Musashi–2 attenuates AHR signalling to expand human haematopoietic stem cells

Stefan Rentas1, Nicholas T. Holzapfel1a, Muluken S. Belew1a, Gabriel A. Pratt2,3a, Veronique Voisin4, Brian T. Wilhelm5, Gary D. Bader4, Gene W. Yeo2,3,6 & Kristin J. Hope1

Umbilical cord blood-derived haematopoietic stem cells (HSCs) are essential for many life-saving regenerative therapies. However, despite their advantages for transplantation, their clinical use is restricted because HSCs in cord blood are found only in small numbers1. Small molecules that enhance haematopoietic stem and progenitor cell (HSPC) expansion in culture have been identified2,3, but in many cases their mechanisms of action or the nature of the pathways they impinge upon are poorly understood. A greater understanding of the molecular circuitry that underpins the self-renewal of human HSCs will facilitate the development of targeted strategies that expand HSCs for regenerative therapies. Whereas transcription factor networks have been shown to influence the self-renewal and lineage decisions of human HSCs2,5, the post-transcriptional mechanisms that guide HSC fate have not been closely investigated. Here we show that overexpression of the RNA-binding protein Musashi-2 (MSI2) induces multiple pro-self-renewal phenotypes, including a 17-fold increase in short-term repopulating cells and a net 23-fold ex vivo expansion of long-term repopulating HSCs. By performing a global analysis of MSI2–RNA interactions, we show that MSI2 directly attenuates aryl hydrocarbon receptor (AHR) signalling through post-transcriptional downregulation of canonical AHR pathway components in cord blood HSPCs. Our study gives mechanistic insight into RNA networks controlled by RNA-binding proteins that underlie self-renewal and provides evidence that manipulating such networks ex vivo can enhance the regenerative potential of human HSCs.

Control of translation by RNA-binding proteins in human HSCs and its potential to regulate HSC self-renewal remain underexplored. MSI2 is known to regulate mouse HSCs6–8 and has been predicted to influence mRNA translation2, so we investigated the role of MSI2 in post-transcriptionally controlling self-renewal of human HSPCs. The expression of MSI2 mRNA was elevated in primitive cord blood HSPCs and decreased during differentiation, whereas the MSI2 parologue, MSJ1, was not expressed (Extended Data Fig. 1a–f). Lentiviral overexpression of MSI2 resulted in a 1.5-fold increase in colony-forming units (CFU) relative to control cells, principally due to a 3.7-fold increase in the most primitive CFU-granulocyte erythrocyte monocyte macrokaryocyte (GEMM) colony type (Extended Data Fig. 2a and Fig. 1a). Remarkably, 100% of MSI2-overexpressing CFU-GEMMs generated secondary colonies compared to only 40% of control CFU-GEMMs. In addition, MSI2 overexpression yielded three times as many colonies per re-seeded CFU-GEMM (Fig. 1b, c and Extended Data Fig. 2b). During in vitro culture, MSI2-overexpressing cells were 2.3- and 6-fold more abundant than control cells at the 7- and 21-day time points, respectively (Extended Data Fig. 2c, d). Moreover, after 7 days in culture, MSI2-overexpressing cells showed a cumulative 9.3-fold increase in colony-forming cells in the absence of changes in cell cycling or death (Extended Data Fig. 2e–h). Together, our data demonstrate that enforced expression of MSI2 has potent self-renewal-inducing effects on early progenitors and promotes their in vitro expansion.

Short-term repopulating cells (STRCs) produce a transient multilineage graft in non-obese diabetic (NOD)/SCID Il2rgnull (NSG) mice10, and in patients these cells reconstitute granulocytes and platelets that are essential for preventing post-transplantation infection and bleeding1. MSI2-overexpressing STRGs yielded 1.8-fold more primitive CD34+ cells post-infection and a 17-fold increase in functional STRCs relative to control STRGs, as determined by limiting dilution analysis (LDA) of human chimaerism in mice 3 weeks after transplantation (Fig. 1d–f and Extended Data Fig. 3a, b). Furthermore, at a protracted engraftment

**Figure 1** MSI2 overexpression enhances in vitro cord blood progenitor activity and increases the number of STRCs. a. CFU output from transduced Lin− cord blood (n = 9 control and 10 MSI2-overexpressing (MSI2) cultures from 5 experiments). b. CFU-GEMM secondary CFU repopulating potential (n = 24 control and 30 MSI2-overexpressing cultures from 2 experiments) and images of primary GEMMs (scale bar, 200 μm). c. Number of secondary colonies per replated CFU-GEMM from b. d. CD34 expression in STRCs before transplantation (n = 3 experiments). e. Human chimaerism at 3 weeks in mice transplanted with varying doses of transduced STRCs. Dashed line indicates engraftment cut-off (n = 3 experiments). f. STRC frequency determined by LDA from e. Dashed lines indicate 95% confidence intervals. Data shown as mean ± s.e.m. Unpaired t-test, *P < 0.05; **P < 0.01; ***P < 0.001.
readout time of 6.5 weeks at non-limiting transplant doses, 100% of mice transplanted with MSI2-overexpressing STRCs were engrafted compared to only 50% of mice transplanted with control STRCs, indicating that MSI2 overexpression extended the duration of STRC-mediated engraftment (Extended Data Fig. 3c).

We next explored the effect of short hairpin (sh)RNA-induced MSI2 knockdown on HSPC function. MSI2 knockdown did not alter the clonogenic potential of HSPCs but did decrease CFU replating threefold (Extended Data Fig. 4a–c). In more primitive culture-initiating cells, MSI2 knockdown significantly decreased cell numbers over culture (Extended Data Fig. 4d, e) independent of increased death or altered cell cycling (data not shown). Upon transplantation, engrafted MSI2 knockdown GFP+ cells showed no evidence of lineage skewing, but the frequency of cells was markedly reduced relative to the percentage of GFP+ cells initially transplanted (Extended Data Fig. 4f–h). Combined, our in vitro and in vivo data show that MSI2 knockdown reduces self-renewal in early progenitors and HSCs.

