© 2018. Published by The Company of Biologists Ltd | Development (2018) 145, dev162990. doi:10.1242/dev.162990

RESEARCH ARTICLE

A novel role for *Lyl1* in primitive erythropoiesis

Sung K. Chiu[1], Jesslyn Saw[1], Yizhou Huang[2,3], Stefan E. Sonderegger[1], Nicholas C. Wong[4,5], David R. Powell[5], Dominic Beck[2,3], John E. Pimanda[2,6,7], Cedric S. Tremblay[1,*] and David J. Curtis[1,8,*,‡]

ABSTRACT

Stem cell leukemia (*Scl* or *Tal1*) and lymphoblastic leukemia 1 (*Lyl1*) encode highly related members of the basic helix-loop-helix family of transcription factors that are co-expressed in the erythroid lineage. Previous studies have suggested that *Scl* is essential for primitive erythropoiesis. However, analysis of single-cell RNA-seq data of early embryos showed that primitive erythroid cells express both *Scl* and *Lyl1*. Therefore, to determine whether *Lyl1* can function in primitive erythropoiesis, we crossed conditional *Scl* knockout mice with mice expressing a Cre recombinase under the control of the Epo receptor, active in erythroid progenitors. Embryos with 20% expression of *Scl* from E9.5 survived to adulthood. However, mice with reduced expression of *Scl* and absence of *Lyl1* (double knockout; DKO) died at E10.5 because of progressive loss of erythropoiesis. Gene expression profiling of DKO yolk sacs revealed loss of *Gata1* and many of the known target genes of the SCL-GATA1 complex. ChIP-seq analyses in a human erythroleukemia cell line showed that LYL1 exclusively bound a small subset of SCL targets including *GATA1*. Together, these data show for the first time that *Lyl1* can maintain primitive erythropoiesis.

KEY WORDS: Primitive erythropoiesis, Transcription factors, SCL (TAL1), LYL1, Mouse

INTRODUCTION

Hematopoietic stem cells (HSCs) give rise to all mature blood cells, including red blood cells (RBCs), by a process known as definitive hematopoiesis. However, the first hematopoietic cells emerge in blood islands within the yolk sac at around embryonic day (E) 7.25 (Palis, 2014). This arrival of yolk sac-derived erythroid progenitors gives rise to a wave of maturing nucleated erythroid cells that sustain the embryo between E9.5 and E12.5. Primitive erythrocytes can be distinguished from adult RBCs in the circulation by the expression of primitive globins and enucleation (Palis, 2014). In contrast, definitive or adult erythrocytes, which express adult globins, are first detected in the fetal bloodstream at around E11.5 and become the major source of circulating RBCs beginning at E13.5 (Kingsley et al., 2004).

A number of transcription factors (TFs) are known to be essential for primitive erythropoiesis. Genetic knockout models demonstrate that the TFs stem cell leukemia (Scl, also known as Tal1) and Gata1, along with co-factors Lmo2 and Ldb1, are responsible for the initiation of primitive erythropoiesis (Dore and Crispino, 2011). These factors form a large multimeric complex in erythroid cells that activates downstream targets for erythroid cell growth and differentiation. The master regulatory role for this multimeric complex is further emphasized by the ability of Scl, Lmo2 and Gata1 to directly reprogram fibroblasts to erythroid cells that resemble primitive erythroid progenitors in the yolk sac (Capellera-Garcia et al., 2016).

Scl is a hematopoietic-specific basic helix-loop-helix (bHLH) TF that is essential for the initial specification of hematopoiesis and subsequent production of RBCs and platelets (Porcher et al., 2017). It is known to form complexes with different binding partners to regulate erythropoiesis (Tripic et al., 2009; Wadman et al., 1997). Its bHLH region is highly homologous to that of the related family member known as lymphoblastic leukemia 1 (Lyl1), which is also expressed in the erythroid lineage (Visvader et al., 1991). During hematopoietic development *Scl* is essential, with *Scl* knockout mice dying *in utero* because of the absence of primitive erythropoiesis and vascular defects (Robb et al., 1995; Shivdasani et al., 1995). In contrast, *Lyl1* knockout mice are viable, with a mild defect in stress erythropoiesis (Capron et al., 2011). These distinct knockout phenotypes can be explained by a crucial role for *Scl* in hematopoietic specification that cannot be replaced by *Lyl1* (Chan et al., 2007). However, once hematopoiesis is established, *Lyl1* can compensate for the loss of *Scl* in HSCs to maintain their function. It is only with the loss of both factors that there is rapid loss of HSCs due to increased apoptosis (Souroullas et al., 2009). This functional redundancy is consistent with genome-wide chromatin immunoprecipitation (ChIP)-seq studies showing shared binding of Scl and Lyl1 along with other crucial hematopoietic TFs in HSCs (Beck et al., 2013; Chacon et al., 2014; Wilson et al., 2010).

The role for *Scl* in primitive erythropoiesis has been difficult to delineate because of its crucial role in hematopoietic specification. Nevertheless, two studies suggest that, unlike in adult HSCs, *Lyl1* cannot adequately compensate for *Scl* in primitive erythropoiesis. First, deletion of *Scl* after the formation of hematopoiesis using *Scl* conditional knockout mice that express the *Tie2Cre* transgene revealed the presence of circulating primitive erythrocytes, but mice died at E13.5-14.5 with edema, hemorrhage and dysplastic immature primitive erythroblasts (Schlaeger et al., 2005). A second study overcame hematopoietic specification by generating *Scl* knock-in mice that carry a germline DNA-binding mutation, in which DNA binding is not required for specification of hematopoiesis (Kassouf et al., 2008). Here, most embryos succumbed at E14.5-17.5 with anemia and defects in erythroid maturation. However, several observations raise the possibility that

[1]Australian Centre for Blood Diseases, Monash University, Melbourne, VIC 3004, Australia. [2]Lowy Cancer Research Centre and the Prince of Wales Clinical School, University of New South Wales, Sydney, NSW 2052, Australia. [3]Centre for Health Technologies, School of Biomedical Engineering and the School of Software, University of Technology, Sydney, NSW 2007, Australia. [4]Central Clinical School, Monash University, Melbourne, VIC 3004, Australia. [5]Bioinformatics Platform, Monash University, Melbourne, VIC 3800, Australia. [6]Department of Haematology, Prince of Wales Hospital, Sydney, NSW 2031, Australia. [7]Department of Pathology, School of Medical Sciences, University of New South Wales, Sydney, NSW 2052, Australia. [8]Department of Clinical Haematology, The Alfred Hospital, Melbourne, VIC 3004, Australia.
*These authors contributed equally to this work

‡Author for correspondence (david.curtis@monash.edu)

D.J.C., 0000-0001-9497-0996

Received 4 January 2018; Accepted 28 August 2018

Development (2018) 145, dev162990. doi:10.1242/dev.162990

Lyl1 can compensate for *Scl* in erythropoiesis. First, one quarter of homozygous Scl DNA-binding mutant mice survive to adulthood with a mild anemia (Kassouf et al., 2008). Second, loss of one *Scl* allele leads to early perinatal death of *Lyl1* knockout mice (Chan et al., 2007). Finally, deletion of *Scl* in adult HSCs led to an acute loss of erythropoiesis, but with recovery over subsequent weeks (Hall et al., 2005).

To more clearly delineate the functional redundancy of Scl and Lyl1 in erythropoiesis, we targeted *Scl* in erythropoiesis using the erythroid-specific erythropoietin receptor Cre-recombinase (*EpoRCre*) transgenic line (Heinrich et al., 2004), which is active in yolk sac blood islands from E8.5 (Drogat et al., 2010). Using this model, we show that markedly reduced expression of *Scl* in committed erythroid progenitors from E8.5 was compatible with embryonic development and resulted in only a mild normocytic anemia in adult mice. However, when generated on a *Lyl1* knockout background, *Scl/Lyl1* double knockout embryos died after E10.5 owing to the failure of primitive erythropoiesis. Gene expression profiling and ChIP-seq suggested that this functional Scl redundancy was due to shared Lyl1 and Scl target genes, including the gene encoding the crucial factor Gata1. Thus, reduced expression of *Scl* revealed a previously unsuspected role for *Lyl1* in primitive erythropoiesis.

