Supplemental Methods

Flow cytometry and cell sorting

For sorting of the High, Low and Negative populations of *Gata2*$^{+/gfp}$::Tom and 3q21q26-*EVI1*::*Gata2*$^{+/gfp}$::Tom mice (Figure 1), bone marrow mononuclear cells were stained with PI to remove dead cells. Gating strategy for identification of High, Low and Negative populations is shown in Supplemental Figure 9.

To characterize the High, Low and Negative populations of *Gata2*$^{+/gfp}$::Tom and 3q21q26-*EVI1*::*Gata2*$^{+/gfp}$::Tom mice (Figure 2), the bone marrow cells were stained with following sets of antibodies. For analyses of Mk lineage (Figure 2A), cells were stained with allophycocyanin (APC)-conjugated anti-CD41 and Brilliant Violet (BV) 421-conjugated anti-CD61 antibodies. For analyses of myeloid lineage (Figure 2B), cells were stained with allophycocyanin-cyanin-7 (APC-Cy7)-conjugated anti-B220, APC-Cy7-conjugated anti-CD3, BV510-conjugated anti-Ly6G and APC-conjugated anti-CD115 antibodies. For analyses of B lineage (Figure 2C), cells were stained with phycoerythrin-cyanin-7 (PE-Cy7)-conjugated anti-B220 and BV510-conjugated anti-CD19 antibodies. For analyses of erythroid lineage (Figure 2D), cells were stained with APC-conjugated anti-Ter119 and BV421-conjugated anti-CD71 antibodies. For analyses of progenitors (Figure 2E), cells were stained with PE-Cy7-conjugated anti-Sca1 and APC-eFluor780-conjugated anti-c-Kit antibodies. Gating strategy for these analyses is shown in Supplemental Figure 9. We showed isotype controls of representative flow cytometry in Supplemental Figure 10.

For analyses of LK and LSK of *Gata2*$^{+/gfp}$::Tom and 3q21q26-*EVI1*::*Gata2*$^{+/gfp}$::Tom mice (Supplemental Figure 5), mononuclear cells with positive expression of lineage markers (Ter119, CD4, CD8, CD3e, Mac1, Gr1, CD19) were removed (gated out) by staining with biotin-conjugate antibodies and peridinin chlorophyll protein (PerCP)-conjugated streptavidin. The biotin-conjugated antibodies were used as follows: anti-Ter119, anti-CD4, anti-CD8, anti-CD3e, anti-Mac1, anti-Gr1 and anti-CD19 antibodies. Cells negative for lineage markers were stained with APC-eFluor780-conjugated anti-c-Kit, PE-Cy7-conjugated anti-Sca1, BV421-conjugated anti-Flt3, APC-conjugated anti-CD150 and BV510-conjugated anti-CD48 antibodies.

To characterize LSK fraction of the WT, *Gata2*$^{+/gfp}$, 3q21q26-*EVI1* and 3q21q26-*EVI1*::*Gata2*$^{+/gfp}$ mice (Figure 3), the bone marrow mononuclear cells with positive expression of lineage markers (Ter119, CD4, CD8, CD3e, Mac1, Gr1, CD19) were removed by staining with biotin-conjugate antibodies and PerCP-conjugated streptavidin. Cells negative for lineage markers were stained with APC-eFluor780-conjugated anti-c-Kit, PE-Cy7-conjugated anti-Sca1, BV421-conjugated anti-Flt3, APC-conjugated anti-CD150 and phycoerythrin (PE)-conjugated anti-CD48 antibodies.

To characterize bone marrow cells of the WT, *Gata2*$^{+/gfp}$, 3q21q26-*EVI1* and 3q21q26-*EVI1*::*Gata2*$^{+/gfp}$ mice (Figure 4), the cells were stained with following sets of antibodies. For analyses of MEP (Figure 4A), mononuclear cells with positive expression of lineage markers (Ter119, CD4, CD8, CD3e, Mac1, Gr1, CD19) were removed by staining with biotin-conjugated antibodies and PerCP-conjugated streptavidin. Cells negative for lineage markers were stained with APC-eFluor780-conjugated anti-c-Kit, PE-Cy7-conjugated anti-Sca1, BV421-conjugated anti-CD34, BV510-conjugated anti-FcγR II/III, PE-conjugated anti-Flt3 antibodies. For analyses of Mk lineage (Figure 4B and 4C), cells were stained with PE-conjugated anti-CD41, BV421-conjugated anti-CD61 and APC-eFluor780-conjugated anti-c-Kit antibodies. For analyses of erythroid lineage (Figure 4E), cells were stained with APC-conjugated anti-Ter119 and BV421-conjugated anti-CD71 antibodies. For analyses of myeloid lineage (Figure 4G and 4H), cells were stained with APC-Cy7-conjugated anti-B220, APC-Cy7-conjugated anti-CD3, BV510-conjugated anti-Ly6G and APC-conjugated anti-CD115 antibodies. For analyses of B lineage (Figure 4I), cells were stained with PE-Cy7-conjugated anti-B220 and BV510-conjugated anti-CD19 antibodies. For analyses of pDC (Figure 4J), cells were stained with APC-Cy7-conjugated anti-B220, PE-Cy7-conjugated anti-CD11c, BV510-conjugated anti-Gr1, biotin-conjugated anti-CD19 antibodies and BV421-conjugated streptavidin.