To characterize the earliest transcriptional changes induced by modulating MSI2 expression, we performed RNA sequencing (RNA-seq) on CD34+ MSI2-overexpressing and knockdown cells immediately after transduction (Supplementary Tables 1 and 2). MSI2 overexpression-induced transcriptional changes showed an inverse correlation with those induced by MSI2 knockdown, suggesting that overexpression and knockdown had opposite effects (Extended Data Fig. 5a). When compared to transcriptome data from 38 human haematopoietic cell subpopulations, genes that were significantly upregulated by MSI2 overexpression and downregulated upon MSI2 knockdown were exclusively enriched in those highly expressed in HSCs and other primitive CD34+ populations (Extended Data Fig. 5b).

As MSI2 overexpression conferred an HSC gene expression program, we hypothesized that it could facilitate HSC expansion ex vivo. MSI2 overexpression induced a fourfold increase in CD34+CD133+ phenotypic HSCs relative to control cells after 7 days of culture (Fig. 2a). We next performed an LDA to define functional HSC frequency before (day 3 post-transduction, D3) and after 7 days of ex vivo culture (day 10, D10; Extended Data Fig. 6a). Mice transplanted with HSCs on D3 displayed no altered engraftment as a result of MSI2 overexpression; however, recipients of MSI2-overexpressing D10-expanded cells displayed multiple phenotypes of enhanced reconstitution relative to recipients of control cells, including a twofold increase in bone marrow GFP+ levels without changes to lineage output, an increase in the proportion of GFP+ cells within the human graft relative to pre-transplant D10 levels, an increase in GFP mean fluorescence intensity and enrichment of CD34 expression in GFP+ cells (Fig. 2b, c and Extended Data Fig. 6b–h). As the lentiviral construct design ensures that levels of GFP mirror those of MSI2, these findings indicate that high levels of MSI2 impart enhanced competitiveness and are conducive to in vivo HSPC activity. Importantly, D10 MSI2-overexpressing cultures contained more CD34+CD133+ cells before transplantation than did control cultures (Extended Data Fig. 6i) and, accordingly, the HSC frequency in D10 MSI2-overexpressing cultures was increased twofold relative to that in D3 MSI2-overexpressing cultures. By contrast, control cultures displayed a threefold decrease in HSC frequency. These results demonstrate that MSI2 overexpression ex vivo facilitated a net sixfold increase in HSC frequency relative to control cultures (Fig. 2g, h and Supplementary Tables 3, 4).

Secondary LDA transplants were performed to explore fully the effects of MSI2 overexpression and culturing on self-renewal and long-term HSCs (LT-HSCs). Robust engraftment with MSI2-overexpressing cells did not induce altered myelo-lymphopoiesis or leukemic development (Fig. 2e). Secondary LDA measurements revealed that the percentage of GFP+ cells in the bone marrow was 4.6-fold higher after transplantation of MSI2-overexpressing cells than after control transplantation, and LT-HSC frequency was 3.5-fold higher (Fig. 2d, f and Supplementary Table 5). The increase in LT-HSC frequency corresponds to MSI2-overexpressing GFP+ HSCs having expanded in primary mice 2.4-fold over input as compared to a 1.5-fold decrease for control HSCs (Fig. 2g). The level of in vivo expansion induced by MSI2 overexpression reflects the behaviour of uncultured HSCs, which undergo similarly controlled expansion during passage in mice4,11,12. Finally, when accounting for the total change in GFP+ HSCs upon ex vivo culture, MSI2 overexpression induced a cumulative 23-fold expansion of secondary LT-HSCs relative to control (Fig. 2g, h), indicating that elevated MSI2 expression provides a considerable self-renewal advantage to functional HSCs during ex vivo culture.

To gain mechanistic insight into this process, we examined genes that were differentially expressed in MSI2-overexpressing cells and found that the gene encoding cytochrome P450 1B1 oxidase (CYP1B1), an effector of AHR signalling5, was among the most repressed (Supplementary Table 1). Pathway analysis revealed that many predicted targets of AHR were enriched in the gene sets that were downregulated by MSI2 overexpression (Fig. 3a) and upregulated by MSI2 knockdown (Extended Data Fig. 7a, b). Binding of
percentage of CD34 expression expands HSPCs by attenuating AHR signalling (Fig. 3d and e). Furthermore, FICZ-treated MSI2-overexpressing cultures displayed greater losses of phenotypic HSPCs compared to treated controls, which showed no change (Extended Data Fig. 8d, e). Together, these results show that agonist-induced restoration of AHR activity reduces the self-renewal-promoting effects of MSI2 overexpression and strongly supports the idea that MSI2 overexpression promotes HSPC expansion through downregulation of AHR signalling.

To identify key RNA targets that underlie MSI2 function, we analysed global MSI2 protein–RNA interactions using cross-linking immunoprecipitation followed by sequencing (CLIP–seq)15 (Extended Data Fig. 9a, b). Replicates were highly correlated via gene RPKMs (reads per kilobase of transcript per million mapped reads) and 5,552 protein-coding genes were bound in both replicates (Extended Data Fig. 9c and Fig. 4a, b). Within the top 40% of reproducible clusters, MSI2 bound predominantly to the 3' untranslated regions (3'UTRs) of mature mRNAs (Fig. 4c). Importantly, 9% of annotated protein-coding gene mRNAs were reproducible MSI2 targets, compared to 0.2% of long non-coding RNAs (Extended Data Fig. 9d), suggesting that MSI2 controls the stability or translation of coding mRNAs. Motif analysis identified a consensus pentamer (U/G)UAGU resembling the known mouse Msi1-binding sequence9,16 within binding sites in all genic regions; additionally, MSI2-binding sites were generally significantly more conserved than background and tended to occur after the stop codon (Fig. 4d and Extended Data Fig. 9e–h). The presence of MSI2 binding sites within Msi1 targets9 across species indicates that Musashi proteins may bind the same genes through 3'UTR-encoded motifs (Extended Data Fig. 9i). Finally, gene ontology analysis revealed 186 biological processes categories (Supplementary Table 6), among the most significant of which were electron transport, oestrogen receptor signalling regulation and metabolism of small molecules, all processes known to be transcriptionally influenced by AHR signalling17.

