LETTER

Haematopoietic stem cell induction by somite–derived endothelial cells controlled by *meox1*

Phong Dang Nguyen[1]*, Georgina Elizabeth Hollway[2,3]*, Carmen Sonntag[1], Lee Barry Miles[1], Thomas Edward Hall[1], Silke Berger[1], Kristine Joy Fernandez[2], David Baruch Gurevich[1], Nicholas James Cole[4,5], Sara Alaei[1,6], Mirana Ramialison[1], Robert Lyndsay Sutherland[2,3]‡, Jose Maria Polo[1,6], Graham John Lieschke[1] & Peter David Currie[1,7]*

Haematopoietic stem cells (HSCs) are self-renewing stem cells capable of replenishing all blood lineages. In all vertebrate embryos that have been studied, definitive HSCs are generated initially within the dorsal aorta (DA) of the embryonic vasculature by a series of poorly understood inductive events[1–3]. Previous studies have identified that signalling relayed from adjacent somites coordinates HSC induction, but the nature of this signal has remained elusive[4]. Here we reveal that somite specification of HSCs occurs via the deployment of a specific endothelial precursor population, which arises within a sub-compartment of the zebrafish somite that we have defined as the endotome. Endothelial cells of the endotome are specified within the nascent somite by the activity of the homeobox gene *meox1*. Specified endotomal cells consequently migrate and colonize the DA, where they induce HSC formation through the deployment of chemokine signalling activated in these cells during endotome formation. Loss of *meox1* activity expands the endotome at the expense of a second somitic cell type, the muscle precursors of the dermomyotomal equivalent in zebrafish, the external cell layer. The resulting increase in endotome-derived cells that migrate to colonize the DA generates a dramatic increase in chemokine-dependent HSC induction. This study reveals the molecular basis for a novel somite lineage restriction mechanism and defines a new paradigm in induction of definitive HSCs.

Cell lineage studies have demonstrated two cellular sources for the amniote DA, the splanchnic and somitic mesoderm[5–12]. The zebrafish DA, by contrast, is believed to originate *in toto* from cells of the lateral plate mesoderm, an assumption based on imaging using lateral plate mesoderm-specific transgenes[13–18]. Consequently, these analyses do not exclude a somitic origin for vasculature progenitors in zebrafish. Our analyses of the *choker* (*cho*) mutation[19] revealed defects in somite lineages. Secondary trunk myogenesis is reduced in *cho* mutants (Fig. 1a–h, k, l, Extended Data Fig. 1i–l, o), as are appendicular and hypaxial muscles and their progenitors (Fig. 1i–n, Extended Data Fig. 1a–h, m, n, Supplementary Videos 1 and 2). These cell types derive from the external cell layer (ECL)[20,21] and ECL cell numbers are reduced in *cho* homozygotes (Fig. 1o–s). The *cho* mutation results from a null mutation in the *meox1* gene (Extended Data Fig. 2a–h) and *meox1* is expressed within the early somite, becomes restricted to the ECL and consequently to appendicular muscle populations, which are disrupted in *cho* mutants (Fig. 1t–z, Extended Data Fig. 2i–n). *meox1* is also expressed in vascular-associated cells (VACs, arrowheads Fig. 1 x′′, y, z, Extended Data Fig. 2m) located adjacent to the DA.

DA marker expression was variably expanded in *meox1* mutants (Fig. 2a–j, Extended Data Fig. 3a–j). The most informative of these markers was *cxcl12b* (also known as *sdf1b*), which is localized initially to a central region of newly formed somites and then in cells positioned in a salt and pepper manner within the DA[22] (Fig. 2a–j). Somite and DA-associated expression of *cxcl12b* is expanded in *meox1* mutants (n = 57) with the global DA marker genes *notch3* (n = 16) and *ephB2* (n = 13)[23] more modestly expanded (Extended Data Fig. 3i, j, and data not shown). Furthermore, *pax3a* GFP transgenic embryos, which express green fluorescent protein (GFP) within the early somite, exhibit perduring GFP within the ECL and endothelial cells (ECs) of the DA (Fig. 2k–p, Extended Data Fig. 3m–o′, v–x′), with the majority of the *pax3a*-GFP-positive cells expressing *cxcl12b* (Fig. 2n–n′, arrows). Furthermore, anti-Meox1 antibody staining reveals DA cells contain perduring Meox1 protein in a similar pattern to *cxcl12b*, despite the DA not being a site of *meox1* mRNA expression (Fig. 2q–q′′, Extended Data Fig. 3k–l′′′). Collectively, these studies suggest a somitic origin for a portion of cells of the zebrafish DA.

To test this hypothesis, somitic cell fate was analysed using the Kaede photoconvertible protein expressed globally (n = 23) or from the somite-specific mesogenin (*msgn*) promoter[24] (n = 29, Fig. 3a–f′′, Supplementary Videos 3 and 4, Extended Data Fig. 3p–u, Extended data Fig. 4). This analysis revealed that the anterior somite gives rise to the ECL and its derivatives as previously documented[20,21], but also contributes to the DA, posterior cardinal vein (PCV) and inter-segmental blood vessels (ISVs), (Fig. 3a–d′′, Extended data Fig. 3p–u, Supplementary Video 5, Extended Data Fig. 4b–h) and VACs (Fig. 3d′′, arrows). Single-cell lineage analysis of the entire somite (Fig. 3h–i′′′, o, n = 96) revealed that cells residing at the anterior/posterior midpoint of the somite provide progenitors to both the vasculature (Fig. 3o) and VACs (Fig. 3i–i′′′). Furthermore, a single-cell Cre/loxP fate mapping technique was developed that demonstrated that central somitic cells are unipotent progenitors that undergo little cell division in contributing to DA formation (n = 7, Fig. 3j–j′′, n). Next we generated a transgenic line that expressed tamoxifen-inducible Cre from the somite-specific *msgn1* promoter Tg(*msgn1*:CreERT2). Crossing this line to Tg(*βactin2*: loxP AcGFP1-STOP pA loxP mCherry pA; *βactin2* is also known as *actb2*) transgenics resulted in DA ECs expressing mCherry upon tamoxifen addition at early somite stages (n = 5, Fig. 3g–g′′). Thus, four independent fate mapping strategies have demonstrated that DA cells can derive from a central somitic region we term the 'endotome'.

Fate mapping further revealed that *meox1* mutants (74%, n = 14, Fig. 3k, k′, m–m′′, o) but not their wild-type siblings (0%, n = 45) can generate DA or VACs from anterior-most somite cells. Reciprocally, labelling of the anterior-most somite cells in *meox1* mutants resulted in fewer ECL cells and their derivatives (26%, n = 5). However, the generation of primary muscle is unaffected in *meox1* mutants (n = 5, Fig. 3l, l′, o) and loss of *meox1* did not alter the proliferation of somite cells either globally or in a lineage-specific manner (Extended Data Fig. 6),

[1]Australian Regenerative Medicine Institute, Level 1, Building 75, Monash University, Wellington Road, Clayton, Victoria 3800, Australia. [2]The Kinghorn Cancer Centre & Cancer Research Program, Garvan Institute of Medical Research, Victoria Street, Darlinghurst, New South Wales 2010, Australia. [3]St Vincent's Clinical School, Faculty of Medicine, University of New South Wales, Kensington, New South Wales 2052, Australia. [4]Anatomy Department, University of Sydney Medical School Anderson Stuart Building University of Sydney, Camperdown, New South Wales 2006, Australia. [5]MND Research Program, Australian School of Advanced Medicine, Macquarie University, New South Wales 2109, Australia. [6]Department of Anatomy and Developmental Biology, Monash University, Wellington Road, Clayton Victoria 3800, Australia. [7]EMBL Australia, Level 1, Building 75, Monash University, Wellington Road, Clayton, Victoria 3800, Australia.
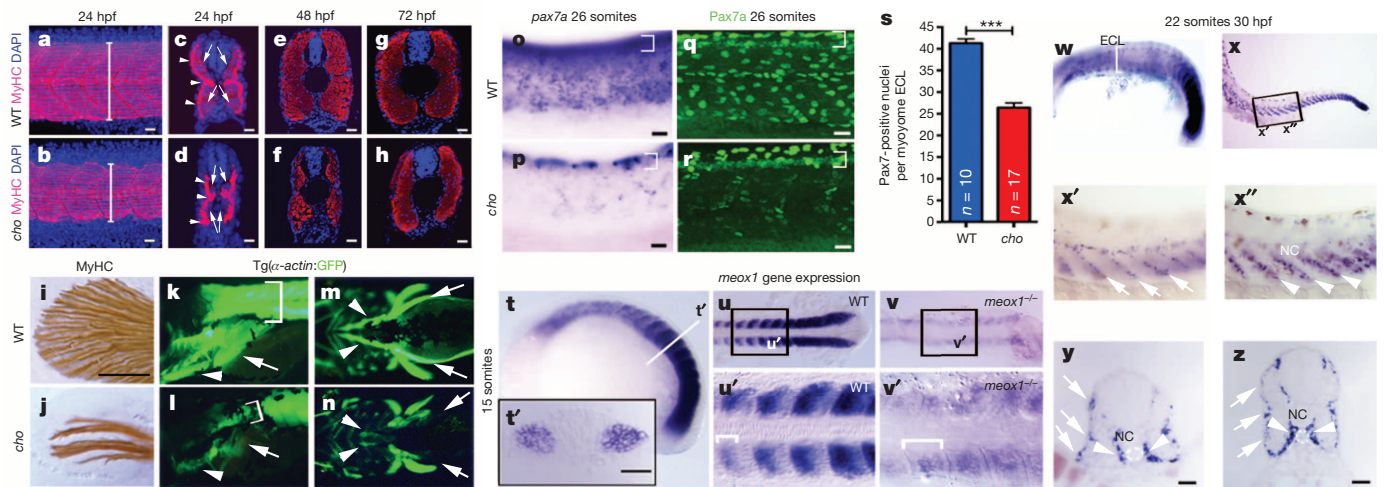*These authors contributed equally to this work.
‡Deceased.

Figure 1 | cho mutants possess defects in the ECL and the muscles derived from it. a–h, Anti-myosin heavy chain (MyHC) staining reveals reduced muscle (brackets, a, b) in cho mutants. Arrowheads (c, d) slow muscle, arrows (c, d) fast muscle. i–n, Reduced trunk (100%, n = 60, brackets) fin (66%, n = 32), hypaxial (65%, n = 60, arrows) and sternohyoids (100% n = 60, arrowheads) muscle in cho mutants. Lateral (k, l) and ventral (m, n) views of 72 hours post fertilisation (hpf) Tg(alpha-actin:GFP) larvae. o–s, pax7a mRNA (o, p) and protein expressing cells (q, r, s) are reduced in number within the ECL but normal in the dorsal neural tube (bracket). s, ECL quantitation: mean ± s.e.m.; significance (***P < 0.0001) in unpaired t-test. t–v. meox1 expression 15 somites. t′, Cross-section as in t. u′, area boxed in u (bracket). v, v′, meox1 transcript levels suggests nonsense mediated decay in cho mutants. v′, region boxed in v. cho mutants lose anterior restriction of remaining meox1 transcripts (compare brackets u′ and v′). w, 22 somites, meox1 is ECL localized. x–z, 30 hpf. meox1 in the ECL, and myosepta of tail somites (arrows, x′, x″). y, z, cross-section within pre-yolk (y) and post–yolk (z) extension somites. Arrows, ECL; arrowheads, VACs; dashed circle, dorsal aorta. Scale bars, 20 μm.

but did result in a transitory increase in the size of the PCV (Extended Data Fig. 5). Thus, meox1 is required to partition the fate of the somite and loss of meox1 expands the endotome at the expense of ECL progenitors. Reciprocally, overexpression of meox1 induces expression of the ECL marker Pax7 in cells destined to be ECs (Extended Data Fig. 7).

Studies in chick embryos have suggested that the colonization of the DA with somite-derived ECs could negatively regulate the haematopoietic competence of the DA, as the haemogenic endothelium of the DA is replaced entirely by somite-derived ECs during embryogenesis[9,10]. We therefore examined whether expression of HSC markers was altered in meox1 mutants. This analysis surprisingly revealed an expansion of HSC marker expression, rather than the reduction predicted by an expanded colonization of haematopoietically refractory, somite-derived ECs in the DA of meox1 mutants.

