

## ORIGINAL ARTICLE

## Overexpression of SOX4 correlates with poor prognosis of acute myeloid leukemia and is leukemogenic in zebrafish

J-W Lu<sup>1,4</sup>, M-S Hsieh<sup>1</sup>, H-A Hou<sup>2</sup>, C-Y Chen<sup>2</sup>, H-F Tien<sup>2</sup> and L-I Lin<sup>1,3</sup>

The SOX4 transcription factor is a key regulator of embryonic development, cell-fate decision, cellular differentiation and oncogenesis. Abnormal expression of SOX4 is related to malignant tumor transformation and cancer metastasis. However, no reports are available regarding the clinical significance of SOX4 in acute myeloid leukemia (AML) and the role of SOX4 in leukemogenesis. In the current study, we found that AML patients with low bone marrow (BM) SOX4 expression had higher remission rates and longer overall survival than those with high SOX4 expression, regardless of age, white blood cell count at diagnosis, karyotype profile and *NPM1/FLT3-ITD* status. To elucidate the role of SOX4 in leukemogenesis, we generated a transgenic zebrafish model that overexpressed human SOX4 in the myeloid lineage Tg(*spi1-SOX4-EGFP*). These transgenic zebrafish showed, at 5 months of age, increased myelopoiesis with dedifferentiation in kidney marrow. At 9 months of age, their kidney structure was significantly effaced and distorted by increased infiltration of myeloid progenitor cells. These results suggest that SOX4 is not only an independent prognostic factor of AML, but also an important molecular factor in leukemogenesis.

Blood Cancer Journal (2017) 7, e593; doi:10.1038/bcj.2017.74; published online 25 August 2017

## INTRODUCTION

The SOX4 belongs to the SOX (Sry-related high-mobility group box) family and has been characterized as a transcription factor.<sup>1,2</sup> Over the past decade, multiple functions of SOX4 have been unveiled, and the protein is now known to play important roles in embryonic development, cell-fate decision and cellular differentiation.<sup>3–6</sup> Overexpression and amplification of SOX4 have been implicated in various cancers and are correlated with poor prognosis.<sup>7–16</sup> In mouse models, previous studies demonstrated that the upregulation of *Sox4* can be induced by and then cooperate with the aberrant expression of *AML1-ETO*,<sup>17</sup> *NUP98-DDX10*<sup>(ref. 18)</sup> and *PML-RAR $\alpha$* ;<sup>19</sup> the overexpression of *HOXA9*<sup>(ref. 20)</sup> *CREB*<sup>21</sup> and *Evi1*;<sup>22</sup> and the haplosufficiency of *PU.1*<sup>(ref. 23)</sup> to trigger leukemogenesis. Conversely, mice lacking *Sox4* have shown pro-B-lymphocyte expansion and defects in cardiac outflow tract formation.<sup>24</sup> Furthermore, a previous study that employed retroviral transduction of *Sox4* and bone marrow transplantation techniques revealed that increased *Sox4* expression may cooperate with the deregulation of *Mef2c* expression to induce myeloid leukemia in recipient mice.<sup>25</sup>

Recently, *Sox4* gene was reported to be a direct target of C/EBP $\alpha$  (CCAAT/enhancer-binding protein- $\alpha$ ). C/EBP $\alpha$  is known to inhibit the self-renewal of leukemic cells and to restore cellular differentiation. The overexpression of *Sox4* that results from C/EBP $\alpha$  inactivation contributes to the development of a type of leukemia that is characterized by a distinct leukemia-initiating cell phenotype. This recent work further indicated that *Sox4* is a key oncogenic target and critical mediator of C/EBP $\alpha$  mutants in acute myeloid leukemia (AML), suggesting a potential novel therapeutic approach to the treatment of this disease.<sup>26,27</sup> However, the

clinical implications of SOX4 expression and its role of AML leukemogenesis are not well understood.

The current study investigated the relationship between bone marrow (BM) SOX4 expression and clinicopathological parameters of *de novo* AML and also evaluated the prognostic value of SOX4 expression for AML patients. This is the first study to report on the prognostic implications of SOX4 expression for AML patients. Our immunohistochemical staining results illustrated that high BM SOX4 protein content is an independent unfavorable prognostic factor for overall survival. A meta-analysis that we conducted using an online data cohort retrieved from PrognScan (a new database for meta-analysis of the prognostic value of genes; <http://www.abren.net/PrognScan/>) revealed similar findings. Finally, as several zebrafish models have been proposed for study of hematopoiesis and myeloid malignancies in recent years,<sup>28–32</sup> we used a transgenic SOX4 zebrafish line to demonstrate, for the first time, that SOX4 overexpression driven by *spi-1* (a myeloid-specific promoter) can lead to expanded myelopoiesis with leukemic phenotype.

## MATERIALS AND METHODS

## Study patients

From March 2009 to December 2011, a total number of 112 adult patients were enrolled in this study. This study was approved by the institutional review board of the National Taiwan University Hospital and written informed consent was obtained from all participants in accordance with the Declaration of Helsinki. See Supplementary Materials and Methods for details.

<sup>1</sup>Department of Clinical Laboratory Sciences and Medical Biotechnology, National Taiwan University, Taipei, Taiwan; <sup>2</sup>Division of Hematology, Department of Internal Medicine, National Taiwan University, Taipei, Taiwan and <sup>3</sup>Laboratory Medicine, National Taiwan University Hospital, Taipei, Taiwan. Correspondence: Professor L-I Lin, Department of Clinical Laboratory Sciences and Medical Biotechnology, College of Medicine, National Taiwan University, No. 1, Chang-Te St., Taipei, Taiwan. E-mail: lillin@ntu.edu.tw

<sup>4</sup>Current address: Department of Biological Sciences, National University of Singapore, Singapore.

Received 5 July 2017; accepted 10 July 2017

### Immunocytochemical staining of SOX4 protein

To assess SOX4 expression in leukemic cells, immunocytochemical staining was performed.<sup>33</sup> See Supplementary Materials and Methods for details.

### Generation and husbandry of transgenic zebrafish

Zebrafish (*Danio rerio*) embryos, larvae and adult fish were maintained at 28 °C under continuous flow and a 14 h light/10 h dark cycle.<sup>31</sup> All experiments involving zebrafish were approved by the institutional animal care and use committee of the National Taiwan University. The myeloid-specific *spi-1*-driven SOX4 transgenic fish was generated as previously described.<sup>29,34</sup> The primer sequences are listed in Supplementary Table 1. See Supplementary Materials and Methods for details.

### Isolation of RNA as well as reverse transcription-PCR

Total RNA from various tissues or a total number of 30 embryos was isolated using NucleoSpin and TRIzol (Invitrogen, Carlsbad, CA, USA), then reverse transcribed into complementary DNA (cDNA) with the High-Capacity RNA-to-cDNA Kit (Applied Biosystems, Carlsbad, CA, USA) and finally amplified by PCR with KOD-FX Taq polymerase (Toyobo, Osaka, Japan), in accordance with the manufacturer's instructions. See Supplementary Materials and Methods for details.

### Quantitative reverse transcription-PCR

Quantitative reverse transcription-PCR was carried out in an ABI 7500 Fast Real-Time PCR system (Applied Biosystems, Foster City, CA, USA) using SYBR green as the detection dye (Power SYBR Green PCR Master Mix, Applied Biosystems). PCR conditions and subsequently cycle thresholds (Ct) calculations were according to standard protocols. The primer sequences of various target genes are listed in Supplementary Table 1. See Supplementary Materials and Methods for details.

