SUPPLEMENTAL INFORMATION

Hijacking a key chromatin modulator creates epigenetic vulnerability for MYC-driven cancer

Zhenhua Yang, Kushani Shah, Theodore Busby, Keith Giles, Alireza Khodadadi-Jamayran, Wei Li, and Hao Jiang

SUPPLEMENTAL METHODS

Flow cytometry analyses and sorting

Single cell suspensions were prepared from bone marrow, spleen or peripheral blood. Red blood cells were removed using ACK lysis buffer (0.15 M NH_4Cl, 10 mM $KHCO_3$, 0.1 nM EDTA, pH 7.2). For bone marrow and splenocyte analyses, cells were stained with antibodies for CD45R (RA3-6B2)-APC; IgM (RMM-1)-PE and IgD (11-26C.2A)-Pacific blue. FACS analysis was performed on LSRFortessa (Becton Dickinson), and data were analyzed using FlowJo software (Tree Star). B220+ spleen cells were enriched with magnetic CD45R (B220) microbeads (Miltenyi Biotec) following manufacturer's instructions. Proliferation and apoptosis were determined by BrdU incorporation and Annexin V staining assays, respectively, as previously described (1). Briefly, for BrdU analysis, 4-5 weeks old mice were intraperitoneally injected with 1.5mg BrdU at 6 hours prior to sacrifice. Cells were stained with cell surface antibodies as described above and then processed with BrdU staining using the FITC-BrdU Flow kit (BD Pharmingen) following the manufacturer's instructions. For apoptosis assays, cells were harvested and stained with antibodies as described above. After washing twice with cold PBS containing 3% heat-inactivated FBS, the cells were then incubated with FITC-Annexin V (BD Pharmingen) and 7-amino-actinomycin D (7-AAD) for 15 minutes in binding buffer (10mM HEPES, 140 mM NaCl and 2.5mM $CaCl_2$) at room temperature in dark. The stained cells were analyzed immediately by flow cytometry. Intracellular ROS levels were detected by ROS-ID™ Total ROS detection kit (Enzo Life Sciences, Cat# ENZ-51011) following manufacturer's instructions. Briefly, cells were stained with ROS detection reagent at 37°C for 15 min after being stained with the B220 antibody described as above. ROS levels in the gated population were quantified using flow cytometry. Proliferation and apoptosis assays for cultured P493-6 cells and transduced MEFs were performed similarly for the splenocytes, except that cells were incubated with BrdU for 30 min before the proliferation assays.

RNA extraction, chromatin immunoprecipitation (ChIP), ATAC assay, DNase I hypersensitivity assay, and quantitative PCR (qPCR)

Total RNAs were isolated using RNeasy Plus Mini Kit (Qiagen). Some of RNAs were reverse transcribed with SuperScript III (Invitrogen). ChIP assays were performed as described (2) using anti-Myc (Santa Cruz Biotechnology, sc-764x) and anti-H3K4me3 (Millipore 07-473). ATAC (3) and DNase I hypersensitivity (4) assays were performed following the published protocols with modifications. For ATAC assays, primary MEFs from *Dpy30*$^{F/F}$ and *CAG-CreER; Dpy30*$^{F/F}$ embryos (1) were treated with 4-hydroxytamoxifen at 0.5 μM for 4 days. 70,000 primary MEFs were used as input and for reaction with transposase. Cells were incubated with transposase for 45 minutes at 37°C with gentle agitation. Input samples were treated same as the transposase-incubated cells up until the transposase reaction step. Input samples were then sonicated for 45 minutes (instead of incubating with the transposase). qPCR was performed to detect change in accessibility upon *Dpy30* KO. Input was used to normalize transposase-treated samples. For DNase I hypersensitivity assays, 5 million spontaneously immortalized MEFs were used for each reaction. (The large number of cells required for DNase I hypersensitivity assay made it difficult to use primary early passage MEFs.) DNase I enzyme and reaction buffer were purchased from NEB (Cat# M0303L). Undigested and digested samples were

purified using PCR purification kit commercially available from Qiagen (Cat# 28106). qPCR was performed to detect change in accessibility upon Dpy30 KD. Undigested samples were used for normalization. qPCR was performed with SYBR Advantage qPCR Premix (Clontech) on a ViiA7 Real-Time PCR System (Applied Biosystems). Primers used are listed in Supplemental Table 7. Relative expression levels were normalized to *Actb*. For ChIP results, percent input was first calculated from ChIP qPCR results as described (2), and ChIP enrichment fold was calculated as the ratio of the percent input value for each locus over that for the indicated sites and samples.

Microarray, RNA-seq, ChIP-seq, and data analyses

For microarray analyses, total RNAs were labeled using the Illumina TotalPrep RNA Amplification Kit (Life Technologies, Cat# AMIL1791) following the manufacturer's protocol, and submitted to the Genomics Resource Center at the Rockefeller University for hybridization to the Illumina 12-sample BeadChip. Genes with Detection P value above 0.05 in all samples were filtered. Signals were normalized against the median in each sample.

For RNA-seq with spike-in control, the ERCC RNA Spike-In Mix (Ambion, 4456740) was used following instructions of the product. The amount of spike-in added was calibrated to the RNA yield to ensure the spike-in signal was in the appropriate dynamic range. Specifically, 2µl of a 1:200 diluted ERCC Spike-In Mix 1 was added to one of the samples that had a total RNA of 465ng. Between 1.8µl to 5.4µl of 1:200 diluted ERCC Spike-In Mix 1 were added to other samples (between 240 to 800ng total RNAs) at the same ratio of Spike-in/cell number.

Library preparation and sequencing for RNA and ChIP (and their corresponding input) DNA were performed at the Genomic Services Lab at HudsonAlpha Institute for Biotechnology (Huntsville, AL). Briefly, the quality of the total RNA and DNA was assessed using the Agilent 2100 Bioanalyzer. Two rounds of polyA+ selection was performed for RNA samples, followed by conversion to cDNAs. The mRNA library generation kits (Agilent, Santa Clara, CA) and TruSeq ChIP Sample Prep Kit (Illumina) were used generate sequencing libraries per manufacturer's instructions. The indexed DNA libraries were quantitated using qPCR in a Roche LightCycler 480 with the Kapa Biosystems kit for library quantitation (Kapa Biosystems, Woburn, MA) prior to cluster generation. Clusters were generated to yield approximately 725 K to 825 K clusters/mm2. Cluster density and quality were determined during the run after the first base addition parameters were assessed. Sequencing was performed on Illumina HiSeq2500 with sequencing reagents and flow cells providing up to 300 Gb per flow cell.

For RNA-seq, we obtained 35-55 million of 51bp paired end reads for each RNA sample. All the reads were mapped to the mouse reference genome (GRCm38/mm10) using TopHat (v2.0.13). The alignment was guided using a Gene Transfer File (GTF version GRCm38.85). Low quality mapped reads (MQ<30) were removed from the analysis. Read count tables were generated using HTSeq (v.0.6.0) based on the Ensembl gene annotation file (Ensembl GTF version GRCm38.85) (5). All the read count tables were then normalized based on their ERCC RNA Spike-In Mix size factors calculated using the DESeq R package (v.3.0), and Deferential Expression (DE) analyses were performed using DESeq (v3.0) (6). All of the downstream statistical analyses and generating plots were performed in R (v3.1.1) (http://www.r-project.org/). To generate the heat map, upregulated and downregulated genes were clustered separately using R dist function in Euclidian mode and then merged to maintain the clustered order of the genes. The values in the matrix were scaled based on columns and not rows using heatmap.2 R function to make the two different data comparable in a single heat map. KEGG pathway analysis was performed using clusterProfiler R package (v3.0.0) (7). Gene ontology analysis was performed with DAVID 6.7 (https://david-d.ncifcrf.gov/), GSEA was performed at http://software.broadinstitute.org/gsea/index.jsp. The gene sets were from GSE51011 (8). Myc bound genes were selected if the Myc ChIP signal was positive in either pre-tumor or tumor samples, and ranked by q value (pre-tumor vs. control) from lowest to highest. Myc-bound UP gene set contains the

top 500 genes that were upregulated with a q value less than 0.001, and Myc-bound DN gene set contains 891 genes that were downregulated with a q value less than 0.01, from control to pre-tumor samples. Genes were rank-ordered in descending order according to the fold change. The list of pre-ranked genes was analyzed with the gene set matrix composed file (.gmx) from curated data. Significant gene sets enriched by Dpy30 KD were identified using nominal P value of 0.05. All analyses were performed using GSEA v2.0 software with pre-ranked list and 1000 data permutations.

For ChIP-seq, we obtained 25-35 million of 50bp single end reads for each DNA sample. Fastq files were first quality filtered using the FASTX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html Key: citeulike:9103573) "fastq_quality_filter" tool, with the following arguments: -q 20 and –p 80, which outputs reads that have a mapq score of at least 20 at least 80% of the bases. The quality filtered reads were then collapsed to eliminate PCR duplicates using the FASTX toolkit "fastx_collapser". The resultant fasta file was aligned using STAR ver. 2.4 (9). The SAM file outputs were converted to BAM using samtools "view" (10) and the BAM converted to BED using BEDtools "bamToBed" (11). The coverage of H3K4me3 and Myc plus/minus 5kb of all transcription start sites (profiles) were generated using deepTools2.0 (12). The workflow was carried out as follows: BAM files were input into bamCoverage to generate a bigwig "bw" file using the –scale argument to normalize coverage according to reads per million. The bw file was then used as an input into the computeMatrix tool using the reference-point argument. The output file was then input into the plotProfile to generate the profiles. For each analysis the GENCODE hg19 genome was used with 52170 genes. Screen captures of coverage were generated from loading a tdf file into the integrative genomic viewer (IGV) (13). The tdf files were generated from bedgraph files using IGVtools "totdf". The bedgraph files were generated from BED files using the BEDTools "genomeCoverageBed" with the –bg option set and the "scale" set to reads per million. The correlations plots comparing Myc and H3K4me3 were done using bedTools coverageBed –counts to quantify the number of total reads that were present within 5kb of each transcription start site. The reads were then normalized to reads per million (RPM) and a scatterplot was generated in Rstudio. The R-squared and p-values of the correlation were generated using the Wilcox.test within Rstudio. Genes were sorted according to their overall MYC binding levels and binned into quintiles (first = top 20%, second = 21-40%, etc.). Box and Whisker plots were also generated in Rstudio using "boxplots".

Protein preparation, binding, and immunoblotting assays

MBP-tagged C-MYC was cloned into pMal-c2x (New England BioLabs, N8076S), and induced by Isopropyl β-D-1-thiogalactopyranoside (IPTG) in BL21 STAR DE3 *E. coli* cells (Invitrogen). Cells were lysed in 300mM KCl, 50mM Tris (pH 7.5), 20% Glycerol, 0.1% NP40, protease inhibitor cocktail (Roche), 0.5mg/ml lysozyme, sonicated and cleared by centrifugation. The supernatant was incubated with Amylose beads (NEB) at 4°C for 2 hours and extensively washed with BC500 (50 mM Tris [pH 7.4], 500 mM KCl, 20% glycerol, 0.2mM EDTA) and 0.5% NP40, followed by BC100 and 0.1% NP40. Bound protein was checked by SDS-PAGE and coomassie blue staining. To express ASH2L and DPY30, Sf9 insect cells (Invitrogen) were infected with baculoviruses expressing FLAG-ASH2L (F-ASH2L) and FLAG-HA-DPY30 (FH-DPY30)(2) for 72 hours. Cells were lysed in 500mM KCl, 50mM Tris (pH 7.5), 20% Glycerol, 0.1% NP40, protease inhibitor cocktail (Roche), and cleared by centrifugation. The supernatant was incubated with anti-FLAG M2 resin (Sigma) at 4°C for 6 hours and extensively washed with BC500 and 0.5% NP40, followed by BC100 and 0.1% NP40. Bound proteins were eluted with 0.4mg/ml FLAG peptide (Sigma) in BC100 and 0.1% NP40. For in vitro binding assays, 10μg MBP-MYC or 10 μg MBP on resin was pre-incubated with 200μl binding buffer (BC300, 0.05% NP40, and protease inhibitor cocktail), and further incubated at 4°C overnight after addition of 4 μg purified F-ASH2L or FH-DPY30. Resin was extensively washed with the binding buffer and was checked for bound proteins by SDS-PAGE and immunoblotting.

