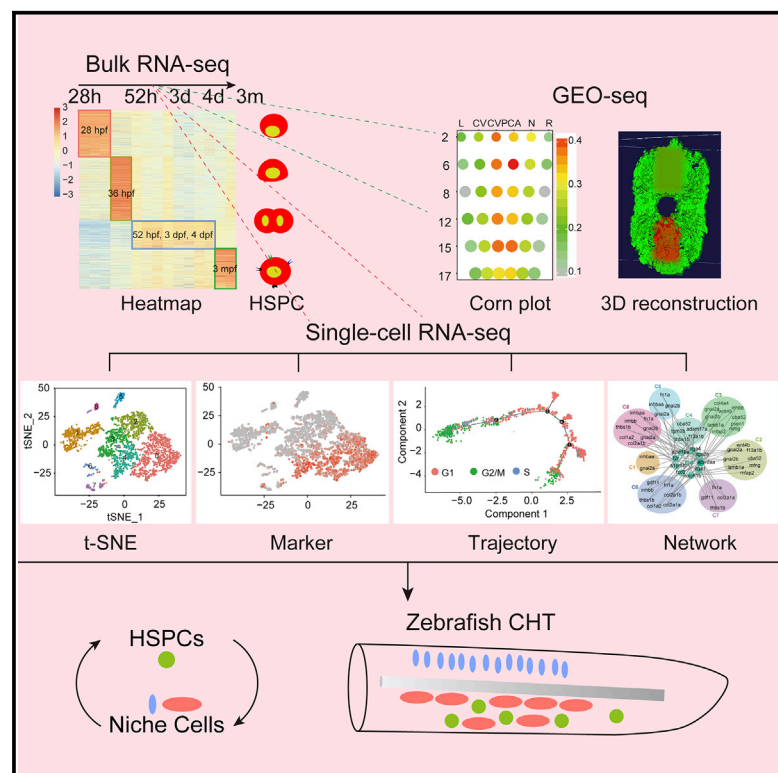


Cell Reports

A 3D Atlas of Hematopoietic Stem and Progenitor Cell Expansion by Multi-dimensional RNA-Seq Analysis

Graphical Abstract



Authors

Yuanyuan Xue, Denghui Liu, Guizhong Cui, ..., Naihe Jing, Jing-Dong J. Han, Feng Liu

Correspondence

jdhan@picb.ac.cn (J.-D.J.H.), liuf@ioz.ac.cn (F.L.)

In Brief

In this work, Xue et al. establish a 3D atlas of zebrafish caudal hematopoietic tissue and reveal mechanisms underlying HSPC expansion in an intact hematopoietic organ by multi-dimensional RNA-seq.

Highlights

- A spatiotemporal transcriptome analysis of HSPCs and their niche in development
- Generation of a 3D transcriptional atlas of a zebrafish hematopoietic organ
- Characterization of regulatory network among hematopoietic niches and progenitors
- Analysis of cell cycle-associated differentiation of progenitor cells



A 3D Atlas of Hematopoietic Stem and Progenitor Cell Expansion by Multi-dimensional RNA-Seq Analysis

Yuanyuan Xue,^{1,2,5,6} Denghui Liu,^{3,5,6} Guizhong Cui,^{4,5} Yanyan Ding,^{1,2,5} Daosheng Ai,³ Suwei Gao,¹ Yifan Zhang,^{1,2,5} Shengbao Suo,^{3,5} Xiaohan Wang,^{1,2,5} Peng Lv,^{1,2,5} Chunyu Zhou,^{3,5} Yizhou Li,³ Xingwei Chen,^{3,5} Guangdun Peng,⁴ Naihe Jing,⁴ Jing-Dong J. Han,^{3,5,*} and Feng Liu^{1,2,5,7,*}

¹State Key Laboratory of Membrane Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China

²Institute for Stem Cell and Regeneration, Chinese Academy of Sciences, 100101 Beijing, China