### Whole-mount *in situ* hybridization and whole mount immunohistochemical staining

Whole-mount *in situ* hybridization analysis was performed essentially as previously described<sup>32</sup> using the antisense digoxigenin-labeled *myeloperoxidase* (*mpo*) RNA probe that was generated from the partial cDNA (901 bp) sequence of *mpo* was amplified by PCR with T3-*mpo*-F and T7-*mpo*-R primers (Supplementary Table 1). See Supplementary Materials and Methods for details.

For whole-mount immunohistochemical staining analysis, embryos were fixed, dehydrated and rehydrated as procedures for whole-mount *in situ* hybridization. Embryos were then incubated in the blocking solution, reacted with rabbit anti-GFP antibodies (1:100 dilutions; GTX113617, GeneTex, San Antonio, TX, USA), and then finally reacted with Alexa Fluor 488-labeled goat-anti rabbit secondary antibodies (1:100 dilutions; A-11008, Invitrogen). See Supplementary Materials and Methods for details.

### Tissue collection and histochemical analysis

Tissue collection and histochemical analysis were performed as previously described.<sup>29</sup> See Supplementary Materials and Methods for details.

### Cytological analysis of kidney marrow and peripheral blood

Collection and cytological analysis of blood cells from kidney marrow (KM) and peripheral blood were performed as previously described.<sup>29</sup> See Supplementary Materials and Methods for details.

### Flow cytometric analysis

Collection and flow cytometric analysis of blood cells from KM and peripheral blood were performed as previously described.<sup>29</sup> See Supplementary Materials and Methods for details.

### Sudan black staining

Sudan black staining was performed as previously described.<sup>35</sup> See Supplementary Materials and Methods for details.

### Myeloperoxidase staining

KM smears were fixed with pH 6.6 formalin-acetone (Muto Pure Chemicals Co., Ltd, Tokyo, Japan) for 30 s, washed with double-distilled water and then dried. Slides were subsequently treated with 3,3'-diaminobenzidine-tetrahydrochloride (Sigma Chem. Co., St Louis, MO, USA) for 15 min at room temperature. After being washed and dried, slides were counterstained with hematoxylin for 3 min.

### Statistical analysis

All statistical analyses performed for this study involved comparison between experimental and control groups using two-tailed Student's *t*-tests, Mann-Whitney *U*-tests, one-way analysis of variance,  $\chi^2$  test or Fisher's exact test and multivariate analysis with Cox proportional hazards regression models as previously described.<sup>36</sup> We used Kaplan-Meier estimation techniques to plot survival curves and log-rank tests to examine difference between groups as previously described.<sup>37</sup> *P*-values of < 0.05 were considered statistically significant. See Supplementary Materials and Methods for details.

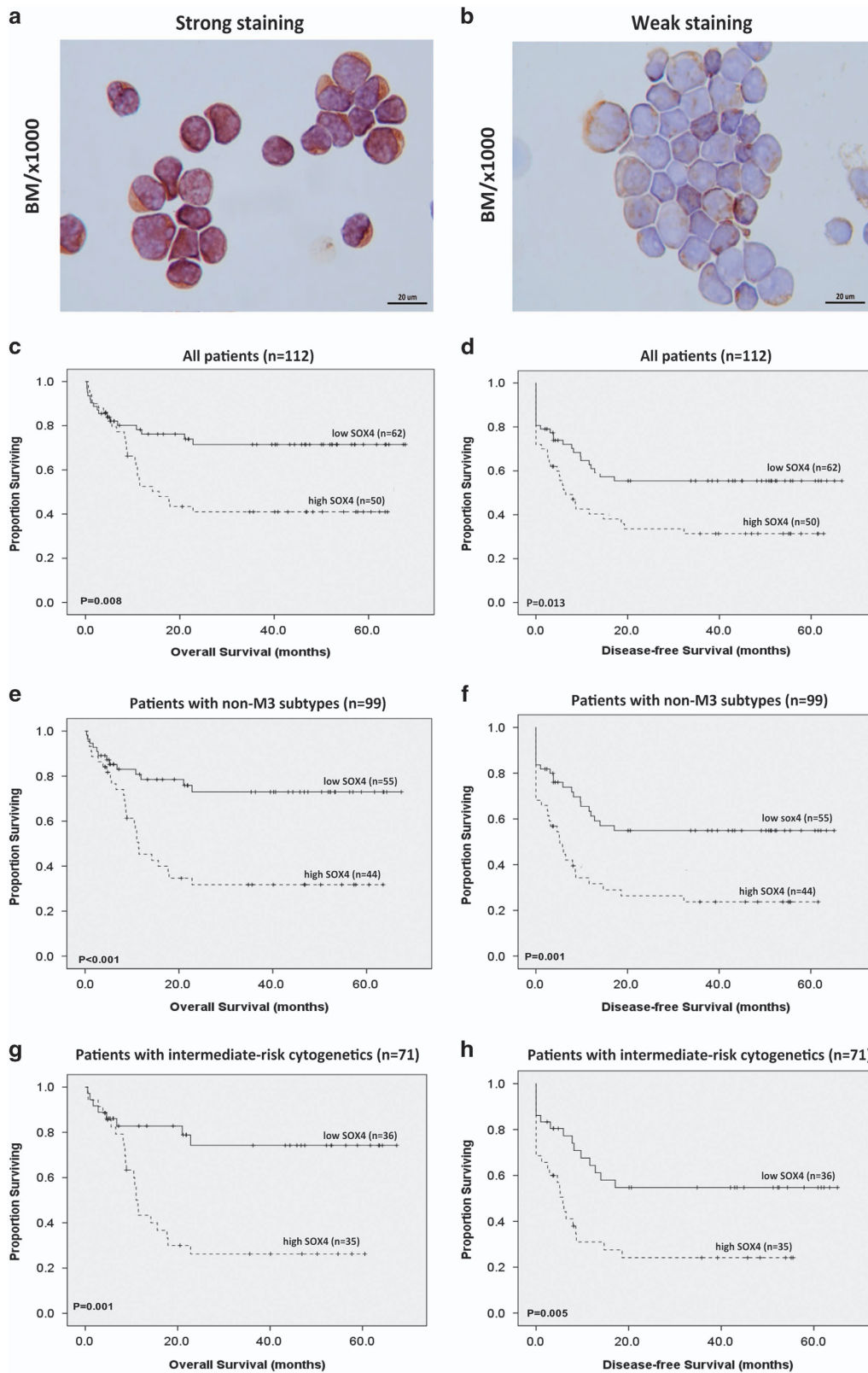
## RESULTS

BM SOX4 expression as an independent prognostic factor of AML Results of immunocytochemical staining revealed that BM SOX4 expression varied greatly in AML patients. We divided patients into two groups according to the intensity and extent of SOX4 expression as follows: low expression group (score 0–2, *n* = 62, Figure 1a) and high expression group (score 3–4, *n* = 50, Figure 1b), respectively. The various clinical manifestations of AML did not show significant differences in terms of SOX4 expression (Supplementary Table 2); except that high SOX4 expression was somehow correlated with CD11b expression in leukemia cells (Supplementary Table 3), and AML patients with low SOX4 expression tended to have favorable-risk karyotyping (*P* = 0.0866, Supplementary Table 2). We did not observe significant differences between the high and low expression groups in terms of age, gender, hemograms, *nucleophosmin-1* (*NPM1*) mutation and *FLT3*-ITD (internal tandem duplication of the *fms*-like tyrosine kinase-3). In addition, of the 112 AML patients who underwent conventional intensive induction chemotherapy, 85 (75.9%) achieved complete remission, and the high and low expression groups showed similar probabilities of achieving first complete remission (36/50, 72 vs 49/62, 79%, *P* = 0.3219). However, high SOX4 expression was associated with increased relapse rates compared with low SOX4 expression (19/36, 52.8 vs 13/49, 26.5%, *P* = 0.028).

