RESEARCH ARTICLE

TOX Regulates Growth, DNA Repair, and Genomic Instability in T-cell Acute Lymphoblastic Leukemia

Riadh Lobbardi[1,2,3], Jordan Pinder[4,5], Barbara Martinez-Pastor[2], Marina Theodorou[1,2,3], Jessica S. Blackburn[6], Brian J. Abraham[7], Yuka Namiki[8], Marc Mansour[9,10], Nouran S. Abdelfattah[1,2], Aleksey Molodtsov[1,2], Gabriela Alexe[9,11], Debra Toiber[2,12], Manon de Waard[13], Esha Jain[1,14], Myriam Boukhali[2], Mattia Lion[8], Deepak Bhere[15], Khalid Shah[15], Alejandro Gutierrez[16], Kimberly Stegmaier[9,11,16], Lewis B. Silverman[9,16], Ruslan I. Sadreyev[1,14], John M. Asara[17], Marjorie A. Oettinger[8], Wilhelm Haas[2], A. Thomas Look[9], Richard A. Young[7], Raul Mostoslavsky[2,14], Graham Dellaire[4,5], and David M. Langenau[1,2,3,14]

ABSTRACT T-cell acute lymphoblastic leukemia (T-ALL) is an aggressive malignancy of thymocytes. Using a transgenic screen in zebrafish, thymocyte selection–associated high mobility group box protein (TOX) was uncovered as a collaborating oncogenic driver that accelerated T-ALL onset by expanding the initiating pool of transformed clones and elevating genomic instability. TOX is highly expressed in a majority of human T-ALL and is required for proliferation and continued xenograft growth in mice. Using a wide array of functional analyses, we uncovered that TOX binds directly to KU70/80 and suppresses recruitment of this complex to DNA breaks to inhibit nonhomologous end joining (NHEJ) repair. Impaired NHEJ is well known to cause genomic instability, including development of T-cell malignancies in KU70- and KU80-deficient mice. Collectively, our work has uncovered important roles for TOX in regulating NHEJ by elevating genomic instability during leukemia initiation and sustaining leukemic cell proliferation following transformation.

SIGNIFICANCE: TOX is an HMG box–containing protein that has important roles in T-ALL initiation and maintenance. TOX inhibits the recruitment of KU70/KU80 to DNA breaks, thereby inhibiting NHEJ repair. Thus, TOX is likely a dominant oncogenic driver in a large fraction of human T-ALL and enhances genomic instability. *Cancer Discov; 7(11); 1336–53. ©2017 AACR.*

INTRODUCTION

T-cell acute lymphoblastic leukemia (T-ALL) affects thousands of children and adults each year in the United States (1). There are many molecular subtypes of T-ALL, with those expressing the master transcriptional regulators *TAL1* and *LMO1/LMO2* comprising 60% of new diagnoses (2). Despite their wide differences in transcription factor activation and arrest at different stages of T-cell development (1, 3), T-ALL subtypes also commonly rely on similar oncogenic pathways for transformation. For example, MYC is now widely recognized as the dominant oncogenic driver in human T-ALL and is often activated downstream of NOTCH1 in a variety of T-ALL subtypes (4). Additional oncogenic drivers also likely collaborate with these potent oncogenes to elevate proliferation, alter tumor initiation, and expand early transformed T-cell pools. For example, T-ALLs often harbor genomic deletion of $p16^{INK4A}$, focal amplification of *cMYB*, and aberrant activation of TAL1 by genomic deletion of the *SIL* locus (1, 2). Chromosomal translocations are also common in T-ALL and often result in misexpression of oncogenic transcription factors or the creation of oncogenic gene fusions that drive transformation and growth (2, 5). Finally, recent studies have uncovered mechanisms to deregulate oncogenes in T-ALL through acquired small genomic insertions and deletions that drive high transcript expression and are acquired through aberrant DNA repair (6, 7). Despite the high number of genomic aberrations found in human T-ALL and the supposition that these are driven by errant DNA repair, it has yet to be fully established how these genetic lesions are acquired and what, if any, DNA repair pathways may be altered in human T-ALL.

Nonhomologous end joining (NHEJ) is important for restoring genomic integrity to cells and provides a rapid and robust mechanism to religate broken DNA strands to protect cells from undergoing apoptosis. Following the creation of a double-strand break (DSB), NHEJ repair can be initiated by recruitment of dimeric KU70/KU80 to sites of DNA damage (8). Other NHEJ factors are then recruited to the DNA lesion, including DNA-dependent protein kinase catalytic subunit (DNA-PKcs), Artemis, XRCC4, Ligase IV, and XLF/Cernunnos (9). These proteins facilitate repair by ligating the DNA ends and restoring DNA integrity. Not unexpectedly, impaired NHEJ often results in DNA translocations,

[1]Department of Pathology, Massachusetts General Hospital, Boston, Massachusetts. [2]Massachusetts General Hospital Cancer Center, Harvard Medical School, Boston, Massachusetts. [3]Harvard Stem Cell Institute, Cambridge, Massachusetts. [4]Departments of Pathology and Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia, Canada. [5]Beatrice Hunter Cancer Research Institute, Halifax, Nova Scotia, Canada. [6]College of Medicine, University of Kentucky, Lexington, Kentucky. [7]Whitehead Institute for Biomedical Research, Cambridge, Massachusetts. [8]Department of Molecular Biology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts. [9]Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts. [10]Department of Haematology, UCL Cancer Institute, University College London, London, United Kingdom. [11]Broad Institute of MIT and Harvard, Cambridge, Massachusetts. [12]Department of Life Sciences, Ben-Gurion University of the Negev, Beer Sheva, Israel. [13]Institute of Biology Leiden, University of Leiden, Leiden, the Netherlands. [14]Center for Regenerative Medicine, Massachusetts General Hospital, Boston, Massachusetts. [15]Center for Stem Cell Therapeutics and Imaging, Department of Neurosurgery, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts. [16]Division of Pediatric Hematology-Oncology, Boston Children's Hospital, Boston, Massachusetts. [17]Division of Signal Transduction, Beth Israel Deaconess Medical Center and Department of Medicine, Harvard Medical School, Boston, Massachusetts.

Note: Supplementary data for this article are available at Cancer Discovery Online (http://cancerdiscovery.aacrjournals.org/).

J. Pinder and B. Martinez-Pastor contributed equally to this article.

Corresponding Author: David M. Langenau, Massachusetts General Hospital, Department of Pathology, Boston, MA 02129. Phone: 617-643-6508; Fax: 617-726-5684; E-mail: dlangenau@mgh.harvard.edu

doi: 10.1158/2159-8290.CD-17-0267

inversions, and deletions that are characteristic of cancer (10). In fact, mice deficient in KU70 or KU80 have elevated genomic instability and develop T-cell malignancies (11, 12). Yet, KU70/KU80 and other NHEJ repair factors are not commonly deleted, mutationally inactivated, or hypermethylated in human T-cell lymphoma or T-ALL, leading investigators to conclude that these pathways may not be dominant oncogenic drivers of T-ALL nor have important roles in regulating aberrant DNA changes seen at leukemia initiation. To date, mechanisms that regulate elevated genomic instability and/or alter NHEJ repair in human T-ALL have not been established.

Using a transgenic screen in zebrafish, we have identified thymocyte selection–associated high mobility group box protein (TOX) as a collaborating oncogene that synergized with MYC and intracellular NOTCH1 to initiate early-onset T-ALL, to expand the number of transformed clones, and to elevate genomic instability. TOX contains a single HMG box motif and has been suggested to be a transcription factor based solely on its homology with other HMG-containing proteins. Yet, TOX shares remarkable conservation in protein sequence within its HMG box domain with well-known chromatin remodeling proteins that lack DNA-binding specificity. Importantly, TOX regulates various aspects of T-cell development (13–19) and is also genomically amplified in a subset of mouse and human T-ALL (20). Work presented here has uncovered that TOX is transcriptionally activated by well-known T-ALL oncogenic transcription factors, including TAL1 and LMO1/2. Functional studies revealed that TOX regulates both proliferation and NHEJ in human T-ALL, and unexpectedly these functions do not require binding of TOX to the chromatin. Rather, the HMG box of TOX binds directly to and inhibits KU70/KU80, impairing its recruitment to sites of DNA damage. Thus, one function of TOX is to lock human T-ALL cells in a state of dampened NHEJ repair. In total, our results provide a plausible cellular mechanism for elevated genomic DNA aberrations observed in human T-ALL and reveal additional roles for TOX in regulating proliferation after leukemic cell transformation.

RESULTS

TOX Accelerates Onset and Malignant Transformation in Zebrafish MYC-Induced T-ALL

To identify genes that collaborate with *MYC* to accelerate time to T-ALL onset, we completed a transgenic screen in zebrafish (Fig. 1A). Importantly, the zebrafish transgenic MYC-induced T-ALL model provides a robust and powerful tool to assess oncogenic drivers required to initiate T-ALL (21–24). This model shares remarkably similar molecular mechanisms of transformation with those found in mouse and human (21, 23–25). Moreover, zebrafish T-ALLs also harbor similar genomic DNA amplifications and deletions when compared with both human and mouse T-ALL (26, 27), supporting conserved roles for genomic instability in driving T-ALL onset. In total, 27 genes were included in our screen based on being highly expressed, amplified, and/or mutationally activated in human T-ALL (Supplementary Table S1; Supplementary Fig. S1A). Fish with fluorescent-labeled thymocytes were followed for disease progression after 21 days of life, with leukemia being defined by >50% of the animal being overtaken by fluorescent-labeled T-ALL cells (24, 26). From this analysis, 8 genes were identified that significantly shortened time to leukemia onset and increased disease penetrance ($P < 0.05$, Log-rank statistic, Fig. 1B). As has been previously found using mouse models of T-ALL, *BMI1* (28), *HIF1α* (29), *IL7R* mutations (30), and *GFI1* (31) were identified as collaborating oncogenes in our screen, validating that similar genes and pathways drive leukemogenesis in zebrafish, mouse, and human (Fig. 1B). Our work also uncovered *TOX* as a collaborating oncogene that synergized with both *MYC* and activated intracellular *NOTCH1* to induce T-ALL (Fig. 1B and C; Supplementary Fig. S1B). Importantly, *TOX* alone was insufficient to transform thymocytes over the 150 days of observation.

To define a role for TOX in modulating T-ALL initiation and growth, we next undertook a cellular and molecular characterization of MYC- and MYC+TOX-expressing T-ALL. Both MYC- and MYC+TOX-expressing T-ALL arose exclusively within the thymus (Fig. 1C), had similar lymphoblast morphology (Fig. 1F), and expressed T-cell markers, including higher TOX expression in double transgenic animals (Fig. 1G). Zebrafish *MYC*- and *MYC+TOX*-expressing T-ALLs expressed both *tal1/scl* and *lmo2*, confirming that TOX did not alter the subtype of T-ALL induced in zebrafish (Fig. 1G). EdU labeling experiments showed that proliferation rates also did not differ between MYC- or MYC+TOX-expressing T-ALL ($P = 0.34$, Student t test, $n > 9$ T-ALL assessed per genotype, Fig. 1D). However, TOX-expressing T-ALL had slightly fewer apoptotic cells when assessed by Annexin-V staining (Fig. 1E, $P = 0.012$, ANOVA test). Finally, clonal analysis uncovered that MYC+TOX-expressing T-ALLs had a dramatic 3-fold increase in TCRβ-recombined clones (MYC+TOX, 7.4 ± 1.5 clones compared with MYC, 2.4 ± 0.3; $P = 0.0001$, Mann–Whitney test, Fig. 1H; Supplementary Table S2). We conclude that TOX has prominent roles in transforming early thymic precursor cells and acts by expanding the overall pool of transformed clones at leukemia initiation.

TOX Induces Genomic Instability in Zebrafish T-ALL and Mouse Embryonic Fibroblasts

Human T-ALLs often harbor genomic deletions and amplifications indicative of genome instability and can be assessed clinically by quantitative changes in DNA content following propidium iodide (PI) staining and flow cytometry (32). This same methodology has been previously adapted to zebrafish T-ALL and is a rapid assay to identify leukemias with genomic copy-number changes (25). As previously reported, 17% of zebrafish MYC-induced T-ALL had increased DNA content ($n = 13$ of 75; ref. 25). By contrast, 47% of MYC+TOX-expressing T-ALLs had altered DNA content ($n = 8$ of 17, $P = 0.013$, Fisher exact test; Fig. 2A and B). TOX-expressing T-ALLs also comprised a substantial fraction of leukemias that were hypodiploid, contrasting starkly with zebrafish MYC-induced T-ALL where hypodiploidy has never been observed ($P = 0.006$, Fisher exact test). Together, these results show that TOX is a critical driver of genomic instability in zebrafish MYC-induced T-ALL.

