

1 ***TG-Interacting Factor 1* expression quantitatively impacts survival in acute myeloid**  
2 **leukemia**

3  
4 Ling Yan<sup>1,◇</sup>, Julie A. Means-Powell<sup>2,◇</sup>, Danko Martincic<sup>2</sup>, Vladimir D. Kravtsov<sup>2</sup>, Yu Shyr<sup>3,4,7</sup>,  
5 John P. Greer<sup>2</sup>, Utpal P. Davé<sup>2,7,8</sup>, Mark J. Koury<sup>2,7,8</sup>, David Wotton<sup>9</sup>, Rizwan Hamid<sup>1,6,7,\*</sup>, and  
6 Stephen J. Brandt<sup>2,5,6,7,8\*</sup>

7  
8 Departments of Pediatrics<sup>1</sup>, Medicine<sup>2</sup>, Biostatistics<sup>3</sup>, Biomedical Informatics<sup>4</sup>, Cell and  
9 Developmental Biology<sup>5</sup>, Cancer Biology<sup>6</sup>, Vanderbilt-Ingram Cancer Center<sup>7</sup> of Vanderbilt  
10 University Medical Center, Nashville, TN, 37232, and VA Tennessee Valley Healthcare  
11 System<sup>8</sup>, Nashville, TN, 37212; Department of Biochemistry and Molecular Genetics, University  
12 of Virginia School of Medicine, Charlottesville, VA<sup>9</sup>.

13  
14 Mailing address: Rizwan Hamid, Division Medical Genetics and Genomic Medicine,  
15 Department of Pediatrics, Room DD2205 MCN, Vanderbilt University Medical Center,  
16 Nashville, TN 37232. Phone: 615-936-1809. Fax: 615-936-3853. E-mail:

17 [rizwan.hamid@vumc.org](mailto:rizwan.hamid@vumc.org)

18

19 <sup>◇</sup> and <sup>\*</sup> denote equal contributions to this paper

20

21 **Abstract**

22 Applying transcriptional profiling analysis to myeloblasts from 59 adult patients with acute  
23 myeloid leukemia (AML) treated at our institution, we found that expression of the three-amino  
24 acid loop extension (TALE) homeobox gene *TG-Interacting Factor 1 (TGIF1)* correlated with  
25 overall and relapse-free survival, which was then confirmed in two other cohorts of patients.  
26 Moreover, *TGIF1* expression correlated with survival for all cytogenetic risk groups and was an  
27 independent prognostic factor in multivariate analysis. To elucidate the mechanism, we used  
28 *Tgif1* knockout mice in which acute or chronic myeloid leukemia was induced through retroviral  
29 transfer of the *MLL-AF9* or *BCR-ABL* fusion genes into bone marrow cells. Loss of *Tgif1*  
30 accelerated disease progression, shortened survival, attenuated the response to chemotherapy,  
31 and doubled the frequency of leukemia-initiating cells. RNA-based sequencing analysis showed  
32 that genes associated with transforming growth factor- $\beta$  (TGF- $\beta$ ) and retinoic acid signaling  
33 pathways were differentially affected in *Tgif1*<sup>-/-</sup> compared to *Tgif1*<sup>+/+</sup> leukemia cells.

## 34 **Introduction**

35 Acute myeloid leukemia (AML) is an aggressive and heterogeneous disease that most commonly  
36 occurs in the elderly. Although a majority of patients achieve a complete remission (CR) with  
37 induction and consolidation chemotherapy, most will eventually relapse, with overall survival  
38 (OS) ranging from 10-35% <sup>1,2</sup>. As a result, clinical indicators have been sought that could inform  
39 post-remission therapy and improve long-term survival <sup>3,4</sup>. Established predictors of outcome  
40 include patient age, white blood cell (WBC) count, antecedent myelodysplastic syndrome, and,  
41 of greatest significance, the presence of specific chromosomal rearrangements in leukemia cells  
42 <sup>5</sup>. Patients with t(15;17), t(8;21), or inv(16) enjoy the longest overall survival (OS), while t(6;9),  
43 11q23, monosomy of chromosomes 5 or 7, and complex karyotypes confer the worst prognoses,  
44 with 80% of these individuals ultimately relapsing <sup>6</sup>. The largest group of patients, however, has  
45 either a normal karyotype or specific cytogenetic abnormalities that put them at some  
46 intermediate risk of relapse and death <sup>7,8</sup>.

47 Patients in the favorable risk group can be successfully treated with intensive chemotherapy  
48 only <sup>8</sup>, while those in the high-risk group require an allogeneic hematopoietic stem cell transplant  
49 (SCT) in remission for long term survival <sup>9</sup>. In contrast, the optimum therapy for intermediate-  
50 risk patients is unclear. Given the substantial number of individuals in whom a cytogenetic  
51 marker is not present or prognostic, additional, especially molecular, predictors of survival have  
52 been sought. One of the most extensively characterized molecular abnormalities is an internal  
53 tandem duplication (ITD) of the gene for Fms-Related Tyrosine Kinase 3 (FLT3). While found  
54 to impact survival in some analyses <sup>10,11</sup>, no difference in that of *FLT3* ITD-positive and ITD-  
55 negative patients was detected in other studies <sup>12,13</sup> as assessed independently of other mutations,  
56 in particular those in the *nucleophosmin* (*NPM1*) gene. More than one ITD <sup>10</sup>, a high ratio of

57 mutant to wild-type DNA <sup>12, 13</sup>, and loss of the wild-type allele have been more clearly linked to a  
58 poor prognosis.

59 Rearrangement and/or altered expression of other genes altered or not have been found to  
60 have prognostic value. Increased expression of the oncogene *Ecotopic Viral Integration Site 1*  
61 (*EVII*), for example, was associated with inferior survival in adults with *de novo* AML <sup>14</sup>, while  
62 reduced expression of the *Deleted in Colorectal Carcinoma (DCC)* gene was an independent  
63 prognostic factor for CR and OS <sup>15</sup>. Further, overexpression of the *Brain And Acute Leukemia,*  
64 *Cytoplasmic (BAALC)* gene had an adverse impact on overall survival (OS) and disease free  
65 survival (DFS) in a study of younger adults with cytogenetically normal *de novo* AML <sup>16</sup>, and  
66 mutation of the gene encoding the myeloid transcription factor C/EBP $\alpha$  (*CEBPA*) was associated  
67 with a better outcome in patients in the intermediate-risk group <sup>17, 18</sup>. Increasingly, expression  
68 profiling and deep sequencing approaches have been used to distinguish groups of patients  
69 differing in biology as well as survival.

70 Analyzing global gene expression profiles from a cohort of patients treated at our institution,  
71 we discovered that the abundance of mRNA for the three-amino acid loop extension (TALE)  
72 superclass homeobox gene *TG-Interacting Factor1 (TGIF1)* <sup>19-22</sup> correlated very significantly  
73 with measures of survival. Analysis of its actions in *Tgif1* knockout mice in which acute or  
74 chronic myeloid leukemia was induced revealed that *Tgif1* abundance affected the number and/or  
75 function of leukemia-initiating cells (LICs), also referred to as leukemia stem cells (LSCs). This  
76 observation has important prognostic and therapeutic implications.

77

78 **Results**

79 ***TGIF1 expression is associated with clinical outcomes in AML.*** To identify new predictors of  
80 outcome, we carried out microarray analysis of leukemia cellular RNA from 59 of the 60 adult  
81 patients with AML treated at our institution (Supplemental Table S1). Modified Cox regression  
82 analysis was then used to identify genes whose expression correlated with OS. A group of 38  
83 genes were predictive with a  $P < 0.001$  and of 5 genes with a  $P < 0.0001$ , while the expression of  
84 a single gene, *TGIF1*, correlated with OS with the greatest significance ( $P < 0.00001$ ).

85 Kaplan-Meier analysis of patient survival in this cohort showed that only 8% (3/39) of  
86 individuals whose leukemia cells contained lower levels of *TGIF1* mRNA than the reference  
87 remained alive at last follow-up (median survival = 181 days, confidence interval (CI) = 112-249  
88 days), while 60% (12/20) of patients who expressed *TGIF1* at higher levels survived. Since more  
89 than half of patients with higher *TGIF1* expression were living at last analysis, median survival  
90 could not be determined. Log-rank analysis revealed that individuals with lower *TGIF1*  
91 expression had significantly shorter OS (Figure 1A) and relapse-free survival (RFS) (Figure 1B)  
92 than the group with higher expression. Although a commercial RNA preparation was used to  
93 stratify patients into lower and higher expressers, the best cut-point in *TGIF1* expression for  
94 discriminating survival determined by receiver operating characteristic analysis was close to that  
95 in the commercial reference (data not shown). When patients were segregated by *TGIF1*  
96 expression in quartiles, a concentration-dependent relationship between *TGIF1* expression and  
97 outcome was suggested (Figure 2A). Finally, *TGIF1* expression correlated with OS with even  
98 greater statistical significance when analyzed as a continuous variable ( $P < 0.0001$ ).

99

100 ***TGIF1 expression correlated with survival for risk groups defined by cytogenetics analysis.***

101 Seventeen percent (1/6) and 86% (18/21) of patients with good- and poor-risk cytogenetic

102 findings, respectively, had lower *TGIF1* expression. In contrast, *TGIF1* expression was more  
103 evenly distributed in patients with a normal karyotype. Specifically, 41% (11/27) had higher and  
104 59% (16/27) lower *TGIF1* mRNA than reference, and individuals in the intermediate-risk group  
105 with lower *TGIF1* survived for significantly shorter times (median survival = 201 days, CI =  
106 138-263 days) than those with higher expression (median survival = 1566 days). *TGIF1*  
107 expression also correlated with shorter OS for patients in poor-risk patients (data not shown) and  
108 even for the good-risk group, although the number of patients was small. Thus, *TGIF1* levels  
109 correlated with clinical outcome in adult AML patients from all three cytogenetic risk groups.

110

111 ***TGIF1 expression correlated with survival in independent cohorts of AML patients.*** To  
112 substantiate the association between *TGIF1* expression and survival observed in the discovery  
113 group, we reanalyzed hybridization data from a published study<sup>23</sup> and related leukemia cell  
114 *TGIF1* expression to survival in this larger (n = 285) group of patients. When patients were  
115 segregated into 5 groups according to myeloblast *TGIF1* RNA abundance and trend and log-rank  
116 analysis of survival in these patients was carried out, a concentration- or dose-dependence in  
117 survival was observed (Figure 2B). We also investigated whether *TGIF1* expression differed in  
118 five patient clusters, 5, 9, 10, 12, and 13<sup>23</sup>. Cluster 5 consisted of FAB class M4 or M5  
119 leukemias, cluster 10 expressed the *EVII* oncogene, and clusters 9, 12 and 13 were defined by  
120 the presence of (inv)16, t(15;17), and t(8;21), respectively. Consistent with the association of  
121 t(15;17), t(8;21), and inv(16) with a better outcome and the inverse association between *EVII*  
122 expression and survival<sup>14</sup>, OS was shorter in patients in clusters 5 and 10 compared to 9, 12, and  
123 13<sup>23</sup> (Figure 2B). Mean *TGIF1* levels differed significantly between these clusters (P < 0.0001),

124 with significant inter-group differences noted between clusters 5 and 9 ( $P = 0.0001$ ), 5 and 12 ( $P$   
125  $< 0.0001$ ), 5 and 13 ( $P < 0.0001$ ), 10 and 12 ( $P < 0.0001$ ), and 10 and 13 ( $P = 0.0003$ ).

126 Finally, OS was significantly worse in AML patients from The Cancer Genome Atlas  
127 (TCGA) whose *TGIF1* expression was below the median than those with expression above the  
128 median for both probe set 1566901\_at ( $P < 0.0001$ ) and 203313\_s\_at ( $P = 0.000254$ )  
129 ([http://servers.binf.ku.dk/bloodspot/?gene=TGIF1&dataset=normal\\_human\\_v2\\_with\\_AMLs](http://servers.binf.ku.dk/bloodspot/?gene=TGIF1&dataset=normal_human_v2_with_AMLs))  
130 (Supplemental Figure S1). This association was specific for *TGIF1*, as there was no difference in  
131 survival according to expression of the related *TGIF2* gene. Thus, analysis of multiple,  
132 independent groups of patients confirmed the correlation between *TGIF1* gene expression and  
133 outcome and, for the European cohort, the quantitative nature of this relationship.

134  
135 ***TGIF1* expression was unrelated to the presence of *FLT3* gene mutations.** To determine  
136 whether decreased *TGIF1* expression was related to *FLT3* mutation, a widely used prognostic  
137 factor in adults with AML, RNA from patient cells and the MV4-11 cell line, used as a positive  
138 control for the *FLT3* ITD <sup>24</sup>, were characterized by RT-PCR analysis. Of the 55 patients from  
139 whom material was available for analysis, a *FLT3* ITD was detected in 15 (27%) and the D835Y  
140 missense mutation in two (4%). In two patients with a *FLT3* ITD, no wild-type *FLT3* RNA could  
141 be detected at the sensitivity of this method, and in accord with published data <sup>25</sup>, these  
142 individuals had very short survivals. In contrast, the OS of patients expressing mutant and wild-  
143 type *FLT3* mRNA did not differ significantly from patients expressing only wild-type message in  
144 our analysis (not shown). Importantly, no association between level of *TGIF1* expression and  
145 *FLT3* mutation could be discerned (data not shown). These data thus argue against a model in  
146 which mutational activation of *FLT3* signaling represses *TGIF1* expression.

147

148 ***Reduced TGIF1 expression in myeloid leukemia cells was not associated with sequence***

149 ***variants.*** The association between decreased *TGIF1* expression and shorter survival could be

150 compatible with a tumor suppressor function for this transcription factor. Since well-

151 characterized loss-of-function mutations in this gene have been described in a human genetic

152 disorder, we undertook analysis of *TGIF1* coding and 5' untranslated sequence for these or other

153 mutations in 49 patients in the discovery group. Although non-synonymous as well as

154 synonymous sequence variants were detected in 18 patients (37%), with previously unreported

155 single nucleotide polymorphisms found in five (Supplemental Table S2), none of these

156 corresponded to the mutations observed in patients with holoprosencephaly. Likewise, Southern

157 blot analysis did not reveal *TGIF1* gene rearrangements in any patient studied (data not shown).

158 Thus, no definite abnormality in either gene sequence or structure was detected in any individual

159 with AML.

160 When *TGIF1* mRNA abundance in myeloblasts was examined in aggregate, expression

161 appeared to vary in continuous fashion (Supplemental Figure S2). *TGIF1* expression in 100

162 lymphoblastoid cell lines derived from normal individuals likewise appeared to be a continuous

163 variable, with an approximately 200-fold difference between the lowest and highest expresser

164 (Supplemental Figure S3).

165

166 ***Tgif1 gene loss decreased survival in an experimental model of acute myeloid leukemia.*** Given

167 our finding that *TGIF1* expression correlated significantly with survival in AML, we sought to

168 determine if myeloblast *TGIF1* abundance affected any aspect of leukemia cell biology. Our

169 previous study has shown that *Tgif1* inactivity promotes the quiescence and the self-renewal

170 activity of hematopoietic stem cells (HSC) without disturbing steady-state hematopoiesis using a  
171 *Tgif1* knockout model<sup>26</sup>. To understand the role of *Tgif1* in leukemia cells, in the present study,  
172 AML was induced by introduction of an *MLL-AF9* fusion cDNA. Specifically, lineage-negative  
173 bone marrow cells (BMCs) from *Tgif1*<sup>+/+</sup> and *Tgif1*<sup>-/-</sup> mice were transduced with the MLL-AF9-  
174 GFP retrovirus and transduced cells transplanted into sub-lethally irradiated recipients *via* tail  
175 vein injection. Development and progression of leukemia in recipient mice was then monitored  
176 by flow cytometry analysis of green fluorescent protein (GFP)-expressing nucleated cells in  
177 peripheral blood. Six weeks post-transplant, mice transplanted with *Tgif1*<sup>-/-</sup> BMCs showed higher  
178 numbers of GFP<sup>+</sup> myeloid (CD11b<sup>+</sup> and Gr-1<sup>+</sup>) cells (Figure 3A and 3B) but lower numbers of B  
179 cells (Figure 3C) compared to mice transplanted with *Tgif1*<sup>+/+</sup> BMCs. Necropsy showed that all  
180 mice with circulating donor-derived leukemia cells had splenomegaly, with massive infiltration  
181 of myeloblasts observed in both spleen and bone marrow (data not shown).