RESULTS

Co-expression of *Scl* and *Lyl1* in normal primitive progenitors and erythroid cells

We first interrogated single-cell RNA-seq data obtained from E7.75 embryos to determine whether *Scl* and *Lyl1* were co-expressed in primitive blood progenitors and erythroid cells (Scialdone et al., 2016). Four populations of cells could be defined by *Scl* and *Lyl1* RNA expression in both blood progenitors and primitive erythroid cells, with the majority of cells (166 of 271 cells, 61%) co-expressing *Scl* and *Lyl1* (Fig. 1A). In these cells, there was a weak correlation between *Scl* and *Lyl1* expression (Pearson correlation coefficient r=0.17). Smaller numbers of cells expressing only *Scl* (‘*Scl*-only’) (14%), cells expressing only *Lyl1* (‘*Lyl1*-only’) (16%) and cells apparently expressing neither (9%) were present. Overall, read counts for cells that lacked *Scl* or *Lyl1* transcripts were lower than for cells that expressed *Scl* and/or *Lyl1* (Fig. S1A), suggesting that the apparent lack of expression might be explained by expression below the sensitivity of sequencing. The distribution of these sub-populations was different between blood progenitors and primitive erythroid cells, with a significant increase in the proportion of *Scl*-only and *Lyl1*-only cells in the primitive erythroid cell population (χ^2=47.2, P<0.00001; Fig. 1B). *Gata1* and primitive globin (*Hbb-bh1*) expression were not significantly different between the four *Scl;Lyl1*-expressing subpopulations (Fig. 1C), although gene set enrichment analysis (GSEA) of the erythroid cells that lacked both *Scl* and *Lyl1* revealed a reduction in the expression of Gata1 targets and genes that are involved in heme metabolism (Fig. S1B). Finally, erythroid cells without demonstrable *Scl* and *Lyl1* transcripts had a reduced expression of cell cycle-related E2F target genes (Fig. 1D and Fig. S1C). Overall, single-cell gene expression analysis demonstrated the co-expression of *Scl* and *Lyl1* expression in the majority of blood progenitors and primitive erythroid cells. *Scl;Lyl1* non-expressing cells were also identified, but these may be explained by technological factors such as low read counts for TFs that are expressed at low levels.

Markedly reduced expression of *Scl* during development is sufficient for survival to adulthood

To determine the role of *Scl* in developmental erythropoiesis, we crossed *Scl* conditional knockout mice (*SclcKO*) (Hall et al., 2003) with *EpoRCre* mice, which express Cre after the initiation of erythropoiesis from E8.5 (Drogat et al., 2010). Mating of *EpoRCre* mice with *EYFP* reporter mice confirmed the presence of Cre activity from the colony forming unit-erythroid stage (CFU-e) onwards (Fig. S2). Deletion of *Scl* from E8.5 did not lead to embryonic lethality: *EpoRCreSclcKO* mice (*ESclcKO*) were born at the expected Mendelian ratio with a mild normocytic anemia (Table 1). Analysis of whole yolk sacs from *ESclcKO* E9.5 embryos demonstrated a reduced but detectable (19%) expression of *Scl* (Fig. 2A). Expression of *Scl* in the absence of *Lyl1* was not increased. Fluorescence-activated cell sorting (FACS) analysis of adult *ESclcKO* bone marrow demonstrated an increase in immature Ter119^{lo}CD71^{+} erythroblasts (Fig. 2B) similar to that seen in *MxCreSclcKO* mice (Hall et al., 2005). As expected, absence of Lyl1 had no effect on steady state bone marrow erythropoiesis (Fig. 2B). Similar to the *ESclcKO* E9.5 embryos, quantitative

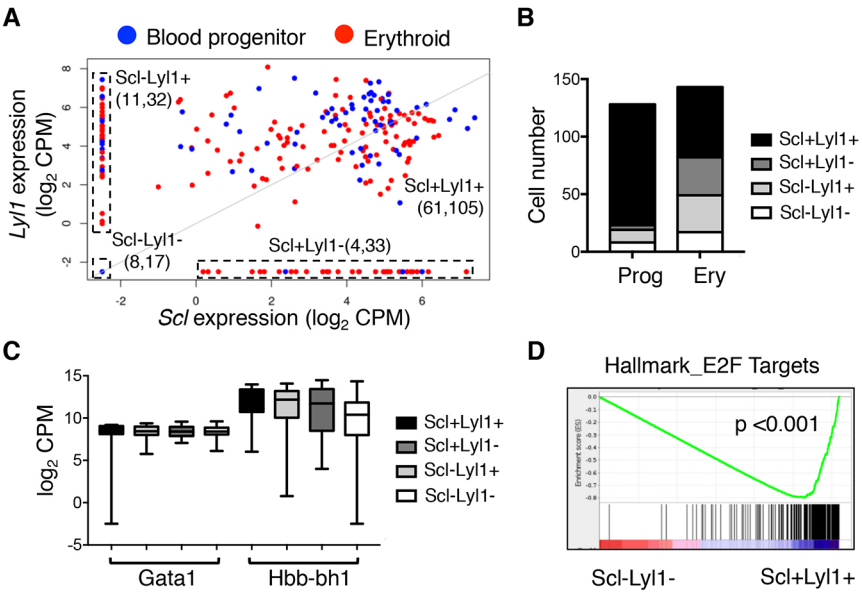

Fig. 1. *Scl* and *Lyl* expression in wild-type embryonic-derived progenitors and RBCs. (A) *Scl* and *Lyl1* expression [log$_2$ counts per million (CPM)] in single-cell blood progenitors and primitive erythroid cells from E7.75 wild-type embryos (GSE74994). The numbers in parentheses designate the numbers of progenitor and erythroid cells in the four subpopulations, respectively: *Scl*$^{+}$*Lyl1*$^{+}$; *Scl*$^{+}$*Lyl1*$^{-}$; *Scl*$^{-}$*Lyl1*$^{+}$ and *Scl*$^{-}$*Lyl1*$^{-}$. Note the *Scl*$^{-}$*Lyl1*$^{-}$ population is designated by a single dot as they are overlaid. (B) Proportion of the four subpopulations in blood progenitors (Prog) and primitive erythroid cells (Ery). (C) Expression of embryonic globin (*Hbb-bh1*) and *Gata1* in the four subpopulations shown as a box plot with median. Whiskers show maximum and minimum values. (D) GSEA analysis of differentially expressed genes between cells expressing *Scl* and *Lyl1* and cells not expressing *Scl* and *Lyl1* that demonstrates enrichment in E2F gene targets.

Development (2018) 145, dev162990. doi:10.1242/dev.162990

Table 1. RBC parameters of adult mice

Parameter	Wild-type	*EpoRCreSclcKO*	Lyl1KO	*P**
RBCs×10^{12}/l	10.37±0.51	8.44±0.28*	9.79±0.44	0.01
Hemoglobin, g/dl	15.13±0.65	12.03±0.53*	14.83±0.64	0.004
Hematocrit (%)	48.20±2.40	39.03±2.06*	48.02±2.20	0.02
MCV, fl	46.31±0.48	45.68±0.98	49.3±0.56*	0.01
MCH, pg	14.54±0.16	14.03±0.25	15.41±0.17*	0.02
MCHC, g/dl	31.30±0.41	30.68±0.37	31.16±0.34	ns
RDW (%)	19.08±0.45	20.61±0.33*	18.95±0.22	0.01

Mean±s.e.m. calculated from 10 mice/genotype. MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; MCV, mean cell volume; RDW, RBC distribution width. **P*-value calculated versus wild type using one-way ANOVA with Tukey's multiple comparison; $P<0.05$ considered significant.

real-time PCR (qPCR) of sorted bone marrow Ter119^{hi}CD71^{+} erythroblasts demonstrated a 20% residual expression of *Scl* mRNA in *ESclcKO* but no compensatory increase in *Lyl1* expression (Fig. 2C). The reduced expression of *Scl* in *ESclcKO* erythroid cells had no effect on the growth of CFU-e *in vitro* (Fig. 2D). Taken together, these results show that the significantly reduced expression of *Scl* after E8.5 was compatible with embryonic survival to adulthood.

