Acute loss of TET function results in aggressive myeloid cancer in mice. Nat. Commun. *6*, 10071.
Atallah-Yunes, S.A., Murphy, D.J., and Noy, A. (2020). HIV-associated Burkitt lymphoma. Lancet Haematol. *7*, e594–e600.
Beausoleil, S.A., Villén, J., Gerber, S.A., Rush, J., and Gygi, S.P. (2006). A probability-based approach for high-throughput protein phosphorylation analysis and site localization. Nat. Biotechnol. *24*, 1285–1292.
Bolotin, D.A., Poslavsky, S., Davydov, A.N., Frenkel, F.E., Fanchi, L., Zolotareva, O.I., Hemmers, S., Putintseva, E.V., Obraztsova, A.S., Shugay, M., et al. (2017). Antigen receptor repertoire profiling from RNA-seq data. Nat. Biotechnol. *35*, 908–911.
Brayton, C.F., Treuting, P.M., and Ward, J.M. (2012). Pathobiology of aging mice and GEM: background strains and experimental design. Vet. Pathol. *49*, 85–105.
Chen, S., Miyazaki, M., Chandra, V., Fisch, K.M., Chang, A.N., and Murre, C. (2016). Id3 Orchestrates Germinal Center B Cell Development. Mol. Cell. Biol. *36*, 2543–2552.
Clavería, C., Giovinazzo, G., Sierra, R., and Torres, M. (2013). Myc-driven endogenous cell competition in the early mammalian embryo. Nature *500*, 39–44.
Coussens, L.M., and Werb, Z. (2002). Inflammation and cancer. Nature *420*, 860–867.
de la Cova, C., Abril, M., Bellosta, P., Gallant, P., and Johnston, L.A. (2004). Drosophila myc regulates organ size by inducing cell competition. Cell *117*, 107–116.
Dang, C.V. (2012). MYC on the Path to Cancer. Cell *149*, 22–35.
Di Gregorio, A., Bowling, S., and Rodriguez, T.A. (2016). Cell Competition and Its Role in the Regulation of Cell Fitness from Development to Cancer. Dev. Cell *38*, 621–634.
Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinforma. Oxf. Engl. *29*, 15–21.
Elias, J.E., and Gygi, S.P. (2007). Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat. Methods *4*, 207–214.
Elias, J.E., and Gygi, S.P. (2010). Target-decoy search strategy for mass spectrometry-based proteomics. Methods Mol. Biol. Clifton NJ *604*, 55–71.
Elias, H.K., Bryder, D., and Park, C.Y. (2017). Molecular mechanisms underlying lineage bias in aging hematopoiesis. Semin. Hematol. *54*, 4–11.
Esteller, M., Corn, P.G., Baylin, S.B., and Herman, J.G. (2001). A gene hypermethylation profile of human cancer. Cancer Res. *61*, 3225–3229.
Fane, M., and Weeraratna, A.T. (2020). How the ageing microenvironment influences tumour progression. Nat. Rev. Cancer *20*, 89–106.
Gabay, M., Li, Y., and Felsher, D.W. (2014). MYC activation is a hallmark of cancer initiation and maintenance. Cold Spring Harb. Perspect. Med. *4*.
Genovese, G., Kähler, A.K., Handsaker, R.E., Lindberg, J., Rose, S.A., Bakhoum, S.F., Chambert, K., Mick, E., Neale, B.M., Fromer, M., et al. (2014). Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. N. Engl. J. Med. *371*, 2477–2487.
Gladyshev, V.N. (2016). Aging: progressive decline in fitness due to the rising deleteriome adjusted by genetic, environmental, and stochastic processes. Aging Cell *15*, 594–602.

Greenough, A., and Dave, S.S. (2014). New clues to the molecular pathogenesis of Burkitt lymphoma revealed through next-generation sequencing. Curr. Opin. Hematol. *21*, 326–332.
Greer, C., Lee, M., Westerhof, M., Milholland, B., Spokony, R., Vijg, J., and Secombe, J. (2013). Myc-dependent genome instability and lifespan in Drosophila. PloS One *8*, e74641.
Greten, F.R., and Grivennikov, S.I. (2019). Inflammation and Cancer: Triggers, Mechanisms, and Consequences. Immunity *51*, 27–41.
Grivennikov, S.I., Greten, F.R., and Karin, M. (2010). Immunity, inflammation, and cancer. Cell *140*, 883–899.
GTEx Consortium (2013). The Genotype-Tissue Expression (GTEx) project. Nat. Genet. *45*, 580–585.
Gygi, J.P., Yu, Q., Navarrete-Perea, J., Rad, R., Gygi, S.P., and Paulo, J.A. (2019). Web-Based Search Tool for Visualizing Instrument Performance Using the Triple Knockout (TKO) Proteome Standard. J. Proteome Res. *18*, 687–693.
Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sadda, S., Klotzle, B., Bibikova, M., Fan, J.-B., Gao, Y., et al. (2013). Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol. Cell *49*, 359–367.
Hao, Y., O'Neill, P., Naradikian, M.S., Scholz, J.L., and Cancro, M.P. (2011). A B-cell subset uniquely responsive to innate stimuli accumulates in aged mice. Blood *118*, 1294–1304.
Harris, A.W., Pinkert, C.A., Crawford, M., Langdon, W.Y., Brinster, R.L., and Adams, J.M. (1988). The E mu-myc transgenic mouse. A model for high-incidence spontaneous lymphoma and leukemia of early B cells. J. Exp. Med. *167*, 353–371.
Herman, J.G., and Baylin, S.B. (2003). Gene silencing in cancer in association with promoter hypermethylation. N. Engl. J. Med. *349*, 2042–2054.
Hofmann, J.W., Zhao, X., De Cecco, M., Peterson, A.L., Pagliaroli, L., Manivannan, J., Hubbard, G.B., Ikeno, Y., Zhang, Y., Feng, B., et al. (2015). Reduced expression of MYC increases longevity and enhances healthspan. Cell *160*, 477–488.
Horvath, S. (2013). DNA methylation age of human tissues and cell types. Genome Biol. *14*, R115.
Howlader, N., Noone, A.M., Krapcho, M., Miller, D., Brest, A., Yu, M., Ruhl, J., Tatalovich, Z., Mariotto, A., and Lewis, D.R. (2020). SEER Cancer Statistics Review, 1975-2017, National Cancer Institute. Bethesda MD April.
Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P.V., Mar, B.G., Lindsley, R.C., Mermel, C.H., Burtt, N., Chavez, A., et al. (2014). Age-related clonal hematopoiesis associated with adverse outcomes. N. Engl. J. Med. *371*, 2488–2498.
John, C.R., Watson, D., Russ, D., Goldmann, K., Ehrenstein, M., Pitzalis, C., Lewis, M., and Barnes, M. (2019). M3C: Monte Carlo reference-based consensus clustering. BioRxiv 377002.
Kimmel, J.C., Penland, L., Rubinstein, N.D., Hendrickson, D.G., Kelley, D.R., and Rosenthal, A.Z. (2019). Murine single-cell RNA-seq reveals cell-identity- and tissue-specific trajectories of aging. Genome Res. *29*, 2088–2103.
Knox, J.J., Buggert, M., Kardava, L., Seaton, K.E., Eller, M.A., Canaday, D.H., Robb, M.L., Ostrowski, M.A., Deeks, S.G., Slifka, M.K., et al. (2017). T-bet+ B cells are induced by human viral infections and dominate the HIV gp140 response. JCI Insight *2*.
Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinforma. Oxf. Engl. *27*, 1571–1572.
Krug, K., Mertins, P., Zhang, B., Hornbeck, P., Raju, R., Ahmad, R., Szucs, M., Mundt, F., Forestier, D., Jane-Valbuena, J., et al. (2019). A Curated Resource for Phosphosite-specific Signature Analysis. Mol. Cell. Proteomics MCP *18*, 576–593.
Labrie, J.E., Sah, A.P., Allman, D.M., Cancro, M.P., and Gerstein, R.M. (2004). Bone marrow microenvironmental changes underlie reduced RAG-mediated recombination and B cell generation in aged mice. J. Exp. Med. *200*, 411–423.
Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with

