Article

# A *Prox1* enhancer represses haematopoiesis in the lymphatic vasculature

https://doi.org/10.1038/s41586-022-05650-9

Received: 3 April 2019

Accepted: 13 December 2022



Jan Kazenwadel[1], Parvathy Venugopal[1,2], Anna Oszmiana[1], John Toubia[1,3], Luis Arriola-Martinez[1,2], Virginia Panara[4,5], Sandra G. Piltz[6,7,8], Chris Brown[9], Wanshu Ma[10], Andreas Schreiber[3,11], Katarzyna Koltowska[4,5], Samir Taoudi[12,13], Paul Q. Thomas[6,7,8], Hamish S. Scott[1,2,3,6] & Natasha L. Harvey[1,6]✉

Transcriptional enhancer elements are responsible for orchestrating the temporal and spatial control over gene expression that is crucial for programming cell identity during development[1–3]. Here we describe a novel enhancer element that is important for regulating the expression of *Prox1* in lymphatic endothelial cells. This evolutionarily conserved enhancer is bound by key lymphatic transcriptional regulators including GATA2, FOXC2, NFATC1 and PROX1. Genome editing of the enhancer to remove five nucleotides encompassing the GATA2-binding site resulted in perinatal death of homozygous mutant mice due to profound lymphatic vascular defects. Lymphatic endothelial cells in enhancer mutant mice exhibited reduced expression of genes characteristic of lymphatic endothelial cell identity and increased expression of genes characteristic of haemogenic endothelium, and acquired the capacity to generate haematopoietic cells. These data not only reveal a transcriptional enhancer element important for regulating *Prox1* expression and lymphatic endothelial cell identity but also demonstrate that the lymphatic endothelium has haemogenic capacity, ordinarily repressed by *Prox1*.

Transcriptional enhancers impart exquisite spatial and temporal control over gene expression to direct the programming of cell identity during development[1–3]. These elements vastly outnumber protein-coding genes, can be located at substantial distances from gene promoters and coordinate chromatin looping events, which bring transcriptional machinery into the proximity of target gene promoters[1–3]. The importance of enhancer elements in precisely controlling the amplitude of gene expression is underscored by the suite of variants in non-coding regions of the genome that underlie human disease[4].

*Prox1* encodes a homeobox transcription factor that has crucial roles during the development of tissues, including the lens, retina, liver, pancreas, heart and lymphatic vasculature[5–12]. *Prox1* is required to both specify lymphatic endothelial cell (LEC) fate and maintain LEC identity; *Prox1*$^{-/-}$ mouse embryos are devoid of lymphatic vessels[11] and the reduction of *Prox1* levels in specified LECs results in a reversion of LEC identity to one resembling blood vascular endothelial cells (BECs)[13]. The perinatal death of many *Prox1*$^{+/-}$ mice demonstrates the sensitivity of the lymphatic vasculature to diminished *Prox1* dosage[13]. Despite the importance of *Prox1* for the specification and maintenance of LEC identity, little is known regarding mechanisms underlying the transcriptional regulation of *Prox1*. Both *Sox18* (ref. [14]) and *Nr2f2* (encoding COUPTFII)[15] are required for the initiation of *Prox1* expression in LEC progenitors, although the dependence of *Prox1* on *Nr2f2* is temporary[15] and the identity of additional transcriptional regulators of *Prox1* remains enigmatic.

We and others have previously determined that *GATA2* mutations underlie Emberger syndrome[16,17], a primary lymphoedema syndrome characterized by lymphoedema, myelodysplasia and predisposition to acute myeloid leukaemia, owing to a key role of *Gata2* in the development and maintenance of lymphovenous and lymphatic vessel valves[18]. A key characteristic of *Gata2*-deficient valve endothelial cells is their failure to increase the levels of PROX1, demonstrating that GATA2 is an important transcriptional regulator of *Prox1* in this context[18]. Investigation of the sites bound by GATA2 in the genome of primary human dermal LECs revealed a potential enhancer element 11 kb upstream of the first non-coding exon of *PROX1* that is also bound by the key transcriptional regulators of lymphatic vascular development and valve development: FOXC2 and NFATC1 (ref. [18]).

Here we reveal that the *Prox1* −11-kb enhancer element comprises a crucial, tissue-specific transcriptional enhancer regulating the expression of *Prox1* in the lymphatic vasculature during development. Mice with homozygous deletions encompassing this enhancer element exhibited perinatal lethality due to profound defects in lymphatic vascular development. *Prox1* mRNA levels were reduced in LECs isolated from enhancer mutant mice and, consequently, markers of lymphatic identity including *Flt4* were lower in expression, whereas

[1]Centre for Cancer Biology, University of South Australia and SA Pathology, Adelaide, Australia. [2]Department of Genetics and Molecular Pathology, SA Pathology, Adelaide, Australia. [3]ACRF Cancer Genomics Facility, Centre for Cancer Biology, University of South Australia and SA Pathology, Adelaide, Australia. [4]Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden. [5]Beijer Gene and Neuro Laboratory and Science for Life Laboratories, Uppsala University, Uppsala, Sweden. [6]Adelaide Medical School, University of Adelaide, Adelaide, Australia. [7]Genome Editing Program, South Australian Health & Medical Research Institute, Adelaide, Australia. [8]South Australian Genome Editing Facility, University of Adelaide, Adelaide, Australia. [9]University of South Australia, Adelaide, Australia. [10]Center for Vascular and Developmental Biology, Feinberg Cardiovascular Research Institute, Northwestern University Feinberg School of Medicine, Chicago, USA. [11]School of Biological Sciences, University of Adelaide, Adelaide, Australia. [12]Molecular Medicine Division, Cancer and Haematology Division, The Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia. [13]Department of Medical Biology, University of Melbourne, Melbourne, Australia. ✉e-mail: natasha.harvey@unisa.edu.au

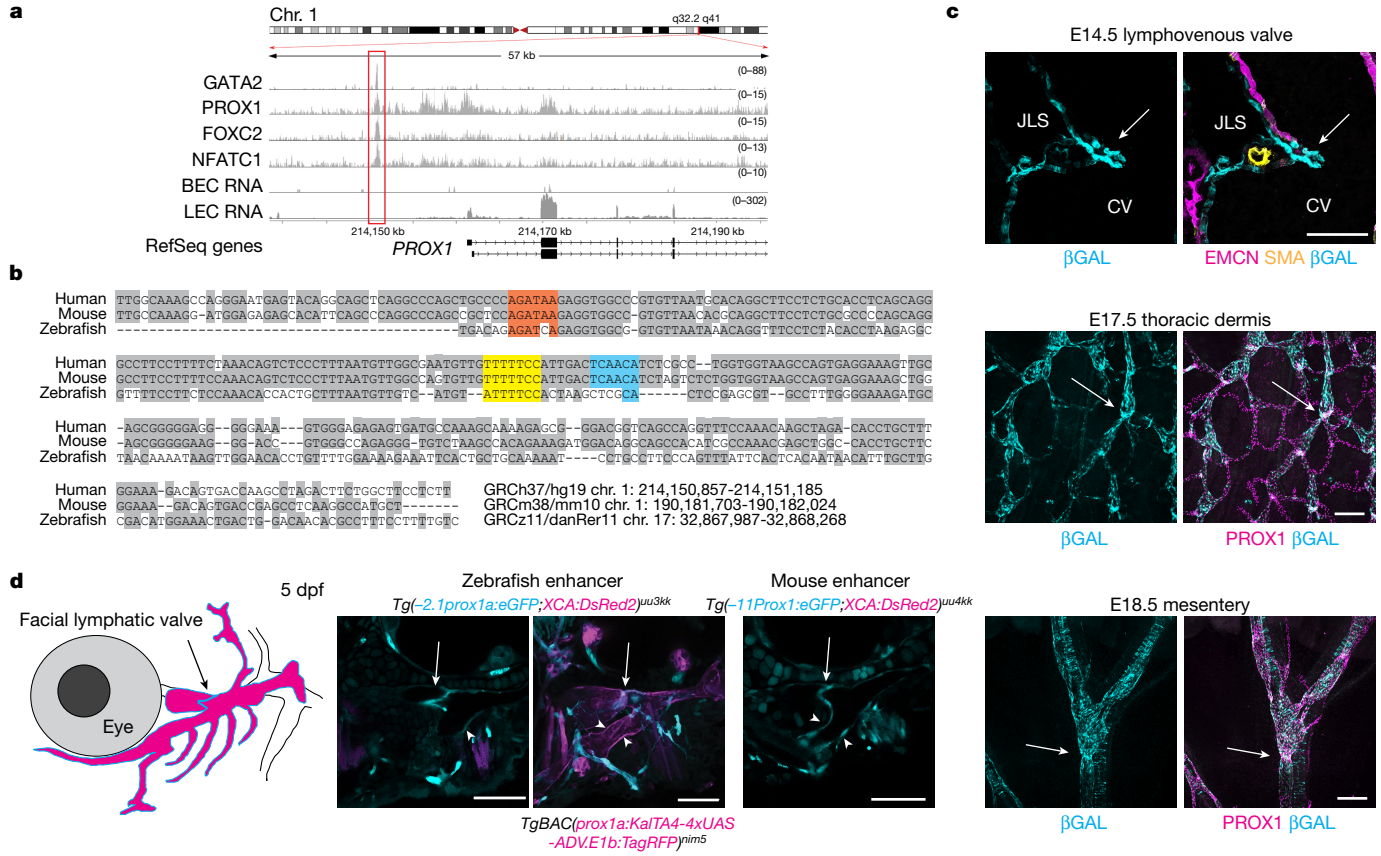



**Fig. 1 | The *Prox1* −11-kb enhancer drives reporter gene expression in LECs.** **a**, ChIP–seq of GATA2, PROX1, FOXC2 and NFATC1 at the *PROX1* −11-kb enhancer in cultured human dermal LECs. Chr. 1, chromosome 1. **b**, Alignment of genomic sequences in the region of the *PROX1* −11-kb enhancer. Conserved nucleotides (grey), consensus binding sites for GATA2 (orange), NFATC1 (yellow) and FOXC2 (blue) are shown. **c**, β-Galactosidase (βGAL) activity (cyan) driven by the mouse *Prox1* −11-kb enhancer in the lymphatic vasculature of transgenic mice. Immunofluorescent staining of coronal sections at E14.5 reveals high levels of β-galactosidase in the lymphovenous valve (arrow). Whole-mount immunostaining of E17.5 skin and E18.5 mesentery demonstrates activity throughout lymphatic vessels and at high levels in valves (arrows). CV, cardinal vein; JLS, jugular lymph sac. **d**, Lateral view of zebrafish facial lymphatic vessels at 5 days post-fertilization (dpf) showing eGFP expression (cyan) in transgenic (Tg) reporter lines driven by the zebrafish *prox1a* −2.1-kb enhancer or mouse *Prox1* −11-kb enhancer. The arrows indicate the facial lymphatic valve and the arrowheads denote the lymphatic endothelium. Scale bars, 100 µm (**c**) and 50 µm (**d**). Representative micrographs from at least three biological replicates are shown.

markers characteristic of haemogenic endothelium including *Runx1* were elevated. LECs from both wild-type and *Prox1* enhancer mutant mice exhibited the capacity to generate haematopoietic cells in vitro and in vivo, and this potential was increased in mice with mutations in the *Prox1* enhancer. Our data demonstrate that GATA2 binding to the *Prox1* −11-kb enhancer is an event crucial for directing the appropriate level of *Prox1* transcription for LEC identity to be specified and maintained, and reveal that LECs have the capacity to generate cells of the haematopoietic lineage.

## The *Prox1* −11-kb enhancer is active in LECs

To understand how *Prox1* expression is regulated during development, chromatin immunoprecipitation followed by sequencing (ChIP–seq) analysis of primary human dermal LECs was undertaken to map regions in the vicinity of the *PROX1* locus bound by key transcriptional regulators of lymphatic vascular development: GATA2 (ref. [18]), FOXC2 (ref. [19]), NFATC1 (ref. [20]) and PROX1 itself[21]. These studies revealed prominent binding of all four transcription factors approximately 11 kb upstream of the *PROX1* promoter (Fig. 1a and Extended Data Fig. 1a). The nucleotide sequence of this prospective transcriptional enhancer is highly conserved between mouse, human and zebrafish, including conservation of consensus binding sites for GATA2 and NFATC1 (Fig. 1b).

To investigate the ability of the *Prox1* −11-kb region to act as an enhancer in vivo, an 832-bp fragment was cloned into a reporter construct consisting of the *LacZ* gene driven by a minimal *hsp68* promoter[22] and was used to generate stable transgenic mice (Extended Data Fig. 1b). Analysis of *LacZ* expression patterns revealed the presence of active β-galactosidase in PROX1-positive LECs, detectable from embryonic day 11.5 (E11.5) onwards (Fig. 1c and Extended Data Fig. 1e–g). Reporter gene activity was particularly prominent in valve endothelial cells, including the lymphovenous valves (Fig. 1c and Extended Data Fig. 1e) and collecting lymphatic vessel valves (Fig. 1c and Extended Data Fig. 1f). Although reporter gene expression was initially widespread in the embryonic lymphatic vasculature of the skin and mesentery (Fig. 1c), expression became progressively restricted to large collecting lymphatic vessels and valves of the lung, skin, mesentery and thoracic duct by postnatal day 4 (P4) (Extended Data Fig. 1f), and staining was not observed in the lymphatic vasculature of adult transgenic mice. These data suggest that activity of the *Prox1* −11-kb enhancer is temporally, as well as spatially, controlled. Although some staining was observed in endothelial cell populations including those comprising the cardiac and venous valves (Extended Data Fig. 1g), BECs were generally negative for β-galactosidase, indicating cell-type-specific activity of the *Prox1* −11-kb enhancer. Of note, in tissues that normally exhibit high levels of PROX1 protein such as the liver, β-galactosidase was not detectable at any of the embryonic time points analysed (Extended Data Fig. 1e).

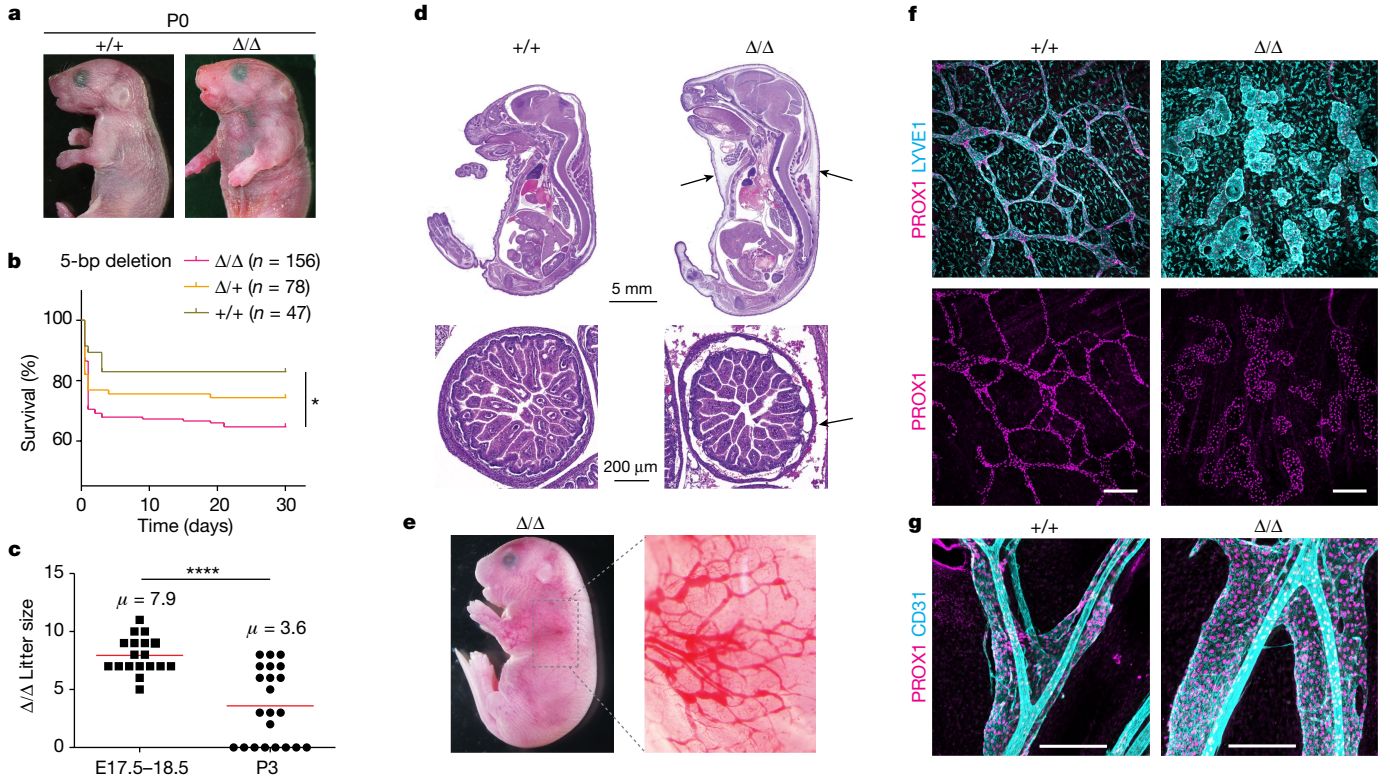



**Fig. 2 | Deletion of the *Prox1* −11-kb enhancer results in perinatal lethality and lymphatic vascular defects. a**, Newborn *Prox1* −11-kb enhancer deletion mutants (Δ/Δ) exhibit pronounced jugular swelling compared with wild-type littermates (+/+). **b**, Thirty-day Kaplan–Meier survival curves show significantly reduced survival of homozygous mutants. **c**, Prenatal and postnatal litter sizes resulting from homozygous 5-bp deletion matings. **d**, Profound interstitial oedema and distended intestinal lymphatic vessels in enhancer mutants at E18.5 (Δ/Δ, arrows). **e**, Blood-filled dermal lymphatic vessels in the axilla of E18.5 mutant embryos. **f**,**g**, Whole-mount immunostaining of skin at E17.5 (**f**) and mesentery at E18.5 (**g**). Mutant littermate exhibits distended and blood-filled dermal lymphatic vessels and no obvious valves. Log-rank (Mantel–Cox) test. $^{*}P = 0.0219$ (**b**). E17.5–E18.5, $n = 18$ litters; P3, $n = 22$ litters. Unpaired two-tailed *t*-test. $^{****}P = 0.000006$ (**c**). Scale bars, 200 μm (**f**,**g**).

To further investigate the functional conservation of this enhancer across species, the capacity of the homologous region in zebrafish, located 2.1 kb upstream of *prox1a*, to drive gene expression in the zebrafish lymphatic vasculature was assessed (Fig. 1b). Reporter gene expression was observed in the facial lymphatics of transgenic zebrafish and at high levels in the recently described facial lymphatic valve[23] (Fig. 1d and Extended Data Fig. 2), strongly suggesting that this enhancer is functionally conserved. The mouse *Prox1* −11-kb enhancer also drove reporter gene expression in the facial lymphatics and lymphatic valve of zebrafish (Fig. 1d and Extended Data Fig. 2). Together, these data provide compelling evidence that the *Prox1* −11-kb enhancer element drives gene expression in the lymphatic vasculature throughout development, and at particularly high levels in valve endothelial cells.

## Enhancer deletion causes perinatal death

To determine the requirement of the *Prox1* −11-kb enhancer for *Prox1* expression and lymphatic vascular development, CRISPR–Cas9-mediated genome editing using a guide RNA targeting the GATA2-binding site of *Prox1* −11 kb was used to generate a series of deletions spanning the enhancer (Extended Data Fig. 1c,d). Crossing of heterozygous mice revealed that a substantial proportion of homozygous pups died at or very soon after birth, with the most severely affected neonates exhibiting profound swelling of the jugular and thoracic regions (Fig. 2a,b). Homozygous mice that survived to 1 week generally continued to thrive and were fertile. The degree of perinatal lethality was consistent across all deletions analysed from 5 bp to 1,068 bp. In contrast to wild-type matings, which generated an average litter size of approximately eight pups at P3, matings set between mice homozygous for 5-bp enhancer deletions averaged approximately four pups per litter at this stage (Fig. 2c), indicating that only 50% of homozygous mutant pups born survived until P3. Histological analysis of E18.5 embryos in which 5 bp spanning the GATA-binding site of the *Prox1* −11-kb enhancer were deleted ($^{\Delta 5/\Delta 5}$) revealed jugular swelling, severely congested lymphatic vessels and gross, generalized interstitial oedema, both in the dermis and in visceral structures including the intestine (Fig. 2d,e).