Lyl1 can substitute for *Scl* in primitive erythropoiesis

As previously described (Capron et al., 2011), mice that lack *Lyl1* (*Lyl1KO*) had a mild RBC phenotype with macrocytosis (Table 1). We bred *ESclcKO* mice with *Lyl1KO* mice to determine whether *Lyl1* was providing the redundancy in primitive erythropoiesis with reduced *Scl* expression. Consistent with this hypothesis, mice with the loss of both *Scl* and *Lyl1* in erythropoiesis did not survive to weaning. Timed mating identified pale embryos from E10.5 onwards (Fig. 3A), with loss of the fetal heartbeat by E11.5 and complete absence of *ESclcKO;Lyl1KO* (DKO) embryos by E12.5 (Table 2). Histological sections of E10.5 DKO yolk sacs demonstrated a decrease in the number of nucleated RBCs (Fig. 3B). Benzidine staining of DKO yolk sac cells showed the progressive loss of benzidine^{+} cells from E10.5, with the reduced intensity of staining indicative of reduced heme (Fig. 3C,D).

To assess the erythroid compartment of the DKO embryos in more detail, E10.5 yolk sacs were digested with collagenase and single-cell suspensions generated for FACS analysis (Fig. 3E). *ESclcKO* erythroblasts had a mildly reduced expression of Ter119, a known transcriptional target of Scl (Kassouf et al., 2008; Lahlil et al., 2004). In contrast, DKO embryos had no detectable expression of surface Ter119. Therefore, normal expression of either *Scl* or *Lyl1* was required to maintain primitive erythropoiesis and survival beyond E10.5.

Scl and *Lyl1* regulate *Gata1* and shared downstream targets in primitive erythropoiesis

To gain insight into the regulation of primitive erythropoiesis by *Scl* and *Lyl1*, we performed RNA-seq analysis on cells from E9.5 yolk sacs, a time before any significant loss of benzidine^{+} erythroblasts in DKO embryos (Fig. 3D). Markedly reduced, but not absent, expression of the *lox-P*-flanked exon 5 of *Scl* (10% of wild-type levels) was observed in *ESclcKO* yolk sacs (Fig. 4A and Fig. S3). DKO yolk sacs also had markedly reduced expression of *Scl*, apart from one yolk sac in which expression was only reduced by ∼50% (Fig. S3). Residual expression of *Scl* might be explained by either incomplete deletion in primitive erythropoiesis or expression in endothelial cells, given that the expression analysis was performed using whole yolk sacs that contain endothelial cells. As expected for a constitutive knockout, *Lyl1* mRNA was essentially absent in *Lyl1KO* and DKO yolk sac cells (Fig. S3). Overall, the RNA-seq analysis at E9.5 confirmed the absence of *Lyl1* expression and the markedly reduced, but not absent, expression of *Scl* exon 5 in the knockout yolk sacs.

Multi-dimensional scaling using the 500 most differentially regulated genes confirmed the clustering of samples according to genotype (Fig. S4). Using a false discovery rate of 0.01, 156 genes were differentially expressed (32 up and 124 down) in *ESclcKO*

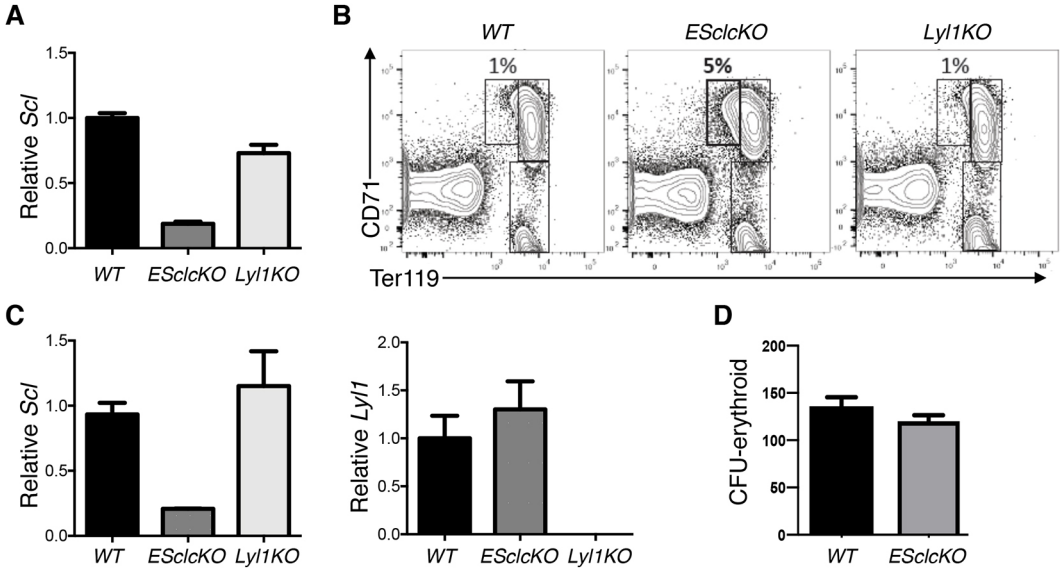

Fig. 2. Erythropoiesis in single knockout mice. (A) Expression of *Scl* in E9.5 yolk sacs from wild-type (WT), *EpoRCreScl$^{f/f}$* (*ESclcKO*) and *Lyl1KO* embryos. (B) Representative FACS plots of WT, *ESclcKO* and *Lyl1KO* adult bone marrow stained with Ter119 and CD71 showing gates that define pro-erythroblasts (CD71^{+}Ter119^{lo}), intermediate erythroblasts (CD71^{+}Ter119^{+}) and late erythroblasts (CD71^{-}Ter119^{+}). (C) Expression of *Scl* (left panel) and *Lyl1* (right panel) in FACS-isolated intermediate erythroblasts from adult mice. (D) The number of erythroid colony forming units (CFU-erythroid) per 50,000 bone marrow cells from adult WT and *ESclcKO* mice. Data are mean±s.e.m. of three mice for each genotype.

