© 2014. Published by The Company of Biologists Ltd | Development (2014) 141, 4018-4030 doi:10.1242/dev.115709

RESEARCH ARTICLE

TECHNIQUES AND RESOURCES

Single-cell analyses of regulatory network perturbations using enhancer-targeting TALEs suggest novel roles for *PU.1* during haematopoietic specification

Adam C. Wilkinson[1,*], Viviane K. S. Kawata[1,2,*], Judith Schütte[1], Xuefei Gao[3], Stella Antoniou[4], Claudia Baumann[4], Steven Woodhouse[1], Rebecca Hannah[1], Yosuke Tanaka[1], Gemma Swiers[4], Victoria Moignard[1], Jasmin Fisher[5,6], Shimauchi Hidetoshi[2], Marloes R. Tijssen[7], Marella F. T. R. de Bruijn[4], Pentao Liu[3] and Berthold Göttgens[1,‡]

ABSTRACT

Transcription factors (TFs) act within wider regulatory networks to control cell identity and fate. Numerous TFs, including Scl (Tal1) and PU.1 (Spi1), are known regulators of developmental and adult haematopoiesis, but how they act within wider TF networks is still poorly understood. Transcription activator-like effectors (TALEs) are a novel class of genetic tool based on the modular DNA-binding domains of *Xanthomonas* TAL proteins, which enable DNA sequence-specific targeting and the manipulation of endogenous gene expression. Here, we report TALEs engineered to target the *PU.1-14kb* and *Scl+40kb* transcriptional enhancers as efficient new tools to perturb the expression of these key haematopoietic TFs. We confirmed the efficiency of these TALEs at the single-cell level using high-throughput RT-qPCR, which also allowed us to assess the consequences of both *PU.1* activation and repression on wider TF networks during developmental haematopoiesis. Combined with comprehensive cellular assays, these experiments uncovered novel roles for *PU.1* during early haematopoietic specification. Finally, transgenic mouse studies confirmed that the *PU.1-14kb* element is active at sites of definitive haematopoiesis *in vivo* and PU.1 is detectable in haemogenic endothelium and early committing blood cells. We therefore establish TALEs as powerful new tools to study the functionality of transcriptional networks that control developmental processes such as early haematopoiesis.

KEY WORDS: Haematopoiesis, Transcription activator-like effectors, Regulatory networks, PU.1

INTRODUCTION

Transcriptions factors (TFs) are key regulators of cell identity and fate. Cell type-specific transcriptional regulation is thought to largely occur by TF binding to distal *cis*-regulatory elements (Heinz et al., 2010). The haematopoietic system provides a well-studied model of mammalian tissue development, in which numerous key TFs have been described [reviewed by Wilkinson and Gottgens (2013)], including Scl (Tal1) and PU.1 (Spi1). The identification of *cis*-regulatory elements that regulate the expression of such TFs has begun to reveal TF circuits that suggest the existence of highly interconnected TF regulatory networks active in the haematopoietic system (Pimanda and Gottgens, 2010; Schutte et al., 2012).

Two well-studied examples of such haematopoietic *cis*-regulatory elements are the *PU.1-14kb* (Rosenbauer et al., 2004; Okuno et al., 2005; Huang et al., 2008; Staber et al., 2013) and *Scl +40kb* (Delabesse et al., 2005; Ogilvy et al., 2007; Ferreira et al., 2013). The *PU.1-14kb* plays a key role in *PU.1* expression in haematopoietic stem/progenitor cells (HSPCs) and mature haematopoietic cell types; its deletion results in an 80% loss of *PU.1* gene expression and acute myeloid leukaemia (AML) in mice (Rosenbauer et al., 2004), while mutation of an (autoregulatory) Ets site within the *PU.1-14kb* causes a 66% reduction in *PU.1* gene expression, which leads to haematopoietic stem cell exhaustion (Staber et al., 2013). Although the *Scl+40kb* element is active during haematopoietic emergence, its deletion causes only a mild erythroid phenotype (Ferreira et al., 2013). The *Scl+40kb* element is additionally thought to regulate expression of the 3′ flanking gene, *Map17* (*Pdzk1ip1*) (Tijssen et al., 2011; Ferreira et al., 2013).

Recent technological advances in microfluidic technology have led to the development of robust protocols for high-throughput quantification of gene expression in single cells (Guo et al., 2010). One of the earliest studies reporting microfluidics-based single-cell gene expression highlighted the potential for heterogeneity of knockdown efficiency within single cells following siRNA-mediated gene silencing (Toriello et al., 2008). However, the ability to accurately assess gene expression in single cells following conventional perturbations, such as retroviral overexpression or shRNA-mediated knockdown, has been limited because the former commonly yields unphysiologically high expression levels with no means to distinguish between the endogenous and ectopically expressed gene, whereas the latter acts post-transcriptionally and can therefore inhibit protein production without affecting transcript abundance. To realise the full potential of analysing perturbation phenotypes by single-cell gene expression profiling, more physiological means to tune gene expression levels are therefore required.

Transcription activator-like effectors (TALEs) are a novel class of TFs identified in the bacterial plant pathogen *Xanthomonas*, where they are secreted as virulence factors to modulate gene expression of the host plant (Boch and Bonas, 2010). TALEs have a unique

[1]Cambridge Institute for Medical Research and Wellcome Trust–MRC Cambridge Stem Cell Institute, University of Cambridge, Hills Road, Cambridge CB2 0XY, UK. [2]Division of Periodontology and Endodontology, Tohoku University Graduate School of Dentistry, Sendai 980-8575, Japan. [3]Wellcome Trust Sanger Institute, Cambridge CB10 1SA, UK. [4]MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DS, UK. [5]Microsoft Research Cambridge, 21 Station Road, Cambridge CB1 2FB, UK. [6]Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, UK. [7]Department of Haematology, University of Cambridge and National Health Service Blood and Transplant, Cambridge CB2 0PT, UK.
*These authors contributed equally to this work

[‡]Author for correspondence (bg200@cam.ac.uk)

Received 25 July 2014; Accepted 18 August 2014

Development (2014) 141, 4018-4030 doi:10.1242/dev.115709

modular DNA-binding domain consisting of 33-35 amino acid repeats, each of which binds a single nucleotide with base recognition specificity (Boch et al., 2009). TALEs fused to transcriptional effector domains have been shown to modulate endogenous gene expression (Zhang et al., 2011; Cong et al., 2012; Gao et al., 2013).

Here, we present the use of TALEs (fused to transcriptional effector domains) designed to target conserved regions within haematopoietic TF *cis*-regulatory elements as an efficient tool to regulate target gene expression. We validated TALEs targeting the *PU.1-14kb* and *Scl+40kb* elements and further assessed the phenotypic effect of modulating the activity of these enhancers on embryoid body (EB) haematopoiesis. We go on to highlight the combination of TALE-mediated endogenous gene expression perturbations with single-cell gene expression studies as a powerful approach to investigate TF regulatory networks. Using these methods in combination with transgenic embryo analysis, we uncover a novel role for PU.1 expression, mediated via *PU.1-14kb*, in haematopoietic specification during development.

RESULTS

Design and validation of TALEs targeting conserved regions within haematopoietic enhancers

We identified regions within the *Scl+40kb* and *PU.1-14kb* elements that were perfectly conserved between human and mouse. TALEs were designed to match these regions and nowhere else in either genome (Fig. 1A-C). TALEs were initially assembled fused to the VP64 (transcriptional activator) domain (Beerli et al., 1998) and an mCherry fluorescent reporter via a 2A peptide (Fig. 1A). TALE constructs were cloned into piggyBac transposon-based plasmids (Wang et al., 2008) for efficient stable genomic integration and under the control of a tetracycline-responsive promoter (TetR) to provide inducible [with doxycycline (dox)] expression (Fig. 1A). We initially validated TALE-VP64 proteins in both human and mouse systems (Fig. 1D). In human K562 cells, the TALE-VP64 targeting *Scl+40kb* (T-VP64-*Scl+40*) upregulated *SCL* expression ~4-fold but had little effect on *MAP17* expression (Fig. 1E). By contrast, in mouse 416B cells T-VP64-*Scl+40* upregulated *Map17* expression ~22-fold but had little effect on *Scl* expression (Fig. 1E). In both the human K562 and mouse 416B cells, expression of the TALE-VP64 targeting *PU.1-14kb* (T-VP64-*PU.1-14*) upregulated *PU.1* expression 3- to 4-fold and *SLC39A13/Slc39a13* expression ~2-fold (Fig. 1F).

Modest (1.5- to 8.5-fold) increases in histone H3 lysine 27 acetylation (H3K27Ac), an epigenetic modification associated with active regions of chromatin (Creyghton et al., 2010), were also seen in 416B cells at the promoters of TALE-VP64 target genes, consistent with increased transcription (supplementary material Fig. S1A,B). H3K27Ac was also enriched 3.8-fold at *Scl+40kb* when the TALE-VP64 targeting this enhancer was expressed (supplementary material Fig. S1A). However, a 50% reduction in H3K27Ac was seen at *PU.1-14kb* when the TALE-VP64 targeting this enhancer was expressed (supplementary material Fig. S1B), perhaps due to nucleosome displacement caused by TALE-VP64 and co-factor DNA binding. In mouse embryonic stem cells (mESCs), in which these enhancers are not active (as determined by H3K27Ac ChIP-seq enrichment; data not shown) and target genes are weakly expressed, TALE-VP64 did not upregulate gene expression (supplementary material Fig. S1C,D).

To determine the specificity of these TALEs, we further determined expression changes to genes within ~100 kb of the target regions (supplementary material Fig. S1E-H). Less than 1.7-fold increases in expression were seen in K562, 416B and mESCs. Reduced expression in some genes (such as *Stil* in 416B cells expressing T-VP64-*Scl+40*) was identified, perhaps due to transcription factory reallocation (Papantonis and Cook, 2013). Additionally, we confirmed by chromatin immunoprecipitation (ChIP) that TALEs bind to their target regions. By ChIP-qPCR, enrichments of 15- to 17-fold were seen at target locations (Fig. 1G,H). To further assess TALE binding specificity genome-wide, we sequenced the HA-T-VP64-*PU.1-14* and 416B control HA antibody ChIP samples. The number of regions across the entire genome that showed enrichment was very small and comparable between the 416B control and HA-T-VP64-*PU.1-14* samples, underlining the high specificity afforded by TAL-mediated targeting reported by others (Mali et al., 2013). Importantly, a clear peak at the *PU.1-14kb* element could be identified in the HA-T-VP64-*PU.1-14* sample, but not in the control (supplementary material Fig. S1I). Manual assessment of enrichment at regions containing similar DNA sequences to the HA-T-VP64-*PU.1-14* target sequence did not identify strong off-target binding, and no other binding events occurred within a 15 Mb window around *PU.1-14kb* (supplementary material Fig. S1I).

We next assessed the ability of TALEs to regulate target gene expression during development using the mouse EB differentiation system, which has been validated as a useful and tractable *in vitro* model of embryonic haematopoiesis (Keller et al., 1993). We transfected the mESC line Ainv18 (Kyba et al., 2002), which constitutively expresses rtTA from the *Rosa26* locus, and expanded stably integrated clones that displayed inducible mCherry expression (data not shown). The data described below are representative of multiple clones tested for each TALE construct. We differentiated TALE-containing ESC lines, induced TALE expression by addition of dox at day 4 (just prior to definitive haematopoiesis in this system) and assessed phenotypic effects after a further 48 h of culture, relative to a culture without dox treatment (Fig. 2A). Initial flow cytometric analysis of the day 6 EBs confirmed pure mCherry^{+} populations in the dox cultures (Fig. 2A). Following this protocol, *Scl* expression was upregulated ~1.9-fold in cells induced to express T-VP64-*Scl+40*, and *Map17* expression was upregulated over 3-fold (Fig. 2B). *PU.1* expression was upregulated over 4-fold by T-VP64-*PU.1-14*, with no significant change in *Slc39a13* expression (Fig. 2B). We additionally generated ESCs containing a *PU.1-14kb*-targeting TALE for *PU.1* repression by swapping the VP64 activation domain for the KRAB repressor domain (T-KRAB-*PU.1-14*). Following the same differentiation protocol as above, we observed efficient repression of *PU.1* by the TALE-KRAB, with expression reduced by over 50% (Fig. 2B). *Slc39a13* was unaffected by TALE expression, suggesting that, at least in this developmental context, *PU.1-14kb* activity is specific to *PU.1*.

Transient expression of a *PU.1* enhancer-targeting TALE alters EB haematopoiesis

We next assessed the phenotypic effect of TALE-mediated modulation of gene expression by haematopoietic colony forming assays using day 6 EB cells (supplementary material Fig. S2A). TALE-VP64-mediated *PU.1* upregulation resulted in a significant loss of colony forming ability in day 6 EBs (Fig. 2C). By contrast, TALE-KRAB-mediated *PU.1* repression caused a doubling in myeloid (CFU-GM/G) and mixed (CFU-Mix) colony numbers (Fig. 2C), although this was not statistically significant. The colony potential of day 6 EBs was largely unaffected by expressing T-VP64-*Scl+40*, except for a slight (but not significant) increase in BFUe frequency (Fig. 2C).

