(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 September 2019 (26.09.2019)

WIPO | PCT

(10) International Publication Number
WO 2019/183508 A1

(51) International Patent Classification:
A61K 35/28 (2015.01) *C12N 15/85* (2006.01)

(21) International Application Number:
PCT/US2019/023637

(22) International Filing Date:
22 March 2019 (22.03.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/647,433 23 March 2018 (23.03.2018) US

(71) Applicant: **THE CHILDREN'S MEDICAL CENTER CORPORATION** [US/US]; 55 Shattuck Street, Boston, Massachusetts 02115 (US).

(72) Inventors: **ZON, Leonard I.**; 105 Naples Road, Brookline, Massachusetts 02446 (US). **HAGEDORN, Elliott J.**; 99 Saint Rose Street, #2, Jamaica Plain, Massachusetts 02130 (US).

(74) Agent: **RESNICK, David S.** et al.; Nixon Peabody LLP, Exchange Place, 53 State Street, Boston, Massachusetts 02109 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(54) Title: ENDOTHELIAL CELL FACTORS AND METHODS THEREOF

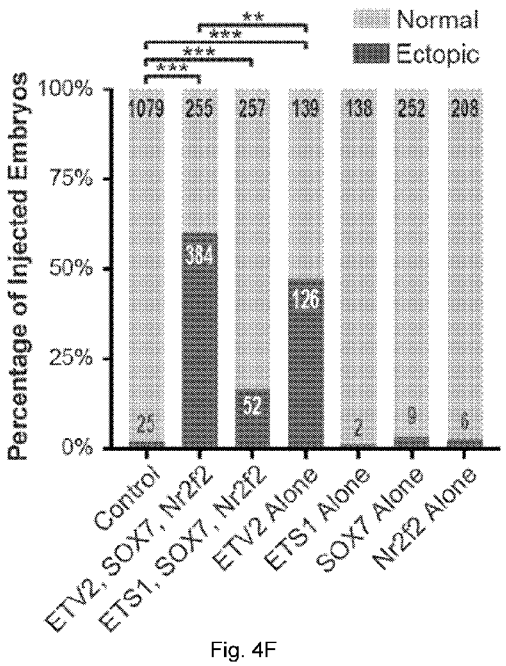


Fig. 4F

(57) Abstract: The technology described herein relates to compositions and methods of generating endothelial niche cells. Embodiments of the technology described herein comprise compositions, kits, vectors, and methods related to generating or engineering endothelial niche cells. One aspect comprises a method to generate/engineer endothelial niche cells, comprising expressing one or more transcription factors in an endothelial cell, wherein the one or more transcription factors are from the Ets family, the Sox family, and/or the Nuclear Hormone (NHR) family.

WO 2019/183508 A1

[Continued on next page]

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

— *with international search report (Art. 21(3))*

ENDOTHELIAL CELL FACTORS AND METHODS THEREOF

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims benefit under 35 U.S.C. § 119(e) of U.S. Provisional Application No. 62/647,433 filed March 23, 2018, the contents of which are incorporated herein by reference in their entirety.

GOVERNMENT SUPPORT

[0002] This invention was made with government support under Grant Nos. DK111790 and HL048801 awarded by the National Institutes of Health. The government has certain rights in the invention.

SEQUENCE LISTING

[0003] The instant application contains a Sequence Listing which has been submitted in ASCII format via EFS-Web and is hereby incorporated by reference in its entirety. Said ASCII copy, created on March 22, 2019, is named 701039-091810-seq_ST25.txt and is 32,536 bytes in size.

TECHNICAL FIELD

[0004] The technology described herein relates to compositions and methods of generating endothelial niche cells.

BACKGROUND

[0005] Haematopoietic stem and progenitor cells (HSPCs) are a rare cell population capable of reconstituting the entire blood system after transplantation. As the functional unit of a bone marrow transplant, these cells offer a curative treatment for many blood and immune diseases. Unfortunately, transplantation is not a viable treatment option for many individuals, particularly those lacking an immune-matched donor. A long-term goal of haematological research has been to culture and expand HSPCs *in vitro*, for use in transplantation and/or genetic modification. While umbilical cord blood-derived HSPCs are somewhat amenable to *in vitro* expansion,

maintaining and inducing self-renewal of adult-derived HSPCs, in the absence of niche signals, has proven challenging.

[0006] Strategies aimed at *in vitro* expansion have co-cultured HSPCs with supportive cells in an effort to recapitulate aspects of the microenvironment or 'niche' that supports HSPCs *in vivo*. In the adult bone marrow, multiple cell types are thought to collectively comprise the HSPC niche, with primary contributors being endothelial cells (ECs) and perivascular mesenchymal stromal cells. Different endothelial cell subtypes in the bone marrow can differentially regulate HSPC homeostasis. Arterial ECs (AECs) are less permeable and are believed to promote HSPC quiescence, while sinusoidal ECs (SECs) are leaky and support the differentiation and mobilization of HSPCs. In addition, during haematopoietic recovery after myelosuppression, ECs play a critical role in niche reconstruction and reconstitution of multi-lineage haematopoiesis. HSPCs can also be supported outside the bone marrow, during embryonic development and under stress conditions that induce extramedullary haematopoiesis in tissues such as the liver, spleen and skull. As in the bone marrow, ECs are thought to function as critical, core components of the HSPC niches in these tissues.

[0007] Researchers have focused on the development of in vitro cultures where HSPCs can be grown in the lab with other cells types that support the maintenance or expansion of the HSPCs for subsequent use in transplantation. To date, however, these *in vitro* cultures have been only modestly successful.

SUMMARY

[0008] In studies of the vascular HSPC niche in the zebrafish embryo a combination of transcription factors (from the Ets, Sox and Nuclear Hormone families) that are normally expressed in the endogenous niche endothelial cells were studied. When human orthologs of these same transcription factors were ectopically expressed, ectopic vascular niches in the zebrafish embryo were generated, to which HSPCs are recruited and maintained.

[0009] As a step towards translating these findings into a clinical application, transcriptions factors (which initially were identified in the zebrafish studies) can be expressed in human endothelial cells to reprogram these cells into an HSPC niche-like identity. These niche endothelial cells can be used in co-cultures with HSPCs in order to expand HSPC numbers or extend culture times, for subsequent use in transplantation.

[0010] For example, the transcription factors known to bind Ets, Sox and Nuclear hormone motifs can be expressed in niche endothelial cells. In the Ets family these factors include *etv2*, *fli1a* and *ets1*; where the corresponding human factors are ETV2, FLI1 and ETS1. In the Sox family these transcription factors include *sox18* and *sox7*; where the corresponding human factors are SOX7 and SOX18. In the Nuclear hormone family these transcription factors include *rxraa* and *nr2f2*; where the corresponding human factors are RXRA and NR2R2.

[0011] The present invention provides a method for making synthetic niche endothelial cells, to stimulate blood stem cells. Transcription factors include Ets family, *etv2*, *fli1a*, *ets1*; SOX family: *sox18*, *sox7*, and Nuclear hormone family: *rxraa*, *nr2f2*) and the corresponding human factors: ETV2, FLI1, ETS1, SOX7, SOX18, RXRA, and NR2F2.

[0012] The method comprises expressing transcription factors in endothelial cells (e.g., human) to reprogram these cells into an HSPC niche-like identity.

[0013] In another embodiment, the niche endothelial cells are used in co-cultures with HSPCs in order to expand HSPC numbers or extend culture times, for subsequent use in transplantation.

[0014] One aspect provides a method to generate/engineer endothelial niche cells, comprising expressing one or more transcription factors in an endothelial cell, wherein the one or more transcription factors are from the Ets family, the Sox family, and/or the Nuclear Hormone Receptor family.

[0015] Another aspect provides an engineered endothelial niche cell comprising one or more exogenous nucleic acid sequences encoding one or more transcription factors, wherein the one or more transcription factors are from the Ets family, the Sox family and/or the Nuclear Hormone Family.

[0016] Another aspect provides a composition comprising the engineered endothelial niche described herein.

[0017] Another aspect provides a method for culturing HSPCs, the method comprising culturing HSPCs in the presence of a population of engineered endothelial niche cells.

[0018] Another aspect provides a method of treating a subject, the method comprising, transplanting a composition comprising HSPCs and a population of engineered endothelial niche-cells into the subject.

[0019] Another aspect provides a method for enhancing engraftment of HSPCs, the method comprising administering a composition comprising HSPCs and a population of engineered endothelial niche cells to a subject in need thereof.

[0020] Another aspect provides a co-culture comprising engineered endothelial niche cells and HSPCs.

[0021] Another aspect provides a kit for culturing HSPCs, the kit comprising: a population of engineered endothelial niche cells, reagents and instructions for use thereof.

[0022] Another aspect provides a kit for generating engineered endothelial niche cells comprising: a vector(s) comprising one or more exogenous nucleic acid sequences encoding one or more transcription factors of the Ets family, the Sox family or the nuclear hormone family and instructions for use thereof.

[0023] Another aspect provides a method for generating an ectopic vascular niche, the method comprising: administering an engineered endothelial niche cell to a target site in a subject in need thereof.

[0024] Another aspect provides a method for extra medullary hematopoiesis, the method comprising transplanting engineered-niche endothelial cells into a subject at a location outside of the bone marrow (e.g., the forearm), thereby creating a synthetic niche.

[0025] Another aspect provides a vector comprising one or more exogenous nucleic acid sequences encoding one or more transcription factors of the Ets family, the Sox family or the nuclear hormone family operably linked to a promoter.

BRIEF DESCRIPTION OF THE DRAWINGS

[0026] **Fig. 1A-Fig. 1E** is a series of images and graphs showing an endothelial expression signature in the fetal HSPC niche. (**Fig. 1A**) Schematic diagram illustrates the haematopoietic tissues of the zebrafish embryo (top) and the sectioning strategy used to perform RNA tomography (tomo-seq) on the CHT (bottom; double transgenic embryo carrying the HSPC markers *cd41*:GFP and *runx1*:mCherry is shown). (**Fig. 1B**) Schematic cross-section and (**Fig. 1B**, cont.) hierarchical clustering heat map reveal clusters of gene expression that correspond to distinct tissues along the dorsal-ventral axis of the zebrafish tail. (**Fig. 1C**) Schematic depicts strategy using *kdrl:*GFP transgenic embryos and FACS to isolate ECs from whole embryos for analysis by RNA-seq. (**Fig. 1D**) Individual tomo-seq expression traces are shown for pan-

endothelial expressed genes (left) and CHT EC-enriched genes (right). (**Fig. 1E**) Images show whole mount in situ hybridization (WISH) for the pan-endothelial gene *kdrl* (top) and CHT EC-enriched genes identified by tomo-seq and tissue-specific RNA-seq. Arrows point to expression in dorsal vasculature and arrowheads point to expression in the CHT. Scale bars represent 250 μm unless noted otherwise.

[0027] **Fig. 2A-Fig. 2C** is a series of images and graphs showing endothelial niche-specific cis-regulatory elements. (**Fig. 2A**) Image and schematic depict the four cell populations that were isolated from *mrc1a*:GFP, *kdrl*:mCherry double positive embryos for analysis by ATAC-seq. (**Fig. 2B**) Gene tracks show regions of chromatin that were uniquely open in the $mCherry^{+}$; GFP^{+} CHT EC fraction (boxes and arrows). (**Fig. 2C**) Images show embryos injected with CHT EC enhancer-GFP reporter constructs corresponding to the boxed regions in Fig. 2B. Arrowheads point to GFP expression in CHT ECs. Scale bars represent 250 μm unless noted otherwise.

[0028] **Fig. 3A-Fig. 3F** is a series of images and graphs showing that Ets, Sox and NHR binding sites are required for selective expression in niche ECs. (**Fig. 3A**) Gene tracks show a region of chromatin (box and lower arrow) upstream of *mrc1a* that is uniquely open in the double positive CHT EC fraction but not the other three cell populations. Bars denote the position of the 125 bp enhancer sequence and the 1.3 kb sequence used to generate the *mrc1a*:GFP reporter transgene. (**Fig. 3B**) Images show transient GFP expression in an F0 embryo injected with the 125 bp enhancer sequence coupled to a minimal promoter and GFP. (**Fig. 3C**) Images show an F0 embryo simultaneously injected with *kdrl*:mCherry and *mrc1a* 125 bp:GFP plasmids. (**Fig. 3D**) Images show an embryo expressing the stably integrated *mrc1a* 125 bp:GFP transgene. (**Fig. 3E**) Wild-type sequence of the 125 bp *mrc1a* enhancer is shown (see e.g., **SEQ ID NO: 12**), annotated highlighting the Ets, Sox and NHR binding motifs. Schematic depicts enhancer-reporter constructs in which each class of motif or control regions was targeted by mutation. X's denote the location of targeted sites. mp-GFP: mouse *Beta-globin* minimal promoter fused to GFP. (**Fig. 3F**) Graphs report the frequency of embryos with GFP expression in CHT ECs after injection with wild-type sequences or mutated variants of the *mrc1a* 125 bp (top) or *sele* 158 bp (bottom) enhancers. Data is normalized to the respective wild-type control for each experiment (44% (155/356) for the *mrc1a* 125 bp enhancer and 23% (176/775) for the *sele* 158 bp enhancer). Mean +/- standard error of the mean (s.e.m.), One-way ANOVA; **P<0.01, ***P<0.001. Scale bars represent 250 μm unless noted otherwise.

[0029] **Fig. 4A-Fig. 4F** is a series of images and graphs showing that the overexpression of defined factors induces ectopic vascular gene expression outside the CHT. (**Fig. 4A**) Schematic depicts the strategy used in transcription factor overexpression experiments. (**Fig. 4B**) Images show embryos that were injected with control DNA (left) or a pool of seven transcription factors (right) from the Ets, Sox and NHR families (FLI1, ETV2, ETS1, SOX7, Sox18, Nr2f2 and RXRA) and then stained by WISH for *mrc1a* (top) or *sele* (bottom). Arrows denote regions of ectopic expression and arrowheads point to normal domains of expression in all panels of Fig. 4A-Fig. 4E. (**Fig. 4C**) Images show *mrc1a*:GFP; *kdrl*:mCherry double transgenic embryos that were injected with the control DNA or the 7-factor pool. (**Fig. 4D**) Injection of a 3-factor pool containing ETV2, SOX7 and Nr2f2 results in ectopic expression of *mrc1a*:GFP (arrows). (**Fig. 4E**) Images show WISH for *mrc1a* in a control embryo (top) or after injection of a 3-factor pool containing ETV2, SOX7 and Nr2f2 (middle) or ETS1, SOX7 and Nr2f2 (bottom). (**Fig. 4F**) Graph reports quantification of the percentage of injected embryos that displayed ectopic *mrc1a* expression after transcription factor overexpression. Chi Square Test; **P<0.01, ***P<0.001. Scale bars represent 250 μm in Fig. 4B and Fig. 4E, and 100 μm in Fig. 4C and Fig. 4D.

[0030] **Fig. 5A-Fig. 5E** is a series of images and graphs showing that HSPCs localize to regions of ectopic niche endothelial gene expression. (**Fig. 5A**) Images show *runx1*:mCherry^{+} HSPCs localized outside the CHT within a dorsal ectopic region of *mrc1a*:GFP expression in an embryo injected with a pool of ETV2, SOX7 and Nr2f2 (top) or a pool of ETS1, SOX7 and Nr2f2 (bottom). Inset magnifications with gray scale images for each channel are shown at right. Arrows point to ectopic expression or localization while arrowheads point to normal expression or localization in all panels in Fig. 5A-Fig. 5E. (**Fig. 5B**) WISH for *runx1* shows HSPC localization in a control (top) and 3-factor injected embryo (bottom). (**Fig. 5C**) ECs ectopically expressing *mrc1a*:GFP are associated with *cxcl12a*:DsRed2^{+} stromal cells, similar to ECs in the CHT. Asterisk denotes notochord expression of *cxcl12a*:DsRed. (**Fig. 5D**) Time-lapse series shows a *runx1*:mCherry^{+} HSPC initially arriving in the CHT and subsequently dividing. Time is shown as hh:mm. (**Fig. 5E**) Time-lapse series from a different embryo than in Fig. 5D shows *runx1*:mCherry^{+} HSPCs dividing and migrating away into circulation. Scale bars represent 100 μm in Fig. 5A-Fig. 5C and 30 μm in Fig. 5D-Fig. 5E.

[0031] **Fig. 6A-Fig. 6B** is a series of images and graphs showing a conserved endothelial expression signature in the HSPC niche. (**Fig. 6A**) Heat map shows the expression of the 29

CHT EC genes in the different cell populations that comprise the adult zebrafish kidney marrow. Spectral scale shows normalized expression between 0 (low) and 1 (high). (**Fig. 6B**) Heat map shows the expression of orthologs of the CHT EC genes in ECs from different organs of the mouse at different stages of development and postnatal transition to adulthood. Arrows denote haematopoietic tissues at the respective stage of development. Black bracket denotes genes enriched in fetal liver ECs at the E14-17 stages and then later in the adult bone marrow. Spectral scale reports z-scores. BM: Bone Marrow.

[0032] **Fig. 7A-Fig. 7B** is a series of images and graphs showing RNA tomography and niche-specific endothelial gene expression. (**Fig. 7A**) Graphs show tomo-seq expression traces for individual tissue-specific genes. Images showing whole mount *in situ* hybridization (WISH) for 35 CHT-enriched genes are available on the world wide web on zfin.org. (**Fig. 7B**) WISH validates the CHT-enriched expression (arrowheads) of CHT EC genes identified using a combination of tomo-seq and tissue-specific RNA-seq. Scale bars represent 250 μm unless noted otherwise.

[0033] **Fig. 8A-Fig. 8E** is a series of images and graphs showing that GFP reporter transgenes selectively label ECs in the HSPC niche. (**Fig. 8A**) Images show a double transgenic embryo carrying the pan-endothelial marker *kdrl*:mCherry and the *mrc1a*:GFP transgene, which is selectively expressed in CHT ECs. Magnification of boxed area is shown on the right. (**Fig. 8B**) Images show *runx1*:mCherry^{+} HSPCs directly interacting with *mrc1a*:GFP^{+} ECs within the CHT niche (arrows). Middle panel shows magnification of boxed area. Additional magnification (bottom) shows an HSPC in a pocket of *mrc1a*:GFP^{+} ECs. (**Fig. 8C**) *cxcl12a*:DsRed2^{+} stromal cells are closely associated with *mrc1a*:GFP^{+} ECs in the CHT. (**Fig. 8D**) Images show a double transgenic embryo carrying the pan-endothelial marker *kdrl*:mCherry and the *sele*:GFP transgene, which is selectively expressed in CHT ECs. Magnification of boxed area is shown on the right. (**Fig. 8E**) Images show *runx1*:mCherry+ HSPCs directly interacting with *sele*:GFP+ ECs within the CHT niche (arrows). Magnification of boxed area is shown on the right. Scale bars represent 250 μm in Fig. 8A and Fig. 8D, and 100 μm in Fig. 8B, Fig. 8C, and Fig. 8E.

[0034] **Fig. 9A-Fig. 9C** is a series of images and graphs showing a pan-endothelial regulatory elements and genome-wide motif enrichment analysis. (**Fig. 9A**) Gene tracks show regions of chromatin that were open in both the mCherry^{+}GFP^{+} (CHT EC) and mCherry^{+}GFP^{-} (non-CHT EC) populations (boxes and straight arrows). (**Fig. 9B**) Images show embryos injected

with pan-endothelial enhancer-GFP reporter constructs corresponding to the boxed regions in Fig. 9A. Arrows point to GFP expression in non-CHT ECs and arrowheads point to expression in CHT ECs. (**Fig. 9C**) Images show the transcription factor binding motifs most enriched in CHT EC regions (top) or pan-endothelial regions (bottom). Scale bars represent 250 μm unless noted otherwise.

[0035] **Fig. 10A-Fig. 10E** is a series of images and graphs showing CHT endothelial cis-regulatory elements. (**Fig. 10A**) Graph reports the anatomical location of endothelial expression in F0 embryos that were injected with *mrc1a* 125 bp:GFP and *kdrl*:mCherry plasmids. (**Fig. 10B**) Gene tracks show a region of chromatin upstream of *sele* that was uniquely open in the double positive CHT EC fraction but not the other three cell populations (box and lower arrow). Bars denote the position of the 158 bp enhancer sequence and the 5.3 kb sequence used to generate the *sele*:GFP reporter transgene. (**Fig. 10C**) Images show transient F0 (top) and stable F2 expression (bottom) of the *sele* 158 bp:GFP construct. (**Fig. 10D**) Wild-type sequence of the 158 bp *sele* enhancer is shown (see e.g., **SEQ ID NO: 13**), annotated highlighting the Ets, Sox and NHR binding motifs (top). Schematic depicts sequence variants in which each class of motif or control regions were targeted by mutation. X's denote the location of targeted sites. mp-GFP: mouse *Beta-globin* minimal promoter fused to GFP. (**Fig. 10E**) Images show electophoretic mobility shift assays with recombinant Nr2f2-GST that was incubated with DNA sequences spanning the NHR motifs present in the 125 bp *mrc1a* (left two gels) or 158 bp *sele* (right gel) enhancer sequences. Arrows point to DNA:protein binding while arrowheads point to super-shifted DNA:protein complexes. Labeled DNA:protein complexes were outcompeted by unlabeled wild-type probe (lane 4) but not by probe in which the NHR motif was disrupted by mutation (arrows with asterisks). Scale bars represent 250 μm unless noted otherwise.

[0036] **Fig. 11A-Fig. 11C** is a series of images and graphs showing that transcription factor overexpression induces ectopic CHT endothelial program. (**Fig. 11A**) Images show WISH for *mrc1a* over the yolk ball in a control (left) and 7-factor injected embryo (right). Arrows point to ectopic expression and arrowheads point to normal domains of expression in all panels of Fig. 11A-Fig. 11C. (**Fig. 11B**) Images show ectopic expression of the *mrc1a*:GFP and *kdrl*:mCherry transgenes over the yolk extension in a 7-factor injected embryo. Magnification of the boxed area is shown at the bottom. (**Fig. 11C**) Images show WISH for *sele*, *gpr182* and *lgmn* in control

embryos (left) and embryos injected with a combination of ETV2, SOX7 and Nr2f2 (right). Scale bars in **Fig. 11A-Fig. 11C** represent 100 μm.

[0037] **Fig. 12A-Fig. 12D** is a series of images and graphs showing that CHT niche endothelial gene expression is induced ectopically by transcription factor overexpression. (**Fig. 12A-Fig. 12B**) Injection of ETV2 alone induces ectopic expression of the endogenous mrc1a gene (**Fig. 12 A**) and the *mrc1a*:GFP transgene (**Fig. 12B**). Arrows point to ectopic expression and black arrowhead point to the normal domain of expression in all panels in Fig. 12A-Fig. 12D. (**Fig. 12C**) Injection of human ETV2 alone induces ectopic expression of zebrafish transcription factors, including *sox7*, *sox18*, *fli1a* and *etv2*. (**Fig. 12D**) Injection of a 3-factor pool containing ETS1, SOX7 and Nr2f2 results in ectopic expression of *mrc1a*:GFP. Scale bar represents 250 μm in Fig. 12A and Fig. 12C and 50 μm in Fig. 12B and Fig. 12D.

[0038] **Fig. 13A-Fig. 13B** is a series of images and graphs showing niche endothelial transgene expression in adult kidney ECs. (**Fig. 13A**) Images show a segment of vasculature (white arrows) dissected from the kidney of a *kdrl*:mCherry, *mrc1a*:GFP double transgenic adult zebrafish. (**Fig. 13B**) Images show sequential sections through an adult kidney isolated from a *sele*:GFP transgenic fish. Sections were stained with H&E (left) and with an antibody against GFP (right). Black arrows point to GFP^+ vascular endothelial cells. Scale bar represents 50 μm in Fig. 13A and 100 μm in Fig. 13B.

[0039] **Fig. 14** is a schematic showing hematopoietic stem cell self-renewal and differentiation. LT-HSC indicates a long-term hematopoietic stem cell. CMP indicates a common myeloid progenitor. MEP indicates a megakaryocyte-erythroid progenitor. GMP indicates a granulocyte-macrophage progenitor. CLP indicates a common lymphoid progenitor.

[0040] **Fig. 15A-Fig. 15B** is a series of images showing visualization of niche colonization by HSPCs *in vivo*. **Fig. 15A** shows visualization of the dorsal aorta and caudal hematopoietic tissue. **Fig. 15B** shows an HSPC surrounded by 5 endothelial cells and attached to stromal cell.

[0041] **Fig. 16** is a series of images showing the use of RNA tomography (tomo-seq) to examine gene expression in the HSPC niche.

[0042] **Fig. 17** is a series of images and graphs showing how Tomo-seq + endothelial RNA-seq identified ~20 genes selectively enriched in niche endothelial cells.

[0043] **Fig. 18** is a series of images showing that *sele* and *mrc1a* promoter-GFP fusions label endothelial cells in the HSPC niche.

[0044] **Fig. 19** is a series of images showing that specific ATAC-seq peaks near original 20 genes can drive expression in niche endothelial cells. 13 out of 19 cloned ATAC-seq peaks can drive GFP expression in CHT endothelial cells (when coupled to a minimal promoter).

[0045] **Fig. 20** is an image showing the TF binding motifs most enriched in the open chromatin of niche endothelial cells. Motif enrichment analysis was performed on 6,710 unique ATAC-seq peaks using HOMER.

[0046] **Fig. 21** is a series of images and graphs showing that Ets, Sox and NHR sites are required for niche expression of a 158 bp *sele* enhancer (see e.g., **SEQ ID NO: 14**).

[0047] **Fig. 22** is a series of images and graphs showing that Ets, Sox and NHR motifs are required for niche expression of a 125 bp *mrc1a* enhancer (see e.g., **SEQ ID NO: 15**).

[0048] **Fig. 23** is a series of images showing that mouse TFs can bind zebrafish sequences *in vitro*. The same mutations that disrupt enhancer:GFP expression abrogate TF binding.

[0049] **Fig. 24** is an image showing that transcription factors known to bind Ets, Sox and NH motifs are expressed in niche endothelial cells.

[0050] **Fig. 25** is a series of images showing the reprogramming of niche endothelial cells.

[0051] **Fig. 26** is a series of images and graphs showing that 3 TF pool injections result in ectopic niche endothelial gene expression.

[0052] **Fig. 27A-Fig. 27C** is a series of images showing that ectopic vascular patches can recruit *runx1*$^{+}$ HSPCs.

[0053] **Fig. 28** is a series of images and heat maps showing that HSPC niche endothelial signature is also in adult marrow.

[0054] **Fig. 29** is a heat map showing that a similar niche endothelial signature found in the mammalian fetal liver and bone marrow.

[0055] **Fig. 30** is a diagram showing that RNA tomography of the HSPC niche led to the following data: ~20 genes are selectively enriched in HSPC niche endothelial cells; ectopic expression of just 3 TFs can induce niche endothelial gene expression and recruit HSPCs; and Ets, Sox and Nuclear Hormone motifs are required for expression in niche endothelial cells.

DETAILED DESCRIPTION

[0056] Embodiments of the technology described herein comprise compositions, kits, vectors, and methods related to generating or engineering endothelial niche cells. One aspect comprises a method to generate/engineer endothelial niche cells, comprising expressing one or more transcription factors in an endothelial cell, wherein the one or more transcription factors are from the Ets family, the Sox family, and/or the Nuclear Hormone Receptor (NHR) family.

[0057] In some embodiments, at least one transcription factor is selected from the Ets family. In some embodiments, at least one transcription factor is selected from the Sox family. In some embodiments, at least one transcription factor is selected from the NHR family.

[0058] In some embodiments, at least one transcription factor is selected from the Ets family, and at least one transcription factor is selected from the Sox family. In some embodiments, at least one transcription factor is selected from the Ets family, and at least one transcription factor is selected from the NHR family. In some embodiments, at least one transcription factor is selected from the Sox family, and at least one transcription factor is selected from the NHR family.

[0059] In some embodiments, at least one transcription factor is selected from the Ets family, at least one transcription factor is selected from the Sox family, and at least one transcription factor is selected from the NHR family. In some embodiments, at least one transcription factor is selected from the Ets family, at least one transcription factor is selected from the Sox family, or at least one transcription factor is selected from the NHR family.

Ets Family

[0060] In some embodiments of any of the aspects, endothelial niche cells express transcription factors from the ETS family. The ETS (E26 transformation-specific or E-twenty-six) family is one of the largest families of transcription factors and is unique to animals. The ETS family is divided into 12 subfamilies: 1) ELF (e.g., ELF1, ELF2/NERF, ELF4/MEF); 2) ELG (e.g., GABPα, ELG); 3) ERG (e.g., ERG, FL1, FEV); 4) ERF (e.g., ERF/PE2, ETV3/PE1); 5) ESE (e.g., ELF3/ESE1/ESX, ELF5/ESE2, ESE3/EHF); 6) ETS (e.g., ETS1, ETS, POINTED); 7) PDEF (e.g., SPDEF/PDEF/PSE); 8) PEA3 (e.g., ETV4/PEA3/E1AF, ETV5/ERM, ETV1/ER81); 9) ERF71 (e.g., ETV2/ER71); 10) SPI (e.g., SPI1/PU.1, SPIB, SPIC); 11) TCF (e.g., ELK1, ELK4/SAP1, ELK3/NET/SAP2, LIN); 12) TEL (e.g., ETV6/TEL, ETV7/TEL2, YAN).

[0061] All ETS family members are identified through a highly conserved DNA binding domain, the ETS domain, which is a winged helix-turn-helix structure that binds to DNA sites with a central GGA(A/T) DNA sequence. DNA motifs for the Ets family can also comprise a central TTCCT sequence (e.g., on DNA strand complementary to the first motif). As well as DNA-binding functions, evidence suggests that the ETS domain is also involved in protein-protein interactions.

[0062] The ETS family is present throughout the body and is involved in a wide variety of functions including the regulation of cellular differentiation, cell cycle control, cell migration, cell proliferation, apoptosis (programmed cell death) and angiogenesis.

[0063] Non-limiting examples of members of the human Ets family that are relevant to endothelial niche cells comprise ETV2, FLI1, and ETS1. The corresponding factors in zebrafish comprise *etv2*, *fli1*, and *ets1*.

[0064] ETV2 can also be referred to herein as ETS Variant 2, ETS Translocation Variant 2, Ets-Related Protein 71, Ets Variant Gene 2, ETSRP71, or ER71.

[0065] Friend leukemia integration 1 transcription factor (FLI1), also known as transcription factor ERGB, is a protein that in humans is encoded by the FLI1 gene. FLI1 can also be referred to herein as Fli-1 Proto-Oncogene, ETS Transcription Factor, Friend Leukemia Integration 1 Transcription Factor, Friend Leukemia Virus Integration 1, Transcription Factor ERGB, Ewing Sarcoma Breakpoint Region, Proto-Oncogene Fli-1, BDPLT21, EWSR2, or SIC-1.

[0066] ETS1 or protein C-ets-1 is a protein that in humans is encoded by the *ETS1* gene. The protein encoded by this gene belongs to the ETS family of transcription factors. ETS1 can also be referred to herein as ETS Proto-Oncogene 1 Transcription Factor, Avian Erythroblastosis Virus E26 (V-Ets) Oncogene Homolog-1, V-Ets Avian Erythroblastosis Virus E26 Oncogene Homolog 1, Protein C-Ets-1, EWSR2, P54, V-Ets Avian Erythroblastosis Virus E2 Oncogene Homolog 1, Ets Protein, C-Ets-1, or ETS-1.

[0067] In some embodiments of any of the aspects, cells are generated which or engineered to express an Ets family member selected from the group consisting of ETV2, FLI1, and ETS1.

[0068] In some embodiments, the Ets gene or protein can be ETV2 or the corresponding zebrafish *etv2*. In some embodiments, the Ets gene or protein can be FLI1 or the corresponding zebrafish *fli1*. In some embodiments, the Ets gene or protein can be ETS1 or the corresponding zebrafish *ets1*.

[0069] In some embodiments, the Ets gene or protein can be ETV2 and FLI1 or the corresponding zebrafish factors. In some embodiments, the Ets gene or protein can be ETV2 and ETS1 or the corresponding zebrafish factors. In some embodiments, the Ets gene or protein can be ETS1 and FLI1 or the corresponding zebrafish factors. In some embodiments, the Ets gene or protein can be ETV2, FLI1, and ETS1 or the corresponding zebrafish factors. In some embodiments, the Ets gene or protein can be ETV2, FLI1, or ETS1 or the corresponding zebrafish factors.

[0070] The amino acid sequences of the polypeptides described herein have been assigned NCBI accession numbers for different species such as human, mouse, rat, and zebrafish. In particular, the NCBI accession numbers for non-limiting examples of the amino acid sequences of human ETV2 (e.g. **SEQ ID NO: 1**), human FLI1 (e.g., **SEQ ID NO: 2**), and human ETS1 (e.g., **SEQ ID NO: 3**) are included herein.

[0071] **SEQ ID NO: 1** (*Homo sapiens* ETV2, NCBI accession number AAI40747, 342 amino acids (aa)):

[0072] MDLWNWDEASPQEVPPGNKLAGLEGAKLGFCFPDLALQGDTPTATAETCW KGTSSSLASFPQLDWGSALLHPEVPWGAEPDSQALPWSGDWTDMACTAWDSWSGASQ TLGPAPLGPGPIPAAGSEGAAGQNCVPVAGEATSWSRAQAAGSNTSWDCSVGPDGDTY WGSGLGGEPRTDCTISWGGPAGPDCTTSWNPGLHAGGTTSLKRYQSSALTVCSEPSPQS DRASLARCPKTNHRGPIQLWQFLLELLHDGARSSCIRWTGNSREFQLCDPKEVARLWGE RKRKPGMNYEKLSRGLRYYYRRDIVRKSGGRKYTYRFGGRVPSLAYPDCAGGGRGAE TQ

[0073] **SEQ ID NO: 2** (*Homo sapiens* FLI1, NCBI accession number AAH10115.1, 452 aa):

[0074] MDGTIKEALSVVSDDQSLFDSAYGAAAHLPKADMTASGSPDYGQPHKINPLP PQQEWINQPVRVNVKREYDHMNGSRESPVDCSVSKCSKLVGGGESNPMNYNSYMDEK NGPPPPNMTTNERRVIVPADPTLWTQEHVRQWLEWAIKEYSLMEIDTSFFQNMDGKEL CKMNKEDFLRATTLYNTEVLLSHLSYLRESSLLAYNTTSHTDQSSRLSVKEDPSYDSVR RGAWGNNMNSGLNKSPPLGGAQTISKNTEQRPQPDPYQILGPTSSRLANPGSGQIQLWQ FLLELLSDSANASCITWEGTNGEFKMTDPDEVARRWGERKSKPNMNYDKLSRALRYYY DKNIMTKVHGKRYAYKFDFHGIAQALQPHPTESSMYKYPSDISYMPSYHAHQQKVNFV PPHPSSMPVTSSSFFGAASQYWTSPTGGIYPNPNVPRHPNTHVPSHLGSYY

[0075] **SEQ ID NO: 3** (*Homo sapiens* ETS1, NCBI accession number CAG47050.1, 441 aa):

[0076] MKAAVDLKPTLTIIKTEKVDLELFPSPDMECADVPLLTPSSKEMMSQALKAT FSGFTKEQQRLGIPKDPRQWTETHVRDWVMWAVNEFSLKGVDFQKFCMNGAALCALG KDCFLELAPDFVGDILWEHLEILQKEDVKPYQVNGVNPAYPESRYTSDYFISYGIEHAQC VPPSEFSEPSFITESYQTLHPISSEELLSLKYENDYPSVILRDPLQTDTLQNDYFAIKQEVVT PDNMCMGRTSRGKLGGQDSFESIESYDSCDRLTQSWSSQSSFNSLQRVPSYDSFDSEDYP AALPNHKPKGTFKDYVRDRADLNKDKPVIPAAALAGYTGSGPIQLRQFLLELLTDKSCQ SFISWTGDGWEFKLSDPDEVARRWGKRKNKPKMNYEKLSRGLRYYYDKNIIHKTAGK RYVYRFVCDLQSLLGYTPEELHAMLDVKPDADE

[0077] In some embodiments, an ETV2 amino acid or DNA sequence can be at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or more, identical to a native or reference sequence. In some embodiments, a FLI1 amino acid or DNA sequence can be at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or more, identical to a native or reference sequence. In some embodiments, an ETS amino acid or DNA sequence can be at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or more, identical to a native or reference sequence.

SOX Family

[0078] SOX genes encode a family of transcription factors that bind to the minor groove in DNA, and belong to a super-family of genes characterized by a homologous sequence called the HMG-box (for high mobility group). This HMG box is a DNA binding domain that is highly conserved throughout eukaryotic species. Homologues have been identified in insects, nematodes, amphibians, reptiles, birds and a range of mammals.

[0079] Sox genes are defined as containing the HMG box of a gene involved in sex determination called SRY, which resides on the Y-chromosome (Sox stands for Sry-related HMG box). There are 20 SOX genes present in humans and mice. The family is divided into subgroups according to homology within the HMG domain and other structural motifs, as well as according to functional assays. In humans the members of the SOX groups comprise: 1) SoxA

(e.g., SRY); 2) SoxB1 (e.g., SOX1, SOX2, SOX3); 3) SoxB2 (e.g., SOX14, SOX21); 4) SoxC (e.g., SOX4, SOX11, SOX12); 5) SoxD (e.g., SOX5, SOX6, SOX13); 6) SoxE (e.g., SOX8, SOX9, SOX10); 7) SoxF (e.g., SOX7, SOX17, SOX18); 8) SoxG (e.g., SOX15); 9) SoxH (e.g., SOX30).

[0080] The developmentally important Sox family has no singular function, and many members possess the ability to regulate several different aspects of development. While many Sox genes are involved in sex determination, some are also important in processes such as neuronal development. Sox proteins bind to the sequence WWCAAW and similar sequences (W=A or T). DNA motifs for the Sox family can also comprise a central ATTGT sequence (e.g., on DNA strand complementary to the first motif).

[0081] Non-limiting examples of members of the human Sox family that are relevant to endothelial niche cells comprise SOX18 and SOX7. The corresponding factors in zebrafish or *Xenopus* comprise *sox18* and *sox7*.

[0082] SOX18 can also be referred to herein as SRY-Box 18, SRY (Sex Determining Region Y)-Box 18, Transcription Factor SOX-18, SRY Box 18, HLTRS, or HLTS.

[0083] SOX7 can also be referred to herein as SRY-Box 7, SRY (Sex Determining Region Y)-Box 7, Transcription Factor SOX-7, or SRY Box 7.

[0084] In some embodiments of any of the aspects, cells are generated which or engineered to express a Sox family member selected from the group consisting of SOX18 and SOX7. In some embodiments, the Sox gene or protein can be SOX18 or the corresponding zebrafish *sox18 or Xenopus* SOX18. In some embodiments, the Ets gene or protein can be SOX7 or the corresponding zebrafish *sox7*. In some embodiments, the Sox gene or protein can be SOX18 and SOX7 or the corresponding zebrafish or *Xenopus* factors. In some embodiments, the Sox gene or protein can be SOX18 or SOX7 or the corresponding zebrafish or *Xenopus* factors.

[0085] The amino acid sequences of the polypeptides described herein have been assigned NCBI accession numbers for different species such as human, mouse, rat, and zebrafish. In particular, the NCBI accession numbers for non-limiting examples of the amino acid sequences of human SOX18 (e.g. **SEQ ID NO: 4**), human SOX7 (e.g., **SEQ ID NO: 5**), and *Xenopus* SOX18 (e.g., **SEQ ID NO: 8**) are included herein.

[0086] **SEQ ID NO: 4** (*Homo sapiens* SOX18, NCBI accession number BAA94874.1, 384 aa):

[0087] MQRSPPGYGAQDDPPARRDCAWAPGHGAAADTRGLAAGPAALAAPAAPAS PPSPQRSPPRSPEPGRYGLSPAGRGERQAADESRIRRPMNAFMVWAKDERKRLAQQNPD LHNAVLSKMLGKAWKELNAAEKRPFVEEAERLRVQHLRDHPNYKYRPRRKKQARKA RRLEPGLLLPGLAPPQPPPEPFPAASGSARAFRELPPLGAEFDGLGLPTPERSPLDGLEPGE AAFFPPPAAPEDCALRPFRAPYAPTELSRDPGGCYGAPLAEALRTAPPAAPLAGLYYGTL GTPGPYPGPLSPPPEAPPLESAEPLGPAADLWADVDLTEFDQYLNCSRTRPDAPGLPYHV ALAKLGPRAMSCPEESSLISALSDASSAVYYSACISG

[0088] **SEQ ID NO: 5** (*Homo sapiens* SOX7, NCBI accession number, CAC84226.1, 388 aa):

[0089] MASLLGAYPWPEGLECPALDAELSDGQSPPAVPRPPGDKGSESRIRRPMNAF MVWAKDERKRLAVQNPDLHNAELSKMLGKSWKALTLSQKRPYVDEAERLRLQHMQD YPNYKYRPRRKKQAKRLCKRVDPGFLLSSLSRDQNALPEKRSGSRGALGEKEDRGEYSP GTALPSLRGCYHEGPAGGGGGGTPSSVDTYPYGLPTPPEMSPLDVLEPEQTFFSSPCQEE HGHPRRIPHLPGHPYSPEYAPSPLHCSHPLGSLALGQSPGVSMMSPVPGCPPSPAYYSPA TYHPLHSNLQAHLGQLSPPPEHPGFDALDQLSQVELLGDMDRNEFDQYLNTPGHPDSAT GAMALSGHVPVSQVTPTGPTETSLISVLADATATYYNSYSVS

[0090] **SEQ ID NO: 8** (*Xenopus tropicalis* SOX18, NCI accession number: AAI67402.1, 362 aa):

[0091] MHRPEPSYCREEPTPCQGVNSTWVPPADTVPETSPTPSSPPAPDSPTPSPQPGY GYSPCEEKPGDPRIRRPMNAFMVWAKDERKRLAQQNPDLHNAVLSKMLGQSWKNLSS AEKRPFVEEAERLRVQHLQDHPNYKYRPRRKKQAKKLKRVDPSPLLRNEGYRGQAMA NLSHFRDLHPLGGSGDLESYGLPTPEMSPLDVVEPSEPAFFPPHMREEADPGPFRTYQHG VDFGQEKTLREISLPYSSSPSHMGGFLRTPTASAFYYNPHGGSPACTPLGQLSPPPEAPAL EAMDHLGPAELWGDFDRNEFDQYLNMSRTQGPGYPFPMSKLGAPRTIPCEESSLISALS DASTAMYYTPCITG

[0092] In some embodiments, a SOX7 amino acid or DNA sequence can be at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or more, identical to a native or reference sequence. In some embodiments, a SOX18 amino acid or DNA sequence can be at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or more, identical to a native or reference sequence.

NHR Family

[0093] The Nuclear Hormone Receptor (NHR) family, also referred to as nuclear receptors, are a class of proteins found within cells that are responsible for sensing steroid and thyroid hormones and certain other molecules. In response, these receptors work with other proteins to regulate the expression of specific genes, thereby controlling the development, homeostasis, and metabolism of the organism. A unique property of nuclear receptors that differentiates them from other classes of receptors is their ability to directly interact with and control the expression of genomic DNA. As a consequence, nuclear receptors play key roles in both embryonic development and adult homeostasis. A non-limiting example of a DNA motifs for NHR family members comprises RRGGTCA, where R denotes a purine (e.g., A or G).

[0094] At least 48 nuclear receptors have been identified in humans, classified into the following subfamilies: 1) Thyroid-Hormone Receptor-like (e.g., Thyroid hormone receptor, Retinoic acid receptor, Peroxisome proliferator-activated receptor, Rev-ErbA, RAR-related orphan receptor, Liver X receptor-like, Vitamin D receptor-like, NRs with two DNA binding domains, RORA); 2) Retinoid X Receptor-like (e.g., Hepatocyte nuclear factor-4, Retinoid X receptor, Testicular receptor, TLX, PNR, COUP, EAR, RXRA, NR2F2); 3) Estrogen Receptor-like (e.g., Estrogen receptor, Estrogen related receptor, 3-Ketosteroid receptors); 4) Nerve Growth Factor IB-like (e.g., NGFIB, NURR1, NOR1); 5) Steroidogenic Factor-like (e.g., SF1, LRH1); 6) Germ Cell Nuclear Factor-like (e.g., GCNF); 7) miscellaneous nuclear receptors (e.g., DX, SHP1).

[0095] Non-limiting examples of members of the human NHR family that are relevant to endothelial niche cells comprise RXRA and NR2F2. The corresponding factors in zebrafish comprise *rxraa* and *nr2f2*.

[0096] RXRA is a nuclear receptor that belongs to the RXR transcription factor group. RXRA can also be referred to herein as Retinoid X Receptor Alpha, Nuclear Receptor Subfamily 2 Group B Member 1, Retinoic Acid Receptor RXR-Alpha, NR2B1, Retinoid X Nuclear Receptor Alpha, or Retinoid X Receptor Alpha.

[0097] The retinoid X receptor (RXR) is a type of nuclear receptor that is activated by 9-cis retinoic acid and 9-cis-13,14-dihydro-retinoic acid, which is likely to be the major endogenous mammalian RXR-selective agonist. There are three retinoic X receptors (RXR): RXR-alpha,

RXR-beta, and RXR-gamma, encoded by the RXRA, RXRB, RXRG genes, respectively. RXR heterodimerizes with subfamily 1 nuclear receptors including CAR, FXR, LXR, PPAR, PXR, RAR, TR, and VDR. As with other type II nuclear receptors, the RXR heterodimer in the absence of ligand is bound to hormone response elements complexed with corepressor protein. Binding of agonist ligands to RXR results in dissociation of corepressor and recruitment of coactivator protein, which, in turn, promotes transcription of the downstream target gene into mRNA and eventually protein.

[0098] NR2F2 is a nuclear receptor that belongs to the COUP transcription factor group. NR2F2 can also be referred to herein as Nuclear Receptor Subfamily 2 Group F Member 2, Apolipoprotein A-I Regulatory Protein 1, COUP Transcription Factor II, COUP Transcription Factor 2, TFCOUP2, ARP-1, ARP1, Chicken Ovalbumin Upstream Promoter Transcription Factor 2, Chicken Ovalbumin Upstream Promoter-Transcription Factor I, Nuclear Receptor Subfamily 2 Group F Member 2, ADP-Ribosylation Factor Related Protein 1, Apolipoprotein AI Regulatory Protein 1, COUP-TF II, COUPTFII, COUP-TF2, COUPTF2, COUPTFB, CHTD4, NF-E3, or SVP40.

[0099] The chicken ovalbumin upstream promoter transcription factor (COUP-TFs) proteins are members of the nuclear receptor family of intracellular transcription factors. There are two variants of the COUP-TFs, labeled as COUP-TFI and COUP-TFII encoded by the NR2F1 and NR2F2 genes respectively. COUP-TFs play critical roles in the development of organisms.

[00100] In some embodiments of any of the aspects, cells are generated which or engineered to express an NHR family member selected from the group consisting of RXRA and NR2F2. In some embodiments, the NHR gene or protein can be RXRA or the corresponding zebrafish *rxraa*. In some embodiments, the NHR gene or protein can be NR2F2 or the corresponding zebrafish *nr2f2*. In some embodiments, the NHR gene or protein can be RXRA and NR2F2 or the corresponding zebrafish factors. In some embodiments, the NHR gene or protein can be RXRA or NR2F2 or the corresponding zebrafish factors.

[00101] The amino acid sequences of the polypeptides described herein have been assigned NCBI accession numbers for different species such as human, mouse, rat, and zebrafish. In particular, the NCBI accession numbers for non-limiting examples of the amino acid sequences of human RXRA (e.g. **SEQ ID NO: 6**), human NR2F2 (e.g., **SEQ ID NO:** 7), and zebrafish *Nr2f2* (e.g., **SEQ ID NO: 9**) are included herein.

[00102] **SEQ ID NO: 6** (*Homo sapiens* RXRA isoform A, NCBI accession number NP_002948.1, 462 aa):

[00103] MDTKHFLPLDFSTQVNSSLTSPTGRGSMAAPSLHPSLGPGIGSPGQLHSPISTL SSPINGMGPPFSVISSPMGPHSMSVPTTPTLGFSTGSPQLSSPMNPVSSSEDIKPPLGLNGV LKVPAHPSGNMASFTKHICAICGDRSSGKHYGVYSCEGCKGFFKRTVRKDLTYTCRDN KDCLIDKRQRNRCQYCRYQKCLAMGMKREAVQEERQRGKDRNENEVESTSSANEDMP VERILEAELAVEPKTETYVEANMGLNPSSPNDPVTNICQAADKQLFTLVEWAKRIPHFSE LPLDDQVILLRAGWNELLIASFSHRSIAVKDGILLATGLHVHRNSAHSAGVGAIFDRVLT ELVSKMRDMQMDKTELGCLRAIVLFNPDSKGLSNPAEVEALREKVYASLEAYCKHKYP EQPGRFAKLLLRLPALRSIGLKCLEHLFFFKLIGDTPIDTFLMEMLEAPHQMT

[00104] **SEQ ID NO: 7** (*Homo sapiens* NR2F2 isoform A, NCBI accession number NP_066285.1, 414 aa):

[00105] MAMVVSTWRDPQDEVPGSQGSQASQAPPVPGPPPGAPHTPQTPGQGGPAST PAQTAAGGQGGPGGPGSDKQQQQQHIECVVCGDKSSGKHYGQFTCEGCKSFFKRSVRR NLSYTCRANRNCPIDQHHRNQCQYCRLKKCLKVGMRREAVQRGRMPPTQPTHGQFAL TNGDPLNCHSYLSGYISLLLRAEPYPTSRFGSQCMQPNNIMGIENICELAARMLFSAVEW ARNIPFFPDLQITDQVALLRLTWSELFVLNAAQCSMPLHVAPLLAAAGLHASPMSADRV VAFMDHIRIFQEQVEKLKALHVDSAEYSCLKAIVLFTSDACGLSDVAHVESLQEKSQCA LEEYVRSQYPNQPTRFGKLLLRLPSLRTVSSSVIEQLFFVRLVGKTPIETLIRDMLLSGSSF NWPYMAIQ

[00106] **SEQ ID NO: 9** (*Danio rerio Nr2f2*; NCBI accession number: AAI62484.1; 428 aa):

[00107] MAMVVWRGSQDDVAETHGTLSSQTQGGLSLPTPQPGQLGLTASQVAPPTPQ TPVQGPPNNNNNTQSTPTNQTTQSQSEKQQPQHIECVVCGDKSSGKHYGQFTCEGCKSF FKRSVRRNLTYTCRANRNCPIDQHHRNQCQYCRLKKCLKVGMRREVSLFTAAVQRGR MPPTQPHHGQFALTNGDPLHCHSYLSGYISLLLRAEPYPTSRYGSQCMQPNNIMGIENIC ELAARMLFSAVEWARNIPFFPDLQITDQVALLRLTWSELFVLNAAQCSMPLHVAPLLAA AGLHASPMSADRVVAFMDHIRIFQEQVEKLKALHVDSAEYSCLKAIVLFTSDACGLSDV AHVESLQEKSQCALEEYVRSQYPNQPTRFGKLLLRLPSLRTVSSSVIEQLFFVRLVGKTPI ETLIRDMLLSGSSFNWPYMSIQ

[00108] In some embodiments, a RXRA amino acid or DNA sequence can be at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at

least 98%, at least 99%, or more, identical to a native or reference sequence. In some embodiments, a NR2F2 amino acid or DNA sequence can be at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or more, identical to a native or reference sequence.

[00109] In some embodiments, the transcription factors can be selected from the group consisting of ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, and NR2F2 or the corresponding zebrafish or *Xenopus* factors. In some embodiments, the transcription factors can be ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2 or the corresponding zebrafish or *Xenopus* factors. In some embodiments, the transcription factors can be ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, and NR2F2 or the corresponding zebrafish or *Xenopus* factors.