Furthermore, with a median follow-up period of 46.7 months (range: 0.3–70.9 months), SOX4 expression was associated with overall survival and disease-free survival in all patients with *de novo* AML (*P* = 0.008 and *P* = 0.013, respectively, Figure 1c), patients with non-M3 subtypes (*P* < 0.001 and *P* = 0.001, respectively, Figure 1d), patients with intermediate-risk cytogenetics (*P* = 0.001 and *P* = 0.005 respectively, Figure 1e) or even in those with normal karyotype profile (*P* = 0.022 and *P* = 0.111, Figure 1f). In multivariate analysis, high SOX4 expression was found to be an independent poor prognostic factor of overall survival (relative risk 1.924, 95% confidence interval 1.020–3.628, *P* = 0.043) irrespective of age, white blood cell count at diagnosis, karyotype profile and *NPM1/FLT3*-ITD status (Table 1).

### Generation of Tg(*spi1:SOX4-EGFP*) zebrafish lines by tol2 transposon system

To further investigate the role of SOX4 in leukemogenesis, we use Multisite Gateway (Invitrogen, Grand Island, NY, USA) and Tol2 transposon methods to generate pTolCG-*spi1:SOX4-EGFP* constructs. In these constructs, SOX4 and EGFP expression is controlled by the myeloid-specific *spi1* (also known as *pu.1*) promoter. The pTolCG-*spi1:SOX4-EGFP* constructs were then coinjected with Tol2 transposase mRNA into one-cell embryos of



**Figure 1.** Kaplan–Meier survival analysis according to SOX4 expression in AML patients. Representative immunocytochemical stain of SOX4 protein in bone marrow specimens from patients with low SOX4 protein (**a**) and with high SOX4 protein (**b**). Overall survival and disease-free survival were analyzed in all patients with *de novo* AML (**c**, **d**), in patients with non-M3 subtypes (**e**, **f**), patients with intermediate-risk cytogenetics (**g**, **h**) and patients with normal cytogenetics (**i**, **j**). Number of patients and median survival time are indicated in each subgroup.

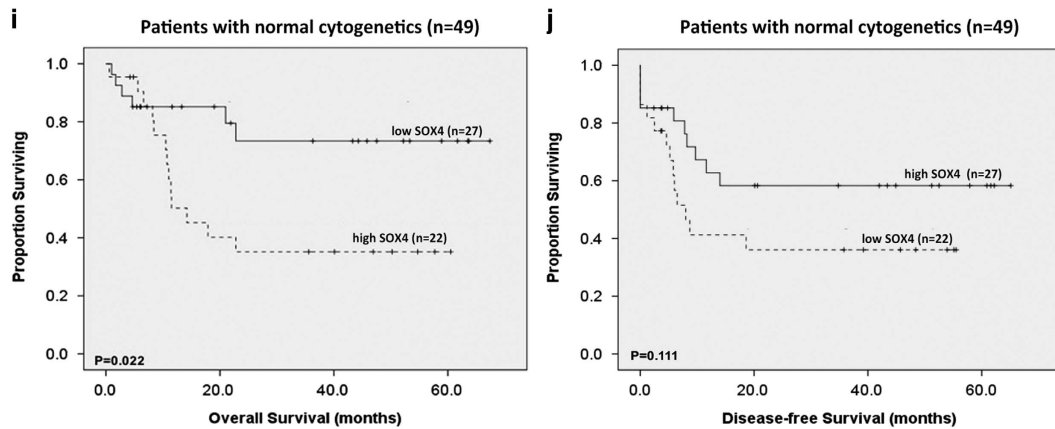


Figure 1. (continued).

**Table 1.** Multivariate analysis (Cox regression) of SOX4 expression on the overall survival and disease-free survival

Variables	Overall survival				Disease-free survival			
	RR	95% CI		P value	RR	95% CI		P value
		Lower	Upper			Lower	Upper	
Age <sup>a</sup>	1.023	1.002	1.044	0.034	1.012	0.994	1.029	0.203
WBC <sup>b</sup>	1.004	1.000	1.007	0.029	1.004	1.001	1.006	0.012
Karyotype <sup>c</sup>	3.038	1.188	7.771	0.020	2.127	1.105	4.096	0.024
<i>NPM1/FLT3-ITD</i> <sup>d</sup>	1.197	0.510	2.811	0.680	1.345	0.651	2.778	0.423
High SOX4 <sup>e</sup>	1.924	1.020	3.628	0.043	1.663	0.983	2.813	0.058

Abbreviations: CI, confidence interval; RR, relative risk; WBC, white blood cell. <sup>a</sup>The risk by 1 year older in age. <sup>b</sup>The risk by 1000/ $\mu$ l increase in WBC. <sup>c</sup>Unfavorable cytogenetics vs others. <sup>d</sup>*NPM1*<sup>mut</sup>/*FLT3-ITD*<sup>neg</sup> vs other subtypes. <sup>e</sup>High SOX4 expression vs low SOX4 expression.

wild-type AB strain zebrafish to generate transgenic zebrafish founder, namely Tg(*spi1:SOX4-EGFP*) (Figure 2a). Two transgenic *SOX4* lines (TG1 and TG2) were generated in parallel. F1 fish were generated by crossing *SOX4* founders and wild-type zebrafish, and *SOX4* transgene expression was demonstrated by semi-quantitative PCR (Figure 2b). The expression of *SOX4-EGFP* fusion proteins was verified using whole-mount immunohistochemical staining with anti-EGFP antibodies. In contrast to wild-type fish, which showed no enhanced green fluorescent protein (EGFP), the Tg(*spi1:SOX4-EGFP*) fish exhibited EGFP in the yolk sac and in caudal hematopoietic tissue after 48 h post fertilization (Figure 2c). We further investigated *SOX4* expression by applying immunohistochemistry staining anti-*SOX4* on tissue sections of 5-month-old Tg(*spi1:SOX4-EGFP*), and found that *SOX4* expression was restricted to a small fraction of cells in the KM (Figure 2d).

Tg(*spi1:SOX4-EGFP*) zebrafish embryos underwent normal hematopoiesis

To assess whether *SOX4* overexpression driven by *spi1* has an effect on early hematopoietic processes in transgenic zebrafish, several primitive and definitive markers including *cebpa*, *csf1r*, *gata1*, *l-plastin*, *mpo*, *mpeg1*, *spi1*, *c-myb* and *runx1* were examined by quantitative reverse transcription-PCR analysis. We found that there were no significant differences of these hematopoiesis-related transcription factors in transgenic zebrafish at both primitive and definitive stages (Figure 2e). Further whole-mount *in situ* hybridization experiments revealed that comparable numbers of *mpo*-positive cells were located in the aorta-gonad-mesonephros and caudal hematopoietic tissue regions in both *SOX4* and wild-type fish at 48 h post fertilization (Figures 2f and h).