## Enhancer mutants have lymphatic defects

To determine at which point lymphatic vascular defects were first obvious in mutants, embryos bearing the complete range of deletions of the *Prox1* −11-kb enhancer were analysed from E11.5 to E18.5. No obvious defects were detected in embryos from E11.5–E13.5; however, at E14.5, many homozygous embryos exhibited visible signs of oedema (Extended Data Fig. 3a,b). In addition, even in the absence of obvious subcutaneous oedema, many mutant embryos exhibited interstitial oedema accompanied by enlarged jugular lymph sacs and dilated dermal lymphatic vessels (Extended Data Fig. 3c). At E17.5 and E18.5, a common phenotype of homozygous embryos was blood-filled lymphatic vessels, particularly in the region of the axilla (Fig. 2e). Each of these phenotypes was consistent across every enhancer deletion line analysed. The skin and mesentery of mutant embryos exhibited striking abnormalities in lymphatic vascular growth and patterning; mutant embryos exhibited extremely distended, tortuous lymphatics and a notable absence of valves (Fig. 2f,g). Immunostaining of embryonic skin revealed reduced levels of PROX1 in dermal lymphatic endothelial cells of *Prox1* enhancer mutants (Fig. 2f). By contrast, PROX1 levels in

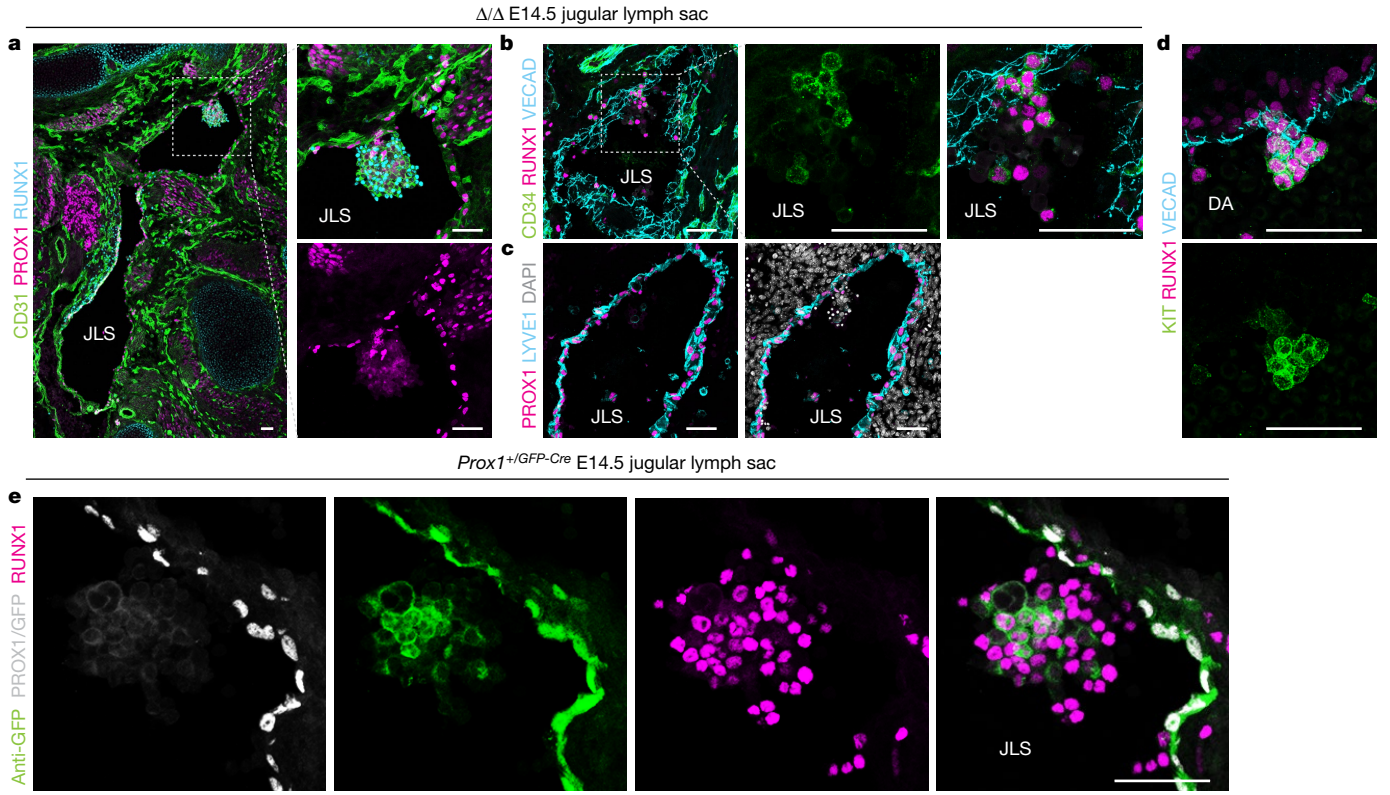



**Fig. 3 | Haematopoietic cell clusters in the jugular lymph sacs of *Prox1* enhancer mutant mice. a**, Coronal section of the E14.5 mutant embryo illustrating a large cluster of RUNX1$^+$CD31$^+$ cells budding from PROX1$^+$ lymphatic endothelium. **b**,**c**, Serial sections immunostained with markers of lymphatic endothelium (PROX1, LYVE1 and VE-cadherin (VECAD)) and haematopoietic stem and progenitor cells (CD34 and RUNX1). DAPI staining indicates various cellular identities in the clusters not detected by antibodies. **d**, Transverse section of a wild-type embryo at E10.5 shows a haematopoietic stem cell cluster in the aorta–gonad–mesonephros (AGM) region. DA, dorsal aorta. **e**, Coronal section of *Prox1*$^{+/GFP\text{-}Cre}$ embryo at E14.5 showing GFP perdurance in RUNX1$^+$PROX1$^-$ clusters. Image is representative of three embryos analysed. Scale bars, 50 μm.

mesenteric lymphatic vessels appeared uniformly high along the length of the vessels (Fig. 2g), consistent with mouse models of dysfunctional lymphatics resulting from defective valve development and, as a result, aberrant lymphatic flow[24]. No defects were observed in the blood vasculature of mutant embryos at any stage analysed.

## Enhancer regulation of *Prox1* mRNA in LECs

To determine whether deletion of the *Prox1* −11-kb enhancer impacted *Prox1* mRNA levels, primary LECs were purified from the skin of E18.5 wild-type, *Prox1* −11 kb$^{\Delta 5/\Delta 5}$ and *Prox1* −11 kb$^{\Delta 1,068/\Delta 1,068}$ embryos. *Prox1* mRNA levels were significantly reduced in LECs purified from embryos with deletions of the *Prox1* enhancer (Extended Data Fig. 4a). *Prox1* mRNA and protein levels were also assessed in the livers of *Prox1* −11 kb$^{\Delta/\Delta}$ embryos and their wild-type counterparts. Although *Prox1* is required for liver development and hepatocytes exhibit high levels of PROX1 protein, reporter gene expression was not obvious in the livers of *Prox1-11kblacZ* mice (Extended Data Fig. 1e) and *Prox1* mRNA and protein levels were not reduced (Extended Data Fig. 4b). These data demonstrate that the *Prox1* −11-kb enhancer acts in a tissue-specific manner.

Previous work demonstrated that reduction in *Prox1* dosage in LECs resulted in reversion of LEC identity towards BEC identity, an event proposed to facilitate the aberrant connection of lymphatic vessels with blood vessels and the filling of lymphatics with blood[13]. To investigate whether this was the case in *Prox1* −11 kb$^{\Delta/\Delta}$ embryos, primary dermal LECs were isolated from the skin, and the levels of markers characteristic of LEC identity (*Prox1* and *Flt4*), together with *Cd34*, a characteristic BEC marker, were assessed. Reduced levels of *Prox1* were associated with significantly reduced levels of *Flt4* and increased levels of *Cd34* in primary LECs isolated from E17.5 *Prox1* −11 kb$^{\Delta/\Delta}$ embryos (Extended Data Fig. 4d). Accordingly, levels of VEGFR3 and LYVE1 appeared reduced in the lymphovenous valve endothelial cells of E13.5 *Prox1* −11 kb$^{\Delta/\Delta}$ embryos (Extended Data Fig. 4e). To further investigate a potential reversion from LEC towards BEC identity, we performed RNA-seq of LECs purified from E14.5 wild-type and *Prox1* −11 kb$^{\Delta/\Delta}$ embryos and compared these gene expression profiles to a list of genes that we established to be the most differentially expressed between E14.5 LECs and BECs. These data demonstrate that the genes elevated in expression in *Prox1* −11 kb$^{\Delta/\Delta}$ LECs correlate significantly with the genes normally expressed more highly in BECs than LECs ($P < 0.0001$; Extended Data Fig. 5). Together, these data provide strong evidence demonstrating that reduced expression of *Prox1* in LECs of *Prox1* −11 kb$^{\Delta/\Delta}$ embryos results in the partial loss of LEC identity and reversion towards BEC identity.

## Deletion of the GATA site ablates transcription factor binding

The severity of the lymphatic vascular phenotype in *Prox1* −11 kb$^{\Delta/\Delta}$ embryos with a 5-bp deletion of the GATA2-binding site prompted us to investigate whether this deletion impacted the recruitment of FOXC2, NFATC1 and PROX1 to the *Prox1* −11-kb enhancer region in mutant LECs. To assess this, primary LECs were purified from the skin of E17.5 *Prox1* −11 kb$^{\Delta/\Delta}$ mouse embryos and wild-type embryos, and ChIP assays were performed to measure the binding of each of these transcription factors to the *Prox1* −11-kb enhancer. Excision of the GATA2-binding site abolished FOXC2, NFATC1 and PROX1 binding to

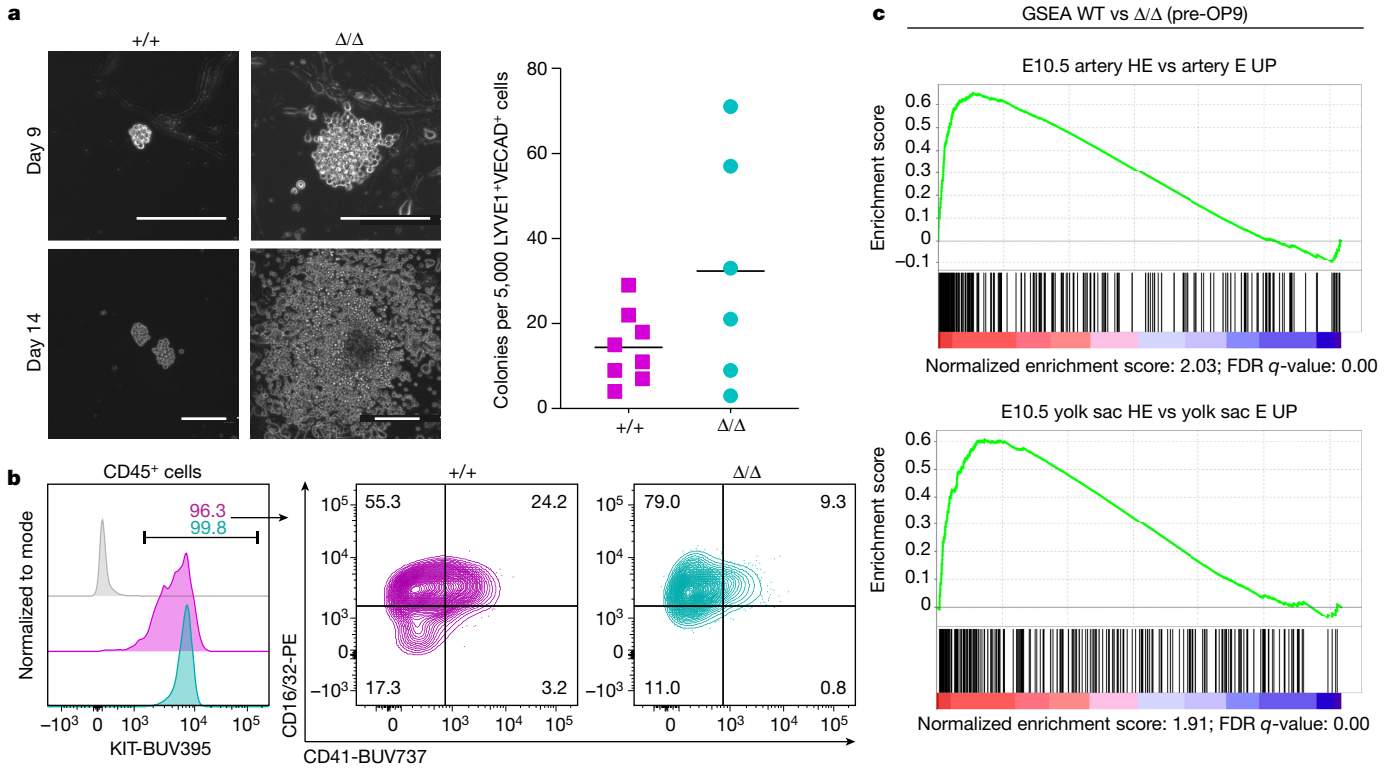



**Fig. 4 | LECs have haemogenic potential that is augmented by mutation of the *Prox1* −11-kb enhancer. a**, Methylcellulose colony-forming assay of LECs isolated from embryos at E14.5 and cultured on OP9 cells for 7 days. Data are colony numbers normalized to 5,000 isolated LECs with means indicated. Representative micrographs from at least six biological replicates are shown. Scale bars, 200 µm. **b**, FACS analysis of day 14 colonies arising from wild-type and mutant LECs shows the presence of cells positive for KIT, CD16/CD32 and CD41. **c**, Gene set enrichment analysis (GSEA) of genes differentially expressed between mutant and wild-type LECs at E14.5 shows significant enrichment when compared with genes upregulated (UP) in haemogenic endothelium (HE) from both the artery and the yolk sac at E10.5. Normalized enrichment scores indicate greater similarity with arterial HECs. E, non-haemogenic endothelium; FDR, false discovery rate.

the *Prox1* −11-kb enhancer in *Prox1* −11 kb$^{\Delta5/\Delta5}$ mouse LECs (Extended Data Fig. 6a). These data suggest that GATA2 acts as a pivotal factor to facilitate the binding of PROX1, FOXC2 and NFATC1 at the *Prox1* −11-kb enhancer, cumulatively increasing *Prox1* transcription, particularly in valve endothelial cells (Extended Data Fig. 6b). Chromosome conformation capture analysis confirmed that the *Prox1* −11-kb enhancer directly interacts with the *PROX1* promoter in human LECs (Extended Data Fig. 6c).

## LECs with haematopoietic potential

Haemogenic endothelial cells (HECs) undergo an endothelial-to-haematopoietic transition to generate haematopoietic stem and progenitor cells during development and reside in discrete embryonic tissues including the yolk sac[25], large arteries[26], heart[27] and placenta[28]. Analyses of lymphatic vascular development in *Prox1* −11 kb$^{\Delta/\Delta}$ embryos revealed that a striking feature of the jugular lymph sacs of *Prox1* −11 kb$^{\Delta/\Delta}$ embryos was the presence of clusters of RUNX1-positive cells (Fig. 3a–c) resembling the haematopoietic clusters found budding from haemogenic endothelium in the mouse dorsal aorta at E10.5 (Fig. 3d). Characterization of these cells revealed that clusters comprised various cellular identities (Fig. 3a–c and Extended Data Fig. 7). Immunostaining with a panel of haematopoietic and endothelial cell markers revealed that, like the haematopoietic clusters that bud from the wall of the dorsal aorta at E10.5, jugular lymph sac clusters exhibit markers characteristic of haematopoietic stem and progenitor cells including CD34, ESAM1, KIT and RUNX1 (Fig. 3 and Extended Data Fig. 7). To determine whether the reduced levels of *Prox1* in LECs of *Prox1* −11 kb$^{\Delta/\Delta}$ embryos might drive the formation of haematopoietic clusters, we investigated whether haematopoietic clusters were present in the jugular lymph sacs of *Prox1*$^{+/GFP\text{-}Cre}$ mice[15]. This was indeed the case; prominent clusters of RUNX1-positive haematopoietic cells were observed in the jugular lymph sacs of E14.5 *Prox1*$^{+/GFP\text{-}Cre}$ embryos (Fig. 3e). Moreover, the perdurance of GFP in PROX1-negative cells within these haematopoietic clusters strongly suggested that these cells were derived from lymphatic endothelium (Fig. 3e).

To further investigate the direct haemogenic capacity of LECs in *Prox1* −11 kb$^{\Delta/\Delta}$ embryos, primary LECs positive for LYVE1 and VE-cadherin but negative for CD45 were purified from the dorso–anterior region of wild-type and *Prox1* −11 kb$^{\Delta/\Delta}$ embryos and cultured on OP9 stromal cells[29] for 7 days. Cells were then harvested, and those positive for CD45, indicative of HECs undergoing a haematopoietic transition, were purified and plated in methylcellulose to assess haematopoietic stem and progenitor cell activity (Extended Data Fig. 8a). Primary LECs from both wild-type and *Prox1* −11 kb$^{\Delta/\Delta}$ embryos generated colonies, although the colonies arising from *Prox1* −11 kb$^{\Delta/\Delta}$ LECs were generally larger and more numerous than those arising from wild-type LECs (Fig. 4a). Close analysis of wild-type embryos revealed rare, small clusters of cells positive for RUNX1 and CD45 closely associated with the jugular lymph sac endothelium, suggesting that wild-type LECs may generate haematopoietic cells normally during development (Extended Data Fig. 7b). FACS analysis of cells within colonies revealed that they resembled the erythromyeloid progenitor cells found in the mouse yolk sac at approximately E9.5 (ref. [30]). Cells were positive for CD41, KIT, CD45 and CD16/CD32 and various combinations of these markers (Fig. 4b and Extended Data Figs. 7a,b and 9). RNA-seq analysis of LECs purified from wild-type and *Prox1* −11 kb$^{\Delta/\Delta}$ embryos revealed

that LECs purified from *Prox1* −11 kb$^{\Delta/\Delta}$ embryos expressed lower levels of characteristic lymphatic genes including *Prox1*, *Reln*, *Sema3d* and *Gja1* and higher levels of genes characteristic of HEC identity including *Kit*, *Emcn*, *Tal1*, *Runx1* and *Esam* before OP9 culture (Extended Data Figs. 4c and 10a). Gene set enrichment analysis of genes elevated in *Prox1* −11 kb$^{\Delta/\Delta}$ compared with wild-type LECs demonstrated significant similarity to the profiles of genes elevated in HECs compared with non-HECs in the E10.5 artery and yolk sac[31] (Fig. 4c). Moreover, LECs purified from wild-type embryonic skin exhibited higher expression of genes important for haematopoiesis, including *Gata2*, *Runx1*, *Kit*, *Hhex* and *Myb*, than BECs (Extended Data Fig. 10b). Together, these data demonstrate that wild-type LECs are poised to adopt HEC identity and that this switch is enhanced in LECs upon reduction of PROX1 levels.

