FACS analysis
For FACS analysis cells were stained with the following rat anti-mouse monoclonal antibodies (all from BD Bioscience Pharmingen): FITC-conjugated CD3 (145-2C11), PE-conjugated CD8 (53-5.8), allophycocyanin (APC)-conjugated CD4 (L3T4), PE-conjugated CD25 (PC61) or PercP-conjugated CD25 (7D4), APC–conjugated CD44 (IM7) for T cells; PE-conjugated CD19 (85657), FITC-conjugated IgM, APC-conjugated CD117 (2B8), FITC-conjugated Gr (RB6-8C5), PE-conjugated Mac (M1/70), APC-conjugated NK1.1 ($NKR^{-}P1C^{+}NKR^{-}P1B$), APC-conjugated B220, PE-conjugated CD45.1 (A20), biotinylated CD45.2 (104) and CD45.1(104) for B cells. For the analysis of BM subpopulations, FITC-conjugated $Sca1^{+}$ (E13-161.7), PE-conjugated BP1 (6C3/BP1 antigen), PE-conjugated CD43 (S7) and biotin-conjugated IL7Rα (B12-1), CD24 (30-F1), CD38 (CALTAG) and CD93 (493) were used, with subsequent staining with PercP-conjugated streptavidin.

Array-comparative genomic hybridization analysis
DNA copy number alteration analysis of murine tumors was performed using a customized mouse CGH platform (Agilent) comprising 244,000 probes, in which the sex chromosome probes were replaced with 18,000 probes. The latter provided dense coverage of gene targets with recurrent copy number alterations in B-progenitor ALL. The raw two-color array data were background-subtracted and LOESS-normalized using the Linear Models for Microarray Analysis software package. The circular binary segmentation algorithm was applied to normalized log2 ratio data and thus identify copy number alterations. A threshold of five consecutive markers per segment and a copy number threshold of less than 1.7 or greater than 2.3 copies was used to identify lesions. All putative segments were manually curated to exclude regions of noise and the technical artifacts known to be inherent in this array design. Thus, the lesions reported here correspond to a stringently filtered list of alterations and should be viewed as the lower limits for copy number alterations.

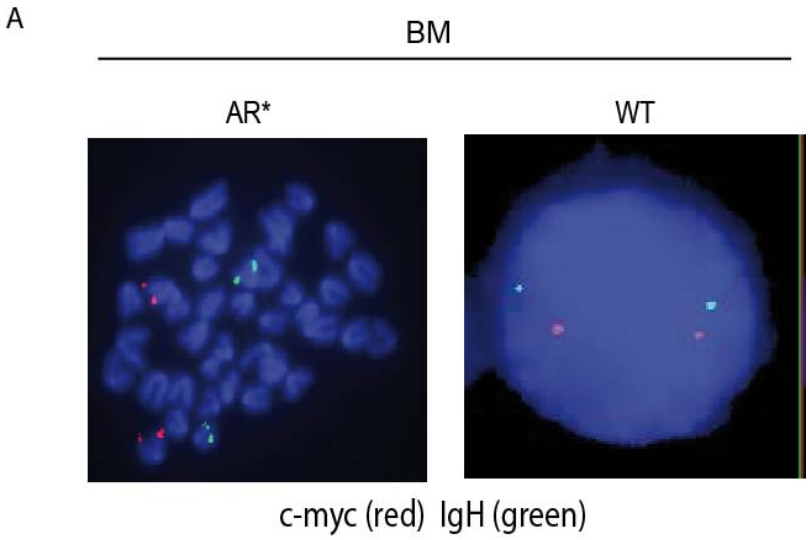
Array comparative expression analysis
Bone marrow from C57BL6 mice (Jackson Laboratories, Bar Harbor, Maine) was flushed from femora and tibiae, red cells lysed, and stained with a panel of antibodies against CD43, B220, CD24 and BP1 to define populations corresponding to common lymphoid progenitors, and B cell populations corresponding to Hardy Fractions A through F. Bone marrow was sorted using FACSVantage SE (with DiVa option) flow cytometers (BD Biosciences, San Jose, CA). Sort purities were greater than 90%. Bone marrow of diseased AR* mice was flushed from femora and tibiae with lymphoblast content greater 80%. Total RNA was extracted using the Arcturus PicoPure kit (Molecular Devices, Sunnyvale, CA), quantitated using a Nanodrop spectrophotometer (Thermo Scientific), and RNA integrity assessed using an Agilent Bioanalyzer. RNA was amplified using the WT-Ovation™ Pico RNA Amplification System (NuGEN Technologies, San Carlos, CA). Fragmented and labeled cRNA was hybridized to 430v2 gene expression microarrays (Affymetrix, Santa Clara, CA), and processed using MAS 5.0 (Affymetrix). GAPDH and Actin 5′:3′ ratios were less than 5 for all arrays. Gene expression data was processed in R and GeneMaths XT v2.1 (Applied Maths, Austin TX) and clustered using log transformed data, Pearson's Correlation coefficient and the Unweighted Pair Group Method with Arithmatic Mean (UPGMA).