Supplemental Table 1

Primer sequences for BAC recombination

Primers	Sequences
TomatoFrtAmp(f)	TATCGATACCGTCGAAGGGGCGGATCCTGAAGTT
TomatoFrtAmp(r)	CCCCCTCGAGGTCGAGCTCAAGCTTCGAA
3q21q26tdTomato(f)	GAGTAGTTGCAGACATTGCGCCGGGAGAGGAGCTTCTGCTGTTC ATGGTGAGCAAGGGCGAGG
3q21q26tdTomato(r)	GCGATGATAAGGTGATAAGGAGGGTGGCGTGAGTGGTACTAAC GCTCAAGCTTCGAATTC

Supplemental Table 2

Antibodies for flow cytometric analysis

Epitope	Conjugate	Clone	Company
B220	PE-Cy7	RA3-6B2	TONBO
B220	APC-Cy7	RA3-6B2	BD
CD19	Biotin	1D3	TONBO
CD19	BV510	1D3	BD
Gr1	BV510	RB6-8C5	BioLegend
Gr1	Biotin	RB6-8C5	eBioscience
Ly6G	BV510	1A8	BD
CD115	APC	AFS98	BioLegend
CD41	PE	MWReg30	BD
CD41	APC	MWReg30	BioLegend
CD61	BV421	2C9.G2	BD
CD71	BV421	RI7217	BioLegend
Ter119	APC	TER-119	BD, eBioscience
Ter119	Biotin	TER-119	eBioscience
c-Kit	APC-eFluor 780	2B8	eBioscience
Sca1	PE-Cy7	D7	BD
Flt3	BV421	A2F10	BioLegend

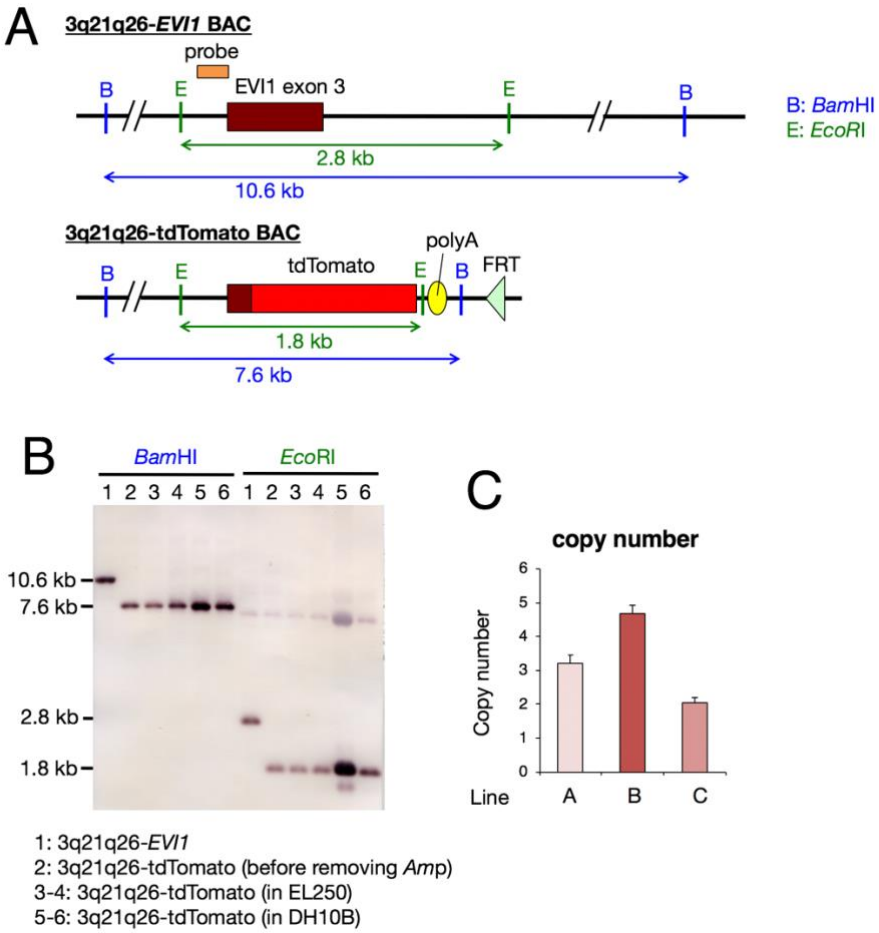
Flt3	PE	A2F10	BioLegend
CD150	APC	TC15-12F12.2	BioLegend
CD48	PE	HM48-1	BioLegend
CD48	BV510	HM48-1	BD
CD3e	Biotin	145-2C11	eBioscience
CD3	APC-Cy7	17A2	BioLegend
CD4	Biotin	GK1.5	BioLegend
CD8	Biotin	53-6.7	BioLegend
Mac1	Biotin	M1/70	eBioscience
CD34	BV421	RAM34	BD
FcγR II/III	BV510	2.4G2	BD
CD11c	PE-Cy7	N418	BioLegend
CD45.1	BV421	A20	BD
CD45.2	APC-Cy7	104	TONBO
streptavidin	PerCP	−	BioLegend
streptavidin	BV421	−	BioLegend

Supplemental Table 3

Primer sequences for q-RT-PCR using SYBR green

Gene	Forward primer	Reverse primer
EVI1	ACAAGCCAAGACCAGCCCCTGG	TGGTTTTTTCGAGGCTCAGTC
Gata2	CCTGGTTCCCAAGACACAGTAG	GGGTGCTGCGCATTCAATAC
Gapdh	TGCACCACCAACTGCTTAG	GGATGCAGGGATGATGTTC