BWA-MEM. ArXiv13033997 Q-Bio.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinforma. Oxf. Engl. *25*, 2078–2079.
Li, J., Maruyama, T., Zhang, P., Konkel, J.E., Hoffman, V., Zamarron, B., and Chen, W. (2010). Mutation of inhibitory helix-loop-helix protein Id3 causes γδ T-cell lymphoma in mice. Blood *116*, 5615–5621.
Li, J., Van Vranken, J.G., Pontano Vaites, L., Schweppe, D.K., Huttlin, E.L., Etienne, C., Nandhikonda, P., Viner, R., Robitaille, A.M., Thompson, A.H., et al. (2020). TMTpro reagents: a set of isobaric labeling mass tags enables simultaneous proteome-wide measurements across 16 samples. Nat. Methods *17*, 399–404.
Love, C., Remple, R., Li, G., and Dave, S. (2015). ID3 Loss In Vivo Synergizes with MYC-Driven Oncogenesis in Burkitt Lymphoma. Blood *126*, 3906–3906.
Martincorena, I., Fowler, J.C., Wabik, A., Lawson, A.R.J., Abascal, F., Hall, M.W.J., Cagan, A., Murai, K., Mahbubani, K., Stratton, M.R., et al. (2018). Somatic mutant clones colonize the human esophagus with age. Science *362*, 911–917.
Mayle, A., Yang, L., Rodriguez, B., Zhou, T., Chang, E., Curry, C.V., Challen, G.A., Li, W., Wheeler, D., Rebel, V.I., et al. (2015). Dnmt3a loss predisposes murine hematopoietic stem cells to malignant transformation. Blood *125*, 629–638.
McAlister, G.C., Huttlin, E.L., Haas, W., Ting, L., Jedrychowski, M.P., Rogers, J.C., Kuhn, K., Pike, I., Grothe, R.A., Blethrow, J.D., et al. (2012). Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. Anal. Chem. *84*, 7469–7478.
McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. *40*, 4288–4297.
McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P., and Cunningham, F. (2016). The Ensembl Variant Effect Predictor. Genome Biol. *17*, 122.
Meer, M.V., Podolskiy, D.I., Tyshkovskiy, A., and Gladyshev, V.N. (2018). A whole lifespan mouse multi-tissue DNA methylation clock. ELife *7*.
Muñoz-Martín, N., Sierra, R., Schimmang, T., Villa Del Campo, C., and Torres, M. (2019). Myc is dispensable for cardiomyocyte development but rescues Mycn-deficient hearts through functional replacement and cell competition. Dev. Camb. Engl. *146*.
Navarrete-Perea, J., Yu, Q., Gygi, S.P., and Paulo, J.A. (2018). Streamlined Tandem Mass Tag (SL-TMT) Protocol: An Efficient Strategy for Quantitative (Phospho)proteome Profiling Using Tandem Mass Tag-Synchronous Precursor Selection-MS3. J. Proteome Res. *17*, 2226–2236.
Niccoli, T., and Partridge, L. (2012). Ageing as a Risk Factor for Disease. Curr. Biol. *22*, R741–R752.
Ott, G., Rosenwald, A., and Campo, E. (2013). Understanding MYC-driven aggressive B-cell lymphomas: pathogenesis and classification. Blood *122*, 3884–3891.
Paulo, J.A., O'Connell, J.D., Everley, R.A., O'Brien, J., Gygi, M.A., and Gygi, S.P. (2016a). Quantitative mass spectrometry-based multiplexing compares the abundance of 5000 S. cerevisiae proteins across 10 carbon sources. J. Proteomics *148*, 85–93.
Paulo, J.A., O'Connell, J.D., and Gygi, S.P. (2016b). A Triple Knockout (TKO) Proteomics Standard for Diagnosing Ion Interference in Isobaric Labeling Experiments. J. Am. Soc. Mass Spectrom. *27*, 1620–1625.
Paulo, J.A., Navarrete-Perea, J., Erickson, A.R., Knott, J., and Gygi, S.P. (2018). An Internal Standard for Assessing Phosphopeptide Recovery from Metal Ion/Oxide Enrichment Strategies. J. Am. Soc. Mass Spectrom. *29*, 1505–1511.
Petkovich, D.A., Podolskiy, D.I., Lobanov, A.V., Lee, S.-G., Miller, R.A., and Gladyshev, V.N. (2017). Using DNA Methylation Profiling to Evaluate Biological Age and Longevity Interventions. Cell Metab. *25*, 954-960.e6.