## Discussion

The specification and maintenance of LEC identity is dependent on *Prox1* (refs. [11,13,32]). Here we identify a novel, tissue-specific enhancer element that is responsible for regulation of *Prox1* transcription and LEC identity. This enhancer is bound by four key transcriptional regulators of lymphatic vessel valve development—GATA2, FOXC2, NFATC1 and PROX1—and is particularly active in the endothelial cells that comprise lymphovenous and lymphatic vessel valves. Activity of the *Prox1* −11-kb enhancer was abolished by removal of only five nucleotides with a GATA2-binding site, revealing that GATA2 is pivotal in orchestrating enhancer activation and lymphatic vascular development. In considering the mechanism by which GATA2 might pioneer the assembly of a transcriptional activation complex at the *Prox1* −11-kb enhancer, it could be envisaged that in response to the disturbed flow that is characteristic of valve-forming regions[33], the levels of GATA2 are elevated in valve-forming territories[18]. Recruitment of GATA2 to the *Prox1* −11-kb enhancer may act to open chromatin, facilitating the binding of FOXC2 and NFATC1 (which have been established to bind to one another to regulate transcription in endothelial cells[20]), PROX1 and potentially additional transcription factors, cumulatively increasing *Prox1* transcription. On the basis of our previous work, in which we demonstrated that *Prox1* levels are reduced, but not ablated in *Gata2*-deficient embryos[18], we hypothesize that the *Prox1* −11-kb enhancer does not constitute an ON/OFF switch for *Prox1* transcription, but that it functions in a rheostat capacity to increase *Prox1* transcription, particularly in valve endothelial cells. Identification of additional enhancer elements that work together with the *Prox1* −11-kb enhancer to control *Prox1* transcription will further inform our understanding of the mechanisms important for orchestrating *Prox1* expression. We hypothesize that the perinatal lethality of approximately 50% of pups homozygous for *Prox1* −11-kb enhancer deletions reflects a critical threshold of *Prox1* that is required for lymphatic vascular development and that dropping below this threshold results in an extent of lymphatic vascular dysfunction incompatible with survival. The impact of reduced levels of *Prox1* is demonstrated by the survival of only a percentage of mice with one functional *Prox1* allele[11]. Our identification of non-coding regions of the genome in which *Prox1* enhancers reside will enable these regions to be interrogated for variants that might underlie human lymphatic vascular disorders, including lymphatic vascular malformations and primary lymphoedema.

Our discovery that the lymphatic endothelium of *Prox1* −11-kb enhancer mutant and *Prox1*$^{+/GFP\text{-}Cre}$ embryos exhibits augmented haemogenic capacity is, to our knowledge, the first report of lymphatic endothelium being competent to generate haematopoietic cells and suggests that PROX1 normally represses HEC identity. Temporal regulation of *Prox1* levels in LECs may endow them with the capacity to generate haematopoietic cells in times of stress or need. Our data are consistent with a previous report demonstrating that short hairpin RNA-mediated reduction of *Prox1* levels in mouse haematopoietic stem cells resulted in increased haematopoietic stem cell self-renewal and improved repopulation of irradiated recipients[34]. Moreover, our data suggest that the prominent phenotype of blood-filled lymphatic vessels in *Prox1* −11-kb enhancer mutant mice may, at least in part, be due to the release of haemogenic capacity in the lymphatic vasculature, such that LECs generate blood cells autonomously, rather than vessels filling with blood solely due to lymphovenous valve defects or aberrant connections between the blood and lymphatic vascular compartments. In conclusion, our data reveal that *Prox1* is crucial not only for programming LEC identity but also for repressing haemogenic cell identity in the lymphatic vasculature. Modulating these functions of PROX1 might prove valuable for purposes including stem cell programming/reprogramming and ex vivo generation/expansion of haematopoietic cells for regenerative medicine therapies.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-022-05650-9.


1. de Laat, W. & Duboule, D. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* **502**, 499–506 (2013).
2. Spitz, F. Gene regulation at a distance: from remote enhancers to 3D regulatory ensembles. *Semin. Cell Dev. Biol.* **57**, 57–67 (2016).
3. Rickels, R. & Shilatifard, A. Enhancer logic and mechanics in development and disease. *Trends Cell Biol.* **28**, 608–630 (2018).
4. Maurano, M. T. et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190–1195 (2012).
5. Oliver, G. et al. Prox1, a prospero-related homeobox gene expressed during mouse development. *Mech. Dev.* **44**, 3–16 (1993).
6. Wigle, J. T., Chowdhury, K., Gruss, P. & Oliver, G. Prox1 function is crucial for mouse lens-fibre elongation. *Nat. Genet.* **21**, 318–322 (1999).
7. Dyer, M. A., Livesey, F. J., Cepko, C. L. & Oliver, G. Prox1 function controls progenitor cell proliferation and horizontal cell genesis in the mammalian retina. *Nat. Genet.* **34**, 53–58 (2003).
8. Sosa-Pineda, B., Wigle, J. T. & Oliver, G. Hepatocyte migration during liver development requires Prox1. *Nat. Genet.* **25**, 254–255 (2000).
9. Wang, J. et al. Prox1 activity controls pancreas morphogenesis and participates in the production of "secondary transition" pancreatic endocrine cells. *Dev. Biol.* **286**, 182–194 (2005).
10. Risebro, C. A. et al. Prox1 maintains muscle structure and growth in the developing heart. *Development* **136**, 495–505 (2009).
11. Wigle, J. T. & Oliver, G. Prox1 function is required for the development of the murine lymphatic system. *Cell* **98**, 769–778 (1999).
12. Harvey, N. L. et al. Lymphatic vascular defects promoted by Prox1 haploinsufficiency cause adult-onset obesity. *Nat. Genet.* **37**, 1072–1081 (2005).
13. Johnson, N. C. et al. Lymphatic endothelial cell identity is reversible and its maintenance requires Prox1 activity. *Genes Dev.* **22**, 3282–3291 (2008).
14. Francois, M. et al. Sox18 induces development of the lymphatic vasculature in mice. *Nature* **456**, 643–647 (2008).
15. Srinivasan, R. S. et al. The nuclear hormone receptor Coup-TFII is required for the initiation and early maintenance of Prox1 expression in lymphatic endothelial cells. *Genes Dev.* **24**, 696–707 (2010).
16. Kazenwadel, J. et al. Loss-of-function germline GATA2 mutations in patients with MDS/AML or monoMAC syndrome and primary lymphedema reveal a key role for GATA2 in the lymphatic vasculature. *Blood* **119**, 1283–1291 (2012).
17. Ostergaard, P. et al. Mutations in GATA2 cause primary lymphedema associated with a predisposition to acute myeloid leukemia (Emberger syndrome). *Nat. Genet.* **43**, 929–931 (2011).
18. Kazenwadel, J. et al. GATA2 is required for lymphatic vessel valve development and maintenance. *J. Clin. Invest.* **125**, 2979–2994 (2015).
19. Petrova, T. V. et al. Defective valves and abnormal mural cell recruitment underlie lymphatic vascular failure in lymphedema distichiasis. *Nat. Med.* **10**, 974–981 (2004).
20. Norrmen, C. et al. FOXC2 controls formation and maturation of lymphatic collecting vessels through cooperation with NFATc1. *J. Cell Biol.* **185**, 439–457 (2009).
21. Srinivasan, R. S. & Oliver, G. Prox1 dosage controls the number of lymphatic endothelial cell progenitors and the formation of the lymphovenous valves. *Genes Dev.* **25**, 2187–2197 (2011).
22. Kothary, R. et al. Inducible expression of an hsp68-lacZ hybrid gene in transgenic mice. *Development* **105**, 707–714 (1989).
23. Shin, M. et al. Valves are a conserved feature of the zebrafish lymphatic system. *Dev. Cell* **51**, 374–386.e5 (2019).
24. Sweet, D. T. et al. Lymph flow regulates collecting lymphatic vessel maturation in vivo. *J. Clin. Invest.* **125**, 2995–3007 (2015).
25. Sabin, F. R. Preliminary note on the differentiation of angioblasts and the method by which they produce blood-vessels, blood-plasma and red blood-cells as seen in the living chick. 1917. *J. Hematother. Stem Cell Res.* **11**, 5–7 (2002).

26. de Bruijn, M. F., Speck, N. A., Peeters, M. C. & Dzierzak, E. Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo. *EMBO J.* **19**, 2465–2474 (2000).
27. Nakano, H. et al. Haemogenic endocardium contributes to transient definitive haematopoiesis. *Nat. Commun.* **4**, 1564 (2013).
28. Gekas, C., Dieterlen-Lievre, F., Orkin, S. H. & Mikkola, H. K. The placenta is a niche for hematopoietic stem cells. *Dev. Cell* **8**, 365–375 (2005).
29. Nakano, T., Kodama, H. & Honjo, T. Generation of lymphohematopoietic cells from embryonic stem cells in culture. *Science* **265**, 1098–1101 (1994).
30. McGrath, K. E. et al. Distinct sources of hematopoietic progenitors emerge before HSCs and provide functional blood cells in the mammalian embryo. *Cell Rep.* **11**, 1892–1904 (2015).
31. Gao, L. et al. RUNX1 and the endothelial origin of blood. *Exp. Hematol.* **68**, 2–9 (2018).
32. Wigle, J. T. et al. An essential role for Prox1 in the induction of the lymphatic endothelial cell phenotype. *EMBO J.* **21**, 1505–1513 (2002).
33. Sabine, A. et al. Mechanotransduction, PROX1, and FOXC2 cooperate to control connexin37 and calcineurin during lymphatic-valve formation. *Dev. Cell* **22**, 430–445 (2012).
34. Hope, K. J. et al. An RNAi screen identifies Msi2 and Prox1 as having opposite roles in the regulation of hematopoietic stem cell activity. *Cell Stem Cell* **7**, 101–113 (2010).

# Methods

### Animal studies

All experiments using mice were approved by the University of Adelaide, University of South Australia or SA Pathology/CALHN Animal Ethics Committees and conducted in accordance with the Australian code for the care and use of animals for scientific purposes. Mice used in this study were provided with water and standard chow ad libitum, and housed in a pathogen-free facility under the following conditions: 12–12 dark–light cycle that includes 30-min dusk and dawn cycles that run from 6:30–7:00 to 18:30–19:00, at 20.5–23.5 °C and humidity between 50% and 60%. Adult female mice subjected to timed pregnancies were scored by the presence of vaginal plugs, with 9:00 on the day of plug detection designated as 0.5 days post-coitum. *Prox1*$^{+/GFP\text{-}Cre}$ mice[15], C57BL/6 background, male and female, were analysed at E14.5. Prox1enh-hsp-LacZ transgenic mice (generated for this study), C57BL/6J background, male and female, were analysed at embryonic stages E11.5, E12.5, E14.5, E11.5, E17.5, E18.5 and P4. Prox1enh-CRISPR mice (generated for this study), C57BL/6J background, male and female, were analysed at embryonic stages E10.5, E14.5, E17.5, E18.5 and P0. Randomization and blinding were not performed. Owing to animal ethics considerations, sample size was determined according to the minimal number of independent biological replicates that significantly identified an effect. For most analyses, at least three sets of biological samples (litters of mice or individual embryos) were assessed.

Zebrafish work was carried out under ethical approval from the Swedish Board of Agriculture (5.2.18-7558/14). The fish were kept at the Genome Engineering Zebrafish National Facility (SciLifeLab, Uppsala, Sweden). Adults and embryos were housed according to the standard procedure. Previously published lines used in this work are *Tg(–5.2lyve1b:DsRed2)*$^{nz101}$ (ref. [35]), *TgBAC(prox1a:KalTA4-4xUAS-ADV.E1b:TagRFP)*$^{nim5}$ (refs. [36,37]) and *Tg(kdr-l:ras-Cherry)*$^{s916}$ (ref. [38]). The *Tg(–11Prox1:eGFP;XCA:DsRed2)*$^{uu4kk}$ and *Tg(–2.1prox1a:eGFP;XCA:DsRed2)*$^{uu3kk}$ lines were generated for this study.

### Transgenic reporter mice

A genomic fragment encompassing the mouse *Prox1* −11 kb region GRCm38/mm10 chr1:190,181,703-190,182,534 (832 bp) was generated by PCR using primers forward (5′- GGCAAGCATGGGCATGGTGGAT-3′) and reverse (5′-AGCATGGCCTTGAGGCTCGGT-3′) and cloned into the polylinker of pKS-hsp-lacZpA[22] (a gift from J.Rossant). A 5.05 kb *SalI* fragment containing genomic DNA, *hsp68* promoter and *LacZ* gene was purified and used for pronuclear injection of fertilized C57BL/6N embryos. Following implantation in pseudo-pregnant females, resulting pups were screened for presence of the *LacZ* transgene by PCR using primers (LacZ-forward 5′-GAACCATCCGCTGTGGTACA-3′) and (LacZ-reverse 5′-TTAGCGAAACCGCCAAGACT-3′) and allowed to develop to maturity to establish stable lines. F1 progeny were screened by PCR and X-gal staining as previously described[12] and subsequently bred into a C57BL/6J background for at least ten generations.

### Transgenic reporter zebrafish

A 282-bp element showing conservation with the mouse enhancer sequence was identified on the basis of an mVista non-coding DNA conservation analysis[39,40] and is located approximately 2.1 kb upstream of *prox1a* (GRCz11/danRer11 chr.17: 32,867,987–32,868,268). The analysis included the upstream and downstream non-coding regions of the *prox1a* locus of selected Osteichthyes species and sequences were aligned using the LAGAN alignment program[41]. The mouse *–11Prox1* and the zebrafish *–2.1prox1a* elements were cloned into the ZED construct as previously described[42], using the following primers:

(*–2.1prox1a*_Fwd 5′-GGGGACAAGTTTGTACAAAAAAGCAGGCTGACAAAAGGAAAGGCGTGTTG-3′); (*–2.1prox1a*_Rev 5′-GGGGACCACTTTGTACAAGAAAGCTGGGTGTGACAGAGATCAGAGGTGG-3′); (*–11Prox1*_Fwd 5′-GGGGACAAGTTTGTACAAAAAAGCAGGCTAGCATGGCCTTGAGGCTCGGT-3′); (*–11Prox1*_Rev 5′-GGGGACCACTTTGTACAAGAAAGCTGGGTGGCAAGCATGGGCATGGTGGAT-3′).

Each construct was injected with 1 nl of 40 ng µl$^{-1}$ plasmid DNA and 100 ng µl$^{-1}$ *tol2* transposase mRNA into the one-cell stage wild-type zebrafish embryos and raised to adulthood. Offspring from ten F0 fish per transgenic line were screened to confirm the expression pattern, and a positive founder for each construct was used to generate stable transgenic lines.

### CRISPR-mutant mice

Guide RNA was designed using the online CRISPR tool developed by the Zhang laboratory at MIT (http://crispr.mit.edu) and synthesized as overlapping oligonucleotides with appropriate overhangs. The target sequence was as follows: 5′-GCCCAGCCGCTCCAGATAAG-3′ GRCm38/mm10 chr. 1: 190,181,971–190,181,990. Overlapping oligonucleotides were phosphorylated and annealed, then cloned into the *BbsI* sites in pX330-U6-chimeric_BB-CBh-hSpCas9 (pX330, Addgene plasmid #42230). C57BL/6J embryos were injected cytoplasmically with CRISPR reagents, transferred into pseudo-pregnant recipients on the same day and allowed to develop to term. Founder pups were screened for insertions or deletions (indels) by two PCR amplification reactions across the targeted region: a 493-bp wild-type amplicon (shortF-5′-CTGGGCCTGTGGTGAGTAAT-3′ and shortR-5′- GGTCACTGTCTTTCCGAAGC-3′) and a 1,532-bp wild-type amplicon (longF-5′-AGAGCTTCTGGGAAAGCAGC-3′ and longR-5′-TGCTTCCCGGTCAGTTTTCA-3′). PCR products from indel-carrying founders were Sanger sequenced to identify specific mutations. Six separate founder lines with deletions ranging from 5 bp encompassing the GATA-binding site to 1,068 bp were further analysed (Extended Data Fig. 1c,d). All founders were backcrossed to wild-type mice to select for individual mutant alleles in F1 progeny and further backcrossed for at least three generations to eliminate potential off-target artefacts. CRISPR-mutant mice were screened by PCR as above. In the case of the smallest 5-bp deletion, the short amplicon was purified and digested with *Xcm1*. This restriction site is present in wild-type mice and absent in mice carrying a 5-bp deletion. Genotypes were periodically confirmed by Sanger sequencing of PCR products.

### Histology

Embryos were removed from pregnant females at E18 and washed in PBS. A small amount of Bouin's solution was injected into the thorax and abdomen to assist fixation and the embryos were further fixed in Bouin's solution for 48 h at room temperature. This was followed by extensive washing in 70% ethanol at room temperature. Embryos were then placed in 4% paraformaldehyde in PBS and delivered to the Australian Phenomics Network for paraffin embedding, sectioning (5 µm) and haematoxylin and eosin staining.

### Immunostaining

For frozen sections and whole-mount staining of skin, embryos were fixed in 4% paraformaldehyde overnight at 4 °C. For whole-mount staining of the mesentery, embryos were dissected, and mesenteries were removed and fixed in 4% paraformaldehyde for 10 min at room temperature. Sections and tissues were immunostained and imaged using confocal microscopy as previously described[16,18]. Images were captured at room temperature using a Carl Zeiss LSM 700 Axio Observer Z1 confocal microscope equipped with four solid lasers (near UV 405, green 488 nm, red 555 nm and far-red diode 637 nm) or a Carl Zeiss LSM 800 Axio Observer 7 confocal microscope with Airyscan, equipped with 405-nm, 488-nm, 561-nm and 640-nm lasers. Images were compiled using ZEN 2.5 (blue edition; Zeiss) and Adobe Photoshop CC (version 21.1.1) software.

### Zebrafish imaging

Transgenic embryos were mounted laterally and ventral–laterally in 1% low-melting agarose and imaged in the face or trunk using a Leica TCS

SP8 DLS microscope with a Fluotar VISR 25X water objective (objective number: 11506375). Images were processed using ImageJ 2.0.0.

### Antibodies

For immunofluorescent immunostaining, the primary antibodies used were rabbit anti-GATA2 (1 in 500; NBP1-82581, Novus), rabbit anti-PROX1 (1 in 1,000; ab101851, Abcam), rabbit anti-LYVE1 (1 in 1,000; 11-034, AngioBio), goat anti-PROX1 (1 in 250; AF2727, R&D Systems), rat anti-CD31 (1 in 500; 553370, BD Pharmingen), rat anti-CD34 (1 in 250; 14-0341, eBioscience), rat anti-CD117/Kit (1 in 250; 14-1171, eBioscience), goat anti-ESAM (1 in 250; AF2827, R&D Systems), rat anti-endomucin (1 in 500; sc-65495, Santa Cruz), goat anti-VE-cadherin (1 in 250; AF1002, R&D Systems), Cy3-conjugated mouse monoclonal anti-α-smooth muscle actin (1 in 1,000; C6198, Sigma), rat anti-FOXC2 (ref. [43]) (1 in 1,000), goat anti-VEGFR3 (1 in 250; AF743, R&D Systems), rabbit anti-RUNX1 (1 in 1,000; ab92336, Abcam), rabbit anti-β-galactosidase (1 in 5,000; #55976, MP Biomedicals), chicken anti-β-galactosidase (1 in 1,000; ab9361, Abcam) and rabbit anti-GFP (1 in 500; A-11122, Thermo Fisher Scientific). The Alexa Fluor fluorochrome-conjugated antibodies used for detection were donkey anti-rat IgG (H+L) Alexa Fluor 488 (1 in 500; A-21208, Thermo Fisher Scientific), donkey anti-goat IgG (H+L) Alexa Fluor 488 (1 in 500; A-11055, Thermo Fisher Scientific), donkey anti-syrian hamster IgG (H+L) Alexa Fluor 488 (1 in 500; A-21110, Thermo Fisher Scientific), donkey anti-rabbit IgG (H+L) Alexa Fluor 488 (1 in 500; A-21206, Thermo Fisher Scientific), donkey anti-rabbit IgG (H+L) Alexa Fluor 555 (1 in 500; A-31572, Thermo Fisher Scientific), donkey anti-goat IgG (H+L) Alexa Fluor 555 (1 in 500; A-21432, Thermo Fisher Scientific), donkey anti-chicken IgG (H+L) Alexa Fluor 594 (1 in 500; 703-585-155, Jackson ImmunoResearch), donkey anti-chicken IgG (H+L) Alexa Fluor 647 (1 in 500; 703-606-155, Jackson ImmunoResearch), donkey anti-goat IgG (H+L) Alexa Fluor 647 (1 in 500; A-32879, Thermo Fisher Scientific), donkey anti-rabbit IgG (H+L) Alexa Fluor 647 (1 in 500; A-31573, Thermo Fisher Scientific), and chicken anti-rat IgG (H+L) Alexa Fluor 647 (1 in 500; A-31573, Thermo Fisher Scientific). For ChIP, the antibodies used were rabbit anti-GATA2 (sc9008X, Santa Cruz), goat anti-FOXC2 (ab5060, Abcam), rabbit anti-NFATC1 (sc13033X, Santa Cruz), goat anti-human PROX1 (AF2727, R&D Systems) and rabbit IgG (#2729, Cell Signaling).