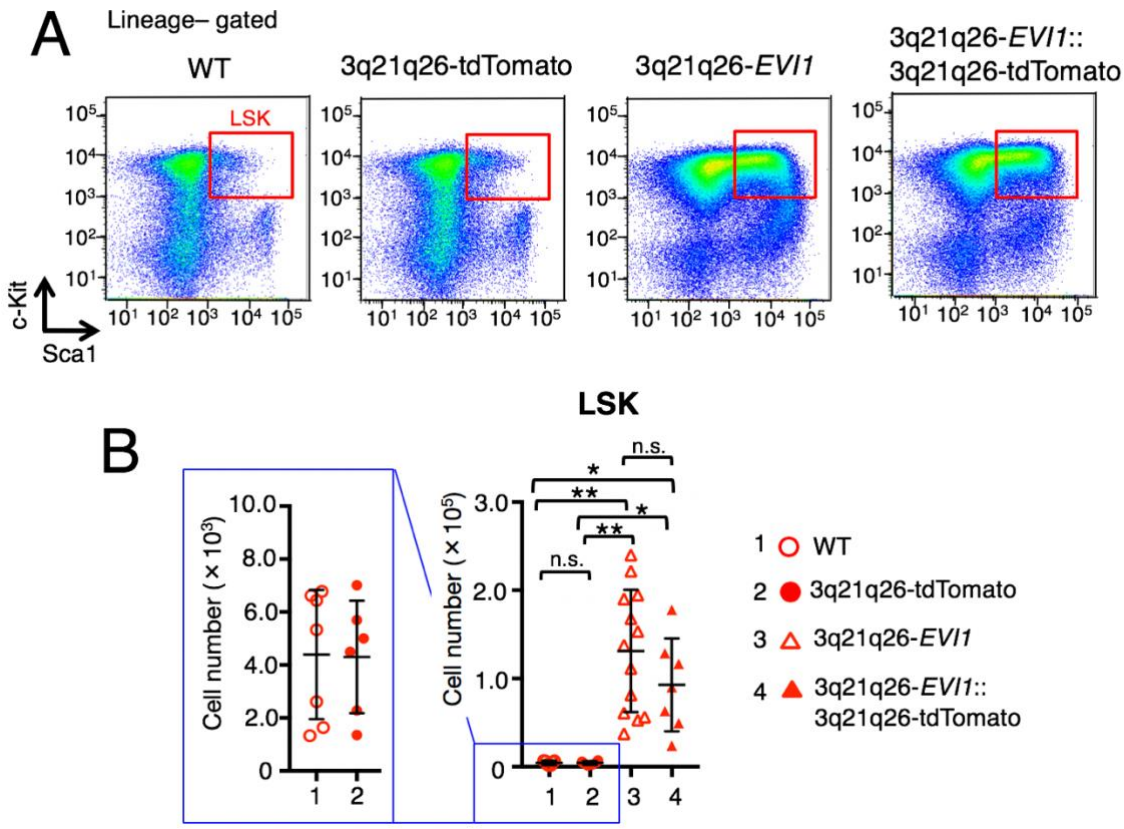
Supplemental Figure 1



Supplemental Figure 1. Validation of 3q21q26-tdTomato BAC construct.

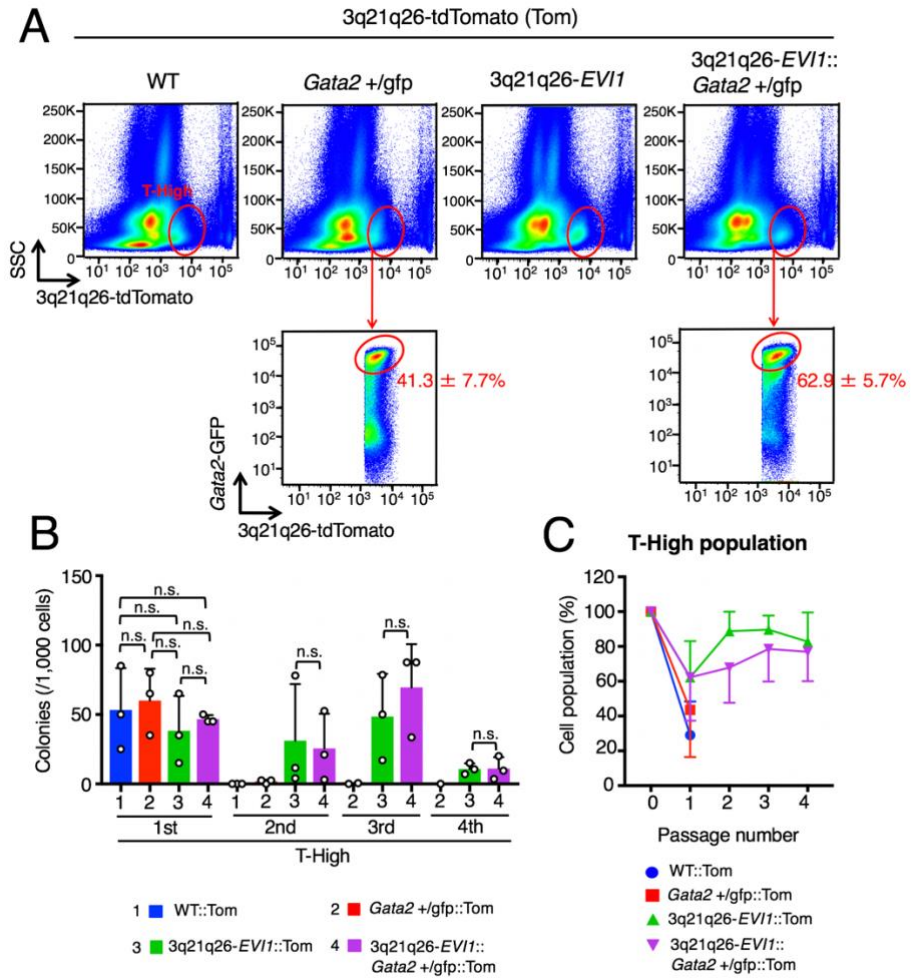
(A) Designs of validations of 3q21q26-tdTomato BAC construct. tdTomato gene was inserted into a translation start site in *EVI1* 3rd exon. Conformation of region around *EVI1* 3rd exon was confirmed in 3q21q26-tdTomato BAC clone by Southern blotting. The positions of probes (an orange box) and restriction enzyme sites (B: *Bam*HI, E: *Eco*RI) are shown. Green and blue double-headed arrows represent fragments of the expected sizes of *Eco*RI and *Bam*HI, respectively. (B) Southern blot analysis of the 3q21q26-tdTomato BAC clone. 3q21q26-EVI1 (lane 1), 3q21q26-tdTomato before removing *Amp* (lane 2), 3q21q26-tdTomato after removing *Amp* in EL250 strain (lanes 3-4) and 3q21q26-tdTomato after removing *Amp* in DH10B strain (lanes 5-6) were digested and the fragments were detected using the probe. (C) Copy numbers of the 3q21q26-tdTomato transgenic mice of lines A, B and C.

