**Cancer Cell, Volume *22***

# Supplemental Information

# Synergy between PI3K Signaling and MYC in Burkitt Lymphomagenesis

**Sandrine Sander, Dinis P. Calado, Lakshmi Srinivasan, Karl Köchert, Baochun Zhang, Maciej Rosolowski, Scott J. Rodig, Karlheinz Holzmann, Stephan Stilgenbauer, Reiner Siebert, Lars Bullinger, and Klaus Rajewsky**

## Inventory of Supplemental Information

## Supplemental Data

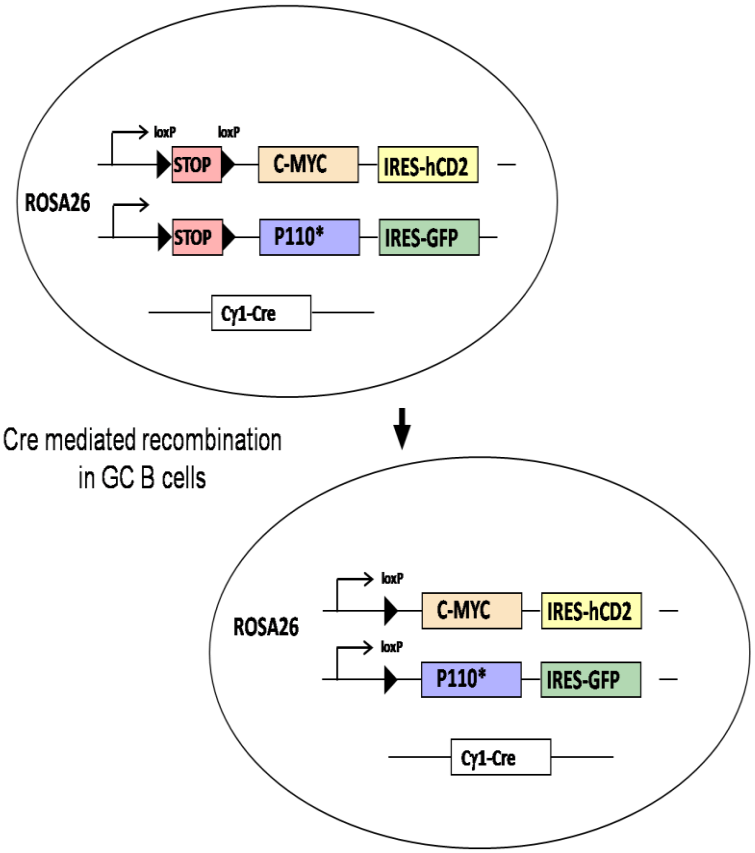


**Figure S1, related to Figure 1.**

**Schematic overview of conditional mice.**

Conditional transgenic mice were generated by ROSA26 locus targeting. These mice carry loxP flanked stop cassettes in front of either human *c-MYC* cDNA or a cDNA encoding a constitutively active form of PI3K, called *P110**. *MYC* transgene expression is controlled by the CAG promoter whereas *P110** transcription is induced by the endogenous ROSA26 promoter. To facilitate transgene detection IRES linked reporter genes succeed the cDNAs (*MYC: hCD2; P110*: GFP*). These mice were bred to *Cγ1-cre* animals in which an IRES coupled Cre coding sequence was introduced into the *Cγ1* locus. In GC B cells undergoing sterile *Cγ1* transcription, recombinase mediated excision of the STOP cassettes occur resulting in *MYC* and *P110** transgene expression.

**Table S1, related to Figure 2.**

**Tumor origin in MYC and P110* co-expressing animals.**

| tumor | origin | Macroscopical liver infiltration |
|---|---|---|
| #6 | PP | + |
| #7 | PP | - |
| #8 | PP | + |
| #10 | PP | + |
| #11 | PP | + |
| #18 | mediastinal LN | - |
| #19 | PP | + |
| #20 | peripheral LN | + |
| #77 | spleen | + |
| #82 | PP | + |
| #87 | PP | - |
| #88 | PP | + |
| #78 | PP | + |
| #80 | spleen | + |
| #81 | spleen | + |
| #83 | mediastinal LN | + |
| #84 | mesenteric LN | + |
| #85 | PP | + |
| #86 | liver | na |
| #89 | peripheral LN | + |
| #91 | PP | + |