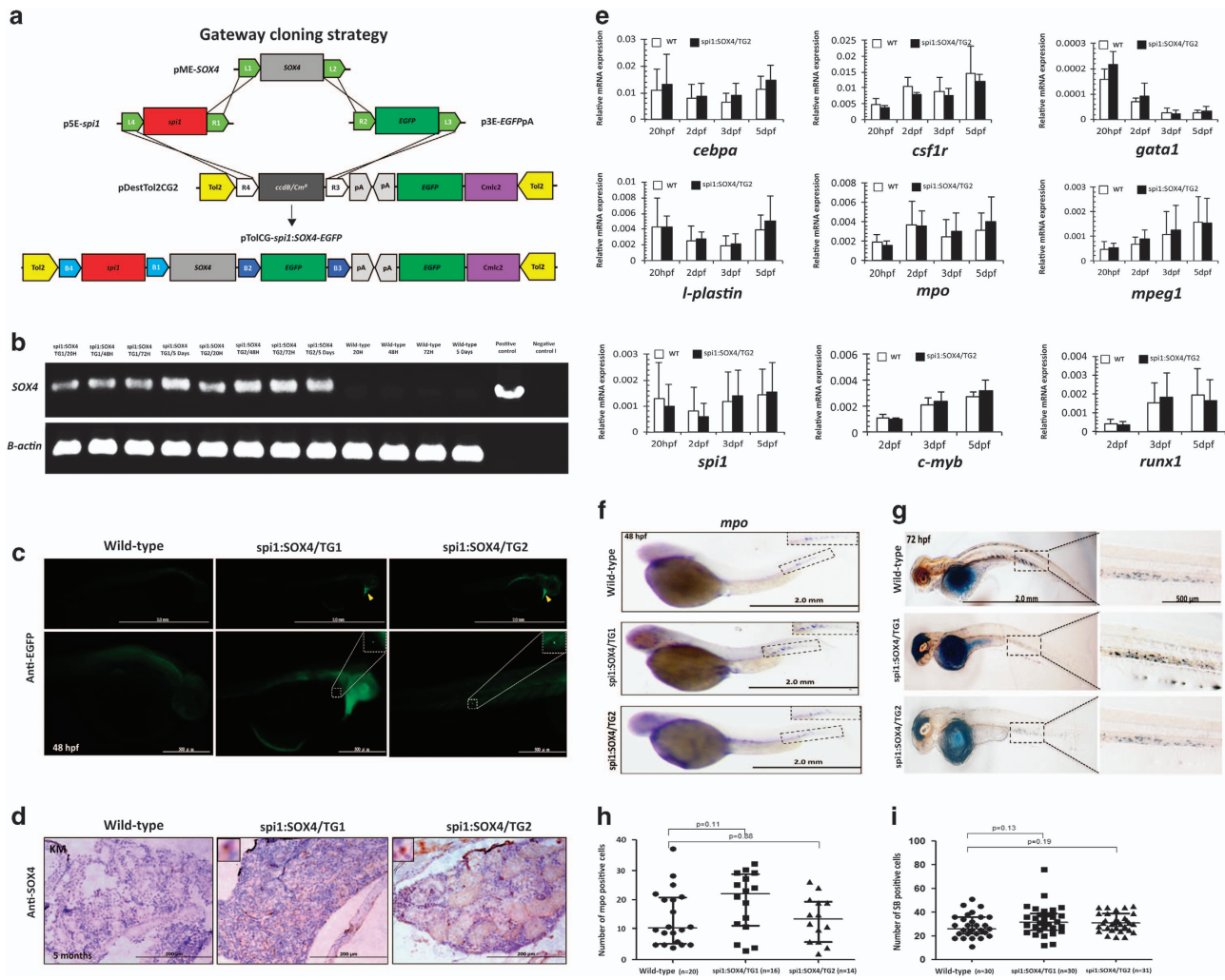
Sudan black staining revealed similar results for granulocyte colonized in caudal hematopoietic tissue at 72 h post fertilization (Figures 2g and i). These results indicate that *SOX4* transgenic zebrafish were undergoing normal hematopoietic processes during the larval stage.

Myeloid-specific expression of *SOX4* results in expanded myelopoiesis with poor differentiation in adult zebrafish

Myeloid malignancies are known to result from multiple long-term processes. Therefore, long-term examination of adult zebrafish was reasonable even though the overexpression of *SOX4* did not influence early hematopoiesis. In this study, biopsy sections of kidney tissues showed mild vacuoles in renal tubule structure in 5-month-old *SOX4* fish; moreover, kidneys showed mild or moderate effacement, a distorted structure and increased infiltration of myeloid cells in 9- and 12-month-old *SOX4* fish (Supplementary Figure 1).

Microscopic observation of KM smears from Tg(*spi1:SOX4-EGFP*) fish were analyzed at the ages of 5, 9, 12 and 15 months. The KM of 5-month-old *SOX4* zebrafish was comparable with the KM of age-matched 5-month-old AB wild-type fish; however, KM of *SOX4* transgenic fish demonstrated a greater number of myeloid progenitors and an excess of blast cells with focal aggregation at 9, 12 and 15 months, implying that myeloid transformation is age dependent (Figure 3a). These blast cells in KM were generally of medium to large size and were characterized by a mildly basophilic cytoplasm with scanty or fewer granule as well as a high nucleus-to-cytoplasm ratio (Figure 3a). In addition, a few typical blast cells in peripheral blood were observed in 9-month-old *SOX4* fish (Figure 3b).