Development (2018) 145, dev162990. doi:10.1242/dev.162990

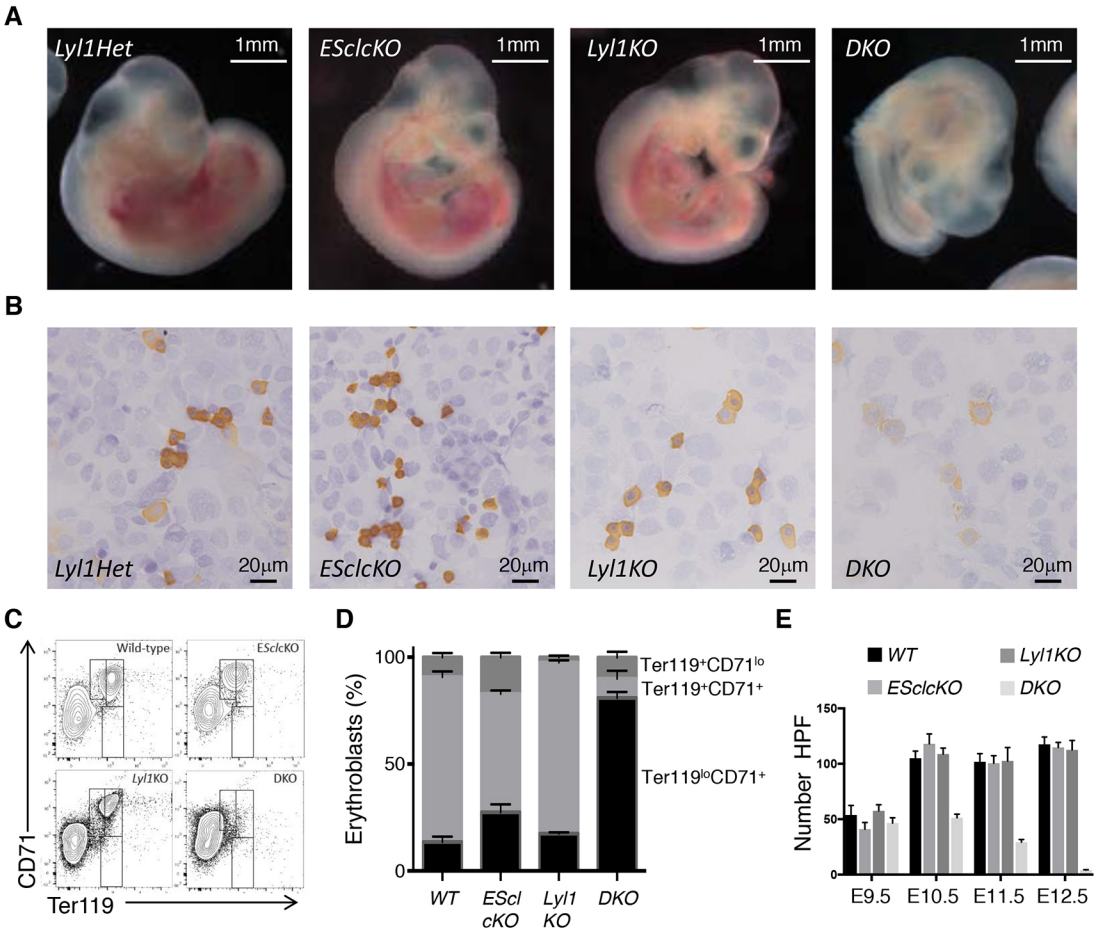

Fig. 3. Embryos lacking *Lyl1* and reduced *Scl* die owing to failure of erythropoiesis. (A) Representative E10.5 *EpoRCre*$^{tl/+}$*Scl*$^{f/f}$*Lyl1*$^{k/k}$ embryo (DKO) showing pale embryo compared with a *Lyl1* heterozygous (*Lyl1Het*) embryo and single knockout embryos (*EScIcKO* and *Lyl1KO*). (B) Benzidine staining of yolk sac cells demonstrating the decreased number of globin-expressing cells (orange) in E10.5 DKO embryos. (C) Representative FACS profiles of wild-type, *EScIcKO*, *Lyl1KO* and DKO E10.5 yolk sacs. (D) Proportion of erythroid fractions are mean±s.e.m. of four mice for each genotype. (E) Number of benzidine^{+} cells per high-power field (HPF) from wild-type (WT), *EScIcKO*, *Lyl1KO* and DKO embryos from E9.5 to E12.5. Data are mean±s.e.m. of four mice for each genotype.

yolk sacs and 310 genes (101 up and 209 down) in *Lyl1KO* yolk sacs (Fig. 4B and Table S1). GSEA showed downregulation of collagen and extracellular matrix genes in single knockouts (Fig. 4C and Fig. S4B). In yolk sacs from DKO embryos, 568 genes were differentially expressed (168 up and 400 down), with more than 70% (401) unique to the DKO embryos (Fig. 4B). Analysis of DKO cells also revealed a reduction in extracellular matrix organization genes, which was further reduced compared with *Lyl1KO* but not with *EScIcKO* cells (Fig. S4B). Unlike single knockouts, GSEA analysis of DKO cells revealed the downregulation of genes that are involved in heme metabolism (Fig. 4D,E). Of note, there was no evidence of increased apoptosis genes in DKO cells (Fig. S4C).

Table 2. Numbers of DKO embryos

Embryonic day	Total embryos	Expected	Observed
E9.5	45	12	14
E10.5	38	9	8
E11.5	25	6	8
E12.5	22	6	0
E13.5	20	5	0
E15.5	8	2	0

Numbers of *EpoRCreScIcKO;Lyl1KO* (DKO) embryos generated from *EpoRCreScIcKO;Lyl1* heterozygous mice mated with *ScIcKO;Lyl1KO* mice.

Gata1 is a known target gene of Scl, with both factors participating in a multimeric complex regulating erythropoiesis (Tripic et al., 2009; Wadman et al., 1997). *Gata1* expression was normal in *EScIcKO* cells but reduced twofold in *Lyl1KO* cells and, most markedly, reduced tenfold in DKO cells (Fig. 4F). Consistent with the loss of *Gata1*, the expression of reported targets of the SCL-GATA1 complex, including Kruppel-like factor 1 (*Klf1*), FOG family member 1 (*Zfpm1*) and transferrin receptor (*Tfrc*) were reduced in DKO cells (Fig. 4G). However, loss of both *Scl* and *Lyl1* did not affect the expression of genes encoding other members of the SCL-GATA1 complex, including *Tcf3* (also known as *E47*), *Ldb1*, *Lmo2* and the co-repressor *Cbfa2t3* (also known as *Eto-2*) (Fig. S5A). Of note, expression of both *Gata2* and *Runx1* was increased in DKO cells (Fig. S5B).

Previous studies have implicated a role for *Gata1* in survival of erythroblasts (Weiss and Orkin, 1995), which might explain the loss of benzidine^{+} erythroblasts between E9.5 and E10.5 (Fig. 2D). However, p53 target genes were not increased in DKO cells (Fig. S6A). In contrast, and similar to wild-type E7.75 primitive erythroid cells with no detectable *Scl* and *Lyl1* transcripts (Fig. 1D), there was a reduction in E2F target genes in DKO cells (Fig. S6B and S6C). In summary, gene expression analyses before the loss of benzidine-staining cells in E9.5 yolk sac demonstrated the reduced expression of Gata1 and its targets in DKO cells.

Development (2018) 145, dev162990. doi:10.1242/dev.162990

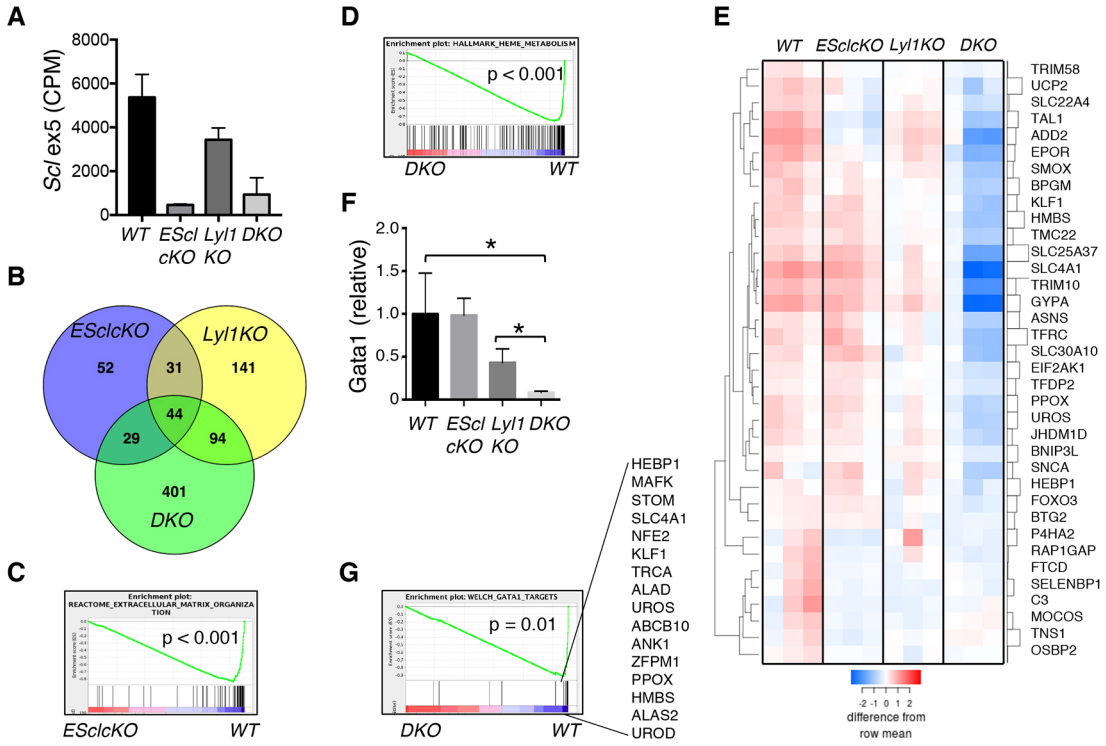

Fig. 4. RNA-seq analysis of E9.5 yolk sacs. (A) Expression of exon 5 of *Scl* in wild-type (WT), *EScIcKO*, *Lyl1KO* and DKO E9.5 yolk sacs. Data are mean± s.e.m. of three mice for each genotype. (B) Venn diagram showing numbers of differentially expressed genes in *EScIcKO*, *Lyl1KO* and DKO E9.5 yolk sacs compared with wild-type E9.5 yolk sacs using a false discovery rate of 0.01. (C) GSEA of genes involved in extracellular matrix organization in *EScIcKO* cells. (D) GSEA of genes involved in heme metabolism in DKO cells. (E) Heat map showing expression of leading edge heme-related genes from panel C. (F) qPCR of *Gata1* in WT, *EScIcKO*, *Lyl1KO* and DKO cells. Expression normalized for *Hprt* and expressed relative to WT. Data are mean±s.e.m. of three pools of yolk sacs for each genotype. **P*<0.05 (Student's *t*-test). (G) GSEA of Gata1 target genes in DKO cells with leading edge genes listed.