Genes	Primer sequences
Pu1	A: CCAACGTCCAATGCATGACT B: GCATGTAGGAAACCTGGTGA C: AGAACTTCCCTGAGAACCACT
E2a	A: AGCGATCTCCATGGGCTTTT B: GTCGGCTACTGATGCGATTT C: GGCCAGTCTTTTGCATAACC
Il7rα	A: CCCTCTGACCTGAAAGTCGT B: GGTAGAACTTGGACTCCACT C: GTAAAGCATGATGTGGCCTAC
Gata1	A: CACTGGCCTACTACAGAGAA B: TATGGCAAGACGGCACTCTA C: ATGGCAGGCTTCCATGAAAC
Runx1	A: GCAGCATGGTGGAGGTACTA B: CTGCCGAGTAGTTTTCATCG C: ACTTCCTCTGCTCCGTGCTA
Lmo2	A: AATGTCCTCGGCCATCGAAA B: GATAGTCTCTCCTGCACAAT C: CGCTACTTCCTGAAAGCCAT
C-myb	A: CAGAAAGTGCTGAACCCTGA B: GGTGGAATTCCAGTGGTTCTT C: ATTGGAAAGCAGTGTCGGGA
Notch1	A: GCTACGAATGTGCCTGTGAA B: CATACGTAGCCACTGGTCAT C: CAACGAGTGCAACAGTAACC

Table S1. Multiplex PCR primer list.

Genes	Primer sequences	annealing
P19Arf	p19 FW AAGTTCGTGCGATCCCGGAG INK4a RV CCAAGAGCGGGGACATCAA	63°C
MDM2	MDM2 FW ATC TCC TAG GAG ATG TGT TTG G MDM2 RV CGT AAG TGA GCA TTC TGG TGA T	60°C
P16/INK4a	P16 FW AACTCTTTCGGTCGTACCCC INK4a RV CCAAGAGCGGGGACATCAA	60°C
Hes1	Hes1 FW CCAGCCAGTGTCAACACGA Hes1 RV AATGCCGGGAGCTATCTTTCT	60°C