Supplemental Figure 2



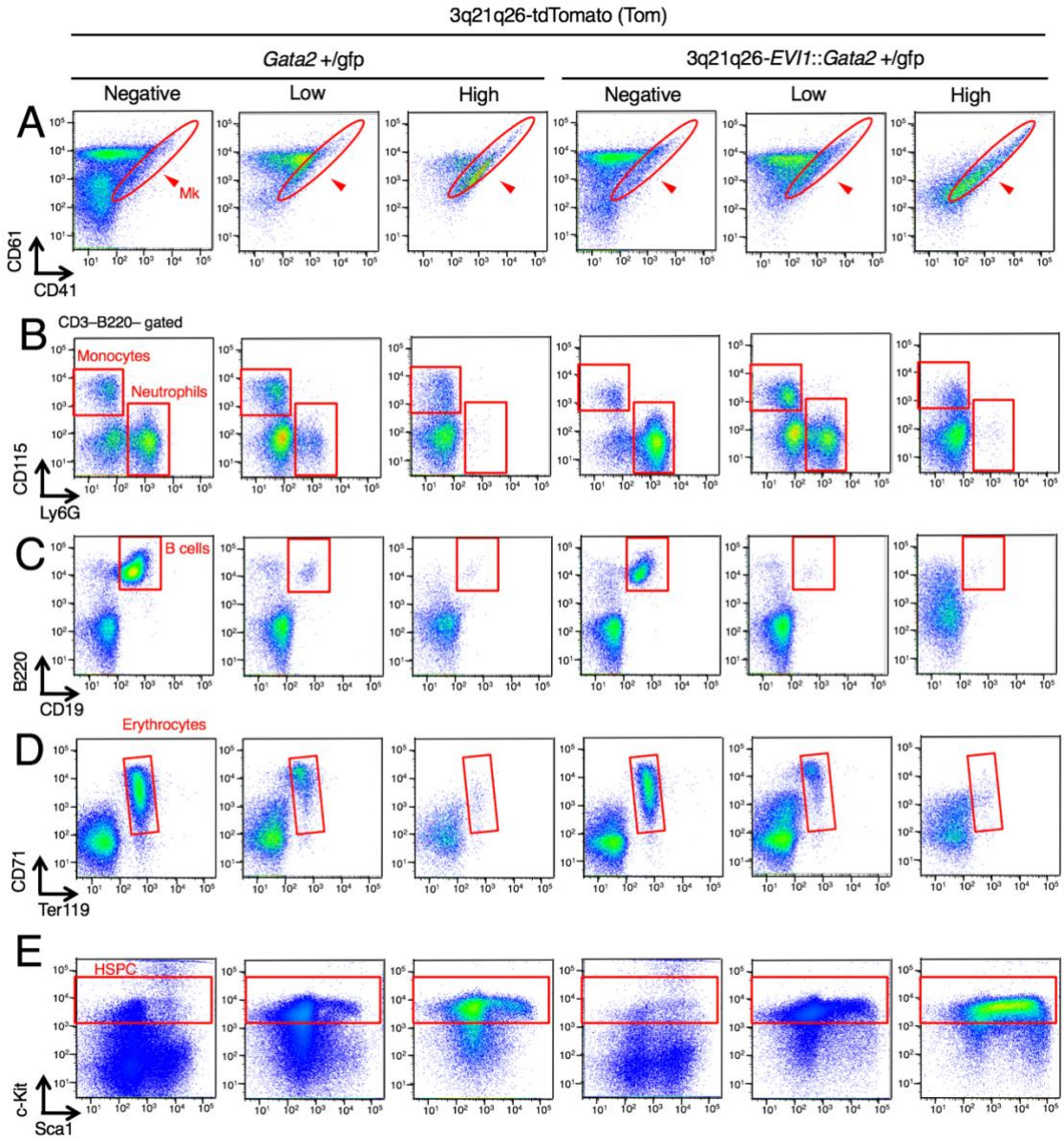
Supplemental Figure 2. 3q21q26-tdTomato transgene does not affect LSK fractions. (A) Representative flow cytometric profiles of lineage-negative fractions from WT, 3q21q26-tdTomato, 3q21q26-*EVI1* and 3q21q26-*EVI1*::3q21q26-tdTomato mouse bone marrows. (B) Absolute numbers of LSK cells in the bone marrows from WT, 3q21q26-tdTomato, 3q21q26-*EVI1* and 3q21q26-*EVI1*::3q21q26-tdTomato mouse bone marrows. Values represent the means ± SD. *, $P<0.05$; **, $P<0.01$; n.s., not significant (one-way ANOVA).