182 Importantly, Kaplan-Meier analysis showed that mice transplanted with MLL-AF9-  
183 transduced *Tgif1*<sup>-/-</sup> BMCs survived a significantly shorter time than mice transplanted with  
184 transduced *Tgif1*<sup>+/+</sup> BMCs (Figure 4). As mice transplanted with either *Tgif1*<sup>+/+</sup> or *Tgif1*<sup>-/-</sup> BMCs  
185 succumbed to leukemia with very short latency, the aggressive nature of this leukemia almost  
186 certainly underestimated the potential effect of *Tgif1* gene loss on survival and LSC number (see  
187 below). Regardless, these results are consistent with the patient data (Figures 1 and 2) and  
188 indicate that *Tgif1* abundance impacts progression of MLL-AF9-induced AML.

189  
190 ***Tgif1* gene loss increased leukemia initiating cell frequency in MLL-AF9-induced AML.** LICs  
191 are thought to be crucial for the maintenance of AML *in vivo*<sup>27-29</sup>. Therefore, one explanation for  
192 the shorter survival seen with *Tgif1* loss could be higher numbers of leukemia initiating cells. To

193 investigate, we carried out limiting dilution analysis of LIC number, transplanting serial numbers  
194 of MLL-AF9-transduced *Tgifl*<sup>+/+</sup> or *Tgifl*<sup>-/-</sup> spleen cells from primary recipients into non-  
195 conditioned C57BL/6J mice. Indeed, there was a two-fold higher frequency of LICs in *Tgifl*<sup>-/-</sup>  
196 mice with AML (1 in 125) (95% CI: lower; 1 in 279 and higher; 1 in 56.5) than in *Tgifl*<sup>+/+</sup> mice  
197 (1 in 250) (95% CI: lower; 1 in 603 and higher; 1 in 103.6). These data suggest that *Tgifl*<sup>-/-</sup>  
198 leukemic populations were enriched in LIC function and/or number. Consistent with this notion,  
199 mice transplanted with *Tgifl*<sup>-/-</sup> leukemia cells over a range of doses had inferior survival  
200 compared to mice transplanted with *Tgifl*<sup>+/+</sup> cells (Figure 5 A-D).

201  
202 ***Tgifl* loss resulted in earlier relapse and more aggressive disease.** Since leukemia cell  
203 preparations from *Tgifl*<sup>-/-</sup> mice contained higher frequencies of LICs, we predicted that disease  
204 induced by these cells would relapse earlier and progress more rapidly after treatment than  
205 disease in mice receiving *Tgifl*<sup>+/+</sup> leukemia cells. To simulate the clinical setting, we  
206 transplanted *Tgifl*<sup>+/+</sup> and *Tgifl*<sup>-/-</sup> MLL-AF9-transformed-spleen cells from mice with established  
207 leukemia into sub-lethally irradiated recipients and treated these mice with cytarabine and  
208 doxorubicin chemotherapy as described<sup>30</sup>. For mice transplanted with *Tgifl*<sup>-/-</sup> leukemia cells,  
209 GFP-expressing cells appeared earlier and in higher numbers in peripheral blood than in mice  
210 transplanted with *Tgifl*<sup>+/+</sup> leukemia cells (Figure 6A). While treatment extended survival of both  
211 *Tgifl*<sup>+/+</sup> and *Tgifl*<sup>-/-</sup> mice with AML (compare Figure 6B to Figure 5D), mice transplanted with  
212 *Tgifl*<sup>-/-</sup> leukemia cells still had shorter survival than mice transplanted with *Tgifl*<sup>+/+</sup> leukemia  
213 cells (Figure 6B). Although selection at the level of a stem cell is not proven by these data, the  
214 differences in latency of disease and response to chemotherapy, we posit, are due to the higher

215 frequency and/or greater fitness of LICs and more rapid expansion of leukemia in *Tgifl*<sup>-/-</sup> mice  
216 (see Figure 6A).  
217  
218 ***Tgifl* loss decreased survival in a model of chronic myeloid leukemia.** Given data suggesting  
219 *Tgifl* expression impacted survival in AML, we went on to investigate whether *Tgifl* loss also  
220 affected leukemia cell biology in a mouse model of CML, a disease for which there is arguably  
221 the greatest evidence for a malignant stem cell. To that end, lineage marker-negative (Lin<sup>-</sup>)  
222 BMCs from *Tgifl*<sup>+/+</sup> and *Tgifl*<sup>-/-</sup> mice were transduced with the BCR-ABL-GFP retrovirus,  
223 which were transplanted into sub-lethally or lethally irradiated recipient mice. After transplant,  
224 leukemia progression was monitored from the kinetics of GFP-expressing myeloid cells in  
225 peripheral blood (PB) of recipient mice. Six weeks post-transplant, *Tgifl*<sup>-/-</sup> mice showed higher  
226 numbers of GFP<sup>+</sup> cells in PB and BM, including the LSK population, compared to *Tgifl*<sup>+/+</sup> mice  
227 (Supplemental Figure S4 A, B & C, respectively). Kaplan-Meier analysis revealed that mice  
228 transplanted with BCR-ABL-GFP-transduced *Tgifl*<sup>-/-</sup> BMCs showed significantly shorter  
229 survival than mice transplanted with BCR-ABL-GFP transduced *Tgifl*<sup>+/+</sup> BMCs (Figure 7A).  
230 When recipient mice were conditioned with a lower dose of irradiation, leukemia developed with  
231 a longer latency; however, *Tgifl* genotype still significantly impacted survival (Figure 7B).  
232 Interestingly, mice transplanted with BCR-ABL-transduced haploinsufficient (*Tgifl*<sup>+/-</sup>) BMCs  
233 also had shorter survivals (Supplemental Figure S5), consistent with the greater long-term  
234 repopulating ability of *Tgifl*<sup>+/-</sup> HSCs compared to *Tgifl*<sup>+/+</sup> HSCs after transplantation<sup>26</sup>. In sum,  
235 these data are compatible with the patient and mouse AML data and suggest *Tgifl* expression  
236 could also impact survival in CML.  
237

238 ***Tgif1* gene loss affected multiple transcriptional networks in AML.** To gain insight into the  
239 biological and molecular pathways affected by *Tgif1* gene loss, we compared global gene  
240 expression profiles in *Tgif1*<sup>-/-</sup> and *Tgif1*<sup>+/+</sup> myeloid leukemia cells using RNA-seq analysis. We  
241 identified at least 45 genes that were differentially expressed in these two populations – 33 of  
242 these were significantly upregulated and 12 were down-regulated in *Tgif1*<sup>-/-</sup> relative to *Tgif1*<sup>+/+</sup>  
243 leukemia cells (Table 1).

244 Ingenuity Pathway Analysis (IPA) allowed us to interrogate which upstream regulators  
245 contributed to gene expression changes. In this analysis, genes involved in TGF-β signaling were  
246 significantly enriched (Figure 8). Other regulators identified by this analysis included all-*trans*  
247 retinoic acid (ATRA) ( $P = 1.02 \times 10^{-8}$ ) and serum response factor (SRF) ( $P = 5.34 \times 10^{-6}$ ). Like  
248 the TGF-β targets, genes activated by ATRA were upregulated by *Tgif1* loss, consistent with this  
249 transcription factor's action as a corepressor of retinoic acid receptor (RAR)-dependent  
250 transcription.

251 Genes important for embryonic stem cell pluripotency were highlighted in IPA canonical  
252 pathway analysis ( $P = 5.88 \times 10^{-4}$ ) (Supplemental Figure S6), while hematological system  
253 development and function ( $P = 1.05 \times 10^{-7}$ ) was enriched in IPA physiological functions,  
254 unsurprising given the cellular source of these RNAs. Other functions impacted by *Tgif1*  
255 expression included cellular movement ( $P = 1.99 \times 10^{-9}$ ), leukocyte function ( $P = 3.87 \times 10^{-8}$ ),  
256 cell death and survival ( $P = 5.38 \times 10^{-8}$ ), myeloid cell function ( $P = 3.24 \times 10^{-8}$ ), and cancer ( $P =$   
257  $2.63 \times 10^{-7}$ ). In summary, pathway analysis in *Tgif1* knockout leukemia cells revealed  
258 upregulation of TGF-β signaling, with potentially important consequences for LSC function.

259

260 **Discussion**

261 Our studies revealed that expression of the transcriptional regulator *TGIF1* correlates  
262 quantitatively with survival in adults with AML treated with standard induction and  
263 consolidation chemotherapy, a finding that was validated in two other, independent cohorts of  
264 patients. *TGIF1* expression impacted outcome in each cytogenetic risk group, and, at least for  
265 patients treated at our institution, was an independent risk factor in multivariate analysis.  
266 Consistent with these results, we found that loss or haploinsufficiency of this gene accelerated  
267 leukemia development, impaired response to chemotherapy, and reduced survival in mouse  
268 models of acute and chronic myeloid leukemia. Whether quantification of *TGIF1* expression will  
269 have clinical utility in risk stratification in patients with AML, however, requires further  
270 investigation.

271 *TGIF1* belongs to the TALE family of homeodomain proteins<sup>21</sup> and functions as a  
272 corepressor of retinoic acid and TGF- $\beta$ -stimulated transcription. It does so by interfering with  
273 retinoid X receptor (RXR) binding to DNA and recruiting general corepressors<sup>31</sup> and promoting  
274 recruitment of histone deacetylases to the TGF- $\beta$  signaling intermediate Smad2<sup>20-22, 32</sup>,  
275 respectively. *TGIF1* can also inhibit transcription directly through binding to a TG-rich sequence  
276 element *via* its homeobox domain. Inactivating mutations in *TGIF1* cause autosomal dominant  
277 holoprosencephaly<sup>33-36</sup>, the most common inherited defect in forebrain development in humans,  
278 while inactivation of the *Tgif1* gene in mice, complete or partial, increases quiescence in bone  
279 marrow HSCs and enhances long-term repopulating activity without an effect on steady-state  
280 hematopoiesis<sup>26</sup>.

281 In light of studies demonstrating that LSC frequency at diagnosis in AML correlates with  
282 increased minimal residual disease and poor survival<sup>37</sup>, the finding that reduced or absent  
283 expression of *Tgif1* in mouse models of myeloid leukemia enhanced LSC function provides at

284 least one explanation for how it affects recurrence. Since effective therapy after relapse is  
285 lacking in AML, except for hematopoietic stem cell transplantation in the case when a second  
286 remission can be induced, this explains how patient survival would be impacted.

287 We found using a mouse model of AML that loss of *Tgif1* resulted in a shorter latency in  
288 development of leukemia, earlier relapse after chemotherapy, and inferior survival. In addition,  
289 we found in a model of CML that heterozygous as well as homozygous loss of *Tgif1* shortened  
290 survival. These results underscore the quantitative relationship between *TGIF1* abundance and  
291 survival in AML patients and stem cell frequency in mice. They are also compatible with a  
292 recent report demonstrating that enforced expression of *TGIF1* decreased leukemia cell  
293 proliferation, induced terminal differentiation, and increased survival in MLL-AF9 leukemia <sup>38</sup>.  
294 Altogether, these data establish a biological basis for the association between *TGIF1* expression  
295 and clinical outcome.

296 In accord with its function as transcriptional repressor or corepressor, RNA expression  
297 profiling identified a number of genes that were upregulated in *Tgif1*-null leukemia cells, with  
298 almost half the genes differentially expressed in knockout and wild-type leukemia cells involved  
299 directly or indirectly in TGF- $\beta$  signaling (Supplemental Figure 6). This result also corroborates  
300 the enhanced TGF- $\beta$  signaling and/or target gene expression noted in epithelial <sup>32</sup> and myeloid  
301 leukemia cells <sup>39</sup> in which *TGIF1* expression was reduced by RNA interference. Importantly,  
302 expression of the prototypic TGF- $\beta$  target gene, *plasminogen activator inhibitor-1*, in human  
303 myeloid leukemia cells correlated significantly with expression of a type I TGF- $\beta$  receptor,  
304 activin receptor-like kinase-5 (ALK-5), and presumably downstream signaling strength. Likely  
305 for the reason offered here, *ALK-5* expression, in turn, correlated inversely with achievement of  
306 CR and with EFS, RFS, and OS <sup>40</sup>.

307 Relevant to this work, TGF- $\beta$  was shown to promote self-renewal in human HSCs <sup>41</sup> and  
308 quiescence in murine HSCs <sup>42</sup> and to enhance the function of leukemia stem cells <sup>41, 43-55</sup>, which  
309 could also explain the curative potential of a small-molecule inhibitor of the type I TGF- $\beta$   
310 receptor <sup>56, 57</sup> added to a BCR-ABL tyrosine kinase inhibitor in CML in mice <sup>51</sup>. Together with  
311 the gene expression results, these studies predict, therefore, that a TGF- $\beta$  signaling inhibitor,  
312 most logically used in combination with post-remission chemotherapy, would have the greatest  
313 impact in patients with the lowest *TGIF1* expression.

314 In addition to an effect on TGF- $\beta$  target gene expression, loss of TGIF1 caused apparent  
315 derepression of retinoic acid receptor (RAR) target genes, consistent with its ability to act as a  
316 corepressor of RAR/RXR-mediated transcription. Finally, our results do not exclude direct  
317 effects of TGIF1 on gene expression or indirect effects through interaction with other TALE  
318 homeodomain proteins, MEIS1 for example <sup>38</sup>. As chemical inhibitors are available for TGF- $\beta$  or  
319 RAR signaling, it should be possible to determine which of these signaling pathways is most  
320 important for TGIF1's actions in leukemia cells. It will also be important to determine whether  
321 inhibition of these pathways, and TGF- $\beta$  in particular, would have equivalent biological effects  
322 on leukemic compared to normal stem cells, which would have therapeutic implications.

323 Although loss-of-function mutations in *TGIF1* have been recognized in holoprosencephaly <sup>35</sup>,  
324 <sup>36</sup>, we did not observe any coding or promoter mutations or any structural rearrangements of the  
325 gene in AML patients. Furthermore, individuals with holoprosencephaly and *Tgif1*-null mice are  
326 not at increased risk for development of myeloid leukemia, all of which argues against a function  
327 of this protein as a tumor suppressor. Its graded expression in hematopoietic cells and dose-  
328 dependent effect on cell function, most clearly demonstrated for normal HSCs <sup>26</sup>, suggest that  
329 *TGIF1* acts as a quantitative trait locus in normal and malignant stem cell biology. In addition,

330 *TGIF1* exemplifies a type of cancer modifier gene that affects disease progression and  
331 persistence rather than pathogenesis.

332 In summary, by relating transcriptional profiles in myeloblasts from patients with AML to  
333 survival, we have identified a new prognostic marker relevant, potentially, to all myeloid  
334 leukemias. In addition, *TGIF1* may be the prototype of a stem cell modifier gene, operating not  
335 in the initiation of leukemia but in disease progression and persistence.

