

U2AF1 mutations induce oncogenic IRAK4 isoforms and activate innate immune pathways in myeloid malignancies

Molly A. Smith^{1,2,#}, Gaurav S. Choudhary^{3,#}, Andrea Pellagatti⁴, Kwangmin Choi¹, Lyndsey C. Bolanos¹, Tushar D. Bhagat³, Shanisha Gordon-Mitchell³, Dagny Von Ahrens³, Kith Pradhan³, Violetta Steeples⁴, Sanghyun Kim⁵, Ulrich Steidl³, Matthew Walter⁵, Iain D.C. Fraser⁶, Aishwarya Kulkarni⁷, Nathan Salomonis^{7,8}, Kakajan Komurov^{1,8}, Jacqueline Boulton^{4*}, Amit Verma^{3*}, Daniel T. Starczynowski^{1,2,8*}

1. Division of Experimental Hematology and Cancer Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH
2. Department of Cancer Biology, University of Cincinnati, Cincinnati, OH
3. Albert Einstein College of Medicine, Montefiore Medical Center, Bronx, NY
4. Nuffield Division of Clinical Laboratory Sciences, Radcliffe Department of Medicine, University of Oxford, and Oxford BRC Haematology Theme, Oxford, UK
5. Department of Medicine, Washington University, St. Louis, MO
6. Laboratory of Immune System Biology, National Institute of Allergy and Infectious Diseases, National Institute of Health, Bethesda, MD
7. Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH
8. Department of Pediatrics, University of Cincinnati, Cincinnati, OH
- #. Equal contribution
- *. Co-correspondence

Correspondence:

Daniel Starczynowski
Division of Experimental Hematology and Cancer Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA
513-803-5317
Daniel.Starczynowski@cchmc.org

Amit Verma, MD
Albert Einstein College of Medicine, Montefiore Medical Center,
Bronx, NY, USA
718-430-8761
amit.verma@einstein.yu.edu

Jacqueline Boulton
University of Oxford, Radcliffe Department of Medicine, John Radcliffe Hospital,
Oxford, United Kingdom
jacqueline.boulton@ndcls.ox.ac.uk

Words: 4055
Abstract: 146

49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64

Abstract

Spliceosome mutations are common in MDS and AML, yet the oncogenic changes due to these mutations have not been identified. A global analysis of exon usage in AML samples revealed distinct molecular subsets containing alternative spliced isoforms of inflammatory and immune genes. IRAK4 was the dominant alternatively spliced isoform in MDS/AML and is characterized by a longer isoform that retains exon 4, encoding a protein, IRAK4-Long (L) that assembles with the Myddosome, results in maximal activation of NF- κ B, and is essential for leukemic cell function. Expression of IRAK4-L is mediated by mutant U2AF1 and is associated with oncogenic signaling in MDS/AML. Inhibition of IRAK4-L abrogates leukemic growth, particularly in AML cells with higher expression of the IRAK4-L isoform. Collectively, mutations in U2AF1 induce expression of therapeutically targetable "active" IRAK4 isoforms and provide a genetic link to activation of chronic innate immune signaling in MDS and AML.

65 Introduction

66 The majority of human genes express more than one RNA isoform resulting from alternative
67 sites of splicing, and transcriptional initiation and termination. Alternative splicing can generate
68 diversity at the RNA and protein levels, which is important for many biological processes.¹
69 Aberrations in alternative splicing are also implicated in human cancers, resulting from somatic
70 mutations in genes encoding RNA splicing proteins or in splicing regulatory motifs within exons
71 and introns, and/or altered expression of RNA-binding proteins. The importance of optimal RNA
72 splicing and isoform expression in chronic lymphocytic leukemia, myelodysplastic syndromes
73 (MDS), and acute myeloid leukemia (AML) is evident by the frequent occurrence of mutations in
74 genes encoding RNA splicing factors.²⁻⁹ Although relatively uncommon in solid tumors, ~50% of
75 MDS patients harbor mutations in RNA splicing factors.^{10,11} Despite significant advances in
76 understanding the role of RNA splicing factors in hematopoiesis, how RNA splicing gene mutations
77 contribute to cancer is unclear.

78 A subset of leukemia-associated genes is regulated at the expression level, which results
79 from the coordinated change in all of a gene's individual isoforms, typically due to usage of
80 alternative transcriptional start sites. However, we posited that a subset of genes is regulated
81 exclusively at the isoform level in leukemia, which results in anti-correlated expression of the
82 individual RNA isoforms. Herein, a global analysis of exon usage in AML revealed enrichment of
83 genes associated with inflammatory and immune pathways that are regulated by RNA isoform
84 changes. Of the inflammatory and immune pathway genes undergoing RNA isoform switching in
85 MDS/AML, isoform expression of interleukin-1 receptor-associated kinase 4 (IRAK4) was the most
86 significantly altered. This isoform encodes a longer protein (IRAK4-L) that physically associates
87 with MyD88 and results in activation of NF- κ B and MAPK. We also demonstrated that expression
88 of IRAK4-L is directly mediated by the mutant U2AF1 splicing factor in MDS/AML. Inhibition of
89 IRAK4 abrogates leukemic growth and is more efficacious in AML cells with U2AF1 mutations
90 and/or higher expression of IRAK4-L. Taken together, we determined that mutations in U2AF1
91 induce expression of therapeutically targetable "active" IRAK4 isoforms and provide a genetic link
92 to activation of chronic innate immune signaling in MDS and AML.

93

94 Results

95 RNA isoforms define a subset of AML associated with innate immune pathway activation.

96 Molecularly distinct subsets of AML have been largely defined on the basis of somatic
97 mutations and changes in global gene expression.^{10,12,13} As a complementary approach, we

investigated whether differential RNA isoform expression and alternative splicing is sufficient to distinguish clinically-relevant molecular subtypes of AML and uncover oncogenic signaling dependencies (**Supplemental Figure 1A**). To determine the prevalence of differential RNA isoform usage in AML, we analyzed the RNA sequencing data from The Cancer Genome Atlas (TCGA). First, we identified genes that were explicitly regulated at the level of mRNA isoform switching. Regulation at the level of mRNA isoform switching for a gene was defined when at least 2 of its RNA isoforms exhibited a mutually exclusive expression pattern (expression correlation < -0.4) and higher isoform-level expression variance as compared to gene-level variance (**Supplemental Figure 1B**). Based on these criteria, we identified 887 genes that showed patterns of RNA isoform switching (**Figure 1A, Supplemental Table 1**). A semi-unsupervised hierarchical cluster analysis revealed 3 groupings of AML characterized by distinct patterns of differential isoform usage for these genes. The three groups have distinct clinical outcomes, with Group 2 having the worst prognosis (**Figure 1B, Supplemental Figure 1C**). To specify alternative exon-level expression events that confer worse prognosis, individual exon inclusion/exclusion events for these genes were correlated with the overall survival using a multivariate Cox regression analysis. We identified a subset of genes undergoing mRNA isoform switching with a least one differential exon inclusion/exclusion event that significantly correlated with AML survival in patients from Group 2 (**Supplemental Table 2**). Functional enrichment analysis revealed that the genes whose differential exon usage correlates with worse AML prognosis is enriched for RNA splicing pathways and innate immune and NF- κ B signaling (**Figure 1C, Supplemental Table 3**).

Dysregulation of immune-related genes is widely reported in MDS and AML, however the precise alterations that drive innate immune signaling in MDS or AML HSC remain largely unknown¹⁴. Therefore, we decided to focus on RNA isoform switching of genes that belong to innate immune and NF- κ B signaling. Among these genes, IRAK4 emerged as the leading candidate since the magnitude of IRAK4 isoform switching was highly significant among AML samples and the inclusion of exon 4 alone correlated with worse outcome (**Figure 1D, Supplemental Figure 1D**). IRAK4 is a serine/threonine kinase that mediates signaling downstream of the Toll-like receptor (TLR) superfamily, resulting in NF- κ B and MAPK activation¹⁵. IRAK4 protein consists of the death domain (aa 1-125), a hinge domain (aa 140-150), and kinase domain (aa 150-460). IRAK4 in AML is encoded by two major protein isoforms, defined by the inclusion or the exclusion of exon 4 in the mRNA (based on NCBI Refseq annotation, or exon 6 according to AltAnalyze Ensembl 72 annotation)(**Figure 1E**). Inclusion of exon 4 is predicted to encode a longer 460 aa IRAK4 protein, which contains the 3 main functional domains (IRAK4-L).

131 In contrast, exclusion of exon 4 is predicted to result in utilization of an alternative translational
132 start site and encode a shorter 336 aa IRAK4 protein lacking the N-terminal death domain (IRAK4-
133 S) (**Figure 1E**).

134 **Expression of IRAK4-L is associated with AML.**

135 We next compared IRAK4 exon 4 usage in AML and normal hematopoietic cells utilizing
136 RNA-sequencing data obtained from AML samples and normal BM samples. Although the
137 cumulative expression of both IRAK4 isoforms did not significantly differ among the patients, >50%
138 of AML patients preferentially expressed *IRAK4-L* (>1.25-fold), while 18% preferentially expressed
139 *IRAK4-S* (**Figure 1F**). As compared to normal BM cells, AML samples exhibit a significant
140 proportion of *IRAK4-L* expression relative to *IRAK4-S* (**Figure 1G**). Breast, colon, and lung cancer
141 cells also express a higher ratio of *IRAK4-L* relative to *IRAK4-S* as compared to normal tissues,
142 indicating that expression of IRAK4-L is associated with oncogenic states (**Supplemental Figure**
143 **1E**). To independently verify IRAK4 RNA isoform expression in AML, we examined human AML
144 and MDS cell lines. Sanger sequencing confirmed that both IRAK4 RNA isoforms are expressed
145 (**Figure 1H, Supplemental Figure 1F**). Consistent with the finding that AML patients express 2
146 major isoforms of IRAK4; MD5L, TF1, and THP1 exhibit preferential expression of *IRAK4-L*, while
147 HL60 and F36P preferentially express *IRAK4-S* (**Figure 1H**). In contrast, healthy cord-blood and
148 BM CD34+ cells preferentially express *IRAK4-S* (**Figure 1H**). As such, a subset of AML is
149 associated with expression of *IRAK4-L* resulting from inclusion of exon 4, while normal BM cells
150 primarily express *IRAK4-S* as a result of exon 4 exclusion.

151 To determine whether both IRAK4 RNA isoforms encode proteins in AML, we examined
152 IRAK4 protein expression in MDS/AML cell lines using an antibody that recognizes a C-terminal
153 epitope present on both IRAK4 protein isoforms. In agreement with the RNA isoform expression
154 analysis, MD5L, TF1, and THP1 exhibit preferential expression of the longer IRAK4 protein isoform
155 (IRAK4-L at ~50 kDa) while HL60 and F36P primarily express a shorter IRAK4 protein isoform
156 (IRAK4-S at ~37 kDa), that isn't recognized by the N-terminal antibody (**Figure 1I**). F36P cells
157 express a moderately smaller IRAK4 protein (~32 kDa) that corresponds to an IRAK4 isoform
158 utilizing a minor second downstream open reading frame. Although expression of IRAK4-L differed
159 between the AML cell lines, expression of MyD88 and IRAK1 was comparable across the AML cell
160 lines (**Figure 1I**). Primary AML cells also express IRAK4-L, while normal BM-derived CD34+
161 hematopoietic cells and mononuclear cells predominantly express IRAK4-S (**Figure 1J**). These
162 findings indicate that a significant proportion of AML patients primarily express the longer IRAK4
163 RNA and protein isoforms.

IRAK4-L results in maximal activation of innate immune signaling.