Among the top 2% of enriched CLIP–seq targets (Supplementary Table 7) were the 3'UTRs of the genes for two AHR pathway components: heat shock protein 90 (HSP90) and CYP1B1. Each exhibited multiple MSI2-binding motifs correlating with overlapping clusters of CLIP–seq reads (Fig. 4e and Extended Data Fig. 10a). To investigate the ability of MSI2 to post-transcriptionally regulate these genes during HSPC expansion, we looked for instances of uncoupled transcript and protein expression. HSP90 displayed uncoupling of transcript (1.6-fold up) and protein (1.6-fold down) expression early in culture, but after 7 days showed further upregulated transcript expression (2.5-fold) and variable protein levels (Fig. 4f and Extended Data Fig. 10b). As AHR–HSP90 binding is essential for ligand-dependent transcriptional activity13, downregulation of HSP90 protein at the outset of HSPC culture would be expected to reduce latent AHR complex formation and attenuate AHR signalling (Fig. 3a). Indeed, CYP1B1 transcript and protein expression displayed twofold reductions early in culture, consistent with decreased AHR pathway activity; however, at day 7, CYP1B1 transcripts were upregulated 1.7-fold and uncoupled from protein expression, which was downregulated twofold (Fig. 4g and Extended Data Fig. 10c). To test whether MSI2 directly mediates post-transcriptional repression of these targets, the 3'UTRs of CYP1B1 and HSP90 were coupled to luciferase. MSI2 overexpression induced significant reductions in luciferase signal from both reporters, and this effect was mitigated when the core CLIP–seq-identified UAG motifs were mutated (Extended Data Fig. 10d, e). As MSI2 overexpression-mediated post-transcriptional downregulation of the AHR pathway converged on CYP1B1 protein repression throughout culture, we explored the effects on HSPCs of inhibiting CYP1B1 independently with (E)-2,3',4,5'-tetramethoxystilbene (TMS). During culture, TMS

![Figure 3](image-url)
which to enhance the regenerative potential of not only human HSCs but also other stem-cell types.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 19 June 2015; accepted 15 March 2016.

1. Miller, P. H., Knapp, D. J. & Eaves, C. J. Heterogeneity in hematopoietic stem cell populations: implications for transplantation. Curr. Opin. Hematol. 20, 257–264 (2013).
2. Boitano, A. E. et al. Aryl hydrocarbon receptor antagonists promote the expansion of human hematopoietic stem cells. Science 329, 1345–1348 (2010).
3. Fares, I. et al. Pyrimidinolide derivatives are agonists of human hematopoietic stem cell self-renewal. Science 345, 1509–1512 (2014).
4. Noresenher, N. et al. Densely interconnected transcriptional circuits control cell states in human hematopoiesis. Cell 144, 296–309 (2011).
5. Laurenti, E. et al. The transcriptional architecture of early human hematopoiesis identifies multilevel control of lymphoid commitment. Nature Immunol. 14, 756–763 (2013).
6. Hope, K. J. et al. An RNAi screen identifies Msi2 and Proxl as having opposite roles in the regulation of hematopoietic stem cell activity. Cell Stem Cell 7, 101–113 (2010).
7. de Andrés-Aguayo, L. et al. Musashi 2 is a regulator of the HSC compartment identified by a retroviral insertion screen and knockout mice. Blood 118, 554–564 (2011).
8. Park, S. M. et al. Musashi-2 controls cell fate, lineage bias, and TGF-β signaling in HSCs. J. Exp. Med. 211, 71–87 (2014).
9. Ohyama, T. et al. Structure of Musashi1 in a complex with target RNA: the role of aromatic stacking interactions. Nucleic Acids Res. 40, 3218–3231 (2012).
10. Glimm, H. et al. Previously undetected human hematopoietic cell populations with short-terminpopulating activity selectively engraft NOD/SCID–1,2 microglobulin-null mice. J. Clin. Invest. 107, 199–206 (2001).
11. Cashman, J. D. & Eaves, C. J. Human growth factor-enhanced regeneration of transplantable human hematopoietic stem cells in nonobese diabetic/severe combined immunodeficient mice. Blood 93, 481–487 (1999).
12. Holyoake, T. L., Nicolini, F. E. & Eaves, C. J. Functional differences between transplantable human hematopoietic stem cells from fetal liver, cord blood, and adult marrow. Exp. Hematol. 27, 1418–1427 (1999).
13. Mimura, J. & Fuji-Kiyuriyama, Y. Functional role of AR in the expression of toxic effects by TCD3. Biochim. Biophys. Acta 1619, 263–268 (2003).
14. Lo, R. & Matthews, J. High-resolution genome-wide mapping of AHR and ARNT binding sites by ChiP-seq, Toxicol. Sci. 130, 349–361 (2012).
15. Yeo, G. W. et al. An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. Nature Struct. Mol. Biol. 16, 130–137 (2009).
16. Katz, Y. et al. Musashi proteins are post-translational regulators of the epithelial-luminal cell state. Elite 3, e03915 (2014).
17. Tijet, N. et al. Aryl hydrocarbon receptor regulates distinct dioxin-dependent and dioxin-independent gene batteries. Mol. Pharmacol. 69, 140–153 (2006).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank E. Lechman and P. Van Galen for experimental advice and for providing H1 and pSMALB vectors. The MA overexpression vector was a gift from L. Naldini. We also thank the SCC-RI core flow cytometry staff, the Obstetrics and Gynecology Unit at McMaster Children’s Hospital for cord blood, D. Bode and M. Bhatia for critical assessment of this work and all members of the Hope laboratory for support and advice.