[00110] In some embodiments, the transcription factors can be ETV2, SOX7, and NR2F2 or the corresponding zebrafish or *Xenopus* factors. In some embodiments, the transcription factors can be ETS1, SOX7, and NR2F2 or the corresponding zebrafish or *Xenopus* factors. In some embodiments, the transcription factors can be ETV2 alone or the corresponding zebrafish or *Xenopus* factor.

[00111] In some embodiments, the transcription factors can be at least 1 factor, at least 2 factors, at least 3 factors, at least 4 factors, at least 5 factors, at least 6 factors, or at least 7 factors selected from the group consisting of ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, and NR2F2 or the corresponding zebrafish or *Xenopus* factors.

[00112] In some embodiments, transcription factors can be at ETV2 and at least 1 factor selected from the group consisting of FLI1, ETS1, SOX18, SOX7, RXRA, and NR2F2 or the corresponding zebrafish or *Xenopus* factors. In some embodiments, transcription factors can be at ETS1 and at least 1 factor selected from the group consisting of ETV2, FLI1, SOX18, SOX7, RXRA, and NR2F2 or the corresponding zebrafish or *Xenopus* factors.

[00113]

Hematopoietic System Development

[00114] The development of the haematopoietic system, including the cell populations and molecular pathways, is highly conserved between fish and mammals. HSPCs are born in the aorta-gonad-mesonephros (AGM) region and then migrate to a transient fetal niche, the fetal liver in mammals or a vascular plexus in the tail of the fish called the caudal haematopoietic

tissue (CHT). HSPCs reside and expand in these developmental sites for several days before migrating to the adult niche – the bone marrow in mammals or the kidney marrow in fish.

[00115] The CHT is comprised primarily of low-flow sinusoids surrounded by mesenchymal stromal cells. HSPCs initially colonize the CHT niche by lodging within the vascular plexus and interacting directly with *cxcl12a*$^{+}$ stromal cells. In a characteristic vascular remodeling step, endothelial cells (ECs) reorganize to form a supportive pocket around the HSPCs, which together with stromal cells and possibly other cell types, forms a niche for the stem cells (the endothelial cells surrounding the HSPCs can be referred to herein as endothelial niche cells). In mammals and zebrafish, specific signaling molecules, adhesion proteins and transcription factors have been implicated in mediating communication and physical interaction between stem cells and ECs in the niche. Collectively, these studies suggest that ECs within the vascular niches of haematopoietic organs express niche-specific gene programs. To date, however, a comprehensive investigation of the transcriptional circuitry that specifies the niche identity of ECs in the HSPC niche has not been undertaken. Understanding this regulation guides new strategies to improve the efficacy and availability of bone marrow transplantation therapies.

Endothelial Niche Cells

[00116] As described herein, endothelial niche cells are endothelial cells that provide an instructive niche for the differentiation of HSPCs. Endothelial niche cells are typically found in the bone marrow. However, as described herein, exogenous expression of specific transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, NR2F2) can cause endothelial niche cells to be found in non-bone marrow tissues, thus providing for extramedullary hematopoiesis.

[00117] In some embodiments of any of the aspects, endothelial niche cells comprise cells that express one are more genes selected from the group consisting of *sele*, *exoc312a*, *snx8a*, *cltca*, *aqp7*, *ap1b1*, *lgmn*, *prcp*, *cldn11a*, *lyve1b*, *adra1d*, *hya12a*, *hya12b*, *tll1*, *i113ra2*, *glu1a*, *hexb*, *slc16a9a*, and *sepp1a*. In some embodiments, the endothelial cells are human.

[00118] In some embodiments of any of the aspects, the endothelial niche cells are generated or engineered to express transcription factors, comprising at least one of the human transcription factors ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2. In some embodiments, the transcription factor comprises at least one transcription factor from the Ets family, at least one

transcription factor from the Sox family, and at least one transcription factor from the Nuclear Hormone Receptor family. In some embodiments, the transcription factors comprise ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, and NR2F2.

[00119] In some embodiments of any of the aspects, the transcription factors are expressed from at least one vector. In some embodiments, the vector comprises an exogenous nucleic acid sequence or sequences encoding the one or more transcription factors. In some embodiments, the exogenous nucleic acid sequences are incorporated into the genome of the endothelial cell. As a non-limiting example, the exogenous nucleic acid sequences can be incorporated into the genome using viral vectors (e.g., AAV, lentivirus) or CRISPR technologies.

[00120] One aspect provides for a composition comprising an engineered endothelial niche cell comprising one or more exogenous nucleic acid sequences encoding one or more transcription factors, wherein the one or more transcription factors are from the Ets family, the Sox family and/or the Nuclear Hormone Family. In some embodiments of any of the aspects, the composition can comprise engineered endothelial niche cells. In some embodiments, the composition is a therapeutic agent or the composition further comprises a pharmaceutically acceptable carrier. In some embodiments, the composition further comprises a culture dish, 3D cell system, or suspension system. In some embodiments, the composition comprises a scaffold.

[00121] Another aspect provides a method for culturing HSPCs, the method comprising culturing HSPCs in the presence of a population of engineered endothelial niche cells. In some embodiments of any of the aspects, the method is performed *in vitro*. In some embodiments, the engineered endothelial niche cells secrete a factor (e.g., growth factors) that affects the growth and/or expansion of the HSPC cells.

[00122] In some embodiments, the HSPCs cultured in the presence of the engineered endothelial niche cells can be cultured for at least 3 days longer than HSPCs that are cultured in the absence of such engineered endothelial niche cells. In, some embodiments, the HSPCs cultured in the presence of the engineered endothelial niche cells can be cultured for at 1 day longer, at least 2 days longer, at least 3 days longer, at least 4 days longer, at least 5 days longer, at least 6 days longer, at least 7 days longer, at least 8 days longer, at least 9 days longer, at least 10 days longer, at least 11 days longer, at least 12 days longer, at least 13 days longer, or at least 14 days longer than HSPCs that are cultured in the absence of such engineered endothelial niche cells.

[00123] In some embodiments, the cells are cultured on a biologically compatible scaffold. Non-limiting examples of a biologically compatible scaffold comprise: a hydrogel, biopolymers, or another biomaterial with the ability to grow cells *in vitro* in preparation for transplantation. In some embodiments, the HSPCs cultured in the presence of the engineered endothelial niche cells have increased engraftment when administered to a subject compared to the engraftment of substantially similar HSPCs that were not cultured with engineered endothelial niche cells. As used herein, "engraftment" refers to the process wherein transplanted HSPCs begin to grow and produce healthy blood cells. Engraftment is a critical milestone in recovery from an HSPC transplant.

[00124] Another aspect provides a method of treating a subject, the method comprising, transplanting a composition comprising a population of engineered endothelial niche-cells into the subject. As a non-limiting example, the method can be used to treat myelofibrosis or other hematopoietic diseases where the endogenous bone marrow niche is compromised, non-limiting examples of which are disclosed herein. In some embodiments, the method can comprise transplanting a composition comprising a population of HSPCs into the subject. In some embodiments, the method can comprise transplanting a composition comprising a population of HSPCs and engineered endothelial niche-cells into the subject.

[00125] Another aspect provides a method for enhancing engraftment of HSPCs, the method comprising administering a composition comprising HSPCs and a population of engineered endothelial niche cells to a subject in need thereof. In some embodiments of any of the aspects, engraftment of the HSPCs is increased by at least 10% compared to the engraftment of substantially similar HSPCs in the absence of engineered endothelial niche cells. In some embodiments of any of the aspects, engraftment of the HSPCs is increased by at least 1%, at least 2%, at least 3%, at least 4%, at least 5%, at least 6%, at least 7%, at least 8%, at least 9%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or at least 100% compared to the engraftment of substantially similar HSPCs in the absence of engineered endothelial niche cells.

[00126] Another aspect provides a co-culture comprising engineered endothelial niche cells and HSPCs. In some embodiments of any of the aspects, the endothelial cells are made by a method described herein.

[00127] Another aspect provides a method for generating an ectopic vascular niche, the method comprising: administering an engineered endothelial niche cell to a target site in a subject in need thereof. As used herein "ectopic vascular niche" refers to an atypical site for endothelial niche cells. For example, the vascular niche can be found outside of the bone marrow. The ectopic vascular niche comprising generated or engineered endothelial niche cells can be anywhere in the body. The ectopic vascular niche can be found in a location where HSPCs, generated or engineered endothelial niche cells, and/or their associated transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) have been injected.

[00128] Another aspect provides a method for extra medullary hematopoiesis, the method comprising transplanting engineered-niche endothelial cells into a subject at a location outside of the bone marrow (e.g., the forearm), thereby creating a synthetic niche. As used herein, "extra medullary hematopoiesis" refers to hematopoiesis occurring in organs outside of the bone marrow. In some embodiments of any of the aspects, the endothelial cells are made by a method described herein.

Myelofibrosis

[00129] In some embodiments of any of the aspects, generated or engineered endothelial niche cells or their associated transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) can be used to treat myelofibrosis.

[00130] Myelofibrosis is an uncommon type of chronic leukemia. Myelofibrosis belongs to a group of diseases called myeloproliferative disorders, often of a chronic form. Chronic myeloproliferative disorders are a group of slow-growing blood cancers in which the bone marrow makes too many abnormal red blood cells, white blood cells, or platelets, which accumulate in the blood. Non-limiting examples of chronic myeloproliferative neoplasms comprise: Chronic myelogenous leukemia, Polycythemia vera, Primary myelofibrosis (also called chronic idiopathic myelofibrosis), Essential thrombocythemia, Chronic neutrophilic leukemia, and Chronic eosinophilic leukemia.

[00131] Myelofibrosis is a serious bone marrow disorder that disrupts the body's normal production of blood cells. The result is extensive scarring in bone marrow, leading to severe anemia, weakness, fatigue and often an enlarged spleen. Many subjects or patients with myelofibrosis get progressively worse, and some subjects or patients may eventually develop a

more serious form of leukemia. Myelofibrosis can occur when blood stem cells (e.g., HSPCs) develop a genetic mutation. Several specific gene mutations have been identified in people with myelofibrosis. The most common is the Janus kinase 2 (JAK2) gene.

[00132] Although the cause of myelofibrosis often isn't known, certain factors are known to increase risk. Increased age can be associated with the development of myelofibrosis. Myelofibrosis can affect anyone, but it's most often diagnosed in people older than 50. Patients with another blood cell disorder are at higher risk for developing myelofibrosis. A small portion of people with myelofibrosis develop the condition as a complication of essential thrombocythemia or polycythemia vera. Exposure to certain chemicals can increase the risk for myelofibrosis. Myelofibrosis has been linked to exposure to industrial chemicals such as toluene and benzene. Exposure to radiation can increase the risk for myelofibrosis. People exposed to high levels of radiation, such as survivors of atomic bomb attacks, have an increased risk of myelofibrosis. Some people who received a radioactive contrast material called Thorotrast, used until the 1950s, have developed myelofibrosis.

[00133] Multiple complications can result from myelofibrosis. A complication of myelofibrosis can include increased pressure on blood flowing into a patient's liver. Normally, blood flow from the spleen enters the liver through a large blood vessel called the portal vein. Increased blood flow from an enlarged spleen can lead to high blood pressure in the portal vein (e.g., portal hypertension). This in turn can force excess blood into smaller veins in the stomach and esophagus, potentially causing these veins to rupture and bleed. Pain can be another complication of myelofibrosis. A severely enlarged spleen can cause abdominal pain and back pain. Myelofibrosis can lead to growths in other areas of the body. Myelofibrosis can be associated with bleeding complications. As the disease progresses, platelet count tends to drop below normal (thrombocytopenia) and platelet function becomes impaired. An insufficient number of platelets can lead to easy bleeding. Myelofibrosis can also be associated with painful bones and joints. Myelofibrosis can lead to hardening of bone marrow and inflammation of the connective tissue that is found around the bones. This may cause bone and joint pain. Myelofibrosis can also be associated with development of acute leukemia. Some patients with myelofibrosis develop acute myelogenous leukemia, a type of blood and bone marrow cancer that progresses rapidly.

[00134] Bone marrow transplantation is currently the only approved treatment for myelofibrosis. Additional treatments can only ameliorate the symptoms of myelofibrosis (e.g., anemia, enlarged spleen). Ruxolitinib, a JAK inhibitor which targets the gene mutation found in most cases of myelofibrosis, can be used to reduce symptoms of an enlarged spleen.

Treatment Methods

[00135] As described herein, levels of functional hematopoiesis can be decreased in myelofibrosis and/or in subjects with myelofibrosis. As used herein, "functional hematopoiesis" refers to hematopoiesis that produces normal levels and proportions of blood cells (e.g., red blood cells, white blood cells, platelets). In some embodiments of any of the aspects, the level of hematopoiesis can be decreased in myelofibrosis or a myeloproliferative disorder and/or in subjects with myelofibrosis or a myeloproliferative disorder. Accordingly, in one aspect of any of the embodiments, described herein is a method of treating myelofibrosis or a myeloproliferative disorder in a subject in need thereof, the method comprising administering HSPCs, engineered endothelial cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) to a subject determined to have a level of functional hematopoiesis that is decreased relative to a reference. In one aspect of any of the embodiments, described herein is a method of treating myelofibrosis or a myeloproliferative disorder in a subject in need thereof, the method comprising: a) determining the level of functional hematopoiesis in a sample obtained from a subject; and b) administering HSPCs, engineered endothelial cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) to the subject if the level of functional hematopoiesis is decreased relative to a reference.

[00136] In some embodiments of any of the aspects, the method comprises administering HSPCs, engineered endothelial cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) to a subject previously determined to have a level of functional hematopoiesis that is decreased relative to a reference. In some embodiments of any of the aspects, described herein is a method of treating myelofibrosis or a myeloproliferative disorder in a subject in need thereof, the method comprising: a) first determining the level of functional hematopoiesis in a sample obtained from a subject; and b) then administering HSPCs, engineered endothelial cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18,

SOX7, RXRA, or NR2F2) to the subject if the level of functional hematopoiesis is decreased relative to a reference.

[00137] In one aspect of any of the embodiments, described herein is a method of treating myelofibrosis or a myeloproliferative disorder in a subject in need thereof, the method comprising: a) determining if the subject has a decreased level of hematopoiesis; and b) administering HSPCs, engineered endothelial cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) to the subject if the level of functional hematopoiesis is decreased relative to a reference. In some embodiments of any of the aspects, the step of determining if the subject has a decreased level of functional hematopoiesis can comprise i) obtaining or having obtained a sample from the subject and ii) performing or having performed an assay on the sample obtained from the subject to determine/measure the level of hematopoiesis in the subject. In some embodiments of any of the aspects, the step of determining if the subject has a decreased level of functional hematopoiesis can comprise performing or having performed an assay on a sample obtained from the subject to determine/measure the level of hematopoiesis in the subject. In some embodiments of any of the aspects, the step of determining if the subject has a decreased level of functional hematopoiesis can comprise ordering or requesting an assay on a sample obtained from the subject to determine/measure the level of hematopoiesis in the subject. In some embodiments of any of the aspects, the step of determining if the subject has a decreased level of functional hematopoiesis can comprise receiving the results of an assay on a sample obtained from the subject to determine/measure the level of functional hematopoiesis in the subject. In some embodiments of any of the aspects, the step of determining if the subject has a decreased level of functional hematopoiesis can comprise receiving a report, results, or other means of identifying the subject as a subject with a decreased level of functional hematopoiesis.

[00138] In one aspect of any of the embodiments, described herein is a method of treating myelofibrosis or a myeloproliferative disorder in a subject in need thereof, the method comprising: a) determining if the subject has a decreased level of functional hematopoiesis; and b) instructing or directing that the subject be administered HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) if the level of functional hematopoiesis is decreased relative to a reference. In some embodiments of any of the aspects, the step of determining if the subject has a decreased level of functional

hematopoiesis can comprise i) obtaining or having obtained a sample from the subject and ii) performing or having performed an assay on the sample obtained from the subject to determine/measure the level of functional hematopoiesis in the subject. In some embodiments of any of the aspects, the step of determining if the subject has a decreased level of functional hematopoiesis can comprise performing or having performed an assay on a sample obtained from the subject to determine/measure the level of functional hematopoiesis in the subject. In some embodiments of any of the aspects, the step of determining if the subject has a functional level of functional hematopoiesis can comprise ordering or requesting an assay on a sample obtained from the subject to determine/measure the level of functional hematopoiesis in the subject. In some embodiments of any of the aspects, the step of instructing or directing that the subject be administered a particular treatment can comprise providing a report of the assay results. In some embodiments of any of the aspects, the step of instructing or directing that the subject be administered a particular treatment can comprise providing a report of the assay results and/or treatment recommendations in view of the assay results.

Administration

[00139] In some embodiments, the methods described herein relate to treating a subject having or diagnosed as having myelofibrosis or a myeloproliferative disorder. Subjects having myelofibrosis or a myeloproliferative disorder can be identified by a physician using current methods of diagnosing myelofibrosis or a myeloproliferative disorder. Symptoms and/or complications of myelofibrosis or a myeloproliferative disorder which characterize these conditions and aid in diagnosis are well known in the art and include but are not limited to anemia, splenomegaly (i.e. an enlarged and painful spleen), fatigue, weak or short of breath, pain or fullness below the ribs on the left side, easy bruising, easy bleeding, excessive sweating during sleep (night sweats), fever, and/or bone pain. Tests that may aid in a diagnosis of myelofibrosis or a myeloproliferative disorder include but are not limited to a blood test (e.g., a complete blood count) or a bone marrow biopsy. Myelofibrosis or a myeloproliferative disorder can also be detected with a physical exam, imaging tests, or genetic tests. A family history of myelofibrosis or a myeloproliferative disorder, or exposure to risk factors for myelofibrosis or a myeloproliferative disorder (e.g. industrial chemicals, radiation) can also aid in determining if a

subject is likely to have myelofibrosis or a myeloproliferative disorder or in making a diagnosis of myelofibrosis or a myeloproliferative disorder.

[00140] The compositions and methods described herein can be administered to a subject having or diagnosed as having myelofibrosis or a myeloproliferative disorder. In some embodiments, the methods described herein comprise administering an effective amount of compositions described herein, e.g. HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) to a subject in order to alleviate a symptom of myelofibrosis or a myeloproliferative disorder. As used herein, "alleviating a symptom of myelofibrosis or a myeloproliferative disorder" is ameliorating any condition or symptom associated with the myelofibrosis or a myeloproliferative disorder. As compared with an equivalent untreated control, such reduction is by at least 5%, 10%, 20%, 40%, 50%, 60%, 80%, 90%, 95%, 99% or more as measured by any standard technique. A variety of means for administering the compositions described herein to subjects are known to those of skill in the art. Such methods can include, but are not limited to oral, parenteral, intravenous, intramuscular, subcutaneous, transdermal, airway (aerosol), pulmonary, cutaneous, topical, injection, or intratumoral administration. Administration can be local or systemic.

[00141] The term “effective amount" as used herein refers to the amount of HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) needed to alleviate at least one or more symptom of the disease or disorder, and relates to a sufficient amount of pharmacological composition to provide the desired effect. The term "therapeutically effective amount" therefore refers to an amount of HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) that is sufficient to provide a particular anti-myelofibrosis or anti-myeloproliferative disorder effect when administered to a typical subject. An effective amount as used herein, in various contexts, would also include an amount sufficient to delay the development of a symptom of the disease, alter the course of a symptom disease (for example but not limited to, slowing the progression of a symptom of the disease), or reverse a symptom of the disease. Thus, it is not generally practicable to specify an exact “effective amount". However, for any given case, an appropriate “effective amount" can be determined by one of ordinary skill in the art using only routine experimentation.

[00142] Effective amounts, toxicity, and therapeutic efficacy can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dosage can vary depending upon the dosage form employed and the route of administration utilized. The dose ratio between toxic and therapeutic effects is the therapeutic index and can be expressed as the ratio LD50/ED50. Compositions and methods that exhibit large therapeutic indices are preferred. A therapeutically effective dose can be estimated initially from cell culture assays. Also, a dose can be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (*i.e.*, the concentration of HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) which achieves a half-maximal inhibition of symptoms) as determined in cell culture, or in an appropriate animal model. Levels in plasma can be measured, for example, by high performance liquid chromatography. The effects of any particular dosage can be monitored by a suitable bioassay. The dosage can be determined by a physician and adjusted, as necessary, to suit observed effects of the treatment.

[00143] In some embodiments, the technology described herein relates to a pharmaceutical composition comprising HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) as described herein, and optionally a pharmaceutically acceptable carrier. In some embodiments, the active ingredients of the pharmaceutical composition comprise HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) as described herein. In some embodiments, the active ingredients of the pharmaceutical composition consist essentially of HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) as described herein. In some embodiments, the active ingredients of the pharmaceutical composition consist of HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) as described herein.

[00144] Pharmaceutically acceptable carriers and diluents include saline, aqueous buffer solutions, solvents and/or dispersion media. The use of such carriers and diluents is well known in the art. Some non-limiting examples of materials which can serve as pharmaceutically-acceptable carriers include: (1) sugars, such as lactose, glucose and sucrose; (2) starches, such as

corn starch and potato starch; (3) cellulose, and its derivatives, such as sodium carboxymethyl cellulose, methylcellulose, ethyl cellulose, microcrystalline cellulose and cellulose acetate; (4) powdered tragacanth; (5) malt; (6) gelatin; (7) lubricating agents, such as magnesium stearate, sodium lauryl sulfate and talc; (8) excipients, such as cocoa butter and suppository waxes; (9) oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; (10) glycols, such as propylene glycol; (11) polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol (PEG); (12) esters, such as ethyl oleate and ethyl laurate; (13) agar; (14) buffering agents, such as magnesium hydroxide and aluminum hydroxide; (15) alginic acid; (16) pyrogen-free water; (17) isotonic saline; (18) Ringer's solution; (19) ethyl alcohol; (20) pH buffered solutions; (21) polyesters, polycarbonates and/or polyanhydrides; (22) bulking agents, such as polypeptides and amino acids (23) serum component, such as serum albumin, HDL and LDL; (22) C_2-C_{12} alcohols, such as ethanol; and (23) other non-toxic compatible substances employed in pharmaceutical formulations. Wetting agents, coloring agents, release agents, coating agents, sweetening agents, flavoring agents, perfuming agents, preservative and antioxidants can also be present in the formulation. The terms such as "excipient", "carrier", "pharmaceutically acceptable carrier" or the like are used interchangeably herein. In some embodiments, the carrier inhibits the degradation of the active agent, e.g. HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) as described herein.

[00145] In some embodiments, the pharmaceutical composition comprising HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) as described herein can be a parenteral dose form. Since administration of parenteral dosage forms typically bypasses the patient's natural defenses against contaminants, parenteral dosage forms are preferably sterile or capable of being sterilized prior to administration to a patient. Examples of parenteral dosage forms include, but are not limited to, solutions ready for injection, dry products ready to be dissolved or suspended in a pharmaceutically acceptable vehicle for injection, suspensions ready for injection, and emulsions. In addition, controlled-release parenteral dosage forms can be prepared for administration of a patient, including, but not limited to, DUROS®-type dosage forms and dose-dumping.

[00146] Suitable vehicles that can be used to provide parenteral dosage forms of HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) as disclosed within are well known to those skilled in the art. Examples include, without limitation: sterile water; water for injection USP; saline solution; glucose solution; aqueous vehicles such as but not limited to, sodium chloride injection, Ringer's injection, dextrose Injection, dextrose and sodium chloride injection, and lactated Ringer's injection; water-miscible vehicles such as, but not limited to, ethyl alcohol, polyethylene glycol, and propylene glycol; and non-aqueous vehicles such as, but not limited to, corn oil, cottonseed oil, peanut oil, sesame oil, ethyl oleate, isopropyl myristate, and benzyl benzoate.

[00147] In some embodiments of any of the aspects, the HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) described herein is administered as a monotherapy, e.g., another treatment for the myelofibrosis or a myeloproliferative disorder is not administered to the subject.

[00148] In some embodiments of any of the aspects, the methods described herein can further comprise administering a second agent and/or treatment to the subject, e.g. as part of a combinatorial therapy. Non-limiting examples of a second agent and/or treatment can include radiation therapy, surgery, gemcitabine, cisplastin, paclitaxel, carboplatin, bortezomib, AMG479, vorinostat, rituximab, temozolomide, rapamycin, ABT-737, PI-103; alkylating agents such as thiotepa and CYTOXAN® cyclosphosphamide; alkyl sulfonates such as busulfan, improsulfan and piposulfan; aziridines such as benzodopa, carboquone, meturedopa, and uredopa; ethylenimines and methylamelamines including altretamine, triethylenemelamine, trietylenephosphoramide, triethiylenethiophosphoramide and trimethylolomelamine; acetogenins (especially bullatacin and bullatacinone); a camptothecin (including the synthetic analogue topotecan); bryostatin; callystatin; CC-1065 (including its adozelesin, carzelesin and bizelesin synthetic analogues); cryptophycins (particularly cryptophycin 1 and cryptophycin 8); dolastatin; duocarmycin (including the synthetic analogues, KW-2189 and CB1-TM1); eleutherobin; pancratistatin; a sarcodictyin; spongistatin; nitrogen mustards such as chlorambucil, chlornaphazine, cholophosphamide, estramustine, ifosfamide, mechlorethamine, mechlorethamine oxide hydrochloride, melphalan, novembichin, phenesterine, prednimustine, trofosfamide, uracil mustard; nitrosureas such as carmustine, chlorozotocin, fotemustine, lomustine, nimustine, and ranimnustine; antibiotics such as the enediyne antibiotics (*e.g.*,

calicheamicin, especially calicheamicin gamma1I and calicheamicin omegaI1 (see, *e.g.*, Agnew, Chem. Intl. Ed. Engl., 33: 183-186 (1994)); dynemicin, including dynemicin A; bisphosphonates, such as clodronate; an esperamicin; as well as neocarzinostatin chromophore and related chromoprotein enediyne antiobiotic chromophores), aclacinomysins, actinomycin, authramycin, azaserine, bleomycins, cactinomycin, carabicin, caminomycin, carzinophilin, chromomycinis, dactinomycin, daunorubicin, detorubicin, 6-diazo-5-oxo-L-norleucine, ADRIAMYCIN® doxorubicin (including morpholino-doxorubicin, cyanomorpholino-doxorubicin, 2-pyrrolino-doxorubicin and deoxydoxorubicin), epirubicin, esorubicin, idarubicin, marcellomycin, mitomycins such as mitomycin C, mycophenolic acid, nogalamycin, olivomycins, peplomycin, potfiromycin, puromycin, quelamycin, rodorubicin, streptonigrin, streptozocin, tubercidin, ubenimex, zinostatin, zorubicin; anti-metabolites such as methotrexate and 5-fluorouracil (5-FU); folic acid analogues such as denopterin, methotrexate, pteropterin, trimetrexate; purine analogs such as fludarabine, 6-mercaptopurine, thiamiprine, thioguanine; pyrimidine analogs such as ancitabine, azacitidine, 6-azauridine, carmofur, cytarabine, dideoxyuridine, doxifluridine, enocitabine, floxuridine; androgens such as calusterone, dromostanolone propionate, epitiostanol, mepitiostane, testolactone; anti-adrenals such as aminoglutethimide, mitotane, trilostane; folic acid replenisher such as frolinic acid; aceglatone; aldophosphamide glycoside; aminolevulinic acid; eniluracil; amsacrine; bestrabucil; bisantrene; edatraxate; defofamine; demecolcine; diaziquone; elformithine; elliptinium acetate; an epothilone; etoglucid; gallium nitrate; hydroxyurea; lentinan; lonidainine; maytansinoids such as maytansine and ansamitocins; mitoguazone; mitoxantrone; mopidanmol; nitraerine; pentostatin; phenamet; pirarubicin; losoxantrone; podophyllinic acid; 2-ethylhydrazide; procarbazine; PSK® polysaccharide complex (JHS Natural Products, Eugene, Oreg.); razoxane; rhizoxin; sizofuran; spirogermanium; tenuazonic acid; triaziquone; 2,2',2''-trichlorotriethylamine; trichothecenes (especially T-2 toxin, verracurin A, roridin A and anguidine); urethan; vindesine; dacarbazine; mannomustine; mitobronitol; mitolactol; pipobroman; gacytosine; arabinoside ("Ara-C"); cyclophosphamide; thiotepa; taxoids, *e.g.*, TAXOL® paclitaxel (Bristol-Myers Squibb Oncology, Princeton, N.J.), ABRAXANE® Cremophor-free, albumin-engineered nanoparticle formulation of paclitaxel (American Pharmaceutical Partners, Schaumberg, Ill.), and TAXOTERE® doxetaxel (Rhone-Poulenc Rorer, Antony, France); chloranbucil; GEMZAR® gemcitabine; 6-thioguanine;

mercaptopurine; methotrexate; platinum analogs such as cisplatin, oxaliplatin and carboplatin; vinblastine; platinum; etoposide (VP-16); ifosfamide; mitoxantrone; vincristine; NAVELBINE.RTM. vinorelbine; novantrone; teniposide; edatrexate; daunomycin; aminopterin; xeloda; ibandronate; irinotecan (Camptosar, CPT-11) (including the treatment regimen of irinotecan with 5-FU and leucovorin); topoisomerase inhibitor RFS 2000; difluoromethylornithine (DMFO); retinoids such as retinoic acid; capecitabine; combretastatin; leucovorin (LV); oxaliplatin, including the oxaliplatin treatment regimen (FOLFOX); lapatinib (Tykerb.RTM.); inhibitors of PKC-alpha, Raf, H-Ras, EGFR (*e.g.*, erlotinib (Tarceva®)) and VEGF-A that reduce cell proliferation and pharmaceutically acceptable salts, acids or derivatives of any of the above.

[00149] In addition, the methods of treatment can further include the use of radiation or radiation therapy. Further, the methods of treatment can further include the use of surgical treatments.

[00150] The methods described herein can further comprise administering a second agent and/or treatment to the subject, e.g. as part of a combinatorial therapy. By way of non-limiting example, if a subject is to be treated for pain or inflammation according to the methods described herein, the subject can also be administered a second agent and/or treatment known to be beneficial for subjects suffering from pain or inflammation. In some embodiments, the second agent is an anti-inflammation agent. Examples of such agents and/or treatments include, but are not limited to, non-steroidal anti-inflammatory drugs (NSAIDs - such as aspirin, ibuprofen, or naproxen); corticosteroids, including glucocorticoids (e.g. cortisol, prednisone, prednisolone, methylprednisolone, dexamethasone, betamethasone, triamcinolone, and beclometasone); methotrexate; sulfasalazine; leflunomide; anti-TNF medications; cyclophosphamide; pro-resolving drugs; mycophenolate; or opiates (e.g. endorphins, enkephalins, and dynorphin), steroids, analgesics, barbiturates, oxycodone, morphine, lidocaine, and the like.

[00151] In certain embodiments, an effective dose of a composition comprising HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) as described herein can be administered to a patient once. In certain embodiments, an effective dose of a composition comprising HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) can be administered to a patient repeatedly. For systemic administration, subjects can be

administered a therapeutic amount of a composition comprising HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2), such as, e.g. 0.1 mg/kg, 0.5 mg/kg, 1.0 mg/kg, 2.0 mg/kg, 2.5 mg/kg, 5 mg/kg, 10 mg/kg, 15 mg/kg, 20 mg/kg, 25 mg/kg, 30 mg/kg, 40 mg/kg, 50 mg/kg, or more.

[00152] In some embodiments, after an initial treatment regimen, the treatments can be administered on a less frequent basis. For example, after treatment biweekly for three months, treatment can be repeated once per month, for six months or a year or longer. Treatment according to the methods described herein can reduce levels of a marker or symptom of a condition, e.g. HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) by at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80 % or at least 90% or more.

[00153] The dosage of a composition as described herein can be determined by a physician and adjusted, as necessary, to suit observed effects of the treatment. With respect to duration and frequency of treatment, it is typical for skilled clinicians to monitor subjects in order to determine when the treatment is providing therapeutic benefit, and to determine whether to increase or decrease dosage, increase or decrease administration frequency, discontinue treatment, resume treatment, or make other alterations to the treatment regimen. The dosing schedule can vary from once a week to daily depending on a number of clinical factors, such as the subject's sensitivity to HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2). The desired dose or amount of activation can be administered at one time or divided into subdoses, e.g., 2-4 subdoses and administered over a period of time, e.g., at appropriate intervals through the day or other appropriate schedule. In some embodiments, administration can be chronic, e.g., one or more doses and/or treatments daily over a period of weeks or months. Examples of dosing and/or treatment schedules are administration daily, twice daily, three times daily or four or more times daily over a period of 1 week, 2 weeks, 3 weeks, 4 weeks, 1 month, 2 months, 3 months, 4 months, 5 months, or 6 months, or more. A composition comprising HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) can be administered over a period of time, such as over a 5 minute, 10 minute, 15 minute, 20 minute, or 25 minute period.

[00154] The dosage ranges for the administration of HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2), according to the methods described herein depend upon, for example, its form, its potency, and the extent to which symptoms, markers, or indicators of a condition described herein are desired to be reduced, for example the extent to which, for example, myelofibrosis or a myeloproliferative disorder is desired to be reduced functional hematopoiesis is desired to be induced. The dosage should not be so large as to cause adverse side effects, such as excessive hematopoiesis or excessive extramedullary hematopoiesis. Generally, the dosage will vary with the age, condition, and sex of the patient and can be determined by one of skill in the art. The dosage can also be adjusted by the individual physician in the event of any complication.

[00155] The efficacy of HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) in, e.g. the treatment of a condition described herein, or to induce a response as described herein (e.g. functional hematopoiesis) can be determined by the skilled clinician. However, a treatment is considered "effective treatment," as the term is used herein, if one or more of the signs or symptoms of a condition described herein are altered in a beneficial manner, other clinically accepted symptoms are improved, or even ameliorated, or a desired response is induced e.g., by at least 10% following treatment according to the methods described herein. Efficacy can be assessed, for example, by measuring a marker, indicator, symptom, and/or the incidence of a condition treated according to the methods described herein or any other measurable parameter appropriate, e.g. blood cell counts. Efficacy can also be measured by a failure of an individual to worsen as assessed by hospitalization, or need for medical interventions (i.e., progression of the disease is halted). Methods of measuring these indicators are known to those of skill in the art and/or are described herein. Treatment includes any treatment of a disease in an individual or an animal (some non-limiting examples include a human or an animal) and includes: (1) inhibiting the disease, e.g., preventing a worsening of symptoms (e.g. pain or inflammation); or (2) relieving the severity of the disease, e.g., causing regression of symptoms. An effective amount for the treatment of a disease means that amount which, when administered to a subject in need thereof, is sufficient to result in effective treatment as that term is defined herein, for that disease. Efficacy of an agent can be determined by assessing physical indicators of a condition or desired response. It is well within the ability of one skilled in the art to monitor efficacy of administration and/or treatment

by measuring any one of such parameters, or any combination of parameters. Efficacy can be assessed in animal models of a condition described herein, for example treatment of myelofibrosis or a myeloproliferative disorder. When using an experimental animal model, efficacy of treatment is evidenced when a statistically significant change in a marker is observed.

Kits

[00156] One aspect described herein provides a kit for culturing HSPCs, the kit comprising: a population of engineered endothelial niche cells, reagents and instructions for use thereof. Another aspect provides for a kit for generating engineered endothelial niche cells comprising: a vector(s) comprising one or more exogenous nucleic acid sequences encoding one or more transcription factors of the Ets family, the Sox family or the nuclear hormone family and instructions for use thereof. Described herein are kit components that can be included in one or more of the kits described herein.

[00157] In some embodiments, the kit comprises an effective amount of reagents for culturing HSPCs and/or endothelial niche cells. As will be appreciated by one of skill in the art, reagents can be supplied in a lyophilized form or a concentrated form that can diluted prior to use with cultured cells. Preferred formulations include those that are non-toxic to the cells and/or does not affect growth rate or viability etc. reagents can be supplied in aliquots or in unit doses.

[00158] In some embodiments the kit further comprises a vector comprising a nucleic acid encoding a gene to one or more transcription factors of the Ets family, the Sox family or the nuclear hormone family under the control of a promoter.

[00159] In some embodiments, the components described herein can be provided singularly or in any combination as a kit. The kit includes the components described herein, e.g., a composition comprising HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2), a composition(s) that includes a vector comprising e.g., a gene to one or more transcription factors of the Ets family, the Sox family or the nuclear hormone family under the control of a promoter as described throughout the specification. In addition, the kit optionally comprises informational material. The kit can also contain culture dishes and/or a substrate for coating culture dishes, such as laminin, fibronectin, Poly-L-Lysine, or methylcellulose.

[00160] In some embodiments, the compositions in the kit can be provided in a watertight or gas tight container which in some embodiments is substantially free of other components of the kit. For example, a HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) composition can be supplied in more than one container, e.g., it can be supplied in a container having sufficient reagent for a predetermined number of experiments, e.g., 1, 2, 3 or greater. One or more components as described herein can be provided in any form, e.g., liquid, dried or lyophilized form. It is preferred that the components described herein are substantially pure and/or sterile. When the components described herein are provided in a liquid solution, the liquid solution preferably is an aqueous solution, with a sterile aqueous solution being preferred.

[00161] The informational material can be descriptive, instructional, marketing or other material that relates to the methods described herein. The informational material of the kits is not limited in its form. In one embodiment, the informational material can include information about production of endothelial niche cells and/or HSPCs, concentration, date of expiration, batch or production site information, and so forth. In one embodiment, the informational material relates to methods for using or administering the components of the kit.

[00162] The kit can include a component for the detection of a marker for HSPC differentiation and/or endothelial niche cell differentiation. In addition, the kit can include one or more antibodies that bind a cell marker, or primers for an RT-PCR or PCR reaction, e.g., a semi-quantitative or quantitative RT-PCR or PCR reaction. Such components can be used to assess the activation of maturation markers or the loss of immature cell markers of endothelial niche cells and/or HSPCs. If the detection reagent is an antibody, it can be supplied in dry preparation, e.g., lyophilized, or in a solution. The antibody or other detection reagent can be linked to a label, e.g., a radiological, fluorescent (e.g., GFP) or colorimetric label for use in detection. If the detection reagent is a primer, it can be supplied in dry preparation, e.g., lyophilized, or in a solution.

[00163] The kit will typically be provided with its various elements included in one package, e.g., a fiber-based, e.g., a cardboard, or polymeric, e.g., a Styrofoam box. The enclosure can be configured so as to maintain a temperature differential between the interior and the exterior, e.g., it can provide insulating properties to keep the reagents at a preselected temperature for a preselected time.

Vectors

[0001] In some embodiments, one or more of the factors described herein is expressed in a recombinant expression vector or plasmid. As used herein, the term "vector" refers to a polynucleotide sequence suitable for transferring transgenes into a host cell. The term “vector” includes plasmids, mini-chromosomes, phage, naked DNA and the like. See, for example, U.S. Pat. Nos. 4,980,285; 5,631,150; 5,707,828; 5,759,828; 5,888,783 and, 5,919,670, and, Sambrook et al, *Molecular Cloning: A Laboratory Manual,* 2nd Ed., Cold Spring Harbor Press (1989). One type of vector is a "plasmid," which refers to a circular double stranded DNA loop into which additional DNA segments are ligated. Another type of vector is a viral vector, wherein additional DNA segments are ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" is used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

[0002] A cloning vector is one which is able to replicate autonomously or integrated in the genome in a host cell, and which is further characterized by one or more endonuclease restriction sites at which the vector may be cut in a determinable fashion and into which a desired DNA sequence can be ligated such that the new recombinant vector retains its ability to replicate in the host cell. In the case of plasmids, replication of the desired sequence can occur many times as the plasmid increases in copy number within the host cell such as a host bacterium or just a single time per host before the host reproduces by mitosis. In the case of phage, replication can occur actively during a lytic phase or passively during a lysogenic phase.

[0003] An expression vector is one into which a desired DNA sequence can be inserted by restriction and ligation such that it is operably joined to regulatory sequences and can be expressed as an RNA transcript. Vectors can further contain one or more marker sequences

suitable for use in the identification of cells which have or have not been transformed or transformed or transfected with the vector. Markers include, for example, genes encoding proteins which increase or decrease either resistance or sensitivity to antibiotics or other compounds, genes which encode enzymes whose activities are detectable by standard assays known in the art (e.g., β-galactosidase, luciferase or alkaline phosphatase), and genes which visibly affect the phenotype of transformed or transfected cells, hosts, colonies or plaques (e.g., green fluorescent protein). In certain embodiments, the vectors used herein are capable of autonomous replication and expression of the structural gene products present in the DNA segments to which they are operably joined.

[0004] As used herein, a coding sequence and regulatory sequences are said to be "operably" joined when they are covalently linked in such a way as to place the expression or transcription of the coding sequence under the influence or control of the regulatory sequences. If it is desired that the coding sequences be translated into a functional protein, two DNA sequences are said to be operably joined if induction of a promoter in the 5′ regulatory sequences results in the transcription of the coding sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region to direct the transcription of the coding sequences, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a promoter region would be operably joined to a coding sequence if the promoter region were capable of effecting transcription of that DNA sequence such that the resulting transcript can be translated into the desired protein or polypeptide.

[0005] When the nucleic acid molecule that encodes any of the factors/polypeptides described herein is expressed in a cell, a variety of transcription control sequences (e.g., promoter/enhancer sequences) can be used to direct its expression. The promoter can be a native promoter, i.e., the promoter of the gene in its endogenous context, which provides normal regulation of expression of the gene. In some embodiments the promoter can be constitutive, i.e., the promoter is unregulated allowing for continual transcription of its associated gene. A variety of conditional promoters also can be used, such as promoters controlled by the presence or absence of a molecule.

[0006] The precise nature of the regulatory sequences needed for gene expression can vary between species or cell types, but in general can include, as necessary, 5′ non-transcribed and 5′

non-translated sequences involved with the initiation of transcription and translation respectively, such as a TATA box, capping sequence, CAAT sequence, and the like. In particular, such 5′ non-transcribed regulatory sequences will include a promoter region which includes a promoter sequence for transcriptional control of the operably joined gene. Regulatory sequences can also include enhancer sequences or upstream activator sequences as desired. The vectors of the invention may optionally include 5′ leader or signal sequences. The choice and design of an appropriate vector is within the ability and discretion of one of ordinary skill in the art.

[0007] Expression vectors containing all the necessary elements for expression are commercially available and known to those skilled in the art. See, e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual,* Second Edition, Cold Spring Harbor Laboratory Press, 1989. Cells are genetically engineered by the introduction into the cells of heterologous DNA (RNA). That heterologous DNA (RNA) is placed under operable control of transcriptional elements to permit the expression of the heterologous DNA in the host cell.

[0008] In some embodiments, the vector is pME Gateway vector (Invitrogen™). In some embodiments, the vector is p5E Gateway™ vector. In some other embodiments, the vector is pGEX2TK™ vector. In some other embodiments, the vector is TOPO-TA™ vector.

[0009] Without limitations, the genes described herein can be included in one vector or separate vectors. For example, at least one gene from the Ets family (e.g., ETV2, FLI1, ETS1) and at least one gene from the SOX family (e.g., SOX18, SOX7) and at least one gene from the NHR family (e.g., RXRA, NR2F2) can be included in the same vector.

[0010] In some embodiments, at least one gene from the Ets family (e.g., ETV2, FLI1, ETS1) and at least one gene from the SOX family (e.g., SOX18, SOX7) can be included in a first vector, and at least one gene from the NHR family (e.g., RXRA, NR2F2) can be included in a second vector.

[0011] In some embodiments, at least one gene from the NHR family (e.g., RXRA, NR2F2) and at least one gene from the SOX family (e.g., SOX18, SOX7) can be included in a first vector, and at least one gene from the Ets family (e.g., ETV2, FLI1, ETS1) can be included in a second vector.

[0012] In some embodiments, at least one gene from the Ets family (e.g., ETV2, FLI1, ETS1) and at least one gene from the NHR family (e.g., RXRA, NR2F2) can be included in a

first vector, and at least one gene from the SOX family (e.g., SOX18, SOX7) can be included in a second vector.

[0013] In some embodiments, at least one gene from the Ets family (e.g., ETV2, FLI1, ETS1) can be included in a first vector, at least one gene from the SOX family (e.g., SOX18, SOX7) can be included in a second vector, and at least one gene from the NHR family (e.g., RXRA, NR2F2) can be included in a third vector.

[0014] In some embodiments, the promoter operably linked to the gene(s) can be zebrafish *ubi* promoter.

[0015] In some embodiments, one or more of the recombinantly expressed gene can be integrated into the genome of the cell.

[0016] A nucleic acid molecule that encodes the enzyme of the claimed invention can be introduced into a cell or cells using methods and techniques that are standard in the art. For example, nucleic acid molecules can be introduced by standard protocols such as transformation including chemical transformation and electroporation, transduction, particle bombardment, etc. Expressing the nucleic acid molecule encoding the enzymes of the claimed invention also may be accomplished by integrating the nucleic acid molecule into the genome.

Definitions

[00164] For convenience, the meaning of some terms and phrases used in the specification, examples, and appended claims, are provided below. Unless stated otherwise, or implicit from context, the following terms and phrases include the meanings provided below. The definitions are provided to aid in describing particular embodiments, and are not intended to limit the claimed invention, because the scope of the invention is limited only by the claims. Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. If there is an apparent discrepancy between the usage of a term in the art and its definition provided herein, the definition provided within the specification shall prevail.

[00165] For convenience, certain terms employed herein, in the specification, examples and appended claims are collected here.

[00166] The terms “decrease”, “reduced”, “reduction”, or “inhibit” are all used herein to mean a decrease by a statistically significant amount. In some embodiments, “reduce,” “reduction" or

"decrease" or "inhibit" typically means a decrease by at least 10% as compared to a reference level (e.g. the absence of a given treatment or agent) and can include, for example, a decrease by at least about 10%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 98%, at least about 99% , or more. As used herein, "reduction" or "inhibition" does not encompass a complete inhibition or reduction as compared to a reference level. "Complete inhibition" is a 100% inhibition as compared to a reference level. A decrease can be preferably down to a level accepted as within the range of normal for an individual without a given disorder.

[00167] The terms "increased", "increase", "enhance", or "activate" are all used herein to mean an increase by a statically significant amount. In some embodiments, the terms "increased", "increase", "enhance", or "activate" can mean an increase of at least 10% as compared to a reference level, for example an increase of at least about 20%, or at least about 30%, or at least about 40%, or at least about 50%, or at least about 60%, or at least about 70%, or at least about 80%, or at least about 90% or up to and including a 100% increase or any increase between 10-100% as compared to a reference level, or at least about a 2-fold, or at least about a 3-fold, or at least about a 4-fold, or at least about a 5-fold or at least about a 10-fold increase, or any increase between 2-fold and 10-fold or greater as compared to a reference level. In the context of a marker or symptom, a "increase" is a statistically significant increase in such level.

[00168] As used herein, a "subject" means a human or animal. Usually the animal is a vertebrate such as a primate, rodent, domestic animal or game animal. Primates include chimpanzees, cynomologous monkeys, spider monkeys, and macaques, e.g., Rhesus. Rodents include mice, rats, woodchucks, ferrets, rabbits and hamsters. Domestic and game animals include cows, horses, pigs, deer, bison, buffalo, feline species, e.g., domestic cat, canine species, e.g., dog, fox, wolf, avian species, e.g., chicken, emu, ostrich, and fish, e.g., trout, catfish and salmon. In some embodiments, the subject is a mammal, e.g., a primate, e.g., a human. The terms, "individual," "patient" and "subject" are used interchangeably herein.

[00169] Preferably, the subject is a mammal. The mammal can be a human, non-human primate, mouse, rat, dog, cat, horse, or cow, but is not limited to these examples. Mammals

other than humans can be advantageously used as subjects that represent animal models of myelofibrosis or a myeloproliferative disorder. A subject can be male or female.

[00170] A subject can be one who has been previously diagnosed with or identified as suffering from or having a condition in need of treatment (e.g. myelofibrosis or a myeloproliferative disorder) or one or more complications related to such a condition, and optionally, have already undergone treatment for myelofibrosis or a myeloproliferative disorder or the one or more complications related to myelofibrosis or a myeloproliferative disorder. Alternatively, a subject can also be one who has not been previously diagnosed as having myelofibrosis or a myeloproliferative disorder or one or more complications related to myelofibrosis or a myeloproliferative disorder. For example, a subject can be one who exhibits one or more risk factors for myelofibrosis or a myeloproliferative disorder or one or more complications related to myelofibrosis or a myeloproliferative disorder or a subject who does not exhibit risk factors.

[00171] A "subject in need" of treatment for a particular condition can be a subject having that condition, diagnosed as having that condition, or at risk of developing that condition.

[00172] As used herein, the terms "protein" and "polypeptide" are used interchangeably herein to designate a series of amino acid residues, connected to each other by peptide bonds between the alpha-amino and carboxy groups of adjacent residues. The terms "protein", and "polypeptide" refer to a polymer of amino acids, including modified amino acids (e.g., phosphorylated, glycated, glycosylated, etc.) and amino acid analogs, regardless of its size or function. "Protein" and "polypeptide" are often used in reference to relatively large polypeptides, whereas the term "peptide" is often used in reference to small polypeptides, but usage of these terms in the art overlaps. The terms "protein" and "polypeptide" are used interchangeably herein when referring to a gene product and fragments thereof. Thus, exemplary polypeptides or proteins include gene products, naturally occurring proteins, homologs, orthologs, paralogs, fragments and other equivalents, variants, fragments, and analogs of the foregoing.

[00173] In the various embodiments described herein, it is further contemplated that variants (naturally occurring or otherwise), alleles, homologs, conservatively modified variants, and/or conservative substitution variants of any of the particular polypeptides described are encompassed. As to amino acid sequences, one of skill will recognize that individual

substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid and retains the desired activity of the polypeptide. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles consistent with the disclosure.

[00174] A given amino acid can be replaced by a residue having similar physiochemical characteristics, e.g., substituting one aliphatic residue for another (such as Ile, Val, Leu, or Ala for one another), or substitution of one polar residue for another (such as between Lys and Arg; Glu and Asp; or Gln and Asn). Other such conservative substitutions, e.g., substitutions of entire regions having similar hydrophobicity characteristics, are well known. Polypeptides comprising conservative amino acid substitutions can be tested in any one of the assays described herein to confirm that a desired activity, e.g. transcription factor activity and specificity of a native or reference polypeptide is retained.

[00175] Amino acids can be grouped according to similarities in the properties of their side chains (in A. L. Lehninger, in Biochemistry, second ed., pp. 73-75, Worth Publishers, New York (1975)): (1) non-polar: Ala (A), Val (V), Leu (L), Ile (I), Pro (P), Phe (F), Trp (W), Met (M); (2) uncharged polar: Gly (G), Ser (S), Thr (T), Cys (C), Tyr (Y), Asn (N), Gln (Q); (3) acidic: Asp (D), Glu (E); (4) basic: Lys (K), Arg (R), His (H). Alternatively, naturally occurring residues can be divided into groups based on common side-chain properties: (1) hydrophobic: Norleucine, Met, Ala, Val, Leu, Ile; (2) neutral hydrophilic: Cys, Ser, Thr, Asn, Gln; (3) acidic: Asp, Glu; (4) basic: His, Lys, Arg; (5) residues that influence chain orientation: Gly, Pro; (6) aromatic: Trp, Tyr, Phe. Non-conservative substitutions will entail exchanging a member of one of these classes for another class. Particular conservative substitutions include, for example; Ala into Gly or into Ser; Arg into Lys; Asn into Gln or into His; Asp into Glu; Cys into Ser; Gln into Asn; Glu into Asp; Gly into Ala or into Pro; His into Asn or into Gln; Ile into Leu or into Val; Leu into Ile or into Val; Lys into Arg, into Gln or into Glu; Met into Leu, into Tyr or into Ile; Phe into Met, into Leu or into Tyr; Ser into Thr; Thr into Ser; Trp into Tyr; Tyr into Trp; and/or Phe into Val, into Ile or into Leu.

[00176] In some embodiments, the polypeptide described herein (or a nucleic acid encoding such a polypeptide) can be a functional fragment of one of the amino acid sequences described

herein. As used herein, a "functional fragment" is a fragment or segment of a peptide which retains at least 50% of the wildtype reference polypeptide's activity according to the assays described below herein. A functional fragment can comprise conservative substitutions of the sequences disclosed herein.