Specifically, the itga2b:GFP transgenic line (also known as cd41:GFP), which drives low-level GFP expression in DA-associated HSCs, and the genes scl (also known as tal1), cmyb and runx1, which mark HSCs in the DA, all showed expanded expression in meox1 mutants or morphants (scl n = 42, Fig. 4a–h; itga2b:GFP, Extended Data Fig. 8a–i′, v; runx1 n = 26, Extended Data Fig. 8j–o; cmyb n = 32, Extended Data Fig. 8p–u).
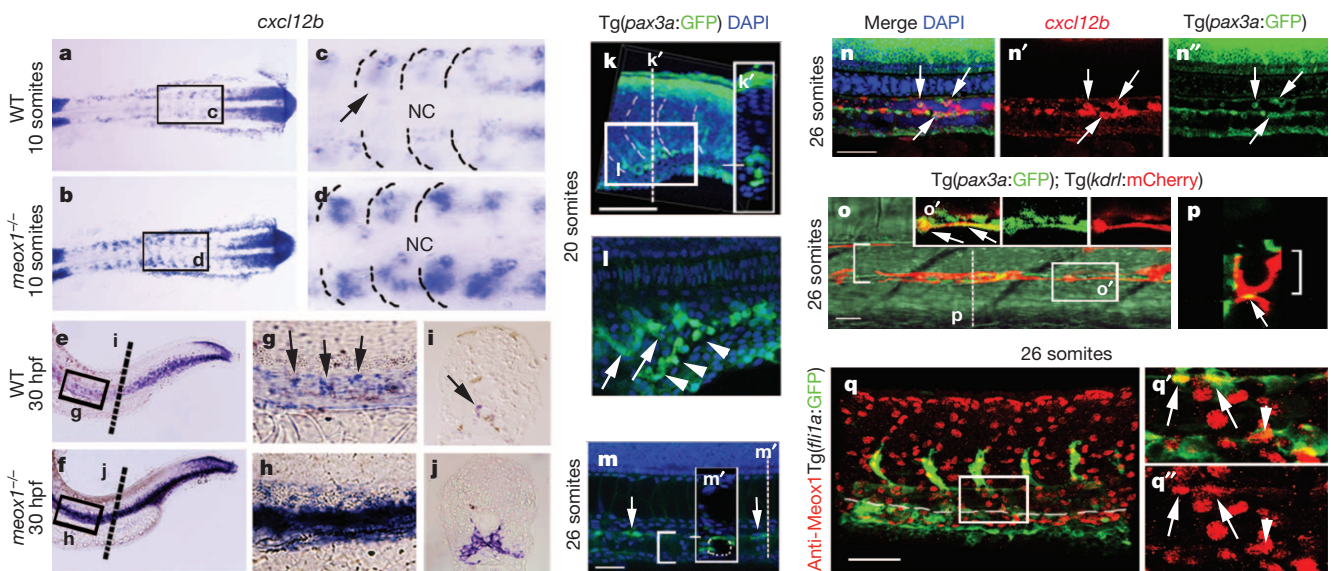

Figure 2 | Endotome markers persist in the DA and are expanded in meox1 mutants. cxcl12b is restricted to the mediolateral somite in wild type (WT) (a, c) but expands into the anterior somite in meox1−/− embryos. (b, d) at 10 somites. Salt and pepper cxcl12b expression in the DA (arrows in g and i) is expanded in meox1−/− embryos (f, h, j) at 30 hpf. a–d, dorsal views; e–h, lateral views; i, j, cross-section. k, l, Tg(pax3a:GFP) (green) marks ventral (arrows, l) and migrating (arrowheads, l) somitic cells. DAPI, blue. k′, Transverse section as in k. l, Confocal section of the region boxed in k. m, 26 somites, GFP in DA cells (arrows). Confocal section at mid-DA. m′, Transverse section as in m. n–n″, Confocal section mid-DA reveals pax3a-GFP cells express cxcl12b (red, arrow). o–p, pax3a-GFP cells co-express Tg(kdrl:mCherry) (arrow, DA brackets). o′, Area boxed in o. p, Transverse section as in o. q–q″, Anti-Meox1 (red) and anti-GFP (green) co-localize in the dorsal (arrows, q′, q″) and ventral (arrowheads, q′–q″) DA. Scale bars, 50 μm; o, 20 μm.

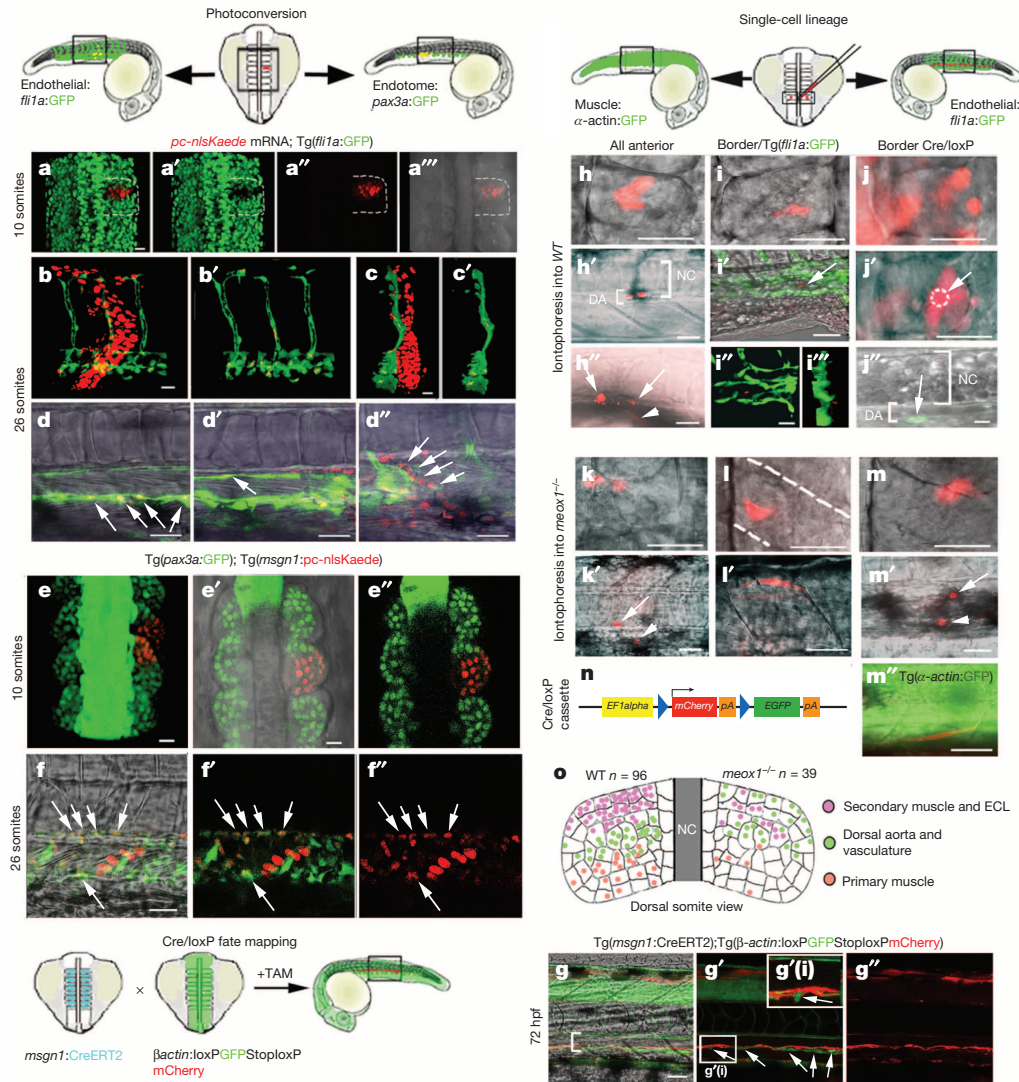**Figure 3 | Endotome cells contribute to DA formation and are expanded in *meox1* mutants.** a–f′′, a–a′′′, nlsKaede (green) photoconverted (red) in the anterior somite (*n* = 23, 10 somites). b–d′′, Same embryo at 26 somites in *fli1a*–GFP. b–b′, *fli1a*–GFP (green) and photoconverted endotomal cells. c, c′, Segmentation for vasculature colocalization (yellow). Scale Bar 20 μm. d–d′′, Endotome-derived nuclei (arrows) in ECs (green) of ventral (d) and dorsal DA (d′) and VACS (d′′). e–f′′, 10 somites TgBAC(*pax3a*:GFP) (green), Tg(*msgn1*:nlsKaede) photoconverted embryos (e–e′′). f–f′′, At 26 somites photoconverted nuclei co-localize with GFP in Tg(*pax3a*:GFP) (100%, *n* = 29, arrows). g–g′′, Tamoxifen added at 10 somites in Tg(*msgn1*:CreERT2); Tg(*βactin2*: loxP AcGFP1-STOP pA loxP mCherry pA) animals results in mCherry in DA cells (*n* = 5, bracket) at 72 hpf. g′(i) region boxed in g′. h–h′′, Anterior somite tetramethylrhodamine dextran (TMRD)-labelled cells at 10 somites (h, RD, red) localize within the DA (h′), ECL (h′′, arrows) and

secondary fast muscle cells (h′′, arrowhead) at 36 hpf. i, TMRD mid-somite label at 10 somites generates VACs lateral to the vasculature (i′, i′′, green) (i′, lateral view; i′′′, transverse view). j–j′′, mCherry mosaically expressed from Cre/loxP 'traffic light' cassette. j′, Cre RNA iontophoretically delivered into a single cell (circle, arrow) and the deleted cell within the DA (j′′, green, arrow, (*n* = 7)). k–m′, Anterior cell in a *meox1*−/− embryo (k) ectopically generates DA (k′, arrow) and VACs (arrowhead). l, *meox1*−/− posterior somite cell generates fast muscle at 36 hpf (l′). m, Cells either side of the somite boundary in *meox1*−/− generate a DA cell (arrow) and a VAC (arrowhead) at 36 hpf (m′) and a fast muscle fibre marked with *alpha-actin*-GFP (m′′). n, Cre reporter 'traffic light' cassette. o, Schematic of approximate positions of labelled cells colour coded as to the subsequent fate of the cell in WT and *meox1*−/− genotypes. Scale bars, 20 μm.

We hypothesized that three mechanisms could generate the increase of HSCs evident in *meox1* mutants. First, endotome cells could directly contribute to HSC formation and the excess HSCs evident in *meox1* mutants could derive from increased endotome cell numbers. However, time-lapse and fate-mapping analyses failed to detect direct contribution of somite-derived cells to HSC formation in either wild-type (*n* = 23) or *meox1*-deficient contexts (*n* = 35), (Extended Data Fig. 9a–n′). Furthermore, somite cells marked using the *msgn1*:CreERT2 line failed to label cells budding from the ventral floor of the DA into the PCV, the known position of HSC cells, despite being intimately associated with these cells (*n* = 5, Fig. 3g–g′′, arrows). In line with this observation, the DA expression of the *pax3a* transgene occurs adjacent to HSCs, but does not co-localize (Extended Data Fig. 3v–x′). Therefore, we concluded

that somite-derived ECs do not directly contribute to HSC formation in wild-type or *meox1* mutant embryos.

Second, we examined whether *meox1* could influence the expression of *wnt16*, which has been previously postulated to indirectly regulate HSC generation through an undescribed 'relay signal'[4]. However, rather than exhibiting the upregulation of *wnt16* expression that this model predicts, *wnt16* expression was severely downregulated in *meox1* mutants (Extended Data Fig. 9o–x). Thus, the expansion of HSC markers evident in *meox1* mutants does not occur through an upregulation of *wnt16*, which appears dispensable for HSC induction.

Lastly, we examined if somite-derived ECs could induce resident lateral-mesoderm-derived ECs of the DA to become HSCs. In support of such a model, somite-derived ECs are found adjacent to HSC clusters in the
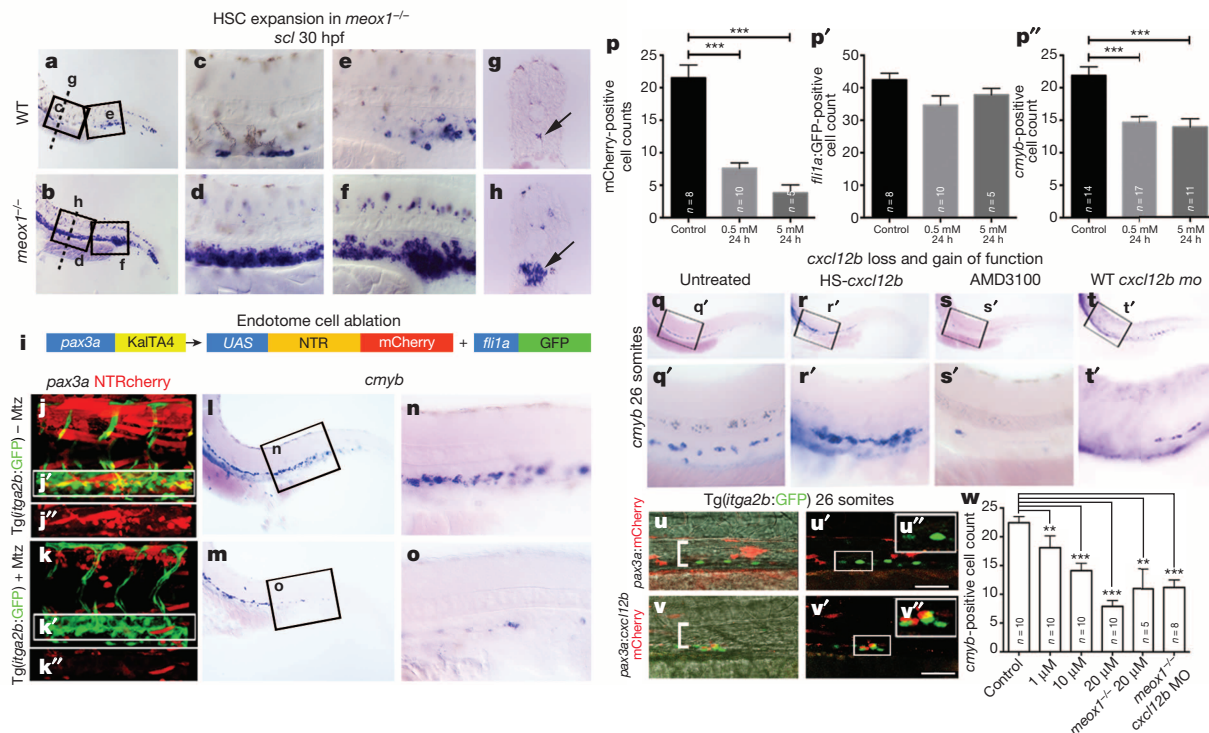
Figure 4 | Somite-derived endothelial cells are required for HSC induction. **a–h**, *scl* expression in 30 hpf wild-type (WT) (**a, c, e, g**) and *meox1*⁻/⁻ (*n* = 42, **b, d, f, h**) embryos. **c, e**, Regions boxed in **a**. **d, f**, Regions boxed in **b**. **g**, Section as in **a**. **h**, Section as in **b**. **i**, *pax3a*:KalTA4 injected into Tg(*uas*:E1b:Eco.NfsB-mCherry) (abbreviated to Tg(*uas*:NTRmCherry)) and Tg(*fli1a*:GFP) incubated without (−) (**j–j″, l, n**) or with (+) (**k–k″, m, o**) metronidazole (Mtz). Mtz addition ablates *pax3a*:NTRmCherry cells from the DA (**k–k″, m, o**) compared to controls (**j–j″, l, n**) and results in a reduction in *cmyb*+ve HSCs at 30 hpf (**m, o, p″**) compared to controls (**l, n, p″**).

p–p″, Quantification: mean ± s.e.m. significance (***P < 0.0001) in unpaired *t*-test. **q–r′**, Untreated (**q, q′**) and heat shocked (HS) (*n* = 25, **r, r′**) HS *cxcl12b* embryos stained for *cmyb*. **s, s′**, AMD3100 treatment (20 μM) reduces HSC numbers in a dose-dependent manner (**w**). **t, t′**, *cxcl12b* morpholino injection reduces HSC numbers **u–v″**, *pax3a*:cxcl12b mCherry fusion injected into the Tg(*itga2b*:GFP) (*n* = 27) reveals HSCs (green) cluster with cxcl12mCherry-expressing cells (**v–v″**) but not mCherry alone cells (**u–u″**). **w**, mean ± s.e.m.; significance (***P < 0.0001, **P < 0.005) from unpaired *t*-test. **r′–t′(i)**, Scale bars, 20 μm.