To investigate whether IRAK4-L expression is correlated with distinct gene regulatory networks in AML, RNA-sequencing data was examined in AML patients stratified based on the expression of *IRAK4-L* and *IRAK4-S*. Gene ontology and pathway analysis revealed that AML samples expressing *IRAK4-L* (top 25%) are significantly enriched in inflammatory and TLR signaling signatures as compared to AML samples expressing *IRAK4-S* (bottom 25%), suggesting that IRAK4-L could be a feed-forward activation node of oncogenic innate immune signaling (**Figure 2A, Supplemental Figure 2A**). To test whether the IRAK4 isoforms exert differential effects on innate immune signaling, IRAK4-L or IRAK4-S were expressed in HEK293 cells and evaluated for NF- κ B and MAPK activation. Expression of IRAK4-L resulted in increased phosphorylation of IRAK1, p65, p38, and JNK (**Figure 2B,C and Supplemental Figure 2B,C**). In contrast, expression of IRAK4-S resulted in phosphorylation of p38, and JNK, but was less efficient at inducing phosphorylation of IRAK1 and p65. IRAK4-L expression also induced NF- κ B- and AP1-dependent transcriptional activation as compared to control cells, while IRAK4-S induced AP1-dependent activation, but was less efficient at inducing NF- κ B-dependent activation (**Figure 2D,E**). Expression of IRAK4-L in HL60, which primarily express IRAK4-S, resulted in activation of IRAK1 and p65, suggesting that IRAK4-L is sufficient to activate NF- κ B in AML cells (**Supplemental Figure 2D**). To determine whether IRAK4-L and/or IRAK4-S are required to activate NF- κ B activation, THP1 cells expressing an NF- κ B reporter were transduced with shRNAs targeting IRAK4-L (shIRAK4-L) or IRAK4-L and IRAK4-S (shIRAK4-L/S) (**Figure 2F**). Knockdown of IRAK4-L resulted in suppression of NF- κ B activation at levels similar to knockdown of IRAK4-L/S, indicating that IRAK4-S does not contribute significantly to NF- κ B activation in AML (**Figure 2G**). These findings show that MAPK signaling is mediated by the C-terminal domain of IRAK4-S and IRAK4-L, while NF- κ B signaling is mediated by the N- and C-terminal domains of IRAK4-L (**Figure 2H**).

To further explore NF- κ B signaling differences between IRAK4-L and IRAK4-S, we performed a sequence-based prediction of protein-protein and domain-domain interactions.¹⁶ IRAK4-L protein isoforms, which retain the N-terminal death-domain, are predicted to directly bind MyD88 and IRAK1 (**Figure 2I**). Death domain interactions between MyD88 and IRAK4 initiate Myddosome formation, which promotes IRAK4 trans-autophosphorylation and activation.^{17,18} IRAK4 then activates IRAK1, enabling recruitment of TRAF6 and NF- κ B activation. Based on these predictions, IRAK4-S protein isoforms, which lack the N-terminal death-domain, are predicted not to bind MyD88 (**Figure 2I**). To confirm this, FLAG-IRAK4-L or FLAG-IRAK4-S were expressed in

196 HA-MyD88-expressing HEK293 cells followed by immunoprecipitation of IRAK4-L or IRAK4-S
197 (anti-FLAG). MyD88 binds IRAK4-L, however is less efficient at binding to IRAK4-S (**Figure 2J**).
198 The interaction between IRAK4 isoforms and MyD88 was also confirmed in AML cells that primarily
199 express IRAK4-L (TF1) or IRAK4-S (HL60). Immunoprecipitation of endogenous MyD88 from TF1
200 and HL60 cells resulted in binding of MyD88 to IRAK4-L (**Figure 2K**). In contrast, IRAK4-S was not
201 immunoprecipitated with MyD88 in HL60 cells (**Figure 2K, Supplemental Figure 2E**). Therefore,
202 IRAK4-L mediates maximal activation of NF- κ B signaling through stochastic assembly of the
203 Myddosome complex in AML, while IRAK4-S exhibits reduced NF- κ B activation resulting from
204 diminished MyD88 binding.

205 **IRAK4-L expression and kinase activity is required for leukemic cell function.**

206 The requirement of IRAK4-L in AML was determined by utilizing shRNAs that target IRAK4-
207 L (shIRAK4-L) or both isoforms (shIRAK4-L/S) (**Figure 3A**). THP1 cells, which express both IRAK4
208 isoforms, were transduced with shIRAK4-L or shIRAK4-L/S (**Figure 2F**). Knockdown of IRAK4-L is
209 sufficient to impair the leukemic progenitor function of THP1 cells as evident by formation of fewer
210 colonies in methylcellulose (**Figure 3B**). Moreover, THP1 cells expressing shIRAK4-L or shIRAK4-
211 L/S were xenografted into NSG mice and examined for leukemic burden (**Figure 3A**). Knockdown
212 of IRAK4-L in THP1 cells is sufficient to reduce leukemic burden in xenografted mice (**Figure**
213 **3C,D**). Importantly, the anti-leukemic effects of knocking down IRAK4-L/S was similar to
214 knockdown of IRAK4-L alone, suggesting that IRAK4-L is the dominant isoform in leukemic cells
215 (**Figure 3B-D**). As an orthogonal approach, all IRAK4 isoforms were deleted in THP1 cells using
216 CRISPR/Cas9 gene editing (**Supplemental Figure 3A,B**). Deletion of IRAK4-L/S in THP1 cells
217 (THP1-IRAK4KO) resulted in reduction of leukemic colony formation in methylcellulose-based
218 media (**Supplemental Figure 3C**). Although knockdown or knockout of IRAK4 isoforms did not
219 affect the growth and/or viability of THP1 cells when cultured in liquid media (**Supplemental**
220 **Figure 3D-F**), deletion of IRAK4 resulted in increased expression of the monocyte differentiation
221 marker CD14 and a corresponding decrease in the hematopoietic stem/progenitor marker CD34
222 (**Supplemental Figure 3G**), suggesting that IRAK4-L is required for maintaining a primitive
223 hematopoietic state in THP1 cells. Reduced expression of IRAK4-L (shIRAK4-L) or in combination
224 with IRAK4-S (shIRAK-L/S) did not impair the progenitor function of normal CD34⁺ cells, but rather
225 resulted in a moderate increase in myeloid and erythroid progenitor colonies (**Figure 3E**). These
226 findings demonstrate the requirement of IRAK4-L specifically in leukemia-propagating/initiating
227 cells.

228 To corroborate the requirement of IRAK4 isoforms in AML, IRAK4-L or IRAK4-S were re-
 229 expressed into THP1-IRAK4KO cells (**Figure 3F**). Expression of IRAK4-L in THP1-IRAK4KO cells
 230 restored the leukemic progenitor function as comparable to control THP1 cells (THP1 +
 231 vector)(**Figure 3G,H**). In contrast, IRAK4-S was unable to rescue the leukemic progenitor defect of
 232 THP1-IRAK4KO cells (**Figure 3G,H**). As a complimentary approach, restoration of IRAK4-L in
 233 THP1 cells expressing shRNAs targeting both IRAK4 isoforms (shIRAK4-L/S) resulted in leukemic
 234 colony formation comparable to control-transduced THP1 cells, while IRAK4-S was unable to
 235 rescue the leukemic progenitor defect of THP1 cells expressing shIRAK4-L/S (**Supplemental**
 236 **Figure 4A-C**). We next asked whether IRAK4 kinase activity was required for leukemic cells
 237 expressing IRAK4-L by using an ATP-competitive IRAK4 kinase inhibitor (CA-4948) approved for
 238 clinical use (NCT0332878)¹⁹. MDSL, TF1, and THP1 cells, which preferentially express IRAK4-L,
 239 formed fewer leukemic progenitor colonies with CA-4948 (**Figure 3I**). In contrast, HL60 and F36P,
 240 or normal CD34+ cells, which express IRAK4-S, were not affected by treatment with CA-4948
 241 (**Figure 3I**). Thus, inhibition of IRAK4 kinase activity in MDS/AML cell lines that express IRAK4-L
 242 resulted in diminished leukemic function *in vitro* as compared to cells that preferentially express
 243 IRAK4-S (**Figure 3J**). To demonstrate that IRAK4 inhibition can suppress leukemic cell function *in*
 244 *vivo*, THP1 were xenografted into NSG mice. Following treatment with IRAK4 inhibitor (**Figure 3K**),
 245 THP1 infiltration in the BM, spleen, and liver was reduced as compared to mice receiving vehicle
 246 control (**Figure 3L,M**). Collectively, these results establish that IRAK4-L, but not IRAK4-S,
 247 signaling programs are required for leukemia-propagating cells, and that inhibition of IRAK4
 248 abrogates leukemic growth in AML cells expressing IRAK4-L.

249 **IRAK4-L expression is associated with U2AF1 mutations in MDS/AML.**

250 We next sought to identify the genetic alterations associated with IRAK4 exon 4 inclusion in
 251 AML. Among all AML patients examined in TCGA, inclusion of IRAK4 exon 4 was significantly
 252 correlated with mutations in U2 small nuclear RNA auxiliary factor 1 (U2AF1), the gene encoding
 253 an RNA-binding protein critical for 3' splice-acceptor site recognition in pre-mRNAs (**Figure 4A**).
 254 Given the prevalence of U2AF1 mutations in MDS, we examined IRAK4 exon 4 usage in MDS
 255 CD34+ cells. All of the MDS patients with U2AF1 mutations (S34F, R156H, and Q157P/R)
 256 expressed *IRAK4-L* (6/6 by RNA-sequencing; and 5/5 by RT-PCR), while 9 of 17 (52%) MDS
 257 patients without splicing factor mutations expressed *IRAK4-L* (MDS-SF wild-type) (**Figure 4B-D**).
 258 In contrast, less than 20% of healthy age-matched BM CD34+ cells expressed *IRAK4-L* (**Figure**
 259 **4B-D**). As observed in AML, MDS patients expressing increased levels of *IRAK4-L* had a worse

260 prognosis as compared to MDS patients with lower levels of *IRAK4*-L (**Supplemental Figure**
261 **4D**).²⁰

262 The consensus sequence flanking the adenine-guanine dinucleotide at the 3' splice
263 acceptor site is recognized by U2AF1 and impacts exon usage by the U2AF1 mutants.^{21,22} A
264 cytosine (C) or adenine (A) at the -3 position correlates with exon inclusion by U2AF1-S34F as
265 compared to wild-type U2AF1.^{23,24} The -3 position of *IRAK4* exon 4 contains an adenine,
266 supporting the model in which U2AF1-S34F mediates exon inclusion (**Figure 4E**). Moreover, the
267 +1 position of *IRAK4* exon 4 contains a guanine, which is associated with exon inclusion of U2AF1-
268 Q157X as compared to wild-type U2AF1.²¹ These findings suggest that mutant U2AF1 correlates
269 with inclusion of exon 4 and preferential expression of *IRAK4*-L in AML and MDS CD34+ cells.

270 **U2AF1-S34F induces *IRAK4*-L and innate immune pathway activation.**

271 To investigate whether mutant U2AF1 directly regulates *IRAK4* exon 4 usage and
272 expression of *IRAK4*-L, RNA-sequencing was performed on healthy CD34+ cells expressing wild-
273 type or U2AF1-S34F (**Figure 5A**). Importantly, expression of U2AF1-S34F resulted in increased
274 reads at the *IRAK4* exon 3-4 junction as compared to U2AF1-WT, indicating that mutant U2AF1
275 mediates inclusion of exon 4 (**Figure 5A**).²⁵ To gain further insight into the regulation of *IRAK4*
276 exon 4, we generated an *IRAK4* exon 4 cassette splicing reporter (HEK293-*IRAK4*exon4)(**Figure**
277 **5B**). A mutant version of the *IRAK4* exon 4 cassette ("TAG") was also generated containing a
278 thymine (T) in place of adenine (A) at the -3 position ("AAG") since U2AF1-S34F mediates exon
279 inclusion in the presence of AAG but not TAG (**Figure 5B**). Transfection of U2AF1-S34F into
280 HEK293-*IRAK4*exon4 cells resulted in inclusion of *IRAK4* exon 4 from the WT "AAG" splicing
281 reporter (11.7% exon 4 inclusion) compared to U2AF1-WT (1.3 % exon 4 inclusion) or vector-
282 transfected HEK293-*IRAK4*exon4 cells (1.4% exon 4 inclusion)(**Figure 5C**). Conversely, U2AF1-
283 S34F did not induce inclusion of *IRAK4* exon 4 from the mutant "TAG" splicing reporter (0.4% exon
284 4 inclusion; n = 2). Therefore, U2AF1-S34F likely mediates inclusion of *IRAK4* exon 4 by direct
285 binding. To extend these observations to human AML cells, we utilized K562 erythroleukemia cells
286 that express doxycycline-inducible FLAG-tagged U2AF1-S34F or U2AF1-WT.²⁶ Expression of
287 U2AF1-WT with doxycycline did not affect *IRAK4* exon 4 usage or protein expression (**Figure**
288 **5D,E**). In contrast, expression of U2AF1-S34F after doxycycline resulted in robust and dose-
289 dependent *IRAK4* exon 4 inclusion (**Figure 5D**), which coincided with increased expression of
290 *IRAK4*-L protein (**Figure 5E**).

