SUPPLEMENTARY APPENDIX

# The SKI proto-oncogene enhances the *in vivo* repopulation of hematopoietic stem cells and causes myeloproliferative disease

Sofie Singbrant,[1,2] Meaghan Wall,[3] Jennifer Moody,[1] Göran Karlsson,[1] Alistair M. Chalk,[2] Brian Liddicoat,[2] Megan R. Russell,[2] Carl R. Walkley,[2*] and Stefan Karlsson[1*]

[1]Department of Molecular Medicine and Gene Therapy, Lund Stem Cell Center, Lund Sweden; [2]St. Vincent's Institute of Medical Research and Department of Medicine, University of Melbourne, Fitzroy, Australia; and [3]Victorian Cancer Cytogenetics Service, St Vincent's Hospital, Fitzroy, Australia

 *doi:10.3324/haematol.2013.093971.*
*Manuscript received on June 28, 2013. Manuscript accepted on January 9, 2014.*
*Correspondence: sofie.singbrant@med.lu.se*

## Supplementary Methods

### Virus and infection of primary cells

The SKI and ARPG mutant constructs were cloned into the murine stem cell (MSCV) based MIG vector. Viruses were produced by transient transfection of the cell line Phoenix Eco (ecotropic envelope). BM enriched for stem and progenitor cells (5-FU treated 4 days before BM harvest or sorted LSK cells) was pre-stimulated in either (a) X-VIVO10 media (BioWhittaker, Walkersville, MD) with 1% bovine serum albumin (BSA; Stem Cell Technologies), 1% penicillin/streptomycin (P/S), 1% L-glutamine and $10^{-4}$mM 2-mercaptoethanol, 50ng/ml murine Stem Cell Factor (mSCF; Amgen, Thousand Oaks, CA), 10ng/ml murine Interleukin 3 (mIL-3; PeproTech, London UK), and 50ng/ml human IL-6 (PeproTech) or (b) StemPro-34 SFM Complete Medium (Gibco, VIC, Australia) containing 10ng/ml mSCF (PeproTech, Rocky Hill, NJ), 20ng/ml murine thrombopoietin (mTPO; PeproTech) 20ng/ml murine insulin-like growth factor-2 (mIGF2; R&D Systems) and 10ng/ml human fibroblast growth factor (hFGF1; R&D Systems) (1) for 48 hours. Pre-stimulated cells were resuspended in (a) IMDM with 20% FCS, 1% P/S, 1% L-glut, $10^{-4}$mM 2-mercaptoethanol, 6μg/ml protamine sulphate and the above described cytokines or (b) in fresh supplemented StemPro-34 medium as above with the addition of 8μg/ml Polybrene (Sigma), and plated onto retronectin-coated (Takara Bioscience, Clontech) plates pre-loaded with viral supernatant (1-2 hits), followed by culture at 37°C, 5% $CO_2$ for 48 hours.

### Mice and transplantations

Mice were bred and maintained in the barrier facility at Lund University, Sweden, or obtained from A.R.C., WA, Australia and kept at B.R.C at St Vincent's Hospital, Australia. All experiments were performed under ethical guidelines set by Lund University or with the approval of St Vincent's Health Melbourne Animal Ethics Committee.

0.1-5 x $10^5$ unsorted transduced cells from 8-12w-old C57Bl/6 (BalbC mice in Figure S3) were injected i.v. with 1-2 x $10^5$ fresh whole BM cells into lethally irradiated recipients (8-12w-old, 1x900 cGy, or 2x500 cGy 3 hours apart). Donor, recipient, and support cells were tracked using CD45.1/CD45.2. Cells from multiple transduction (n>8 for MIG and SKI and n=3 for ARPG) with varying initial

transduction efficiencies (MIG 5.0-50.4%, SKI 3.9-41.2%, ARPG 5.5-41.3%) were transplanted into 4-6 recipients per transduction. PB was collected at several time points. Recipients were sacrificed after 16 weeks (12 weeks for BalbC recipients) and half a femur–equivalent of BM was transplanted into secondary and tertiary recipients. One cohort of mice was kept for 18 months.