[00177] In some embodiments, the polypeptide described herein can be a variant of a sequence described herein. In some embodiments, the variant is a conservatively modified variant. Conservative substitution variants can be obtained by mutations of native nucleotide sequences, for example. A "variant," as referred to herein, is a polypeptide substantially homologous to a native or reference polypeptide, but which has an amino acid sequence different from that of the native or reference polypeptide because of one or a plurality of deletions, insertions or substitutions. Variant polypeptide-encoding DNA sequences encompass sequences that comprise one or more additions, deletions, or substitutions of nucleotides when compared to a native or reference DNA sequence, but that encode a variant protein or fragment thereof that retains activity. A wide variety of PCR-based site-specific mutagenesis approaches are known in the art and can be applied by the ordinarily skilled artisan.

[00178] A variant amino acid or DNA sequence can be at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or more, identical to a native or reference sequence. The degree of homology (percent identity) between a native and a mutant sequence can be determined, for example, by comparing the two sequences using freely available computer programs commonly employed for this purpose on the world wide web (e.g. BLASTp or BLASTn with default settings).

[00179] Alterations of the native amino acid sequence can be accomplished by any of a number of techniques known to one of skill in the art. Mutations can be introduced, for example, at particular loci by synthesizing oligonucleotides containing a mutant sequence, flanked by restriction sites enabling ligation to fragments of the native sequence. Following ligation, the resulting reconstructed sequence encodes an analog having the desired amino acid insertion, substitution, or deletion. Alternatively, oligonucleotide-directed site-specific mutagenesis procedures can be employed to provide an altered nucleotide sequence having particular codons altered according to the substitution, deletion, or insertion required. Techniques for making such alterations are very well established and include, for example, those disclosed by Walder et al. (Gene 42:133, 1986); Bauer et al. (Gene 37:73, 1985); Craik (BioTechniques, January 1985, 12-

19); Smith et al. (Genetic Engineering: Principles and Methods, Plenum Press, 1981); and U.S. Pat. Nos. 4,518,584 and 4,737,462, which are herein incorporated by reference in their entireties. Any cysteine residue not involved in maintaining the proper conformation of the polypeptide also can be substituted, generally with serine, to improve the oxidative stability of the molecule and prevent aberrant crosslinking. Conversely, cysteine bond(s) can be added to the polypeptide to improve its stability or facilitate oligomerization.

[00180] As used herein, the term "nucleic acid" or "nucleic acid sequence" refers to any molecule, preferably a polymeric molecule, incorporating units of ribonucleic acid, deoxyribonucleic acid or an analog thereof. The nucleic acid can be either single-stranded or double-stranded. A single-stranded nucleic acid can be one nucleic acid strand of a denatured double- stranded DNA. Alternatively, it can be a single-stranded nucleic acid not derived from any double-stranded DNA. In one aspect, the nucleic acid can be DNA. In another aspect, the nucleic acid can be RNA. Suitable DNA can include, e.g., genomic DNA or cDNA. Suitable RNA can include, e.g., mRNA.

[00181] The term "expression" refers to the cellular processes involved in producing RNA and proteins and as appropriate, secreting proteins, including where applicable, but not limited to, for example, transcription, transcript processing, translation and protein folding, modification and processing. Expression can refer to the transcription and stable accumulation of sense (mRNA) or antisense RNA derived from a nucleic acid fragment or fragments of the invention and/or to the translation of mRNA into a polypeptide.

[00182] In some embodiments, the expression of a biomarker(s), target(s), or gene/polypeptide described herein is/are tissue-specific. In some embodiments, the expression of a biomarker(s), target(s), or gene/polypeptide described herein is/are global. In some embodiments, the expression of a biomarker(s), target(s), or gene/polypeptide described herein is systemic.

[00183] "Expression products" include RNA transcribed from a gene, and polypeptides obtained by translation of mRNA transcribed from a gene. The term "gene" means the nucleic acid sequence which is transcribed (DNA) to RNA in vitro or in vivo when operably linked to appropriate regulatory sequences. The gene may or may not include regions preceding and following the coding region, e.g. 5' untranslated (5'UTR) or "leader" sequences and 3' UTR or

"trailer" sequences, as well as intervening sequences (introns) between individual coding segments (exons).

[00184] "Marker" in the context of the present invention refers to an expression product, e.g., nucleic acid or polypeptide which is differentially present in a sample taken from subjects having myelofibrosis or a myeloproliferative disorder, as compared to a comparable sample taken from control subjects (e.g., a healthy subject). The term "biomarker" is used interchangeably with the term "marker."

[00185] In some embodiments, the methods described herein relate to measuring, detecting, or determining the level of at least one marker. As used herein, the term "detecting" or "measuring" refers to observing a signal from, e.g. a probe, label, or target molecule to indicate the presence of an analyte in a sample. Any method known in the art for detecting a particular label moiety can be used for detection. Exemplary detection methods include, but are not limited to, spectroscopic, fluorescent, photochemical, biochemical, immunochemical, electrical, optical or chemical methods. In some embodiments of any of the aspects, measuring can be a quantitative observation.

[00186] In some embodiments of any of the aspects, a polypeptide, nucleic acid, or cell as described herein can be engineered. As used herein, "engineered" refers to the aspect of having been manipulated by the hand of man. For example, a polypeptide is considered to be "engineered" when at least one aspect of the polypeptide, e.g., its sequence, has been manipulated by the hand of man to differ from the aspect as it exists in nature. As is common practice and is understood by those in the art, progeny of an engineered cell are typically still referred to as "engineered" even though the actual manipulation was performed on a prior entity.

[00187] In some embodiments of any of the aspects, the HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) described herein is exogenous. In some embodiments of any of the aspects, the HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) described herein is ectopic. In some embodiments of any of the aspects, the HSPCs, engineered endothelial niche cells, and/or transcription factors (e.g., ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2) described herein is not endogenous.

[00188] The term "exogenous" refers to a substance present in a cell other than its native source. The term "exogenous" when used herein can refer to a nucleic acid (e.g. a nucleic acid

encoding a polypeptide) or a polypeptide that has been introduced by a process involving the hand of man into a biological system such as a cell or organism in which it is not normally found and one wishes to introduce the nucleic acid or polypeptide into such a cell or organism. Alternatively, “exogenous” can refer to a nucleic acid or a polypeptide that has been introduced by a process involving the hand of man into a biological system such as a cell or organism in which it is found in relatively low amounts and one wishes to increase the amount of the nucleic acid or polypeptide in the cell or organism, e.g., to create ectopic expression or levels. In contrast, the term "endogenous" refers to a substance that is native to the biological system or cell. As used herein, “ectopic” refers to a substance that is found in an unusual location and/or amount. An ectopic substance can be one that is normally found in a given cell, but at a much lower amount and/or at a different time. Ectopic also includes substance, such as a polypeptide or nucleic acid that is not naturally found or expressed in a given cell in its natural environment.

[00189] In some embodiments, a nucleic acid encoding a polypeptide as described herein (e.g. a ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2 polypeptide) is comprised by a vector. In some of the aspects described herein, a nucleic acid sequence encoding a given polypeptide as described herein, or any module thereof, is operably linked to a vector. The term "vector", as used herein, refers to a nucleic acid construct designed for delivery to a host cell or for transfer between different host cells. As used herein, a vector can be viral or non-viral. The term “vector” encompasses any genetic element that is capable of replication when associated with the proper control elements and that can transfer gene sequences to cells. A vector can include, but is not limited to, a cloning vector, an expression vector, a plasmid, phage, transposon, cosmid, chromosome, virus, virion, etc.

[00190] In some embodiments of any of the aspects, the vector is recombinant, e.g., it comprises sequences originating from at least two different sources. In some embodiments of any of the aspects, the vector comprises sequences originating from at least two different species. In some embodiments of any of the aspects, the vector comprises sequences originating from at least two different genes, e.g., it comprises a fusion protein or a nucleic acid encoding an expression product which is operably linked to at least one non-native (e.g., heterologous) genetic control element (e.g., a promoter, suppressor, activator, enhancer, response element, or the like).

[00191] In some embodiments of any of the aspects, the vector or nucleic acid described herein is codon-optimized, e.g., the native or wild-type sequence of the nucleic acid sequence has been altered or engineered to include alternative codons such that altered or engineered nucleic acid encodes the same polypeptide expression product as the native/wild-type sequence, but will be transcribed and/or translated at an improved efficiency in a desired expression system. In some embodiments of any of the aspects, the expression system is an organism other than the source of the native/wild-type sequence (or a cell obtained from such organism). In some embodiments of any of the aspects, the vector and/or nucleic acid sequence described herein is codon-optimized for expression in a mammal or mammalian cell, e.g., a mouse, a murine cell, or a human cell. In some embodiments of any of the aspects, the vector and/or nucleic acid sequence described herein is codon-optimized for expression in a human cell. In some embodiments of any of the aspects, the vector and/or nucleic acid sequence described herein is codon-optimized for expression in a yeast or yeast cell. In some embodiments of any of the aspects, the vector and/or nucleic acid sequence described herein is codon-optimized for expression in a bacterial cell. In some embodiments of any of the aspects, the vector and/or nucleic acid sequence described herein is codon-optimized for expression in an *E. coli* cell.

[00192] As used herein, the term "expression vector" refers to a vector that directs expression of an RNA or polypeptide from sequences linked to transcriptional regulatory sequences on the vector. The sequences expressed will often, but not necessarily, be heterologous to the cell. An expression vector may comprise additional elements, for example, the expression vector may have two replication systems, thus allowing it to be maintained in two organisms, for example in human cells for expression and in a prokaryotic host for cloning and amplification.

[00193] As used herein, the term “viral vector" refers to a nucleic acid vector construct that includes at least one element of viral origin and has the capacity to be packaged into a viral vector particle. The viral vector can contain the nucleic acid encoding a polypeptide as described herein in place of non-essential viral genes. The vector and/or particle may be utilized for the purpose of transferring any nucleic acids into cells either *in vitro* or *in vivo*. Numerous forms of viral vectors are known in the art.

[00194] It should be understood that the vectors described herein can, in some embodiments, be combined with other suitable compositions and therapies. In some embodiments, the vector is episomal. The use of a suitable episomal vector provides a means of maintaining the nucleotide

of interest in the subject in high copy number extra chromosomal DNA thereby eliminating potential effects of chromosomal integration.

[00195] As used herein, the terms "treat," "treatment," "treating," or "amelioration" refer to therapeutic treatments, wherein the object is to reverse, alleviate, ameliorate, inhibit, slow down or stop the progression or severity of a condition associated with a disease or disorder, e.g. myelofibrosis or a myeloproliferative disorder. The term "treating" includes reducing or alleviating at least one adverse effect or symptom of a condition, disease or disorder associated with myelofibrosis or a myeloproliferative disorder. Treatment is generally "effective" if one or more symptoms or clinical markers are reduced. Alternatively, treatment is "effective" if the progression of a disease is reduced or halted. That is, "treatment" includes not just the improvement of symptoms or markers, but also a cessation of, or at least slowing of, progress or worsening of symptoms compared to what would be expected in the absence of treatment. Beneficial or desired clinical results include, but are not limited to, alleviation of one or more symptom(s), diminishment of extent of disease, stabilized (*i.e.*, not worsening) state of disease, delay or slowing of disease progression, amelioration or palliation of the disease state, remission (whether partial or total), and/or decreased mortality, whether detectable or undetectable. The term "treatment" of a disease also includes providing relief from the symptoms or side-effects of the disease (including palliative treatment).

[00196] As used herein, the term "pharmaceutical composition" refers to the active agent in combination with a pharmaceutically acceptable carrier e.g. a carrier commonly used in the pharmaceutical industry. The phrase "pharmaceutically acceptable" is employed herein to refer to those compounds, materials, compositions, and/or dosage forms which are, within the scope of sound medical judgment, suitable for use in contact with the tissues of human beings and animals without excessive toxicity, irritation, allergic response, or other problem or complication, commensurate with a reasonable benefit/risk ratio. In some embodiments of any of the aspects, a pharmaceutically acceptable carrier can be a carrier other than water. In some embodiments of any of the aspects, a pharmaceutically acceptable carrier can be a cream, emulsion, gel, liposome, nanoparticle, and/or ointment. In some embodiments of any of the aspects, a pharmaceutically acceptable carrier can be an artificial or engineered carrier, e.g., a carrier that the active ingredient would not be found to occur in in nature.

[00197] As used herein, the term "administering," refers to the placement of a compound as disclosed herein into a subject by a method or route which results in at least partial delivery of the agent at a desired site. Pharmaceutical compositions comprising the compounds disclosed herein can be administered by any appropriate route which results in an effective treatment in the subject. In some embodiments, administration comprises physical human activity, e.g., an injection, act of ingestion, an act of application, and/or manipulation of a delivery device or machine. Such activity can be performed, e.g., by a medical professional and/or the subject being treated.

[00198] As used herein, “contacting" refers to any suitable means for delivering, or exposing, an agent to at least one cell. Exemplary delivery methods include, but are not limited to, direct delivery to cell culture medium, perfusion, injection, or other delivery method well known to one skilled in the art. In some embodiments, contacting comprises physical human activity, e.g., an injection; an act of dispensing, mixing, and/or decanting; and/or manipulation of a delivery device or machine.

[00199] The term “statistically significant" or “significantly" refers to statistical significance and generally means a two standard deviation (2SD) or greater difference.

[00200] Other than in the operating examples, or where otherwise indicated, all numbers expressing quantities of ingredients or reaction conditions used herein should be understood as modified in all instances by the term “about.” The term “about” when used in connection with percentages can mean ±1%.

[00201] As used herein, the term “comprising” means that other elements can also be present in addition to the defined elements presented. The use of “comprising” indicates inclusion rather than limitation.

[00202] The term "consisting of" refers to compositions, methods, and respective components thereof as described herein, which are exclusive of any element not recited in that description of the embodiment.

[00203] As used herein the term "consisting essentially of" refers to those elements required for a given embodiment. The term permits the presence of additional elements that do not materially affect the basic and novel or functional characteristic(s) of that embodiment of the invention.

[00204] As used herein, the term "corresponding to" refers to an amino acid or nucleotide at the enumerated position in a first polypeptide or nucleic acid, or an amino acid or nucleotide that is equivalent to an enumerated amino acid or nucleotide in a second polypeptide or nucleic acid. Equivalent enumerated amino acids or nucleotides can be determined by alignment of candidate sequences using degree of homology programs known in the art, e.g., BLAST.

[00205] As used herein, the term "specific binding" refers to a chemical interaction between two molecules, compounds, cells and/or particles wherein the first entity binds to the second, target entity with greater specificity and affinity than it binds to a third entity which is a non-target. In some embodiments, specific binding can refer to an affinity of the first entity for the second target entity which is at least 10 times, at least 50 times, at least 100 times, at least 500 times, at least 1000 times or greater than the affinity for the third nontarget entity. A reagent specific for a given target is one that exhibits specific binding for that target under the conditions of the assay being utilized.

[00206] The singular terms "a," "an," and "the" include plural referents unless context clearly indicates otherwise. Similarly, the word "or" is intended to include "and" unless the context clearly indicates otherwise. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of this disclosure, suitable methods and materials are described below. The abbreviation, "e.g." is derived from the Latin exempli gratia, and is used herein to indicate a non-limiting example. Thus, the abbreviation "e.g." is synonymous with the term "for example."

[00207] Groupings of alternative elements or embodiments of the invention disclosed herein are not to be construed as limitations. Each group member can be referred to and claimed individually or in any combination with other members of the group or other elements found herein. One or more members of a group can be included in, or deleted from, a group for reasons of convenience and/or patentability. When any such inclusion or deletion occurs, the specification is herein deemed to contain the group as modified thus fulfilling the written description of all Markush groups used in the appended claims.

[00208] Unless otherwise defined herein, scientific and technical terms used in connection with the present application shall have the meanings that are commonly understood by those of ordinary skill in the art to which this disclosure belongs. It should be understood that this invention is not limited to the particular methodology, protocols, and reagents, etc., described

herein and as such can vary. The terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention, which is defined solely by the claims. Definitions of common terms in immunology and molecular biology can be found in The Merck Manual of Diagnosis and Therapy, 20th Edition, published by Merck Sharp & Dohme Corp., 2018 (ISBN 0911910190, 978-0911910421); Robert S. Porter et al. (eds.), The Encyclopedia of Molecular Cell Biology and Molecular Medicine, published by Blackwell Science Ltd., 1999-2012 (ISBN 9783527600908); and Robert A. Meyers (ed.), Molecular Biology and Biotechnology: a Comprehensive Desk Reference, published by VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8); Immunology by Werner Luttmann, published by Elsevier, 2006; Janeway's Immunobiology, Kenneth Murphy, Allan Mowat, Casey Weaver (eds.), W. W. Norton & Company, 2016 (ISBN 0815345054, 978-0815345053); Lewin's Genes XI, published by Jones & Bartlett Publishers, 2014 (ISBN-1449659055); Michael Richard Green and Joseph Sambrook, Molecular Cloning: A Laboratory Manual, 4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., USA (2012) (ISBN 1936113414); Davis et al., Basic Methods in Molecular Biology, Elsevier Science Publishing, Inc., New York, USA (2012) (ISBN 044460149X); Laboratory Methods in Enzymology: DNA, Jon Lorsch (ed.) Elsevier, 2013 (ISBN 0124199542); Current Protocols in Molecular Biology (CPMB), Frederick M. Ausubel (ed.), John Wiley and Sons, 2014 (ISBN 047150338X, 9780471503385), Current Protocols in Protein Science (CPPS), John E. Coligan (ed.), John Wiley and Sons, Inc., 2005; and Current Protocols in Immunology (CPI) (John E. Coligan, ADA M Kruisbeek, David H Margulies, Ethan M Shevach, Warren Strobe, (eds.) John Wiley and Sons, Inc., 2003 (ISBN 0471142735, 9780471142737), the contents of which are all incorporated by reference herein in their entireties.

[00209] Other terms are defined herein within the description of the various aspects of the invention.

[00210] All patents and other publications; including literature references, issued patents, published patent applications, and co-pending patent applications; cited throughout this application are expressly incorporated herein by reference for the purpose of describing and disclosing, for example, the methodologies described in such publications that might be used in connection with the technology described herein. These publications are provided solely for their disclosure prior to the filing date of the present application. Nothing in this regard should be

construed as an admission that the inventors are not entitled to antedate such disclosure by virtue of prior invention or for any other reason. All statements as to the date or representation as to the contents of these documents is based on the information available to the applicants and does not constitute any admission as to the correctness of the dates or contents of these documents.

[00211] The description of embodiments of the disclosure is not intended to be exhaustive or to limit the disclosure to the precise form disclosed. While specific embodiments of, and examples for, the disclosure are described herein for illustrative purposes, various equivalent modifications are possible within the scope of the disclosure, as those skilled in the relevant art will recognize. For example, while method steps or functions are presented in a given order, alternative embodiments may perform functions in a different order, or functions may be performed substantially concurrently. The teachings of the disclosure provided herein can be applied to other procedures or methods as appropriate. The various embodiments described herein can be combined to provide further embodiments. Aspects of the disclosure can be modified, if necessary, to employ the compositions, functions and concepts of the above references and application to provide yet further embodiments of the disclosure. Moreover, due to biological functional equivalency considerations, some changes can be made in protein structure without affecting the biological or chemical action in kind or amount. These and other changes can be made to the disclosure in light of the detailed description. All such modifications are intended to be included within the scope of the appended claims.

[00212] Specific elements of any of the foregoing embodiments can be combined or substituted for elements in other embodiments. Furthermore, while advantages associated with certain embodiments of the disclosure have been described in the context of these embodiments, other embodiments may also exhibit such advantages, and not all embodiments need necessarily exhibit such advantages to fall within the scope of the disclosure.

[00213] The technology described herein is further illustrated by the following examples which in no way should be construed as being further limiting.

[00214] Some embodiments of the technology described herein can be defined according to any of the following numbered paragraphs:

1. A method to generate/engineer endothelial niche cells, comprising expressing one or more transcription factors in an endothelial cell, wherein the one or more transcription

factors are from the Ets family, the Sox family, and/or the Nuclear Hormone Receptor family.

2. The method of any one of the above paragraphs, wherein the endothelial niche cells express one are more genes comprising; *sele*, *exoc312a*, *snx8a*, *cltca*, *aqp7*, *ap1b1*, *lgmn*, *prcp*, *cldn11a*, *lyve1b*, *adra1d*, *hya12a*, *hya12b*, *tll1*, *i113ra2*, *glu1a*, *hexb*, *slc16a9a*, or *sepp1a*.
3. The method of any one of the above paragraphs, wherein the endothelial cells are human.
4. The method of any one of the above paragraphs, wherein the transcription factors comprises at least one of the human transcription factors ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2.
5. The method of any one of the above paragraphs, wherein the transcription factor includes at least one transcription factor from the Ets family, at least one transcription factor from the Sox family, and at least one transcription factor from the Nuclear Hormone Receptor family.
6. The method of any one of the above paragraphs, wherein the transcription factors include ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, and NR2F2.
7. The method of any one of the above paragraphs, wherein the transcription factors are expressed from at least one vector.
8. The method of any one of the above paragraphs, wherein the vector comprises an exogenous nucleic acid sequence(s) encoding the one or more transcription factors.
9. The method of any one of the above paragraphs, wherein the exogenous nucleic acid sequences are incorporated into the genome of the endothelial cell.
10. An engineered endothelial niche cell comprising one or more exogenous nucleic acid sequences encoding one or more transcription factors, wherein the one or more transcription factors are from the Ets family, the Sox family and/or the Nuclear Hormone Family.
11. A composition comprising the engineered endothelial niche cells of paragraph 10.
12. The composition of any one of the above paragraphs , wherein the composition is a therapeutic agent or the composition further comprises a pharmaceutically acceptable carrier.

13. The composition of any one of the above paragraphs, wherein the composition further comprises a culture dish, 3D cell system, or suspension system.
14. The composition of any one of the above paragraphs, wherein the composition comprises a scaffold.
15. A method for culturing HSPCs, the method comprising culturing HSPCs in the presence of a population of engineered endothelial niche cells.
16. The method of any one of the above paragraphs, wherein the method is performed *in vitro*.
17. The method of any one of the above paragraphs, wherein the engineered endothelial niche cells secrete a factor that affects the growth and/or expansion of the HSPC cells.
18. The method of any one of the above paragraphs, wherein the HSPCs cultured in the presence of the engineered endothelial niche cells can be cultured for at least 3 (e.g., at least 4, at least 5, at least 6, at least 7) days longer than HSPCs that are cultured in the absence of such engineered endothelial niche cells.
19. The method of any one of the above paragraphs, wherein the cells are cultured on a biologically compatible scaffold.
20. The method of any one of the above paragraphs, wherein the HSPCs cultured in the presence of the engineered endothelial niche cells have increased engraftment when administered to a subject compared to the engraftment of substantially similar HSPCs that were not cultured with engineered endothelial niche cells.
21. A method of treating a subject, the method comprising, transplanting a composition comprising HSPCs and a population of engineered endothelial niche-cells into the subject.
22. A method for enhancing engraftment of HSPCs, the method comprising administering a composition comprising HSPCs and a population of engineered endothelial niche cells to a subject in need thereof.
23. The method of any one of the above paragraphs, wherein engraftment of the HSPCs is increased by at least 10% compared to the engraftment of substantially similar HSPCs in the absence of engineered endothelial niche cells.
24. A co-culture comprising engineered endothelial niche cells and HSPCs.

25. The co-culture of paragraph 24 wherein the endothelial cells are made by the method of any one of the above paragraphs.
26. A kit for culturing HSPCs, the kit comprising: a population of engineered endothelial niche cells, reagents and instructions for use thereof.
27. A kit for generating engineered endothelial niche cells comprising: a vector(s) comprising one or more exogenous nucleic acid sequences encoding one or more transcription factors of the Ets family, the Sox family or the nuclear hormone family and instructions for use thereof.
28. A method for generating an ectopic vascular niche, the method comprising: administering an engineered endothelial niche cell to a target site in a subject in need thereof.
29. A method for extra medullary hematopoiesis, the method comprising transplanting engineered-niche endothelial cells into a subject at a location outside of the bone marrow (e.g., the forearm), thereby creating a synthetic niche.
30. The method of any one of the above paragraphs, wherein the endothelial cells are made by any of the methods in the above paragraphs.
31. A vector comprising one or more exogenous nucleic acid sequences encoding one or more transcription factors of the Ets family, the Sox family or the nuclear hormone family operably linked to a promoter.

EXAMPLES

Example 1

[00215] *Transcription factor induction of vascular blood stem cell niches in vivo.*

[00216] The haematopoietic niche is a supportive *in vivo* microenvironment comprised of distinct cell types, including specialized vascular endothelial cells that directly interact with haematopoietic stem and progenitor cells (HSPCs) to facilitate stem cell function. The molecular factors that specify niche endothelial cells and their pro-haematopoietic activity remain largely unknown. Using multi-dimensional gene expression analyses and a chromatin accessibility assay, defined herein is a conserved gene expression signature and cis-regulatory landscape unique to sinusoidal endothelial cells in the HSPC niche. Using enhancer mutagenesis and transcription factor overexpression, a transcriptional code was elucidated involving members of

the Ets, Sox and Nuclear Hormone Receptor families that is sufficient to induce ectopic niche endothelial cells that recruit HSPCs and support their homeostasis *in vivo*. Together, these studies have important implications for generating more efficient synthetic vascular niches for blood stem cells or for modulating the niche in a therapeutic context.

Results

[00217] *An endothelial gene expression signature unique to the fetal HSPC niche.*

[00218] ECs from different organs express distinct genes, but whether this is regulated by organ-specific transcriptional programs remains poorly understood. To investigate the regulation of gene expression in the CHT niche, RNA tomography (tomo-seq) was performed on the zebrafish tail at 72 hours post fertilization (hpf; see e.g., **Fig. 1A**). This tomo-seq analysis revealed clusters of gene expression corresponding to specific tissues along the dorsal-ventral axis of the tail, including spinal cord, notochord, muscle and epidermis, as well as specific blood and immune cell populations (see e.g., **Fig. 1B**, **Fig. 7A**, **Fig. 7B**). In total, 144 genes displayed enriched expression within the few cryosections spanning the CHT (see e.g., **Table 3**). Using whole mount *in situ* hybridization (WISH), CHT expression of 35 of these genes was confirmed (see e.g., **Table 3**; images available on the world wide web at zfin.org). To determine whether any of the 144 genes were expressed by ECs, the pan-endothelial transgene *kdrl*:GFP and fluorescence activated cell sorting (FACS) was used to isolate ECs for bulk and single cell RNA-seq (see e.g., **Fig. 1C**). In addition, these genes were cross-referenced with macrophage and neutrophil RNA-seq datasets (see e.g., Theodore et al. Distinct Roles for Matrix Metalloproteinases 2 and 9 in Embryonic Hematopoietic Stem Cell Emergence, Migration, and Niche Colonization. Stem cell reports 8, 1226-1241, 2017). Twenty-nine genes were identified that were selectively enriched in CHT ECs (see e.g., **Table 1**). In contrast to pan-endothelial genes, the tomo-seq expression traces for these CHT endothelial genes lacked the strong peak corresponding to expression in dorsal vasculature (see e.g., **Fig. 1D**). For 25 out of 29 genes the CHT EC-enriched expression was confirmed by WISH (see e.g., **Fig. 1E**, **Fig. 7A**, **Fig. 7B**, and **Table 1**).

[00219] To selectively isolate CHT ECs, transgenic lines were designed to label these cells. 1.3 or 5.3 kb upstream regulatory sequences were cloned for two CHT endothelial genes, *mrc1a* and *sele* to generate GFP reporter transgenes that were then crossed to the pan-endothelial marker *kdrl*:mCherry. For both the *mrc1a*:GFP and *sele*:GFP transgenes, the highest levels of

expression were observed in the venous sinusoids of the CHT, while low levels of GFP expression were detected in the posterior cardinal vein above the yolk extension and in a small number of vessels in the head (see e.g., **Fig. 8A-Fig. 8E**). Although GFP expression was also observed in mesenchymal cells in the outer tail fin, these cells are a frequent site of ectopic transgene expression and do not represent the endogenous expression of these genes. Notably, robust GFP expression was observed in sinusoidal ECs that directly interacted with HSPCs, confirming that the *mrc1a*:GFP and *sele*:GFP transgenes labeled ECs in the HSPC niche (see e.g., **Fig. 8A-Fig. 8E**). These GFP^+ ECs closely associated with *cxcl12a*:$DsRed2^+$ stromal cells and were observed to form pockets around HSPCs – a cellular behavior characteristic of ECs in the CHT niche (see e.g., **Fig. 8A-Fig. 8E**).

[00220] *Endothelial niche-specific cis-regulatory elements.*

[00221] To investigate the transcriptional control of niche-specific gene expression within CHT ECs, double positive *mrc1a*:GFP/*kdrl*:mCherry embryos were dissociated and FACS was used to isolate four different populations for RNA-seq and ATAC-seq analyses: $GFP^+/mCherry^+$ (CHT ECs), $GFP^-/mCherry^+$ (ECs outside the CHT), $GFP^+/mCherry^-$ (mesenchymal cells in the tail fin), and $GFP^-/mCherry^-$ (negative remainder of the embryo; see e.g., **Fig. 2A**). By comparing regions of chromatin accessibility across the four populations, 6,848 regions uniquely open were identified in CHT ECs. Of the 29 CHT EC genes, 26 out of 29 had an ATAC-seq element within 100 kb of the transcriptional start site that was found only in CHT ECs (see e.g., **Fig. 2B**, **Table 1**). To test whether these regions might be tissue-specific enhancers, the sequences were cloned for 15 of the elements, which were then fused to a minimal promoter and GFP, and then injected into zebrafish embryos. 12 out of 15 constructs showed GFP expression enriched in CHT ECs at 60-72 hpf (see e.g., **Fig. 2C**, **Table 2**). As a control, regions of chromatin were cloned, which were predicted to contain pan-endothelial regulatory elements based on proximity to previously annotated vascular-specific genes and their accessibility in both the CHT and non-CHT EC fractions (see e.g., **Fig. 9A-Fig. 9C**, **Table 2**). A number of these regions were previously shown to contain endothelial enhancers (see e.g., Quillien et al. Robust Identification of Developmentally Active Endothelial Enhancers in Zebrafish Using FANS-Assisted ATAC-Seq. Cell reports 20, 709-720, 2017). For 6 out of 6 pan-endothelial regions, mosaic GFP expression in ECs were observed throughout the embryo that was not restricted to

the CHT (see e.g., **Fig. 9A-Fig. 9C**). Thus, the zebrafish system was able to rapidly validate, *in vivo*, niche-specific endothelial enhancers predicted by ATAC-seq analysis.

[00222] *Transcription factor binding sites for niche endothelial expression.*

[00223] To identify transcription factors that might bind the CHT EC enhancers, a motif enrichment analysis was performed on the 6,848 regions of chromatin that were uniquely accessible in CHT ECs. This analysis revealed that Ets, Sox (SoxF factors, specifically) and Nuclear Hormone Receptor (NR2F2/RORA/RXRA factors, specifically, abbreviated hereafter as NHR) binding motifs were most enriched in the 6,848 regions (see e.g., **Fig. 9A-Fig. 9C**). In contrast, there were 4,522 pan-endothelial elements (i.e., regions of chromatin accessible in both CHT and non-CHT ECs) across the genome enriched for Ets sites, but not SoxF or NHR binding motifs (see e.g., **Fig. 9A-Fig. 9C**). Of note, the 12 CHT-EC elements that drove GFP expression in the *in vivo* reporter assay all harbored Ets, SoxF and NHR sites, whereas one of the three CHT-EC elements that failed to drive GFP expression and three of the six pan-endothelial regions lacked a NHR binding site (see e.g., **Table 2**).

[00224] To determine a minimal sequence sufficient to drive CHT EC expression, a 125 base pair (bp) sequence upstream of *mrc1a* and a 158 bp sequence upstream of *sele* were cloned; these sequences had been included the our original transgenes and corresponded to the strongest ATAC-seq signal in these regions (see e.g., **Fig. 3A**, **Fig. 10A-Fig. 10E**). When coupled to a minimal promoter, these elements drove GFP expression in CHT ECs in 44% (125 bp *mrc1a* sequence; 155 out of 356) and 23% (158 bp *sele* sequence; 176 out of 775) of injected embryos (see e.g., **Fig. 3B**, **Fig. 10A-Fig. 10E**). Compared to injection of a *kdrl*:mCherry construct, in which mosaic expression was observed in ECs throughout the embryo, expression of the *mrc1a* and *sele* enhancer-GFP constructs was restricted to CHT ECs (see e.g., **Fig. 3C**, **Fig. 10A-Fig. 10E**). Moreover, when stable transgenic lines were established using these short constructs, GFP was specifically expressed in CHT ECs (see e.g., **Fig. 3D**, **Fig. 10A-Fig. 10E**). Mammalian *Mrc1* is prominently expressed by macrophages and venous sinusoidal ECs. Zebrafish have two homologous genes, *mrc1a* and *mrc1b*. A recent study of *mrc1a* reported expression in both macrophages and ECs when the promoter was coupled to an intronic enhancer that showed conservation with *mrc1b* (see e.g., Jung et al. Development of the larval lymphatic system in zebrafish. Development 144, 2070-2081, 2017). The 125 bp enhancer element, in contrast, drove expression specifically in CHT ECs, illustrating the specificity in the enhancers.

[00225] Both the 125 bp *mrc1a* and 158 bp *sele* regulatory sequences contained Ets, SoxF and NHR motifs (see e.g., **Fig. 3E**, **Fig. 10A-Fig. 10E**). To test whether these transcription factor binding sites were required for expression, variants were generated in which each class of motif was disrupted by mutation. In each case, disruption of the Ets, SoxF or NHR motifs led to a significant reduction or complete loss of GFP expression in CHT ECs (see e.g., **Fig. 3F**). In control constructs where mutations were targeted to intervening sequences between the Ets, SoxF and NHR motifs, GFP expression was unperturbed (see e.g., **Fig. 3F**). Studies of arterial-venous specification in zebrafish have shown that the NHR Nr2f2, also known as COUP-TFII, promotes venous endothelial cell fate (see e.g., Aranguren et al. Transcription factor COUP-TFII is indispensable for venous and lymphatic development in zebrafish and *Xenopus laevis*. Biochemical and biophysical research communications 410, 121-126, 2011). To test whether Nr2f2 could directly bind the enhancer sequences, an *in vitro* gel electrophoretic mobility shift assay was performed. Incubation of murine NR2F2-GST protein with labeled probe from either the 125 bp *mrc1a* or 158 bp *sele* zebrafish enhancers led to DNA:protein complexes that were super-shifted upon addition of an NR2F2 antibody and could be outcompeted by an unlabeled competitor probe (see e.g., **Fig. 10A-Fig. 10E**). However, unlabeled probes in which the NHR motifs were mutated were unable to outcompete the wild-type probes, demonstrating that NR2F2 binds these NHR motifs in a sequence-specific manner.

[00226] *Defined factors induce niche endothelial expression.*

[00227] To determine which transcription factors are expressed in CHT ECs and might bind the Ets, Sox and NHR motifs *in vivo*, RNA-seq data from CHT ECs was examined. The most highly expressed factors were *fli1a*, *etv2*, *ets1*, *sox18*, *sox7*, *nr2f2* and *rxraa* (see e.g., **Table 4**). To test whether these seven factors could induce niche endothelial gene expression outside the CHT, constructs were generated in which the orthologs for each transcription factor was under the control of a ubiquitous (ubi) promoter (see e.g., **Fig. 4A**). A pool of the seven ubi-driven factors were then injected into one cell-stage zebrafish embryos, and *mrc1a* and *sele* expression was examined by WISH at 60-72 hpf. Strikingly, 17% (12 out of 69) of these embryos had ectopic vascular patches of *mrc1a* expression outside of the CHT, dorsally within the trunk and tail, and over the yolk (see e.g., **Fig. 4B**, **Fig. 11A-Fig. 11C**). Control-injected embryos did not show ectopic expression (0 out of 56). Similar results were obtained with WISH for *sele* or when factors were injected into *mrc1a*:GFP/*kdrl*:mCherry double transgenic embryos (see e.g., **Fig.**

4B-Fig. 4C, **Fig. 11A-Fig. 11C**). Vessels ectopically expressing *mrc1a* were often larger than their normal counterparts in these regions and had a sinusoidal-like morphology similar to the CHT (see e.g., **Fig. 4C**). Using DIC microscopy, blood flow through these regions was readily visualized (data not shown). These data suggest that a small number of transcription factors are sufficient to ectopically induce the niche endothelial gene program.

[00228] This mutational analysis of the 125 bp *mrc1a* and 158 bp *sele* enhancers indicated that a factor from each of the Ets, Sox and NHR families was required for expression, which led to the experiment of whether a combination of just three factors (one from each family) was sufficient to induce niche endothelial gene expression. ETV2 is a pioneer factor that is essential for specification of early mesodermal progenitors into vascular cell fates. Forced expression of ETV2 in nonvascular cells induces reprogramming towards an early endothelial fate that can generate many types of vasculature. Previous work in zebrafish has shown the importance of SoxF factors (*sox7* and *sox18*) and *nr2f2* during arterial-venous specification (see e.g., Swift et al. SoxF factors and Notch regulate *nr2f2* gene expression during venous differentiation in zebrafish. Developmental biology 390, 116-125, 2014). A combination of three of these factors – ETV2, SOX7 and Nr2f2 – could be sufficient to induce ectopic niche endothelial gene expression. Consistent with this, when these three factors were injected, significant ectopic *mrc1a* expression was observed (see e.g., **Fig. 4D-Fig. 4F**). The frequency of ectopic vessels generated with three factors was higher than the 7-factor pool, suggesting factor concentration has functional significance. Injected embryos similarly showed ectopic expression of *sele*, *gpr182* and *lgmn*, indicating that the 3-factor pool induced the niche endothelial program (see e.g., **Fig. 11A-Fig. 11C**).

[00229] To evaluate the contribution of the individual transcription factors, each factor was injected alone, and WISH was performed for *mrc1a*. Embryos injected with SOX7 or Nr2f2 alone showed minimal ectopic expression (see e.g., **Fig. 4F**). Injection of ETV2 alone led to ectopic expression of *mrc1a*, although at a lower frequency than when ETV2 was injected in combination with SOX7 and Nr2f2, indicating the additional factors increase ectopic induction (see e.g., **Fig. 4F**, **Fig. 12-Fig. 12D**). Injecting ETV2 alone induced ectopic expression of the endogenous zebrafish *sox7* gene, as well as *sox18*, *fli1a* and *etv2*, demonstrating that human ETV2 can induce several zebrafish endothelial gene programs, including artery, vein and niche endothelial genes (see e.g., **Fig. 12-Fig. 12D**). Because endogenous *ets1* and *nr2f2* are expressed

broadly outside the CHT (e.g. in the spinal cord region), it was difficult to visualize whether they were similarly induced by ETV2 overexpression (see e.g., **Fig. 12-Fig. 12D**). ETS1 has the capacity to bind to the consensus ETS motif enriched in the peaks and is expressed at an appreciable level in the niche ECs. In contrast to ETV2, overexpression of ETS1 alone did not lead to widespread vessel induction or ectopic mrc1a expression (see e.g., **Fig. 4F**). Injection of ETS1, SOX7 and Nr2f2, however, led to significant ectopic expression of *mrc1a*, illustrating the combinatorial activity of these transcription factors (see e.g., **Fig. 4E**, **Fig. 4F**, **Fig. 12-Fig. 12D**). Notably, the zebrafish genes encoding each of the seven initial transcription factors had regions of chromatin associated with them that were uniquely accessible in the CHT EC fraction and harbored Ets, SoxF and NHR sites (see e.g., **Table 5**), indicating that these factors regulate one another similar to an auto-regulatory loop established during reprogramming. Thus, overexpression of the three factor combinations, including either ETV2 or ETS1 with Sox and NHR factors, mimics the endogenous expression of these factors in niche ECs at 72 hpf and drives robust specification of the niche endothelial gene program.

[00230] *Ectopic vascular regions recruit HSPCs.*

[00231] It was next asked whether the ectopic regions of CHT EC gene expression were capable of recruiting and supporting HSPCs. The ETV2, SOX7 and Nr2f2 pool was injected into *mrc1a*:GFP/*runx1*:mCherry double positive transgenic embryos, and the localization of *runx1*:mCherry^{+} HSPCs was examined. Strikingly, in 12 out of 22 embryos that had ectopic vascular patches of *mrc1a*:GFP, HSPCs were observed localizing to these regions (see e.g., **Fig. 5A**). In contrast, 0 out of 27 control embryos had ectopically localized HSPCs. Similar ectopic localization of *runx1*$^{+}$ HSPCs was observed by *runx1* WISH and in embryos injected with the ETS1, SOX7 and Nr2f2 combination (see e.g., **Fig. 5A**, **Fig. 5B**).

[00232] To evaluate whether these ectopic regions provide a supportive environment for HSPCs, high-resolution live cell confocal microscopy was used to more closely examine HSPC behaviors and interactions with ECs at these sites. This analysis revealed that HSPCs directly associated with the *mrc1a*:GFP^{+} ECs within ectopic sites and could be found in both intraluminal and extravascular spaces. The ectopic *mrc1a*:GFP^{+} ECs were often associated with *cxcl12a*:DsRed2^{+} stromal cells and formed pockets around the HSPCs, similar to what is observed in the CHT (see e.g., **Fig. 5A**, **Fig. 5C**). Notably, using time-lapse microscopy, the initial recruitment, lodging and division of HSPCs was observed within these sites (see e.g., **Fig.**

5D). When the HSPCs divided, daughter cells migrated away from the ectopic site and entered circulation, presumably traveling to subsequent niches (see e.g., **Fig. 5E**). Together, these data demonstrate that reprogramming by the three factors that drive the niche endothelial gene program leads to functional ectopic niches that can recruit and sustain HSPCs.

[00233] *A conserved endothelial signature in the HSPC niche.*

[00234] To determine whether a similar endothelial expression signature was present in the adult zebrafish niche, the *mrc1a*:GFP and *sele*:GFP transgenes were examined, and it was found that both lines had GFP expression in ECs in the kidney marrow (see e.g., **Fig. 13A**, **Fig. 13B**). In addition, a single cell RNA-seq dataset for zebrafish kidney marrow was examined (see e.g., Tang et al. Dissecting hematopoietic and renal cell heterogeneity in adult zebrafish at single-cell resolution using RNA sequencing. The Journal of experimental medicine 214, 2875-2887, 2017). This analysis found that 23 out of 29 of the CHT EC genes were strongly enriched in vascular ECs in the kidney marrow (see e.g., **Fig. 6A**). To determine whether this pro-haematopoietic vascular niche signature is conserved in mammals, an RNA-seq dataset was probed, which comprises gene expression for ECs from multiple organs of the mouse, including the heart, kidney, liver, lung and bone marrow, from five stages of development (E11-13, E14-15, E16-17, P2-P4 and adult). Orthologs for 21 out of 29 CHT EC genes were enriched in the ECs of a mammalian haematopoietic organ – the fetal liver and/or adult bone marrow – relative to their expression in ECs from non-haematopoietic organs at the same stage of development (see e.g., **Fig. 6B**). Notably, a subset of expression patterns mirrored the temporal dynamics of HSPC ontogeny, showing robust expression in fetal liver ECs at the E14-17 stages and then later displaying elevated expression in adult bone marrow ECs with a concomitant and expected reduction in liver EC expression (P2-P4 and adult; see e.g., **Fig. 6B**). To determine if a transcriptional program similar to the one uncovered in the zebrafish could control these genes in the mouse, the expression of transcription factors from the Ets, Sox and NHR families was examined using the same RNA-seq datasets. In both fetal liver ECs (E14-17) and adult bone marrow ECs, the most highly expressed factors were Ets1, Sox18 and Nr2f2 (see e.g., **Table 5**). Taken together, these data indicate that a conserved transcriptional program regulates the pro-haematopoietic niche identity of ECs in the fetal and adult HSPC niches.

Discussion

[00235] These data support a model in which a transcriptional program comprised of factors from the Ets, SoxF and NHR families specifies the identity and unique capacity of vascular niche ECs to maintain and expand blood stem cells. This is a conserved feature of the haematopoietic niche at multiple stages of development and across species. The niche endothelial expression signature identified herein comprises genes with known niche functions (e.g., the adhesion receptor E-selectin and the cysteine proteinase Cathepsin L, which has been linked to Cxcl12 turnover in the bone marrow). In addition, there are numerous genes not previously associated with the HSPC niche, including several genes with scavenger functions or activities related to endocytosis and intracellular uptake, including *mrc1a*, *stab1/2*, *dab2*, *ap1b1*, *pxk* and *snx8*. These molecules could regulate ligand and receptor turnover in the niche or could function to clear potentially harmful agents, such as waste products, modified proteins, or viral, bacterial or fungal related material, from the niche microenvironment. CHT ECs share gene expression with lymphatic ECs, including genes such as *lyve1b*, which is consistent with their shared origin. In a recent study of the mouse bone marrow, differential gene expression was examined between SECs and AECs. The niche EC signature disclosed herein aligns with the venous SECs of the marrow. Although AECs may also support haematopoiesis, the work herein illustrates the sufficiency of SECs to instruct vascular niche formation and foster stem cell expansion. Extramedullary haematopoiesis that occurs during stress could involve the local induction of this SEC niche program. Other transcription factors have been shown to drive CHT gene expression that maintains haematopoietic cells, including *tfec* and *klf6a*, although enriched binding motifs were not observed for these factors. It is plausible that their targets are in different cells of the CHT or that these factors act downstream of our program to specify diverse populations of tissue-specific vascular niches.

[00236] Experiments with parabiotic mice have indicated that the size of the niche determines the number of HSPCs (see e.g., Chen et al. Mobilization as a preparative regimen for hematopoietic stem cell transplantation. Blood 107, 3764-3771, 2006). Functional ectopic niches, termed ossicles, have been used to assemble a bone marrow equivalent when transplanted into mice, and it is likely that SECs are present in these structures. Together, these studies and the work presented herein support a method of increasing HSPC numbers *in vivo* by generating ectopic vascular niches at new safe harbor locations in the body. This approach lays the

foundation for novel treatments for diseases where the endogenous bone marrow niche is compromised (e.g., myelofibrosis). At a broader level, this work advances the fundamental understanding of the vascular niche that choreographs homeostasis and regeneration of blood stem cells, which can guide new therapeutic strategies to culture and expand HSPCs for transplantation.

Methods

[00237] *Animal models.*

[00238] Wild-type AB, *casper* or *casper*-EKK, and transgenic lines *cd41*:EGFP, *runx1*:mCherry [*runx1+23*:NLS-mCherry], *kdrl*(*flk1*):GFP [*kdrl*:GRCFP], *kdrl*:mCherry [*kdrl*:Hsa.hras-mCherry], and *cxcl12a*(*sdf1a*):DsRed2 were used in this study. Alternative gene names are listed in parenthesis and full transgene names are listed in brackets.

[00239] *Genomic analyses.*

[00240] For RNA tomography (tomo-seq), 72 hpf embryos were euthanized by tricaine overdose and the portion of the tail containing the CHT was manually dissected using a scalpel. The tissue was oriented in optimal cutting temperature (OCT) tissue freezing media in a cryomold with the ventral side facing the bottom of the mold. After snap freezing on dry ice, 40 individual 8 μm-thick cryosections were collected along the dorsal-ventral axis using a cryostat. The RNA from individual cryosections was extracted using TRIzol™ and then barcoded during a reverse transcription step prior to pooling for library preparation and sequencing (see e.g., Junker et al. Genome-wide RNA Tomography in the zebrafish embryo. Cell 159, 662-675, 2014). For single cell and bulk RNA-seq, *kdrl*:GFP embryos were dissociated using Liberase™ and GFP^+ cells were isolated by FACS. For bulk RNA-seq, total RNA was isolated using TRIzol™ and GenElute LPA™ carrier. Libraries were prepared from 50 ng of total RNA/sample as input using Ribogone™ and a SMARTer Universal Low Input RNA Kit™. For single-cell sequencing, approximately 2,000 cells were encapsulated, and libraries were prepared for sequencing. For ATAC-seq, embryos were dissociated using Liberase™ and a minimum of 12,000 cells (max 50,000) were isolated by FACS. Cells were subsequently lysed and isolated nuclei were incubated in a transposition reaction. All sequencing was done using an Illumina Hiseq 2500™. For RNA-seq, quality control was performed by Fast QC™ and Cutadapt™ to remove adaptor sequences and low quality regions. High-quality reads were aligned to UCSC build danRer7 of the zebrafish genome using Tophat™ 2.0.1158 without novel splicing form calls. Transcript

abundance and differential expression were calculated with Cufflinks™ 2.2.159. FPKM values were used to normalize and quantify each transcript. For ATAC-seq, reads were aligned to UCSC build danRer7 of the zebrafish genome using Bowtie2™ (version 2.2.1) with the following parameters: --end-to-end, -N0-, -L20. The MACS2™ (version 2.1.0) peak finding algorithm was used to identify regions of ATAC-seq peaks with the following parameters: --nomodel --shift -100 --extsize 200. An initial q-value threshold of enrichment of 0.05 was used for peak calling and a more stringent q-value of 14 was used to identify peaks that were distinct between different samples. Genome-wide motif enrichment analysis was performed using HOMER™ and motif annotation was done using Consite™. Gene expression analysis of the adult kidney marrow was performed data available on the world wide web at molpath.shinyapps.io/zebrafishblood/.

[00241] *Whole mount in situ hybridization (WISH).*

[00242] *In situ* hybridization was performed using a standard protocol. Embryos were subsequently transferred to glycerol for scoring and imaging. *In situ* probes were generated by PCR amplification using a cDNA or plasmid (for transcription factors from other species) template followed by reverse transcription with digoxigenin-linked nucleotides. Primer sequences for all WISH probes used herein are provided in **Table** 7. WISH images for the 35 CHT-enriched genes identified by tomo-seq can be found on the world wide web at zfin.org.

[00243] *Transgenesis and enhancer-GFP reporter assays.*

[00244] Transgenic lines were established. For the *mrc1a*:GFP and *sele*:GFP transgenes, 1.3 kb and 5.3 kb sequences, respectively, upstream of the transcriptional start site were PCR amplified off of genomic DNA and then TOPO-TA™ cloned into a p5E Gateway™ vector, which was then recombined with GFP and a polyA tail, all flanked by Tol2 sites. For the 125 bp *mrc1a* and 158 bp *sele* enhancers, the elements were PCR amplified off of genomic DNA, TOPO-TA™ cloned into a p5E Gateway™ vector and then recombined with the mouse *Beta-globin* minimal promoter fused to GFP with a polyA tail, all flanked by Tol2 sites. Embryos were injected at the one cell-stage with Tol2 RNA and at least two independent lines showing similar expression were established for each construct: (Tg(*mrc1a(1.3kb)*:GFP); Tg(*sele(5.3kb)*:GFP); Tg(*mrc1a(125bp)*:GFP); and Tg(*sele(158bp)*:GFP). The CHT EC and pan-EC ATAC-seq elements were similarly amplified by PCR using genomic DNA and then fused to the *Beta-globin* minimal promoter and GFP. Mutational variants of 125 bp *mrc1a* and 158 bp

sele were generated by annealing overlapping oligos followed by a T4 DNA polymerase reaction to generate blunt-ended products, which were subsequently cloned into p5E Gateway™ vectors (following A-tailing with Klenow Fragment using the same work flow as for the ATAC-seq elements). Transcription factor binding motifs were disrupted by changing nucleotides in the core binding sites, purines for pyrimidines and vice versa. Injected F0 embryos were scored between 60-72 hpf. Control and experimental groups were blinded prior to scoring and all experiments were performed at least three times, with independent clutches. GFP expression in CHT ECs or pan-EC expression was scored as significant if it was observed in at least 10% of F0 injected embryos. Embryos scored as negative had either no GFP expression or had only sparse ectopic expression in muscle cells. The sequences for primers used to amplify the *mrc1a* and *sele* regulatory elements, as well as the 15 CHT EC and 6 pan-EC ATAC-seq elements, are provided in **Table 7**. The sequences for the overlapping oligos that were used to generate the enhancer variants are provided in **Table 8**. The fidelity of all constructs was confirmed by sequencing prior to injection.