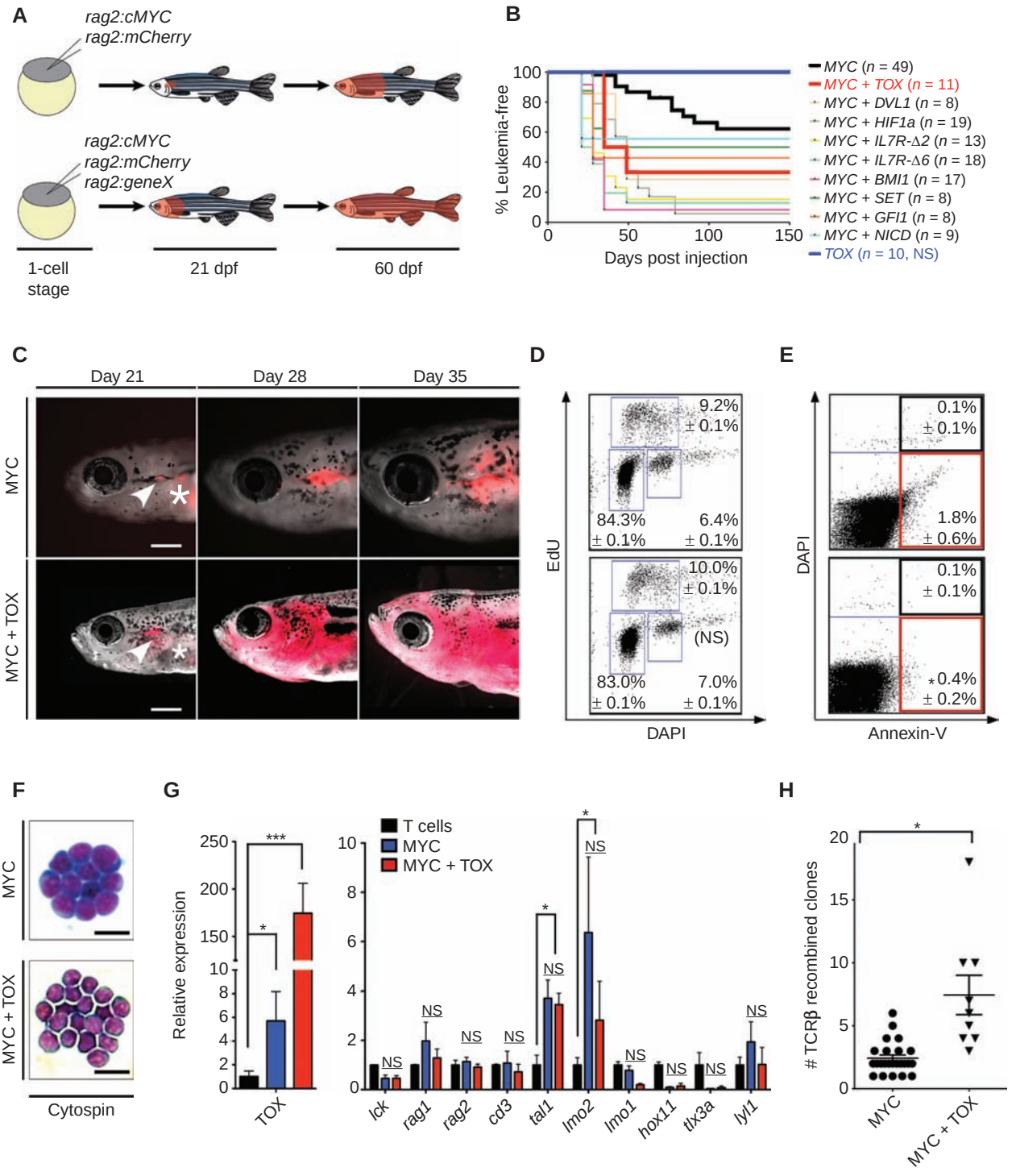


Figure 1. TOX collaborates with MYC to accelerate T-ALL onset. **A,** Schematic of transgenic screen. **B,** Kaplan–Meier analysis ($P < 0.05$, log-rank statistic). Leukemic fish have >50% of their body overtaken by T-ALL cells. Number of animals analyzed per genotype is shown in parenthesis. **C,** Images of T-ALL transgenic fish at 21, 28, and 35 days postfertilization (dpf). Asterisks denote autofluorescence. Arrowheads show leukemias initiating in the thymus. Scale bar, 2 mm. **D,** EdU proliferation analysis of zebrafish T-ALL. NS, not significant. ±STD noted. **E,** Annexin-V apoptosis staining as assessed by flow cytometry. *, $P < 0.05$, two-tailed Student *t* test. ±STD noted. Note that ≥9 primary T-ALLs were analyzed per genotype in **D** and **E**. **F,** Cytospin showing lymphoblast morphology ($n > 5$ leukemias/genotype analyzed). Scale bar, 20 μm. **G,** qRT-PCR gene expression comparing *MYC*- and *MYC+TOX*-expressing T-ALL ($n = 5$ per genotype, run in triplicate) with sorted thymocytes isolated from *rag2*:GFP transgenic fish (purity >95%, viability >95%). *, $P < 0.05$ and ***, $P < 0.001$, two-tailed Student *t* test. NS, not significant. Error bars denote SEM. **H,** Number of TCRβ clones per primary leukemia (*MYC+TOX*: $n = 9$ animals; *MYC*: $n = 24$ animals). *, $P < 0.05$, two-tailed Student *t* test. Error bars, SEM.

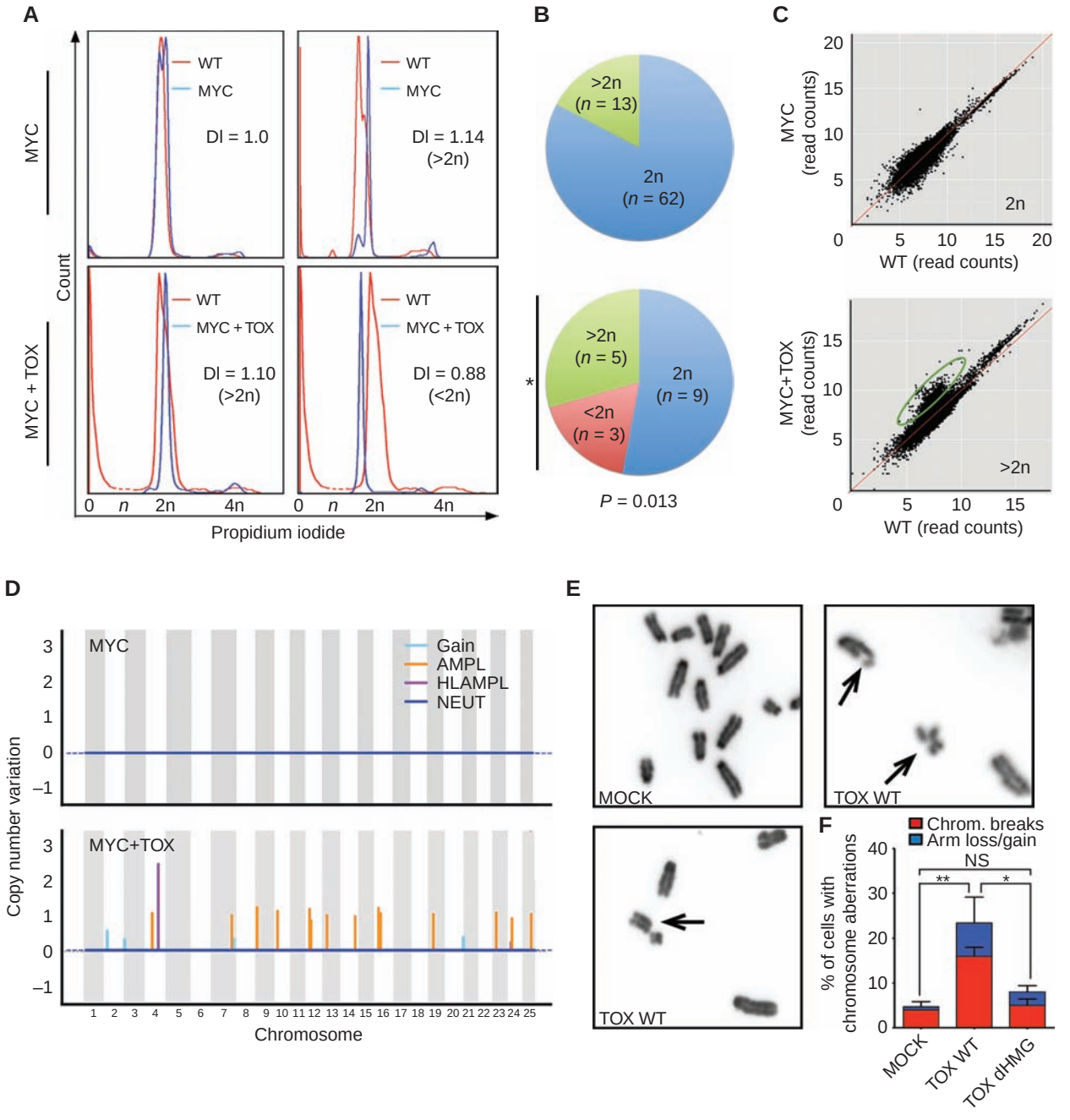


Figure 2. TOX promotes genomic instability. **A,** Flow cytometry analysis showing DNA content of zebrafish T-ALL (blue) compared with normal blood-derived DNA (red). Representative leukemias are shown with changes in DNA index (DI) noted. 2n has a DNA index of 1. WT, wild-type. **B,** Quantization of DNA content. *, $P = 0.013$, Fisher exact test. **C,** Genomic DNA alterations identified by whole-genome sequencing. The log–log scatter plot represents read counts within nonoverlapping window size of 10 kb across the genome comparing leukemia and control tissue from the same, representative animal. Amplifications found in the MYC+TOX leukemia are denoted by the green circle. **D,** Manhattan plot representing the copy-number variation across the genome of a representative MYC-expressing (top) and MYC+TOX-expressing (bottom) T-ALL. Regions of significant gain, amplification (AMPL), high-level amplification (HLAMPL), and neutral 2N copy number (NEUT) noted. **E–F,** Analysis of metaphase spreads from MEF cells infected with control (MOCK), full-length wild-type TOX (TOX WT), or TOX that lacks the HMG box domain (TOX dHMG). Arrows denote chromosome abnormalities. **F,** Quantification of cells with genomic abnormalities. More than 50 nuclei were counted per condition and replicated 3 independent times. Error bars denote SD. *, $P < 0.05$ and **, $P < 0.005$, Student t test. NS, not significant.

To validate that TOX-expressing T-ALLs have elevated genomic instability and harbored alterations in genomic DNA, low-pass whole-genome sequencing (WGS) was completed using Illumina HiSeq (0.5-1.0X genome coverage). Low-pass WGS can effectively detect deletions and amplifications akin to array comparative genomic hybridization. When compared with matched control tissue isolated from the same leukemic fish, all three MYC+TOX leukemias showed large-scale genomic variations including deletions, gains, and amplifications (Fig. 2C and D; Supplementary Table S3). No recurrent lesions were identified in all three of the TOX+MYC-expressing T-ALLs (Supplementary Table S3), indicating that TOX does not collaborate with a specified set of amplifications or deletions to drive transformation, but rather that TOX is likely a general modulator of genomic instability. Three MYC-induced T-ALLs were also analyzed by WGS, confirming DNA content changes identified by FACS and overall fewer DNA copy-number aberrations when compared with TOX-expressing T-ALLs (Fig. 2D; Supplementary Table S3). Taken together, these results independently confirm that TOX imparts elevated genomic instability to zebrafish T-ALL cells.