*PP = Peyer's patches*
*LN = lymph node*
*na = not applicable*

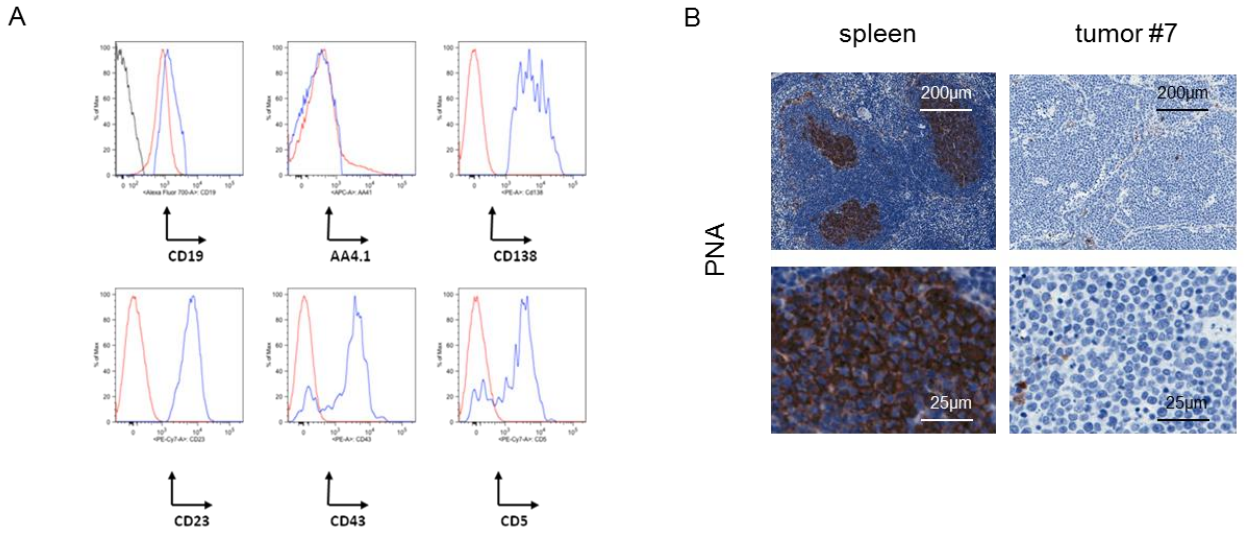


**Figure S2, related to Figure 3**

**Representative FACS analysis and PNA binding on tumor cells.**

**(A)** Expression of B cell characteristic markers on MYC and P110* co-expressing tumor cells (red) and the below listed controls [CD19 control (blue)= mature B cells (spleen)**;** CD19 control (black)= non-B cells (spleen)**;** AA4.1 control (blue)= mature B cells (spleen)**;** CD138 control (blue)= plasma cells (spleen)**;** CD23 control (blue)= marginal zone B cells (spleen)**;** CD43 control (blue)= B1 cells (peritoneal cavity)**;** CD5 control (blue)= B1a cells (peritoneal cavity)]. **(B)** Representative immunohistochemical staining for PNA in tumor #7 and control spleen (derived from a *Cγ1-cre, R26Stop$^{FL}$eYFP* animal 10 days after SRBC immunization). In total 7 tumors were analyzed.

