HHS Public Access

Author manuscript

Oncogene. Author manuscript; available in PMC 2016 July 08.

Published in final edited form as:

Oncogene. 2016 June 30; 35(26): 3454–3464. doi:10.1038/onc.2015.405.

Sleeping Beauty transposon screen identifies signaling modules that cooperate with STAT5 activation to induce B cell acute lymphoblastic leukemia

Lynn M. Heltemes-Harris[1,2,3], **Jon D. Larson**[3,5], **Timothy K. Starr**[3,4,5,6], **Gregory K. Hubbard**[1,2,3], **Aaron L. Sarver**[3], **David A. Largaespada**[3,6,7], and **Michael A. Farrar**[1,2,3]

[1]Center for Immunology, University of Minnesota, Minneapolis, MN

[2]Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, MN

[3]Masonic Cancer Center, University of Minnesota, Minneapolis, MN

[4]Department of Obstetrics, Gynecology & Women's Health, University of Minnesota, Minneapolis, MN

[5]Department of Genetics, Cell Biology, and Development, University of Minnesota, Minneapolis, MN

[6]Center for Genome Engineering, University of Minnesota, Minneapolis, MN

[7]Department of Pediatrics, University of Minnesota, Minneapolis, MN

Abstract

STAT5 activation occurs frequently in human progenitor B cell acute lymphoblastic leukemia (B-ALL). To identify gene alterations that cooperate with STAT5 activation to initiate leukemia we crossed mice expressing a constitutively active form of STAT5 (*Stat5b-CA*) to mice in which a mutagenic *Sleeping Beauty* transposon (*T2/Onc*) was mobilized only in B cells. *Stat5b-CA* mice typically do not develop B-ALL (<2% penetrance); in contrast, 89% of *Stat5b–CA* mice in which

CORRESPONDENCE: Michael A. Farrar, University of Minnesota, Center for Immunology, Masonic Cancer Center, 2101 6th St. SE, 2-116 WMBB, Minneapolis, MN 55455; farra005@umn.edu.

CONFLICT-OF-INTEREST DISCLOSURE: D.A.L. has ownership interest (including patents) in Discovery Genomics, Inc., and NeoClone Biotechnologies International and is also a consultant/advisory board member of Discovery Genomics, Inc., and NeoClone Biotechnologies International.

Authorship

Contribution: LMHH, DAL and MAF designed research; LMHH, JDL, TKS, and GKH performed experiments and analyzed data; ALS and MAF analyzed data, LMHH and MAF wrote the paper; and all authors critically reviewed and edited the paper.

URLs.

Gene centric CIS analysis, http://ias.eng.uiowa.edu/uploader/

Supplemental Table 1

Genomic regions containing transposon insertions and detailed information from TAPDANCE CIS analysis

Supplemental Table 2

Table containing CIS information from gCIS analysis.

Supplemental Table 3

Ingenuity pathway analysis of genes identified using gCIS (top table) or TAPDANCE (bottom table) analyses.

Supplemental Table 4

Reagents used in experiments.

the *T2/Onc* transposon had been mobilized died of B-ALL by 3 months of age. High-throughput sequencing approaches were used to identify genes frequently targeted by the *T2/Onc* transposon; these included *Sos1* (74%), *Kdm2a* (35%), *Jak1* (26%), *Bmi1* (19%), *Prdm14* or *Ncoa2* (13%), *Cdkn2a* (10%), *Ikzf1* (8%), *Caap1* (6%) and *Klf3* (6%). Collectively, these mutations target three major cellular processes: (i) the JAK/STAT5 pathway (ii) progenitor B cell differentiation and (iii) the CDKN2A tumor suppressor pathway. Transposon insertions typically resulted in altered expression of these genes, as well as downstream pathways including STAT5, ERK and p38. Importantly, expression of *Sos1* and *Kdm2a*, and activation of p38, correlated with survival, further underscoring the role these genes and associated pathways play in B-ALL.

Keywords

STAT5; Sleeping Beauty; Acute Lymphoblastic Leukemia

INTRODUCTION

B cell acute lymphoblastic leukemia (B-ALL) affects both children and adults. The survival rate in children is currently approximately 80–90%; however, relapsed B-ALL remains the most common cause of death due to cancer in children.[17, 35] Moreover, outcomes in infants and adults are still poor.[17] B-ALL is characterized by chromosomal rearrangements and gene mutations that perturb lymphoid development, cell proliferation, and epigenetic regulation.[31] However, discriminating between mutations in key driver genes versus passenger genes remains difficult. Recent studies have identified the transcription factor STAT5 as a key driver of B-ALL; STAT5 activation drives B-ALL in mouse models and increased STAT5 activation in human patient samples correlates with significantly worse outcomes.[15] A key question is what pathways cooperate with STAT5 in developing B cells to complete transformation. Previous studies demonstrated that targeted deletions in genes encoding the transcription factors EBF1 or PAX5, or the adaptor protein BLNK, cooperate with a weak constitutively active allele of STAT5 called *Stat5b*-CA to initiate transformation.[15, 32] However, it is unlikely that these are the only mechanisms that cooperate with STAT5 activation to initiate leukemia. Moreover, even *Stat5b-CAxEbf1*$^{+/-}$, *Stat5b-CAxPax5*$^{+/-}$ and *Stat5b-CAxBlnk*$^{+/-}$ mice are not born with leukemia,[15, 32] indicating that additional alterations are required for complete progenitor B cell transformation.

We utilized a *Sleeping Beauty* transposon mutagenesis screen to identify genetic alterations that cooperate with STAT5 activation to initiate transformation.[12, 13] Sleeping beauty mutagenesis screens have been useful in identifying cancer gene drivers in multiple cancers including colorectal cancer and T-ALL.[1, 3, 40] In our screen we targeted the mutagenic activity of *Sleeping Beauty* to progenitor B cells, which led to highly penetrant leukemia similar to human progenitor B-ALL. Using two bioinformatic pipelines we identified 65 gene targets that cooperate with STAT5 activation to promote progenitor B cell transformation. Subsequent analysis of this dataset identified three modules critical to progenitor B-ALL initiation including (1) disruption of B cell development, (2) enhanced JAK/STAT5 signaling, and (3) modification of the CDKN2A tumor suppressor pathway.

RESULTS

Sleeping Beauty mutagenesis screen identifies genetic alterations that cooperate with STAT5 to initiate transformation

To determine what genes cooperate with STAT5 to induce leukemia we employed sleeping beauty mutagenesis in a strain of mice that expresses a weak constitutively activated form of STAT5b (referred to as *Stat5b-CA* mice). *Stat5b-CA* mice were crossed to *Cd79a-Cre* mice, which express CRE recombinase only in developing B cells.[16] *Stat5b-CAxCd79a-Cre mice* were then crossed to mice homozygous for a concatamer of mutagenic SB transposon vectors (T2/Onc) on chromosome 1 or 15 and a Cre-inducible SB transposase gene (Rosa26LSL-SB11)[13](*T2/OncxRosa26*$^{LSL\text{-}SB11}$ combination referred to as *SB* hereafter). The Rosa26$^{LSL\text{-}SB11}$ transposase transgene also contains a *Gfp* cDNA that is removed upon CRE-mediated recombination and therefore allows identification of cells in which CRE recombinase is active and the *SB11* transposase enzyme expressed. Mobilization of the *T2/Onc* transposon in *Stat5b-CAxCd79a-CrexSB* mice was efficient and specific as demonstrated by loss of GFP expression in progenitor B cells (Fig. 1a). We generated 80 *Stat5b-CAxCd79a-CrexSB* mice and monitored them for illness; 71 of these mice (89%) developed leukemia with a median survival of 81 days (Fig. 1b). The mice presented with enlarged lymph nodes and spleen (Fig. 1c). These leukemias were characterized as progenitor B cells as they typically expressed CD19, B220, CD43 and BP-1 but lacked expression of IgM (Fig. 1d). We also generated 31 *Cd79a-CrexSB* mice as controls. These mice mobilize the transposon in developing B cells in the absence of constitutively active STAT5b. As seen for *Stat5b-CAxCd79a-CrexSB* mice, mobilization of the T2/Onc transposon was induced (Fig. 1a) but none of these mice developed leukemia (Fig. 1b).

To identify transposon **C**ommon **I**nsertion **S**ites (CISs) that cooperate with STAT5 to initiate transformation we used a PCR-based approach to amplify *T2/Onc* insertion sites followed by high throughput sequencing. We used both TAPDANCE and gene-centric (gCIS) analyses to find CIS genetic loci in 65 of our leukemias.[4, 36] The high confidence CIS identified by the TAPDANCE analysis included ten annotated genes and two regions of the genome that lack any annotated gene feature within 20 kB. The *Sos1* gene was the most frequent gene target with integrations identified in over 70% of the leukemias. Additional high-confidence targets included *Kdm2a, Jak1*, *Bmi1*, *Cdkn2a, Ikzf1*, *Prdm14/Ncoa2*, *Snora17, Caap1* and *Klf3* (Table 1 and Supplemental Table 1). The gene-centric CIS (gCIS) analysis identified a total of 65 genes, including all of the annotated genes identified in the TAPDANCE analysis (Table 2 and Supplemental Table 2). Ingenuity Core pathway analysis of genes identified by TAPDANCE or gCIS analysis established that these target genes were involved in cancer, hematopoiesis and differentiation (Supplemental Table 3). Thus, our analysis has revealed a number of genes that cooperate, either as tumor suppressors or oncogenes, with STAT5 activation to initiate progenitor B cell leukemia.

JAK/STAT5 signaling is uniformly increased in SB-initiated leukemias

The genes identified by our SB screen segregate into a number of distinct modules. One of the most surprising targets from our screen was the gene encoding the tyrosine kinase *Jak1*. Despite increased STAT5 activity from a constitutively active STAT5b transgene, the JAK/

STAT5 signaling pathway was one of the most frequently targeted pathways in our sleeping beauty screen (Table I & II). Sequence mapping results suggest that the *T2/Onc* transposon insertions drive expression of *Jak1*, as the insertions were found almost exclusively in the first intron of JAK1 and were oriented in a direction that should drive *Jak1* transcription (Fig. 2a). To determine if JAK1 expression levels were increased, we carried out immunoblotting studies. JAK1 levels were significantly higher in all leukemic mice, not just those with *T2/Onc* insertions in the *Jak1* gene locus (Fig. 2b). Consistent with this observation we found that STAT5 tyrosine phosphorylation (and hence activation) was also increased in almost all samples, not just those with *T2/Onc* insertions in *Jak1* (Fig 2c,d,f). Similar to the protein expression, *Jak1* mRNA transcripts from leukemic samples were also expressed at elevated levels when compared to non-leukemic *Stat5b-CA* control samples (Fig. 2e). However, once again no difference in *Jak1* gene expression was seen when comparing samples harboring a transposon insertion in *Jak1* and those with no insertion (Fig. 2e). Thus, while a *Stat5b-CA* allele was present in all of our leukemias, JAK/STAT5 signaling was further augmented in the vast majority of leukemias from our screen.