DA (Extended Data Figs 3v–x′, 9a–f) with individual somite-derived ECs associated with the generation of multiple HSCs (Extended Data Fig. 9g–n′). Furthermore, a small but significant increase in proliferation of HSCs occurs in the context of *meox1* loss of function (Extended Data Fig. 8w). To ablate endotome-derived ECs we injected a *pax3a*:KalTA4 construct into a Tg(*uas*: E1b:Eco.NfsB-mCherry) fish line which results in the expression of the nitroreductase enzyme specifically within somite-derived ECs. Addition of metronidazole (Mtz) to these embryos ablated somite-derived ECs and led to a reduction in HSCs (Fig. 4j–p″).

Genetic deletion of the genes encoding *Cxcl12* and its receptor *Cxcr4* has highlighted multiple roles for Cxcl12 in induction, maintenance, proliferation and mobilization of murine HSCs within the bone marrow[25–27], roles that could be provided embryonically by *cxcl12b* expression in somite-derived ECs. To test this hypothesis we both genetically and chemically interfered with *cxcl12b* signalling during HSC induction. Loss of *cxcl12b* function resulted in a reduction of HSCs, a deficit that can be rescued by endotome-specific expression of *cxcl12b* (Fig. 4q–q′, s–t′, w, Extended Data Fig. 10e) and reduced *cxcl12b* signalling rescued the HSC expansion evident in *meox1* mutants (Fig. 4w, Extended Data Fig. 10a–c′). Reciprocally, overexpression of *cxcl12b* via induction of the heat-shock promoter (Fig. 4q–r′) or from the endotome-specific *pax3a* promoter (*n* = 25, Fig. 4u–v′, Extended Data Fig. 10d–d′) resulted in clumps of HSC cells irregularly positioned along the DA. Furthermore, injection of *pax3a*:cxcl12b mCherry into the *itga2b*:GFP line revealed that HSC clumps are specifically associated with cxcl12b:mCherry expressing cells (*n* = 27, Fig. 4u–v″).

To study the molecular basis of *meox1* regulation of *cxcl12b* signalling in more detail, we transfected *meox1* into human ECs, which severely suppressed the level of *Cxcl12* expression in these cells (Extended Data

Fig. 10f). Furthermore, Meox1-mediated chromatin immunoprecipitation of the *cxcl12b* locus from zebrafish embryos indicates that this regulation is direct (Extended Data Fig. 10g, h).

Much emphasis has been placed on attempting to understand the cellular and molecular mechanisms underlying HSC induction, due to the obvious and far reaching implications this process holds for HSC-derived therapies. The existence of somite-derived ECs in the DA of both chick and mouse embryos suggests that the mechanism we describe here could be reproduced during amniote HSC induction. We suggest that the difficulty in inducing HSCs from pluripotent cells *in vitro*[28] may in some part be due to a requirement for somitically-derived ECs to participate in this process and provide an embryonic niche for HSC induction (Extended Data Fig. 10i). An obvious imperative is to determine what molecular signals are produced by somite-derived ECs, apart from the ones we have defined here, to coordinate HSC induction, an understanding that could lead to the efficient generation of well-defined, and clinically applicable HSCs *in vitro* and *in vivo*.

1. Davidson, A. J. & Zon, L. I. The 'definitive' (and 'primitive') guide to zebrafish hematopoiesis. *Oncogene* **23,** 7233–7246 (2004).
2. Kissa, K. & Herbomel, P. Blood stem cells emerge from aortic endothelium by a novel type of cell transition. *Nature* **464,** 112–115 (2010).
3. Bertrand, J. Y. *et al.* Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature* **464,** 108–111 (2010).

4. Clements, W. K. *et al.* A somitic Wnt16/Notch pathway specifies haematopoietic stem cells. *Nature* **474,** 220–224 (2011).
5. Noden, D. M. Embryonic origins and assembly of blood vessels. *Am. Rev. Respir. Dis.* **140,** 1097–1103 (1989).
6. Pardanaud, L. & Dieterlen-Lievre, F. Does the paraxial mesoderm of the avian embryo have hemangioblastic capacity? *Anat. Embryol. (Berl.)* **192,** 301–308 (1995).
7. Pardanaud, L. *et al.* Two distinct endothelial lineages in ontogeny, one of them related to hemopoiesis. *Development* **122,** 1363–1371 (1996).
8. Wilting, J. *et al.* Angiogenic potential of the avian somite. *Dev. Dyn.* **202,** 165–171 (1995).
9. Jaffredo, T., Gautier, R., Eichmann, A. & Dieterlen-Lievre, F. Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. *Development* **125,** 4575–4583 (1998).
10. Pouget, C., Gautier, R., Teillet, M. A. & Jaffredo, T. Somite-derived cells replace ventral aortic hemangioblasts and provide aortic smooth muscle cells of the trunk. *Development* **133,** 1013–1022 (2006).
11. Esner, M. *et al.* Smooth muscle of the dorsal aorta shares a common clonal origin with skeletal muscle of the myotome. *Development* **133,** 737–749 (2006).
12. Ben-Yair, R. & Kalcheim, C. Notch and bone morphogenetic protein differentially act on dermomyotome cells to generate endothelium, smooth, and striated muscle. *J. Cell Biol.* **180,** 607–618 (2008).
13. Childs, S., Chen, J. N., Garrity, D. M. & Fishman, M. C. Patterning of angiogenesis in the zebrafish embryo. *Development* **129,** 973–982 (2002).
14. Jin, S. W., Beis, D., Mitchell, T., Chen, J. N. & Stainier, D. Y. Cellular and molecular analyses of vascular tube and lumen formation in zebrafish. *Development* **132,** 5199–5209 (2005).
15. Kohli, V., Schumacher, J. A., Desai, S. P., Rehn, K. & Sumanas, S. Arterial and venous progenitors of the major axial vessels originate at distinct locations. *Dev. Cell* **25,** 196–206 (2013).
16. Lawson, N. D. & Weinstein, B. M. *In vivo* imaging of embryonic vascular development using transgenic zebrafish. *Dev. Biol.* **248,** 307–318 (2002).
17. Williams, C. *et al.* Hedgehog signaling induces arterial endothelial cell formation by repressing venous cell fate. *Dev. Biol.* **341,** 196–204 (2010).
18. Zhang, X. Y. & Rodaway, A. R. SCL-GFP transgenic zebrafish: *in vivo* imaging of blood and endothelial development and identification of the initial site of definitive hematopoiesis. *Dev. Biol.* **307,** 179–194 (2007).
19. Svetic, V. *et al.* Sdf1a patterns zebrafish melanophores and links the somite and melanophore pattern defects in choker mutants. *Development* **134,** 1011–1022 (2007).
20. Hollway, G. E. *et al.* Whole-somite rotation generates muscle progenitor cell compartments in the developing zebrafish embryo. *Dev. Cell* **12,** 207–219 (2007).
21. Stellabotte, F., Dobbs-McAuliffe, B., Fernandez, D. A., Feng, X. & Devoto, S. H. Dynamic somite cell rearrangements lead to distinct waves of myotome growth. *Development* **134,** 1253–1257 (2007).
22. Cha, Y. R. *et al.* Chemokine signaling directs trunk lymphatic network formation along the preexisting blood vasculature. *Dev. Cell* **22,** 824–836 (2012).
23. Bussmann, J., Bakkers, J. & Schulte-Merker, S. Early endocardial morphogenesis requires Scl/Tal1. *PLoS Genet.* **3,** e140 (2007).
24. Yabe, T. & Takada, S. Mesogenin causes embryonic mesoderm progenitors to differentiate during development of zebrafish tail somites. *Dev. Biol.* **370,** 213–222 (2012).
25. Ding, L. & Morrison, S. J. Haematopoietic stem cells and early lymphoid progenitors occupy distinct bone marrow niches. *Nature* **495,** 231–235 (2013).
26. Greenbaum, A. *et al.* CXCL12 in early mesenchymal progenitors is required for haematopoietic stem-cell maintenance. *Nature* **495,** 227–230 (2013).
27. Zou, Y. R., Kottmann, A. H., Kuroda, M., Taniuchi, I. & Littman, D. R. Function of the chemokine receptor CXCR4 in haematopoiesis and in cerebellar development. *Nature* **393,** 595–599 (1998).
28. Cerdan, C. & Bhatia, M. Novel roles for Notch, Wnt and Hedgehog in hematopoesis derived from human pluripotent stem cells. *Int. J. Dev. Biol.* **54,** 955–963 (2010).

Author Contributions P.D.C. (fate mapping, mutant analyses), P.D.N. (fate mapping, micro injections, mutant analyses, transgenic construct generation) and G.E.H. (mutant cloning, construct generation, *in situ* hybridization, micro injections, mutant analyses and cell transfections) designed and performed experiments; C.S. (*in situ* hybridizations, microinjection), L.M. (confocal analyses), T.E.H. (generated transgenic constructs), S.B. (histology), K.J.F. (*in situ* hybridization, cell transfections), D.B.G. (generated transgenic constructs) and S.A. (ChIP) performed experiments; N.J.C. and R.L.S. provided reagents; G.J.L., M.R. and J.M.P. provided reagents and assisted with revisions; P.D.C., P.D.N. and G.E.H. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.D.C. (peter.currie@monash.edu).

METHODS

Linkage mapping and candidate sequencing. *Choker* was mapped to linkage group 12 using mutant and wild-type DNA pools of WIK and Tu wild-type strains. Genomic DNA was extracted from 1,211 homozygous mutant choker embryos and used to fine map the mutation to a region of less than 1Mb (as defined by Ensembl release 49), by standard techniques with repeat markers from (http://zebrafish. mgh.harvard.edu/zebrafish/index.htm). To genotype *cho* embryos following *in situ* hybridization or antibody staining, the head was dissected, processed for DNA extraction and genotyped by sequencing using the primer set Fwd 5′-CATCCTTG AGGACTCCACTA-3′ and Rev 5′-CTCATTTGTTCTTGCCTACAGTG-3′. The trunk was mounted for microscopy and photography.

Statistical analysis. All experiments were performed with a minimum of 3 independent biological replicates; exact numbers have been indicated in figures. Statistical analysis was performed with Prizm (GraphPad) using student's unpaired two-tailed *t*-test when comparing two conditions. ANOVA with Tukey's post-hoc analysis was used when comparing multiple samples with each other. Animals were allocated to treatment groups by simple randomization.

Zebrafish strains and embryological manipulation. The *choker* allele *cho^{tm26}* has been previously described[29]. Existing transgenic lines used were Tg(*actc1b*: GFP)[zf13] (ref. 30) Tg(*itga2b*: GFP)[la2] (ref. 31), Tg(*fli1a*: EGFP)[y1] (ref. 32), Tg(*kdr1*: mCherry)[ci5] (ref. 33), Tg(*uas*:E1b:Eco.NfsB-mCherry)[c264] (ref. 34), TgBAC(*pax3a*: GFP)[i150] (ref. 35) and Tg(*hsp70l*: cxcl12b-EGFP)[mik1Tg] (ref. 36). Transgenic line and construct generation. Unless stated otherwise, all gateway compatible constructs were obtained from the tol2kit[37]. Tg(*pax3a*:GFP)[pc7]: The *pax3a*:GFP construct was generated using p5E *pax3a*[38], pME GFP, p3E polyA and pDestTol2pA2. Tg(*msgn1*: nlsKaede)[pc8]: The mCherry sequence from the pSK-tol2-3.0kb *msgn1*: mCherry construct[24] was replaced with a nls-tagged Kaede sequence. Tg(*mgsn1*: CreERT2)[pc9] was generated by replacing the mCherry sequence from pSK-tol2-3.0kb *msgn1*: mCherry construct with a CreERT2 sequence. Tg(βactin2: loxP AcGFP1-STOP loxP pA mCherry pA)[pc18] was generated by combining p5E bactin2, pME loxP AcGFP1 loxP pA (loxP flanked AcGFP pA cassette inserted into pME empty vector), p3E mCherry and pDestTol2pA2. Tg(*pax3a*: KalTA4, *alpha-crystallin:* GFP)[pc19] was generated with the construct described below. nlsKaede RNA construct was made by subcloning the nls-tagged Kaede sequence into the pCS2+ vector. The traffic light cassette was made by combining p5E EF1a, pME loxP mCherry pA, p3E GFP and pDestTol2pA2. The *pax3a: KalTA4, alpha-crystallin:* GFP construct was made by combining p5E *pax3a*, pME KalTA4[39], p3E polyA and pDestTol2pA; *alpha-crystallin:* GFP[40]. The *fli1a:* mCherry pA and *fli1a:* meox1-mCherry pA constructs were made by combining p5E *fli1aep*[41], pME mCherry, pME meox1 (full length ORF without stop codon cloned into a pME empty vector), p3E polyA, p3E mCherry and pDestTol2pA. The *pax3a:* cxcl12b mCherry pA and *pax3a:* mCherry pA was made by using p5E pax3a, pME cxcl12b (full length ORF without stop codon cloned into pME empty vector), pME mCherry, p3E mCherry, p3E polyA and pDestTol2pA in the appropriate combinations.