Supplemental Figure 3



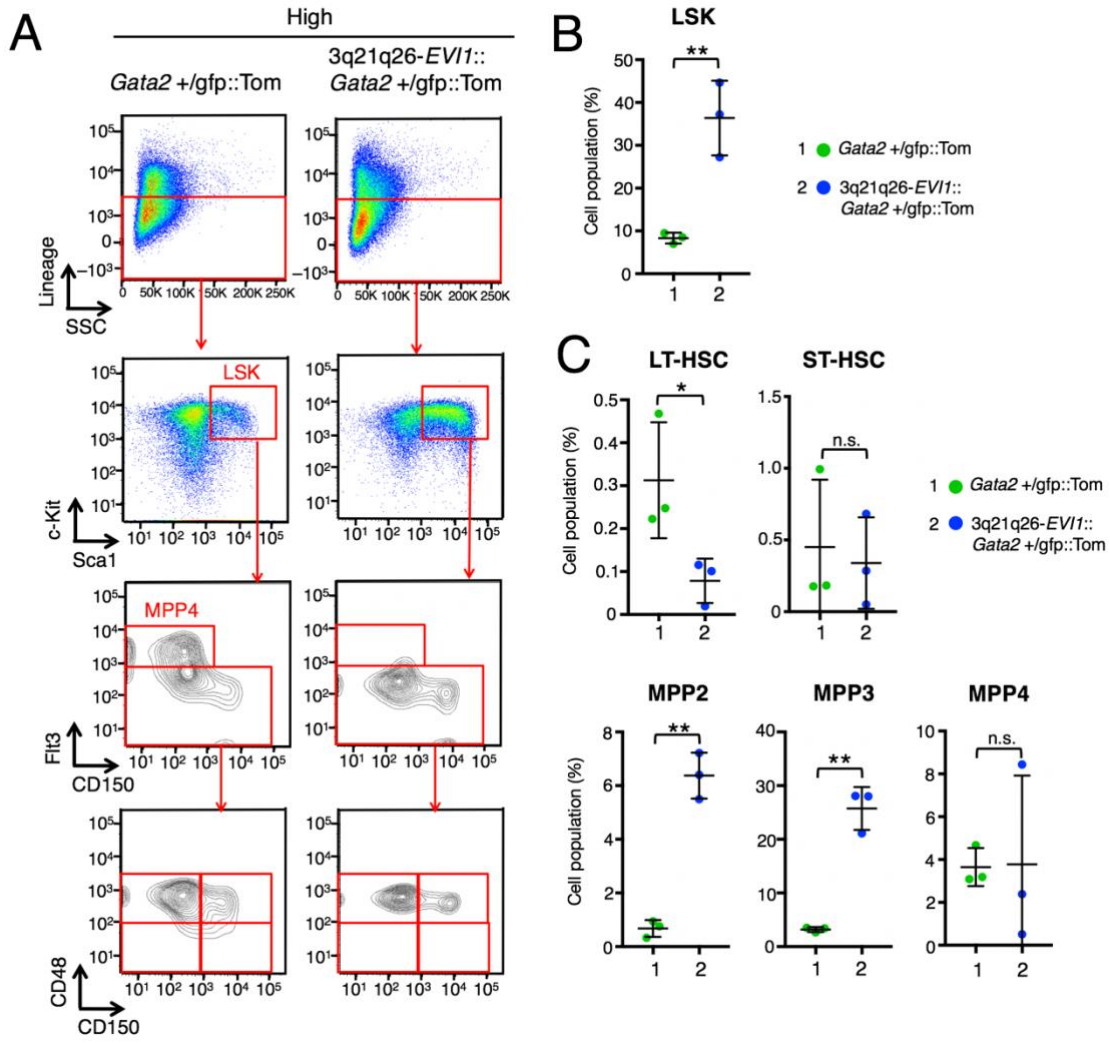
Supplemental Figure 3. Colony-forming analysis of 3q21q26-tdTomato-high (T-High) populations. (A) Representative flow cytometric profiles of the bone marrow cells from WT::Tom, *Gata2*$^{+/gfp}$::Tom 3q21q26-*EVI1*::Tom and 3q21q26-*EVI1*::*Gata2*$^{+/gfp}$::Tom mice. (B) Colony-forming ability of T-High populations of WT::Tom, *Gata2*$^{+/gfp}$::Tom 3q21q26-*EVI1*::Tom and 3q21q26-*EVI1*::*Gata2*$^{+/gfp}$::Tom mice. (C) Percentages of T-High population in the colonies derived from T-High populations of WT::Tom, *Gata2*$^{+/gfp}$::Tom 3q21q26-*EVI1*::Tom and 3q21q26-*EVI1*::*Gata2*$^{+/gfp}$::Tom mice. Values represent the means ± SD. n.s., not significant (one-way ANOVA).

Supplemental Figure 4



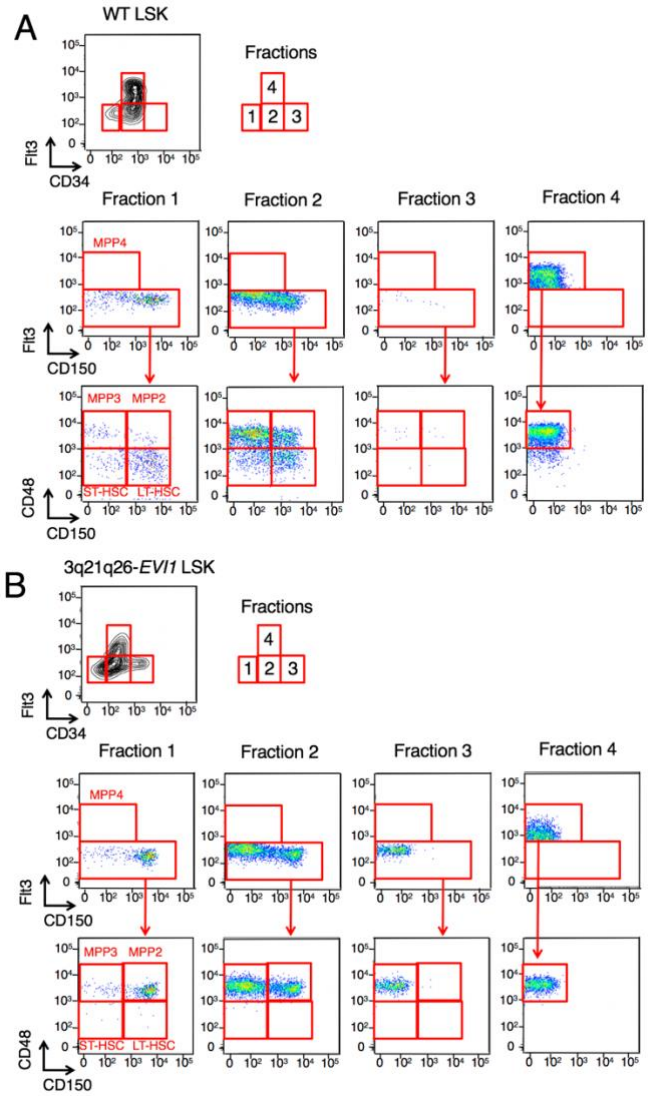
Supplemental Figure 4. *EVI1* and *Gata2*-highly induced fraction contains HSPC and Mks. (A-E) Representative flow cytometric profiles of High, Low and Negative populations in *Gata2*$^{+/gfp}$::Tom and 3q21q26-*EVI1*::*Gata2*$^{+/gfp}$::Tom mouse bone marrows. Red circles and boxes in each panel represent populations of Mks (A), monocytes and neutrophils (B), B cells (C), erythrocytes (D) and HSPC (E).