336

### 337 **Materials and Methods**

338 ***Patients.*** Pre-treatment samples from 60 patients with AML treated at Vanderbilt University  
339 Hospital or Nashville VA Medical Center from 1996-2000 were obtained with the informed  
340 consent of patients and approval of the responsible Institutional Review Board (Supplemental  
341 Table 1)<sup>47</sup>. A diagnosis of AML was made by examination of blood and/or bone marrow with  
342 cytochemical stains and with cytogenetics and flow cytometry analysis. Mononuclear cells were  
343 enriched by Ficoll-Hypaque centrifugation, depleted of T-cells and monocytes with antibodies to  
344 CD2 and CD14 (Dynal Biotech, Brown Deer, WI), respectively, and stored frozen at -80°C.  
345 Anti-CD2-conjugated beads alone were used for leukemias of FAB classes M4 and M5.

346

347 ***Classification by cytogenetic risk group.*** Patients were risk-stratified according to the results of  
348 cytogenetics analysis. The good-risk group was defined by the presence of t(8;21), t(15;17), or  
349 inv(16) and poor-risk group by del(5), del(7), t(6,9), alterations in 11q23, or multiple cytogenetic  
350 abnormalities. Patients with normal karyotypes were assigned to the intermediate-risk group.

351

352 **Survival analysis in patients.** Gene expression profiles were related to individual patient  
353 survival. Modified Cox PH regression analysis was used to rank genes by ability to predict OS,  
354 with the importance of individual genes weighted in accordance with the data available. Survival  
355 analysis in the discovery group was carried out using SPSS version 22 (Chicago, IL). Patients  
356 were enrolled between 1996 and 2001 and their outcomes last updated on June 6, 2004. OS was  
357 calculated from the time of initial diagnosis (n = 53) or relapse (n = 7) to last follow-up or death.  
358 Survival was analyzed as a function of *TGIF1* expression using both normalized hybridization  
359 and real-time PCR data. For the original microarray analysis, *TGIF1* expression in leukemia cells  
360 was divided by that in a commercial reference, log<sub>2</sub> transformed, and related to patient survival.  
361 Kaplan-Meier plots were analyzed with a log-rank test. P values were based on two-sided  
362 comparisons, with P < 0.05 considered significant. RFS was calculated from the time of  
363 diagnosis to documented relapse, SCT, death, or last follow-up. Patients who failed to achieve a  
364 CR were considered to have an RFS of 0 and all untreated patients were excluded. As described  
365 below, the patient cohort described in the study by Valk *et al*<sup>23</sup> was subdivided into five groups  
366 according to leukemia cell *TGIF1* expression and the actuarial probabilities of OS and event-free  
367 survival (EFS) were estimated from Kaplan-Meier analysis (Stata Statistical Software, release  
368 9.2; College Station, TX). Trend and log-rank analysis were carried out using Microsoft Excel  
369 (Redmond, WA). Hybridization data were normalized with MAS 5.0, log<sub>2</sub>-transformed *TGIF1*  
370 expression ratios were median-centered, and differences in *TGIF1* expression between clusters 5,  
371 9, 10, 12, and 15 were assessed by one-way analysis of variance (SAS, Cary, NC).

372

373 **Treatment.** Ninety four percent (50/53) of newly diagnosed patients received induction  
374 chemotherapy, with 82% (41/50) treated with 100 mg/m<sup>2</sup>/day of cytarabine by continuous

375 infusion from days 1-7 and 12 mg/m<sup>2</sup> of idarubicin or 45 mg/m<sup>2</sup> daunorubicin intravenously on  
376 days 1-3. Two other patients received cytarabine for five days instead of seven. Most of the  
377 remaining patients received cytarabine and an anthracycline together with other agents. Five  
378 patients received an induction regimen of dexamethasone, cytarabine, thioguanine, etoposide,  
379 and daunorubicin. Four patients with *de novo* disease did not receive induction chemotherapy,  
380 while three of seven patients with relapsed leukemia received salvage therapy.

381  
382 ***Quantitative real-time PCR analysis.*** Total cellular RNA (1 µg) was transcribed into cDNA  
383 with a mixture of random hexamer and oligo(dT) primers (iScript cDNA Synthesis Kit, Bio-Rad,  
384 Hercules, CA). Real-time PCR analysis was carried out using SYBR Green I detection (iTaq  
385 SYBR Green Supermix and MyiQ Single-Color Real-Time PCR Detection System, Bio-Rad).  
386 Amplification parameters consisted of an initial denaturation at 95°C for 3 min. This was  
387 followed by 35 cycles of denaturation at 94°C for 30 seconds, annealing at 56.7°C for 30  
388 seconds, and extension at 75°C for 10 seconds. *TGIF* expression in leukemia samples was  
389 quantified with the same reference RNA employed for the microarray studies, and relative  
390 quantification of *glyceraldehyde phosphate dehydrogenase (GAPDH)* gene expression was  
391 performed identically. Threshold cycles ( $C_T$ ) were determined for *GAPDH* and *TGIF*, and if the  
392 *GAPDH*  $C_T$  differed by three cycles or more between blast cell and reference cDNA, the  
393 concentration and quality of RNA in that patient's sample was reevaluated. Four of the five  
394 samples to which this applied were subsequently used after repeat determination of RNA  
395 concentration. Primers were designed using the MacVector software package (Accelrys, San  
396 Diego, CA). The sequences of the primers used in this analysis were: forward *TGIF* primer  
397 CAGATTCTTCGGGATTGGCTG, reverse *TGIF* primer

398 GGCGTTGATGAACCAGTTACAGAC, forward *GAPDH* primer

399 AATCCCCAGTGCTCCTTTTC, reverse *GAPDH* primer CAAAGTTGTCATCGATGACC.

400

401 ***TGIF* mutational analysis.** cDNA synthesis was carried out with the Superscript First-Strand  
402 System (Invitrogen Life Technologies, Carlsbad, CA) using 1 µg of total cellular RNA and an  
403 oligo(dT) primer. One-tenth volume of the first-strand reaction was then used as template for  
404 PCR amplification (Elongase Amplification System, Invitrogen). The sequences of the primers  
405 used in this analysis, corresponding to nucleotides 25-45 and 1489-1509, respectively, of *TGIF*  
406 RNA sequence, were: forward primer GCAGGAGCAGGGAACAAAG, reverse primer  
407 CAGTATGTGGGCATCCTGTTC. Amplification parameters consisted of an initial denaturation  
408 at 94°C for 3 min, annealing at 56°C for 30 seconds, and extension at 72°C for 1 minute. This  
409 was followed by 35 cycles of denaturation at 94°C for 30 seconds, annealing at 56°C for 30  
410 seconds, and extension at 68°C for 90 seconds and, terminally, extension at 68°C for 5 minutes.  
411 The resulting 1,477 bp cDNA product was purified in a Microcon 50 microconcentrator  
412 (Amicon, Danvers, MA) and an aliquot visualized by ethidium bromide staining after  
413 electrophoresis in a 1% agarose gel. Automated sequencing of purified PCR products was carried  
414 out with a BigDye Terminator cycle sequencing kit (v3.1, Applied Biosystems) using sets of  
415 primer encompassing the complete *TGIF* coding sequence. All sequence variants were confirmed  
416 by sequencing of both strands.

417

418 ***FLT3* mutational analysis.** The presence of an ITD or D835Y kinase domain mutation in *FLT3*  
419 was determined by reverse transcription-polymerase chain (RT-PCR) analysis of leukemic cell  
420 RNA. Total cellular RNA was purified using the RNeasy purification system (Qiagen) from

421 blasts obtained as described above. RNA was reverse transcribed using an iScript cDNA  
422 Synthesis kit according to the manufacturer's (Bio-Rad) instructions, and one-tenth volume of  
423 the RT mixture was used for PCR amplification. The sequences of the primers used in this  
424 analysis were: forward ITD TGGTGTGGTCTCCTCTTCATTGTCG, reverse ITD  
425 TCTTTCAGCATTTTGACGGCAACC, forward D835Y  
426 GAAGAAGAGGAGGACTTGAATGTGC, reverse D835Y  
427 AGGTCGCCTGTTTTGGTAGGTG. PCR was carried out with TaKaRa LA Taq polymerase  
428 (TaKaRa Mirus Bio, Madison, WI) in a 50  $\mu$ l reaction volume containing 1 $\times$  PCR buffer, 2mM  
429 MgCl<sub>2</sub>, 0.4 mM of each dNTP, 0.2  $\mu$ M of each primer, and 2.5 U of enzyme. Amplification  
430 parameters consisted of an initial denaturation at 94°C for 3 min, annealing at 56°C for 30  
431 seconds, and extension at 72°C for 1 minute, followed by 30 cycles of denaturation at 94°C for  
432 30 seconds, annealing at 56°C for 30 seconds, and extension at 72°C for 1 minute. PCR products  
433 were digested with *EcoRV* for 1 hour at 37°C for detection of the D835Y mutation. Amplified  
434 fragments and *EcoRV*-digested products were analyzed on a 3% NuSieve-agarose gel (Cambrex,  
435 Rutherford, NJ), with a *FLT3* ITD recognized from an increase in the size of the amplified  
436 fragment (>300 bp) and the D835Y mutation by the absence of *EcoRV* digestion.

437  
438 ***Evaluation of remission.*** Bone marrow aspiration and biopsy were carried out on days 14 and 28  
439 of therapy to assess response. CR was defined as the disappearance of leukemia cells from the  
440 bone marrow accompanied by return of normal blood counts. For patients with antecedent MDS,  
441 CR was defined as recovery of pre-leukemic numbers, with blasts making up  $\leq$  5% of nucleated  
442 bone marrow cells.

443

444 **Mice and plasmids.** Mice with a *Tgif1* null mutation have been previously described <sup>26</sup>.  
445 C57BL/6J mice were purchased from the Jackson Laboratories (Bar Harbor, ME). All animal  
446 experiments were approved by the Animal Care and Use Committees of Vanderbilt University  
447 and University of Virginia. MSCV-MLL-AF9-IRES-GFP vector was provided by Dr. Scott  
448 Armstrong (Memorial Sloan-Kettering Cancer Center, New York, NY) and MSCV-BCR/ABL-  
449 IRES-GFP vector was provided by Dr. Michael Engel (University of Utah, Salt Lake City, UT).

450 **Retrovirus production.** Retroviral vectors were transfected into Phoenix ecotropic retroviral  
451 packaging cells using Lipofectamine 2000 according to the manufacturer's instructions. Forty-  
452 eight hours after transfection, viral supernatants were collected, filtered, and stored at -80°C.  
453

454 **Induction of myeloid leukemia in mice.** The MSCV-MLL-AF9-IRES-GFP retrovirus was used  
455 to induce AML in mice. Bone marrow cells were flushed from the tibias and femurs of *Tgif1*<sup>+/+</sup>  
456 or *Tgif1*<sup>-/-</sup> mice and lineage marker-negative (Lin<sup>-</sup>) cells were enriched using the Mouse Lineage  
457 Cell Depletion kit (Miltenyi Biotec, San Diego, CA). Lin<sup>-</sup> cells purified from *Tgif1*<sup>+/+</sup> or *Tgif1*<sup>-/-</sup>  
458 mice were then transduced using low-speed centrifugation, or spinoculation, with MSCV-MLL-  
459 AF9-IRES-GFP retrovirus in the presence of 8 µg/ml polybrene at 1350g × 45 min (32°C).  
460 Transduced *Tgif1*<sup>+/+</sup> or *Tgif1*<sup>-/-</sup> Lin<sup>-</sup> cells were resuspended in phosphate-buffered saline (PBS)  
461 and 1 × 10<sup>5</sup> cells were injected intravenously into sub-lethally irradiated (4.5 Gy) *Tgif1*<sup>+/+</sup> or  
462 *Tgif1*<sup>-/-</sup> recipient mice, respectively. Reconstitution of transduced cells in recipient mice was  
463 evaluated by monitoring GFP expression in flow cytometry (FACS) analysis at 1-2 week  
464 intervals following transplantation. Recipient mice were euthanized for analysis when they  
465 developed palpable splenomegaly or appeared ill. Bone marrow and spleen cells from leukemic  
466 mice were isolated and stored frozen for later use.

467 The MSCV-BCR/ABL-IRES-GFP retrovirus was used to induce CML in mice. Bone marrow  
468 cells were flushed from the tibias and femurs of *Tgifl*<sup>+/+</sup>, *Tgifl*<sup>+/-</sup>, or *Tgifl*<sup>-/-</sup> mice and single-cell  
469 suspensions of Lin<sup>-</sup> c-Kit<sup>+</sup> cells were obtained using the Mouse Lineage Cell Depletion kit and  
470 CD117 microbeads (Miltenyi Biotec, San Diego, CA). Lin<sup>-</sup> c-Kit<sup>+</sup> cells were spinoculated with  
471 MSCV-BCR/ABL-IRES-GFP retrovirus in the presence of 8 µg/ml polybrene for 45 minutes at  
472 1350g (32°C). Transduced cells were resuspended in PBS and 8 × 10<sup>4</sup> cells were injected  
473 intravenously into lethally or sub-lethally irradiated C57BL/6J recipient mice. Reconstitution of  
474 transduced cells in recipient mice was evaluated by FACS analysis of GFP expression two weeks  
475 after transplantation. Mice were monitored for development of disease as described above, at  
476 which time they were euthanized.

477  
478 ***Quantification of leukemia-initiating cell frequency.*** Leukemia cells from spleens of three  
479 *Tgifl*<sup>+/+</sup> or three *Tgifl*<sup>-/-</sup> mice with AML induced with the MSCV-MLL-AF9-IRES-GFP  
480 retrovirus were pooled, and 1 × 10<sup>6</sup>, 1 × 10<sup>5</sup>, 1 × 10<sup>4</sup>, 1 × 10<sup>3</sup>, 100, and 50 cells were transplanted  
481 into sub-lethally irradiated recipients. Animals were euthanized when they became visibly ill and  
482 the development of leukemia was confirmed. LIC frequency was determined by limiting dilution  
483 analysis with the ELDA (for Extreme Limiting Dilution Analysis) software package <sup>58</sup>.