This work was supported by an Ontario Institute for Cancer Research New Investigator Award (IA-033), an Ontario Institute for Cancer Research Cancer Stem Cell Program Team Grant (FCSC.005) and a Canadian Institutes of Health Research (MOP-126030) grant to K.J.H. N.T.H was supported in part by a CIHR MD/PhD Studentship. M.S.B. was supported by an NSERC Alexander Graham Bell Doctoral Fellowship, S.R. is supported by a Canadian Blood Services Graduate Fellowship, M.W.Y. was supported by the SCC-RI core flow cytometry staff, the Obstetrics and Gynecology Unit at McMaster Children’s Hospital for cord blood, D. Bode and M. Bhatia for critical assessment of this work and all members of the Hope laboratory for support and advice.

This work was supported by an Ontario Institute for Cancer Research New Investigator Award (IA-033), an Ontario Institute for Cancer Research Cancer Stem Cell Program Team Grant (FCSC.005) and a Canadian Institutes of Health Research (MOP-126030) grant to K.J.H. N.T.H was supported in part by a CIHR MD/PhD Studentship. M.S.B. was supported by an NSERC Alexander Graham Bell Doctoral Fellowship, S.R. is supported by a Canadian Blood Services Graduate Fellowship, M.W.Y. was supported by the SCC-RI core flow cytometry staff, the Obstetrics and Gynecology Unit at McMaster Children’s Hospital for cord blood, D. Bode and M. Bhatia for critical assessment of this work and all members of the Hope laboratory for support and advice.

The views expressed herein do not necessarily represent the view of the federal government of Canada. This work was partially supported by grants from the National Institute of Health (HG004659 and NS075449) and the California Institute of Regenerative Medicine (PB3-05219) to G.W.Y. which was supported by a National Science Graduate Fellowship, G.W.Y. is an Alfred P. Sloan Research Fellow. We thank the UCSD Institute for Genomic Medicine’s Genomics Center for providing access to high-throughput sequencing facilities.

Author Contributions S.R. designed and performed experiments, analysed data and wrote the manuscript. N.T.H. constructed CLIP–seq libraries. M.S.B. helped perform cord blood experiments. G.A.P. and G.W.Y. advised on CLIP–seq library construction, performed CLIP–seq bioinformatic analyses and wrote the manuscript. B.T.W. performed RNA-seq analyses. V.V. and G.D.B performed RNA-seq bioinformatic analyses. K.J.H. conceived the project, supervised the study, analysed data, interpreted results and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.W.Y. (geneyeo@ucsd.edu) or K.J.H. (kristin@mcmaster.ca).
METHODS

Mice. NOD/SCID IL2rgnull mice (Jackson Laboratory) were bred and maintained in the Stem Cell Unit animal barrier facility at McMaster University. All procedures were approved by the Animal Research Ethics Board at McMaster University.

Isolation of primitive human haematopoietic cells and flow cytometry. All patient samples were obtained with informed consent and with the approval of local human subject research ethics boards at McMaster University. Human umbilical cord blood mononuclear cells were collected by centrifugation with Ficoll-Paque Plus (GE), followed by red blood cell lysis with ammonium chloride (StemCell Technologies). Cells were then incubated with a cocktail of lineage-specific antibodies (CD2, CD3, CD11b, CD11c, CD14, CD16, CD19, CD24, CD56, CD61, CD66b, and GlyA; StemCell Technologies) for negative selection of Lin− cells using an EasySep immunomagnetic column (StemCell Technologies). Live cells were discriminated on the basis of cell size, granularity and, as needed, absence of viability dye 7-AAD (BD Biosciences) uptake. All flow cytometry analysis was performed using a BD LSR II instrument (BD Biosciences). Data acquisition was conducted using BD FACSDiva software (BD Biosciences) and analysis was performed using FlowJo software (Tree Star).

HSPC sorting and qRT–PCR analysis. To quantify MSI2 expression in human HSPCs, Lin− cord blood cells were stained with the appropriate antibody combinations to resolve HSC (CD34−HSPCs, Lin−; CD34+, Lin+) and EP (CD34+, Lin+) fractions as similarly described previously,8,9 with all antibodies from BD Biosciences: CD45RA (HI100), CD45 (SE10), CD34 (SBI), CD13 (H7B) and CD71 (M-A712). Cell viability was assessed using the viability dye 7AAD (BD Biosciences). All cell subsets were isolated using a BD FACSaria II cell sorter (BD Biosciences) or a MoFlo XDP cell sorter (Beckman Coulter). HemaExplorer software was used to confirm MSI2 expression in human HSPCs across the hierarchy. For all qRT–PCR determinations total cellular RNA was isolated with TRIzol LS reagent according to the manufacturer’s instructions (Invitrogen) and cDNA was synthesized using the qScript cDNA Synthesis Kit (Quanta Biosciences). qRT–PCR was performed in triplicate with PerfeCTa qPCR SuperMix Low ROX (Quanta BioSciences) with gene-specific probes (Universal Probe Library (UPL), Roche) and primers: MS1 UPL-26, F-GGCAAAAGAAGATCGG, R-CCTGGAGATGCGCGGACA; HSP90 UPL-46, F-GGGCAAACAGCTACAAGGA, R-CTGGGCTCTGTGGTCTTTCCT; CYPB1 UPL-20, F-ACGTCACCTCCATCTAC, R-CTTCAGTCTGGCCTACGAGGA; GAPDH UPL-60, F-AGGACCACGTGTCAGACGAC, R-GGCCAAATCGACATTB (UPL Set Reference Gene Assays, Roche). The mRNA content of samples compared by qRT–PCR was normalized based on the amplification of GAPDH or ACTB.