To determine protein expression levels, cells were lysed by lysis buffer (1% SDS, 10mM EDTA and 50mM Tris-Cl, pH 7.5) with fresh added Protease Inhibitor Cocktail (Roche) and boiled for 5 min. Proteins were resolved on SDS-PAGE gel followed by immunoblotting using antibodies described before (2) for Dpy30 and H3K4me3, from Santa Cruz Biotechnology for c-Myc (sc-764), β-Actin (sc-47778), and p53 (sc-6243), or from other commercial sources: anti-Gapdh (Chemicon, MAB374); anti-Flag (Sigma, A8592); anti-Flag (Sigma, A2220, M2 beads); anti-phospho-Histone H2A.X (Ser139) (Millipore, 05-636), and anti-MBP (New England BioLabs, E8032S). Key signals were quantified using ImageJ program followed by subtraction of a blank area and normalization to the indicated reference signal intensity.

SUPPLEMENTAL TABLE LEGENDS

Table S1. Microarray analyses of global gene expression in P493-6 cells.
P493-6 cells were stably infected with viruses expressing scramble control (Scr) shRNA, or two different shRNAs for *DPY30* (shRNA #2, D2 and shRNA #5, D5). Cells were subject to experimental scheme in Figure 2D. Microarray assays were performed using RNAs from cells treated with tetracycline for 3 days (to turn off Myc) and right before washing away tetracycline (0h) as well as from cells 4h after culture in Tetracycline-free medium after tetracycline withdrawal (4h).

Table S2. Genome-wide MYC binding in control and DPY30 KD P493-6 cells.
P493-6 cells were stably infected with viruses expressing scramble control (Scr) shRNA, or *DPY30* shRNA #2. MYC ChIP-seq results from these cells were analyzed. Genes were ranked in the order of the Control/KD ratio from highest to lowest.

Table S3. Genome-wide H3K4me3 in control and DPY30 KD P493-6 cells.
P493-6 cells were stably infected with viruses expressing scramble control (Scr) shRNA, or *DPY30* shRNA #2. H3K4me3 ChIP-seq results from these cells were analyzed.

Table S4. Curation of Myc targets for GSEA.
This file shows curation of genes from GSE51011 for use as gene sets in GSEA. Genes in the tab "Myc bound" are genes that have positive Myc ChIP signals in any of the P or T samples in the tab "All genes". Genes in the tab "Myc bound, up in EuMyc" are genes that that are upregulated from non-transgenic control to pre-tumor samples, and ranked from lowest (most significant) to highest (least significant) q value comparing pre-tumor and non-transgenic control. The top 500 genes were used in the "Myc bound UP" gene set for GSEA. Genes in the tab "Myc bound, down in EuMyc" are genes that that are downregulated from non-transgenic control to pre-tumor samples, and ranked from lowest (most significant) to highest (least significant) q value comparing pre-tumor and non-transgenic control. All genes with q value < 0.01 were used in the "Myc bound DN" gene set for GSEA.

Table S5. Clustered genes.
A total of 305 Genes that were down- or upregulated over 4 fold in the average of 4 *Eμ-Myc*; *Dpy30*$^{+/-}$ compared to *Eμ-Myc* animal-derived splenic B cells. Their Myc ChIP-seq signals at TSSs in pre-tumors of 3 *Eμ-Myc* animals from GSE51011 (samples GSM1234472, GSM1234473, and GSM1234474) are shown. The RNA-seq results of 4 different animals of each indicated genotype (with distinctive color labeling) were normalized to Spike-In RNA and shown after logarithmic transformation.

Table S6. Analyses of gene expression change by *Dpy30* heterozygosity.

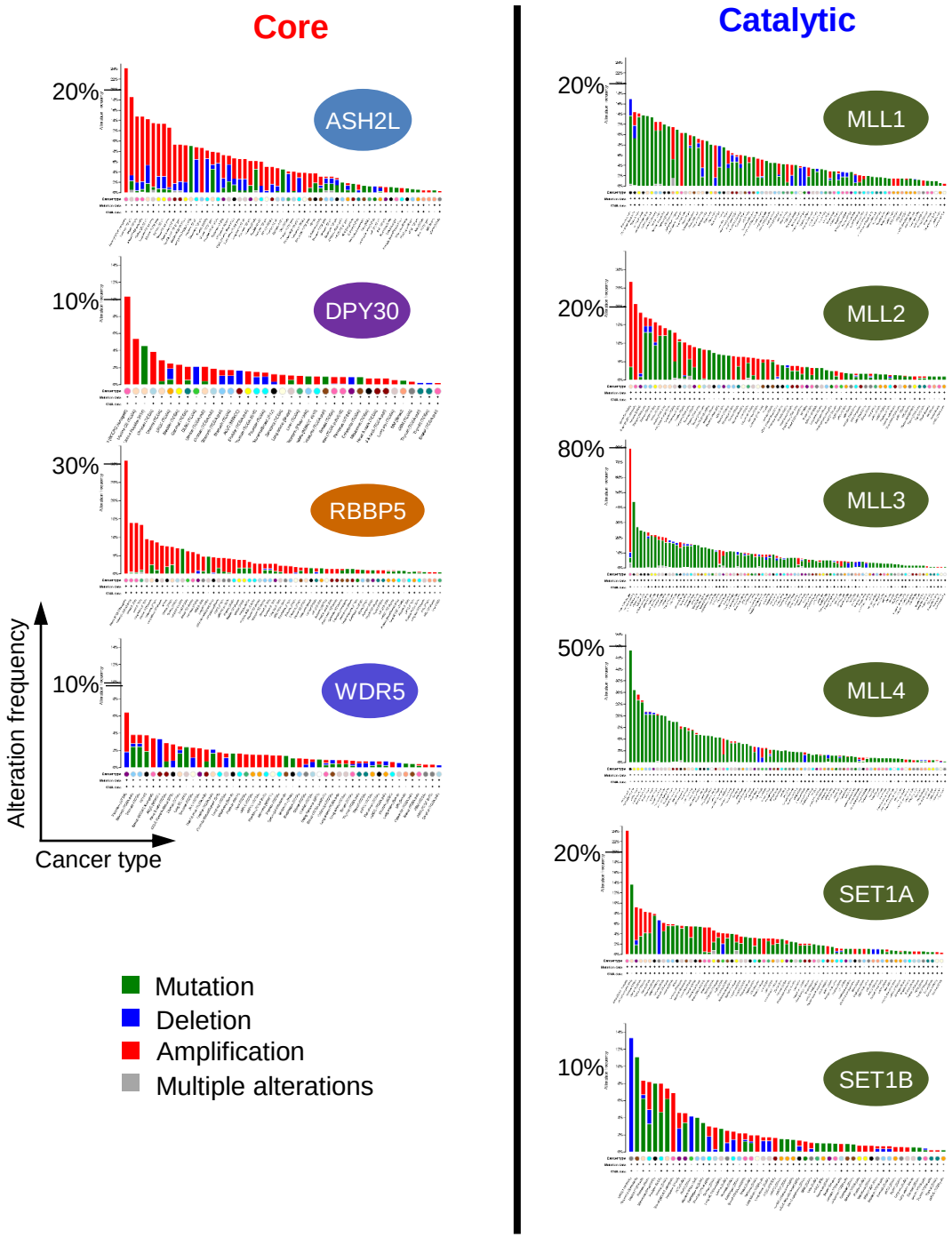
The tab "down in at least 3 litters" shows genes that were downregulated over 1.2 fold in *Eμ-Myc; Dpy30*$^{+/-}$ compared to *Eμ-Myc* B cells in 3 out of the 4 sequenced litters (each litter has one mouse for each genotype). The tab "up in at least 3 litters" shows genes that were upregulated over 1.2 fold in *Eμ-Myc; Dpy30*$^{+/-}$ compared to *Eμ-Myc* B cells in 3 out of the 4 sequenced litters (each litter has one mouse for each genotype). The tab "KEGG and DAVID GO" shows the results of KEGG and DAVID GO analyses of genes in the other tabs.

Table S7. Primers used.

SUPPLEMENTAL REFERENCES

1. Yang Z, Shah K, Khodadadi-Jamayran A, and Jiang H. Dpy30 is critical for maintaining the identity and function of adult hematopoietic stem cells. *The Journal of experimental medicine.* 2016;213(11):2349-64.
2. Jiang H, Shukla A, Wang X, Chen WY, Bernstein BE, and Roeder RG. Role for Dpy-30 in ES Cell-Fate Specification by Regulation of H3K4 Methylation within Bivalent Domains. *Cell.* 2011;144(4):513-25.
3. Buenrostro JD, Wu B, Chang HY, and Greenleaf WJ. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Current protocols in molecular biology / edited by Frederick M Ausubel [et al].* 2015;109:21.9.1-9.
4. John S, Sabo PJ, Canfield TK, Lee K, Vong S, Weaver M, et al. Genome-scale mapping of DNase I hypersensitivity. *Current protocols in molecular biology / edited by Frederick M Ausubel [et al].* 2013;Chapter 27:Unit 21.7.
5. Anders S, Pyl PT, and Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics (Oxford, England).* 2015;31(2):166-9.
6. Anders S, and Huber W. Differential expression analysis for sequence count data. *Genome biology.* 2010;11(10):R106.
7. Yu G, Wang LG, Han Y, and He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics : a journal of integrative biology.* 2012;16(5):284-7.
8. Sabo A, Kress TR, Pelizzola M, de Pretis S, Gorski MM, Tesi A, et al. Selective transcriptional regulation by Myc in cellular growth control and lymphomagenesis. *Nature.* 2014;511(7510):488-92.
9. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England).* 2013;29(1):15-21.
10. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25(16):2078-9.
11. Quinlan AR, and Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England).* 2010;26(6):841-2.
12. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic acids research.* 2016;44(W1):W160-5.
13. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nature biotechnology.* 2011;29(1):24-6.

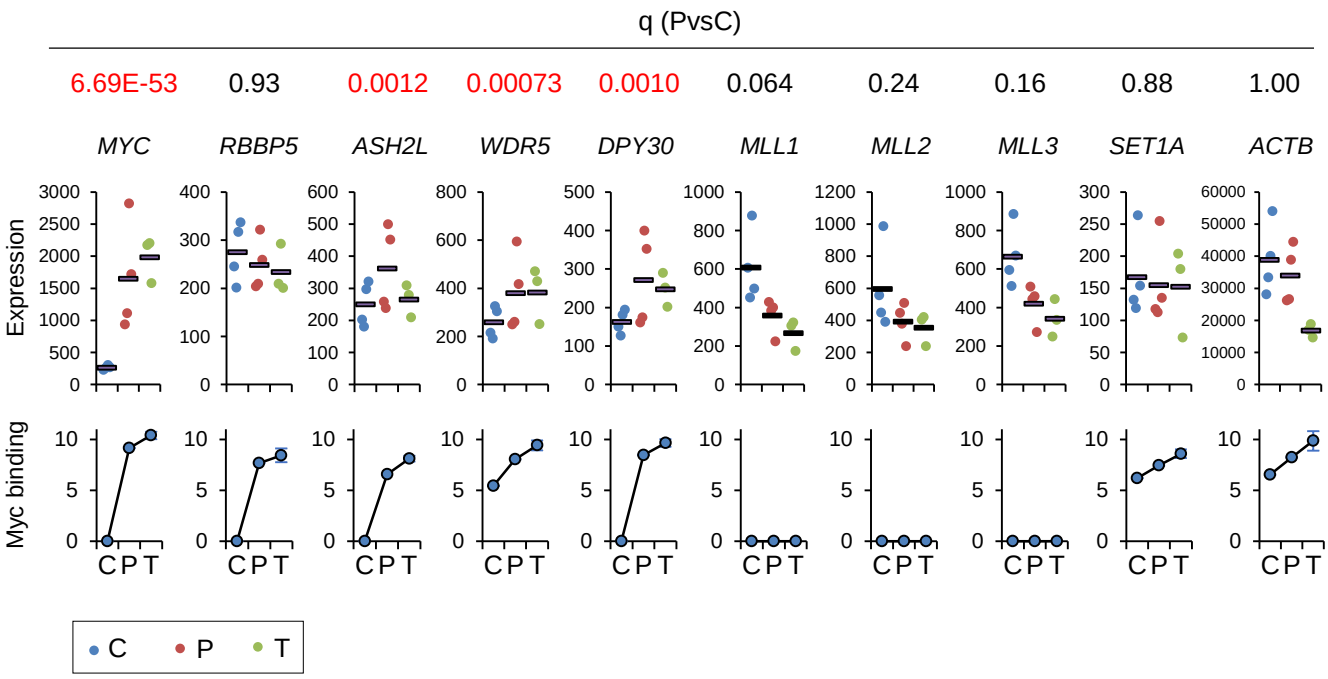
Supplemental Figure 1



Supplemental Figure 1. Divergent alterations of the core versus catalytic subunits of the SET1/MLL complexes in human cancers.