**Cell preparations and flow cytometry**

PB, BM and spleen were analyzed on a blood cell analyzer (Sysmex, Boule Medonice CA 530-16, or Sysmex KX-21N, Roche Diagnostics, Australia). Erythroid cells in PB were lysed before FACS analysis. For lineage depletion, lineage-positive cells were removed with a magnetic particle concentrator (MPC-6; Dynal Biotech) and sheep anti-rat Fc–conjugated immunomagnetic beads (Dynal Biotech, Oslo, Norway), or BM was enriched for c-kit$^+$ cells using positive magnetic selection (Miltenyi Biotec, Bergisch Gladbach, Germany). Antibodies against murine CD45.1, CD45.2, Mac1, Gr1, F4/80, CD3, CD4, CD5, CD8, B220, IgM, CD43, CD19, Ter119, CD71, Sca1, c-kit, CD34, Flk2, FcγR (CD16/32), ILR7a, CD41, biotinylated or conjugated with FITC, PE, PeCy5, PerCPCy5.5, PECy7, APC, APC Alexa 750, and Pac Blue were all obtained from BD Biosciences Pharmingen (Franklin Lakes, NJ) or eBioscience (San Diego, CA). CD105 and CD150 (clone TC15-12F12.2) were from Biolegend (San Diego, CA), and biotinylated antibodies were detected with Streptavidin conjugated with Alexa 405 or Qdot 605 (Invitrogen, VIC, Australia). Dead cells were excluded using 7-aminoactinomycin D (7-AAD; Sigma-Aldrich, St Louis, MO). Analysis was performed on a FACS Calibur or LSRFortessa, while cell sorting was performed on a FACS Diva or FACS Aria (BD, San Jose CA). Results were analyzed with FlowJo software, version 8.8.6 (Tree Star, Ashland, OR).

**Hematopoietic progenitor assays**

To analyze TGF-β sensitivity 200 GFP$^+$ cells/ml were seeded in methylcellulose (M3231, Stem Cell Technologies), supplemented with 50ng/ml mSCF, 10ng/ml mIL-3, and 10ng/ml hIL-6, with or without 10ng/ml TGF-β (all from PeproTech, Rocky Hill, NJ) (n=2 separate transductions). Unsorted BM cells from 16 weeks post-transplantation were plated for colonies; 30,000 cells/ml as above for myeloid and 50,000 cells/ml for Pre-B colonies (M3630, Stem Cell Technologies). For HGF-

analysis 250 sorted $GFP^+$ cells/ml were seeded into serum-free 1% methylcellulose (Methocel MC #64630; Fluka, IMDM; Invitrogen) supplemented with 20% BIT Serum replacement (Stem Cell Technologies, BC, Canada), P/S, GlutaMAX (Gibco), 50ng/ml mSCF, 10ng/ml mIL-3, 50ng/ml hIL6, and 3U/ml rhEpo (Jansen-Cilag) together with 0, 100nm or 300nm of c-Met Inhibitor III (Calbiochem cat #448105) (n=3 separate transductions). All colonies were plated in duplicates and scored day 7.

**Morphology**

PB smears and cytospins from BM cell suspensions were stained with May Grünwald (Merck) and Giemsa (BDH). Two hundred cells/slide and 1-2 slides/mouse were scored. Megakaryocytes were scored in 10 consecutive high power fields per mouse in paraffin-embedded sections of decalcified tibiae stained with Mayer's hematoxylin and eosin (n=4 for MIG and 5 for SKI). Pictures were taken using an Olympus BH2 microscope, DPlan Apo10 UV PL 10x/0.40NA, DPlan Apo20 UV PL 20x/0.70NA and SPlan 40 40x/0.70NA Olympus objective lenses, and a MagnaFire-SP Optronics Digital Camera. Magnafire2.1C was used to acquire the images digitally, and brightness was adjusted using Adobe Photoshop CS5.

**Gene Expression Profiling**

Sorted LKS cells from C57Bl6 mice were cultured for 48 hours, infected and $GFP^+$ cells were isolated after 48 hours (n=6 independent transductions). Total RNA was extracted using Trizol followed by RNeasy Micro kit (Qiagen). RNA was hybridized to the Illumina Mouse WG-6 v2.0 Expression BeadChip at the Ramaciotti Centre for Gene Function Analysis at the University of New South Wales, Australia. Illumina Data was processed in GenePattern (GP) (1) using IlluminaExpressionFileCreator (GP, v2.1) with background correction using manifest MouseWG-6_V2_0_R3_11278593_A.txt (GPL6887). Detection p-values were calculated as the proportion of negative control probes with expression greater than the regular probe in question using the calculateDetected function in the bioconductor beadarray package (v2.8.1) (2). The data was log transformed in GenePattern (logTransform, GP, v3.1), quantile normalized using NormalizeColulms (GP, v4.2.1) and filtered to remove 6,324 “Bad” probes as defined by (3) (leaving 38,958 probes). For differential expression, probes that were not detected in at least 5/6 (75%) of either SKI or MIG samples were discarded, leaving 9,780 probes (7,514 unique gene symbols).