Morpholinos and RNA were injected using standard techniques[42]. Morpholinos (GeneTools) used were *meox1* (CTGGCTGACTGTTCCATACTGAAGA) and *cxcl12b* ATG (CGCTACTACTTTGCTATCCATGCCA) *cxcl12b* 5′ UTR (TTGATCACA GCGAGTGTCCAGAGCT). Experimental procedures were approved by the Monash Animal Services Animal Ethics committee, Monash University.

Cell ablation, photoconversion, lineage analyses, heat shock and drug treatments. Cell ablation: 24-h incubations of 0.5–5 mM metronidazole (Sigma) were started at 15 somites according to ref. 43. Photoconversion: 10-somite embryos were photoconverted on a LSM710 confocal using the 405-nm diode laser. Iontophoretic labelling was conducted as previously described[20]. Cre/loxP fate mapping: The "traffic light" cassette was injected at the one-cell stage and mosaically inherited in the G0-injected fish. iCre RNA[44] was delivered directly into mCherry-positive somitic cells by iontophoresis in 10-somite embryos. Heat shock: 10-somite embryos were immersed in 37 °C E3 water for 30 min. Drug treatments: 10-somite embryos were incubated in 1 μM–20 μM AMD3100 (Sigma) in E3 water for 24 h. Tamoxifen induced recombination: 10-somite embryos were immersed in 2.5 μM 4-hydroxy tamoxifen (Sigma) for 24 h.

***In situ* hybridization and immunohistochemistry, time-lapse analysis and microscopy.** *In situ* hybridization was conducted as described in ref. 45. Antisense probes used were: *cxcr12b*[20], *lbx* and *meox1*[46], *myod1* and *myogenin*[47], *cmyb*[48], *scl*[49], *runx1*[50] and *pax7a*[51]. For analysis of gene expression in *meox1* mutants a minimum of 30 embryos were analysed at the time points indicated. The following primer sets were PCR amplified and cloned into pGEMTeasy (Promega) *meox2a* Fwd 5′-AT GGATCACACTCTTTTCGGC-3′ and Rev 5′-TCATAAGTGAGCGTGTTCA-3′, *meox2b* Fwd 5′-ATGGATCACTCTCTGTTCGGGATGC-3′ and Rev 5′-TCATAA ATGTGCATGATCG-3′. *foxc1a*, *wnt16* and *Notch3* were obtained from Source Bioscience Lifescience Accession #BC053129 and #BC066432 and IRAKp961A02145Q respectively.

Antibody staining was performed as described in ref. 20 using antibodies anti Pax7 (DSHB), anti phospho histone H3 (Upstate) anti MyHC (A4.10.25, DSHB), anti GFP, RFP and secondary Alexa Fluor-coupled antibodies (Invitrogen). The Zf Meox1 antibody was made by MATF, Monash University, against synthetic peptide encompassing amino acids 142 to 165 of the zebrafish Meox1 protein, accession number NM_001002450.

Iontophoretic labelling and time-lapse analyses were conducted as previously described[19,20]. Time-lapse analyses were performed on Zeiss Z1, confocal images were captured on a Zeiss LSM710 and Nikon C1 upright microscopes, and image segmentation was performed using Bitplane Imaris software.

Cell culture transfections, bioinformatic predictions and chromatin immunoprecipitation analyses. Human aortic endothelial cells (HAECs) were maintained in EBM-2 media (Lonza) at 37 °C, 5% CO_2. Cells were plated into 10-cm dishes 24 h before transfection. Media was changed to Optimem (Gibco) 24 h after cells were plated into 10-cm dishes. DNA was transfected into HAECs using Lipofectamine 2000 (Invitrogen) as per the manufacturer's instructions, and media was changed back to EBM-2, 5 h after transfection. 24 h after transfection the cells were processed to a single cell suspension and sorted using fluorescence-activated cell sorting (FACS) for GFP expression. Samples were sorted on a FACSDiva Option; (BD Biosciences) running BD FACSDiVa software version 5.0.3 (BD Biosciences). RNA was extracted from the sorted cells using RNeasy (Qiagen) according to the manufacturer's instructions. cDNA was synthesized using an avian myeloblastosis virus (AMV) reverse transcriptase system (Promega) as per protocol. Quantitative real-time PCRs were carried out using the Roche Universal Probe Library System on a Roche LightCycler480, 384 well platform, with primers and UPL probes as follows: *cxcl12*: Fwd 5′-CCAAACTGTGCCCTTCAGAT-3′ Rev 5′-CTTTAGCTTCGGG TCAATGC-3′, ribosomal protein, large, P0 (*RPLP0*): Fwd 5′-GATGCCCAGGGA AGACAG-3′, Rev 5′-TCTGCTCCCACAATGAAACAT-3′. Quantitative PCR was then used to determine the relative level of expression of cxcl12 to RPLPO. Fold expression changes between GFP-positive and GFP-negative cells were calculated from three separate transfection experiments. In order to identify putative regulatory regions around the *cxcl12b* locus, UCSC genome browser (ZV9/danRer7)[52] was used to identify 7 conserved non-coding regions surrounding the *cxcl12b* locus, more specifically the identified regions were between exons and approximately 6.5kb upstream of the ATG start site. *meox1* consensus binding sites (in positional weight matrix format) were extracted from the Jaspar database. Possum; http:// zlab.bu.edu/~mfrith/possum/ was used to predict Meox1 binding sites within the identified conserved regions. Only 4 conserved regions showed a prediction score above the threshold 7. As a positive control, we searched for experimentally validated and published Meox1 binding targets to guide with interpreting the prediction score. ChIP-qPCR analysis was then conducted on these 4 regions.

Chromatin Immunoprecipitation was conducted as described in ref. 53 with the following modifications: 500 embryos staged at 15 somites within their chorions were dissociated by passing embryos several times through a 21 G syringe. Sonication was conducted for 6 × 20 s with 20 s pauses between sonication rounds using Bioruptor (Diagenode). The zebrafish Meox1 antibody described above was used for the chromatin immunoprecipitation. A PCR mix containing SYBR green (Roche) was made according to manufacturer's specifications and qPCR was conducted on a 7500 Real-Time PCR machine (Applied Biosystems). 3 non-overlapping primer sets were designed over the 4 identified conserved regions and sequences in 5′- to -3′ format were: 1_1_F: ACACACACCGAACTTTACAGT, 1_1_R: TGGGATC TCTACTTCACTCCAGA, 1_2_F: TCATTACACTGTGGCGACCC, 1_2_R: TG TGCATGCTTGAGTAACAGT, 1_3_F: TGTCAGCTGGAAAGGAATGCT, 1_3_R: TCTCACACAACAGAAATGACAGA, 2_1_F: TCCTGAAGAAAACTCAGTTA TCCT, 2_1_R: GTGTGTGTGTCTTCATGTCACT, 2_2_F: AAAGGAACTCAC ACTGTTTTTGT, 2_2_R: TCTTTGTGTTTAACAGCAGACACA, 2_3_F: TGT GTGTCTGCTGTTAAACACAAA, 2_3_R: TGACAAATTGTTTGCAATCTCA AAA, 6_1_F: TTGCATGAGTCTGTCGGTGG, 6_1_R: ACAGCCACCTAGCC TTTTCC, 6_2_F: TCCGCCACACACAGAGAAAA, 6_2_R: GCAACGTGAGC CGCTAATTC, 6_3_F: GGGGTACAAATGGTCCGGTT, 6_3_R: CTCTGTGT TGTGAGCGACCT, 7_1_F: ACAACAATCACAAATCTAAACTCGA, 7_1_R: TAAGTGACTTTCCGCAACCG, 7_2_F: TGAGATTGCACAAATTCATACG AA, 7_2_R: CAGGAGGGGTAGCAAAATCA, 7_3_F: TCCTCACTTATCCAA TCACTATTAGT and 7_3_R: TGTGTTAACAACCAATTTGAGTAGCT. C_t values were obtained and fold enrichments were calculated over IgG using the $\Delta\Delta C_t$ method: fold enrichment = $2^{-(\Delta\Delta Ct)}$, where $\Delta\Delta C_t = (C_{t(meox1)} - C_{t(Input)}) - (C_{t(IgG)} - C_{(Input)})$.

29. van Eeden, F. J. *et al.* Mutations affecting somite formation and patterning in the zebrafish, Danio rerio. *Development* **123,** 153–164 (1996).
30. Higashijima, S., Okamoto, H., Ueno, N., Hotta, Y. & Eguchi, G. High-frequency generation of transgenic zebrafish which reliably express GFP in whole muscles or the whole body by using promoters of zebrafish origin. *Dev. Biol.* **192,** 289–299 (1997).

31. Ma, D., Zhang, J., Lin, H. F., Italiano, J. & Handin, R. I. The identification and characterization of zebrafish hematopoietic stem cells. *Blood* **118,** 289–297 (2011).
32. Lawson, N. D. & Weinstein, B. M. *In vivo* imaging of embryonic vascular development using transgenic zebrafish. *Dev. Biol.* **248,** 307–318 (2002).
33. Proulx, K., Lu, A. & Sumanas, S. Cranial vasculature in zebrafish forms by angioblast cluster-derived angiogenesis. *Dev. Biol.* **348,** 34–46 (2010).
34. Pisharath, H., Rhee, J. M., Swanson, M. A., Leach, S. D. & Parsons, M. J. Targeted ablation of beta cells in the embryonic zebrafish pancreas using *E. coli* nitroreductase. *Mech. Dev.* **124,** 218–229 (2007).
35. Seger, C. *et al.* Analysis of Pax7 expressing myogenic cells in zebrafish muscle development, injury, and models of disease. *Dev. Dyn.* **240,** 2440–2451 (2011).
36. Li, Q. *et al.* Chemokine signaling guides axons within the retina in zebrafish. *J. Neurosci.* **25,** 1711–1717 (2005).
37. Kwan, K. M. *et al.* The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev. Dyn.* **236,** 3088–3099 (2007).
38. Johnson, J.-L. F. A. *et al.* Scube activity is necessary for Hedgehog signal transduction *in vivo*. *Dev. Biol.* **368,** 193–202 (2012).
39. Distel, M., Wullimann, M. F. & Koster, R. W. Optimized Gal4 genetics for permanent gene expression mapping in zebrafish. *Proc. Natl Acad. Sci. USA* **106,** 13365–13370 (2009).
40. Berger, J. & Currie, P. D. 503unc, a small and muscle-specific zebrafish promoter. *Genesis* **51,** 443–447 (2013).
41. Villefranc, J. A., Amigo, J. & Lawson, N. D. Gateway compatible vectors for analysis of gene function in the zebrafish. *Dev. Dyn.* **236,** 3077–3087 (2007).
42. Westerfield, M. *The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (Danio rerio)* 4th edn, (Univ. of Oregon Press, 2000).
43. Curado, S., Stainier, D. Y. & Anderson, R. M. Nitroreductase-mediated cell/tissue ablation in zebrafish: a spatially and temporally controlled ablation method with applications in developmental and regeneration studies. *Nature Protocols* **3,** 948–954 (2008).
44. Shimshek, D. R. *et al.* Codon-improved Cre recombinase (iCre) expression in the mouse. *Genesis* **32,** 19–26 (2002).
45. Thisse, C. & Thisse, B. High-resolution *in situ* hybridization to whole-mount zebrafish embryos. *Nature Protocols* **3,** 59–69 (2008).
46. Neyt, C. *et al.* Evolutionary origins of vertebrate appendicular muscle. *Nature* **408,** 82–86 (2000).
47. Weinberg, E. S. *et al.* Developmental regulation of zebrafish MyoD in wild-type, no tail and spadetail embryos. *Development* **122,** 271–280 (1996).
48. Thompson, M. A. *et al.* The cloche and spadetail genes differentially affect hematopoiesis and vasculogenesis. *Dev. Biol.* **197,** 248–269 (1998).
49. Liao, E. C. *et al.* SCL/Tal-1 transcription factor acts downstream of cloche to specify hematopoietic and vascular progenitors in zebrafish. *Genes Dev.* **12,** 621–626 (1998).
50. Kalev-Zylinska, M. L. *et al.* Runx1 is required for zebrafish blood and vessel development and expression of a human RUNX1-CBF2T1 transgene advances a model for studies of leukemogenesis. *Development* **129,** 2015–2030 (2002).
51. Seo, H. C., Saetre, B. O., Havik, B., Ellingsen, S. & Fjose, A. The zebrafish Pax3 and Pax7 homologues are highly conserved, encode multiple isoforms and show dynamic segment-like expression in the developing brain. *Mech. Dev.* **70,** 49–63 (1998).
52. Karolchik, D. *et al.* The UCSC Genome Browser database: 2014 update. *Nucleic Acids Res.* **42,** D764–D770 (2014).
53. Lindeman, L. C., Vogt-Kielland, L. T., Alestrom, P. & Collas, P. Fish'n ChIPs: chromatin immunoprecipitation in the zebrafish embryo. *Methods Mol. Biol.* **567,** 75–86 (2009).
54. Mathelier, A. *et al.* JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **42,** D142–D147 (2014).
55. Whitesell, T. R. *et al.* An alpha-smooth muscle actin (acta2/alphasma) zebrafish transgenic line marking vascular mural cells and visceral smooth muscle cells. *PLoS ONE* **9,** e90590 (2014).
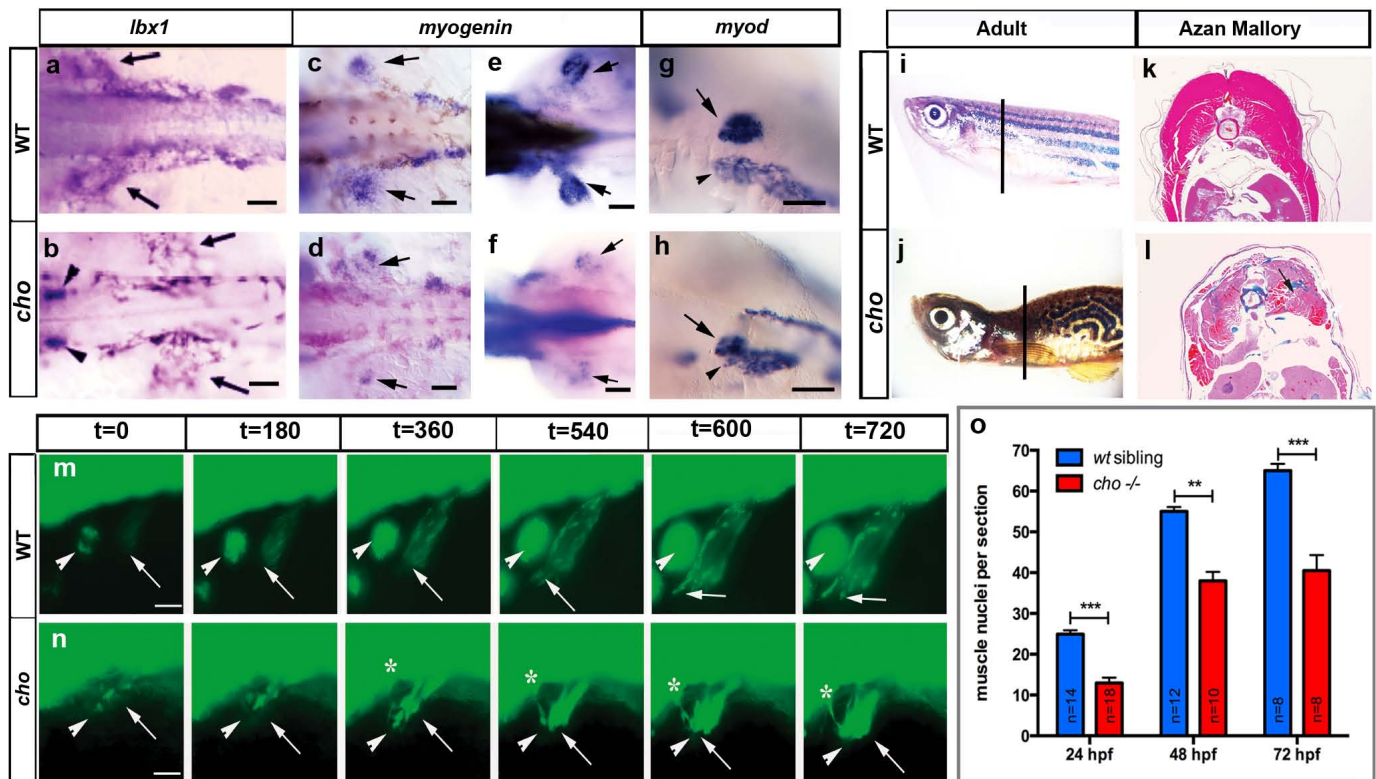
Extended Data Figure 1 | *cho* mutants exhibit defects in appendicular muscle and secondary muscle growth. a–h, Appendicular myoblast markers are reduced in *cho* mutants. **a, b,** Expression of the migratory myoblast marker *lbx1a*, which is expressed in migratory fin (arrows) is significantly reduced in *cho* homozygotes (**b**) compared to wild-type siblings (**a**). **b** is a more anterior view than **a**, displaying the hindbrain expression of *lbx1a* (arrowheads), which is unaffected in *cho* mutants. The muscle myoblast markers *myogenin* (**c–f**) and *myod* (**g, h**) are severely reduced in the fin and hypaxial muscle of *cho* mutants (**d, f, h**) compared to wild-type siblings (**c, e, g**). The posterior hypaxial muscle (arrowhead **g, h**), also exhibits an altered trajectory and reduced size in *cho* homozygotes. **a–d,** 30 hpf; **e–h,** 48 hpf. Dorsal views anterior to the left. **i–l,** Deficits in muscle formation are permanent in *cho* mutants, a small proportion of which are adult viable. Adult mutant fish (**j**) exhibit a severe