Table S2. RT-PCR primer list



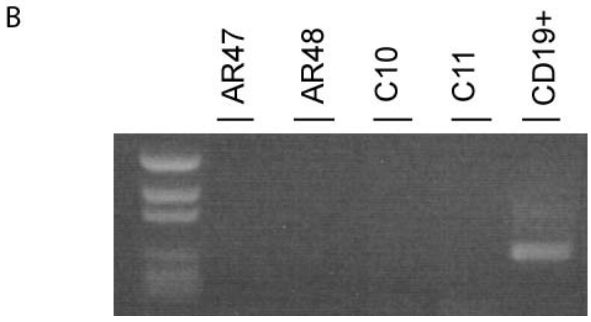


Figure S1. No residual Rag1 activity is detectable in pro–B-cell leukemia
(A) Metaphases were prepared directly from "AR" pro–B-cell leukemia and "WT" mononuclear bone marrow cells after treatment with colcemid. (10 µl per 1×10^6 cells (Karyomax 10mg/ml, Gibco) and incubated at 37°C for 3 hrs at a concentration of 1×10^6 cells/ml in GIBCO RPMI 1640 Glutamax (Invitrogen) and 5% GIBCO FBS (Invitrogen). The cells were then incubated for 20 min with 5.6 g/l KCl and subsequently fixed in methanol/acetic acid (3:1). The FISH probes were purchased from the BACPAC Resource Center at Children's Hospital Oakland Research Institute (Oakland, CA, USA): IgH-chromosome 12 (RP23-65B9) and c-myc–chromosome 15 (RP23-331E20). FISH probes targeted the c-myc locus (red signal) and the IgH locus (green signal). Shown is one representative experiment out of three. (B) Genomic DNA was prepared from "AR" B-cell precursor leukemia or from "WT" $CD19^+$ sorted splenocytes as positive control. Fragments containing the introns downstream of the $V_HD_HJ_H$ elements were amplified using the J558Fr3 (5′CAGCCTGACATCTGAGGACTCTGC-3′) and JH4-int2 primers (5′GCTCCACCAGACCTCTCTAGACAGC-3′), which anneal within the framework 3 region of most V_HJ558 elements and within the intron downstream of J_H4, respectively. Annealing was performed at 71°C for 35 cycles (Schenten et al. 2009).