To further assess a role for TOX in regulating genomic instability, *TOX* was transfected into mouse embryonic fibroblast (MEF) cells and chromosomal abnormalities were assessed by metaphase spread (Fig. 2E and F). This analysis revealed that TOX significantly elevated genomic instability that was indicative of impaired NHEJ including elevation in unrepaired chromosome breaks and loss/gain of chromosome arms ($P = 0.005$, two-tailed Student *t* test; Fig. 2F). MEF cells that expressed a mutant form of TOX that lacked the HMG box, as a result of deletion of amino acids 261–339, also did not exhibit elevated genomic instability when compared with control-treated cells ($P = 0.15$, Student *t* test). These data further support a role for TOX in elevating genomic instability and suggest prominent roles for the HMG box domain in regulating this process.

TOX Is Expressed in a Majority of Human T-ALL and Is Regulated by the TAL1/MYB Transcriptional Complex

To determine the extent of TOX expression in human T-ALL, we next analyzed transcript expression in a wide array of human cancer cell lines and uncovered that *TOX* was highly expressed in human T-ALL (Fig. 3A). This observation was extended to primary T-ALL patient samples, showing that *TOX* transcripts were highly expressed in human primary T-ALL when compared with normal marrow, irrespective of molecular subtype classification ($n = 157$ of 165 primary T-ALL express high transcript levels of *TOX*; Fig. 3B; Supplementary Fig. S2A and S2B). TOX protein expression was also assessed by Western blot analysis and showed that TOX was expressed in both human T-ALL cell lines and primary patient samples ($n = 10$, Fig. 3C; Supplementary Fig. S2C and S2D). FACS analysis confirmed high expression of TOX at single-cell resolution in a large fraction of primary human T-ALLs (Fig. 3D and E). By contrast, TOX protein was detected at lower levels in human thymocytes and B-ALL ($n = 5$).

Given that *TOX* is highly expressed in a large fraction of human T-ALL and yet amplified only in a small subset of human leukemias (20), we next investigated how *TOX* might be transcriptionally regulated in malignant T cells. High levels of H3K27ac occupancy are associated with recruitment of mediator and master transcription factors, which define stretch/superenhancers that ultimately drive high transcript expression (33, 34). Genomic analysis uncovered elevated H3K27ac occupancy near the *TOX* locus in human T-ALLs but not immature and mature T-cell subsets or $CD34^+$ marrow cells (Fig. 3F). Superenhancer analysis revealed two prominent genomic regions that drive high transcript expression in human T-ALL cells but were not found in normal thymocytes or mature T-cell subsets (Supplementary Table S4). Further analysis of H3K27ac peaks revealed that superenhancer #2 contained both the mediator complex and master T-ALL transcription factors including *TAL1, MYB,* and *GATA3* (Fig. 3G). Knockdown of these factors and additional members of the TAL1 complex, including *HEB* and *E2A*, led to significant reduction in *TOX* transcript expression in human T-ALL cells (Fig. 3H). We conclude that high *TOX* transcription is regulated, at least in part, by superenhancer regulation in human T-ALL.

TOX Regulates Human T-ALL Growth, Proliferation, and Apoptosis

To assess the consequences of TOX loss of function, HPB-ALL, CCRF-CEM, and MOLT-4 human T-ALL cells were transduced with shRNAs against *TOX* and compared with scramble control shRNA. Following stable knockdown of TOX using independent shRNAs, all three T-ALL cell lines had reduced viability and overall growth when assessed by CellTiter-Glo (Fig. 4A and B). EdU proliferation analysis showed that sh*TOX* knockdown T-ALL cells exhibited a marked disruption of cell cycle, with arrest of cells in S-phase and overall reductions in cycling G_2–M cells (Fig. 4C). *TOX* shRNA knockdown cells also had higher levels of apoptosis when assessed by Annexin-V/PI staining when compared with control cells ($P < 0.02$, Student *t* test, Fig. 4D; Supplementary Fig. S3). These results show that TOX has major roles in regulating human T-ALL cell growth and leukemia maintenance, affecting both cell cycle and apoptosis.

To extend our findings to the *in vivo* setting, human HPB-ALL and CCRF-CEM knockdown cells were transduced with luciferase and assessed for growth in NOD/SCID/IL2Rγ-null mice. TOX knockdown cells exhibited significantly reduced xenograft growth by 21 days when compared with control shRNA–treated cells (Fig. 4E–J). Together, these data show that TOX also has important roles in regulating continued growth and maintenance of human T-ALL cells *in vivo*.

TOX Binds Directly with KU70/KU80 through the HMG Box Domain

To develop hypotheses about how TOX might be regulating transformation, we sought to identify TOX binding partners. Specifically, endogenous TOX antibody pull-down experiments were performed in human HPB-ALL, MOLT-4, and Jurkat T-ALL cells and interacting factors identified through LC/MS-MS. Following separation of proteins on an SDS-PAGE gel, prominent protein bands were detected in the TOX pull-down cell lysates from

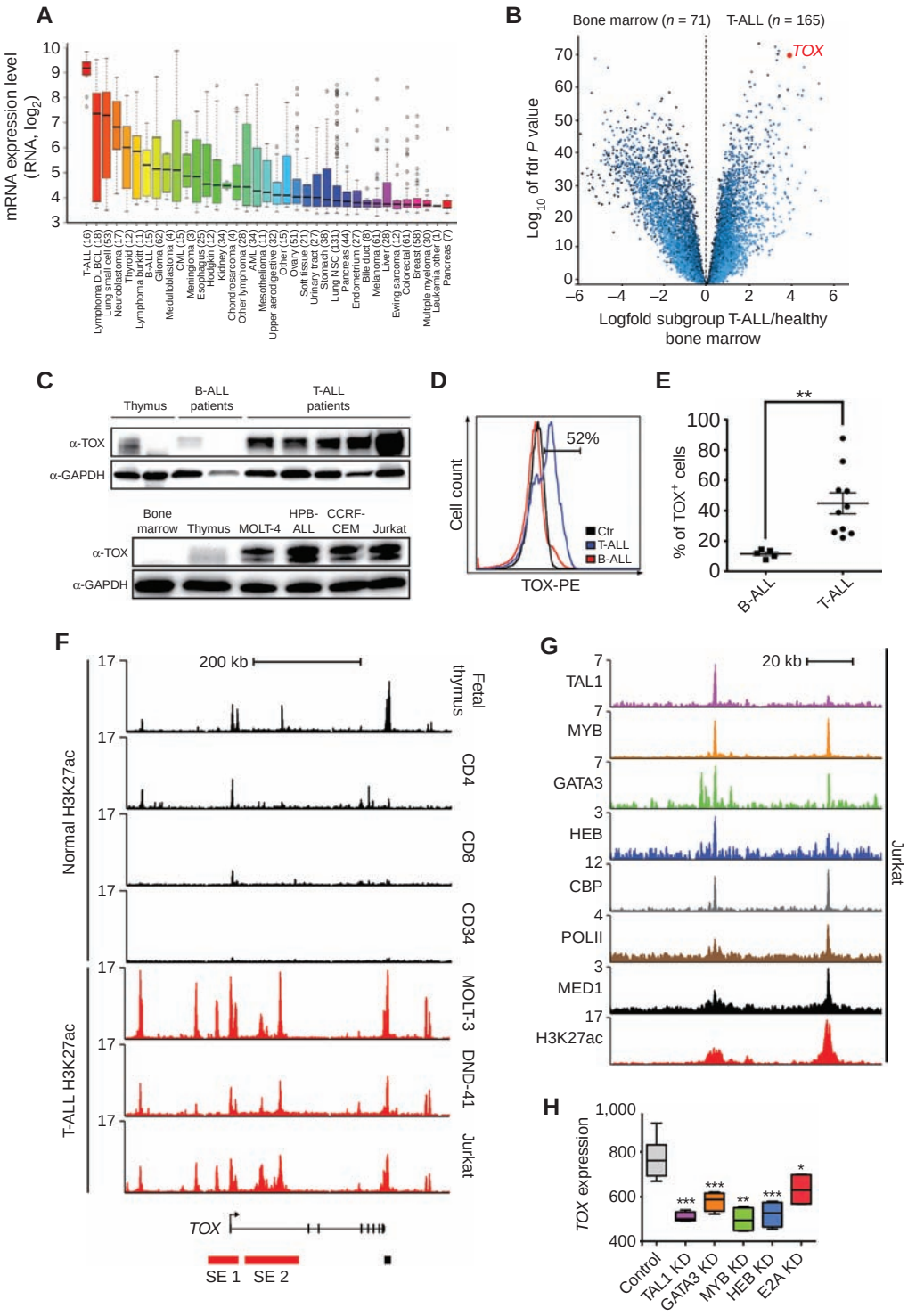


Figure 3. TOX is highly expressed in human T-ALL and transcriptionally regulated in a subset of leukemias by the TAL1/MYB complex. **A,** TOX microarray gene expression from the Cancer Cell Line Encyclopedia. Box-and-whisker plots denote median *TOX* expression (black line), the interquartile range (box), and 1.5 × the interquartile range (bars). **B,** Volcano plot comparing gene expression between human T-ALL patient samples and bone marrow. **C,** Western blot analysis of primary human leukemia (top) and T-ALL cell lines (bottom). The same thymus sample was run in lane 1 (top) and lane 2 (bottom). **D,** Flow cytometry analysis showing TOX expression in primary human T-ALL compared with B-ALL. Control is T-ALL cells stained with conjugated IgG control antibody (ctr). **E,** Quantification showing the percentage of TOX$^+$ cells found in each patient sample (right, **, $P = 0.006$, two-tailed Student *t* test). Mean and SEM are denoted. **F–H,** TOX is associated with two distinct, H3K27 acetylated superenhancers (SE) in human T-ALL. **F,** Chromatin immunoprecipitation sequencing of human T cells, CD34$^+$ progenitor cells from the marrow (black peaks), and T-ALL cell lines (red peaks). Superenhancer 1 (SE 1) is found in MOLT-3 and Jurkat, whereas SE 2 is found in all three human T-ALL cell lines. **G,** Magnified view of SE 2 in Jurkat T-ALL cells showing superenhancer occupancy by the TAL1/MYB complex. **H,** *TOX* gene expression in Jurkat cells following knockdown of T-ALL transcription factors found within the H3K27 acetylated superenhancer (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$, dataset from GSE29179; ± SEM denoted).