484  
485 ***Chemotherapy studies in AML mice.*** For treatment studies, 2 × 10<sup>3</sup> splenic leukemia cells from  
486 *Tgifl*<sup>+/+</sup> or *Tgifl*<sup>-/-</sup> mice were transplanted into sub-lethally irradiated recipients by tail vein  
487 injection. Two weeks after transplant, mice were treated by intraperitoneal injection with 100  
488 mg/kg cytarabine once per day for five days and 3 mg/kg doxorubicin for three days. Mice were

489 monitored closely for disease development and GFP expression was evaluated weekly by FACS  
490 analysis. Animals were euthanized when they appeared ill and transfer of disease was verified.  
491

492 **RNA sequencing (RNA-seq) analysis of AML cells.** RNA was extracted from sorted populations  
493 of GFP-expressing leukemia cells isolated from the bone marrow or spleen of *Tgif1*<sup>+/+</sup> or *Tgif1*<sup>-/-</sup>  
494 mice with AML (3 replicates each). RNA-seq and bioinformatics analysis was performed in the  
495 Vanderbilt Technologies for Advanced Genomics (or VANTAGE) Shared Resource. Sequence  
496 data were annotated with the University of California, Santa Cruz Genome Browser  
497 (<https://genome.ucsc.edu/>), and the TopHat and Cufflinks software packages were used for  
498 transcript alignment and quantification of gene expression. Gene set enrichment analysis<sup>59</sup> was  
499 carried out using the default weighted enrichment statistic to test whether differentially expressed  
500 genes (between *Tgif1*<sup>+/+</sup> and *Tgif1*<sup>-/-</sup> LSK cells) were randomly distributed or enriched at the top  
501 or bottom of the gene list. A false discovery rate of  $\leq 0.25$  was considered significant. Ingenuity  
502 Pathway Analysis (<http://www.ingenuity.com/>) was performed to group functionally related  
503 genes and assign them to biological pathways.  
504

505 **RNA preparation and cDNA microarray analysis.** Leukemic blasts from heparinized bone  
506 marrow or peripheral blood were enriched by Ficoll-Hypaque gradient centrifugation, depleted of  
507 T lymphocytes and monocytes using magnetic beads (DynaL Biotech, Brown Deer, WI)  
508 conjugated with monoclonal antibodies to CD2 and CD14, respectively, and immediately  
509 cryopreserved. In the case of myelomonocytic (FAB M4) or monocytic (M5) leukemias, only the  
510 CD2 antibody was used. All cell populations contained at least 95% blasts and had an initial  
511 viability of  $\geq 90\%$ . Total cellular RNA was prepared from thawed cells on silica-gel membranes

512 with on-column DNase treatment (RNeasy Total RNA Isolation Kit, Qiagen, Valencia,  
513 CA). RNA integrity was assessed with an Agilent Bioanalyzer (Agilent Technologies, Palo Alto,  
514 CA), and samples showing evidence of degradation were discarded. Fluorescent labeling of  
515 probes was carried out in the Microarray Shared Resource of the Vanderbilt-Ingram Cancer  
516 Center, with incorporation of amino-allyl dUTP during reverse transcription and reaction of the  
517 amine-modified cDNA with monofunctional forms of Cy3 and Cy5 dyes. Dye-labeled cDNAs  
518 were then hybridized to glass slides that had been robotically spotted with an 11,000-element  
519 human cDNA library, and fluorescence intensity was measured with an AXON or GSI Lumonics  
520 dual laser scanner. Leukemic cell cDNAs were competitively hybridized with cDNA generated  
521 from a commercial reference RNA (Universal Human Reference RNA, Stratagene, La Jolla, CA)  
522 derived from a mixture of ten human cell lines. The ratio of transcript abundance in experimental  
523 samples to the RNA reference was first transformed by  $\log_2$ . Normalization of hybridization  
524 signals was then carried out with GeneTraffic software (Iobion, La Jolla, CA) using the  
525 following parameters: signal to background intensity  $<1.3$ , Cy3 intensity  $<150$ , and Cy5 intensity  
526  $<200$ .

527

## 528 **Acknowledgements**

529 This work was supported by National Institutes of Health grants K08 HL089903 (to R.H.) and  
530 P30 CA068485 (to Vanderbilt-Ingram Cancer Center) and American Cancer Society Scholar  
531 Award RSG LIB-118462 (to R.H). We thank Drs. Ruud Delwel, Sanne Lugthart, and Bob  
532 Löwenberg for making available and analyzing microarray data from the Erasmus University  
533 study, Scott Boswell and Scott Armstrong for provision of materials, and Craig Jordan and Scott  
534 Hiebert for helpful discussions.

535

## 536 **Author Contributions**

537 L.Y., J.A.M.-P., R.H. and S.J.B. conceived the studies. V.D.K., J.P.G., U.P.D. and D.W.

538 provided study materials or patients. L.Y., J.A.M.-P., R.H., and D.M., collected and/or

539 assembled data. L.Y., R.H., Y.S., J.P.G., U.P.D., M.J.K., D.W. and S.J.B. contributed to data

540 analysis and interpretation. L.Y., J.A.M.-P., R.H. and S.J.B. drafted the manuscript. All authors

541 gave final approval of the manuscript.

542 **Conflict of Interest** The authors declare that they have no conflict of interest.

543

## 544 **References**

545 [1] Kadia, T. M., Ravandi, F., O'Brien, S., Cortes, J., and Kantarjian, H. M. (2015) Progress in  
546 acute myeloid leukemia, *Clin Lymphoma Myeloma Leuk* 15, 139-151.

547 [2] Lengfelder, E., Hofmann, W. K., and Nolte, F. (2013) Management of elderly patients with  
548 acute promyelocytic leukemia: progress and problems, *Ann Hematol* 92, 1181-1188.

549 [3] Estey, E. H. (2014) Acute myeloid leukemia: 2014 update on risk-stratification and  
550 management, *Am J Hematol* 89, 1063-1081.

551 [4] Liersch, R., Muller-Tidow, C., Berdel, W. E., and Krug, U. (2014) Prognostic factors for  
552 acute myeloid leukaemia in adults--biological significance and clinical use, *Br J*  
553 *Haematol* 165, 17-38.

554 [5] Giles, F. J., Keating, A., Goldstone, A. H., Avivi, I., Willman, C. L., and Kantarjian, H. M.  
555 (2002) Acute myeloid leukemia, *Hematology Am Soc Hematol Educ Program*, 73-110.

556 [6] Lowenberg, B. (2001) Prognostic factors in acute myeloid leukaemia, *Best Pract Res Clin*  
557 *Haematol* 14, 65-75.

558 [7] Ghanem, H., Tank, N., and Tabbara, I. A. (2012) Prognostic implications of genetic  
559 aberrations in acute myelogenous leukemia with normal cytogenetics, *Am J Hematol* 87,  
560 69-77.

561 [8] Slovak, M. L., Kopecky, K. J., Cassileth, P. A., Harrington, D. H., Theil, K. S., Mohamed,  
562 A., Paietta, E., Willman, C. L., Head, D. R., Rowe, J. M., Forman, S. J., and Appelbaum,  
563 F. R. (2000) Karyotypic analysis predicts outcome of preremission and postremission  
564 therapy in adult acute myeloid leukemia: a Southwest Oncology Group/Eastern  
565 Cooperative Oncology Group Study, *Blood* 96, 4075-4083.

566 [9] Drobyski, W. R. (2004) The role of allogeneic transplantation in high-risk acute myelogenous  
567 leukemia, *Leukemia* 18, 1565-1568.

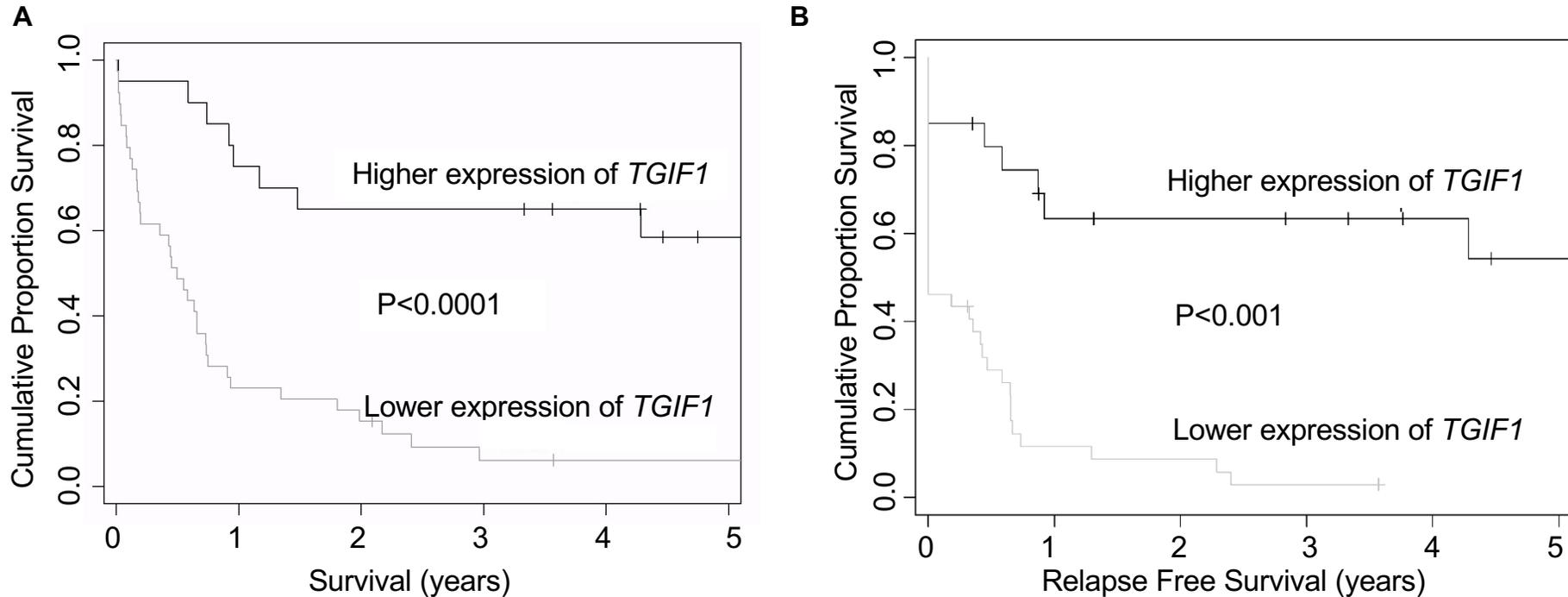
568 [10] Kottaridis, P. D., Gale, R. E., Frew, M. E., Harrison, G., Langabeer, S. E., Belton, A. A.,  
569 Walker, H., Wheatley, K., Bowen, D. T., Burnett, A. K., Goldstone, A. H., and Linch, D.  
570 C. (2001) The presence of a FLT3 internal tandem duplication in patients with acute

- 571 myeloid leukemia (AML) adds important prognostic information to cytogenetic risk  
572 group and response to the first cycle of chemotherapy: analysis of 854 patients from the  
573 United Kingdom Medical Research Council AML 10 and 12 trials, *Blood* 98, 1752-1759.
- 574 [11] Kottaridis, P. D., Gale, R. E., and Linch, D. C. (2003) Flt3 mutations and leukaemia, *Br J*  
575 *Haematol* 122, 523-538.
- 576 [12] Frohling, S., Schlenk, R. F., Breitruck, J., Benner, A., Kreitmeier, S., Tobis, K., Dohner, H.,  
577 Dohner, K., and leukemia, A. M. L. S. G. U. A. m. (2002) Prognostic significance of  
578 activating FLT3 mutations in younger adults (16 to 60 years) with acute myeloid  
579 leukemia and normal cytogenetics: a study of the AML Study Group Ulm, *Blood* 100,  
580 4372-4380.
- 581 [13] Thiede, C., Steudel, C., Mohr, B., Schaich, M., Schakel, U., Platzbecker, U., Wermke, M.,  
582 Bornhauser, M., Ritter, M., Neubauer, A., Ehninger, G., and Illmer, T. (2002) Analysis of  
583 FLT3-activating mutations in 979 patients with acute myelogenous leukemia: association  
584 with FAB subtypes and identification of subgroups with poor prognosis, *Blood* 99, 4326-  
585 4335.
- 586 [14] Barjesteh van Waalwijk van Doorn-Khosrovani, S., Erpelinck, C., van Putten, W. L., Valk,  
587 P. J., van der Poel-van de Luytgaarde, S., Hack, R., Slater, R., Smit, E. M., Beverloo, H.  
588 B., Verhoef, G., Verdonck, L. F., Ossenkoppele, G. J., Sonneveld, P., de Greef, G. E.,  
589 Lowenberg, B., and Delwel, R. (2003) High EVI1 expression predicts poor survival in  
590 acute myeloid leukemia: a study of 319 de novo AML patients, *Blood* 101, 837-845.
- 591 [15] Inokuchi, K., Yamaguchi, H., Hanawa, H., Tanosaki, S., Nakamura, K., Tarusawa, M.,  
592 Miyake, K., Shimada, T., and Dan, K. (2002) Loss of DCC gene expression is of  
593 prognostic importance in acute myelogenous leukemia, *Clin Cancer Res* 8, 1882-1888.
- 594 [16] Baldus, C. D., Tanner, S. M., Ruppert, A. S., Whitman, S. P., Archer, K. J., Marcucci, G.,  
595 Caligiuri, M. A., Carroll, A. J., Vardiman, J. W., Powell, B. L., Allen, S. L., Moore, J. O.,  
596 Larson, R. A., Koltz, J. E., de la Chapelle, A., and Bloomfield, C. D. (2003) BAALC  
597 expression predicts clinical outcome of de novo acute myeloid leukemia patients with  
598 normal cytogenetics: a Cancer and Leukemia Group B Study, *Blood* 102, 1613-1618.
- 599 [17] Barjesteh van Waalwijk van Doorn-Khosrovani, S., Erpelinck, C., Meijer, J., van  
600 Oosterhoud, S., van Putten, W. L., Valk, P. J., Berna Beverloo, H., Tenen, D. G.,  
601 Lowenberg, B., and Delwel, R. (2003) Biallelic mutations in the CEBPA gene and low  
602 CEBPA expression levels as prognostic markers in intermediate-risk AML, *Hematol J* 4,  
603 31-40.
- 604 [18] Frohling, S., and Dohner, H. (2004) Disruption of C/EBPalpha function in acute myeloid  
605 leukemia, *N Engl J Med* 351, 2370-2372.
- 606 [19] Hamid, R., Patterson, J., and Brandt, S. J. (2008) Genomic structure, alternative splicing and  
607 expression of TG-interacting factor, in human myeloid leukemia blasts and cell lines,  
608 *Biochim Biophys Acta* 1779, 347-355.
- 609 [20] Wotton, D., Knoepfler, P. S., Laherty, C. D., Eisenman, R. N., and Massague, J. (2001) The  
610 Smad transcriptional corepressor TGIF recruits mSin3, *Cell Growth Differ* 12, 457-463.
- 611 [21] Wotton, D., Lo, R. S., Lee, S., and Massague, J. (1999) A Smad transcriptional corepressor,  
612 *Cell* 97, 29-39.
- 613 [22] Wotton, D., Lo, R. S., Swaby, L. A., and Massague, J. (1999) Multiple modes of repression  
614 by the Smad transcriptional corepressor TGIF, *J Biol Chem* 274, 37105-37110.
- 615 [23] Valk, P. J., Verhaak, R. G., Beijen, M. A., Erpelinck, C. A., Barjesteh van Waalwijk van  
616 Doorn-Khosrovani, S., Boer, J. M., Beverloo, H. B., Moorhouse, M. J., van der Spek, P.