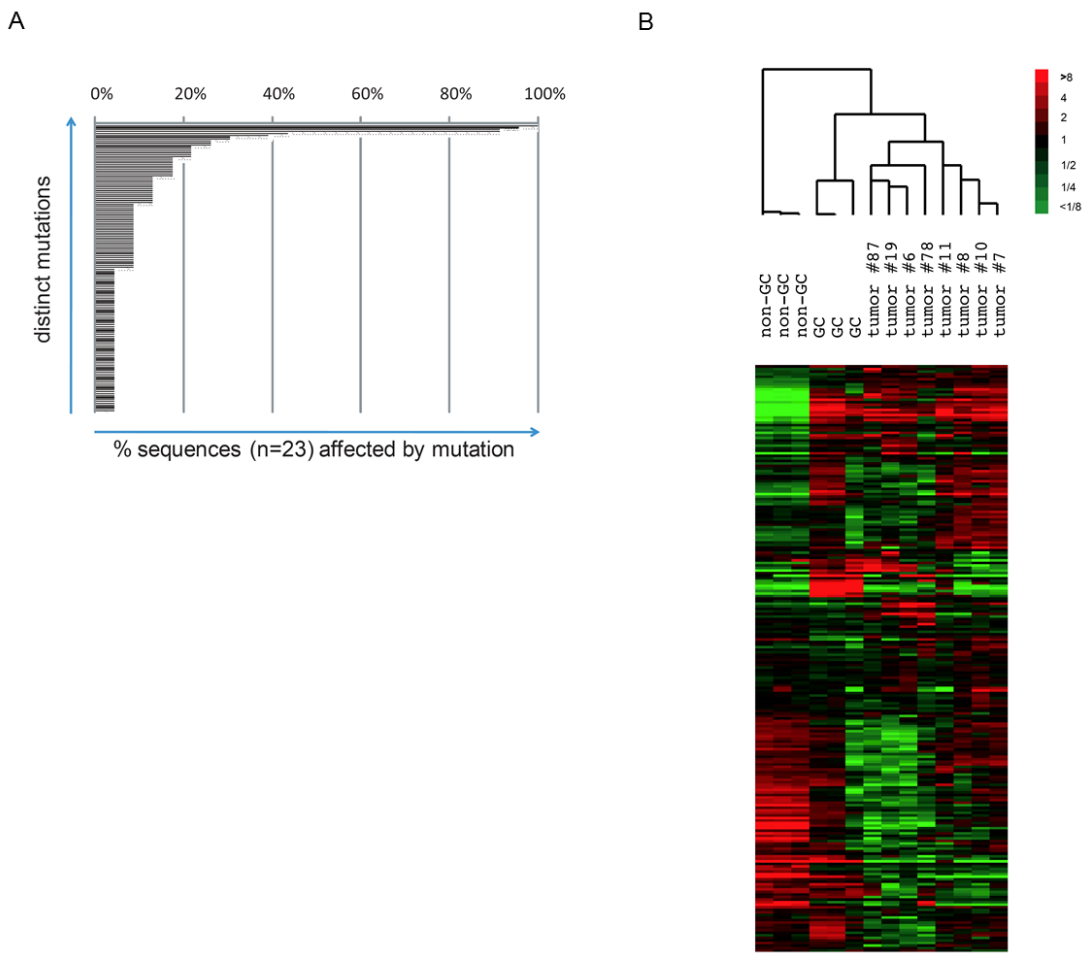


**Figure S3, related to Figure 4.**

**MYC and P110* co-expressing tumors show ongoing somatic hypermutation (SHM) and express GC B cell specific genes.**

**(A)** Mutations identified in the $V_H$ gene region of tumor#7 are listed on the y axis. The frequency how often the same $V_H$ mutation was found in the 23 sequences analyzed is depicted on the x axis. While the majority of mutations is seen only once (1/23: 4.3%), some mutations are present in all sequences (23/23:100%). This intraclonal heterogeneity of $V_H$ mutations was detectable in 3/3 tumors.

**(B)** Hierarchical cluster analysis based on relative transcript levels of 233 genes comprised in a GC B cell signature (Green et al., 2011) in GC B cells (n=3), MYC and P110* co-expressing tumor samples (n=8) and non-GC B cells (n=3); mean-centered $\log_2$ gene expression ratios are depicted by color scale.