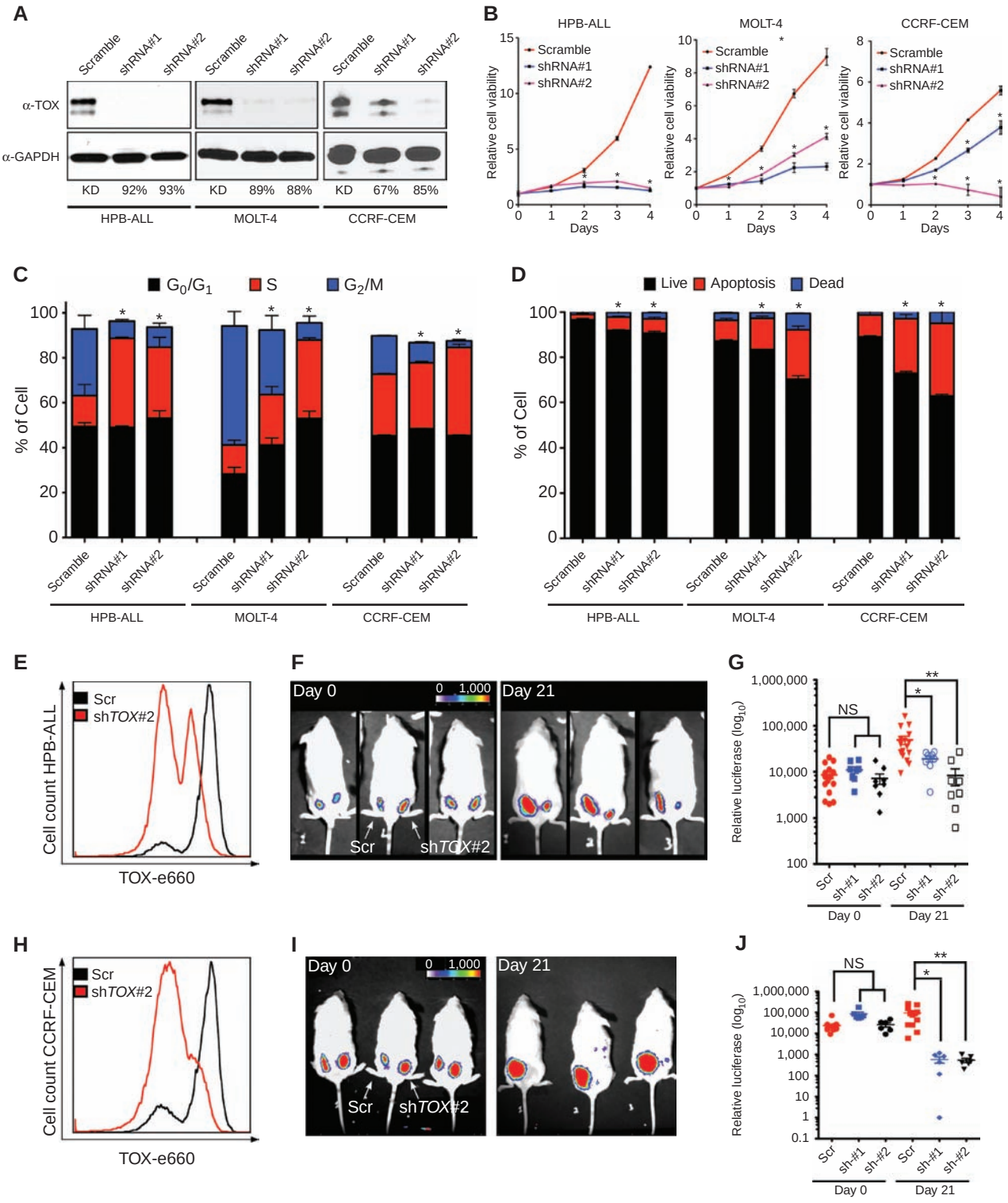


Figure 4. TOX is required for continued human T-ALL growth and maintenance. **A,** Western blot analysis following stable shRNA knockdown in human T-ALL cells. Percent knockdown noted. **B,** Cell viability following knockdown as assessed by CellTiter-Glo. **C,** EdU proliferation analysis. **D,** Annexin-V staining. Asterisks in **B** to **D** denote $P < 0.05$, Student *t* test. Xenograft studies performed with human HPB-ALL (**E–G**) and CCRF-CEM (**H–J**). **E** and **H,** Flow cytometry analysis showing efficiency of TOX knockdown. **F** and **I,** Luciferase bioluminescent imaging of representative animals engrafted at 0 day compared with 21 days. Scramble shRNA control (left flank) or sh*TOX#*2 (right flank). **G** and **J,** Quantification of xenograft growth using two independent shRNAs. NS, not significant; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$, ANOVA test.

HPB-ALL cells but not the IgG control cells (Supplementary Fig. S4). These bands were excised and analyzed by LC/MS-MS. From this analysis, KU70 and KU80 were identified as TOX binding factors (Supplementary Table S5A). Additional LC/MS-MS experiments performed in MOLT-4 and Jurkat cells confirmed the KU70/KU80 interaction with TOX using analysis of total protein eluates obtained following endogenous TOX immunoprecipitation (Supplementary Table S5B). In these experiments, KU70 and KU80 exhibited a remarkable >22-fold enrichment over IgG control pull-down cells, independently verifying TOX interaction with KU70/KU80.

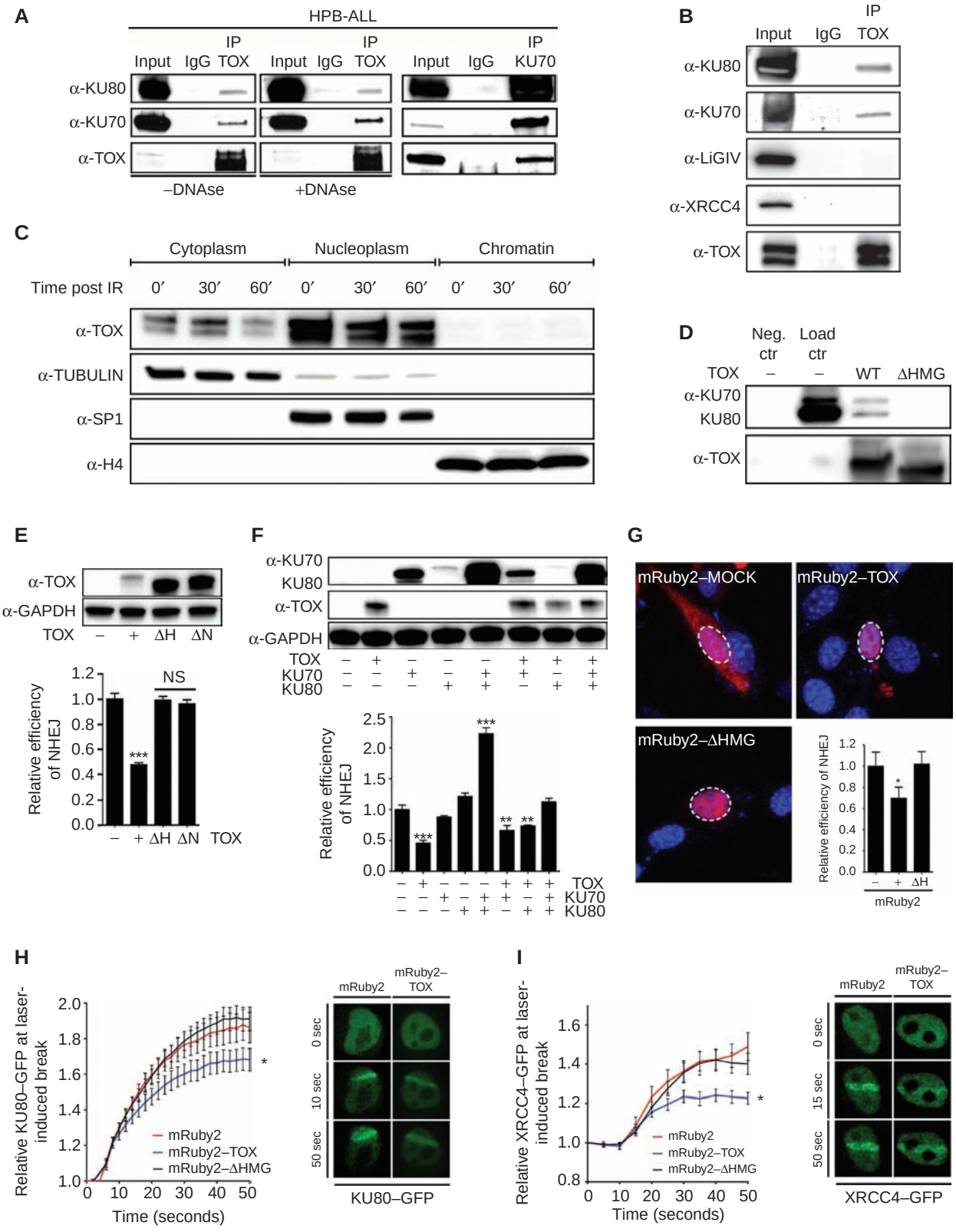
To confirm KU70/KU80 protein interactions identified by mass spectrometry, TOX immunoprecipitation was performed in the presence and the absence of DNase I, and Western blot analysis was completed with KU70- and KU80-specific antibodies in all three human T-ALL cell lines. DNase I treatment ensured that interactions of TOX and KU70/KU80 did not result from binding common genomic DNA fragments (Fig. 5A). As was seen by the IP–mass spectrometry experiments, TOX interacted with KU70/KU80 in human HPB-ALL, CCRF-CEM, and MOLT-4 human T-ALL cell lines (Fig. 5A; Supplementary Fig. S5A). Reciprocal immunoprecipitation using the KU70 antibody confirmed specificity of the interaction (Fig. 5A, right). Consistent with our LC/MS-MS experiments, Western blot analysis also revealed that late-acting NHEJ factors including XRCC4 and DNA Ligase IV did not pull down with TOX (Fig. 5B), showing that the interaction of TOX with KU70/KU80 was limited to the dimeric initiating complex that is first recruited to DNA DSBs. *In vitro* binding assays using purified protein and performed in the absence of DNA confirmed a direct interaction between TOX, KU70, and KU80 (Fig. 5D; Supplementary Fig. S5B). By contrast, HMG-deleted TOX was unable to bind KU70/KU80 in *in vitro* binding assays (Fig. 5D), confirming a requirement for this domain in regulating the TOX protein:protein interaction with dimeric KU70/KU80.

Endogenous antibody pull-down and *in vitro* binding assays showed the HMG box motif was responsible for protein:protein interactions with KU70 and KU80 and did not require DNA binding; thus, we wondered if, like other HMG group box–containing proteins, TOX bound chromatin. Cell fractionation studies were performed in human HPB-ALL T-ALL cells and identified that TOX was highly expressed in the nuclear fraction and to a lesser degree to the cytoplasm, yet TOX did not bind to chromatin (Fig. 5C). Irradiation of cells induces large amounts of DNA damage and facilitates recruitment of repair enzymes to DNA breaks. Despite high-dose irradiation of human T-ALL cells, TOX was still not recruited to chromatin even after 60 minutes following 3 Gy irradiation treatment (Fig. 5C; Supplementary Fig. S5C). We conclude that TOX binds directly to KU70/KU80 through its HMG box domain and that this protein:protein interaction does not require DNA binding.

TOX Inhibits DNA Repair by Suppressing Recruitment of NHEJ Factors to DSBs

Given the prominent roles for KU70/KU80 in regulating NHEJ and subsequently genomic instability (11, 12) and that TOX overexpression in zebrafish T-ALL and MEFs leads to genomic instability, we hypothesized that TOX might be an inhibitor of KU70/KU80 function and thus affect NHEJ repair. Full-length TOX or deletion mutants that lack the nuclear localization signal (NLS) or HMG box domain were stably expressed in 3T3 fibroblast cells (Fig. 5E; Supplementary Fig. S6A). Cells were then transfected with linearized plasmid that contained a DNA DSB between the promoter and the GFP coding sequence, where GFP is expressed only following successful DNA repair. Cells expressing full-length TOX exhibited a remarkable 54% ± 4.5% reduction in DNA repair when compared with control cells (two-tailed Student *t* test, $P = 0.0004$, ±STD, Fig. 5E), whereas 3T3 cells expressing TOX deletion mutants that lack the HMG box or nuclear localization signal did not have altered NHEJ. Epistasis experiments revealed that when both KU70 and KU80 were expressed together, NHEJ was fully restored to TOX-expressing cells (Fig. 5F). By contrast, expression of either KU70 or KU80 alone was not sufficient to restore NHEJ in TOX-expressing cells (Fig. 5F). Suppressed NHEJ was also observed in 3T3 cells transfected with red fluorescent protein (RFP) fusions of mRuby2 with full-length TOX but not when fused with deletion mutants that lack the HMG box domain (Fig. 5G). Imaging studies verified that both full-length TOX and HMG box–deleted TOX were predominantly localized to the nucleus (Fig. 5G), whereas the NLS mutant was localized to the cytoplasm (Supplementary Fig. S6B).

Figure 5. TOX binds directly to KU70/KU80 and inhibits NHEJ by suppressing recruitment of KU70/KU80 to sites of DNA damage. **A,** TOX immunoprecipitation followed by Western blot analysis in HPB-ALL cells in the presence or absence of DNaseI treatment. The right plot shows the reciprocal immunoprecipitation using anti-KU70. **B,** TOX immunoprecipitation followed by Western blot analysis for members of the NHEJ pathway. **C,** Western blot analysis following cell fractionation in nonirradiated and 3 Gy-irradiated HPB-ALL cells. TUBULIN, SP1, and Histone H4 (H4) are controls for assessing cytoplasmic, nucleoplasmic, and chromatin fractions, respectively. **D,** *In vitro* binding of TOX (WT) or HMG-box (ΔHMG) deletion mutant with KU70/KU80 followed by Western blot analysis. Negative control was KU70/KU80 added to beads and then purified in the absence of TOX (Neg. ctr). Loading control contains only purified KU70/KU80 proteins (Load ctr). **E,** Western blot analysis of 3T3 cells following transfection with full-length TOX (+), ΔHMG (ΔH), or ΔNLS (ΔN). Below is the quantification of the NHEJ assay (NS, not significant; ***, $P < 0.0001$, two-tailed Student *t* test comparing experimental samples with control, ± STD noted). **F,** Western blot analysis of 3T3 cells following transfection with full-length TOX, KU70, and/or KU80. Below is shown the quantification of the NHEJ assay (**, $P < 0.001$; ***, $P < 0.0001$, two-tailed Student *t* test, ± STD noted). **G,** Confocal imaging of 3T3 cells transfected with mRuby2-fused constructs. Dashed lines denote nucleus as assessed by Hoechst stain ($n > 100$ cells/construct analyzed). The bottom right plot shows the fluorescence-based NHEJ assay completed in 3T3 (*, $P = 0.03$, two-tailed Student *t* test). ± STD noted. **H** and **I,** Quantitative assessment of recruitment of KU80–GFP (**H**) or XRCC4–GFP (**I**) to sites of UV laser-induced DNA damage in 3T3 cells. (*, $P < 0.05$, two-tailed Student *t* test). ± STD noted. Representative fluorescent images of cells following laser-induced damage are shown to the right.