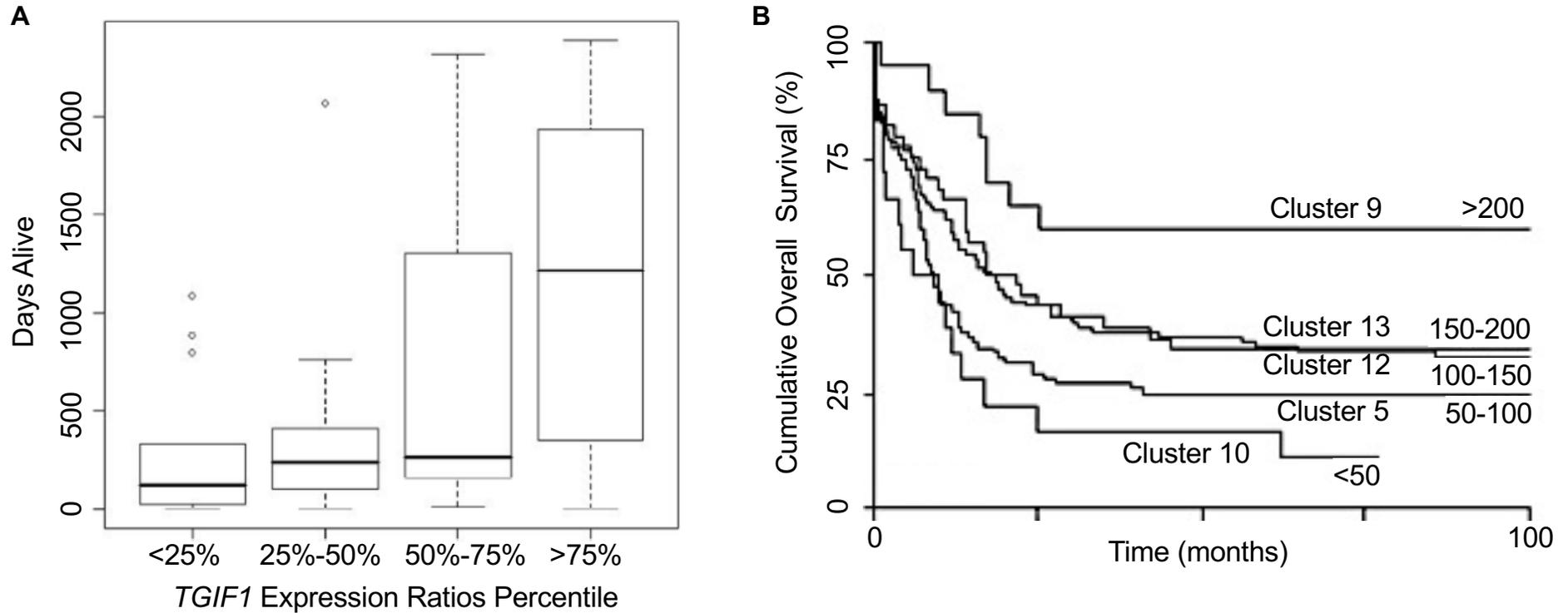
- 617 J., Lowenberg, B., and Delwel, R. (2004) Prognostically useful gene-expression profiles  
618 in acute myeloid leukemia, *N Engl J Med* 350, 1617-1628.
- 619 [24] Quentmeier, H., Reinhardt, J., Zaborski, M., and Drexler, H. G. (2003) FLT3 mutations in  
620 acute myeloid leukemia cell lines, *Leukemia* 17, 120-124.
- 621 [25] Whitman, S. P., Archer, K. J., Feng, L., Baldus, C., Becknell, B., Carlson, B. D., Carroll, A.  
622 J., Mrozek, K., Vardiman, J. W., George, S. L., Kolitz, J. E., Larson, R. A., Bloomfield,  
623 C. D., and Caligiuri, M. A. (2001) Absence of the wild-type allele predicts poor  
624 prognosis in adult de novo acute myeloid leukemia with normal cytogenetics and the  
625 internal tandem duplication of FLT3: a cancer and leukemia group B study, *Cancer Res*  
626 61, 7233-7239.
- 627 [26] Yan, L., Womack, B., Wotton, D., Guo, Y., Shyr, Y., Dave, U., Li, C., Hiebert, S., Brandt,  
628 S., and Hamid, R. (2013) Tgfl1 regulates quiescence and self-renewal of hematopoietic  
629 stem cells, *Mol Cell Biol* 33, 4824-4833.
- 630 [27] Kirstetter, P., Schuster, M. B., Bereshchenko, O., Moore, S., Dvinge, H., Kurz, E.,  
631 Theilgaard-Monch, K., Mansson, R., Pedersen, T. A., Pabst, T., Schrock, E., Porse, B. T.,  
632 Jacobsen, S. E., Bertone, P., Tenen, D. G., and Nerlov, C. (2008) Modeling of  
633 C/EBPalpha mutant acute myeloid leukemia reveals a common expression signature of  
634 committed myeloid leukemia-initiating cells, *Cancer Cell* 13, 299-310.
- 635 [28] Shlush, L. I., Zandi, S., Mitchell, A., Chen, W. C., Brandwein, J. M., Gupta, V., Kennedy, J.  
636 A., Schimmer, A. D., Schuh, A. C., Yee, K. W., McLeod, J. L., Doedens, M., Medeiros,  
637 J. J., Marke, R., Kim, H. J., Lee, K., McPherson, J. D., Hudson, T. J., Consortium, H. P.-  
638 L. G. P., Brown, A. M., Yousif, F., Trinh, Q. M., Stein, L. D., Minden, M. D., Wang, J.  
639 C., and Dick, J. E. (2014) Identification of pre-leukaemic haematopoietic stem cells in  
640 acute leukaemia, *Nature* 506, 328-333.
- 641 [29] Taussig, D. C., Vargaftig, J., Miraki-Moud, F., Griessinger, E., Sharrock, K., Luke, T.,  
642 Lillington, D., Oakervee, H., Cavenagh, J., Agrawal, S. G., Lister, T. A., Gribben, J. G.,  
643 and Bonnet, D. (2010) Leukemia-initiating cells from some acute myeloid leukemia  
644 patients with mutated nucleophosmin reside in the CD34(-) fraction, *Blood* 115, 1976-  
645 1984.
- 646 [30] Zuber, J., Radtke, I., Pardee, T. S., Zhao, Z., Rappaport, A. R., Luo, W., McCurrach, M. E.,  
647 Yang, M. M., Dolan, M. E., Kogan, S. C., Downing, J. R., and Lowe, S. W. (2009)  
648 Mouse models of human AML accurately predict chemotherapy response, *Genes Dev* 23,  
649 877-889.
- 650 [31] Bertolino, E., Reimund, B., Wildt-Perinic, D., and Clerc, R. G. (1995) A novel homeobox  
651 protein which recognizes a TGT core and functionally interferes with a retinoid-  
652 responsive motif, *J Biol Chem* 270, 31178-31188.
- 653 [32] Seo, S. R., Lallemand, F., Ferrand, N., Pessah, M., L'Hoste, S., Camonis, J., and Atfi, A.  
654 (2004) The novel E3 ubiquitin ligase Tiul1 associates with TGIF to target Smad2 for  
655 degradation, *EMBO J* 23, 3780-3792.
- 656 [33] Ming, J. E., Kaupas, M. E., Roessler, E., Brunner, H. G., Golabi, M., Tekin, M., Stratton, R.  
657 F., Sujansky, E., Bale, S. J., and Muenke, M. (2002) Mutations in PATCHED-1, the  
658 receptor for SONIC HEDGEHOG, are associated with holoprosencephaly, *Hum Genet*  
659 110, 297-301.
- 660 [34] Ming, J. E., and Muenke, M. (2002) Multiple hits during early embryonic development:  
661 digenic diseases and holoprosencephaly, *Am J Hum Genet* 71, 1017-1032.

- 662 [35] Wallis, D., and Muenke, M. (2000) Mutations in holoprosencephaly, *Hum Mutat* 16, 99-  
663 108.
- 664 [36] Wallis, D. E., and Muenke, M. (1999) Molecular mechanisms of holoprosencephaly, *Mol*  
665 *Genet Metab* 68, 126-138.
- 666 [37] van Rhenen, A., Feller, N., Kelder, A., Westra, A. H., Rombouts, E., Zweegman, S., van der  
667 Pol, M. A., Waisfisz, Q., Ossenkoppele, G. J., and Schuurhuis, G. J. (2005) High stem  
668 cell frequency in acute myeloid leukemia at diagnosis predicts high minimal residual  
669 disease and poor survival, *Clin Cancer Res* 11, 6520-6527.
- 670 [38] Willer, A., Jakobsen, J. S., Ohlsson, E., Rapin, N., Waage, J., Billing, M., Bullinger, L.,  
671 Karlsson, S., and Porse, B. T. (2014) TGIF1 is a negative regulator of MLL-rearranged  
672 acute myeloid leukemia, *Leukemia*.
- 673 [39] Hamid, R., and Brandt, S. J. (2009) Transforming growth-interacting factor (TGIF)  
674 regulates proliferation and differentiation of human myeloid leukemia cells, *Mol Oncol* 3,  
675 451-463.
- 676 [40] Otten, J., Schmitz, L., Vettorazzi, E., Schultze, A., Marx, A. H., Simon, R., Krauter, J.,  
677 Loges, S., Sauter, G., Bokemeyer, C., and Fiedler, W. (2011) Expression of TGF-beta  
678 receptor ALK-5 has a negative impact on outcome of patients with acute myeloid  
679 leukemia, *Leukemia* 25, 375-379.
- 680 [41] Pierelli, L., Marone, M., Bonanno, G., Rutella, S., de Ritis, D., Mancuso, S., Leone, G., and  
681 Scambia, G. (2002) Transforming growth factor-beta1 causes transcriptional activation of  
682 CD34 and preserves haematopoietic stem/progenitor cell activity, *Br J Haematol* 118,  
683 627-637.
- 684 [42] Yamazaki, S., Iwama, A., Takayanagi, S., Eto, K., Ema, H., and Nakauchi, H. (2009) TGF-  
685 beta as a candidate bone marrow niche signal to induce hematopoietic stem cell  
686 hibernation, *Blood* 113, 1250-1256.
- 687 [43] Falk, L. A., De Benedetti, F., Lohrey, N., Birchenall-Roberts, M. C., Ellingsworth, L. W.,  
688 Faltynek, C. R., and Ruscetti, F. W. (1991) Induction of transforming growth factor-beta  
689 1 (TGF-beta 1), receptor expression and TGF-beta 1 protein production in retinoic acid-  
690 treated HL-60 cells: possible TGF-beta 1-mediated autocrine inhibition, *Blood* 77, 1248-  
691 1255.
- 692 [44] Imai, Y., Kurokawa, M., Izutsu, K., Hangaishi, A., Maki, K., Ogawa, S., Chiba, S., Mitani,  
693 K., and Hirai, H. (2001) Mutations of the Smad4 gene in acute myelogenous leukemia  
694 and their functional implications in leukemogenesis, *Oncogene* 20, 88-96.
- 695 [45] Kim, S. J., and Letterio, J. (2003) Transforming growth factor-beta signaling in normal and  
696 malignant hematopoiesis, *Leukemia* 17, 1731-1737.
- 697 [46] Krause, D. S., Fulzele, K., Catic, A., Sun, C. C., Dombkowski, D., Hurley, M. P., Lezeau,  
698 S., Attar, E., Wu, J. Y., Lin, H. Y., Divieti-Pajevic, P., Hasserjian, R. P., Schipani, E.,  
699 Van Etten, R. A., and Scadden, D. T. (2013) Differential regulation of myeloid leukemias  
700 by the bone marrow microenvironment, *Nat Med* 19, 1513-1517.
- 701 [47] Kurokawa, M., Mitani, K., Imai, Y., Ogawa, S., Yazaki, Y., and Hirai, H. (1998) The  
702 t(3;21) fusion product, AML1/Evi-1, interacts with Smad3 and blocks transforming  
703 growth factor-beta-mediated growth inhibition of myeloid cells, *Blood* 92, 4003-4012.
- 704 [48] Kurokawa, M., Mitani, K., Irie, K., Matsuyama, T., Takahashi, T., Chiba, S., Yazaki, Y.,  
705 Matsumoto, K., and Hirai, H. (1998) The oncoprotein Evi-1 represses TGF-beta  
706 signalling by inhibiting Smad3, *Nature* 394, 92-96.

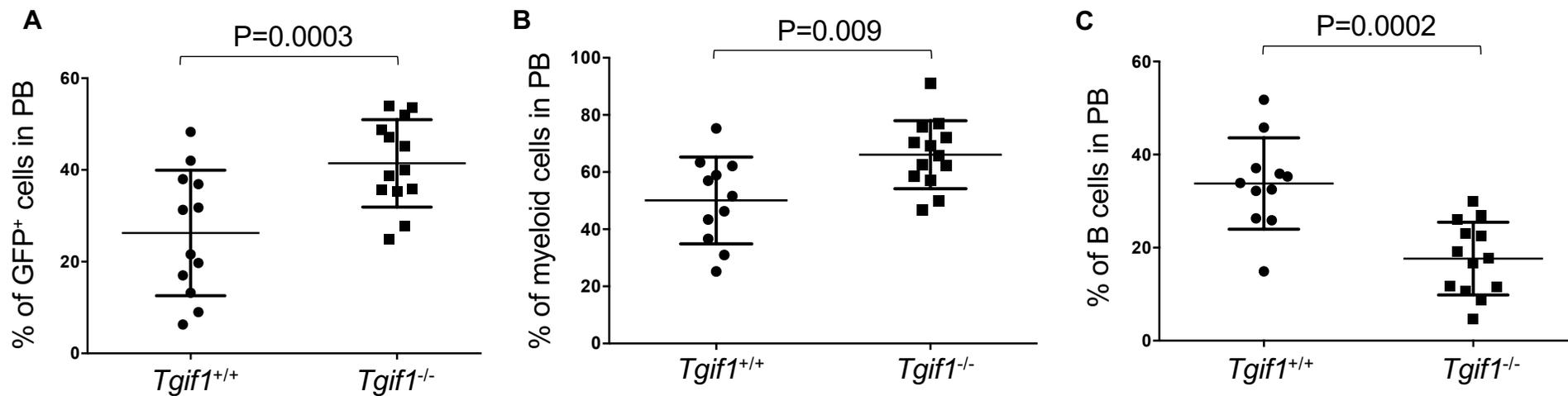
- 707 [49] Marone, M., Scambia, G., Bonanno, G., Rutella, S., de Ritis, D., Guidi, F., Leone, G., and  
708 Pierelli, L. (2002) Transforming growth factor-beta1 transcriptionally activates CD34 and  
709 prevents induced differentiation of TF-1 cells in the absence of any cell-cycle effects,  
710 *Leukemia* 16, 94-105.
- 711 [50] Murohashi, I., Endho, K., Nishida, S., Yoshida, S., Jinnai, I., Bessho, M., and Hirashima, K.  
712 (1995) Differential effects of TGF-beta 1 on normal and leukemic human hematopoietic  
713 cell proliferation, *Exp Hematol* 23, 970-977.
- 714 [51] Naka, K., Hoshii, T., Muraguchi, T., Tadokoro, Y., Ooshio, T., Kondo, Y., Nakao, S.,  
715 Motoyama, N., and Hirao, A. (2010) TGF-beta-FOXO signalling maintains leukaemia-  
716 initiating cells in chronic myeloid leukaemia, *Nature* 463, 676-680.
- 717 [52] Siegel, P. M., and Massague, J. (2003) Cytostatic and apoptotic actions of TGF-beta in  
718 homeostasis and cancer, *Nat Rev Cancer* 3, 807-821.
- 719 [53] Sitnicka, E., Ruscetti, F. W., Priestley, G. V., Wolf, N. S., and Bartelmez, S. H. (1996)  
720 Transforming growth factor beta 1 directly and reversibly inhibits the initial cell divisions  
721 of long-term repopulating hematopoietic stem cells, *Blood* 88, 82-88.
- 722 [54] Soucek, K., Pachernik, J., Kubala, L., Vondracek, J., Hofmanova, J., and Kozubik, A.  
723 (2006) Transforming growth factor-beta1 inhibits all-trans retinoic acid-induced  
724 apoptosis, *Leuk Res* 30, 607-623.
- 725 [55] Tabe, Y., Shi, Y. X., Zeng, Z., Jin, L., Shikami, M., Hatanaka, Y., Miida, T., Hsu, F. J.,  
726 Andreeff, M., and Konopleva, M. (2013) TGF-beta-Neutralizing Antibody 1D11  
727 Enhances Cytarabine-Induced Apoptosis in AML Cells in the Bone Marrow  
728 Microenvironment, *PLoS One* 8, e62785.
- 729 [56] Akhurst, R. J., and Hata, A. (2012) Targeting the TGFbeta signalling pathway in disease,  
730 *Nat Rev Drug Discov* 11, 790-811.
- 731 [57] Yingling, J. M., Blanchard, K. L., and Sawyer, J. S. (2004) Development of TGF-beta  
732 signalling inhibitors for cancer therapy, *Nat Rev Drug Discov* 3, 1011-1022.
- 733 [58] Hu, Y., and Smyth, G. K. (2009) ELDA: extreme limiting dilution analysis for comparing  
734 depleted and enriched populations in stem cell and other assays, *J Immunol Methods* 347,  
735 70-78.
- 736 [59] Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A.,  
737 Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., and Mesirov, J. P. (2005)  
738 Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide  
739 expression profiles, *Proc Natl Acad Sci U S A* 102, 15545-15550.
- 740