Lentivirus production and western blot validation. MS1 shRNAs were designed with the Dharmacon algorithm (http://www.dharmacon.com). Predicted sequences were synthesized as complimentary oligonucleotides, annealed and cloned downstream of the H1 promoter of the modified cptt-PGK-EGFP-ires-pac-WPRE lentiviral expression vector19. Sequences for the MS1 targeting and control RFP targeting shRNAs were as follows: shMS1, 5′-GACAGCATCCACTACGAGA-3′; shRFP, 5′-GTGGGAAAGGGCTGTGATGATGAC-3′; MS1 UPL-60, F-AGGACCACGTGTCAGACGAC, R-GGCCAAATCGACATTB (UPL Set Reference Gene Assays, Roche). The mRNA content of samples compared by qRT–PCR was normalized based on the amplification of GAPDH or ACTB.

Lentivirus production and western blot validation. MS1 shRNAs were designed with the Dharmacon algorithm (http://www.dharmacon.com). Predicted sequences were synthesized as complimentary oligonucleotides, annealed and cloned downsteam of the H1 promoter of the modified cptt-PGK-EGFP-ires-pac-WPRE lentiviral expression vector19. Sequences for the MS1 targeting and control RFP targeting shRNAs were as follows: shMS1, 5′-GACAGCATCCACTACGAGA-3′; shRFP, 5′-GTGGGAAAGGGCTGTGATGATGAC-3′; MS1 UPL-60, F-AGGACCACGTGTCAGACGAC, R-GGCCAAATCGACATTB (UPL Set Reference Gene Assays, Roche). The mRNA content of samples compared by qRT–PCR was normalized based on the amplification of GAPDH or ACTB.

Intracellular flow cytometry. Lin− cord blood cells were initially stained with anti-CD34 PE (S81) and anti-CD34 APC (H7B) antibodies (BD Biosciences) followed by secondary antibodies and cell permeabilization with a Cytofix/Cytoperm kit (BD Biosciences). Intracellular FSC−CD34+ cells were subsequently stained with mouse anti-MSI2 rabbit IgG (EP1305Y, Epitomics) and Alexa-488 goat anti-rabbit IgG antibody (Invitrogen).

RNA-seq data processing. CD34− cells were transduced with an MS1-overexpression or MS1-knockdown lentivirus along with their corresponding controls and sorted for GFP expression 3 days later. Transductions for MS1-overexpression or knockdown were each performed on two independent cord blood samples. Total RNA from transduced cells (>×10^6) was isolated using TRIzol LS as recommended by the manufacturer (Invitrogen), and then further purified using RNeasy columns (Qiagen). Sample quality was assessed using Bioanalyzer RNA Nano chips (Agilent). Paired-end, barcoded RNA-seq sequencing libraries were generated using the TruSeq RNA Sample Prep Kit (Illumina) according to the manufacturer’s protocols starting from 1 μg total RNA. The quality of library generation was then assessed using a BioAnalyzer platform (Agilent) and Illumina MiSeq QC run was performed or quantified by qPCR using KAPA quantification kit (KAPA Biosystems). Sequencing was performed using an Illumina HiSeq2000 using TruSeq SBS v3 chemistry at the Institute for Research in Immunology and Cancer’s Genomics Platform (University of Montreal) with cluster density targeted at 750,000 clusters per mm² and paired-end 2 × 100-bp read lengths. For each sample, 90–95 million reads were produced and mapped to the hg19 (GRCh37) human genome assembly using CASAVA (version 1.8). Raw counts generated by CASAVA were then aggregated into a single data file using the DataPrep module of the Broad Institute’s Genome Analysis Toolkit (GATK) and reads mapped to the reference genome were extracted. This normalized dataset was then passed to DESeq analysis using the edgeR R package (Ref. 32, 33). Differentially expressed genes were filtered using a log2 fold change threshold of 0.5 and FDR < 0.05. Statistical analysis was performed using DESeq2 with the Benjamini–Hochberg method. The fold change of log2 expression ratio was calculated for each gene and the list was ranked by the magnitude of fold change.
significantly differentially expressed genes. RNA-seq data have been deposited in NCBI's Gene Expression Omnibus (GEO) and are accessible through GEO Series accession number GSE70685.

**GSEA and iRegulon AHR target prediction.** iRegulon was used to retrieve the top 100 AHR predicted targets with a minimum occurrence count threshold of 5. The data were analysed using GSEA with ranked data as input with parameters set to 2.000 gene-set permutations.

**GSEA and StemRegenin 1 (SR1) gene sets.** The GEO dataset GSE28359, which contains Affymetrix Human Genome U133 Plus 2.0 Array gene expression data for CD34+ cells treated with SR1 at 30 nM, 100 nM, 300 nM, and 1,000 nM was used to obtain lists of genes differentially expressed in the treated samples compared to the control ones (nM3). Data were background corrected using Robust Multi-Array Average (RMA) and quantile normalized using the expresso() function of the affy Bioconductor package (affy_3.6.1, R 3.0.1). Lists of genes were created from the 150 top upregulated and downregulated genes from the SR1-treated samples at each dose compared to the non-treated samples (0 nM). The data were analysed using GSEA with ranked data as input with parameters set to 2.000 gene-set permutations. The normalized enrichment score (NES) and false discovery rate (FDR) were calculated for each comparison.

**Differentiation Map of Haematopoiesis (DMAP) population comparisons.** The GEO data set GSE24759, which contains Affymetrix GeneChip HT−HG_U133A Data Early Access Array gene expression data for 38 distinct haematopoietic cell states, was compared to the MS2 overexpression and knockdown data. GSE24759 data were background corrected using Robust Multi-Array Average (RMA), quantile normalized using the expresso() function of the affy Bioconductor package (affy_3.6.1, R 3.0.1), batch corrected using the ComBat() function of the sva package (sva_3.6.0) and scaled using the standard score. Bar graphs were calculated by creating for significantly differentially expressed genes the number of scaled data that were above (>0) or below (<0) the mean for each population. Percentages indicating how long the observed value (set of up- or downregulated genes) was better represented in that population than random values were calculated from 1,000 trials.