Development (2014) 141, 4018-4030 doi:10.1242/dev.115709

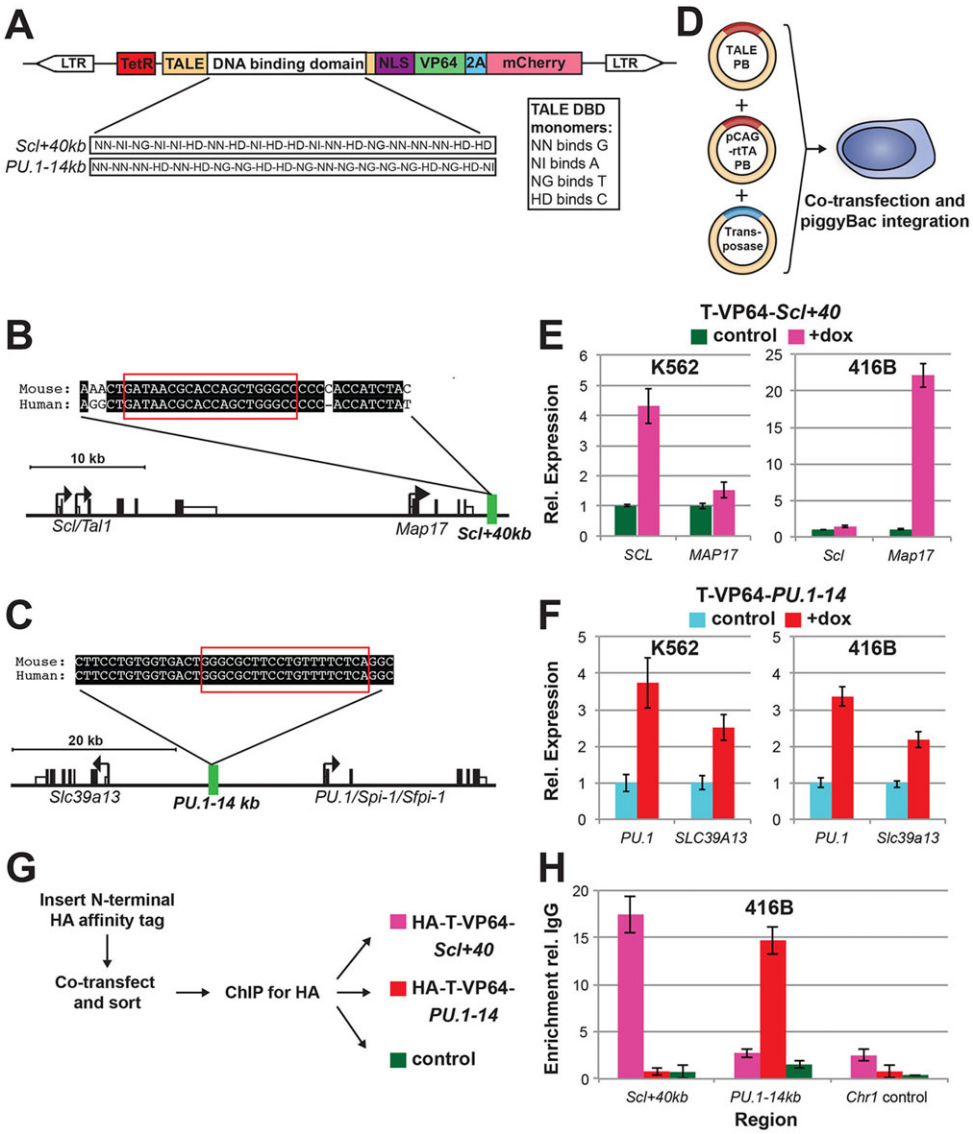


Fig. 1. Experimental approach and validation. (A) Structure of the TALE-expressing piggyBac construct. TALE cDNA consists of the TALE sequence followed by a nuclear localisation domain (NLS), a VP64 domain, 2A (peptide sequence cleaved after translation) and mCherry fluorescent protein. TALE cDNA was cloned downstream of a tetracycline-responsive promoter (TetR), and within piggyBac long terminal repeats (LTRs) for stable transposase-mediated genomic integration. The DNA-binding domain (DBD) within the TALE sequence consists of twenty monomers. Monomers contain two hypervariable amino acids that determine nucleotide-binding specificity: NN, NI, NG or HD. (B,C) Schematics of the mouse *Scl* (*Tal1*) (B) and *PU.1* (*Spi1*) (C) genomic loci, with the *Scl+40kb* and *PU.1-14kb* elements highlighted in green. TALE target sites within conserved (between human and mouse) sequences are highlighted in red. (D) Experimental approach to express TALEs in cell lines. K562 and 416B cells were co-transfected with the TALE-expressing piggyBac (TALE-PB) from A, a constitutively expressing rtTA piggyBac vector (pCAG-rtTA-PB) and a piggyBac transposase, to create inducible TALE-expressing cells. (E) Effect of expressing TALE-VP64 targeting *Scl+40kb* (T-VP64-*Scl+40*) in human K562 (left) and mouse 416B (right) cells on neighbouring gene expression. T-VP64-*Scl+40* was expressed for 48 h by addition of doxycycline (dox) and gene expression in mCherry^{+} cells was determined relative to mCherry^{-} control cells. Error bars indicate s.d. of technical triplicates, and are representative of two biological replicates. (F) As in E, but for TALE-VP64 targeting *PU.1-14kb* (T-VP64-*PU.1-14*). (G) ChIP approach for TALE-VP64 proteins in H. An HA affinity tag was inserted at the N-terminus of the TALE-VP64 (HA-T-VP64); 416B cells were co-transfected as in D, sorted and ChIP performed 48 h after dox addition. (H) ChIP-qPCR enrichment of HA-tagged TALE-VP64 (HA-T-VP64) relative to IgG in HA-T-VP64-*Scl+40* (pink), HA-T-VP64-PU.1-14 (red) and untransfected 416B control (green) cells at *Scl+40kb*, *PU.1-14kb* and a control region on chromosome 1 (*Chr1*). Error bars indicate s.d. of technical triplicates from one biological experiment.

To ensure that TALE-VP64 expression alone was not affecting CFU frequency, we assessed ESC lines carrying a non-functional TALE-VP64 [generated previously (Gao et al., 2013)], which did not affect CFU frequency (supplementary material Fig. S2B). To correlate the changes in haematopoietic progenitors/CFUs with changes in the cellular composition of the day 6 EBs, we analysed day 6 EBs by flow cytometry. Consistent with the modest effects in colony forming assays, expression of T-VP64-*Scl+40* minimally affected haematopoietic cell populations present in day 6 EBs (supplementary material Fig. S2C). Expression of TALEs targeting the *PU.1-14kb* partially, but not significantly, reduced the total cell numbers recovered (supplementary material Fig. S2D). However, this was not due to increased apoptosis, as assessed by Annexin V and DAPI staining of day 6 EBs (supplementary material Fig. S2D).

Although TALE-VP64-mediated upregulation of *PU.1* caused an increase in the relative size of the CD41^{+} population (Fig. 2D;

Development (2014) 141, 4018-4030 doi:10.1242/dev.115709

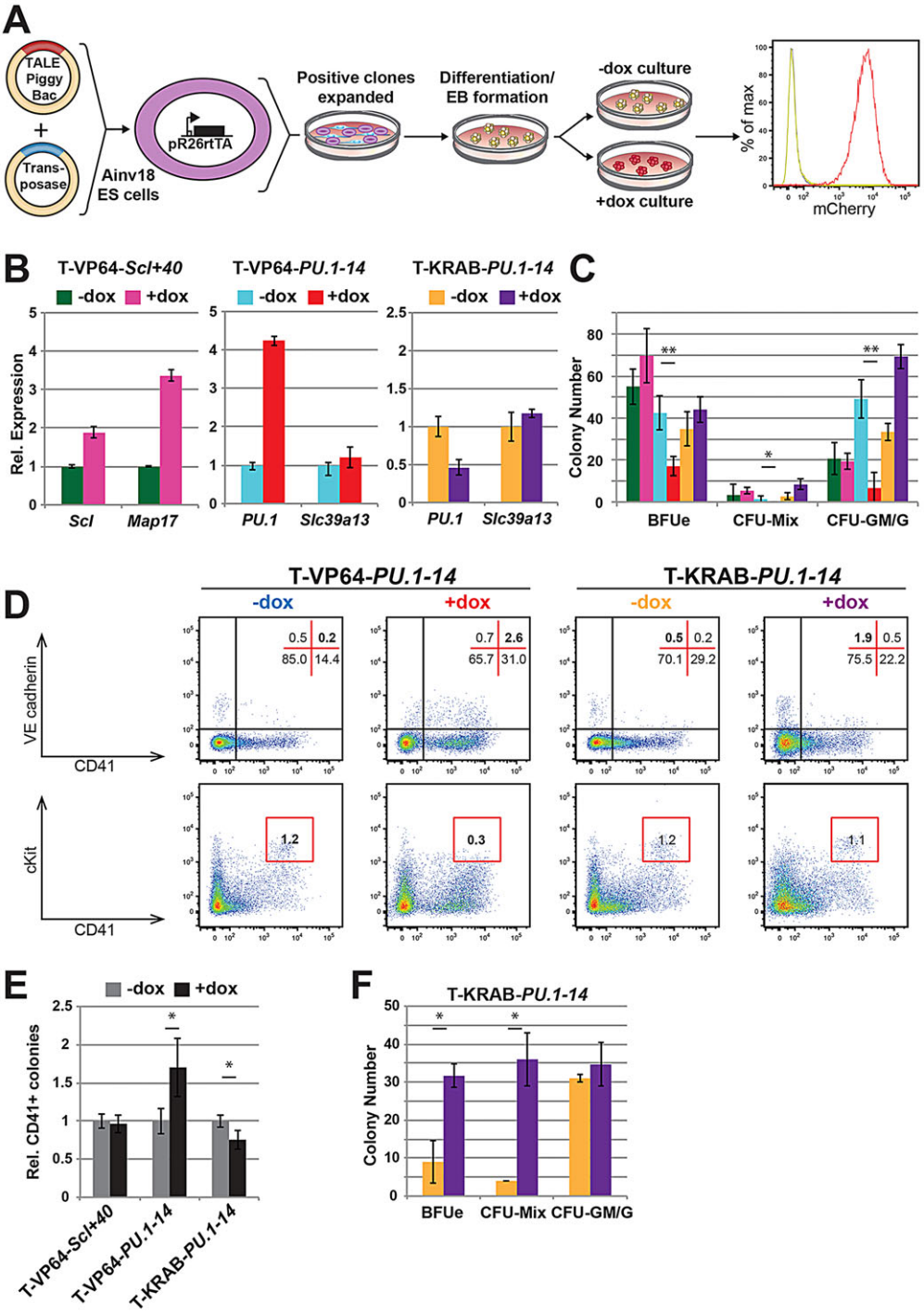


Fig. 2. Transient TALE expression affects haematopoietic cell fate decisions. (A) Experimental approach using Ainv18 ESC differentiation to study TALE-mediated gene expression perturbations in haematopoiesis. Mouse Ainv18 ESCs constitutively expressing rtTA from the *Rosa26* locus (pR26-rtTA) were co-transfected with the inducible TALE-PB construct and transposase. Targeted ESCs were differentiated into embryoid bodies (EBs), and TALE expression induced at day 4 by addition of dox. Changes in gene expression, colony potential and surface marker phenotype were analysed at day 6 in the +dox EBs as compared with –dox controls. (B) Gene expression changes in day 6 EBs after induction of T-VP64-*Scl+40* (left), T-VP64-*PU.1-14* (middle) and T-KRAB-*PU.1-14* (right). Error bars indicate s.e.m. of three biological replicates. (C) Representative haematopoietic colony numbers from 1×10^5 day 6 EB cells (colour scheme as in B). Colonies were grown in methylcellulose supplemented with SCF, IL-3, IL-6 and Epo. See supplementary material Fig. S2A for images of representative colony forming units (CFUs) scored. Error bars indicate s.d. of technical triplicates. **P*<0.05, ***P*<0.01 (Student's *t*-test), from three biological replicates. (D) Flow cytometry plots of day 6 EB cells showing Flk1 versus CD41 (top) and VEcad versus CD41 (bottom). Representative staining patterns are shown for T-VP64-*PU.1-14* (left) and T-KRAB-*PU.1-14* (right). The distribution of cells within quadrants/gates is shown by percentage. (E) Relative number of day 4 $Flk1^+$ EB-derived colonies containing $CD41^+$ haematopoietic cells, grown on OP9 stromal cells for 84 h (dox added after 36 h). See supplementary material Fig. S2G for representative image of scored colony. Error bars indicate s.e.m. from biological triplicates. **P*<0.01 (Student's *t*-test), from three biological replicates. (F) Average numbers of haematopoietic colonies from 1×10^5 day 6 EB T-KRAB-*PU.1-14* cells plated onto confluent OP9 stromal cells for 24 h before CFU assay initiated by addition of methylcellulose supplemented with SCF, IL-3, IL-6 and Epo. Colour scheme as in B. Error bars indicate s.d. of three biological replicates. **P*<0.05 (Student's *t*-test), from three biological replicates.

Development (2014) 141, 4018-4030 doi:10.1242/dev.115709

supplementary material Fig. S2E), when combined with total cell numbers recovered from the EBs this did not result in a significant increase in the absolute number of CD41$^+$ cells (supplementary material Fig. S2F). Interestingly, TALE-VP64-mediated *PU.1* expression caused a loss of the Flk1$^+$ (mesoderm) population (supplementary material Fig. S2E,F), and significantly increased the CD41$^+$ VE-cadherin (VEcad)$^+$ (committing haemogenic endothelial) population (Fig. 2D; supplementary material Fig. S2F). Additionally, TALE-VP64-mediated *PU.1* upregulation caused a loss of the CD41$^+$ cKit^{hi} ('early definitive haematopoietic progenitor') population, which might help explain the loss of colony forming potential described above (Fig. 2C,D). Combined with the CFU assays, these data suggested that PU.1 upregulation might push differentiating cells towards a haematopoietic fate but then inhibits proliferation of the resulting blood cells. Consistent with this hypothesis, TALE-mediated *PU.1* induction modestly increased (1.5-fold) day 4 EB-derived colonies containing budding CD41$^+$ haematopoietic cells (Fig. 2E; supplementary material Fig. S2G), whereas *PU.1* repression modestly reduced their frequency.

By contrast, the major change caused by downregulation of *PU.1* by T-KRAB-*PU.1-14* was an almost complete loss of the CD45$^+$ (Ptprc$^+$) population (committed definitive haematopoietic cells; supplementary material Fig. S2E), which led us to speculate that the delayed haematopoiesis caused by *PU.1* repression might be masking an increase in haematopoietic CFU frequency. To test this further, we allowed day 6 EB cells to mature on OP9 stromal cells for 24 h before assessing CFU frequency (Fig. 2F). This led to a significant 3-fold and 9-fold increase in BFUe and CFU-Mix colonies, respectively, which is consistent with published data suggesting that PU.1 expression restricts haematopoietic cells to a myeloid fate (Mak et al., 2011).

Combined, these data suggest that upregulation of *PU.1* drives haematopoietic commitment, but causes loss of proliferative ability within the haematopoietic population, whereas temporary downregulation of *PU.1* inhibits the maturation and differentiation of early haematopoietic cells.

Single-cell gene expression analysis of TALE-mediated *PU.1* perturbation

Having determined the phenotypic effects of TALE-mediated *PU.1* expression perturbations, we next asked what effects *PU.1* modulation might have on TF regulatory networks. We assessed the effect of T-VP64-*PU.1-14* induction on the expression of 44 haematopoietic, mesodermal and endothelial TFs and surface markers as well as four control housekeeping genes in single day-6 EB VEcad$^+$ cells using the Fluidigm Biomark platform. At this time point, VEcad expression marks endothelium and haemogenic endothelium, which were not expected to express robust levels of *PU.1*. To provide an internal control, we differentiated a chimeric mixture of wild-type (WT) and TALE-inducible ESCs, and sorted VEcad$^+$ cells from mCherry$^-$ and mCherry$^+$ populations at day 6 (48 h after dox addition; Fig. 3A). We assessed the expression of all 48 genes in 160 single cells for each population, which, after quality control, resulted in expression data for 136 and 147 cells, respectively.

PU.1 was only expressed in 33% (45 of 136) of mCherry$^-$ VEcad$^+$ cells (Fig. 3B). By contrast, TALE-VP64 efficiently induced *PU.1* expression in 84% (124 of 147) of the mCherry$^+$ VEcad$^+$ cells. Moreover, *PU.1*-expressing cells in the mCherry$^+$ VEcad$^+$ population tended to express *PU.1* at a higher level than the *PU.1*-expressing cells in the mCherry$^-$ VEcad$^+$ population (an average of 3.3 ΔCt higher, relative to *Polr2a* and *Ubc* expression; Fig. 3B). This demonstrates that T-VP64-*PU.1-14* can induce gene expression efficiently (but not with complete efficiency), and that the distribution of *PU.1* expression levels within *PU.1*-expressing cells is altered, with a much larger proportion of individual cells expressing high levels of *PU.1*. Single-cell expression analysis therefore reveals both qualitative (a shift towards more cells expressing) and quantitative (a shift towards higher per-cell expression levels) consequences of TALE-mediated activation of *PU.1*.

Importantly, TALE-induced *PU.1* expression was associated with consistent changes in the expression of other genes. mCherry$^+$ cells expressed higher levels of several haematopoietic genes, including *Csf1r*, *Gata3*, *Gfi1b*, *Runx1* and *Etv6* (*Tel*) (Fig. 3B). Interestingly, mCherry$^+$ cells also express lower levels of several genes thought to be important for mesoderm or endothelium, including *Ets1*, *Etv2*, *Flk1* (*Kdr*), *Notch1*, *Sox17* and *VEcad* (*Cdh5*) (Fig. 3B). Moreover, gene expression changes for *Flk1*, *CD41* (*Itga2b*) and *Kit* correlated well with the expression of these surface markers as assessed by flow cytometry (Fig. 2D and Fig. 3B; supplementary material Figs S2 and S3). As *Kit* encodes the receptor for the pro-proliferative cytokine stem cell factor (Scf; Kitl), its downregulation at the transcriptional level and cell surface might partially explain the loss of proliferative ability in T-VP64-*PU.1-14*-expressing day 6 EB cells (Fig. 2E).

Pairwise all-against-all comparisons of the expression of the 44 TFs and surface proteins across all 283 single cells were performed by calculating Spearman rank correlation coefficients, which were displayed using a heatmap to illustrate both positive and negative correlations between pairs of genes. This identified two positively correlated gene clusters: a haematopoietic gene cluster (including *PU.1*) and a mesodermal/endothelial gene cluster (Fig. 3C). Although genes from both clusters can be co-expressed in single cells (Fig. 3D), genes from the haematopoietic cluster predominantly showed negative correlation to genes from the endothelial cluster (Fig. 3C), suggesting an antagonism between these regulatory networks. Pairwise analysis and hierarchical clustering of cells based on their gene expression signatures largely separated the mCherry$^-$ and mCherry$^+$ cells within the VEcad$^+$ population (Fig. 3D). As expected, it was within the mCherry$^+$ population that the positively correlated cluster of haematopoietic genes was more frequently activated.