LYL1 binds a small subset of SCL targets

We used ChIP to identify the common binding targets of Scl and Lyl1 in erythropoiesis. None of the commercially available anti-Lyl1 antibodies were able to pull down mouse Lyl1 and therefore we used the human erythroleukemia cell line K562 for ChIP experiments to compare SCL and LYL1 DNA binding. Overall, 23,547 SCL- and 1409 LYL1-binding peaks were identified in the K562 genome (Fig. 5A and Table S2). Most strikingly, 1405 (99.7%) of the LYL1-binding sites were also bound by SCL, indicating that LYL1 almost exclusively binds a small subset of SCL targets. Importantly, overlapping binding was seen at the *GATA1* gene locus (Fig. 5B) as well as at loci of other erythroid-specific TFs (KLF1 and ZFPM1), heme synthetic enzymes and RBC membrane proteins (Fig. S7A-C), and at the locus control

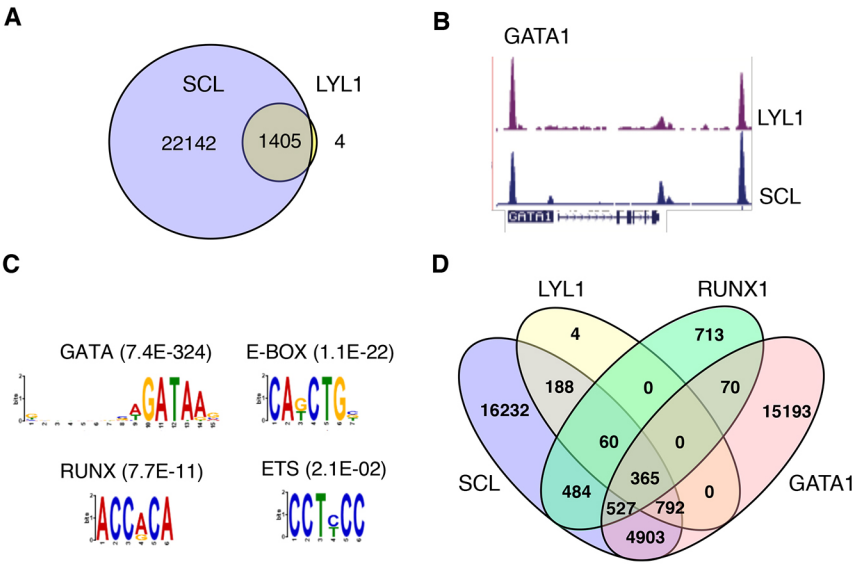

Fig. 5. LYL1 binds a small subset of SCL targets including *GATA1*. (A) Venn diagram showing numbers of common and unique binding peaks for *SCL* and *LYL1* in the human erythroid leukemia cell line, K562. (B) Read peak plot for ChIP-seq reads in K562 cells displayed in the UCSC genome browser (genome.ucsc.edu) shows overlapping binding of SCL and LYL1 near the *GATA1* gene. (C) *De novo* motif discovery of LYL1 and SCL-binding sites in K562 cells has identified significant enrichment for the GATA, E-BOX, RUNX and ETS motifs. (D) Venn diagram showing numbers of shared and unique binding peaks for SCL, LYL1, GATA1 and RUNX1 in K562 cells.

Development (2018) 145, dev162990. doi:10.1242/dev.162990

region of the β-globin locus. The shared SCL- and LYL1-binding sites were enriched for GATA1, E-BOX, RUNX and ETS motifs (Fig. 5C). Furthermore, the majority of SCL and LYL1 shared binding sites (1157/1405, 82%) were also occupied by GATA1 with or without RUNX1 (Fig. 5D and Table S2). In summary, genome-wide ChIP-seq analysis in K562 human erythroleukemic cells shows that LYL1 binds a small subset (6%) of SCL targets that are also bound by GATA1.

Previous studies using mice with a mutant form of Scl that does not bind DNA (Scl DNA-binding mutant mice) have identified 594 genome regions (~20% of all Scl-binding sites) that could bind in the absence of direct Scl binding (Kassouf et al., 2010). To test the possibility that LYL1 might be able to replace SCL at these target regions, LYL1 and SCL common binding sites in K562 cells (1405 regions) were compared with Scl-binding regions in the Scl DNA-binding mutant mice (GSE18720). These 594-mouse genome regions were mapped to 494 human-genome regions. Intersection with the 1405 SCL and LYL1 shared binding sites identified in K562 cells revealed a significant overlap of 83 common sites (Table S3, *z*-score 25.96). Importantly, this overlap was more significant than the intersection of the remaining 2400 (80%) SCL-dependent binding sites with the same 1405 SCL and LYL1 shared binding sites (Table S3, *z*-score 12.61). Pathway analysis of the 75 nearest genes showed enrichment for embryonic hematopoiesis (Table S3) that included *Gata1*, *Runx1*, *Lmo2* and *Gfi1*. Therefore, the previously identified SCL-binding sites that are independent of direct SCL binding are enriched for LYL1-binding sites.

Scl and Lyl1 directly regulate Gata1 and its targets

To identify the functionally relevant shared Scl- and Lyl1-binding targets, we integrated ChIP-seq and RNA-seq data. We mapped the 1405 shared targets from K562 to 1403 mouse homologs (Table S2). Taken together, we found that 127 genes were differentially expressed in either single knockout or DKO mice. Of these, the majority (80 genes, 63%) were differentially expressed in the DKO mice only (Fig. 6A, Fig. S8 and Table S2). The majority of these shared SCL and LYL1 targets were repressed (59 genes, 74%) including *Gata1*, *Klf1* and *Zpfm1*, whereas the remaining genes (21 genes, 26%) were upregulated and included *Runx1*. We confirmed these findings using Scl ChIP-seq data obtained from purified hematopoietic progenitors derived from mouse embryonic stem cells (Goode et al., 2016). Overall, Scl-binding sites could be mapped to 3061 genes (Table S2). Intersection of these genes with our mouse expression data from single knockout or DKO yolk sacs identified 114 genes that were differentially expressed in the DKO mice only (Fig. 6B). Of these, 33 genes were common to the 80 genes that were identified using the K562 ChIP-seq data (Table S2). Importantly, this subset included *Gata1*, *Runx1*, *Zfpm1* and erythroid differentiation genes *Gypc*, *Hmbs* and *Slc4a1*.

Finally, *de novo* motif discovery of the shared SCL- and LYL1-bound sequences identified significant enrichment for GATA1 motifs (Fig. 6C), supporting our hypothesis that either SCL or LYL1 functions in concert with GATA1 to regulate erythroid growth and differentiation. This was further confirmed by targeted motif scanning (Fig. 6D), which verified enrichment for GATA1 (42% of bound sequences), and also identified the relative overrepresentation of LYL1/SCL (32% of bound sequences) and RUNX1 (26% of bound sequences) motifs. Of note, the majority of these sites were also bound by GATA1 with or without RUNX1 (Fig. 6E). Similarly, the 114 differentially expressed Scl targets identified in mouse hematopoietic progenitors (Fig. 6B) were also frequently bound by Lmo2 (97/114), Gata1 (58/114) and Runx1