### Primary dermal endothelial cell isolation

Primary embryonic dermal lymphatic and blood endothelial cells were isolated at E14.5, E16.5 and E18.5 as previously described[44]. Cells for RNA extraction were immediately processed following isolation using RNAeasy Minikit (Qiagen). Cells for immunostaining and ChIP were plated on gelatin-coated culture dishes or chamber slides. Cells were expanded for 2–4 days and either fixed for 10 min with PFA for immunostaining or harvested using 0.05% trypsin–EDTA and processed for ChIP.

### Cell lines for ChIP and 3C analysis

For adult hLECs: HMVEC-dLyAd-Der Lym Endo, Lonza (cat. CC-2810, lot 7F3304 and 0000254463). For adult hBECs: HMVEC-dBlAd, Lonza (cat. CC-2811, lot 0000125028). Cell lines were authenticated by the supplier and confirmed by immunostaining and quantitative PCR (qPCR) analysis using appropriate markers of endothelial cell identity. Cells were not tested for mycoplasma and were used within four passages.

### cDNA synthesis and qRT–PCR analysis

cDNA was synthesised using a QuantiTect Reverse Transcription kit (Qiagen) incorporating a guide DNA clean-up step. qPCR with reverse transcription (qRT–PCR) was performed with RT2 Real-Time SYBR Green/Rox PCR master mix (Qiagen) and analysed on a Rotor-Gene6000 (Qiagen). Data were normalized to the housekeeping gene *Actb* as previously described[44].

Primer sequences (mouse) were as follows: *Actb* Fwd_5′-GATCA TTGCTCCTCCTGAGC-3′ and *Actb* Rev_5′-GTCATAGTCCGCCTAGA AGCAT-3′; *Cd34* Fwd_5′-TCCCCATCAGTTCCTACCAA-3′ and *Cd34* Rev_5′-CAGTTGGGGAAGTCTGTGGT-3′; *Flt4* Fwd_5′-CTGGCCAG AGGCACTAAGAC-3′ and *Flt4* Rev_5′-CAGGGTGTCCTCTGGGAA TA-3′; *Gata2* Fwd_5′-ATGGGCACCCAGCCTGCAAC-3′ and *Gata2* Rev_5′-GTGGCCCGTGCCATCTCGTC-3′; *Prox1* Fwd_5′-CTGGGCCA ATTATCACCAGT-3′ and *Prox1* Rev_5′-GCCATCTTCAAAAGCTCGTC-3′; and *Runx1* Fwd_5′-TTTCGCAGAGCGGTGAAAGAA-3′ and *Runx1* Rev_5′-CAGCGCCTCGCTCATCTT-3′.

### ChIP

Cells were harvested and processed for ChIP using a truChIP Low Cell Chromatin Shearing Kit with SDS Shearing Buffer (Covaris). In brief, 10 million cells per millilitre were crosslinked using 1% formaldehyde for 5 min, neutralized with glycine, lysed and nuclei washed. For transcription factor ChIP in cultured hLECs, chromatin from 3 million cells was sheared using a 130-µl microtube in a Covaris sonicator at the recommended settings for 8 min and 5 µg antibody or IgG control was used to immunoprecipitate sheared DNA as previously described[18]. DNA was purified using a Qiagen MinElute PCR Purification Kit and recovery of *PROX1* −11-kb enhancer sequences analysed by qPCR using a Roche LightCycler 480 and Universal Probe Library (human) probe #43 with specific primers, as follows: forward 5′-AGCCAGGGAATGAGTACAGG-3′ and reverse 5′- AGGAAGCCTGTGCATTAACAC-3′. Recovery of *PROX1* promoter sequences was analysed using Universal Probe Library (human) probe #82 with specific primers, as follows: forward 5′-AA TAGTTGGAGGTGTGAGTGGTG-3′ and reverse 5′-GCGTCTATCACG GAAGCAA-3′. For ChIP in embryonic mouse LECs, chromatin from 0.5 to 1 million primary cells was sheared and 1 µg antibody or IgG control was used. Following washing and reversal of crosslinks, DNA was purified by phenol/chloroform extraction and ethanol precipitation using linear acrylamide and glycogen as carriers. Recovery of *Prox1* −11-kb enhancer sequences was analysed by SYBR Green qRT–PCR using primers as follows: mouse forward 5′-CTTGCCAAAGGATGGAGAGA-3′ and mouse reverse 5′-TGGCCAACATTAAAGGGAGA-3′.

### 3C analysis

Detection of physical interaction between the *Prox1* promoter and −11-kb enhancer was performed and quantified following published protocols[45]. In brief, $10^7$ hLECs were crosslinked and chromatin was digested with EcoRI before religation and reversal of crosslinks. PCR was performed using an anchor primer with primers specific for each of 12 fragments upstream of the promoter and PCR products were measured using standard agarose gel quantification. Interaction frequencies were calculated relative to a control library generated from BAC clone RP11-783K13. Primer sequences are listed in Supplementary Data.

### Cell isolation for RNA-seq and colony-forming assays

Litters consisting of 6–8 pooled embryos of a single genotype (wild-type or homozygous mutant) were used for each isolation. In brief, at E14.5 the dorsal–anterior regions of embryos, as indicated in Extended Data Fig. 8, were dissected at room temperature in HHF (5% FCS and 10 mM HEPES in Hanks balanced salt solution). Care was taken to eliminate the livers, lungs, heart and thymus from torsos before rinsing briefly with DMEM/20% FCS. Tissue was digested in 10 ml DMEM/20% FCS containing 25 mg collagenase type II, 25 mg collagenase type IV and 10 mg deoxyribonuclease I (Worthington) for 30 min at 37 °C while mixing gently with a wide-bore transfer pipette every 5 min to assist tissue dissociation. Cell suspensions were filtered through a 40-mm cell strainer. Filtrates were centrifuged at 200*g* for 10 min and resuspended in 5 ml HHF at room temperature. Cells were counted (generally approximately $5–10 \times 10^6$ cells per embryo) and centrifuged for a further 5 min at 300*g*. The resulting pellet was resuspended in 1 ml HHF containing 1:100 dilution of F4/80 monoclonal antibody (clone BM8, Thermo Fisher), incubated at room temperature for 5 min, and F4/80-positive cells were depleted using anti-rat MACS beads (Miltenyi Biotech) according to

Q21

the manufacturer's instructions. Following F4/80 MACS depletion, the cells were lineage depleted using biotinylated lineage antibodies and biotin binder Dynabeads (11047, Thermo Fisher). Lineage-depleted cells were resuspended in sort buffer (2% FBS, 5 μM EDTA, 25 mM HEPES pH 7, and 10 U $ml^{-1}$ DNAse I in phenol red-free HBSS) and incubated for 10 min at room temperature before addition of fluorochrome-conjugated monoclonal antibodies: anti-CD144 BV421 (1 in 100; clone 11D4.1, 747749, BD Biosciences), LYVE-1 PE (1 in 100; clone 223322, FAB2125P, R&D Systems) and CD45 APC-Cy7 (1 in 100; clone 30-F11, 557659, BD Biosciences). Cells were incubated with antibodies for 20 min at room temperature, washed with 3 ml of sort buffer and resuspended in sort buffer with SYTOX Red Dead Cell Stain (5 nM; S34859, Invitrogen). Samples were sorted using MoFlo Astrios EQ cell sorter (70-μm nozzle; Beckman Coulter). For RNA-seq analysis pre-OP9 culture, half of the sorted cells were pelleted at 300*g* for 5 min, resuspended in 1 ml TRIzol reagent (Thermo Fisher) and stored at −80 °C. Remaining sorted cells were plated on OP9 feeder cells (approximately 70% confluent) in 10% MEM-α containing cytokines FLT3L, IL-3 and SCF (Peprotech) at a concentration of 100 ng $ml^{-1}$ each. Fresh media containing murine cytokines was topped up on day 2 and day 4. Cells were harvested on day 7 and stained with fluorochrome-conjugated antibodies as described above before cell sorting. For RNA-seq analysis post-OP9, $LYVE1^+VECAD^+$ cells were sorted directly in 1 ml TRIzol reagent and stored at −80 °C. For haematopoietic colony assays, sorted $CD45^+$ cells were seeded in MethoCult (M3434, Stem Cell Technologies) and incubated in a humidified chamber at 37 °C before colonies were enumerated on day 9. Colonies were harvested on day 14 and the cells were stained with antibodies: anti-CD117 BUV395 (1 in 100; clone 2B8, 564011, BD Biosciences), anti-CD11b APC (1 in 100; clone M1/70, 101211, BioLegend), anti-CD71 BV510 (1 in 100; clone C2, 563112, BD Biosciences), anti-CD144 BV421 (1 in 100; clone 11D4.1, 747749, BD Biosciences), anti-CD45 APC-Cy7 (1 in 100; clone 30-F11, 557659, BD Biosciences), anti-CD41 BUV737 (1 in 100; clone MWReg30, 741759, BD Biosciences), anti-Ly-6G PECy7 (1 in 100; clone 1A8, 560601, BD Biosciences) and anti-CD16/32 PE (1 in 100; clone 93, 101307, BioLegend). Cells were incubated with antibodies for 20 min at room temperature, washed with 3 ml of sorting buffer and resuspended in sorting buffer before data acquisition on a BD LSR Fortessa flow cytometer. Unstained cells were used as a negative control. Flow cytometry data were analysed using FlowJo software (version 10.7.1, Becton Dickinson). The gates used to identify populations of interest are shown in Extended Data Fig. 8a.

### RNA-seq and bioinformatic analysis

Sorted cells pre-OP9 and post-OP9 culture were stored in TRIzol at −80 °C until ready for processing. RNA was prepared using Direct-Zol Microprep (Zymo Research) according to the manufacturer's instructions and eluted in a final volume of 10 μl, and RNA quality was assessed using a Bioanalyser PicoChip (Agilent Technologies). RNA was submitted to the ACRF Cancer Genomics Facility (Adelaide) and sequenced using a Smart-seq Stranded Kit (Takara Bio). Single replicate total RNA-seq libraries for pre-OP9 and post-OP9 wild-type and homozygous mutant samples were multiplexed and sequenced on the Illumina NextSeq 500 platform using the stranded, paired-end protocol with a read length of 75. Raw data, averaging 139 million reads per sample, were analysed and quality checked using the FastQC program (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Reads were mapped against the mouse reference genome (mm10) using the STAR spliced alignment algorithm[46] (version 2.5.3a with default parameters and --chimSegmentMin 20, --quantMode GeneCounts), returning an average unique alignment rate of 83%. Alignments were visualized and interrogated using the Integrative Genomics Viewer v2.8.9 (ref. [47]). Duplicate reads derived as a consequence of the ultra-low amounts of starting RNA and the SMART-seq protocol were removed using the markdup function of Sambamba[48] (v0.6.7 with settings; --remove-duplicates, --nthreads 16, --overflow-list-size 600000), retaining an average of 57 million deduplicated reads per sample. Deduplicated read counts in annotated genes were tabulated using the htseq-count function of HTSeq[49] (v0.11.2 with settings; --format bam --order pos --stranded yes --minaqual 10). Differential expression was evaluated from TMM-normalized gene counts using R (version 4.1.1) and edgeR (version 3.34.0)[50], following protocols as described in section 2.12 of the edgeR User's Guide (http://www.bioconductor.org/packages/release/bioc/vignettes/edgeR/inst/doc/edgeRUsersGuide.pdf). In brief, the data were normalized and filtered to remove genes with low counts across all libraries, descriptive analyses were performed (MDS and scatter plots) before changes in gene expression between libraries were calculated ($log_2$ fold change). Gene-set enrichment analysis (GSEA) was subsequently performed to look for coordinate expression to groups of genes in the Molecular Signatures Database (MSigDB)[51] and likewise, but in a focused manner, to select gene sets from HECs[31] relevant to this experiment. Genes were ranked for the GSEA analysis using the fold-change measurements between libraries. Heatmaps were generated using the heatmap.2 function from the gplots package in R, using log-transformed and mean-subtracted counts to improve visualization.

### RNA microarrays

Embryonic mouse LECs and BECs were purified from E14.5, E16.5 and E18.5 dermis as detailed in ref. [44] and were obtained from at three independent litters containing 5–7 embryos at each time point. Between 0.5 and 3 mg of total RNA was isolated using TRIzol reagent (Thermo Fisher) according to the manufacturer's directions. RNA quality was assessed using a Bioanalyser (Agilent Technologies); all samples achieved an RNA integrity number score of more than 9.5. Triplicate samples were submitted to the ACRF Cancer Genomics Facility (Adelaide, South Australia, Australia) and hybridized to GeneChip Mouse Gene 1.0 ST arrays (Affymetrix) for gene expression profiling. Microarray data analysis was performed using Partek Genomics Suite version 6.4 software (Partek Incorporated). Microarray data analysis was performed using R (version 4.1.1) and the Bioconductor package Oligo (version 1.60.0). Data were preprocessed using the RMA method for background correction and normalization[52]. Heatmaps were generated using the heatmap.2 function from the gplots package in R.

### Statistics and reproducibility

For survival curves, *P* values were calculated using the log-rank (Mantel–Cox) test. For all other statistical analyses, *P* values were calculated using the two-tailed Student's *t*-test unless otherwise noted in figure legends. *P* values of less than 0.05 were considered statistically significant. Representative micrographs from at least three biological replicates are shown.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

The datasets and material generated during the current study are available from the corresponding author on reasonable request (N.L.H.). GATA2 ChIP–seq data have been deposited in the European Nucleotide Archive under accession number PRJEB9436. PROX1, FOXC2 and NFATC1 ChIP–seq and human LEC and BEC RNA-seq data have been submitted to the Gene Expression Omnibus (GEO) under accession number GSE129634. Mouse LEC (pre-OP9 or post-OP9) RNA-seq and mouse LEC and BEC microarray data have been submitted to the GEO under accession number GSE184046. The HE and E RNA-seq data used for GSEA (Fig. 4) were generated in a published study[31] and were obtained from the GEO database under the accession number GSE103813. Source data are provided with this paper.

35. Okuda, K. S. et al. lyve1 expression reveals novel lymphatic vessels and new mechanisms for lymphatic vessel development in zebrafish. *Development* **139**, 2381–2391 (2012).
36. Dunworth W. P. et al. Bone morphogenetic protein 2 signaling negatively modulates lymphatic development in vertebrate embryos. *Circ. Res.* **114**, 56–66 (2014).
37. van Impel, A. et al. Divergence of zebrafish and mouse lymphatic cell fate specification pathways. *Development* **141**, 1228–1238 (2014).
38. Hogan, B. M. et al. Ccbe1 is required for embryonic lymphangiogenesis and venous sprouting. *Nat. Genet.* **41**, 396–398 (2009).
39. Dubchak, I. et al. Active conservation of noncoding sequences revealed by three-way species comparisons. *Genome Res.* **10**, 1304–1306 (2000).
40. Frazer, K. A., Pachter, L., Poliakov, A., Rubin, E. M. & Dubchak, I. VISTA: computational tools for comparative genomics. *Nucleic Acids Res.* **32**, W273–W279 (2004).
41. Brudno, M. et al. LAGAN and multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. *Genome Res.* **13**, 721–731 (2003).
42. Bessa, J. et al. Zebrafish enhancer detection (ZED) vector: a new tool to facilitate transgenesis and the functional analysis of cis-regulatory regions in zebrafish. *Dev. Dyn.* **238**, 2409–2417 (2009).
43. Furumoto, T. A. et al. Notochord-dependent expression of MFH1 and PAX1 cooperates to maintain the proliferation of sclerotome cells during the vertebral column development. *Dev. Biol.* **210**, 15–29 (1999).
44. Kazenwadel, J., Michael, M. Z. & Harvey, N. L. Prox1 expression is negatively regulated by miR-181 in endothelial cells. *Blood* **116**, 2395–2401 (2010).
45. Naumova, N., Smith, E. M., Zhan, Y. & Dekker, J. Analysis of long-range chromatin interactions using chromosome conformation capture. *Methods* **58**, 192–203 (2012).
46. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
47. Thorvaldsdottir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform.* **14**, 178–192 (2013).
48. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032–2034 (2015).
49. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
50. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
51. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
52. Irizarry, R. A. et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249–264 (2003).

**Acknowledgements** We thank the staff at the UniSA CRI Core Animal Facility for animal husbandry and D. Lawrence for assistance with initial bioinformatic analyses. This study utilized the Australian Phenomics Network Histopathology and Organ Pathology Service (University of Melbourne) and the South Australian Genome Editing Service (University of Adelaide. Confocal microscopy and flow cytometry were performed at the Detmold Imaging and Flow Cytometry Facility (UniSA). This work was supported by grants from the NHMRC (1061365 and 1146706 to N.L.H. and H.S.S.) and ARC (DP210103351 to N.L.H.). K.K. and V.P. were supported by Wallenberg Fellowships (2017.0144), Vetenskapsådet (VR-MH-2016-01437) and the Kjell and Märta Beijer Foundation. P.V. was supported by a fellowship from Maddie Riewoldt's Vision (MRV0017). N.L.H. was supported by an ARC Future Fellowship (FT130101254). H.S.S. was supported by an NHMRC Principal Research Fellowship (1023059) and Cancer Council SA's Beat Cancer Project on behalf of its donors and the State Government of South Australia through the Department of Health.

Q25

**Author contributions** J.K., H.S.S. and N.L.H. conceived the study. P.V., J.K. and A.O. performed the haematopoiesis assays. V.P. and K.K. designed, generated and assessed transgenic zebrafish. S.G.P. and P.Q.T. advised on design and generated transgenic and genome-edited mice. C.B. analysed mouse phenotypes. W.M. generated and provided *Prox1*$^{+/GFP\text{-}Cre}$ embryos. J.K. performed ChIP, ChIP–seq and RNA-seq experiments and assessed mouse lymphatic phenotypes. J.K., J.T., L.A.-M. and A.S. analysed the bioinformatic data. S.T. advised on embryonic haematopoiesis assays. J.K. and N.L.H. wrote the manuscript. All authors edited and approved the manuscript.

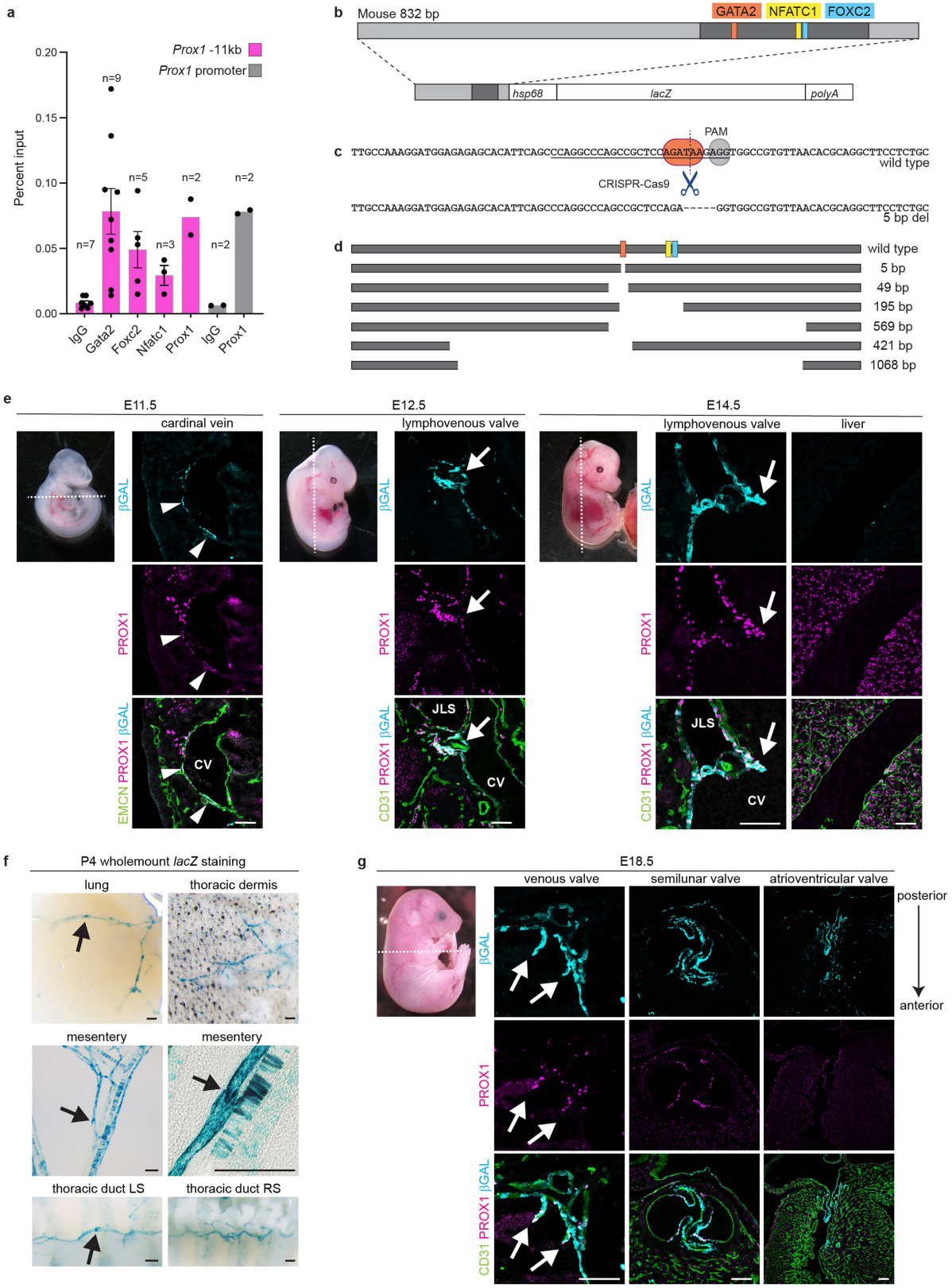
**Competing interests** The authors declare no competing interests.