Frequent targeting of genes and pathways involved in pre-BCR differentiation

We identified *T2/Onc* insertions in several genes potentially involved in pre-BCR signaling, including *Sos1*, *Ikzf1* and *Klf3*. *Sos1* was the most commonly targeted gene in our SB mutagenesis screen. *T2/Onc* integrations in the *Sos1* locus clustered between exons 6 and 10 and were oriented in the same direction, suggesting transposon-mediated expression of an N-terminally truncated form of SOS1 (Fig. 3a). Confirming this hypothesis we identified transcripts that contained 5' sequence originating from *T2/Onc* that spliced into downstream exonic sequence from *Sos1* (data not shown). qRT-PCR analysis of *Sos1* expression from the 5' and 3' ends of the *Sos1* gene only showed increased expression when using primers near the 3' end of the gene (Fig. 3b). Furthermore, immunoblotting studies showed selective expression of a truncated form of *Sos1* in leukemias with *T2/Onc* integrations in the *Sos1* locus (Fig. 3c). Importantly, the amount of 3' *Sos1* transcripts correlated with reduced overall survival in our model. In contrast, 5' *Sos1* transcripts did not correlate with survival (Fig. 3d). Thus, *T2/Onc* insertions into the *Sos1* locus result in a truncated *Sos1* transcript and production of a truncated form of SOS1 whose expression correlated inversely with overall survival.

To study the effects of this truncated form of SOS1 on potential downstream pathways, we examined a number of downstream targets of the SOS1/RAS signaling cascade in *Stat5b-CAxCd79a-CrexSB* leukemia's, including ERK, S6K, AKT and p38. ERK phosphorylation levels were significantly reduced compared to that seen in *Stat5b-CA* B cells (Fig. 4a and c). We next compared the level of pERK in samples with or without transposon integration in the *Sos1* locus. In contrast, p38 phosphorylation levels (pp38) were significantly increased (Fig. 4b and c). Other downstream targets, such as S6K or Akt, were not notably altered (data not shown). To investigate if these changes correlated with overall survival in our model we plotted both pERK and pp38 levels versus overall survival. pERK levels were relatively uniformly reduced relative to *Stat5b-CA* control progenitor B cells and the modest differences in pERK levels did not correlate with decreased survival. In contrast, pp38 levels varied over a greater range and had a more complex relationship with survival. *Stat5b-*

CAxCd79a-CrexSB leukemias with intermediate levels of pp38 appeared to have the shortest survival compared to those with very low or very high levels of pp38 (Fig. 4d,e). Finally, total p38 levels directly correlated with survival (Fig. 4f).

The gene encoding the transcription factor *Ikzf1* was also targeted in our screen. IKZF1 is required for normal pre-B cell differentiation [37], and is commonly deleted in human B-ALL [30]. *T2/Onc* insertions were scattered throughout the *Ikzf1* gene and were oriented in both directions suggesting that the *T2/Onc* insertions lead to loss-of-function mutations in *Ikzf1* (Fig. 5a). Consistent with the insertion data, we observed that *Ikzf1* transcripts in *Stat5b-CAxCd79a-CrexSB* leukemias were significantly reduced relative to *Stat5b-CA* control progenitor B cells. However, the reduction in *Ikzf1* transcripts was also observed in *Stat5b-CAxCd79a-CrexSB* leukemias without a *T2/Onc* insertion in the *Ikzf1* gene (Fig. 5b) suggesting strong selective pressure to reduce *Ikzf1* transcript levels via multiple mechanisms.

IKZF1 is a member of a transcription factor network including PAX5 and EBF1 that collectively regulate B cell development. Moreover, deletions in *IKZF1* and *PAX5* frequently occur together in human progenitor B-ALL.[27, 28] Therefore, we examined whether *Pax5* transcripts were also reduced in *Stat5b-CAxCd79a-CrexSB*-initiated leukemias. As shown in figure 5c, *Pax5* levels were significantly reduced in all *Stat5b-CAxCd79a-CrexSB* leukemias. Expression levels of *Ikzf1* and *Pax5* correlated with each other (Fig. 5d), providing additional evidence that they functioned as part of a coherent signaling network. Finally, both *Ikzf1* and *Pax5* transcripts were only ~2-fold lower than in *Stat5b-CA* control progenitor B cells, suggesting that reduced but not absent expression of these transcription factors is important for progenitor B cell transformation.

KDM2A upregulation occurs frequently in SB-driven leukemias

The second most common target in our *Stat5b-CAxCd79a-CrexSB* mutagenesis screen was *Kdm2a*. KDM2A is a H3K36 demethylase and plays a global role in altering gene transcription.[8] More recently, KDM2A has been shown to demethylate NFκB at K218/221 and thereby reduce NFκB function.[24, 25] *T2/Onc* insertions were all located at the 5' end of the *Kdm2a* gene and largely oriented in the same direction, suggesting they should drive transcription (Fig 6a). Consistent with this idea, we observed that *Kdm2a* transcripts were significantly increased in *Stat5b-CAxCd79a-CrexSB* leukemias with *T2/Onc* insertions in the *Kdm2a* locus when compared to *Stat5b-CAxCd79a-CrexSB* leukemias that lacked such insertions in *Kdm2a* (Fig. 6b). Importantly, we observed that the amount of *Kdm2a* transcripts correlated inversely with overall survival (Fig. 6c). Thus, overexpression of *Kdm2a* was frequently observed in *Stat5b-CAxCd79a-CrexSB*-driven leukemias and the degree of overexpression correlated with reduced survival.

T2/Onc insertions that target the *Cdkn2a* locus occur frequently

The final major module identified in our *Stat5b-CAxCd79a-CrexSB* screen involves the tumor suppressor gene *Cdkn2a*. *Cdkn2a* is frequently deleted in human B-ALL[29, 44] and was targeted directly in 10% of our leukemias. The *T2/Onc* insertions in the *Cdkn2a* locus are distributed throughout the gene and found in both orientations suggesting that this should

disrupt the *Cdkn2a* allele (Fig. 7a). In fact, *Cdkn2a* levels in fully transformed *Stat5b-CAxCd79a-CrexSB* leukemias were much higher than that observed in WT and *Stat5b-CA* control progenitor B cells (Fig. 7b), suggesting a negative feedback loop in which genetic alterations that promote full transformation feed back on *Cdkn2a*, and induce higher abundance of *Cdkn2a* transcript. Further, although *Cdkn2a* has not been described as a haploinsufficient tumor suppressor we typically did not observe complete loss of *Cdkn2a* transcripts in *Stat5b-CAxCd79a-CrexSB* leukemias; this was true for *Cdkn2a* transcripts encoding both p19ARF and p16INK4a (data not shown).

In addition to direct targeting of the *Cdkn2a* locus by *T2/Onc*, other mechanisms could also affect the *Cdkn2a* locus. For example, previous studies have shown that BMI1 overexpression reduces *Cdkn2a* levels.[20] Along these lines, we observed *T2/Onc* insertions in the *Bmi1* gene in 19% of SB-driven leukemias. These insertions sites were all oriented in the same direction and localized at the 5' end of the gene suggesting that they increase *Bmi1* transcription (Fig. 7a). When assessed by qRT-PCR, *Bmi1* transcripts were significantly increased in *Stat5b-CAxCd79a-CrexSB* leukemias with *T2/Onc* insertions in *Bmi1* when compared to *Stat5b-CAxCd79a-CrexSB* leukemias that lacked *T2/Onc* insertions in the *Bmi1* gene locus (Fig. 7c). Consistent with the idea that CDKN2A and BMI1 might mediate similar functions, *T2/Onc* insertions in *Bmi1* and *Cdkn2a* were mutually exclusive. Finally we compared the level of *Cdkn2a* transcripts in *Stat5b-CAxCd79a-CrexSB* leukemias with or without *T2/Onc* insertions in the *Bmi1* gene locus (*Stat5b-CAxCd79a-CrexSB* leukemias with *T2/Onc* insertions in *Cdkn2a* were excluded from this analysis) (Fig. 7d). Although the mean and median levels did not differ significantly between these two populations, the variance between populations was significantly different ($p<0.0001$); *Stat5b-CAxCd79a-CrexSB* leukemias with *Bmi1* insertions exhibited a much narrower range of *Cdkn2a* expression levels and lacked the high levels of *Cdkn2a* that were occasionally seen in *Stat5b-CAxCd79a-CrexSB* leukemias that lacked *T2/Onc* insertions in *Bmi1*. Thus, increased *Bmi1* expression may act to limit hyperinduction of the *Cdkn2a* tumor suppressor gene.

DISCUSSION

In this study we carried out an unbiased sleeping beauty transposon mutagenesis screen to identify genes and signaling pathways that cooperate with STAT5 activation to initiate progenitor B-ALL. This screen identified a large number of genes that likely cooperate with STAT5 to drive transformation. Although the exact mechanism by which all these SB-identified gene targets cooperate with STAT5 is not known, we have identified three primary signaling modules that affect transformation. First, despite the presence of a constitutively active STAT5 driver mutation in our screen, we observed that transformation in this model almost always involved increased JAK1 protein and increased STAT5 phosphorylation (and hence activation). Interestingly, this occurred in all *Stat5b-CAxCd79a-CrexSB* leukemias, not just those with *T2/Onc* insertions in the *Jak1* locus. This result suggests that there is strong selective pressure to increase JAK1 levels, either via direct *T2/Onc*-induced upregulation, or via alternative pathways. Importantly, these findings demonstrate that transformation of progenitor B cells requires very high levels of JAK/STAT5 activation and

suggest that even partial inhibition of JAK/STAT5 signaling may be sufficient to limit leukemia cell growth.