Data were generated from cBioPortal (http://www.cbioportal.org/).

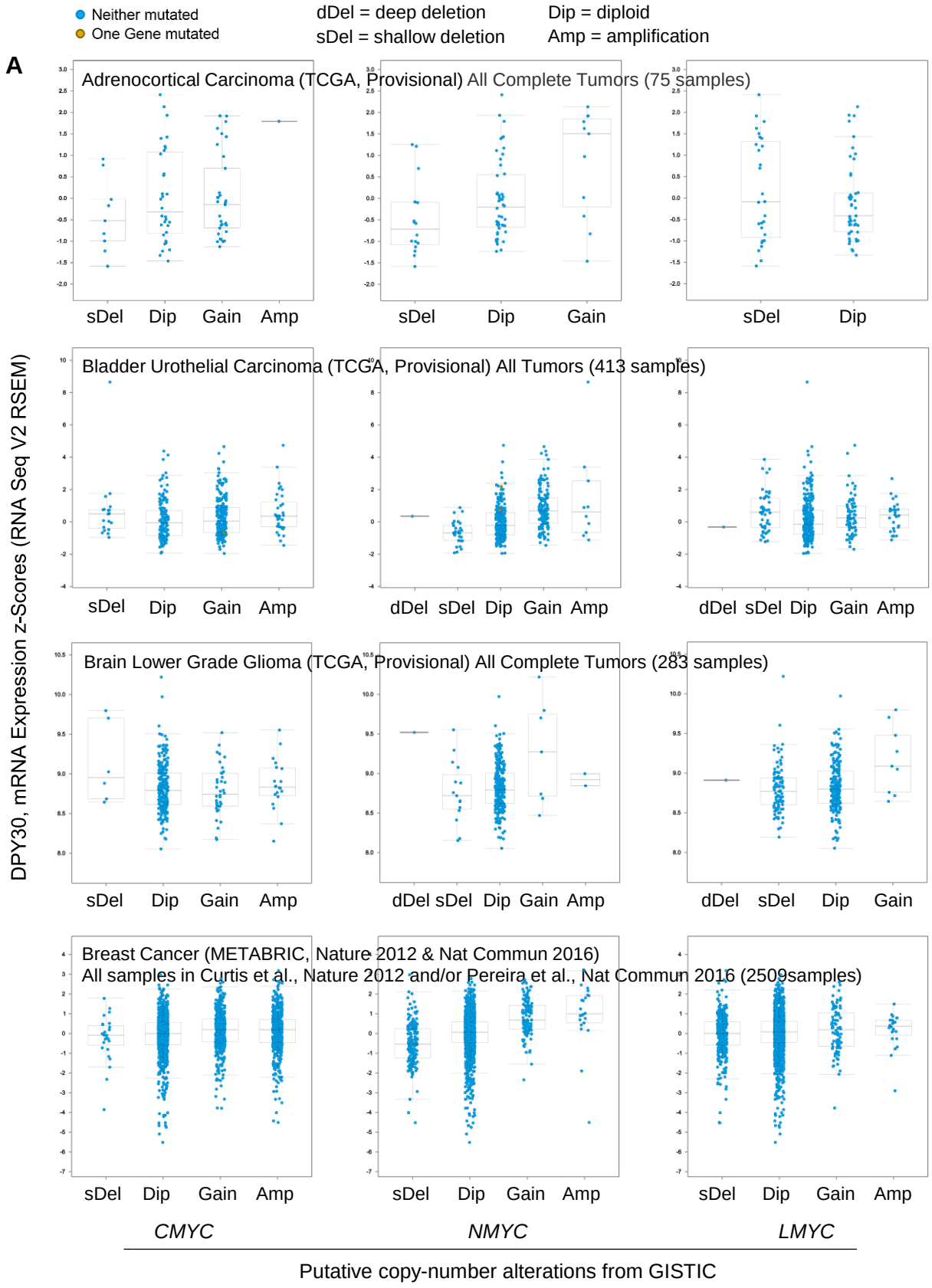
Supplemental Figure 2



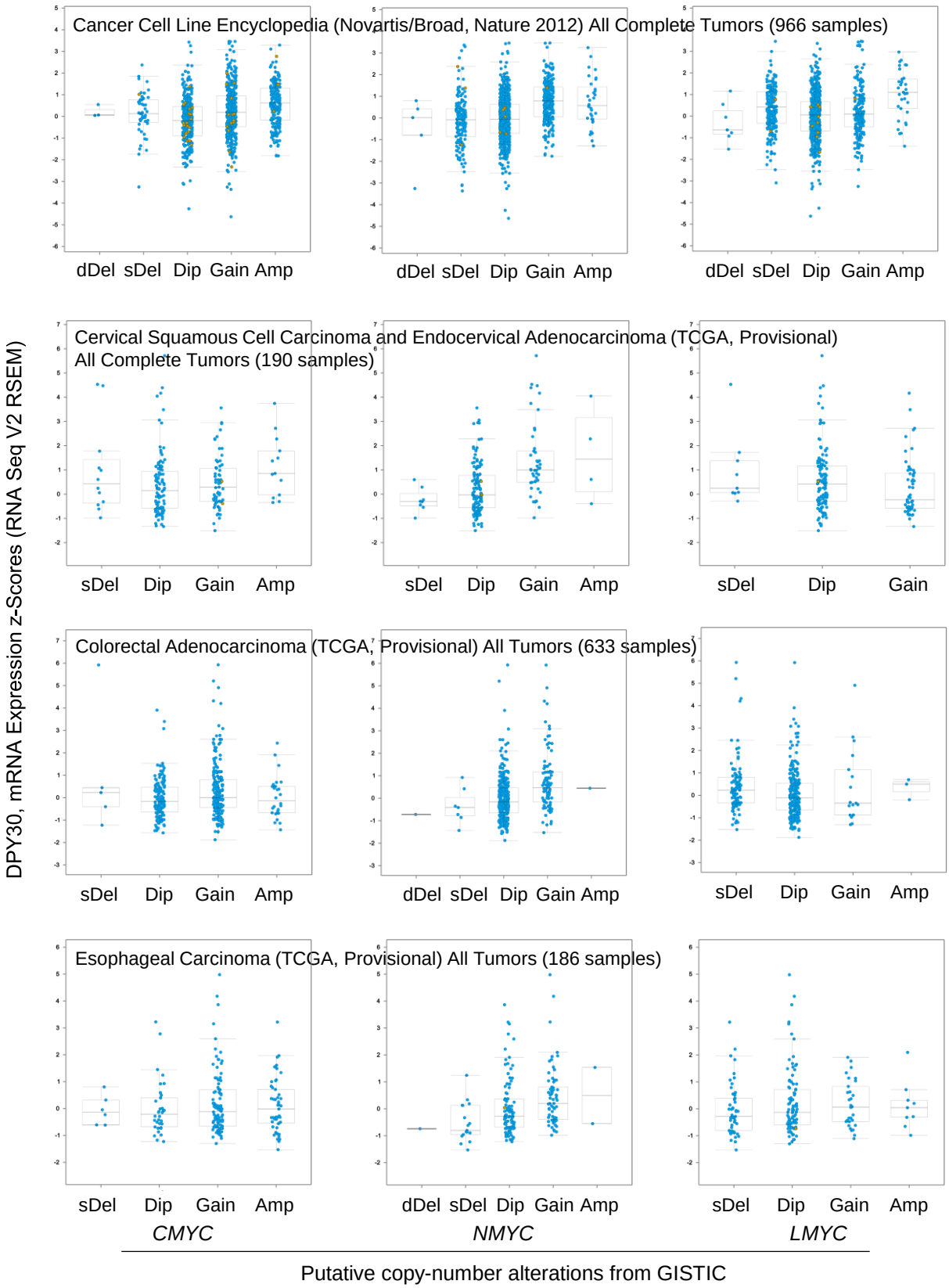
Supplemental Figure 2. MYC directly promotes the expression of the core subunits of SET1/MLL complexes.

Curated from GSE51011. Data on SET1B and MLL4 are not available. Top panels, expression levels of indicated genes from 4 control (C), 3 pre-tumor (P), and 3 tumor (T) samples normalized by mean expression level of all genes in each sample. False Discovery Rates (q value) between P vs. C are shown at the top, those < 0.05 in red. Bottom panels, MYC ChIP signal peak enrichment values at the promoters of indicated genes from one control, one pre-tumor, and 3 tumor samples shown as mean ± SD.