[00245] *Transcription factor overexpression studies.*

[00246] For transcription factor overexpression studies, the open reading frames for the human (FLI1, ETV2, ETS1, SOX7 and RXRA), xenopus (Sox18) or zebrafish (Nr2f2) genes were cloned into a pME Gateway vector (Invitrogen™) and then recombined with the zebrafish *ubi* promoter and a polyA tail, all flanked by Tol2 sites. The fidelity of all constructs was confirmed by sequencing prior to injection. Embryos were injected with pools of *ubi*-driven transcription factors (1 nl at 25 ng/µl total DNA, plus Tol2 RNA) at the one cell-stage and then screened between 24-72 hpf for ectopic niche endothelial gene expression or ectopic HSPC localization. For control and single-factor injections, the empty Tol2 Gateway™ destination vector was used as filler DNA in the injection mix. Expression of the transcription factors was confirmed by WISH using species-specific *in situ* probes. Ectopic expression was scored as vascular staining or vascular GFP expression outside the normal domain of gene expression. Control and experimental samples were blinded prior to scoring and all experiments were performed at least three times, with independent clutches.

[00247] *Microscopy and image analysis*

[00248] Time-lapse microscopy was performed using a Yokogawa CSU-X1™ spinning disk mounted on an inverted Nikon Eclipse Ti™ microscope equipped with dual Andor iXon™

EMCCD cameras and a climate controlled (maintained at 28.5°C) motorized x-y stage to facilitate tiling and imaging of multiple specimens simultaneously. Screening of injected enhancer-GFP constructs and imaging of WISH embryos was performed using a Nikon SMZ18™ stereomicroscope equipped with a Nikon DS-Ri2™ camera. All images were acquired using NIS-Elements™ and processed using Imaris™ or Adobe Photoshop™ software. Embryos were mounted for imaging. Briefly, specimens were mounted in 0.8% LMP agarose with tricaine (0.16 mg/ml) in glass bottom 6-well plates and covered with E3 media containing tricaine (0.16 mg/ml).

[00249] *Flow cytometry, kidney marrow dissection, dissociation and histology.*

[00250] Embryos were prepared for FACS. Briefly, embryos were chopped with a razor blade in cold PBS and then incubated in Liberase™ for 20 minutes at 37˚C before filtering the dissociated cells through a 40 μm mesh filter and transferring to 2% FBS. FACS was performed using a FACS Aria™ machine. Gates were set using transgene positive and negative control embryos, and SYTOX Blue™ was used as a live/dead stain. At least 12,000 (50,000 max) cells were collected per sample for ATAC-seq experiments and at least 10,000 (300,000 max) cells per sample were collected for RNA-seq experiments. Kidney marrow was harvested from adult zebrafish by manual dissection and then fixed in 4% PFA (for histology) or dissociated by gentle pipetting (for live cell imaging). For histology the kidney marrow was embedded in paraffin prior to sectioning; alternating sections were stained with H&E or with an antibody to GFP. Mouse EC populations were sorted as $Cd45^{-}Pdpn^{-}Cd31^{+}$ cells.

[00251] *Electrophoretic mobility shift assay.*

[00252] The Nr2f2 fragment was cloned into the pGEX2TK™ vector to generate GST-tagged Nr2f2 and fidelity was verified by sequencing. The pGEX2TK-Nr2f2 protein plasmid was transformed into *E. coli* BL21 competent cells. Proteins were expressed and purified, and purified proteins were quantified against BSA. EMSAs were performed. Probes were generated by annealing 100 pmol of sense and antisense oligonucleotides, and 1-2 pmol of probe was used in each reaction. All primer and probe sequences are provided in **Table 9**. Gel shift reactions were conducted at 4°C in 20% glycerol, 20 mM Tris (pH 8.0), 10 mM KCl, 1 mM DTT, 12.5 ng poly dI/C, 6.25 pmol of random, single-stranded oligonucleotides, BSA and the probe in the amount specified above. Samples involving the Nr2f2 protein were loaded on a 6% gel to resolve protein-DNA complexes. In reactions with cold competitors, 20x unlabeled probes were included

in the reactions. Anti-NR2F2 antibody (R&D Biosystems™; cat # PP-H7147-00) was at the same amount of the Nr2f2 protein to obtain super-shifts.

[00253] The GEO accession number for the mammalian genomic data reported herein is GSE100910. The zebrafish genomic data reported herein was submitted to the NCBI™ Gene Expression Omnibus.

[00254] Real-time imaging can show blood circulation through region of ectopic niche endothelial gene expression. At 72 hpf show blood cells can be seen circulating through a region of vessels in the dorsal trunk that are ectopically expressing *mrc1a*:GFP in a 72 hpf embryo that had been injected with a pool of seven transcription factors at the one cell-stage.

[00255] Real-time imaging can show initial recruitment of HSPC to region of ectopic niche endothelial expression. A *runx1*$^{+}$ HSPC can be seen initially lodging in a dorsal vessel that is ectopically expressing mrc1a:GFP (see e.g., **Fig. 5D** arrows) in a 72 hpf embryo that had been injected with a pool of ETV2, SOX7 and Nr2f2 at the one cell-stage. Black arrowhead points to HSPC localization in the CHT. The duration of the time-lapse is 6.5 hours; time intervals are 2 min. Time-series is shown in **Fig. 5D**.

[00256] Real-time imaging can show proliferation of HSPCs and egress from ectopic region of niche endothelial gene expression. *Runx1*$^{+}$ HSPCs can be localized to a vessel ectopically expressing mrc1a:GFP (see e.g., **Fig. 5E** arrows) in a 72 hpf embryo that had been injected with a pool of ETV2, SOX7 and Nr2f2 at the one cell stage. HSPCs divide several times and migrate away into circulation. Black arrowhead points to HSPC localization in the CHT. The duration of the time-lapse is 2.6 hours; time intervals are 2 min. Time-series shown in **Fig. 5E**.

Table 1: CHT EC-enriched genes.

Gene	Full Gene Name	CHT Expression Confirmed by WISH	Associated with CHT EC ATAC-seq Element*	Function
adra1d	*adrenoreceptor alpha 1D*	No†	Yes	G-protein coupled receptor; mitogenic response activation
ap1b1	*adaptor-related protein complex 1, beta 1 subunit*	Yes	Yes	Coated vesicle clathrin recruitment
cldn11a	*claudin 11a*	Yes	Yes	Tight junction strand component
cltca	*clathrin, heavy chain a (Hc)*	Yes	Yes	Major coated vesicle and coated pit component
ctsh	*cathepsin H*	Yes	No	Lysosomal cysteine proteinase
ctsla	*cathepsin La*	Yes	Yes	Lysosomal cysteine proteinase
dab2	*Dab, mitogen-responsive phosphoprotein, homolog 2 (Drosophila)*	n/a‡	Yes	Mitogen-responsive phosphoprotein; clathrin-mediated endocytosis
exoc3l2a	*exocyst complex component 3-like 2a*	Yes	Yes	SNARE binding
glula	*glutamate-ammonia ligase (glutamine synthase) a*	Yes	Yes	Glutamine synthesis
gpr182	*G protein-coupled receptor 182*	Yes	Yes	G-protein coupled receptor; vasodilation
hexb	*hexosaminidase B (beta polypeptide)*	Yes	Yes	Degradation of N-acetyl hexosamine containing molecules
hyal2a	*hyaluronidase 2a*	n/a§	Yes	Hyaluronan degradation
hyal2b	*hyaluronidase 2b*	Yes	Yes	Hyaluronan degradation
ifi30	*interferon, gamma-inducible protein 30*	Yes	Yes	Disulfide bond reduction, MHC class II-restricted antigen processing
il13ra2	*interleukin 13 receptor, alpha 2*	Yes	No	Interleukin 13 binding
lgmn	*legumain*	Yes	Yes	Hydrolysis of asparaginyl bonds
lyve1b	*lymphatic vessel endothelial hyaluronic receptor 1b*	Yes	Yes	Hyaluronan receptor
man2b2	*mannosidase, alpha, class 2B, member 2*	Yes	Yes	Mannose glycosylase
mrc1a	*mannose receptor, C type 1a*	Yes	Yes	Glycoprotein endocytosis
npl	*N-acetylneuraminate pyruvate lyase (dihydrodipicolinate synthase)*	n/a‡	Yes	N-acetylneuraminic acid cleavage
prcp	*prolylcarboxypeptidase (angiotensinase C)*	n/a‡	Yes	C-terminal proline linked amino acid cleavage
pxk	*PX domain containing serine/threonine kinase*	No	Yes	Synaptic transmission
sele	*selectin E*	Yes	Yes	Endothelial cell adhesion to blood leukocytes
sepp1a	*selenoprotein P*	Yes	No	Selenium binding
slc16a9a	*solute carrier family 16, member 9a*	No	Yes	Symporter activity
snx8a	*sorting nexin 8a*	Yes	Yes	Phosphatidylinositol binding
stab1	*stabilin 1*	Yes	Yes	Scavenger receptor activity
stab2	*stabilin 2*	Yes	Yes	Scavenger receptor activity, hyaluronan receptor
tll1	*tolloid-like 1*	Yes	Yes	Procollagen C-propeptide processing

[00257] Table 2: *In vivo* screening of predicted enhancer elements.

Type of Element	Gene Name	Genomic Coordinates of ATAC-seq Element*	Relative to TSS (kb)	Amplicon Size (bp)	Showed Predicted GFP Expression Pattern†	Element Contains Ets, SoxF and NHR Motifs
CHT EC Element	*ap1b1*	chr5:26463217-26463695	17	750	Yes	Yes
	cltca	chr10:29,047,274-29,047,619	2.8	404	Yes	Yes
	dab2	chr5:33,980,000-33,980,306	-3.5	394	Yes	Yes
	exoc3l2a	chr5:38359097-38359903	5.9	901	Yes	Yes
	giula	chr2:19,458,704-19,459,047	4.8	446	No	Yes
	gpr182	chr23:36701205-36701682	-4.9	481	Yes	Yes
	gpr182	chr23:36694073-36694476	-2.8	398	Yes	Yes
	gpr182	chr23:36696363-36696656	1.6	577	Yes	Yes
	lgmn	chr13:36,448,465-36,448,818	2.9	414	Yes	Yes
	prcp	chr15:10,400,588-10,400,868	23	334	No	No‡
	sele	chr20:34,010,027-34,010,326	-9.7	398	Yes	Yes
	sele	chr20:34,011,251-34,011,563	-8.5	360	Yes	Yes
	snx8a	chr3:42,090,805-42,091,062	5.5	395	No	Yes
	stab1	chr22:10467346-10467937	-2.8	874	Yes	Yes
	stab2	chr4:9790795-9791116	4.3	422	Yes	Yes
Pan-EC Element	*cdh5*	chr7:45457842-45458791	13	823	Yes	Yes
	clec14a	chr17:10362325-10362844	-3.1	455	Yes	Yes
	dll4	chr20:28219013-28219619	-55	452	Yes	No‡
	fli1a	chr18:47039842-47040466	47	800	Yes	No‡
	lmo2	chr18:36722030-36722527	-3.6	367	Yes	No‡
	nrp1b	chr2:43535098-43535801	-34	552	Yes	Yes

[00258] **Table 1** shows CHT EC genes identified by tomo-seq and tissue-specific RNA-seq. Asterisks (*) indicates that gene is within 100 kb of TSS; some genes are associated with multiple elements. Cross (†) indicates that no expression was observed by WISH. Double cross (‡) indicates that WISH was not attempted but CHT expression was reported on the world web wide available at zfin.org. Double S (§) indicates that WISH was not attempted.

[00259] **Table 2** shows CHT EC-specific and pan-EC ATAC-seq elements that were fused to a minimal promoter and GFP and injected into one-cell stage zebrafish embryos. Asterisks (*) indicates coordinates of MACS2 peak. Cross (†) indicates expression in CHT ECs for CHT EC elements and in vessels throughout the embryo for pan-EC elements. Double cross (‡) indicates lack of NHR motif.

[00260] Table 3: CHT-enriched genes identified by tomo-seq

Gene	Full Gene Name	CHT Expression Confirmed by WISH
abi3bp	*ABI family, member 3 (NESH) binding protein*	Yes
ACKR3	*atypical chemokine receptor 3*	No*
adam8a	*ADAM metallopeptidase domain 8a*	Yes
adra1d	*adrenoceptor alpha 1D*	No*
adrb3b	*adrenoceptor beta 3b*	No*
agrp	*agouti related neuropeptide*	No*
ANGPT4	*angiopoietin 4*	No*
ap1b1	*adaptor-related protein complex 1, beta 1 subunit*	Yes
aplnra	*apelin receptor a*	Yes
aqp7	*aquaporin 7*	Yes
atp1a1a.2	*ATPase Na+/K+ transporting subunit alpha 1a, tandem duplicate 2*	n/a†
atp1b1b	*ATPase, Na+/K+ transporting, beta 1b*	n/a†
ba1	*ba1 globin*	n/a†
blf	*bloody fingers*	n/a†
BX005069.4	*leukocyte cell-derived chemotaxin-2-like*	n/a†
BX323861.1	*SLAM family member 9-like isoform X2*	n/a†
C10H8orf4		n/a†
ca15a	*carbonic anhydrase XVa*	n/a†
CABZ01049362.1	*PREDICTED: GTPase IMAP family member 4-like [Danio rerio], or 8-like*	n/a†
CABZ01058863.1		n/a†
CABZ01066695.1	*PREDICTED: protein lyl-1-like isoform X1 [Danio rerio]./..lymphocytic leukemia protein*	n/a†
ccdc88b	*coiled-coil domain containing 88B*	n/a†
ccr9a	*chemokine (C-C motif) receptor 9a*	Yes
CD209	*CD209 molecule*	n/a†
cd28	*CD28 molecule*	Yes
ceacam1	*carcinoembryonic antigen-related cell adhesion molecule 1*	Yes
ch25hl2	*cholesterol 25-hydroxylase like 2*	n/a†
cldn11a	*claudin 11a*	Yes
cldng	*claudin g*	n/a†
cltca	*clathrin, heavy chain a (Hc)*	Yes
cmklr1	*chemokine-like receptor 1*	n/a†
cndp2	*carnosine dipeptidase 2*	Yes
cnn1a	*calponin 1, basic, smooth muscle, a*	Yes
COL19A1	*collagen, type XIX, alpha 1*	Yes
col28a1	*collagen, type 28, alpha 1*	n/a†
coro1a	*coronin, actin binding protein, 1A*	n/a†
cpa5	*carboxypeptidase A5*	n/a†
CR381673.2	*Natural killer cell receptor 2B4-like isoform x1/2 or SLAM family member7/9-*	n/a†
ctsh	*cathepsin H*	Yes
ctsla	*cathepsin La*	Yes
CU463790.1		Yes
CU861664.1	*PREDICTED: zinc finger protein 521-like [Danio rerio]*	n/a†
CU915778.1	*CU915778.1*	No*
cyp24a1	*cytochrome P450, family 24, subfamily A, polypeptide 1*	No*
cysltr1	*cysteinyl leukotriene receptor 1*	Yes
dab2	*Dab, mitogen-responsive phosphoprotein, homolog 2 (Drosophila)*	n/a†
drl	*draculin*	n/a†
ela2	*elastase 2*	Yes
ENSDARG00000075833	*lymphatic vessel endothelial hyaluronic receptor 1a/b*	Yes
entpd2a.1	*ectonucleoside triphosphate diphosphohydrolase 2a, tandem duplicate 1*	n/a†
exoc3l2a	*exocyst complex component 3-like 2a*	Yes
f2r	*coagulation factor II (thrombin) receptor*	Yes
foxi3b	*forkhead box I3b*	n/a†
frem1b	*Fras1 related extracellular matrix 1b*	Yes
gcm2	*glial cells missing homolog 2 (Drosophila)*	Yes
glud1a	*glutamate dehydrogenase 1a*	Yes
glula	*glutamate-ammonia ligase (glutamine synthase) a*	Yes
GMIP	*GEM interacting protein*	n/a†

gpr182	*G protein-coupled receptor 182*	Yes
grap2b	*GRB2-related adaptor protein 2b*	n/a†
gsto1	*glutathione S-transferase omega 1*	Yes
havcr1	*hepatitis A virus cellular receptor 1*	n/a†
hbaa1	*hemoglobin, alpha adult 1*	n/a†
hdr	*hematopoietic death receptor*	n/a†
hexb	*hexosaminidase B (beta polypeptide)*	Yes
hyal2a	*hyaluronidase 2*	n/a†
hyal2b	*hyaluronidase 2*	Yes
HYAL2	*hyaluronidase 2*	Yes
ifi30	*interferon, gamma-inducible protein 30*	Yes
il10ra	*interleukin 10 receptor, alpha*	n/a†
il13ra2	*interleukin 13 receptor, alpha 2*	Yes
il6r	*interleukin 6 receptor*	Yes
ITGAE	*integrin, alpha E, tandem duplicate 1/2*	No*
itgb2	*integrin, beta 2*	Yes
kcnj1a.3	*potassium inwardly-rectifying channel, subfamily J, member 1a, tandem duplicate 3*	n/a†
kcnj1a.5	*potassium inwardly-rectifying channel, subfamily J, member 1a, tandem duplicate 5*	n/a†
lamp2	*lysosomal-associated membrane protein 2*	Yes
lgals9l1	*lectin, galactoside-binding, soluble, 9 (galectin 9)-like 1*	n/a†
lgmn	*legumain*	Yes
lpar5a	*lysophosphatidic acid receptor 5a*	n/a†
mafbb	*v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog Bb*	n/a†
man2b2	*mannosidase, alpha, class 2B, member 2*	Yes
marco	*macrophage receptor with collagenous structure*	n/a†
MCOLN2	*mucolipin 2[WM2]*	n/a†
mir142a	*micro RNA 142a*	n/a†
mmp13a	*matrix metallopeptidase 13a*	n/a†
MOV10L1	*putative helicase Mov10l1 [Danio rerio]. 95% ident.*	n/a†
mpx	*myeloid-specific peroxidase*	n/a†
mrc1a	*mannose receptor, C type 1a*	Yes
mrc1b	*mannose receptor, C type 1b*	n/a†
myh11a	*myosin, heavy chain 11a, smooth muscle*	Yes
myha	*myosin, heavy chain a*	n/a†
myo1f	*myosin IF*	Yes
ncf1	*neutrophil cytosolic factor 1*	n/a†
npl	*N-acetylneuraminate pyruvate lyase (dihydrodipicolinate synthase)*	n/a†
ostf1	*osteoclast stimulating factor 1*	n/a†
parvg	*parvin, gamma*	Yes
pdia2	*protein disulfide isomerase family A, member 2*	n/a†
PLCXD1	*phosphatidylinositol-specific phospholipase C, X domain containing 1*	n/a†
plek	*pleckstrin*	n/a†
polm	*polymerase (DNA directed), mu*	n/a†
prcp	*prolylcarboxypeptidase (angiotensinase C)*	n/a†
pxk	*PX domain containing serine/threonine kinase*	No*
rasal3	*RAS protein activator like 3*	n/a†
RNF223	*ring finger protein 223*	n/a†
s1pr4	*sphingosine-1-phosphate receptor 4*	n/a†
sele	*selectin E*	Yes
sepp1a	*selenoprotein P*	Yes
setx	*senataxin*	Yes
si:ch1073-429i10.1	*si:ch1073-429i10.1*	Yes
si:ch211-214p16.1	*si:ch211-214p16.1*	No*
si:ch211-214p16.2	*si:ch211-214p16.2*	Yes
si:ch211-250g4.3	*PREDICTED: nesprin-1 isoform X4 [Danio rerio]*	n/a†
si:ch211-284o19.8	*si:ch211-284o19.8*	n/a†
si:ch211-285f17.1	*si:ch211-285f17.1*	n/a†
si:ch211-67e16.2	*cd28-like molecule*	n/a†
si:ch73-248e21.7	*si:ch73-248e21.7*	n/a†
si:ch73-27e22.6	*si:ch73-27e22.6*	n/a†
si:dkey-102g19.3	*si:dkey-102g19.3*	n/a†
si:dkey-188i13.7	*interferon alpha inducible protein 46*	n/a†

si:dkey-237j10.2	*si:dkey-237j10.2*	n/a†
si:dkey-33i11.4	*si:dkey-33i11.4*	Yes
si:dkey-69c1.1	*si:dkey-69c1.1*	n/a†
skap2	*src kinase associated phosphoprotein 2*	n/a†
sla1	*src-like-adaptor 1*	n/a†
slc16a9a	*solute carrier family 16, member 9a*	No*
slc4a11	*solute carrier family 4, sodium borate transporter, member 11*	n/a†
snx8a	*sorting nexin 8a*	Yes
srgn	*serglycin*	Yes
stab1	*stabilin 1*	Yes
stab2	*stabilin 2*	Yes
syk	*spleen tyrosine kinase*	n/a†
tagapb	*T-cell activation RhoGTPase activating protein b*	n/a†
tll1	*tolloid-like 1*	Yes
tmem106a	*transmembrane protein 106a*	Yes
tnfsf12	*TNF superfamily member 12*	n/a†
tnni1b	*troponin I type 1b (skeletal, slow)*	Yes
tubb1	*tubulin, beta 1 class VI*	n/a†
wasa	*Wiskott-Aldrich syndrome (eczema-thrombocytopenia) a*	n/a†
wasb	*Wiskott-Aldrich syndrome (eczema-thrombocytopenia) b*	n/a†
WIPF1	*WAS/WASL interacting protein family, member 1a/b*	n/a†
zgc:158446	*complement factor b, like*	n/a†
zgc:174945	*zgc:174945*	n/a†
zgc:198419	*ferritin, heavy polypeptide-like 28*	Yes

[00261] Asterisk (*) in **Table 3** indicates that no CHT expression was observed by WISH. Dagger symbol (†) in **Table 3** indicates WISH was not attempted.

[00262] **Table 4: Transcription factor expression in CHT ECs**

Transcription Factor	Family	FPKM	Associated with CHT EC ATAC-seq Element Containing Ets, Sox and NHR Sites*	Genomic Coordinates of Representative Element
fli1a	Ets	480.4	Yes	chr18:46966409-46966698
etv2	Ets	192.3	Yes	chr16:44782409-44782895
ets1	Ets	183	Yes	chr18:46883643-46884100
sox18	SoxF	206.4	Yes	chr23:8886011-8886744
sox7	SoxF	125.1	Yes	chr20:19158376-19158663
nr2f2	NHR	84.6	Yes	chr18:23728906-23729747
rxraa	NHR	45.9	Yes	chr21:16411020-16411531

[00263] **Table 4** shows FPHM expression values in CHT ECS for highly expressed members of the Ets, Sox and NHR transcription factor families. Asterisk (*) in **Table 4** indicates within 100 kb of TSS; some genes are associated with multiple elements.

[00264] Table 5: Transcription factor expression in mouse hematopoetic niche

Transcription Factor	Family	Mouse E14-E15 Liver EC FPKM	Mouse E16-E17 Liver EC FPKM	Mouse Adult Bone Marrow EC FPKM
Ets1	Ets	218.4666	251.9493	153.2657
Erg	Ets	46.64156	78.53131	45.14673
Elk4	Ets	9.369453	11.4226	22.83457
Elk1	Ets	7.003965	9.08418	6.8779
Etv1	Ets	2.203135	3.10327	1.488542
Etv2	Ets	0.235977	0	0
Sox18	SoxF	127.1509	262.44	130.1783
Sox7	SoxF	49.94503	46.9365	19.80563
Sox17	SoxF	33.01219	68.37438	90.24645
Sox11	SoxF	12.01665	11.1584	0.67509
Sox12	SoxF	11.81507	21.5267	0.556478
Sox6	SoxF	1.741399	1.158524	0.51182
Sox5	SoxF	0.193041	0.289841	0.437005
Sox9	SoxF	0.119527	0.072563	0
Nr2f2	NHR	58.97832	103.5558	63.17458
Rxra	NHR	23.98264	33.0942	22.08392
Rara	NHR	19.29841	27.37294	13.93433
Nr4a2	NHR	10.13413	3.130986	30.30394
Esrrb	NHR	6.219864	7.884516	0.586381
Rora	NHR	1.219604	1.086872	5.922048

[00265] **Table 6: Primers used for WISH probe synthesis**

Category	Gene	Forward*	Reverse†
CHT EC enriched	*adra1d*	GCTCCATAGTATCGTCTGAACC	AAACCATTGCCATTTTGCCA
CHT EC enriched	*ap1b1*	GGGAGTTCTTCGGGTGACTG	GCTTGCAACAAAAAGCGCAG
CHT EC enriched	*cldn11a*	TGTGTGATCTCAACTGCGCT	GGTGCAATCTAGTCTGATCGGT
CHT EC enriched	*citca*	CCAGCAAACCCCATGGATCT	AACCGAGTACAGGACACACG
CHT EC enriched	*ctsh*	CGACTGGAGAACCAAGGGAC	TGGAGGCTAATCGAGTGTGC
CHT EC enriched	*ctsla*	CCATGCAACAGAGGAAGGGT	TACTGGGCGGGTCTCCTTTA
CHT EC enriched	*exoc3l2a*	AAGTTCCGCAGGATGGACTG	TCGCTTGTGTGATCAAGTATGAC
CHT EC enriched	*glula*	AGTTATGCCAGCTCAGTGGG	GGCCTCCCCAAGAAACCATT
CHT EC enriched	*gpr182*	CTTCCCACAGCAGCACAAAC	GAAAGTTGTTGTTGAAGTGAACG
CHT EC enriched	*hexb*	GAATTTGCTCGCATGAGGGG	CGGCAGTGGCCAACAAATAG
CHT EC enriched	*hyal2b*	ATGGAGGTCTACCACAGGCT	AGTGCAGGTATGTGTCCGTG
CHT EC enriched	*ifi30*	TTCGGCTTTAACCTGTGCGT	CCTGACGCGAGTAGTGTTGT
CHT EC enriched	*il13ra2*	AGTTAGAATGGGCGCCACC	GGCAAGACCACTGGCATTTG
CHT EC enriched	*igmn*	AACTTGAGCCACCGAGGATTT	CCCTAACTCCAGCACACACT
CHT EC enriched	*lyve1b*	GCTACAGTCTGCGTAGCAT	TGGAAGCAGCTCTAAGTGACAG
CHT EC enriched	*man2b2*	TACCCAATGGTTCGAGTGGC	GCTTAGGTGATCAATTTTGGGACA
CHT EC enriched	*mrc1a*	GTGTCCCCTCATCAATGCCA	ACGGCATTCCACAAACCAGA
CHT EC enriched	*sele*	TGCCCAGCCCTTGATAATCT	ACCCAAACTGACTTTATATGTGC
CHT EC enriched	*sepp1a*	AGGCAGCACTGGACTTTAGC	AGGTACAAATGCAAGTACAACACT
CHT EC enriched	*snx8a*	ACAAAGAGATCTGCATTCCAAGC	AGCCTGTCAGCTCACTTTATT
CHT EC enriched	*stab1*	AAGGCGTACTATGTCCTCAGGC	CGCCGTTCTATAATGCACCG
CHT EC enriched	*stab2*	TTGTGGATTACGGGGTTCGG	AAAGAGAGCTGCACCGACT
CHT EC enriched	*tll1*	GAGCTTTACTCTGCTGGCGA	ACAAATGATGTCTGTCTCCGCT
CHT enriched	*abi3bp*	CTGTTTTCCCCCACCAGTGA	CAAAGGATTGGCAGGGACCA
CHT enriched	*ackr3*	TGGGATTTATTTGTAACTACACGGA	TTTTAAGCACATTTCTGAAGCACA
CHT enriched	*adam8a*	CCAGGAAGCGCAAAGAACAG	ACATTAGGCGGGCAAAACAAA
CHT enriched	*adrb3b*	GCAGCAAACGACTGCTACAA	CCCACTTCGCTGCTCTTTAC
CHT enriched	*agrp*	TCATCCACACCTGAGACGCA	ACACCTTAAAACCGCAGCC
CHT enriched	*angpt4*	ATCCGACTGCTGGAATGGAC	GCTTTGAGGAGCTTAAGAGGC
CHT enriched	*aplnra*	GTGCTGGTCAACATGTACGC	CGTCACTTTTCACCCCCAGA
CHT enriched	*aqp7*	TCCACTGGGAAAAGCTGGAAT	TTTCAGATGCAGCACAGGCA
CHT enriched	*bnip3lb*	ATGGGGCTGACGGATACC	GCACAGGAAACGCACATGAT
CHT enriched	*ccr9a*	TTGTCCAGACTACCAAGGCG	TTACTTCACTGCCAGTCGGC
CHT enriched	*cd28*	ATCCAAACTGAGGCCGGAAG	AGAAAATACAGTGCATACATGTCA
CHT enriched	*ceacam1*	GGCCCAAGCATGGCAGAAAC	CCTACAAGCCTCATTCAGACAGT
CHT enriched	*cndp2*	ACATGGGACATGGAGCGAAG	ACACTAGAAAACCGATCGTGTCA
CHT enriched	*cnn1a*	GACTCTCTGCGGATGTCAGG	GGTCATGCCCTTTTGGCTTG
CHT enriched	*col19a1*	CATGTCCACCCCTGAAGCTG	GGGTTCTGTTGTGGAGTGCT
CHT enriched	*cu463790.1*	GGCGTCTCTTTTTCTGCTGC	TGACGCTTAAACAGAGCGGT
CHT enriched	*cu915778.1*	CCCTAGTGTCCGAGGTCTCA	TTTCCCCTGTGTGGATGAGC
CHT enriched	*cyp24a1*	GATACCGTGCTGGGCGATTA	CCACCACTCACTCATTCAGACA
CHT enriched	*cysltr1*	TCCCGGTGCAAAATCTGAGG	AGTCATGCACAAAATCTGCGG

CHT enriched	*ela2*	GTTTATTGCTGGCGCCTACG	TTCTTGGGGTAGTTGCAGCC
CHT enriched	*f2r*	GCTGCCGAACAACGAAACAT	TAGGACGCGTCATTGTGCTT
CHT enriched	*frem1b*	AGTACACTCCGGACCCAAGA	CACCAGAAAGAATGTCACCGT
CHT enriched	*gcm2*	TCCAGAGCGATTCAGCATCA	CAGTCCCTCAGTATTCCCCG
CHT enriched	*glud1a*	AGTCTCCTACTTCGAGTGGCT	ACGCCTGAGATTCATCCTGC
CHT enriched	*il6r*	AACTGTTCTTTCTCCCGGTCCC	CCTCTGGCTGAACAGGAAGG
CHT enriched	*itgae*	ACTGGTCAACCACCTCCTCT	ACACAATCAGGCAAGGTCTC
CHT enriched	*itgb2*	TGCCTTTCAAAGTGGACCGT	ACCAGTCACACCAGCCATTC
CHT enriched	*lamp2*	AGCCTGTTCCTGGACCATTG	AGCTACAACCATTGAGGGCT
CHT enriched	*myh11a*	GGTTCGCCAGAAGGACAAGA	AGCATCCAAAAGTACTCGGTGT
CHT enriched	*myo1f*	AAGCTGTCATCAAAGCCGGA	TTCTCGACCTGTCAGCTGTT
CHT enriched	*parvg*	TGAAAGCCCTGAACGAGACC	CGTCAGCATCCAAACGCAAT
CHT enriched	*setx*	AGGAGTTTGGCTTCGACCAG	GTGACGCTGGAATATCCCGT
CHT enriched	*si:ch1073-429i10.1*	TCGCTCTGATGCTCAGCTTG	CACTCGGCGACAGTATTCCC
CHT enriched	*si:ch211-214p16.1*	TACACATTTTCTGCCCCACTGA	AATGGGGCAAGAGTCCATCT
CHT enriched	*si:ch211-214p16.2*	CTCACCCTCGGTCCAGAACT	ACAGACACACTTGCCAGTCA
CHT enriched	*si:dkey-33i11.4*	ACAGCCATCAGTTCCTCTGC	AGCTTTGCATCCCCATCACT
CHT enriched	*srgn*	GGAAGCCACTCCTGATACGG	GTACAACATTTACTTGCTGTCCA
CHT enriched	*tmem106a*	GGTCACGCACCAAATGAACC	AACAGTTCTGATTGGATTTTGCTCA
CHT enriched	*tnni1b*	TCTGCATCTCGCAAGCTGAT	CATGTGTAGTGCAGACAGAACA
CHT enriched	*zgc:198419*	AGAACTACGACAGCGACTGC	GGTTTTGGATAAGAGCTGTGTCA
Transcription Factor	*ets1*	ACAGACTCTGTACGTTTGAATGCGT	GTCCAGACTTTACTCGTCCGTGTC
Transcription Factor	*etv2*	TATGACTGCAGTGGTGAAGACC	CTTTCCCGCCGTTTTGTGAA
Transcription Factor	*fli1a*	CAGACCCGTCTCTGTGGTC	CCAGTATGGGGTTGTGGGAC
Transcription Factor	*nr2f2*	ACCCCCGAACAACAATAACA	AGAGGGCAAGCGCAGTAATA
Transcription Factor	*sox7*	TATAGCCCTTCGTTCCCCCA	ACCGAAACCGGCTAAACTGA
Transcription Factor	*sox18*	TCCTTGGACGCTGTGGACCAAC	TCAAAGCGCTGCTTTCCTCGC

[00266] Asterisk (*) in **Table 6** indicates that the T3 sequence CATTAACCCTCATAAAGGGAA (**SEQ ID NO: 10**) was added to the 5' end of each forward primer. Dagger symbol (†) in **Table 6** indicates that The T7 sequence TAATACGACTCACTATAGGG (**SEQ ID NO: 11**) was added to the 5' end of each reverse primer.

[00267] Table 7: Primers used to clone promoter and enhancer elements

Type of Element	Gene	Genomic Coordinates of ATAC-seq Element	Amplicon Size (bp)	Forward	Reverse
5' upstream of TSS	*mrc1a* (1.3 kb)	chr7:65,468,213-65,469,565	1353	CTTTTGGCCATTACTGCCG	TTCTGTCTTTTAATCAGCAATCC
CHT EC element	*mrc1a* (125 bp)	chr7:65469086-65469210	125	GCTCTCAGTTCCTGGTATTTTTCT	TGAAGCTTGTACCTTTCATTTCC
5' upstream of TSS	*sele* (5.3 kb)	chr20:34,001,481-34,006,781	5301	TCGTTACTGCACTTGAAAGCGT	TATCAGTGATGTTCTGCAGTGGTC
CHT EC element	*sele* (158 bp)	chr20:34004805-34004962	158	CCATGAAACTGGGAAGATGAA	CAGGAAGAAATAATGGCAAAAA
CHT EC element	*ap1b1*	chr5:26463217-26463695	750	GAAGCTCTCCAGCAGCTCA	CATTTCCACCAGCTGTCTGAT
CHT EC element	*cltca*	chr10:29,047,274-29,047,619	404	GCTGTCAGCACATTCTTTTCC	CCCTGCTGATCACACATGAC
CHT EC element	*dab2*	chr5:33,980,000-33,980,306	394	ACTGCTCCTCACCAATCGTC	TGCACTAAATCTGTGCCAAGTC
CHT EC element	*exoc3l2a*	chr5:38359097-38359903	901	TTTATATAATCGGAAGGAACCTTTTT	TCCTGTCAGCTGTTTTCATCC
CHT EC element	*glula*	chr2:19,458,704-19,459,047	446	GGCAAAATGCTTAGATGCAGA	TGCGAGGAGGACATAAACAA
CHT EC element	*gpr182*	chr23:36701205-36701682	481	TAGCCTTGTGCAATGCTTGT	TGCTGAATTCAAAAGCCACTT
CHT EC element	*gpr182*	chr23:36694073-36694476	398	CACTTCTGGTACCAAATGATCAAC	GAGGGTTAAACGTGGCCTTA
CHT EC element	*gpr182*	chr23:36696363-36696656	577	GCGGCAAACTTTTTGAGTGT	GCCAGCCTCAAAGTTTGTTCT
CHT EC element	*lgmn*	chr13:36,448,465-36,448,818	414	CGCGTGATGAGGATCTGATT	GGTGTTGAAAGGTGATGCTG
CHT EC element	*prcp*	chr15:10,400,588-10,400,868	334	AAAATTAAGAGCGGGCAGACT	TGGAAACAACAACAGCCTGA
CHT EC element	*sele*	chr20:34,010,027-34,010,326	398	AAAGCACTTGATTGAGAATTGC	TGTTTGGTTCAGTTACACGTTTT
CHT EC element	*sele*	chr20:34,011,251-34,011,563	360	CAGTTTCCCAAGCTTCAAGG	TGTGATTACACATTCCCACACAT
CHT EC element	*snx8a*	chr3:42,090,805-42,091,062	395	AATGGTTGCAGCATTGTGTT	GCTTTTGTTTGGTGATGTGC
CHT EC element	*stab1*	chr22:10467346-10467937	874	GTTACCTGGCAACCACCAAC	TGGTCAGAATAAGCACGTTTCA
CHT EC element	*stab2*	chr4:9790795-9791116	422	ACGTTAACAAAGGCGATGTTTT	TCTAAACAATTTTTAAGGTAAACCAAA
Pan-EC element	*cdh5*	chr7:45457842-45458791	823	TGACAGGACTCATCAGCACG	AATAGTCTCTGGTCTGCTGTTAAA
Pan-EC element	*clec14a*	chr17:10362325-10362844	455	TGGGAAAAATACCAGGAAGCGT	AAGCAGCGAGCTCTCATAATAAA
Pan-EC element	*dll4*	chr20:28219013-28219619	452	AGATCAATGAGAGCGAGGCG	GGAGCAGATGAGGTTAAGTCCT
Pan-EC element	*fli1a*	chr18:47039842-47040466	800	CGGACAGTAATGTCTGGATGG	CCACAACTCCATACTGGGAAA
Pan-EC element	*lmo2*	chr18:36722030-36722527	367	TCATCATGGCCAACAGAATG	GTGCAGGAAATGAGCACAGA
Pan-EC element	*nrp1b*	chr2:43535098-43535801	552	TGACTCAACCAATCAATCAGCCT	TAGCAAAGCTCTCAGGCCC

[00268] Table 8: Sequences and primers for mutational variants of the 125 bp *mrc1a* and 158 bp *sele* enhancer elements

Gene	Fragment Name	Total Fragment Sequence	Forward Primer	Reverse Primer
mrc1a	Wild-type	CCATGAAACTGGGAAGATGAAAGCATT AGTTGAATTGTTACTGGCAACATCTTCT CTGTAATGCCCCCTGTGACCCATATTG TCTCGCTCTTTCCTTTATAAACAGAGCT GTAGATATCCACAGGAAATGGGGGTGT TTTTGCCATTATTTCTTCCTG	TGAAGCTTGTACCTTTCATTTCCTTTTTGC TGAGCTTTATTTTCTCTAGAATTGCCATT GTGTTTCCATTCTAG	GCTCTCAGTTCCTGGTATTTTTCTTTCAGCTG AAAAAAAAAATGCTGATTTGCTAGAATGGAA ACACAATGGCAAT
mrc1a	Ets mutant	TGAAGCTTGTACCTTTCATTTaaaTTTTG CTGAGCTTTATTTTCTCTAGAATTGCCA TTGTGTTTCCATTCTAGCAAATCAGCAT TTTTTTTTCAGCTGAAAGAAAAATACCA tttACTGAGAGC	TGAAGCTTGTACCTTTCATTTaaaTTTTGC TGAGCTTTATTTTCTCTAGAATTGCCATT GTGTTTCCATTCTAG	GCTCTCAGTaaaTGGTATTTTTCTTTCAGCTG AAAAAAAAAATGCTGATTTGCTAGAATGGAA ACACAATGGCAAT
mrc1a	Sox mutant	TGAAGCTTGTACCTTTCATTTCCTTTTTG CTGAGCgggcggggaTCTAGAATTGCacg gtgGTTTCCATTCTAGCAAATCAGCcggg gggTTTCAGCTGAAAGAAAAATACCAG	TGAAGCTTGTACCTTTCATTTCCTTTTTGC TGAGCgggcggggaTCTAGAATTGCacggtg GTTTCCATTCTAG	GCTCTCAGTTCCTGGTATTTTTCTTTCAGCTG AAAccccccgGCTGATTTGCTAGAATGGAAAC caccgtGCAAT
mrc1a	NHR mutant	GAACTGAGAGC attAGCagatTtaaTTTCATTTCCTTTTTGCa ttaaTTTATTTTCTCTAGAATTGCCATTGT GTTTCCATTCTAGCAAATCAGCATTTTT TTTTCAGCTGAAAGAAAAATACCAGGA	attAGCagatTtaaTTTCATTTCCTTTTTGCatt aaTTTATTTTCTCTAGAATTGCCATTGTGT TTCCATTCTA	GCTCTCAGTTCCTGGTATTTTTCTTTCAGCTG AAAAAAAAAATGCTGATTTGCTAGAATGGAA ACACAATGGCAAT
mrc1a	Control mutant	ACTGAGAGC TGAAGCTTGTACCTTTCATTTCCTTTTTG CTGAGCTTTATTTTCTCTAGAATTGCCA TTGTGTTTCCATTCcgGCAAATCAGCAT TTTTTTTTCAGCTGAccGAAAAATACCA	TGAAGCTTGTACCTTTCATTTCCTTTTTGC TGAGCTTTATTTTCTCTAGAATTGCCATT GTGTTTCCATTCTAG	GCTCTCAGTTCCTGGTATTTTTCggTCAGCTG AAAAAAAAAATGCTGATTTGCcgGAATGGAAA CACAATGGCAAT
sele	Wild-type	GGAACTGAGAGC CCATGAAACTGGGAAGATGAAAGCATT AGTTGAATTGTTACTGGCAACATCTTCT CTGTAATGCCCCCTGTGACCCATATTG TCTCGCTCTTTCCTTTATAAACAGAGCT GTAGATATCCACAGGAAATGGGGGTGT TTTTGCCATTATTTCTTCCTG	CCATGAAACTGGGAAGATGAAAGCATTA GTTGAATTGTTACTGGCAACATCTTCTCT GTAATGCCCCCTGTGACCCATATTGTCT CGCTCT	CAGGAAGAAATAATGGCAAAAACACCCCC ATTTCCTGTGGATATCTACAGCTCTGTTTATA AAGGAAAGAGCGAGACAATATGGGTCACAG
sele	Ets mutant	CCATGAAACTGGGAAGATGAAAGCATT AGTTGAATTGTTACTGGCAACATCTTCT CTGTAATGCCCCCTGTGACCCATATTG TCTCGCTCTTTaaaTTATAAACAGAGCT GTAGATATCCACAtttAATGGGGGTGTTT TTGCCATTATTTCTaaaTG	CCATGAAACTGGGAAGATGAAAGCATTA GTTGAATTGTTACTGGCAACATCTTCTCT GTAATGCCCCCTGTGACCCATATTGTCT CGCTCT	CAtttAGAAATAATGGCAAAAACACCCCCATT aaaTGTGGATATCTACAGCTCTGTTTATAAtttA AAGAGCGAGACAATATGGGTCACAG
sele	Sox mutant	CCATGAAACTGGGAAGATGAAAGCATT AGTTGAAggtggACTGGCAACATCTTCT CTGTAATGCCCCCTGTGACCCATAggtg aTCGCTCTTTCCTTTATAAACAGAGCTG TAGATATCCACAGGAAATGGGGGTGTT TTTGCCATTATTTCTTCCTG	CCATGAAACTGGGAAGATGAAAGCATTA GTTGAAggtggACTGGCAACATCTTCTCTG TAATGCCCCCTGTGACCCATAggtgaTCG CTCT	CAGGAAGtttattaGGCAAAAACACCCCCATTT CCTGTGGATATCTACAGCTCTGTTTATAAAG GAAAGAGCGAtcaccTATGGGTCACAG
sele	NHR mutant	CCATGAAACTGGGAAGATGAAAGCATT AGTTGAATTGTTACTGGCAACATCTTCT CTGTAATGCCCCCTGattaaCATATTGTC TCGCTCTTTCCTTTATAAACAGAGCTGT AGATATCCACAGGAAATGGGGGTGTTT TTGCCATTATTTCTTCCTG	CCATGAAACTGGGAAGATGAAAGCATTA GTTGAATTGTTACTGGCAACATCTTCTCT GTAATGCCCCCTGattaaCATATTGTCTCG CTCT	CAGGAAGAAATAATGGCAAAAACACCCCC ATTTCCTGTGGATATCTACAGCTCTGTTTATA AAGGAAAGAGCGAGACAATATGttaatCAG
sele	Control mutant	CCATGAAACTGGGAAtcTGAAAGCATT AGTTGAATTGTTACTGGCAACATCTTCT CTGTAATGCCCCCTGTGACCCATATTG TCTCGCTCTTTCCTTTATAAACAGAGag GTAGATATCCACAGGAAATGGGGGacT TTTTGCCATTATTTCTTCCTG	CCATGAAACTGGGAAtcTGAAAGCATTA GTTGAATTGTTACTGGCAACATCTTCTCT GTAATGCCCCCTGTGACCCATATTGTCT CGCTCT	CAGGAAGAAATAATGGCAAAAAgtCCCCCA TTTCCTGTGGATATCTACctCTCTGTTTATAAA GGAAAGAGCGAGACAATATGGGTCACAG

[00269] In **Table 8**, lowercase letters indicate base pair changes used to disrupt transcription factor binding motifs.

[00270] **Table 9: Primers used for cloning and EMSA probe synthesis**

Category	Primer Name	Forward	Reverse	Comment
Cloning	*Nr2f2*	CGGGATCCatggca atggtagtca gcacg	CCGGAATTCCGGttgaattgccatatatggc	
Probe synthesis	*mrc1a* site 1 wild-type	tttaTGAAGCTTGTACCTTTCATTTCCTTTTTG	CAAAAAGGAAATGAAAGGTACAAGCTTCAtaaa	
Probe synthesis	*mrc1a* site 1 mutation	TTTAattAGCagatTtaaTTTCATTTCCTTTTTG	CAAAAAGGAAATGAAAttaAatctGCTaatTAAA	1st NHR site mutated like *in vivo* GFP reporter experiment
Probe synthesis	*mrc1a* site 2 wild-type	TTCATTTCCTTTTTGCTGAGCTTTATTTTC	GAAAATAAAGCTCAGCAAAAAGGAAATGAA	
Probe synthesis	*mrc1a* site 2 mutation	TTCATTTCCTTTTTGCattaaTTTATTTTC	GAAAATAAAttaatGCAAAAAGGAAATGAA	2nd NHR site mutated like *in vivo* GFP reporter experiment
Probe synthesis	*sele* wild-type	GTAATGCCCCCTGTGACCCATATTGTCTCGCTCTTTCCTTTATA	TATAAAGGAAAGAGCGAGACAATATGGGTCACAGGGGGCATTAC	
Probe synthesis	*sele* mutation	GTAATGCCCCCTGattaaCATATTGTCTCGCTCTTTCCTTTATA	TATAAAGGAAAGAGCGAGACAATATGttaatCAGGGGGCATTAC	NHR site mutated like *in vivo* GFP reporter experiment

[00271] **Table 9** shows primers used for cloning mouse Nr2f2 into the pGEX2TK vector and DNA probes from the zebrafish *mrc1a* and *sele* enhancers.

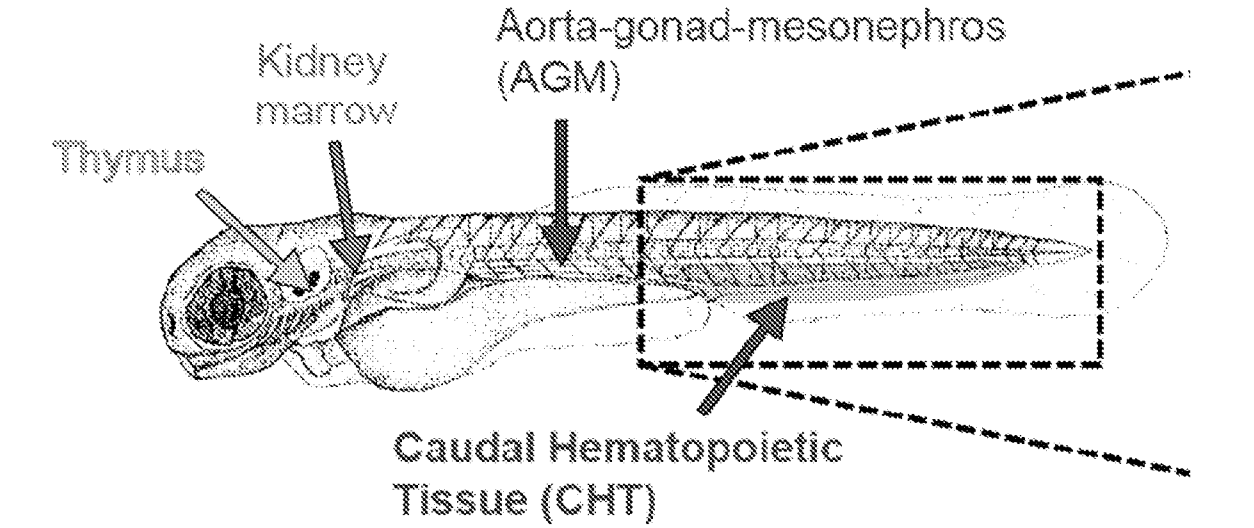
CLAIMS

What is claimed herein is:

1. A method to generate/engineer endothelial niche cells, comprising expressing one or more transcription factors in an endothelial cell, wherein the one or more transcription factors are from the Ets family, the Sox family, and/or the Nuclear Hormone Receptor family.
2. The method of claim 1, wherein the endothelial niche cells express one are more genes comprising: *sele*, *exoc312a*, *snx8a*, *cltca*, *aqp7*, *ap1b1*, *lgmn*, *prcp*, *cldn11a*, *lyve1b*, *adra1d*, *hya12a*, *hya12b*, *tll1*, *i113ra2*, *glu1a*, *hexb*, *slc16a9a*, or *sepp1a*.
3. The method of claim 1, wherein the endothelial cells are human.
4. The method of any one of the above claims, wherein the transcription factors comprises at least one of the human transcription factors ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, or NR2F2.
5. The method of claim 1, wherein the transcription factor includes at least one transcription factor from the Ets family, at least one transcription factor from the Sox family, and at least one transcription factor from the Nuclear Hormone Receptor family.
6. The method of claim 1, wherein the transcription factors include ETV2, FLI1, ETS1, SOX18, SOX7, RXRA, and NR2F2.
7. The method of claim 1, wherein the transcription factors are expressed from at least one vector.
8. The method of claim1, wherein the vector comprises an exogenous nucleic acid sequence(s) encoding the one or more transcription factors.
9. The method of claim 1, wherein the exogenous nucleic acid sequences are incorporated into the genome of the endothelial cell.
10. An engineered endothelial niche cell comprising one or more exogenous nucleic acid sequences encoding one or more transcription factors, wherein the one or more transcription factors are from the Ets family, the Sox family and/or the Nuclear Hormone Family.

11. A composition comprising the engineered endothelial niche cells of claim 10.
12. The composition of claim 11, wherein the composition is a therapeutic agent or the composition further comprises a pharmaceutically acceptable carrier.
13. The composition of claim 11, wherein the composition further comprises a culture dish, 3D cell system, or suspension system.
14. The composition of claim 11, wherein the composition comprises a scaffold.
15. A method for culturing HSPCs, the method comprising culturing HSPCs in the presence of a population of engineered endothelial niche cells.
16. The method of claim 15, wherein the method is performed *in vitro*.
17. The method of claim 15, wherein the engineered endothelial niche cells secrete a factor that affects the growth and/or expansion of the HSPC cells.
18. The method of claim 15, wherein the HSPCs cultured in the presence of the engineered endothelial niche cells can be cultured for at least 3 (e.g., at least 4, at least 5, at least 6, at least 7) days longer than HSPCs that are cultured in the absence of such engineered endothelial niche cells.
19. The method of claim 15, wherein the cells are cultured on a biologically compatible scaffold.
20. The method of claim 15, wherein the HSPCs cultured in the presence of the engineered endothelial niche cells have increased engraftment when administered to a subject compared to the engraftment of substantially similar HSPCs that were not cultured with engineered endothelial niche cells.
21. A method of treating a subject, the method comprising, transplanting a composition comprising HSPCs and a population of engineered endothelial niche-cells into the subject.
22. A method for enhancing engraftment of HSPCs, the method comprising administering a composition comprising HSPCs and a population of engineered endothelial niche cells to a subject in need thereof.
23. The method of claim 22, wherein engraftment of the HSPCs is increased by at least 10% compared to the engraftment of substantially similar HSPCs in the absence of engineered endothelial niche cells.
24. A co-culture comprising engineered endothelial niche cells and HSPCs.