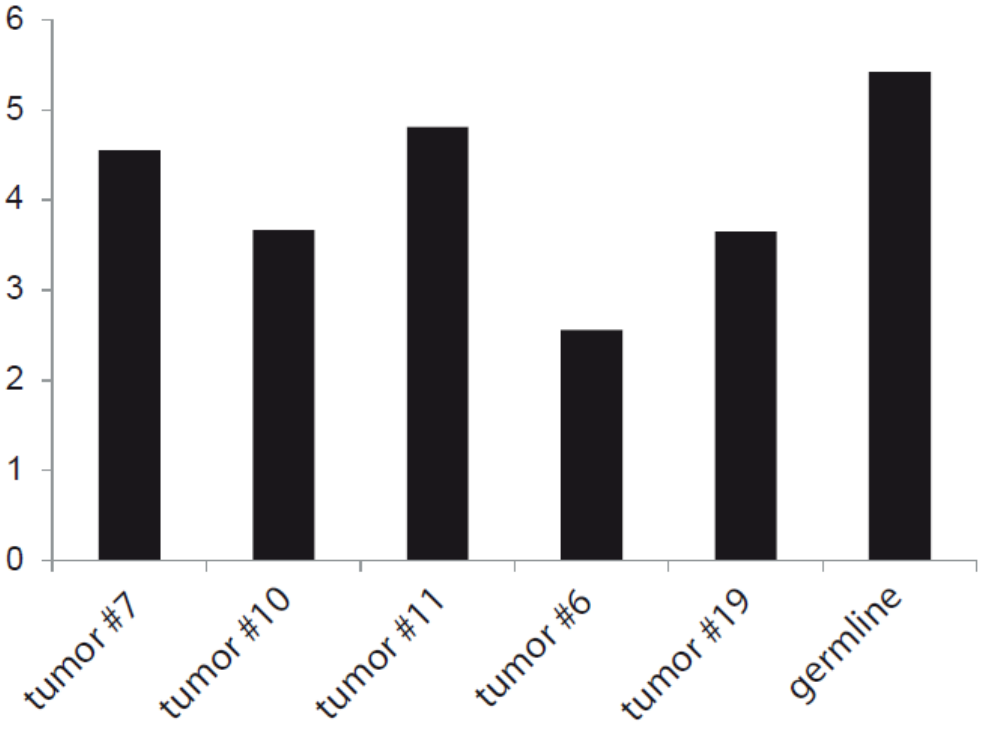


**Figure S4, related to Figure 6.**

**Somatic mutations in MYC and P110* co-expressing tumors infrequently affect AID mutation hot spots.**

AID hot spot hit frequency of mutations identified in MYC and P110* co-expressing tumors (n=5) and germline DNA.

**Table S2, related to Figure 6.**

**Somatically acquired mutations in MYC and P110* tumors.**

| MGI gene symbol | number of affected tumors |
|---|---|
| *Kifc5b* | 5 |
| *4930427A07Rik* | 4 |
| *Camk2d* | 4 |
| *Ndufs6* | 4 |
| *Serpina12* | 4 |
| *EU599041* | 3 |
| *Herpud2* | 3 |
| *Ppig* | 3 |
| *Serpina3k* | 3 |
| *Serpina3m* | 3 |
| *Sfrs18* | 3 |
| *Ston2* | 3 |
| *Vmn2r114* | 3 |
| *Vmn2r115* | 3 |
| *Dmd* | 2 |
| *Dyrk1b* | 2 |
| *Eif2s3y* | 2 |
| *Foxo1* | 2 |
| *Kifc1* | 2 |
| *Lrp1b* | 2 |
| *Muc19* | 2 |
| *Pcdhb22* | 2 |
| *Rps18* | 2 |
| *Serpina3n* | 2 |
| *Slc16a14* | 2 |
| *Tex16* | 2 |
| *Vmn2r117* | 2 |
| *Zfp386* | 2 |
| *Sept9* | 1 |
| *Sept14* | 1 |
| *1700125H20Rik* | 1 |
| *2210018M11Rik* | 1 |
| *3110082D06Rik* | 1 |
| *4930468A15Rik* | 1 |
| *4930480E11Rik* | 1 |
| *9830107B12Rik* | 1 |
| *A330021E22Rik* | 1 |
| *A630007B06Rik* | 1 |
| *A730008H23Rik* | 1 |
| *AA467197* | 1 |

| MGI gene symbol | number of affected tumors |
|---|---|
| *Aars* | 1 |
| *Actb* | 1 |
| *Adamts16* | 1 |
| *Adamts20* | 1 |
| *Ambp* | 1 |
| *Ank1* | 1 |
| *Ankrd7* | 1 |
| *Anks1b* | 1 |
| *Aplp2* | 1 |
| *Arap2* | 1 |
| *Arhgap39* | 1 |
| *Arid4b* | 1 |
| *Asb18* | 1 |
| *Atad1* | 1 |
| *Atp1b4* | 1 |
| *B3gnt2* | 1 |
| *Bank1* | 1 |
| *BC003965* | 1 |
| *BC021891* | 1 |
| *BC053749* | 1 |
| *BC057079* | 1 |
| *Bcor* | 1 |
| *Brp44* | 1 |
| *C130074G19Rik* | 1 |
| *C430048L16Rik* | 1 |
| *Calr4* | 1 |
| *Ccdc40* | 1 |
| *Ccnd3* | 1 |
| *Cd79b* | 1 |
| *Cdh10* | 1 |
| *Cdh23* | 1 |
| *Cebpz* | 1 |
| *Cep110* | 1 |
| *Cep192* | 1 |
| *Chpt1* | 1 |
| *Chrd* | 1 |
| *Cit* | 1 |
| *Cnpy2* | 1 |
| *Cpne2* | 1 |
| *Cpne8* | 1 |

