Supporting Information

Roy et al. 10.1073/pnas.1211405109

SI Experimental Procedures

Sample Processing. Samples were red cell- and granulocyte-depleted by density-gradient separation and $CD34^+$ cell separation was carried out on freshly isolated mononuclear cells (MNC) using StemSep $CD34^+$ selection antibody mixture (Stem Cell Technologies) and Miltenyi MS magnetic columns (1). Postseparation $CD34^+$ purity was >95% in all cases. MNC and $CD34^+$ fractions were frozen in 90% FCS/10% (vol/vol) DMSO in liquid nitrogen. The study was approved by Hammersmith and Queen Charlotte's Hospital Research Ethics Committee (ref. 04/Q0406/145).

Flow Cytometric Analysis and Sorting. Cells (10^4 to 5×10^5) were stained with up to eight fluorophore-conjugated monoclonal antibodies from BD (Becton-Dickinson) unless otherwise stated: CD34PECy7 (8G12), CD38 Pacific Blue (HIT2: Exbio), CD45RA FITC (HI100), CD45RA APC (HI100), CD7PE (M-T701), CD2PECy5 (RPA-2.10), CD127 PE (hIL-7R-21), CD10FITC (W8E7), CD10PE Cy5 (HI10a), CD19APC (HIB19), CD123PE (9F5), CD33PE (WM-53; eBiosciences), CD14FITC (M5E2), CD15FITC (H198; eBiosciences), CD16FITC (eBioCB16; eBiosciences), CD11bPE (ICR F44), CD61FITC (RUU-PL7F12), CD135PE (BV10A4H2), CD235PE (AME-1; Invitogen), CD90 PECy5 (5E10), CD117PE (YB5B8), hCD45 Alexaflour 700 (F10-89-4; AbD Serotec), mCD45.1 FITC (A20). Samples were analyzed using a BD LSR Fortessa or FACSAria II (Becton Dickinson). Gates were set with unstained controls gating on viable cells using DAPI. Data were analyzed on FlowJo software (Tree Star).

Clonogenic Assays. Methylcellulose hematopoietic progenitor assays were performed using Methocult H4230 (Stem Cell Technologies), as previously described (1). Progenitor populations were sorted directly to achieve a final concentration of 100 cells/mL into Eppendorfs containing Methocult H4230 with cytokines [IL-3 20 ng/mL, IL-6 10 ng/mL, IL-11 10 ng/mL, stem cell factor (SCF) 10 ng/mL, FLT3 10 ng/mL, GM-CSF 50 ng/mL, thrombopoietin (TPO) 50 ng/mL (all Peprotech), and erythropoietin (EPO) 4 U/mL (R&D Systems)] and aliquoted into flat-bottomed 24-well plates (Becton Dickinson), incubated at 37 °C, and counted after 7, 14, 21, and 28 d. Individual colonies were identified morphologically and plucked under direct microscopy, resuspended in 1 mL Robosep buffer (Stem Cell Technologies) or into Methocult with cytokines, and vortexed for 1 min. Cells in Methocult were replated to check for secondary replating ability; cells in PBS were stained for FACS analysis or used for cytospin preparations. Cytospins on cultured fetal cells were carried out using a Shandon Cytospin 2 (Fisher Scientific). 10^4 cells were suspended in Robosep buffer (Stem Cell Technologies) and stained using May–Grunwald Giemsa after methanol fixation.

Immunohistochemistry. Archived fetal liver (FL) samples were used in accordance with the Human Tissue Act 2004 and the Hammersmith and Queen Charlotte's Hospital Research Ethics Committee (ref. 04/Q0406/145). FL were fixed in 10% (vol/vol) neutral buffered formalin, paraffin-embedded, and 5-µm sections prepared. De-waxed, rehydrated sections were subjected to antigen retrieval by microwaving in 0.01 M citrate (pH 6.0) buffer and treated with 3% (vol/vol) H_2O_2 (Sigma) in methanol for 30 min, and washed in PBS for 10 min, as previously described (2). Endogenous biotin was blocked using an avidin/biotin blocking kit (Vector Laboratories). Sections were incubated with normal goat serum [diluted 1:5 with 5% (wt/vol) BSA (Sigma)] for 30 min, incubated overnight with either CD20 (L26; DAKO) diluted 1.250; CD34 (QBEND10; Dako) diluted 1.100; CD42b (NCL-CD42b; Novocastra) diluted 1.200 in serum in a humidified chamber at 4 °C. After washing, sections were incubated for 30 min with goat anti-mouse biotinylated secondary antibody, washed 2× in PBS, and incubated for 30 min with Streptavidin–horseradish peroxidase at 1:1,000 (Dako). Visualization was performed using 3,3-diaminobenzidine tetrahydrochloride (Vector Laboratories) counterstained with hematoxylin (Vector Laboratories). For negative controls, primary antibody was replaced with normal goat serum. Images were captured using a Nikon Elipse E400 microscope and a Nikon DN100 digital camera.