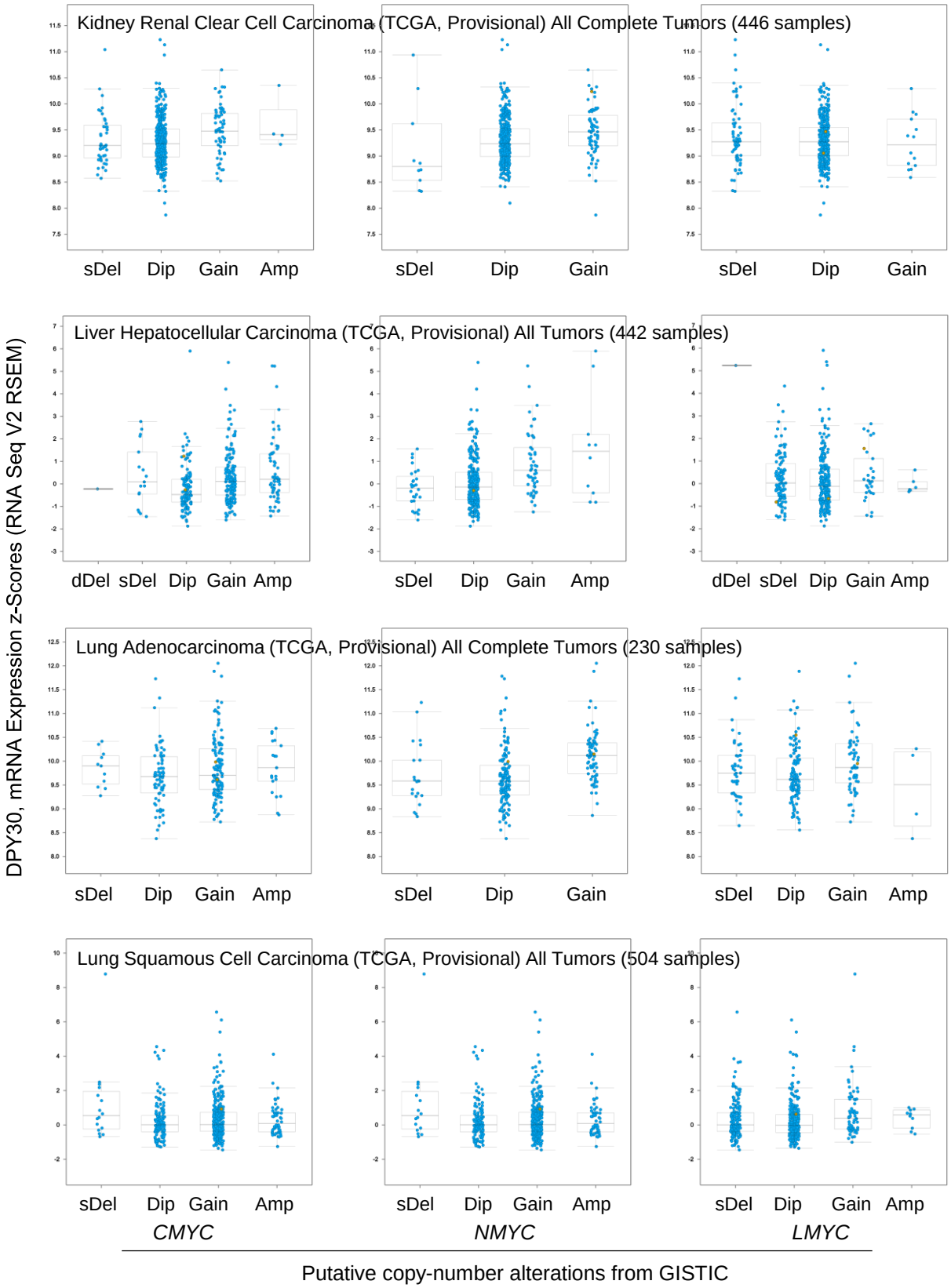
Supplemental Figure 3



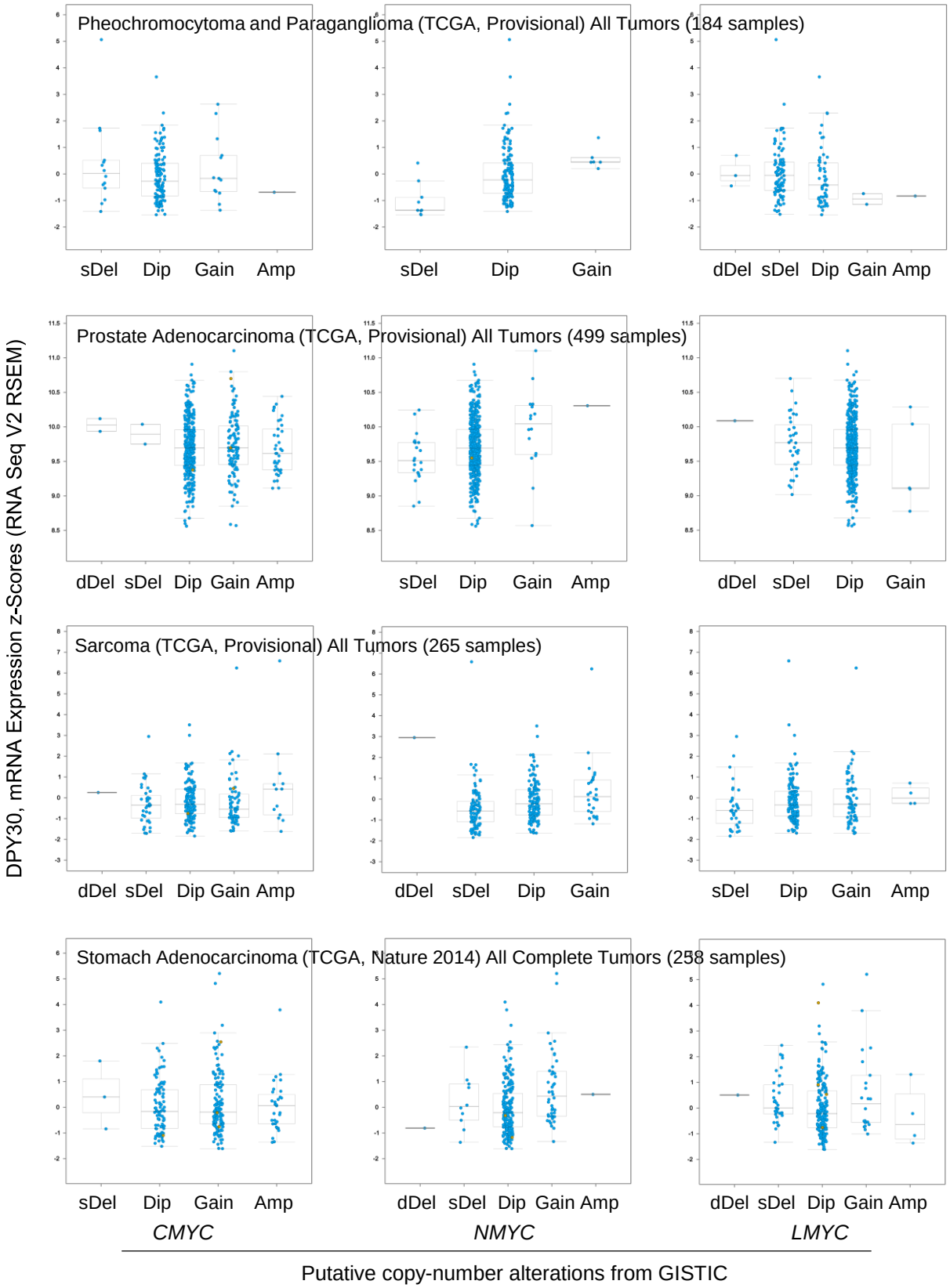
Supplemental Figure 3, continued 1



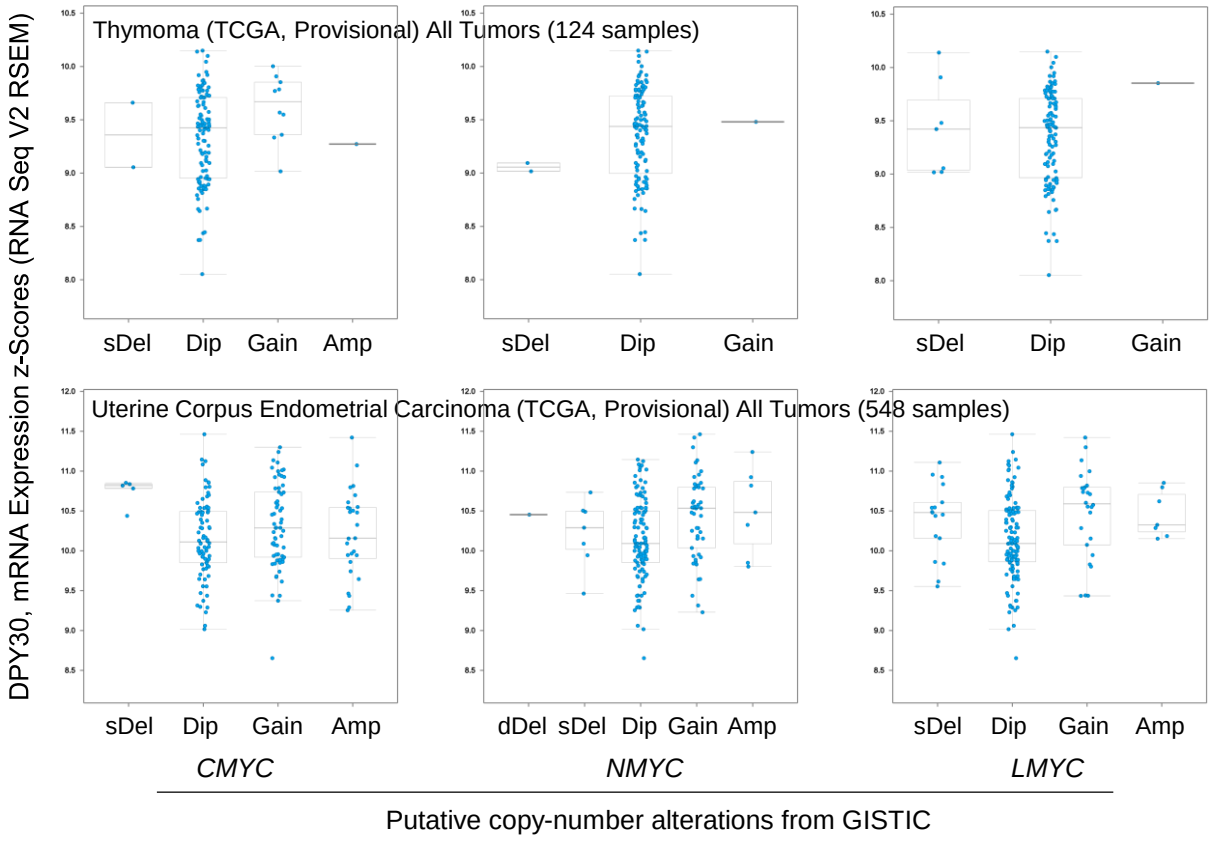
Supplemental Figure 3, continued 2



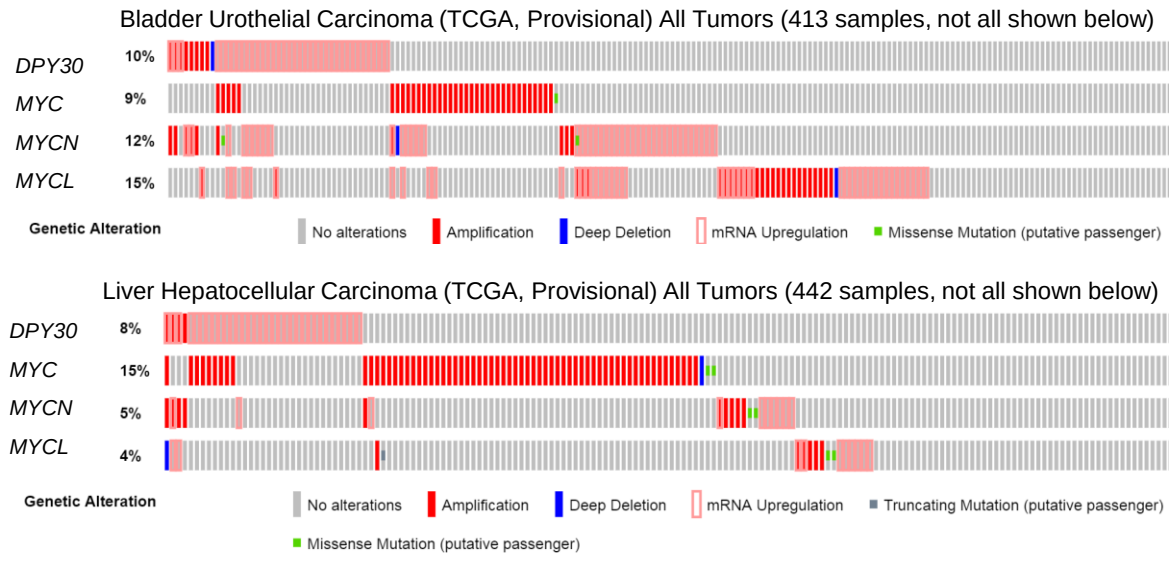
Supplemental Figure 3, continued 3



Supplemental Figure 3, continued 4

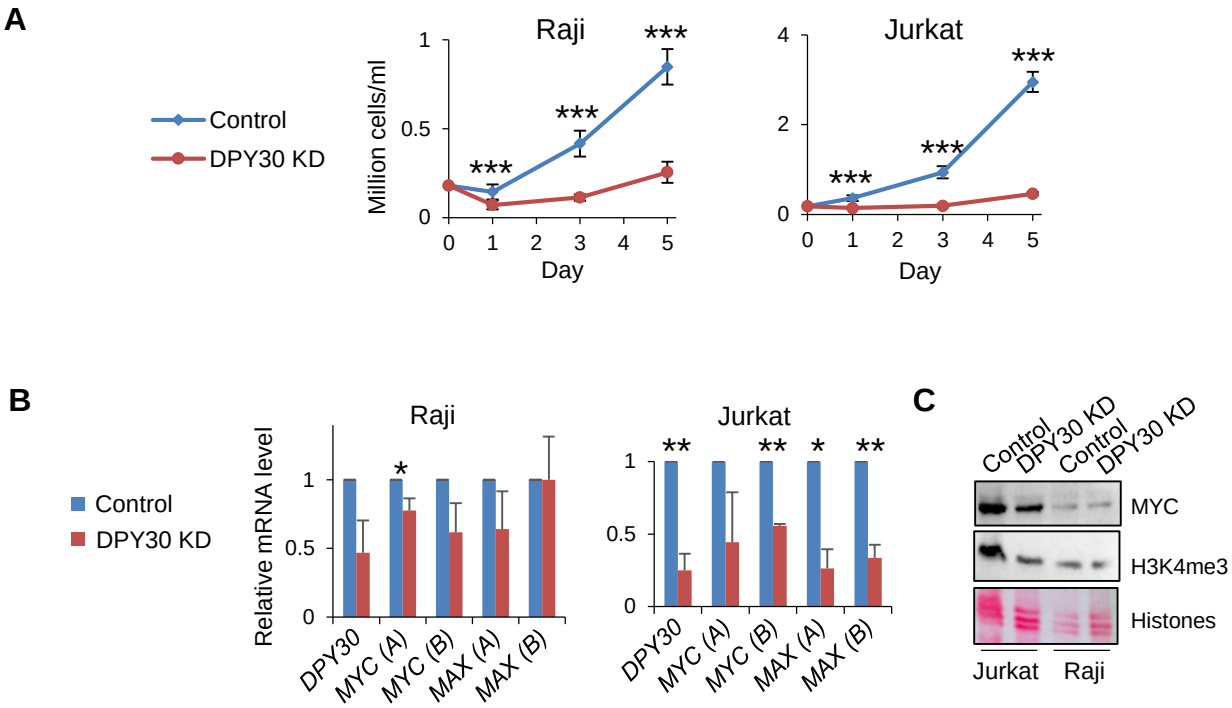


B



Supplemental Figure 3. Association of *MYC* gene amplification with *DPY30* expression in cancers.
(A) Plots of *DPY30* mRNA levels against copy-number alterations of *CMYC*, *NMYC*, and *LMYC* in a total of 18 cancer studies from cBioPortal that show notable elevation of *DPY30* expression with copy number increase (gain or Amplification) of one of more of the *MYC* genes. In all 169 studies in cBioPortal, 38 have over 50 samples and both copy number and mRNA expression results available. Note that increase of *Dpy30* level is associated with *CMYC* copy number increase in 7 studies, *NMYC* in 14, and *LMYC* in 3 studies.
(B) Oncoprint showing *DPY30* and *MYC* gene alterations in two different cancers. *MYC* = *CMYC*. Each small box represents a patient sample. All data in this figure are outputs from the cBioPortal.