**Additional information**

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41586-022-05650-9.

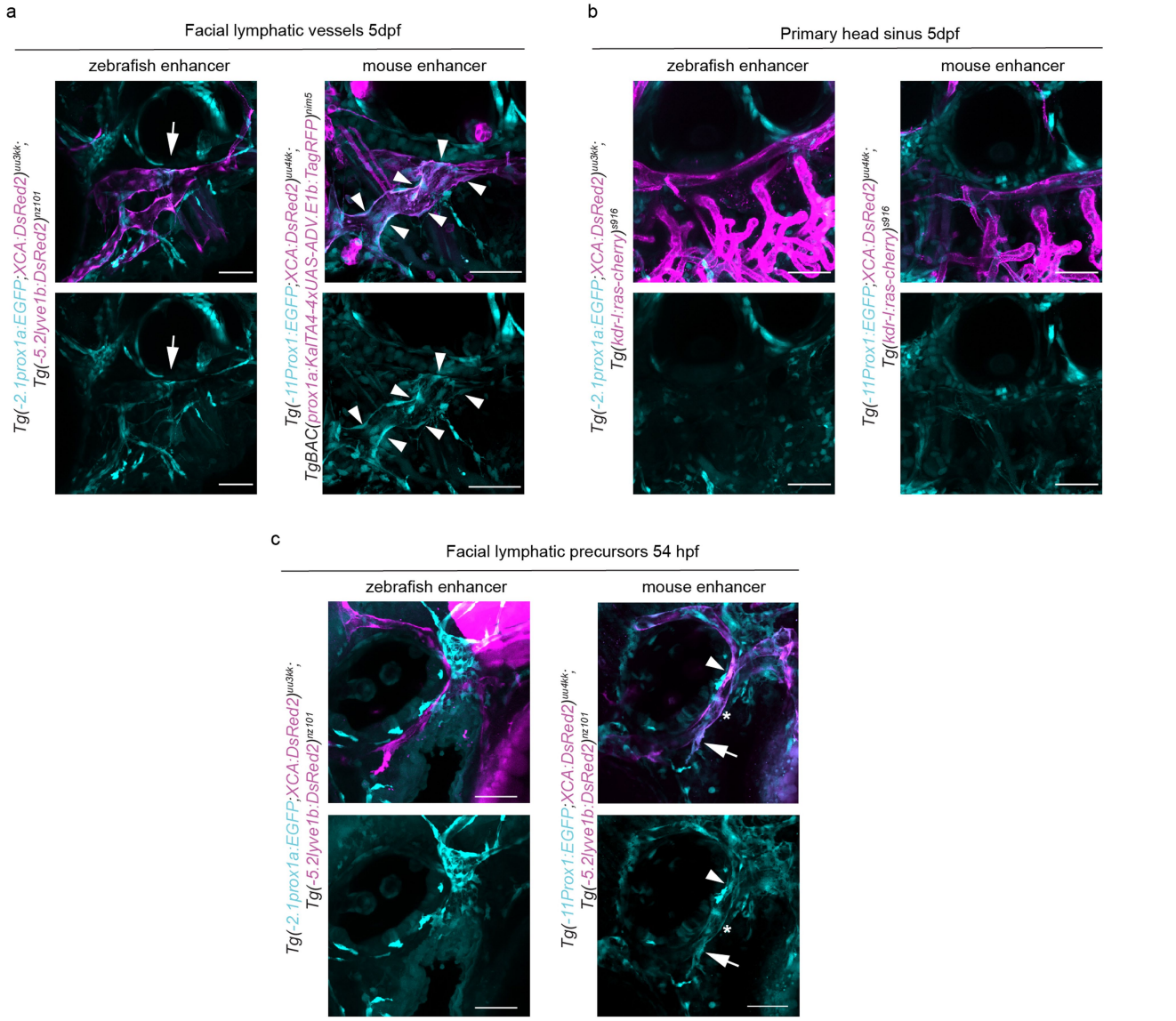
**Correspondence and requests for materials** should be addressed to Natasha L. Harvey.

**Peer review information** *Nature* thanks Hanna Mikkola and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

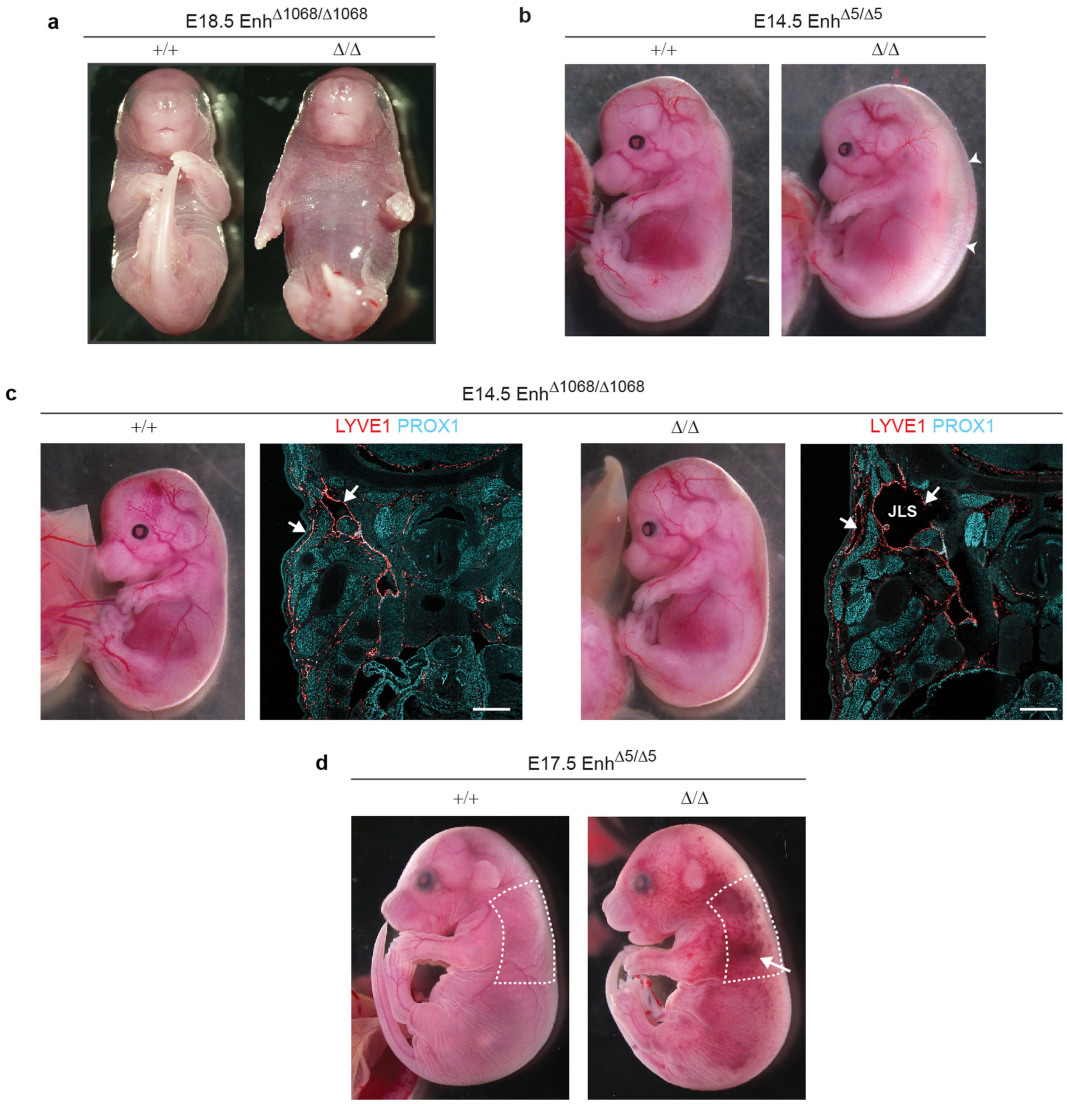



**Extended Data Fig. 1** | See next page for caption.

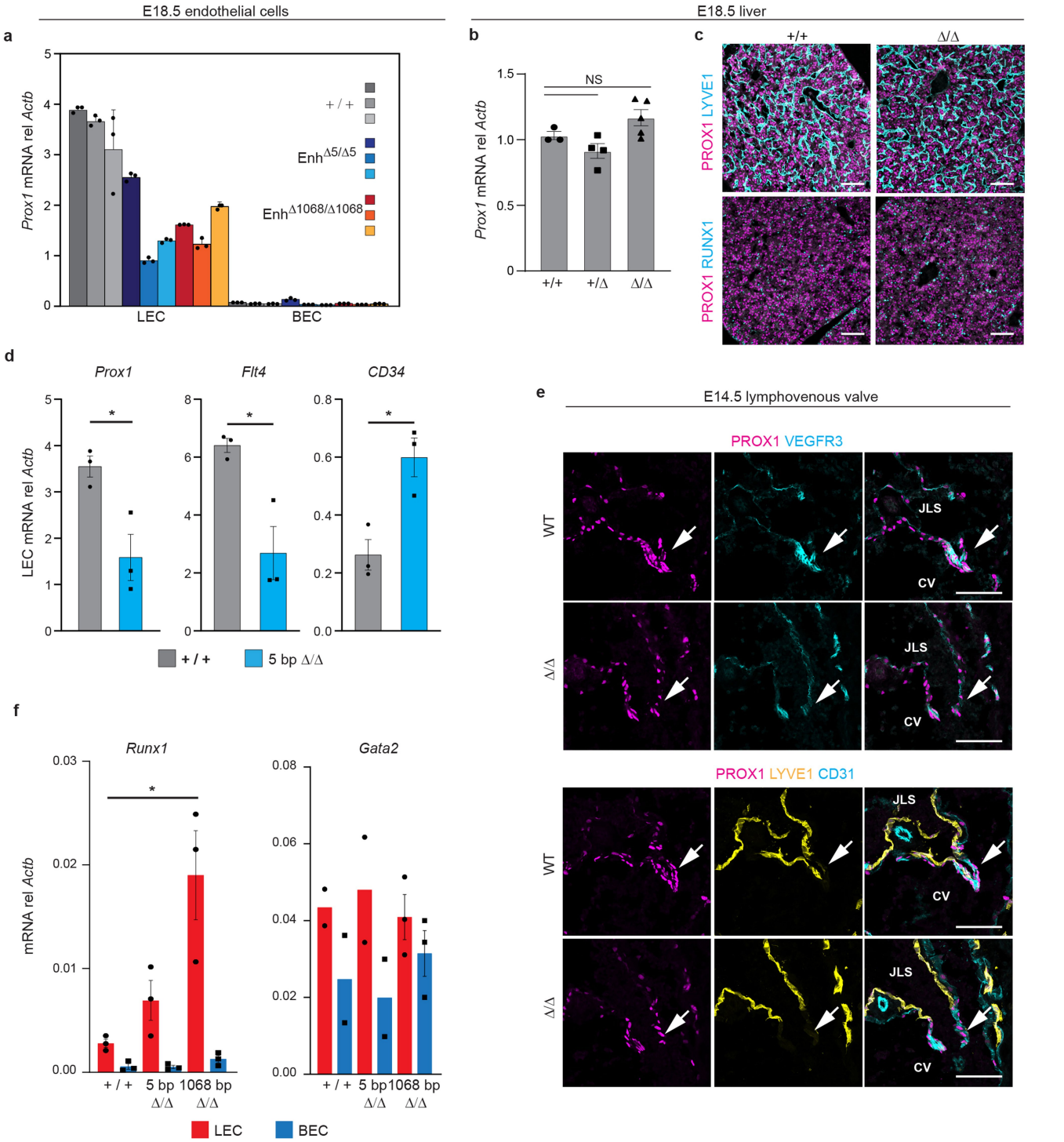
**Extended Data Fig. 1 | The *Prox1* −11kb enhancer drives reporter gene expression in lymphatic endothelial cells and at high levels in valves.** **a**, ChIP in hLECs demonstrates binding of GATA2, FOXC2, NFATC1 and PROX1 at the Prox1 −11 kb enhancer and promoter regions. Data are independent experiments and shown as mean ± SEM when n > 2. **b**, Schematic of construct used to generate stable transgenic reporter mice. **c**, Strategy for CRISPR-Cas9 mediated deletion of the *Prox1* −11 kb element. Guide RNA sequence targeting GATA2 binding site (underlined) and resulting 5 bp deletion are indicated. **d**, CRISPR-Cas9 mediated deletion series. **e**, Immunofluorescent analysis of mouse embryos carrying the *Prox1* −11 kb enhancer driven *LacZ* reporter transgene. Transverse sections at E11.5 show β-galactosidase activity is detected in PROX1$^+$ endothelial cells lining the cardinal vein (arrowheads). Coronal sections in the jugular region at E12.5 and E14.5, reveal high levels of reporter activity in lymphovenous valves (arrows) while there is no detectable reporter gene expression in PROX1$^+$ hepatocytes in E14.5 liver. **f**, Wholemount X-gal staining of tissues from transgenic mouse pups at post-natal day 4. β-galactosidase activity is present in the lung, dermis, thoracic duct, and mesenteric lymphatic vasculature. In lymphatic vasculature at early post-natal stages, reporter activity is restricted to larger collecting vessels and is not observed in lymphatic vessels in tissues analysed by wholemount X-gal staining at adult stages >P28. Black arrows indicate valves; left side (LS); right side (RS). **g**, The *Prox1* −11 kb enhancer drives reporter gene expression in venous and cardiac valves. Transverse sections of transgenic mouse embryos at E18.5 show reporter activity in venous valves (arrows), semilunar and atrioventricular cardiac valves. Scale bars, 100 μm (**e, g**), 200 μm (**f**).



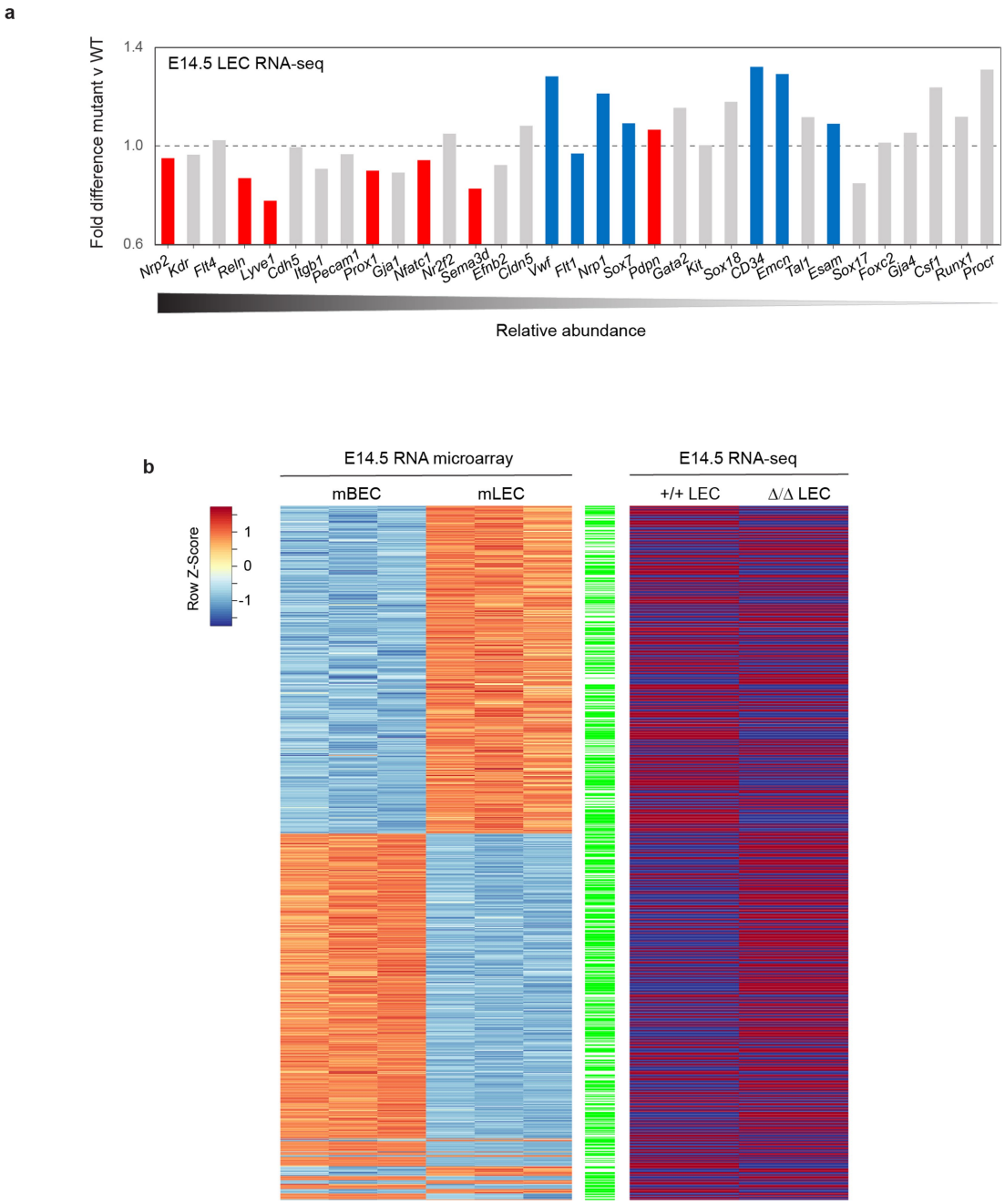

**Extended Data Fig. 2 | Enhancer driven reporter activity in zebrafish. a**, At 5 days post-fertilization (dpf) zebrafish reporter lines for the zebrafish *–2.1 kb prox1a* and mouse *–11 kb Prox1* enhancer elements drive GFP expression (cyan) in the facial lymphatic endothelium (arrowheads), including the facial valve (arrow) as demonstrated by overlapping expression with *Tg(–5.2lyve1b:DsRed2)*$^{nz101}$ or *TgBAC(prox1a:KalTA4-4xUAS-ADV.E1b:TagRFP)*$^{nim5}$ (both magenta). **b**, In both enhancer reporter lines at 5dpf, endothelial GFP signal (cyan) is restricted to the lymphatic vessels in the face, as shown by lack of co-expression with *Tg(kdr-l:ras-cherry)*$^{s916}$, which marks the venous endothelium of the primary head sinus (magenta). Additional domains of non-endothelial expression in the face appear to be induced ectopically by both enhancers. **c**, At 54 h post-fertilization, after sprouting of the facial lymphatics has commenced, co-expression of mouse enhancer-driven GFP (cyan) with lyve1b is observed in the lymphatic progenitors coming from two venous niches, the common cardinal vein lymphangioblasts (CCV-L, arrowhead) and the primary head sinus lymphangioblasts (PHS-LP, arrow). Asterisk indicates co-expression in the underlying PHS. Scale bars, 50 µm.