PU.1 can promote haematopoietic commitment of haemogenic endothelial precursors

The data described above suggest that precocious *PU.1* expression in haematopoietic precursors can drive haematopoietic commitment through the activation of a TF network. To investigate this further, we performed additional single-cell gene expression analyses for the CD41$^+$ cKit^{hi} population from WT and T-KRAB-*PU.1-14* differentiated ESCs. As above, 160 mCherry$^+$ and mCherry$^-$ cells were sorted from day 6 EBs, from which 142 and 141 single cells, respectively, passed quality control. Within the CD41$^+$ cKit^{hi} mCherry$^-$ (WT Ainv18) population, over 90% expressed *PU.1* (132 of 141), and clearly had acquired a committed haematopoietic gene expression pattern (including *Runx1*, *Myb*, *Ikaros*) with only a few cells expressing mesoderm/endothelium-associated genes (e.g. *Sox7*, *Sox17*, *Etv2*) (Fig. 4A; supplementary material Fig. S4). By contrast, less than 60% (85 of 142) of the CD41$^+$ cKit^{hi} mCherry$^+$ (TALE-KRAB-*PU.1-14*) cells expressed detectable *PU.1* transcript, and *PU.1* was expressed at lower levels in those that did (an average of 2.8 ΔCt lower; Fig. 4A), demonstrating that TALE-KRAB efficiently repressed *PU.1* expression in CD41$^+$ cKit^{hi} cells. The expression of *Csf1r*, a known downstream target of PU.1, is tightly correlated with *PU.1* expression and *Csf1r* is not expressed in cells lacking *PU.1*

Development (2014) 141, 4018-4030 doi:10.1242/dev.115709

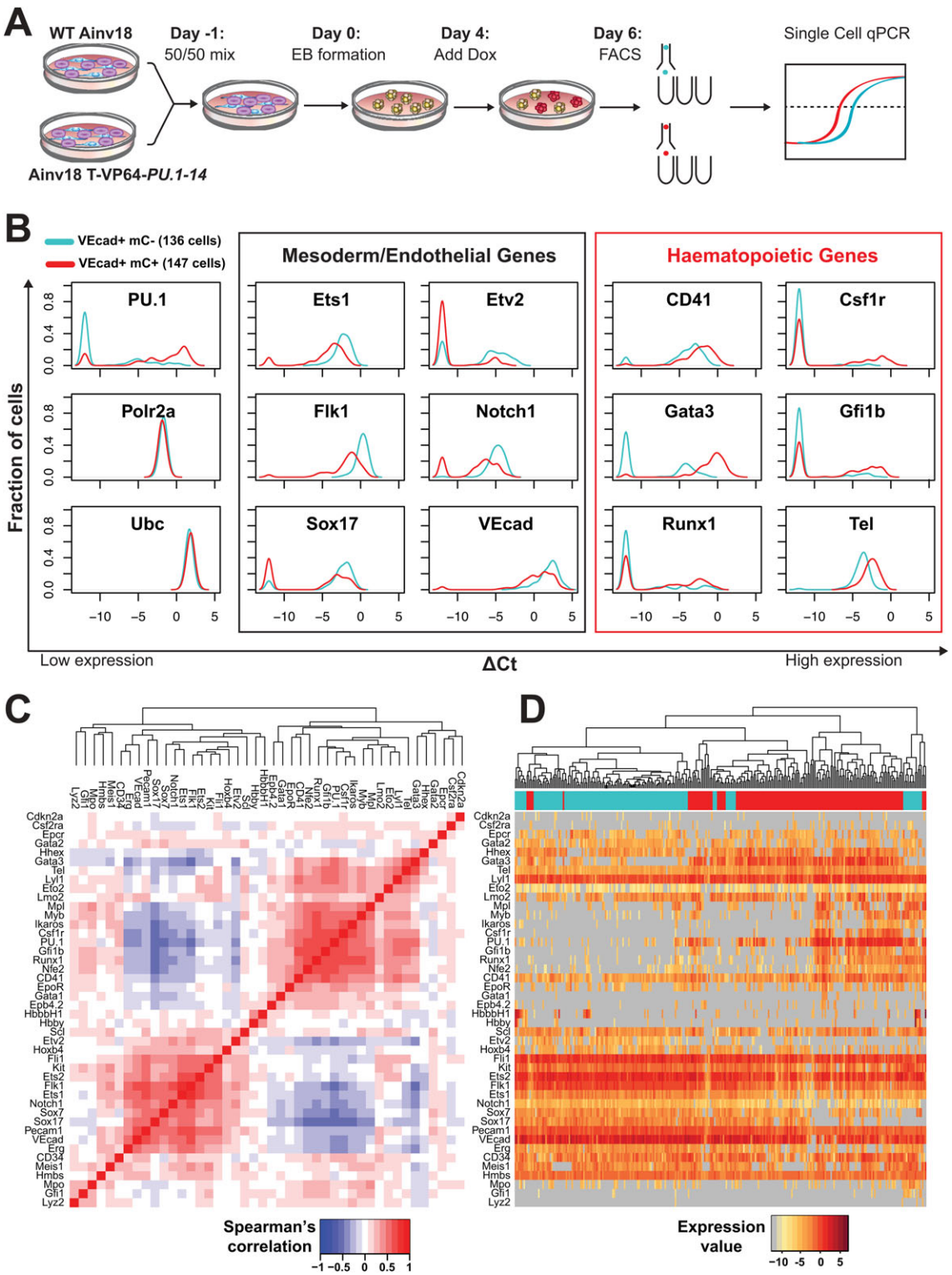


Fig. 3. Single-cell analysis of TALE-mediated *PU.1* expression in haematopoietic precursors. (A) Strategy for single-cell gene expression analysis of TALE-mediated perturbations. Wild-type (WT) Ainv18 and T-VP64-*PU.1-14*-targeted ESCs were passaged once as a 1:1 mix before EB formation. Dox was added at day 4 and EBs disaggregated at day 6. Single VEcad^{+} cells (mCherry^{+} and mCherry^{-} sorted as T-VP64-*PU.1-14*-expressing and WT, respectively) were sorted into lysis buffer. Single-tube reverse transcription and targeted pre-amplification were undertaken, followed by multiplexed qPCR gene expression analysis using the Fluidigm Biomark platform. (B) Density plots of gene expression in day 6 EB VEcad^{+} mCherry^{-} (136 WT Ainv18; cyan) and VEcad^{+} mCherry^{+} (147 T-VP64-*PU.1-14* expressing; red) cells. The density indicates the fraction of cells at each expression level, relative to housekeeping genes (*Polr2a* and *Ubc*). Cells with non-detected gene expression were set to –12. See supplementary material Fig. S3 for density plots for all 48 genes analysed in these two populations. (C) Hierarchical clustering of Spearman rank correlations between all pairs of genes (excluding housekeepers) from all 283 VEcad^{+} cells (red, positive correlation; blue, negative correlation). (D) Hierarchical clustering of the 283 VEcad^{+} cells according to gene expression, with genes ordered according to C (dark red, highly expressed; grey, non-expressed). Top bar indicates cell type: cyan, mCherry^{-}; red, mCherry^{+}.

Development (2014) 141, 4018-4030 doi:10.1242/dev.115709

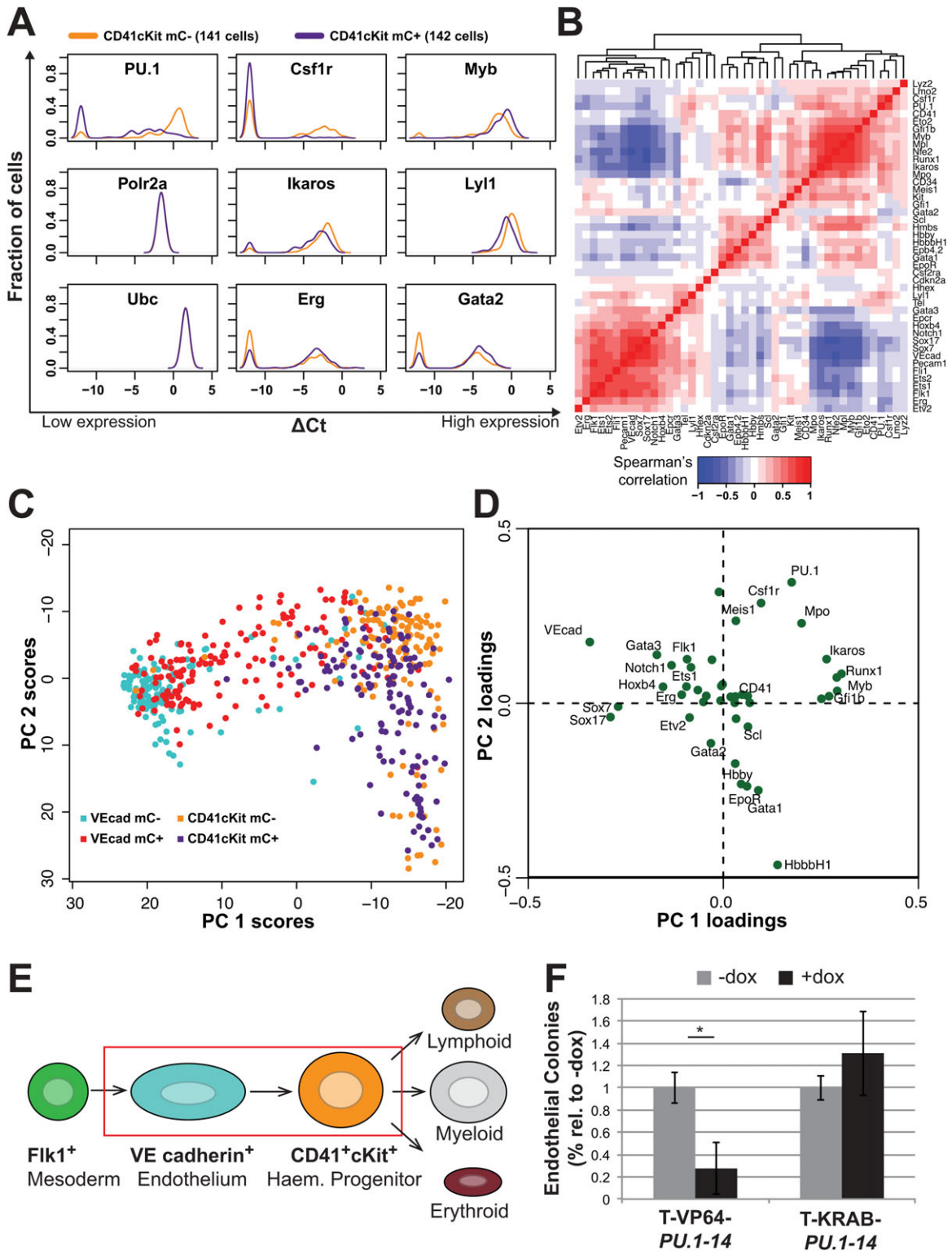


Fig. 4. TALE-mediated expression perturbations suggest transcriptional interactions during blood specification. (A) Density plots of gene expression in day 6 EB CD41^{+} cKit^{hi} (CD41cKit) mCherry^{-} (141 WT Ainv18; orange) and CD41cKit mCherry^{+} (142 Ainv18 expressing T-KRAB-*PU.1-14*; purple) cells. The density indicates the fraction of cells at each expression level, relative to housekeeping genes (*Polr2a* and *Ubc*). Cells with non-detected gene expression were set to –12. See supplementary material Fig. S4 for density plots for all 48 genes analysed in these two populations. (B) Hierarchical clustering of Spearman rank correlations between all pairs of genes (excluding housekeepers) using gene expression data from all 566 cells (VEcad^{+} and CD41cKit). (C) Principal component analysis (PCA) of the 566 VEcad^{+} and CD41cKit cells, in the first and second components, from the expression of all 44 genes (excluding the four housekeeping genes). (D) Principal component loadings indicate the extent to which each gene contributes to the separation of cells along each component in C. (E) Current model of definitive haematopoietic specification from Flk1^{+} mesoderm through a haemogenic endothelial precursor to a haematopoietic stem/progenitor that can differentiate into lymphoid, myeloid or erythroid lineages. (F) Endothelial potential of TALE-expressing VEcad^{+} cells, as a percentage of –dox control cells. *P<0.01 (Student's t-test), from three biological replicates.

Development (2014) 141, 4018-4030 doi:10.1242/dev.115709

(Fig. 4A). Other genes affected by repression of *PU.1* in CD41^{+} cKit^{hi} cells included the downregulation of *Ikaros* and *Lyl1*, as well as the upregulation of *Erg*, *Gata2* and *Myb* or an increase in the fraction of cells expressing the respective genes (Fig. 4A).

Having generated a total of 566 single-cell expression profiles from the TALE-VP64 and TALE-KRAB perturbation experiments, we next combined all the expression data to explore the potential of this substantial dataset for the identification of possible regulatory relationships. Pairwise all-against-all comparisons were performed as before by calculating Spearman rank correlation coefficients (Fig. 4B). This analysis placed *PU.1* next to a cluster of haematopoietic genes containing, among others, *Myb*, *Runx1* and *Ikaros*. A second cluster of strongly correlating genes consisted of endothelial genes (e.g. *Sox7*, *VEcad*, *Pecam1*). *Gata2* was adjacent to a third and somewhat smaller cluster consisting of erythroid genes such as *Gata1*, *Epb4.2* and globin genes. Of note, *PU.1* showed negative correlation with *Gata2*, as expected from the results in Fig. 4A, but not with core erythroid genes such as *Gata1*. Additionally, we observed negative correlation of *PU.1* with many genes within the 'endothelial' cluster, suggesting that *PU.1* might antagonise endothelial fate.

To further assess possible effects of *PU.1* expression perturbations on the entire multi-dimensional gene expression dataset from all 566 cells, we performed principal component analysis (PCA). PCA separated the two mCherry^{-} populations (VEcad^{+} and CD41^{+} cKit^{hi}) into distinct groups across principal component 1 (PC1), consistent with the notion of two developmentally distinct populations (Fig. 4C). This separation is driven by the expression of endothelial genes in the VEcad^{+} population (including *VEcad*, *Sox17*, *Sox7*) and haematopoietic TFs in the CD41^{+} cKit^{hi} population (including *Runx1*, *Myb*, *Gfi1b*, *Ikaros* and *PU.1*) (Fig. 4D). The CD41^{+} cKit^{hi} population is resolved into two populations by PC2, by expression of myeloid genes (including *PU.1* and *Csf1r*) and erythroid genes (including *Hbb-bh1*, *Gata1*, *Epor*), suggesting that the CD41^{+} cKit^{hi} population contains myeloid- and erythroid-biased CD41^{+} cKit^{hi} progenitor cells. PCA of our dataset therefore provided good resolution of early developmental populations based on current models of developmental haematopoietic specification [Fig. 4E, based on Medvinsky et al. (2011)]. Interestingly, T-VP64-*PU.1-14* mCherry^{+} VEcad^{+} cells bridge the separation between the control VEcad^{+} and CD41^{+} cKit^{hi} populations (Fig. 4C), consistent with the notion that *PU.1* expression pushes VEcad^{+} cells to haematopoietic commitment but is unable to drive the transition completely. By contrast, the separation of the T-KRAB-*PU.1-14* mCherry^{+} CD41^{+} cKit^{hi} population from the mCherry^{-} CD41^{+} cKit^{hi} population is less striking, although more *PU.1* repressed cells are closer to the VEcad^{+} population and none form part of the most distant group of cells in the top right-hand part of the plot (Fig. 4C), consistent with the block observed in haematopoietic maturation.

Both the pairwise correlation analysis and PCA suggested that *PU.1* expression contributes to a haematopoietic fate in VEcad^{+} cells. We therefore assessed the effect of *PU.1* perturbation on the endothelial potential of the day 6 VEcad^{+} cells. TALE-VP64-mediated *PU.1* activation inhibited endothelial colony formation, whereas *PU.1* repression did not (Fig. 4F; supplementary material Fig. S5). Combined, these data suggest that activation of *PU.1* expression during developmental haematopoiesis plays a role in driving a haematopoietic rather than endothelial transcriptional programme, and that activation of *PU.1* expression in haemogenic endothelium might be an important molecular decision in haematopoietic commitment.