The second module we identified involves gene targets with known roles in pre-BCR signaling. The most common target gene in this pre-BCR signaling module was *Sos1*. Integration of *T2/Onc* into the *Sos1* locus resulted in the expression of a truncated SOS1 protein lacking its N-terminal H and DBL-homology domains. Western blot experiments demonstrated reduced ERK phosphorylation but increased p38 phosphorylation levels when examining all of our leukemic samples. Overexpression of truncated SOS1 reduced the level of pERK further. One potential explanation for this result is that efficient ERK activation requires allosteric interaction of Ras-GTP with SOS1.[21] This interaction requires the Dbl-homology domain of SOS1,[21] which is missing in our truncated form of SOS1. In contrast, p38 activation by SOS1 does not require this Dbl-homology domain.[21] Based on this model, expression of our truncated SOS1 should increase p38 activation but inhibit ERK activation. Such a result fits with studies that have shown that ERK activation is required for progenitor B cell differentiation and thus truncated SOS1 could result in a block in B cell differentiation at the highly proliferative progenitor B cell stage.[26] The effect of p38 activation on progenitor-B cells has not been as well characterized. However, ectopic p38 activation in pro-T cells leads to a differentiation block and an increase in cells in the S/G2/M phase of the cell cycle,[11] suggesting a similar effect might occur in progenitor-B cells. Finally, it is possible that SOS1 also increases STAT5b activity as previous studies have shown that RAS activation in hematopoietic stem cells can lead to increased STAT5 activation.[23] Taken together, expression of a truncated form of SOS1 correlates with changes in expression or activation of several downstream targets that are predicted to inhibit pre-B cell differentiation.

Mutations in the Ras signaling pathway are frequent at ALL diagnosis with the highest incidence of these mutations being found in the high hyperdiploidy group.[46] More recently, Ras mutations have been identified in “high-risk” ALL and relapsed ALL patients.[18] These mutations commonly activate the Ras pathway as demonstrated by increased pErk activity. In contrast to the studies identifying Ras activation and its role in ALL, Shojaee and colleagues found increases in negative feedback regulation of Erk signaling in pre-B ALL.[38] In fact, they found that reduced ERK levels were required for permissive oncogenic signaling. This study fits with our data where we see decreased levels of ERK phosphorylation in pre-B cells. Finally, there are differences in Ras signaling during different stages of early B cell development and the Ras signal is dependent on the pre-BCR.[26] Mandal and colleagues demonstrated that RAS signaling was at its peak during the large Pre-B cell stage of differentiation. This suggests that differences in Ras signaling could be dependent on the stage of the progenitor B cell that transforms into leukemia. Therefore, Ras signaling is more complicated than originally suggested and more research is required to determine what mechanisms are required for leukemic transformation.

An important part of our studies is the observation that expression levels of truncated *Sos1* and pp38 correlated with overall survival in our model. This suggests that SOS1 inhibitors[14] may be useful in treating forms of B-ALL with amplified SOS1 expression. The relationship between p38 activation and survival in our model is more complex. Two- to three-fold

increases in p38 phosphorylation were linked to decreased overall survival. In contrast, higher levels of p38 activation were linked to longer survival. The mechanism that accounts for this biphasic response is unclear. However, p38 activation has been linked to oncogene-induced senescence[5, 19, 43] and thus the results we see may reflect this behavior - modest increases in p38 activation may block progenitor B cell differentiation and enhance proliferation while higher increases may initiate senescence. This hypothesis is consistent with recent work by Muschen and colleagues who demonstrated that hyper-activation of pre-BCR signals could activate the pathway that normally mediates negative selection of autoreactive B cells.[7] In that study increasing pre-BCR signaling above a threshold that activates the negative selection checkpoint in leukemic pre-B cells led to apoptotic cell death. This idea illustrates the importance of signaling thresholds and that changes that result in either too high or low levels of critical cellular signals may influence leukemic cell survival.

In addition to effects on SOS1, multiple other gene targets identified in our SB screen could affect progenitor B cell differentiation. These include KDM2A, whose overexpression has been shown to reduce NFκB-dependent signals and therefore could inhibit NFκB-dependent progenitor B cell differentiation.[24, 25] In addition, our gene centric CIS analysis identified *T2/Onc* insertions in several other genes that should also interfere with progenitor B cell differentiation. These include T2/Onc insertions that should drive *Ptpn6* expression, which encodes the phosphatase SHP1 that has been shown to block BCR-dependent signals[10], and the kinase MAP3K8, which plays an important role in both ERK and NFκB activation.[41] The *T2/Onc* integration sites we identified in *Map3k8* should disrupt gene function, suggesting that this modification could contribute to both reduced ERK and NFκB activity observed in *Stat5b-CAxSB* leukemias. Importantly, MAP3K8 levels are reduced in human B-ALL[41], suggesting this may be an important modification to promote transformation.

The transcription factors IKZF1 and KLF3 both play important roles in B cell development. IKZF1 is required for normal pre-B cell differentiation.[37] Less is known about the role of KLF3, although mice defective in KLF3 do have defects in B cell development.[42] Importantly, *IKZF1* deletions are frequently observed in human leukemia[30]; our data provide independent evidence that this alteration is in fact a key driver of B-ALL. In addition, the fact that we observed on average a 2-fold decrease in IKZF1 as opposed to complete loss of IKZF1 suggests that reduced, but not absent IKZF1, is critical for transformation. This idea is consistent with the observation that human leukemias frequently lack one but not both alleles of *IKZF1*.[27] Taken together, *T2/Onc* insertions in genes affecting pre-B cell differentiation were observed in 95% of our *Stat5b-CAxCd79a-CrexSB* leukemias.

The third module we identified involves the tumor suppressor CDKN2A. *Cdkn2a* encodes ARF and INK4a, which target the p53 and RB tumor suppressor pathways, respectively. We observed that *Cdkn2a* was targeted in 10% of *Stat5b-CAxCd79a-CrexSB* leukemias. In addition, BMI1, which inhibits expression of *Cdkn2a*,[20] was targeted in an additional 19% of *Stat5b-CAxCd79a-CrexSB* leukemias and appeared to limit high-level *Cdkn2a* transcription. Importantly, since *Cdkn2a* was typically not eliminated in *Stat5b-CAxCd79a-CrexSB* leukemias with *T2/Onc* insertions in *Cdkn2a* (or *Bmi1*), this suggests that *Cdkn2a* can act as a haploinsufficient tumor suppressor. Taken together, 29% of *Stat5b-CAxCd79a-*

CrexSB leukemias had *T2/Onc* insertions that should affect *Cdkn2a*. This correlates well with human B-ALL in which the *Cdkn2a* locus is deleted ~40% of the time.[27]

The polycomb family protein BMI1 plays a critical role in maintaining self-renewal capacity of both normal and leukemic cells.[22, 34] In addition, a recent genome-wide association study identified SNP variants in a region linked to *BMI1-PIP4K2* that were associated with B-ALL susceptibility.[45] This same study also identified *IKZF1* and *CDKN2A* susceptibility variants associated with childhood ALL, genes that have been well studied in leukemia. This underscores the importance of the *BMI1* gene and further validates an additional target of our screen. In related studies, Oguro and colleagues demonstrated the importance of BMI1 in lineage specification and more specifically in regulating EBF1 and PAX5.[33] In this study, they found that loss of BMI1 resulted in the premature expression of PAX5 and EBF1 in hematopoietic stem cells, suggesting that BMI1 is capable of regulating PAX5 and EBF1 expression. Further, a recent report demonstrated the ability of BMI1 to regulate *Ikzf1* expression.[2] This suggests that increases in BMI1 expression could lead to downregulation of *Ikzf1* expression thereby interfering with normal pre-B cell differentiation. Previous links with BMI1 and hematological malignancies[22] have been demonstrated but the mechanism by which BMI1 drove transformation was not well established. Taken together, these findings suggest that one possible mechanism by which BMI1 promotes leukemic transformation may be through down regulation of IKZF1 expression.[2] Therefore, BMI1 plays an important role in two modules targeted in our screen, B cell development and CDKN2A tumor suppressor pathway; this is consistent with the observation that *Bmi1* was targeted twice as frequently as either *Ikzf1* or *Cdkn2a* alone in our screen.

In conclusion, our SB screen identified four highly penetrant features of B-ALL. First, JAK/STAT5 signals were uniformly elevated in all leukemias, suggesting that this is a critical step for transformation. Second, reduced expression of IKAROS and PAX5 transcription factors was also a consistent feature of B-ALL. Third, disruption of pathways involved in pre-BCR signaling occurred in virtually all B-ALL. Fourth, perturbation of the CDKN2A pathway occurred in a substantial fraction of B-ALL, suggesting that effects on this pathway can significantly contribute to transformation. Finally, our observation that Sos1, *Kdm2a* and pp38 levels correlate with survival suggests that inhibitors of these enzymes may be useful for treating B-ALL.

MATERIALS AND METHODS

Mice and Cells

Stat5b-CA mice have been previously described.[6] *Cd79a-Cre* mice were provided by Drs. Tim Bender and Michel Reth.[16] The two strains of *T2/Onc* transgenic mice (*T2/Onc(chr1)* and *T2/Onc(chr15)*) were previously described.[9] The Cre-inducible SB transposase ($Rosa26^{LSL\text{-}SB11}$) mouse was described previously.[13] Mice were monitored for up to 400 days for leukemia. Lymph nodes were isolated from leukemia-bearing mice and used for further experiments. The University of Minnesota IACUC approved all animal experiments.

CD19+ control cells were purified from 6–12 week old *Stat5b-CA* and C57BL/6 mice using anti-CD19 microbeads (Miltenyi) as described by the manufacturer. Purification of

progenitor B cells was done using a two-step process. Cells were labeled with FITC-labeled α-Igλ (Southern Biotech), α-Igκ (Southern Biotech), α-Ly6G (GR-1, eBioscience) and α-Ter119 (eBioscience) followed by labeling with anti-FITC microbeads (Miltenyi) to deplete mature and non-B cells. Cells were labeled with anti-CD19 microbeads (Miltenyi) and cells bound to the column were collected and used in further experiments.

Transposon Insertion Analysis

Genomic DNA was isolated using the HighPure DNA kit (Roche). DNA was subjected to linker-mediated PCR as previously described,[40] except that primer sequences were changed to include 12 bp barcodes and Illumina HiSeq 2000 platform-specific sequences (sequences available upon request). PCR amplicons were subjected to sequencing using the Illumina HiSeq 2000 platform following manufacturer's protocol.

Sequences were processed and analyzed using the Transposon Annotation Poisson Distribution Association Network Connectivity Environment (TAPDANCE)[36] and gene-centric Common Insertion Site (CIS) analysis software.[4]

Flow Cytometry

Single cell suspensions were prepared from the lymph nodes and stained with antibodies listed in Supplemental Table 4 and run on a LSRII flow cytometer (BD Biosciences); data was analyzed using FlowJo software (Treestar).