reduction of muscle, which is most evident within anterior myotomes when compared to wild type (**i**). In cross sections stained with Azan Mallory, adult *cho* mutants (**l**) exhibit fibrotic collagen deposits (blue, arrow) and a reduction in muscle (red) when compared to wild type (**k**). **m, n,** Time-lapse analyses of wild type (**m**) and *cho* mutant (**n**) transgenic for Tg(α-actin:GFP). This analysis reveals that fin muscle (arrowheads) fails to form correctly in *cho* mutants and that the posterior hypaxial muscle (PHM) (arrows) fails to extend and is much reduced in size. Furthermore, differentiating muscle cells join the PHM from anterior somites (denoted *), which are not normally fated to do so. Oblique lateral views anterior to the left. Time (t) is in minutes. Scale bars, 50 µm. **o,** Quantitation of fibre number reduction in *cho* mutants through embryonic and early larval growth. Data are mean ± s.e.m.; significance (***$P < 0.0001$, **$P < 0.005$) is from an unpaired t-test.
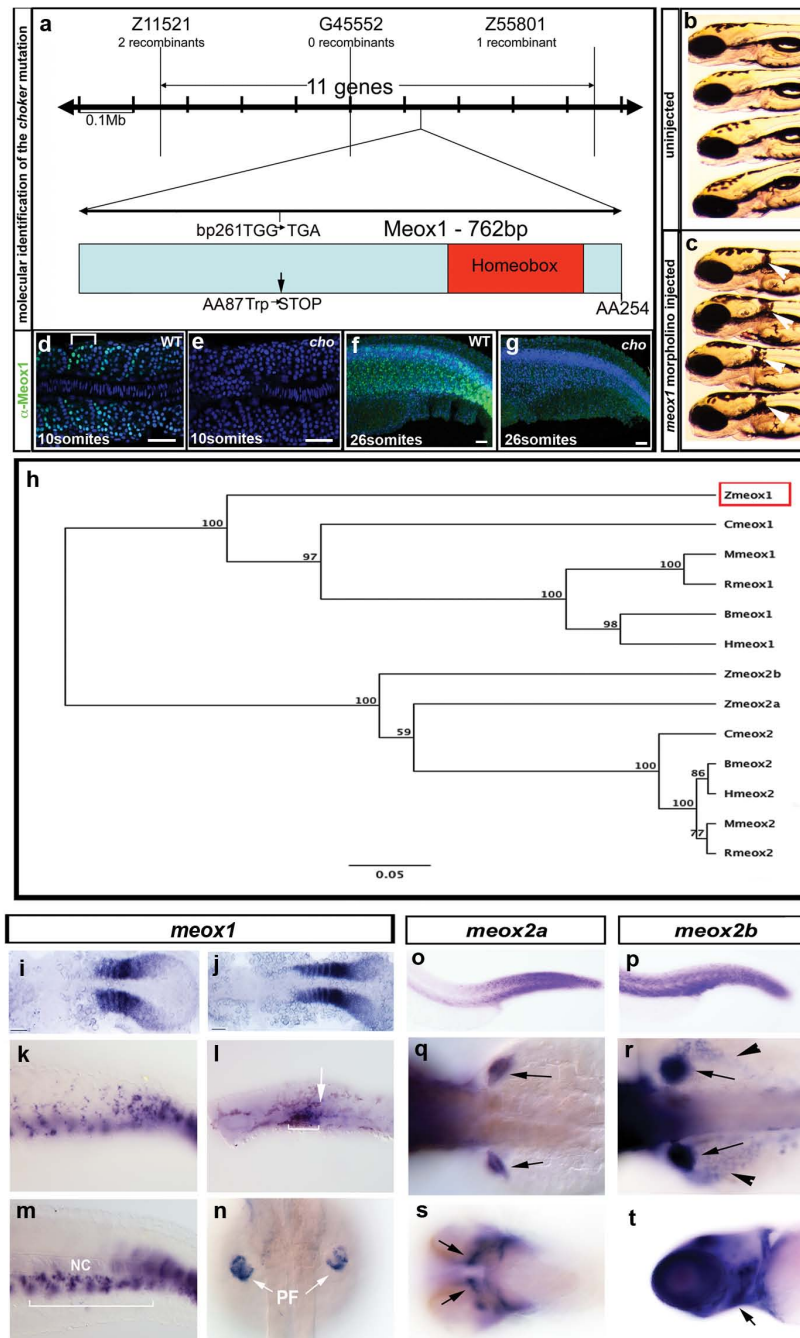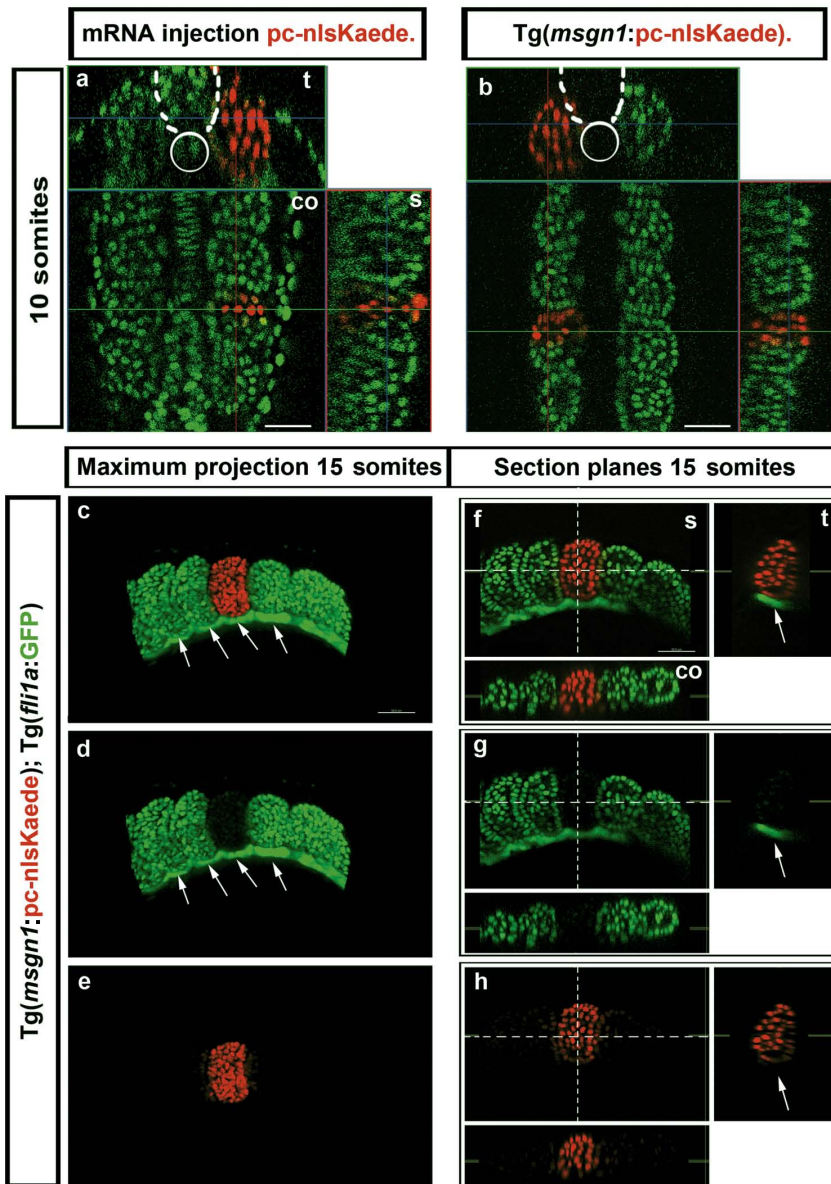
Extended Data Figure 2 | The *cho* phenotype results from a mutation in the homeobox gene *meox1*, which is expressed in somites. a, The *cho* mutation was mapped to linkage group 12, recombinants are per 2422 meioses. A stop codon (arrow) within the *meox1* ORF was identified at amino acid position 87. **b, c**, Antisense morpholinos against *meox1* resulted in a phenocopy of the *cho* phenotype, including inducing the characteristic melanocyte "choker" (*n* = 40/45, **c**, arrowhead) that is absent from uninjected siblings (**b**). **d–g**, An antibody to Meox1 recapitulates the mRNA expression of 10 and 26 somite embryos in genotyped wild-type (*meox1*+/+) siblings (**d, f**) and is restricted to nuclei of the anterior most cells of rostral somites in 10-somite embryos (bracket **d**) and the ECL in 26-somite embryos (**f**). Meox1 protein is undetectable in similarly staged genotyped *meox1*−/− homozygotes (*n* = 18, **e, g**). Blue, DAPI nuclear counterstain, scale bars 50 μm. **h**, Three different *meox* orthologues exist in the zebrafish genome. A phylogenetic tree was assembled using the Geneious program with UPGMA tree build method and bootstrap resampling using *meox* sequences in the human (H), bovine (B), mouse (M), rat (R), chicken (C) and zebrafish (Z) genomes. The gene mutated in *choker* mutants (red box) is clearly aligned as the solitary zebrafish *meox1* orthologue. Two zebrafish *meox2* orthologues are identified by this analysis.