291 Since U2AF1-S34F expression results in expression of *IRAK4*-L, and expression of *IRAK4*-L
292 alone is sufficient to mediate activation of innate immune and NF- κ B signaling, we next

investigated the consequences of U2AF1-S34F on NF- κ B and MAPK activation in AML cells. Expression of U2AF1-S34F in K562 cells resulted in activation of MAPK signaling (p-p38 and p-ERK) and NF- κ B signaling (p-IRAK1, p-IKK, and p-p65) as compared to wild-type U2AF1-expressing K562 cells (**Figure 5F**). U2AF1-S34F-expressing K562 cells are also hypersensitive to IL1 β stimulation as indicated by increased expression of NF- κ B (*TNFAIP3*, *IL6*, *CXCL2*, and *CXCL8*) and MAPK (*Jun* and *Egr1*) target genes (**Figure 5G**). As such, U2AF1-S34F is sufficient to induce expression of IRAK4-L and consequently innate immune pathway activation.

U2AF1-S34F AML cells are sensitive to IRAK4 inhibitors.

We determined whether inhibiting IRAK4 kinase activity was sufficient for suppressing NF- κ B and/or MAPK signaling in U2AF1-mutant AML cells expressing IRAK4-L. To inhibit IRAK4 kinase activity, we evaluated two structurally-unrelated IRAK4 inhibitors, IRAK1/4-Inh²⁷ and CA-4948.¹⁹ When K562 cells expressing U2AF1-S34F (K562-U2AF1-S34F) were treated with IRAK1/4-Inh, the activation of NF- κ B signaling was inhibited as indicated by reduced p-IRAK1 and p-IKK (**Figure 6A,B, and Supplemental Figure 5**). In contrast, MAPK signaling was not diminished following treatment with IRAK1/4-Inh. These findings suggest that MAPK activation not only occurs downstream of IRAK4-L, but also secondary to RNA splicing changes.²⁸ Based on these observations, we reasoned that inhibition of IRAK4-L catalytic function will suppress pro-survival signaling via NF- κ B, thus exposing a therapeutic vulnerability in U2AF1-mutant AML/MDS. K562-U2AF1-S34F cells treated with IRAK4 inhibitors (IRAK1/4-Inh or CA-4948) exhibited increased cytotoxicity (Trypan blue-positive cells) as compared to vehicle-treated cells (**Figure 6C**), which correlated with elevated AnnexinV/PI-positive cells (**Figure 6D,E**). In contrast, K562-U2AF1 were not sensitive to the IRAK4 inhibitors (**Figure 6C-E**). In addition, K562-U2AF1-S34F cells formed fewer leukemic progenitor colonies when treated with IRAK4 inhibitors as compared to vehicle-treated cells (**Figure 6F**). Next, primary patient-derived cells from MDS and AML patients with U2AF1 mutations were grown in clonogenic assays after siRNA mediated knockdown or pharmacologic inhibition of IRAK4. Both IRAK4 knockdown and inhibition with CA-4948 led to enhanced erythroid and myeloid differentiation (**Figure 6G-I; Supplemental Figure 6A-C**). In contrast, normal CD34⁺ cells treated with CA-4948 did not exhibit altered differentiation (**Supplemental Figure 6D**).

To determine the efficacy of IRAK4 inhibitors on primary U2AF1-mutant cells *in vivo*, xenografts from MDS patients with U2AF1 mutations were established in NSG mice and treated with CA-4948 (12.5 mg/kg/d) or vehicle (**Figure 6G**). While vehicle-treated mice exhibited a 3.2-

fold increased expansion of xenografted MDS cells in the BM, IRAK4 inhibition led to a 50% decrease in MDS cell engraftment after 3 weeks of CA-4948 treatment (**Figure 6J,K**). To establish that inhibition of IRAK4 suppresses the function of disease-propagating MDS cells, BM cells were isolated from primary xenografted mice treated with vehicle or CA-4948 for 3 weeks (from Figure 6J,K), and then transplanted into secondary recipient NSG mice. Mice transplanted with vehicle-treated MDS cells engrafted into secondary recipient mice (**Figure 6L,M**). In contrast, the engraftment of CA-4948-treated MDS cells was negligible in secondary recipients (**Figure 6L,M**), suggesting that inhibition of IRAK4-L impairs U2AF1-mutant MDS propagating cells *in vivo*. In parallel, we evaluated potential hematologic toxicity after repeated administration of CA-4948 in mice. Administration of CA-4948 for 12 weeks in NSG mice did not affect body weight (not shown) and hematologic blood parameters (**Supplemental Figure 6E**). These findings indicate that U2AF1-S34F-expressing MDS/AML cells have acquired a dependency on IRAK4-L and are sensitive to IRAK4 inhibition.

Discussion

Clinically-relevant molecular subtypes of AML were identified based on differential RNA isoform usage. One major subtype of AML was enriched for RNA isoform changes affecting innate immune pathway genes. Among these genes, an isoform of IRAK4 that retains exon 4 and encodes IRAK4-L is expressed in AML and U2AF1-mutant MDS/AML. While consistent with the current paradigm of Myddosome signaling, our findings uncovered a mechanism by which IRAK4-L is generated by alternative splicing in MDS/AML to mediate maximal activation of NF- κ B through MyD88 binding. In contrast, healthy hematopoietic cells primarily express IRAK4-S and have diminished binding to MyD88 and lower NF- κ B activation potential. Thus, IRAK4-S is normally expressed to mitigate spurious and/or excessive NF- κ B activation as a result of its diminished ability to bind MyD88. This observation may reveal an evolutionarily-conserved mechanism to regulate innate immune responses. Overexpression of immune-related genes is widely reported in MDS and AML,^{29,30} and chronic innate immune pathway activation, predominantly via TLRs, increases the risk of developing myeloid malignancies.³¹ Evidence that innate immune pathway activation is causal in MDS or AML is supported by work related to miR-146a and TIFAB, which reside within the deleted region in del(5q) MDS/AML. Deletion of miR-146a and/or TIFAB results in derepression of their targets, TRAF6 and IRAK1, and impaired hematopoiesis and BM failure in mice³²⁻³⁴. Phosphorylation of IRAK1 is observed in non-del(5q) MDS/AML and other cancers, however the mechanism of IRAK1 activation remains unidentified.³⁵⁻⁴¹ In some of these cases, we

358 propose that isoform switching from IRAK4-S to IRAK4-L is the molecular basis of IRAK1
359 activation. While there are examples of indirect mechanisms that result in activation of innate
360 immune signaling in MDS/AML, the present study provides direct genetic evidence, via mutations
361 in the splicing factor U2AF1, that links oncogenic innate immune signaling and IRAK4 activation in
362 MDS/AML. Although all U2AF1-mutant MDS and AML exhibit expression of IRAK4-L,
363 approximately 50% of remaining patients express IRAK4-L indicating that other genetic and/or
364 molecular mechanisms regulate alternative splicing of IRAK4 and expression of IRAK4-L. IRAK4-L
365 correlated with mutations in U2AF1, but it did not correlate with other splicing factor mutations in
366 MDS or AML. For example, mutations in SRSF2 also result in activation of NF- κ B signaling, but
367 these effects are independent of IRAK4 isoform regulation⁴².

368 Mutations in RNA splicing factors are the most frequent class of mutation in MDS and
369 common in AML.^{2,5,9,43-45} Not surprisingly, mouse models with splicing factor gene mutations
370 develop features of MDS.^{3,26,46-49} Mutations affecting the genes encoding core spliceosome
371 machinery, such as in SF3B1, U2AF1, and SRSF2 do not result in loss-of-function phenotypes, but
372 rather in selective splicing changes due to aberrant splice site recognition.^{3,5,21,25,48-50} Although
373 many RNA splicing changes in U2AF1-mutant cells result in aberrant RNA splicing, non-functional
374 transcripts, and/or nonsense-mediated decay, IRAK4 is an example in which mutant U2AF1
375 generates an isoform with oncogenic activity. Given the prevalence of RNA splicing gene
376 mutations in hematologic malignancies, effective strategies to therapeutically target spliceosome
377 mutant malignancies are needed. Small molecule modulators of the spliceosome target normal
378 splicing by interfering with the splicing machinery and altering 3' splice site recognition. Based on
379 our reported findings, an alternative strategy to target U2AF1-mutant MDS/AML is by exposing the
380 therapeutic vulnerability of IRAK4-L expression with IRAK1/4 inhibitors. However, elucidating the
381 mechanism of IRAK4 isoform control during normal development, inflammation, and in non-splicing
382 factor mutant cancers will better inform of disease conditions and patients most suitable for
383 treatment with IRAK1.4 inhibitors. Taken together, we report that mutations in U2AF1 induce
384 expression of targetable "active" IRAK4 isoforms and provide a genetic link to activation of
385 oncogenic innate immune signaling in MDS and AML.

386
387

388 **Conflicts of Interest**

389 A.V. has received research funding from GlaxoSmithKline, Incyte, MedPacto, Novartis, Curis and
390 Eli Lilly and Company, has received compensation as a scientific advisor to Novartis, Stelexis

391 Therapeutics, Acceleron Pharma, and Celgene, and has equity ownership in Stelexis
392 Therapeutics.

393
394 **Acknowledgments**

395 This work was supported by Cincinnati Children’s Hospital Research Foundation, Leukemia
396 Lymphoma Society, and National Institute of Health (R35HL135787, RO1DK102759,
397 RO1DK113639), and Edward P. Evans Foundation grants to DTS. AV is supported by National
398 Institute of Health (R01HL139487 and R01DK103961), Leukemia and Lymphoma Society, and
399 EvansMDS grants. AP and JB are supported by Bloodwise (UK, grant 13042). MAS is supported
400 by a National Institute of Health Research Training and Career Development Grant
401 (F31HL132420). IDCF is supported by the Intramural Research Program of NIAID. We also thank
402 Robert Boohar (Curis Inc.) for suggestions. We thank Melissa Scott and Yueh-Chiang Hu from the
403 Gene Editing Core at CCHMC for their assistance.

404
405
406
407
408
409
410
411
412
413
414

Figure Legends

Figure 1. Differential RNA isoform usage correlates with AML prognosis and oncogenic IRAK4 isoforms. (A) Scatterplot of cumulative and isoform variances of genes in AML samples from TCGA (n = 160). Every point represents a gene, colored by the expression correlation of the most negatively correlated pair of isoforms for that gene. Blue-colored genes have at least one pair of isoforms with mutually exclusive expression pattern. (B) Hierarchical clustering analysis and relative expression of mRNA isoforms in AML samples. Relative expression is such that a value of 1 indicates that the isoform is the only isoform expressed for the given gene, and 0 indicates that the isoform is not expressed. (C) Pathway analysis of genes in Group 2 (from Figure 1B and Supplemental Figure 1C) associated with worse clinical outcome (n = 347 genes) determined by hypergeometric distribution test. (D) Kaplan-meier analysis of AML patients stratified on IRAK4-L (exon 4 included) or IRAK4-S (exon 4 excluded) expression. (E) Exon architecture of IRAK4 and protein domains (below). Sashimi plots represent junction reads in representative AML samples. Exon 4 reads were normalized to the total number of reads in each sample and plotted as a ratio (inclusion/total and exclusion/total). (F) Cumulative expression of IRAK4-L and IRAK4-S in individual AML patients (TCGA). Patients are ordered according the relative expression of IRAK4-L versus IRAK4-S. (G) Relative expression of IRAK4-L to IRAK4-S in normal BM (n = 4) and AML samples (TCGA; n = 160) is shown as box plots including the center (mean) and top and bottom quartiles. Wilcoxon Test, P = 0.07. (H) RT-PCR of IRAK4-L/S using primers flanking exon 4. Densitometric quantification of IRAK4 exon 4 inclusion calculated as the ratio of IRAK4-L versus both isoforms is shown below (three independent experiments). (I) Immunoblot of IRAK4 using an N-terminal antibody that recognizes IRAK4-L and a C-terminal antibody that recognizes IRAK4-L and IRAK4-S (three independent experiments). (J) Immunoblot of IRAK4 in AML patient samples (AML1 and AML2) and healthy samples (cord blood (CB), bone marrow CD34+ (BM-CD34+), BM mononuclear cells (BM-MNC)). The healthy samples are collected from independent donors. The percent of IRAK4-L of total IRAK4 isoforms is shown below.