Pfammatter, S., Bonneil, E., McManus, F.P., and Thibault, P. (2019). Accurate Quantitative Proteomic Analyses Using Metabolic Labeling and High Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS). J. Proteome Res. *18*, 2129–2138.
Richter, J., Schlesner, M., Hoffmann, S., Kreuz, M., Leich, E., Burkhardt, B., Rosolowski, M., Ammerpohl, O., Wagener, R., Bernhart, S.H., et al. (2012). Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. Nat. Genet. *44*, 1316–1320.
Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47.
Rubtsov, A.V., Rubtsova, K., Fischer, A., Meehan, R.T., Gillis, J.Z., Kappler, J.W., and Marrack, P. (2011). Toll-like receptor 7 (TLR7)-driven accumulation of a novel $CD11c^+$ B-cell population is important for the development of autoimmunity. Blood *118*, 1305–1315.
Rubtsova, K., Rubtsov, A.V., van Dyk, L.F., Kappler, J.W., and Marrack, P. (2013). T-box transcription factor T-bet, a key player in a unique type of B-cell activation essential for effective viral clearance. Proc. Natl. Acad. Sci. U. S. A. *110*, E3216-3224.
Saba, J., Bonneil, E., Pomiès, C., Eng, K., and Thibault, P. (2009). Enhanced sensitivity in proteomics experiments using FAIMS coupled with a hybrid linear ion trap/Orbitrap mass spectrometer. J. Proteome Res. *8*, 3355–3366.
Saghafinia, S., Mina, M., Riggi, N., Hanahan, D., and Ciriello, G. (2018). Pan-Cancer Landscape of Aberrant DNA Methylation across Human Tumors. Cell Rep. *25*, 1066-1080.e8.
Schindelin, J., Rueden, C.T., Hiner, M.C., and Eliceiri, K.W. (2015). The ImageJ ecosystem: An open platform for biomedical image analysis. Mol. Reprod. Dev. *82*, 518–529.
Schweppe, D.K., Eng, J.K., Yu, Q., Bailey, D., Rad, R., Navarrete-Perea, J., Huttlin, E.L., Erickson, B.K., Paulo, J.A., and Gygi, S.P. (2020a). Full-Featured, Real-Time Database Searching Platform Enables Fast and Accurate Multiplexed Quantitative Proteomics. J. Proteome Res. *19*, 2026–2034.
Schweppe, D.K., Rusin, S.F., Gygi, S.P., and Paulo, J.A. (2020b). Optimized Workflow for Multiplexed Phosphorylation Analysis of TMT-Labeled Peptides Using High-Field Asymmetric Waveform Ion Mobility Spectrometry. J. Proteome Res. *19*, 554–560.
Shugay, M., Bagaev, D.V., Turchaninova, M.A., Bolotin, D.A., Britanova, O.V., Putintseva, E.V., Pogorelyy, M.V., Nazarov, V.I., Zvyagin, I.V., Kirgizova, V.I., et al. (2015). VDJtools: Unifying Post-analysis of T Cell Receptor Repertoires. PLoS Comput. Biol. *11*, e1004503.
Simpson, P., and Morata, G. (1981). Differential mitotic rates and patterns of growth in compartments in the Drosophila wing. Dev. Biol. *85*, 299–308.
Tate, P.H., and Bird, A.P. (1993). Effects of DNA methylation on DNA-binding proteins and gene expression. Curr. Opin. Genet. Dev. *3*, 226–231.
Turturro, A., Duffy, P., Hass, B., Kodell, R., and Hart, R. (2002). Survival characteristics and age-adjusted disease incidences in C57BL/6 mice fed a commonly used cereal-based diet modulated by dietary restriction. J. Gerontol. A. Biol. Sci. Med. Sci. *57*, B379-389.
Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr. Protoc. Bioinforma. *43*, 11.10.1-11.10.33.
Wang, M., and Lemos, B. (2019). Ribosomal DNA harbors an evolutionarily conserved clock of biological aging. Genome Res. *29*, 325–333.
Wang, Z., Hu, J., Johnson, W.E., and Campbell, J.D. (2019). scruff: an R/Bioconductor package for preprocessing single-cell RNA-sequencing data. BMC Bioinformatics *20*, 222.
Wirbel, J., Cutillas, P., and Saez-Rodriguez, J. (2018). Phosphoproteomics-Based

Profiling of Kinase Activities in Cancer Cells. Methods Mol. Biol. Clifton NJ *1711*, 103–132.
Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. *19*, 15.
Wu, C., Fu, Q., Guo, Q., Chen, S., Goswami, S., Sun, S., Li, T., Cao, X., Chu, F., Chen, Z., et al. (2019). Lupus-associated atypical memory B cells are mTORC1-hyperactivated and functionally dysregulated. Ann. Rheum. Dis. *78*, 1090–1100.
Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S., Das, P.K., Kivioja, T., Dave, K., Zhong, F., et al. (2017). Impact of cytosine methylation on DNA binding specificities of human transcription factors. Science *356*.
Yizhak, K., Aguet, F., Kim, J., Hess, J.M., Kübler, K., Grimsby, J., Frazer, R., Zhang, H., Haradhvala, N.J., Rosebrock, D., et al. (2019). RNA sequence analysis reveals macroscopic somatic clonal expansion across normal tissues. Science *364*.
Yokoyama, A., Kakiuchi, N., Yoshizato, T., Nannya, Y., Suzuki, H., Takeuchi, Y., Shiozawa, Y., Sato, Y., Aoki, K., Kim, S.K., et al. (2019). Age-related remodelling of oesophageal epithelia by mutated cancer drivers. Nature *565*, 312–317.
Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. Omics J. Integr. Biol. *16*, 284–287.
Zenin, A., Tsepilov, Y., Sharapov, S., Getmantsev, E., Menshikov, L.I., Fedichev, P.O., and Aulchenko, Y. (2019). Identification of 12 genetic loci associated with human healthspan. Commun. Biol. *2*, 41.
Zhang, W., and Liu, H.T. (2002). MAPK signal pathways in the regulation of cell proliferation in mammalian cells. Cell Res. *12*, 9–18.
Zhao, Z., Chen, L., Dawlaty, M.M., Pan, F., Weeks, O., Zhou, Y., Cao, Z., Shi, H., Wang, J., Lin, L., et al. (2015). Combined Loss of Tet1 and Tet2 Promotes B Cell, but Not Myeloid Malignancies, in Mice. Cell Rep. *13*, 1692–1704.
Zintzaras, E., Voulgarelis, M., and Moutsopoulos, H.M. (2005). The risk of lymphoma development in autoimmune diseases: a meta-analysis. Arch. Intern. Med. *165*, 2337–2344.