Supplemental Figure 4



Supplemental Figure 4. DPY30 regulates expression of *MYC* in two blood cancer cell lines.

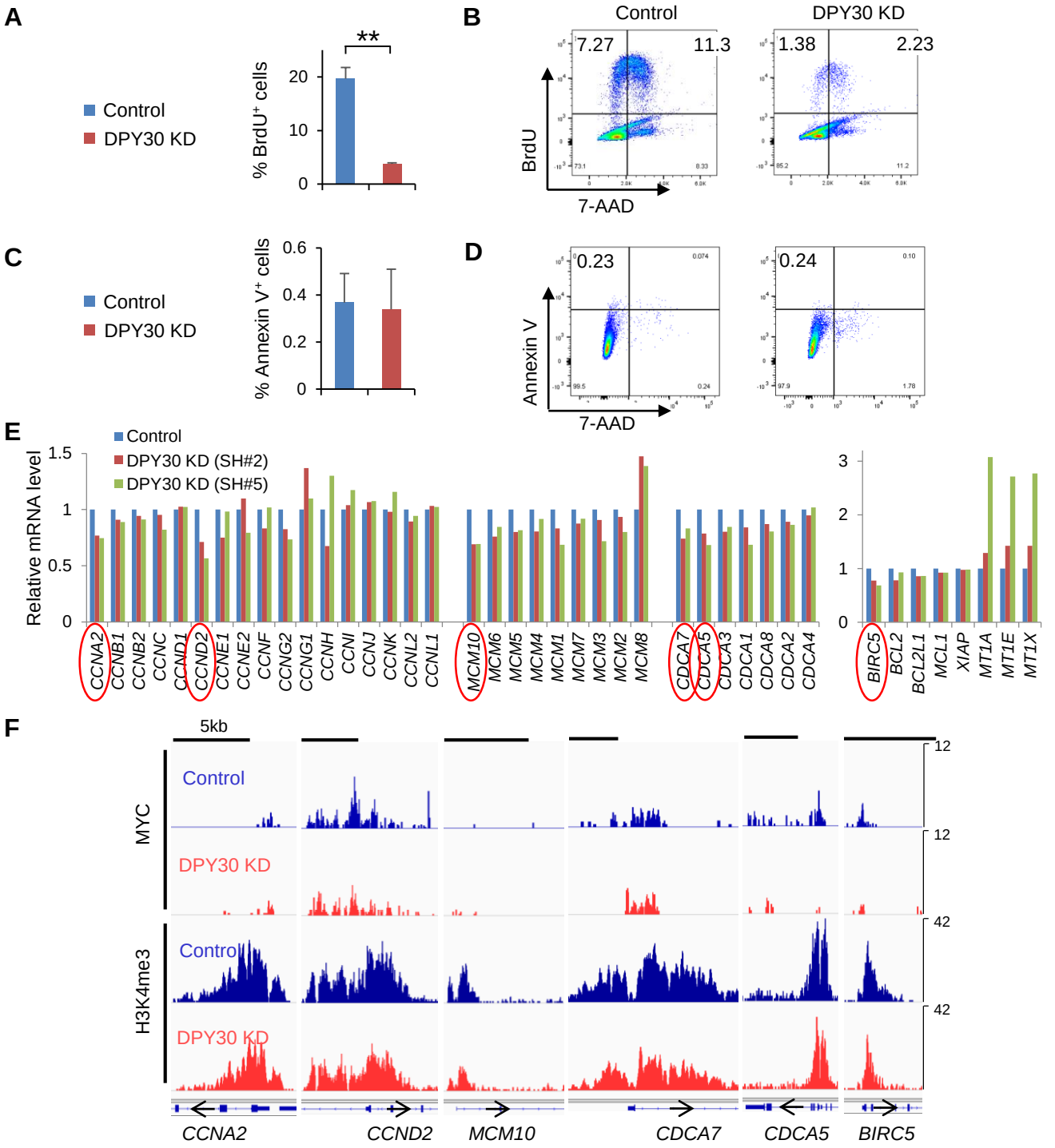
(A) Growth of Raji and Jurkat cells following DPY30 KD. Cell numbers were counted and mean ± SD from 3 independent KD assays are plotted.

(B) *DPY30*, *MYC,* and *MAX* mRNA levels were determined by qPCR using two different primers for *MYC* and *MAX* and normalized to *ACTB*, shown as mean + SD from 3 independent KD assays in Raji and Jurkat cells.

(C) Total cell lysates from control and DPY30 KD Raji and Jurkat cells were used for immunoblotting by indicated antibodies and for Ponceau S staining of histones.

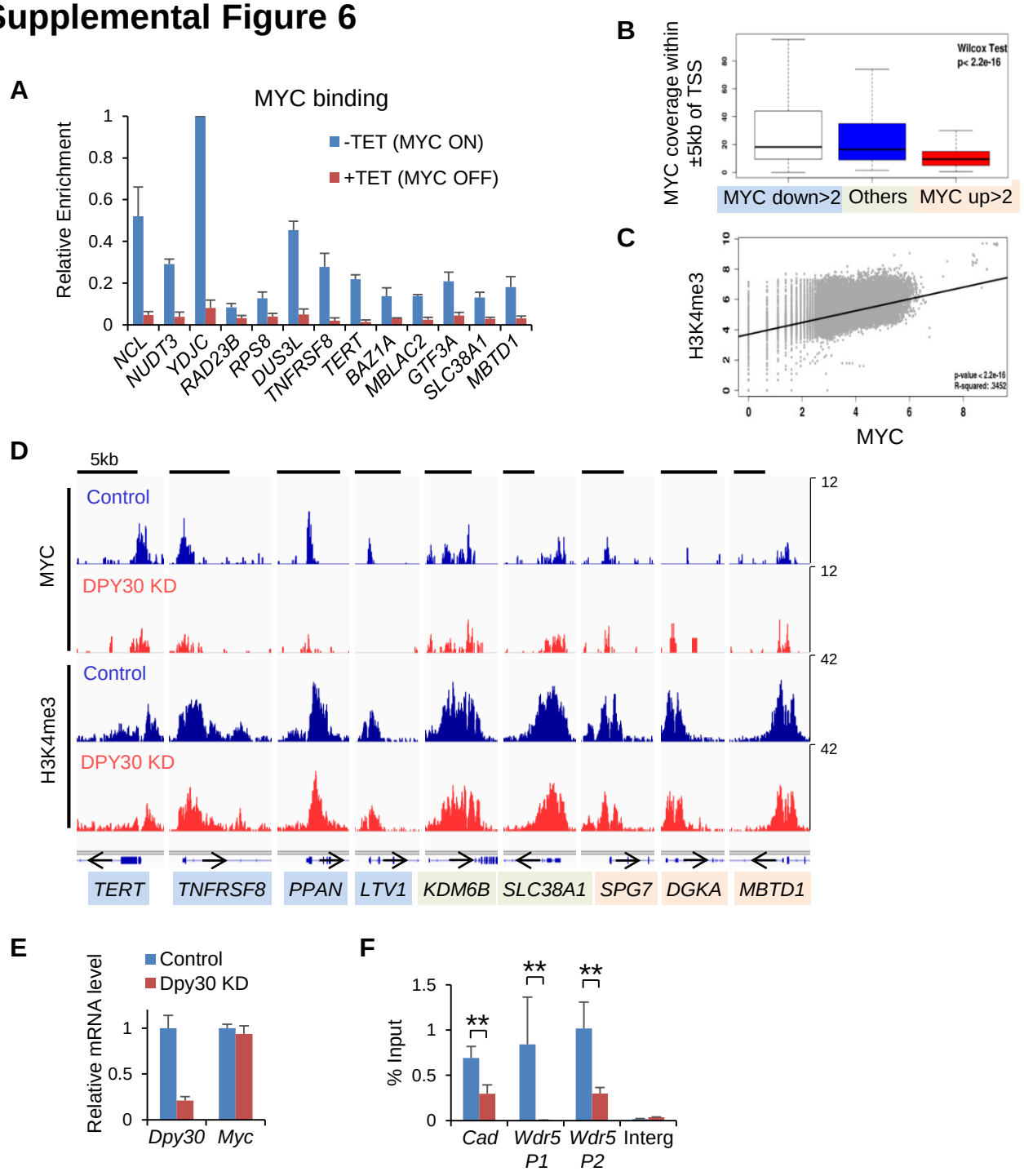
Data represent the mean ± SD (A) or + SD (B). **P* < 0.05, ***P* < 0.01, and ****P* < 0.001 by Student's *t*-test.

Supplemental Figure 5



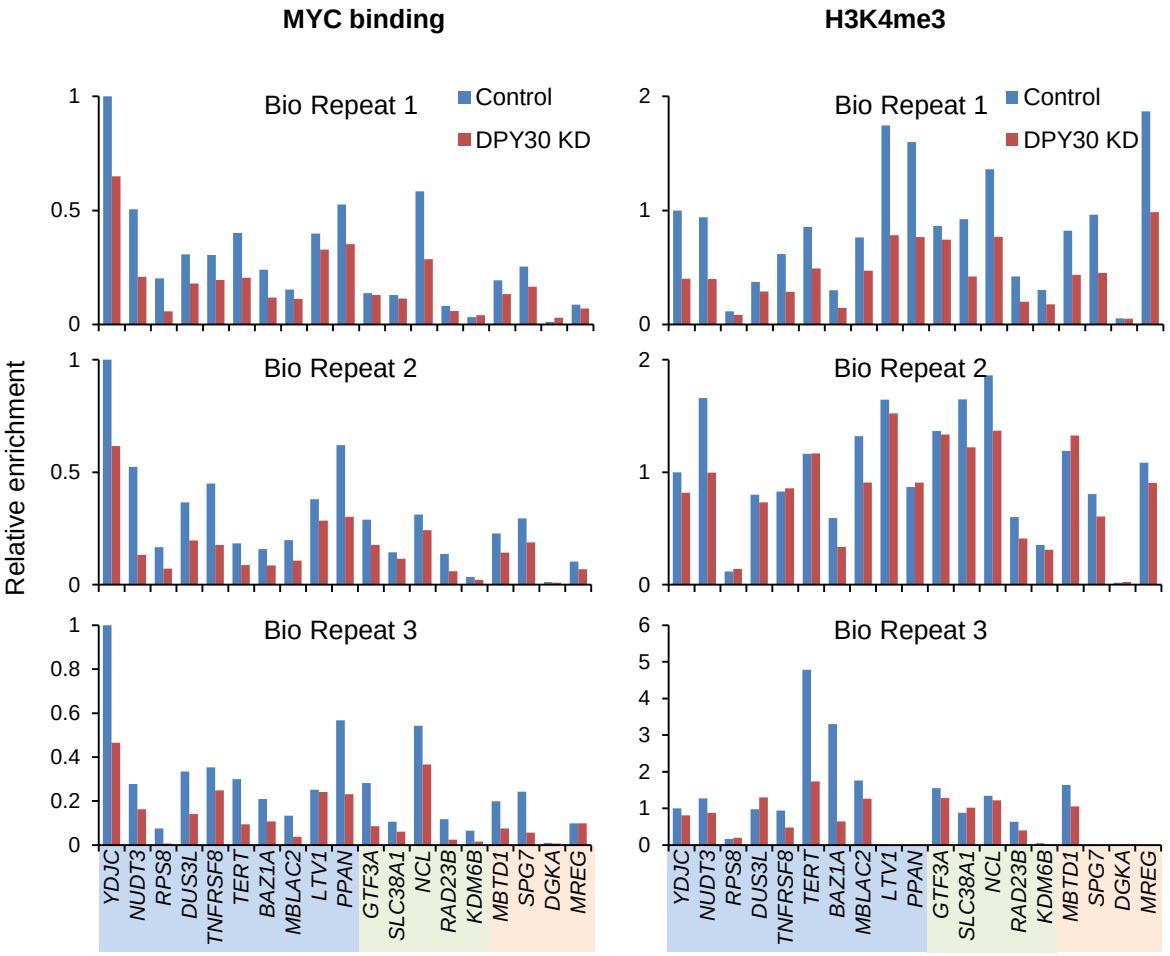
Supplemental Figure 5. Effects of DPY30 KD on proliferation and apoptosis of P493-6 cells.
All assays in this figure used control and DPY30 KD P493-6 cells. DPY30 KD was by DPY30 shRNA #2 except that both shRNA #2 and #5 were used in (E).
(A) BrdU$^+$ percentage in proliferation assays. A and C were from 5 repeats.
(B) Representative flow cytometry analysis results for proliferation assays.
(C) Annexin V$^+$ percentage in apoptosis assays.
(D) Representative flow cytometry analysis results for apoptosis assays.
(E) Relative mRNA levels of cell cycle regulator and pro-survival genes were analyzed by microarray assay. Genes in red circle were consistently downregulated by both DPY30 shRNAs.
(F) MYC binding and H3K4me3 profiles for circled genes in (E).
Data represent the mean + SD (A and C). $^{**}P < 0.01$ by Student's *t*-test.