i–t, Expression analysis of *meox* genes. **i–n**, *meox1* expression initiates in early somitogenesis (**i**, 5 somites) is confined to anterior presomitic mesoderm (PSM), the entirety of newly formed somites and is rapidly confined to anterior of mature somites (**j**, 10 somites). **k**, At 30 hpf expression in the tail is still evident within the entire ECL, and also with VACs ventral to the notochord (**m**, bracket). **l**, Expression is also evident in a large cluster of cells (bracket) immediately adjacent to somite 1 (arrow) at 30 hpf and in the pectoral fin (PF, arrows) at 48 hpf (**n**). *meox2a* is expressed at the end of somitogenesis in formed myotomes (**o**) as is *meox2b* (**p**). At 48 hpf, both *meox2a* (**q**) and *meox2b* (**r**) are expressed in fin myoblasts (arrows) and *meox2b* is expressed in the posterior hypaxial muscle (arrowhead, **r**). **s**, At 72 hpf, *meox2a* is expressed in ventral extraocular muscles (arrow). **t**, At 72 hpf, *meox2b* is expressed in specific muscles of the head (arrow). Accession numbers used: NM_001045124 Hmeox1, NM_004527 variant 1 Cmeox2, NM_001005427 Cmeox1 NM_204765, Zmeox1 NM_001002450, Zmeox2a XM_679832, Zmeox2b NM_001045124, Hmeox2 NP_005915, Bmeox2 NP_001091514, Mmeox2 NP_032610, Rmeox2 NP_058845, Bmeox1 NP_001030453, Mmeox1 NP_034921, Rmeox1 NP_001102307.
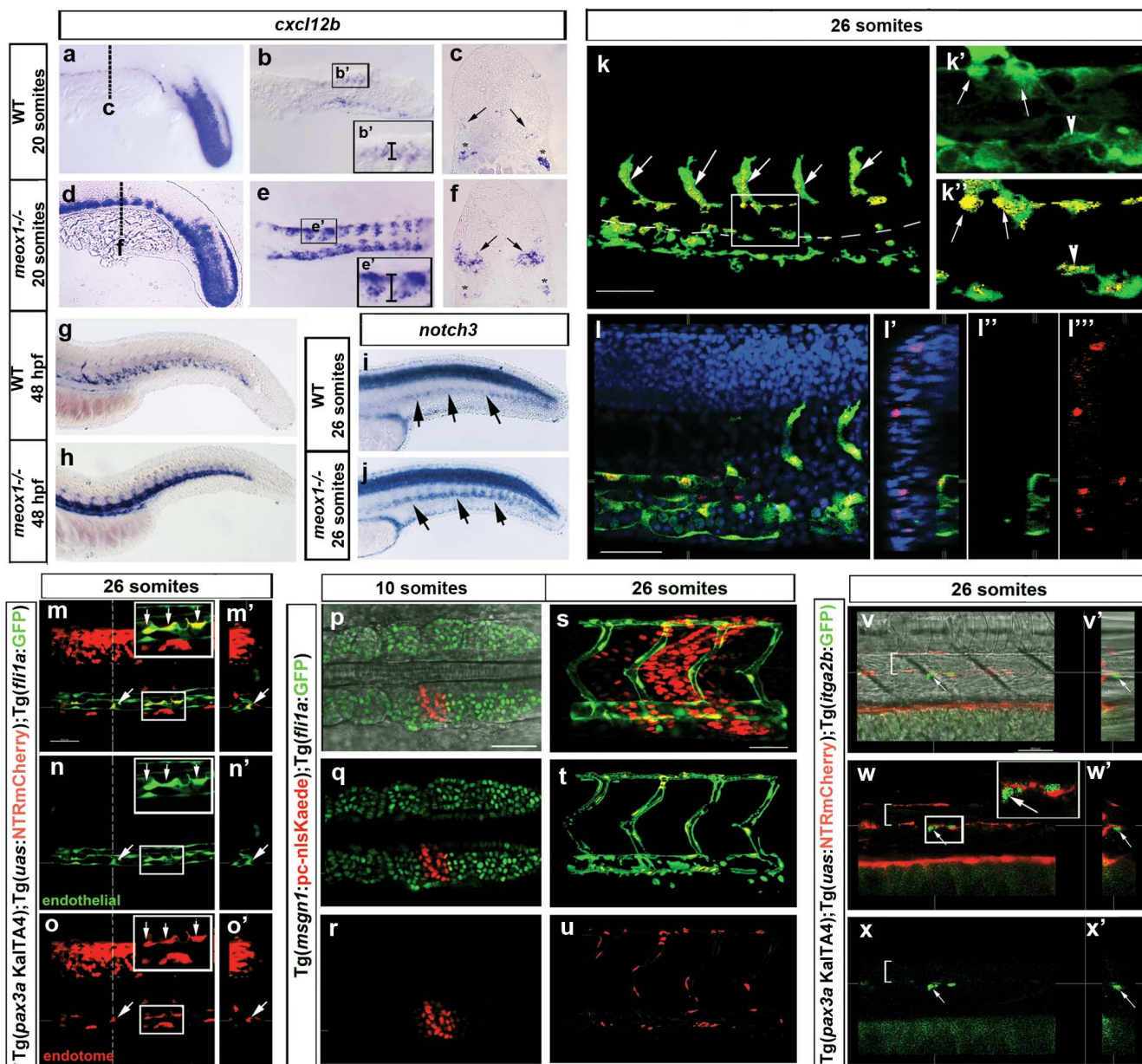
Extended Data Figure 3 | The endotome contributes to the endothelial cells of the embryonic vasculature. a–f, In 20 somite WT embryos (a–c), *cxcl12b* is expressed within cells of the ventral aspect of the somites (arrows, c). In *meox1*[−/−] mutants, the ventral, somitic expression of *cxcl12b* is expanded though out the mediolateral, anterior/posterior and dorsoventral extent of the ventral somite (compare brackets b′ to e′, ventral views of yolk extension somites and arrows in c, f; asterisk, expression in the pronephros). g, h, At 48 hpf, *cxcl12b* expression in WT (g) embryos diminishes within the embryonic vasculature, while it remains high in *meox1*[−/−] mutants (h). i, j, *notch3* expression, which marks the dorsal aorta (arrows), is moderately expanded in *meox1* mutants (j) compared to wild type siblings (i). a, d, g, h, i, j, Lateral view, anterior to the left, c, f, cross-section dorsal to top at the level of dotted lines in a and d respectively, b, b′, e, e′ ventral view anterior to left. k–k′′, Segmentation reveals computationally annotated co-localization of Meox1 protein and vascular GFP in yellow. Notably, in the dorsal aspect of the DA, co-localization occurs in DA cells immediately adjacent to the base of ISVs (arrows k′–k′′) and in the initially formed cells of the vessels themselves (arrows k–k′′) as well as in cells positioned ventrally in the DA (arrowheads, k′–k′′). l–l′′′, Single confocal slices in sagittal (l) and transverse (l′–l′′′) views of the same Meox1 positive nuclei co-localizing with *fli1a*–GFP positive endothelial cells in the DA (yellow), blue is DAPI, green is vasculature, red is Meox1 protein. Scale bars, 50 μm. m–o, Zebrafish triply transgenic for *pax3a*:KalTA4, *uas*:NTRmCherry and *fli1a*:GFP reveals that cells of the endotome, (marked by mCherry expression, m, m′, o, o′) contribute endothelial cells (marked by *fli1a*:GFP expression, m, m′, n, n′, arrows) to the embryonic vasculature including the dorsal aorta and PCV (n = 3). m′, n′, o′, Cross sections at the level of the dotted line in m, n, o, respectively. p–u, Somite-specific photo conversion reveals a contribution of the endotome to the endothelial cells of the nascent vasculature. p–r, Zebrafish carrying transgene for the photoconvertible nlsKaede protein expressed from the somite-specific mesogenin promoter (Tg(*msgn1*:nlsKaede)) that has undergone regional photoconversion in the anterior somite at the 10 somite stage encompassing the endotome (n = 11). s–u, Same embryo as in p–r in which red represents photoconverted somite nuclei and green represents *fli1a*–GFP cytoplasmic expression in the embryonic vasculature. s, The total anterior somitic contribution, including muscle nuclei, is shown in red and endothelial cells of the vasculature in green. t, Same image as in s, but with non-vasculature photoconverted nuclei computationally removed, showing the nuclei of the endotome contributing to the nascent vasculature in yellow. u, Computationally generated somite-derived nuclei that contribute to the vasculature shown alone in red. v–x′, Zebrafish triply transgenic for *pax3a*:KalTA4, *uas*:NTRmCherry and *itga2b*:GFP reveal a close association between somite-derived endothelial cells marked in red, and HSCs in green but no colocalization (arrow) (n = 4). v′, w′, x′ are transverse sections in v, w and x, respectively. w, Higher magnification view of the region boxed in w. Brackets highlight DA, lines reveal transverse section level. Scale bars, 50 μm.

mRNA injection pc-nlsKaede.

Tg(msgn1:pc-nlsKaede).

10 somites

Maximum projection 15 somites

Section planes 15 somites

Tg(msgn1:pc-nlsKaede); Tg(fli1a:GFP)

**Extended Data Figure 4 | Fidelity of somite-specific photoconversion.**
a, Views in 3 dimensions (t, transverse; co, coronial; s, sagittal) of a 10-somite
stage embryo into which mRNA encoding Kaede protein has been injected
and photoconverted reveals somite-specific photoconversion. **b**, A similar
analysis of the Tg(msgn1: nlsKaede) line reveals the somite-specific nature of
the photoconversion. Dotted lines mark the neural tube, solid circle is
notochord. **c–e**, Maximum projection of a photoconversion of a 15-somite
stage embryo carrying both the msgn1:nlsKaede and the fli1a:GFP transgenes.

This shows that there is no photoconversion (red) of any cells in the nascent
vasculature (marked by the solid line of cytoplasmic GFP ventral to the somites,
arrows) and that photoconversion is somite-specific in this line. **f–h**, Views
in 3 dimensions (t, transverse, co, coronial, s sagittal) of the same embryo in
c–e showing the somite-specific nature of the photoconversion and the lack of
any photoconverted cells in the endothelial cells of the nascent vasculature
(arrow). Scale bars, 50 µm.
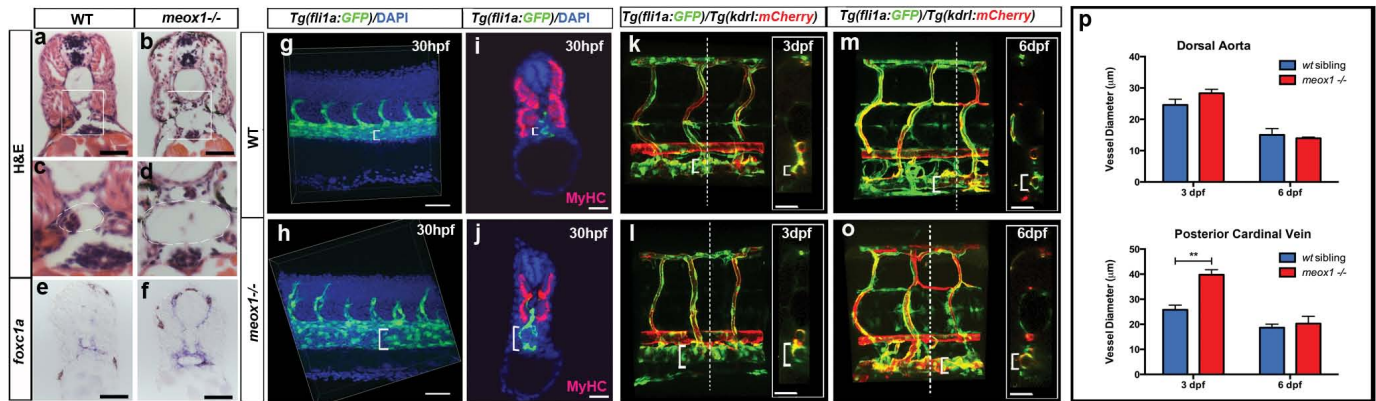
Extended Data Figure 5 | The posterior cardinal vein is transiently expanded in *meox1* mutants. a–d, Haematoxylin and eosin stained sections of wild-type (WT, $n = 3$) sibling (**a**, **c**) and *meox1* mutant embryos ($n = 5$, **b**, **d**) at 30 hpf show an increase in size of the posterior cardinal vein (PCV, boxed, dotted circle). **c**, Higher magnification of region boxed in **a**. **d**, higher magnification of the region boxed in **b**. **e–f**. *In situ* hybridization with *foxc1a* at 30 hpf, which marks the perivascular space, highlights the increase in cross sectional area of the PCV in *meox1* mutants ($n = 10$, **f**) compared to WT siblings ($n = 30$, **e**). **g**, **h**, Lateral views of *fli1a*–GFP transgenic zebrafish counterstained with DAPI (blue) reveals an expansion of the PCV (brackets) in *meox1* mutants ($n = 7$, **h**) compared to WT ($n = 9$, **g**). **i**, **j**, Cross sections at the level of the yolk extension of *fli1a*–GFP transgenic zebrafish stained with an antibody against Myosin Heavy Chain (MyHC, red), where the reduction in