***GATA1* Mutation Analysis.** DNA was extracted using QiAmp DNA Mini Kit (Qiagen). *GATA1* mutations were identified by WAVE dHPLC (Transgenomics) and PCR/cloning followed by sequencing. Exons 2 and 3 of *GATA1* were amplified by PCR using some or all of the following primers and conditions, as previously described (1, 3): exon 1 forward 5′-CAGGAAGACGCACATACACAGGA-3′ and reverse 5′-GATGGAGCTAGGGTTTGGCAGAT-3′ (amplicon 965 bp), annealing temperature 61 °C; exon 2 forward 5′-AAAGGAGGAAGAGGAGCAG-3′ and reverse 5′-AAGCTTCCAGCCATTTCTGA-3′ (amplicon 432 bp), annealing temperature 60 °C; Exon 2 forward 5′-GGATTTCTGTGTCTGAGG-3′ and reverse 5′-CCAACAGCACTCAGCCAA-3′ (amplicon 327 bp), annealing temperature 55 °C; Exon 2 forward 5′-GATGGGGGAGAGGGAGATAAGGT-3′ and reverse 5′-GGCAACCACCACATACT-TCCAGT-3′ (amplicon 1,103 bp), annealing temperature 61 °C; exon 3.1 forward 5′GGAACTTGGCCACCATGTTGG-3′ and reverse 5′-AGCCGCTCTGTCTTCAAAG-TCTC-3′ (amplicon 310 bp), annealing temperature 58 °C; exon 3.2 forward 5′-CTGGATGGAAAAGGCAGCACCA-3′ and reverse 5′-GAGCTAGGCTCAGCTCAGCT-TTAC-3′ (amplicon 304 bp), annealing temperature 58 °C. PCR conditions were: 5–10 min at 95 °C, 35 cycles of 1 min at 95 °C, 1 min at the annealing temperature given and 1 min at 72 °C. After the last cycle, an additional step of 5–10 min at 72 °C was performed. PCR products were verified by gel electrophoresis.

MS5 Stromal Cocultures for B-Lymphoid Differentiation. MS5 stromal cells were grown to confluency, passaged as described previously (4), and stromal layers prepared in 96- or 24-well plates, 24–48 h before sorted cells were plated. Next, 100 cells from the sorted $CD34^+$ subpopulation were seeded onto 80% confluent wells in MS5 medium (α-MEM; Invitrogen) with 10% (vol/vol) FCS, FLT3 (10 ng/mL), SCF (20 ng/mL), IL-2 (10 ng/mL), IL-7 (5 ng/mL), GM-CSF (20 ng/mL), and G-CSF (10 ng/mL) (Peprotech). Each well was regularly examined under an inverted microscope to detect proliferation of plated human hematopoietic cells. Cocultures were disaggregated by vigorous pipetting, passaged through a 70-µm filter to eliminate contaminating MS5 before replating on fresh MS5 cells every 4 d. FACS analysis was used to assess lymphoid differentiation weekly from day 7–21.

Statistical Comparison of Gene Expression Between Trisomy 21 and Normal FL Hematopoietic Stem Cells and Progenitors. ***Data description.*** Gene expression levels [for both trisomy 21 (T21) and normal samples] relative to GAPDH are smaller than 1.2354. If the relative gene expression level is smaller than 2.8×10^{-10}, exact values are not reported and we have to use the cutoff as an upper bound.

In the following, a sample of size N from this dataset will be denoted by $x = (y_1, \ldots, y_n, k)$ where $(y_1, \ldots y_n)$ are the values of the

sample larger than 2.8×10^{-10} and k is the number of values below the cutoff; i.e., $N = n + k$.