Differentially expressed genes were determined using the LimmaGP module (4) with the most informative probe for each gene used (**Supplementary file 1**). For GSEA analysis GSEAPreRanked (5) was run against C2 (curated gene sets) using the ranked gene list from LimmaGP on all "good" probes (38,958 probes), with the number of permutations=1,000, against the c2.all.v3.0.symbols.gmt gene sets database (**Supplementary file 2**). The Gene Pattern modules LimmaGP and GSEAPreRanked were developed by Mark Cowley (m.cowley@garvan.org.au) from Peter Wills Bioinformatics Centre, Garvan Institute of Medical Research, Sydney Australia. The heatmap was generated using ComparativeMarkerSelectionViewer (6) selecting the top 25 up and down regulated genes by score determined by LimaGP using the 9,780 probes. Raw and processed microarray data is available in GEO (http://www.ncbi.nlm.nih.gov/geo/) under accession GSE39457.

**Western blot analysis**

Whole cell extracts prepared from GFP-sorted BM LSK cells transduced with MIG, SKI or ARPG were separated on 4-12% NuPAGE Bis-Tris mini gels and blotted using iBlot gel transfer stacks (Life technologies). Membranes were blocked in PBS Tween20 containing ECL prime blocking agent (GE Healthcare, UK) before incubation with antibodies (SKI; G8, Santa Cruz Biotechnology, Santa Cruz, CA, actin; Ab5, BD Bioscience Pharmingen, San Diego, CA). The binding of primary antibody was detected by a horseradish peroxidase conjugated secondary anti-mouse antibody (Amersham Biosciences), followed by a chemiluminescence ECL blot kit (GE Healthcare).

**Statistical analysis**

The significance of results was analyzed using a paired or unpaired two-tailed student T-test, $p<0.05$ was considered significant. All data is presented as mean ± S.E.M.