Figures and Legends

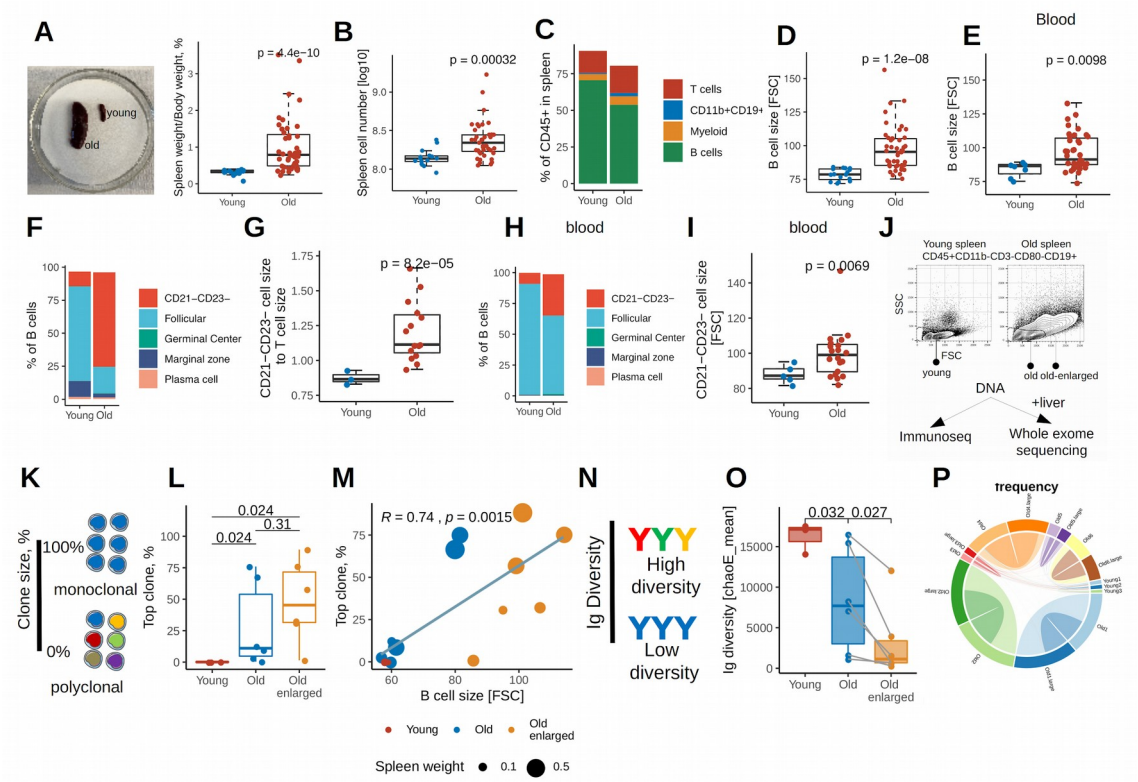


Figure 1. Aged splenic and blood B cells increase cell size and acquire premalignant phenotypes. (A) Representative spleens from young and old C57BL/6 mice (left) and spleen weights of young and old C57BL/6 mice (right). (B) Total nucleated cell number in spleens of young and old mice. (C) Percentages of lymphoid ($CD45^+$) cells in spleens from young and old mice measured by FACS. Myeloid cells were gated as $CD45^+CD3^-CD11b^+CD19^-$, B cells as $CD45^+CD11b^-CD3^-CD19^+$, T cells as $CD45^+CD3^+CD11b^-CD19^-$, enlarged B cells as $CD45^+CD11b^-CD3^-CD19^+FSC^{hi}$, and $CD11b^+CD19^+$ cells as $CD45^+CD11b^+CD3^-CD19^+$. (D) Mean B cell size measured by FACS in spleens of young and old mice. (E) Mean B cell size measured by FACS in the blood of young and old mice. (F) Percentages of B-cell populations to total B cells in spleens of young and old mice measured by FACS. (G) Mean cell size of $CD21^-CD23^-$ B cells measured by FACS in spleens of young and old mice. (H) Percentage of B cells populations to total B cells in the blood of young and old mice measured by FACS. (I) Mean $CD21^-CD23^-$ B cells size in the blood of young and old mice measured by FACS. (J) Experiment schematic to estimate clone sizes of old B cells. Gating strategy for enlarged (FSC^{hi}) and regular B cells, and representative FACS of splenic B cells from young and old mice. (K) Scheme explaining clone size. (L) Largest Ig clone sizes of young, old and enlarged old B cells FACS-sorted from spleens and measured by sequencing of CDR3 regions of Ig heavy chains. (M) Correlation between B cell size measured by FACS and size of top B cell clone measured by sequencing of CDR3 regions of Ig heavy chains in FACS-sorted B cells. (N) Scheme explaining Ig diversity. (O) Ig diversity in chaoE_mean in young, old, and enlarged old B cells sorted from mouse spleens and estimated from Immunoseq data with VDJtools software. P-value for young-old comparison was calculated with two-sided Student t-test. P-value for old and old-enlarged comparison was calculated with two-sided paired Student t-test. (P) Circular plot showing frequency of top clones in a given sample. CDR3 clones with the same amino acid sequences are connected. All mice in the experiments were female C57BL/6 mice; old mice were 26-27 months old, and young mice were 6 months old. All comparisons were calculated with two-tailed Student t-test, unless otherwise specified.