secondary muscle is also apparent in *meox1* mutants as well as the expansion in the PCV, brackets show diameter of PCV. **k–o**, The expansion in the size of the PCV is transient. Transgenic zebrafish carrying both the *fli1a*–GFP and the *kdrl*–mCherry transgenes, which allows the PCV versus aorta to be more clearly distinguished due to differential level of expression of these transgenes, reveals that at 3 days post fertilisation (dpf) (**k**, **l**) the expansion in size of the PCV remains evident in *meox1* mutants (**l**) compared to WT (**k**) but is not detectable in *meox1* mutants (**o**) compared to WT (**m**) at 6dpf. **p**, Quantitation of the diameter size of the dorsal aorta and PCV at 3 and 6 dpf reveals a statistically significant increase in the size of the PCV at 3dpf but not at 6 dpf and no change in dorsal aorta size over a similar time. ($n = 6$ for each time point and genotype, insets **k–o**, cross sections at the level of the dotted line, brackets mark the PCV) (**$P < 0.005$). Scale bars, 50 μm.
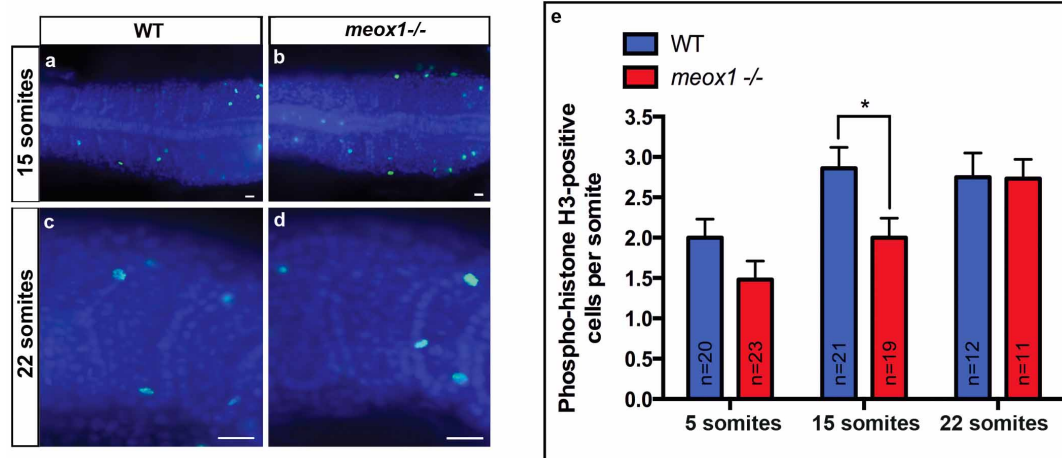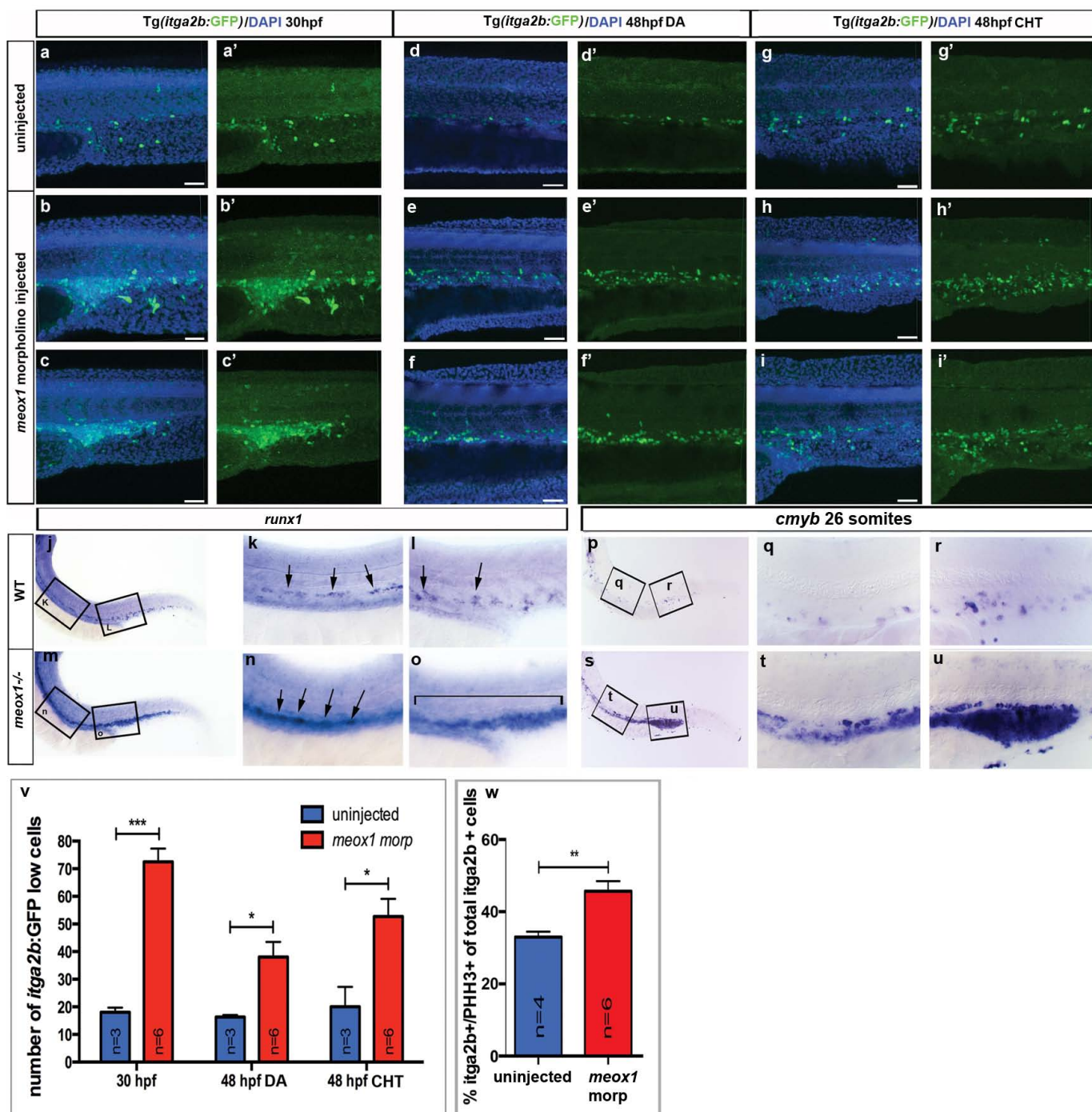
Extended Data Figure 6 | *meox1* mutant embryos do not exhibit an increase in the number of cells expressing phosphohistoneH3 (PHH3). a–e, An examination of PHH3 positive nuclei at 15 somites (**a**, **b**) and 22 somites (**c**, **d**) in wild type (**a**, **c**) and *meox1* mutants (**b**, **d**) reveals no increase in proliferation within the somites of *meox1* mutants (**e**). There was in fact a significant drop in somite proliferation in *meox1* mutants at 15 somites, which may reflect the exit of the excess endotomal cells from the somite at this stage. Data are mean ± s.e.m.; significance (*$P < 0.05$) is from an unpaired *t*-test. Blue, DAPI nuclear counterstain; scale bars, 20 µm. We also retrospectively examined the number of cells that generated DA cells and VACs in *meox1* mutant and wild-type embryos when labelled iontophoretically. This analysis also showed there was no significant difference in the number of cells generated from individual single somitic cell labels in either wild-type (1 somite cell generates 1.167 ± 0.1 s.e.m. clonal progeny $n = 38$) or *meox1* mutant/ morphant embryos (1 somite cell generates 1.20 ± 0.93 s.e.m. clonal progeny $n = 34$) despite their different sized 'vascular competent' endotome.
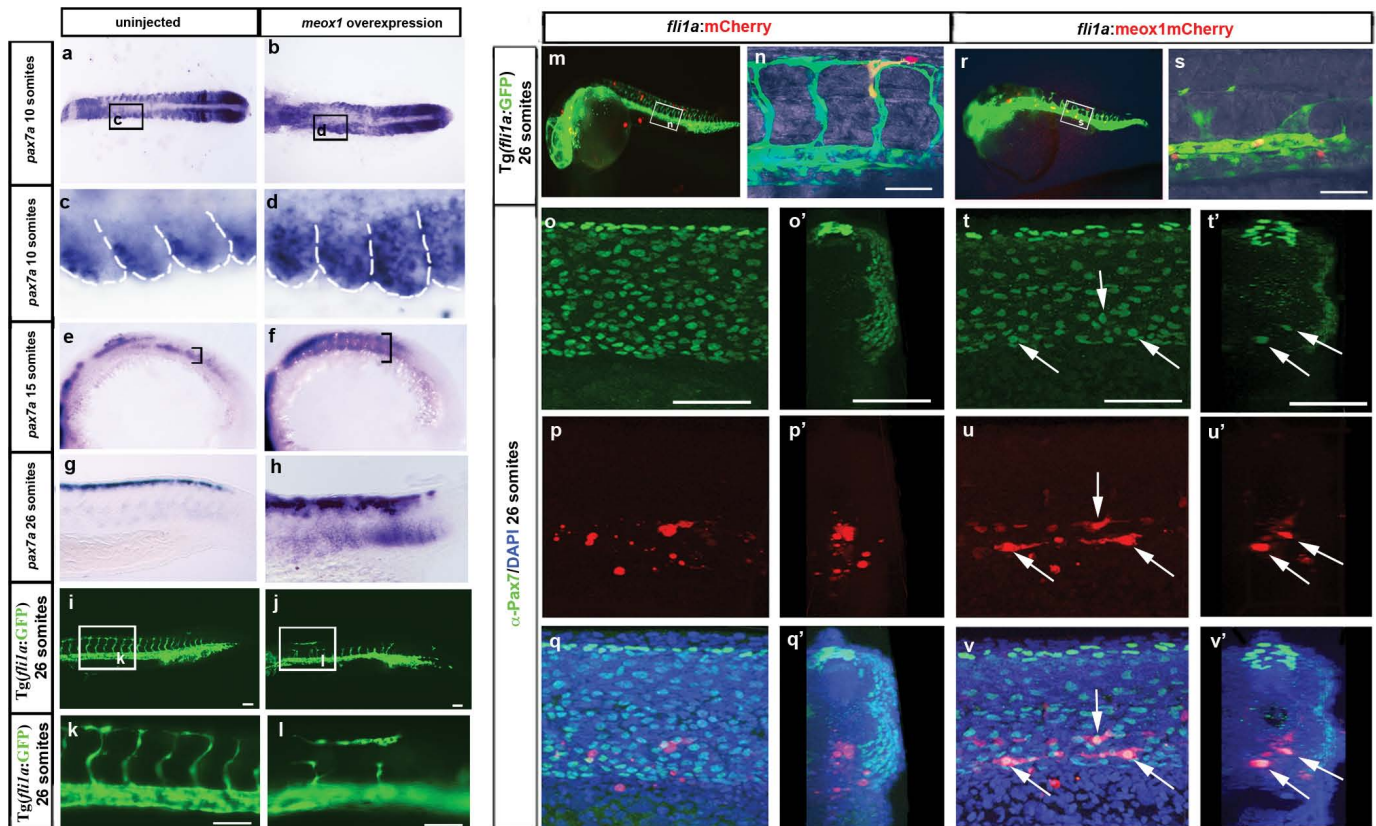
Extended Data Figure 7 | Sustained *meox1* activity promotes muscle progenitor formation and suppresses endothelial fate. **a–l**, Injection of *meox1* mRNA. **a–h**, *meox1* mRNA injection induces expression of *pax7a*. At 10 somites, *pax7a* transcript localization is expanded from its wild-type expression in anterior border cells (**a**) to the entire AP extent of the somite (**b**) (38% *n* = 26). **c**, High magnification view of the region boxed in **a**. **d**, High magnification view of the region boxed in **b**. **e**, **f**, Lateral views, anterior to the left, of 15-somite stage embryos injected with *meox1* mRNA show a dorsal-ventral expansion of *pax7a* expression (bracket, **f**) when compared to uninjected sibling (bracket, **e**). **g**, **h**, *meox1* injection induces ectopic *pax7a* expression in 26 somite injected embryos (**h**) when compared to wild type (**g**) (26%, of injected embryos *n* = 31). **i–l**, Injection of *meox1* mRNA into the Tg(*fli1a*–GFP) line specifically inhibits ISV sprouting (*n* = 37, **j**, **l**) when

compared to uninjected siblings (**i**, **k**). **k** is a high magnification view of the region boxed in **i**. **l** is a high magnification view of the region boxed in **j**. **m–v′**, Ectopic expression of *meox1* in the embryonic vasculature can cell autonomously induce Pax7 in a subset of cells and disrupt vasculature formation. **m–q′**, Expression of mCherry (red) from the *fli1a* promoter does not induce vascular malformations when injected in the *fli1a*–GFP line (**m**, **n**) and does not induce ectopic Pax7 (Green) **o–q′**. **r–v′**, By contrast expression of Meox1–mCherry fusion (red) but not mCherry alone from the *fli1a* promoter induced vascular malformations, specifically inhibiting ISV formation when injected in the Tg(*fli1a*–GFP) line (**r–s**) and induces ectopic Pax7 within vascular cells (arrows, *n* = 15, 18%). **t–v′**, Blue, DAPI nuclear counterstain, scale bars 50 μm.

Extended Data Figure 8 | Meox1 deficient embryos possess a marked increase in haematopoietic stem cells. a–i′, Injection of a *meox1* morpholino results in an increase in *itga2b*–GFP positive cells at 30 hpf (**b–c′**) and 48 hpf, both in Dorsal Aorta (DA) and the Caudal Haematopoietic Tissue (CHT) (**e–f′, h–i′**), compared to uninjected embryos (**a, a′** 30 hpf, **d, d′, g, g′** 48 hpf). Scale bars 50 μm. **j–o,** The number of *runx1* positive cells (arrows, **k, l, n,** bracket **o**) is expanded in *meox1* mutants (*n* = 26, **m–o**) compared to wild-type siblings in 48 hpf embryos (**j–l**). **k** and **l,** high magnification views of the areas boxed in **j. n** and **o,** high magnification views of the areas boxed in **m. p–u,** At 30 hpf, *meox1* deficient embryos (*n* = 32, **s–u**) exhibit an expanded population of DA associated HSC marker *cmyb* when compared to WT siblings (**p–r**). **q, r,** higher magnification of regions boxed in **p. t** and **u** higher magnification of regions