**References**

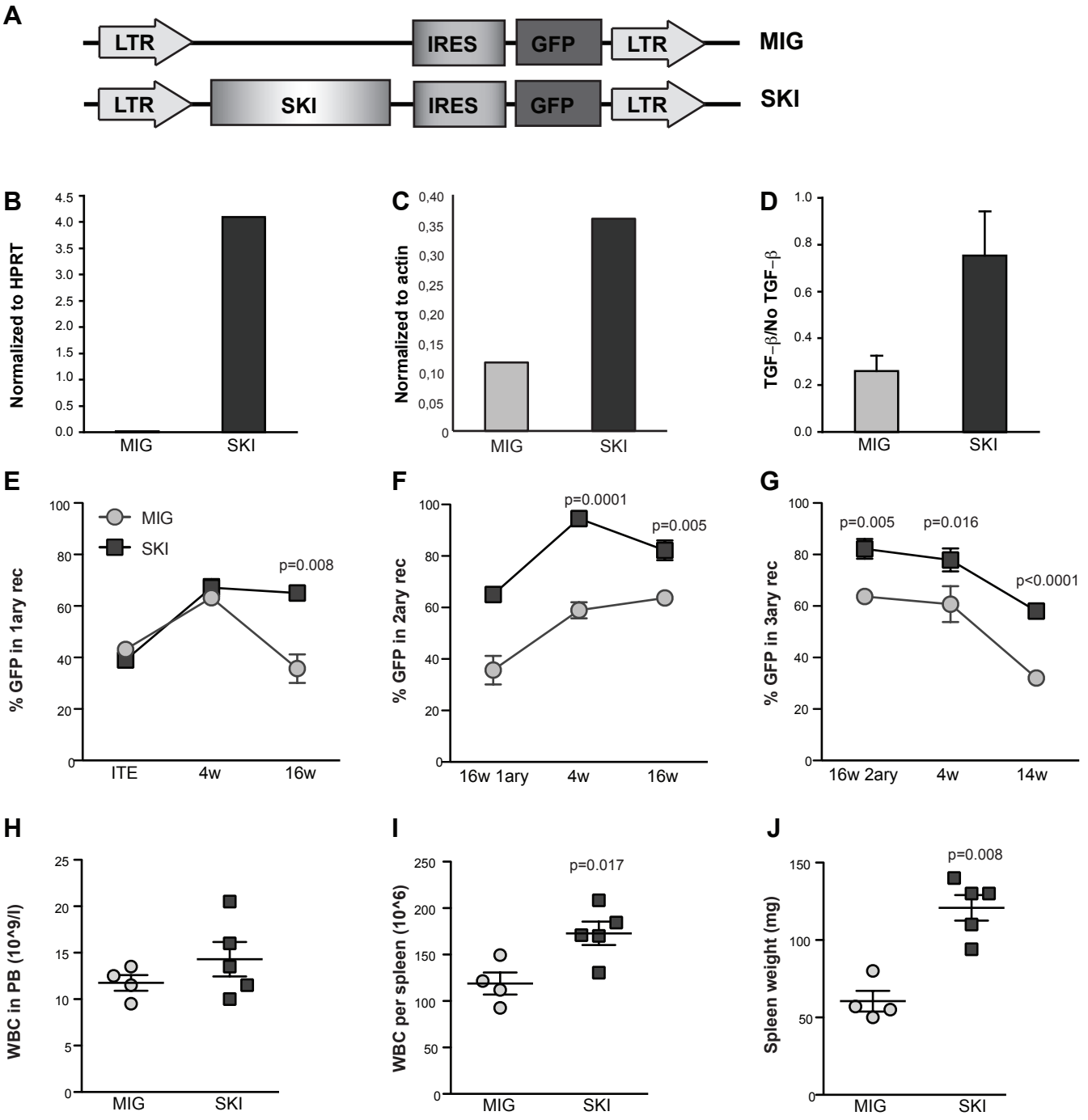

1. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. GenePattern 2.0. Nat Genet 2006;38(5):500-1.
2. Dunning MJ, Smith ML, Ritchie ME, Tavare S. beadarray: R classes and methods for Illumina bead-based data. Bioinformatics 2007;23(16):2183-4.
3. Barbosa-Morais NL, Dunning MJ, Samarajiwa SA, Darot JF, Ritchie ME, Lynch AG, et al. A re-annotation pipeline for Illumina BeadArrays: improving the interpretation of gene expression data. Nucleic Acids Res 2010;38(3):e17.

4. Smyth GK. Limma: linear models for microarray data.
5. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102(43):15545-50.
6. Gould J, Getz G, Monti S, Reich M, Mesirov JP. Comparative gene marker selection suite. Bioinformatics 2006;22(15):1924-5.

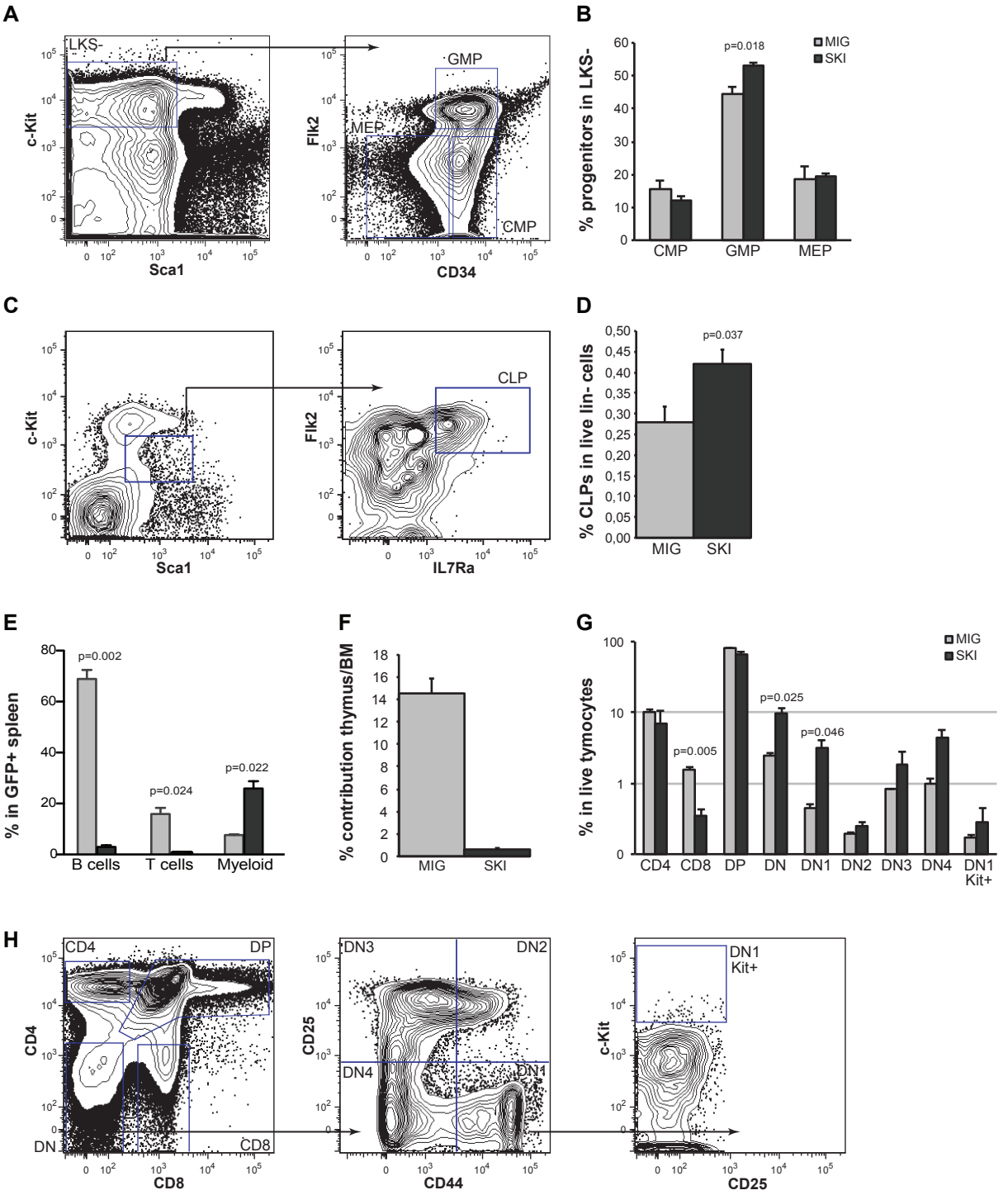
**Supplemetary Table 1. Over-expression of SKI in mice results in a phenotype highly consistent with a myeloproliferative neoplasm**