25. The co-culture of claim 24 wherein the endothelial cells are made by the method of claims 1-9.
26. A kit for culturing HSPCs, the kit comprising: a population of engineered endothelial niche cells, reagents and instructions for use thereof.
27. A kit for generating engineered endothelial niche cells comprising: a vector(s) comprising one or more exogenous nucleic acid sequences encoding one or more transcription factors of the Ets family, the Sox family or the nuclear hormone family and instructions for use thereof.
28. A method for generating an ectopic vascular niche, the method comprising: administering an engineered endothelial niche cell to a target site in a subject in need thereof.
29. A method for extra medullary hematopoiesis, the method comprising transplanting engineered-niche endothelial cells into a subject at a location outside of the bone marrow (e.g., the forearm), thereby creating a synthetic niche.
30. The method of claim 29, wherein the endothelial cells are made by any of the methods in the above claims.
31. A vector comprising one or more exogenous nucleic acid sequences encoding one or more transcription factors of the Ets family, the Sox family or the nuclear hormone family operably linked to a promoter.


Aorta-gonad-mesonephros
(AGM)
Kidney
marrow
Thymus
Caudal Hematopoietic
Tissue (CHT)

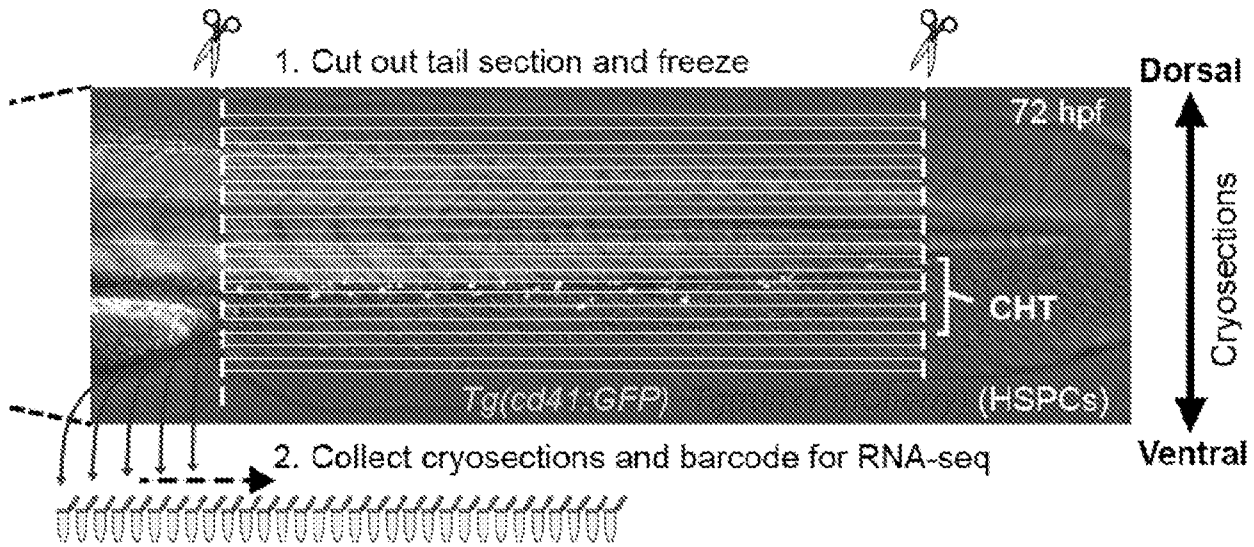

1. Cut out tail section and freeze
Dorsal
72 hpf
CHT
Cryosections
Tg(cd41:GFP)
(HSPCs)
2. Collect cryosections and barcode for RNA-seq
Ventral

Fig. 1A

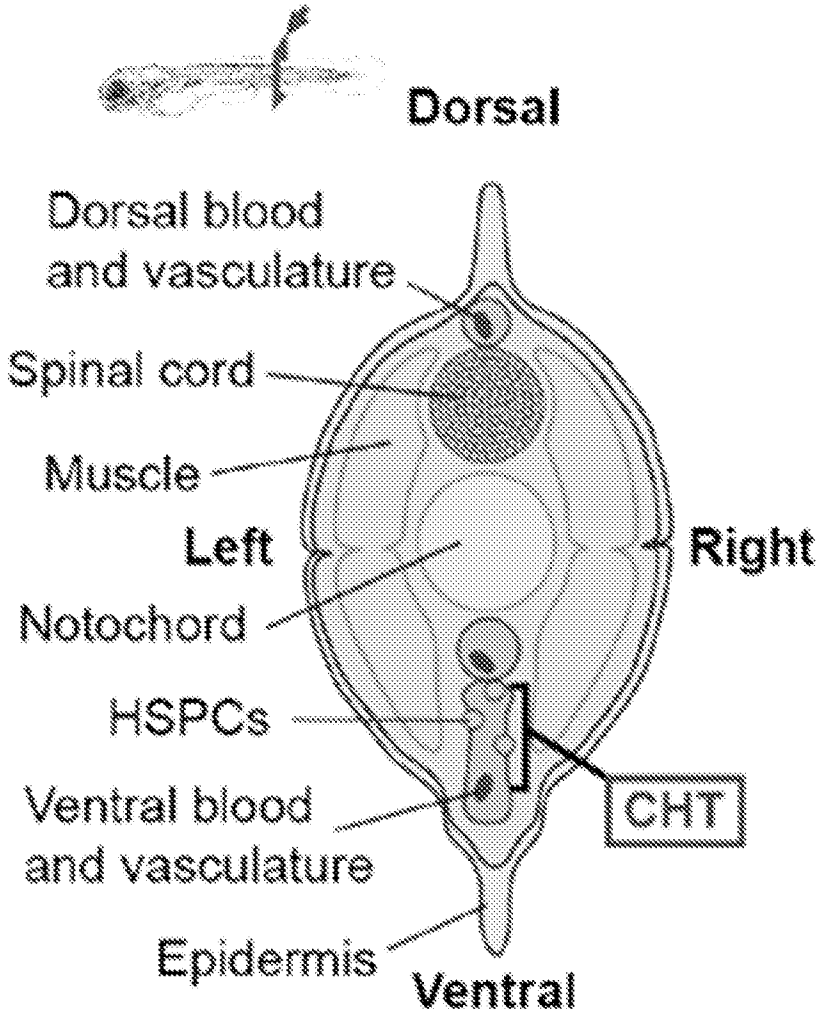

Dorsal
Dorsal blood and vasculature
Spinal cord
Muscle
Left
Right
Notochord
HSPCs
CHT
Ventral blood and vasculature
Epidermis
Ventral

Fig. 1B

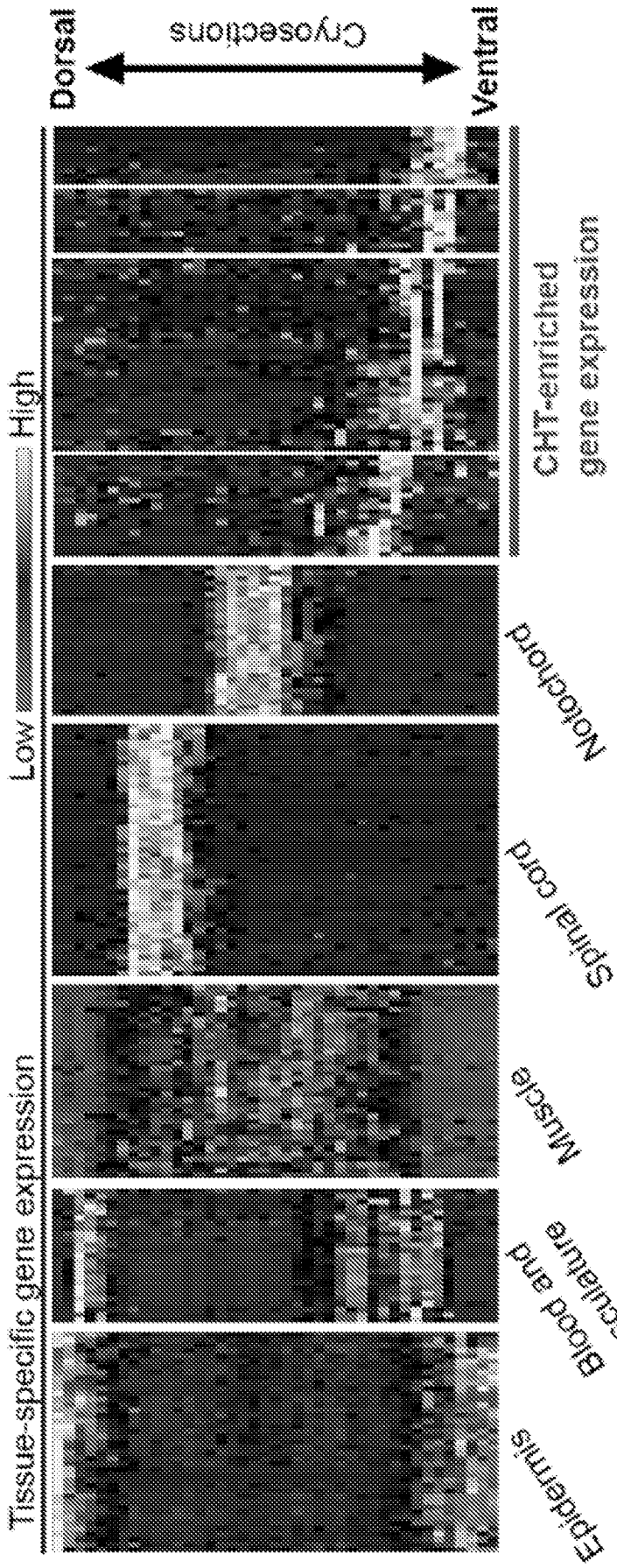

Dorsal
Cryosections
Ventral
Tissue-specific gene expression
Low
High
CHT-enriched gene expression
Notochord
Spinal cord
Muscle
Blood and vasculature
Epidermis

Fig. 1B, cont.

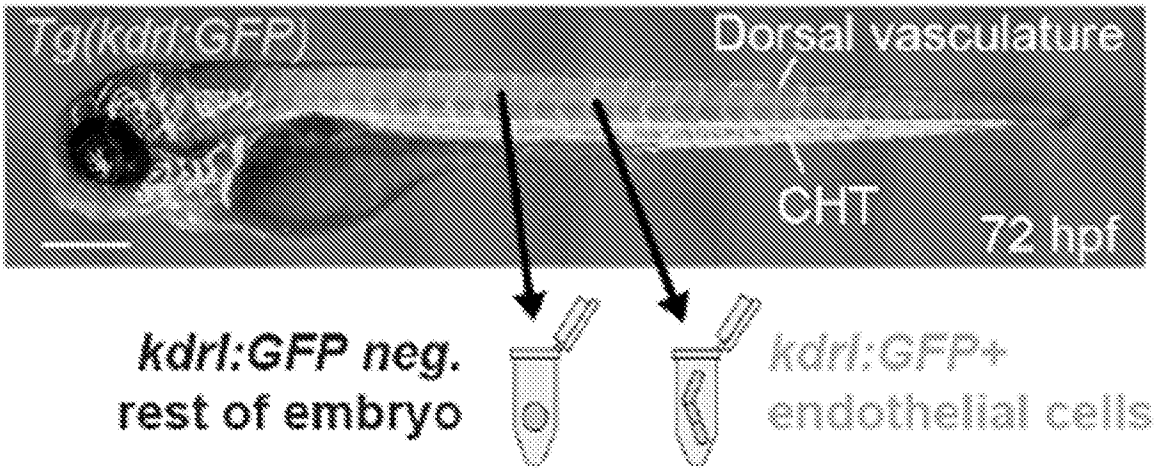

Tg(kdrl:GFP)
Dorsal vasculature
CHT
72 hpf
kdrl:GFP neg.
rest of embryo
kdrl:GFP+
endothelial cells

Fig. 1C

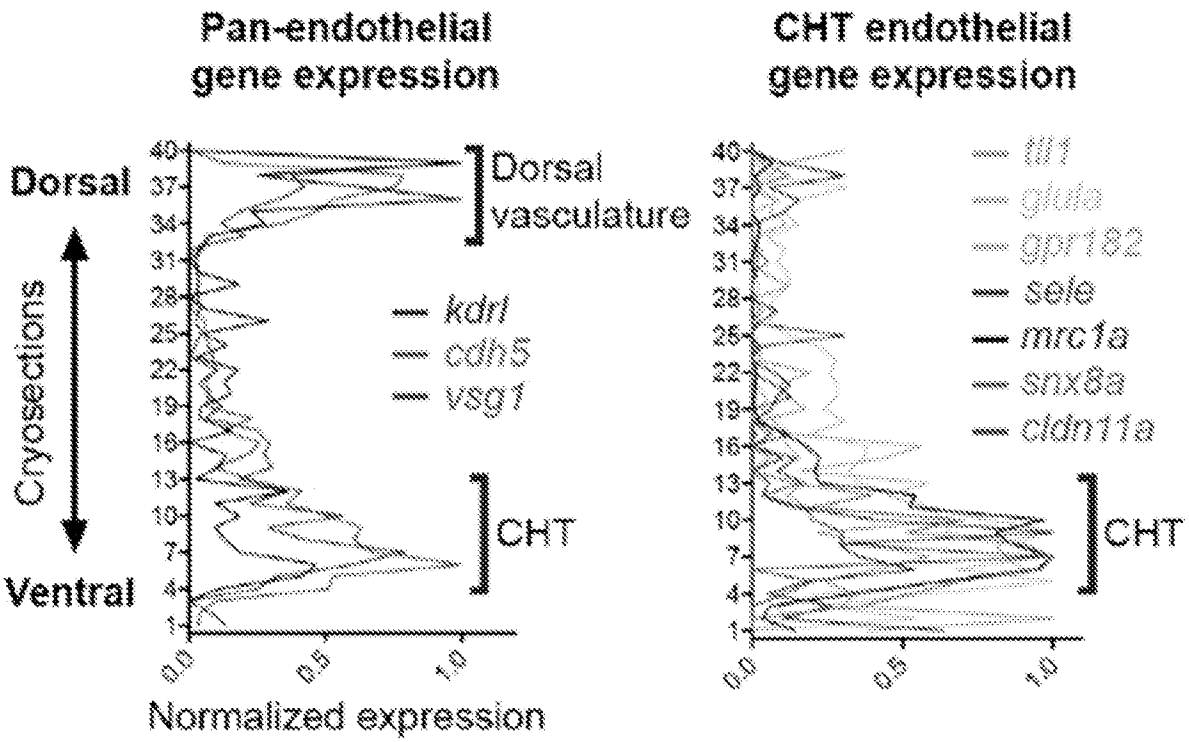

Pan-endothelial gene expression
CHT endothelial gene expression
Dorsal
Ventral
Cryosections
Dorsal vasculature
CHT
kdrl
cdh5
vsg1
tll1
glula
gpr182
sele
mrc1a
snx8a
cldn11a
Normalized expression

Fig. 1D

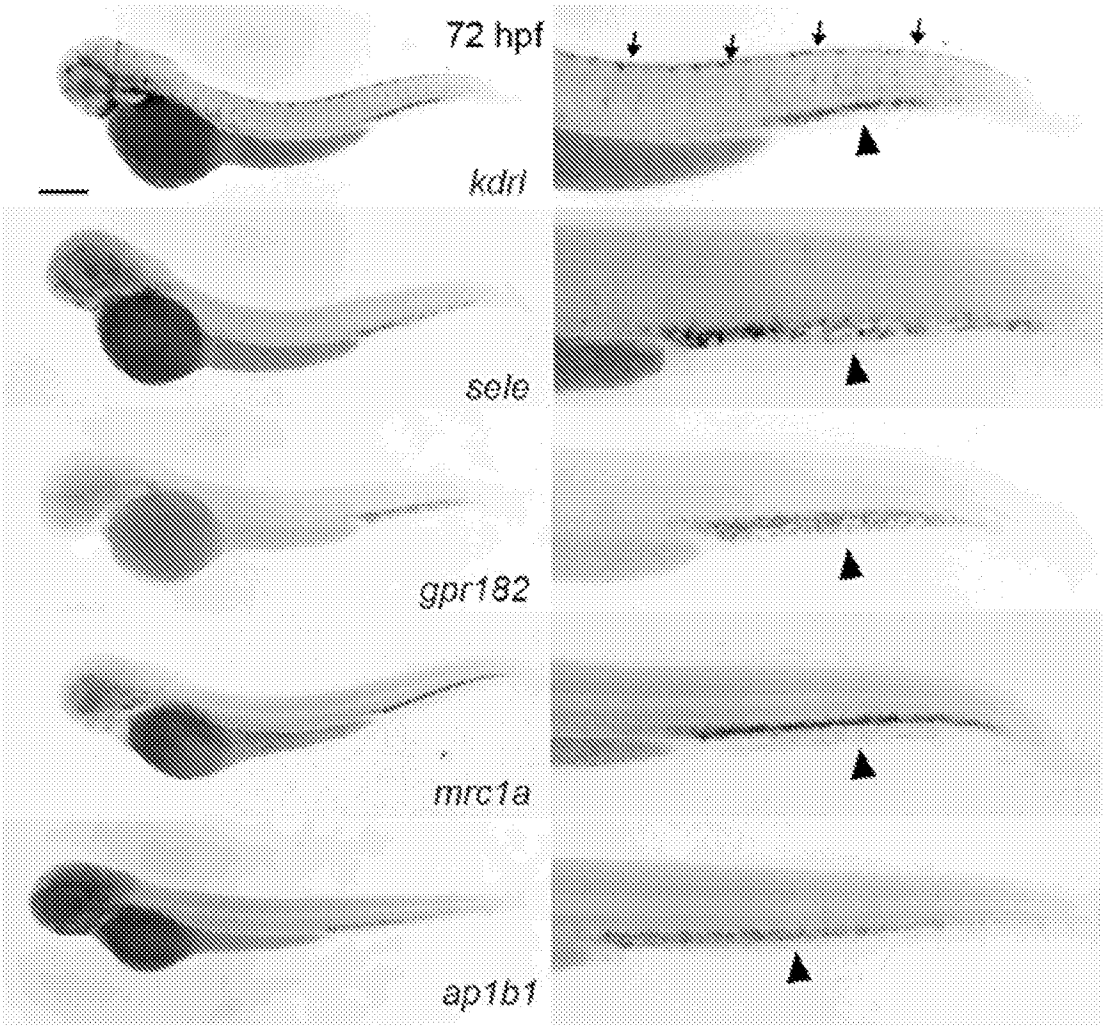


Fig. 1E

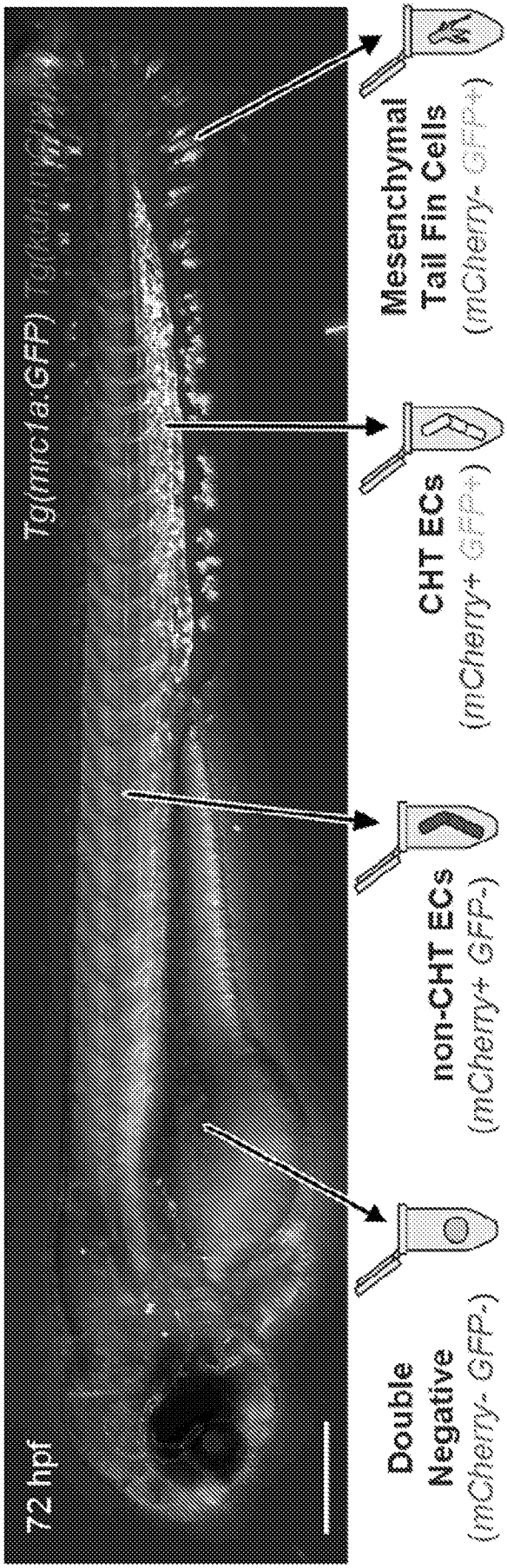

Tg(mic1a:GFP)
72 hpf
Double Negative (mCherry- GFP-)
non-CHT ECs (mCherry+ GFP-)
CHT ECs (mCherry+ GFP+)
Mesenchymal Tail Fin Cells (mCherry- GFP+)

Fig. 2A

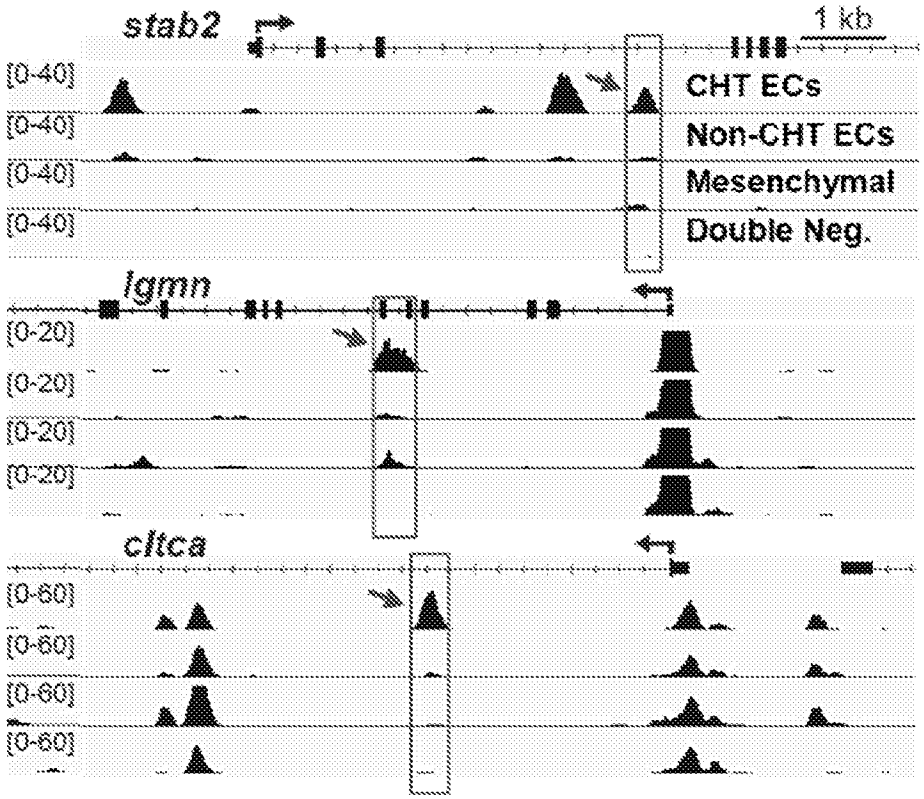

stab2
1 kb
[0-40]
[0-40]
[0-40]
[0-40]
CHT ECs
Non-CHT ECs
Mesenchymal
Double Neg.
lgmn
[0-20]
[0-20]
[0-20]
[0-20]
cltca
[0-60]
[0-60]
[0-60]
[0-60]

Fig. 2B

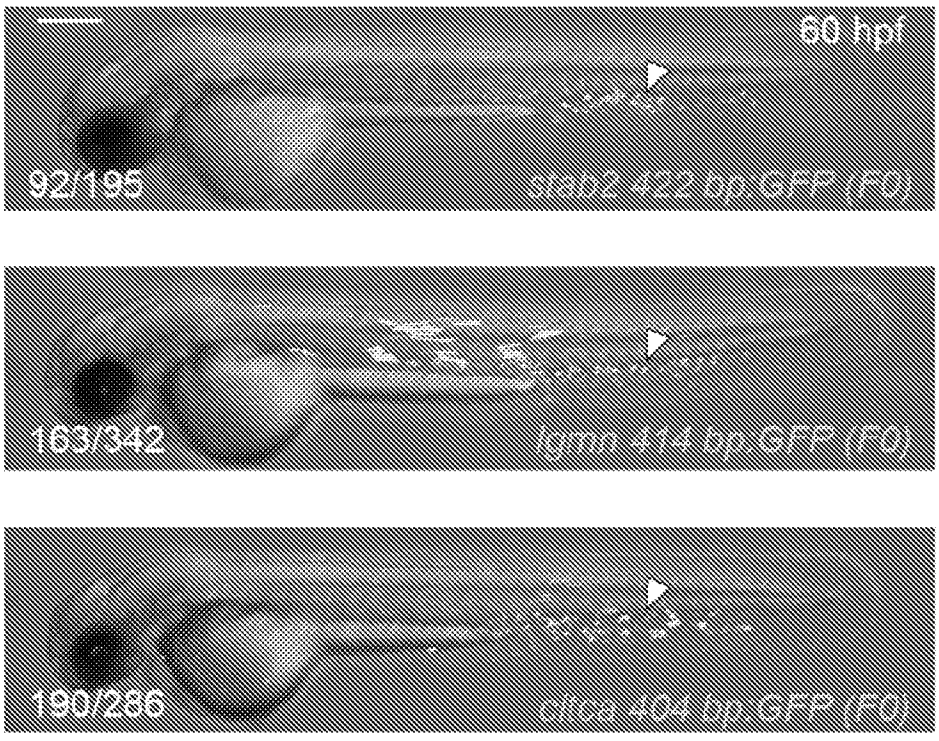

60 hpf
92/195
stab2 -22.2 kb:GFP (F0)
163/342
lgmn 3'4 kb:GFP (F0)
190/286
cltca -40.4 kb:GFP (F0)

Fig. 2C

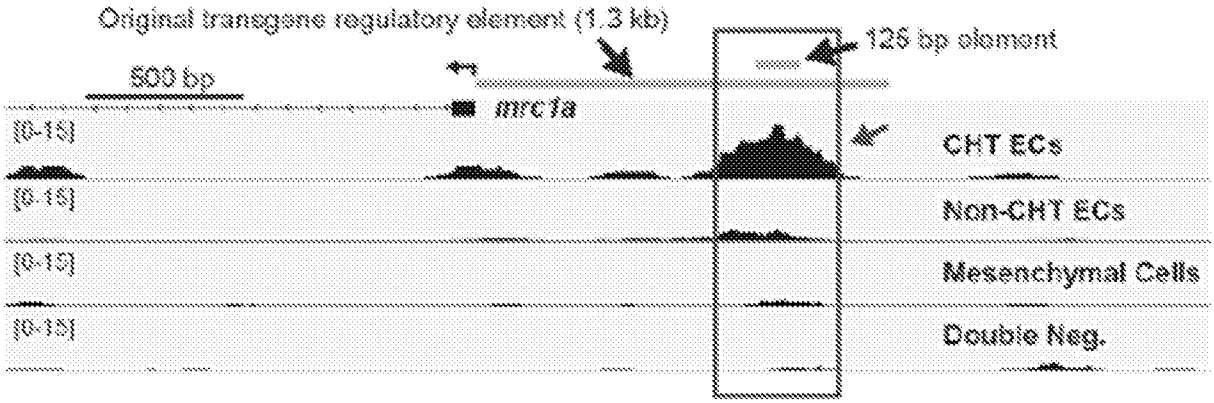

Original transgene regulatory element (1.3 kb)
125 bp element
500 bp
mrc1a
[0-15]
CHT ECs
[0-15]
Non-CHT ECs
[0-15]
Mesenchymal Cells
[0-15]
Double Neg.

Fig. 3A

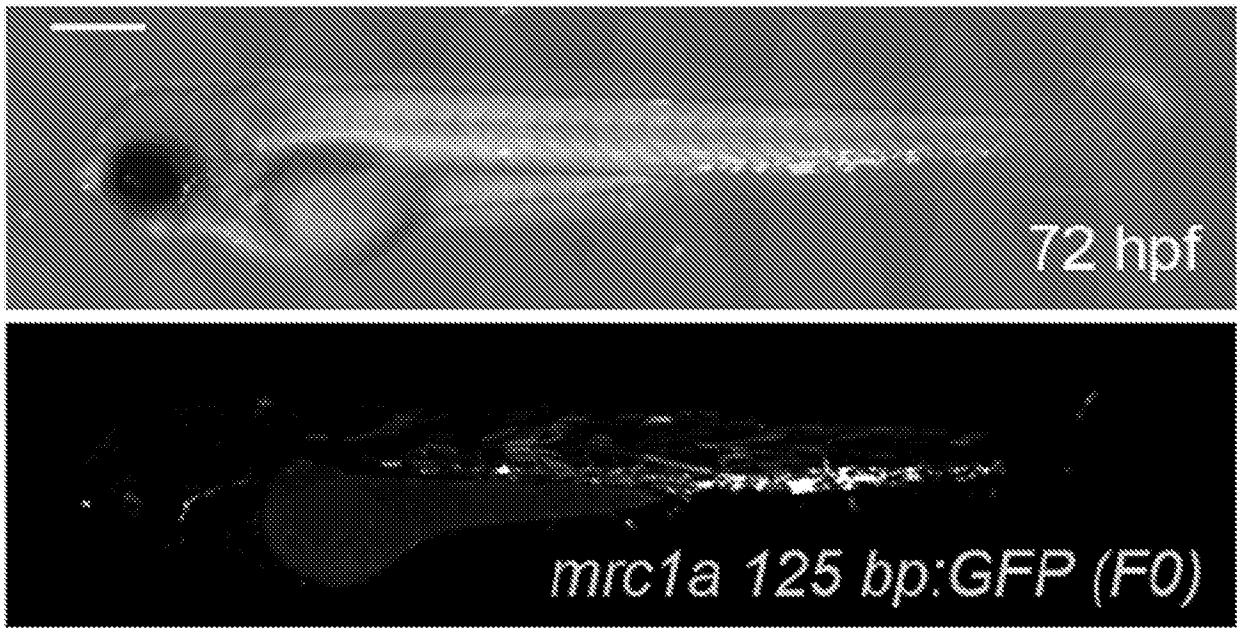

72 hpf
mrc1a 125 bp:GFP (F0)

Fig. 3B

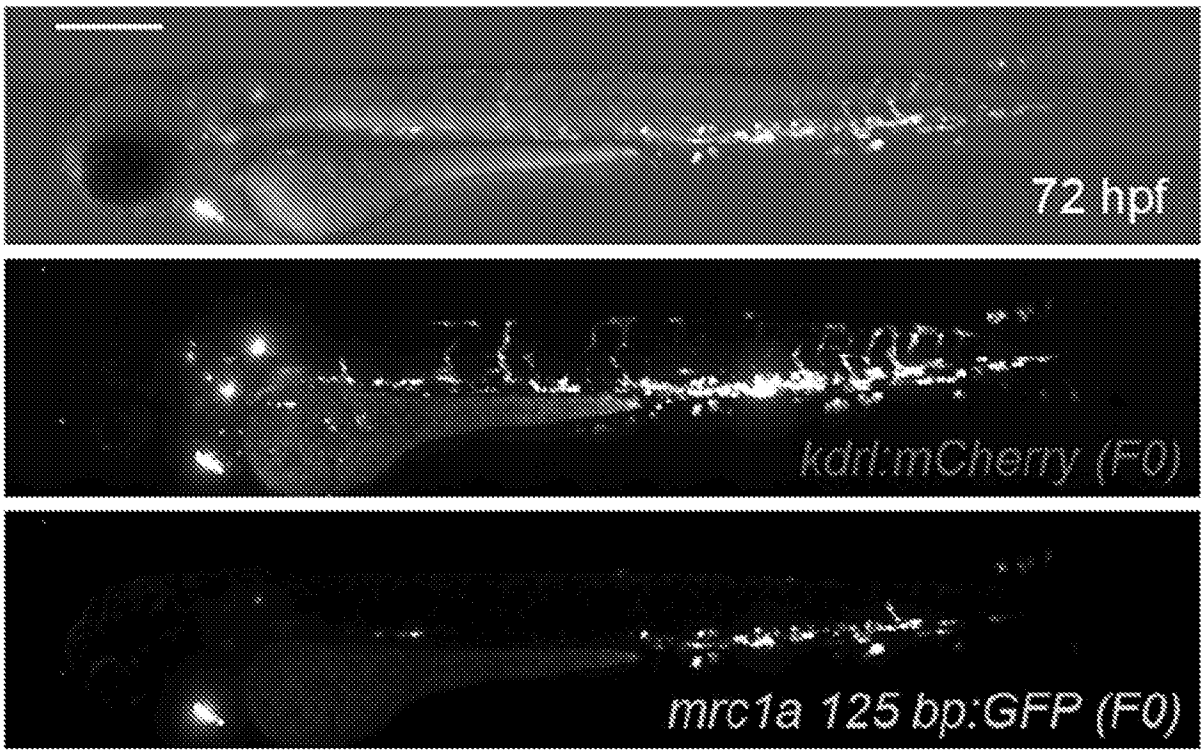

72 hpf
kdrl:mCherry (F0)
mrc1a 125 bp:GFP (F0)

Fig. 3C

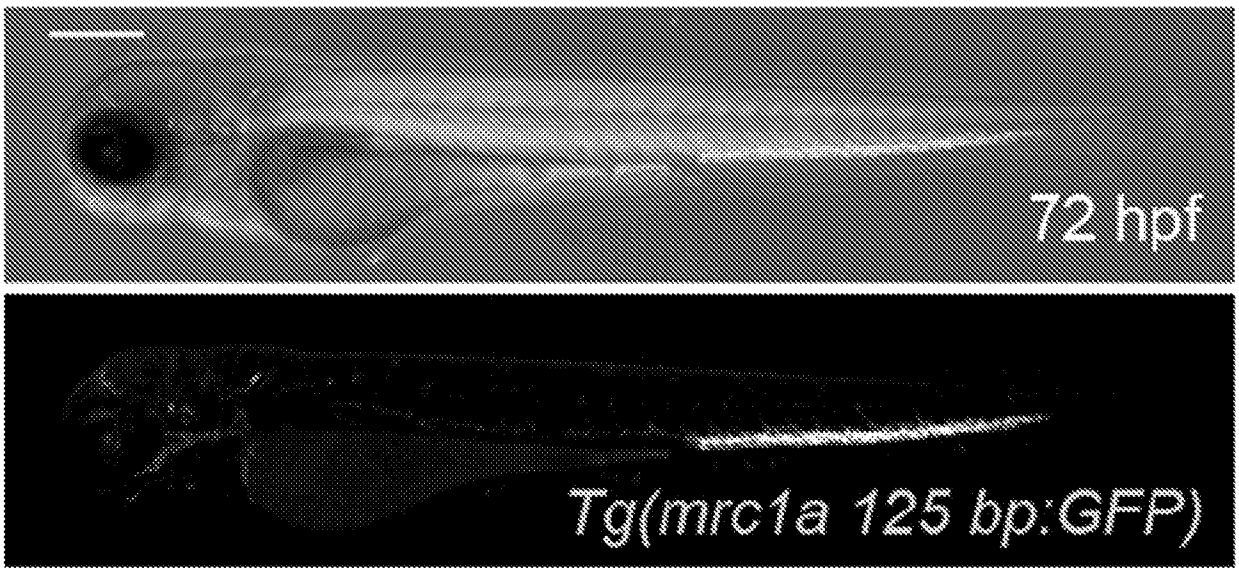

72 hpf
Tg(mrc1a 125 bp:GFP)

Fig. 3D

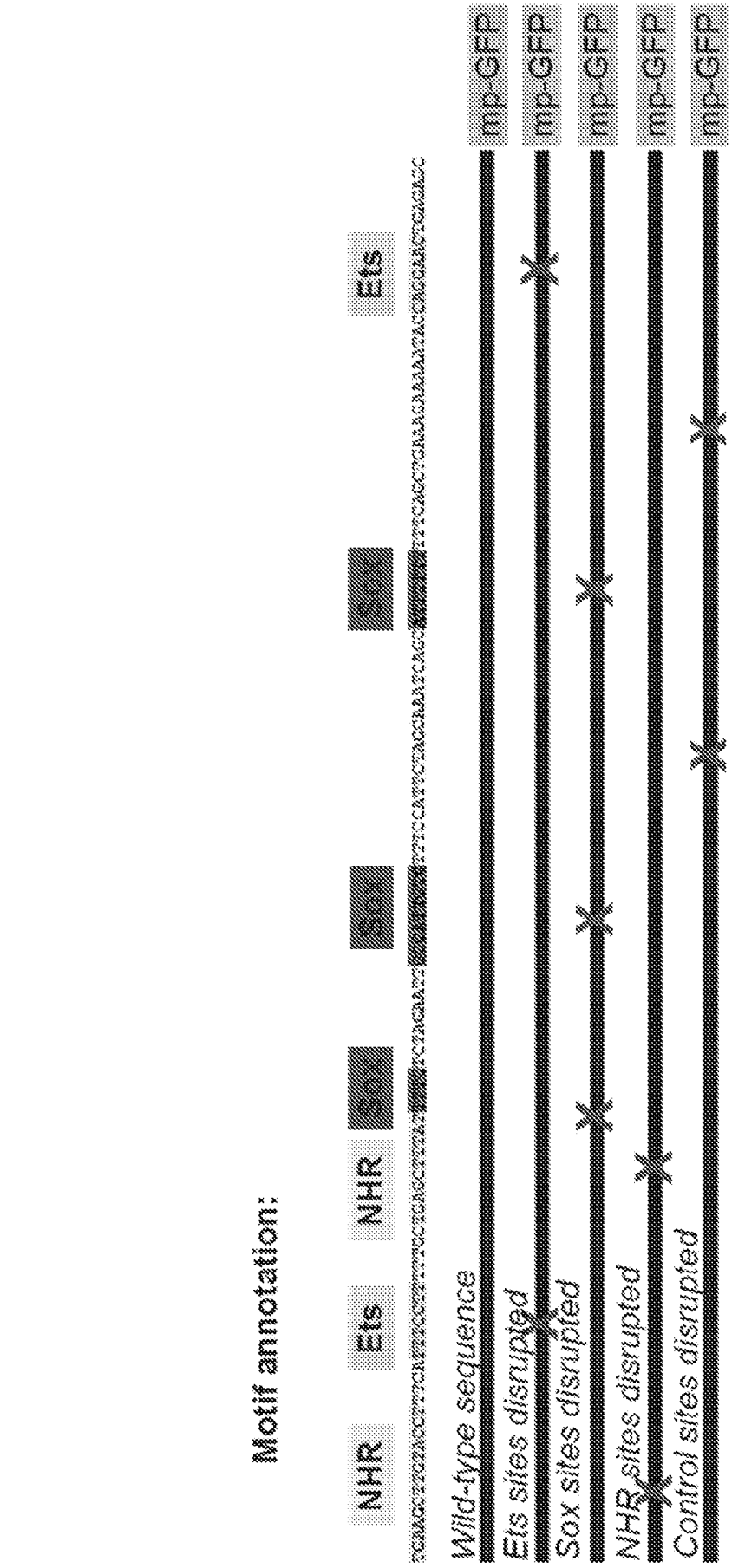

Motif annotation:
NHR
Ets
NHR
Sox
Sox
Ets
Wild-type sequence
Ets sites disrupted
Sox sites disrupted
NHR sites disrupted
Control sites disrupted
mp-GFP
mp-GFP
mp-GFP
mp-GFP
mp-GFP

Fig. 3E

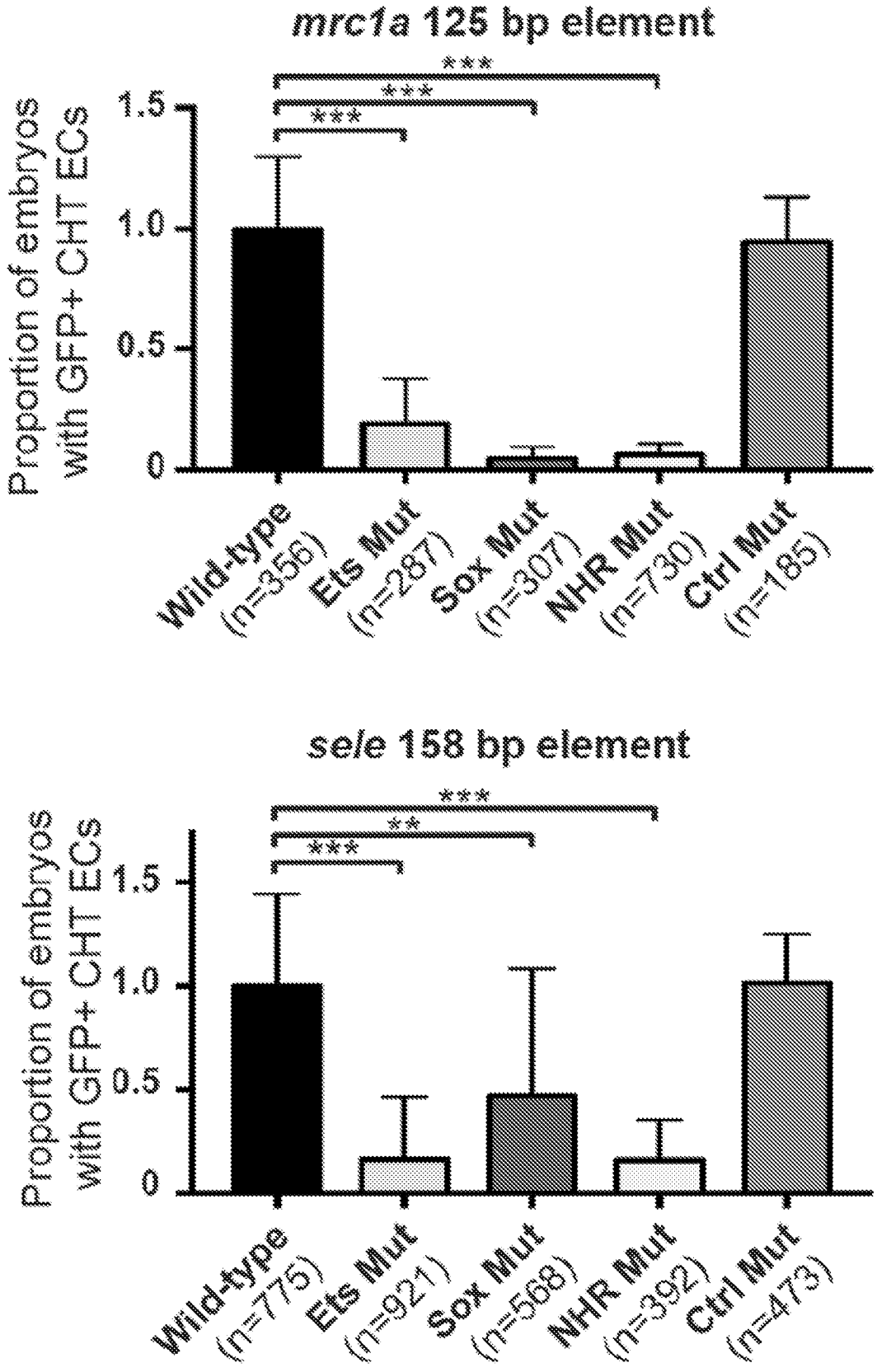
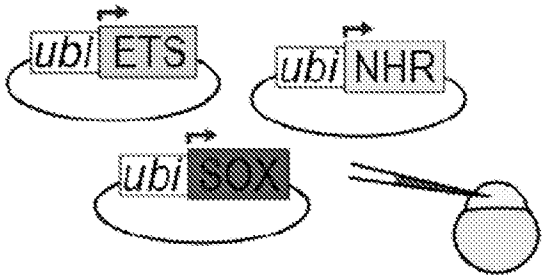
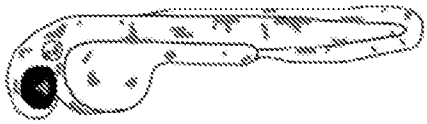

mrc1a 125 bp element
Proportion of embryos with GFP+ CHT ECs
1.5
1.0
0.5
0
***
***
***
Wild-type (n=356)
Ets Mut (n=287)
Sox Mut (n=307)
NHR Mut (n=730)
Ctrl Mut (n=185)
sele 158 bp element
Proportion of embryos with GFP+ CHT ECs
1.5
1.0
0.5
0
***
**
***
Wild-type (n=775)
Ets Mut (n=921)
Sox Mut (n=568)
NHR Mut (n=392)
Ctrl Mut (n=473)

Fig. 3F



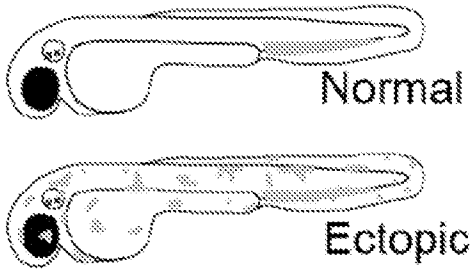
Inject human factors driven by *ubi* promoter

(pools of Ets, Sox and NHR TFs)



Mosaic expression of human factors

(confirmed by WISH for human genes)



Screen embryos for niche endothelial gene expression

(by *mrc1a*:GFP or WISH for *mrc1a*)

Fig. 4A

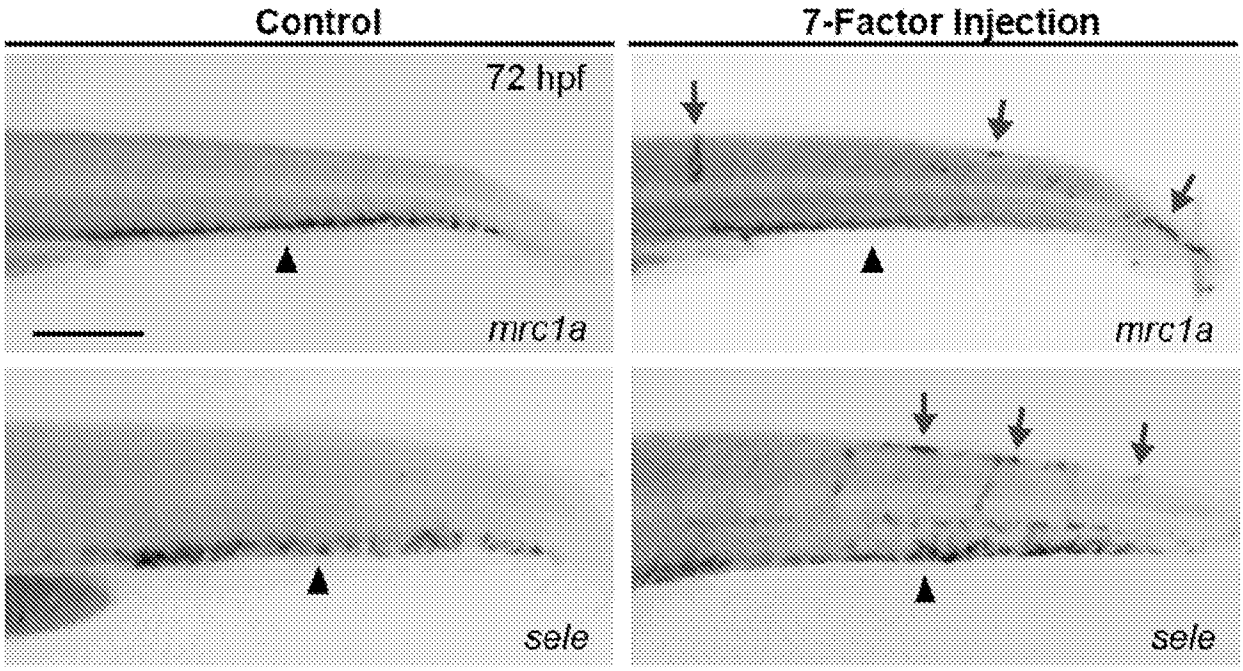

Control
7-Factor Injection
72 hpf
mrc1a
mrc1a
sele
sele

Fig. 4B

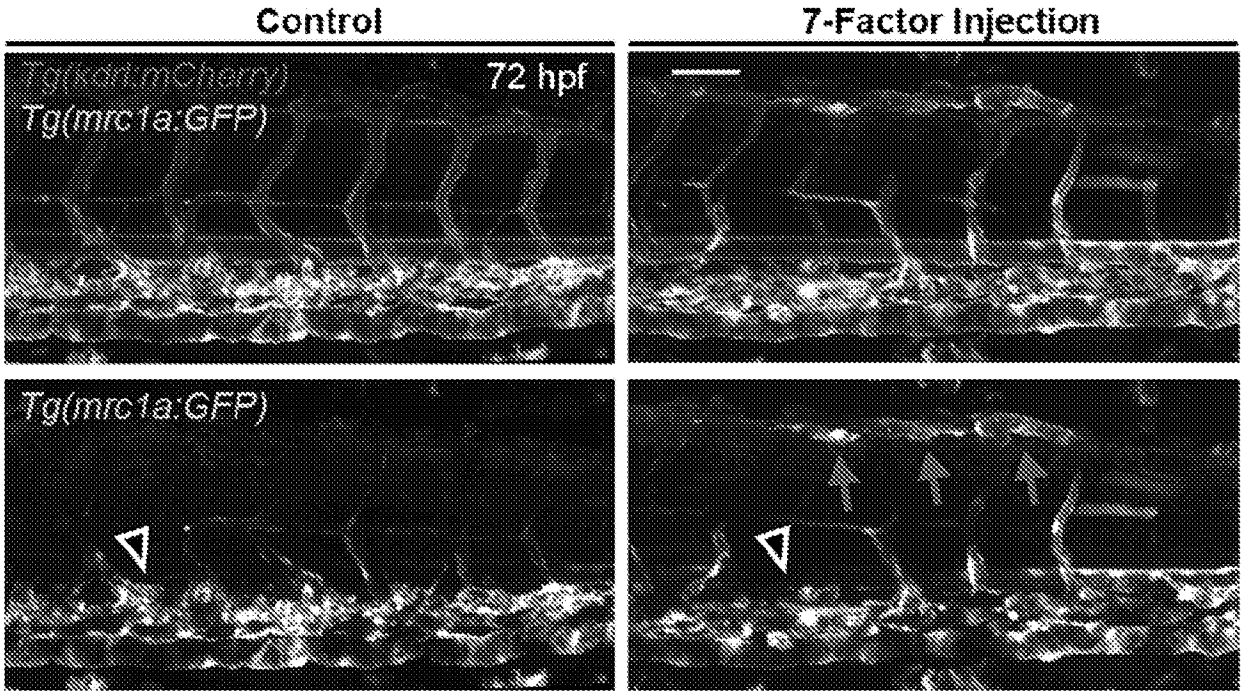

Control
7-Factor Injection
Tg(kdrl:mCherry)
Tg(mrc1a:GFP)
72 hpf
Tg(mrc1a:GFP)