Figure 2. IRAK4-L results in Myddosome assembly and maximal activation of innate immune signaling. (A) Pathway analysis of enriched genes in AML patients preferentially expressing IRAK4-L (top) or IRAK4-S (bottom). The pathway scores were calculated by NetWalker⁵¹ based on a random walk method⁵². The scores directly reflect the enrichment of the indicated pathways for high (positive scores) or low (negative scores) expression of the IRAK4 Long isoform. The pathways shown here were the highest-scoring pathways. (B) HEK293 cells

transfected with vector (pcDNA3.1), or FLAG-tagged IRAK4-L or IRAK4-S for 24 hours and immunoblotted for the indicated NF- κ B and MAPK proteins (three independent experiments). (C) Densiometric analysis of panel (B). One-sided t-tests were used for statistical analyses. (D-E) NF- κ B (D) and AP1 (E) activation was measured by κ B-site and AP1-site containing reporter assays in HEK293 cells transfected with empty vector, IRAK4-L, or IRAK4-S. Values are normalized to Renilla-luciferase and empty vector (1.0) (3 or 6 independent experiments, respectively). Two-sided t-tests were used for statistical analyses. (F) Representative PCR and immunoblot of THP1 cells following knockdown of IRAK4-L and IRAK4-S (shIRAK4-L/S) or just IRAK4-L (shIRAK4-L) (three independent experiments). (G) NF- κ B activation was measured in THP1 cells stably expressing a κ B-site containing reporter and shRNAs targeting IRAK4-L (left) or IRAK4-L/S (right) after stimulation with PAM3CSK4 (+, 500 ng/ml or ++, 1000 ng/ml). (3 or 5 independent experiments, respectively). Two-sided t-tests were used for statistical analyses. (H) Overview of IRAK4-L and IRAK4-S downstream signaling based on Figure 2A-E. (I) Sequence-based prediction of protein-protein and domain-domain interactions using all isoforms of IRAK1, IRAK4, and MyD88. Superscripts indicate distinct RNA isoforms encoding the indicated long or short proteins. (J) HEK293 cells transfected with HA-MyD88 and either FLAG-IRAK4-L or FLAG-IRAK4-S were immunoblotted for HA-MyD88 after immunoprecipitation of FLAG. (representative of two independent experiments). (K) TF1 and HL60 cells were immunoblotted for IRAK4 with C-terminal antibody after immunoprecipitation of MyD88 (representative of two independent experiments). *, indicate IgG bands. All data represent the mean $\pm$ s.e.m.

Figure 3. IRAK4-L is required for leukemic cell function. (A) Overview of experimental design. (B) Colony formation of THP1 cells expressing shIRAK4-L, shIRAK4-L/S, or a control shRNA (shControl) (three independent experiments). Two-sided t-tests were used for statistical analyses. (C-D) THP1 cells expressing shIRAK4-L (n = 9 animals), shIRAK4-L/S (n = 5 animals), or a control shRNA (shControl; n = 6 animals) were engrafted into sublethally-irradiated NSG mice. After 6 weeks, leukemic burden (GFP+ cells) in the BM (C) and PB (D) was determined. Two-sided t-tests were used for statistical analyses. (E) Cord-blood CD34+ cells were transduced with shIRAK4-L, shIRAK4-L/S, or a control shRNA (shControl) and examined for IRAK4 knockdown (left; two independent experiments) and progenitor colony formation in methylcellulose (right; six independent experiments). Two-sided t-tests were used for statistical analyses. (F) THP1 cells transduced with empty retroviral vector (MSCV-IRES-GFP as a control) or THP1 IRAK4-KO cells transduced with retroviral vectors encoding IRAK4-L (left panel) or IRAK4-S (right panel) and then

immunoblotted with the N-terminal IRAK4 antibody (left panel) and C-terminal IRAK4 antibody (right panel). **(G-H)** Representative images of parental and IRAK4-KO THP1 cells examined for leukemic progenitor function in methylcellulose (three biological replicates). Scale bar, 700 microns. Two-sided t-tests were used for statistical analyses. **(I)** Colony formation of MDS/AML cell lines and control CD34+ cells (2 independent donors) treated with DMSO or CA-4948 for 7-10 days (three independent experiments). Two-sided t-tests were used for statistical analyses. **(J)** Graph representing the IC50 relative to the ratio of IRAK4-L to IRAK4-S expression in the indicated cell lines and CD34+ cells summarized from Figure 1I and 3I. **(K)** THP1 cells were xenografted in NSG mice (n = 10) and treated with IRAK4 inhibitor (CA-4948) or vehicle at 12.5 mg/kg/5d/week for 5 weeks. **(L)** Mice xenografted with THP1 cells were assessed for leukemic engraftment in BM, and spleen and liver weight. Two-sided t-tests were used for statistical analyses. **(M)** Representative spleens are shown isolated from mice xenografted with THP1 cells at time of death after treatment with vehicle or CA-4948. All data represent the mean $\pm$ s.e.m.

Figure 4. IRAK4-L expression is associated with U2AF1 mutations in MDS and AML. **(A)** Heatmap was generated based on the Z-score correlation of individual exon's expression to all somatic mutations. Each indicated mutation was regressed against the relative expression (after adjusting for total IRAK4 expression) of each IRAK4 exon (RPKM) in a multiple linear regression model. Shown are z-scores of partial correlations. A z-score of greater than |1.96| corresponds to $P < 0.05$. Six out of 160 AML patients exhibit U2AF1 mutations ($Z = 2.43$; *, $P = 0.01$). **(B)** Sashimi plots representing IRAK4 exon 4 inclusion or exclusion in healthy CD34+ cells (n = 7 samples), CD34+ cells isolated from MDS patients with no splicing factor (SF^{WT}) mutations (MDS- SF^{WT} ; n = 6 samples), and from patients with mutation (mut) in U2AF1 (MDS-U2AF1^{mut}; n = 6 samples) based on RNA-sequencing junction reads. Quantification of the junction reads for exon 4 are shown for healthy controls (n = 7), MDS- SF^{WT} (n = 6), and MDS-U2AF1^{mut} (n = 6). Exon 4 inclusion and exclusion reads were normalized to the total number of reads in each sample and plotted as a ratio (inclusion/total and exclusion/total). Two-sided t-tests were used for statistical analyses. **(C)** RT-PCR analysis of healthy CD34+ cells (n = 7 samples), CD34+ cells isolated from MDS patients with no splicing factor mutations (MDS- SF^{WT} ; n = 17 samples), and from patients with mutation in U2AF1 (MDS-U2AF1^{mut}; n = 5 samples) using primers flanking IRAK4 exon 4. Densitometric quantification of IRAK4 exon 4 inclusion calculated as the ratio of the long isoform versus both isoforms is shown below each sample. **(D)** Densitometric quantification of IRAK4 exon 4 inclusion is summarized from panel (C). The number in parentheses indicates the number of samples

exhibiting >10% expression of IRAK4-L relative to all IRAK4 isoforms. Two-sided t-tests were used for statistical analyses. (E) Consensus sequence motifs identified at the distal 3' splice sites of exon exclusion events in U2AF1-S34F cells as compared to U2AF1 using publicly available data and visualized with weblogo software. Above the U2AF1-S34F consensus motifs is the sequence of the distal 3' splice site of IRAK4 exon 4. All data represent the mean $\pm$ s.e.m.

Figure 5. U2AF1-S34F induces expression of IRAK4-L and increased innate immune pathway activation. (A) Experimental design to analyze RNA splicing changes in normal CD34+ cells expressing U2AF1 or U2AF1-S34F (left). Exon 3-4 junction reads of IRAK4 in U2AF1 and U2AF1-S34F CD34+ cells (three independent samples) (right). Two-sided t-tests were used for statistical analyses. (B) Schematic of IRAK4 exon 4 splicing reporter. Wild-type IRAK4 exon 4 (AAG at -3 position) or mutant IRAK4 exon 4 (TAG at -3 position) and 100 bp of flanking introns were cloned into a splicing reporter (pFlare5A)⁵³ and expressed into HEK293 cells (HEK293-IRAK4exon4). (C) Representative PCR of IRAK4 exon 4 splicing in HEK293-IRAK4exon4 expressing U2AF1 and U2AF1-S34F (five independent experiments for AAG; two independent experiments for TAG). (D) (Top) Experimental design to measure IRAK4 RNA and protein isoform expression in K562 cell lines that have stably integrated doxycycline-inducible FLAG-U2AF1 or FLAG-U2AF1-S34F. (Bottom) IRAK4 RNA isoform expression was measured using primers flanking exon 4 in doxycycline (DOX)-induced K562 cells expressing FLAG-U2AF1 or FLAG-U2AF1-S34F. Top band indicates IRAK4-L and bottom band indicates IRAK4-S (five independent experiments). (E) IRAK4 protein isoform expression analyzed with a C-terminal IRAK4 antibody in DOX-induced K562 cell expressing FLAG-U2AF1 or FLAG-U2AF1-S34F (three independent experiments). (F) The indicated NF- κ B and MAPK protein expression was measured in DOX-induced K562 cells expressing FLAG-U2AF1 or FLAG-U2AF1-S34F by immunoblotting (two independent experiments). (G) Expression of the indicated NF- κ B (*TNFAIP3*, *IL6*, *CXCL2*, and *CXCL8*) and MAPK (*Jun* and *Egr1*) target genes was measured in DOX-induced K562 cells expressing FLAG-U2AF1 or FLAG-U2AF1-S34F after IL-1 β stimulation (10 ng/ml) for 30 or 60 min (three independent experiments). Two-sided t-tests were used for statistical analyses. All data represent the mean $\pm$ s.e.m.

Figure 6. U2AF1-S34F AML cells are sensitive to IRAK4 inhibitors. (A) K562-U2AF1-S34F cells were treated with 10 μ M IRAK1/4-inhibitor for 1 or 2 hours and immunoblotted for NF- κ B and

546 MAPK activation. **(B)** Densitometric analysis of panel (A) summarized from three independent
547 biological replicates. Two-sided t-tests were used for statistical analyses. **(C)** K562-U2AF1 cells
548 were treated with DMSO, 10 μ M IRAK1/4-Inh, or 10 μ M CA-4948 for 7 days and assessed for
549 viability by Trypan Blue exclusivity (three independent experiments). One-sided t-tests were used
550 for statistical analyses. **(D)** Representative images of K562 cells expressing wild-type U2AF1 or
551 U2AF1-S34F were treated with DMSO or 10 μ M CA-4948 for 48 hours days and then analyzed by
552 flow cytometry for AnnexinV and Propidium Iodide (PI) staining (three independent experiments).
553 **(E)** Summary of (D). Two-sided t-tests were used for statistical analyses.**(F)** K562 cells treated with
554 10 μ M IRAK1/4-Inh (five independent experiments) or CA-4948 (three independent experiments)
555 were evaluated after 7 days for colony formation in methylcellulose. Two-sided t-tests were used
556 for statistical analyses. **(G)** Schematic of experimental design. **(H)** MDS patient-derived BM cells
557 were evaluated after 7 days for colony formation in methylcellulose treated with 0.5 μ M CA-4948 or
558 vehicle (two independent experiments). **(I)** MDS patient-derived BM cells were transfected with
559 siIRAK4 or control siRNA and then evaluated after 7 days for colony formation in methylcellulose
560 (two independent experiments). **(J)** NSG mice were xenografted with U2AF1-mutant MDS BM
561 cells. After engraftment, mice were treated with CA-4948 or vehicle and human cell engraftment
562 was assessed by flow cytometry on BM aspirates. **(K)** Summary of BM cell engraftment (hCD45)
563 for independent mice and patient-derived samples. Two-sided t-tests were used for statistical
564 analyses. **(L)** MDS BM cells were isolated from mice treated with vehicle (PBS) or CA-4948
565 treatment after 3 weeks (panel J,K) and then xenografted into secondary NSG mice. Human cell
566 engraftment was assessed by flow cytometry on BM aspirates. **(M)** Summary of BM cell
567 engraftment (hCD45) for mice and 2 independent patient-derived samples. Data represent the
568 mean $\pm$ s.e.m. for all panels except **H** and **I** where data represent the mean $\pm$ s.d.