| | | Myeloproliferative neoplasms | | Myelodysplastic/myeloprolierative neoplasms | |
|---|---|---|---|---|---|
| **FEATURE** | **SKI over-expression** | **CNL** | **CML-CP** | **Atypical CML** | **CMML** |
| Splenomegaly | Y | Y | Y | Y/N | Y/N |
| Progression to acc.phase/blast phase/AML | N | N/Y | Y | Y | Y/N |
| **BLOOD** | | | | | |
| Leukocytosis | Y/N | Y | Y | Y | Y/N |
| Neutrophil leucocytosis | Y | Y | Y | Y | Y/N |
| Left shift in granulopoiesis | minor | minor | Y | Y | minor |
| Eosinophilia | N | N | Y | N | N/Y |
| Basophilia | N | N | Y | N/Y | N/Y |
| Anemia | Y | N | Y | Y | Y |
| Thrombocythemia | Y | N/Y | Y | N | N |
| Monocytosis | Y | N | Y/N | Y | Y |
| Dysplasia | N | N | N | Y | Y |
| **BONE MARROW** | | | | | |
| Hypercelluar | Y | Y | Y | Y | Y |
| Myeloid hyperplasia | Y | Y | Y | Y | Y |
| Reduced erythropoiesis | Y | N/Y | Y | N/Y | N |
| Increased megakaryocytes | Y | N | N/Y | N/Y | N |
| Monolobular megakaryocytes | N | N | Y | Y | N/Y |
| Dysplasia | N | N | N | Y | Y |
| **PERFECT MATCH** | | 10 of 17 | 9 of 17 | 6 of 17 | 6 of 17 |