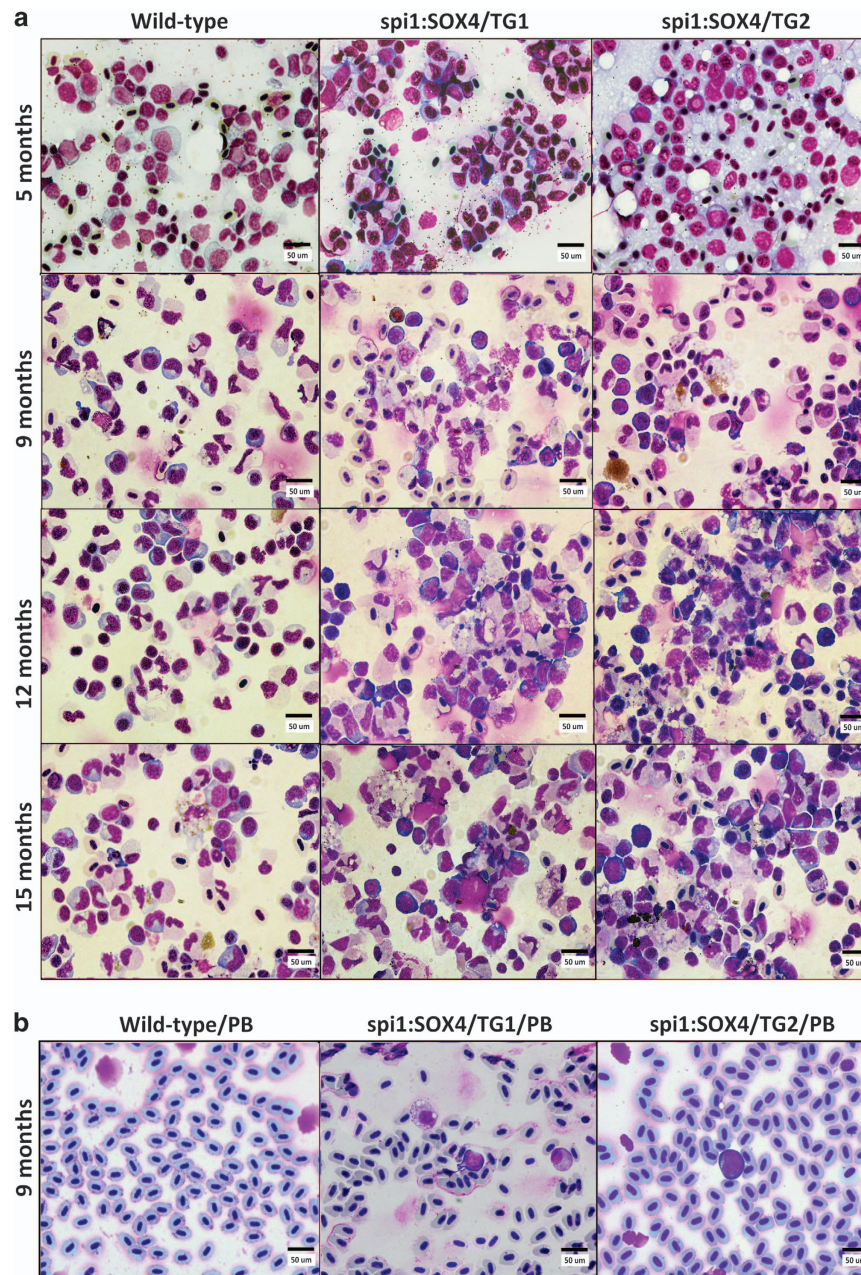
**Figure 2.** Generation and characterization of the Tg(*spi1:SOX4-EGFP*) transgenic zebrafish. **(a)** Schematic diagram showing the LR recombination reaction used to generate the expression constructs, including three entry clones (p5E-*spi1*, pME-SOX4 and p3E-EGFPpA) and a destination vector (pDestToi2CG2) that contains the *cm1c2:EGFP-pA* expression cassette. The final construct (pToiCG-*spi1:SOX4-EGFP*) is shown at the bottom of the figure. **(b)** Results of semiquantitative PCR showing the expression of SOX4 in TGs and wild-type embryos at 20, 48 and 72 h post fertilization (HPF) and at 5 days. Negative control: non-template; Positive control: pToiCG-*spi1:SOX4-EGFP* plasmid. **(c)** Whole-mount immunohistochemical staining (IHC) showed EGFP-positive cells in the heart (yellow arrow) as well as around the yolk sac and caudal hematopoietic tissue (CHT) at 48 HPF. The panel shows GFP+ cells at a higher magnification. **(d)** Tissue sections showing SOX4-positive cells in the KM of Tg(*spi1:SOX4-EGFP*) fish but not in those of wild-type fish. The panel shows SOX4-positive cells at a higher magnification. **(e)** Results from quantitative reverse transcription-PCR (RT-PCR) analysis of hematopoietic marker genes in SOX4 transgenic zebrafish embryos and in wild-type fish embryos. Data are presented as the mean  $\pm$  s.e.m. from three independent experiments. **(f, g)** Results from whole-mount *in situ* hybridization (WISH) of *mpo* and Sudan black (SB) staining showing *mpo*-positive cells in CHT at 48 HPF and SB-positive cells in CHT at 72 HPF, respectively (magnification:  $\times 40$ ). The panel shows positive cells at a higher magnification ( $\times 100$ ), respectively. Quantification of *mpo*-positive cells **(h)** and SB-positive cells **(i)**. Differences among variables were assessed using Student's t-test.

On microscopic examination of blood cells in KM smears, we found that compared with age-matched 9-month-old AB wild-type fish, the KM of 9-month-old SOX4 zebrafish showed a greater myeloid-to-erythroid ratio, a higher number of myeloid progenitors and decreased number of both immature and mature erythroid cells. Additional significant differences in myeloid-to-erythroid ratios and in all blood components of KM were found in 12- and 15-month-old transgenic SOX4 zebrafish (Table 2). Similar results were demonstrated using an additional flow cytometric analysis (Supplementary Figure 2). In addition, MPO staining further confirmed that there were significantly more myeloid cells in SOX4 fish compared with wild-type control fish of the same age (Figure 4). Taken together, our findings reveal that continuous overexpression of SOX4 led to an expansion of myeloid cells in the

KM of adult fish, and that the majority of accumulated myeloid cells were immature.

## DISCUSSION

To the best of our knowledge, this is the first study to report that increased BM SOX4 expression is an independent poor prognostic factor of overall survival in AML patients, irrespective of age, white blood cell count, karyotype profile and other genetic markers. The poor prognosis associated with high BM SOX4 expression was also demonstrated in CN-AML patients. These results are consistent with an online transcriptome data (GSE12417-GPL96 cohort) from AML patients (Supplementary Figure 3). Their data were analyzed by PrognScan, a new database in which meta-analyses to



**Figure 3.** Morphological analysis of various blood cell types harvested from the KM and peripheral blood of Tg(*spi1:SOX4-EGFP*) zebrafish at indicated age. Tg(*spi1:SOX4-EGFP*) KM smears showed myeloid hyperplasia with increased production of myeloblasts in an age-dependent manner (a) and their peripheral blood smears showed immature myeloid cells (b). Conversely, wild-type fish showed normal hematopoiesis with adequate maturation (magnification:  $\times 1000$ ).

elucidate the prognostic value of genes can be performed. It demonstrated that higher *SOX4* mRNA expression led to a worse overall survival in AML patients (hazard ratio 1.21, 95% confidence interval 1.01–1.47,  $P=0.043$ ). Zhang *et al.*<sup>26</sup> previously reported that *SOX4* is a direct target and crucial mediator of *C/EBPα* mutants in AML. It is therefore possible that poor prognosis is associated with *C/EBPα* mutants or *C/EBPα* methylation. However, in our previous report, we found that AML patients with *C/EBPα* double mutations<sup>38</sup> or higher *C/EBPα* methylation<sup>39</sup> had a superior response to chemotherapy treatment, and similar findings were reported in studies by other researchers as well.<sup>40–42</sup> The different prognoses associated with *SOX4* expression and *C/EBPα* gene status may be explained by the fact that *SOX4* expression is also associated with other genetic abnormalities including *MOZ-TIF2*,