Western Blot

Protein samples were isolated from 50×10^6 leukemic cells or *Stat5b-CA* CD19+ splenic B cells using RIPA buffer (Thermo Scientific) with HALT protease and phosphatase inhibitors (ThermoScientific) following standard procedures. Protein concentration was determined using a Bradford Assay (Sigma). The gels were run according to the BioRad Criterian instruction manual and application guide with 50 ug of protein per well. Reagents are listed in Supplemental Table 4. Blots were visualized on a LiCor Odyssey (Li-Cor Biosciences) and analyzed using Image Studio Lite software (Li-Cor Biosciences).

qRT PCR Analysis

RNA was extracted from 50×10^6 leukemic cells or progenitor B cells purified from Stat5b-CA or C57BL/6 mice using an RNeasy Mini kit (Qiagen) following the recommended procedures. RNA concentrations were determined in a Nanodrop (ThermoScientific). 250 ng of RNA was converted to cDNA with the ABI High-Capacity cDNA Reverse Transcription Kit (#4368814) according to the manufacturer's instructions. Gene specific primers are listed in Supplemental Table 4. PCR was carried out in an ABI 7900HT system in triplicate using the iTaq Universal SYBR Green (Bio-Rad) in 20 ul reactions containing 200 nm of each forward and reverse primer, 2 ul of cDNA diluted mix (~10 ng) and 10 ul of 2X SYBR Master Mix.

Data Analysis

Number of mice used was based on previous work and suggested that 60 mice would be sufficient to identify CIS that showed up in 5% of leukemias.[39] Mice with SB transposition

were generated in a random manner; investigators were not blinded. Data was analyzed using Prism 6 (Graphpad). Data were not normally distributed so they were analyzed using either MannU-Whitney or Kruskal-Wallace non-parametric tests. Kaplan-Meier Survival curves were analyzed by Log-rank (Mantel-Cox) Test. The pp38 data in figure 4D was fit to a second order smoothing polynomial curve. Ingenuity Pathway Analysis was done using standard settings for a Core Analysis. Integrated Genomics Viewer was used for mapping of sequences (Broad Institute).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors thank Peter Schoettler, Chris Reis, Alyssa Kne, Amy Mack, and Emilea Sykes for assistance with animal husbandry, Maya Raghunandan for technical assistance, Paul Champoux for assistance with Flow cytometry and Dr. Richard Williams for helpful comments.

FINANCIAL SUPPORT: This work was supported by a Brainstorm grant from the University of Minnesota Masonic Cancer Center (MAF and DAL), NIH grants R01 CA154998 and CA151845 to MAF, NIH grant 5R00CA151672-04 to TKS, NIH R01 CA113636 to DAL, an NIH Institutional Shared Resource grant to the Masonic Cancer Center P30-CA77598, and a Leukemia and Lymphoma Society Scholar award (MAF).

REFERENCES

1. Abbott KL, Nyre ET, Abrahante J, Ho YY, Isaksson Vogel R, Starr TK. The Candidate Cancer Gene Database: a database of cancer driver genes from forward genetic screens in mice. Nucleic acids research. 2015; 43:D844–D848. [PubMed: 25190456]
2. Arranz L, Herrera-Merchan A, Ligos JM, de Molina A, Dominguez O, Gonzalez S. Bmi1 is critical to prevent Ikaros-mediated lymphoid priming in hematopoietic stem cells. Cell Cycle. 2012; 11:65–78. [PubMed: 22185780]
3. Berquam-Vrieze KE, Nannapaneni K, Brett BT, Holmfeldt L, Ma J, Zagorodna O, et al. Cell of origin strongly influences genetic selection in a mouse model of T-ALL. Blood. 2011; 118:4646–4656. [PubMed: 21828136]
4. Brett BT, Berquam-Vrieze KE, Nannapaneni K, Huang J, Scheetz TE, Dupuy AJ. Novel molecular and computational methods improve the accuracy of insertion site analysis in Sleeping Beauty-induced tumors. PloS one. 2011; 6:e24668. [PubMed: 21931803]
5. Bulavin DV, Kovalsky O, Hollander MC, Fornace AJ Jr. Loss of oncogenic H-ras-induced cell cycle arrest and p38 mitogen-activated protein kinase activation by disruption of Gadd45a. Molecular and cellular biology. 2003; 23:3859–3871. [PubMed: 12748288]
6. Burchill MA, Goetz CA, Prlic M, O'Neil JJ, Harmon IR, Bensinger SJ, et al. Distinct effects of STAT5 activation on CD4+ and CD8+ T cell homeostasis: development of CD4+CD25+ regulatory T cells versus CD8+ memory T cells. Journal of immunology. 2003; 171:5853–5864.
7. Chen Z, Shojaee S, Buchner M, Geng H, Lee JW, Klemm L, et al. Signalling thresholds and negative B-cell selection in acute lymphoblastic leukaemia. Nature. 2015; 521:357–361. [PubMed: 25799995]
8. Cheng Z, Cheung P, Kuo AJ, Yukl ET, Wilmot CM, Gozani O, et al. A molecular threading mechanism underlies Jumonji lysine demethylase KDM2A regulation of methylated H3K36. Genes & development. 2014; 28:1758–1771. [PubMed: 25128496]
9. Collier LS, Carlson CM, Ravimohan S, Dupuy AJ, Largaespada DA. Cancer gene discovery in solid tumours using transposon-based somatic mutagenesis in the mouse. Nature. 2005; 436:272–276. [PubMed: 16015333]

10. Cornall RJ, Cyster JG, Hibbs ML, Dunn AR, Otipoby KL, Clark EA, et al. Polygenic autoimmune traits: Lyn, CD22, and SHP-1 are limiting elements of a biochemical pathway regulating BCR signaling and selection. Immunity. 1998; 8:497–508. [PubMed: 9586639]
11. Diehl NL, Enslen H, Fortner KA, Merritt C, Stetson N, Charland C, et al. Activation of the p38 mitogen-activated protein kinase pathway arrests cell cycle progression and differentiation of immature thymocytes in vivo. J Exp Med. 2000; 191:321–334. [PubMed: 10637276]
12. Dupuy AJ, Akagi K, Largaespada DA, Copeland NG, Jenkins NA. Mammalian mutagenesis using a highly mobile somatic Sleeping Beauty transposon system. Nature. 2005; 436:221–226. [PubMed: 16015321]
13. Dupuy AJ, Rogers LM, Kim J, Nannapaneni K, Starr TK, Liu P, et al. A modified sleeping beauty transposon system that can be used to model a wide variety of human cancers in mice. Cancer Res. 2009; 69:8150–8156. [PubMed: 19808965]
14. Evelyn CR, Duan X, Biesiada J, Seibel WL, Meller J, Zheng Y. Rational design of small molecule inhibitors targeting the Ras GEF, SOS1. Chemistry & biology. 2014; 21:1618–1628. [PubMed: 25455859]
15. Heltemes-Harris LM, Willette MJ, Ramsey LB, Qiu YH, Neeley ES, Zhang N, et al. Ebf1 or Pax5 haploinsufficiency synergizes with STAT5 activation to initiate acute lymphoblastic leukemia. The Journal of experimental medicine. 2011; 208:1135–1149. [PubMed: 21606506]
16. Hobeika E, Thiemann S, Storch B, Jumaa H, Nielsen PJ, Pelanda R, et al. Testing gene function early in the B cell lineage in mb1-cre mice. Proc Natl Acad Sci U S A. 2006; 103:13789–13794. [PubMed: 16940357]
17. Inaba H, Greaves M, Mullighan CG. Acute lymphoblastic leukaemia. Lancet. 2013; 381:1943–1955. [PubMed: 23523389]
18. Irving J, Matheson E, Minto L, Blair H, Case M, Halsey C, et al. Ras pathway mutations are prevalent in relapsed childhood acute lymphoblastic leukemia and confer sensitivity to MEK inhibition. Blood. 2014; 124:3420–3430. [PubMed: 25253770]
19. Iwasa H, Han J, Ishikawa F. Mitogen-activated protein kinase p38 defines the common senescence-signalling pathway. Genes to cells : devoted to molecular & cellular mechanisms. 2003; 8:131–144. [PubMed: 12581156]
20. Jacobs JJ, Kieboom K, Marino S, DePinho RA, van Lohuizen M. The oncogene and Polycomb-group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus. Nature. 1999; 397:164–168. [PubMed: 9923679]
21. Jun JE, Yang M, Chen H, Chakraborty AK, Roose JP. Activation of extracellular signal-regulated kinase but not of p38 mitogen-activated protein kinase pathways in lymphocytes requires allosteric activation of SOS. Molecular and cellular biology. 2013; 33:2470–2484. [PubMed: 23589333]
22. Lessard J, Sauvageau G. Bmi-1 determines the proliferative capacity of normal and leukaemic stem cells. Nature. 2003; 423:255–260. [PubMed: 12714970]
23. Li Q, Bohin N, Wen T, Ng V, Magee J, Chen SC, et al. Oncogenic Nras has bimodal effects on stem cells that sustainably increase competitiveness. Nature. 2013; 504:143–147. [PubMed: 24284627]
24. Lu T, Jackson MW, Wang B, Yang M, Chance MR, Miyagi M, et al. Regulation of NF-kappaB by NSD1/FBXL11-dependent reversible lysine methylation of p65. Proc Natl Acad Sci U S A. 2010; 107:46–51. [PubMed: 20080798]
25. Lu T, Yang M, Huang DB, Wei H, Ozer GH, Ghosh G, et al. Role of lysine methylation of NF-kappaB in differential gene regulation. Proc Natl Acad Sci U S A. 2013; 110:13510–13515. [PubMed: 23904479]
26. Mandal M, Powers SE, Ochiai K, Georgopoulos K, Kee BL, Singh H, et al. Ras orchestrates exit from the cell cycle and light-chain recombination during early B cell development. Nat Immunol. 2009; 10:1110–1117. [PubMed: 19734904]
27. Mullighan CG, Goorha S, Radtke I, Miller CB, Coustan-Smith E, Dalton JD, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature. 2007; 446:758–764. [PubMed: 17344859]