Supplemental Figure 5



Supplemental Figure 5. High population of 3q21q26-*EVI1*::*Gata2*$^{+/gfp}$::Tom mouse contains high number of HSPC. (A) Representative flow cytometric profiles of High populations in *Gata2*$^{+/gfp}$::Tom and 3q21q26-*EVI1*::*Gata2*$^{+/gfp}$::Tom mouse bone marrows. (B) Percentages of LSK cells in High populations. (C) Percentages of LT-HSC, ST-HSC, MPP2, MPP3 and MPP4 in High populations.

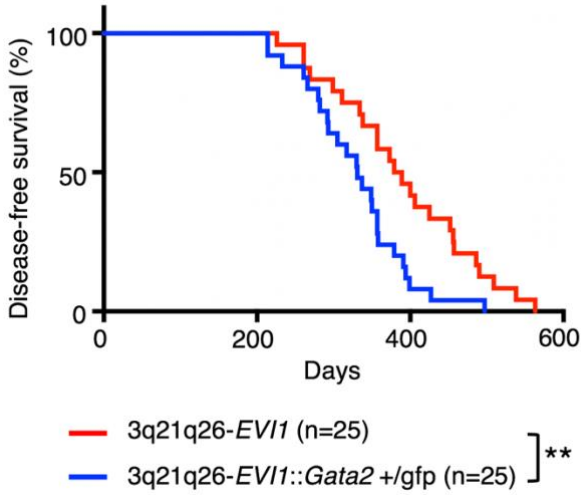
Supplemental Figure 6



Supplemental Figure 6. LSK CD34$^-$ fraction in 3q21q26-*EVI1* mice contains MPP2. (A-B) Representative flow cytometric profiles of LSK in WT (A) and 3q21q26-*EVI1* (B) mice. Analyses of Fractions 1-4 in upper panel in top panels are shown in the middle and bottom panels. In our previous study[1], we described LSK Flt3$^-$CD34$^-$ cells as LT-HSC based on a classical method[2]. In the present study, we analyzed LSK cells of 3q21q26-*EVI1* mice using different, and relatively newer characterization (SLAM-marker based evaluation). We compared the two analyses and found that in WT mice, hematopoietic stem cells in the LSK Flt3$^-$CD34$^-$ fraction (fraction 1, LT-HSC of the previous CD34-based classification) are mainly contained in LSK Flt3$^-$CD150$^+$CD48$^-$ fraction (LT-HSC of the current SLAM-based classification), showing very good coincidence.

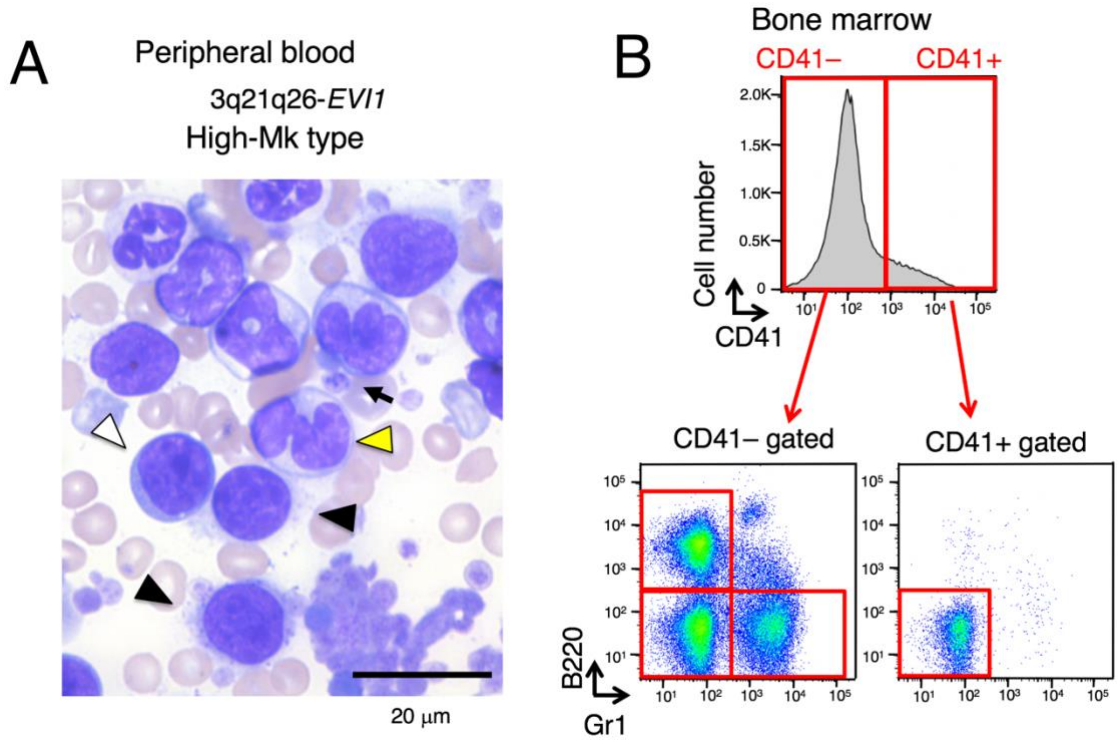
Unexpectedly, however, most of cells in the LSK Flt3^{-}CD34^{-} fraction in 3q21q26-*EVI1* mice are found in the LSK Flt3^{-}CD150^{+}CD48^{+} fraction, which is denoted as the MPP2 fraction in the SLAM classification. Thus, since CD34^{-} MPP2 cells are increased substantially in 3q21q26-*EVI*1 mice, LSK Flt3^{-}CD34^{-} cells were determined to be increased in our previous study. Therefore, we decided to use the SLAM-based classification in this study.