To directly assess if TOX modulates recruitment of KU70/KU80 to the sites of DNA repair, 3T3 fibroblasts were transfected with KU80–GFP and mRuby2, mRuby2–TOX, or mRuby2–ΔHMG, subjected to UV laser microirradiation-induced DNA damage, and assessed for KU80–GFP recruitment to DSBs (Fig. 5H). Cells expressing full-length TOX exhibited significantly reduced recruitment of KU80–GFP to DNA breaks ($P < 0.05$, Student t test), whereas cells that expressed mRuby2–ΔHMG efficiently recruited KU80–GFP to sites of DNA damage with similar kinetics to control cells. mRuby2–TOX and fusions with HMG-deleted TOX were not recruited to microirradiation-induced breaks and remained nuclear localized throughout these experiments. Impaired recruitment of NHEJ pathway regulators to sites of DNA damage was independently confirmed using XRCC4–GFP ($P < 0.05$, Student t test, Fig. 5I). These experiments show that TOX sequesters KU70/KU80 away from sites of active DNA repair.

TOX Loss of Function in Human T-ALL Leads to Elevated NHEJ Repair

We hypothesized that TOX might lock human T-ALL cells in a state of dampened DNA repair, with the prediction that loss of TOX should elevate NHEJ repair. Stable knockdown or shRNA-expressing control cells were transfected with linearized plasmid that contained a DNA DSB between the promoter and the mCherry coding sequence, with fluorescence being detected only following successful DNA repair. Plasmid-based repair assays showed that HPB-ALL, CCRF-CEM, and MOLT-4 T-ALL cells all exhibited enhanced NHEJ repair following stable TOX knockdown using independent shRNAs (Fig. 6A–F; Supplementary Fig. S7; $P < 0.05$, Student t test). Similar results were obtained using stable integration of the Traffic Light Reporter (TLR). The TLR assay reads out NHEJ repair via RFP expression following induction of a DSB using tamoxifen-inducible I-SceI restriction enzyme (35). These experiments showed that NHEJ repair was greatly increased following TOX knockdown in CCRF-CEM cells, exhibiting a striking increase in NHEJ following TOX depletion ($P = 0.0001$, two-tail Student t test, Fig. 6G and H).

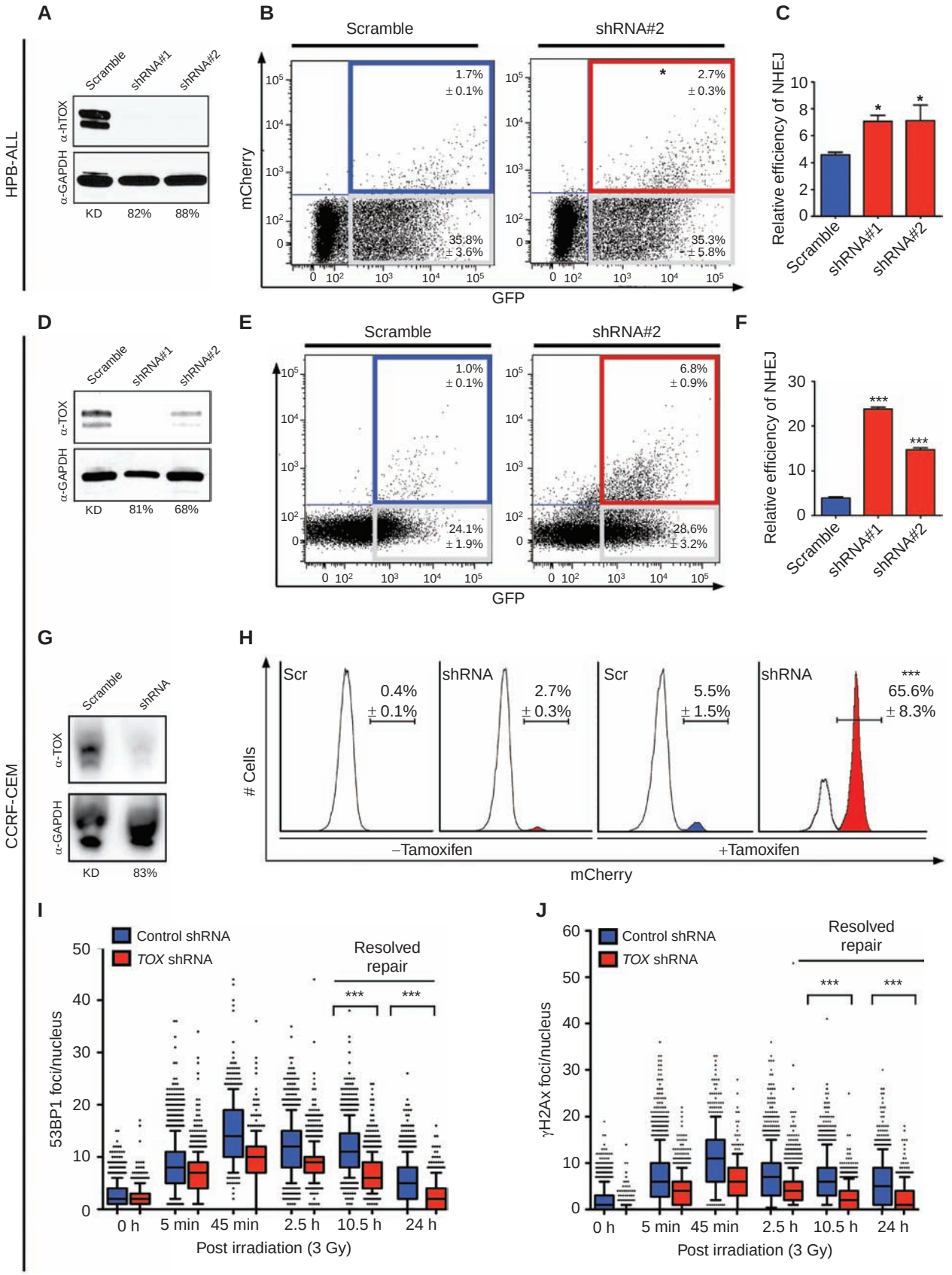
To independently confirm a role for TOX in modulating DNA DSB repair, we next assessed the kinetics and localization of 53BP1 and γH2A.X in irradiated CCRF-CEM and HPB-ALL T-ALL cells following stable knockdown of TOX. 53BP1 and γH2A.X are recruited to DNA following induction of DSBs, and their kinetics can be measured to quantify resolution of the breaks following γ-irradiation (36). Using this strategy, we found that TOX knockdown led to significantly faster DNA break resolution in human T-ALL cells following 3 Gy irradiation (Fig. 6I and J; Supplementary Fig. S8; $P < 0.001$, Student t test). The overall numbers of foci were also reduced at many early time points, reflecting ongoing and faster repair in TOX-depleted cells following irradiation-induced DSB damage. Similar results were also seen in irradiated HPB-ALL cells (Supplementary Fig. S8; $P < 0.001$, Student t test). Together, our data indicate that TOX negatively regulates the NHEJ pathway in human T-ALL and that human T-ALLs are locked in a state of dampened DSB DNA repair (Supplementary Fig. S9).

DISCUSSION

Our work has uncovered TOX as a collaborating oncogenic driver in T-ALL with important roles in both leukemia initiation and maintenance. *TOX* is amplified in a small subset of mouse and human T-ALL and was included in our screen based on these results (20). Yet, our work uncovered that most human T-ALLs express high levels of TOX and that this high transcript expression largely results from superenhancer regulation by well-known T-ALL transcription factors including TAL1 and LMO1/2. Thus, TOX does not commonly meet the criteria for being a classically defined oncogene that is activated by translocation, amplification, or activating mutations, but rather TOX is expressed in developing T-ALL cells by oncogenic transcription factors, locking leukemia cells in a state of both heightened proliferation and dampened DNA repair. Importantly, TOX also had unexpected roles in regulating leukemia initiation by expanding the pool of initiating clones and elevating genomic instability in the zebrafish T-ALL model. Experiments completed in MEF cells confirmed roles for TOX in inducing genomic instability and generating chromosomal aberrations. Finally, studies of human T-ALL and xenograft studies revealed additional roles for TOX in regulating continued leukemia maintenance by specifically regulating cell-cycle proliferation and apoptosis. These results are in keeping with known roles for TOX in regulating cell cycle in primary Sezary cells and cutaneous T-cell lymphoma where *TOX* knockdown led to cell-cycle arrest and secondarily cell death (37). Collectively, our work has identified TOX as a collaborating oncogenic driver in T-ALL which likely exerts important and diverse functional effects in a large fraction of human T-ALLs.

Because TOX had been previously shown to regulate proliferation in primary Sezary cells and cutaneous T-cell

Figure 6. TOX loss of function increases DNA repair and accelerates the time to break resolution in human T-ALL cells. **A–F,** Transient repair assays in HPB-ALL (**A–C**) and CCRF-CEM (**D–F**) cells that have shRNA knockdown. **A** and **D,** Western blot analysis with percent knockdown noted. **B** and **E,** Flow cytometry analysis of knockdown and control cells following transient fluorescence-based NHEJ assay. **C** and **F,** Graphical summary of data. *, $P < 0.05$; ***, $P < 0.001$ when compared with control shRNA–treated cells, two-tailed Student t test. **G,** Western blot analysis of TOX expression in CCRF-CEM cells with stable *TOX* shRNA #2 or scramble control shRNA knockdown used in the Traffic Light Reporter assay. **H,** Flow cytometry analysis of CCRF-CEM cells stably integrated with the Traffic Light Reporter. Cells were nucleofected with Tamoxifen-inducible I-SceI meganuclease and induced as noted in the figure. Percentages of cells that undergo NHEJ repair and express mCherry are noted (±1STD). ***, $P = 0.0001$, two-tailed Student t test. **I** and **J,** Kinetics of the DNA repair in CCRF-CEM control and stable knockdown T-ALL cells (*TOX* shRNA #2). The numbers of 53BP1 (**I**) and γH2A.X (**J**) foci per nucleus following 3 Gy irradiation are denoted. Each point represents data from a single cell, and the black bars denote the median foci per cell (***, $P < 0.0001$, two-tailed Student t test, >150 cells counted/condition). Box-and-whisker plots denote expression with the median 75% of samples and bars 90%.



lymphoma (37), our studies focused on uncovering additional roles for TOX in regulating DNA repair pathways. Remarkably, our endogenous TOX pull-downs and subsequent LC/MS mass spectrometry analysis uncovered potent interactions between TOX and KU70/KU80, prompting further investigation into roles for TOX in regulating DNA repair and NHEJ. Loss of NHEJ repair genes in mice is well known to cause elevated genomic instability and cancer. For example, KU70 and KU80 deficiency in mice results in genomic instability and predisposition to T-cell malignancy (11, 12). Yet, genetic loss of KU70 or KU80 or disruption of other NHEJ factors has not been implicated as a major tumor suppressor pathway in human T-ALL. Rather, our work has shown that TOX acts as a negative regulator of KU70/KU80 function to alter NHEJ repair and to elevate genomic instability. Remarkably, TOX remains highly expressed in fully malignant human T-ALL and locks leukemia cells in a state of dampened NHEJ repair, an observation confirmed using a wide array of DNA repair assays performed directly in human T-ALL cells. Even more remarkable is that TOX performs this function by directly binding KU70/KU80 and does not require binding to chromatinized DNA. Perturbations in NHEJ result in the creation of deletions, amplifications, transversions, and translocations that impart growth advantages to early transformed cells and drive cancer initiation. Thus, it is not surprising that mutations in DNA repair proteins can alter NHEJ and are predisposing to cancer. Our work strongly suggests that TOX may be such a factor in T-ALL.