Supplemental Figure 6



Supplemental Figure 6. DPY30 is important for efficient binding of MYC to its genomic targets.
(A) MYC ChIP-qPCR at indicated TSSs in P493-6 cells with or without tetracycline (Tet) treatment from 3 independent Tet treatments. Significant (P<0.01) for all genes.
(B) Box-and-Whisker plot of the MYC ChIP coverages in the control cells for the three groups of genes in Figure 3C. The box represents the 25th, 50th, and 75th percentiles of the data, and the whiskers represent the 5th and 95th percentiles. P value (=2.2e-16) for the last two gene groups was shown as calculated by Wilcoxon signed-rank test.
(C) Correlation of genome-wide MYC binding and H3K4me3, from ChIP results of control cells. R^2=0.3452.
(D) Representative MYC binding and H3K4me3 profiles for genes within each group shown in Figure 3C.
(E) Relative mRNA levels of Dpy30 and Myc in MEFs expressing scramble (control) or Dpy30 shRNA (KD), normalized to *Actb* from triplicate measurements.
(F) Myc ChIP results of at indicated loci from triplicate measurements in MEFs in (E).
Data represent the mean + SD (A, E, and F). ***P* < 0.01 by Student's *t*-test.

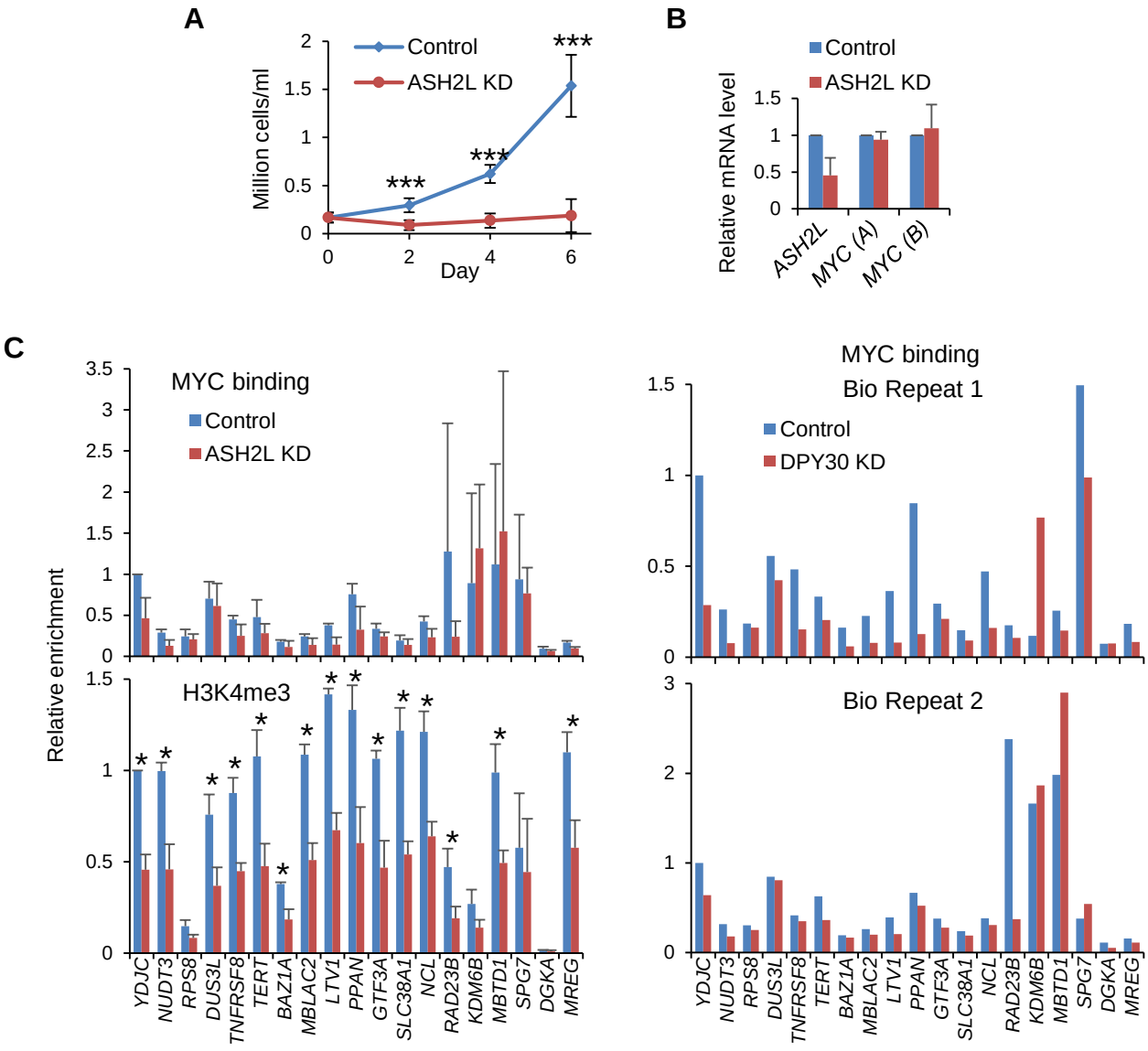
Supplemental Figure 7



Supplemental Figure 7. Effects of DPY30 KD on MYC binding and H3K4me3 at selected genes in P493-6 cells.

MYC and H3K4me3 ChIP at TSSs by qPCR, calculated from the ratio of the percent input value for each locus over that for the *YDJC* site in the control sample in 3 individual biological repeats. Each repeat is an independent KD followed by ChIP assay in P493-6 cells.

Supplemental Figure 8



Supplemental Figure 8. Effects of ASH2L KD on P493-6 cells.

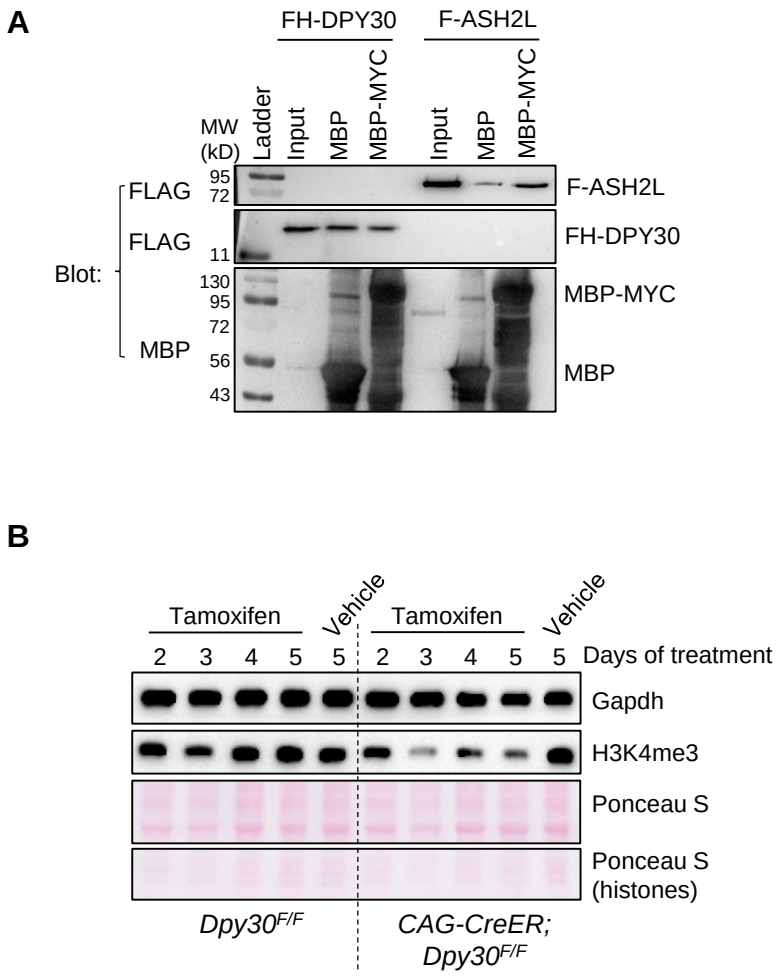
(A) Growth of P493-6 cells following ASH2L KD. Cell numbers counted from 4 independent KD assays are plotted.

(B) *ASH2L*, *MYC,* and *MAX* mRNA levels were determined by qPCR using two different primers for *MYC* and *MAX* and normalized to *GAPDH* from 4 independent KD assays.

(C) The same genes shown for DPY30 KD in Figure 3E were validated for MYC and H3K4me3 ChIP at TSSs by qPCR, calculated from the ratio of the percent input value for each locus over that for the *YDJC* site in the control sample. Shown on the left are from 2 (MYC) or 3 (H3K4me3) independent KD and ChIP assays in P493-6 cells, and shown on the right are the individual 2 independent KD and MYC-ChIP assays.

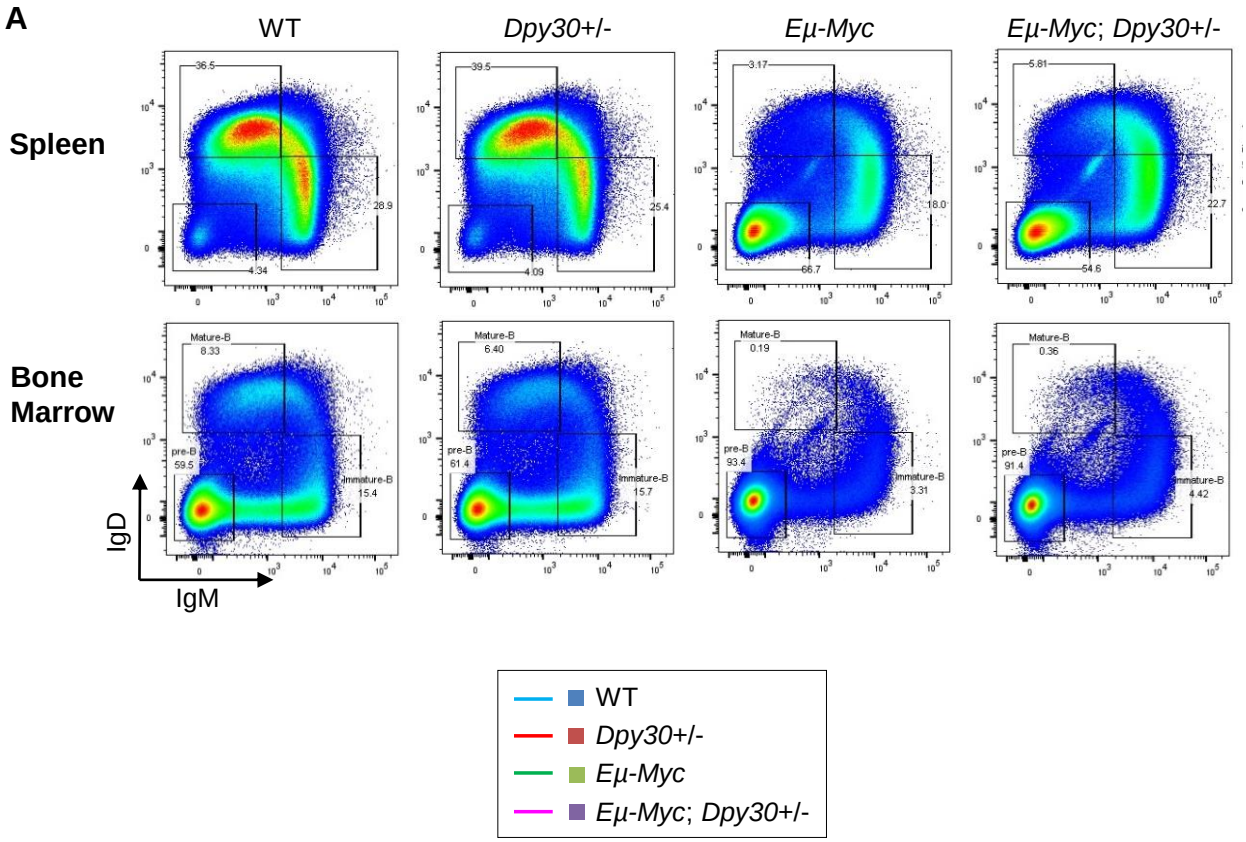
Data represent the mean ± SD (A) or + SD (B and C). **P* < 0.05, and ****P* < 0.001 by Student's *t*-test.