**AHR ChIP−seq comparison with downregulated gene sets.** A unique list of genes close to AHR-bound regions previously identified from TCD−treated MCF7 ChIP−seq data was used to calculate the overlap with genes showing >1.5-fold downregulation in response to treatment with UM171 (35 nM) or SR1 (500 nM) relative to DMSO−treated samples. The three gene lists were uploaded into oPOSSUM-3 and the AHR:ARNT transcription factor binding site profile was used with the matrix score threshold set at 80% to analyse the region 1,500bp upstream and 1,000bp downstream of the transcription start site. The percentage of downregulated genes with AHR-bound regions was then plotted for each gene set. P values were generated with Fisher's exact test for comparisons between gene lists.

**oPOSSUM analysis for promoter AHR binding sites in downregulated gene sets.** AHR transcription factor binding sites in downregulated gene sets were identified with oPOSSUM-3. Genes showing >1.5-fold downregulation in response to treatment with UM171 (35 nM) or SR1 (500 nM) relative to DMSO−treated samples were used along with significantly downregulated genes (FDR < 0.05) with EdgeR-analysed MS2-overexpressing versus control-treated samples. The three gene lists were uploaded into oPOSSUM-3 and the AHR:ARNT transcription factor binding site profile was used with the matrix score threshold set at 80% to analyse the region 1,500bp upstream and 1,000bp downstream of the transcription start site. The percentage of downregulated genes with AHR-binding sites in their promoters was then plotted for each gene set. Fisher's exact test was used to identify significant overrepresentation of AHR-binding sites in gene lists relative to background.