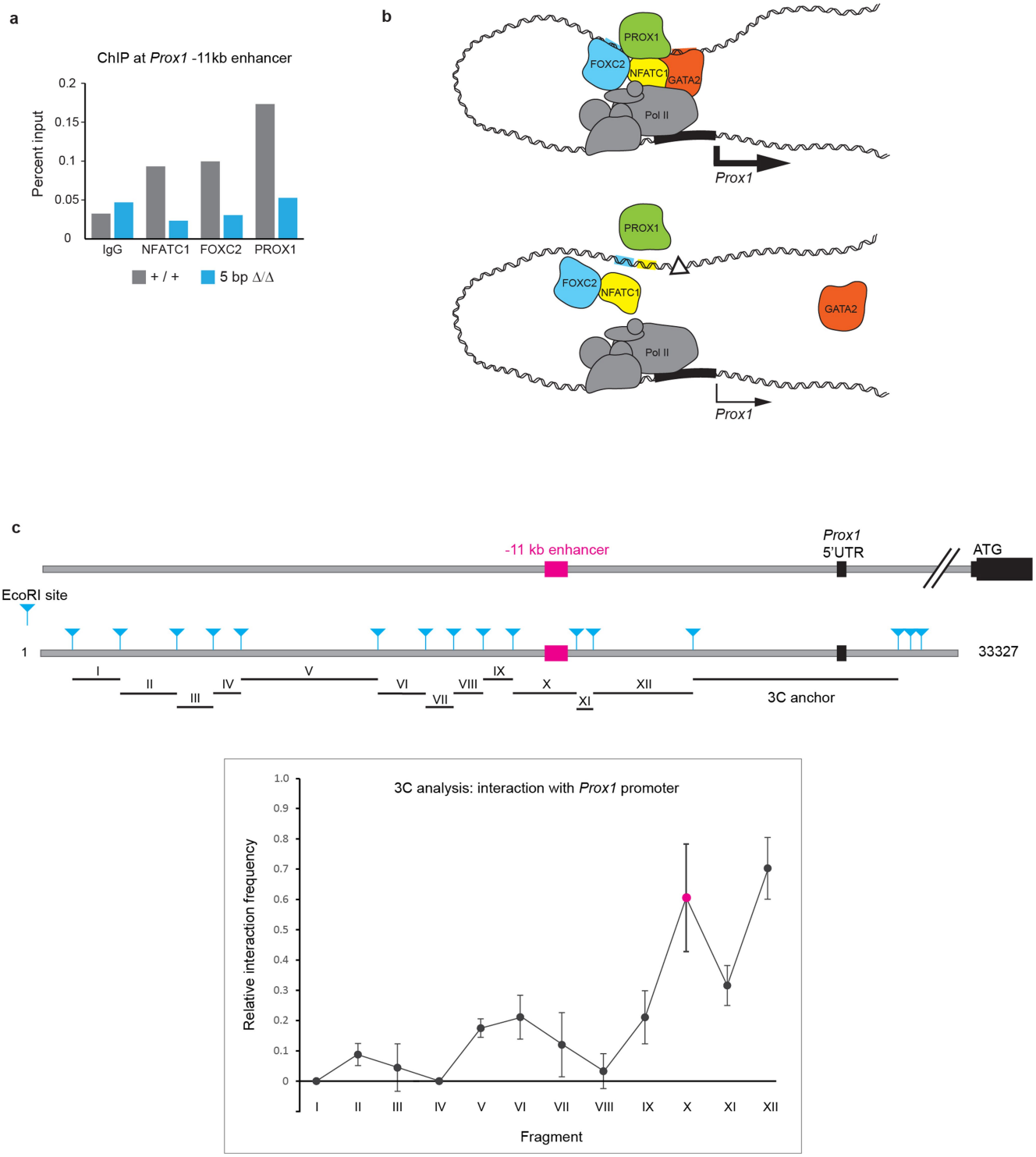

**Extended Data Fig. 3 | Oedema and lymphatic vascular defects are similar in mutant embryos with a 1068 bp or 5 bp deletion of the *Prox1* −11 kb enhancer. a**, Ventral view illustrating jugular oedema in enhancer mutant compared with wildtype littermate. **b**, Mutant and wild type littermates at E14.5. Arrowheads indicate subcutaneous oedema. Oedema was observed in 17% (6/35) of homozygous E14.5 *Prox1* −11 kb$^{\Delta5/\Delta5}$ embryos and 11.7% (7/60) of heterozygous E14.5 *Prox1* −11 kb$^{+/\Delta5}$ embryos. **c**, Compared with wildtype littermates, mutant embryos exhibit enlarged jugular lymph sacs and dilated dermal lymphatic vessels (arrows) in the absence of overt oedema. **d**, E17.5 embryos subjected to wholemount immunostaining of skins and mesenteries in Fig. 2f, g. showing the region of dermis used for staining. Blood filled vessels in the region of the axilla are indicated (arrow). Scale bars, 400 µm.



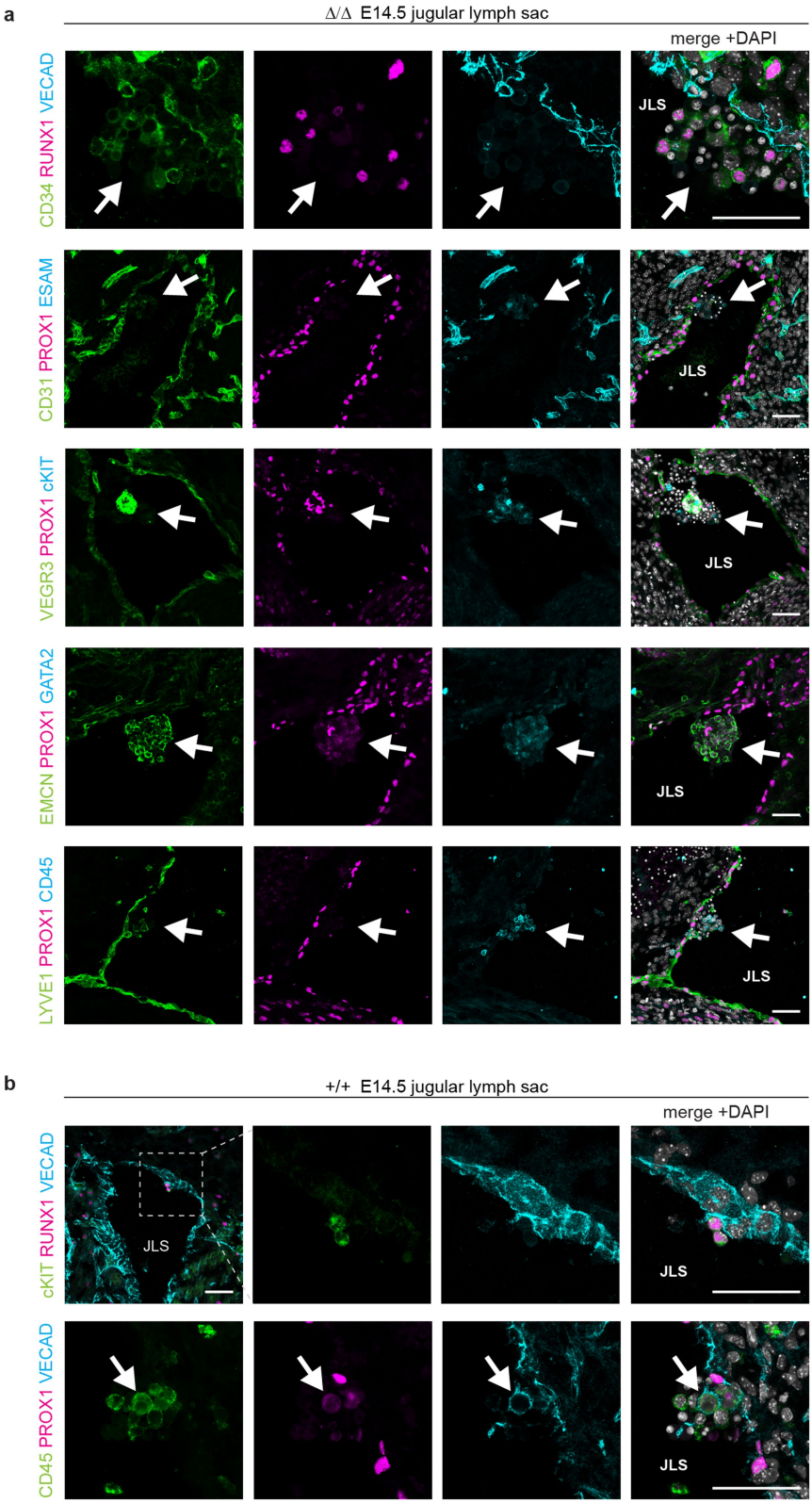

**Extended Data Fig. 4 | The *Prox1* −11 kb enhancer controls Prox1 mRNA levels in lymphatic endothelial cells. a**, *Prox1* mRNA levels in E18.5 primary dermal LECs and BECs. Data represent individual litters (5–7 embryos) of each genotype and are shown as mean values ± SD (n = 3 replicates for each litter). **b**, *Prox1* mRNA levels in livers isolated from embryos of each genotype at E18.5. Data shown as mean ± SD, unpaired two-tailed *t* test with no adjustment for multiple comparisons. **c**, Immunostaining of liver sections taken from embryos at E14.5 show no differences in PROX1 levels (red). **d**, Reduction of *Prox1* mRNA in LECs (*$P$ = 0.02310) is accompanied by reduced *Flt4* (*$P$ = 0.01706) and increased *CD34* expression (*$P$ = 0.01705). Data represent average expression in LECs from three independent litters and are shown as mean ± SEM., unpaired two-tailed *t* test with no adjustment for multiple comparisons. **e**, Immunostaining of coronal sections of E14.5 embryos demonstrates reduced levels of PROX1 (magenta), LYVE1 (yellow) and VEGFR3 (cyan) in lymphovenous valves (arrows) of mutant embryos compared with wild type littermates. **f**, *Runx1* and *Gata2* mRNA levels in primary dermal LECs and BECs isolated at E18.5. Data represent average expression from three independent litters and are shown as mean values ± SEM. *$P$ = 0.0101, ordinary one-ANOVA with Dunnett's multiple comparisons (where error bars are not shown n = 2). Scale bars, 100 μm (**c, e**).



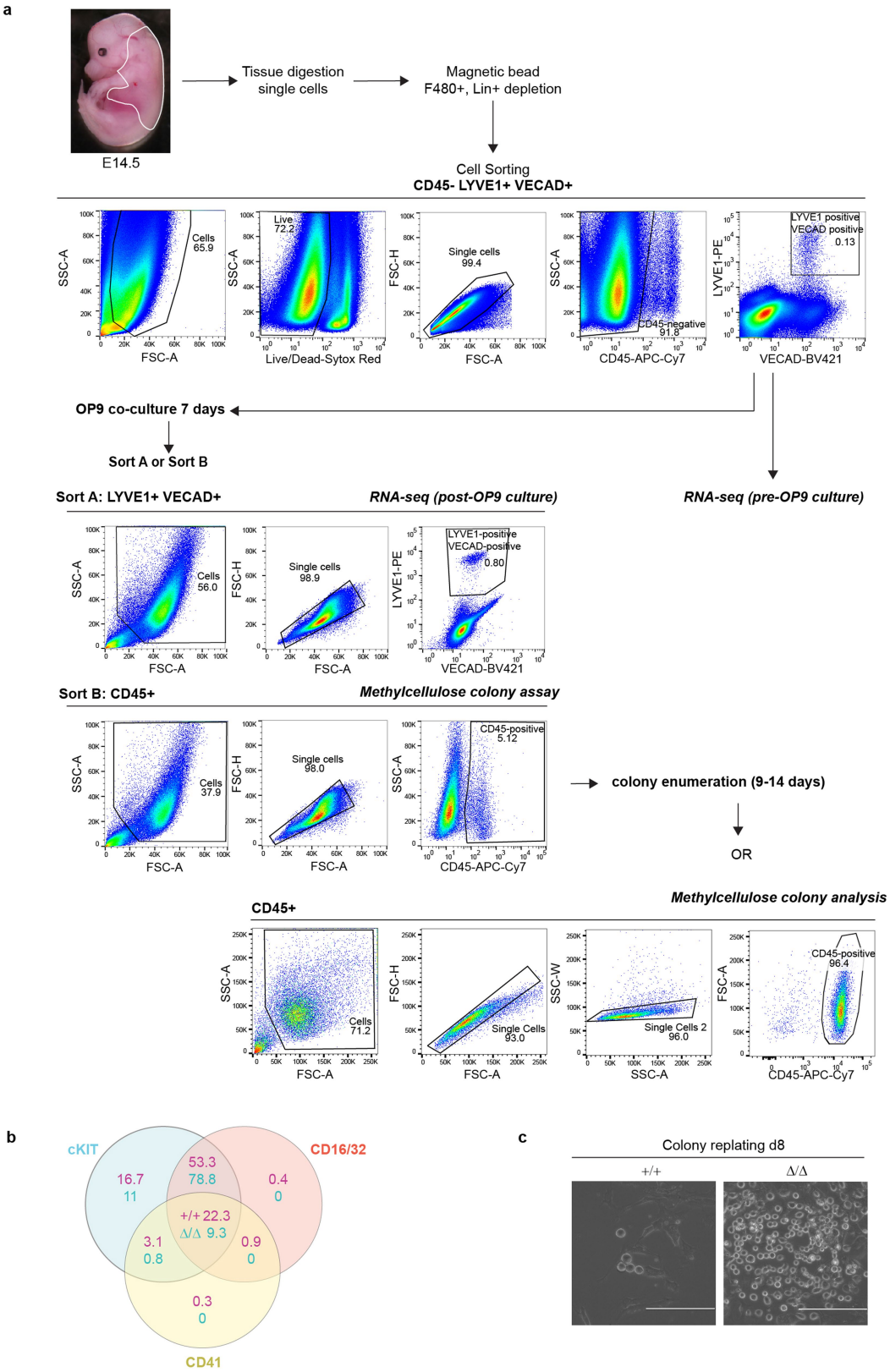

**Extended Data Fig. 5 | Transcriptomic analysis of lymphatic endothelial cells indicates a shift in identity of enhancer mutant LECs towards that of blood endothelial cells. a**, Differential gene expression in RNA-seq analysis of LECs isolated from mutant and wildtype embryos at E14.5. Selected genes are ranked in order of expression level in wildtype (highest to lowest, left to right) with markers of LEC identity (red) and BEC identity (blue) highlighted. **b**, Heatmap comparing microarray analysis of E14.5 LEC and BEC RNA with RNA-seq data shows genes up-regulated in enhancer mutant LECs correlate with genes expressed at higher levels in BECs and vice versa. Green bars indicate genes with a positive correlation to a shift in identity of LEC to BEC in mutant versus wildtype. A Fisher's Exact Test shows the association between microarray and RNA-seq outcomes is significant, two-tailed p value < 0.0001.




**Extended Data Fig. 6 | Deletion of the GATA site ablates binding of PROX1, FOXC2 and NFATC1 to the *Prox1* −11 kb enhancer. a**, ChIP assays using primary LECs isolated from mutant embryos at E17.5 show no enrichment over IgG control when the GATA2 site is absent. **b**, Model proposing that GATA2 acts as a pivotal factor at the *Prox1* −11 kb enhancer to promote recruitment of transcriptional componentry responsible for driving *Prox1* expression in LECs. **c**, Quantification of interaction frequency of the *Prox1* −11 kb enhancer with the *Prox1* promoter in cultured human LEC. Chromosome Conformation and Capture (3C) analysis of regions proximal to the *Prox1* gene demonstrates increased interaction frequency relative to a BAC control, of the anchor fragment (containing the promoter) and fragments X (containing the enhancer) and XII. Data are shown as mean values ± SD, n = 3. Primer sequences and quantification are available in Source Data Extended Data Fig. 6.

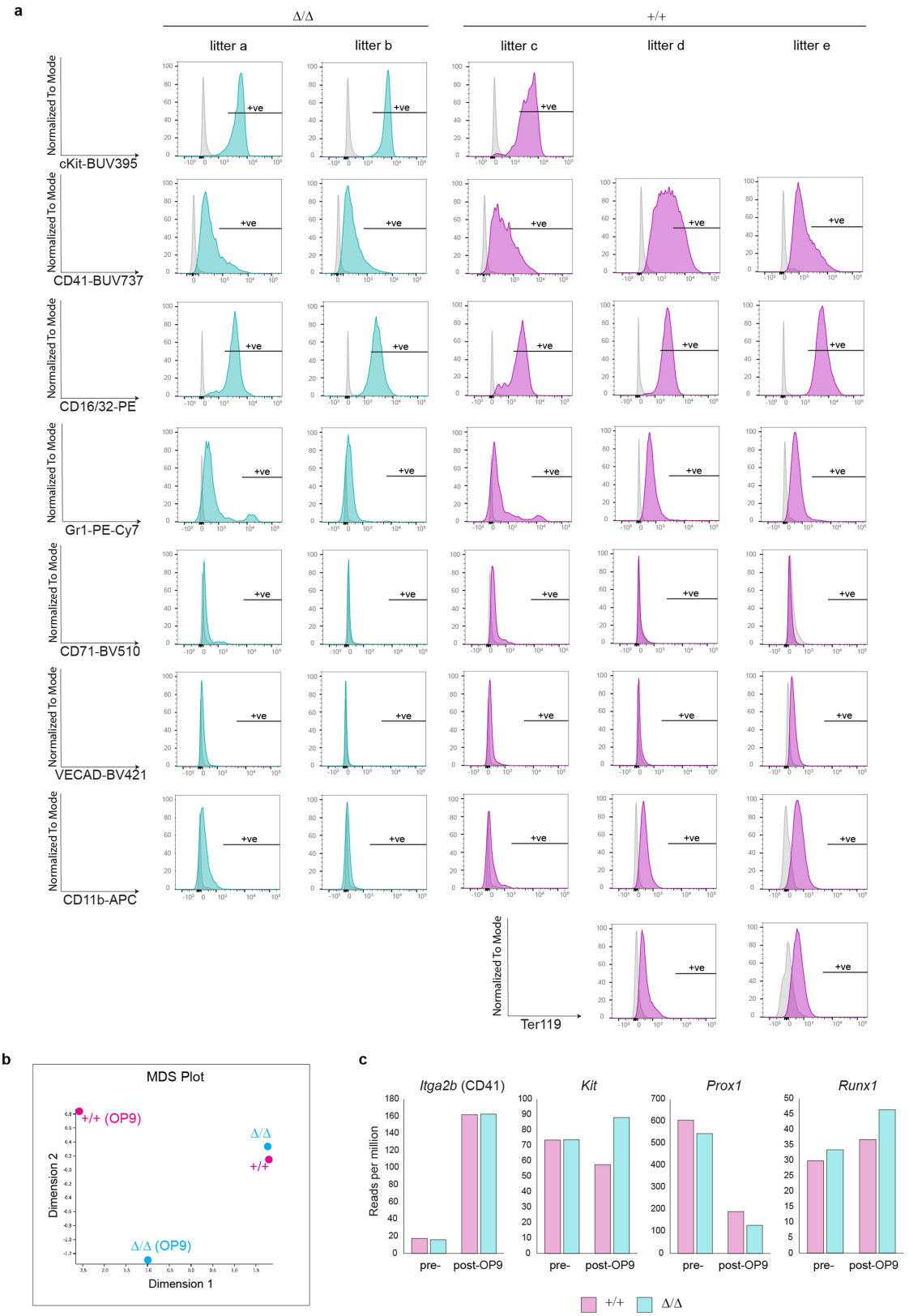



**Extended Data Fig. 7 | Cell clusters in the jugular lymph sacs of *Prox1* enhancer mutant mice express range of hematopoietic and endothelial markers. a**, Coronal sections of E14.5 mutant embryos illustrating cells within clusters budding from PROX1+ LYVE1+ CD31+VECAD+ lymphatic endothelium are heterogenous in identity and are variously positive for RUNX1, cKIT, CD45, GATA2 and CD34. Data are representative of clusters observed in 7 of 15 embryos analyzed at E14.5. **b**, Rare, small clusters were also observed budding from the jugular lymph sacs of wildtype embryos at E14.5. These clusters are positive for RUNX1, cKIT and CD45. A CD45+ VECAD+ PROX1+ cell is indicated (arrow). Data are representative of two independent embryos. JLS, jugular lymph sac. Scale bars, 50 μm.