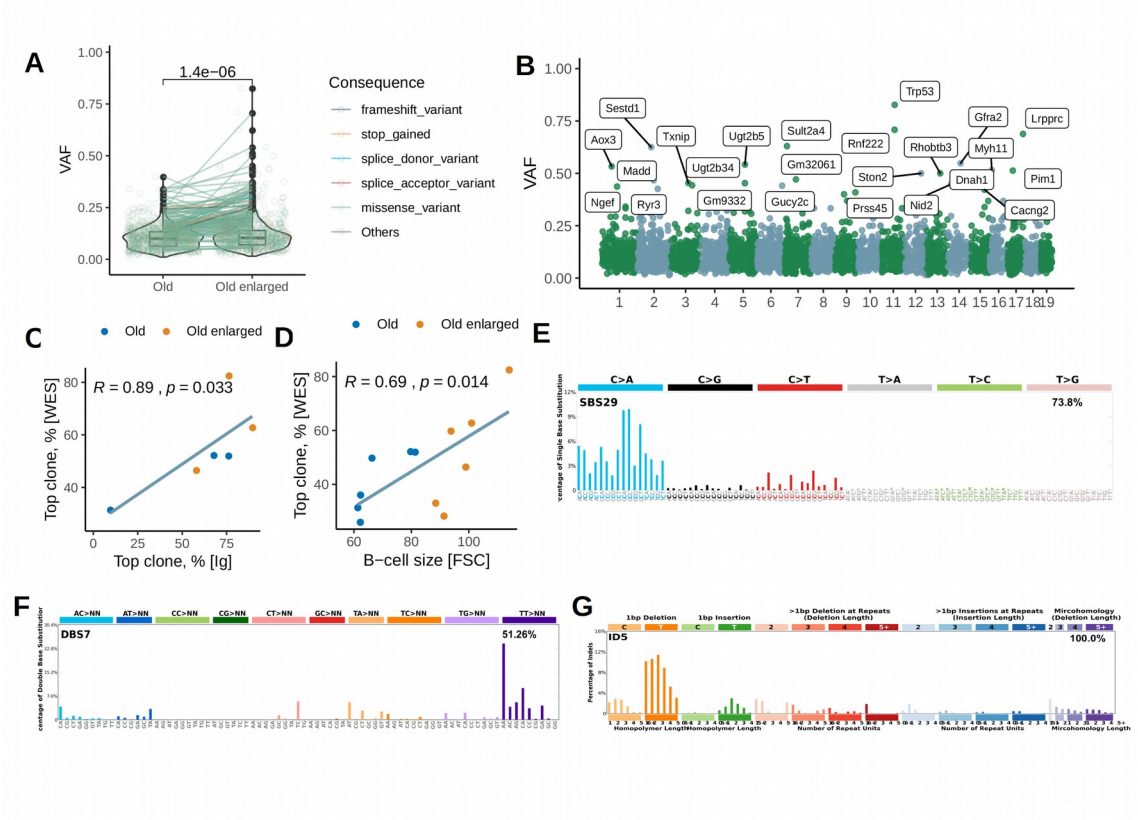


Figure 2. Mouse B cells undergo clonal expansions with age driven by somatic mutations. (A) Variant allele fractions for somatic mutations detected from whole exome sequences of old and old enlarged B cells, seven mice per group. Lines are colored according to the predicted consequence of the variant, as listed in the legend. Dots are connected if variants are identical in old and old enlarged B cells from the same animal. Only variants predicted to have high or moderate impact are shown. (B) Manhattan plot for somatic mutations in all samples with variant allele fraction (VAF) on Y-axis and chromosome position on X-axis. Labeled genes are harboring variants with VAF over 0.4. (C) Correlation between top clone size (Ig) estimated from Immunoseq data and top clone size estimated by VAF from whole exome sequences. (D) Correlation between B cell size measured by FACS and top clone size estimated by VAF from whole exome sequences. (E), (F) and (G) are results of mutational signature deconvolution of total somatic mutations in regular and enlarged B cells. Correlations between two variables were evaluated using Pearson's correlation coefficient, and all means comparisons were calculated with two-tailed Student t-test, unless otherwise specified. All mice in the experiments were 27 months old female C57BL/6 mice.

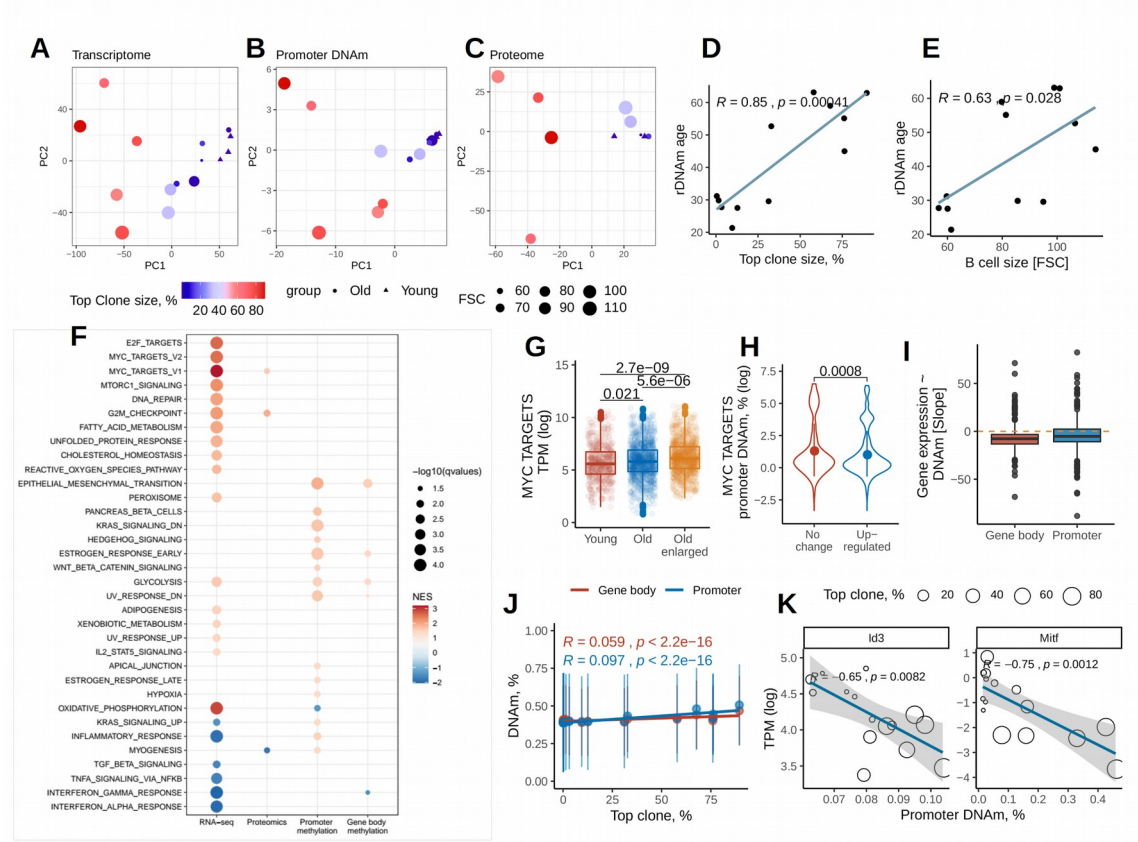


Figure 3. Aged clonal B cells overexpress oncogene c-Myc and exhibit aberrant DNA methylation. PCA analyses of (A) transcriptome, (B) promoter DNA methylation and (C) proteome from young, old and old enlarged B cells, colored according to top Ig clone size in a sample, shaped according to the group, and sized according to the mean B cell size of a sample. (D) Correlation between top Ig clone size estimated by Immunoseq and the rDNA methylation age of B cells. (E) Correlation between mean B cell size and the rDNA methylation age. (F) GSEA plot for gene expression, proteome, promoter and gene body DNA methylation changes correlated with top clone Ig size of a sample (referred here as clonality). (G) Gene expression levels of c-Myc targets in young, old and old enlarged B cells. Myc targets were taken from the hallmark set of genes 'MYC_TARGETS_V1' from msigdb database. (H) Mean DNA promoter methylation levels for genes encoding Myc targets which expression increase or did not change in clonal B cells. (I) Slope values from correlations between gene expressions and mean promoter DNA methylation or gene body DNA methylation. The dashed line indicates slope value when there is no correlation. (J) Correlation of top Ig clone size and mean global DNA methylation in promoters and gene bodies. (K) Correlation between mean promoter DNA methylation and gene expression of two tumor suppressor genes, *Id3* and *Mitf*. Dots are sized according to the size of the top Ig clone in a sample. Correlations between two variables were evaluated using Pearson's correlation coefficient, and all p-values for group means comparisons were calculated with two-tailed Student t-test.