Fig. 4C

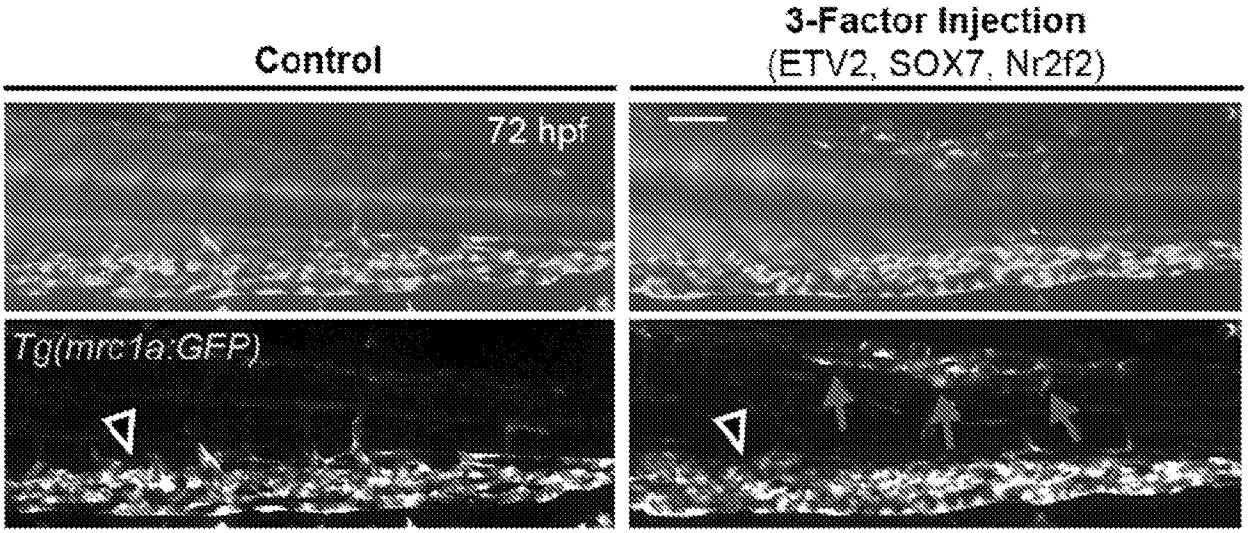

Control
3-Factor Injection
(ETV2, SOX7, Nr2f2)
72 hpf
Tg(mrc1a:GFP)

Fig. 4D

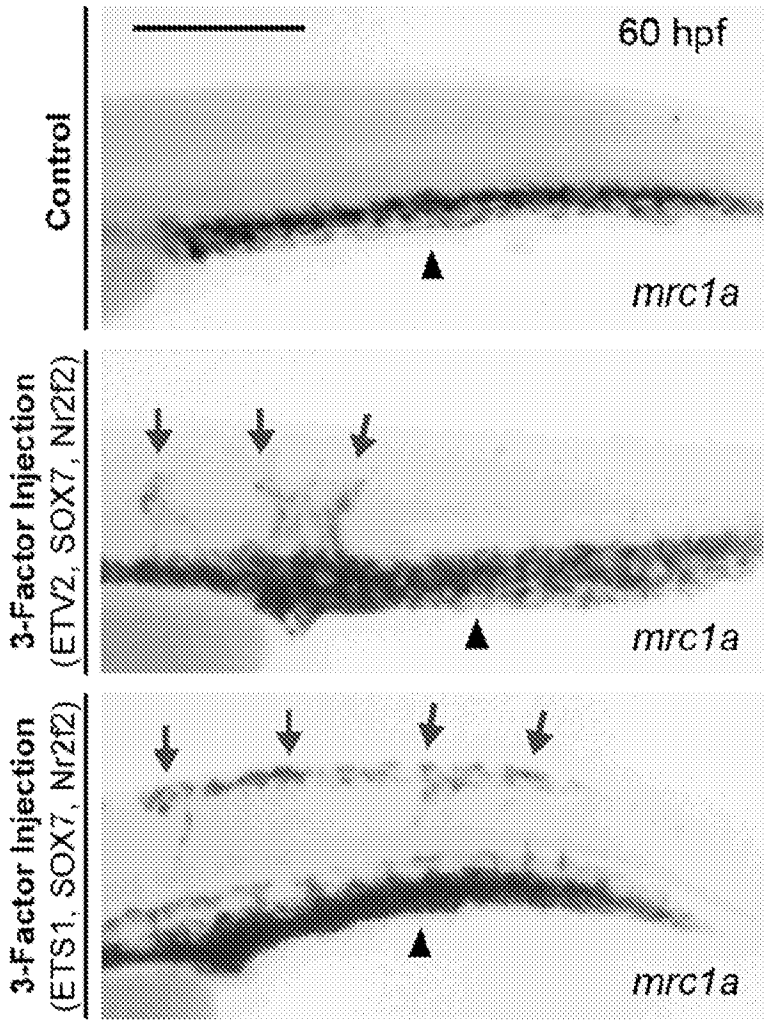

Control
60 hpf
mrc1a
3-Factor Injection
(ETV2, SOX7, Nr2f2)
mrc1a
3-Factor Injection
(ETS1, SOX7, Nr2f2)
mrc1a

Fig. 4E

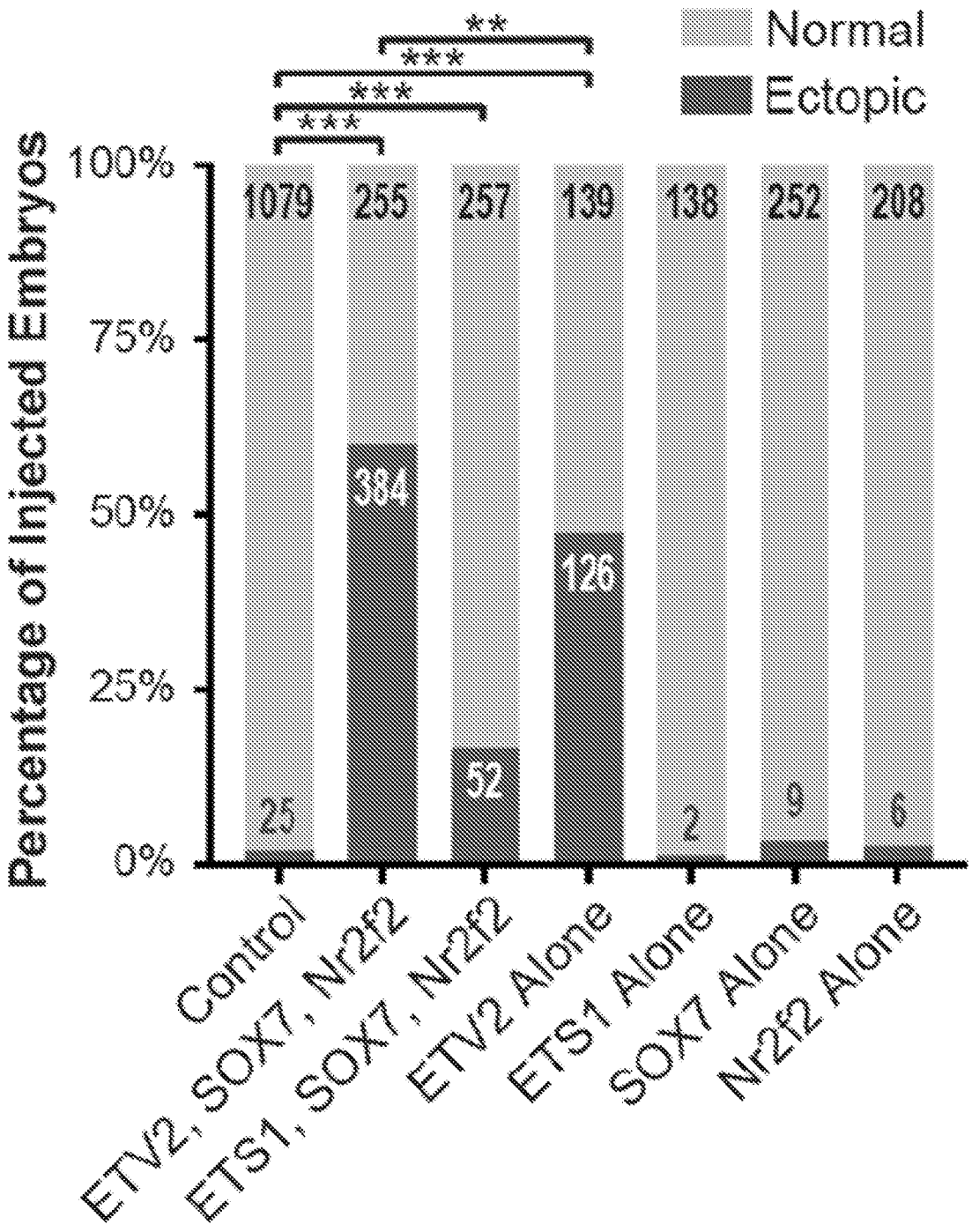

Normal
Ectopic
**
***
***
***
Percentage of Injected Embryos
100%
75%
50%
25%
0%
1079
255
257
139
138
252
208
25
384
52
126
2
9
6
Control
ETV2, SOX7, Nr2f2
ETS1, SOX7, Nr2f2
ETV2 Alone
ETS1 Alone
SOX7 Alone
Nr2f2 Alone

Fig. 4F

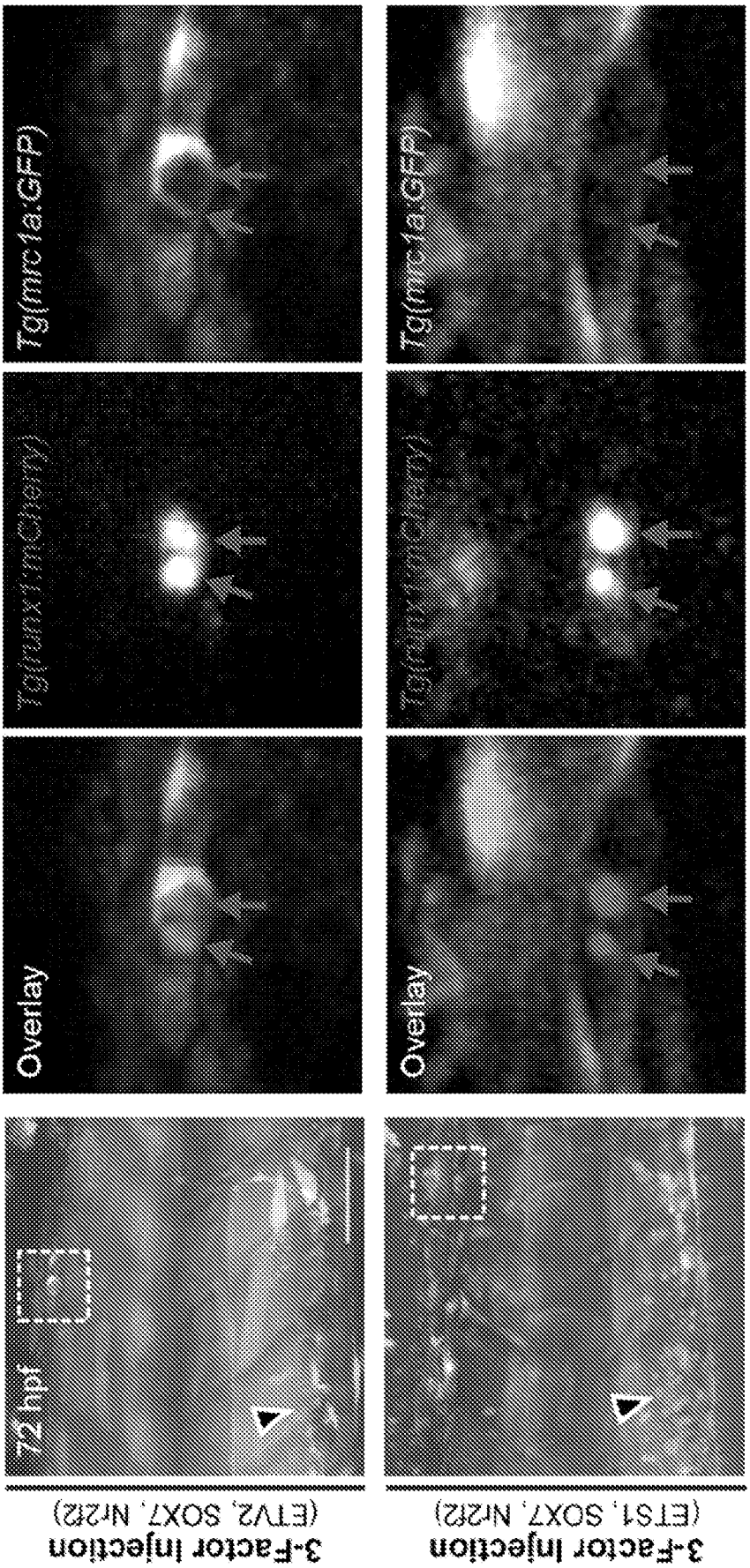

72 hpf
Overlay
Tg(runx1:mCherry)
Tg(mrc1a:GFP)
3-Factor injection
(ETV2, SOX7, Nr2f2)
3-Factor injection
(ETS1, SOX7, Nr2f2)

Fig. 5A

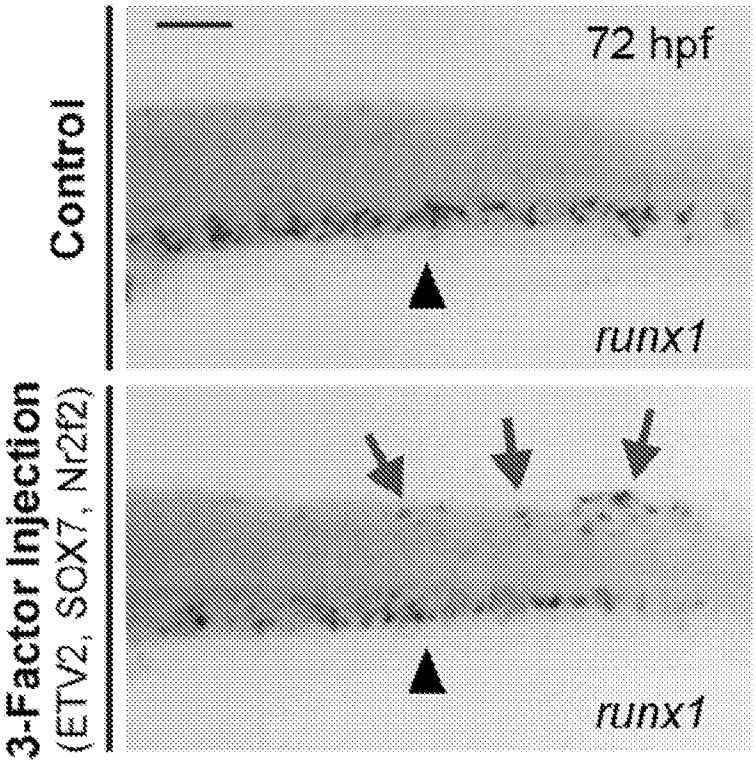

Control
72 hpf
runx1
3-Factor Injection
(ETV2, SOX7, Nr2f2)
runx1

Fig. 5B

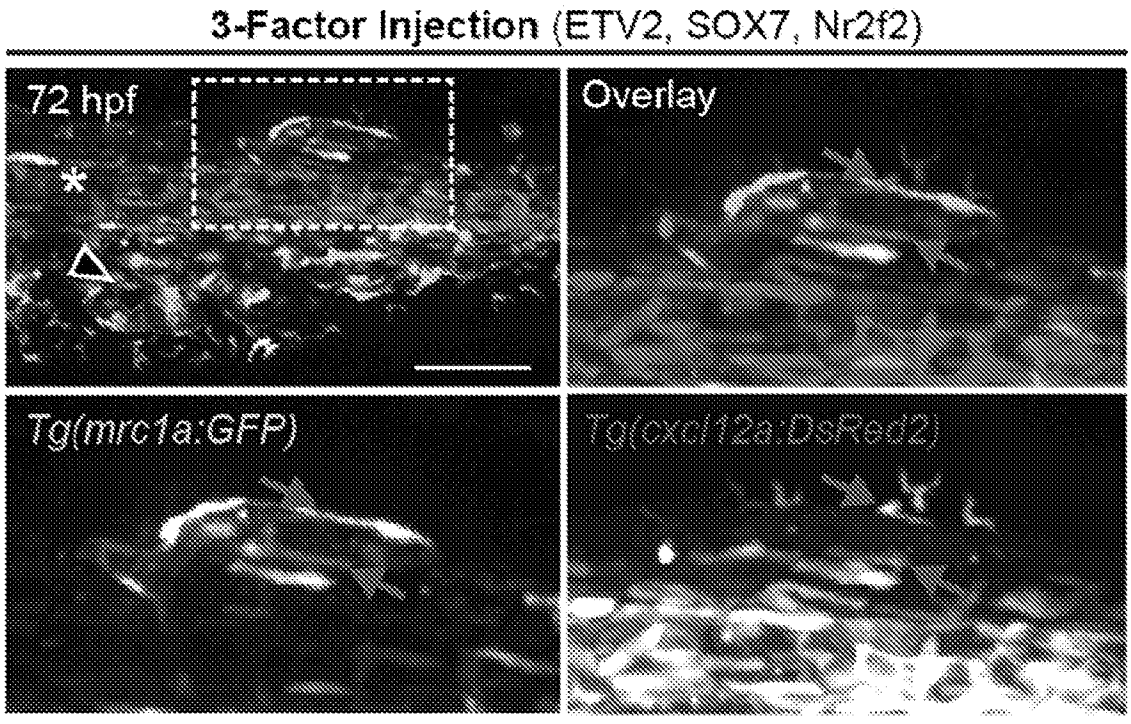

3-Factor Injection (ETV2, SOX7, Nr2f2)
72 hpf
Overlay
Tg(mrc1a:GFP)
Tg(cxcl12a:DsRed2)

Fig. 5C

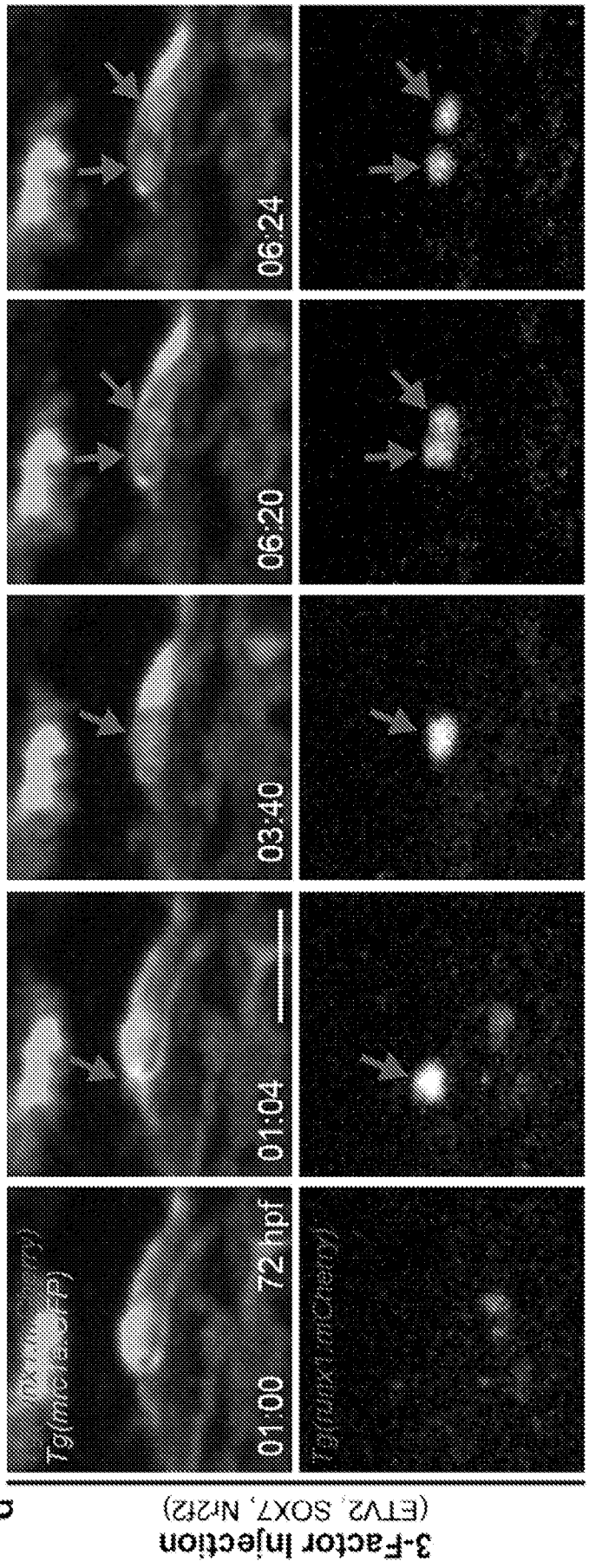

d
3-Factor injection
(ETV2, SOX7, Nr2f2)
Tg(myl7:nucGFP)
Tg(fli1a:nEGFP)
Tg(flk1:mCherry)
72 hpf
01:00
01:04
03:40
06:20
06:24

Fig. 5D

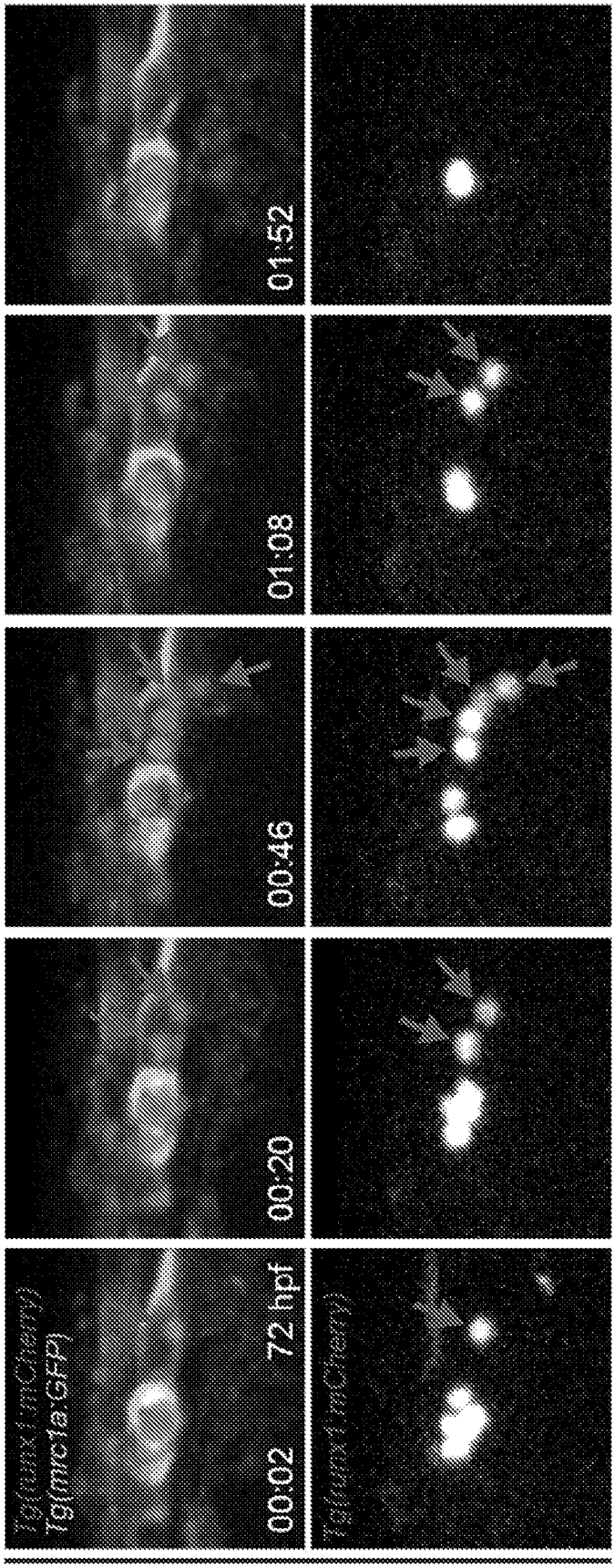

Tg(runx1:mCherry)
Tg(mrc1a:GFP)
00:02
72 hpf
00:20
00:46
01:08
01:52
Tg(runx1:mCherry)
3-Factor injection
(ETV2, SOX7, Notch2)

Fig. 5E

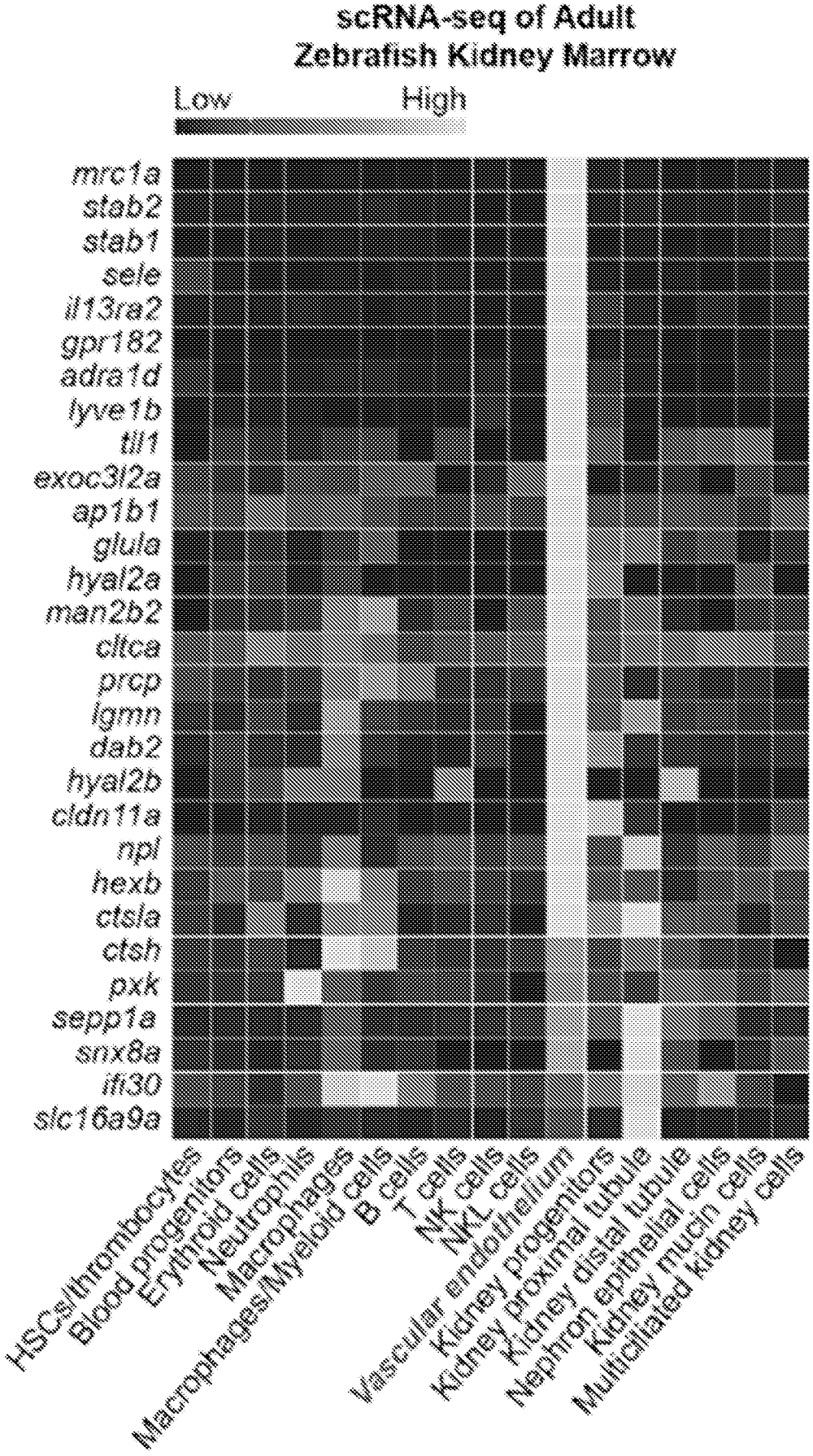

scRNA-seq of Adult
Zebrafish Kidney Marrow
Low
High
mrc1a
stab2
stab1
sele
il13ra2
gpr182
adra1d
lyve1b
tll1
exoc3l2a
ap1b1
glula
hyal2a
man2b2
cltca
prcp
lgmn
dab2
hyal2b
cldn11a
npl
hexb
ctsla
ctsh
pxk
sepp1a
snx8a
ifi30
slc16a9a
HSCs/thrombocytes
Blood progenitors
Erythroid cells
Neutrophils
Macrophages
Macrophages/Myeloid cells
B cells
T cells
NK cells
NKL cells
Vascular endothelium
Kidney progenitors
Kidney proximal tubule
Kidney distal tubule
Nephron epithelial cells
Kidney mucin cells
Multiciliated kidney cells

Fig. 6A

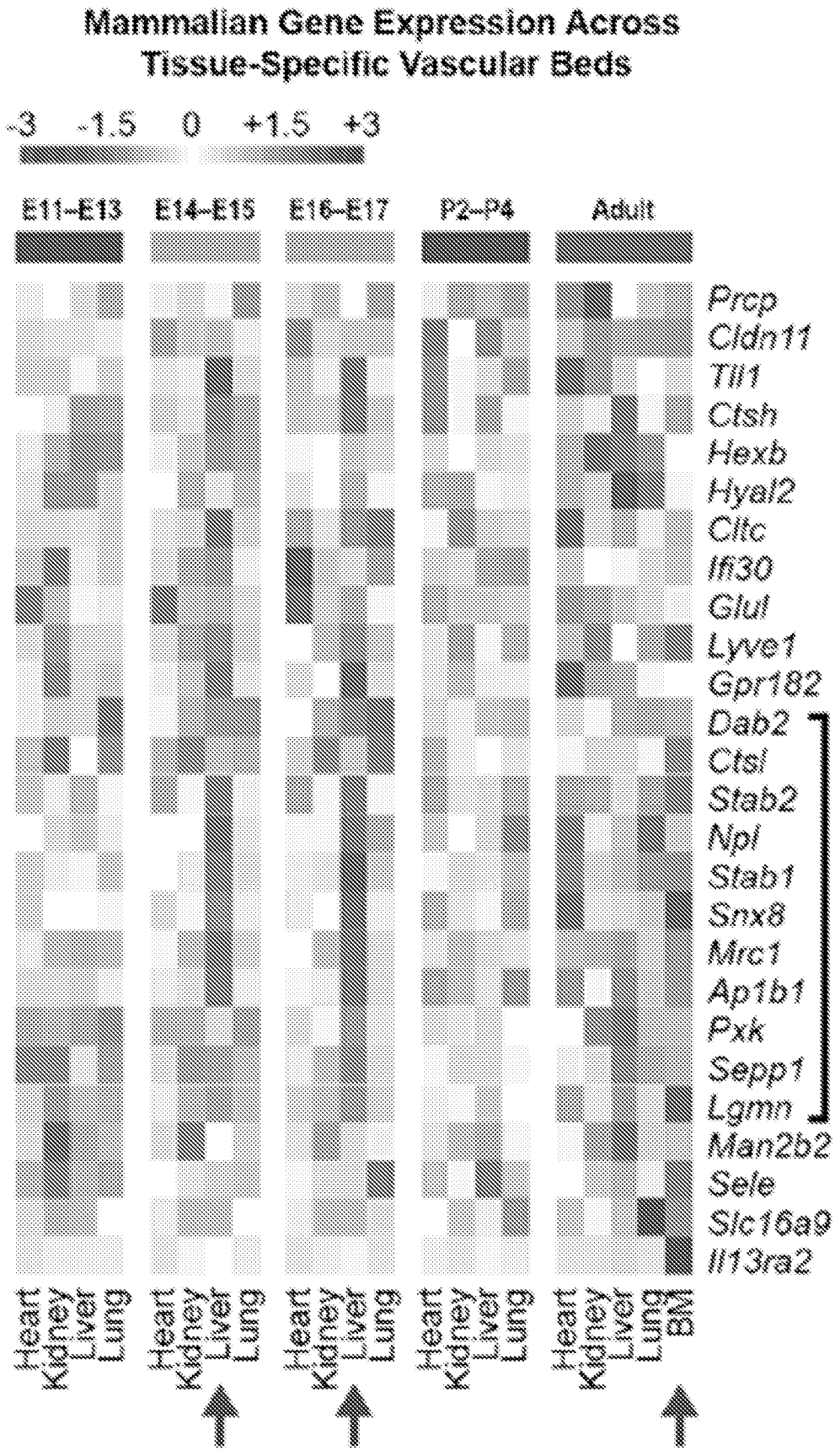

Mammalian Gene Expression Across
Tissue-Specific Vascular Beds
-3 -1.5 0 +1.5 +3
E11–E13
E14–E15
E16–E17
P2–P4
Adult
Prcp
Cldn11
Tll1
Ctsh
Hexb
Hyal2
Cltc
Ifi30
Glul
Lyve1
Gpr182
Dab2
Ctsl
Stab2
Npl
Stab1
Snx8
Mrc1
Ap1b1
Pxk
Sepp1
Lgmn
Man2b2
Sele
Slc16a9
Il13ra2
Heart
Kidney
Liver
Lung
BM

Fig. 6B

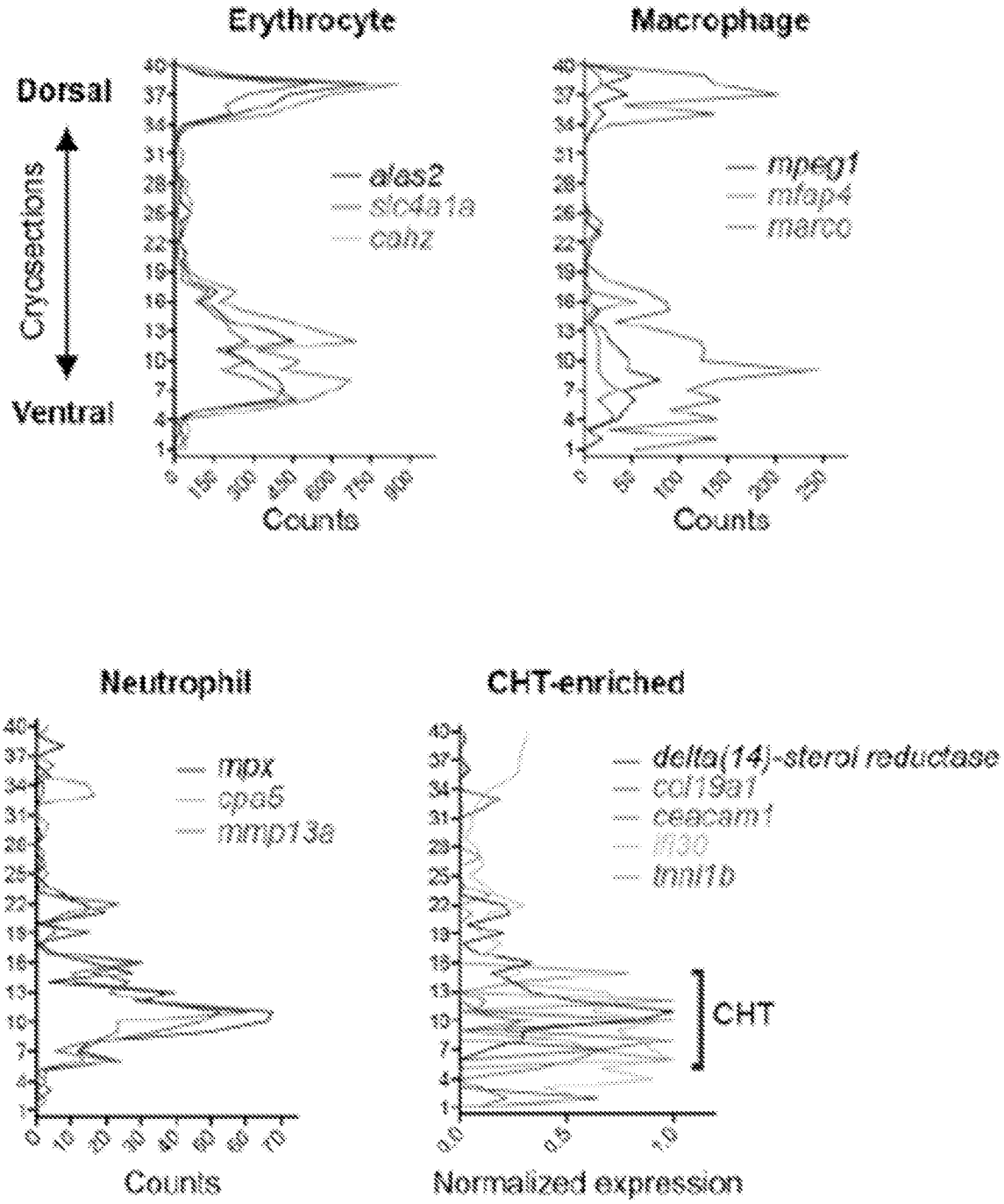

Erythrocyte
Macrophage
Dorsal
Ventral
Cryosections
alas2
slc4a1a
cahz
mpeg1
mfap4
marco
Counts
Counts
Neutrophil
CHT-enriched
mpx
cpa5
mmp13a
delta(14)-sterol reductase
col19a1
ceacam1
tnfrsf1b
CHT
Counts
Normalized expression

Fig. 7A

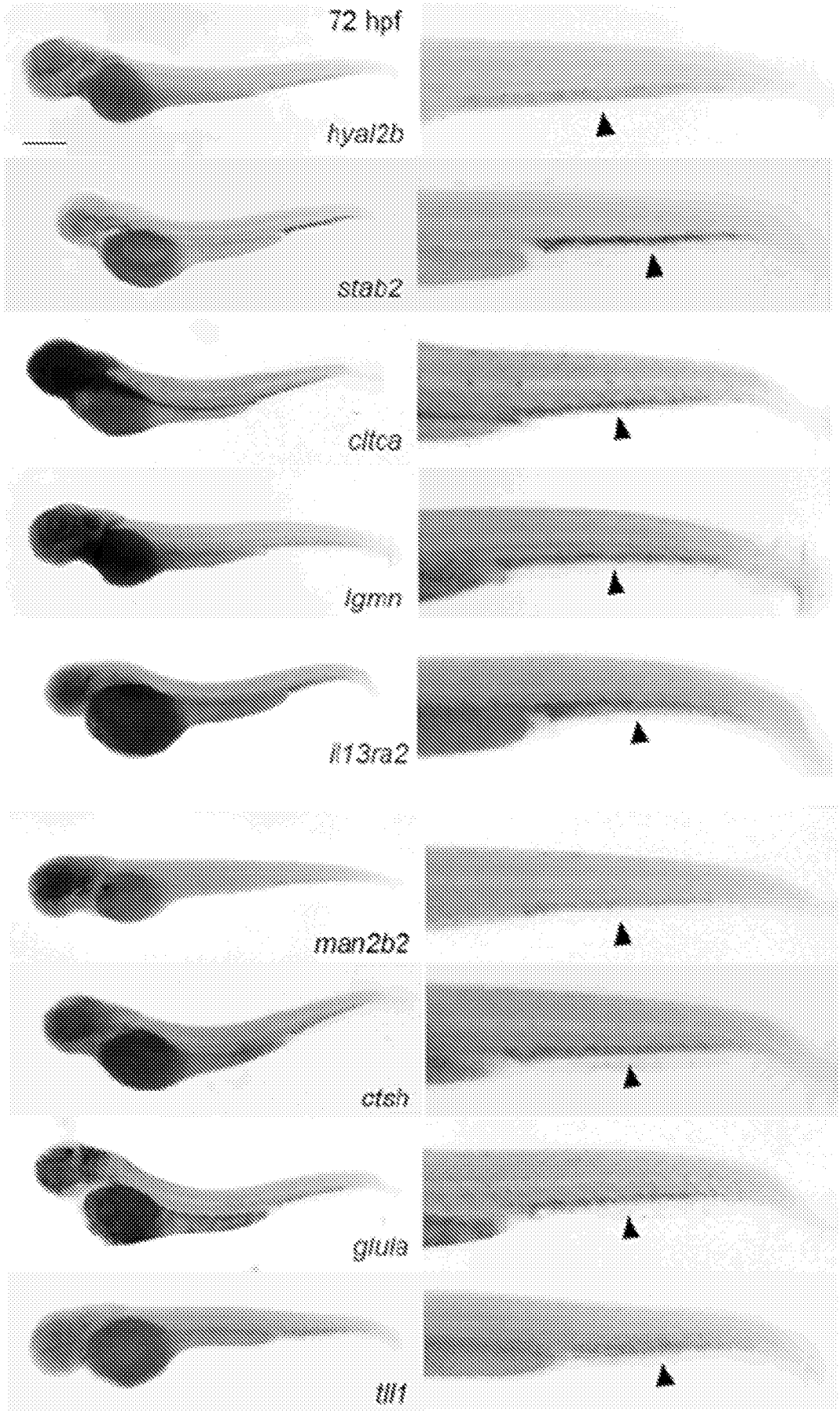

72 hpf
hyal2b
stab2
cltca
lgmn
il13ra2
man2b2
ctsh
glula
tll1

Fig. 7B

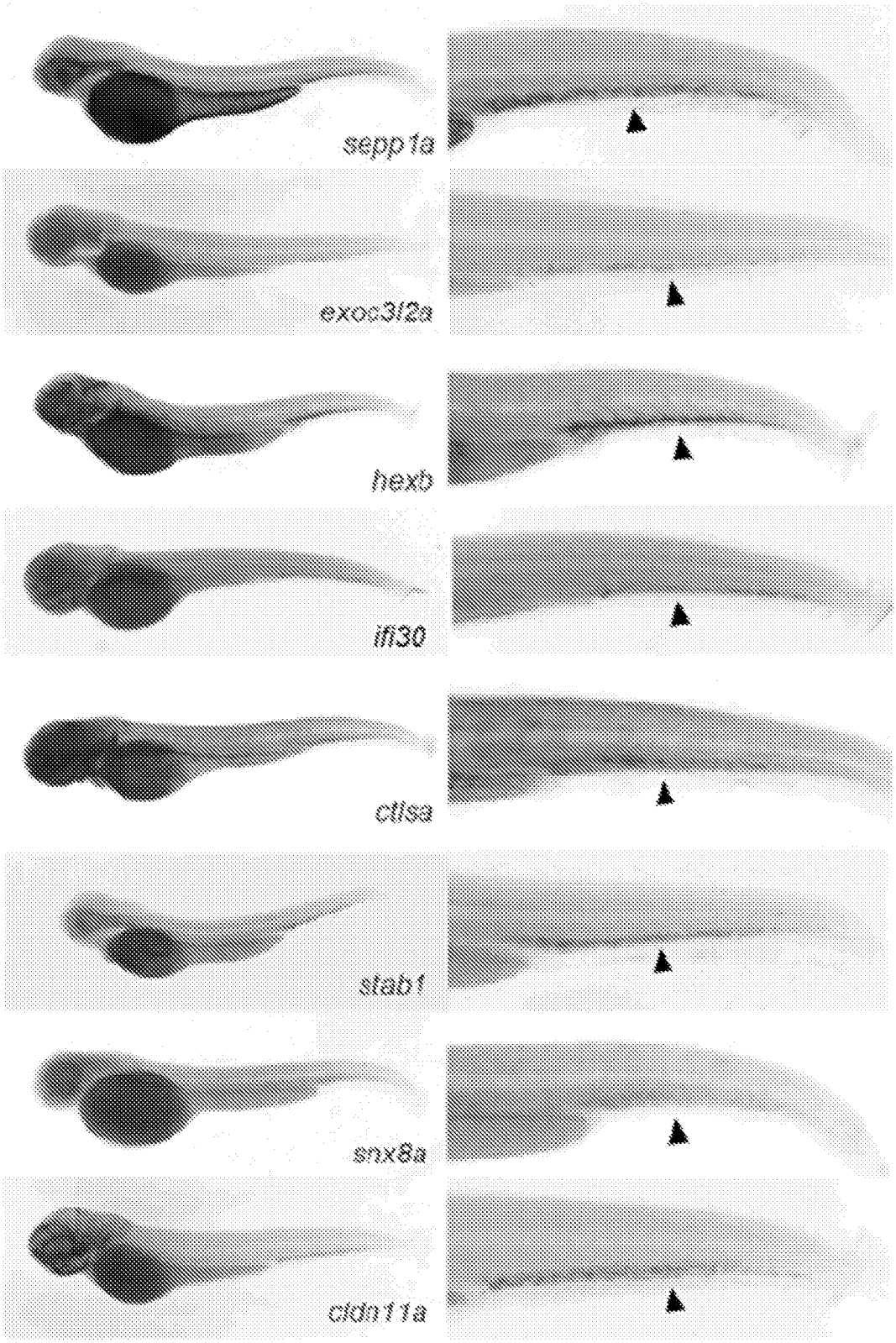


Fig. 7B, cont.