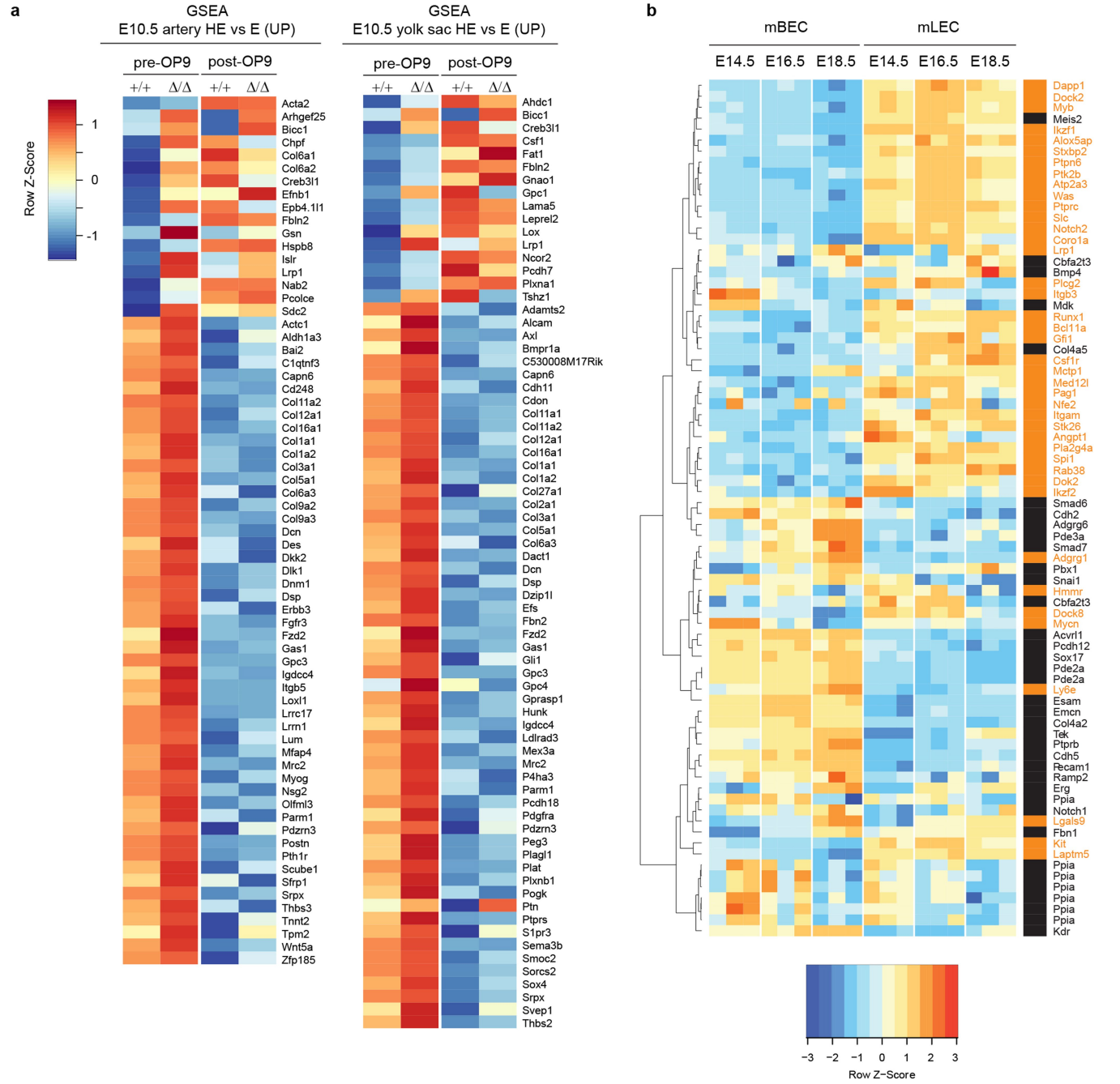

**Extended Data Fig. 8** | See next page for caption.

**Extended Data Fig. 8 | Gating strategies used for sorting and characterization of lymphatic and haemogenic endothelial cells. a**, Dorso-anterior regions of E14.5 embryos from wildtype (+/+) or homozygous mutant (Δ/Δ) litters were dissected as indicated, taking care to remove liver, lungs hearts and thymus. 6–8 torsos from a single litter were pooled and digested to generate a single cell suspension. Following F480+ Lin+ depletion, FACS sorted CD45-LYVE1+VECAD+ cells were plated on OP9 feeder layers with cytokines for 7 days. In the case of transcriptomic analyses (Extended Data Figs. 5a,b, 10a), half of the cells from each genotype were processed for RNA (pre-OP9) while the other half were grown on OP9 and then sorted to purify LYVE1+VECAD+ cells (post-OP9). For methylcellulose colony assays all CD45+ cells were FACS purified from OP9 co-cultures, plated into Methocult and cultured for 9–14 days (Fig. 4a). Colonies were enumerated and harvested for FACS analysis (Fig. 4b and Extended Data Fig. 9a). **b**, FACS analysis of colonies arising after 14 days demonstrates differences and overlap between enhancer mutant and wildtype LEC derived colonies. Venn diagram shows percentages of CD45+ cells also positive for CD41, CD16/32 and cKIT in each genotype. Data are representative of 5 independent experiments. **c**, Methylcellulose colonies derived from wildtype and enhancer mutant LECs have replating capacity. Colonies were harvested 14 days after initial plating, replated in Methocult™ and assessed at d8. Cells from enhancer mutant colonies demonstrate enhanced proliferation. Data are representative of three independent experiments. Scale bars, 200 μm.




**Extended Data Fig. 9** | See next page for caption.

**Extended Data Fig. 9 | Hematopoietic colonies arising from lymphatic endothelial cells express markers characteristic of erythromyeloid progenitor cells which originate from the yolk sac. a**, Cells harvested from methylcellulose cultures and positive for CD45 were assessed by staining with a range of antibodies for hematopoietic markers. Colonies from both wildtype and enhancer mutant LECs include populations of cells positive for CD41, CD16/32 and cKIT, and negative for other markers analysed, except for a small population of Gr1+ cells observed in two litters (D/D litter a; +/+ litter c). **b**, Multidimensional scaling (MDS) plot of RNA-seq data from wildtype and mutant LECs isolated at E14.5. While both genotypes show similarity prior to co-culture with OP9 cells, RNA sequence analysis reveals a divergent response of the transcriptome of wildtype compared with enhancer mutant LECs post-OP9 culture. **c**, Comparison of RNA levels (expressed as reads per million) in wildtype and enhancer mutant LEC pre- and post-OP9 culture, demonstrates an increase in *Itga2b* transcripts in both genotypes, consistent with the levels of CD41 observed in FACS analysis. Expression of *Kit* is increased in enhancer mutant LEC but decreased in wildtype, which is also reflected in FACS analysis of cKIT levels, while levels of *Prox1* are reduced and *Runx1* levels are increased in enhancer mutant cells.




**Extended Data Fig. 10 | Transcriptomic analyses show mutation of the *Prox1* −11 kb enhancer results in a shift towards a hemogenic endothelial identity. a**, RNA-seq analysis of genes differentially expressed between mutant and wildtype LECs isolated at E14.5, pre- and post-OP9 coculture. Heatmaps show relative expression of genes identified in gene set enrichment analysis (GSEA) in Fig. 4c. **b**, Microarray analysis of RNA isolated from wild-type E14.5, E16.5 and E18.5 dermal LEC and BEC shows that LECs express higher levels of hematopoietic and hemogenic endothelial genes than do BECs. Heatmap of gene expression highlighting a selection of genes. Haematopoietic genes are marked in orange while those marked in black are endothelial genes. These data suggest that LECs are poised to acquire hemogenic endothelial cell identity.

# Author Queries

Journal: **Nature**
Paper: **s41586-022-05650-9**
Title: **A Prox1 enhancer represses haematopoiesis in the lymphatic vasculature**

**AUTHOR:**
The following queries have arisen during the editing of your manuscript. Please answer by making the requisite corrections directly in the e.proofing tool rather than marking them up on the PDF. This will ensure that your corrections are incorporated accurately and that your paper is published as quickly as possible.

| Query Reference | Reference |
|---|---|
| Q1 | A single sentence summarizing your paper has been provided (editor's summary in the eproof), which will appear online on the table of contents and in e-alerts. Please check this sentence for accuracy and appropriate emphasis. |
| Q2 | Please check your article carefully, coordinate with any co-authors and enter all final edits clearly in the eproof, remembering to save frequently. Once corrections are submitted, we cannot routinely make further changes to the article. |
| Q3 | Note that the eproof should be amended in only one browser window at any one time; otherwise changes will be overwritten. |
| Q4 | Author surnames have been highlighted. Please check these carefully and adjust if the first name or surname is marked up incorrectly. Note that changes here will affect indexing of your article in public repositories such as PubMed. Also, carefully check the spelling and numbering of all author names and affiliations, and the corresponding email address(es). |
| Q5 | You cannot alter accepted Supplementary Information files except for critical changes to scientific content. If you do resupply any files, please also provide a brief (but complete) list of changes. If these are not considered scientific changes, any altered Supplementary files will not be used, only the originally accepted version will be published. |
| Q6 | If applicable, please ensure that any accession codes and datasets whose DOIs or other identifiers are mentioned in the paper are scheduled for public release as soon as possible, we recommend within a few days of submitting your proof, and update the database record with publication details from this article once available. |
| | |

# Author Queries

Journal: **Nature**
Paper: **s41586-022-05650-9**
Title:**A *Prox1* enhancer represses haematopoiesis in the lymphatic vasculature**

**AUTHOR:**
The following queries have arisen during the editing of your manuscript. Please answer by making the requisite corrections directly in the e.proofing tool rather than marking them up on the PDF. This will ensure that your corrections are incorporated accurately and that your paper is published as quickly as possible.

| Query Reference | Reference |
|---|---|
| Q7 | Please provide the states for all Australian affiliations and the state codes for US affiliations to comply with style. |
| Q8 | Your paper has been copy edited. (1) Please review every sentence to ensure that it conveys your intended meaning; if changes are required, please provide further clarification rather than reverting to the original text. Please note that formatting (including hyphenation, Latin words, and any reference citations that might be mistaken for exponents) and usage have been made consistent with our house style.(2) Check the title and the first paragraph with care, as they may have been re-written to aid accessibility for non-specialist readers.(3) Check the symbols for affiliations with care, and check all author names and Acknowledgements carefully to ensure that they are correct; check the email address of the corresponding author and the Competing Interests statement.(4) Check that there has been no corruption of mathematical symbols. Only single-letter variables are set in italics (but not their subscripts unless these are also variables); multi-letter variables are set in roman. Vectors are set as bold; matrices are set as italic only. We do not use italics for emphasis. Genetic material is set in italic and gene products are set upright. Please check that italicization and bolding are correct throughout. (5) Ensure that, where practicable, all figures, tables and other discrete elements of Supplementary Information are referred to at least once in the paper at an appropriate place in the text or figure legends.(6) Please note, we reserve 'significant' and its derivatives for statistical significance. Please reword where this is not the intended meaning (for example, to important, notable, substantial). |
| Q9 | Please check that the display items are as follows (ms no: 2019-04-05000D): Figs 0(black & white); 4 (colour); Tables: None; Boxes: None; Extended Data display items: 10 figures; SI: yes. The eproof contains the main-text figures edited by us and (if present) the Extended Data items (unedited except for minor formatting) and the Supplementary Information (unedited). Please check the edits to all main-text figures (and tables, if any) very carefully, and ensure that any error bars in the figures are defined in the figure legends. Extended Data items may be revised only if there are errors in the original submissions. If you need to revise any Extended Data items please upload these files when you submit your corrections to this eproof, and include a list of what has been changed. |

# Author Queries

Journal: **Nature**
Paper: **s41586-022-05650-9**
Title: **A Prox1 enhancer represses haematopoiesis in the lymphatic vasculature**

**AUTHOR:**
The following queries have arisen during the editing of your manuscript. Please answer by making the requisite corrections directly in the e.proofing tool rather than marking them up on the PDF. This will ensure that your corrections are incorporated accurately and that your paper is published as quickly as possible.

| Query Reference | Reference |
|---|---|
| Q10 | Please ensure that genes are correctly distinguished from gene products: for genes, official gene symbols (e.g., NCBI Gene) for the relevant species should be used and italicized; gene products such as proteins and noncoding RNAs should not be italicized. |
| Q11 | Please note that we reserve 'deficient' for protein; therefore, in the sentence beginning "A key characteristic of *Gata2*-deficient valve endothelial cells...", please clarify whether your intended meaning is 'null', 'knockout' or *Gata2*$^{-/-}$'. Please check and change where necessary throughout. |
| Q12 | We avoid the use of claims of priority as manuscripts of similar complexity or in different language may be published around the same time; therefore 'the first' was removed throughout. |
| Q13 | (1) Please ensure that the following information is included in the figure legends where relevant. Sample size (exact n number); a statement of replicability (how many times was experiment replicated in the lab); description of sample collection (clarify whether technical or biological replicates and include how many animals, litters, cultures, etc.); state the statistical test used and give P values; define centre values (median or average) and error bars. (2) For figures/images that are reproduced or adapted from a third party, it is important that you confirm that permission has been obtained and that appropriate acknowledgement of the copyright holder is given. (3) Please note that we edit the main figures (but not the Extended Data figures) in house. There is no need to resupply any of the main figures to make minor changes to text labels to match the changes made in the text, as figures will have been edited accordingly. If you wish to make changes to any Extended Data figures, however, please resupply these, and please let us know what has changed. |
| Q14 | Please check the panels and captions of all figures carefully, as these have been edited for style and clarity. Please make sure that any abbreviations listed only in the figure are defined in the caption. |
| Q15 | Please clarify what the red lines indicate in Fig. 2c. |

# Author Queries

Journal: **Nature**
Paper: **s41586-022-05650-9**
Title:**A *Prox1* enhancer represses haematopoiesis in the lymphatic vasculature**

**AUTHOR:**
The following queries have arisen during the editing of your manuscript. Please answer by making the requisite corrections directly in the e.proofing tool rather than marking them up on the PDF. This will ensure that your corrections are incorporated accurately and that your paper is published as quickly as possible.

| Query Reference | Reference |
|---|---|
| Q16 | Scale bars are missing from Fig. 2a,e. If scale bars are necessary, please resupply the affected panels by attaching to the eproof (paperclip symbol in the top right corner). |
| Q17 | For all *t*-tests, please indicate whether a Student's *t*-test or non-parametric equivalent was used. |
| Q18 | Gene nomenclature correct in the sentence beginning "Chromosome conformation capture analysis confirmed that..."? If not, please amend. |
| Q19 | Please confirm that the definition of 'E' at the end of Fig. 4 caption is correct. If not, please amend. |
| Q20 | (1) Please ensure that the following information is provided in the Methods section where relevant. Animal experiments require: statement about randomization; statement about blinding; statement of sex, age, species and strain of animals; statement of IRB approval for live vertebrate experimentation. For experiments involving humans: statement of IRB approval; statement of informed consent; statement of consent to publish any photos included in figures. Randomized clinical trials require trial registration. (2) We recommend that detailed protocols are deposited in Protocol Exchange, or a similar repository. (3) If custom computer code has been used and is central to the conclusions of this paper, please insert a section into the Methods titled 'Code availability' and indicate within this section whether and how the code can be accessed, including any restrictions to access. (4) If unpublished data are used, please obtain permission. (5) Please state whether statistical methods were used to predetermine sample size. (6) Please state whether blinding and randomization were used. (7) To address the issue of cell line misidentification and cross-contamination, for any cell lines mentioned in the paper please provide source of the cell lines and indicate whether the cell lines have been correctly identified/authenticated (if so, by what methods). Also, please state whether cell lines have been tested for mycoplasma contamination. |
| | |

# Author Queries

Journal: **Nature**
Paper: **s41586-022-05650-9**
Title: **A Prox1 enhancer represses haematopoiesis in the lymphatic vasculature**

**AUTHOR:**
The following queries have arisen during the editing of your manuscript. Please answer by making the requisite corrections directly in the e.proofing tool rather than marking them up on the PDF. This will ensure that your corrections are incorporated accurately and that your paper is published as quickly as possible.

| Query Reference | Reference |
|---|---|
| Q21 | In the sentence beginning "Care was taken to eliminate the livers...", please clarify what '/' indicates in 'DMEM/20% FCS'. Please also check and amend throughout. |
| Q22 | In the sentence beginning "Fresh media containing murine...", please change 'murine' to 'mouse' (or rat) when only mice are meant, as murine indicates both mice and rats, unless murine is meant. |
| Q23 | In the sentence beginning "Differential expression was evaluated from TMM...", please define TMM. |
| Q24 | In the sentence beginning "In brief, the data were normalized and filtered...", please define MDS. |
| Q25 | (1) The acknowledgements have been edited for style and brevity; note, we do not include titles or affiliations and only include the initial(s) of the first name(s) of colleagues mentioned here. Please check. (2) Please check that all funders have been appropriately acknowledged and that all grant numbers are correct. |
| | |

nature portfolio

Corresponding author(s): Natasha Harvey

Last updated by author(s): Nov 21, 2022

# Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☒ | ☐ | A description of all covariates tested |
| ☒ | ☐ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | Enhancer identification/mutation:<br>Mouse - CRISPR gRNA designed online (Zhang laboratory MIT, http://crispr.mit.edu)<br>Zebrafish - mVISTA (https://genome.lbl.gov/vista/index.shtml)<br><br>Flow cytometry:<br>MoFlo Astrios: Summit Software version 6.2.4.15830 (Beckman Coulter, Miami, FL, USA)<br>Becton Dickinson LSR Fortessa: FACS Diva Software version 8.0.3 (BD Biosciences, San Diego, CA, USA) |
| Data analysis | Microscopy:<br>ZEN lite 2011 (blue edition) version 1.0 (Carl Zeiss)<br>ZEN 2.5 (blue edition; Zeiss)<br>Adobe Photoshop CC version 19.0 (Adobe)<br>ImageJ 2.0.0<br><br>3C gel quantification: ImageQuant TL 1D version 8.1 (GE Healthcare)<br><br>Flow cytometry: FlowJo version 10.7.1 (Becton Dickinson)<br><br>RNA sequencing bioinformatic analysis:<br>FastQC, http://www.bioinformatics.babraham. ac.uk/projects/fastqc (raw data analysis and quality checking)<br>STAR spliced alignment algorithm, version 2.5.3a, with default parameters and --chimSegmentMin 20, --quantMode GeneCounts (mapping of reads against the mouse reference genome mm10)<br>Integrative Genomics Viewer v2.8.9 (visualization and interrogation of alignments)<br>Sambamba v0.6.7; markdup function with settings: --remove-duplicates, --nthreads 16, --overflow-list-size 600000 (removal of duplicate |

reads)
HTSeq v 0.11.2; htseq-count function with settings: --format bam --order pos --stranded yes --minaqual 10 (tabulating of counts)
R, version 4.1.1
edgeR, version 3.34.0 (differential expression analysis)
gplots package in R; heatmap.2 function

RNA microarrays bioinformatic analysis:
Partek Genomics Suite™ version 6.4 software (Partek Incorporated, St. Louis, MO)
oligo (v 1.60.0)

Gene set enrichment analysis (UC San Diego, Broad Insititute) https://www.gsea-msigdb.org

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

GATA2 ChIP-Seq data has been deposited in the European Nucleotide Archive (ENA), accession number PRJEB9436 (http://www.ebi.ac.uk/ena/data/view/PRJEB9436).
PROX1/FOXC2/NFATC1 ChIP-Seq and human LEC/BEC RNA-seq data have been submitted to GEO, accession number GSE129634 (https://www.ncbi.nlm.nih.govgeo/query/acc.cgi?acc=GSE129634).
Mouse LEC (pre/post-OP9) RNA-seq and mouse LEC/BEC microarray data have been submitted to GEO, accession number GSE184046 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE184046)
The HE and E RNA-Seq data used for GSEA (Fig4) were generated in the following study (https://doi.org/10.1016/j.exphem.2018.10.009) and were obtained from the GEO database, accession number GSE103813