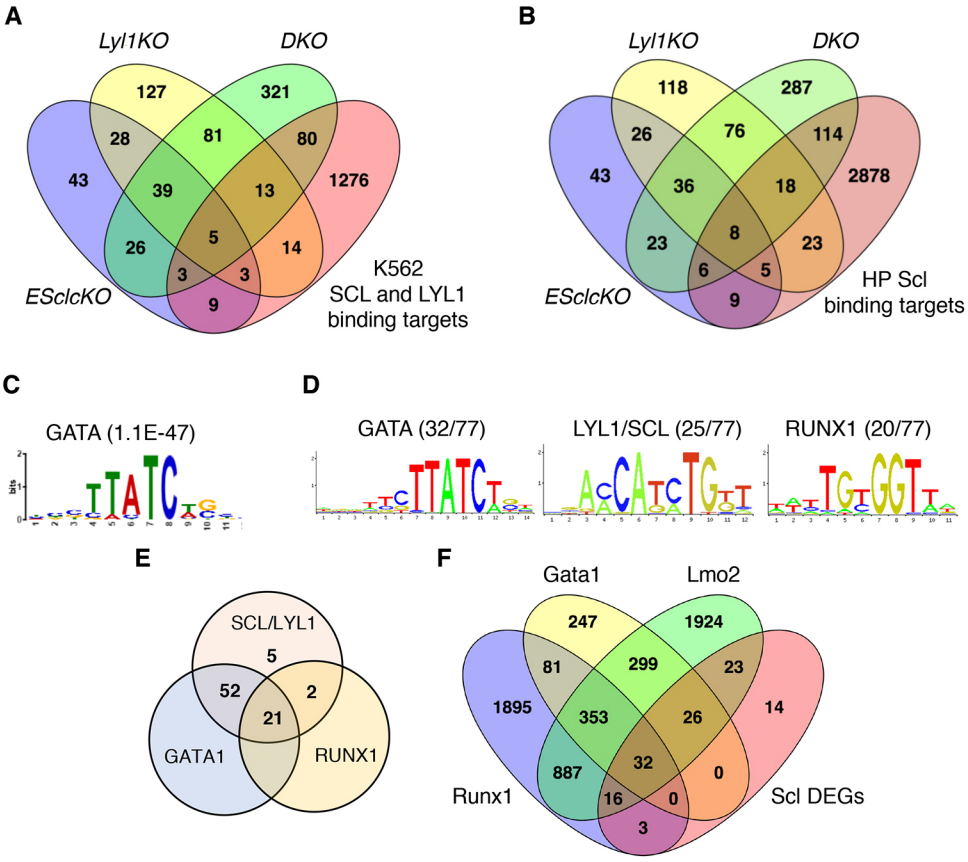

Fig. 6. Integration of ChIP-seq and RNA-seq data. (A) Venn diagram showing overlap of differentially expressed genes for *EScIcKO*, *Lyl1KO* and DKO cells with mouse homologs of the common SCL- and LYL1-binding sites identified in K562 cells. (B) Venn diagram showing overlap of differentially expressed genes with the genes nearest to Scl-binding sites in mouse hematopoietic progenitors. (C) *De novo* motif discovery of the 80 K562 ChIP-seq binding sites has identified significant enrichment for the GATA1 motif. (D) Targeted motif scanning has identified GATA, LYL1/SCL and RUNX1 motifs. (E) Intersection of LYL1, SCL, GATA1 and RUNX1 K562 ChIP-seq binding sites. (F) Venn diagram showing overlap of the 114 Scl-binding sites that have been identified in mouse hematopoietic progenitors with Gata1-, Lmo2- and Runx1-binding sites.

Development (2018) 145, dev162990. doi:10.1242/dev.162990

(51/114) (Fig. 6F). Overall, these results suggest that shared Scl- and Lyl1-binding targets are also targets of Gata1, Lmo2 and Runx1.

DISCUSSION

Functional genetic redundancy is an evolutionary process that limits the impact of mutations on key developmental programs and enables a finer degree of control (Kafri et al., 2009). Robust primitive erythropoiesis is essential for the viability of the rapidly developing embryo and, in this context, we show that either Scl or Lyl1, two closely related bHLH TFs, are capable of maintaining primitive erythropoiesis. However, *Lyl1KO* embryos with a markedly reduced expression of *Scl* die at E10.5 because of the loss of primitive erythropoiesis. An integrated approach using expression analyses and ChIP shows that this functional redundancy is most likely explained by sharing common partner proteins and target genes including the master erythroid regulator *Gata1*.

We have previously shown that deletion of *Scl* in adult HSCs leads to a transient loss of erythroid progenitors that recovers within a few weeks (Hall et al., 2005). We postulated that Lyl1 was the TF that maintained erythropoiesis in the absence of *Scl*. To examine the functional role of *Scl* in primitive erythropoiesis, we used an erythroid-specific *Cre* transgenic mouse strain that can delete *lox-P*-flanked alleles in primitive erythroid cells from E8.5 (Drogat et al., 2010). The RNA-seq analyses performed at E9.5 demonstrated 10% residual *Scl* expression that was sufficient to allow embryonic development to adulthood. However, in the absence of Lyl1, this markedly reduced expression of *Scl* was unable to maintain primitive erythropoiesis, leading to the death of all embryos by E10-E11, a time in development that relies upon primitive erythropoiesis (Palis, 2014). There is, therefore, functional redundancy of Scl and Lyl1 for murine primitive erythropoiesis. Given that the *Cre* transgene is not active until E8.5, our experiments do not address the functional redundancy of Scl and Lyl1 for initiation of primitive erythropoiesis.

In light of our results, previous conclusions that *Scl* is essential for primitive erythropoiesis should be reconsidered. Deletion of *Scl* with the *Tie2Cre* transgene led to embryonic death with systemic edema and hemorrhage at around E13.5 (Schlaeger et al., 2005). Primitive erythroid cells showed a delay in maturation, leading to the conclusion that defects in erythropoiesis contributed to embryonic death. However, our results suggest that the loss of *Scl* in other lineages in which *Tie2Cre* is active, such as megakaryocytes and endothelium, might explain the embryonic death of *Tie2Cre;* $Scl^{\Delta/\Delta}$ mice. A second study examined the phenotype of *Scl* knock-in mice (Scl^{RER}), which carry a germline mutation that prevents DNA binding (Kassouf et al., 2008). These mice had defects of both primitive and definitive erythropoiesis, with the majority (75%) of embryos dying between E14.5 and E17.5. This phenotype led to the conclusion that Scl DNA-binding was crucial for erythropoiesis. However, an alternate explanation for the embryonic lethality of the Scl^{RER} mice is a dominant-negative effect of the Scl DNA-binding mutant, which retains the ability to form a multimeric complex and therefore compete with Lyl1. A competitive model between the Scl DNA-binding mutant and Lyl1 could explain the later death of Scl^{RER} embryos and the partial penetrance in which 25% survived to adulthood with a mild anemia similar to *EScIcKO* mice. The ability of Lyl1 to bind a small subset (10%) of Scl target genes may also explain the ChIP-seq analyses of Scl^{RER} fetal livers, in which 20% of Scl-binding targets were retained in the absence of Scl DNA binding (Kassouf et al., 2010). In support of this hypothesis, we found significant overlap of common Scl/Lyl1-binding sites with the sites that can be bound by the Scl DNA-binding mutant that were identified in Scl^{RER} fetal livers (Table S3). Taken together, our results suggest that *Scl* is not essential for the maintenance of primitive erythropoiesis, as was previously thought.

The early death of DKO mice indicates that Lyl1 is the TF that is responsible for survival of the *EScIcKO* mice. Analysis of single-cell data at E7.75 revealed that 170/187 (91%) of erythroid cells express Scl and/or Lyl1 (Fig. 1A). This provides pools of cells that could explain the ongoing erythropoiesis in the absence of either factor at E10.5 into adulthood. The embryonic lethality of *Lyl1KO* mice that show markedly reduced *Scl* expression suggests that the minor cell population that lacks *Scl* and *Lyl1* at E7.75 is either transient, insufficient to maintain survival at E10.5 or is an experimental artefact of single-cell technologies. The likely mechanism for *in utero* death of DKO embryos is transcriptional loss of *Gata1*, as *Gata1* null embryos die at a similar developmental stage (Fujiwara et al., 1996). Reduced expression of *Gata1* in DKO embryos cannot be explained by reduced numbers of erythroid cells because gene expression profiling was performed before significant loss of primitive erythropoiesis. Studies of embryonic stem cells shows an important role for Gata1 in erythroid survival and maturation (Weiss and Orkin, 1995). Our analyses of E9.5 DKO cells also suggests a role for Scl/Lyl1 in cell proliferation.

We identified 80 other direct target genes shared by Scl and Lyl1 that may also contribute to the embryonic lethality. However, the loss of many of these erythroid genes, including *Klf1* and *Zfpm1*, is likely due to loss of the Scl-Gata1 and Lyl1-Gata1 complexes, given that the majority of common Scl/Lyl1-binding sites also bind *Gata1* (Fig. 6C,D). The common binding sites for SCL, LYL1 and GATA1 is consistent with the presence of a multimeric transcriptional complex, although detailed analyses of the LYL1 complex in erythroid cells has not yet been reported. Given that Gata1 functions within a multiprotein complex that contains Scl, overexpression of Gata1 alone would not be sufficient to rescue the defects in DKO cells.

In summary, we have demonstrated that *Lyl1* can compensate for the loss of *Scl* expression during primitive erythropoiesis, with functional redundancy due to shared DNA-binding targets including the crucial erythroid TF gene *Gata1*.