Supplemental Figure 9



Supplemental Figure 9. On mechanisms by which DPY30 regulates genomic binding of MYC.
(A) Binding assay using all purified proteins. MBP or MBP-MYC protein immobilized on amylose resins were incubated with FLAG-HA-tagged DPY30 or FLAG-tagged ASH2L proteins followed by immunoblotting. Specific binding of ASH2L or DPY30 with MYC is shown by the relative FLAG tag signals bound to MBP-MYC versus MBP.
(B) Primary MEFs derived from *Dpy30$^{F/F}$* (control) or *CAG-CreER; Dpy30$^{F/F}$* (KO) mice were treated with 4-hydroxytamoxifen or vehicle (ethanol) for indicated days followed by immunoblotting.

Supplemental Figure 10



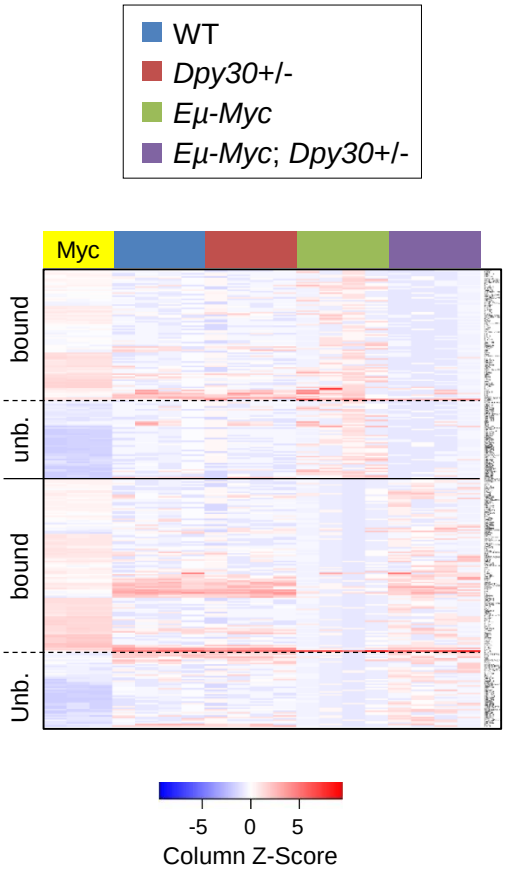
Supplemental Figure 10. Effects of *Dpy30* heterozygosity on B cell development and cellular damage response, with or without Myc hyperactivation.

(A) Representative FACS analysis of cell surface IgD and IgM in B220^{+} bone marrow and spleen cells of indicated genotypes.

(B) Immunoblotting for indicated proteins in B220^{+} cells from individual mice of indicated genotypes from two different litters.

(C) Intracellular ROS levels were determined for B220^{+} cells from different genotypes. A representative FACS analysis from one litter is shown on left, and results from multiple littermates are plotted on right as mean + SD. WT, n = 2; *Dpy30*$^{+/-}$, n = 2; *Eμ-Myc*, n = 4; and *Eμ-Myc; Dpy30*$^{+/-}$, n = 3.

Supplemental Figure 11



Supplemental Figure 11. Impact of *Dpy30* heterozygosity on gene expression in splenic B cells.

Heatmap showing relative expression levels of 305 genes that were down- or up-regulated (divided by the solid horizontal line) over 4 fold in the average of 4 (in each column) *Eμ-Myc*; *Dpy30*$^{+/-}$ compared to *Eμ-Myc* animal-derived splenic B cells. Genes are further clustered into Myc-bound and unbound (divided by the dash lines) based on Myc ChIP-seq signals (Myc, highlighted in yellow) at TSSs in pre-tumors of 3 *Eμ-Myc* animals from GSE51011 (samples GSM1234472, GSM1234473, and GSM1234474). The RNA-seq results of 4 different animals of each indicated genotype were normalized to Spike-In RNA and reflect per cell expression. More information of the clustered genes is in Supplemental Table 5.

Supplemental Figure 12

A

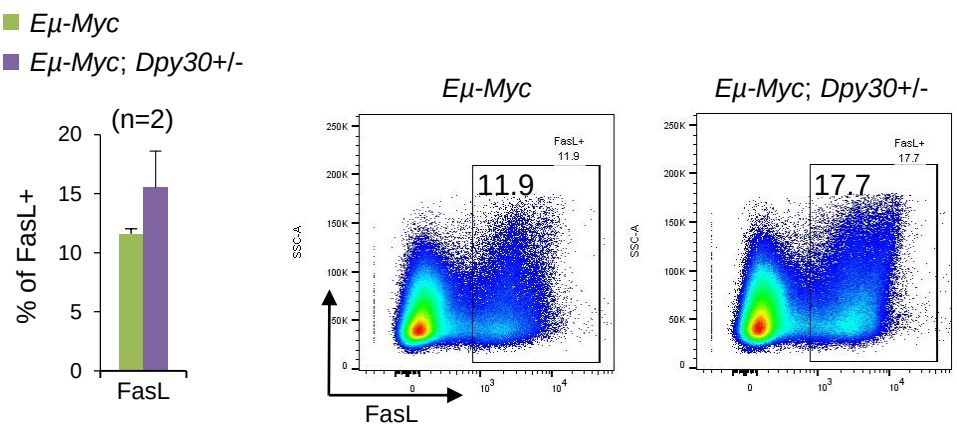
Enrichment of genes **downregulated** in *Eμ-Myc; Dpy30+/-* B cells:

KEGG Pathway Term	q value	Gene
Ribosome biogenesis in eukaryotes	0.0022	Nhp2l1/Gnl3/Nxf1/Nxt1/Nop10/Nop56/Nob1/Mphosph10/Fcf1/Rpp25/Taf9/Wdr3
Spliceosome	0.02	Srsf10/Hspa2/Srsf3/Snrpd1/Nhp2l1/Sf3b6/Magohb/Ncbp2/Ppil1/Prpf3/Snrpd2/Srsf7
RNA transport	0.026	Eif1a/Eif2s1/Eif4e/Sap18/Smn1/Tacc3/Nxf1/Nxt1/Magohb/Ncbp2/Eif3j1/Rpp25/Nup54

Enrichment of genes **upregulated** in *Eμ-Myc; Dpy30+/-* B cells:

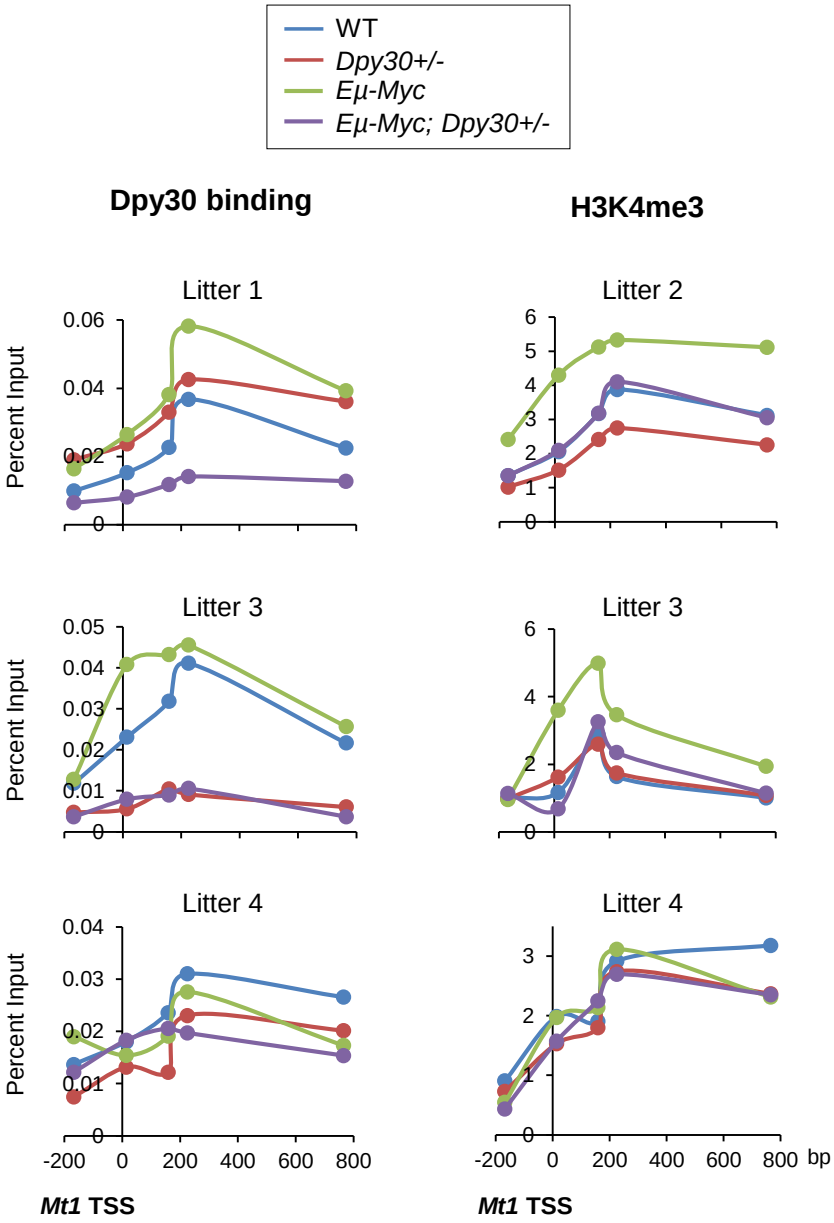
KEGG Pathway Term	q value	Gene
Cytokine-cytokine receptor interaction	4.78E-14	Ccr6/Ccr1/Ccr9/Ccr2/Ccr5/Ccr10/Ackr3/Csf2rb/Csf2rb2/Csf3r/Cx3cr1/Egf/Egfr/Fasl/Flt1/Cxcl10/Il18rap/Il18r1/Il2rb/Il9r/Kdr/Ltb/Pdgfa/Ccl5/Tgfb3/Tnf/Tnfrsf17/Tnfrsf18/Cd27/Xcr1/Ifnlr1
Hematopoietic cell lineage	1.58E-11	Cd3d/Cd3e/Cd3g/Cd4/Cd5/Cd7/Cd8a/Cd8b1/Cd9/Cr2/Csf3r/Fcer2a/Il9r/Itga2/Itgb3/Anpep/Tnf
Natural killer cell mediated cytotoxicity	6.13E-08	Cd247/Fasl/Gzmb/Klra3/Klra4/Klra7/Klra8/Klrc1/Klrd1/Klrb1c/Ncr1/Nfatc2/Prf1/Tnf/Hcst/Klrk1
Antigen processing and presentation	8.91E-05	Cd4/Cd8a/Cd8b1/H2-Q7/H2-Q8/Klrc1/Klrd1/Tnf/H2-Q6

B



Supplemental Figure 12. *Dpy30* heterozygosity suppresses Myc-driven lymphomagenesis.
(A) KEGG analyses for genes that were down- or up-regulated over 1.2 fold in at least 3 out of 4 *Eμ-Myc; Dpy30*$^{+/-}$ compared to littermate *Eμ-Myc* splenic B cell samples, based on the RNA-seq results. The complete results of KEGG analyses are in Supplemental Table 6.
(B) FACS analysis of cell surface expression of FasL from 2 animals of each indicated genotype as mean + data range. Representative plots are shown on the right.