Such a large single-cell gene expression dataset presented the opportunity to investigate underlying TF network interactions active during the endothelial-to-haematopoietic transition (EHT) using partial correlation analysis. This analysis identifies network interactions (edges) by detecting irreducible statistical dependencies between TFs that cannot be otherwise explained by other statistical dependencies within the network (see supplementary material Tables S1 and S2). To visualise the results, we plotted the 34 network edges between the TF nodes with highly significant correlations (P<0.0001; Fig. 5). Although this method of analysis provides positive/negative correlation information, directionality cannot be inferred. Most TF interactions were positive and formed a highly interconnected network, which could be important in network stabilisation. Two types of negative correlations were observed: (1) between haematopoietic genes and endothelial genes, including *Runx1* and *Sox17*; and (2) between haematopoietic lineage-specific genes, including *Nfe2* and *Gata3*. Such TF antagonisms might be important switches in cell fate commitment.

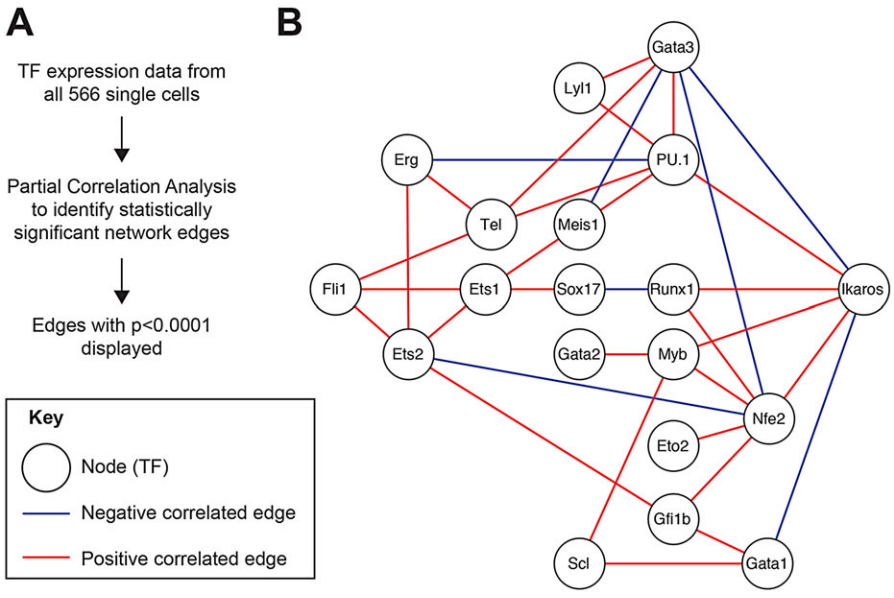


Fig. 5. Partial correlation analysis identifies a highly interconnected TF network that is active during the EHT. (A) Method used to build the TF network model in B. (B) TF network model showing highly statistically significant interactions (P<0.0001) as connections (edges) between TFs (nodes).

Development (2014) 141, 4018-4030 doi:10.1242/dev.115709

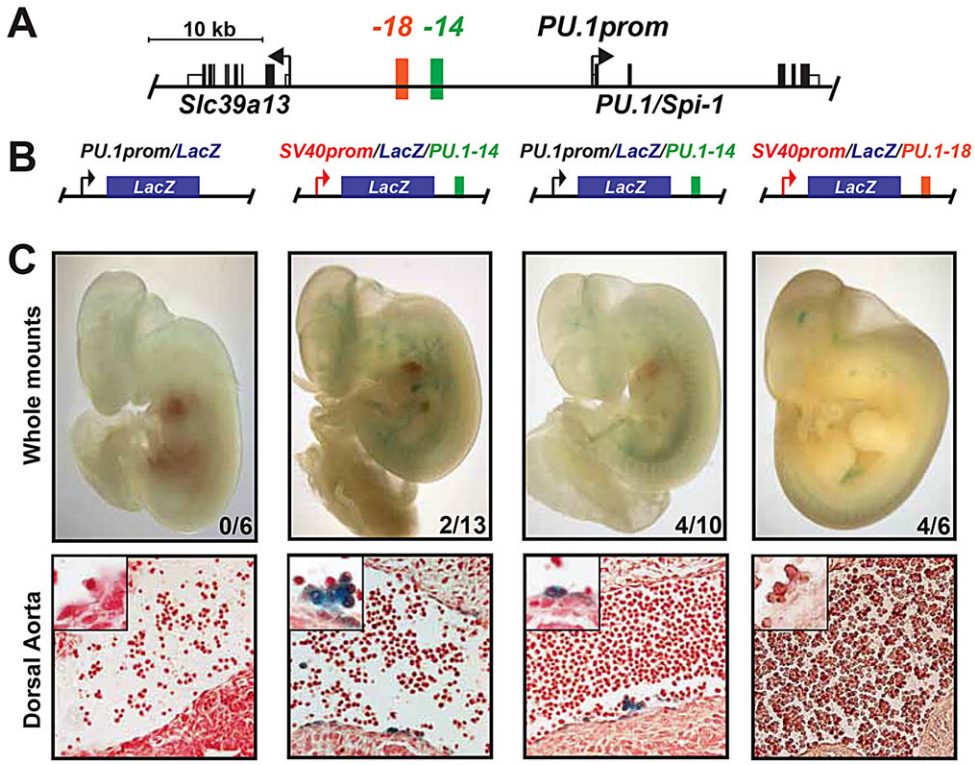


Fig. 6. The *PU.1-14kb* element is active at sites of mouse definitive haematopoiesis *in vivo*. (A) Schematic of the *PU.1* locus highlighting the relevant *cis*-regulatory elements. (B) Reporter constructs used for transient transgenic embryo generation. (C) Representative *lacZ*$^{+}$ whole-mount images and section images (original magnification: 40×) of the dorsal aorta of E11.5 transgenic embryos carrying the reporters illustrated above in B. Insets (original magnification: 100×) show cell clusters budding from the ventral side of the dorsal aorta. The number of *lacZ*$^{+}$ embryos/number of total PCR^{+} embryos analysed is indicated in each whole-mount image.

The *PU.1-14kb* enhancer is active in the mid-gestation dorsal aorta *in vivo*

The above data suggest that *PU.1* plays a role in driving endothelial cells to a haematopoietic fate. Although *PU.1* expression at sites of embryonic definitive haematopoiesis has been reported (Huang et al., 2008; Swiers et al., 2013), very little is known about its transcriptional regulation during developmental haematopoiesis *in vivo*. We therefore generated E11.5 transient transgenic mouse embryos carrying *lacZ* reporter gene constructs. The *PU.1* promoter alone (*PU.1prom/LacZ*) was unable to drive *lacZ* expression (Fig. 6A-C), suggesting that distal enhancers govern *PU.1* expression at this developmental stage. To determine the activity of the *PU.1-14kb* enhancer, we initially generated transgenic mouse embryos carrying the enhancer downstream of a *lacZ* reporter driven by the SV40 minimal promoter (*SV40prom/LacZ/PU.1-14*; Fig. 6B). In E11.5 mouse embryos, *PU.1-14kb* drove *lacZ* expression in cell clusters budding from the dorsal aorta (DA), a site of definitive haematopoiesis (Fig. 6C). *lacZ* expression was also seen in a minority of endothelial cells within the DA and rare circulating blood cells (Fig. 6C). These data demonstrate that *PU.1-14kb* is activated *in vivo* at embryonic sites of the EHT. By comparison, previous analysis of *Scl+40kb* failed to identify activity in the endothelium or budding clusters of the DA at E11.5, although primitive blood cells and HSPCs residing in the foetal liver later in development were *lacZ*$^{+}$ (Ogilvy et al., 2007).

To confirm the ability of the *PU.1-14kb* enhancer to drive expression from the *PU.1* promoter, we also generated transient transgenic embryos carrying *PU.1-14kb* downstream of the *PU.1* promoter driving *lacZ* (*PU.1prom/LacZ/PU.1-14*; Fig. 6B). Embryos analysed at E11.5 also showed expression in cell clusters budding from the DA and endothelial patches of the DA, but expression in circulating cells was less obvious than with *SV40prom/LacZ/PU.1-14* (Fig. 6C). As a control, we generated transgenic embryos carrying a different putative *PU.1 cis*-regulatory element, *PU.1-18kb* (Chou et al., 2009; Zarnegar and Rothenberg, 2012), downstream of *lacZ* driven by the SV40 minimal promoter (*SV40/LacZ/PU.1-18*; Fig. 6A,B). E11.5 *SV40/LacZ/PU.1-18* embryos failed to display staining in the DA or circulating blood (Fig. 6C), although non-haematopoietic *lacZ* expression was detected (such as in the neural tube). We conclude that *PU.1-14kb* activates the *PU.1* promoter during the EHT, although other *cis*-regulatory elements are likely to be involved in regulating PU.1 expression during this process as *PU.1-14kb* is unable to confer physiological PU.1 expression in *PU.1*$^{-/-}$ mice (Leddin et al., 2011).

PU.1 is detectable within endothelium of the aorta-gonad-mesonephros region and vitelline and umbilical arteries, and is upregulated during the EHT *in vivo*

Single-cell gene expression analysis has recently been used to study the EHT in the developing embryo (Swiers et al., 2013) using a *Runx1+23kb* enhancer-reporter (Runx1+23GFP) transgenic mouse line. Combined with surface markers, this reporter line allows further resolution of the cell types involved in the EHT: endothelium (VEcad^{+} Runx1+23GFP^{-} Ter119^{-} CD41^{-} CD45^{-}), haemogenic endothelium (VEcad^{+} Runx1+23GFP^{+} Ter119^{-} CD41^{-} CD45^{-}), committing HSPCs (VEcad^{+} Runx1+23GFP^{+} Ter119^{-} CD41^{+} CD45^{-}) and mature HSPCs (VEcad^{+} Runx1 +23GFP^{+} Ter119^{-} CD41^{+} CD45^{+}). We re-analysed this dataset to specifically study *PU.1* expression during haematopoietic commitment in the aorta-gonad-mesonephros region and vitelline and umbilical arteries (AGM+VUA) from E10.5 embryos (Fig. 7A). As controls, we also re-analysed gene expression for *Fli1* (a TF expressed in endothelium and blood) and *Myb* (a haematopoietic TF). Low expression of *PU.1* was detectable in ~5% of endothelium and ~19% of haemogenic endothelium. Over 80% of early committing HSPCs expressed higher levels of *PU.1*, and by the mature HSPC stage 97% of cells expressed robust levels of *PU.1*. This expression profile was consistent at E8.5, E9.5 and E11.5 (data not shown) (Swiers et al., 2013). These expression dynamics were similar to those of the haematopoietic gene *Myb*,

Development (2014) 141, 4018-4030 doi:10.1242/dev.115709

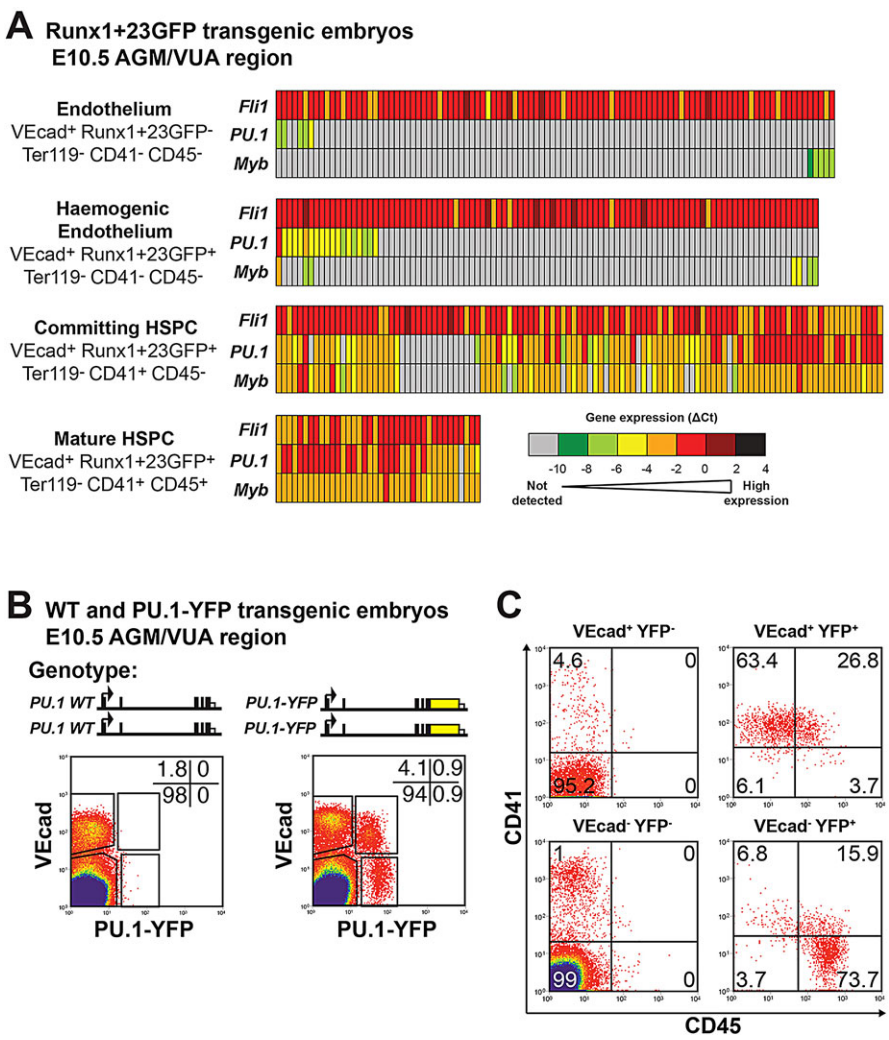


Fig. 7. PU.1 expression is induced during the EHT *in vivo*. (A) *Fli1*, *PU.1* and *Myb* single-cell RT-qPCR gene expression data from Swiers et al. (2013) in the AGM region and vitelline and umbilical artery (AGM +VUA) for VEcad$^+$ cell populations from E10.5 embryos carrying a Runx1+23kb-GFP enhancer reporter. Gene expression levels are displayed as a heatmap of ΔCt relative to housekeeping genes (*Ubc* and *Atp5a1*). (B) Genotype schematics and flow cytometry plots displaying VEcad versus PU.1-YFP expression for the Ter119$^-$ (Ly76$^-$) cell population from the AGM+VUA of E10.5 WT (left) and homozygous PU.1-YFP transgenic (right) embryos. Data were collected from pooled embryos and are representative of two E10-10.5 embryo litters. (C) Flow cytometry plots displaying CD41 versus CD45 expression for the populations within the VEcad/PU.1-YFP gates in B for the homozygous PU.1-YFP transgenic embryos.

although it is interesting to note that, within the haemogenic endothelium population, only 7% of cells expressed *Myb*, which is lower than the 19% that were *PU.1* expressing. By the committing and mature HSPC states, *Myb* and *PU.1* expression almost entirely overlap. These data suggest that *in vivo*, *PU.1* can be expressed in the early haemogenic endothelial stages of the EHT, and is upregulated during this haematopoietic cell fate commitment, concomitant with CD41 surface expression.

Finally, to confirm this early expression of PU.1 at the protein level during the EHT, we analysed PU.1 expression in the AGM +VUA region of E10-10.5 *PU.1-YFP* knock-in homozygous mouse embryos (Kirstetter et al., 2006) by flow cytometry (Fig. 7B). PU.1-YFP was detectable in 0.6% of VEcad$^+$ Ter119$^-$ CD41$^-$ CD45$^-$ endothelium, 82% of VEcad$^+$ Ter119$^-$ CD41$^+$ CD45$^-$ committing HPSCs and 97% of VEcad$^+$ Ter119$^-$ CD41$^+$ CD45$^+$ mature HSPCs (Fig. 7B,C; supplementary material Fig. S6). These results correlate well with the *in vivo* and *in vitro* single-cell qPCR data (Fig. 3, Fig. 4 and Fig. 7A). Taken together, our results establish PU.1 and its *–14kb* enhancer as likely key building blocks of the gene regulatory network that drives early blood specification.

DISCUSSION

Here, we demonstrate a novel use of TALEs in combination with single-cell gene expression profiling to investigate the consequences of transcriptional perturbation of developmental regulatory networks. High-throughput RT-qPCR coupled with comprehensive cellular assays uncovered a previously unrecognised role for *PU.1* in the EHT. Transgenic mouse studies confirmed that the *PU.1-14kb* enhancer is active at mid-gestation in the DA, where the EHT occurs *in vivo*, and that PU.1 expression can be detected from the endothelium stage of this process.