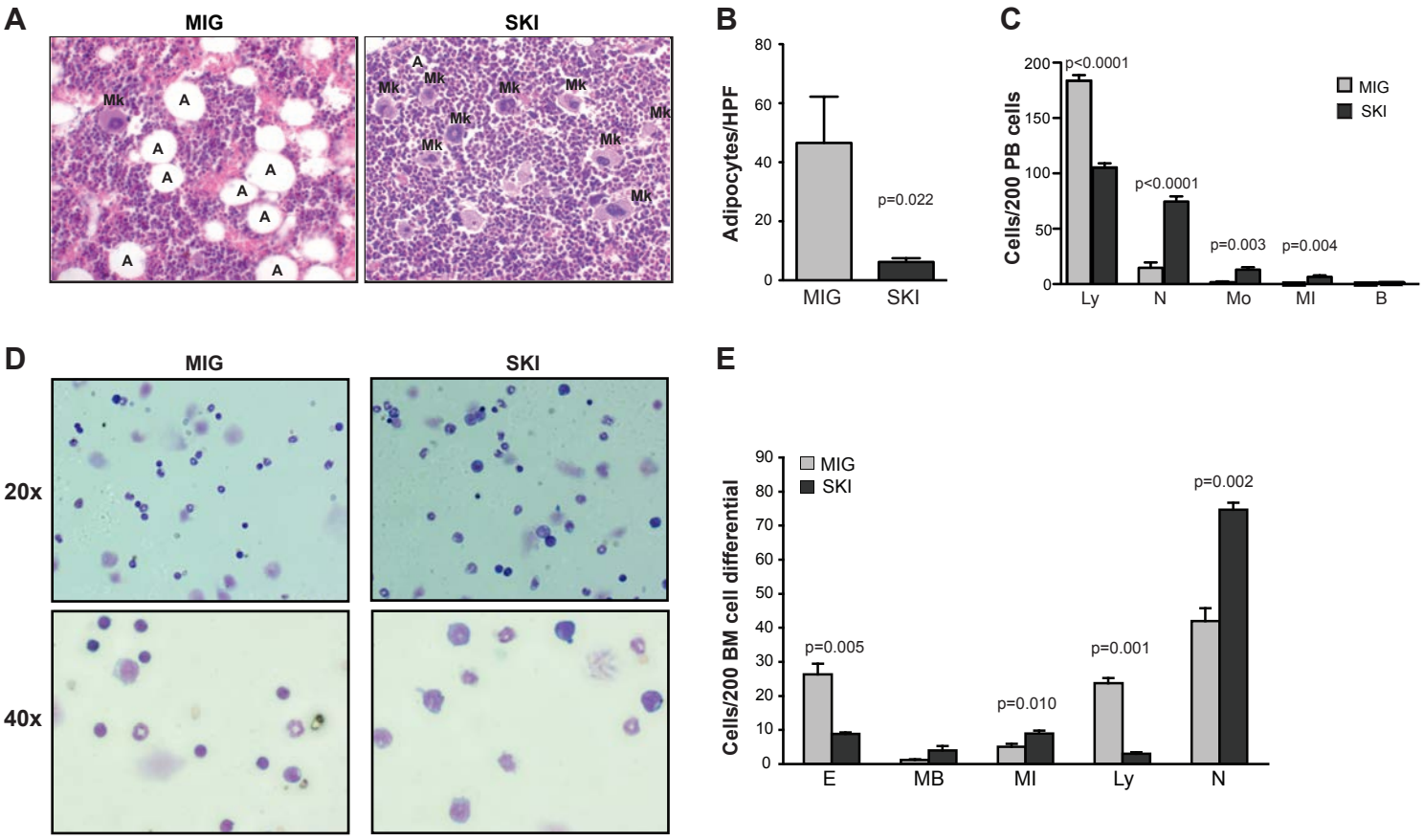
CNL=Chronic Neutrophilic Leukemia, CML=Chronic Myeloid Leukemia, CP=chronic phase, CMML=Chronic Myelomonocytic Leukemia
Acc=accelerated, Y=Yes, N=No, Y/N=More commonly Yes than No, N/Y=More commonly No than Yes