Supplemental Figure 7



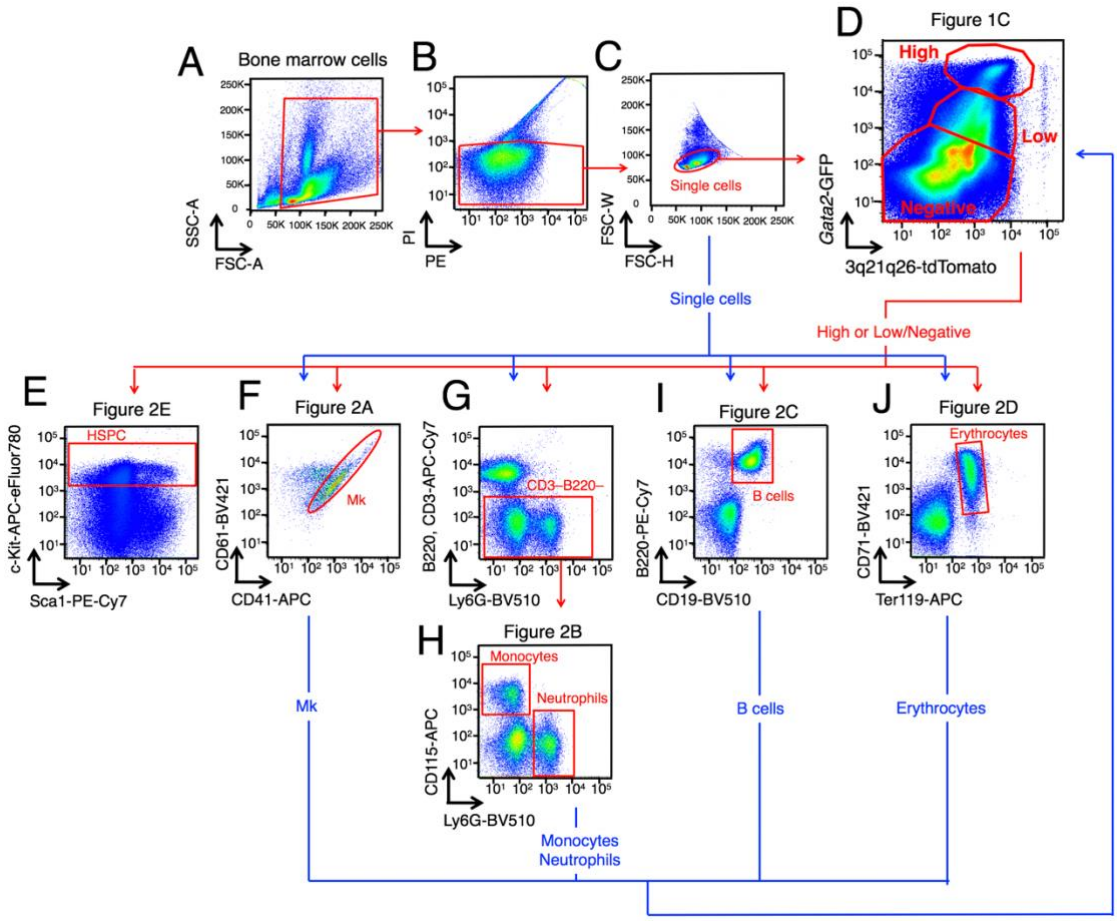
Supplemental Figure 7. Survival curves of 3q21q26-*EVI1* and 3q21q26-*EVI1*::*Gata2*$^{+/gfp}$ mice. Kaplan-Meier survival curves of 3q21q26-*EVI1* (red line) and 3q21q26-*EVI1*::*Gata2*$^{+/gfp}$ mice (blue line). **, $P<0.01$ (log-rank test).