569
570

References

1. Vitting-Seerup, K. & Sandelin, A. The Landscape of Isoform Switches in Human Cancers. *Mol Cancer Res* **15**, 1206-1220 (2017).
2. Inoue, D., Bradley, R.K. & Abdel-Wahab, O. Spliceosomal gene mutations in myelodysplasia: molecular links to clonal abnormalities of hematopoiesis. *Genes Dev* **30**, 989-1001 (2016).
3. Kim, E., et al. SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition. *Cancer Cell* **27**, 617-630 (2015).
4. Zhang, J., et al. Disease-associated mutation in SRSF2 misregulates splicing by altering RNA-binding affinities. *Proc Natl Acad Sci U S A* **112**, E4726-4734 (2015).
5. Saez, B., Walter, M.J. & Graubert, T.A. Splicing factor gene mutations in hematologic malignancies. *Blood* **129**, 1260-1269 (2017).
6. Graubert, T.A., et al. Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. *Nat Genet* **44**, 53-57 (2011).
7. Hirsch, C.M., et al. Molecular features of early onset adult myelodysplastic syndrome. *Haematologica* **102**, 1028-1034 (2017).
8. Papaemmanuil, E., et al. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. *N Engl J Med* **365**, 1384-1395 (2011).
9. Pellagatti, A. & Boultonwood, J. Splicing factor gene mutations in the myelodysplastic syndromes: impact on disease phenotype and therapeutic applications. *Adv Biol Regul* **63**, 59-70 (2017).
10. Papaemmanuil, E., et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood* **122**, 3616-3627; quiz 3699 (2013).
11. Haferlach, T., et al. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia* **28**, 241-247 (2014).
12. Lindsley, R.C., et al. Acute myeloid leukemia ontogeny is defined by distinct somatic mutations. *Blood* **125**, 1367-1376 (2015).
13. Ley, T.J., et al. DNMT3A mutations in acute myeloid leukemia. *N Engl J Med* **363**, 2424-2433 (2010).
14. Barreyro, L., Chlon, T.M. & Starczynowski, D.T. Chronic immune response dysregulation in MDS pathogenesis. *Blood* **132**, 1553-1560 (2018).
15. Patra, M.C. & Choi, S. Recent Progress in the Molecular Recognition and Therapeutic Importance of Interleukin-1 Receptor-Associated Kinase 4. *Molecules* **21**(2016).
16. Yang, X., et al. Widespread Expansion of Protein Interaction Capabilities by Alternative Splicing. *Cell* **164**, 805-817 (2016).
17. Dossang, A.C., et al. The N-terminal loop of IRAK-4 death domain regulates ordered assembly of the Myddosome signalling scaffold. *Sci Rep* **6**, 37267 (2016).
18. Ferrao, R., et al. IRAK4 dimerization and trans-autophosphorylation are induced by Myddosome assembly. *Mol Cell* **55**, 891-903 (2014).
19. Booher RN, S.M., Xu G, Cheng H, Tuck DP. Efficacy of the IRAK4 inhibitor CA-4948 in patient-derived xenograft models of diffuse large B cell lymphoma. in *American Association for Cancer Research* (Washington DC, 2017).
20. Gerstung, M., et al. Combining gene mutation with gene expression data improves outcome prediction in myelodysplastic syndromes. *Nat Commun* **6**, 5901 (2015).
21. Ilagan, J.O., et al. U2AF1 mutations alter splice site recognition in hematological malignancies. *Genome Res* **25**, 14-26 (2015).
22. Brooks, A.N., et al. A pan-cancer analysis of transcriptome changes associated with somatic mutations in U2AF1 reveals commonly altered splicing events. *PLoS One* **9**, e87361 (2014).

23. Fei, D.L., *et al.* Wild-Type U2AF1 Antagonizes the Splicing Program Characteristic of U2AF1-Mutant Tumors and Is Required for Cell Survival. *PLoS Genet* **12**, e1006384 (2016).
24. Yip, B.H., *et al.* The U2AF1S34F mutation induces lineage-specific splicing alterations in myelodysplastic syndromes. *J Clin Invest* **127**, 2206-2221 (2017).
25. Okeyo-Owuor, T., *et al.* U2AF1 mutations alter sequence specificity of pre-mRNA binding and splicing. *Leukemia* **29**, 909-917 (2015).
26. Shirai, C.L., *et al.* Mutant U2AF1-expressing cells are sensitive to pharmacological modulation of the spliceosome. *Nat Commun* **8**, 14060 (2017).
27. Powers, J.P., *et al.* Discovery and initial SAR of inhibitors of interleukin-1 receptor-associated kinase-4. *Bioorg Med Chem Lett* **16**, 2842-2845 (2006).
28. Maimon, A., *et al.* Mnk2 alternative splicing modulates the p38-MAPK pathway and impacts Ras-induced transformation. *Cell Rep* **7**, 501-513 (2014).
29. Hofmann, W.K., *et al.* Characterization of gene expression of CD34+ cells from normal and myelodysplastic bone marrow. *Blood* **100**, 3553-3560 (2002).
30. Pellagatti, A., *et al.* Deregulated gene expression pathways in myelodysplastic syndrome hematopoietic stem cells. *Leukemia* **24**, 756-764 (2010).
31. Kristinsson, S.Y., *et al.* Chronic immune stimulation might act as a trigger for the development of acute myeloid leukemia or myelodysplastic syndromes. *J Clin Oncol* **29**, 2897-2903 (2011).
32. Starczynowski, D.T., *et al.* Identification of miR-145 and miR-146a as mediators of the 5q-syndrome phenotype. *Nat Med* **16**, 49-58 (2010).
33. Varney, M.E., *et al.* Epistasis between TIFAB and miR-146a: neighboring genes in del(5q) myelodysplastic syndrome. *Leukemia* **31**, 1659 (2017).
34. Varney, M.E., *et al.* Loss of Tifab, a del(5q) MDS gene, alters hematopoiesis through derepression of Toll-like receptor-TRAF6 signaling. *J Exp Med* **212**, 1967-1985 (2015).
35. Rhyasen, G.W., *et al.* Targeting IRAK1 as a therapeutic approach for myelodysplastic syndrome. *Cancer Cell* **24**, 90-104 (2013).
36. Beverly, L.J. & Starczynowski, D.T. IRAK1: oncotarget in MDS and AML. *Oncotarget* **5**, 1699-1700 (2014).
37. Dussiau, C., *et al.* Targeting IRAK1 in T-cell acute lymphoblastic leukemia. *Oncotarget* **6**, 18956-18965 (2015).
38. Li, Z., *et al.* Inhibition of IRAK1/4 sensitizes T cell acute lymphoblastic leukemia to chemotherapies. *J Clin Invest* **125**, 1081-1097 (2015).
39. Goh, J.Y., *et al.* Chromosome 1q21.3 amplification is a trackable biomarker and actionable target for breast cancer recurrence. *Nat Med* **23**, 1319-1330 (2017).
40. Adams, A.K., *et al.* IRAK1 is a novel DEK transcriptional target and is essential for head and neck cancer cell survival. *Oncotarget* **6**, 43395-43407 (2015).
41. Wee, Z.N., *et al.* IRAK1 is a therapeutic target that drives breast cancer metastasis and resistance to paclitaxel. *Nat Commun* **6**, 8746 (2015).
42. Lee, S.C., *et al.* Synthetic Lethal and Convergent Biological Effects of Cancer-Associated Spliceosomal Gene Mutations. *Cancer Cell* **34**, 225-241 e228 (2018).
43. Yoshida, K., *et al.* Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478**, 64-69 (2011).
44. Makishima, H., *et al.* Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis. *Blood* **119**, 3203-3210 (2012).
45. Ogawa, S. Splicing factor mutations in AML. *Blood* **123**, 3216-3217 (2014).
46. Kon, A., *et al.* Physiological Srsf2 P95H expression causes impaired hematopoietic stem cell functions and aberrant RNA splicing in mice. *Blood* **131**, 621-635 (2018).

47. Komeno, Y., *et al.* SRSF2 Is Essential for Hematopoiesis, and Its Myelodysplastic Syndrome-Related Mutations Dysregulate Alternative Pre-mRNA Splicing. *Mol Cell Biol* **35**, 3071-3082 (2015).
48. Obeng, E.A., *et al.* Physiologic Expression of Sf3b1(K700E) Causes Impaired Erythropoiesis, Aberrant Splicing, and Sensitivity to Therapeutic Spliceosome Modulation. *Cancer Cell* **30**, 404-417 (2016).
49. Mupo, A., *et al.* Hemopoietic-specific Sf3b1-K700E knock-in mice display the splicing defect seen in human MDS but develop anemia without ring sideroblasts. *Leukemia* **31**, 720-727 (2017).
50. Shirai, C.L., *et al.* Mutant U2AF1 Expression Alters Hematopoiesis and Pre-mRNA Splicing In Vivo. *Cancer Cell* **27**, 631-643 (2015).
51. Komurov, K., Dursun, S., Erdin, S. & Ram, P.T. NetWalker: a contextual network analysis tool for functional genomics. *BMC Genomics* **13**, 282 (2012).
52. Komurov, K., White, M.A. & Ram, P.T. Use of data-biased random walks on graphs for the retrieval of context-specific networks from genomic data. *PLoS Comput Biol* **6**(2010).
53. Stoilov, P., Lin, C.H., Damoiseaux, R., Nikolic, J. & Black, D.L. A high-throughput screening strategy identifies cardiotonic steroids as alternative splicing modulators. *Proc Natl Acad Sci U S A* **105**, 11218-11223 (2008).

1 **Methods**

2 **RNA isoform switching, pathway enrichment, and mutation analysis in AML**

3 To identify genes undergoing isoform switching in the TCGA database, we defined four gene-level
4 metrics: (1) Cumulative Expression (CE) is the sum of individual isoform expression levels for a
5 gene in a given sample; (2) Cumulative Abundance (CA) measures the average gene CE across
6 samples; (3) Cumulative Variance (CV) measures the variance in the CE; (4) Isoform Variance
7 (IV) measures variance in the expression of an individual mRNA isoform; and (5) Isoform
8 Divergence (ID) measures the most negative correlation (Pearson's r) between the expressions of
9 mRNA isoforms for a given gene. A strong negative ID (e.g. < -0.5) indicates that at least two
10 isoforms of a gene have a mutually exclusive expression pattern, and therefore the gene is at least
11 partially regulated at the level of isoform switching where the expression of one mRNA isoform is
12 substituted by another. A 3-dimensional plot relating these measures to each other for all genes
13 expressed in AML is shown in Figure 1A. For pathway enrichment of genes that are regulated at
14 the level of isoform switching, hypergeometric distribution test was performed. For pathway over-
15 representation profiles of genes that correlate with the expression of IRAK4-L or IRAK4-S in TCGA
16 AML dataset was generated using NetWalker.¹ For analysis of somatic mutations enriched in AML
17 patients with distinct IRAK4 isoforms, correlations of individual exon expression with the indicated
18 mutations in AML was calculated and the represented by Z-score in a heatmap.

19 **Cell lines and human samples**

20 AML cell lines, HL60, THP1, F36P, and TF-1 were purchased from the American Type Culture
21 Collection. MDS-L was provided by Dr. Kaoru Tohyama (Kawasaki Medical School, Okayama,
22 Japan)^{2,3}. Cell-lines were cultured in RPMI 1640 medium with 10% FBS and 1% penicillin-
23 streptomycin. MDSL and TF-1 cell lines were cultured in the presence of 10 ng/mL human
24 recombinant IL-3 (Stem Cell Technologies). Doxycycline-inducible FLAG-tagged-WT U2AF1 or
25 FLAG-tagged U2AF1-S34F lentiviral expression plasmids were previously described⁴. THP1-
26 IRAK4-KO (referred to as clone 14) and the corresponding wild-type THP1 cells were generated as
27 previously described⁵. Human CD34+ umbilical cord blood (UCB) and adult BM-derived
28 mononuclear cells (MNCs) were obtained from the Translational Research Development Support
29 Laboratory of CCHMC under an approved Institutional Review Board protocol. Human CD34+
30 UCB cells and primary MDS/AML cells were maintained in StemSpan Serum-Free Expansion
31 Media (Stem Cell Technologies) supplemented with 10 ng/mL of recombinant human stem cell

factor (SCF), Flt3 ligand (Flt3L), thrombopoietin (TPO), IL-3, and IL-6 (Stem Cell Technologies). The MDS patient samples used in this study were obtained from patients diagnosed with MDS after written informed consent under approval by the IRB at Albert Einstein College of Medicine.