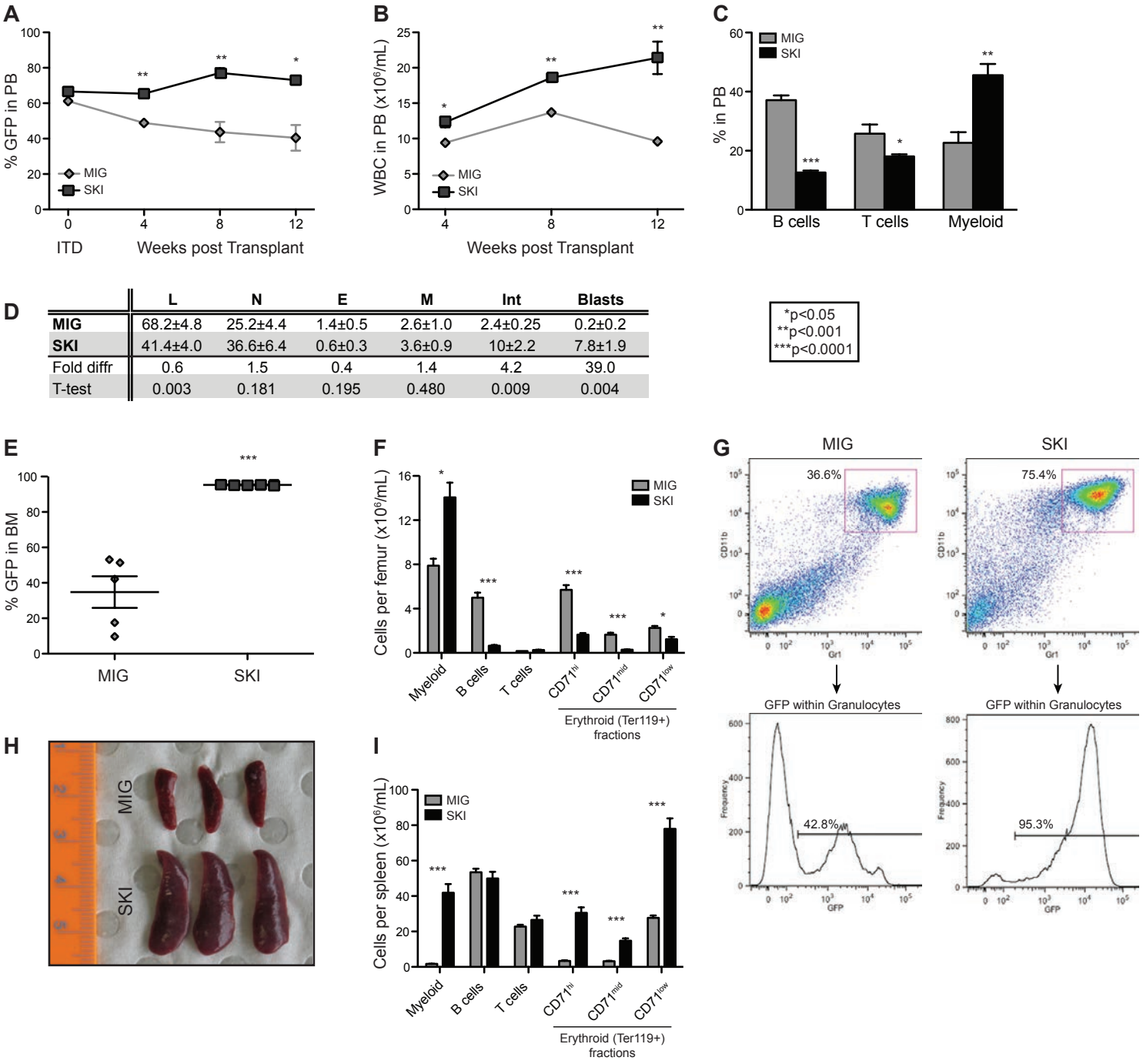

**Supplementary Figure 1**. **SKI is robustly and functionally over-expressed in primary cells.** (A) Schematic of MIG control and SKI over-expressing vectors. (B) Real-time PCR on GFP+ BM cells confirming expression of human SKI mRNA after transduction. (C) Western blot on GFP-sorted LSK+ cells detected with an anti-SKI antibody identifyting both endogenous murine and over-expressed human SKI confirming over-expression of SKI at the protein level. (D) Colony assay measuring TGF-β sensitivity, using GFP-sorted cells plated with or without 10 ng/ml TGF-β (n=2). Donor mreconstitution in (E) primary, (F) secondary, and (G) tertiary recipients at indicated timepoints from one representative transduction with high initial transduction efficiency (ITE, n=4-5 per group). (H) White blood cell count (WBC) in PB and (I) spleen. (J) Spleen weight. (n=4 for MIG and 5 for SKI).The data is collected from primary recipients at 16 weeks post transplantation if not stated otherwise and is presented as mean±S.E.M.




**Supplementary Figure 2**. **SKI over-expressing cells greatly contribute to CLPs and are able to form all stages during T-cell development.** Phenotypic analysis of (A-B) the myelo-erythroid progenitors common myeloid progenitor (CMP) cells, granulocyte-monocyte progenitor (GMP), and megakaryocyte-erythrocyte progenitor (MEP) and (C-D) common lymphoid progenitor (CLP) using FACS and the indicated surface markers in lineage- bone marrow (BM). (E) Lineage distribution in spleen as determined by FACS. (F) The contribution of transduced cell to thymus compared to BM. (G-H) T-cell development in thymus including double positive (DP) and double negative (DN) cells. The data is presented as mean±S.E.M. (n=4 for MIG and 3-5 for SKI).