Our results also suggest the existence of a new class of HMG group box proteins that function without binding chromatinized DNA. Indeed, endogenous chromatin immunoprecipitation studies and DNA-free *in vitro* binding assays show that TOX does not bind to chromatin. Rather, the HMG box domain of TOX is required for protein:protein associations with KU70/KU80 and functionally impairs recruitment of NHEJ factors to breaks. TOX contains a single highly conserved HMG box motif that is structurally distinct from both class I and II HMG box group–containing proteins (38). Class I HMG box group proteins are transcription factors and contain a single HMG box that binds DNA in a sequence-specific manner. These transcription factors include the TCF/LEF1 and SOX family members and have important roles in regulating a wide array of developmental and cancer processes (39). By contrast, class II group proteins contain two DNA-binding motifs, bind chromatin in a sequence agnostic manner, and have important roles in modifying DNA structure. For example, HMGB1 modulates nucleosome structure and alters transcription by looping chromatinized DNA to facilitate interaction of transcription factors with enhancer elements (40). HMGB1 also regulates a multitude of DNA repair pathways (41–43). For example, HMGB1 can enhance *in vitro* DNA ligase activity, by bringing DNA ends into close proximity (44). Yet, to date, functional roles for HMGB1 and other class II HMG group box proteins in regulating DNA binding and repair necessarily require binding to chromatin.

Together, our data suggest that TOX may represent the first of a class of HMG proteins that function, at least in part, by binding regulatory proteins through their HMG box and regulating NHEJ repair. Given the prominent role TOX has in T-cell development and its coordinated regulation coincident with TCRβ and TCRα rearrangement in mice, it is also possible that TOX may have roles in development including regulation recombination-activating gene (RAG)–mediated recombination. Moreover, TOX is highly and specifically expressed in human T-cell malignancies, whereas its related family members TOX2, TOX3, and TOX4 are differentially regulated in a variety of human cancers including breast cancer, Ewing sarcoma, multiple myeloma, and small-cell lung cancer (45–47). Many of these tumors are well known to have elevated genomic instability and harbor characteristic lesions that are associated with impaired NHEJ. It will be important to assess whether these closely related TOX family members share molecular functions in regulating cell proliferation, transformation, and genomic instability in a wider range of human cancers.

MATERIALS AND METHODS

Transgenic DNA Expression Constructs

DNA constructs used to generate transgenic zebrafish included *rag2:mCherry* (23), *rag2:MYC* (21), and *rag2:NICD* (24). All genes, except *FNBP1, GFI1, GITR, BMI1,* and *TOX,* used in this screen were obtained from hORFeome (48). *IL7R* constructs were created using site-directed mutagenesis. Full-length open reading frames for human *FNBP1, GFI1,* and *GITR* were obtained by PCR amplification of cDNA derived from human T-ALL cells. Zebrafish *tox* was amplified from the cDNA of 1- to 5-day-old Tu/AB zebrafish and subcloned into pENTR gateway system (Life Technologies). PCR primer sequences can be found in Supplementary Table S6. All genes were transferred into the *rag2* promoter destination vector using LR clonase II (Life Technologies).

Creation of Zebrafish T-ALL

Plasmids were linearized with *NotI* or *XhoI* and purified. Mosaic transgenic animals were generated as previously described (22). Note that 40 ng/μL *rag2:mCherry* was mixed with 40 ng/μL of *rag2:Myc* or *rag2:NICD* and 40 ng/μL *rag2:experimental gene* and microinjected into one-cell stage Tu/AB embryos. Animals were scored for fluorescent-labeled thymi at 21 and 28 days of life and then followed weekly for disease onset. Leukemic fish were defined as >50% of their body being infiltrated with fluorescent T-ALL cells, as previously described (24, 26). Zebrafish experiments were approved under animal protocol 2011N000127 (Massachusetts General Hospital).

Clonality, Proliferation, Apoptosis, and DNA Content

Proliferation was assessed using the Click-IT EdU Kit (Life Technologies) according to the manufacturer's protocol. A pulse of EdU was performed following cell harvest and completed for 1 hour at 28.5°C. Zebrafish T-ALL cells were assessed for apoptosis using DAPI and Alexa Fluor 647–conjugated Annexin-V (Life Technologies). Numbers indicate the mean of ≥9 biological replicates ± the SEM. For DNA content, 1×10^6 zebrafish T-ALL cells were permeabilized in 70% ice-cold ethanol for >2 hours at −20°C. Cells were then incubated in 500 μL of a solution containing 50 μg/mL PI, 1 mg/mL PureLink RNase, and PBS-1% BSA for 30 minutes at 25°C. EdU-stained cells, Annexin-V–stained cells, and DNA content were then analyzed using LSR II or Fortessa Flow Cytometer (BD Biosciences). The DNA index quantifies differences in total DNA in the test cell population in relation to that in normal diploid cells, comparing maximal G_0–G_1 peak heights between samples when assessed by FACS. A DNA index of 1.0 is indicative of normal diploid cells.

The zebrafish TCRB recombination assay was first described by Blackburn and colleagues (24). Briefly, RNA was extracted from

FACS-sorted T-ALL cells and made into complementary DNA, and PCR was performed utilizing each Vb and Cb primer (51 variable spanning PCR primers and 2 constant regions). A semi-nested PCR was completed using 1 μL of the PCR product and resolved on a 2% agarose gel. From this analysis, we can quantify the overall numbers of clones contained within the leukemia (24, 26).

qRT-PCR was performed on bulk leukemias ($n = 5$ fish/genotype) and compared with FACS-sorted *rag2:GFP*+ thymocytes ($n = 2$ samples). Samples were run in triplicate, with error bars representing the SEM of compiled data from all replicates and experimental samples.

Genomic DNA Sequencing of Zebrafish T-ALL

Zebrafish genomic DNA was extracted from T-ALL cells using a solution containing 10 mmol/L Tris-EDTA, 0.2% SDS, 200 mmol/L NaCl, and 200 μg/mL Proteinase K following incubation at 50°C for >16 hours. The next day, the Proteinase K was heat-inactivated at 95°C for 10 minutes. The gDNA extract was then purified using genomic DNA clean and concentrator (Zymo). Purified zebrafish gDNA was sheared and made into a library by the Massachusetts General Hospital next-generation sequencing core. Next-generation sequencing of zebrafish gDNA from MYC, MYC+TOX, and wild-type (WT) samples was performed using the Illumina HiSeq Instrument. Following alignment to the Zv9 genome using BWA (49), analysis of large-scale genomic variants was done using HMMCopy (50). In addition, a direct comparison of mutant and WT coverage was performed by calculating read coverage within nonoverlapping 10 Kb windows across the genome. The log–log scatter plots of the resulting read counts from leukemias were compared with the corresponding WT tissue isolated from the same fish.

Metaphase Spread on Mouse MEF Cells

MEF cells were transfected with control, full-length TOX, and dHMG box–containing TOX. After 48 hours of transfection, MEF cells were incubated with 100 ng/μL KaryoMax Colcemid for 90 minutes (Life Technologies), washed, and incubated for 18 minutes with 0.075 mol/L KCl. Note that 1 mL of a 3:1 methanol:glacial acetic acid (fixative solution) was added to the chromosome preparation and centrifuged for 10 minutes (400 g, 4°C). The pellet was washed 2 times with 3 mL of fixative solution and resuspended in 1 mL of fixative solution. Chromosomes were spread on microscope slides and dried 10 minutes at room temperature. DAPI Fluoromount-G (Southern Biotech) was added to each slide and imaged. At least 50 nuclei are counted per condition and per experiment. Experiments were replicated 3 independent times.

Gene Expression Analysis of TOX in Human T-ALL

TOX transcript expression was assessed in human cell lines using the Cancer Cell Line Encyclopedia (ref. 45; GSE36139) and in human T-ALL patient samples using the program R2: Genomics Analysis and Visualization Platform (http://r2.amc.nl). Datasets included normalized gene expression for healthy bone marrow (GSE13159) and human leukemia samples comprising the "Mixed Leukemia – MILE – 2004" (GSE13159) and "Tumor ALL (T) – Meijerink – 124" datasets (GSE26713). Statistical analysis and visualization were completed using Prism6.0. Whisker plots have been utilized to visualize the data and are represented using the Tukey method. Primary human samples used in this work were collected under the IRB DFCI 05-001.

Protein lysates were extracted using RIPA buffer, quantified using Coomassie Plus Bradford protein assay (Pierce), and analyzed by Western blot analysis. Specifically, denatured proteins were loaded on a 4% to 20% gradient gel (Bio-Rad) and then transferred to a PVDF membrane (Bio-Rad). Antibody staining was visualized using Amersham ECL Prime reagent (GE Healthcare) followed by analysis using a Fluor-S MultiImager (Bio-Rad). Band intensity was quantified using the ImageJ software package. Human T-ALL cells were assessed by Western blot analysis using primary antibodies for α-TOX (1:1,000 dilution; eBioscience), α-GAPDH (1:2,500, #2118; Cell Signaling Technology), α-TAL1 (1:200, BTL73; EMD Millipore), α-LMO2 (1:200, sc-65736; Santa Cruz Biotechnology), and α-cleaved NOTCH1 (1:1,000, #4147; Cell Signaling Technology). The secondary antibodies used were horseradish peroxidase (HRP)–conjugated anti-rat IgG (1:2,000, #7077; Cell Signaling Biotechnology), HRP-conjugated anti-rabbit IgG (1:2,500, #7074; Cell Signaling Biotechnology), and HRP-conjugated anti-mouse IgG (1:5,000, G21040; Life Technologies).

Mouse Xenografts and Luciferase Imaging

Human T-ALL cells were luciferized using the pLKO.1-*luc-mKate* (gift from Drs. Matthijssens and Van Vlierberghe, Ghent University, Belgium), and equal numbers of viable cells were injected into the flanks of 6-week-old *NOD/SCID/IL2rg*-null female mice (1×10^6 of viable cells per 200 μL). Mice were anesthetized by isoflurane, and leukemia growth was monitored by bioluminescence imaging following subcutaneous injection into the loose tissue over the neck of 75 mg/kg D-luciferin (Perkin Elmer) in 100 μL of PBS. Comparisons of leukemia size used an ANOVA test and Student *t* test comparisons between control and treated cells. Mouse experiments were approved under animal protocol 2013N000038 (Massachusetts General Hospital).

Chromatin Immunoprecipitation Sequencing Analysis

Reads were aligned to the hg19 version of the human genome using bowtie (51) with parameters –k2 –m 2 –best –sam and –l set to read length. Wiggle files to display the density of reads relative to genomic loci were created with MACS (52) with parameters –w –S –space=50 –nomodel –shiftsize=200 –keep-dup=1. The wiggle files were subsequently normalized to the millions of unique positions covered by reads. Wiggle files were displayed in the UCSC genome browser. Array file name used in this study is indicated in Supplementary Table S6.

Superenhancer Identification

Superenhancers were identified in two stages as described previously (34, 53). First, constituent enhancers were defined as peaks of H3K27ac using MACS with two sets of parameters to account for focal amplifications: -p 1e-9 –keep-dup=1 and –p 1e-9 –keep-dup=all. Second, the union of these sets of peaks was used as input for ROSE, which stitches proximal enhancer constituents and ranks stitched enhancers by signal. ROSE was run with parameters –s 12500 and –t 1000 to stitch constituents within 12.5 kb of each other and exclude constituents fully contained within ±1 kb from promoters of RefSeq genes.

Stitched enhancers were each assigned to the single expressed RefSeq gene locus whose transcription start site (TSS) is most proximal to the center of the stitched enhancer. Genes were considered expressed if they were in the top 2/3 of all TSS when ranking by promoter H3K27ac density. RPM-normalized promoter H3K27ac density was calculated in 1 kb windows center on TSSs using bamToGFF with parameters –e 200 –m 1 –r –d.