28. Mullighan CG, Miller CB, Radtke I, Phillips LA, Dalton J, Ma J, et al. BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. Nature. 2008; 453:110–114. [PubMed: 18408710]
29. Mullighan CG, Williams RT, Downing JR, Sherr CJ. Failure of CDKN2A/B (INK4A/B-ARF)-mediated tumor suppression and resistance to targeted therapy in acute lymphoblastic leukemia induced by BCR-ABL. Genes & development. 2008; 22:1411–1415. [PubMed: 18519632]
30. Mullighan CG, Su X, Zhang J, Radtke I, Phillips LA, Miller CB, et al. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. N Engl J Med. 2009; 360:470–480. [PubMed: 19129520]
31. Mullighan CG. Genomic characterization of childhood acute lymphoblastic leukemia. Seminars in hematology. 2013; 50:314–324. [PubMed: 24246699]
32. Nakayama J, Yamamoto M, Hayashi K, Satoh H, Bundo K, Kubo M, et al. BLNK suppresses pre-B-cell leukemogenesis through inhibition of JAK3. Blood. 2009; 113:1483–1492. [PubMed: 19047679]
33. Oguro H, Yuan J, Ichikawa H, Ikawa T, Yamazaki S, Kawamoto H, et al. Poised lineage specification in multipotential hematopoietic stem and progenitor cells by the polycomb protein Bmi1. Cell Stem Cell. 2010; 6:279–286. [PubMed: 20207230]
34. Park, I-k; Qian, D.; Kiel, M.; Becker, MW.; Pihalja, M.; Weissman, IL., et al. Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. Nature. 2003; 423:302–305. [PubMed: 12714971]
35. Roberts KG, Mullighan CG. Genomics in acute lymphoblastic leukaemia: insights and treatment implications. Nat Rev Clin Oncol. 2015
36. Sarver AL, Erdman J, Starr T, Largaespada DA, Silverstein KA. TAPDANCE: an automated tool to identify and annotate transposon insertion CISs and associations between CISs from next generation sequence data. BMC Bioinformatics. 2012; 13:154. [PubMed: 22748055]
37. Schwickert TA, Tagoh H, Gultekin S, Dakic A, Axelsson E, Minnich M, et al. Stage-specific control of early B cell development by the transcription factor Ikaros. Nat Immunol. 2014; 15:283–293. [PubMed: 24509509]
38. Shojaee S, Caeser R, Buchner M, Park E, Swaminathan S, Hurtz C, et al. Erk Negative Feedback Control Enables Pre-B Cell Transformation and Represents a Therapeutic Target in Acute Lymphoblastic Leukemia. Cancer Cell. 2015; 28:1–15. [PubMed: 26175407]
39. Starr TK, Largaespada DA. Cancer gene discovery using the Sleeping Beauty transposon. Cell Cycle. 2005; 4:1744–1748. [PubMed: 16294016]
40. Starr TK, Allaei R, Silverstein KA, Staggs RA, Sarver AL, Bergemann TL, et al. A transposon-based genetic screen in mice identifies genes altered in colorectal cancer. Science. 2009; 323:1747–1750. [PubMed: 19251594]
41. Vougioukalaki M, Kanellis DC, Gkouskou K, Eliopoulos AG. Tpl2 kinase signal transduction in inflammation and cancer. Cancer letters. 2011; 304:80–89. [PubMed: 21377269]
42. Vu TT, Gatto D, Turner V, Funnell AP, Mak KS, Norton LJ, et al. Impaired B cell development in the absence of Kruppel-like factor 3. Journal of immunology. 2011; 187:5032–5042.
43. Wang W, Chen JX, Liao R, Deng Q, Zhou JJ, Huang S, et al. Sequential activation of the MEK-extracellular signal-regulated kinase and MKK3/6-p38 mitogen-activated protein kinase pathways mediates oncogenic ras-induced premature senescence. Molecular and cellular biology. 2002; 22:3389–3403. [PubMed: 11971971]
44. Williams RT, Sherr CJ. The INK4-ARF (CDKN2A/B) locus in hematopoiesis and BCR-ABL-induced leukemias. Cold Spring Harb Symp Quant Biol. 2008; 73:461–467. [PubMed: 19028987]
45. Xu H, Yang W, Perez-Andreu V, Devidas M, Fan Y, Cheng C, et al. Novel susceptibility variants at 10p12.31-12.2 for childhood acute lymphoblastic leukemia in ethnically diverse populations. Journal of the National Cancer Institute. 2013; 105:733–742. [PubMed: 23512250]
46. Zhang J, Mullighan CG, Harvey RC, Wu G, Chen X, Edmonson M, et al. Key pathways are frequently mutated in high-risk childhood acute lymphoblastic leukemia: a report from the Children's Oncology Group. Blood. 2011; 118:3080–3087. [PubMed: 21680795]

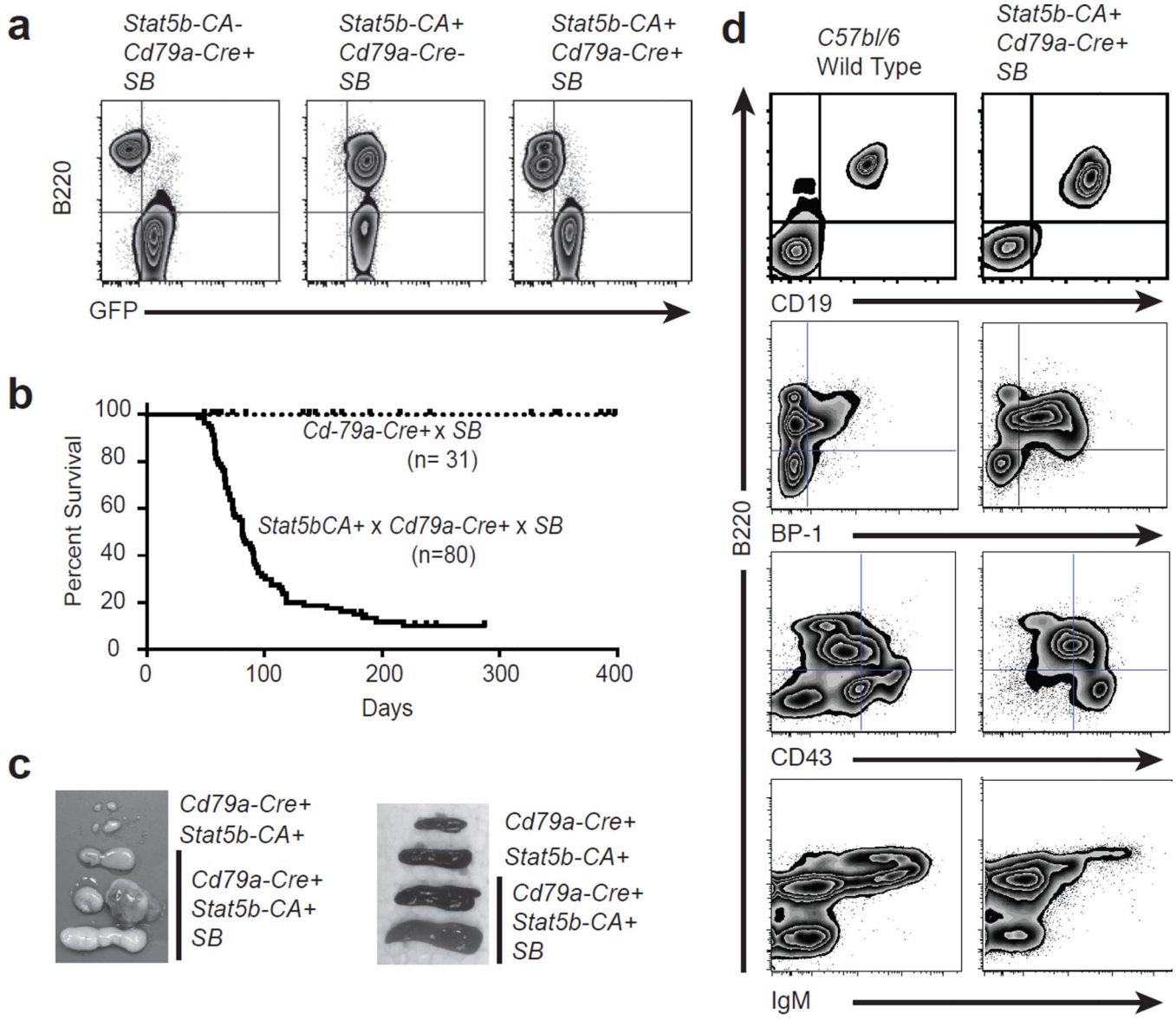

Figure 1.

Sleeping Beauty mutagenesis screen induces leukemia in *Stat5b-CA* mice. (**a**) Flow cytometry of bone marrow from *Cd79a-Cre*x*SB* and *Stat5b-CAxSB* control mice versus *Stat5b-CAxCd79a-Cre*x*SB* leukemic mice, showing efficient deletion of GFP in B lineage cells. (**b**), Kaplan-Meier survival analysis of mice comparing *Stat5b-CAxCd79a-Cre*x *SB*(n=80) to control mice*Cd79a-Cre*x*SB*(n=31). (**c**) Pictures display a comparison of lymph node and spleen from *Cd79a-Cre* and*Stat5b-CA* control mice and *Stat5b-CAxCd79a-Cre*x*SB* tumor mice. (**d**), Flow cytometric analysis of lymph node cells from C57Bl/6 control mice and *Stat5b-CAxCd79a-Cre*x*SB* tumor mice. Representative flow cytometric analysis of B220, GFP, CD19, BP-1, CD43 (S7), and IgM expression on lymph node cells is shown. Doublets were gated out and a lymphocyte gate was set based on side and forward scatter properties. All gates shown are based on cells from control animals.