Bayesian analysis. A Bayesian approach was used to compare relative gene expression levels between T21 and normal FL populations because of the small sample sizes and censored data (see above). To do so, we assume that the log (base 2)-transformation of any sample x is normally distributed with unknown mean m and SD s. The Bayesian approach consists of determining the posterior probability distribution over the parameter space (i.e., all possible values for the mean and the SD). The posterior probability of a pair (m, s) is proportional to the likelihood of the data, $x = (y_1, \ldots, y_n, k)$, given the parameter (m, s) multiplied by the prior distribution $p(m, s)$ of the parameter:

$$p(m,s|x) \propto p(x|m,s)p(m,s)$$

The likelihood $p(x|m, s)$ is the probability that $[\log_2(y_1), \ldots, \log_2(y_n)]$ is an independent and identically distributed sample from a normal distribution with mean m and SD s, and we assume that this model, $\mathcal{N}(m, s)$, has generated k values smaller than the censor cutoff, $T = \log_2(2.8 \times 10^{-10})$; therefore,

$$p(x|m,s) = \left(\int_{-\infty}^{T} \phi(t,m,s)dt \right)^k \prod_{i=1}^{n} (\phi(\log_2(y_i)), m, s),$$

where $\phi(t, m, s)$ denotes the probability density function of a normal distribution with mean m and SD s. We use independent Jeffrey priors for both parameters m and s.

Significance of the difference in relative gene expression between two samples. To assess whether the relative gene expression level x of a sample is higher than the relative gene expression level x' of another sample, we compute a Bayesian P value, which is the probability under the posterior distribution of both samples that the mean m of the first sample is at least as large as the mean m' of the second one:

$$p = \iint dsds' \iint \mathbb{1}(m < m')p(m,s|x)p(m',s'|x')dmdm'.$$

Because this integral cannot be evaluated in closed form, importance sampling was used to evaluate the value of this quadruple integral.

1. Tunstall-Pedoe O, et al. (2008) Abnormalities in the myeloid progenitor compartment in Down syndrome fetal liver precede acquisition of GATA1 mutations. *Blood* 112(12): 4507–4511.
2. Naresh KN, et al. (2006) Optimal processing of bone marrow trephine biopsy: The Hammersmith Protocol. *J Clin Pathol* 59(9):903–911.
3. Alford KA, et al.; International Myeloid Leukemia-Down Syndrome Study Group (2011) Analysis of GATA1 mutations in Down syndrome transient myeloproliferative disorder and myeloid leukemia. *Blood* 118(8):2222–2238.
4. Goardon N, et al. (2011) Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia. *Cancer Cell* 19(1):138–152.

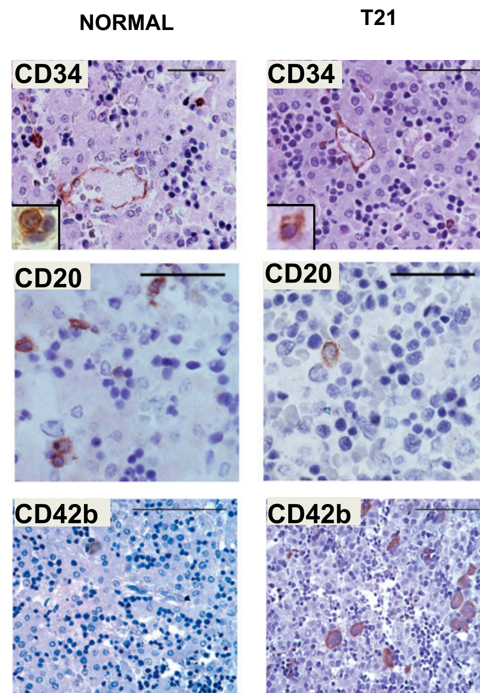

Fig. S1. Immunohistochemical staining of FL sections. Representative images from second trimester normal and T21 FL paraffin-embedded sections stained with antibodies directed against CD34, CD20, and CD42b (brown). (Scale bars, 100 μm.)