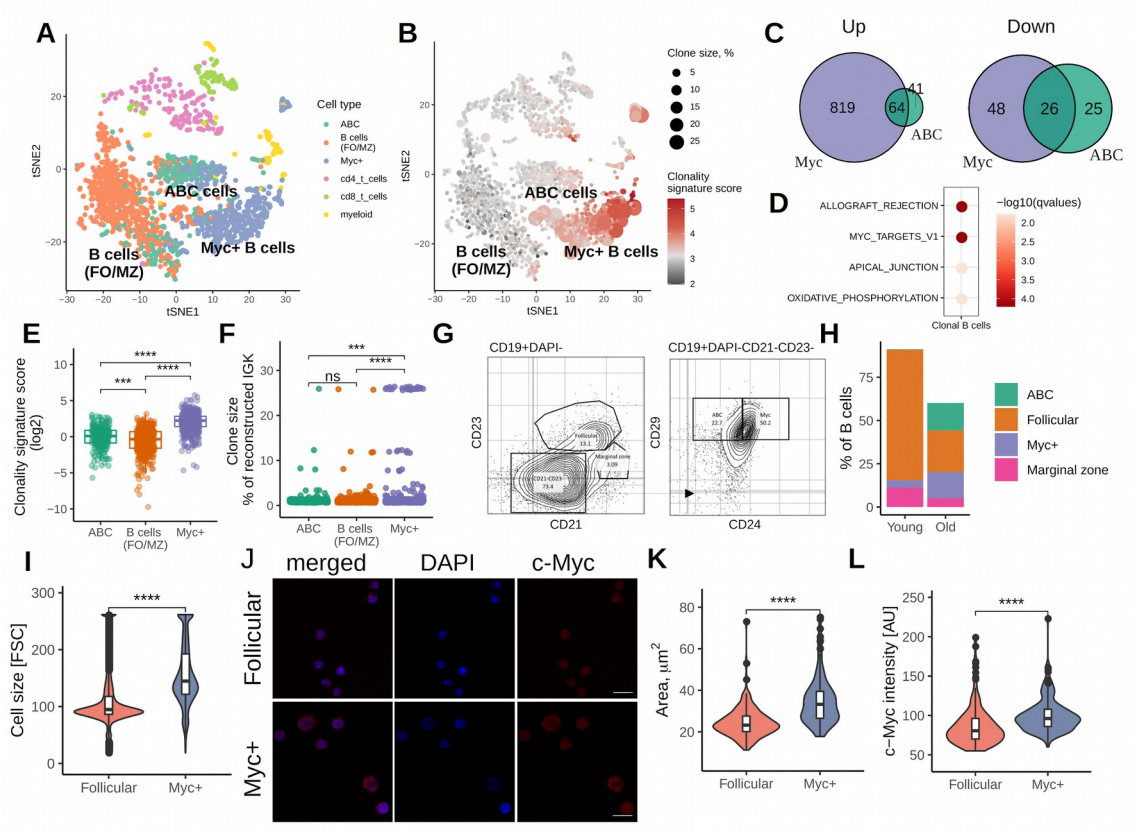


Figure 4. Single-cell RNA-seq reveals a novel B cell population that is clonal and may originate from ABC cells. (A) tSNE plot for single-cell RNA-seq data of aged mouse spleen and colored according to the cell type. Each dot is a cell. (B) tSNE for single-cell RNA-seq data from aged mouse spleen colored according to clonality signature score. Dot size is a size of the Ig clone. Each dot represents a cell. (C) Overlap between up-regulated and down-regulated genes (FDR<0.05) in ABC and Myc+ B cells compared to B cells (FO/MZ). (D) GSEA for genes that differentially expressed in B cells from the top Ig clone compared to the gene levels in smallest clones. (E) Clonality signature scores for ABCs, Myc+ cells and the rest of B cells (FO/MZ). Each dot represents a cell. (F) Clone sizes to which B cells (FO/MZ), ABC cells and Myc+ cells belong. Each dot is a cell. P-value was calculated with the Wilcoxon signed-rank test. (G) Gating strategy for FACS of ABC and Myc+ cells in spleens. (H) Percentages of ABC, Myc+, follicular and marginal zone B cells to total B cells in spleens of young and old animals. (I) Cell sizes of follicular and Myc+ B cells in four young and ten old spleens. (J) Representative images of follicular and Myc+ B cells immunolabeled for c-Myc and for DNA with DAPI. Scale bar, 10 µm. (K) Cell sizes and (L) mean c-Myc intensity for follicular and Myc+ B cells estimated from confocal images for two independent experiments. P-values were calculated with two-sided Student t-test, unless otherwise specified, **** p-value < 0.0001.