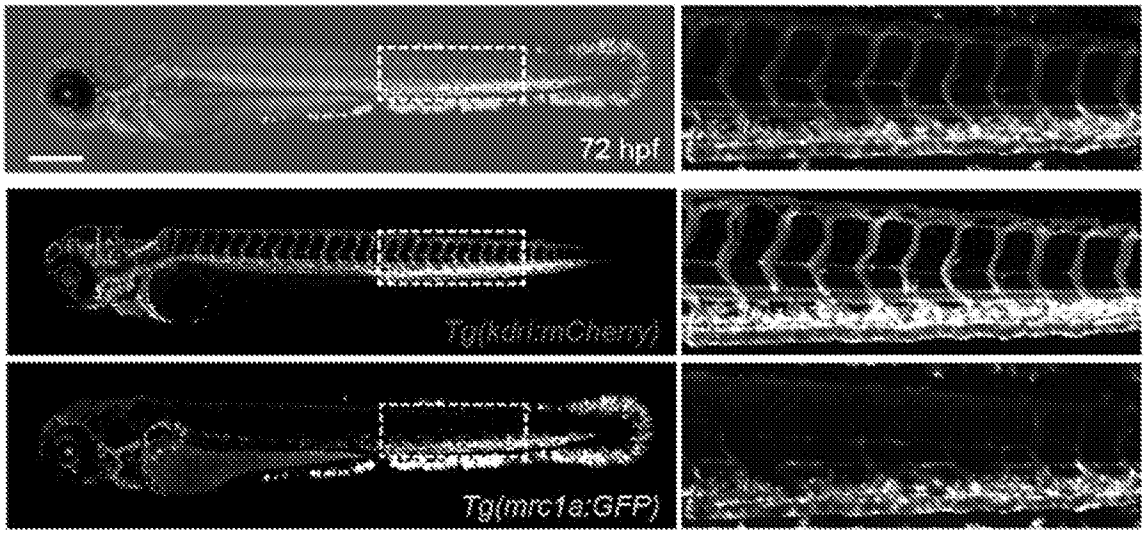

72 hpf
Tg(kdrl:mCherry)
Tg(mrc1a:GFP)

Fig. 8A

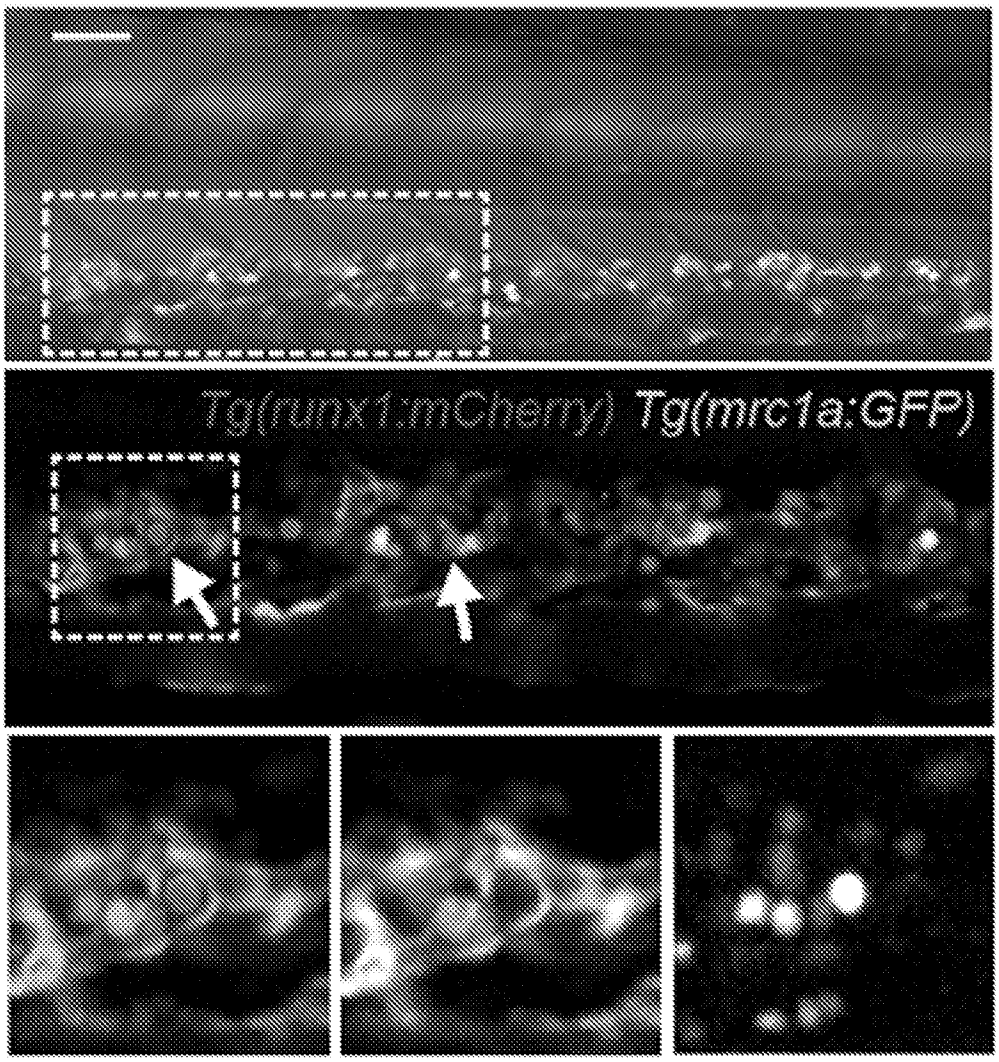

Tg(runx1:mCherry) Tg(mrc1a:GFP)

Fig. 8B

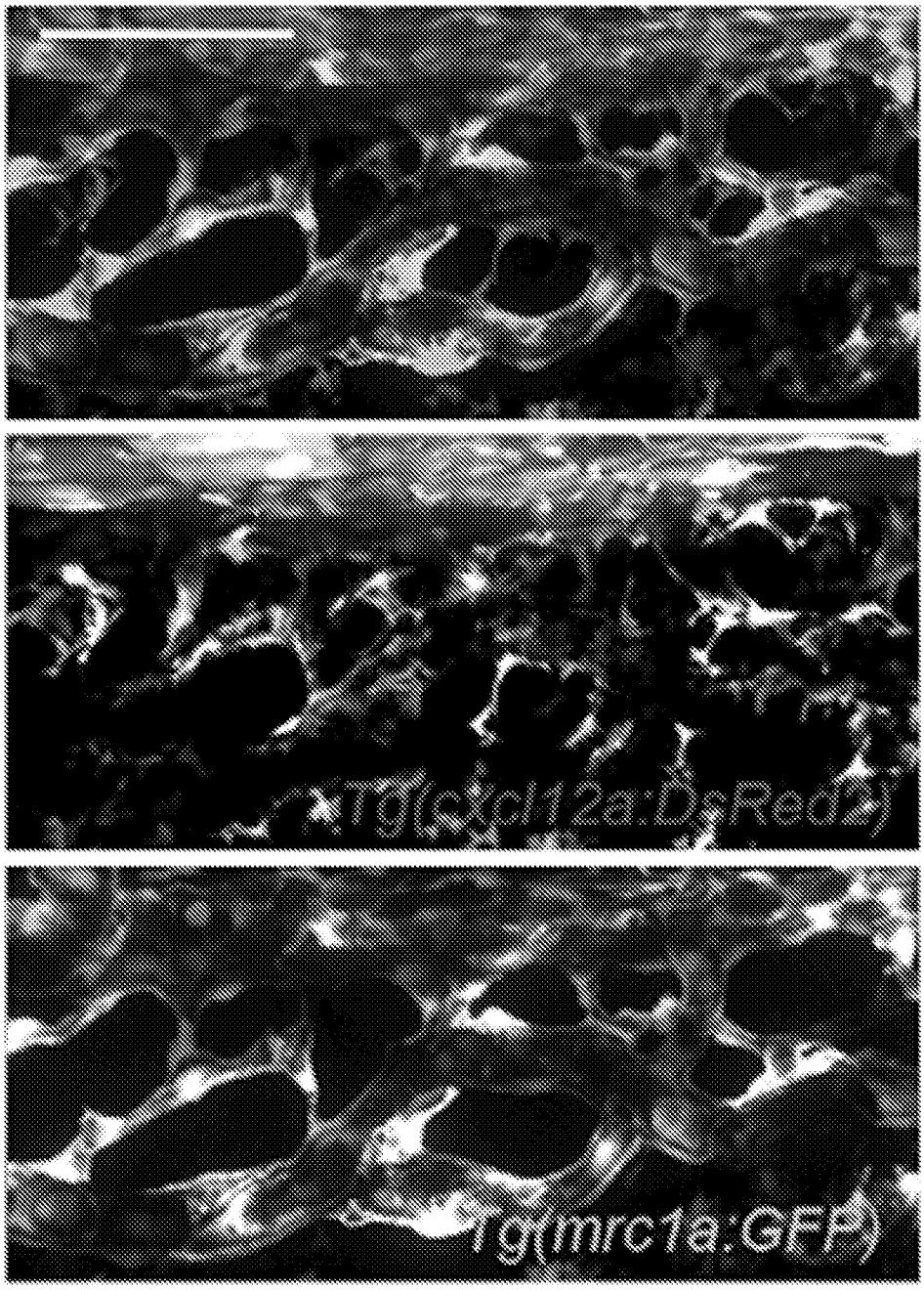


Fig. 8C

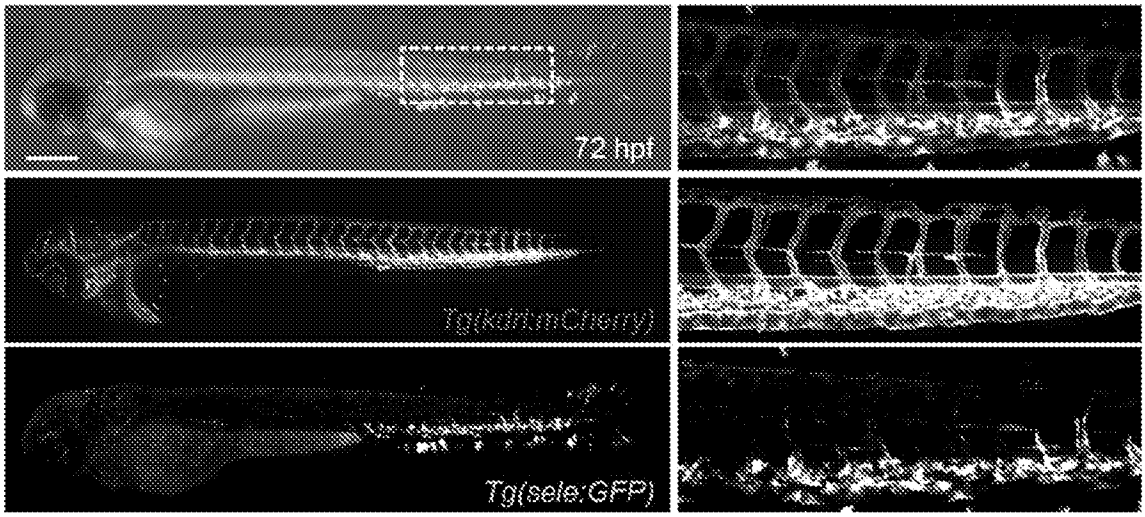

72 hpf
Tg(kdrl:mCherry)
Tg(sele:GFP)

Fig. 8D

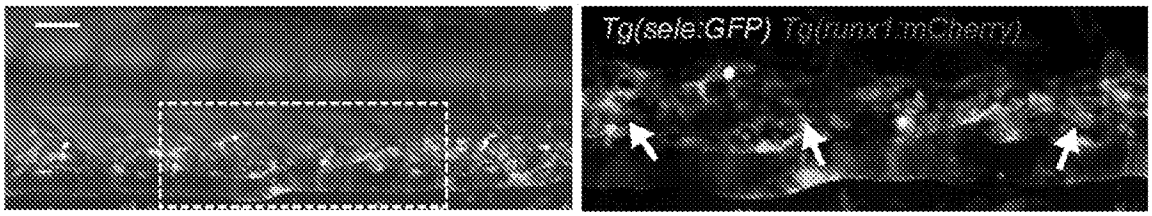

Tg(sele:GFP) Tg(runx1:mCherry)

Fig. 8E

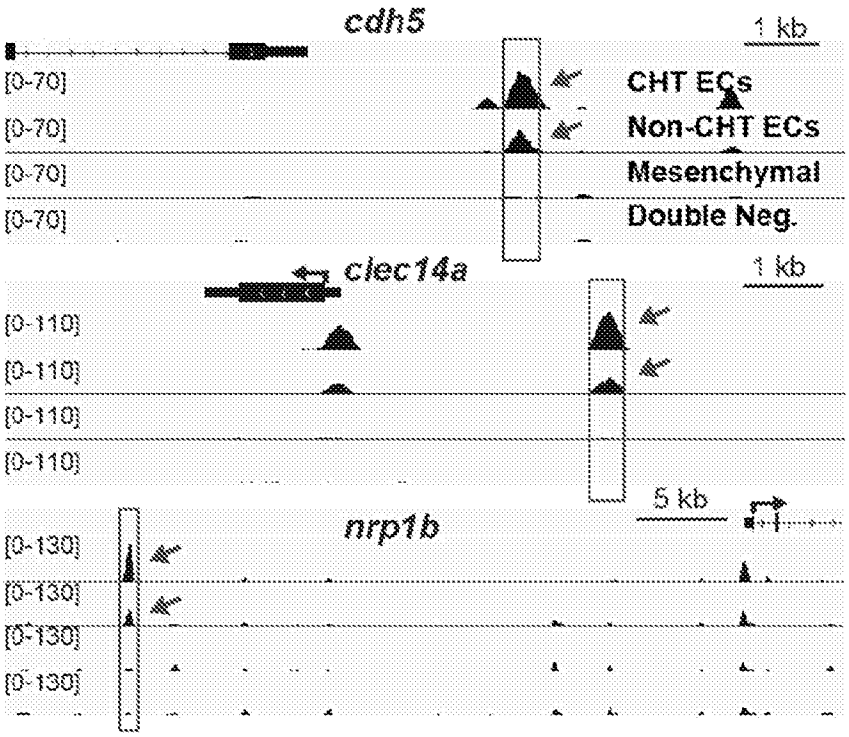

cdh5
1 kb
[0-70]
[0-70]
[0-70]
[0-70]
CHT ECs
Non-CHT ECs
Mesenchymal
Double Neg.
clec14a
1 kb
[0-110]
[0-110]
[0-110]
[0-110]
5 kb
nrp1b
[0-130]
[0-130]
[0-130]
[0-130]

Fig. 9A

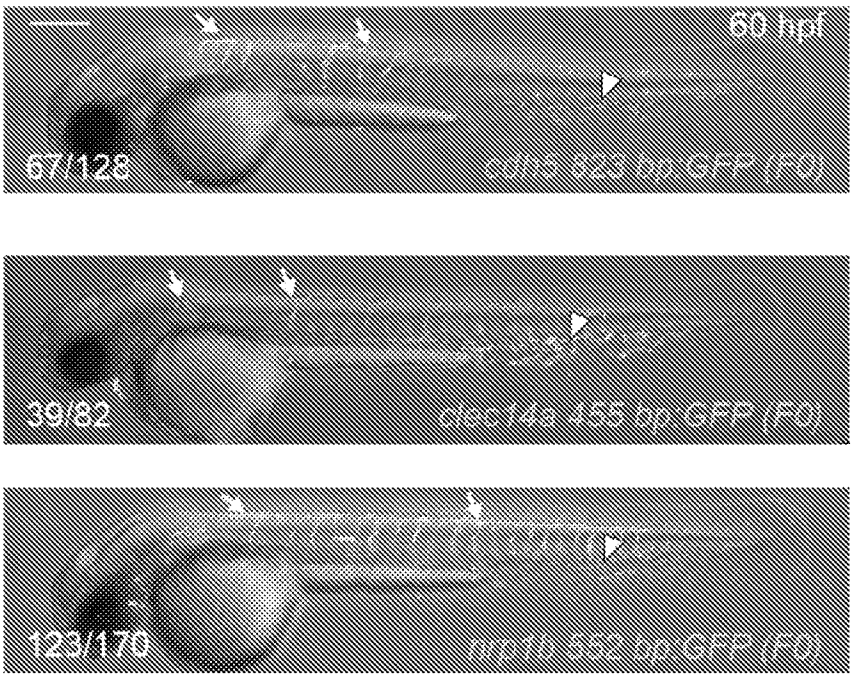

60 hpf
67/128
39/82
123/170

Fig. 9B

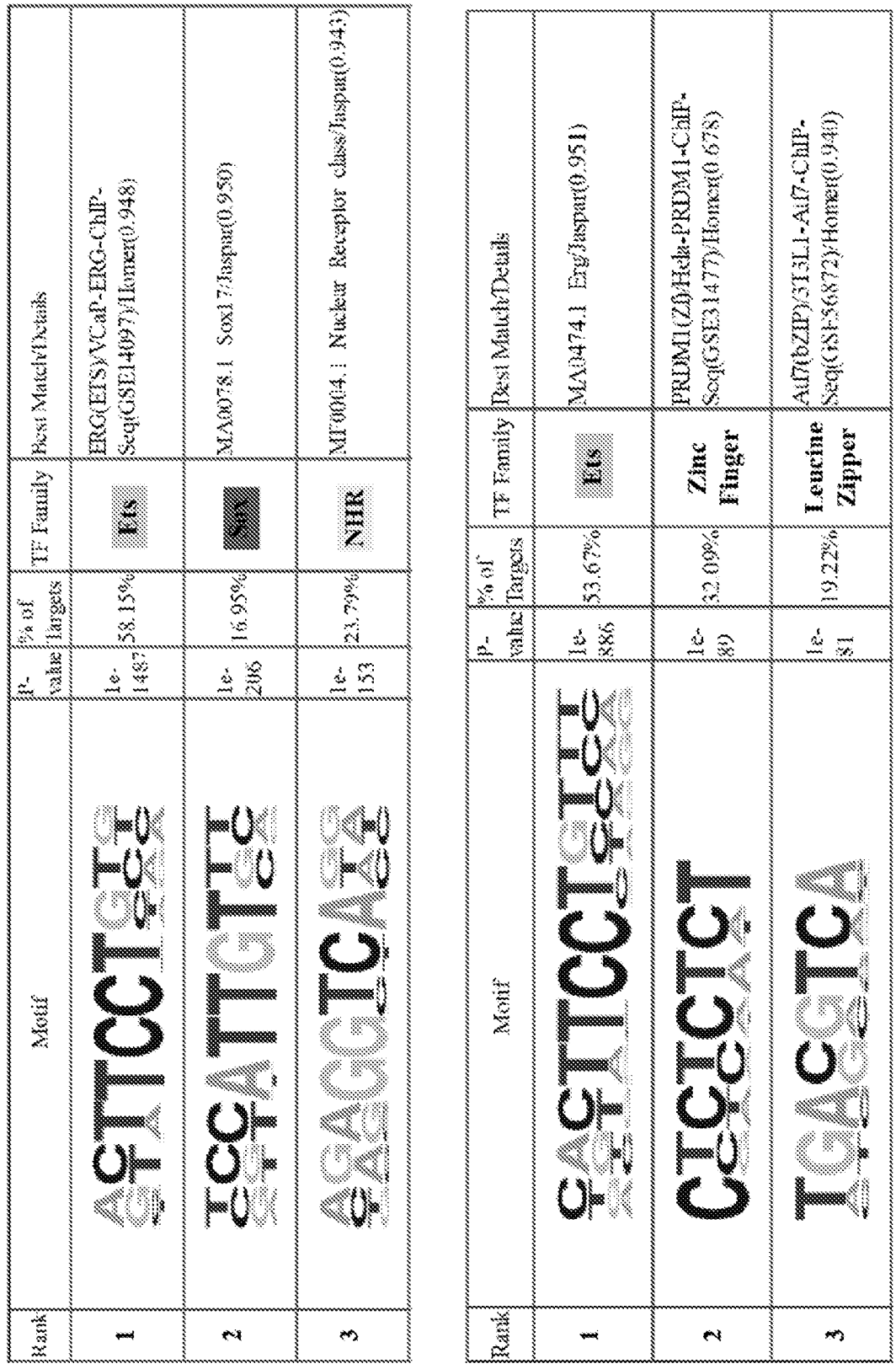

Motifs over-represented in pan-endothelial elements
Rank
Motif
P-value
% of Targets
TF Family
Best Match/Details
1
1e-886
53.67%
Ets
MA0474.1 Erg/Jaspar(0.951)
2
1e-80
32.09%
Zinc Finger
PRDM1(ZnF)/Hela-PRDM1-ChIP-Seq(GSE31477)/Homer(0.678)
3
1e-81
19.22%
Leucine Zipper
Atf7(bZIP)/3T3L1-Atf7-ChIP-Seq(GSE56872)/Homer(0.940)
Motifs over-represented in CHT endothelial elements
Rank
Motif
P-value
% of Targets
TF Family
Best Match/Details
1
1e-1487
58.15%
Ets
ERG(ETS)/VCaP-ERG-ChIP-Seq(GSE14097)/Homer(0.948)
2
1e-206
16.95%
Sox
MA0078.1 Sox17/Jaspar(0.950)
3
1e-153
23.79%
NHR
MF0004.1 Nuclear Receptor class/Jaspar(0.943)

Fig. 9C

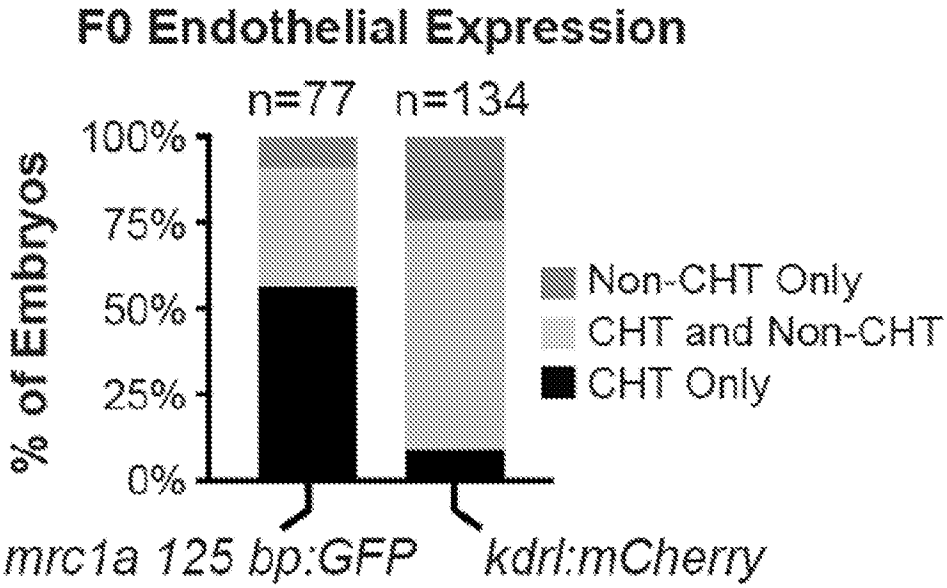


Fig. 10A

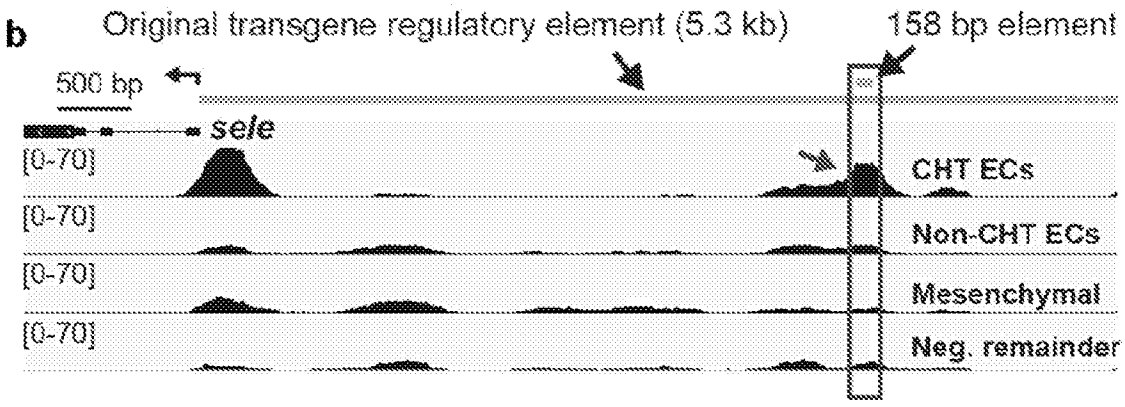


Fig. 10B

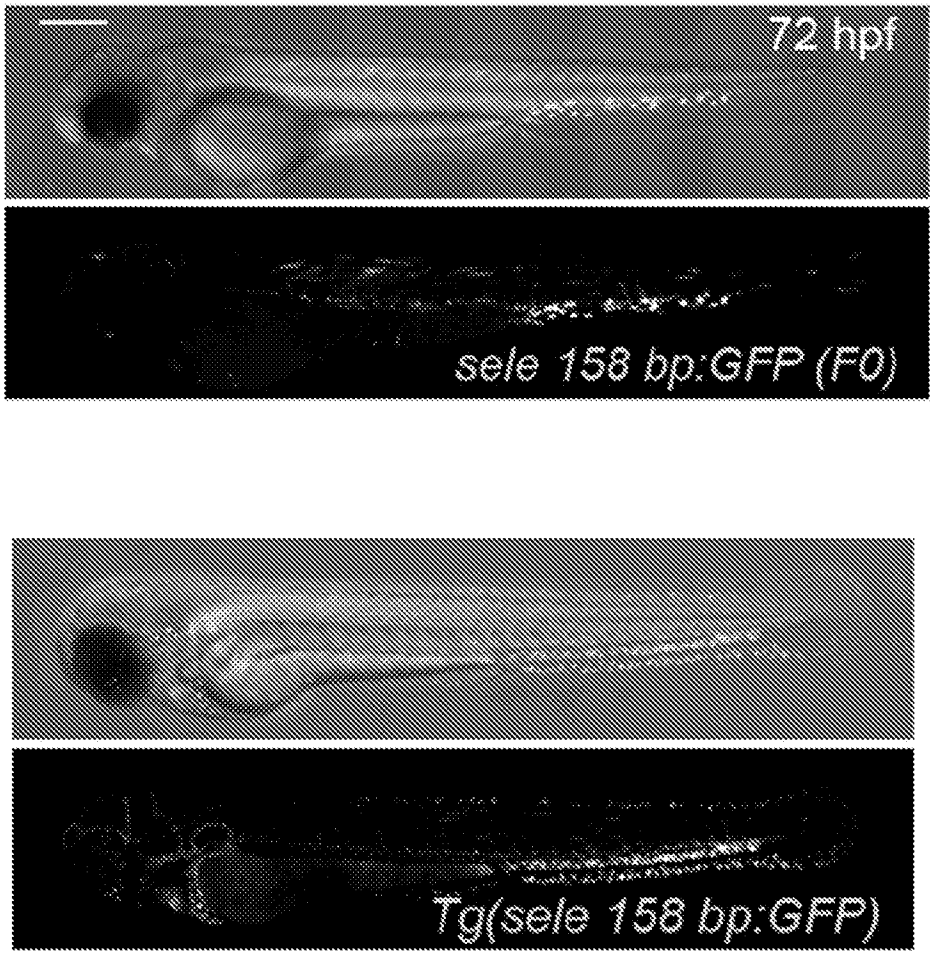


Fig. 10C

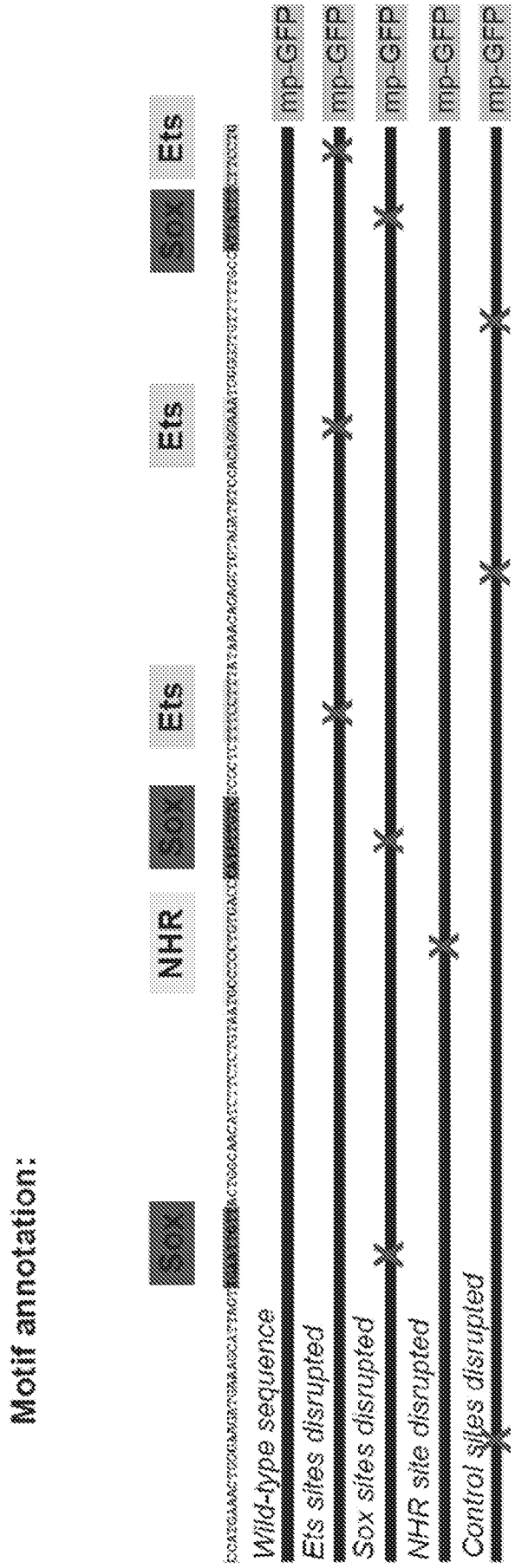

Motif annotation:
Sox
NHR
Sox
Ets
Ets
Sox
Ets
Wild-type sequence
Ets sites disrupted
Sox sites disrupted
NHR site disrupted
Control sites disrupted
mp-GFP
mp-GFP
mp-GFP
mp-GFP
mp-GFP

Fig. 10D

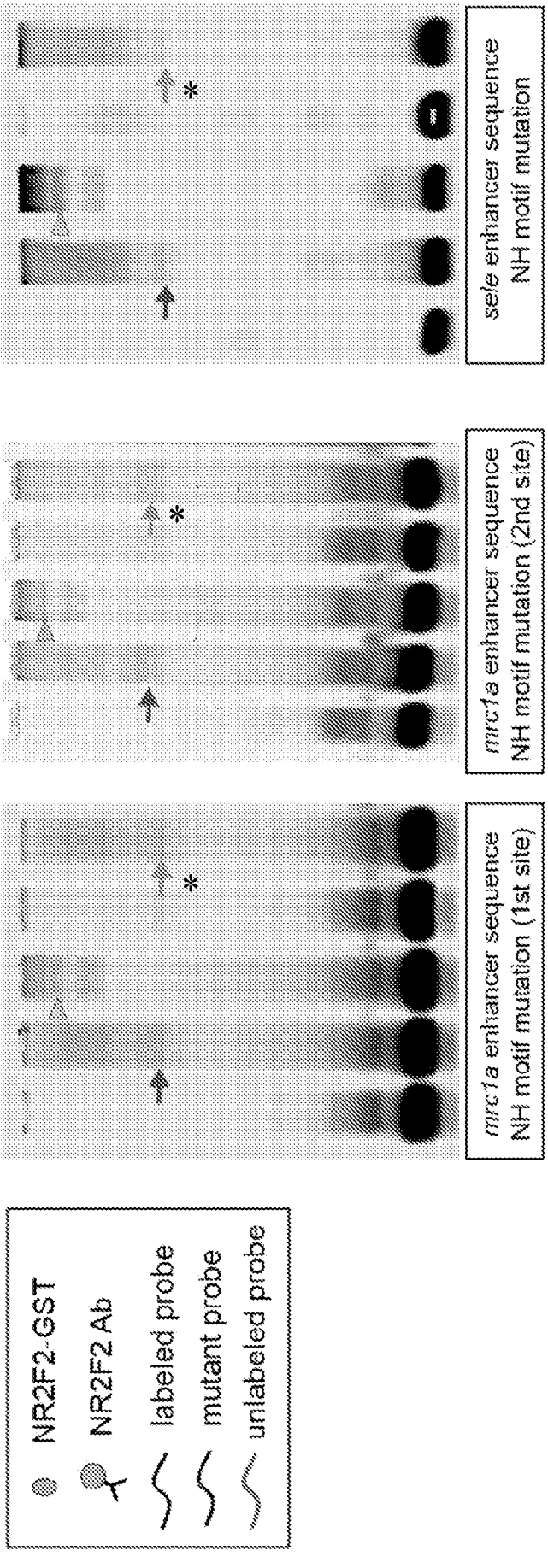

NR2F2-GST
NR2F2 Ab
labeled probe
mutant probe
unlabeled probe
mrc1a enhancer sequence
NH motif mutation (1st site)
mrc1a enhancer sequence
NH motif mutation (2nd site)
sele enhancer sequence
NH motif mutation
*
*
*

Fig. 10E

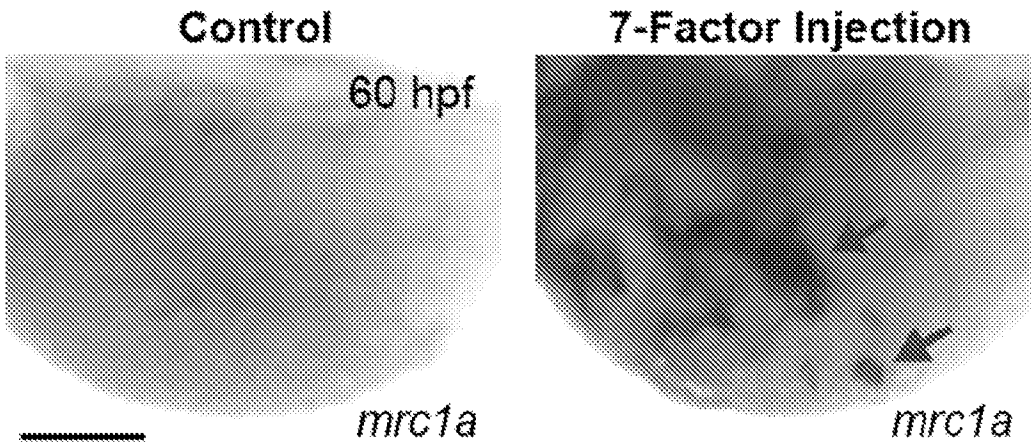


Fig. 11A

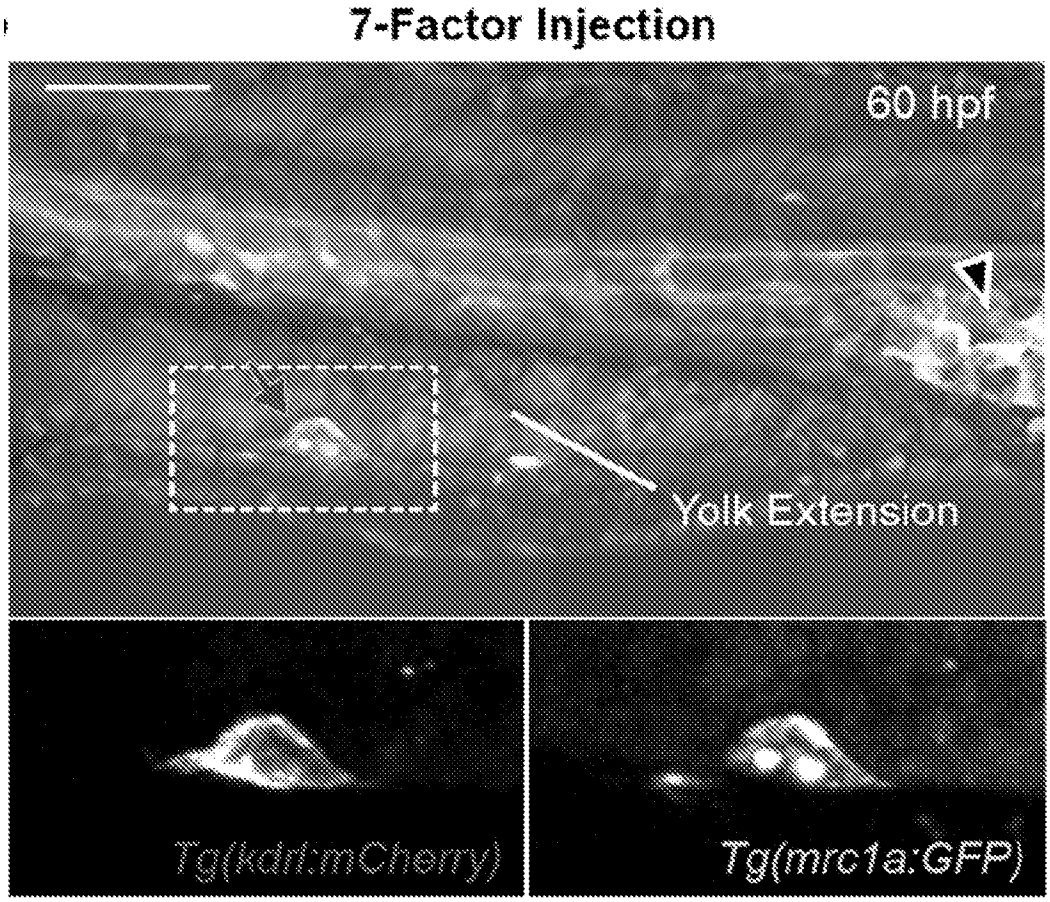


Fig. 11B

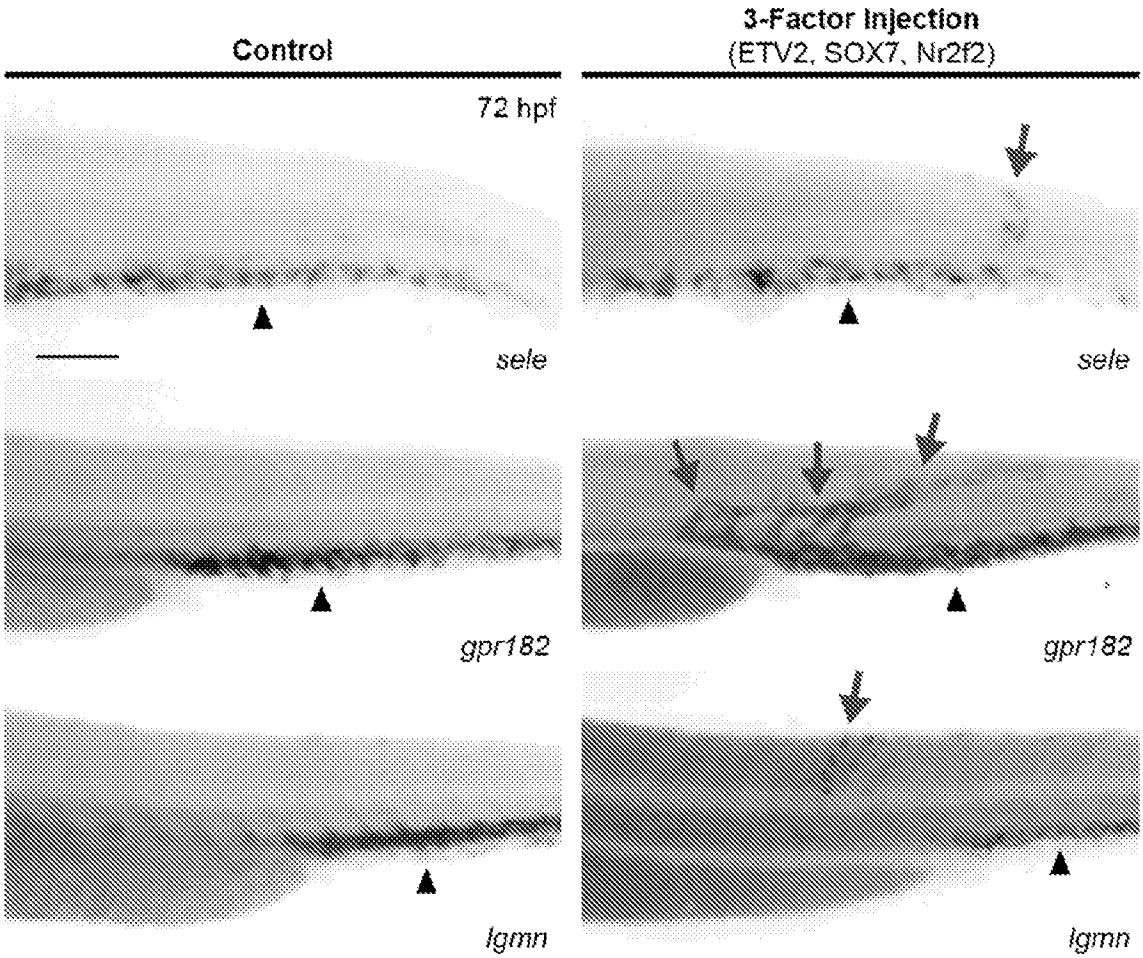

Control
3-Factor Injection
(ETV2, SOX7, Nr2f2)
72 hpf
sele
sele
gpr182
gpr182
lgmn
lgmn

Fig. 11C

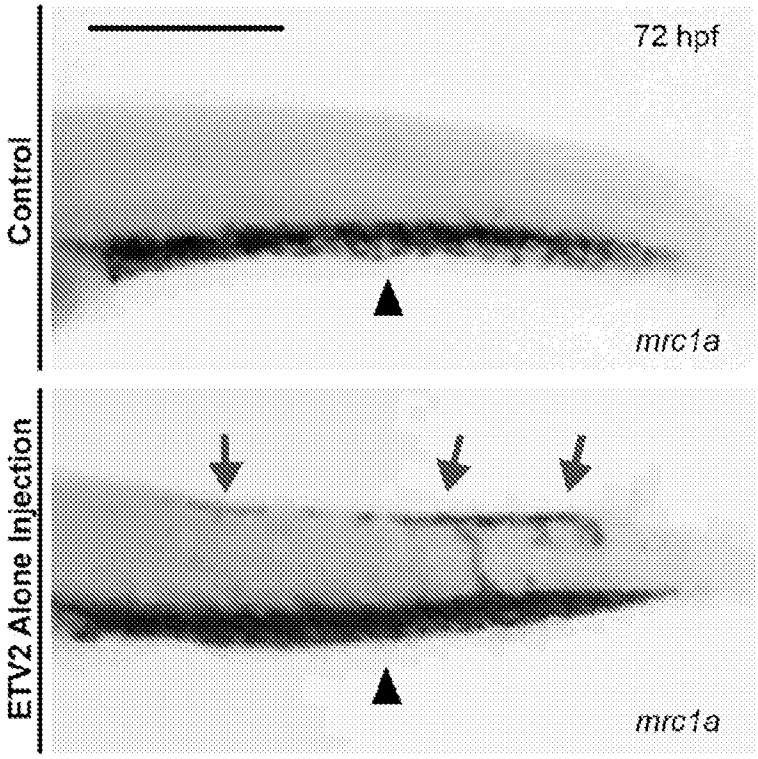


Fig. 12A

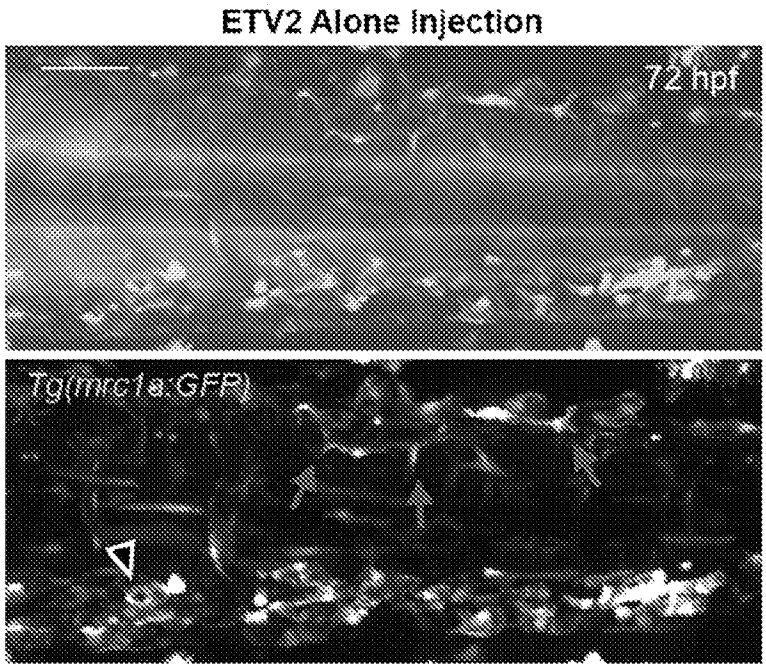


Fig. 12B

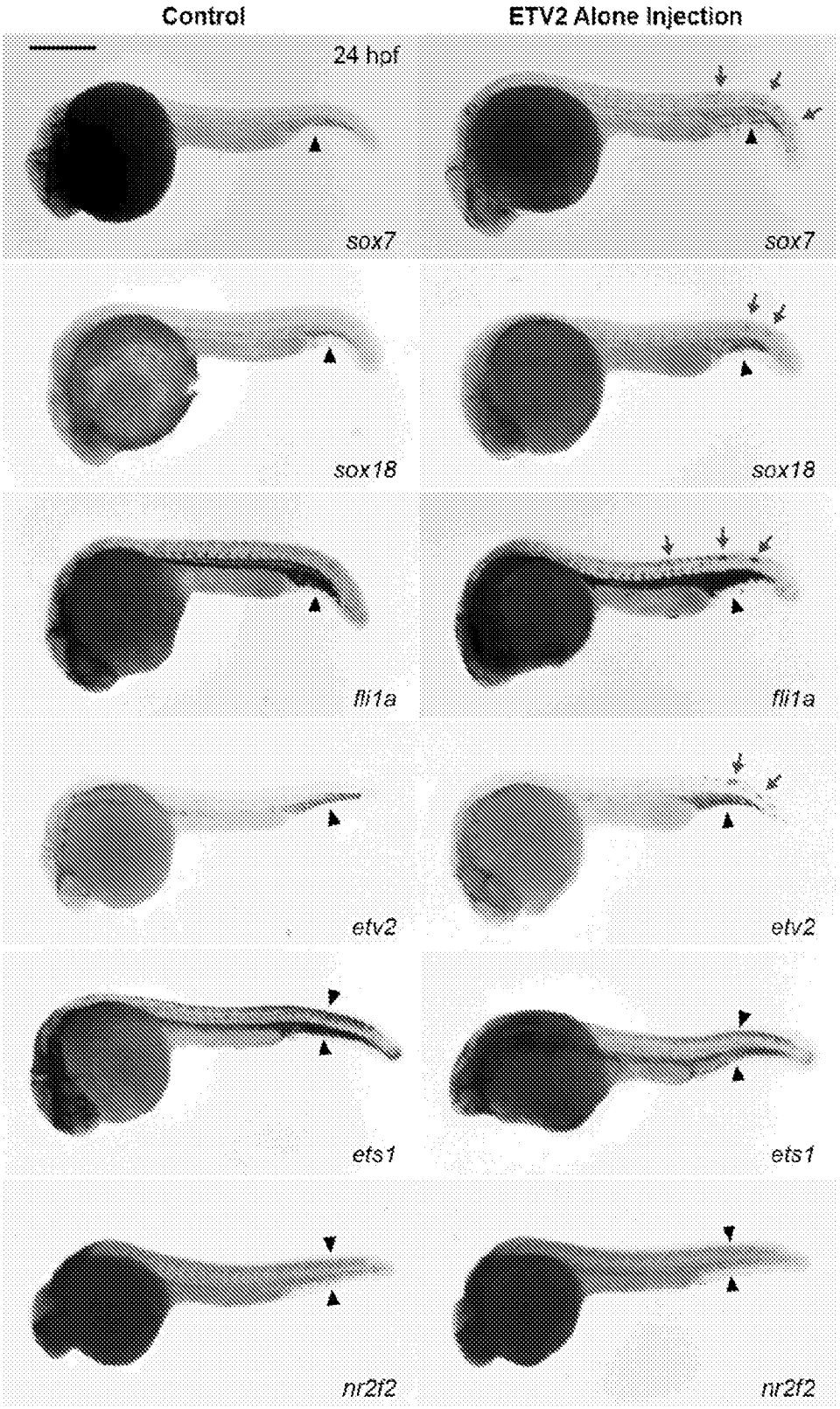

Control
ETV2 Alone Injection
24 hpf
sox7
sox7
sox18
sox18
fli1a
fli1a
etv2
etv2
ets1
ets1
nr2f2
nr2f2

Fig. 12C

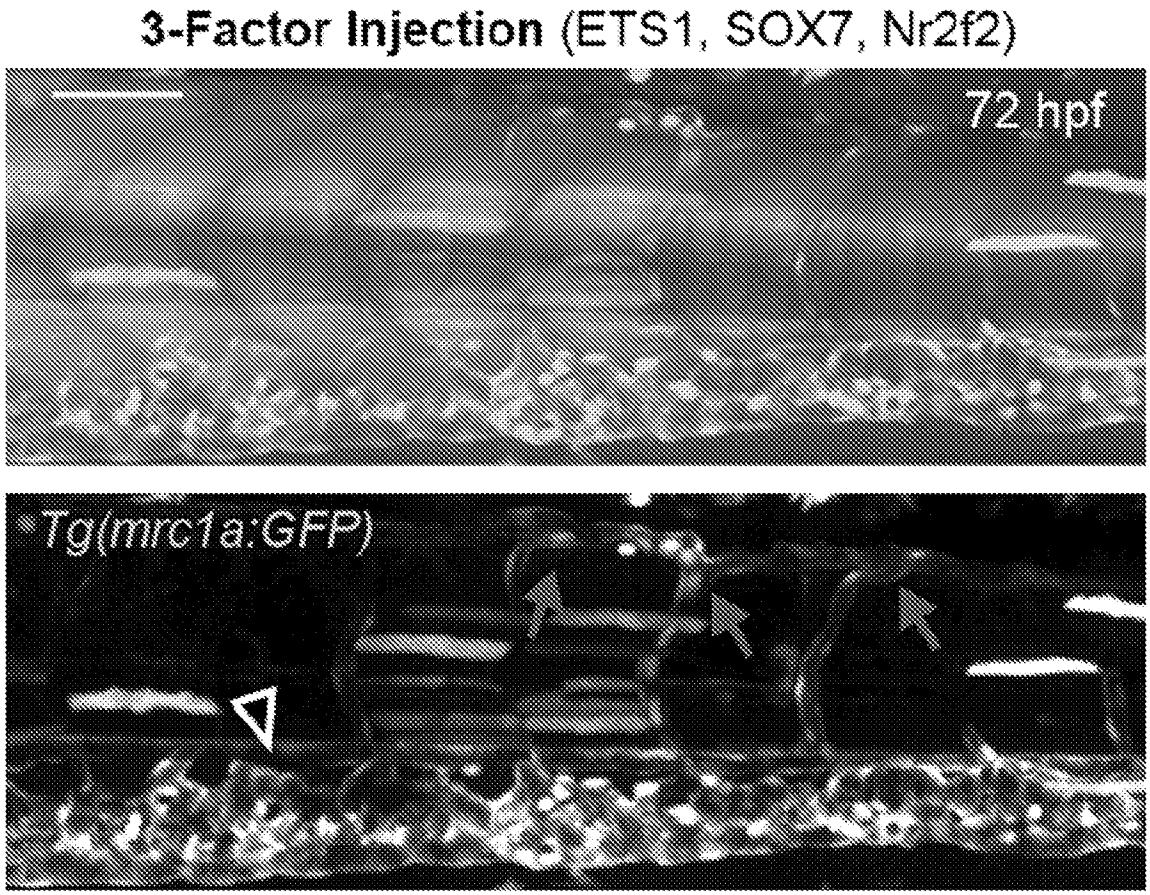


Fig. 12D

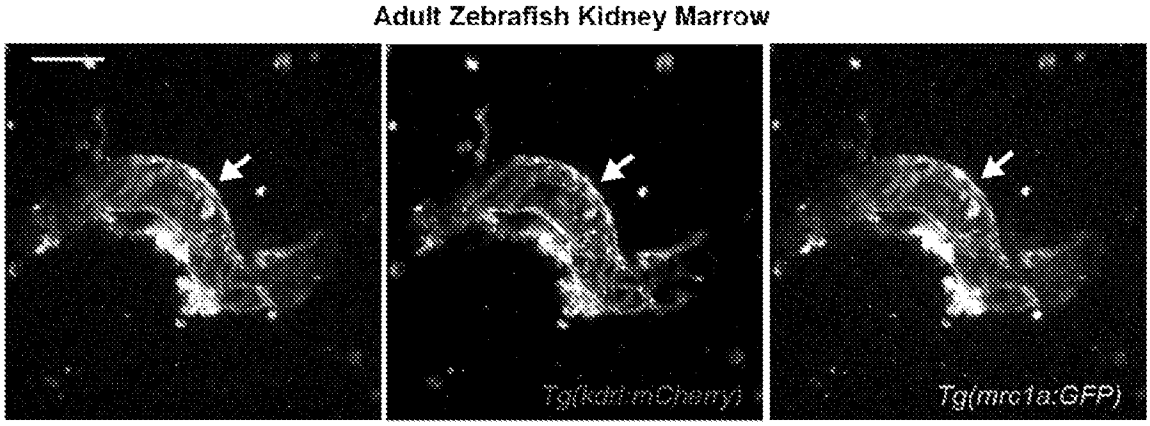


Fig. 13A

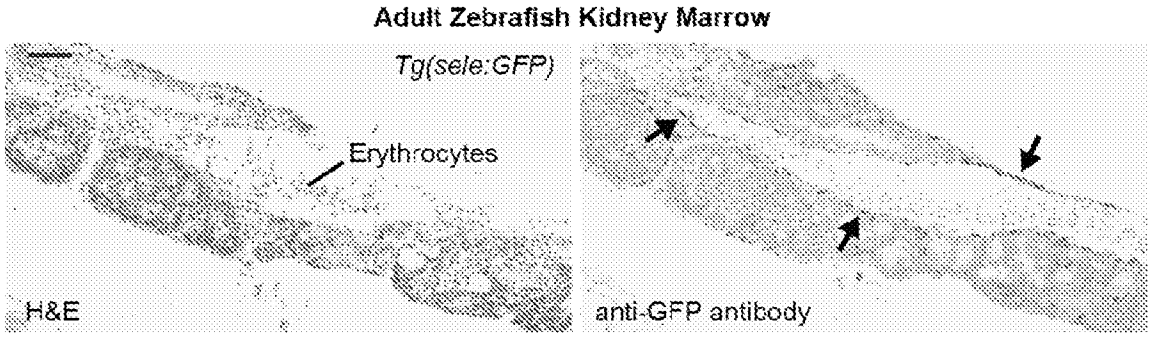


Fig. 13B

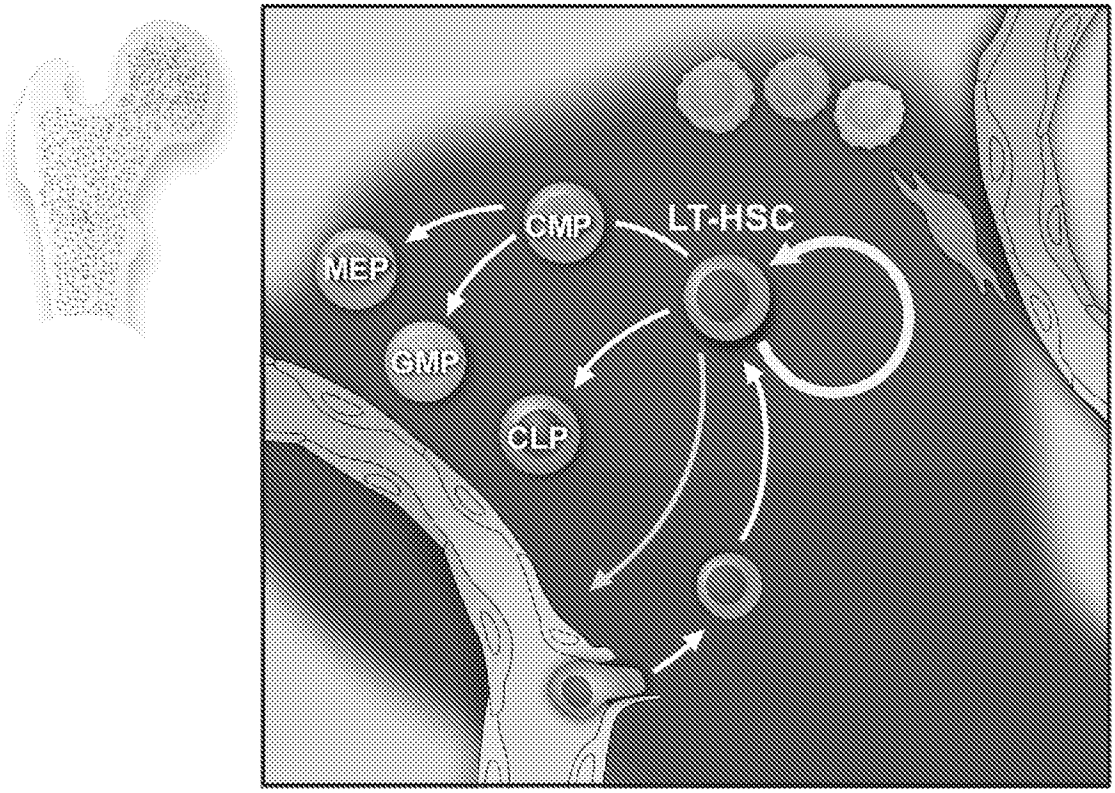


Fig. 14

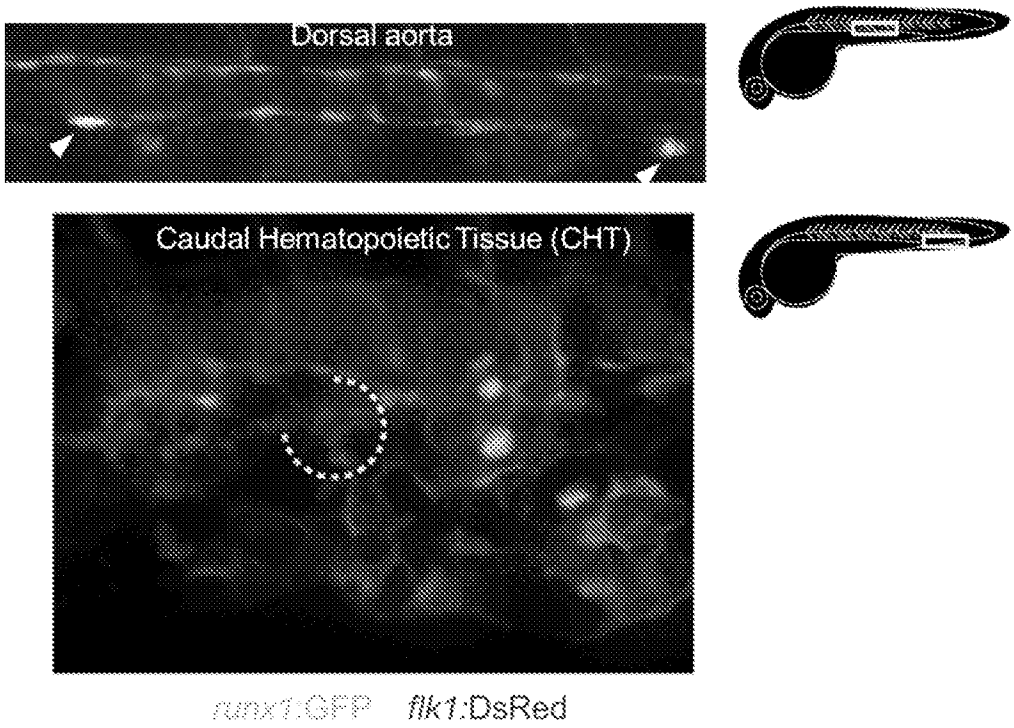


Fig. 15A

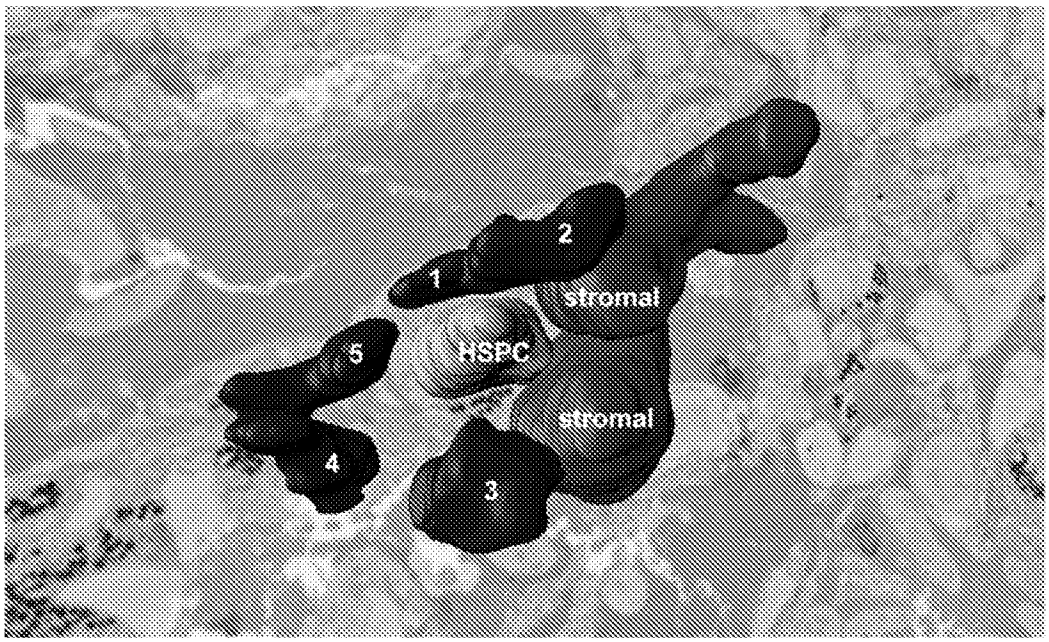


Fig. 15B

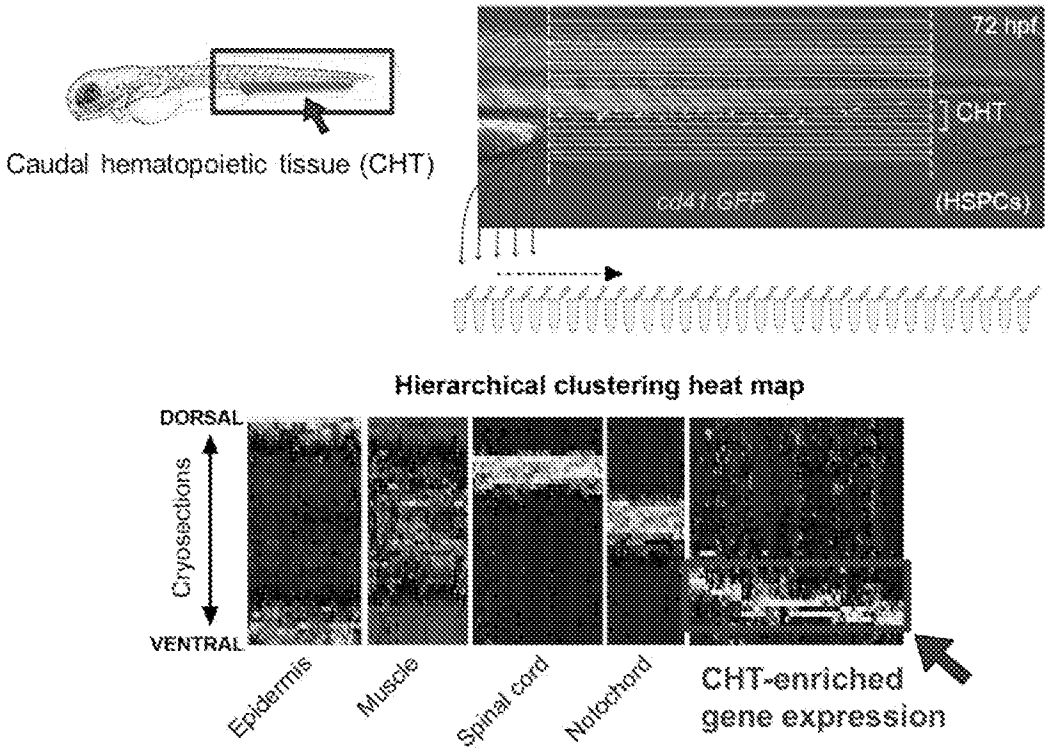

Caudal hematopoietic tissue (CHT)
72 hpf
CHT
(HSPCs)
Hierarchical clustering heat map
DORSAL
Cryosections
VENTRAL
Epidermis
Muscle
Spinal cord
Notochord
CHT-enriched
gene expression