MATERIALS AND METHODS

Mice

Mice with homozygous *lox-P*-targeted *Scl* allele (*SclcKO*) (Hall et al., 2003) were crossed with erythropoietin receptor Cre-recombinase (*EpoRCre*) mice (Drogat et al., 2010) to generate *EpoRCreSclcKO* mice with lineage-specific deletion of *Scl*. These mice were then crossed with homozygous *Lyl1* knockout mice (*Lyl1KO*) (Capron et al., 2006) to generate *EpoRCreSclcKOLyl1KO* mice, with both genes deleted in RBCs. The Animal Ethics Committee of the Alfred Medical Research and Education Precinct approved all animal experiments.

Blood counts

Mice were bled by puncturing the submandibular vein. Blood was then placed into Sarstedt Microvette collection tubes containing EDTA and full blood counts were performed on a Drew Scientific Hemavet system.

Embryo dissection and yolk sac collection

Embryos were collected at E9.5 to E15.5 and placed in PBS at room temperature and dissected with fine forceps under a dissecting microscope (Nikon SMZ1500). Photographs of embryos were taken with a Zeiss AxioCam MRC5 camera. Yolk sac cells were obtained by digestion at 37°C for 30 min in PBS with 7% fetal calf serum (FCS) and 1% penicillin/streptomycin collagenase/dispase (Sigma-Aldrich). Ice-cold PBS with 7% FCS and 1 mM EDTA solution was added to stop the digestion and the cells were then spun at 1500 rpm (300 ***g***) for 5 min at 4°C. The yolk sac pellet was

Development (2018) 145, dev162990. doi:10.1242/dev.162990

re-suspended in ice-cold PBS with 7% FCS and 1 mM EDTA, homogenized with vigorous pipetting and then filtered through a 40 µm nylon strainer. Samples were then used for cytospin, benzidine staining, RNA extraction and flow cytometry.

Flow cytometry

Flow cytometry and cell sorting was performed as previously described (Tremblay et al., 2010). Briefly 2×10^6 bone marrow cells were stained using BD Pharmingen antibodies against mouse CD71 (C2 ; 553266) and Ter119 (Ter-119; 561033) and BioLegend antibodies against mouse CD105 (Mj7/18; 120412) and CD150 (TC15-12F12.2; 115904). All antibodies were used at 1:500 except CD150 which was used at 1:100. Dead cells were excluded using propidium iodide staining. Cell sorting was performed using FACSAria (BD Biosciences) and FACS analysis was performed using CantoII, LSRII and LSR-Fortessa (BD Biosciences). Results were analyzed using DIVA and FlowJo software.

RNA extraction and qPCR

RNA was extracted using Trizol and the RNAeasy mini kit (Qiagen). RNA (1 µg) was used to make cDNA with the Transcriptor First Strand cDNA kit (Roche). qPCR was performed on a Lightcycler 480 II machine (Roche). Results were standardized to expression of the *Hprt* housekeeping gene.

RNA-seq

Total RNA from E9.5 yolk sacs was extracted for mRNA sequencing using Illumina HiSeq2500. Library preparation and sequencing was performed by Micromon (Monash University). Data was analyzed using an in-house developed pipeline, RNASik (v1.4.7) (monashbioinformaticsplatform.github.io/RNAsik-pipe). Raw reads were aligned using STAR (v2.5.2b) (Dobin et al., 2013) with GRCm38 reference genome from ENSEMBL. Gene counts were generated using feature counts in subread (v1.5.2) (Liao et al., 2014) with GRCm38.87 gene annotation from ENSEMBL. Differential expression analysis was performed in Degust (degust.erc.monash.edu) using LIMMA/edgeR/Voom. Public single-cell RNA-seq data: raw read count single-cell data were extracted from a recent study (Scialdone et al., 2016; gastrulation.stemcells.cam.ac.uk/scialdone2016) and analyzed using LIMMA(v3.32.2)/edgeR (v3.18.1)/Voom libraries in R Bioconductor v3.5 (bioconductor.org). Raw read counts were Voom-transformed into counts per million.

ChIP-seq analysis

The raw ChIP-seq reads were filtered for adapter contamination and low quality scores, and we also excluded reads in which more than 10% of bases were unknown. The following K562 ChIP-seq datasets were publicly available from GEO (SCL, GSE31477; RUNX1, GSE24777; GATA1, GSE70482). The ChIP-seq data for mouse hematopoietic progenitors was obtained from GSE69101 (Goode et al., 2016). Data processing was performed as previously described (Beck et al., 2013; Chacon et al., 2014; Wilson et al., 2010). In brief, reads were aligned to the human genome (hg19) using the software package BWA with standard parameters, resulting in 13,394,203 mapped reads for LYL1, 29,426,999 mapped reads for SCL, 8,033,900 mapped reads for RUNX1, 22,622,534 mapped reads for GATA1. Three publicly available peak finding programs, MACS (Feng et al., 2012), findPeaks (HOMER) (Heinz et al., 2010) and SPP (Kharchenko et al., 2008), were used to call peaks. Peaks called by at least two algorithms were identified as high-confidence binding sites for downstream analysis. *De novo* motif discovery was performed using MEME (Bailey and Elkan, 1994). High-confidence binding regions were assigned as regulatory regions to at most two genes using annotations provided by the GREAT analysis package (McLean et al., 2010). Homology information for human and mouse was downloaded from MGI (Blake et al., 2017) and we found that 16,824 human genes were associated with 17,196 mouse homologs. 1575 LYL1 and SCL target genes in K562 cells were mapped 1403 mouse homologs. RUNX1, GATA1 and TAL1 motifs were downloaded from the JASPAR database and occurrence searched using the FIMO software from the MEME suite. A bootstrapping approach was used to address the significance of combinatorial binding events between the TFs for all possible binding patterns. A lower estimate of 80,000 binding sites available for TF binding (Tijssen et al., 2011; Wilson et al., 2010) was used to estimate the background distribution of combinatorial binding events. The standardized *z*-score metric was then used to express the deviation of the combinatorial binding events from the expected mean (normalized by the standard deviation) of the background distribution (i.e. a positive *z*-score indicates an overlap more likely than random chance; a higher *z*-score indicates a more significant overlap).

Acknowledgements

The authors acknowledge Jacqueline E. Boyle for genotyping mice; staff at Monash ARL for animal husbandry; Jelena Kezic of Monash Histology Platform for processing and Haemotoxylin and Eosin staining of embryos and yolk sacs; and Geza Paukovics, Phil Donaldson and Eva Orlowski from AMREP flow cytometry facility for their assistance in flow cytometry. The authors would also like to thank Bertie Gottgens, University of Cambridge, for reading the manuscript and providing insightful feedback.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: D.B., J.E.P., C.S.T., D.J.C.; Methodology: C.S.T., D.J.C.; Formal analysis: S.K.C., J.S., Y.H., S.E.S., N.C.W., D.R.P., D.B., J.E.P., C.S.T., D.J.C.; Data curation: S.K.C., J.S., Y.H., S.E.S.; Writing - original draft: S.K.C., D.B., J.E.P.; Writing - review & editing: S.K.C., C.S.T., D.J.C.; Supervision: C.S.T., D.J.C.

Funding

This work was supported by a project grant APP1052313 (D.J.C. and J.E.P.) from the Australian National Health and Medical Research Council (NHMRC), a Leukaemia Foundation scholarship (S.K.C.), a Senior Medical Research Fellowship from the Sylvia and Charles Viertel Charitable Foundation (D.J.C.) and a Cancer Institute NSW Fellowship (D.B.). We also acknowledge funding from the Anthony Rothe Memorial Trust (J.E.P.), Cancer Australia (D.B), Gilead Sciences (D.B.) and a National Health and Medical Research Council Peter Doherty fellowship (D.B.).

Data availability

Data have been deposited in GEO under accession number GSE103789.

Supplementary information

Supplementary information available online at
http://dev.biologists.org/lookup/doi/10.1242/dev.162990.supplemental

References

Bailey, T. L. and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2**, 28-36.

Beck, D., Thoms, J. A. I., Perera, D., Schutte, J., Unnikrishnan, A., Knezevic, K., Kinston, S. J., Wilson, N. K., O'Brien, T. A., Gottgens, B. et al. (2013). Genome-wide analysis of transcriptional regulators in human HSPCs reveals a densely interconnected network of coding and noncoding genes. *Blood* **122**, e12-e22.

Blake, J. A., Eppig, J. T., Kadin, J. A., Richardson, J. E., Smith, C. L. and Bult, C. J. and The Mouse Genome Database Group (2017). Mouse genome database (MGD)-2017: community knowledge resource for the laboratory mouse. *Nucleic Acids Res.* **45**, D723-D729.