Our analysis of several hundred single cells with TALE induction highlights the efficiency of TALE-mediated endogenous gene expression perturbations at the single-cell level, which proved to be comparable to alternative methods of perturbation such as siRNA knockdown (Kouno et al., 2013), and provided more physiologically relevant expression changes when comparing upregulation of gene expression using TALE-VP64 proteins with retroviral cDNA overexpression. Moreover, TALE-mediated perturbation does not require a distinction between exogenous and endogenous cDNAs, allows normal co- and post-transcriptional processing to occur, and allows for detection by gene expression primers that are located in untranslated regions (UTRs).

The CRISPR-Cas9 system has recently been adapted to modulate gene expression by a similar mechanism to TALEs (Gilbert et al., 2013; Maeder et al., 2013; Perez-Pinera et al., 2013; Qi et al., 2013). As CRISPR-Cas9 target specificity is based on guide RNAs rather than a modular protein domain, the generation of these ‘designer’ TFs is faster than the assembly of TALEs (Gaj et al., 2013). However, a recent comparison between CRISPR-Cas9 and TALEs

Development (2014) 141, 4018-4030 doi:10.1242/dev.115709

suggested higher targeting specificity for the latter (Fu et al., 2013; Mali et al., 2013).

The majority of previous research on PU.1 has concerned its role in adult haematopoiesis, where high PU.1 levels promote terminal myeloid differentiation, and reduced PU.1 expression results in proliferation [reviewed by Mak et al. (2011)], consistent with our CFU data. A novel link between PU.1 levels and the cell cycle has been described recently, in which PU.1, by regulating cell cycle lengthening, determines PU.1 protein accumulation within the cell, effecting lympho-myeloid cell fate decisions (Kueh et al., 2013). Our data additionally highlight the importance of tightly regulated *PU.1* expression for early haematopoiesis to occur. Moreover, since our TALEs target conserved DNA sequences within *cis*-regulatory elements, these tools can be directly applied to manipulate human haematopoiesis.

PU.1 has recently been shown to inhibit proliferation by directly controlling cell cycle regulators (Staber et al., 2013), which is consistent with our observed loss of haematopoietic colonies after TALE-VP64-mediated *PU.1* upregulation and increase in haematopoietic colonies after TALE-KRAB-mediated *PU.1* repression. Our gene selection for the single-cell expression analysis was focused on TF networks controlling early haematopoietic development, and it is likely that genes other than those assayed contribute to the phenotypic changes caused by *PU.1* expression perturbations, and this might well include cell cycle regulators. It is worth highlighting that the TALE-mediated perturbations caused consistent gene expression changes in single cells, suggesting that *PU.1* operates within a tightly interconnected haematopoietic TF network.

In adult haematopoiesis, *PU.1-14kb* is a known target of Runx1 (Huang et al., 2008), a TF crucial for definitive haematopoiesis (Okuda et al., 1996; Wang et al., 1996). We now report that the *PU.1-14kb* element is active *in vivo* in mid-gestation AGM endothelium and blood clusters, where definitive HSPCs arise, and PU.1 is expressed in a subset of these haematopoietic precursors. Runx1 has been shown to initiate chromatin unfolding at *PU.1-14kb* early during haematopoietic specification (Hoogenkamp et al., 2009). Such enhancer priming is likely to be important for efficient TALE-VP64-mediated induction of expression. Although Hoogenkamp et al. were unable to determine the frequency of such priming events within the precursor population, our data would be consistent with a model whereby the majority of VEcad$^+$ cells contain primed *PU.1-14kb* enhancers owing to the high efficiency of TALE-mediated *PU.1* expression activation in this cell type. Enhancer priming might also contribute to the low-level expression of *PU.1* prior to haematopoietic commitment, with more robust expression seen later as haematopoietic TF network circuitry is reinforced. Such low-level expression might be analogous to the transcriptional noise of lineage regulators previously seen in adult haematopoietic progenitor cells (Pina et al., 2012).

In summary, we have validated the use of TALEs targeting conserved *cis*-regulatory elements as an efficient, multifaceted tool to modulate endogenous gene expression and study TF regulatory network perturbations in single cells, and in doing so we have uncovered a role for PU.1 in haematopoietic specification.

MATERIALS AND METHODS

TALE design and assembly

TALE sequences were designed to 20 bp regions within the *Scl+40kb* and *PU.1-14kb* elements that were conserved between human and mouse, and were unique within both genomes by BlastN and Blat (Altschul et al., 1990; Kent, 2002). TALEs were assembled and cloned into piggyBac (PB) as described previously (Gao et al., 2013). For ChIP experiments, an HA-tagged TALE-VP64 was used, as described previously (Gao et al., 2013).

Cell culture and transfection

K562 (Lozzio et al., 1981) and 416B (Dexter et al., 1979) cells were cultured and electroporated as described previously (Knezevic et al., 2011; Moignard et al., 2013), as in Fig. 1D. TALE expression was induced by addition of 2 µg/ml dox (Sigma) for 48 h and mCherry$^+$ and mCherry$^-$ K562/416B cells were sorted using a BD Influx (BD Biosciences). Ainv18 ESCs (Kyba et al., 2002) were cultured as described previously (Ismailoglu et al., 2008) and transfected by nucleofection (Lonza). OP9 stromal cells were cultured as described previously (Nakano et al., 1994).

ESC differentiation and haematopoietic colony forming assays

ESCs were differentiated essentially as described previously (Wilkinson et al., 2013), and treated as described in Fig. 2A and Fig. 3A. 100,000 day 6 EB cells were plated in triplicate in 1.1 ml M3434 Methocult (Stem Cell Technologies). For OP9 co-culture colony forming assays, 100,000 day 6 EB cells were plated on confluent OP9 in MEMα supplemented with 10% FCS for 24 h before the medium was replaced with M3434 Methocult. Definitive haematopoietic colonies were counted after 10-12 days (see supplementary material Fig. S2A for representative images of these colonies and counting criteria).

Flow cytometry of EBs

Dissociated EB cells were stained with combinations of the antibodies listed in supplementary material Table S3. Annexin V-APC (BD Biosciences, 550474) antibody and DAPI were used to assess cell apoptosis according to the manufacturers' instructions. Surface marker and mCherry expression data were collected using a five-laser LSRFortessa (BD Biosciences) and analysed using FlowJo software.

OP9 co-culture assays

MACS (Miltenyi) sorted Flk1$^+$ cells from day 4 EBs (over 95% purity) were cultured on confluent OP9s in MEMα supplemented with 10% FCS. Mesodermal colonies were allowed to form for 36 h before dox was added. Haematopoietic cells were not seen before this time point. After 48 h, cells were fixed in 2% PFA overnight, blocked, stained with purified CD41 antibody (BD Biosciences, 553847, MWReg30), visualised by DAB staining, and mesodermal colonies containing (at least two) small rounded budding CD41$^+$ haematopoietic cells were scored. Specific staining was confirmed using an isotype control antibody. VEcad$^+$ cells from day 6 EBs were sorted using a BD Influx and plated on confluent OP9s, and cultured for 4 days in MEMα supplemented with 10% FCS. After 4 days, cells were fixed and stained as above using purified CD31 antibody (BD Biosciences, 553370, MEC13.3), and endothelial sheet colonies were scored.

RNA extraction and RT-qPCR

RT-qPCR was undertaken as described previously (Wilson et al., 2010a) using the primers listed in supplementary material Table S4. All gene expression values are normalised to *ACTB*/*Actb*.

Single-cell gene expression and data analysis

Single-cell sorting and gene expression analysis were undertaken as previously described (Moignard et al., 2013) using a BD Influx and TaqMan assays (see supplementary material Table S5). Partial correlation analysis, as described elsewhere (Bailey, 1995), was applied to the normalised TF gene expression data using Spearman correlation. Correlation values and associated *P*-values are displayed in supplementary material Tables S1 and S2, respectively. Interactions where $P<0.0001$ are displayed as edges in Fig. 6.

ChIP

mCherry$^+$ 416B cells were sorted by FACS (24 h after dox addition), then expanded without dox before HA-T-VP64 expression was induced for 48 h and cells crosslinked using 1% formaldehyde (Sigma). ChIP-qPCR and ChIP-seq were performed and analysed essentially as described previously

Development (2014) 141, 4018-4030 doi:10.1242/dev.115709

(Wilson et al., 2010b) using anti-HA (Sigma, H6908) and anti-H3K27Ac (Abcam, ab4729) antibodies, and the ChIP-qPCR primers listed in supplementary material Table S6. ChIP-seq data are available at GEO with accession number GSE61189.

Transient transgenic embryo generation

Promoter and enhancer reporter pGlac constructs were cloned as described previously (Wilson et al., 2009) using the primers listed in supplementary material Table S7. E11.5 transient transgenic embryos were generated by Cyagen Biosciences (Guangzhou, China) and analysed as described (Ogilvy et al., 2007).

PU.1-YFP embryo generation and analysis

Mice were housed with free access to food and water. All procedures were in compliance with United Kingdom Home Office regulations and approved by the Local Ethical Review Committee and the Home Office. Timed matings were set up overnight between wild-type (CBA×C57BL/6)/F1 male and female and homozygous PU.1-YFP male and female mice (Kirstetter et al., 2006). Dissection and flow cytometry analysis were undertaken as described previously (Swiers et al., 2013).

Acknowledgements

We thank Claus Nerlov and Sten Eirik Jacobsen for PU.1-YFP mice; Reiner Schulte, Veronika Romashova and Chiara Cossetti at the CIMR Flow Cytometry Core for their expertise with cell sorting; and Cheuk-Ho Tsang, Juan Li and David Kent for helpful suggestions.

Competing interests

The authors declare no competing financial interests.

Author contributions

A.C.W., V.K.S.K., Y.T., G.S., M.F.T.R.d.B., P.L. and B.G. designed the experiments. A.C.W., V.K.S.K., J.S., X.G., S.A., C.B., S.W., R.H., G.S. and V.M. performed experiments and analysed data. A.C.W., V.K.S.K. and B.G. wrote the manuscript. All authors, including J.F., S.H. and M.R.T., contributed to editing the manuscript.

Funding

Research in the authors' laboratories was supported by Leukaemia and Lymphoma Research [12029], The Wellcome Trust [098051], Cancer Research UK [C1163/A12765], the Biotechnology and Biological Sciences Research Council [BB/I00050X/1], the National Institute of Health Research [RP-PG-0310-10002], Microsoft Research [2012-023], the Medical Research Council [G0900951] and core support grants by the Wellcome Trust to the Cambridge Institute for Medical Research [100140/Z/12/Z] and Wellcome Trust–MRC Cambridge Stem Cell Institute. V.K.S.K. was supported by a Japan Society for the Promotion of Science (JSPS) Research Fellowship for Young Scientists. Deposited in PMC for immediate release.

Supplementary material

Supplementary material available online at
http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.115709/-/DC1

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410.

Bailey, N. T. J. (1995). *Statistical Methods in Biology*. Cambridge, UK: Cambridge University Press.

Beerli, R. R., Segal, D. J., Dreier, B. and Barbas, C. F., III. (1998). Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc. Natl. Acad. Sci. USA* **95**, 14628-14633.

Boch, J. and Bonas, U. (2010). Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annu. Rev. Phytopathol.* **48**, 419-436.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A. and Bonas, U. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509-1512.

Chou, S. T., Khandros, E., Bailey, L. C., Nichols, K. E., Vakoc, C. R., Yao, Y., Huang, Z., Crispino, J. D., Hardison, R. C., Blobel, G. A. et al. (2009). Graded repression of PU.1/Sfpi1 gene transcription by GATA factors regulates hematopoietic cell fate. *Blood* **114**, 983-994.

Cong, L., Zhou, R., Kuo, Y.-C., Cunniff, M. and Zhang, F. (2012). Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nat. Commun.* **3**, 968.

Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., Hanna, J., Lodato, M. A., Frampton, G. M., Sharp, P. A. et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. USA* **107**, 21931-21936.

Delabesse, E., Ogilvy, S., Chapman, M. A., Piltz, S. G., Gottgens, B. and Green, A. R. (2005). Transcriptional regulation of the SCL locus: identification of an enhancer that targets the primitive erythroid lineage in vivo. *Mol. Cell. Biol.* **25**, 5215-5225.

Dexter, T. M., Allen, T. D., Scott, D. and Teich, N. M. (1979). Isolation and characterisation of a bipotential haematopoietic cell line. *Nature* **277**, 471-474.

Ferreira, R., Spensberger, D., Silber, Y., Dimond, A., Li, J., Green, A. R. and Gottgens, B. (2013). Impaired in vitro erythropoiesis following deletion of the Scl (Tal1) +40 enhancer is largely compensated for in vivo despite a significant reduction in expression. *Mol. Cell. Biol.* **33**, 1254-1266.

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K. and Sander, J. D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **31**, 822-826.

Gaj, T., Gersbach, C. A. and Barbas, C. F., III. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **31**, 397-405.

Gao, X., Yang, J., Tsang, J. C. H., Ooi, J., Wu, D. and Liu, P. (2013). Reprogramming to pluripotency using designer TALE transcription factors targeting enhancers. *Stem Cell Rep.* **1**, 183-197.

Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., Stern-Ginossar, N., Brandman, O., Whitehead, E. H., Doudna, J. A. et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442-451.

Guo, G., Huss, M., Tong, G. Q., Wang, C., Li Sun, L., Clarke, N. D. and Robson, P. (2010). Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. *Dev. Cell* **18**, 675-685.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H. and Glass, C. K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576-589.

Hoogenkamp, M., Lichtinger, M., Krysinska, H., Lancrin, C., Clarke, D., Williamson, A., Mazzarella, L., Ingram, R., Jorgensen, H., Fisher, A. et al. (2009). Early chromatin unfolding by RUNX1: a molecular explanation for differential requirements during specification versus maintenance of the hematopoietic gene expression program. *Blood* **114**, 299-309.

Huang, G., Zhang, P., Hirai, H., Elf, S., Yan, X., Chen, Z., Koschmieder, S., Okuno, Y., Dayaram, T., Growney, J. D. et al. (2008). PU.1 is a major downstream target of AML1 (RUNX1) in adult mouse hematopoiesis. *Nat. Genet.* **40**, 51-60.

Ismailoglu, I., Yeamans, G., Daley, G. Q., Perlingeiro, R. C. R. and Kyba, M. (2008). Mesodermal patterning activity of SCL. *Exp. Hematol.* **36**, 1593-1603.

Keller, G., Kennedy, M., Papayannopoulou, T. and Wiles, M. V. (1993). Hematopoietic commitment during embryonic stem cell differentiation in culture. *Mol. Cell. Biol.* **13**, 473-486.

Kent, W. J. (2002). BLAT – the BLAST-like alignment tool. *Genome Res.* **12**, 656-664.

Kirstetter, P., Anderson, K., Porse, B. T., Jacobsen, S. E. W. and Nerlov, C. (2006). Activation of the canonical Wnt pathway leads to loss of hematopoietic stem cell repopulation and multilineage differentiation block. *Nat. Immunol.* **7**, 1048-1056.

Knezevic, K., Bee, T., Wilson, N. K., Janes, M. E., Kinston, S., Polderdijk, S., Kolb-Kokocinski, A., Ottersbach, K., Pencovich, N., Groner, Y. et al. (2011). A Runx1-Smad6 rheostat controls Runx1 activity during embryonic hematopoiesis. *Mol. Cell. Biol.* **31**, 2817-2826.

Kouno, T., de Hoon, M., Mar, J. C., Tomaru, Y., Kawano, M., Carninci, P., Suzuki, H., Hayashizaki, Y. and Shin, J. W. (2013). Temporal dynamics and transcriptional control using single-cell gene expression analysis. *Genome Biol.* **14**, R118.

Kueh, H. Y., Champhekhar, A., Nutt, S. L., Elowitz, M. B. and Rothenberg, E. V. (2013). Positive feedback between PU.1 and the cell cycle controls myeloid differentiation. *Science* **341**, 670-673.

Kyba, M., Perlingeiro, R. C. R. and Daley, G. Q. (2002). HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors. *Cell* **109**, 29-37.

Leddin, M., Perrod, C., Hoogenkamp, M., Ghani, S., Assi, S., Heinz, S., Wilson, N. K., Follows, G., Schonheit, J., Vockentanz, L. et al. (2011). Two distinct auto-regulatory loops operate at the PU.1 locus in B cells and myeloid cells. *Blood* **117**, 2827-2838.