| MGI gene symbol | number of affected tumors |
|---|---|
| *Creb3l2* | 1 |
| *Csmd1* | 1 |
| *Csmd3* | 1 |
| *Ctdsp1* | 1 |
| *Cts6* | 1 |
| *Cyp2j9* | 1 |
| *Cyp4f18* | 1 |
| *Dcaf12l1* | 1 |
| *Ddx24* | 1 |
| *Dnahc7b* | 1 |
| *Dock11* | 1 |
| *Dpyd* | 1 |
| *Duoxa2* | 1 |
| *Dusp2* | 1 |
| *Dux* | 1 |
| *Eef1b2* | 1 |
| *Emd* | 1 |
| *Eomes* | 1 |
| *Epb4.1l3* | 1 |
| *Epb4.2* | 1 |
| *Ephb3* | 1 |
| *Ermn* | 1 |
| *Exo1* | 1 |
| *Extl3* | 1 |
| *Fads3* | 1 |
| *Fam135a* | 1 |
| *Fam83a* | 1 |
| *Fancc* | 1 |
| *Farp2* | 1 |
| *Fmo9* | 1 |
| *Foxl1* | 1 |
| *Foxr1* | 1 |
| *Fpr3* | 1 |
| *Frmd7* | 1 |
| *Gas2l3* | 1 |
| *Git2* | 1 |
| *Gm14725* | 1 |
| *Gm4861* | 1 |
| *Gm5134* | 1 |
| *Gp5* | 1 |

| MGI gene symbol | number of affected tumors |
|---|---|
| *Gpd1* | 1 |
| *Gramd1b* | 1 |
| *Gulp1* | 1 |
| *Hcfc1* | 1 |
| *Herc1* | 1 |
| *Hhipl1* | 1 |
| *Hist1h1b* | 1 |
| *Hist1h1e* | 1 |
| *Hist1h2bc* | 1 |
| *Hist1h3b* | 1 |
| *Hist1h3d* | 1 |
| *Hist1h4h* | 1 |
| *Hmcn1* | 1 |
| *Hnf1a* | 1 |
| *Inppl1* | 1 |
| *Jund* | 1 |
| *Kat2a* | 1 |
| *Kbtbd12* | 1 |
| *Kcnmb2* | 1 |
| *Kcnn1* | 1 |
| *Kdm6a* | 1 |
| *Krt72-ps* | 1 |
| *Lamc3* | 1 |
| *Lancl3* | 1 |
| *Lim2* | 1 |
| *Lonrf3* | 1 |
| *Lrp12* | 1 |
| *Lrrc27* | 1 |
| *Lrrtm1* | 1 |
| *Man2a2* | 1 |
| *Mcl1* | 1 |
| *Mcm3* | 1 |
| *Mdn1* | 1 |
| *Med1* | 1 |
| *Med23* | 1 |
| *Mllt3* | 1 |
| *Mllt4* | 1 |
| *Murc* | 1 |
| *Myh7b* | 1 |
| *Mynn* | 1 |

| MGI gene symbol | number of affected tumors |
|---|---|
| *Myo18a* | 1 |
| *Myo9a* | 1 |
| *Myof* | 1 |
| *Nav3* | 1 |
| *Ncaph* | 1 |
| *Ncf4* | 1 |
| *Ncoa1* | 1 |
| *Ncoa7* | 1 |
| *Nrg2* | 1 |
| *Nudt16* | 1 |
| *Nvl* | 1 |
| *Oca2* | 1 |
| *Olfr1043* | 1 |
| *Olfr1094* | 1 |
| *Olfr1188* | 1 |
| *Olfr508* | 1 |
| *Ophn1* | 1 |
| *Pafah1b3* | 1 |
| *Paox* | 1 |
| *Pclo* | 1 |
| *Pdcd10* | 1 |
| *Pfdn2* | 1 |
| *Phactr4* | 1 |
| *Phf16* | 1 |
| *Pkhd1* | 1 |
| *Pla2g2f* | 1 |
| *Plin2* | 1 |
| *Pnpt1* | 1 |
| *Ppbp* | 1 |
| *Prdx4* | 1 |
| *Prkca* | 1 |
| *Prl* | 1 |
| *Prl4a1* | 1 |
| *Prr5* | 1 |
| *Prune2* | 1 |
| *Psmc5* | 1 |
| *Rab3gap2* | 1 |
| *Rap1gap* | 1 |
| *Rasgrp3* | 1 |
| *Rbl2* | 1 |