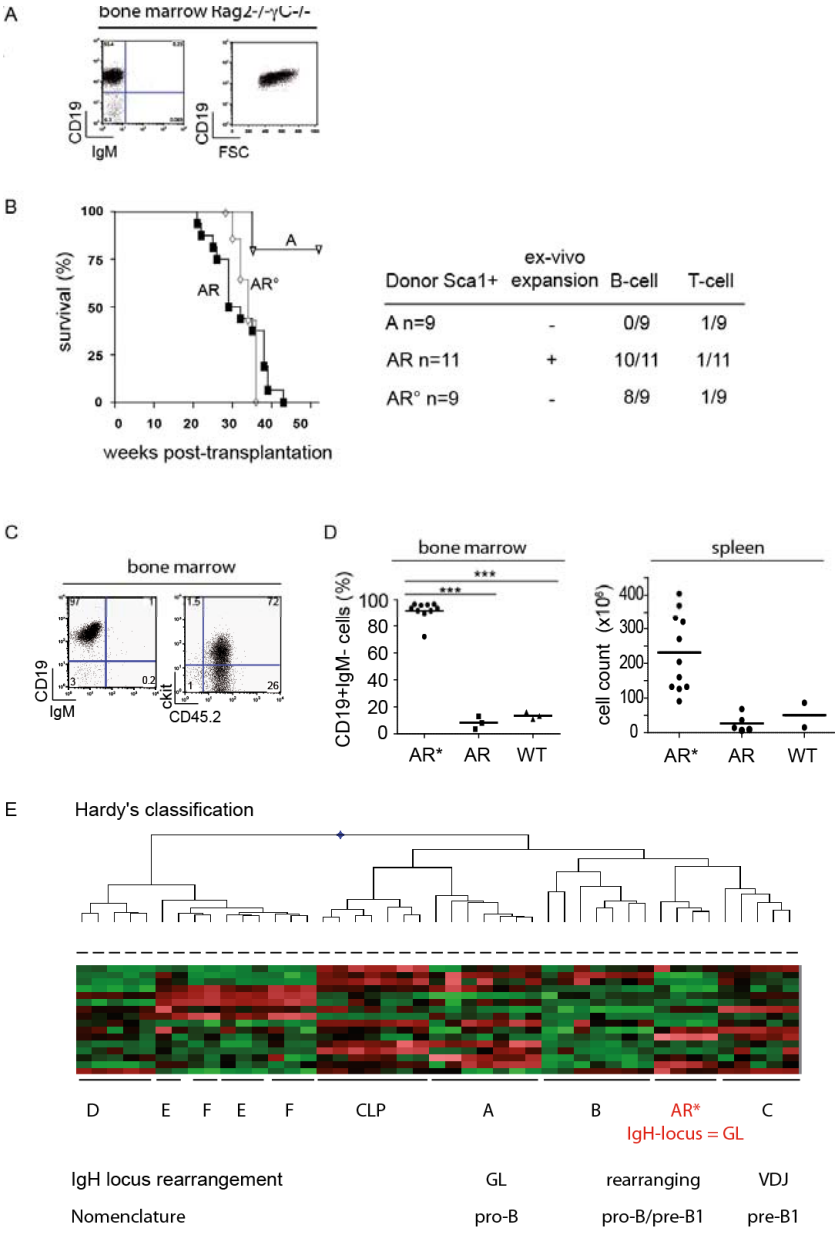


Figure S2. Phenotype of diseased mice transplanted with AR Sca1^{+} cells
(A) FACS analysis of pro–B-cell leukemia occurring in Rag2$^{-/-}$γC$^{-/-}$ mice 14 days after secondary transplantation of 5 × 10^{6} "AR" pro–B-cell leukemia derived blast cells (shown is one representative analysis out of six recipient mice transplanted from two independent mice). (B) Kaplan-Meier survival curves of "WT" mice transplanted with "A"-derived Sca1^{+} mock-cultured (n=9), "AR"-derived Sca1^{+} mock-cultured (n=11) and "AR°" not *ex vivo* cultured cells (n=9). (C) The pro–B-cell leukemia phenotype in the "AR" transplanted group indicates a donor-derived CD45.2^{+}CD19^{+}IgM^{-}ckit^{low} phenotype. (D) Percentage of CD19^{+}IgM^{-} cells of total MNCs in the bone marrow of diseased mice compared to age matched healthy "AR" and "WT" animals Full cell count of spleens of "AR" leukemic mice (AR*) is highly increased compared to spleens from age-matched "AR" and "WT" mice as it is indicated in absolute cell count × (×10^{6}). (E) Expression profile of pro–B-cell leukemia (AR*) in comparison to bone marrow populations corresponding to the common lymphoid progenitor and B-cell populations of Hardy Franctions A through F. Indicated is the time point of BCR-receptor rearrangement in correspondence to the Hardy fraction. GL indicates germline configuration of the IgH locus.