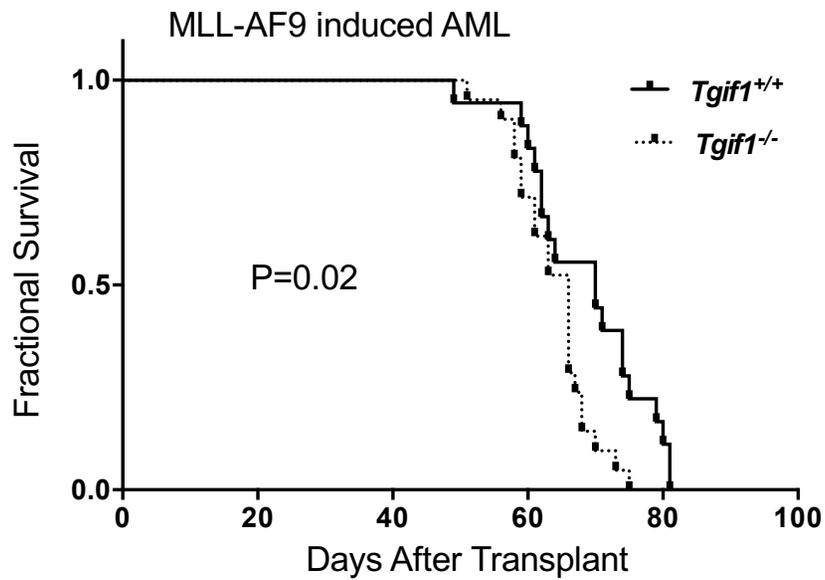
**Figure 1. Lower *TGIF1* expression was associated with adverse and higher *TGIF1* expression associated with favorable clinical outcome in AML. A, OS in patients (n = 20) with higher expression relative to reference is shown in top curve. Median survival in this group was not reached at last follow-up, with 70% of this group still alive. The survival of patients (n = 39) with lower expression is shown in bottom curve. Median survival in this group was 181 days (CI = 112-249 days), with 7.6% still alive. B, RFS in patients (n = 20) with *TGIF1* expression greater than reference is shown in top curve. The survival of patients (n = 39) with lower *TGIF1* expression is shown in bottom curve.**



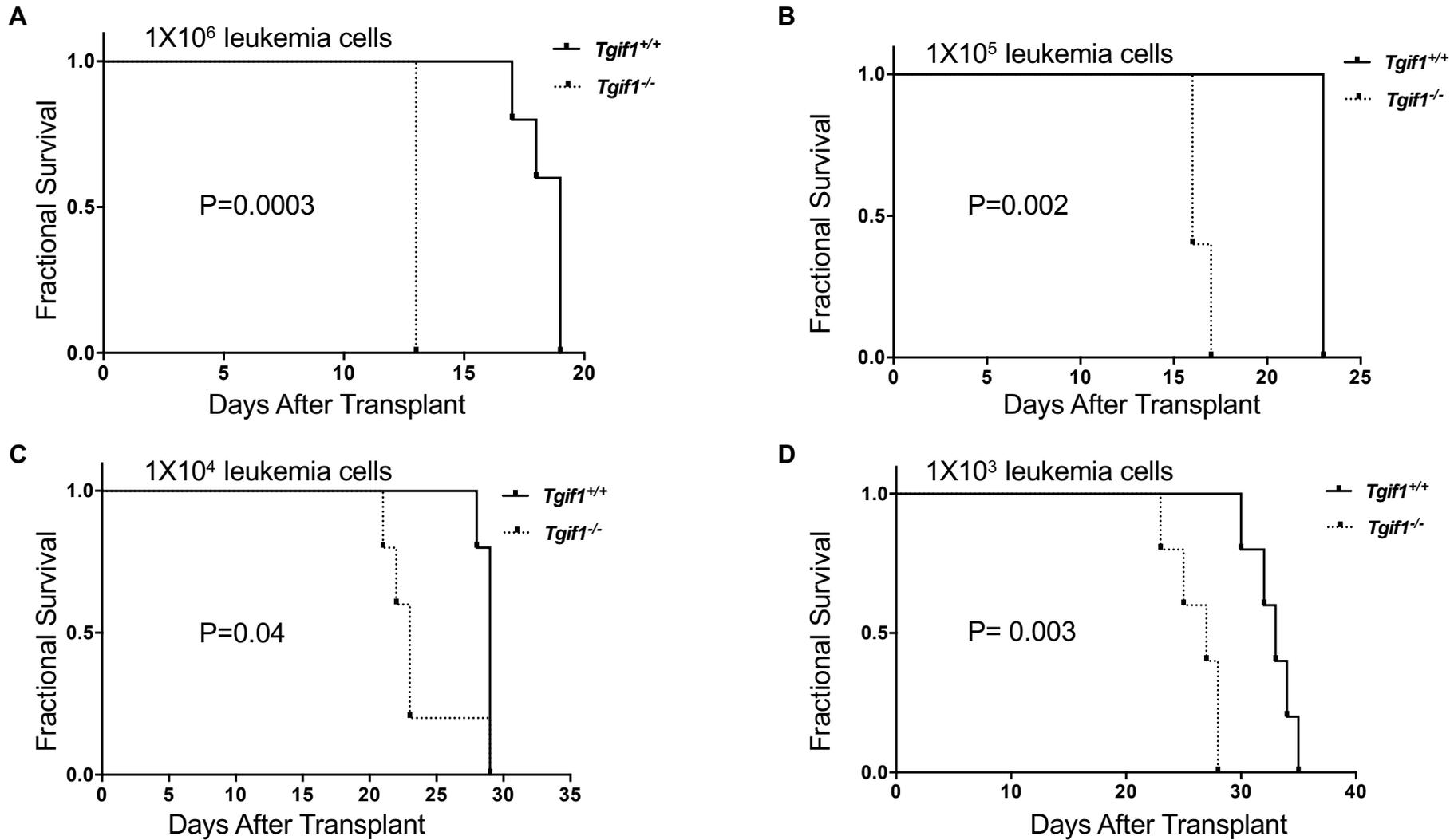
**Figure 2. Analysis of OS in patients as a function of *TGIF1* mRNA expression.** **A**, Box-and-whiskers plot of OS expressed as days alive vs. ratio of *TGIF1* expression in leukemia cells to commercial reference. Cases were subdivided into 4 groups according to *TGIF1* expression ratios as shown. **B**, *TGIF1* expression determined in 285 AML patients with Affymetrix U133A GeneChips (probe set 203313\_s\_at) and cases subdivided according to *TGIF1* expression into 5 groups: < 50 (n = 20), 51-100 (n = 45), 101-150 (n = 107), 151-200 (n = 95), and > 200 (n = 18). Log-rank analysis was carried out using Microsoft Excel software.



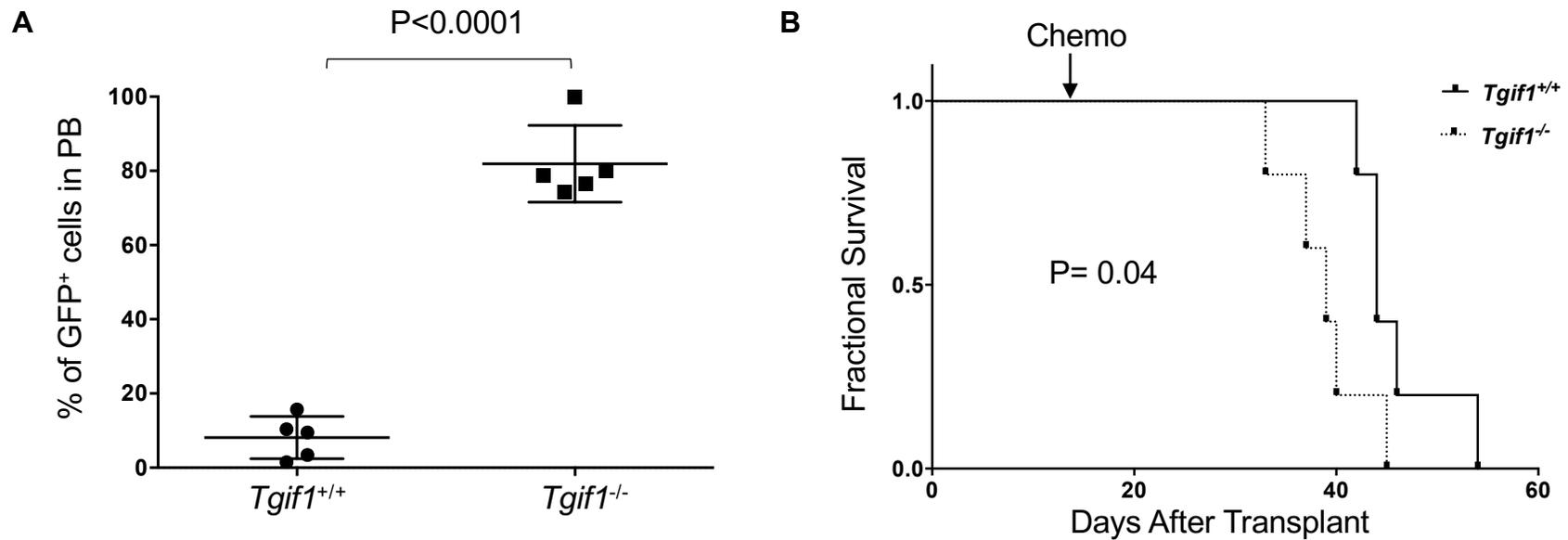
**Figure 3. MLL-AF9-induced AML in *Tgif1*<sup>-/-</sup> mice was more aggressive than in *Tgif1*<sup>+/+</sup> mice.** Lin<sup>-</sup> bone marrow cells from *Tgif1*<sup>+/+</sup> or *Tgif1*<sup>-/-</sup> mice were transduced with MLL-AF9-GFP retrovirus and transduced cells transplanted into sub-lethally irradiated *Tgif1*<sup>+/+</sup> and *Tgif1*<sup>-/-</sup> recipient mice, respectively. Six weeks post-transplant, the percentages of **A**, GFP<sup>+</sup> cells; **B**, myeloid cells (CD11b<sup>+</sup> and Gr-1<sup>+</sup>) and **C**, B-lymphocytes (B220<sup>+</sup>) in peripheral blood (PB) were analyzed by flow cytometry.



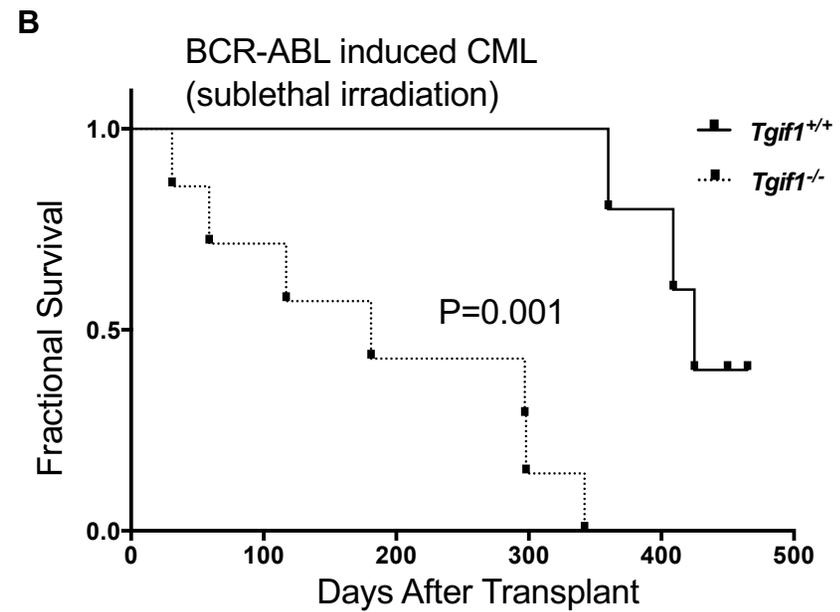
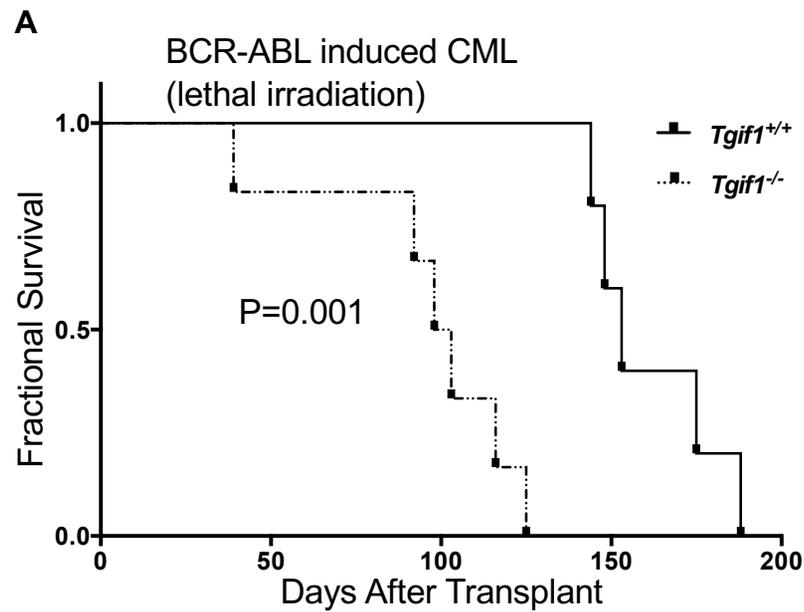
**Figure 4. *Tgif1* loss was associated with shorter survival *in vivo*.** Kaplan-Meier survival curves for mice with the indicated genotype and MLL-AF9-induced AML. Lin<sup>-</sup> bone marrow cells from *Tgif1*<sup>+/+</sup> and *Tgif1*<sup>-/-</sup> mice were transduced with MLL-AF9-GFP retrovirus and transduced cells transplanted into sub-lethally irradiated *Tgif1*<sup>+/+</sup> and *Tgif1*<sup>-/-</sup> recipients, respectively.