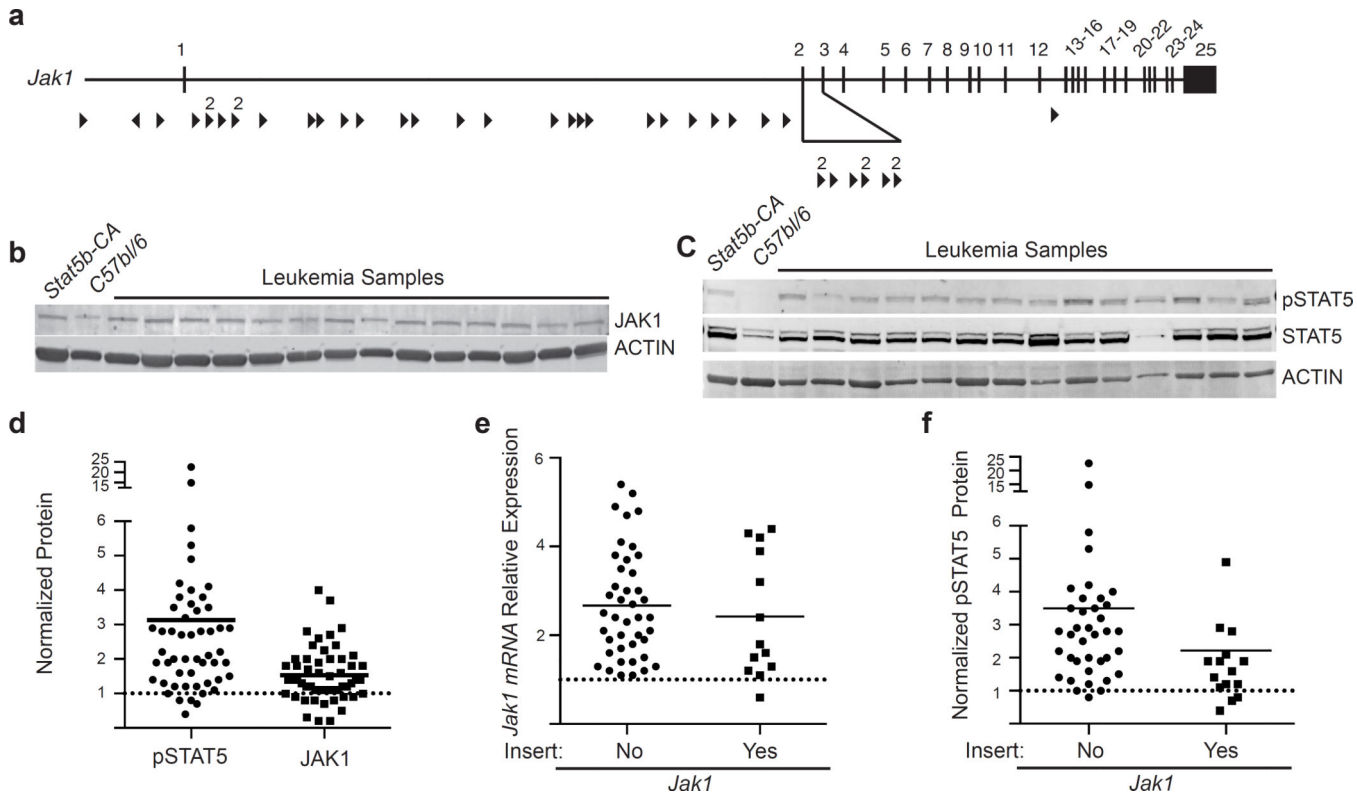

Figure 2.

Increased expression of both Stat5 and Jak1 in the leukemic mice. (**a**) Map of common insertion sites in the *Jak1* gene. Arrow direction indicates the direction of the mapped transposon insertion and the size of the arrow is indicative of the number of mapped insertions at that specific site. The number above the arrow is the number of insertions at that site. (**b**) Representative western blot analysis showing increased levels of total JAK1 expression. (**c**) Representative western blot analysis showing increased levels of phospho-STAT5. (**d**) Scatter plot summary of western blotting data from 53 samples. Control CD19+ splenocytes were isolated from *C57bl/6* or *Stat5b-CA* mice. Values in D were normalized to expression of the gene of interest in *Stat5b-CA* B cells using *Actin* as an internal control. Data were analyzed using a MannU-Whitney test. pSTAT5, and total JAK1 levels were all significantly different from controls, $p<0.05$. (**e**) *JAK1* mRNA expression. Values were normalized to expression of the gene of interest in *Stat5b-CA* progenitor B cells using *Actin* as an internal control. Closed circles represent expression in mice without transposon insert (no) and closed squares represent expression in mice with a transposon insert (yes). (**f**) pSTAT5 expression in samples with or without a *Jak1* insert. Closed circles represent expression in mice without transposon insert (no) and closed squares represent expression in mice with a transposon insert (yes). Lines in panels d, e and f represent means.

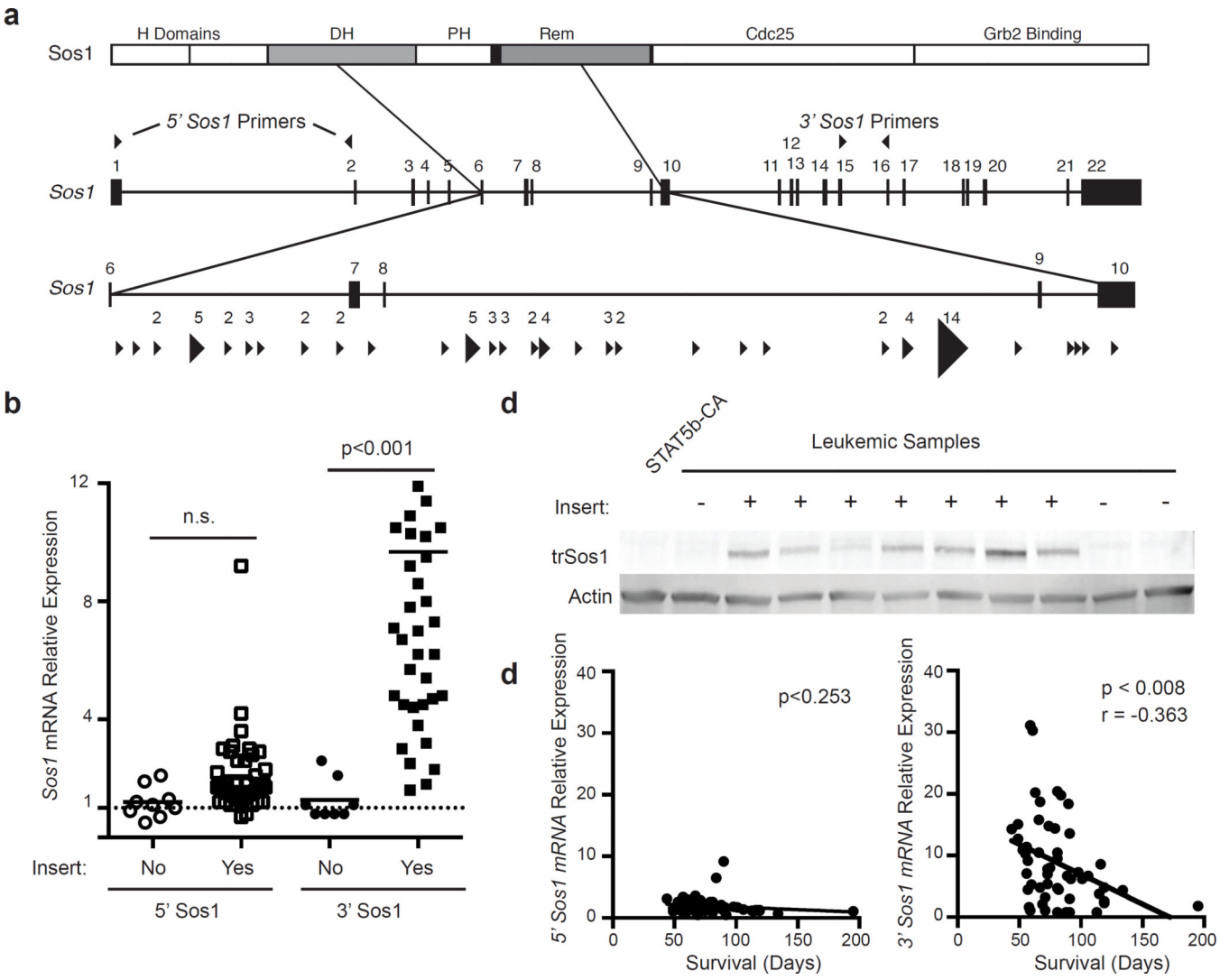

Figure 3.

Identification of a truncated form of *Sos1* that is over-expressed in leukemic cells. (**a**) Map of common insertion sites in the *Sos1* gene. The top diagram illustrates the domains in the SOS1 protein. Arrow direction indicates the direction of the mapped transposon insertion and the size of the arrow is indicative of the number of mapped insertions at that specific site. The number above the arrow is the number of insertions at that site. The 5' and 3' primers used to quantitate 5' and 3' mRNA transcripts are indicated on the gene. (**b**) *Sos1* mRNA levels were measured by real-time RT-PCR with primers from either the 5' or 3' end of the transcript. Open symbols represent primers in the 5' region of *Sos1* (Exons1–2) while filled symbols represent expression near the 3' region of *Sos1* (Exons 15–16). Values were normalized to expression of the gene of interest in *Stat5b-CA* progenitor B cells using *Actin* as an internal control. P-values were determined using a Kruskal-Wallace test. Lines represent means. (**c**) SOS1 truncated protein expression was examined by western blot. Controls are $CD19^{+}$ splenic B cells from *Stat5b-CA* mice. The presence of a transposon insert is indicated by a plus sign. (**d**) Linear regression analysis of either 5' or 3' *Sos1* transcript levels compared to survival.