Creation of inducible 3XFLAG-U2AF1(S34F) and U2AF1 wild-type AML cell lines

A doxycycline-inducible lentiviral expression plasmid was previously described.⁶ For better detection of exogenous U2AF1 expression by western blotting, c-terminus 3XFLAG tagged U2AF1 wild-type or S34F was introduced into the plasmid. Lentivirus was generated in 293T cells with the packaging plasmids pMD-G, pMD-Lg and REV as described previously.⁴ Concentrated virus (multiplicity of infection of 3) was used to transduce K562 cells (ATCC, CCL243). Transduced cells (marked by green fluorescent protein) were isolated by fluorescence-activated cell sorting. Expression of U2AF1 wild-type or S34F was induced using the indicated concentrations of doxycycline hyclate (Sigma, St Louis, MO) in water.

Semi-quantitative RT-PCR and qRT-PCR

Total RNA was extracted and purified using Quick-RNA MiniPrep (Zymo research, R1055) and reverse transcription was carried out using SuperScript cDNA Synthesis Kit (Invitrogen) or High-Capacity cDNA Reverse Transcription Kit (Thermo-Fisher). Long and short isoforms of IRAK4 were amplified using primers flanking IRAK4 exon 4 (set 1 forward primer sequence: ATATGTGCGCTGCCTCAATG, reverse primer sequence: GGAGTGTATTAGCAGTTTTTGGG; set 2 forward primer sequence: GCTGCCTCAATGTTGGACTA, reverse primer sequence: TCTGGACTTGAGGAGTCAGG) in a PCR reaction using maxima hotstart PCR mastermix (Thermo-Fisher) or GeneAmp Fast PCR Master Mix (Applied Biosystems), and PCR products (312 and 166 bp, 426 and 280 bp, respectively) were separated by electrophoresis on a 2% agarose gel. Band intensity was measured with ImageJ. Statistical significance was determined by Kruskal-Wallis one-way analysis of variance followed by Dunn's *post hoc* test. Quantitative (q) RT-PCR was performed with either Taqman Master Mix (Life Technologies) for *TNFAIP3* (*Hs00234713_m1*), *IL-6* (*Hs99999032_m1*), and *GAPDH* (*Hs02758991_g1*) or PowerUP SYBR Green Master Mix (Thermo Fischer Scientific) for *Jun* (forward primer sequence: TCCAAGTGCCGAAAAAGGAAG reverse primer sequence: CGAGTTCTGAGCTTTCAAGGT), *Egr1* (forward primer sequence: CTTCAACCCTCAGGC GGACA reverse primer sequence: GGAAAAGCGGCCAGTATAGGT3), *CXCL2* (forward primer sequence: CGCCCATGGTTAAGAAAATCA, reverse primer sequence: CCTTCTGGTCAGTTGGATTGTC),

63 CXCL8 (forward primer sequence: CTTGGCAGCCTTCCTGATTT, reverse primer sequence:
64 TTCTTTAGCACTCCTTGGCAAAA). qRT-PCR was performed on an Applied Biosystems StepOne
65 Plus Real-Time PCR System.

66 **shRNA-mediated and siRNA-mediated knockdown of IRAK4**

67 The pLKO.1 constructs were obtained from the Lentiviral core at CCHMC. Puromycin resistance
68 gene was replaced by green fluorescent protein (GFP). Two independent and validated pLKO.1-
69 shIRAK4 constructs were obtained: TRCN000000063 and TRCN000000065. For siRNA
70 knockdown, a pool of several siIRAK4 were used.

71 **Plasmids and transfection**

72 FLAG-IRAK4-Long and FLAG-IRAK4-Short cDNAs were created by custom gene synthesis (IDT).
73 FLAG-IRAK4 was cloned into pcDNA3.0 with EcoRI and BamHI, or into pMSCV-pGK-GFP with
74 EcoRI and XhoI. Transfections were performed using TransIT-LT1 transfection reagent (Mirus MIR
75 2306).

76 **Immunoblotting and flow cytometry**

77 Total protein lysates were obtained from cells by lysing the samples in cold RIPA buffer (50mM
78 Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100 and 0.1% SDS), in the presence of PMSF,
79 sodium orthovanadate, protease and phosphatase inhibitors. Protein concentration was evaluated
80 by a BCA assay (32106, Pierce). The bound proteins were separated by SDS-polyacrylamide gel
81 electrophoresis, transferred to nitrocellulose membranes, and analyzed by immunoblotting.
82 Western blot analysis was performed with the antibodies described in Supplemental Table 4. For
83 apoptosis assay, cells were treated with 10 μ M CA-4948 for 48 hours, then washed and
84 resuspended in Annexin V-binding buffer (10 mM HEPES; 140 mM NaCl; 2.5 mM CaCl₂; pH
85 7.4). Annexin V-conjugated antibody (1:100) was added to each sample. After a 15-minute
86 incubation, PI (1:100) was added and cells were analyzed by flow cytometry. For analysis of
87 hematopoietic cell surface markers, primary cells were collected, washed once with 1X PBS (GE
88 Healthcare Life Sciences), and subsequently stained with conjugated antibodies. The antibodies in
89 Supplemental Table 4 were used for flow cytometric analysis of human MDS and AML cells. All the
90 flow data was analyzed on BD LSRII or FACSCanto using BD FACSDiva Software and FlowJo
91 10.4. For shIRAK4-expressing THP1 xenografts, transduced cells were sorted for GFP expression
92 on the SH800S (Sony Biotechnology).

93 **Generation of THP1 CRISPR-Cas9 IRAK4 knockout cells**

THP1 cells were maintained in RPMI-1640 media supplemented with 10% FBS. Media was renewed every 2-3 days. Cells were electroporated using the Invitrogen Neon Transfection system with the following conditions: 1350 Volts, 20ms width, 1 pulse, using 100µl tip and 2.0×10^6 cells/ml in R buffer. Cells were transfected with 1µg pSpCas9(BB)-2A-GFP, a plasmid expressing GFP, Cas9, and a sgRNA targeted to exon 6 of IRAK4. Cells were allowed to recover for 72 hours in antibiotic free RPMI-1640 media supplemented with 10% FBS. GFP expressing cells were FACS sorted into 96 well plates containing THP1 cell conditioned media. Clones were allowed to expand and screened for gene editing by PCR of the targeted region. Clones of interest were sent for Sanger sequencing to confirm a functional knockout of IRAK4 exon 6. THP1-IRAK4 KO cells analyzed are clone 3.

Rescue experiments in THP1 cells

THP-1 cells were transduced with mCherry-tagged shRNAs (pLKO.1-mCherry) targeting both IRAK4-L and shIRAK4-S (shIRAK4-L/S), or empty control, sorted for mCherry expression, and then transduced with vector control (pLEGO-GFP), IRAK4-L-GFP or IRAK4-S-GFP constructs and sorted for mCherry/GFP expression. Concurrently, CRISPR-Cas9 generated THP1 IRAK4KO cells were transduced with vector-control (pMSCV-pGK-GFP), IRAK4-L-GFP or IRAK4-S-GFP constructs and sorted for GFP expression. Transduced cells were then assessed for clonogenic progenitor function in methylcellulose and for proliferation in liquid culture by Trypan blue exclusion.

Luciferase assay

HEK293 cells were transfected for 24 - 48 h with κB-luciferase or AP1-luciferase and TK Renilla plasmids along with IRAK4-Long or IRAK4-Short. Lysates were analyzed for NF-κB reporter or AP1 reporter activity using the dual luciferase reporter assay system (Promega E1910).

NF-κB reporter assay

THP1-Blue NF-κB cells (InvivoGen) carrying a stable integrated NF-κB-inducible secreted embryonic alkaline phosphatase (SEAP) reporter construct were transduced with shIRAK4-GFP and sorted for GFP expression after 48 hours. Cells were then plated at a concentration of 2×10^4 cells/well and stimulated with increasing doses of Pam3CSK4 for 24 hours. Cells were centrifuged and 20uL supernatant was incubated with 180uL QUANTI-Blue reagent at 37 degrees for 30 min - 2 hours. The levels of NF-κB-induced SEAP was measured in a microplate reader at 620 nm.

Protein interaction network

Protein isoform-isoform interaction network were generated based on annotated protein-protein interaction and protein-domain interaction databases. Briefly, all proteins and their amino acid sequences were obtained from UCSC genome browser (<https://genome.ucsc.edu/>). Well-conserved protein domains of all isoforms were predicted using HMMER3 (<http://hmmer.org>, E(xpect)-value < 0.001) and RPS-BLAST (<ftp://ftp.ncbi.nih.gov/toolbox>, % identity > 40% & E-value < 0.001) against the PFAM (<https://www.ebi.ac.uk/interpro/>), CCD (<https://www.ncbi.nlm.nih.gov/Structure/cdd>), and InterPro (<https://www.ebi.ac.uk/interpro/>) databases. Experimentally validated protein-protein interaction information was retrieved from HPRD (<http://www.hprd.org>). DOMINE (<https://manticore.niehs.nih.gov/cgi-bin/Domine>) was used as a source of domain-domain interactions. Red lines represent IRAK4-L interactions based on death-domain sequences of IRAK4 and MyD88. Blue lines represent IRAK4-S interactions based on kinase-domain sequences of IRAK4 and death-domain sequences of IRAK1. Green lines represent IRAK4-L interactions based on kinase-domain and death-domain sequences of IRAK4 and IRAK1.

Immunoprecipitation

HEK293 cells were transfected with 5ug MyD88 and either 5ug vector control, FLAG-IRAK4-L, or FLAG-IRAK4-S for 2 days. Cells were lysed in RIPA. Flag antibody was added to cell lysates (1mg) overnight at 4°C and captured by the addition of Anti-Flag M2 Affinity Gel (A2220; Sigma-Aldrich). The immune complexes were washed 3 times with lysis buffer followed by elution with Laemmli Buffer. The bound proteins were separated by SDS-polyacrylamide gel electrophoresis, transferred to nitrocellulose membranes, and analyzed by IB for HA-MyD88. For endogenous immunoprecipitation of MyD88 and IRAK4, protein lysates from TF1 and HL60 cells were prepared by lysing cell pellets with CHAPS buffer (20 mmol/l Tris-HCl, pH 7.5; 1 mmol/l EDTA; 150 mmol/l NaCl; 2% CHAPS; Sigma) containing protease (Sigma) and phosphatase inhibitors (Sigma) for 1 h on ice. Protein concentration was estimated using Bio-Rad protein Assay (Bio-Rad). Equal protein lysates were incubated with primary antibody overnight at 4 °C, then an equal amount of protein A agarose beads (Rockland Immunochemicals Inc.) was added to all samples, followed by 1 h of incubation at 4 °C. The beads were washed three times with CHAPS, eluted with loading buffer supplemented with 2-mercaptoethanol (Sigma-Aldrich) and immunoblotting was performed.

Clonogenic progenitor assays

Clonogenic progenitor frequencies were determined by plating 1×10^3 cells/mL for cell lines and 3×10^5 cells/mL for normal bone marrow in methylcellulose (Methocult H4236; Stem Cell Technologies). Colonies were scored after 7 days. For clonogenic assays with primary MDS/AML cells, primary patient samples and healthy controls were plated in methylcellulose (Stem Cell Technologies H4435, Vancouver, CA) with the CA-4948 and control, and colonies were counted after 14-17 days.

Xenografts

NOD/SCID IL2Rgamma KO (NSG) mice were injected with 1×10^6 THP-1 cells and treated with 12.5 mg/kg/d of CA-4948 or control for 5 weeks. Mice were sacrificed and assessed for tumor burden measurements. For primary patient-derived xenografts NSG mice were irradiated (200 rads) 24 hours prior to injection. Mononuclear cells from primary MDS patients were isolated by Ficoll separation. $2-5 \times 10^6$ MNCs were administered via tail vein injection. Three weeks later, BM aspiration were performed and analyzed by flow cytometry for the human cell engraftment. Mice were considered to be engrafted if they showed 0.1% or higher human derived CD45+ cells. The engrafted mice were randomized for treatment with CA-4948 or control for 3 weeks. Following treatment, BM aspirations and flow cytometry analysis for the above-mentioned markers were performed. After primary engraftment, the NSG mice were sacrificed. Cells were collected by crushing bones and spleen using mortar and pestle and straining with 70 um nylon mesh filter. Human cells were isolated by using Dyna beads Untouched Mouse T cells (Thermofisher 11413D) and injected in mice previously irradiated with 200 rads. After 8 weeks, the mice were sacrificed and analyzed for engraftment.