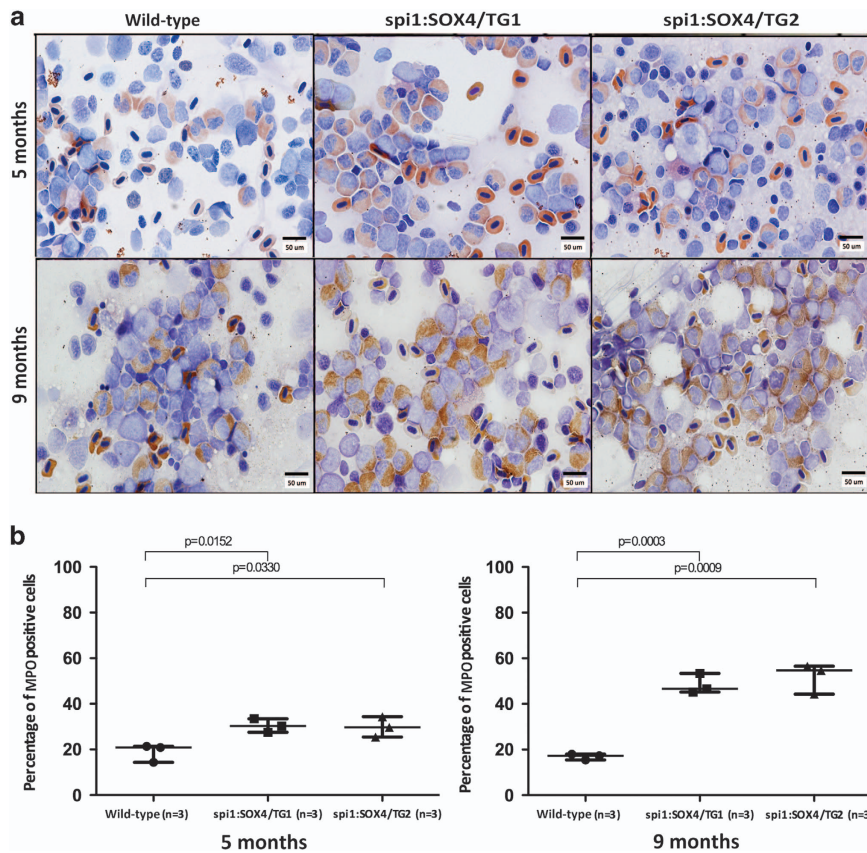
*AML1-ETO*, *NUP-HOXA9* and *FLT3-ITD*.<sup>26,43</sup> Of these abnormalities, *AML1-ETO* was found to be a better prognostic factor for AML,<sup>44,45</sup> and *FLT3-ITD* was found to be a poor prognostic factor.<sup>46</sup> Therefore, the conflicting results could result from the complexity, heterogeneity and multifactorial nature of AML. In addition, results from a cancer microarray database and integrated data (OncoPrint; <https://www.oncoPrint.org/resource/login.html>) showed that leukemia cells of AML patients ( $n=542$ ) expressed significantly more *SOX4* than did normal cells from controls ( $n=74$ ) (GSE13164 cohort; Supplementary Figure 4a). Significant differences between M1 and other FAB (French–American–British) categories (GSE14468 cohort) were also found (Supplementary Figure 4b). The mechanisms underlying the increase in *SOX4* expression independently associated with poor prognosis and



**Table 2.** Morphological analysis of blood cell types in the kidney marrow from *SOX4* transgenic fish compared with wild type

Total number of zebrafish <sup>a</sup> (n = 65)	Myeloid progenitors (blasts/blast-like; %)	Myelomonocyte/neutrophil (%)	Lymphocyte (%)	Immature erythroid (%)	Mature erythroid (%)	M/E ratio
5M-wild-type (n = 5)	5.27 ± 0.74	45.00 ± 5.36	15.33 ± 1.36	10.27 ± 1.82	24.13 ± 3.19	5.08 ± 1.57
9M-wild-type (n = 5)	5.73 ± 0.64	39.67 ± 5.01	14.39 ± 2.97	12.27 ± 0.83	27.93 ± 6.03	3.72:1 ± 0.53
12M-wild-type (n = 5)	7.00 ± 1.10	42.07 ± 3.18	16.20 ± 1.12	10.06 ± 2.55	24.67 ± 4.16	4.61 ± 0.64
15M-wild-type (n = 6)	7.11 ± 1.88	40.95 ± 2.33	15.95 ± 2.04	11.28 ± 1.04	25.44 ± 4.06	4.30 ± 0.47
5M- <i>SOX4</i> (n = 12)#	7.64 ± 2.56	37.53 ± 5.22	13.33 ± 1.65*	11.95 ± 3.02	29.08 ± 6.40	4.10 ± 1.59
9M- <i>SOX4</i> (n = 12)#	14.86 ± 7.07**	50.53 ± 10.33*	9.69 ± 1.65**	7.22 ± 3.46*	17.78 ± 5.19*	11.62 ± 7.25**
12M- <i>SOX4</i> (n = 11)#	16.24 ± 7.25**	58.33 ± 7.14**	6.09 ± 3.61**	6.21 ± 2.16*	13.12 ± 4.38**	13.55 ± 5.45**
15M- <i>SOX4</i> (n = 9)#	14.19 ± 5.16**	50.74 ± 9.15*	9.59 ± 1.38***	7.15 ± 3.29*	16.74 ± 4.42**	10.48 ± 4.08**

Abbreviations: M, month; M/E ratio, myeloid-to-immature erythroid ratio. Significant differences between the *SOX4* transgenic fish and wild-type are indicated (\**P* < 0.05; \*\**P* < 0.01; \*\*\**P* < 0.001) by Mann–Whitney *U*-test. Significant difference of groups between the *SOX4* transgenic fish and wild-type are indicated (#*P* < 0.001) by one-way analysis of variance (ANOVA). <sup>a</sup>A total number of 300 cells per zebrafish was counted to classify the distribution and subtypes of these hematopoietic cells.



**Figure 4.** Myeloperoxidase (MPO) staining analysis of *spi1:SOX4-EGFP* zebrafish. **(a)** Increased myeloperoxidase was detected in *SOX4* transgenic fish at 5 and 9 months (magnification: × 1000). **(b)** Quantification of MPO-positive cells. Differences among variables were assessed using Student's *t*-test.

various genetic aberrations still require further study. Nonetheless, complementing chemotherapy with agents that target *SOX4* could be beneficial treatment strategy for AML patients.

Zebrafish *sox4* proteins (*sox4a*: accession number. NP\_957195.1 and *sox4b*: accession number NP\_998287.1) share a high level of amino acid similarity with human *SOX4* (accession number: NP\_003098.1) and mouse *Sox4* (accession number: NP\_033264.2), particularly within the HMG box (*sox4a*: 100%, *sox4b*: 97.4%) and the TAD/DD (*sox4a*: 91.2%, *sox4b*: 70.6%) functional domains (Supplementary Figure 4a). Phylogenetic analysis revealed that the zebrafish *sox4a/b* is homologous with *SOX4* proteins from other

species (Supplementary Figure 5b). The high conservation of *SOX4* among humans, mice and zebrafish suggests that *SOX4* proteins could share similar functions. We also characterized the mRNA expression of *sox4a* and *sox4b* in early developmental stages of transgenic zebrafish. In so doing, we found that the expression levels of *sox4b* and *sox4a* in transgenic zebrafish were comparable to those of wild-type fish (data not shown).

The kidneys of Tg(*spi1:SOX4-EGFP*) zebrafish began showing mild vacuoles in renal tubule structure at the age of the 5 months. At 9 and 12 months, the kidneys further showed mild or moderate effacement, a distorted structure and increased infiltration of

myeloid cells. Microscopic examination revealed that the KM blood cells of 5-month-old Tg(*spi1:SOX4-EGFP*) zebrafish were comparable with those of age-matched AB wild-type fish. However, at 9, 12 and 15 months, KM cells from Tg(*spi1:SOX4-EGFP*) fish had a greater number of myeloid progenitors and an excess of blast cells with focal aggregation than did KM cells of age-matched AB wild-type fish, implying that myeloid transformation was age dependent. This is the first study to report on transgenic *SOX4* zebrafish with myeloid transformation phenotype.