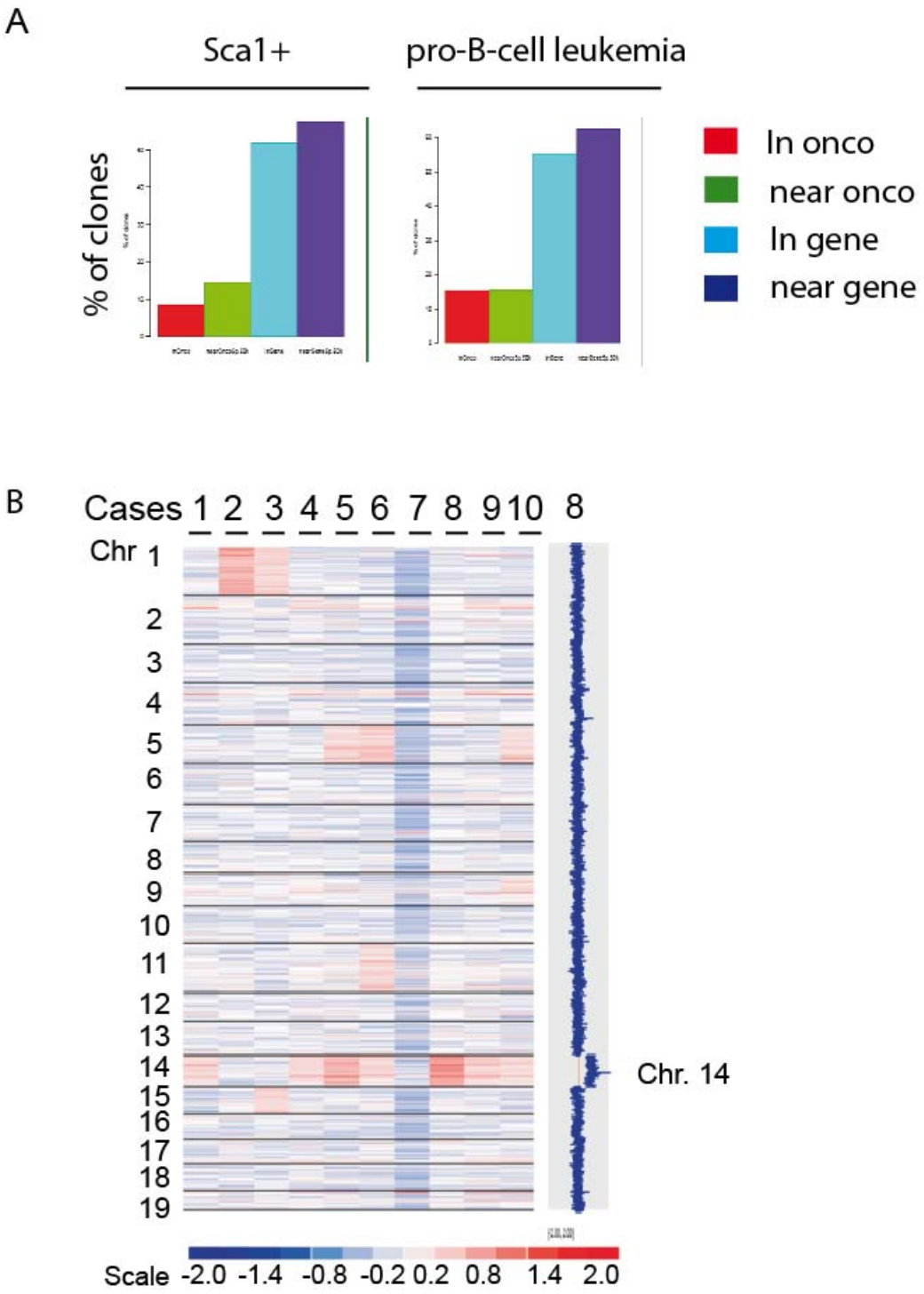


Figure S3
(A) Integration site analysis of AR Sca1^{+} cells after transduction with the gamma–retroviral-GFP vector before transplantation and in the pro–B-cell leukemia after transplantation in recipient mice. Indicated is the percentage of clones with integrations in or near a gene or oncogene. (B) A comparative genomic hybridization analysis (on the Agilent platform) was performed on 10 "AR" pro–B-cell leukemia (cases 1–10). The heat map indicates gains (red) and losses (blue) of genomic regions in the 19 murine chromosomes (Chr. 1–19). Recurrent gains of chromosome 14 are observed, as highlighted for case 8.