| MGI gene symbol | number of affected tumors |
|---|---|
| *Reg2* | 1 |
| *Rhbdl1* | 1 |
| *Rora* | 1 |
| *Rpl18* | 1 |
| *Rpl27a* | 1 |
| *Rrs1* | 1 |
| *Sarm1* | 1 |
| *Sbf1* | 1 |
| *Scaf11* | 1 |
| *Sdpr* | 1 |
| *Senp3* | 1 |
| *Senp5* | 1 |
| *Serpina1d* | 1 |
| *Sez6l* | 1 |
| *Sgol2* | 1 |
| *Sgsm2* | 1 |
| *Siglecg* | 1 |
| *Sipa1l3* | 1 |
| *Sirpb1b* | 1 |
| *Sis* | 1 |
| *Slc5a8* | 1 |
| *Slc6a17* | 1 |
| *Slc7a7* | 1 |
| *Slc8a1* | 1 |
| *Slco6c1* | 1 |
| *Slit1* | 1 |
| *Slit2* | 1 |
| *Slitrk2* | 1 |
| *Slitrk3* | 1 |
| *Smcr7l* | 1 |
| *Smox* | 1 |
| *Snca* | 1 |
| *Snx12* | 1 |
| *Snx30* | 1 |
| *Sorcs3* | 1 |
| *Sptlc2* | 1 |
| *Srcin1* | 1 |
| *Srrt* | 1 |
| *Srsf2* | 1 |
| *Stk38* | 1 |

| MGI gene symbol | number of affected tumors |
|---|---|
| *Suv420h1* | 1 |
| *Svep1* | 1 |
| *Tanc1* | 1 |
| *Taok1* | 1 |
| *Tas2r134* | 1 |
| *Tbx2* | 1 |
| *Tchhl1* | 1 |
| *Tdrd7* | 1 |
| *Tdrd9* | 1 |
| *Tfip11* | 1 |
| *Thap7* | 1 |
| *Tm9sf2* | 1 |
| *Tnn* | 1 |
| *Tom1* | 1 |
| *Trim66* | 1 |
| *Trpc6* | 1 |
| *Trub1* | 1 |
| *Txnl1* | 1 |
| *Uba1* | 1 |
| *Ube3b* | 1 |
| *Ubox5* | 1 |
| *Ulk4* | 1 |
| *Upf2* | 1 |
| *Ush2a* | 1 |
| *Usp25* | 1 |
| *Usp9y* | 1 |
| *Utrn* | 1 |
| *Vcam1* | 1 |
| *Vmn2r107* | 1 |
| *Vmn2r121* | 1 |
| *Wdfy3* | 1 |
| *Wdr35* | 1 |
| *Wdr47* | 1 |
| *Wdr6* | 1 |
| *Wdr7* | 1 |
| *Wee1* | 1 |
| *Whsc1l1* | 1 |
| *Zbtb24* | 1 |
| *Zc3h13* | 1 |

| MGI gene symbol | number of affected tumors |
|---|---|
| *Zdhhc24* | 1 |
| *Zer1* | 1 |
| *Zfp101* | 1 |
| *Zfp316* | 1 |
| *Zfp36l2* | 1 |
| *Zfp598* | 1 |
| *Zfp646* | 1 |
| *Zfp738* | 1 |
| *Zfp87* | 1 |
| *Zic2* | 1 |

## Supplemental Experimental Procedures

### Real time RT-PCR

For the analysis of *Actb* transcript levels we used: (forward primer) 5' TGTTACCAACTGGGACGACA ; (reverse primer) 5' GGGGTGTTGAAGGTCTCAAA). Primers to detect *Aicda* transcript levels were previously described (Sayegh et al., 2003).