Results from previous research in mouse model studies support our findings. For example, in one study, 73% of mice developed myeloid leukemia after receiving a transplant of *MSCV-Sox4*-infected marrow, whereas none of the control mice developed leukemia.<sup>25</sup> Other reports have also found that *Sox4* leads to myeloid malignancies in mouse model as follows. *Sox4* may upregulate *CREB* expression and cooperate with *CREB* to activate downstream *cyclin D1*, *bcl-2* and *c-fos* that in turn increases cell proliferation.<sup>21</sup> *Sox4* may also downregulate *Pu.1*, leading to *Pu.1* haploinsufficiency, which accelerates and increases the penetrance of *Sox4*-induced leukemia.<sup>23</sup> A later study revealed that a novel PAD4/SOX4/PU.1 signaling pathway is involved in the differentiation of leukemic cells into granulocytes.<sup>47</sup> Furthermore, in an AKXD23 mouse model of myeloid leukemia, proviral insertions to *Sox4* and *Evi1* were found to cause increases in *Sox4* and *Evi1* expressions, respectively; and *Sox4* acted alone or cooperated with *Evi1* to induce transcriptional activation and subsequent cell proliferation.<sup>22</sup> Findings from that research also demonstrated that overexpression of *Sox4* in 32Dcl3 cells markedly inhibited cytokine-induced granulocyte maturation and associated cell proliferation.<sup>22</sup> Finally, Fung et al.<sup>27</sup> suggested that activation of the Wnt pathway, activation of the TGF-FOXO pathway and epigenetic aberration of *Sox4* could be responsible for myeloid leukemogenesis.

In conclusion, we reveal that BM *SOX4* expression could serve as an informative new biomarker for the clinical prognosis of AML patients. Moreover, we demonstrate that in zebrafish, the myeloid-specific expression of *SOX4* can induce leukemic phenotype. Therefore, we suggest that treating AML patients with agents that target *SOX4* or its downstream molecules in addition to chemotherapy could be an effective therapeutic strategy. Furthermore, this zebrafish disease model could be useful for studying gene functions involved in myeloid malignancies and to be a valuable *in vivo* platform for screening antileukemic drugs.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

This study was supported by the research grants from Taiwan Health Foundation (to L-IL), the National Taiwan University Hospital (NTUH1055-3065 to C-YC), the Ministry of Science and Technology (MOST-102-2628-B-002-029-MY3 and MOST-105-2320-B-002-051-MY3 to L-IL; MOST-103-2628-B-002-008-MY3 and MOST-104-2314-B-002-128-MY4 to H-AH) and the Ministry of Health and Welfare (MOHW105-TDU-B-211-134005 to H-FT), Taiwan. We acknowledge the services provided by Flow Cytometric Analysing and Sorting Core Facility at NTUH; and we also Taiwan Zebrafish Core Facility (TZeCF) at NHRI for providing fish lines and resources; TZeCF is supported by a grant (MOST-104-2321-B-001-045) from the Ministry of Science and Technology, Taiwan.

## AUTHOR CONTRIBUTIONS

J-WL and L-IL designed and coordinated the study; J-WL and M-SH were responsible for establishing the transgenic fish and designing individual zebrafish experiments; J-WL, M-SH and H-AH were responsible for flow cytometric and cytomorphologic analysis of zebrafish KM and PB; H-AH performed immunocytochemical staining of human bone marrow; H-AH, C-YC and H-FT contributed patient samples and clinical data; and J-WL, M-SH, H-AH and L-IL wrote the manuscript.

## REFERENCES

- Harley V, Lefebvre V. Twenty Sox, twenty years. *Int J Biochem Cell Biol* 2010; **42**: 376–377.
- Jafarnejad SM, Ardekani GS, Ghaffari M, Li G. Pleiotropic function of SRY-related HMG box transcription factor 4 in regulation of tumorigenesis. *Cell Mol Life Sci* 2013; **70**: 2677–2696.
- Jiang Y, Ding Q, Xie X, Libby RT, Lefebvre V, Gan L. Transcription factors SOX4 and SOX11 function redundantly to regulate the development of mouse retinal ganglion cells. *J Biol Chem* 2013; **288**: 18429–18438.
- Vervoort SJ, van Bostel R, Coffey PJ. The role of SRY-related HMG box transcription factor 4 (SOX4) in tumorigenesis and metastasis: friend or foe? *Oncogene* 2013; **32**: 3397–3409.
- Schilham MW, Moerer P, Cumano A, Clevers HC. Sox-4 facilitates thymocyte differentiation. *Eur J Immunol* 1997; **27**: 1292–1295.
- Graham JD, Hunt SM, Tran N, Clarke CL. Regulation of the expression and activity by progestins of a member of the SOX gene family of transcriptional modulators. *J Mol Endocrinol* 1999; **22**: 295–304.
- Aaboe M, Birkenkamp-Demtroder K, Wiuf C, Sorensen FB, Thykjaer T, Sauter G et al. SOX4 expression in bladder carcinoma: clinical aspects and in vitro functional characterization. *Cancer Res* 2006; **66**: 3434–3442.
- Medina PP, Castillo SD, Blanco S, Sanz-Garcia M, Largo C, Alvarez S et al. The SRY-HMG box gene, SOX4, is a target of gene amplification at chromosome 6p in lung cancer. *Hum Mol Genet* 2009; **18**: 1343–1352.
- Wang C, Zhao H, Lu J, Yin J, Zang L, Song N et al. Clinicopathological significance of SOX4 expression in primary gallbladder carcinoma. *Diagn Pathol* 2012; **7**: 41.
- Lin CM, Fang CL, Hseu YC, Chen CL, Wang JW, Hsu SL et al. Clinical and prognostic implications of transcription factor SOX4 in patients with colon cancer. *PLoS One* 2013; **8**: e67128.
- Wang L, Zhang J, Yang X, Chang YW, Qi M, Zhou Z et al. SOX4 is associated with poor prognosis in prostate cancer and promotes epithelial-mesenchymal transition in vitro. *Prostate Cancer Prostatic Dis* 2013; **16**: 301–307.
- Wang W, Zhang J, Zhan X, Lin T, Yang M, Hu J et al. SOX4 is associated with poor prognosis in cholangiocarcinoma. *Biochem Biophys Res Commun* 2014; **452**: 614–621.
- Li L, Li Q, Chen X, Xu M, Li X, Nie L et al. SOX4 is overexpressed in diffusely infiltrating astrocytoma and confers poor prognosis. *Neuropathology* 2015; **35**: 510–517.
- Lu N, Lin T, Wang L, Qi M, Liu Z, Dong H et al. Association of SOX4 regulated by tumor suppressor miR-30a with poor prognosis in low-grade chondrosarcoma. *Tumour Biol* 2015; **36**: 3843–3852.
- Song GD, Sun Y, Shen H, Li W. SOX4 overexpression is a novel biomarker of malignant status and poor prognosis in breast cancer patients. *Tumour Biol* 2015; **36**: 4167–4173.
- Wang D, Hao T, Pan Y, Qian X, Zhou D. Increased expression of SOX4 is a biomarker for malignant status and poor prognosis in patients with non-small cell lung cancer. *Mol Cell Biochem* 2015; **402**: 75–82.
- Tonks A, Pearn L, Musson M, Gilkes A, Mills KI, Burnett AK et al. Transcriptional dysregulation mediated by RUNX1-RUNX1T1 in normal human progenitor cells and in acute myeloid leukaemia. *Leukemia* 2007; **21**: 2495–2505.
- Yassin ER, Abdul-Nabi AM, Takeda A, Yaseen NR. Effects of the NUP98-DDX10 oncogene on primary human CD34+ cells: role of a conserved helicase motif. *Leukemia* 2010; **24**: 1001–1011.
- Omidvar N, Maunakea ML, Jones L, Sevcikova S, Yin B, Himmel KL et al. PML-RARalpha co-operates with Sox4 in acute myeloid leukemia development in mice. *Haematologica* 2013; **98**: 424–427.
- Huang Y, Sitwala K, Bronstein J, Sanders D, Dandekar M, Collins C et al. Identification and characterization of Hoxa9 binding sites in hematopoietic cells. *Blood* 2012; **119**: 388–398.
- Sandoval S, Kraus C, Cho EC, Cho M, Bies J, Manara E et al. Sox4 cooperates with CREB in myeloid transformation. *Blood* 2012; **120**: 155–165.
- Boyd KE, Xiao YY, Fan K, Poholek A, Copeland NG, Jenkins NA et al. Sox4 cooperates with Evi1 in AKXD-23 myeloid tumors via transactivation of proviral LTR. *Blood* 2006; **107**: 733–741.
- Aue G, Du Y, Cleveland SM, Smith SB, Dave UP, Liu D et al. Sox4 cooperates with PU.1 haploinsufficiency in murine myeloid leukemia. *Blood* 2011; **118**: 4674–4681.
- Schilham MW, Oosterwegel MA, Moerer P, Ya J, de Boer PA, van de Wetering M et al. Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking Sox-4. *Nature* 1996; **380**: 711–714.
- Du Y, Spence SE, Jenkins NA, Copeland NG. Cooperating cancer-gene identification through oncogenic-retrovirus-induced insertional mutagenesis. *Blood* 2005; **106**: 2498–2505.
- Zhang H, Alberich-Jorda M, Amabile G, Yang H, Staber PB, Di Ruscio A et al. Sox4 is a key oncogenic target in C/EBPalpha mutant acute myeloid leukemia. *Cancer Cell* 2013; **24**: 575–588.