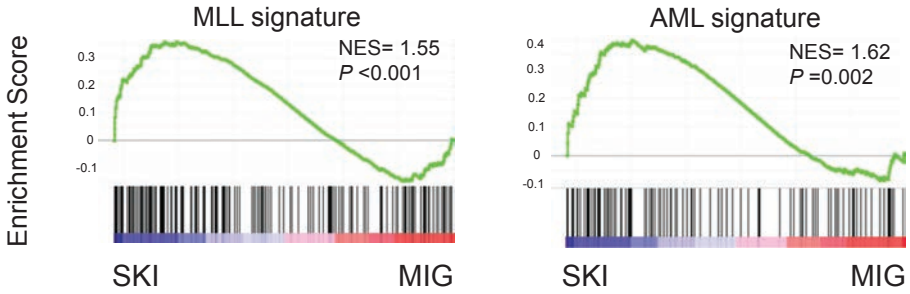



**Supplementary Figure 3. Morphological analysis of SKI over-expressing cells reveals increased myeloid cells and megakaryocytes with a concurrent decrease in erythroid and lymphoid cells.** (A) Hematoxylin and eosin stained bone marrow sections (original magnification 200x). (B) Quantification of adipocytes in bone marrow sections. (C-D) Cytospins of unsorted bone marrow from (C) primary, and (D) secondary recipients 16 weeks post-transplant. A=adipocyte, Mk=megacaryocyte, Ly=lymphocyte, N= neutrophil, Mo=monocyte, M=myeloid, E=erythroid, B=blast, I=intermediate, L=late, HPF=high power field. n=4-5 for MIG and 5 for SKI, all data is presented as mean±S.E.M

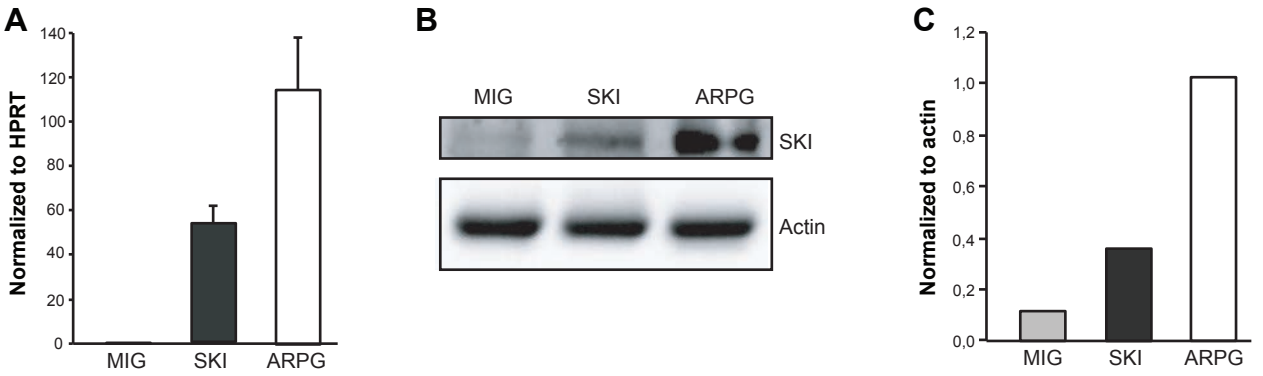



| | L | N | E | M | Int | Blasts |
|---|---|---|---|---|---|---|
| **MIG** | 68.2±4.8 | 25.2±4.4 | 1.4±0.5 | 2.6±1.0 | 2.4±0.25 | 0.2±0.2 |
| **SKI** | 41.4±4.0 | 36.6±6.4 | 0.6±0.3 | 3.6±0.9 | 10±2.2 | 7.8±1.9 |
| Fold diffr | 0.6 | 1.5 | 0.4 | 1.4 | 4.2 | 39.0 |
| T-test | 0.003 | 0.181 | 0.195 | 0.480 | 0.009 | 0.004 |

**Supplementary Figure 4. Over-expression of SKI on a BalbC background results in a more pronounced myeloproliferative disease than on C57Bl6 background.** (A) Donor contribution, (B) white blood cell counts (WBC) and (C) lineage distribution in peripheral blood. (D) Differential counts of blood smears, 100 cells counted per mouse, L=lymphocyte, N=neutrophil, E=erythroid, M=monocyte, Int=intermediate. (E) Donor contribution and (F) lineage distribution in BM. (G) Percentage of granulocytes in total BM and % of GFP within the granulocytic compartment. (H) Spleen size of three representative mice, and (I) lineage distribution in spleen. The data is collected at 12 weeks, and is presented as mean ± S.E.M (n=5)




**Supplementary Figure 5. SKI induces gene signatures associated with myeloid leukemia.** Gene set enrichment analysis (GSEA) determining specific pathways that are positively regulated by SKI, including MLL and AML. q-values<0.25 were considered significant.




**Supplementary Figure 6**. **ARPG mutant is robustly over-expressed in primary cells.**

(A) Real-time PCR (n=3 seperate transductions each of MIG, SKI and ARPG) and (B-C) Western blot on GFP+ BM cells confirming robust over-expression of human ARPG mutant, exceeding the level of SKI over-expression at both the mRNA and protein level. The data is presented as mean±S.E.M.