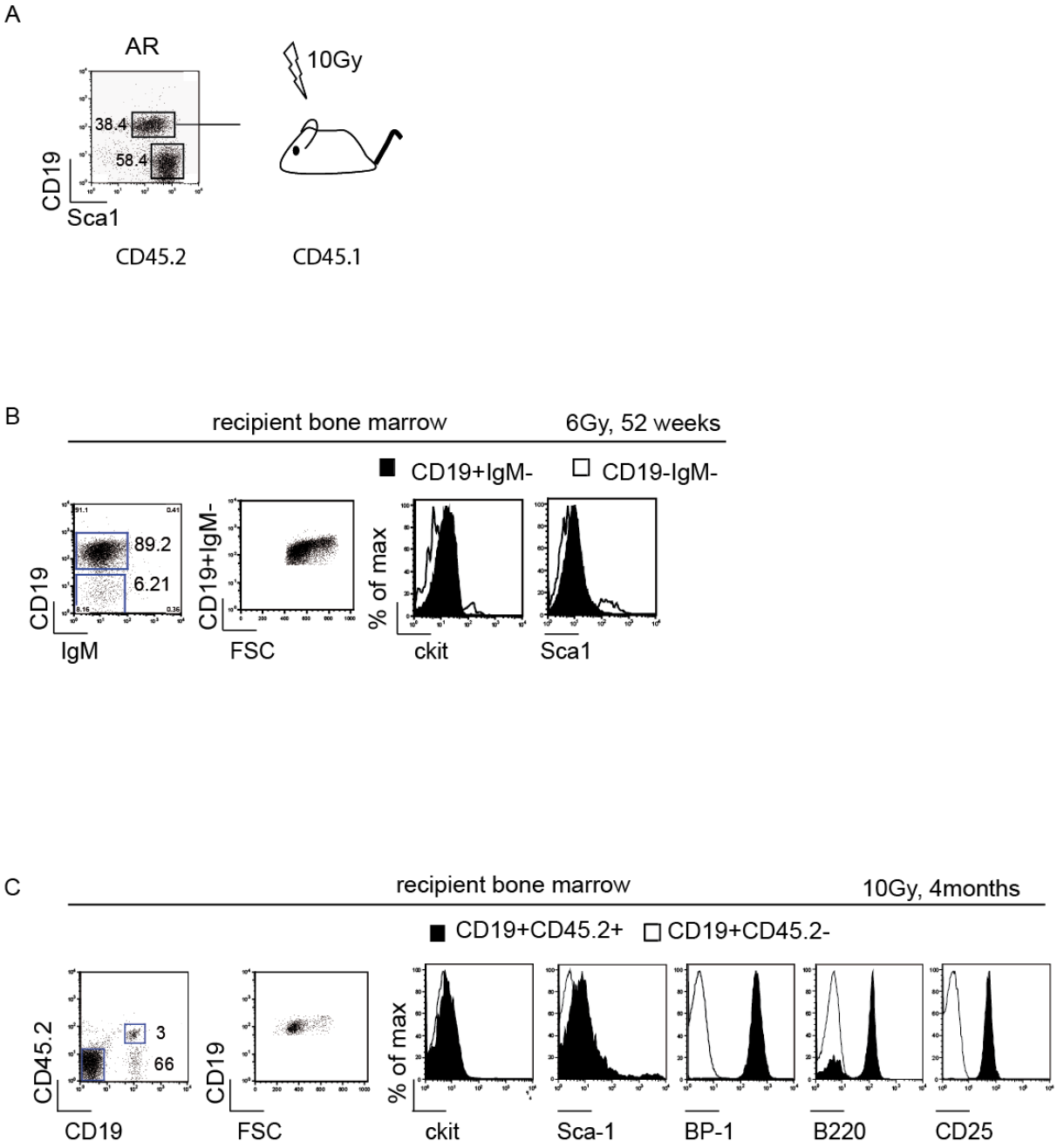
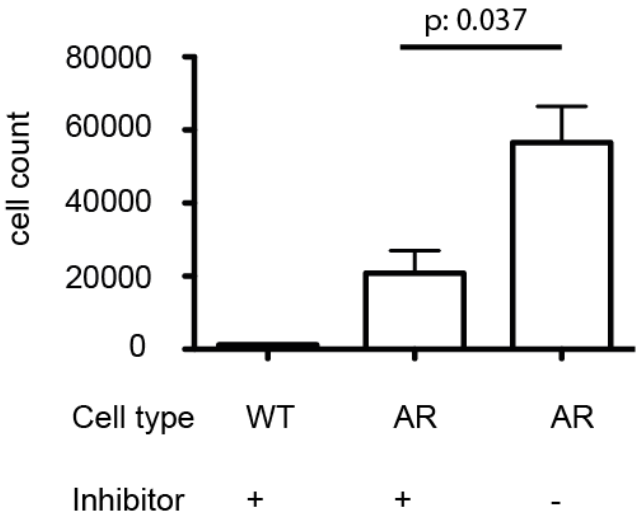