Capellera-Garcia, S., Pulecio, J., Dhulipala, K., Siva, K., Rayon-Estrada, V., Singbrant, S., Sommarin, M. N. E., Walkley, C. R., Soneji, S., Karlsson, G. et al. (2016). Defining the minimal factors required for erythropoiesis through direct lineage conversion. *Cell Rep.* **15**, 2550-2562.

Capron, C., Lecluse, Y., Kaushik, A. L., Foudi, A., Lacout, C., Sekkai, D., Godin, I., Albagli, O., Poullion, I., Svinartchouk, F. et al. (2006). The SCL relative LYL-1 is required for fetal and adult hematopoietic stem cell function and B-cell differentiation. *Blood* **107**, 4678-4686.

Capron, C., Lacout, C., Lécluse, Y., Wagner-Ballon, O., Kaushik, A.-L., Cramer-Bordé, E., Sablitzky, F., Dumenil, D. and Vainchenker, W. (2011). LYL-1 deficiency induces a stress erythropoiesis. *Exp. Hematol.* **39**, 629-642.

Chacon, D., Beck, D., Perera, D., Wong, J. W. H. and Pimanda, J. E. (2014). BloodChIP: a database of comparative genome-wide transcription factor binding profiles in human blood cells. *Nucleic Acids Res.* **42**, D172-D177.

Chan, W. Y. I., Follows, G. A., Lacaud, G., Pimanda, J. E., Landry, J.-R., Kinston, S., Knezevic, K., Piltz, S., Donaldson, I. J., Gambardella, L. et al. (2007). The paralogous hematopoietic regulators Lyl1 and Scl are coregulated by Ets and

GATA factors, but Lyl1 cannot rescue the early Scl–/– phenotype. *Blood* **109**, 1908-1916.
Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21.
Dore, L. C. and Crispino, J. D. (2011). Transcription factor networks in erythroid cell and megakaryocyte development. *Blood* **118**, 231-239.
Drogat, B., Kalucka, J., Gutierrez, L., Hammad, H., Goossens, S., Farhang Ghahremani, M., Bartunkova, S., Haigh, K., Deswarte, K., Nyabi, O. et al. (2010). Vegf regulates embryonic erythroid development through Gata1 modulation. *Blood* **116**, 2141-2151.
Feng, J., Liu, T., Qin, B., Zhang, Y. and Liu, X. S. (2012). Identifying ChIP-seq enrichment using MACS. *Nat. Protoc.* **7**, 1728-1740.
Fujiwara, Y., Browne, C. P., Cunniff, K., Goff, S. C. and Orkin, S. H. (1996). Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *Proc. Natl. Acad. Sci. USA* **93**, 12355-12358.
Goode, D. K., Obier, N., Vijayabaskar, M. S., Lie-A-Ling, M., Lilly, A. J., Hannah, R., Lichtinger, M., Batta, K., Florkowska, M., Patel, R. et al. (2016). Dynamic gene regulatory networks drive hematopoietic specification and differentiation. *Dev. Cell* **36**, 572-587.
Hall, M. A., Curtis, D. J., Metcalf, D., Elefanty, A. G., Sourris, K., Robb, L., Gothert, J. R., Jane, S. M. and Begley, C. G. (2003). The critical regulator of embryonic hematopoiesis, SCL, is vital in the adult for megakaryopoiesis, erythropoiesis, and lineage choice in CFU-S12. *Proc. Natl. Acad. Sci. USA* **100**, 992-997.
Hall, M. A., Slater, N. J., Begley, C. G., Salmon, J. M., Van Stekelenburg, L. J., McCormack, M. P., Jane, S. M. and Curtis, D. J. (2005). Functional but abnormal adult erythropoiesis in the absence of the stem cell leukemia gene. *Mol. Cell. Biol.* **25**, 6355-6362.
Heinrich, A. C., Pelanda, R. and Klingmuller, U. (2004). A mouse model for visualization and conditional mutations in the erythroid lineage. *Blood* **104**, 659-666.
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.
Kafri, R., Springer, M. and Pilpel, Y. (2009). Genetic redundancy: new tricks for old genes. *Cell* **136**, 389-392.
Kassouf, M. T., Chagraoui, H., Vyas, P. and Porcher, C. (2008). Differential use of SCL/TAL-1 DNA-binding domain in developmental hematopoiesis. *Blood* **112**, 1056-1067.
Kassouf, M. T., Hughes, J. R., Taylor, S., McGowan, S. J., Soneji, S., Green, A. L., Vyas, P. and Porcher, C. (2010). Genome-wide identification of TAL1's functional targets: insights into its mechanisms of action in primary erythroid cells. *Genome Res.* **20**, 1064-1083.
Kharchenko, P. V., Tolstorukov, M. Y. and Park, P. J. (2008). Design and analysis of ChIP-seq experiments for DNA-binding proteins. *Nat. Biotechnol.* **26**, 1351-1359.
Kingsley, P. D., Malik, J., Fantauzzo, K. A. and Palis, J. (2004). Yolk sac-derived primitive erythroblasts enucleate during mammalian embryogenesis. *Blood* **104**, 19-25.
Lahlil, R., Lecuyer, E., Herblot, S. and Hoang, T. (2004). SCL assembles a multifactorial complex that determines glycophorin A expression. *Mol. Cell. Biol.* **24**, 1439-1452.
Liao, Y., Smyth, G. K. and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930.
McLean, C. Y., Bristor, D., Hiller, M., Clarke, S. L., Schaar, B. T., Lowe, C. B., Wenger, A. M. and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495-501.
Palis, J. (2014). Primitive and definitive erythropoiesis in mammals. *Front Physiol* **5**, 3.
Porcher, C., Chagraoui, H. and Kristiansen, M. S. (2017). SCL/TAL1: a multifaceted regulator from blood development to disease. *Blood* **129**, 2051-2060.
Robb, L., Lyons, I., Li, R., Hartley, L., Kontgen, F., Harvey, R. P., Metcalf, D. and Begley, C. G. (1995). Absence of yolk sac hematopoiesis from mice with a targeted disruption of the scl gene. *Proc. Natl. Acad. Sci. USA* **92**, 7075-7079.
Schlaeger, T. M., Mikkola, H. K., Gekas, C., Helgadottir, H. B. and Orkin, S. H. (2005). Tie2Cre-mediated gene ablation defines the stem-cell leukemia gene (SCL/tal1)-dependent window during hematopoietic stem-cell development. *Blood* **105**, 3871-3874.
Scialdone, A., Tanaka, Y., Jawaid, W., Moignard, V., Wilson, N. K., Macaulay, I. C., Marioni, J. C. and Gottgens, B. (2016). Resolving early mesoderm diversification through single-cell expression profiling. *Nature* **535**, 289-293.
Shivdasani, R. A., Mayer, E. L. and Orkin, S. H. (1995). Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. *Nature* **373**, 432-434.
Souroullas, G. P., Salmon, J. M., Sablitzky, F., Curtis, D. J. and Goodell, M. A. (2009). Adult hematopoietic stem and progenitor cells require either Lyl1 or Scl for survival. *Cell Stem Cell* **4**, 180-186.
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.
Tremblay, M., Tremblay, C. S., Herblot, S., Aplan, P. D., Hebert, J., Perreault, C. and Hoang, T. (2010). Modeling T-cell acute lymphoblastic leukemia induced by the SCL and LMO1 oncogenes. *Genes Dev.* **24**, 1093-1105.
Tripic, T., Deng, W., Cheng, Y., Zhang, Y., Vakoc, C. R., Gregory, G. D., Hardison, R. C. and Blobel, G. A. (2009). SCL and associated proteins distinguish active from repressive GATA transcription factor complexes. *Blood* **113**, 2191-2201.
Visvader, J., Begley, C. G. and Adams, J. M. (1991). Differential expression of the LYL, SCL and E2A helix-loop-helix genes within the hemopoietic system. *Oncogene* **6**, 187-194.
Wadman, I. A., Osada, H., Grutz, G. G., Agulnick, A. D., Westphal, H., Forster, A. and Rabbitts, T. H. (1997). The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *EMBO J.* **16**, 3145-3157.
Weiss, M. J. and Orkin, S. H. (1995). Transcription factor GATA-1 permits survival and maturation of erythroid precursors by preventing apoptosis. *Proc. Natl. Acad. Sci. USA* **92**, 9623-9627.
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. (2010). Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell* **7**, 532-544.