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | No statistical methods were used to predetermine sample size. Due to animal ethics considerations sample size was determined according to the minimal number of independent biological replicates that significantly identified an effect. For most analyses, 3 sets of biological samples (ie, litters of mice or individual embryos) were assessed. Variability between the three biological replicates was minimal and therefore did not require increasing the sample size.<br>For GATA2 ChIP-seq using cultured human LEC, initially two biological replicates (individual donor cell lines) were sequenced. Peaks called were found to be consistent; thereafter, single biological replicates were used for all other cultured human cell line ChIP-seq/RNA-seq experiments.For 3C experiments using cultured human LEC, one biological replicate (cell line) was analysed and at least three technical replicates were performed. |
| Data exclusions | No data were excluded. |
| Replication | All experiments with the exception of 3C were replicated and were conducted with both biological and technical replicates. Excluding technical malfunctions, all attempts at replication were successful and verified the reproducibility of the findings. All 'n' values are specified in figure legends. |
| Randomization | No experiments were randomized, chiefly due to the requirement for knowing genotypes prior to pooling embryos for cell isolation. Biological variability was controlled for by analyzing multiple embryos or pooling embryos for studies requiring cell isolation. Wild type controls were litter mates or were generated within the same colony to eliminate potential for strain variation. Technical variability was minimized by subjecting all samples to a standardized workflow. |
| Blinding | Investigators were not blinded during data collection and analysis; in most cases, personnel undertaking the experiments included the person responsible for genotyping the mice. Blinding was not relevant for this study since the aim was to quantify discriminating features between already established biological groups. |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☐ | ☒ Antibodies |
| ☐ | ☒ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology and archaeology |
| ☐ | ☒ Animals and other organisms |
| ☒ | ☐ Human research participants |
| ☒ | ☐ Clinical data |
| ☒ | ☐ Dual use research of concern |

## Methods

| n/a | Involved in the study |
|---|---|
| ☐ | ☒ ChIP-seq |
| ☐ | ☒ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

## Antibodies

Antibodies used

(Lot numbers listed where known/relevant)

Flow cytometry:
CD144-BV421, clone 11D4.1, BD Biosciences, Cat# 747749, Lots 0254558 & 2133582, 1 in 100
LYVE1 PE, clone 223322, R&D Systems, Cat# FAB2125P, Lot ACFE0220031, 1 in 100
CD45-APC-Cy7, clone 30-F11, BD Biosciences, Cat# 557659, Lot 1085935, 1 in 100
CD117-BUV395, clone 2B8, BD Biosciences, Cat# 564011, Lot 0337172, 1 in 100
CD11b-APC, clone M1/70, BioLegend Cat# 101211, 1 in 100, Lot B226978
CD71-BV510, clone C2, BD Biosciences, Cat# 563112, Lot 2032056, 1 in 100
CD41-BUV737, clone MWReg30, BD Biosciences, Cat# 741759, Lot 1019895, 1 in 100
Ly-6G-PECy7, clone 1A8, BD Biosciences, Cat# 560601, 1 in 100
CD16/32 PE, clone 93, BioLegend, Cat# 101307, 1 in 100

Immunostaining:
GATA2, rabbit polyclonal, Novus, Cat# NBP1-82581, Lots C76352 and 000035473, 1 in 500
PROX1, rabbit polyclonal, Abcam, Cat# ab101851, 1 in 1000
LYVE1, rabbit polyclonal, AngioBio, Cat# 11-034,1 in 1000
PROX1, goat polyclonal, R&D Systems, Cat# AF2727, Lot VIY0420091, 1 in 250
CD31, clone MEC 13.3, BD Biosciences, Cat# 553370, Lot 2335717, 1 in 500
CD34, clone RAM34, eBioscience, Cat# 14-0341, Lot E019241, 1 in 250
CD117/cKit, clone 2B8, eBioscience, Cat# 14-1171, 1 in 250
ESAM, goat polyclonal, R&D Systems, Cat# AF2827, 1 in 250
Endomucin, clone V.7C7, Santa Cruz, Cat# sc-65495, Lot H0819, 1 in 500
CD144/VECAD, goat polyclonal, R&D Systems, Cat# AF1002, Lot FQI0120041, 1 in 250
α smooth muscle actin-Cy3, clone 1A4, Sigma, Cat# C6198, Lot 059M4797V, 1 in 1000
FOXC2 (a gift from N.Miura) 1 in 1000
VEGFR3, goat polyclonal, R&D Systems, Cat# AF743, 1 in 250
RUNX1, clone EPR3099, Abcam, Cat# ab92336, 1 in 1000
β-galactosidase, rabbit polyclonal, MP Biomedicals, Cat# 55976, 1 in 5000
β-galactosidase, chicken polyclonal, Abcam, Cat# ab9361, 1 in 1000
GFP, rabbit polyclonal, Thermo Fisher Scientific, Cat# A-11122, Lot# 2083201, 1 in 500

Alexa Fluor™-conjugated antibodies used for detection (all used 1 in 500):
Donkey anti-Rat IgG (H+L) Alexa Fluor™ 488, Thermo Fisher Scientific, Cat# A-21208
Donkey anti-Goat IgG (H+L) Alexa Fluor™ 488, Thermo Fisher Scientific, Cat# A-11055
Donkey anti-Syrian Hamster IgG (H+L) Alexa Fluor™ 488, Thermo Fisher Scientific, Cat# A-21110
Donkey anti-Rabbit IgG (H+L) Alexa Fluor™ 488, Thermo Fisher Scientific, Cat# A-21206
Donkey anti-Rabbit IgG (H+L) Alexa Fluor™ 555, Thermo Fisher Scientific, Cat# A-31572
Donkey anti-Goat IgG (H+L) Alexa Fluor™ 555, Thermo Fisher Scientific, Cat# A-21432
Donkey anti-Chicken IgG (H+L) Alexa Fluor™ 594, Jackson ImmunoResearch, Cat# 703-585-155
Donkey anti-Chicken IgG (H+L) Alexa Fluor™ 647, Jackson ImmunoResearch, Cat# 703-606-155
Donkey anti-Goat IgG (H+L) Alexa Fluor™ 647, Thermo Fisher Scientific, Cat# A-32879
Donkey anti-Rabbit IgG (H+L) Alexa Fluor™ 647, Thermo Fisher Scientific, Cat# A-31573
Chicken anti-Rat IgG (H+L) Alexa Fluor™ 647, Thermo Fisher Scientific, Cat# A-31573

ChIP:
GATA2, rabbit polyclonal, Santa Cruz, Cat# sc9008X (discontinued/no longer available)
FOXC2, goat polyclonal, Abcam, Cat# ab5060, Lot H0912
NFATC1, rabbit polyclonal, Santa Cruz, Cat# sc13033X (discontinued/no longer available)
PROX1, goat polyclonal, R&D Systems, Cat# AF2727
Rabbit IgG, Cell Signaling, Cat# 2729

Cell isolation:
F4/80, clone BM8, ThermoFisher, Cat# 14-4801-82, 1 in 100

Validation

All primary antibodies obtained from the indicated commercial vendors were validated for the application by the manufacturer. Titration experiments were performed prior to the study to confirm optimal dilutions. Antibodies for immunostaining have been used extensively by multiple laboratories, or were further validated in cells from GATA2 KO mice (Kazenwadel et al, JCI, 2015) or in cells treated with appropriate siRNAs. Rat anti-FOXC2 has been validated (Furimoto et al, Dev.Biol. 1999). Antibodies used for Chip analysis were validated by confirming enrichment at predicted binding sites (Kazenwadel et al, JCI, 2015).
Validation statements from manufacturers websites can be found using the following links:

CD144-BV421, https://www.bdbiosciences.com/en-au/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/purified-rat-anti-mouse-cd144.550548
LYVE1 PE, https://www.rndsystems.com/products/mouse-lyve-1-pe-conjugated-antibody-223322_fab2125p
CD45-APC-Cy7, https://www.bdbiosciences.com/en-au/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/apc-cy-7-rat-anti-mouse-cd45.557659
CD117-BUV395, https://www.bdbiosciences.com/en-au/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/buv395-rat-anti-mouse-cd117.564011
CD11b-APC, https://www.biolegend.com/it-it/products/apc-anti-mouse-human-cd11b-antibody-345
CD71-BV510, https://www.bdbiosciences.com/en-au/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/bv510-rat-anti-mouse-cd71.563112
CD41-BUV737, https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/buv737-rat-anti-mouse-cd41.741759/
Ly-6G-PECy7, https://www.bdbiosciences.com/en-au/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/pe-cy-7-rat-anti-mouse-ly-6g.560601
CD16/32 PE, https://www.biolegend.com/en-us/products/pe-anti-mouse-cd16-32-antibody-189
GATA2, rabbit polyclonal, https://www.novusbio.com/products/gata-2-antibody_nbp1-82581
PROX1, rabbit polyclonal, https://www.abcam.com/prox1-antibody-bsa-and-azide-free-ab101851.html
LYVE1, rabbit polyclonal, http://www.angiobio.com/new/product.php?pid=8
PROX1, goat polyclonal, https://www.rndsystems.com/products/human-prox1-antibody_af2727
CD31, clone MEC 13.3, https://www.bdbiosciences.com/en-au/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/purified-rat-anti-mouse-cd31.553370
CD34, clone RAM34, https://www.citeab.com/antibodies/2038656-14-0341-82-cd34-monoclonal-antibody-ram34-ebiosci
CD117/cKit, clone 2B8, https://www.thermofisher.com/antibody/product/CD117-c-Kit-Antibody-clone-2B8-Monoclonal/14-1171-82
ESAM, goat polyclonal, https://www.rndsystems.com/products/mouse-esam-antibody_af2827
Endomucin, clone V.7C7, https://www.scbt.com/p/endomucin-antibody-v-7c7
CD144/VECAD, goat polyclonal, https://www.rndsystems.com/products/mouse-ve-cadherin-antibody_af1002
α smooth muscle actin-Cy3, clone 1A4, https://www.sigmaaldrich.com/AU/en/product/sigma/c6198
VEGFR3, goat polyclonal, https://www.rndsystems.com/products/mouse-vegfr3-flt-4-antibody_af743
RUNX1, clone EPR3099, https://www.abcam.com/runx1--aml1--runx3--runx2-antibody-epr3099-ab92336.html
β-galactosidase, rabbit polyclonal, https://www.mpbio.com/au/rabbit-igg-fraction-to-beta-galactosidase
β-galactosidase, chicken polyclonal, https://www.abcam.com/beta-galactosidase-antibody-ab9361.html
GFP, rabbit polyclonal, https://www.thermofisher.com/antibody/product/GFP-Antibody-Polyclonal/A-11122
FOXC2, goat polyclonal, https://www.abcam.com/foxc2-antibody-ab5060.html
F4/80, clone BM8, https://www.thermofisher.com/antibody/product/F4-80-Antibody-clone-BM8-Monoclonal/14-4801-82

## Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

Adult human dermal lymphatic microvascular endothelial cells (hLEC):
HMVEC-dLyAd-Der Lym Endo, Lonza Cat# CC-2810, Lots# 7F3304 and 0000254463

Adult human dermal blood microvascular endothelial cells (hBEC):
HMVEC-dBlAd, Lonza Cat# CC-2811, Lot# 0000125028

Authentication

Primary cells isolated from mice were authenticated as previously described (Kazenwadel et al, Blood, 2010) and used without passaging.
Primary human cells were authenticated by the supplier (Lonza) and confirmed by immunostaining and qPCR analysis of appropriate markers of endothelial cell identity. Cell were used within 4 passages.

Mycoplasma contamination

Cells were not screened for mycoplasma.

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified cell lines were used.

## Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

Mice used in this study were provided with water and standard chow ad libitum, and housed in a pathogen free facility under the following conditions: 12/12 dark/light cycle that includes 30min Dusk and Dawn cycles that run from 6.30-7.00am/pm, 20.5-23.5 degrees Celsius, humidity between 50-60%.

| | |
|---|---|
| | Stains used in this study:<br>Prox1enh-hsp-LacZ transgenic mice (generated for this study), C57BL/6J background, male and female. Embryonic stages E11.5, E12.5, E14.5, E11.5, E17.5, E18.5. Postnatal stage P4<br><br>Prox1enh-CRISPR mice (generated for this study), C57BL/6J background, male and female. Embryonic stages E10.5, E14.5, E17.5, E18.5. Postnatal stage P0<br><br>Prox1+/GFPCre mice (Srinivasan et al, 2010 doi:10.1101/gad.1859310), C57BL/6 background, male and female, embryonic day E14.5<br><br>C57BL/6J mice (used for microarray analysis ExtDataFigure 10b), male and female. Embryonic stages E14.5, E16.5, E18.5<br>C57BL/6N mice (embryos used for pronuclear injection to establish LacZ transgenic line, see Methods) |
| Wild animals | This study did not involve wild animals. |
| Field-collected samples | The study did not involve samples collected from the field. |
| Ethics oversight | All experiments using mice were approved by the University of Adelaide, University of South Australia, SA Pathology/CALHN or Northwestern University Animal Ethics Committees and conducted in accordance with the Australian code for the care and use of animals for scientific purposes.<br>Zebrafish work was carried out under ethical approval from the Swedish Board of Agriculture (5.2.18-7558/14). |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

# ChIP-seq

## Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

| | |
|---|---|
| Data access links<br>*May remain private before publication.* | GATA2 ChIP-Seq data has been deposited in the European Nucleotide Archive (ENA), accession number PRJEB9436 (http://www.ebi.ac.uk/ena/data/view/PRJEB9436).<br>PROX1, FOXC2 and NFATC1 ChIP-Seq data have been submitted to GEO, accession number GSE129634 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE129634). |
| Files in database submission | Files submitted to GEO:<br>Prox1_R1.bwa.hg19.macs1.4_peaks.xls,<br>NfatCL_R1.bwa.hg19.macs1.4_peaks.xls,<br>FoxC2_R1.bwa.hg19.macs1.4_peaks.xls,<br>Prox1_R1.fastq.gz,<br>NfatCL_R1.fastq.gz,<br>FoxC2_R1.fastq.gz,<br>Input35_R1.fastq.gz,<br>Input5_R1.fastq.gz,<br>BEC_R1.trimmed_NEB_PE.tophat2_pe.hg19.htseq-count_gene_id_rev_stranded.tsv,<br>LEC_R1.trimmed_NEB_PE.tophat2_pe.hg19.htseq-count_gene_id_rev_stranded.tsv,<br>BEC_R1.fastq.gz, BEC_R2.fastq.gz,<br>LEC_R1.fastq.gz, LEC_R2.fastq.gz |
| Genome browser session<br>(e.g. UCSC) | No longer applicable |

## Methodology

| | |
|---|---|
| Replicates | None; samples were processed using a standardized work flow and in parallel where possible. |
| Sequencing depth | Read depth is >30million/sample |
| Antibodies | Rabbit anti-Gata2 (Santa Cruz, sc9008X), goat anti-Foxc2 (Abcam, ab5060), rabbit anti-Nfatc1 (Santa Cruz, sc13033X), goat anti-human Prox1 (AF2727; R&D Systems). |
| Peak calling parameters | Peaks were called using MACS v1.4. Samples were run with controls ( input5 for NfatC1 and FoxC2; Input3.5 for ProX1). Command line and Parameters:<br>macs14 callpeaks --nomodel --shift 0 --gsize hs --format BAM --qvalue 0.05 --treatment $TREATMENT --control $CONTROL |
| Data quality | See Methods ChIP-seq analysis. |
| Software | ChIPseq data was mapped to the hg19 reference using BWA aligner v0.7.9a-r786 allowing at most 3 alignments, and ChIPseq Peak calling was performed using MACS v1.4. |

# Flow Cytometry

## Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

## Methodology

| | |
|---|---|
| Sample preparation | For FACS sorting, litters consisting of 6-8 pooled embryos of a single genotype (wildtype or homozygous mutant) were used for each isolation. At E14.5 the dorso-anterior regions of embryos were dissected at room temperature in HHF (5% FCS, 10mM Hepes in Hanks Balanced Salt Solution and rinsed briefly with DMEM/20%FCS. Tissue was digested in 10ml DMEM/20%FCS containing 25 mg Collagenase Type II, 25 mg Collagenase Type IV and 10 mg Deoxyribonuclease I (Worthington) for 30 minutes at 37 degrees C while mixing gently with a wide-bore transfer pipette every 5 min to assist tissue dissociation. Cell suspensions were filtered through a 40 mm cell strainer. Filtrates were centrifuged at 200 g for 10 min and resuspended in 5 ml HHF at room temperature. Cells were counted (generally approximately 5-10 x 106/embryo) and centrifuged for a further 5 minutes at 300g. The resulting pellet was resuspended in 1 ml HHF containing 1:100 dilution F4/80 monoclonal antibody, incubated at room temperature for 5 minutes and F4/80 positive cells were depleted using anti-rat MACS beads. Following F4/80 MACS depletion, the cells were lineage depleted using biotinylated lineage antibodies and Biotin Binder Dynabeads. Lineage depleted cells were resuspended in sort buffer (2% FBS, 5 µM EDTA, 25 mM HEPES pH7, 10 U/mL DNAse I in phenol red-free HBSS) and incubated for 10 min at room temperature prior to addition of fluorochrome-conjugated monoclonal antibodies. Cells were incubated with antibodies for 20 min at room temperature, washed with 3 ml of sort buffer and resuspended in sort buffer with SYTOX Red Dead Cell Stain. |
| Instrument | For sorting: Beckman Coulter MoFlo Astrios EQ High Speed Cell Sorter, equipped with 355 nm (100mW), 405 nm (55 mW), 488 nm (150 mW), 561 nm (200 mW) and 633 nm (100 mW) lasers, enclosed within Baker SterilGuard BSL Class II Biosafety cabinet (The Baker Company, Sanford, Maine, USA).<br>For flow cytometry analysis: Becton Dickinson LSR Fortessa Special Order Research Product, equipped with 355 nm (20 mW), 405 nm (50 mW), 488 nm (50 mW), 561 nm (50 mW) and 633 nm (40 mW) lasers. |
| Software | For sorting on MoFlo Astrios: Summit Software version 6.2.4.15830 (Beckman Coulter, Miami, FL, USA).<br>For data collection using Becton Dickinson LSR Fortessa: FACS Diva Software version 8.0.3 (BD Biosciences, San Diego, CA, USA).<br>For analysis of collected data: FlowJo version 10.7.1 (Becton Dickinson). |
| Cell population abundance | To maximize the numbers of cells available for downstream experiments and reduce the impact on cell viability, purity of sorted samples was not routinely assessed. |
| Gating strategy | For cell isolation for RNA-seq and colony forming assays samples were first gated on a 2D FSC-Area vs SSC-Area plot to exclude debris. Viable cells were then selected as a Sytox Red Dead Cell Marker-negative population on a 2D Sytox Red-Area vs SSC-Area plot. Single cells were selected by plotting FSC-Area vs FSC-Height. CD45-negative cells were then selected by plotting CD45-APC-Cy7-Area vs SSC-Area. VECAD-, Lyve1-double positive population was then selected on a 2D dot plot (VECAD BV421-Area vs Lyve1 PE-Area).<br>After OP9 culture, samples were first gated on a 2D FSC-Area vs SSC-Area plot to exclude debris and single cells were then selected by plotting FSC-Area vs FSC-Height. For RNA-Seq, VECAD-, Lyve1-double positive population was then selected on a 2D dot plot (VECAD BV421-Area vs Lyve1 PE-Area). For methylcellulose colony assay, CD45-positive cells were selected out of single cell population by plotting CD45-APC-Cy7-Area vs SSC-Area.<br>For analysis of colonies harvested from methylcellulose, samples were first gated on a 2D FSC-Area vs SSC-Area plot to exclude debris and single cells were then selected by plotting FSC-Area vs FSC-Height and subsequently SSC-Area vs SSC-Width. CD45-positive cells were then selected by plotting CD45-APC-Cy7-Area vs FSC-Area. CD45-positive cells were then analyzed for VECAD (CD144)-BV421, CD45 APC-Cy7, cKIT(CD117)-BUV395, CD11b -APC, CD71-BV510, CD41-BUV737, Ly-6G-PECy7 and CD16/32-PE.<br>The boundaries of positive and negative gates were established by comparison to unstained samples, single stained samples and, for analysis of CD45-positive cells harvested from methylcellulose, by comparison to CD45-negative populations present within the same samples. |

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.