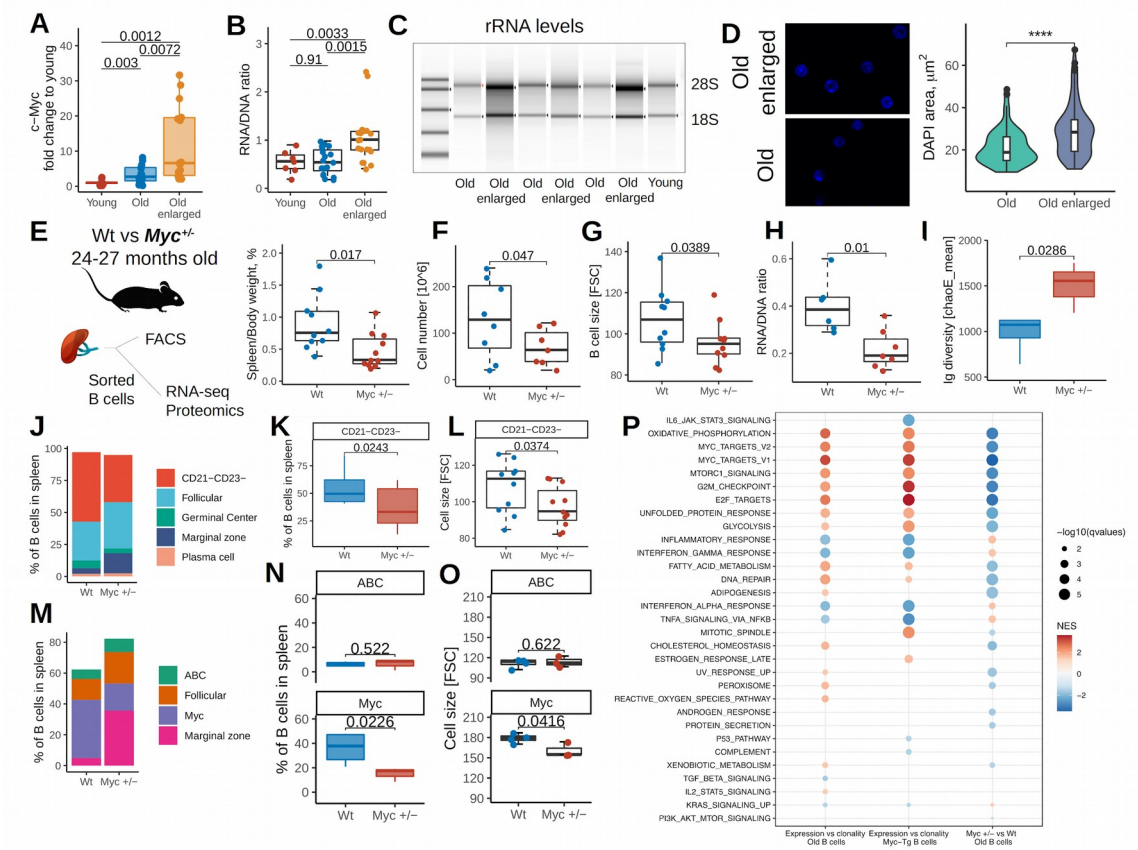


Figure 5. Myc as a driver of B cell size increase and clonal expansion. (A) c-Myc fold change in young, old and old enlarged FACS-sorted B cells normalized for DNA concentration and measured by qRT-PCR, p-value was calculated with a two-sided Student t-test. (B) The total RNA to DNA ratio in young, old and old enlarged B cells measured by Qubit, p-value was calculated with a two-sided Student t-test. (C) Tapestation gel showing rRNA 28S and 18S abundance in young, old and old enlarged B cells. (D) Representative images of old and old enlarged B cell nuclei stained for DNA with DAPI and the measured DAPI area for three independent experiments, and quantification of nuclei size. P-value was calculated with a two-sided Student t-test, **** p-value < 0.0001. Scale bar, 5 µm. (E) Experimental strategy for the analysis of wild type vs *Myc*$^{+/-}$ mice (left). Spleen weight normalized by body weight in 24-27 months old wild type and *Myc*$^{+/-}$ female mice (right). (F) Total number of nucleated cells in spleens of 24-27 months old Wt and *Myc*$^{+/-}$ female mice, counted by automated counter and stained with Trypan Blue. (G) Mean B cell size of 24-26 months old wild type and *Myc*$^{+/-}$ female mice measured by FACS. (H) Total RNA to DNA ratio in splenic B cells sorted from 24-27 months old Wt and *Myc*$^{+/-}$- female mice and measured by Qubit. (I) Ig diversity (chaoE_mean) estimated from RNA-sequencing data of FACS-sorted B cells from 24-27 months old wild type and *Myc*$^{+/-}$ female mice using mixcr and VDJtools software. (J) Percentage of CD21^{-}CD23^{-}, follicular, germinal center, marginal zone and plasma B cells to total B cells in 24-26 months old wild type and *Myc*$^{+/-}$ female mice. CD21^{-}CD23^{-} population was gated as CD19^{+}CD21^{-}CD23^{-}, follicular B cells - as CD19^{+}CD21^{int}CD23^{+}, marginal zone B cells - as CD19^{+}CD21^{+}CD23^{low}, germinal center B cells - as CD19^{+}Fas^{+}, plasma cells - as CD19^{+}CD138^{+}. (K) Percentages of ABCs and Myc+ B cells in spleens of 24-26 months old wild type and *Myc*$^{+/-}$ female mice. ABC cells were gated as CD19^{+}CD21^{-}CD23^{-}CD29^{+}CD24^{lo} and Myc+ B cells were gated as CD19^{+}CD21^{-}CD23^{-}CD29^{+}CD24^{hi}. (L) Mean size of CD21^{-}CD23^{-} B cells in spleens from 24-27 months old wild type and *Myc*$^{+/-}$ female mice measured by FACS. (M) Percentages of CD21^{-}CD23^{-}, ABC, Myc+, follicular and marginal zone B cells in wild type and *Myc*$^{+/-}$ 26 months old female mice. (N) Percentages of ABCs and Myc+ B cells of total splenic B cells (CD19+) in 26 months old wild type and *Myc*$^{+/-}$ female mice. (O)

Cell size of ABCs and Myc+ B cells in 24-26 months old wild type and *Myc*$^{+/-}$ female mice measured by FACS. (P) Gene set enrichment analysis for hallmark pathways in msigdb database for differential gene expressions: 1) correlated with top B cell clone size from 27 months old C57BL/6 mice, 2) correlated with top clone size in B cells of 8 weeks old Emu-Myc female mice, and 3) changed in B cells of 26 months old *Myc*$^{+/-}$ mice compared to their wild type siblings. Dots are sized according to -log10 of q-values and colored according to enrichment scores. P-values were calculated with a one-sided Student t-test, unless otherwise specified.