Supplemental Figure 8



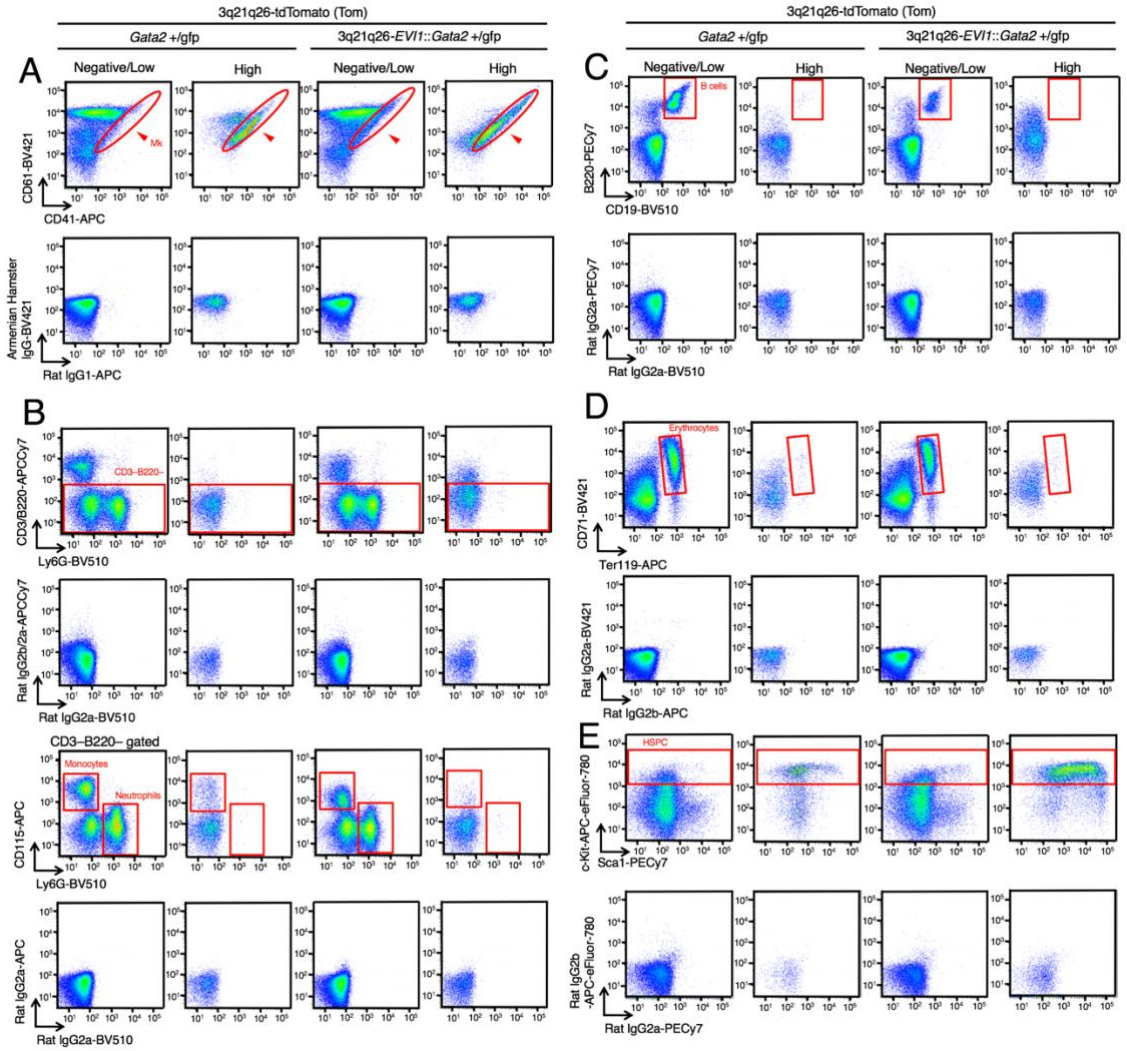
Supplemental Figure 8. High-Mk type leukemia of 3q21q26-*EVI1* mice. (A) Representative smears (Wright-Giemsa staining) of peripheral blood from leukemic 3q21q26-*EVI1* mice (High-Mk type leukemia). Blast cells (white arrowheads), myeloid cells (yellow arrowhead), Mks (black arrowheads) and giant platelets (arrow) are observed. (B) Representative flow cytometric profiles of bone marrow cells from leukemic 3q21q26-*EVI1* mice (High-Mk type leukemia). Note that the flow cytometric profiles of High-Mk type leukemia of 3q21q26-*EVI1* mice are similar to those of 3q21q26-*EVI1*::*Gata2*$^{+/gfp}$ mice.

Supplemental Figure 9



Supplemental Figure 9. Gating strategy for characterization of High, Low and Negative populations in the *Gata2*$^{+/gfp}$::Tom and 3q21q26-*EVI1*::*Gata2*$^{+/gfp}$::Tom mice. Bone-marrow cells are identified by first gating of cell natures by FSC-A versus SSC-A (panel A), of live/death by PI versus PE (panel B), and of singlets by FSC-H versus FSC-W (panel C). High, Low and Negative populations are sub-gated based on the fluorescence of GFP and tdTomato (panel D). To examine which cell types are contained in the High, Low and Negative populations, we gated the High, Low and Negative populations and then examined differentiation-marker expressions of HSPC (panel E), Mk (panel F), myeloid cells (panels G-H), B cells (panel I) and erythrocytes (panel J) in the High, Low and Negative populations (panel D to panels E-J). To calculate frequencies of the High population in each cell type, we gated Mk, monocytes, neutrophils, B cells and erythrocytes and then examined frequencies of the High population in these cells (panels F-J to panel D).

Supplemental Figure 10



Supplemental Figure 10. Isotype controls of flow cytometric analyses. (A-E) Representative flow cytometric profiles of isotype controls of High and Low/Negative populations in *Gata2*$^{+/gfp}$::Tom and 3q21q26-*EVI1*::*Gata2*$^{+/gfp}$::Tom mouse bone marrows. Red circles and boxes in each panel represent populations of Mk (A), monocytes and neutrophils (B), B cells (C), erythrocytes (D) and HSPC (E).

References

1. Yamazaki H, Suzuki M, Otsuki A, et al. A Remote GATA2 Hematopoietic Enhancer Drives Leukemogenesis in inv(3)(q21;q26) by Activating EVI1 Expression. *Cancer Cell*. 2014;25(4):415-427.

2. Osawa M, Hanada K, Hamada H, Nakauchi H. Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell. *Science*. 1996;273(5272):242-245.