Lozzio, B. B., Lozzio, C. B., Bamberger, E. G. and Feliu, A. S. (1981). A multipotential leukemia cell line (K-562) of human origin. *Exp. Biol. Med.* **166**, 546-550.

Maeder, M. L., Linder, S. J., Cascio, V. M., Fu, Y., Ho, Q. H. and Joung, J. K. (2013). CRISPR RNA-guided activation of endogenous human genes. *Nat. Methods* **10**, 977-979.

Mak, K. S., Funnell, A. P. W., Pearson, R. C. M. and Crossley, M. (2011). PU.1 and haematopoietic cell fate: dosage matters. *Int. J. Cell Biol.* **2011**, 808524.

Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., Yang, L. and Church, G. M. (2013). CAS9 transcriptional activators for target

Development (2014) 141, 4018-4030 doi:10.1242/dev.115709

specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* **31**, 833-838.

Medvinsky, A., Rybtsov, S. and Taoudi, S. (2011). Embryonic origin of the adult hematopoietic system: advances and questions. *Development* **138**, 1017-1031.

Moignard, V., Macaulay, I. C., Swiers, G., Buettner, F., Schütte, J., Calero-Nieto, F. J., Kinston, S., Joshi, A., Hannah, R., Theis, F. J. et al. (2013). Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis. *Nat. Cell Biol.* **15**, 363-372.

Nakano, T., Kodama, H. and Honjo, T. (1994). Generation of lymphohematopoietic cells from embryonic stem cells in culture. *Science* **265**, 1098-1101.

Ogilvy, S., Ferreira, R., Piltz, S. G., Bowen, J. M., Gottgens, B. and Green, A. R. (2007). The SCL +40 enhancer targets the midbrain together with primitive and definitive hematopoiesis and is regulated by SCL and GATA proteins. *Mol. Cell. Biol.* **27**, 7206-7219.

Okuda, T., van Deursen, J., Hiebert, S. W., Grosveld, G. and Downing, J. R. (1996). AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell* **84**, 321-330.

Okuno, Y., Huang, G., Rosenbauer, F., Evans, E. K., Radomska, H. S., Iwasaki, H., Akashi, K., Moreau-Gachelin, F., Li, Y., Zhang, P. et al. (2005). Potential autoregulation of transcription factor PU.1 by an upstream regulatory element. *Mol. Cell. Biol.* **25**, 2832-2845.

Papantonis, A. and Cook, P. R. (2013). Transcription factories: genome organization and gene regulation. *Chem. Rev.* **113**, 8683-8705.

Perez-Pinera, P., Kocak, D. D., Vockley, C. M., Adler, A. F., Kabadi, A. M., Polstein, L. R., Thakore, P. I., Glass, K. A., Ousterout, D. G., Leong, K. W. et al. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* **10**, 973-976.

Pimanda, J. E. and Göttgens, B. (2010). Gene regulatory networks governing haematopoietic stem cell development and identity. *Int. J. Dev. Biol.* **54**, 1201-1211.

Pina, C., Fugazza, C., Tipping, A. J., Brown, J., Soneji, S., Teles, J., Peterson, C. and Enver, T. (2012). Inferring rules of lineage commitment in haematopoiesis. *Nat. Cell Biol.* **14**, 287-294.

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P. and Lim, W. A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173-1183.

Rosenbauer, F., Wagner, K., Kutok, J. L., Iwasaki, H., Le Beau, M. M., Okuno, Y., Akashi, K., Fiering, S. and Tenen, D. G. (2004). Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. *Nat. Genet.* **36**, 624-630.

Schütte, J., Moignard, V. and Göttgens, B. (2012). Establishing the stem cell state: insights from regulatory network analysis of blood stem cell development. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **4**, 285-295.

Staber, P. B., Zhang, P., Ye, M., Welner, R. S., Nombela-Arrieta, C., Bach, C., Kerenyi, M., Bartholdy, B. A., Zhang, H., Alberich-Jorda, M. et al. (2013). Sustained PU.1 levels balance cell-cycle regulators to prevent exhaustion of adult hematopoietic stem cells. *Mol. Cell* **49**, 934-946.

Swiers, G., Baumann, C., O'Rourke, J., Giannoulatou, E., Taylor, S., Joshi, A., Moignard, V., Pina, C., Bee, T., Kokkaliaris, K. D. et al. (2013). Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level. *Nat. Commun.* **4**, 2924.

Tijssen, M. R., Cvejic, A., Joshi, A., Hannah, R. L., Ferreira, R., Forrai, A., Bellissimo, D. C., Oram, S. H., Smethurst, P. A., Wilson, N. K. et al. (2011). Genome-wide analysis of simultaneous GATA1/2, RUNX1, FLI1, and SCL binding in megakaryocytes identifies hematopoietic regulators. *Dev. Cell* **20**, 597-609.

Toriello, N. M., Douglas, E. S., Thaitrong, N., Hsiao, S. C., Francis, M. B., Bertozzi, C. R. and Mathies, R. A. (2008). Integrated microfluidic bioprocessor for single-cell gene expression analysis. *Proc. Natl. Acad. Sci. USA* **105**, 20173-20178.

Wang, Q., Stacy, T., Binder, M., Marin-Padilla, M., Sharpe, A. H. and Speck, N. A. (1996). Disruption of the Cbfa2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis. *Proc. Natl. Acad. Sci. USA* **93**, 3444-3449.

Wang, W., Lin, C., Lu, D., Ning, Z., Cox, T., Melvin, D., Wang, X., Bradley, A. and Liu, P. (2008). Chromosomal transposition of PiggyBac in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. USA* **105**, 9290-9295.

Wilkinson, A. C. and Göttgens, B. (2013). Transcriptional regulation of haematopoietic stem cells. *Adv. Exp. Med. Biol.* **786**, 187-212.

Wilkinson, A. C., Goode, D. K., Cheng, Y.-H., Dickel, D. E., Foster, S., Sendall, T., Tijssen, M. R., Sanchez, M.-J., Pennacchio, L. A., Kirkpatrick, A. M. et al. (2013). Single site-specific integration targeting coupled with embryonic stem cell differentiation provides a high-throughput alternative to in vivo enhancer analyses. *Biol. Open* **2**, 1229-1238.

Wilson, N. K., Miranda-Saavedra, D., Kinston, S., Bonadies, N., Foster, S. D., Calero-Nieto, F., Dawson, M. A., Donaldson, I. J., Dumon, S., Frampton, J. et al. (2009). The transcriptional program controlled by the stem cell leukemia gene Scl/Tal1 during early embryonic hematopoietic development. *Blood* **113**, 5456-5465.

Wilson, N. K., Timms, R. T., Kinston, S. J., Cheng, Y.-H., Oram, S. H., Landry, J.-R., Mullender, J., Ottersbach, K. and Gottgens, B. (2010a). Gfi1 expression is controlled by five distinct regulatory regions spread over 100 kilobases, with Scl/Tal1, Gata2, PU.1, Erg, Meis1, and Runx1 acting as upstream regulators in early hematopoietic cells. *Mol. Cell. Biol.* **30**, 3853-3863.

Wilson, N. K., Foster, S. D., Wang, X., Knezevic, K., Schütte, J., Kaimakis, P., Chilarska, P. M., Kinston, S., Ouwehand, W. H., Dzierzak, E. et al. (2010b). Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell* **7**, 532-544.

Zarnegar, M. A. and Rothenberg, E. V. (2012). Ikaros represses and activates PU.1 cell-type-specifically through the multifunctional Sfpi1 URE and a myeloid specific enhancer. *Oncogene* **31**, 4647-4654.

Zhang, F., Cong, L., Lodato, S., Kosuri, S., Church, G. M. and Arlotta, P. (2011). Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat. Biotechnol.* **29**, 149-153.

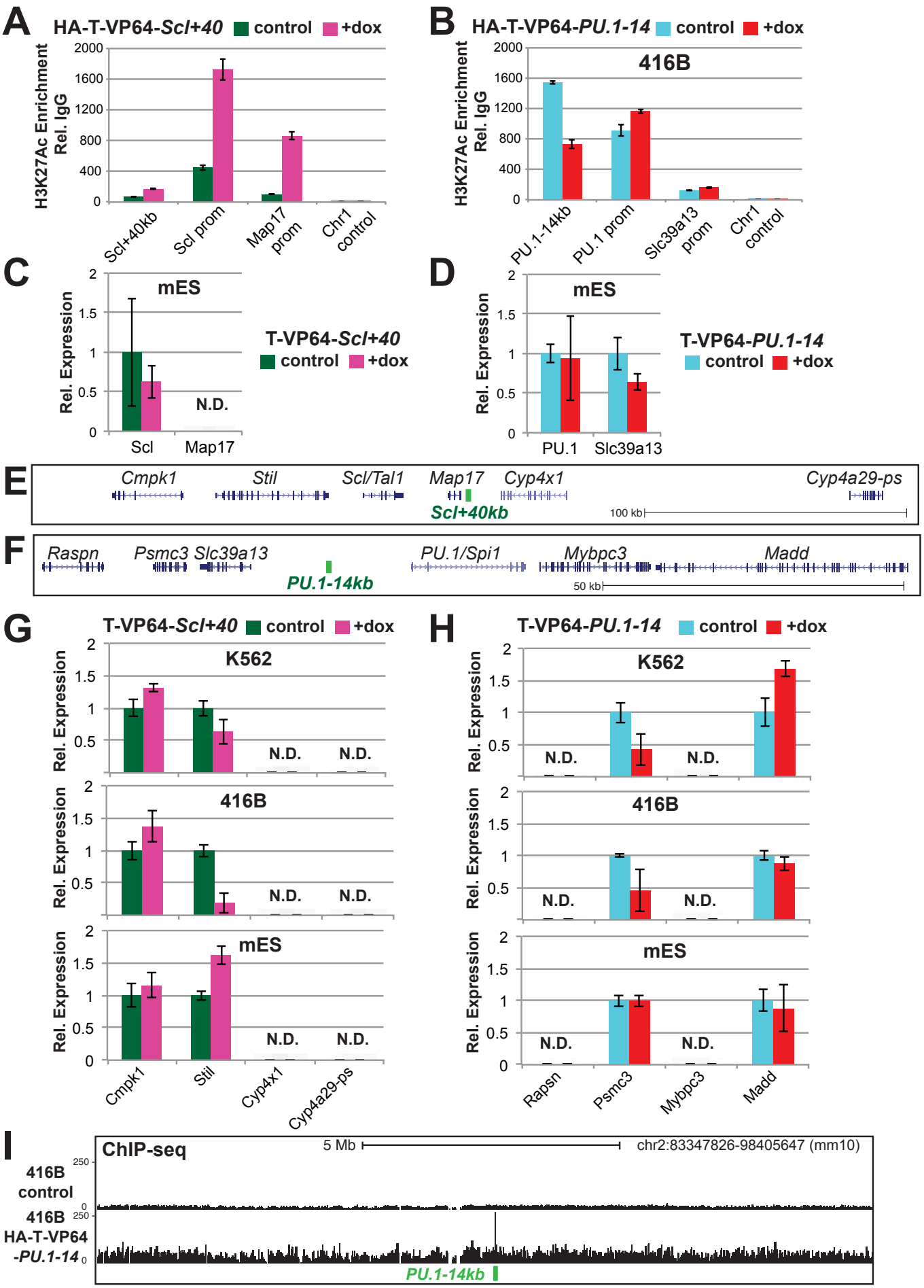
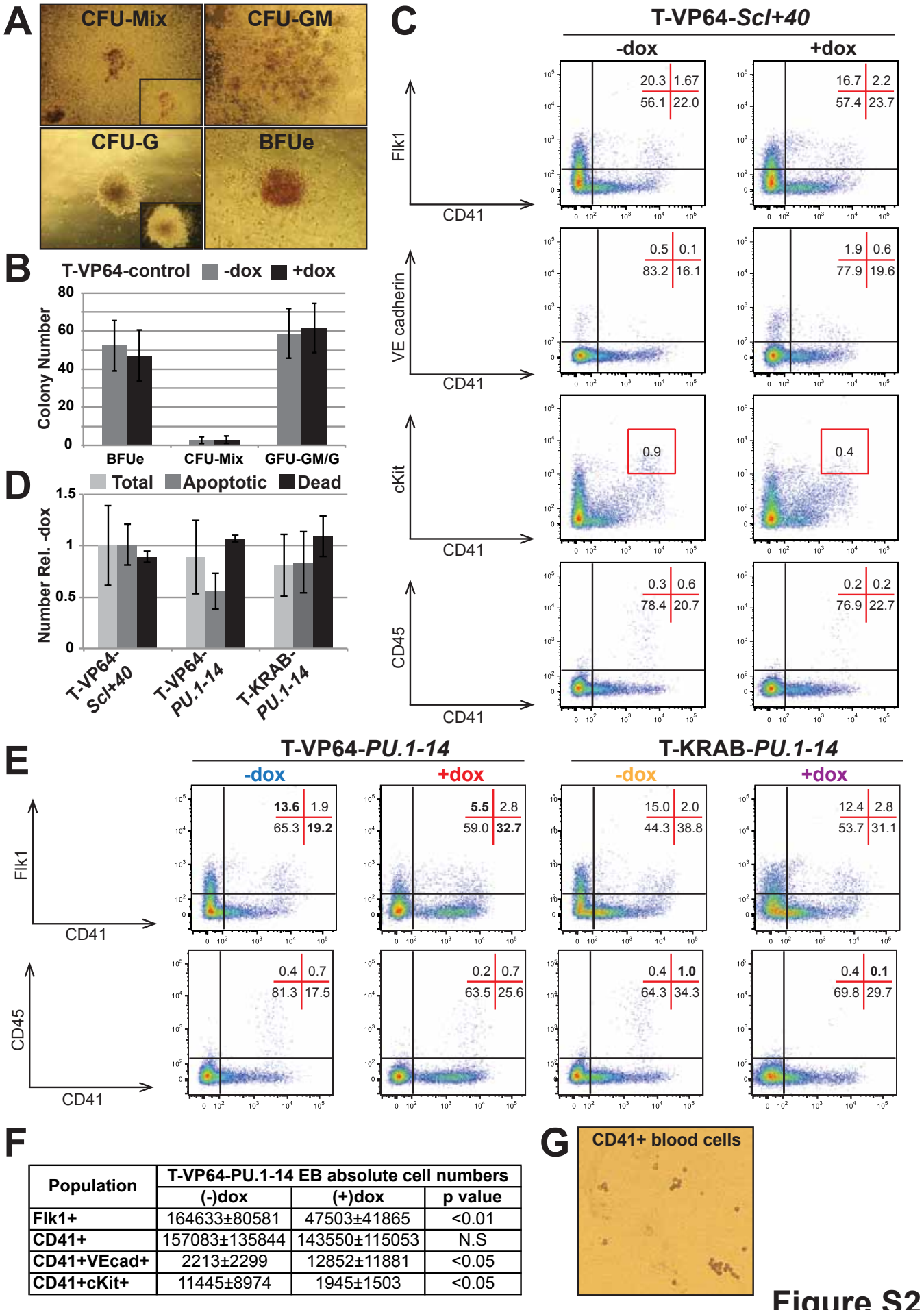

A
HA-T-VP64-Scl+40
control
+dox
H3K27Ac Enrichment Rel. IgG
2000
1600
1200
800
400
0
Scl+40kb
Scl prom
Map17 prom
Chr1 control
B
HA-T-VP64-PU.1-14
control
+dox
416B
H3K27Ac Enrichment Rel. IgG
PU.1-14kb
PU.1 prom
Slc39a13 prom
Chr1 control
C
mES
Rel. Expression
T-VP64-Scl+40
control
+dox
N.D.
Scl
Map17
D
mES
Rel. Expression
T-VP64-PU.1-14
control
+dox
PU.1
Slc39a13
E
Cmpk1
Stil
Scl/Tal1
Map17
Cyp4x1
Cyp4a29-ps
Scl+40kb
100 kb
F
Raspn
Psmc3
Slc39a13
PU.1/Spi1
Mybpc3
Madd
PU.1-14kb
50 kb
G
T-VP64-Scl+40
control
+dox
K562
416B
mES
Rel. Expression
N.D.
Cmpk1
Stil
Cyp4x1
Cyp4a29-ps
H
T-VP64-PU.1-14
control
+dox
K562
416B
mES
Rel. Expression
N.D.
Rapsn
Psmc3
Mybpc3
Madd
I
ChIP-seq
5 Mb
chr2:83347826-98405647 (mm10)
416B control
416B HA-T-VP64 -PU.1-14
250
0
PU.1-14kb