For shIRAK4 xenografts, THP-1 cells were transduced with pLKO-scrambled-GFP, pLKO-shIRAK4-L-GFP or pLKO-shIRAK4-L/S-GFP and sorted for GFP expression 48 hours later. 1×10^6 GFP+ cells from each group were transplanted into NSG mice (n = 10). BM aspirations and bleeds were performed 5 weeks after transplantation and sacrificed when sick and assessed for GFP proportions in BM and peripheral blood. All mice were bred, housed and handled in the Association for Assessment and Accreditation of Laboratory Animal Care-accredited animal facility of Cincinnati Children's Hospital Medical Center and Albert Einstein College of Medicine.

Hematologic toxicity assessment of CA-4948 in vivo

185 NSG mice received 12.5 mg/kg of CA-4948 (n = 9 mice) or vehicle (PBS; n = 7 mice) for 5 days
186 per week for 12 weeks. Complete blood counts were determined at the end of the study.

187 **RNA-sequencing of MDS CD34+ cells**

188 Total RNA from MDS and control BM CD34+ cells was extracted using TRIzol with a linear
189 acrylamide carrier, treated with DNase I (Life technologies) and purified using Agencourt
190 RNAClean XP beads (Beckman Coulter). RNA quality was determined using an RNA 6000
191 Bioanalyzer pico kit (Agilent). A cDNA library was produced using a SMARTer library preparation
192 protocol (Clontech). Sequencing (100 bp paired-end reads) was performed on an Illumina
193 HiSeq4000. The reads were mapped to human genome GRCh37 using HISAT2 version 2.0.0-
194 beta. Uniquely mapped read pairs were counted using htseq-count. The data have been deposited
195 in the NCBI's Gene Expression Omnibus (GEO) repository (accession number: GSE114922).

196 **IRAK4 exon 4 reporter cells**

197 IRAK4 exon 4 (WT: A at -3 position; Mutant: T at -3 position) and its flanking intronic sequences
198 (100 nucleotides upstream and 100 nucleotides downstream from exon 4) were inserted into the
199 EcoRI and BamHI sites of pFlareA vector. IRAK4 exon 4 is based on NCBI Refseq annotation, or
200 exon 6 according to AltAnalyze Ensembl 72 annotation. pFlareA-exon4 vectors were linearized by
201 DraIII and transfected into 293T cells. Cell colonies stably expressing the reporter were selected by
202 1 mg/mL G418 and stable expression of GFP and RFP.

203 **Reagents**

204 The IRAK1/4 inhibitor (IRAK1/4-Inh; Amgen Inc.) was purchased from Sigma-Aldrich (I5409). CA-
205 4948 was received from Curis Inc. Recombinant human IL-1 β was purchased from PeproTech
206 (200-01B).

207 **Growth Curves**

208 Trypan blue exclusion was measured using an automated cell counter (BioRad TC10). Cells were
209 replenished with fresh media, doxycycline, and drug every fifth day.

210 **Statistics and reproducibility.** Multiple independent experiments were performed to verify the
211 reproducibility of all experimental findings. All quantitative data represent the mean $\pm$ s.e.m., unless
212 otherwise stated. Unless otherwise indicated, two-tailed Student's t-tests were used to assess
213 statistical significance for comparisons between two groups using Prism 8 (GraphPad software).
214 The log rank test was used to determine P values for survival curves. All experiments were carried
215 out at least twice, and the findings of all key experiments were reliably reproduced. All replicates

216 and precise P values are provided in the figure and legends state the sample size and number of
217 independent experiments. To knockout the gene of interest, at least three independent sgRNAs
218 were employed to generate cell lines, and similar results were found in all cell lines.

219 **Data Availability Statement**

220 RNA sequencing data that support the findings of this study have been deposited under accession
221 codes GSE114922 (corresponding to data in Figure 4B), GSE58831 (corresponding to data in
222 Supplemental Figure 4), dbGaP: phs000159.v8 (SRX1608652, SRX1608653, SRX1608605,
223 SRX1608609, SRX1608616,SRX1608620)(corresponding to data in Figure 5A), and E-MTAB-
224 1733 (corresponding to data in Figure 1G). The human AML data were derived from the TCGA
225 Research Network: <http://cancergenome.nih.gov/>. Source data for Figures 2-6 and Supplemental
226 Figures 3,4, and 6 have been provided as Supplementary Table 5. All other data supporting the
227 findings of this study are available from the corresponding author on reasonable request.

228

229 **References**

230

- 231 1. Komurov, K., Dursun, S., Erdin, S. & Ram, P.T. NetWalker: a contextual network analysis
232 tool for functional genomics. *BMC Genomics* **13**, 282 (2012).
- 233 2. Matsuoka, A., *et al.* Lenalidomide induces cell death in an MDS-derived cell line with
234 deletion of chromosome 5q by inhibition of cytokinesis. *Leukemia* **24**, 748-755 (2010).
- 235 3. Rhyasen, G.W., *et al.* An MDS xenograft model utilizing a patient-derived cell line. *Leukemia*
236 **28**, 1142-1145 (2014).
- 237 4. Shirai, C.L., *et al.* Mutant U2AF1-expressing cells are sensitive to pharmacological
238 modulation of the spliceosome. *Nat Commun* **8**, 14060 (2017).
- 239 5. Sun, J., *et al.* Comprehensive RNAi-based screening of human and mouse TLR pathways
240 identifies species-specific preferences in signaling protein use. *Sci Signal* **9**, ra3 (2016).
- 241 6. Yoshida, K., *et al.* Frequent pathway mutations of splicing machinery in myelodysplasia.
242 *Nature* **478**, 64-69 (2011).