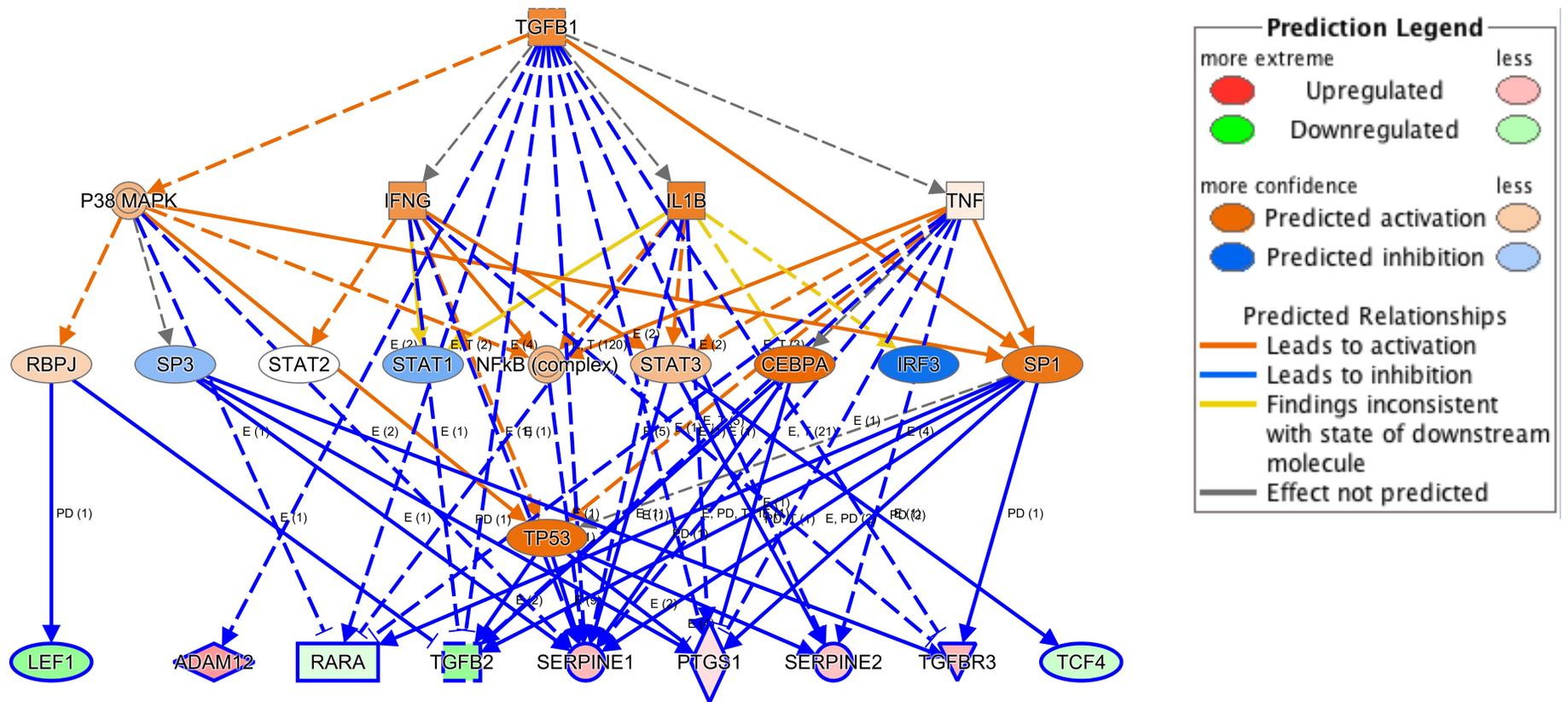
**Figure 5. LSC frequency was increased in *Tgif1*<sup>-/-</sup> mice with AML compared to *Tgif1*<sup>+/+</sup> mice.** Lin<sup>-</sup> bone marrow cells from *Tgif1*<sup>+/+</sup> and *Tgif1*<sup>-/-</sup> mice were transduced with MLL-AF9-GFP retrovirus and transduced cells transplanted into sub-lethally irradiated *Tgif1*<sup>+/+</sup> and *Tgif1*<sup>-/-</sup> recipients, respectively. At a predetermined time, serial dilutions of spleen cells from mice with leukemia then collected and transplanted into secondary recipients. Kaplan-Meier analysis of survival is plotted as a function of number of transplanted spleen cells. **A**, 1x10<sup>6</sup>; **B**, 1x10<sup>5</sup>; **C**, 1x10<sup>4</sup>; **D**, 1x10<sup>3</sup>.



**Figure 6. AML in *Tgif1*<sup>-/-</sup> mice progressed earlier after chemotherapy than in *Tgif1*<sup>+/+</sup> AML mice.** Sub-lethally irradiated C57BL/6 mice were transplanted with spleen cells from *Tgif1*<sup>+/+</sup> or *Tgif1*<sup>-/-</sup> mice with established MLL-AF9-induced AML, and two weeks following transplant, recipient mice were treated with doxorubicin for 3 days and cytarabine for 5 days by intraperitoneal injection. **A**, Percentage of GFP<sup>+</sup> cells in peripheral blood two weeks post chemotherapy. **B**, Kaplan-Meier analysis of survival of *Tgif1*<sup>+/+</sup> and *Tgif1*<sup>-/-</sup> mice with AML after chemotherapy.



**Figure 7. *Tgif1*<sup>-/-</sup> mice with BCR-ABL-induced CML had inferior survival compared to *Tgif1*<sup>+/+</sup> mice with CML.** Lin<sup>-</sup> c-Kit<sup>+</sup> bone marrow cells from *Tgif1*<sup>+/+</sup> or *Tgif1*<sup>-/-</sup> mice were transduced with BCR-ABL-GFP retrovirus and transduced cells transplanted into C57BL/6J recipient mice. Recipient mice were **A**, lethally irradiated; **B**, sub-lethally irradiated.



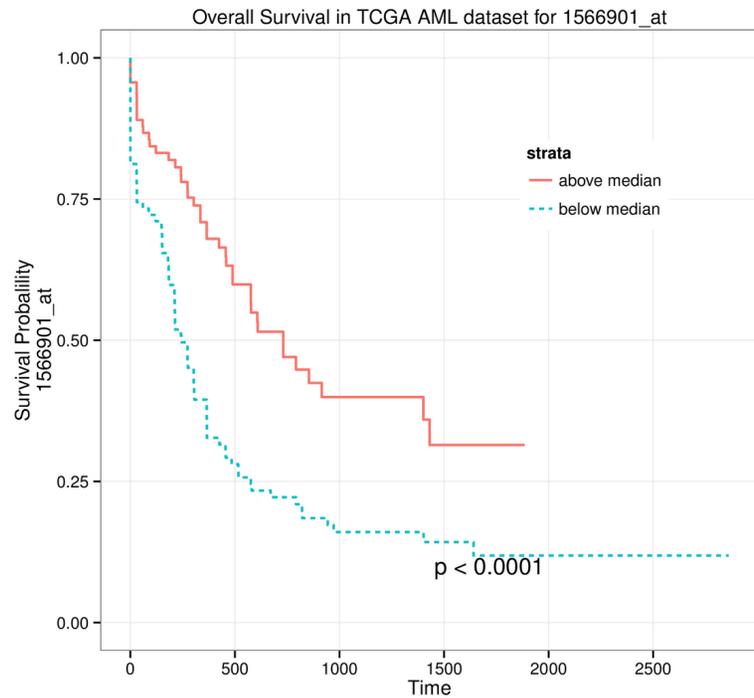
**Figure 8. Upstream regulator analysis of differentially expressed genes.** This analysis identified TGF-b as a major regulator of differentially expressed genes ( $p=8.52 \times 10^{-10}$ ).

**Table 1. Differentially expressed genes in *TGIF1*<sup>-/-</sup> vs. *TGIF1*<sup>+/+</sup> leukemic cells**

ID	Symbol	Location	Log Ratio
<b>Slc15a2</b>	SLC15A2	Plasma Membrane	-7.651
<b>Eps8l1</b>	EPS8L1	Cytoplasm	-7.587
<b>Dlgap1</b>	DLGAP1	Plasma Membrane	-3.719
<b>Gdf3</b>	GDF3	Extracellular Space	-3.658
<b>Camk2b</b>	CAMK2B	Cytoplasm	-3.619
<b>Col19a1</b>	COL19A1	Extracellular Space	-3.488
<b>Kcnq5</b>	KCNQ5	Plasma Membrane	-2.727
<b>Vill</b>	VILL	Cytoplasm	-2.499
<b>Cgn1</b>	CGNL1	Plasma Membrane	-1.938
<b>Cdkn2c</b>	CDKN2C	Nucleus	-1.649
<b>Gm15448</b>	LILRB3	Plasma Membrane	-1.362
<b>Selm</b>	SELM	Cytoplasm	-0.761
<b>Ly75</b>	LY75	Plasma Membrane	0.652
<b>Anxa3</b>	ANXA3	Cytoplasm	0.714
<b>Agap1</b>	AGAP1	Cytoplasm	0.742
<b>Ttc21a</b>	TTC21A	Extracellular Space	0.967
<b>Pip5k1b</b>	PIP5K1B	Cytoplasm	0.976
<b>B4galt6</b>	B4GALT6	Cytoplasm	1.029
<b>Ptgs1</b>	PTGS1	Cytoplasm	1.045
<b>Olfm4</b>	OLFM4	Extracellular Space	1.123
<b>Dsp</b>	DSP	Plasma Membrane	1.212
<b>Gfi1</b>	GFI1	Nucleus	1.278
<b>Plekha6</b>	Plekha6	Other	1.291
<b>Sept5</b>	SEPT5	Cytoplasm	1.301
<b>Optn</b>	OPTN	Cytoplasm	1.325
<b>Kcnh7</b>	KCNH7	Plasma Membrane	1.417
<b>Fcnb</b>	FCN1	Extracellular Space	1.427
<b>S100a8</b>	S100A8	Cytoplasm	1.503
<b>Gca</b>	GCA	Cytoplasm	1.575
<b>Serpine2</b>	SERPINE2	Extracellular Space	1.875
<b>Camp</b>	CAMP	Cytoplasm	1.940
<b>Serpine1</b>	SERPINE1	Extracellular Space	2.160
<b>Retnlg</b>	Retnlg	Extracellular Space	2.167
<b>Cd74</b>	CD74	Plasma Membrane	2.609
<b>C1qb</b>	C1QB	Extracellular Space	3.161
<b>C1qc</b>	C1QC	Extracellular Space	3.269
<b>Adam12</b>	ADAM12	Plasma Membrane	3.167
<b>C1qa</b>	C1QA	Extracellular Space	3.503
<b>Mmp14</b>	MMP14	Extracellular Space	4.032
<b>Thbs4</b>	THBS4	Extracellular Space	4.578
<b>Bmp1</b>	BMP1	Extracellular Space	4.577
<b>Npr1</b>	NPR1	Plasma Membrane	5.607
<b>Mtus2</b>	MTUS2	Other	5.713
<b>Syce1</b>	SYCE1	Nucleus	7.905
<b>4930447C04Rik</b>	C14orf39	Extracellular Space	8.430

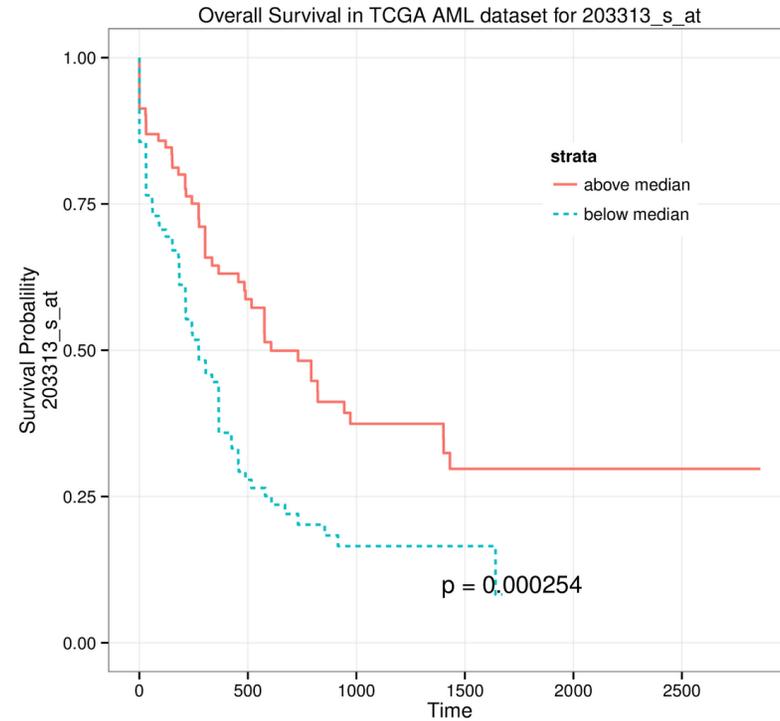
## Supplemental Table S1. Patient characteristics.

Characteristic	Number	Percentage
Total number	60	
Median age	48 (range 3-79)	
Age 3-17	6	10%
Age 18-45	21	35%
Age 46-60	19	32%
Age 60-79	14	23%
Males	28	47%
Females	32	53%
Newly diagnosed	53	88%
Relapsed	7	12%
Prior MDS	21	35%
Cytogenetics		
Good prognosis	6	10%
Intermediate prognosis	27	45%
Poor prognosis	22	37%
Induction therapy	49/53 newly diagnosed	92%
Salvage therapy	3/7 relapsed	43%
Achieved CR	35	58%
No response or PR	17	28%
Remission not evaluated	8	13%
SCT in CR1 or CR2	20	33%
Patients surviving	17	25%



<b>above median</b>	90	36	15	6	0	0
<b>below median</b>	82	24	13	8	4	1

Numbers at risk



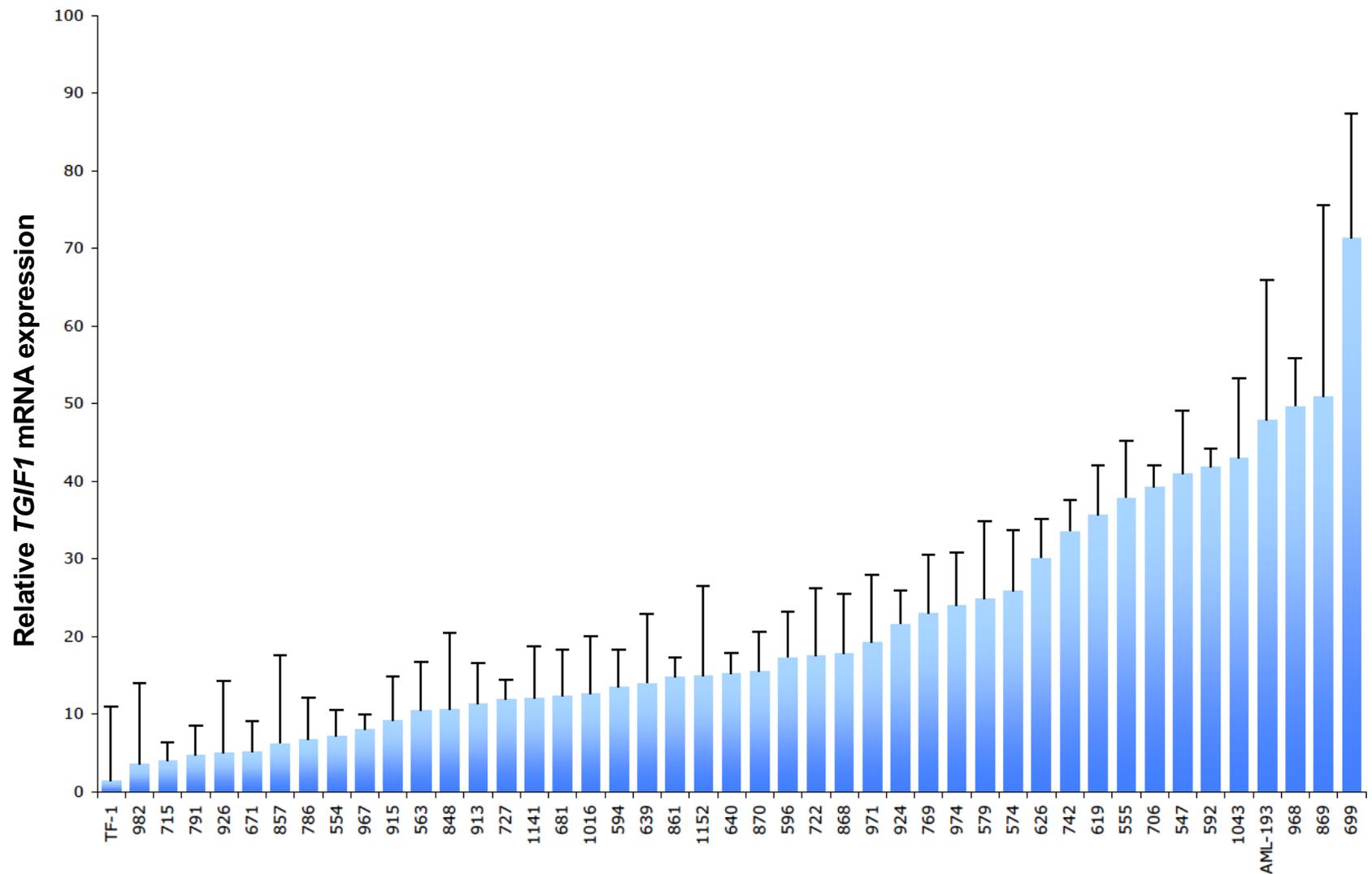
<b>above median</b>	87	40	20	10	4	1
<b>below median</b>	85	20	8	4	0	0