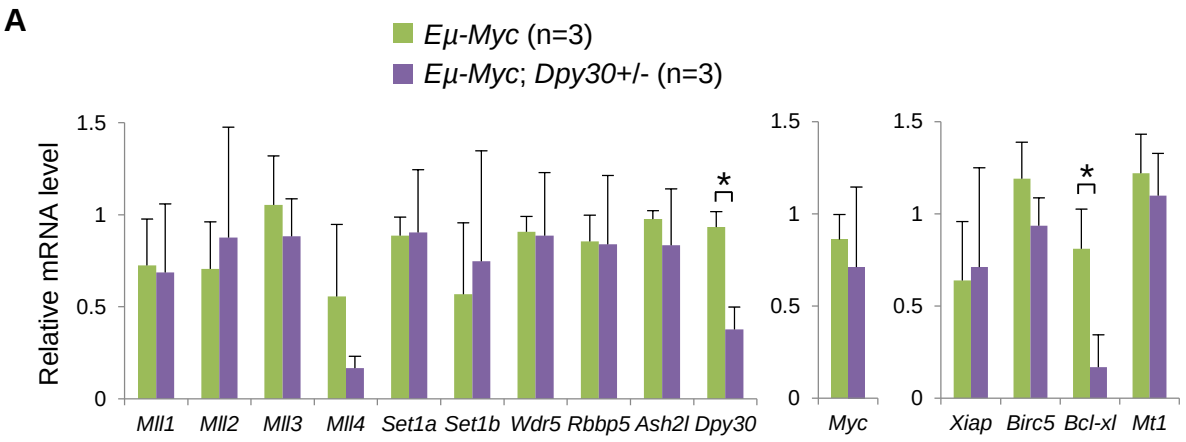
Supplemental Figure 13

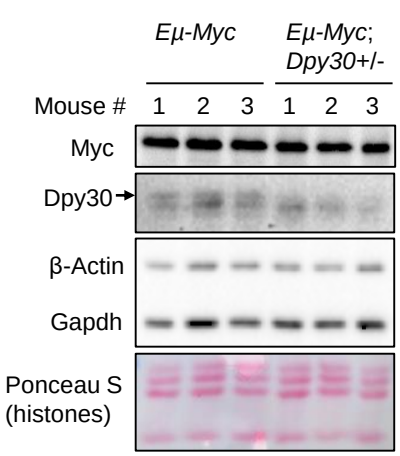


Supplemental Figure 13. Effects of *Dpy30* heterozygosity on Dpy30 binding and H3K4me3 at *Mt1*.

Dpy30 and H3K4me3 ChIP assays were performed using splenic B220+ cells from littermate mice of different genotypes, followed by qPCR using a series of primers around *Mt1* gene TSS. Each dot represents a primer pair with their relative distance to *Mt1* TSS shown on x-axis.

Supplemental Figure 14



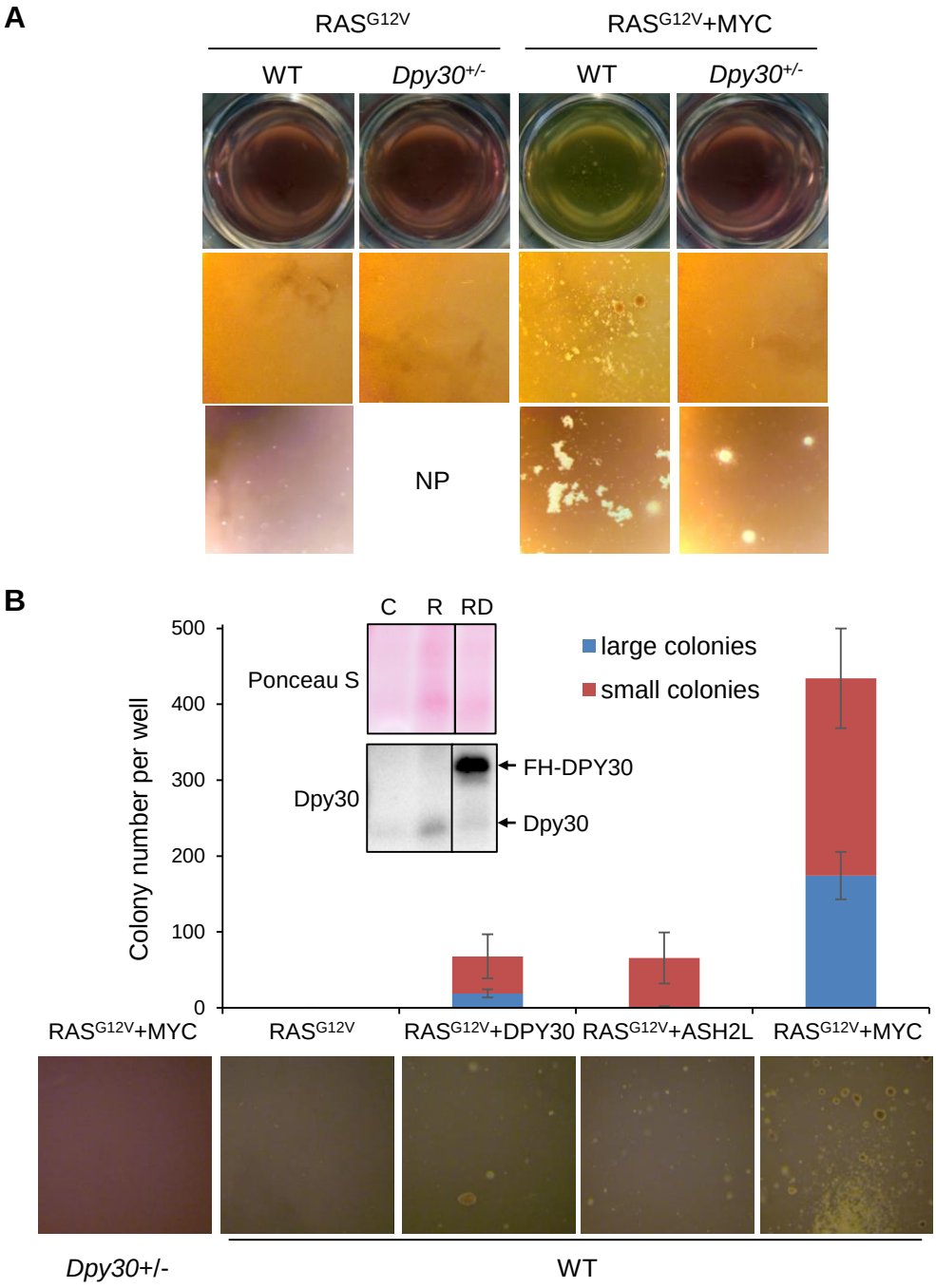


Supplemental Figure 14. Effects of *Dpy30* heterozygosity on gene expression in tumors that eventually form.

(A) Relative mRNA levels of genes encoding subunits of Set1/Mll complexes, *Myc*, and pro-survival genes from tumors that eventually formed in 3 *Eμ-Myc* mice and 3 *Eμ-Myc; Dpy30*$^{+/-}$ mice were determined by qPCR and normalized to *Actb*, shown as mean + SD. **P* < 0.05 by Student's *t*-test.

(B) Levels of indicated proteins in tumors that eventually formed in 3 *Eμ-Myc* mice and 3 *Eμ-Myc; Dpy30*$^{+/-}$ mice were determined by immunoblotting following resolution on SDS-PAGE gel. Note for the Dpy30 blot, the top band is specific for Dpy30 while the bottom band is a non-specific signal.

Supplemental Figure 15

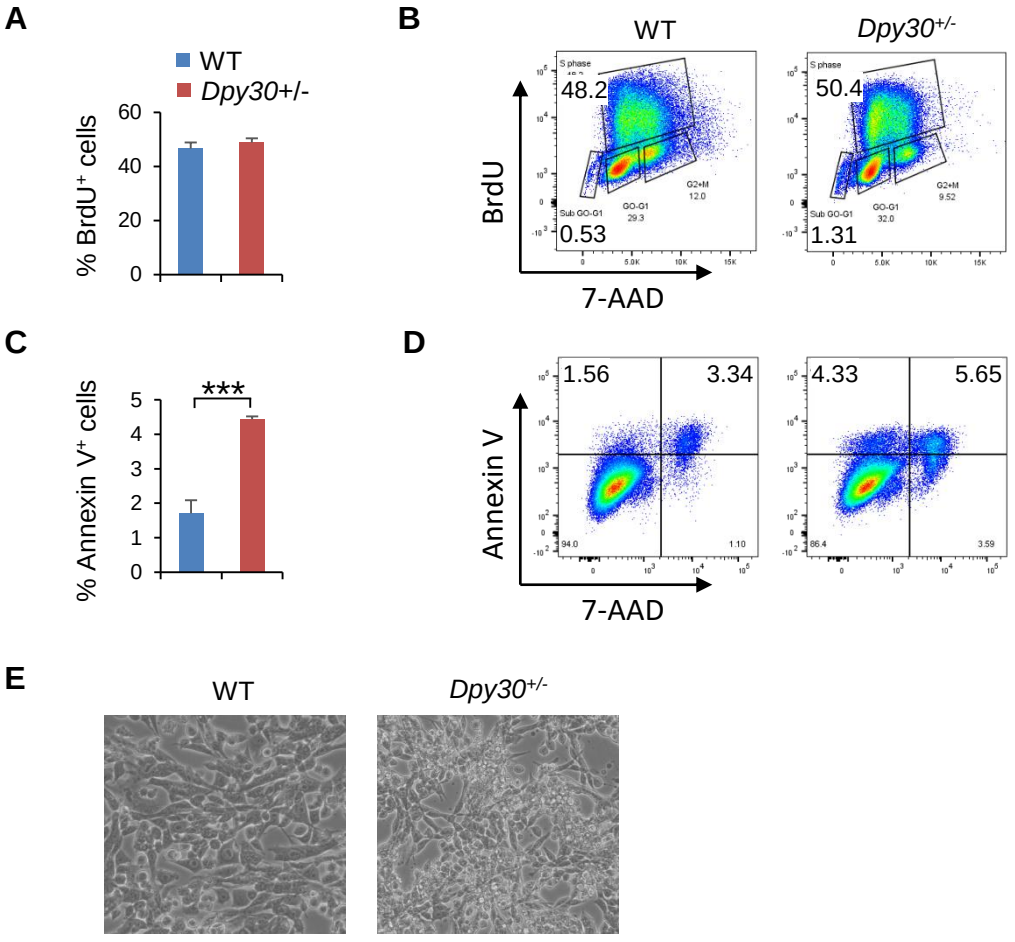


Supplemental Figure 15. Effects of *Dpy30* expression on cellular transformation.

(A) Representative image of wells and colonies from soft agar colony formation assay for $HRAS^{G12V}$ and MYC-mediated oncogenic transformation of WT and *Dpy30*$^{+/-}$ MEFs. NP, no pictures were taken.

(B) WT and *Dpy30*$^{+/-}$ MEFs (indicated at bottom) were transduced with combination of viruses as indicated (above images) and subjected to colony formation assay in soft agar. Colony numbers were plotted as mean ± SD from 6 independently seeded wells in one colony formation assays representative of two independent transduction assays, which showed consistent phenotype. large colonies are those larger than 0.2 mm in diameter. This was arbitrarily set as the cutoff. $P < 0.00001$ for the large colonies between $HRAS^{G12V}$+DPY30 (we used FLAG-HA-tagged DPY30 or FH-DPY30) and $HRAS^{G12V}$+ASH2L. Inserts: Ponceaus S staining and anti-Dpy30 immunoblotting for total lysates of WT MEFs untransduced (C) or transduced with $HRAS^{G12V}$ (R) or with $HRAS^{G12V}$+FH-DPY30 (RD). Overall effect on the colony numbers was analyzed separately for large and small colonies, using one-factor ANOVA with post hoc *t* test.

Supplemental Figure 16



Supplemental Figure 16. **Effects *Dpy30* heterozygosity and proliferation and death of MEFs.** All assays in this figure used WT and *Dpy30*$^{+/-}$ MEFs after transduction with HRAS^{v12} and MYC.

(A) BrdU^{+} percentage in proliferation assays. In A and C, Average ± SD from six repeats are plotted. *** P<0.001 by Student's t-test.

(B) Representative flow cytometry analysis results for proliferation assays.

(C) Annexin V^{+} percentage in apoptosis assays.

(D) Representative flow cytometry analysis results for apoptosis assays.

(E) Representative images of the cells in culture. Note the large number of detached and round-up *Dpy30*$^{+/-}$ MEFs after transduction with HRAS^{v12} and MYC.