- 27 Fung TK, Leung AY, So CW. Sox4you: a new player in C/EBPalpha leukemia. *Cancer Cell* 2013; **24**: 557–559.
- 28 Tregnago C, Manara E, Zampini M, Bisio V, Borga C, Bresolin S *et al*. CREB engages C/EBPdelta to initiate leukemogenesis. *Leukemia* 2016; **30**: 1887–1896.
- 29 Lu JW, Hou HA, Hsieh MS, Tien HF, Lin LI. Overexpression of FLT3-ITD driven by spi-1 results in expanded myelopoiesis with leukemic phenotype in zebrafish. *Leukemia* 2016; **30**: 2098–2101.
- 30 Liu W, Wu M, Huang Z, Lian J, Chen J, Wang T *et al*. c-myc hyperactivity leads to myeloid and lymphoid malignancies in zebrafish. *Leukemia* 2017; **31**: 222–233.
- 31 Lu JW, Hsieh MS, Liao HA, Yang YJ, Ho YJ, Lin LI. Zebrafish as a model for the study of human myeloid malignancies. *Biomed Res Int* 2015; **2015**: 641475.
- 32 He BL, Shi X, Man CH, Ma AC, Ekker SC, Chow HC *et al*. Functions of flt3 in zebrafish hematopoiesis and its relevance to human acute myeloid leukemia. *Blood* 2014; **123**: 2518–2529.
- 33 Hou HA, Chou WC, Lin LI, Tang JL, Tseng MH, Huang CF *et al*. Expression of angiopoietins and vascular endothelial growth factors and their clinical significance in acute myeloid leukemia. *Leuk Res* 2008; **32**: 904–912.
- 34 Kwan KM, Fujimoto E, Grabher C, Mangum BD, Hardy ME, Campbell DS *et al*. The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev Dyn* 2007; **236**: 3088–3099.
- 35 Le Guyader D, Redd MJ, Colucci-Guyon E, Murayama E, Kissa K, Briolat V *et al*. Origins and unconventional behavior of neutrophils in developing zebrafish. *Blood* 2008; **111**: 132–141.
- 36 Hou HA, Lin CC, Chou WC, Liu CY, Chen CY, Tang JL *et al*. Integration of cytogenetic and molecular alterations in risk stratification of 318 patients with de novo non-M3 acute myeloid leukemia. *Leukemia* 2014; **28**: 50–58.
- 37 Tsai CH, Hou HA, Tang JL, Liu CY, Lin CC, Chou WC *et al*. Genetic alterations and their clinical implications in older patients with acute myeloid leukemia. *Leukemia* 2016; **30**: 1485–1492.
- 38 Hou HA, Lin LI, Chen CY, Tien HF. Reply to 'Heterogeneity within AML with CEBPA mutations; only CEBPA double mutations, but not single CEBPA mutations are associated with favorable prognosis'. *Br J Cancer* 2009; **101**: 738–740.
- 39 Lin TC, Hou HA, Chou WC, Ou DL, Yu SL, Tien HF *et al*. CEBPA methylation as a prognostic biomarker in patients with de novo acute myeloid leukemia. *Leukemia* 2011; **25**: 32–40.
- 40 Poiana C, Neamtu MC, Avramescu ET, Carsote M, Trifanescu R, Terzea D *et al*. The poor prognosis factors in G2 neuroendocrine tumor. *Rom J Morphol Embryol* 2013; **54**(3 Suppl): 717–720.
- 41 Preudhomme C, Sagot C, Boissel N, Cayuela JM, Tigaud I, de Botton S *et al*. Favorable prognostic significance of CEBPA mutations in patients with de novo acute myeloid leukemia: a study from the Acute Leukemia French Association (ALFA). *Blood* 2002; **100**: 2717–2723.
- 42 Wouters BJ, Lowenberg B, Erpelinck-Verschueren CA, van Putten WL, Valk PJ, Delwel R. Double CEBPA mutations, but not single CEBPA mutations, define a subgroup of acute myeloid leukemia with a distinctive gene expression profile that is uniquely associated with a favorable outcome. *Blood* 2009; **113**: 3088–3091.
- 43 Kvinlaug BT, Chan WI, Bullinger L, Ramaswami M, Sears C, Foster D *et al*. Common and overlapping oncogenic pathways contribute to the evolution of acute myeloid leukemias. *Cancer Res* 2011; **71**: 4117–4129.
- 44 Bloomfield CD, Lawrence D, Byrd JC, Carroll A, Pettenati MJ, Tantravahi R *et al*. Frequency of prolonged remission duration after high-dose cytarabine intensification in acute myeloid leukemia varies by cytogenetic subtype. *Cancer Res* 1998; **58**: 4173–4179.
- 45 Grimwade D, Walker H, Oliver F, Wheatley K, Harrison C, Harrison G *et al*. The importance of diagnostic cytogenetics on outcome in AML: analysis of 1612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children's Leukaemia Working Parties. *Blood* 1998; **92**: 2322–2333.
- 46 Kiyoi H, Naoe T, Nakano Y, Yokota S, Minami S, Miyawaki S *et al*. Prognostic implication of FLT3 and N-RAS gene mutations in acute myeloid leukemia. *Blood* 1999; **93**: 3074–3080.
- 47 Song G, Shi L, Guo Y, Yu L, Wang L, Zhang X *et al*. A novel PAD4/SOX4/PU.1 signaling pathway is involved in the committed differentiation of acute promyelocytic leukemia cells into granulocytic cells. *Oncotarget* 2016; **7**: 3144–3157.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017

Supplementary Information accompanies this paper on Blood Cancer Journal website (<http://www.nature.com/bcj>)