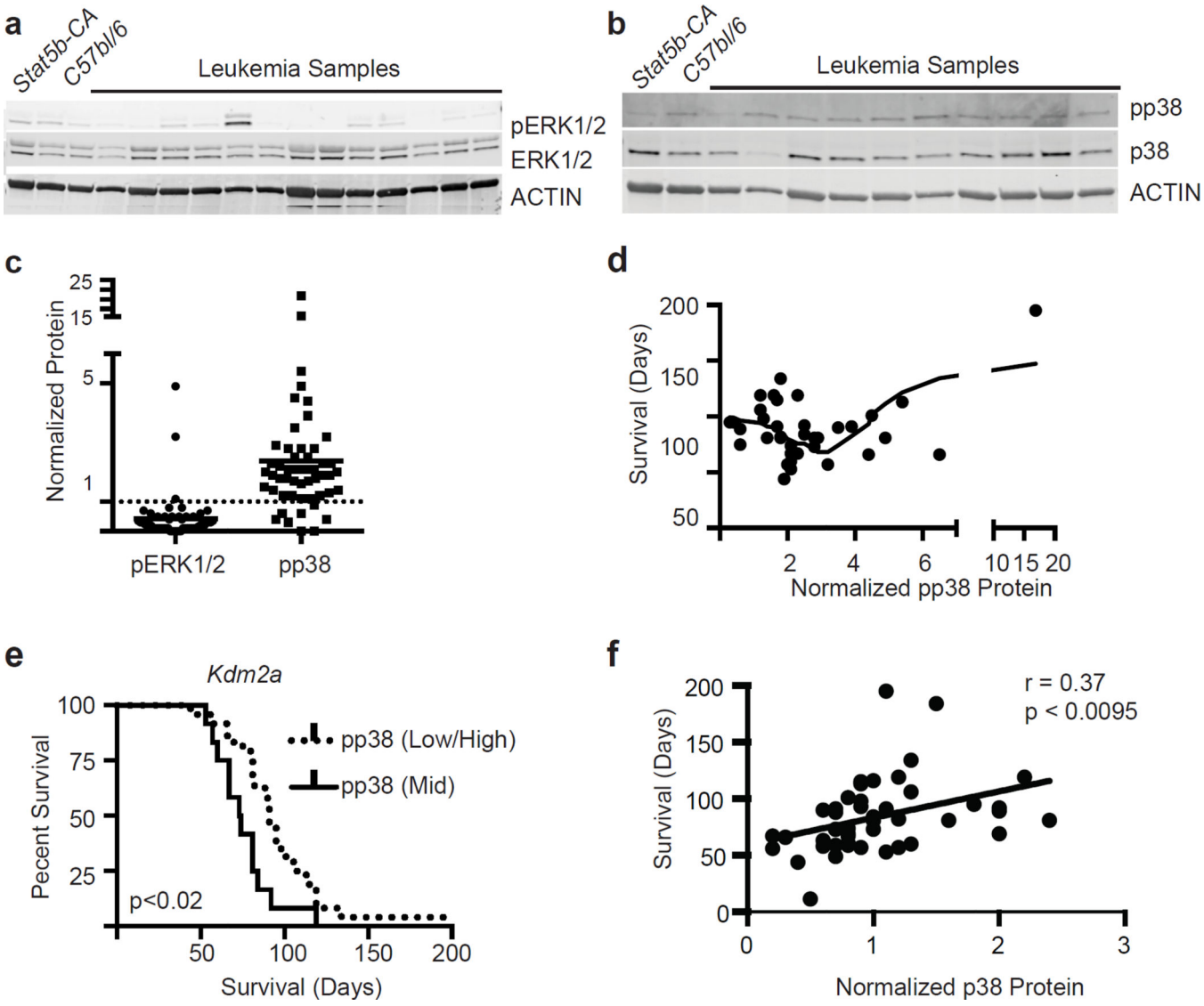

Figure 4.

Increased levels of pp38 but decreased phospho-ERK in leukemic cells. (**a**) Western blot analysis showing levels of pErk. (**b**) Western blot analysis showing levels of pp38. (**c**) Scatter plot summary of all western blotting data. There were 54 samples analyzed for pERK1/2 and 51 samples analyzed for pp38 expression. The samples were normalized to *Stat5b-CA* splenic B cell expression of each protein. Data were analyzed using a MannU-Whitney test; pERK and pp38 levels were all significantly different from controls, $p<0.001$. Lines represent means. (**d**) Plot showing the distribution of pp38 levels and survival. The line represents a second order polynomial fit to the data points. (**e**) Survival curves were generated by splitting mice into two groups based on pp38 expression – (group 1) 2–3 fold increase in pp38 levels (n=12), and (group 2) less than two or greater than 3 pp38 levels (n=24). A log-rank test was used to determine the p value. (f) Correlation of total p38 levels with survival. Correlation was calculated by computing r for p38 versus every survival using Pearson correlation coefficients.

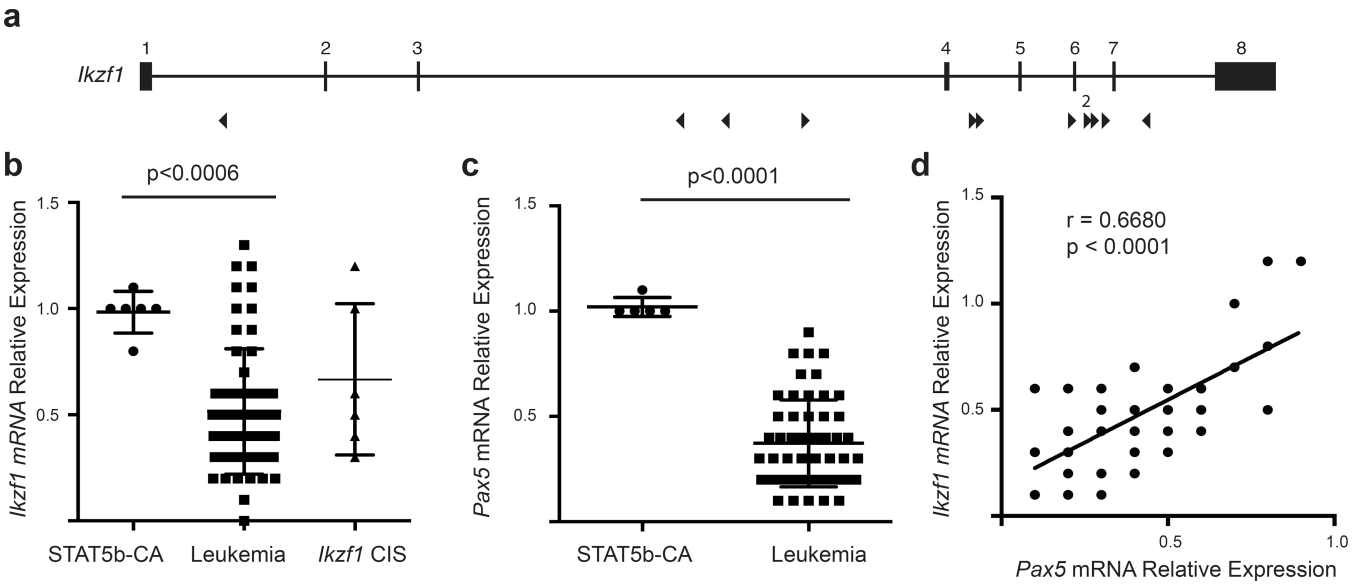

Figure 5.

Decreased expression of *Ikzf1* and *Pax5* transcripts in the leukemic cells. (**a**) Map of common insertion sites in the *Ikzf1* gene. Arrow direction indicates the direction of the mapped transposon insertion and the size of the arrow is indicative of the number of mapped insertions at that specific site. The number above the arrow is the number of insertions at that site. (**b**) *Ikzf1* mRNA expression. These results represent 6 *Stat5b-CA* replicates and 57 leukemic samples. *Ikzf1 Cis* column shows the *Ikzf1* expression for just those samples with transposon insertions. (**c**) *Pax5* mRNA expression. These results represent 5 *Stat5b-CA* replicates and 54 leukemic samples. p-values were calculated using MannU-Whitney test. Lines in panels b and c represent means ± standard deviation. (**d**) Correlation of *Ikzf1* and *Pax5* expression. Correlation was calculated by computing r for *Pax5* versus every *Ikzf1* using Pearson correlation coefficients.

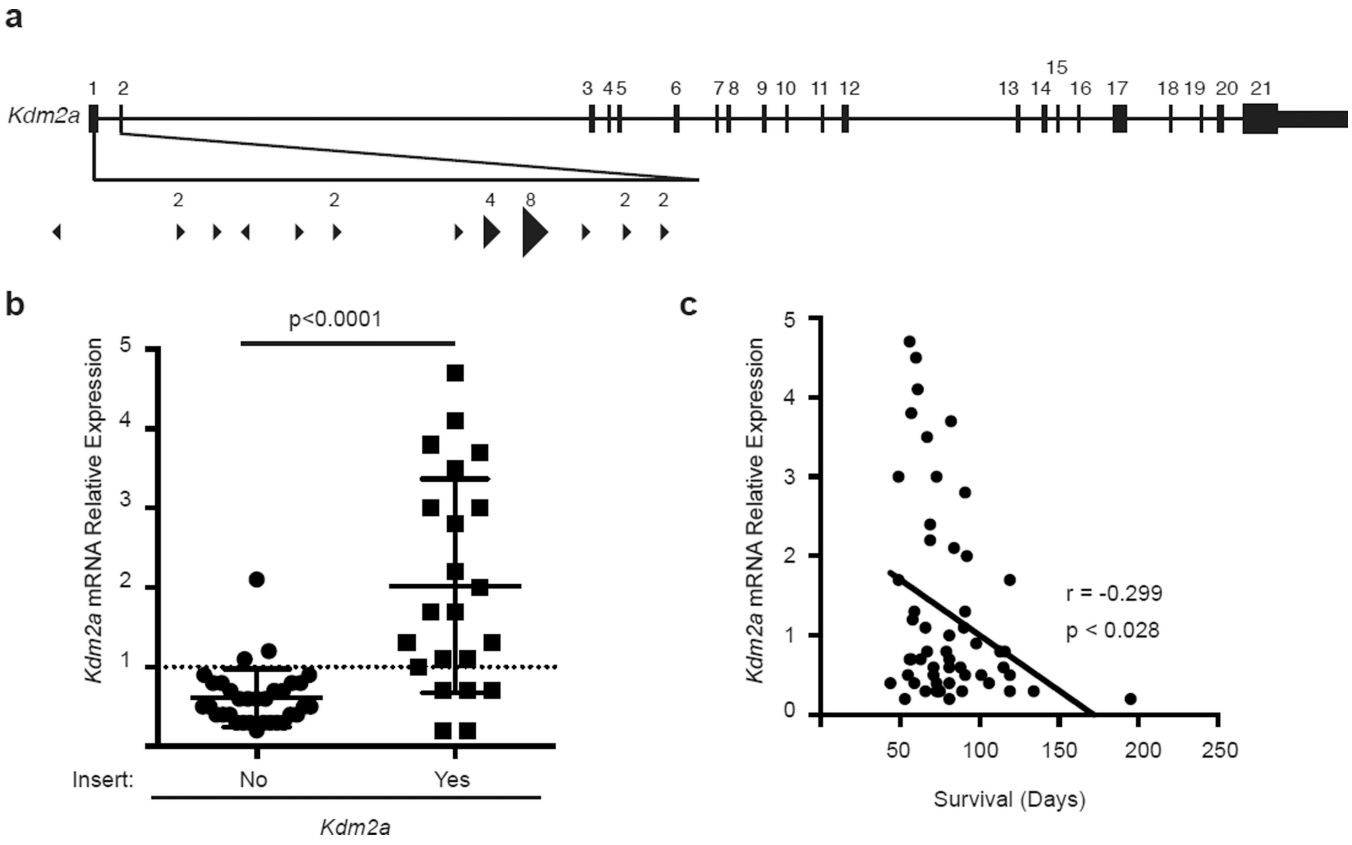

Figure 6.