**Analysis for human chimaerism.** Eight- to 12-week-old male or female NSG mice were sublethally irradiated (315 cGy) one day before intrafemoral injection with transduced cells carried in IMDM 1% FBS at 25 μl per mouse. Injected mice were analysed for human haematopoietic engraftment 12–14 weeks after transplantation or at 3 and 6.5 weeks for STRC experiments. Mice were sacrificed by CO2 asphyxiation and blood sample were collected for immunophenotyping and quantification of engraftment. Aortic root and femur bone marrow were aspirated without dilation and processed for flow cytometric analysis. Bone marrow and spleen cells were removed and stained with CD45 vis and spleen were removed and bones were crushed with a mortar and pestle then filtered into single-cell suspensions. Bone marrow and spleen cells were blocked with mouse Fc block (BD Biosciences) and human IgG (Sigma) and then stained with fluorochrome-conjugated antibodies specific to human haematopoietic cells. For multilocus engraftment analysis, cells from mice were stained with CD45 (H100) (Invitrogen), CD33 (P6.7), CD15 (H98), CD14 (M297), CD19 (HIB19), CD233a/Glya (GA-R2), CD41a (HIP8) and CD34 (581) (BD Biosciences). The data were analysed using FlowJo and the clusters and genes within replicates. CLIP–seq reads to the human genome was performed, sequencing reads from the top 100 AHR predicted targets with a minimal occurrence count threshold of 5. The data were aligned to the genome using 5′-UGGAAGUUCCCGGUGGAGGCAG-3′ as an adapt with CLIP–seq reads to the human genome was performed, sequencing reads from the top 100 AHR predicted targets with a minimal occurrence count threshold of 5. The data were aligned to the genome using 5′-UGGAAGUUCCCGGUGGAGGCAG-3′ as an adapt
libraries were trimmed of polyA tails, adapters, and low quality ends using Cutadapt with parameters–match-read-wildcards–times 2 -e 0 -Q quality-cutoff 6 -m 18 -b TCGTATGGGCTCCTCCTGCTG-b ATCCGTATGGGCTCCTCCTGCTG-b CGACAGTTCAAGGTTCAAGGTTACGACATG-b TGGTATCCTCCGGTCCAGAA-b AAAAAGGCTTACAGATGAGTCATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT...
Extended Data Figure 1 | MSI2 is highly expressed in human haematopoietic stem and progenitor cell populations. a, Schematic of the human haematopoietic hierarchy showing key primitive cell populations and simplified surface marker expression. b, qRT–PCR analysis of MSI1 and MSI2 expression in Lin− cord blood (CB) cell populations (n = 3 independent Lin− CB samples). c, Gating strategy used to sort sub-fractions of Lin− CB HSPCs for MSI2 qRT–PCR expression analysis (n = 2 independent pooled Lin− CB samples). d, MSI2 expression across the human haematopoietic hierarchy. Each circle represents an independent gene expression data set curated by HemaExplorer. e, Intracellular flow cytometry analysis of MSI2 protein levels in Lin− CB. Histograms show background staining with secondary antibody (red) and positive staining with anti-MSI2 antibody plus secondary in Lin− CB (blue). MSI2 fluorescence intensity was divided into quartiles of negative (Q1), low (Q2), mid (Q3) and high (Q4) level expression. f, Plots show cell percentage within each quartile from e that are CD34− CD38− (left) and CD34+ CD38+ (right) (n = 2 independent Lin− CB samples). All data presented as mean ± s.e.m. Unpaired t-test, *P < 0.05; ***P < 0.001.
Extended Data Figure 2 | MSI2 overexpression enhances in vitro culture of primitive CB cells. **a**, Top: schematic of bi-directional promoter lentivirus used to overexpress MSI2. Bottom: western blot and histogram showing intracellular flow validation of enforced MSI2 expression in 293FT cells (left) and Lin− CB (right), respectively. **b**, Representative images of secondary CFU made from replated control and MSI2-overexpressing (MSI2) CFU-GEMMs and types of colonies made. Scale bar, 200 μm. **c**, Fold change in Lin− CB transduced cell number after 7 days in culture following transduction (n = 5 experiments). **d**, Growth curve over 21 days of transduced Lin− CB cells (n = 4 experiments). **e**, Colony output of transduced Lin− CB from day 7 cultures (n = 8 cultures from 4 experiments). **f**, BrdU cell cycle analysis of transduced Lin− CB cells from day 10 cultures (n = 3 experiments). **g**, Ki67 cell cycle analysis of transduced Lin− CB cells from day 4 cultures (n = 4 experiments). **h**, Apoptotic and dead cells in day 7 cultures of transduced Lin− CB by Annexin V staining (n = 3 experiments). Western blot source data are available in Supplementary Fig. 1. All data presented as mean ± s.e.m. Unpaired t-test, **P < 0.01; ***P < 0.001.
Extended Data Figure 3 | MSI2 overexpression does not affect STRC lineage output and extends STRC-mediated engraftment. a, Schematic of STRC LDA experimental setup. b, Left: gating strategy to identify engrafted GFP+ CD45+ progenitor and myelo-lympho lineage-positive cell types or GFP+ CD45− erythroid cells and platelets. Right: summary of lineage output in the injected femur 3 weeks after transplantation (n = 4 mice for control and n = 18 mice for MSI2 overexpressing cells). MK, megakaryocyte; E, erythroid cells; P, platelets. c, Representative flow plots and summary of transduced STRC read out for engraftment with human CD45+ cells at 6.5 weeks post-transplant (n = 4 mice per condition). All data presented as mean ± s.e.m.
Extended Data Figure 4 | MSI2 knockdown impairs secondary CFU replating potential and HSC engraftment capacity. a, Left: schematic of MSI2- and control RFP-targeted shRNA lentiviruses. Right: confirmation of MSI2 protein knockdown (both isoforms that can be detected by western blot) in transduced NB4 cells. b, CFU production by shMSI2- and shControl-transduced Lin<sup>−</sup> CB (n = 8 cultures from 4 experiments). c, Secondary CFU output from shMSI2-transduced Lin<sup>−</sup> CB and images of representative secondary CFUs (scale bar, 200 μm; performed on n = 4 cultures from 2 experiments). d, Fold change in transduced cell number after 7 days in culture (n = 4 experiments). e, Growth curves of cultures initiated with transduced Lin<sup>−</sup> CB cells (n = 4 experiments). f, Experimental design to read out changes in HSC capacity with MSI2 knockdown. g, Left: representative flow analysis of transduced CD34<sup>+</sup> CD38<sup>−</sup>-derived human chimaerism in NSG mouse bone marrow. Right: ratio of the percentage of GFP<sup>+</sup> cells in the CD45<sup>+</sup> population post-transplant to the initial pre-transplant GFP<sup>+</sup> cell percentage. Dotted line indicates that the proportion of GFP<sup>+</sup> cells is unchanged relative to input. One sample t-test, no change = 1; n = 6 mice receiving shControl and n = 8 mice receiving shMSI2-transduced cells pooled from two experiments. h, Representative flow plots and summary of multilineage engraftment with shControl and shMSI2 cells (gated on GFP<sup>+</sup> cells). Western blot source data are shown in Supplementary Fig. 1. Data presented as mean ± s.e.m. Unpaired t-test, *P < 0.05; **P < 0.001.
Extended Data Figure 5 | MSI2 overexpression confers an HSC gene expression signature. a, Genes that are upregulated (21 genes, logFC >0) or downregulated (156 genes, logFC <0) in MSI2-overexpressing (OE) cells relative to control cells with FDR < 0.05 were compared to expression data from MSI2 knockdown cells normalized to shControl expression data. Red circles represent 177 genes that were significantly differentially expressed in MSI2-overexpressing cells. Gray outlined circles represent random genes (equal number of grey circles and red circles). Only genes that were significantly up- or downregulated in MSI2-overexpressing cells showed anti-correlation with MSI2 knockdown cells. b, Genes that were differentially expressed between MSI2-overexpressing and control cells (FDR < 0.05) compared to DMAP populations. Numbers beside each bar indicate the percentage of time for which the observed value (set of up- or downregulated genes) was better represented in that population than random values (equal number of randomly selected genes based on 1,000 trials).
Extended Data Figure 6 | MSI2 overexpression enhances HSC activity after ex vivo culture. a, Experimental procedure for measuring changes in HSC engraftment capacity and frequency with ex vivo culture. b, Representative flow plots of CD45+ GFP reconstitution from mice receiving the highest cell dose transplanted per time point. c, Multilineage engraftment of mice injected with D3 cultures. d, Proportion of the human CD45+ graft containing GFP+ cells from mice receiving the two highest doses of D3 primary grafts relative to pre-transplant levels of GFP+ cells before expansion (n = 8 mice for each dose). e, Proportion of the human CD45+ graft containing GFP+ cells from mice receiving the two highest doses of D10 primary grafts relative to pre-transplant levels of GFP+ cells after expansion (n = 8 mice for each dose, one-sample t-test, no change = 1). f, Multilineage engraftment of mice injected with D10 cultures. g, GFP mean fluorescence intensity (MFI) in D10 primary cell-engrafted mice. Data are from mice transplanted with the highest three doses; n = 11 control and 13 MSI2-overexpressing cell-engrafted mice. h, CD34 expression in GFP high (top 60%) relative to GFP low (bottom 40%) gated cells (set per mouse) from engrafted recipients in e. i, Number of transduced phenotyped HSCs after 7 days of culture from HSC expansion experiment described in a. Symbols represent individual mice and shaded symbols represent mice grafted with MSI2-overexpressing cells. All data presented as mean ± s.e.m. Unpaired t-test, *P < 0.05.
Extended Data Figure 7 | Predicted AHR targets and genes downregulated by SR1 or MSI2 overexpression are upregulated by MSI2 knockdown. 