Figure S1, related to Figure 1:
(A) Effect of expressing HA-TALE-VP64 targeting the *Scl+40kb* (HA-T-VP64-*Scl+40*) for 48 hr in mouse 416B on histone 3 lysine 27 acetylation (H3K27Ac; relative to IgG enrichment) at the *Scl+40kb*, *Scl* promoter (prom), *Map17* prom and a control region on chromosome 1 (*chr1* control). HA-T-VP64-*Scl+40* expressed for 48 hr by addition of doxycycline (+dox; pink bars). Untransfected 416B used as control (green bars). Error bars are standard deviation of technical triplicates from one biological experiment.
(B) Effect of expressing HA-TALE-VP64 targeting the *PU.1-14kb* (HA-T-VP64-*PU.1-14*) in mouse 416B on histone 3 lysine 27 acetylation (H3K27Ac) ChIP-qPCR enrichment (relative to IgG enrichment) at the *PU.1-14kb*, *PU.1* prom, *Slc39a13* prom and *Chr1* control region. HA-T-VP64-*PU.1-14* expressed for 48 hr by addition of dox (+dox; red bars). Untransfected 416B used as control (blue bars). Error bars are standard deviation of technical triplicates from one biological experiment.
(C) Effect of expressing TALE-VP64 targeting the *Scl+40kb* (T-VP64-*Scl+40*) in mouse Ainv18 ES cells on *Scl* and *Map17* gene expression, normalised to *ActB*. T-VP64-*Scl+40* expressed for 48 hr by addition of dox and gene expression in +dox cells (green bars) determined relative to –dox control cells (pink bars). Error bars are standard deviation of technical triplicates from one biological experiment.
(D) Effect of expressing TALE-VP64 targeting the *PU.1-14kb* (T-VP64-*PU.1-14*) in mouse Ainv18 ES cells on *PU.1* and *Slc39a13* gene expression, normalised to *ActB*. T-VP64-*PU.1-14* expressed for 48 hr by addition of dox and gene expression in +dox cells (blue bars) determined relative to –dox control cells (red bars). Error bars are standard deviation of technical triplicates from one biological experiment.
(E) UCSC Genome Browser screen shot of RefSeq annotated genes within a ~300kb genomic window surrounding the *Scl+40kb* enhancer (highlighted in green).
(F) UCSC Genome Browser screen shot of RefSeq annotated genes within a ~150kb genomic winder surrounding the *PU.1-14kb* enhancer (highlighted in green).
(G) Effect of expressing T-VP64-*Scl+40* in human K562 (top), mouse 416B and mouse Ainv18 ES cells on *Cmpk1*, *Stil*, *Cyp4x1* and *Cyp4a29-ps* gene expression, normalised to *ACTB/ActB*. T-VP64-*Scl+40* expressed for 48 hr by addition of dox and gene expression in +dox/mCherry^{+} cells (green bars) determined relative to –dox/mCherry^{-} control cells (pink bars). Error bars are standard deviation of technical triplicates from one biological experiment. Not detected; N.D
(H) Effect of expressing T-VP64-*PU.1-14* in human K562 (top), mouse 416B and mouse Ainv18 ES cells on *Rapsn*, *Psmc3*, *Mybpc3* and *Madd* gene expression, normalised to *ACTB/ActB*. T-VP64-*PU.1-14* expressed for 48 hr by addition of dox and gene expression in +dox/mCherry^{+} cells (blue bars) determined relative to –dox/mCherry^{-} control cells (red bars). Error bars are standard deviation of technical triplicates. Not detected; N.D
(I) UCSC Genome Browser screenshot of HA antibody ChIP-Seq enrichment in a ~15Mb window surrounding the *PU.1-14kb* element in wild type (above) and T-VP64-*PU.1-14* expressing (below) 416B cells.



Population	T-VP64-PU.1-14 EB absolute cell numbers		
	(-)dox	(+)dox	p value
Flk1+	164633±80581	47503±41865	<0.01
CD41+	157083±135844	143550±115053	N.S
CD41+VEcad+	2213±2299	12852±11881	<0.05
CD41+cKit+	11445±8974	1945±1503	<0.05

Figure S2

Figure S2, related to Figure 2:
(A) Representative images of haematopoietic colonies scored in methylcellulose CFU assays in Figures 2C and S2B. Red coloured erythroid colonies of at least ~30 small cells dispersed within small clusters with tight cell-cell junctions were scored as burst forming unit erythroid (BFUe). Colonies of at least ~50 small round bright cells (often tightly packed with grey centre) were scored as colony forming unit-granulocyte (CFU-G). Large colonies of over ~200 cells containing both granulocytes (as described above) and macrophages (large round cells, less bright than granulocytes and often more dispersed) were scored as colony forming unit-granulocyte macrophage (CFU-GM). Large colonies of over ~200 cells, densely packed, including red erythroid cells (similar to those described above) as well as at least two other lineages (usually granulocytes and macrophages described above) or megakaryocytes were scored as colony forming unit-mix (CFU-Mix).
(B) Representative haematopoietic colonies numbers from 1 x 10^5 day 6 EB cells derived from mouse ES cells inducible expressing a non-functional TALE-VP64 (due to mutations within the DNA binding domain), previously generated (Gao et al., 2013). As in Figure 2C, dox added to EBs day 4 to induce TALE-VP64 expression. Colonies grown in methylcellulose supplemented with SCF, IL-3, IL-6 and Epo. Error bars are standard deviation of technical triplicates. No statistically significant changes in CFU numbers were seen from three biological triplicates, as determined by the student t test.
(C) Flow cytometry plots of day 6 EB cells showing (from top to bottom) Flk1 vs. CD41, VE cadherin (VEcad) vs. CD41, cKit vs. CD41, CD45 vs. CD41). Representative staining patterns for a T-VP64-*Scl+40* expressing mouse ES cell line, both uninduced (-dox) and induced (from day 4; +dox). Distribution of cells within quadrant/gates shown as percentages.
(D) Total number (light grey bars), frequency of apoptotic (Annexin V^+ $DAPI^-$; dark grey bars) and frequency of dead (Annexin V^+ $DAPI^+$; black bars) T-VP64-*Scl+40*, T-VP64-*PU.1-14* and T-KRAB-*PU.1-14* expressing EB day 6 cells (+dox from day 4) relative to –dox controls. Error bars are standard deviation of three biological replicates. No statistically significant changes were seen from three biological triplicates, as determined by the student t test.
(E) Flow cytometry plots of day 6 EB cells showing Flk1 vs. CD41 (top) and CD41 vs. CD45 (bottom). Representative staining patterns for TALE-VP64 (left) and TALE-KRAB (right) targeting *PU.1-14kb* clones, both uninduced (-dox) and induced (from day 4; +dox). Distribution of cells within quadrant/gates shown as percentages.
(F) Table displaying absolute cells numbers for cells populations identified by flow cytometry in Figure 2D ± standard deviation from three biological replicates, and p values (using the student t test). N.S; not significant.
(G) Representative image of day 4 EB $Flk1^+$-derived colony containing haematopoietic (round, budding) $CD41^+$ (stained black) cells scored in Figure 2E. Colonies containing endothelial cells that stained weakly $CD41^+$ were not scored unless haematopoietic cells were also present.

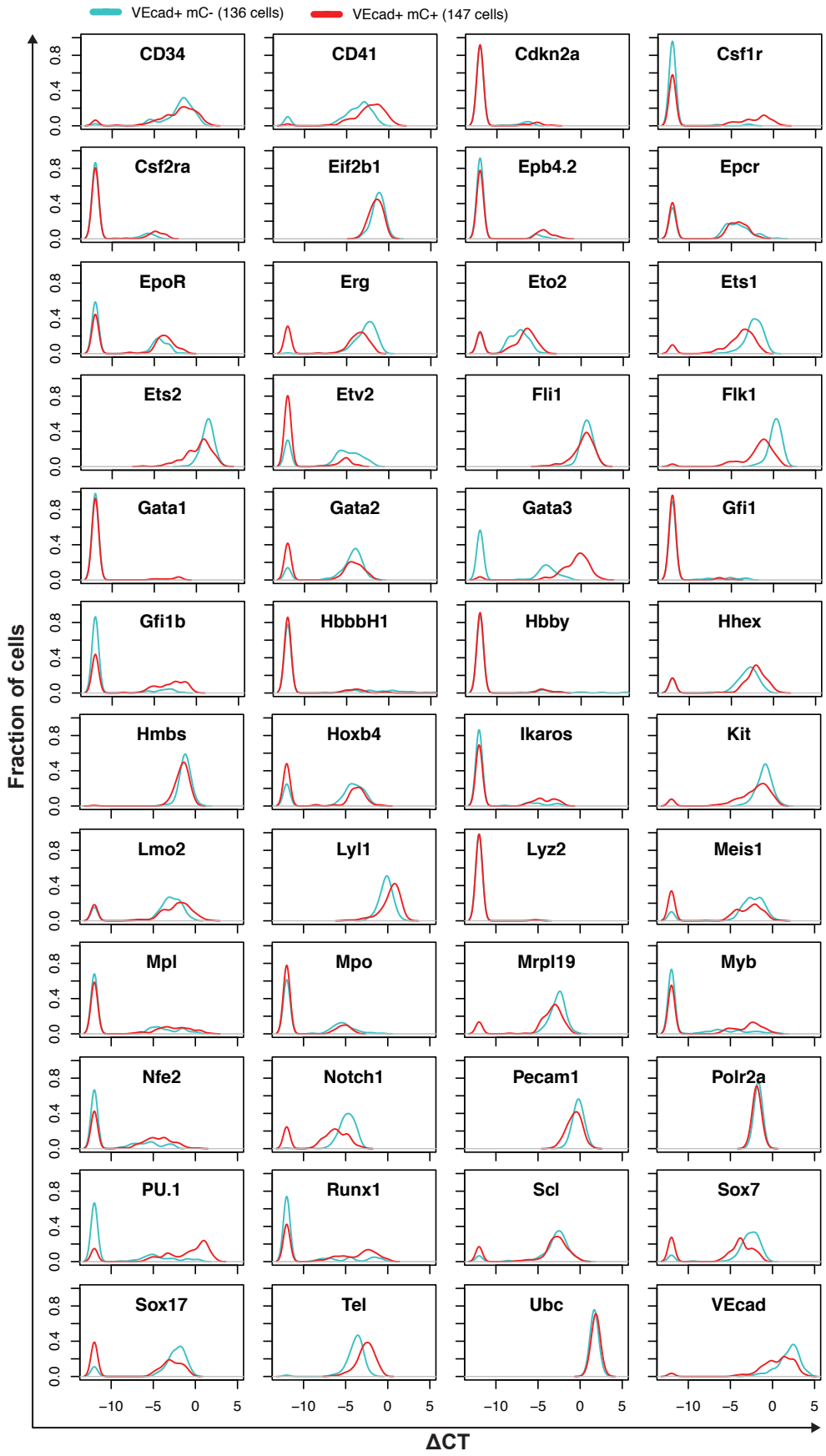


Figure S3

Figure S3, related to Figure 3:
Density plots of gene expression in day 6 EB VEcad^{+} mCherry^{-} (136 WT Ainv18; in cyan) and VEcad^{+} mCherry^{+} (147 Ainv18 expressing T-VP64-*PU.1-14*; in red) for all 48 genes analysed. The density indicates the fraction of cells at each expression level, relative to housekeeping genes (*Polr2a* and *Ubc*). Cells with non-detected gene expression set to -12.

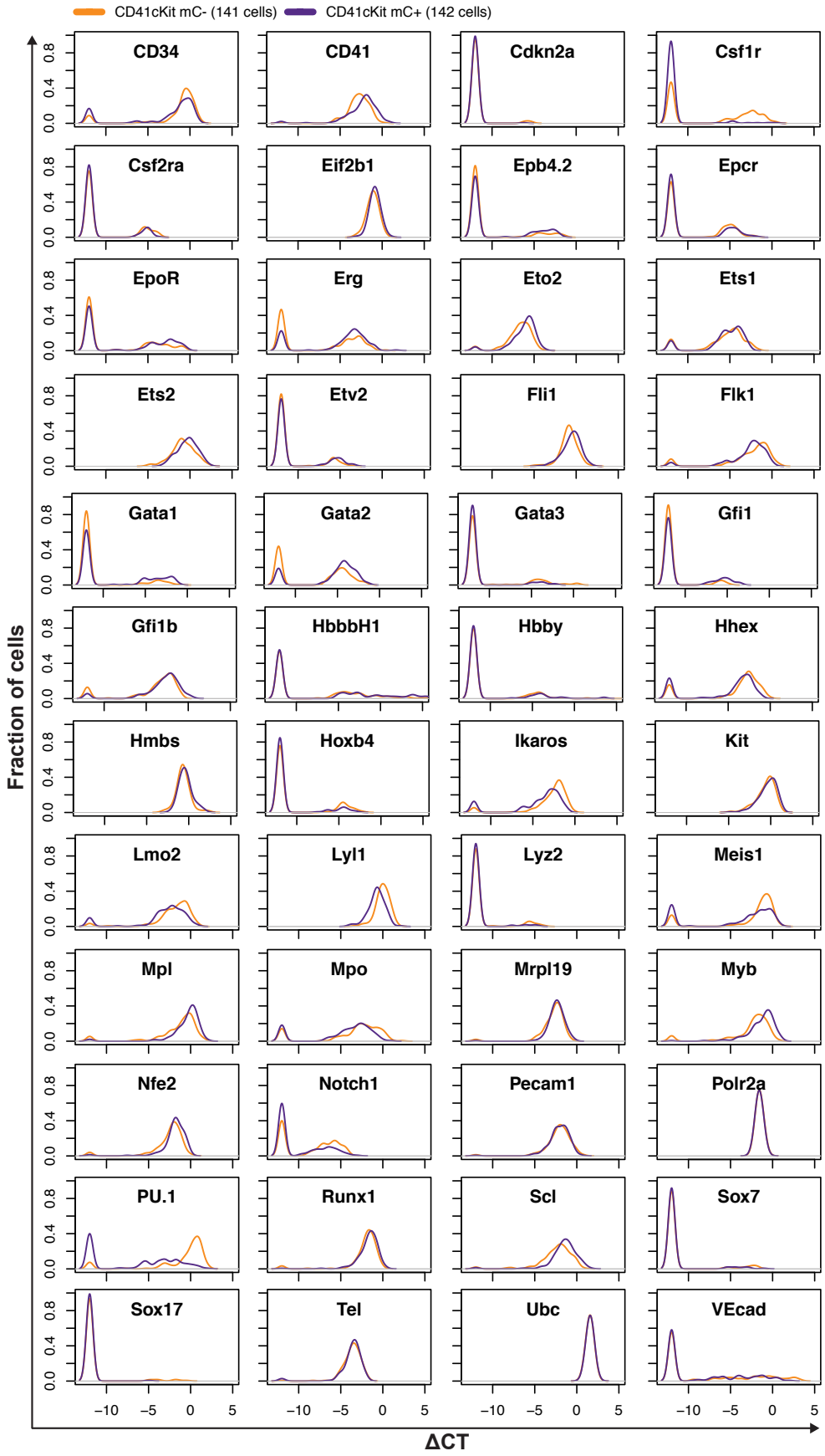


Figure S4

Figure S4, related to Figure 4:
Density plots of gene expression in day 6 EB CD41^{+}cKit^{hi} mCherry^{-} (141 WT Ainv18; in orange) and CD41^{+}cKit^{hi} mCherry^{+} (142 Ainv18 expressing T-KRAB-*PU.1-14*; in purple) for all 48 genes analysed. The density indicates the fraction of cells at each expression level, relative to housekeeping genes (*Polr2a* and *Ubc*). Cells with non-detected gene expression set to -12.