Numbers at risk

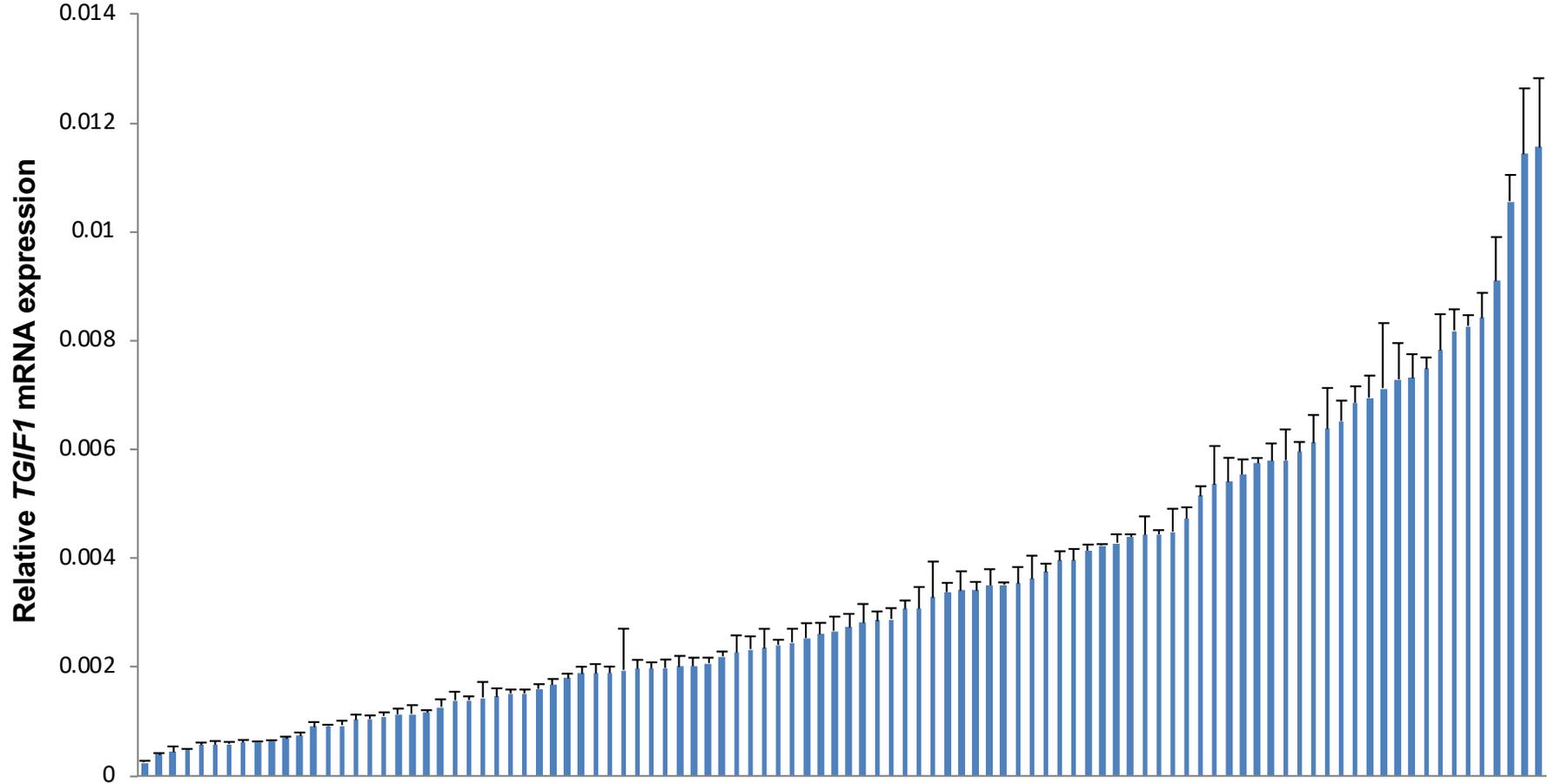
**Supplemental Figure S1. Kaplan-Meier analysis of OS as function of *TGIF1* expression: TCGA AML dataset.** Kaplan-Meier analysis of OS for adult patients with AML from TCGA dataset, with *TGIF1* expression above the median (red continuous lines) or below the median (blue dotted lines) plotted for probe set 203313\_s\_at (*left*) and 1566901\_at (*right*).

**Supplemental Table S2. TGIF1 sequence variants in myeloblasts.**

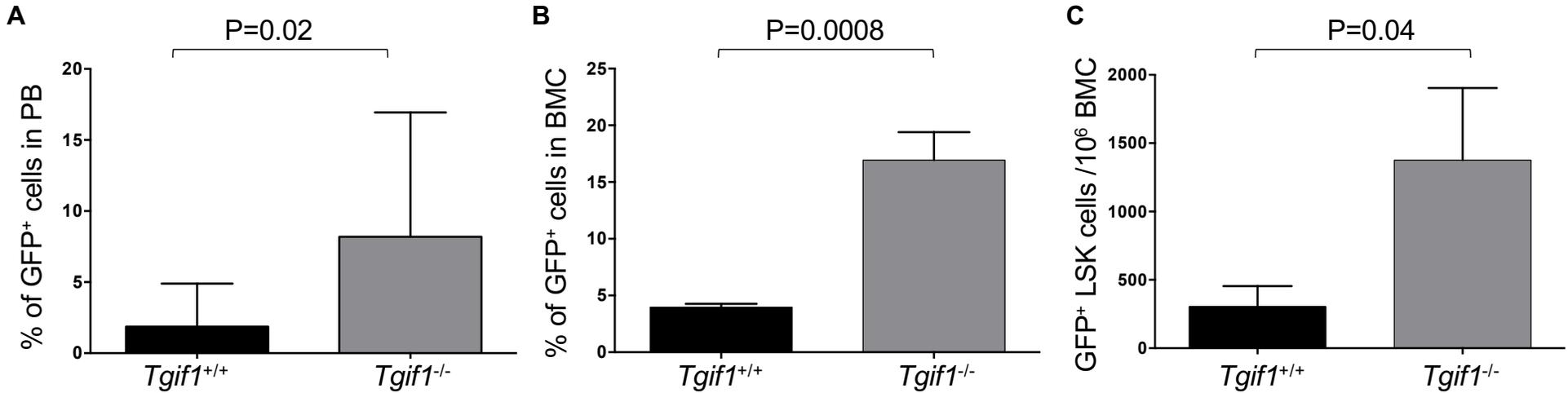
Polymorphism	TGIF region	Previously reported	Number
-33C>A	5' untranslated	Yes	9/49
-9G>A	5' untranslated	No	1/49
127C>G (P43A)	Coding	No	1/49
420A>G (P140P)	Smad2-interacting	Yes	7/49
487C>T (P163S)	Smad2-interacting	Yes	7/49
488C>T (P163L)	Smad2-interacting	Yes	7/49
713C>A (P238H)	Phosphorylation site	No	1/49
723G>A (P241P)	Phosphorylation site	No	1/49
744T>A (S248R)	mSin3A-interacting	No	1/49



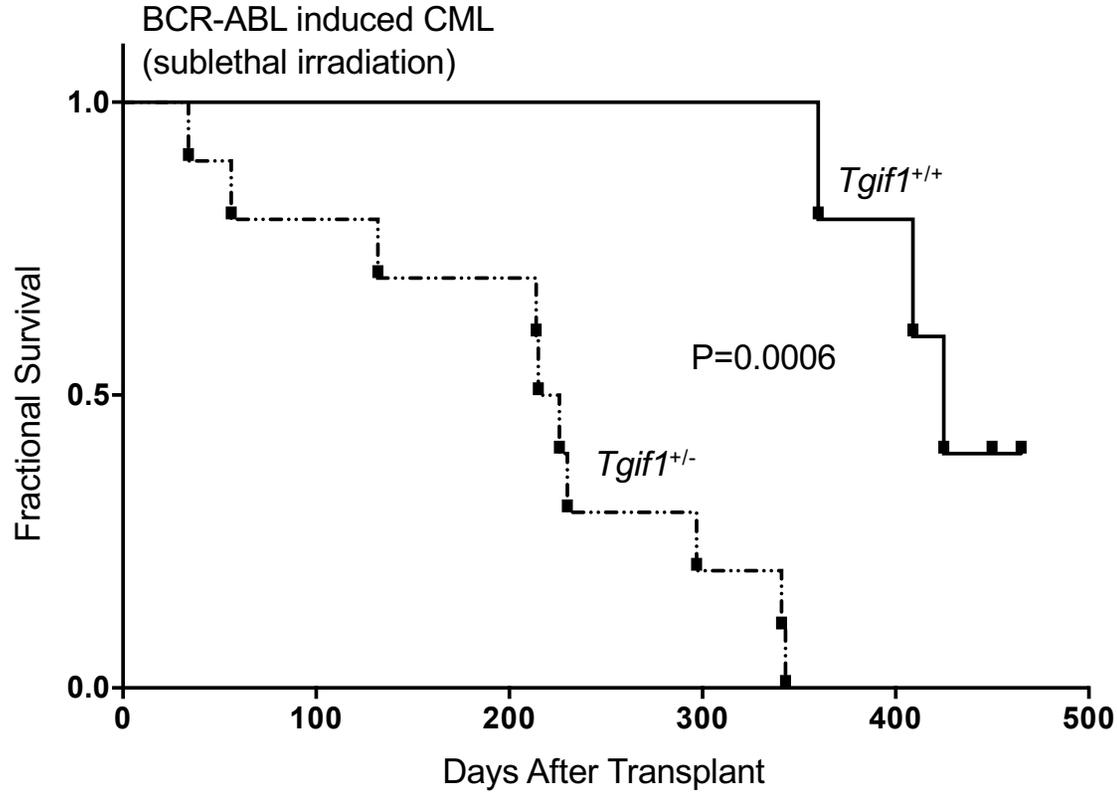
**Supplemental Figure S2. Quantitative RT-PCR analysis of *TGIF1* mRNA in leukemic blasts.** Bar graph shows, in ascending order, mean *TGIF1* mRNA level  $\pm$  SEM in myeloblasts from 43 patients with AML determined by real-time PCR analysis. *TGIF1* expression in human AML cell lines TF-1 and AML-193 are shown for comparison. The level of expression in the TF-1 cell line was assigned a value of 1 and all other measurements normalized to it.



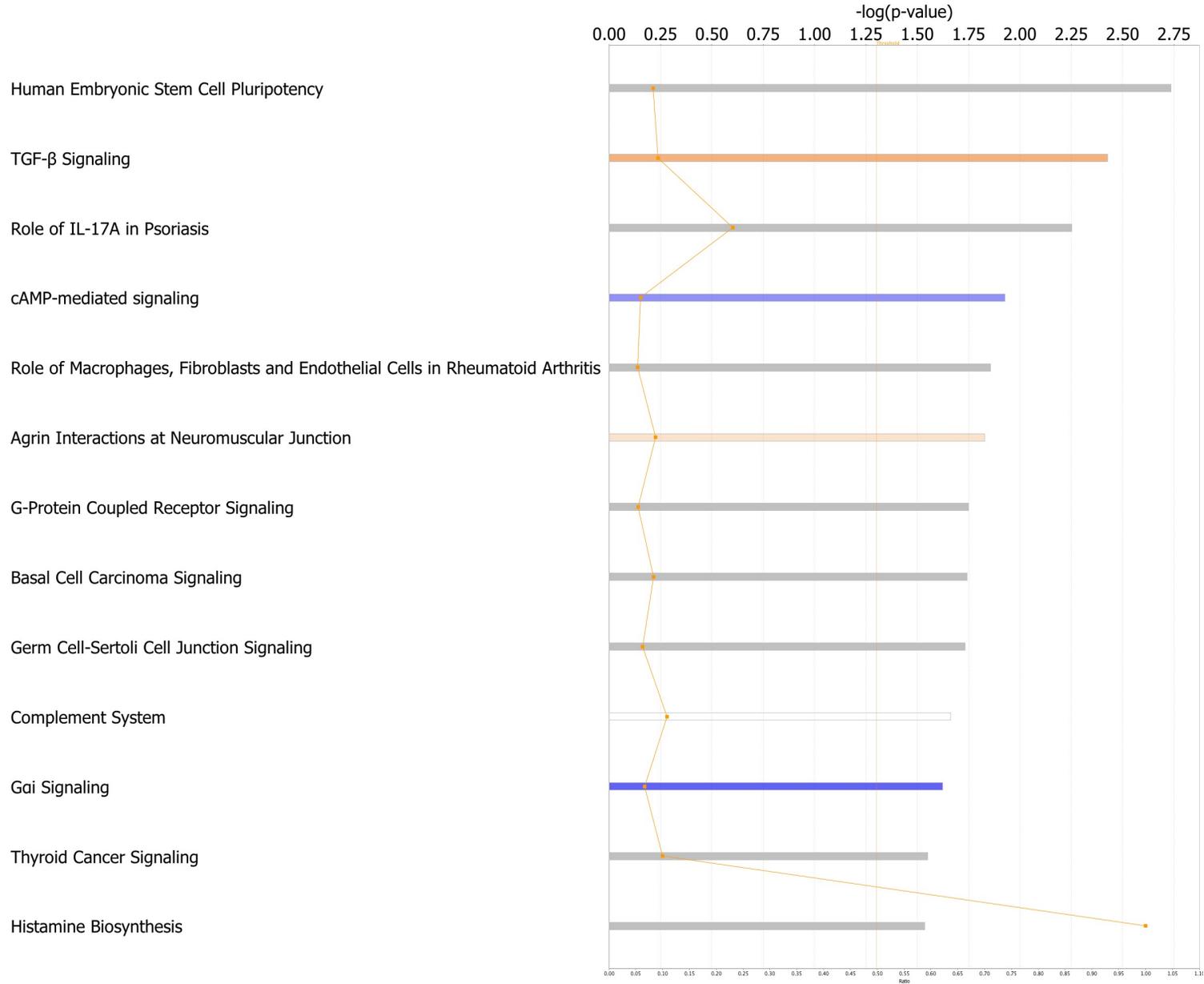
**Supplemental Figure S3. Quantitative RT-PCR analysis of *TGIF1* mRNA in immortalized lymphoblasts.** Bar graph shows mean *TGIF1* mRNA level  $\pm$  SEM in ascending order in immortalized lymphoblast lines derived from 100 normal individuals.



**Supplemental Figure S4. *Tgif1*<sup>-/-</sup> mice with CML have more aggressive disease than in *Tgif1*<sup>+/+</sup> mice.** Enriched Lin<sup>-</sup> c-Kit<sup>+</sup> cells from *Tgif1*<sup>+/+</sup> or *Tgif1*<sup>-/-</sup> bone marrow cells were transduced with BCR-ABL-GFP retrovirus and transduced cells transplanted into lethally irradiated C57BL/6J recipients to induce CML. Six weeks post-transplant, flow cytometry was used to analyze **A**, percentage of GFP<sup>+</sup> cells in PB; **B**, percentage of GFP<sup>+</sup> cells in bone marrow; **C**, number of GFP<sup>+</sup> LSK cells (or LSCs) in bone marrow.



**Supplemental Figure S5. Kaplan-Meier analysis of survival in BCR-ABL-induced CML in mice as a function of *Tgif1* expression level.** Enriched Lin<sup>-</sup> c-Kit<sup>+</sup> cells from *Tgif1*<sup>+/+</sup> or *Tgif1*<sup>+/-</sup> bone marrow cells were transduced with BCR-ABL-GFP retrovirus and transduced cells transplanted into sub-lethally irradiated C57BL/6J recipients to induce CML.



**Supplemental Figure S6. Canonical Pathway Analysis.** Depicted are differentially utilized pathways in leukemia cells from *Tgfl*<sup>-/-</sup> compared to *Tgfl*<sup>+/+</sup> mice.