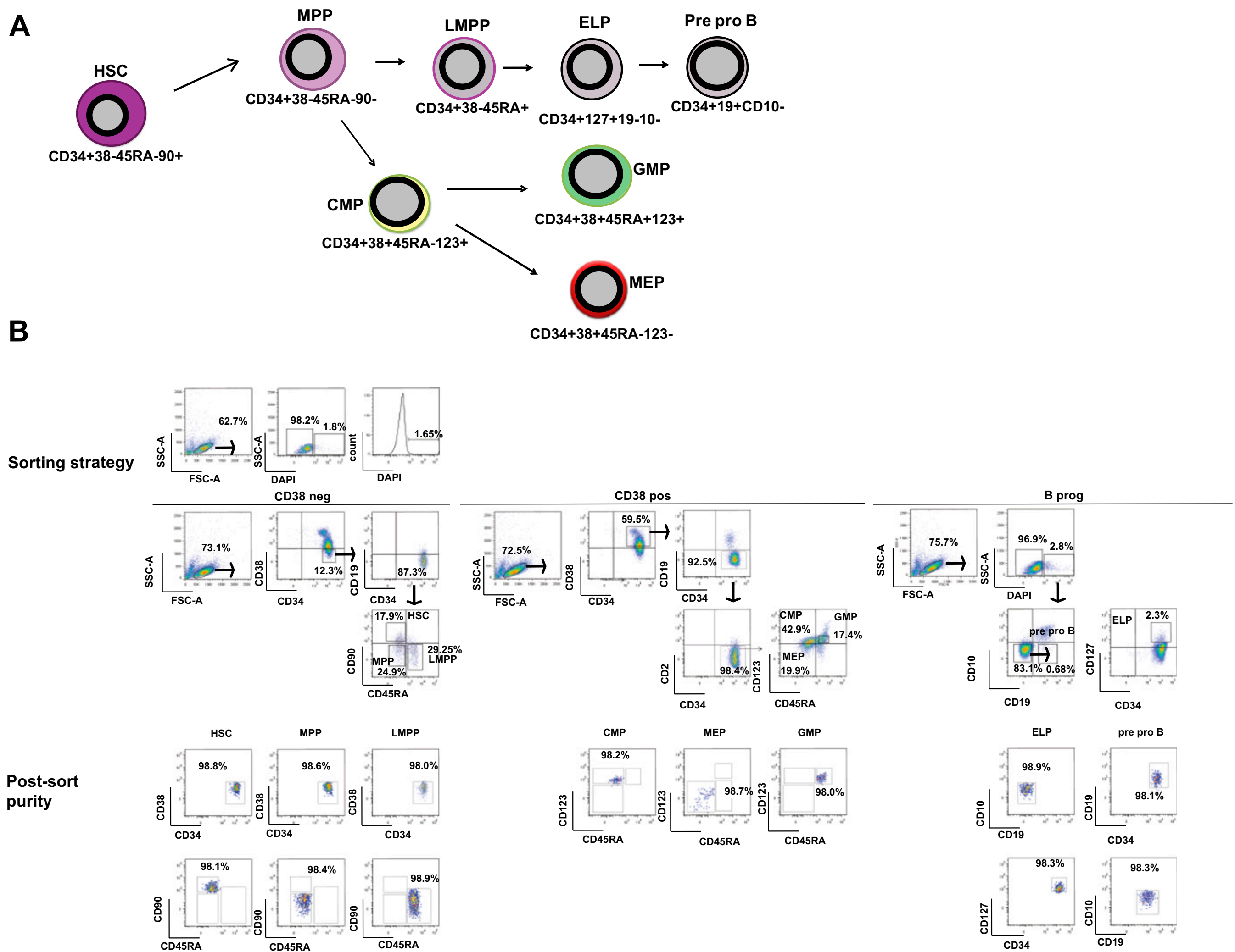

Fig. S2. Sorting strategy and postsort purity for fetal liver hematopoietic stem cells (HSC) and progenitor cells. (*A*) Lineage diagram showing immunophenotypes of sorted populations. (*B*) Strategy used for sorting HSC, multipotential progenitors (MPP), lymphoid-primed multipotential progenitors (LMPP), common myeloid progenitor (CMP), megakaryocyte-erythroid progenitor (MEP), granulocyte-macrophage progenitors (GMP), early lymphoid progenitor (ELP), and PreproB populations. Purity of the hematopoietic progenitor populations was analyzed on a FACSAria II. Cells were sorted to at least 98% purity.