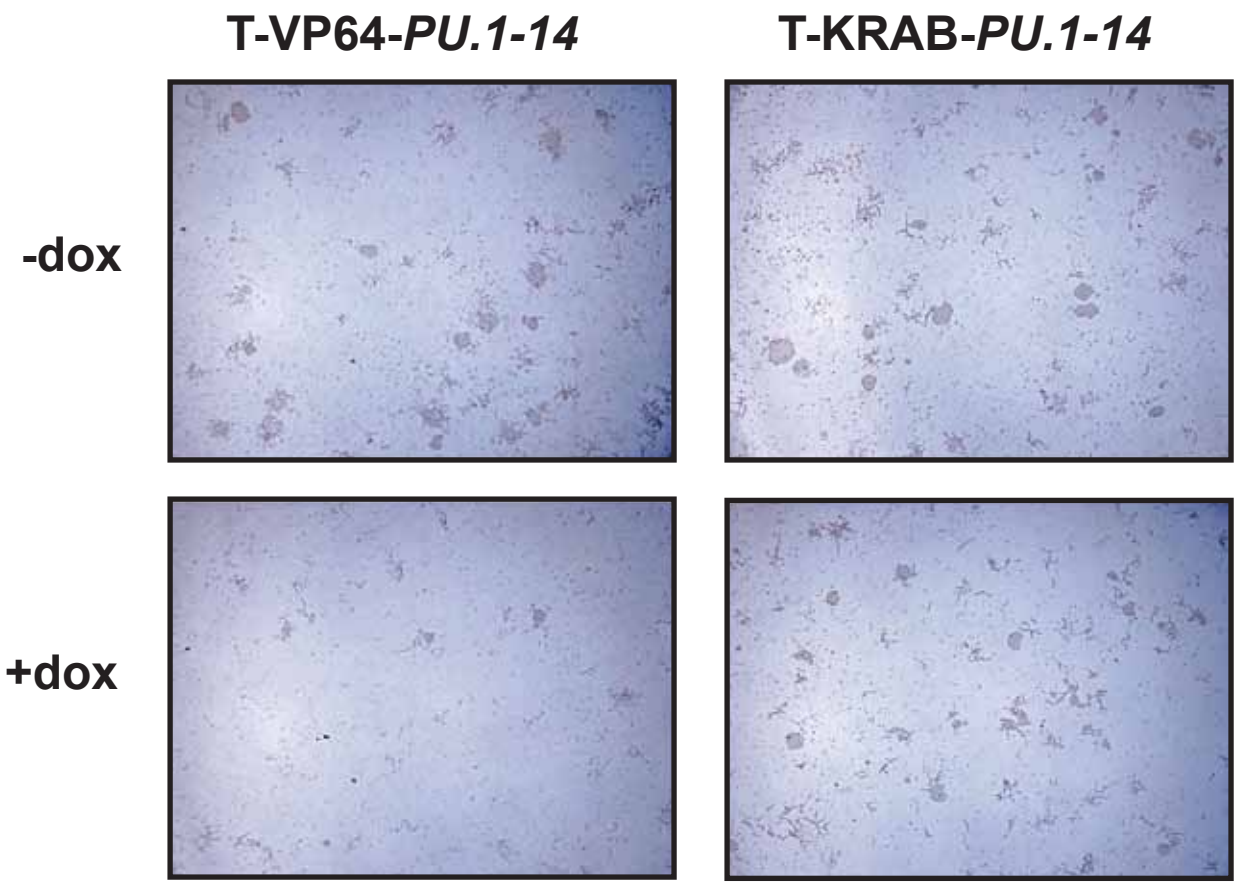


Figure S5, related to Figure 4:
Representative images of endothelial colony assays from day 6 EB VEcad^{+} cells from T-VP64-*PU.1-14* or T-KRAB-*PU.1-14* ES cell lines as described in Figure 4.

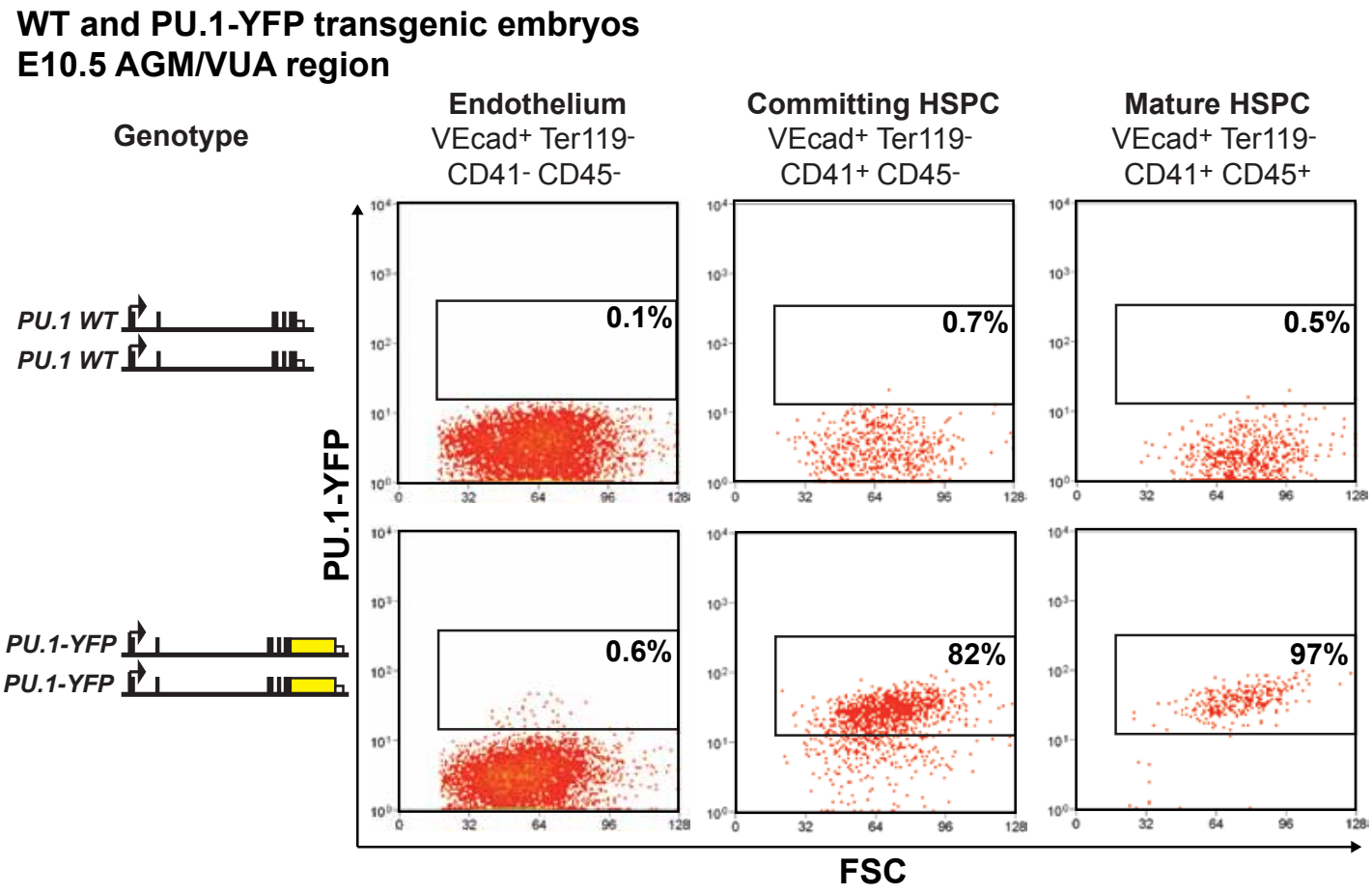


Figure S6, related to Figure 7:
Alternative analysis of flow cytometry data from Figure 6B,C, plots displaying PU.1-YFP expression vs. FSC for VEcad+ Ter119- CD41- CD45- (left), VEcad+ Ter119- CD41+ CD45- (middle) and VEcad+ Ter119- CD41+ CD45+ (right) populations from the AG-M+VUA region of pooled E10.5 WT (above) and PU.1-YFP (below) embryos.

Supplementary Table 1: Partial Correlation Analysis

Download Table S1

Supplementary Table 2: Partial Correlation Analysis

Download Table S2

Supplementary Table 3: Flow Cytometry Antibodies

Antibody	Supplier	Cat. Number	Clone ID	Reactivity
CD16/32 (Fc block)	BD	553142	2.4G2	Mouse
CD41-PE-Cy7	Biolegend	25-0411-82	eBioMWReg30	Mouse
CD41-APC	eBioscience	17-0411	MWReg30	Mouse
CD45-APC	Biolegend	103112	30-F11	Mouse
Flk1-APC	BD	560070	Avas 12a1	Mouse
Flk1-PE	BD	561052	Avas 12a1	Mouse
VEcad-PE-Cy7	Biolegend	92055	VECD1	Mouse
cKit-APC	BD	553356	2B8	Mouse

Supplementary Table 4: RT-qPCR primers

Primer name	Sequence
Human ACTB forward	AGAGCTACGAGCTGCCTGAC
Human ACTB reverse	AGCACTGTGTTGGCGTACAG
Human SCL/TAL1 forward	TTCCCTATGTTCACCACCAA
Human SCL/TAL1 reverse	AAGATACGCCGCACAACTTT
Human MAP17 forward	TGCCTATGAGAATGTGCCG
Human MAP17 reverse	TGGACATCCATCCCATGTGC
Human PU.1/SPI-1 forward	CGGCTGGATGTTACAGGCGTG
Human PU.1/SPI-1 reverse	TCGTGCGTTTGGCGTTGG
Human SLC39A13 forward	TTCCCCTAGAGATGGGGACC
Human SLC39A13 reverse	GGCAGCAGATGCAGAAACAC
Human PSMC3 forward	ACAGACGTACTTCCTTCC
Human PSMC3 reverse	CCAATGAACATCTGCACCAG
Human RAPSN forward	GTACGACTCCGCCATGAGCA
Human RAPSN reverse	ATGGCATCCAGAGCCTTGTCC
Human MYBPC3 forward	GCTCTTCCAGACCCATCTCG
Human MYBPC3 reverse	CAGCGGGATGACAGGAAACA
Human MADD forward	TAGTGATCGTAGGGGCCAGG
Human MADD reverse	GCAGGGGAAACTCAGTGTGA
Human STIL forward	ATGCACATAACGTGGATCACG
Human STIL reverse	TCCATGCTCAAATCCACACC
Human CMPK1 forward	TCTCATGAAGCCGCTGGT
Human CMPK1 reverse	TCCTGCAGAAAGGTGTGTGT
Human CYP4X1 forward	TCAGGACACAAGCGTGGAGGTCTA
Human CYP4X1 reverse	TGCATAAGGATCATGGGTGCTGTT
Human CYP4A29-PS forward	CTGCTTTTCAAGGCAGCACA
Human CYP4A29-PS forward	CTGCAAGCAATGCCCAAAGA
Mouse Actb forward	TCCTGGCCTCACTGTCCAC
Mouse Actb reverse	GTCCGCCTAGAAGCACTTGC
Mouse Scl/Tal1 forward	CATGTTCACCAACAACAACCG
Mouse Scl/Tal1 reverse	GGTGTGAGGACCATCAGAAATCTC
Mouse Map17 forward	GTCCTTGTTGCAATCGTCTTC
Mouse Map17 reverse	GAGGAGTATCTGCCATCCATTC
Mouse PU.1/Sfpi-1 forward	AGAGCATACCAACGTCCAATG C
Mouse PU.1/Sfpi-1 reverse	GTGCGGAGAAATCCCAGTAGTG
Mouse Slc39a13 forward	TTGCTGGTCATTCCCCTGGA
Mouse Slc39a13 reverse	GTCCACCTAAGGCAAAGCTGA
Mouse Psmc3 forward	GACCGTGTGGGATGAAGCTG
Mouse Psmc3 reverse	CGCTGGACAATCTCTTCCGTG
Mouse Rapsn forward	ATATCGGGCCATGAGCCAGT
Mouse Rapsn reverse	TCACAACACTCCATGGCACTGC
Mouse Mybpc3 forward	TGAAGGGTCAGTCTCGGTAACC
Mouse Mybpc3 reverse	TCCTGTGGTCGCATCAGAAA
Mouse Madd forward	AAGAAACTGGGCATCCCTCG
Mouse Madd reverse	GAAGGGCACTGGACTTCTCC
Mouse Stil forward	GGTGATGATCAAGAGCCCGA
Mouse Stil reverse	ACCAGGTTCTTTGCTCTGCT
Mouse Cmpk1 forward	TCAGAAGCGCGTTGTATGCT
Mouse Cmpk1 reverse	AAAACGAACACGACCAACGG
Mouse Cyp4x1 forward	CCTGGACATAATAATGAAATGTGCTT
Mouse Cyp4x1 reverse	CTTCACGTAAGACTCATAGGTGCC
Mouse Cyp4a29-ps forward	CAGTGCACCATCTGGACCTC
Mouse Cyp4a29-ps reverse	GATTACGTAATAGTGGTCCCTCAGG

Supplementary Table 5: Taqman assays used for single cell gene expression

Protein name	Assay ID
Csf1r/c-fms	Mm01266652_m1
CD34	Mm00519283_m1
CD41	Mm00439768_m1
Cdkn2a	Mm00494449_m1
Csf2ra	Mm00438331_g1
Eif2b1	Mm01199614_m1
Epb4.2	Mm00469107_m1
Epcr	Mm00440993_mH
EpoR	Mm00438760_m1
Erg	Mm01214246_m1
Eto2	Mm00486780_m1
Ets1	Mm01175819_m1
Ets2	Mm00468977_m1
Etv2	Mm00468389_m1
Fli1	Mm00484409_m1
Flk1	Mm01222421_m1
Gata1	Mm00484678_m1
Gata2	Mm00492300_m1
Gata3	Mm00484683_m1
Gfi1	Mm00515855_m1
Gfi1b	Mm00492318_m1
Hbb-bH1	Mm00756487_mH
Hbb-y	Mm00433936_g1
hHex	Mm00433954_m1
Hmbs	Mm01143545_m1
Hoxb4	Mm00657964_m1
Ikaros	Mm01187882_m1
Kit	Mm00445212_m1
Lmo2	Mm01281680_m1
Lyl1	Mm01247198_m1
Lyz2	Mm01612741_m1
Meis1	Mm00487659_m1
Mpl	Mm00440310_m1
Mpo	Mm00447886_m1
Mrpl19	Mm03048937_m1
Myb	Mm00501741_m1
Nfe2	Mm00801891_m1
Notch1	Mm00435249_m1
Pecam1	Mm01242584_m1
Polr2a	Mm00839493_m1
PU.1	Mm00488142_m1
Runx1	Mm01213405_m1
Scl/Tal1	Mm01187033_m1
Sox17	Mm04208182_m1
Sox7	Mm00776876_m1
Tel	Mm01261325_m1
Ubc	Mm01201237_m1
VE cadherin	Mm00486938_m1

Supplementary Table 6: ChIP-qPCR primers

Primer name	Sequence
Scl promoter F	CATGCGCACTCCAGCCTC
Scl promoter R	CACACACCGCCCAGAAGC
Map17 promoter F	ACAAGGTAGCACAGGACAGG
Map17 promoter R	TCCAAAACACCACCCCATCC
Scl+40kb F	CTCTTCCTTATCGCCAGCTC
Sc+40kb R	CAGCTGGTGCGTTATCAGTT
PU.1 promoter F	GGGCCATTGGCTTCCTTAGA
PU.1 promoter R	AATAGCTGTTCAGGCCCCAC
Slc39a13 promoter F	AGAGTAGGACGGAAGTGGGTA
Slc39a13 promoter R	CTGTTGATCCCGGTTCCCG
PU.1-14kb F	GCTGTTGGCGTTTTGCAAT
PU.1-14kb R	GGCCGGTGCCTGAGAAA
Chr1 control	CATAGATGAAGCTGCCACATAGGT
Chr1 control	GTGGGCAAGGACAAAGCATTA

Supplementary Table 7: Genomic PCR cloning primers

Primer	Sequence
PU.1prom_F	taaagatctCACCAGAGGGGACTGAGAAG
PU.1prom_R	taaaagcttGCTGAGCTCCAGGTTGGTC
PU.1-14kb_F	taaggatccGCTTGGGTGCTGGACTTAGA
PU.1-14kb_R	taagtcgacGACCTTTCCCTTGGGTGAGC
PU.1-18kb_F	taaggatccGGAGGGCTGTGTCAGTCAGA
PU.1-18kb_R	taagtcgacCTTCTTCCATCCCACAGAGC