boxed in **s. v,** Quantification of *itga2b*–GFP low cells. 30 hpf are counts for the region of the DA underlying the first 6 somites post yolk extension. 48 hpf DA are *itga2b*–GFP positive cell counts for the region of the DA underlying the first 6 somites anterior to the end of the yolk extension. 48 hpf CHT, Increased colonization of HSCs to the CHT is also observed, in line with the expected consequence of increased numbers of DA-specified HSCs evident in *meox1* mutants. CHT counts are for the region underlying the first 6 somites posterior to the yolk extension. **w,** *meox1* morpholino injection results in a significant increase in the percentage of phospho histone H3 positive (PHH3+) *itga2b*–GFP positive HSCs present in the DA at 36 hpf. Data are mean ± s.e.m.; significance (***$P < 0.0001$, **$P < 0.005$, *$P < 0.05$) are from an unpaired *t*-test.
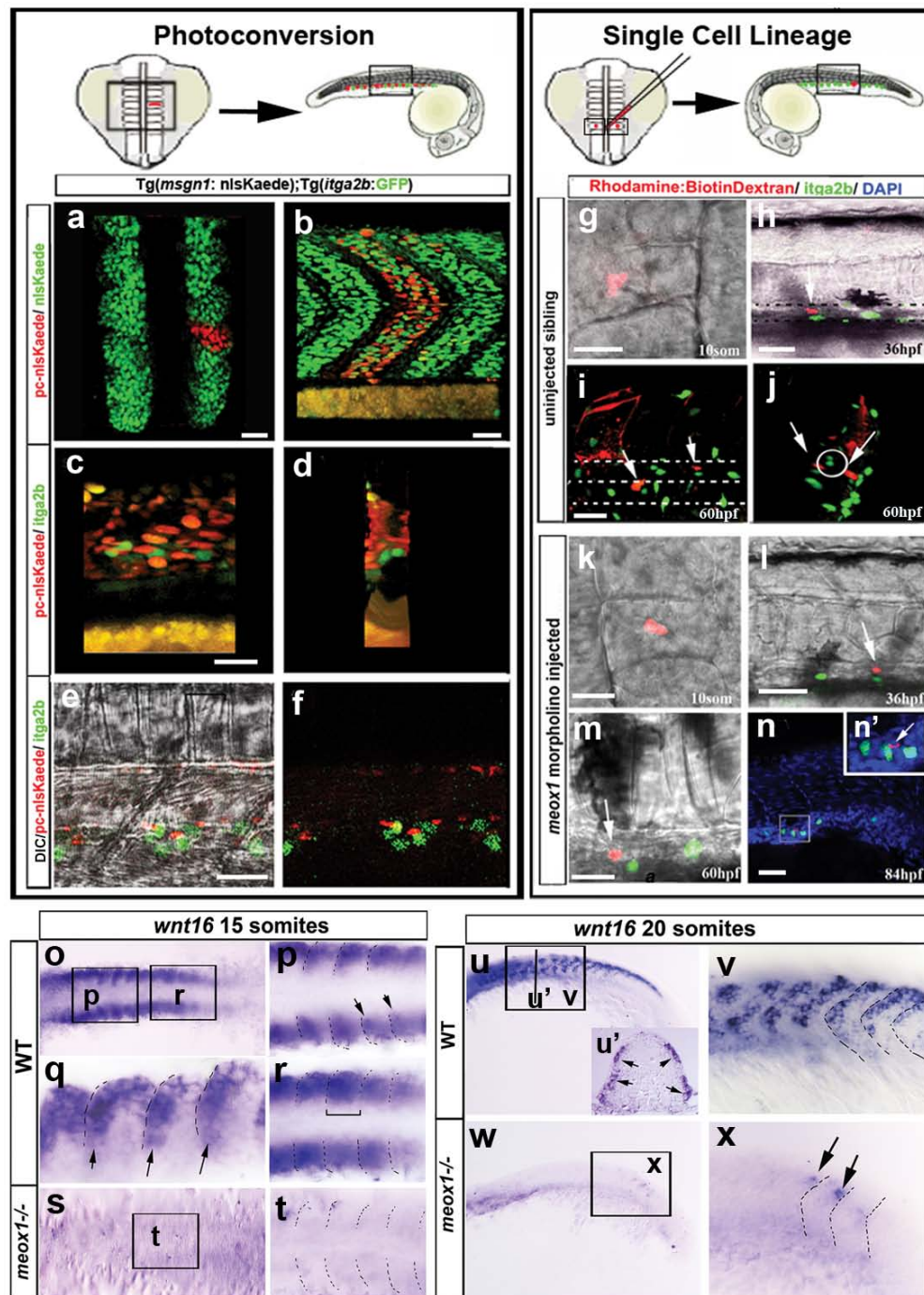
Extended Data Figure 9 | Somite-derived endothelial cells do not contribute to HSCs in wild type and *meox1* mutant embryos.
a–f, Tg(*msgn1*:nlsKaede) crossed to Tg(*itga2b*:EGFP). Photoconversion of nuclear Kaede expressed exclusively within the somites using a *mesogenin* promoter at the 10 somite stage (**a**) generates muscle and vascular associated nuclei at 26 somites (**b**). **c, d**, Maximum projections of the vascular region of the same embryo shown in **b** revealing an association but a lack of co-localization between somite-derived ECs and *itga2b*–EGFP, **c** is lateral view, **d** is transverse view. **e, f**, Single confocal slice with DIC (**e**) or without (**f**), revealing that somite-derived EC are intimately associated with *itga2b*–EGFP positive HSCs but do not co-localize with them (*n* = 40). **g–n′**, Iontophoretic fate mapping in wild-type (*n* = 23, **g–j**) and *meox1* morpholino injected embryos (*n* = 35, **k–n′**) reveals that under time-lapse individually labelled cells derived from the endotome are intimately associated with multiple *itga2b*–EGFP expressing cells, but never themselves express *itga2b*–EGFP. **g, k**, Dorsal view 10 somite stage, tetramethylrhodamine dextran label (red) within the endotome in wild-type (**g**) and *meox1* morpholino injected embryos (**k**). Same embryos at 36 hpf

(**h, l**), 60 hpf (**i, j, m**) and 84 hpf (**n**) lateral view anterior to the left. **g, h, k–m**, Live images. **i, j, n**, Single confocal scans. **i, j, n**, Maximum projections; **i, n**, lateral view; **j**, cross-sectional view, arrows mark labelled ECs. Lines mark vasculature (**i**) and circle represent the position of the DA (**j**). **n′**, high magnification view of the region boxed in **n**. **o–x**, *wnt16* expression is not expanded in *meox1* mutants. *wnt16* expression becomes restricted to anterior dorsal anterior sector (arrows **p, q**) of the zebrafish somite after an initially broad somitic expression at the 15-somite stage (bracket **r**), but is undetectable in *meox1* mutants (**s, t**). **o, p, r–t**, Dorsal views anterior to the left. **q**, Lateral view, anterior to the left. **p, r** and **t**, High magnification views of the regions boxed in **o** and **s**. Dotted lines outline somite boundaries. **u–x**, *wnt16* expression at the 20 somite stage is restricted to a dorsal anterior section of the external cell layer (ECL) in WT embryos. **u′** cross section at the level of the line marked in **u**, ECL arrows. **v**, High magnification view of the region marked in **u**. **w**, In *meox1* mutants, *wnt16* is severely reduced and detected in only a very few cells ectopically located at the posterior medial section of individual somites (arrows, **x**). **x**, High magnification view of the region boxed in **w**.
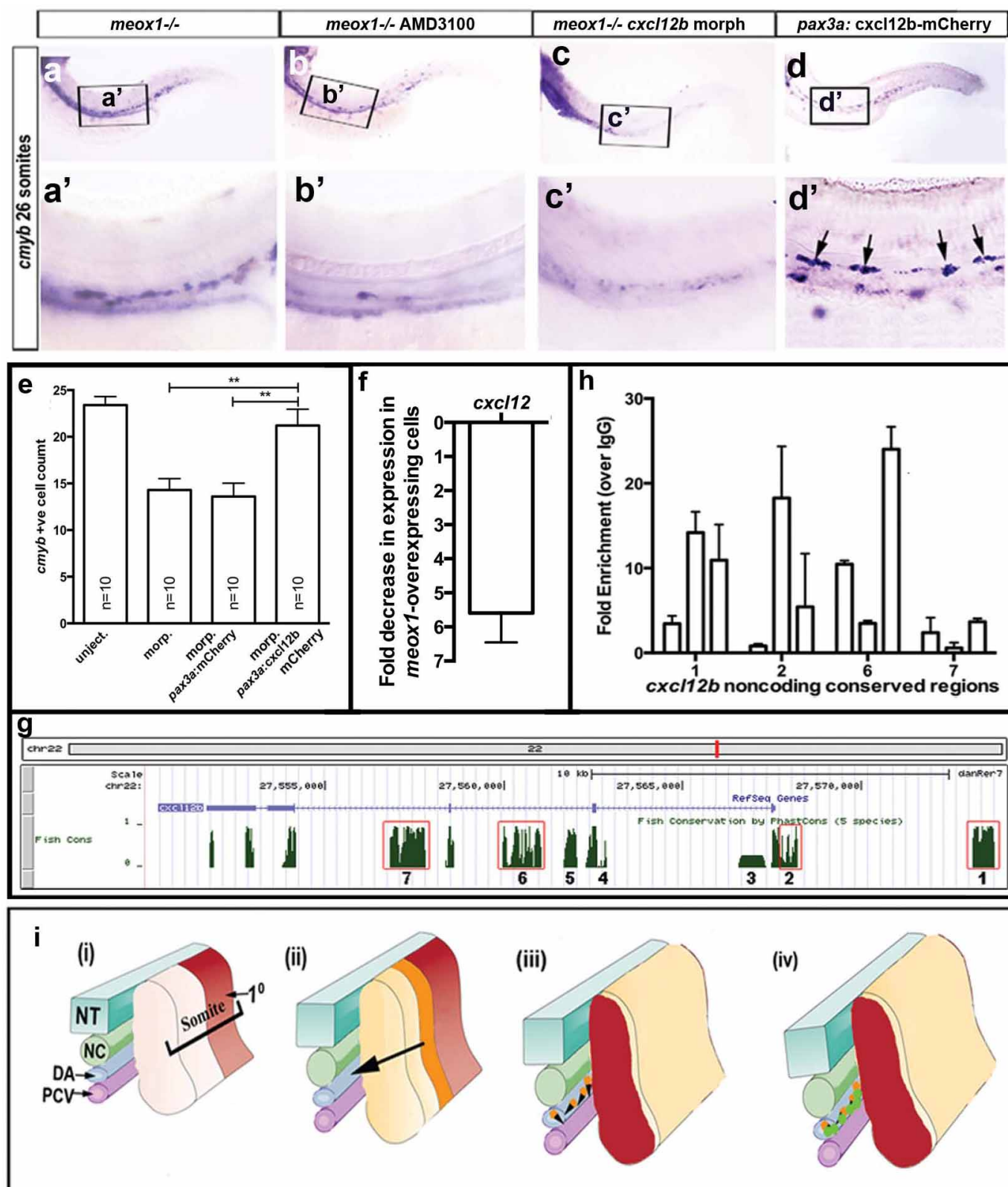
Extended Data Figure 10 | Meox1 acts as a repressor and directly regulates the expression of *cxcl12b* in zebrafish. **a–d′**, Reduction in *cxcl12* signalling rescues the *meox1*[−/−] phenotype. Addition of AMD3100 (20 μM **b**, **b′**) or injection of with a morpholino against the *cxcl12b* gene (**c**, **c′**) prevents the expansion of HSCs evident in *meox1* mutants (**a**, **a′**). Expression of *cxcl12b* from the endotome specific *pax3a* promoter induces clustering of HSC. **d**, **d′**, Clustering of *cmyb* positive HSCs is induced upon expression of a *cxcl12b* protein fused to mCherry from the endotome specific *pax3a* promoter. **e**, *cxcl12b* expressed on endotome-derived endothelial cells rescues the HSC deficit in embryos lacking global Cxcl12b activity. unject, uninjected siblings embryos. morp: Embryos injected with a *cxcl12b* morpholino, designed against the 5′UTR of *cxcl12b* gene. This morpholino knocks down global endogenous Cxcl12b activity in the developing embryo in a similar manner to the original ATG targeted morpholino, but does not target the open reading frame of *cxc12b* present within the *pax3a*:cxcl12b mCherry construct. morp. *pax3a*–mCherry: Embryos injected with the 5′UTR morpholino and pax3a mCherry vector only control that does not contain the cxcl12b open reading frame. morp. *pax3a*;cxcl12bmcherry: Embryos injected with the 5′UTR morpholino and a construct expressing cxcl12b fused to mCherry from the *pax3a* promoter. In these injected animals, endotome-derived endothelial cells expressing *cxcl12b* are able to rescue the deficit of global knock down of *cxcl12b*. Counts are total *cmyb* positive (cmyb+ve) HSCs in the DA, six somites anterior to the end of the yolk extension. Data are mean+ s.e.m., significance **P < 0.005 from an unpaired *t*-test. **f**, Human aortic endothelial cells (HAECs) were transfected with an expression construct in which Meox1 is fused to GFP. Cells were cultured for 24 h and sorted by FACs for GFP expression. Real time PCR was used to determine the relative level of expression of *cxcl12* to *ribosomal protein, large, P0* (*RPLP0*). Fold expression changes between GFP-positive and GFP-negative cells were calculated from three separate transfection experiments. Data are mean ± s.e.m. transfections $n = 3$,

PCR on each transfection performed in triplicate. **g**, UCSC genome browser showing alignment using the Phastcons algorithim of *cxcl12b* across 5 fish species revealing specific conserved non-coding elements (CNEs). All alignment algorithms tested (Phastcons, Phylof, Multiz) reveal the same seven (labelled 1–7) CNEs within introns and approximately 6.5 kb upstream of the ATG start site. Possum software was consequently used to locate Meox1 binding sites within these regions using the Meox1 positional weight matrix (PH0103.1) obtained from the Jaspar database[54]. Of the 7 selected regions only 4 (red boxes) had matching Meox1 binding sites. **h**, Chromatin immunoprecipitation (ChIP)-qPCR analysis of the 4 selected regions was conducted with 3 non-overlapping primer sequences spanning the conserved region, and showed Meox1 occupancy is enriched in regions 1, 2 and 6 but is not found at region 7. Data are mean+ s.e.m. **i**, A model for endotome formation and its contribution to definitive HSC induction. A number of distinct phases contribute to the induction of haematopoietic stem cells by somite-derived endothelial cells. (i), The somite is initially separated into a primary posterior myogenic domain (red) and a non myogenic anterior compartment (pink). (ii), The anterior somitic compartment is further partitioned into either endotome (orange) or dermomyotome (yellow) by the activity of the *meox1* gene. Endotome cells expressing the chemokine *cxcl12b* migrate from the somite. (iii), The dorsal aorta (DA) is colonized by endotome cells, which contribute to endothelial cells and a second set of cells termed vascular associated cells (VACs). A potential fate of VACs could be mural cells, which give rise to vascular associated pericytes and smooth muscle cells of the vasculature. This is unlikely, however, as DA associated expression of smooth muscle actin, which specifically marks both these cell types, does not occur until 3dpf, a phase of development well after the processes we observe here[55]. (iv), HSC (green) induction requires *cxcl12b* activity (arrowheads) secreted from endotome-derived endothelial cells. NT, neural tube; NC, notochord; DA, dorsal aorta; PCV, posterior cardinal vein.