Figure S4. "AR" Sca1^{+}CD19^{+} population is expanding in the bone marrow but not reconstituting the haematopoietic lineages after transplantation

(A) Bone marrow of "AR" mice was harvested and the "AR" Sca1^{+}CD19^{+} population was flow sorted. The "AR" Sca1^{+}CD19^{+} population was directly injected i.v. in three sublethally (6 Gy) and three lethally irradiated (10 Gy) "WT" mice. (B) Follow up of sublethally irradiated mice was 52 weeks. One out of three mice in the sublethally irradiated group showed infiltration of BM and spleen with a "AR" Sca1^{low} CD19^{+} population at 52 weeks after transplantation. After red blood cell lysis, the "AR" CD19^{+}IgM^{-} population (with co-expression of Sca1^{+} and ckit^{+}) represents 89.2% percent of the whole bone marrow compartment. (C) In the lethally irradiated recipients the "AR" Sca1^{+} CD19^{+} population could be tracked 18 weeks after transplantation in the bone marrow in one out of three mice, with a donor derived B-cell phenotype (BP-1^{+}, B220^{+}, CD25^{+}) analysed by FACS.

A



B

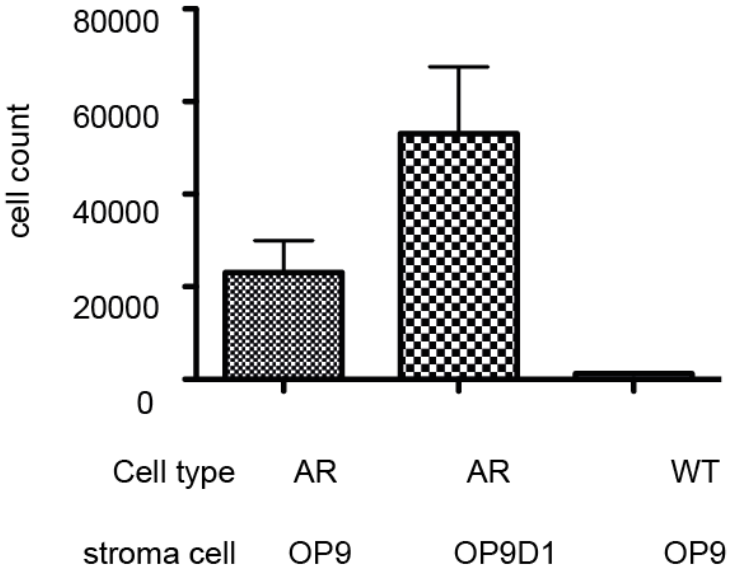


Figure S5
(A) $Sca1^{low}ckit^{low}CD19^{+}$ population harvested from "AR" and "WT" bone marrow was co-cultured on murine OP-9-delta 1 stroma cell line in the presence of FLT3L and IL7. Shown is the cell count at day 32 of co-culture with and without supplementation of γ-secretase inhibitor (N-[N-(3,5-Difluorophenacetyl)-L-alanyl]-S-phenylglycine t-butyl ester, DAPT) at a concentration of 10 µg/ml to co-culture. Controls were treated with the same concentration of DMSO carrier or without any supplement. The cells were replated every 4 to 5 days. (B) "AR" $Sca1^{low}ckit^{low}CD19^{+}$ population was co-cultured on OP-9 and OP-9–delta1 stroma cells. Indicated is the cell count at day 32 of co-culture.