243

Figure 1

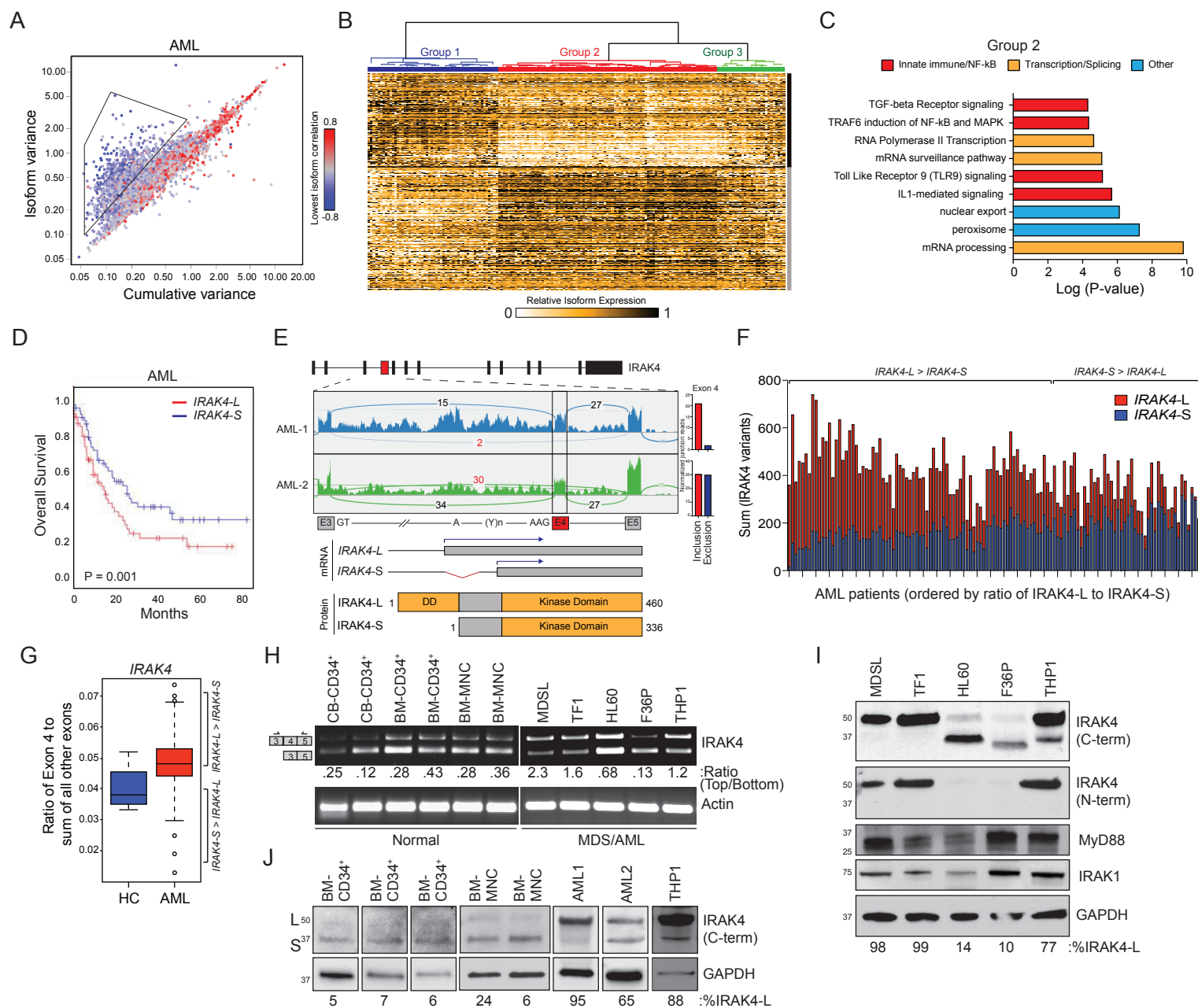


Figure 2

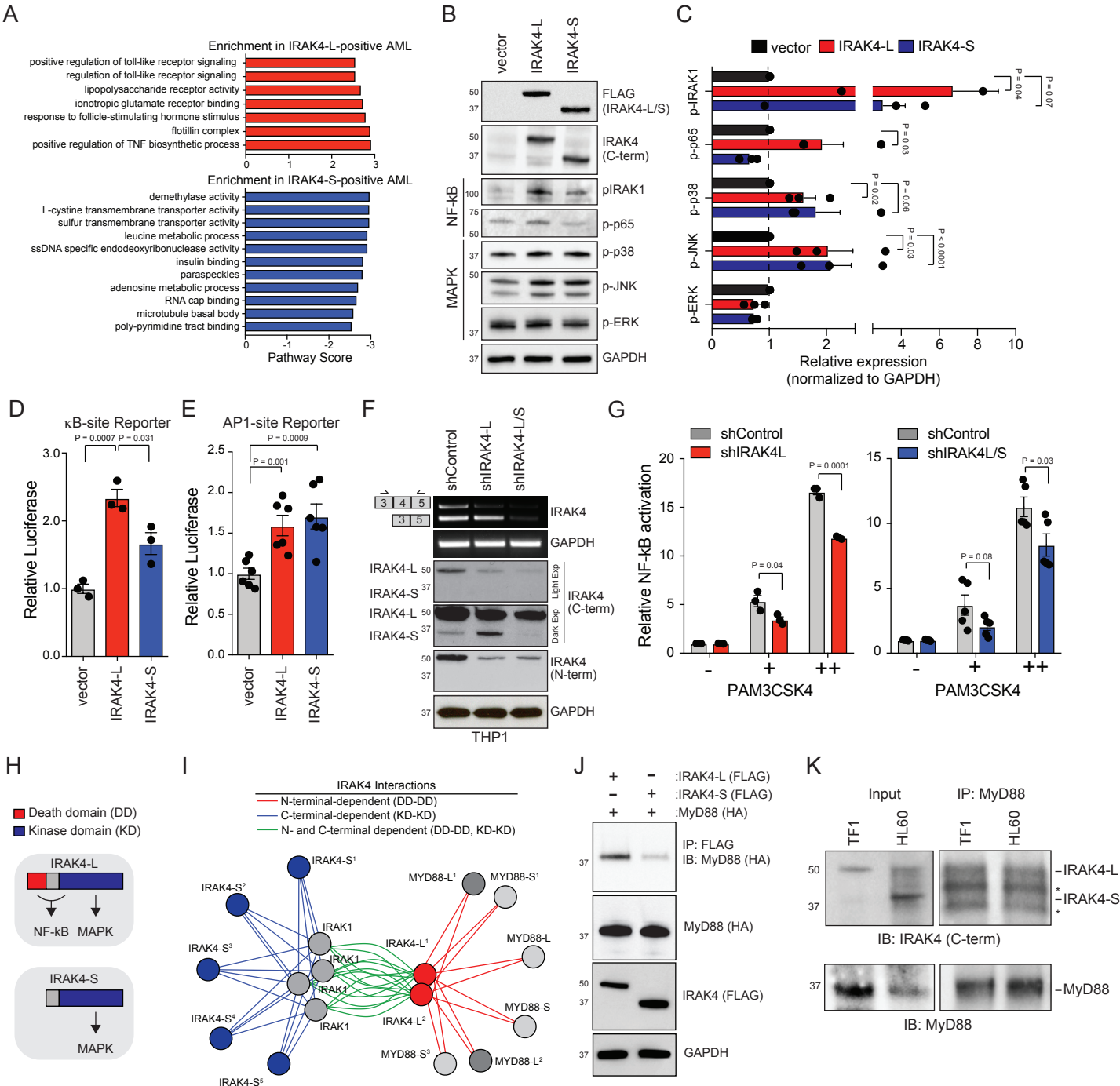


Figure 3

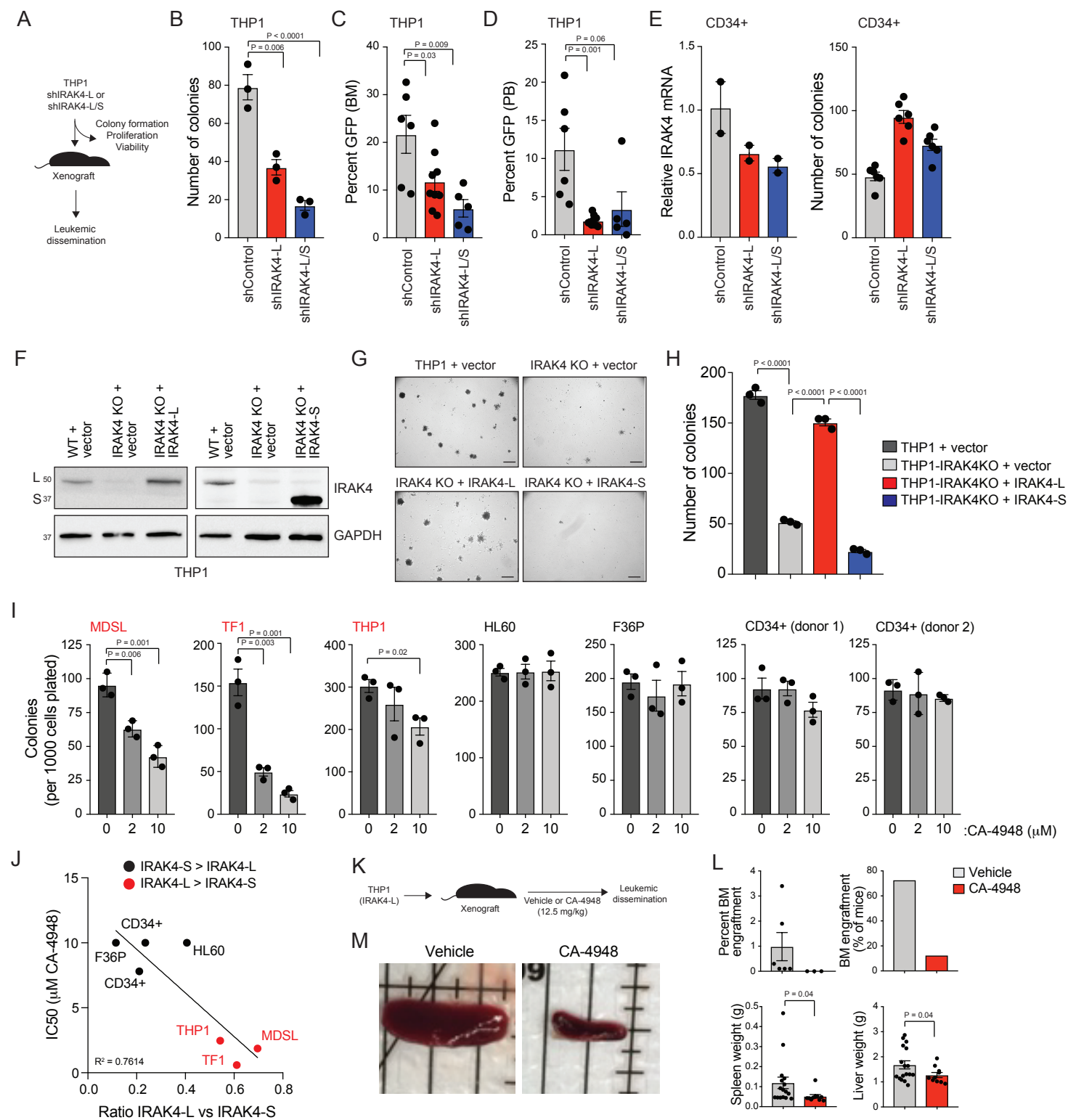


Figure 4

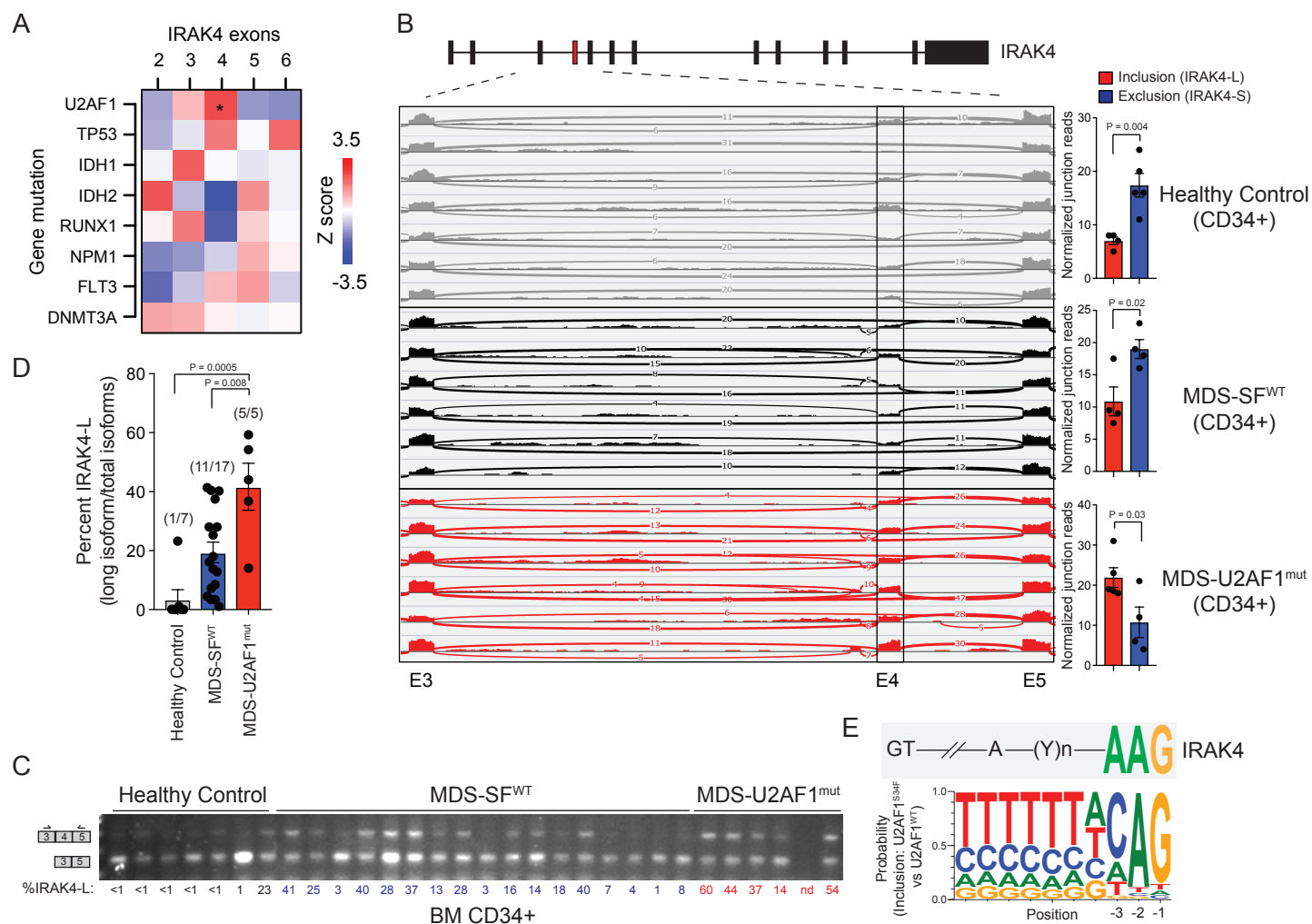


Figure 5

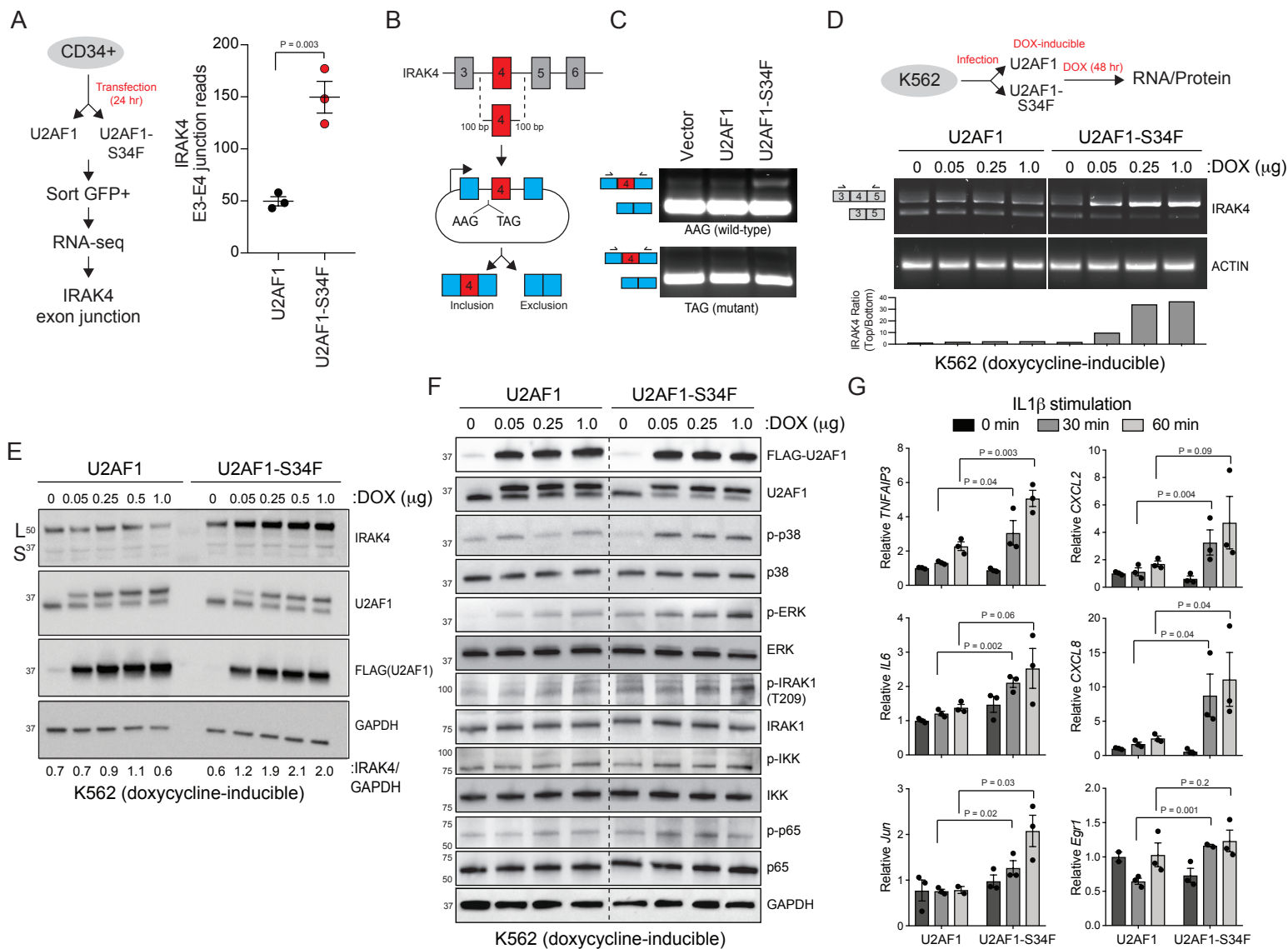


Figure 6