Table S1. Clonogenic progenitor read out of T21 and normal FL HSC and progenitors

No. of colonies per 100 cells		HSC	MPP	LMPP	CMP	MEP	GMP
CFU-MK	Normal	0.2 ± 0.2	1.0 ± 0.45	0 ± 0	1.3 ± 0.5	3.0 ± 0.9	0 ± 0
	T21	9.5 ± 4.1*	5.7 ± 1.5	0 ± 0	14.5 ± 4**	16.8 ± 3.3**	0 ± 0
MkE	Normal	0.4 ± 0.4	1.4 ± 0.75	0 ± 0	1.9 ± 0.7	4.7 ± 1.3	0 ± 0
	T21	12.5 ± 5.6*	7.7 ± 2.2*	0 ± 0	20.8 ± 4.9**	18.8 ± 3.2**	0 ± 0
BFU-E	Normal	1.4 ± 0.9	10.6 ± 2.2	0 ± 0	8.1 ± 2.8	12.4 ± 3.7	0 ± 0
	T21	7.3 ± 1.6**	16.0 ± 0.8	0 ± 0	9.8 ± 2.7	17.8 ± 2.7	0 ± 0
Blast-E	Normal	3.8 ± 1.5	5.2 ± 2.8	0 ± 0	2.1 ± 1.4	2.1 ± 0.8	0 ± 0
	T21	11.5 ± 3.0*	2.7 ± 2.7	0 ± 0	6.0 ± 1.8*	6.8 ± 2.8*	0 ± 0
CFU-GM	Normal	4.0 ± 1.6	6.0 ± 0.6	4.3 ± 1.4	8.4 ± 2.2	0 ± 0	9.7 ± 1.7
	T21	3.8 ± 2.2	2.7 ± 0.8	5.3 ± 0.7	3.8 ± 1.3*	0 ± 0	6.8 ± 3.5
Blast-My	Normal	2.8 ± 1.7	2.2 ± 1.0	0 ± 0	1.1 ± 0.6	0 ± 0	0 ± 0
	T21	16.0 ± 5.9*	2.8 ± 1.4	2.6 ± 1.9	8.0 ± 4.8*	0 ± 0	0 ± 0
CFU-GEMM	Normal	0 ± 0	0.2 ± 0.2	0 ± 0	0.43 ± 0.4	0.43 ± 0.3	0 ± 0
	T21	0 ± 0	1.4 ± 0.7	0 ± 0	1.3 ± 0.5	1.0 ± 0.6	6.8 ± 3.5

Flow-sorted FL HSC or progenitor cells from normal FL (n = 8) and T21 FL (n = 5) were plated in Methocult 4230 with IL-3, IL-6, IL-11, SCF, FLT3, GM-CSF, TPO, and EPO. *$P < 0.05$; **$P < 0.02$ (T21 compared with normal FL). BFU-E, erythroid blast-forming unit; MkE, MK-erythroid.

Table S2. Blast cell colonies in normal cord blood and adult bone marrow (BM)

No. of colonies per 100 cells		HSC	MPP	LMPP	CMP	MEP	GMP
Blast-My	Cord blood (n = 4)	7.7 ± 2.4	9.0 ± 3.2	0 ± 0	2.5 ± 0.6	0 ± 0	0 ± 0
	Adult BM (n = 3)	0.3 ± 0.4	1.6 ± 0	0 ± 0	1.5 ± 0.8	0 ± 0	0 ± 0
Blast-E	Cord blood (n = 4)	0 ± 0	4.0 ± 2	0 ± 0	2.0 ± 1.1	4.2 ± 1	0 ± 0
	Adult BM (n = 3)	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0

Table S3. Taqman gene expression assay ID from Applied Biosystems used for quantitative PCR

Gene name	Assay ID
MPL	Hs00180489_m1
IL3Rα	Hs00608141_m1
IKZF1	Hs00172991_m1
RUNX1	Hs00231079_m1
ERG	Hs01554635_m1
NOTCH1	Hs01062011_m1
FLT3	Hs00174690_m1
CRLF2	Hs00845692_m1
E2A	Hs00413032_m1
ETS1	Hs00901425_m1
MEF2C	Hs00231149_m1
GATA3	Hs00231122_m1
HES1	Hs00172878_m1
DYRK1A	Hs00176369_m1
EBF1	Hs00395519_m1
IGH@	Hs00378230_g1
PAX5	Hs00277134_m1
IL7R	Hs00902334_m1
CD19	Hs00174333_m1
GABPA	Hs01022023_m1
EPOR	Hs00181092_m1
SCL	Hs01097987_m1
GATA2	Hs00231119_m1
GATA1	Hs01085823_m1
VWF	Hs00169795_m1
KLF1	Hs00610592_m1
CEBPA	Hs00269972_s1
PU.1	Hs02786711_m1
CSF3R	Hs01114427_m1
CSF2R	Hs00538900_m1
CSF1R	Hs00911250_m1