Immunoprecipitation and Western Blot Analysis

Immunoprecipitation was performed using 300 μg of protein lysate along with 4 μg of TOX antibody (GeneTex), KU70 antibody (ab3108; Abcam), or control rabbit IgG (Santa Cruz Biotechnology). Immunoprecipitation was completed in the presence of 30 to 40 μL of Dynabeads Protein G (Life Technologies) and incubated for 16 hours at 4°C in Cell Lysis Buffer (Cell Signaling Technology). Beads were then washed 5 times, and proteins were eluted using denaturating Laemmli, at 95°C for 10 minutes. Denatured proteins were loaded on a 4% to 20% gradient gel (Bio-Rad) for Western blot

analysis and then transferred to a PVDF membrane (Bio-Rad). Antibody staining was visualized using Amersham ECL Prime reagent (GE Healthcare) followed by analysis using a Fluor-S MultiImager (Bio-Rad). Band intensity was quantified by ImageJ. For DNase I treatment, EDTA-free lysis buffer was used to prepare cell lysates. $MgCl_2$ was added to the final concentration of 0.4 mmol/L. Note that 2.5 μL of 5 mg/mL DNase I were added to 500 μg of cell lysates. The digestion was performed by rotating at room temperature for 10 minutes. EDTA was then added to the final concentration of 5 mmol/L to stop the nuclease reaction.

Primary antibodies were α-TOX (1:1,000 dilution; eBioscience), α-KU70 (1:1,000, ab92450; Abcam), α-KU80 (1:1,000, #2753; Cell Signaling Technology), α-GAPDH (1:2,500, #2118; Cell Signaling Technology), α-XRCC4 (1:200, sc-365118; Santa Cruz Biotechnology), and α-LIGASE IV (1:200, sc-271299; Santa Cruz Biotechnology). The secondary antibodies used were the same as described above.

In Vitro *Binding Assay*

TOX and Δ*HMG* were subcloned into pT7CFE1-NHis-GST-CHA plasmid using the *BamHI* and *XhoI* restriction enzyme sites. Note that 4 μg of each construct was then transcribed and translated using the 1-Step Human High-Yield mini IVT kit (Pierce). TOX proteins were purified using the MagneGST Protein Purification System (Promega) and interactions with KU70/KU80 assessed on column. Specifically, 0.5 to 3 μg of KU70/KU80 proteins (Abcam) were added to TOX-bound beads. Following 2 hours of binding at 25°C, samples were washed 5 times using wash binding buffer supplied from the manufacturer (Pierce). TOX was cleaved from the beads using HRV3C protease (Pierce), purified away from beads, and the elution loaded on a 4%–20% gradient gel for Western blot analysis. *In vitro* binding assays were performed 3 independent times.

Cell Lines and Authentication

NIH3T3 and MEF cells were obtained from the ATCC in 2012 and 2015, respectively, and used within 3 months of receipt. Human T-ALL cell lines, MOLT-4, CCRF-CEM, HPB-ALL, Jurkat, DND-41, KE-37, KOPTK1, MOLT-13, P12-ICHIKAWA, and PEER, were a gift from A. Thomas Look in 2012. Human HEK293T cells were ordered from the ATCC in 2012. All human cell lines were authenticated at receipt and just prior to use in experiments using Small Tandem Repeat profiling and certified *Mycoplasma*-free (MycoAlert Plus, Lonza, tested every 6 months).

Cell Fractionation

Cell fractionation was completed essentially as described by the manufacturer (Subcellular Protein Fractionation Kit, Pierce; 1×10^7 human T-ALL cells). However, one additional wash was utilized in each cell fraction. Protein fractions were quantified by Bradford assay, and 20 to 45 μg of proteins were loaded on a 4% to 20% gradient gel for Western blot analysis. Primary antibodies were α-TOX, α-TUBULIN (1:500, ab4074; Abcam), α-SP1 (1:1,000, #9389; Cell Signaling Technology), and α-H4 (1:1,000, #2935; Cell Signaling Technology). The secondary antibodies used were the same as described previously. Cell fractionation was performed 3 times with similar results.

Transient Assays for NHEJ in 3T3 Cells and Human T-ALL Cells

For the NHEJ assay, DNA DSB was made in pEGFP-N1 vector or a pCS2-GW-mCherry by double digestion with *XhoI/BamHI* and *EcoRI/BamHI,* respectively, similar to previous published work (54). When 3T3 cells were 70% confluent, they were transfected with control or TOX-expressing vectors with or without KU70 and/or KU80. Cells were also transfected at the same time with linearized NHEJ reporter DNA into a 6-well plate (TransIT-X2 reagent; Mirus Bio). For nucleofection experiments in human T-ALL cells, 5×10^5 cells were nucleofected with control or TOX-expressing vectors along with linearized NHEJ reporter DNA. The Amaxa Cell Line Nucleofector Kit L was used for HPB-ALL and CCRF-CEM, whereas Kit V was used for MOLT-4 (program A-030; Lonza). Forty-eight and 72 hours after transfection/nucleofections, the cells were analyzed by flow cytometry counting the number of GFP or mCherry-positive cells and normalized to the number of fluorescent reporter cells. These assays were performed 3 independent times in triplicate.

Live Cell Imaging Following Microirradiation-Induced DNA Damage

Human *TOX* or the HMG box deletion mutants were gateway-cloned into the pDest26 expression vector (Life Technologies). mRuby was then subcloned in frame using *SalI* and *BglII*. pmRuby2, pmRuby2–TOX, or pmRuby2–ΔHMG constructs were cotransfected into 3T3 cells with KU80–GFP or XRCC4–GFP using electroporation (Neon transfection system; Life Technologies). Eight to 24 hours later, cells were treated with Hoescht 33342 and subjected to 405 nm laser-induced microirradiation and imaging as described previously (55). Images were processed and analyzed using Slidebook 5.0 software (Intelligent Imaging Innovations) and Adobe Photoshop CS6. Quantification of fluorescence intensity was measured within laser track regions and normalized to regions outside the track for each time point (>15 cells assessed for each time point).

Lentivirus Infection and Traffic Light Reporter Assay

Sequences for each shRNA are provided in Supplementary Table S6. Approximately 2 μg of pLKO–sh*TOX* or pLKO-scramble (Addgene plasmid 1864) were cotransfected into HEK 293T cells with 2 μg pCMV-dR8.91, 0.2 μg pVSV-g, and TransIT-LT1 reagent (Mirus Bio). Supernatants containing the lentivirus were collected, filtered, and added to T-ALL cell lines in the presence of 4 μg/mL polybrene (Millipore). T-ALL cells were spinoculated at 2,000 g for 90 minutes at 32°C. Two days after infection, cells were selected by adding 6 μg/mL puromycin (Invivogen). For the Traffic Light reporter assay, 2 μg pCVL-TLR-Ef1a-BFP (Addgene plasmid 31481; ref. 35) was cotransfected in addition to the previous plasmids. BFP-positive cells were FACS sorted on the Aria Flow Cytometer (BD Biosciences). BFP-positive, puromycin-selected cells were then nucleofected with tamoxifen-inducible I-SceI–expressing vector (gift from Dr. Mostoslavsky, Massachusetts General Hospital). Twenty-four hours after nucleofection, 200 nmol/L of tamoxifen was added to the cells and incubated for 48 hours before FACS analysis. These experiments were repeated in triplicate and repeated twice.

LC/MS-MS

Gel-separated samples were alkylated with iodoacetamide and digested overnight at pH 8.3 with sequencing grade trypsin (Promega). Peptide mixtures were analyzed by microcapillary LC/MS-MS using the EASY-nLC nanoflow HPLC (Thermo Fisher Scientific) with a 75 μm inner diameter × 15 cm length C18 capillary column coupled to a hybrid Orbitrap Elite mass spectrometer (Thermo Fisher Scientific). The Elite was operated in data-dependent acquisition mode (1 profile FT-MS spectrum followed by 15 centroided IT/MS-MS spectra). The resolution was 60,000 in FT/MS mode, and MS-MS spectra were read out at low resolution via CID in the Velos ion trap. The gradient consisted of 3% to 38% acetonitrile in 0.1% formic acid (FA) at a flow rate of 300 nL/min for 75 minutes, 38% to 95% acetonitrile in 0.1% FA for 2 minutes, and held at 95% acetonitrile in 0.1% FA at for 7 minutes followed by column re-equilibration for 10 minutes

at 3% acetonitrile in 0.1% FA. MS-MS fragmentation spectra were searched for protein identification using the Mascot search engine against the reversed and concatenated SwissProt protein database (v7_2012). Carbamidomethylation of Cys was set as fixed modification, and variable modifications were oxidation of Met and deamidation of Gln and Asn. Precursor ion mass tolerance was set to 12 ppm, and fragment ion mass tolerance was set to 0.8 Da. Two missed cleavages were allowed, and the minimal length required for a peptide was six amino acids. Relative quantification of proteins was achieved through peptide spectral counting using Scaffold 4 software. The peptide and protein FDRs were set to 1.5%.

To confirm the results from identifying proteins from selected protein bands from SDS-PAGE gels, the unfractionated sample of affinity enriched proteins was reduced and alkylated in solution basically as described previously (56). Proteins then were precipitated with chloroform-methanol, resuspended in 4 mol/L urea/50 mmol/L HEPES (pH 8.5), and digested using first endoproteinase LysC (Wako) and then sequencing grade trypsin (Promega). The generated peptides were analyzed by LC-MS2 in a 70-minute gradient on an Orbitrap Fusion mass spectrometer equipped with an EASY-nLC 1000 autosampler/HPLC pump. The mass spectrometer was operated in a data-dependent mode with a full MS spectrum acquired in the Orbitrap followed by MS2 spectra acquired in the linear ion trap on the most abundant ions detected in the full MS spectrum. MS2 spectra were assigned using a SEQUEST-based (57) proteomics analysis platform (58) by searching against the human Uniprot sequence database. Peptide and protein assignments were filtered to an FDR of <1 % employing the target-decoy database search strategy (59) and using linear discriminant analysis and posterior error histogram sorting. Peptides with sequences contained in more than one protein sequence from the database were assigned to the protein with most matching peptides (58). Interacting proteins were assigned by the following criteria: (1) at least 10 independent peptides must be found in the TOX pull-down samples, and (2) there must be at least a 3-fold enrichment over IgG pull-down controls.

Disclosure of Potential Conflicts of Interest

D. Bhere is a consultant/advisory board member for AMASA Technologies Inc. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: R. Lobbardi, J. Pinder, M.A. Oettinger, G. Dellaire, D.M. Langenau

Development of methodology: R. Lobbardi, M. Theodorou, Y. Namiki, A. Molodtsov, D. Toiber, M. Lion, J.M. Asara

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): R. Lobbardi, J. Pinder, B. Martinez-Pastor, M. Theodorou, J.S. Blackburn, Y. Namiki, A. Molodtsov, M. Lion, D. Bhere, K. Shah, A. Gutierrez, L.B. Silverman, J.M. Asara, W. Haas, R. Mostoslavsky, G. Dellaire

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): R. Lobbardi, J. Pinder, B. Martinez-Pastor, M. Theodorou, B.J. Abraham, Y. Namiki, M. Mansour, N.S. Abdelfattah, G. Alexe, E. Jain, M. Lion, D. Bhere, K. Shah, A. Gutierrez, K. Stegmaier, R.I. Sadreyev, J.M. Asara, M.A. Oettinger, R.A. Young, R. Mostoslavsky, G. Dellaire, D.M. Langenau

Writing, review, and/or revision of the manuscript: R. Lobbardi, J. Pinder, M. Theodorou, B.J. Abraham, M. Mansour, K. Stegmaier, L.B. Silverman, J.M. Asara, M.A. Oettinger, A.T. Look, R.A. Young, D.M. Langenau

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): R. Lobbardi, N.S. Abdelfattah, M. de Waard

Study supervision: R. Lobbardi, R.I. Sadreyev, G. Dellaire, D.M. Langenau

Other (experiments and cell line development): D. Toiber

Other (performing experiments): M. Theodorou, M. de Waard

Other (mass spectrometry experiment, sample processing): M. Boukhali

Acknowledgments

We would like to thank Dr. Ravi Mylvaganam for help with flow cytometry and Dr. Virginie Esain for manuscript advice and support.