Transposon insertion into *Kdm2a* locus results in increased levels of *Kdm2a* and increased *Kdm2a* expression correlates with poor survival. (**a**) Map of common insertion sites in the *Kdm2a* gene. Arrow direction indicates the direction of the mapped transposon insertion and the size of the arrow is indicative of the number of mapped insertions at that specific site. The number above the arrow is the number of insertions at that site. (**b**) *Kdm2a* mRNA expression. Values were normalized to expression of the gene of interest in *Stat5b-CA* progenitor B cells using *Actin* as an internal control. Closed circles represent expression in mice without transposon insert (no) and closed squares represent expression in mice with a transposon insert (yes). A total of 55 samples were analyzed for *Kdm2a* expression with 22 containing a transposon insertion. Error bars represent the standard deviation and p-values were calculated using MannU-Whitney test. (**c**) Linear regression analysis of *Kdm2a* expression and survival.

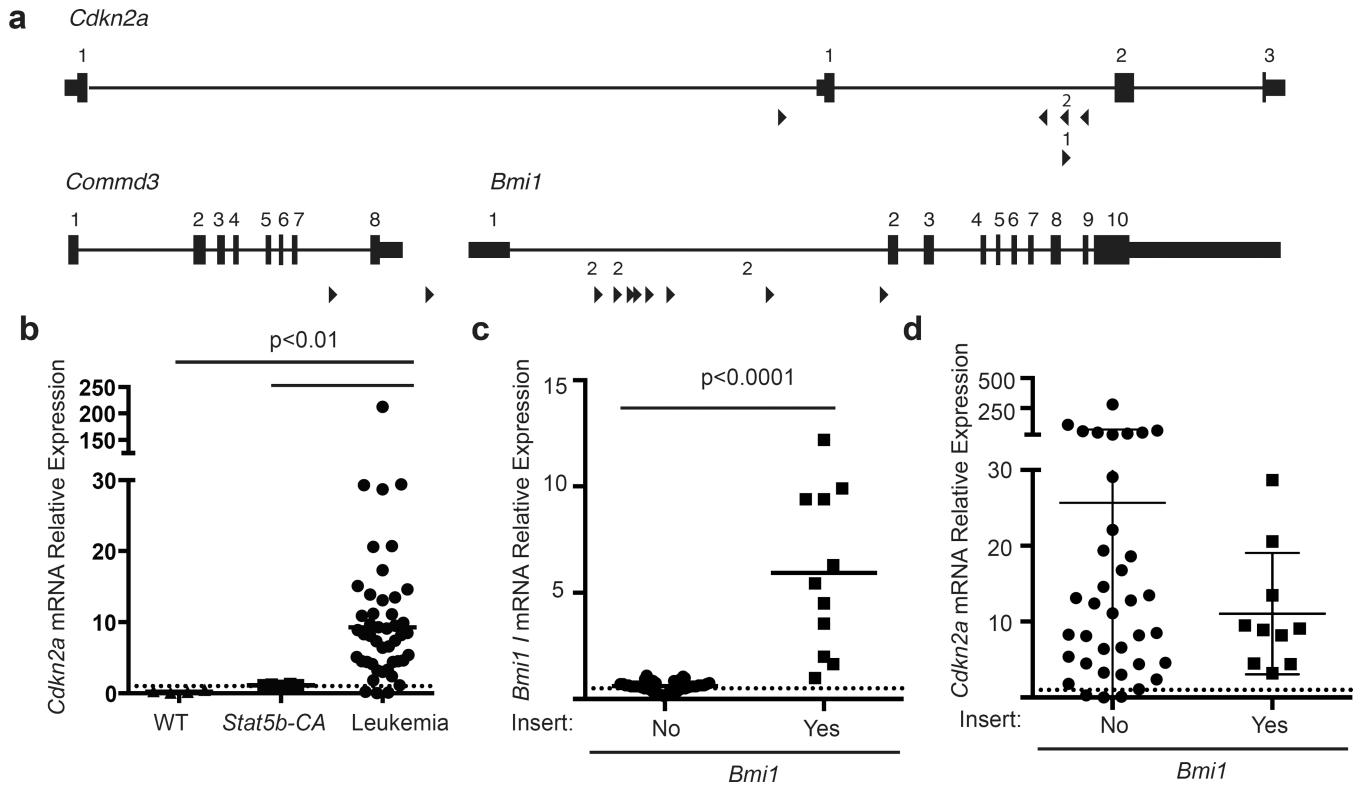

Figure 7.

Role of *Bmi1* and *Cdkn2a* in leukemia. (**a**) Map of common insertion sites in the *Bmi1* and *Cdkn2a* genes. Arrow direction indicates the direction of the mapped transposon insertion and the size of the arrow is indicative of the number of mapped insertions at that specific site. The number above the arrow is the number of insertions at that site. (**b**) *Cdkn2a* mRNA expression in leukemic or progenitor B cells from control mice. Line represents mean of the samples. This analysis includes 4 *C57bl/6* (WT), 6 *Stat5b-CA* and 55 leukemic samples; *Cdkn2a* levels were significantly higher ($p<0.01$) in leukemic samples than WT or *Stat5b-CA* samples; p-values were calculated by Kruskal-Wallace test. (**c**) *Bmi1* mRNA expression. *Bmi1* expression is based on 47 leukemic samples with 10 samples containing a transposon insertion. Line represents mean of the samples and p-values were calculated using MannU-Whitney test. (**d**) Comparison of *Cdkn2a* expression in mice with or without a transposon insertion in the *Bmi1* gene. Error bars represent the standard deviation and p-values were calculated using MannU-Whitney test.

Table 1

Line Number	CIS-Associated Gene	TAPDANCE CIS (%)	TAPDANCE CIS p value library number	Prediction of effect on gene function
1	*Sos1*	74	2.6E-104	Drive
2	*Kdm2a*	35	1.3E-42	Gain
3	*Jak1*	26	2.6E-11	Gain
4	*Bmi1*	19	1.5E-17	Drive
5	*No Results within 20000 bp window 1*	18	1.9E-15	??
6	*Cdkn2a*	10	1.5E-05	Disrupt
7	*Ikzf1*	8	9.3E-04	Loss
8	*Snora 17*	8	5.0E-02	
9	*5830433M19Rik, RP23-456B18.4*	6	5.0E-02	?
10	*Cap1*	6	5.0E-02	Disrupt
11	*Klf3*	6	5.0E-02	Gain
12	*No Results within 20000 bp window 2*	6	5.0E-02	??

Table 2

Line Number	CIS-Associated Gene	Gene Centric CIS (%)	Gene Centric Q-value	Prediction of effect on gene function
1	***Sos1***	73	0	Drive
2	*0610039K10Rik*	52	0	Disrupt
3	***9330182L06Rik***	46	0	Drive
4	***Kdm2a***	32	0	Gain
5	*Pcdh9*	31	0.00031	Drive
6	*Foxf2*	28	0	Disrupt
7	*Tbc1d22b*	24	0	Disrupt
8	*Ccdc53*	23	0	Disrupt
9	*Duox1*	23	0	Disrupt
10	***Jak1***	21	1.06E-50	Gain
11	*Sorcs1*	21	3.80E-05	Disrupt
12	*Prkca*	20	7.07E-10	Drive
13	***Bmi1***	17	0	Drive
14	*Disc1*	15	1.90E-16	Drive
15	*Sfi1*	14	4.25E-46	Drive
16	*Srgap1*	14	3.74E-07	Disrupt
17	*Chl1*	13	0.000145	Drive
18	*Klf17*	13	0	Drive
19	*Xkr7*	13	0	Disrupt
20	*Garnl3*	11	4.12E-16	Drive
21	*Rasgrf1*	11	1.45E-12	Disrupt?
22	***Ikzf1***	10	4.15E-13	Loss
23	*Scn10a*	10	1.81E-12	Drive
24	***Cdkn2a***	8	1.00E-39	Disrupt
25	*Ckap5*	8	1.63E-07	Disrupt
26	*Ptpn6*	8	2.29E-61	Drive
27	*Setd5*	8	3.10E-07	Disrupt
28	*Trim12c*	8	1.42E-50	Disrupt
29	*Col5a1*	7	0.000224	Disrupt
30	*Ift140*	7	4.82E-08	Disrupt
31	*Nlrp1c*	7	2.79E-13	Disrupt
32	*Rxfp2*	7	6.60E-09	Truncated
33	*Dmd*	6	0.000141	Disrupt
34	***Klf3***	6	7.25E-16	Gain
35	*Lingo1*	6	2.76E-08	Disrupt?
36	*Map3k8*	6	2.49E-15	Disrupt

Line Number	CIS-Associated Gene	Gene Centric CIS (%)	Gene Centric Q-value	Prediction of effect on gene function
37	*Myb*	6	5.68E-09	Drive?
38	*Mysm1*	6	1.07E-06	Disrupt
39	*Plekha8*	6	8.26E-06	Disrupt
40	*Rab36*	6	1.04E-24	Drive
41	*Tmem51*	6	1.11E-09	Disrupt
42	*4833419F23Rik*	4	4.76E-12	?
43	*Akirin2*	4	2.15E-08	Disrupt
44	*Anxa7*	4	1.79E-06	Disrupt
45	*Ap2s1*	4	3.86E-20	Disrupt
46	***Cap1***	4	2.90E-06	Disrupt
47	*Chst3*	4	8.36E-07	Disrupt
48	*Dok1*	4	4.75E-60	Drive?
49	*Eci3*	4	4.47E-09	Disrupt
50	*Il20rb*	4	1.17E-06	?
51	*Lass4*	4	3.79E-06	Disrupt?
52	*Lats1*	4	0.000246	?
53	*Loxl3*	4	3.03E-15	Disrupt
54	*Mau2*	4	2.41E-08	Disrupt
55	*Mki67*	4	3.45E-07	Drive?
56	*Nxpe2*	4	1.81E-06	?
57	*Osbpl5*	4	7.79E-05	Disrupt
58	*Prrxl1*	4	0.000206	Disrupt
59	*Sema4f*	4	1.33E-07	?
60	*Snapc3*	4	9.86E-05	Disrupt
61	*Tmem204*	4	3.47E-10	Disrupt?
62	*Ube2i*	4	5.87E-16	Drive
63	*Vars*	4	2.95E-18	Disrupt
64	*Zbtb8b*	4	9.63E-13	Disrupt
65	*Zfp318*	4	0.000293	Disrupt