a, Predicted AHR targets were identified with the iRegulon tool and compared with MSI2 knockdown normalized to shControl-upregulated gene signature by GSEA. 
b, log fold-change of MSI2-overexpression and knockdown shared leading edge AHR target genes from GSEA.
c, GSEA comparing gene sets downregulated by SR1 high and low dose with the MSI2 knockdown upregulated gene signature.
d, Heatmap and log fold-change of shared leading edge genes identified by GSEA from MSI2 overexpression, MSI2 knockdown and SR1 at varying doses. 
e, The percentage of downregulated genes in UM171-treated, SR1-treated and MSI2-overexpressing cells containing at least one AHR-binding site within 1,500bp upstream or downstream of the transcription start site. Dotted line indicates the background percentage of genes with AHR-binding sites. P values were generated relative to background with Fisher’s exact test.
Extended Data Figure 8 | AHR antagonism with SR1 has redundant effects with MSI2 overexpression, and AHR activation with MSI2 overexpression results in a loss of HSPCs. a, Representative flow plots and summary of CD34 expression in MSI2-overexpressing and control transduced CD34+ CB cells grown for 10 days in medium containing SR1 or DMSO vehicle (n = 3 experiments). b, Fold change in CD34 expression from results shown in a. c, Fold increase in CYP1B1 and AHRR transcript levels after FICZ treatment in transduced cultures (n = 3 experiments). d, Transduced CD34+ CB cells grown for 3 days in medium supplemented with FICZ and the corresponding change in CD34 expression. Each coloured pair (DMSO and FICZ) represents a matched CB sample (n = 3 experiments). e, Differences in culture CD34 levels from d. All data presented as mean ± s.e.m. Unpaired t-test, *P < 0.05.
Extended Data Figure 9 | MSI2 preferentially binds mature mRNA within the 3′UTR. **a**, Validation of the capacity of the anti-MSI2 antibody to immunoprecipitate MSI2 compared to IgG control pulldowns (heavy chain, HC; light chain, LC). **b**, Autoradiogram showing anti-MSI2 immunoprecipitated, MNase digested and radiolabelled RNA isolated for CLIP library construction and sequencing (red box). Low levels of MNase show a smearing pattern extending upwards from the modal weight of MSI2. **c**, Scatter plot of total number of uniquely mapped CLIP-seq reads for each gene, comparing both replicates. **d**, Heatmap indicating the number of different classes of Gencode-annotated genes that contain at least one predicted MSI-binding site. **e**, Consensus motifs within MSI2 clusters in the different genic regions. *P* values for the most statistically significant enriched motif are presented for all overlapping clusters between replicates. **f**, Cumulative distribution function of mean conservation score (Phastcons) of MSI2 clusters, compared to a shuffled background control, computed for all overlapping clusters and the top 40% of overlapping clusters. *P* values were obtained by a Kolgomorov–Smirnov two-tailed test comparing the distributions from actual and shuffled locations. **g**, Number of clusters within 200 bases of the annotated stop codon in known mRNA transcripts for all overlapping clusters between replicates and the top 40% of overlapping clusters. **h**, Cumulative distribution function of mean conservation score (Phastcons) of MSI2 clusters, compared to a shuffled background control, computed for overlapping clusters between the replicates and the top 40% of overlapping clusters found in different genic regions. Similarity between the 3′UTR conservation for the top 40% and the shuffled background control is probably due to MSI2 sites being small and not needing structural contexts for conservation. *P* values were obtained by a Kolgomorov–Smirnov two-tailed test comparing the distributions from actual and shuffled locations. **i**, Genome browser views displaying CLIP-seq mapped reads from replicate 1 (blue), predicted clusters (purple), exact matches for the GUAG sequence (black) and mammal conservation scores (PhyloP) in the 3′UTRs for a previously predicted Msi1 target.
Extended Data Figure 10 | MSI2 overexpression represses CYP1B1 and HSP90 3’UTR Renilla Luciferase reporter activity. a, CLIP–seq reads (replicate 1 in blue and replicate 2 in green) and clusters (purple) mapped to the 3’UTR of HSP90. Matches to the GUAG motif are shown in black. Mammal PhyloP score listed in last track. b, c, Representative data of mean per cell fluorescence for HSP90 and CYP1B1 protein in transduced CD34⁺ CB. Protein level in cells during in vitro culture was analysed 3 days (D3) and 7 days (D7) after transduction and sorting for GFP. Corresponding secondary-alone antibody staining is shown for each experiment. Each circle represents a cell, and more than 200 cells were analysed per condition. d, e, Levels of renilla luciferase activity in NIH-3T3 cells co-transfected with control or MSI2 overexpression vectors and the CYP1B1 or HSP90 wild-type or TCC mutant 3’UTR luciferase reporter (dotted line indicates no change in renilla activity; n = 4 CYP1B1 and n = 3 HSP90 experiments). f, Flow plots of co-transduced CD34⁺ CB cells with MSI2 (GFP) and CYP1B1 (BFP) lentivirus. g, GFP⁺ BFP⁺ CFU-GEMMs generated from f were replated into secondary CFU assays and the total number of colonies formed was counted. A total of 24 CFU-GEMMs from MSI2-BFP and MSI2-CYP1B1 were replated (n = 2 experiments). Data presented as mean ± s.e.m. Unpaired t-test, ***P < 0.001, **P < 0.01. h, A model for AHR pathway attenuation by MSI2 post-transcriptional processing. MSI2 mediates the post-transcriptional downregulation of HSP90 at the outset of culture and continuously represses the prominent AHR pathway effector CYP1B1 to facilitate HSPC expansion. The resulting MSI2-mediated repression of AHR signalling enforces a self-renewal program and allows HSPC expansion ex vivo.