Grant Support

This work was supported through funding provided to D.M. Langenau from the American Cancer Society, the Leukemia Research Foundation, the MGH Goodman Fellowship, the Live Like Bella Foundation, and the Alex's Lemonade Stand Foundation. Funding support was also garnered from the Harvey Graham Cancer Research Fund and The Terry Fox Foundation (J. Pinder), the Hope Funds for Cancer Research (B.J. Abraham), the William Lawrence and Blanche Hughes Foundation (A.T. Look), and the Boston Children's Hospital Translational Research Program (A. Gutierrez). We also acknowledge NIH grants 1S10OD010612 (J.M. Asara), 5 P01CA120964 (J.M. Asara), 1K99CA181500 (J.S. Blackburn), CA109901 (A.T. Look and R.A. Young), CA193651 (A. Gutierrez), and CA211734 (D.M. Langenau). G. Dellaire is funded by a Discovery Grant (RGPIN 05616) from the Natural Science and Engineering Research Council (NSERC).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received March 13, 2017; revised June 7, 2017; accepted September 7, 2017; published OnlineFirst October 3, 2017.

REFERENCES

1. Pui CH, Relling MV, Downing JR. Acute lymphoblastic leukemia. N Engl J Med 2004;350:1535–48.
2. Van Vlierberghe P, Ferrando A. The molecular basis of T cell acute lymphoblastic leukemia. J Clin Invest 2012;122:3398–406.
3. Ferrando AA, Neuberg DS, Staunton J, Loh ML, Huard C, Raimondi SC, et al. Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. Cancer Cell 2002;1:75–87.
4. Sanchez-Martin M, Ferrando A. The NOTCH1-MYC highway toward T-cell acute lymphoblastic leukemia. Blood 2017;129:1124–33.
5. Look AT.Oncogenic transcription factors in the human acute leukemias. Science 1997;278:1059–64.
6. Abraham BJ, Hnisz D, Weintraub AS, Kwiatkowski N, Li CH, Li Z, et al. Small genomic insertions form enhancers that misregulate oncogenes. Nat Commun 2017;8:14385.
7. Mansour MR, Abraham BJ, Anders L, Berezovskaya A, Gutierrez A, Durbin AD, et al. Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. Science 2014;346:1373–7.
8. Jeggo PA.Identification of genes involved in repair of DNA double-strand breaks in mammalian cells. Radiat Res 1998;150:S80–91.
9. Dietlein F, Thelen L, Reinhardt HC. Cancer-specific defects in DNA repair pathways as targets for personalized therapeutic approaches. Trends Genet 2014;30:326–39.
10. Bunting SF, Nussenzweig A. End-joining, translocations and cancer. Nat Rev Cancer 2013;13:443–54.
11. Gu Y, Seidl KJ, Rathbun GA, Zhu C, Manis JP, van der Stoep N, et al. Growth retardation and leaky SCID phenotype of Ku70-deficient mice. Immunity 1997;7:653–65.
12. Nussenzweig A, Chen C, da Costa Soares V, Sanchez M, Sokol K, Nussenzweig MC, et al. Requirement for Ku80 in growth and immunoglobulin V(D)J recombination. Nature 1996;382:551–5.

13. Wilkinson B, Chen JY, Han P, Rufner KM, Goularte OD, Kaye J. TOX: an HMG box protein implicated in the regulation of thymocyte selection. Nat Immunol 2002;3:272–80.
14. Aliahmad P, O'Flaherty E, Han P, Goularte OD, Wilkinson B, Satake M, et al. TOX provides a link between calcineurin activation and CD8 lineage commitment. J Exp Med 2004;199:1089–99.
15. Aliahmad P, Kaye J. Development of all CD4 T lineages requires nuclear factor TOX. J Exp Med 2008;205:245–56.
16. Aliahmad P, de la Torre B, Kaye J. Shared dependence on the DNA-binding factor TOX for the development of lymphoid tissue-inducer cell and NK cell lineages. Nat Immunol 2010;11:945–52.
17. Aliahmad P, Kadavallore A, de la Torre B, Kappes D, Kaye J. TOX is required for development of the CD4 T cell lineage gene program. J Immunol 2011;187:5931–40.
18. Seehus CR, Aliahmad P, de la Torre B, Iliev ID, Spurka L, Funari VA, et al. The development of innate lymphoid cells requires TOX-dependent generation of a common innate lymphoid cell progenitor. Nat Immunol 2015;16:599–608.
19. Yun S, Lee SH, Yoon SR, Kim MS, Piao ZH, Myung PK, et al. TOX regulates the differentiation of human natural killer cells from hematopoietic stem cells in vitro. Immunol Lett 2011;136:29–36.
20. Maser RS, Choudhury B, Campbell PJ, Feng B, Wong KK, Protopopov A, et al. Chromosomally unstable mouse tumours have genomic alterations similar to diverse human cancers. Nature 2007;447:966–71.
21. Langenau DM, Traver D, Ferrando AA, Kutok JL, Aster JC, Kanki JP, et al. Myc-induced T cell leukemia in transgenic zebrafish. Science 2003;299:887–90.
22. Langenau DM, Keefe MD, Storer NY, Jette CA, Smith AC, Ceol CJ, et al. Co-injection strategies to modify radiation sensitivity and tumor initiation in transgenic Zebrafish. Oncogene 2008;27:4242–8.
23. Smith AC, Raimondi AR, Salthouse CD, Ignatius MS, Blackburn JS, Mizgirev IV, et al. High-throughput cell transplantation establishes that tumor-initiating cells are abundant in zebrafish T-cell acute lymphoblastic leukemia. Blood 2010;115:3296–303.
24. Blackburn JS, Liu S, Raiser DM, Martinez SA, Feng H, Meeker ND, et al. Notch signaling expands a pre-malignant pool of T-cell acute lymphoblastic leukemia clones without affecting leukemia-propagating cell frequency. Leukemia 2012;26:2069–78.
25. Langenau DM, Feng H, Berghmans S, Kanki JP, Kutok JL, Look AT. Cre/lox-regulated transgenic zebrafish model with conditional myc-induced T cell acute lymphoblastic leukemia. Proc Natl Acad Sci U S A 2005;102:6068–73.
26. Blackburn JS, Liu S, Wilder JL, Dobrinski KP, Lobbardi R, Moore FE, et al. Clonal evolution enhances leukemia-propagating cell frequency in T cell acute lymphoblastic leukemia through Akt/mTORC1 pathway activation. Cancer Cell 2014;25:366–78.
27. Frazer JK, Batchelor LA, Bradley DF, Brown KH, Dobrinski KP, Lee C, et al. Genomic amplification of an endogenous retrovirus in zebrafish T-cell malignancies. Adv Hematol 2012;2012:627920.
28. Haupt Y, Bath ML, Harris AW, Adams JM. bmi-1 transgene induces lymphomas and collaborates with myc in tumorigenesis. Oncogene 1993;8:3161–4.
29. Giambra V, Jenkins CE, Lam SH, Hoofd C, Belmonte M, Wang X, et al. Leukemia stem cells in T-ALL require active Hif1alpha and Wnt signaling. Blood 2015;125:3917–27.
30. Treanor LM, Zhou S, Janke L, Churchman ML, Ma Z, Lu T, et al. Interleukin-7 receptor mutants initiate early T cell precursor leukemia in murine thymocyte progenitors with multipotent potential. J Exp Med 2014;211:701–13.
31. Zornig M, Schmidt T, Karsunky H, Grzeschiczek A, Moroy T. Zinc finger protein GFI-1 cooperates with myc and pim-1 in T-cell lymphomagenesis by reducing the requirements for IL-2. Oncogene 1996;12:1789–801.
32. Look AT, Melvin SL, Williams DL, Brodeur GM, Dahl GV, Kalwinsky DK, et al. Aneuploidy and percentage of S-phase cells determined by flow cytometry correlate with cell phenotype in childhood acute leukemia. Blood 1982;60:959–67.
33. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, et al. Super-enhancers in the control of cell identity and disease. Cell 2013;155:934–47.
34. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell 2013;153:307–19.
35. Certo MT, Ryu BY, Annis JE, Garibov M, Jarjour J, Rawlings DJ, et al. Tracking genome engineering outcome at individual DNA breakpoints. Nat Methods 2011;8:671–6.
36. Toiber D, Erdel F, Bouazoune K, Silberman DM, Zhong L, Mulligan P, et al. SIRT6 recruits SNF2H to DNA break sites, preventing genomic instability through chromatin remodeling. Mol Cell 2013;51:454–68.
37. Huang Y, Su MW, Jiang X, Zhou Y. Evidence of an oncogenic role of aberrant TOX activation in cutaneous T-cell lymphoma. Blood 2015;125:1435–43.
38. O'Flaherty E, Kaye J. TOX defines a conserved subfamily of HMG-box proteins. BMC Genomics 2003;4:13.
39. Ikushima H, Todo T, Ino Y, Takahashi M, Miyazawa K, Miyazono K. Autocrine TGF-beta signaling maintains tumorigenicity of glioma-initiating cells through Sry-related HMG-box factors. Cell Stem Cell 2009;5:504–14.
40. Stros M.HMGB proteins: interactions with DNA and chromatin. Biochim Biophys Acta 2010;1799:101–13.
41. Lange SS, Mitchell DL, Vasquez KM. High mobility group protein B1 enhances DNA repair and chromatin modification after DNA damage. Proc Natl Acad Sci U S A 2008;105:10320–5.
42. Prasad R, Liu Y, Deterding LJ, Poltoratsky VP, Kedar PS, Horton JK, et al. HMGB1 is a cofactor in mammalian base excision repair. Mol Cell 2007;27:829–41.
43. Yuan F, Gu L, Guo S, Wang C, Li GM. Evidence for involvement of HMGB1 protein in human DNA mismatch repair. J Biol Chem 2004;279:20935–40.
44. Stros M, Cherny D, Jovin TM. HMG1 protein stimulates DNA end joining by promoting association of DNA molecules via their ends. Eur J Biochem 2000;267:4088–97.
45. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 2012;483:603–7.
46. Tessema M, Yingling CM, Grimes MJ, Thomas CL, Liu Y, Leng S, et al. Differential epigenetic regulation of TOX subfamily high mobility group box genes in lung and breast cancers. PLoS One 2012;7:e34850.
47. Chung W, Kwabi-Addo B, Ittmann M, Jelinek J, Shen L, Yu Y, et al. Identification of novel tumor markers in prostate, colon and breast cancer by unbiased methylation profiling. PLoS One 2008;3:e2079.
48. Lamesch P, Li N, Milstein S, Fan C, Hao T, Szabo G, et al. hORFeome v3.1: a resource of human open reading frames representing over 10,000 human genes. Genomics 2007;89:307–15.
49. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754–60.
50. Ha G, Roth A, Lai D, Bashashati A, Ding J, Goya R, et al. Integrative analysis of genome-wide loss of heterozygosity and monoallelic expression at nucleotide resolution reveals disrupted pathways in triple-negative breast cancer. Genome Res 2012;22:1995–2007.
51. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 2009;10:R25.
52. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol 2008;9:R137.
53. Christensen CL, Kwiatkowski N, Abraham BJ, Carretero J, Al-Shahrour F, Zhang T, et al. Targeting transcriptional addictions in small cell lung cancer with a covalent CDK7 inhibitor. Cancer Cell 2014;26:909–22.

54. Seluanov A, Mittelman D, Pereira-Smith OM, Wilson JH, Gorbunova V. DNA end joining becomes less efficient and more error-prone during cellular senescence. Proc Natl Acad Sci U S A 2004;101:7624–9.
55. Andrin C, McDonald D, Attwood KM, Rodrigue A, Ghosh S, Mirzayans R, et al. A requirement for polymerized actin in DNA double-strand break repair. Nucleus 2012;3:384–95.
56. Edwards A, Haas W. Multiplexed quantitative proteomics for high-throughput comprehensive proteome comparisons of human cell lines. Methods Mol Biol 2016;1394:1–13.
57. Eng JK, McCormack AL, Yates JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom 1994;5:976–89.
58. Huttlin EL, Jedrychowski MP, Elias JE, Goswami T, Rad R, Beausoleil SA, et al. A tissue-specific atlas of mouse protein phosphorylation and expression. Cell 2010;143:1174–89.
59. Elias JE, Gygi SP. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat Methods 2007;4:207–14.