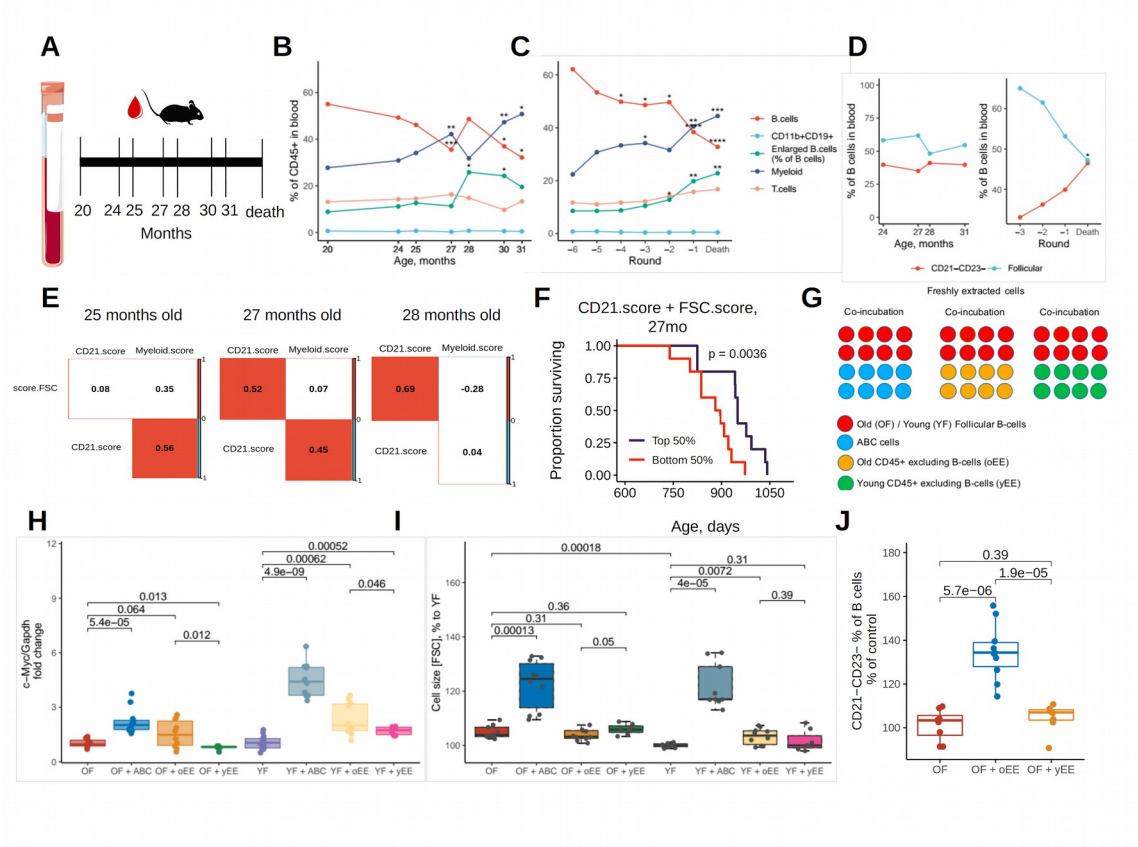
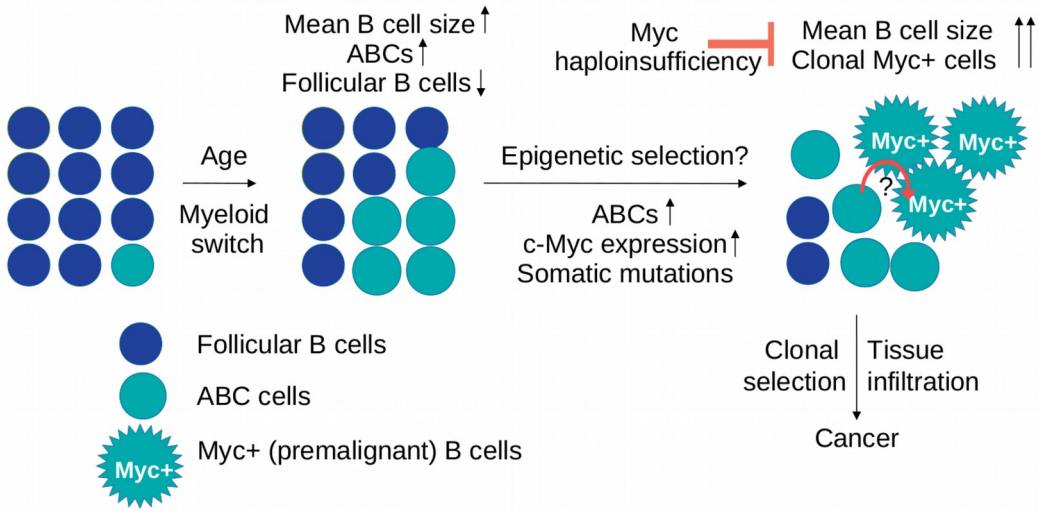


Figure 6. Longitudinal analysis reveals B cell size and CD21-CD23- cells as predictors of lifespan, and co-culture experiments point to their origin. (A) Schematic representation of the longitudinal experiment. Blood was collected from tails of C57BL/6 female mice across seven timepoints. (B) Mean chronological changes in blood cell composition in 25 old mice. Myeloid cells were gated as $CD45^{+}CD3^{-}CD11b^{+}CD19^{-}$, B cells as $CD45^{+}CD11b^{-}CD3^{-}CD19^{+}$, T cells as $CD45^{+}CD3^{+}CD11b^{-}CD19^{-}$, enlarged B cells as $CD45^{+}CD11b^{-}CD3^{-}CD19^{+}FSC^{hi}$, and $CD11b^{+}CD19^{+}$ cells as $CD45^{+}CD11b^{+}CD3^{-}CD19^{+}$. (C) Same as in (B) but aligned to the last blood collection before animal death, -1 is one round before last blood collection, -2 is two rounds before last blood collection, and so on. (D) Chronological changes in B cell composition for 23 mice: Follicular B cells were gated as $CD19^{+}CD21^{int}CD23^{+}$, and $CD21^{-}CD23^{-}$ B cells as $CD19^{+}CD21^{-}CD23^{-}$. (E) Same as in (D) but aligned to the death round which is the last blood collection before the animal died, -1 is one round before last blood collection, -2 is two rounds before last blood collection, and so on. (E) Correlation between the cell size score (FSC.score), myeloid score and $CD21^{-}CD23^{-}$ (CD21.score) calculated for 25, 27 and 28 months old C57BL/6 female mice. Red squares are statistically significant ($p<0.05$) positive correlations; p-values were calculated with ctest. (F) Survival curves for mice with higher and lower combination of $CD21^{-}CD23^{-}$ and FSC scores in 27 months old mice. P-value was calculated with a log-rank test. Higher and lower scores were defined as those above and below the median score in a group. (G) Scheme for co-culture experiment with FACS-sorted cells from spleens of young (6 months old) and old (27-29 months old) C57BL/6 female mice. (H) c-Myc fold change to GAPDH level measured with qRT-PCR in young (YF) and old follicular (OF) B cells incubated with i) ABCs from old spleens, ii) cells from young spleens excluding B cells (yEE), or iii) cells from old spleens excluding B cells (oEE). Fold changes are normalized to OF means and to YF means within OF and YF groups for easier comparisons. Results are from two independent experiments, 3 biological replicas per experiment, 2 technical replicas per biological replica. (I) Mean cell size of follicular B cells after incubation with ABC cells, yEE or oEE measured with FACS.

Results shown are from three independent experiments, 3-4 biological replicas per experiment. (J) Percentages of $CD21^{-}CD23^{-}$ B cells after incubation with yEE and oEE measured by FACS. Due to big variability in cell viability each value is normalized to OF value within the experiment. Results are shown from three independent experiments, 3-4 biological replicas per experiment. P-values in B-D and H-J were calculated with a two-sided Student t-test.



Graphical figure. A mechanism for the development of age-related B-cell lymphoma.