### Gene expression profiling and data analysis

Fluorescence ratios were normalized by applying the RMA algorithm using the BRB Array Tools software package (available at http://linus.nci.nih.gov/BRB-ArrayTools.html). For subsequent analyses, we only included probe sets (n=7039) whose expression varied as previously determined (genes with a log intensity variation of *p* value >0.01 were excluded) (Sander et al., 2008).

For the comparison of our tumors with different B cell subtypes as well as other murine lymphomas a previously published GEP data set was used (GSE26408) (Green et al., 2011) and normalized as described above. To determine the similarity of MYC and P110* expressing tumors to GC B cells, combined gene expression data was analyzed by hierarchical clustering (average linkage clustering, distance measure: correlation uncentered). The clustering was based on GC associated genes as determined by Green and colleagues (Green et al., 2011) and included only probe sets that were well measured in the combined data set (n=233).

The same probe set was used to cluster MYC and P110* expressing tumors and (non-) GC B cells which were hybridized on the same platform (average linkage clustering, distance measure: correlation uncentered). Results were visualized using TreeView (Eisen et al., 1998).

For the comparison of MYC and P110* co-expressing tumors vs. *IμHABcl6* induced lymphomas supervised analysis was carried out using Significance Analysis of Microarrays (SAM; false discovery rate, FDR<0.05) (Tusher et al., 2001). Within the SAM-derived gene list, genes

overlapping with two previously published human BL gene expression signatures (Dave et al., 2006; Hummel et al., 2006) were evaluated with regard to their distribution, i.e. up- or down-regulation, in the MYC and P110* co-expressing tumors vs. the *IμHABcl6* induced tumors (Fisher's exact test).

**Pathway activation indices**

To obtain an index of the relative activity of a pathway in a tumor sample, we computed the sum of the values of the 50 metagenes in this sample weighted by the correlations of the metagenes with the activation of this pathway in the corresponding interventional data set. More specifically, let x(i,j), i=1,..50, j=1,..,n be the value of metagene i in patient j. Further, let b(i) be a statistic computed in an interventional data set and representing the correlation of metagene i with the experimental activation of a pathway. Then the sum of b(i) * x(i,j) over i=1,..,50 is the value of the pathway activation index for patient j. Full details of the method will be published elsewhere (Rosolowski M, Läuter J, Horn F, Siebert R, Loeffler M; manuscript in preparation). We used preprocessed data as available at the Gene Expression Omnibus.

**Exome sequencing**

To capture exonic sequences the in solution oligo bait SureSelect Target Enrichment System (Agilent) was used in conjunction with paired end sequencing library preparation kits (Illumina) according to the manufacturer's protocols. Paired-end sequencing with a 100 bp read length was carried out on a HiSeq 2000 (Illumina). Reads were aligned against the UCSC NCBI37/mm9 reference genome (http://genome.ucsc.edu/) using the Burrows-Wheeler Aligner (BWA; http://bio-bwa.sourceforge.net/). For variant detection of single nucleotide variations (SNVs) and insertion/deletions (Indels) SAMtools (http://samtools.sourceforge.net/) was used. Sequencing and data analysis were performed by Axeq (Axeq Technologies).

To determine the frequency of SNVs that coincided with putative AID hotspots, the sequences surrounding a specific SNV (at arbitrary position 0) from SNV relative positions -2 to +1 were taken from the UCSC NCBI37/mm9 assembly. They were then screened for any of the possible AID hot spot sequence motifs (WRCY or the inverse RGYW: W=A/T; R=A/G; Y=C/T). The SNV-AID hot spot hit frequency in a certain tumor was determined by calculating the ratio between SNV-AID hot spot hits and the total number of SNVs in the tumor.

**Supplemental References**


Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A *95*, 14863-14868.

Sander, S., Bullinger, L., Klapproth, K., Fiedler, K., Kestler, H. A., Barth, T. F., Moller, P., Stilgenbauer, S., Pollack, J. R., and Wirth, T. (2008). MYC stimulates EZH2 expression by repression of its negative regulator miR-26a. Blood *112*, 4202-4212.

Sayegh, C. E., Quong, M. W., Agata, Y., and Murre, C. (2003). E-proteins directly regulate expression of activation-induced deaminase in mature B cells. Nat Immunol *4*, 586-593.

Tusher, V. G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A *98*, 5116-5121.