Fig. 16

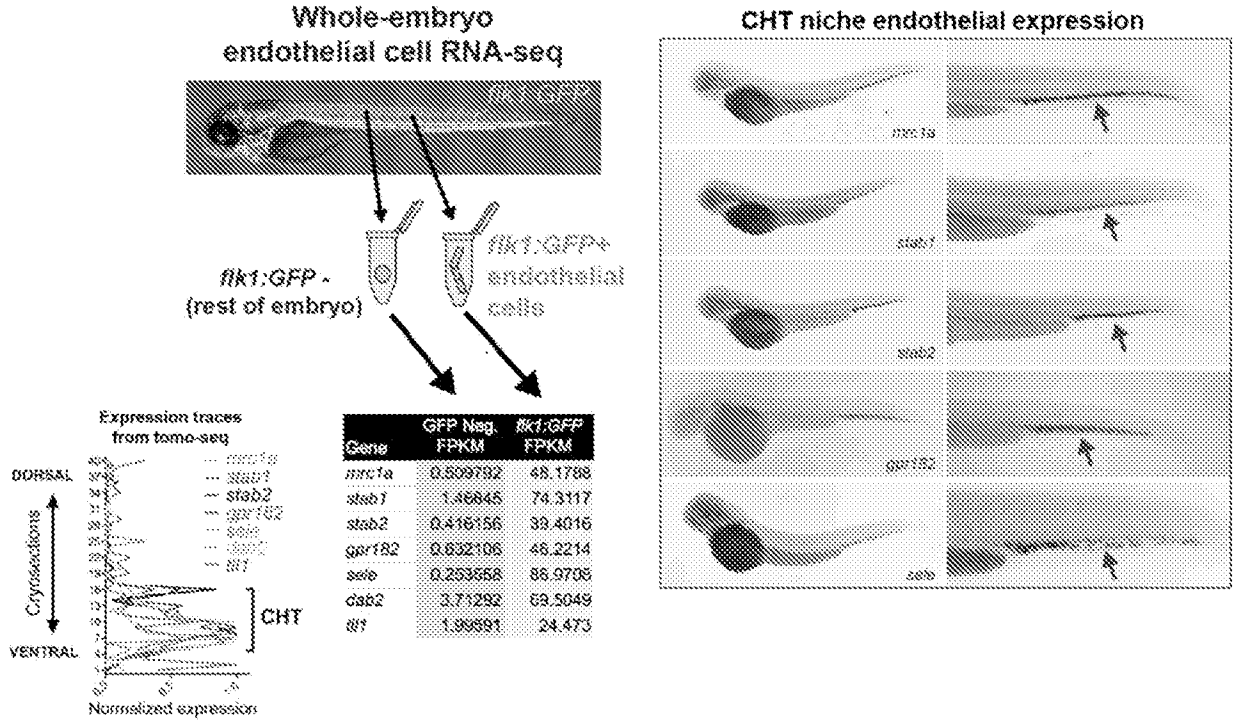

Whole-embryo
endothelial cell RNA-seq
flk1:GFP -
(rest of embryo)
flk1:GFP+
endothelial
cells
Expression traces
from tomo-seq
DORSAL
Cryosections
VENTRAL
CHT
Normalized expression
Gene | GFP Neg. FPKM | flk1:GFP FPKM
mrc1a | 0.509792 | 48.1788
stab1 | 1.46845 | 74.3117
stab2 | 0.416156 | 39.4016
gpr182 | 0.632106 | 46.2214
selo | 0.253658 | 86.9706
dab2 | 3.71292 | 69.5049
tll1 | 1.39691 | 24.473
CHT niche endothelial expression
mrc1a
stab1
stab2
gpr182
selo

Fig. 17

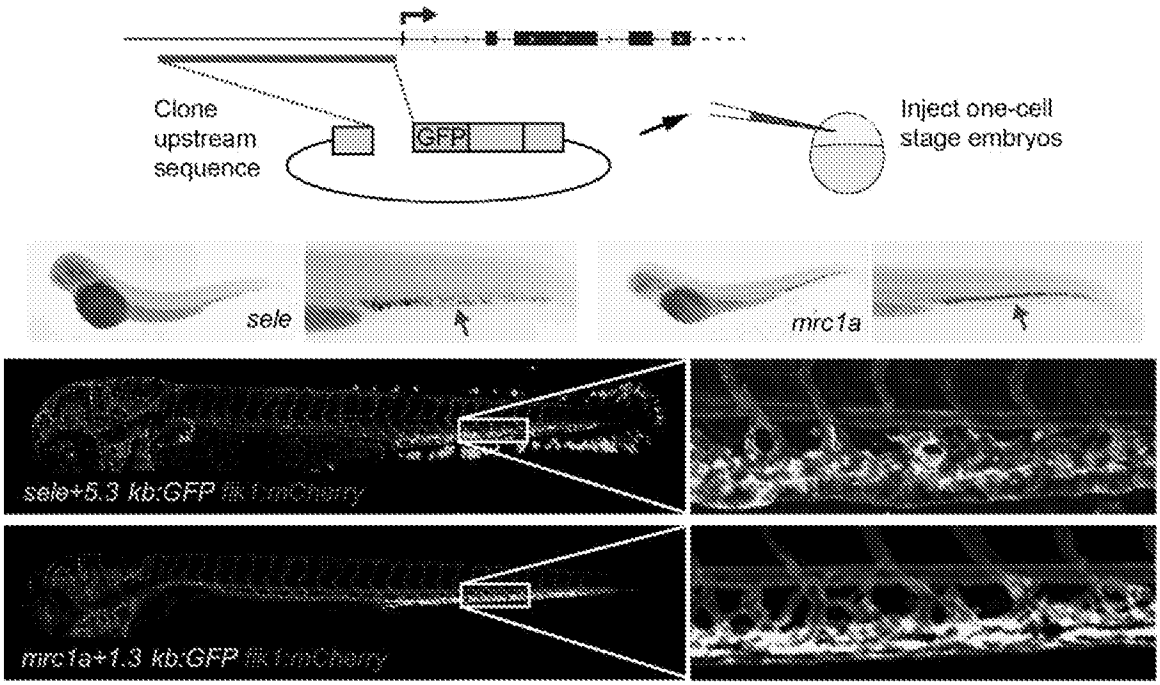

Clone
upstream
sequence
GFP
Inject one-cell
stage embryos
sele
mrc1a
sele+5.3 kb:GFP
mrc1a+1.3 kb:GFP

Fig. 18

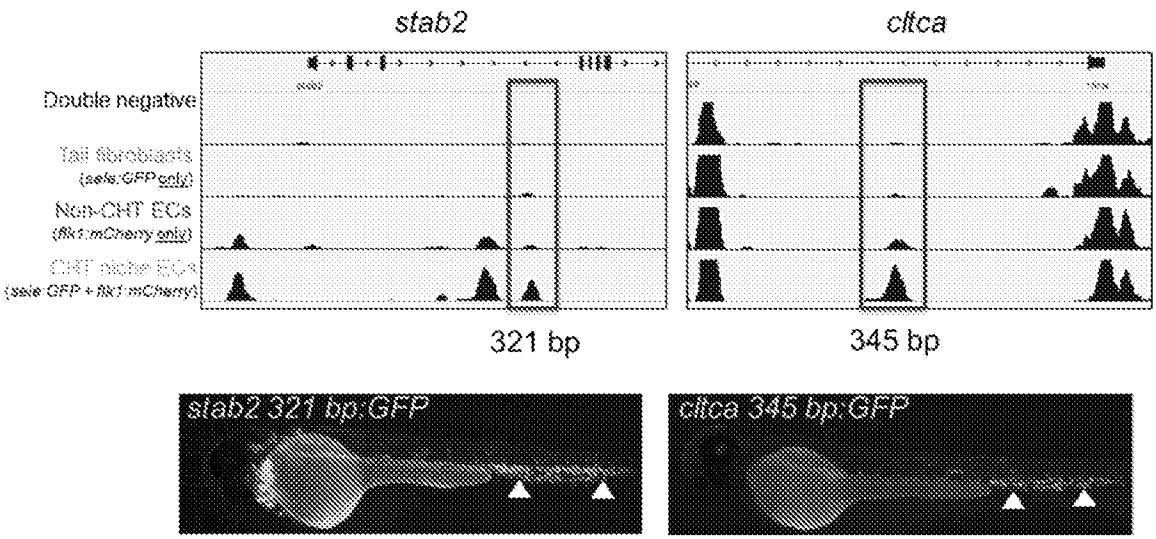

stab2
cltca
Double negative
Non-CHT ECs
321 bp
345 bp
stab2 321 bp:GFP
cltca 345 bp:GFP

Fig. 19



	P-value
Ets family (Fli1, Ets1)	1e-647
Sox family (Sox7, Sox17)	1e-124
Gata factors (Gata4, Gata2)	1e-72
Nuclear Hormone (Nr2f2, Rora)	1e-63

Fig. 20

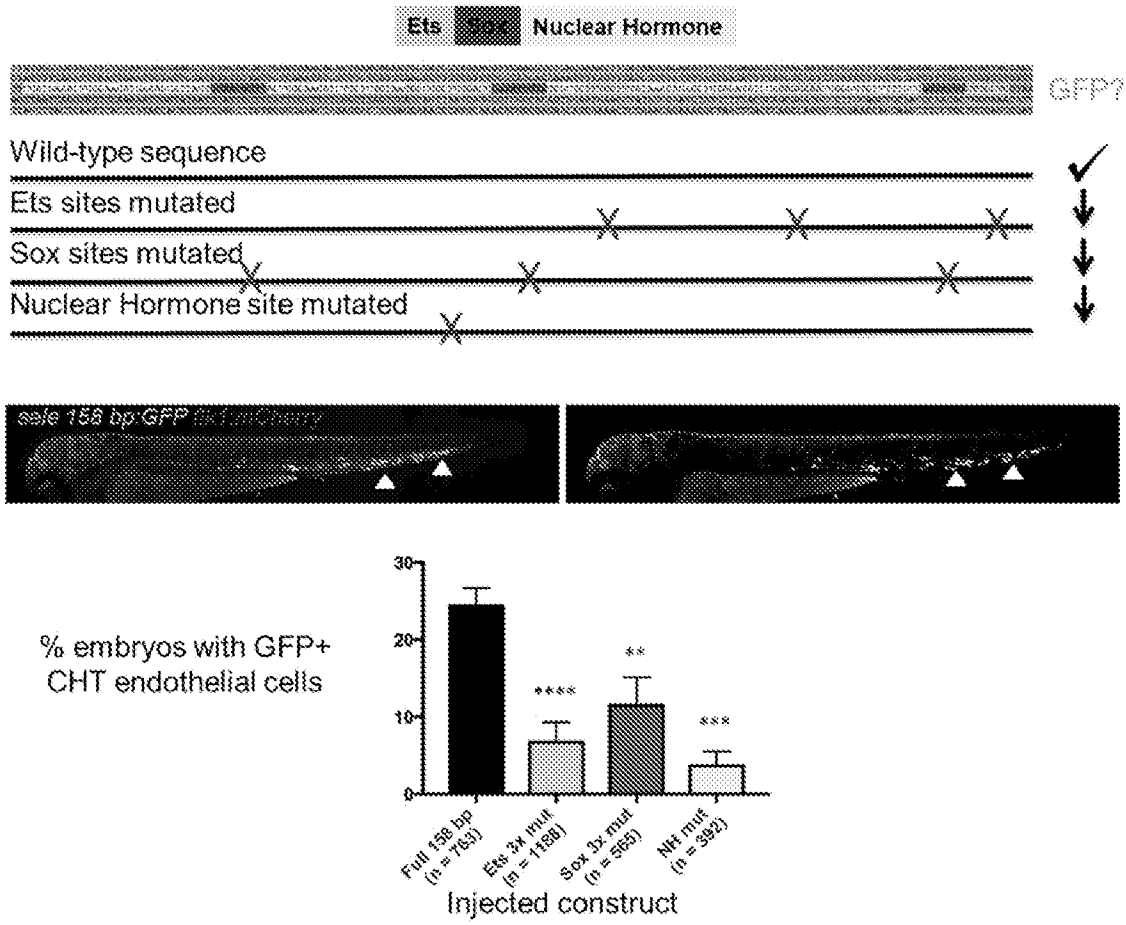


Fig. 21

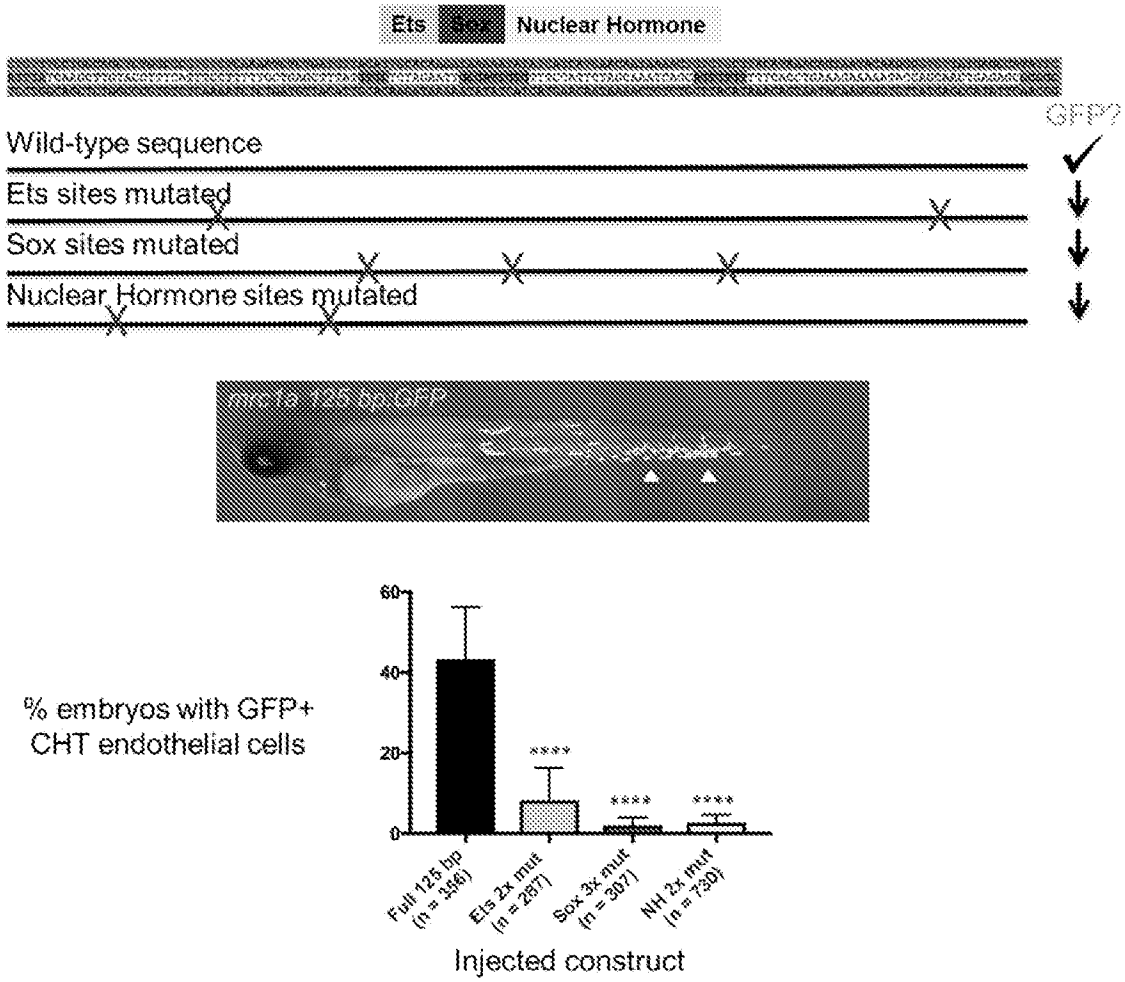

Ets
Sox
Nuclear Hormone
GFP?
Wild-type sequence
Ets sites mutated
Sox sites mutated
Nuclear Hormone sites mutated
% embryos with GFP+
CHT endothelial cells
60
40
20
0
****
****
****
Full 125 bp
(n = 356)
Ets 2x mut
(n = 287)
Sox 3x mut
(n = 307)
NH 2x mut
(n = 730)
Injected construct

Fig. 22

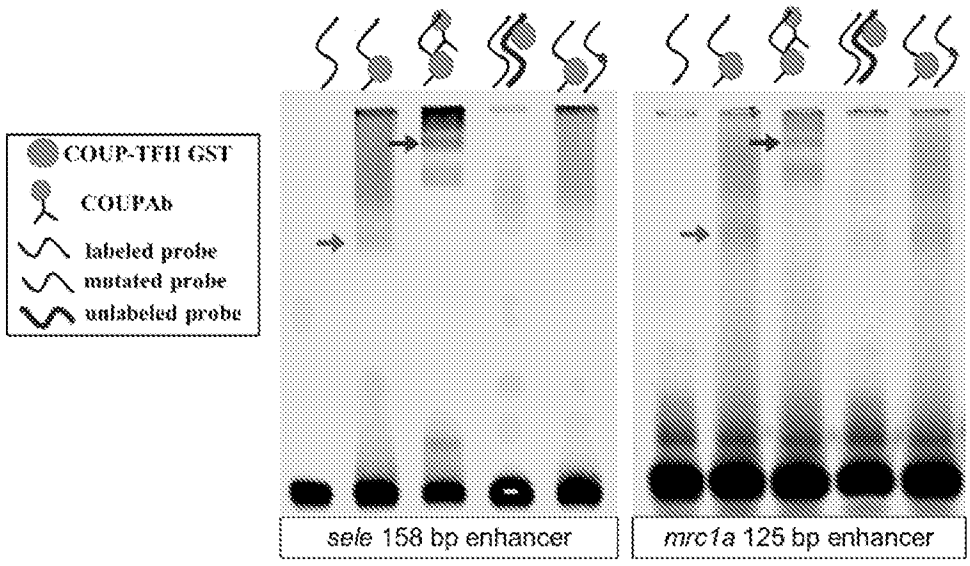

COUP-TFII GST
COUPAb
labeled probe
mutated probe
unlabeled probe
sele 158 bp enhancer
mrc1a 125 bp enhancer

Fig. 23

TF Family	Gene	CHT endothelial cells (GFP+ mCherry+) FPKM
Ets	etv2	192.3
	fli1a	480.4
	ets1	183.0
Sox	sox18	206.4
	sox7	125.1
Nuclear Hormone	rxraa	45.9
	nr2f2	84.6

Are a combination of these transcription factors sufficient to induce an ectopic niche?

Fig. 24

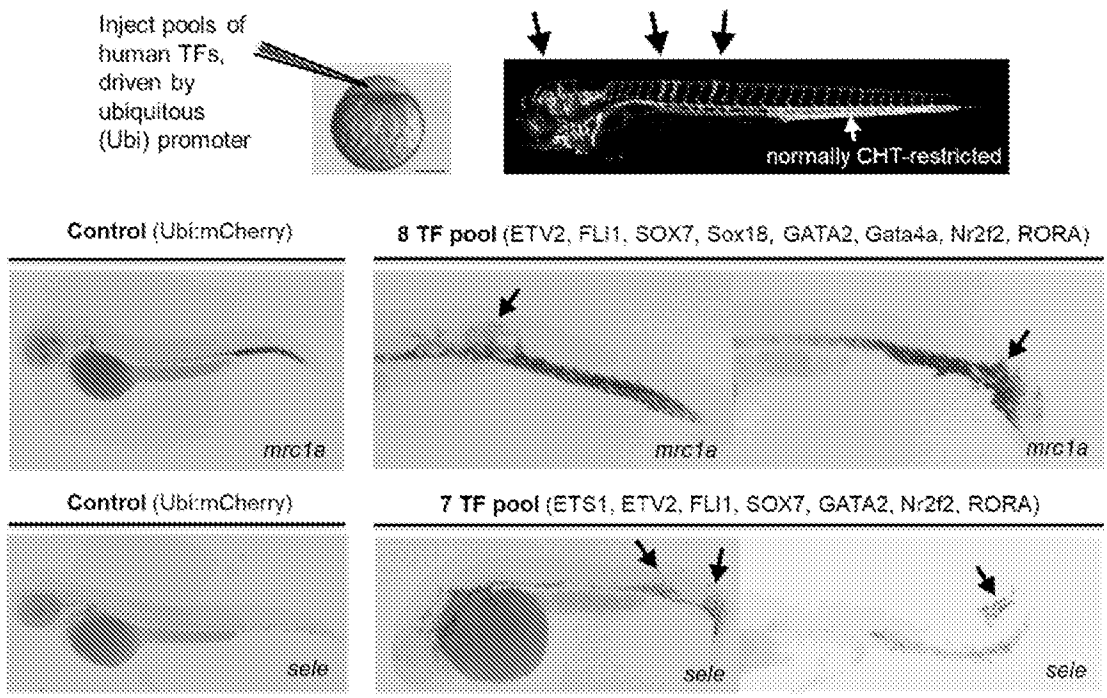


Fig. 25

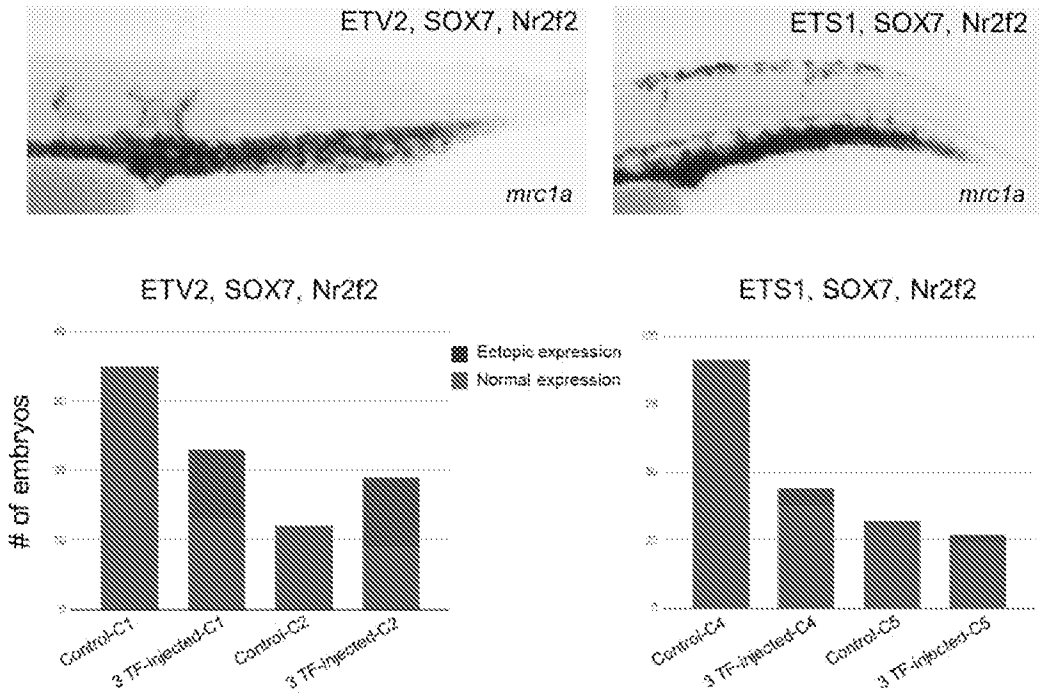


Fig. 26

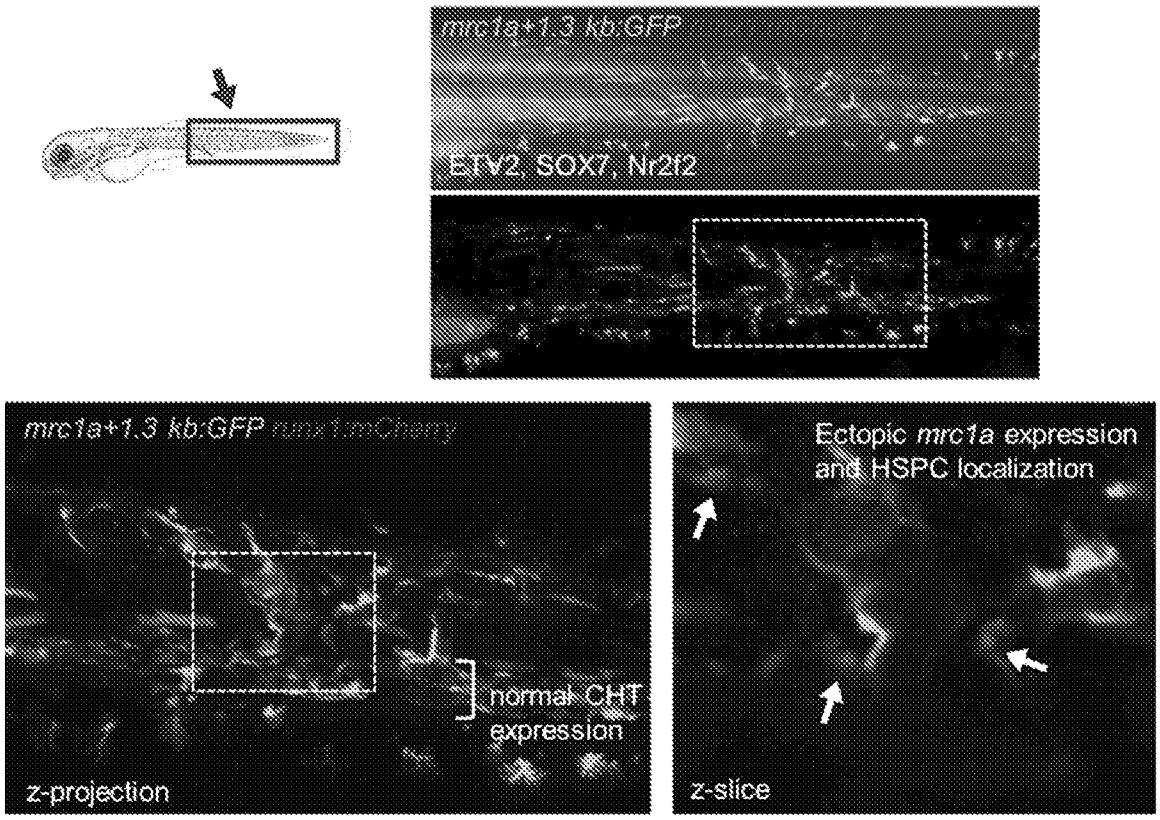


Fig. 27A

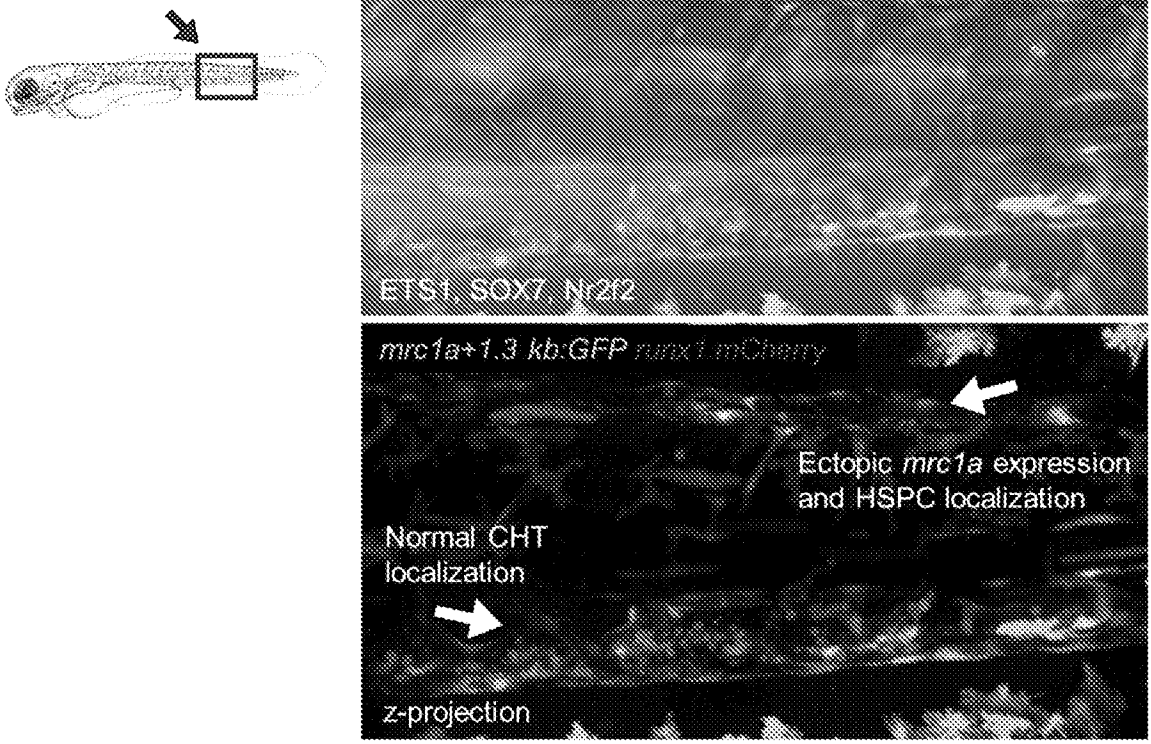


Fig. 27B

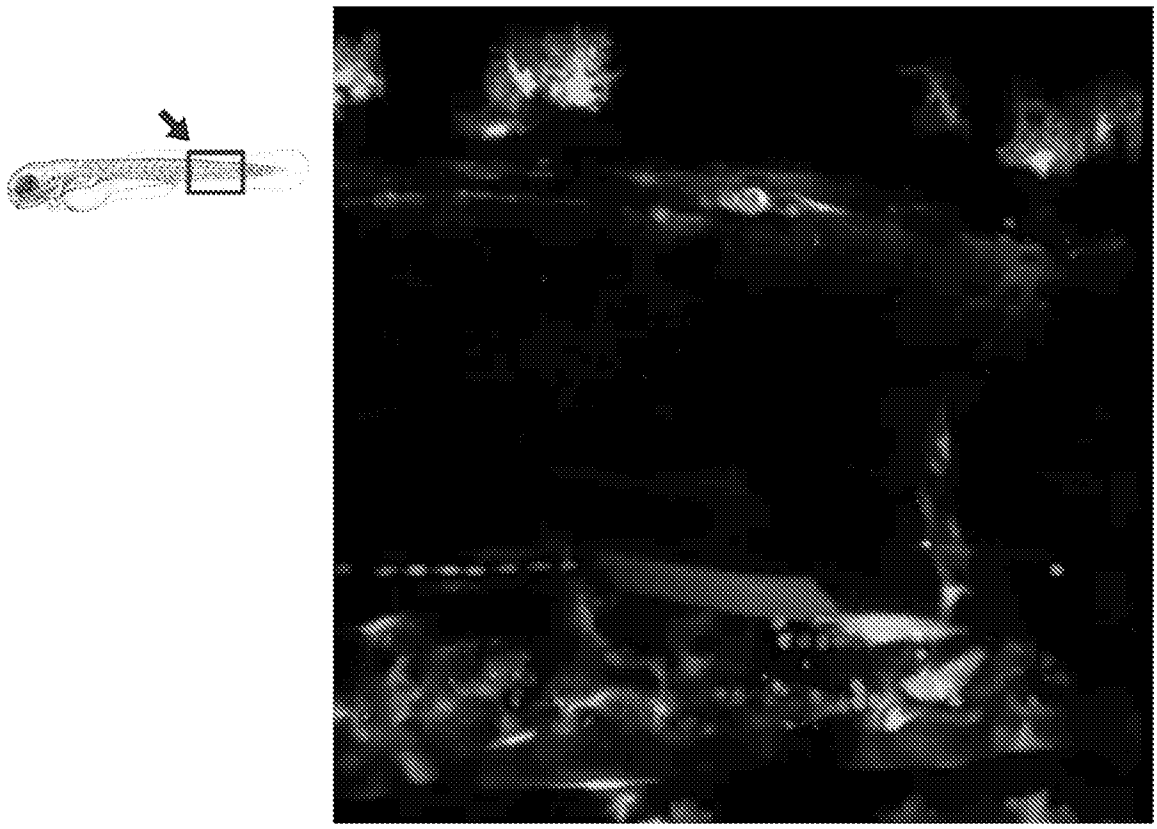


Fig. 27C

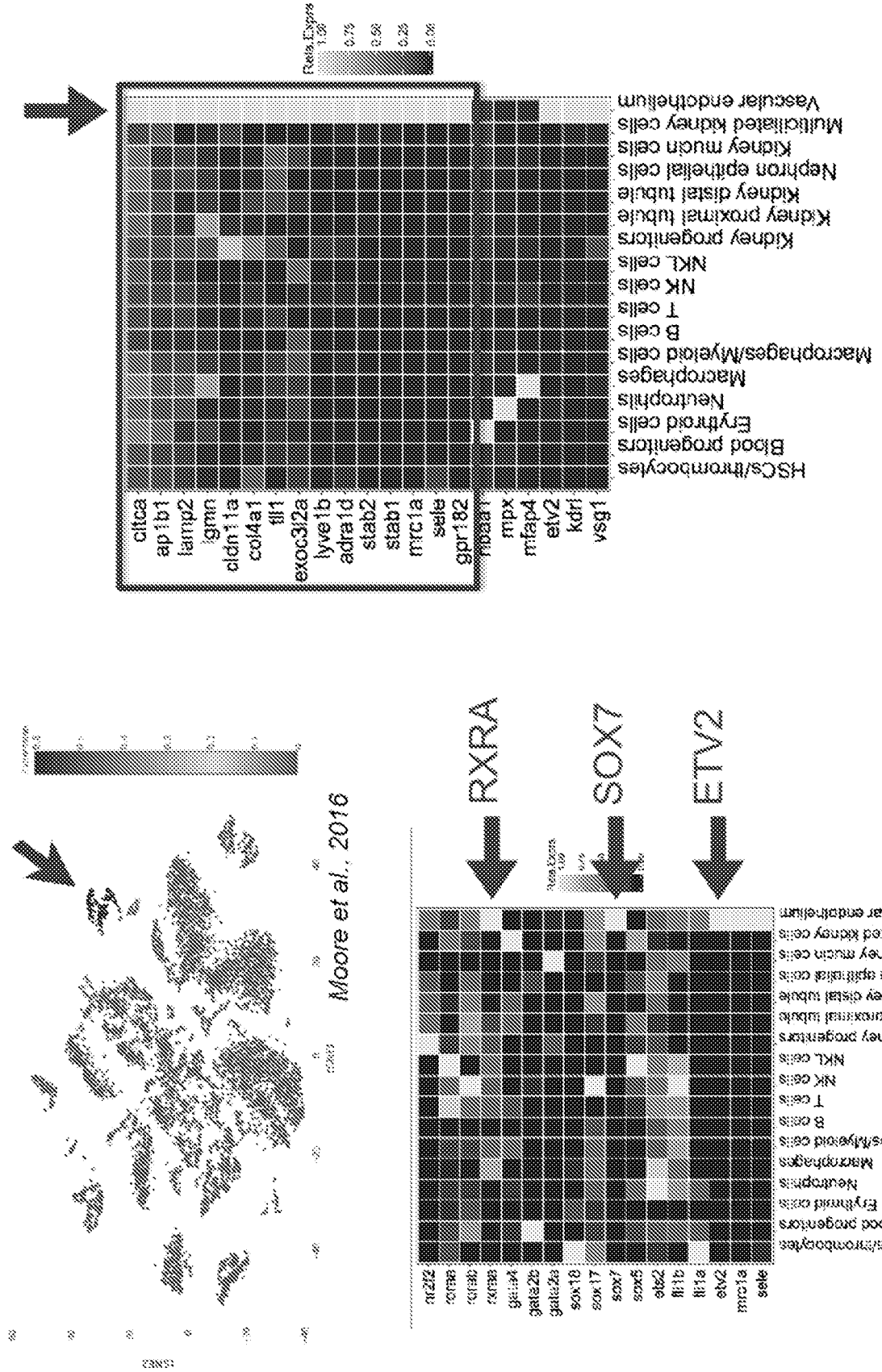

RXRA
SOX7
ETV2
Moore et al., 2016
HSCs/thrombocytes
Blood progenitors
Erythroid cells
Neutrophils
Macrophages
Macrophages/Myeloid cells
B cells
T cells
NK cells
NKL cells
Kidney progenitors
Kidney proximal tubule
Kidney distal tubule
Nephron epithelial cells
Kidney mucin cells
Multiciliated kidney cells
Vascular endothelium
citca
ap1b1
lamp2
gmn
cldn11a
col4a1
fli1
exoc3l2a
lyve1b
adra1d
stab2
stab1
mrc1a
sele
gpr182
mpx
mfap4
etv2
kdrl
vsg1
rxra
rxrab
rxrbb
rxrga
gata1a
gata2b
gata2a
sox18
sox17
sox7
sox5
ets2
fli1b
fli1a
etv2
mrc1a
sele

Fig. 28

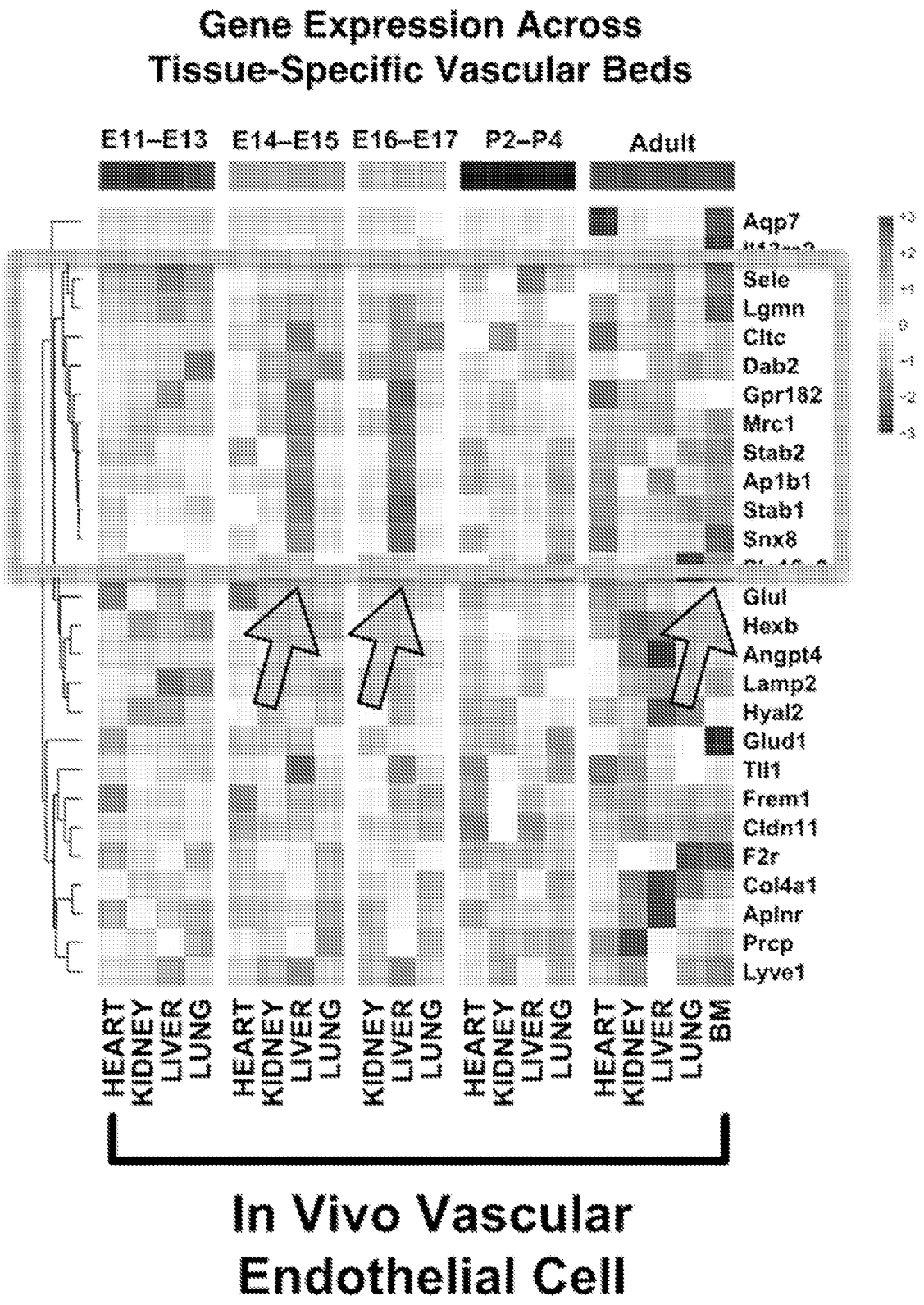

Gene Expression Across
Tissue-Specific Vascular Beds
E11–E13
E14–E15
E16–E17
P2–P4
Adult
Aqp7
Sele
Lgmn
Cltc
Dab2
Gpr182
Mrc1
Stab2
Ap1b1
Stab1
Snx8
Glul
Hexb
Angpt4
Lamp2
Hyal2
Glud1
Tll1
Frem1
Cldn11
F2r
Col4a1
Apinr
Prcp
Lyve1
HEART
KIDNEY
LIVER
LUNG
BM
In Vivo Vascular
Endothelial Cell

Fig. 29

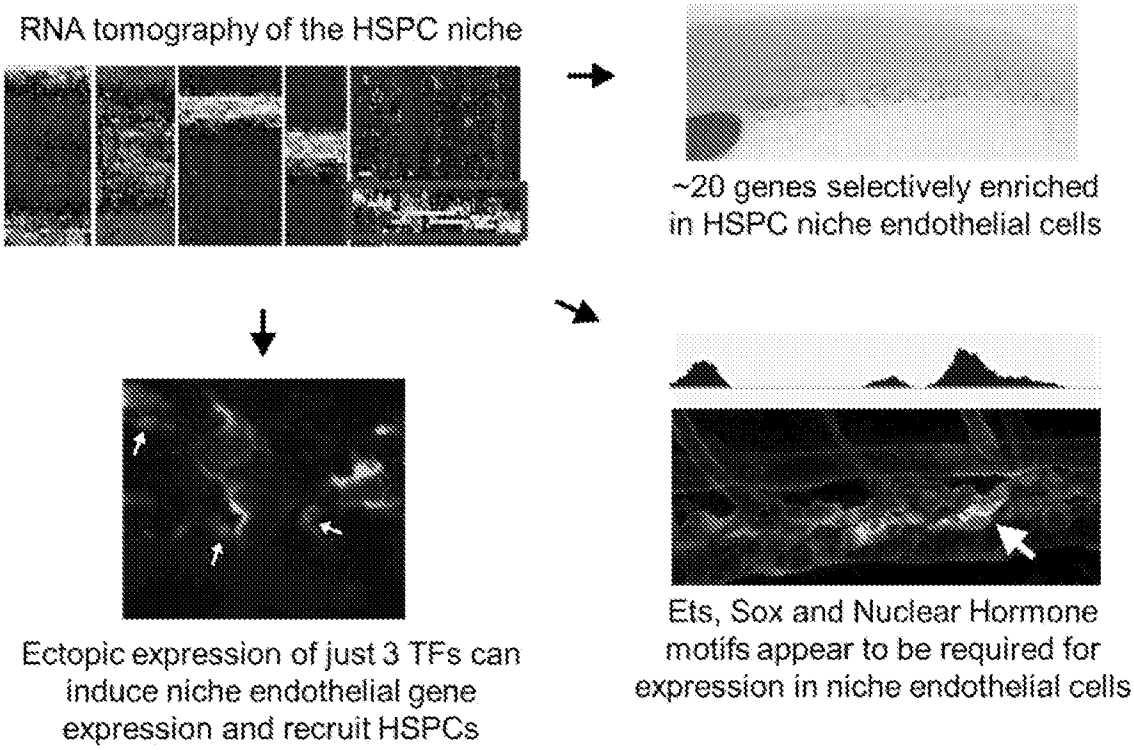

RNA tomography of the HSPC niche
~20 genes selectively enriched in HSPC niche endothelial cells
Ectopic expression of just 3 TFs can induce niche endothelial gene expression and recruit HSPCs
Ets, Sox and Nuclear Hormone motifs appear to be required for expression in niche endothelial cells

Fig. 30

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US19/23637

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. ☐ forming part of the international application as filed:

      ☐ in the form of an Annex C/ST.25 text file.

      ☐ on paper or in the form of an image file.

   b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. ☒ furnished subsequent to the international filing date for the purposes of international search only:

      ☒ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. ☒ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US19/23637

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 25, 30
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US19/23637

A. CLASSIFICATION OF SUBJECT MATTER
IPC - A61K 35/28; C12N 5/069, 5/0647, 15/85 (2019.01)
CPC - A61K 35/28; C12N 5/069, 5/0647, 15/85

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
See Search History document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	WO 2016/201133 A2 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA) 15 December 2016; page 10, 3rd paragraph; page 19, 1st paragraph; page 20, 1st paragraph; page 21, 2nd paragraph; page 52, 1st column; page 52, 1st paragraph; page 56, 1st paragraph; page 57, 2nd paragraph; page 66, 3rd paragraph; page 71, 4th paragraph – page 72, 1st paragraph; page 87, 2nd paragraph – page 88, 3rd paragraph; page 94, 1st paragraph; page 96, 1st paragraph; page 118, 2nd paragraph – page 121, 3rd paragraph; page 120, 2nd paragraph; page 121, 2nd-4th paragraphs; page 122, 5th paragraph; Figures 13, 16	1-3, 4/1-3, 7-17, 19, 24, 31 --- 26-27
X	WO 2017/015245 A1 (ANGIOCRINE BIOSCIENCE, INC.) 26 January 2017; abstract; page 20, lines 12-13; page 20, lines 23-25; page 33, lines 20-22; page 134, lines 17-18; claims 1, 9	15, 20-22, 28
X	(SUN, J et al.) E-Selectin Mediated Adhesion and Migration of Endothelial Colony Forming Cells Is Enhanced by SDF-1[alpha]/CXCR4. PLoS ONE. 2 April 2013, Vol. 8, No. 4, e60890; page e60890, 2nd column, 4th paragraph; DOI: 10.1371/journal.pone.0060890	2, 4/2
X	(LETARTE, M et al.) Reduced endothelial secretion and plasma levels of transforming growth factor-[beta]1 in patients with hereditary hemorrhagic telangiectasia type 1. Cardiovascular Research. 23 May 2005, Vol. 68, pages 155-164; abstract; DOI: 10.1016/j.cardiores.2005.04.028	17
Y	US 2014/0017784 A1 (WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH) 16 January 2014; abstract; paragraphs [0007], [0093]-[0094], [0116]; claims 39, 55	26-27
A	WO 2014/113415 A1 (CORNELL UNIVERSITY) 24 July 2014; paragraph [0009]; claim 1	6, 18, 23

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
29 March 2019 (29.05.2019)	10 JUL 2019
Name and mailing address of the ISA/ Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Authorized officer Shane Thomas PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (January 2015)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US19/23637

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2009/042910 A2 (UNIVERSITY OF SOUTH FLORIDA et al.) 2 April 2009; page 3, lines 5-8	29