³Key Laboratory of Computational Biology, CAS Center for Excellence in Molecular Cell Science, Collaborative Innovation Center for Genetics and Developmental Biology, Chinese Academy of Sciences-Max Planck Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

⁴State Key Laboratory of Cell Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

⁵University of Chinese Academy of Sciences, Beijing 100049, China

⁶These authors contributed equally

⁷Lead Contact

*Correspondence: jdhan@picb.ac.cn (J.-D.J.H.), liuf@ioz.ac.cn (F.L.)
<https://doi.org/10.1016/j.celrep.2019.04.030>

SUMMARY

In vertebrates, hematopoiesis occurring in different niches is orchestrated by intrinsic and extrinsic regulators. Previous studies have revealed numerous linear and planar regulatory mechanisms. However, a multi-dimensional transcriptomic atlas of any given hematopoietic organ has not yet been established. Here, we use multiple RNA sequencing (RNA-seq) approaches, including cell type-specific, temporal bulk RNA-seq, *in vivo* GEO-seq, and single-cell RNA-seq (scRNA-seq), to characterize the detailed spatiotemporal transcriptome during hematopoietic stem and progenitor cell (HSPC) expansion in the caudal hematopoietic tissue (CHT) of zebrafish. Combinatorial expression profiling reveals that, in the CHT niche, HSPCs and their neighboring supporting cells are co-regulated by shared signaling pathways and intrinsic factors, such as integrin signaling and *Smchd1*. Moreover, scRNA-seq analysis unveils the strong association between cell cycle status and HSPC differentiation. Taken together, we report a global transcriptome landscape that provides valuable insights and a rich resource to understand HSPC expansion in an intact vertebrate hematopoietic organ.

INTRODUCTION

Hematopoietic stem and progenitor cells (HSPCs) undergo dynamic and continuous development to meet the demands for lifelong hematopoiesis (Laurenti and Götting, 2018; Zhang et al., 2018). These cells, at various developmental stages, selec-

tively reside in the different hematopoietic niches, which could modulate the state and behavior of HSPCs (Sánchez et al., 1996; Crane et al., 2017). The earliest HSPCs are derived from hemogenic endothelium (HE) in the aorta-gonad mesonephros (AGM) region through the endothelial-to-hematopoietic transition (EHT) (Bertrand et al., 2010; Kissa and Herbomel, 2010; Boisset et al., 2010). Subsequently, nascent HSPCs migrate to the fetal liver (in mammals) or the caudal hematopoietic tissue (CHT) (in zebrafish) for rapid expansion and differentiation and, finally, seed into the bone marrow (BM) (in mammals) or the kidney (in zebrafish) to maintain adult hematopoiesis. Previous studies have mainly focused on the regulatory mechanisms underlying HSPC emergence in AGM and HSPC maintenance in BM. Nonetheless, the molecular regulation of HSPC expansion and differentiation in the mammalian fetal liver or the zebrafish CHT remains largely unclear.

Zebrafish CHT, the counterpart of the mammalian fetal liver, has been identified as a crucial hematopoietic niche that supports nascent HSPC maturation, expansion, and differentiation (Xue et al., 2017; Tamplin et al., 2015; Murayama et al., 2006). HSPCs in CHT undergo dramatic expansion while maintaining their robust functional properties. Therefore, CHT serves as an ideal paradigm for induction and expansion of countless and functional HSPCs *in vitro*. Previous studies suggest that a few epigenetic and transcription factors (TFs) regulate HSPC expansion in a cell-autonomous manner and that niche cells modulate HSPCs in the fetal liver and CHT through secreted growth factors and cytokines and physical interaction with HSPCs (Khan et al., 2016; Gao et al., 2018; Xue et al., 2017; Gao and Liu, 2018). However, a systematic understanding of why and how CHT or fetal liver can support the rapid expansion and differentiation of HSPCs is still lacking.

Transcriptomics based on the bulk of cells has long been employed to investigate the regulatory mechanisms underlying HSPC development (Zhang et al., 2015). Unfortunately, that strategy fails to reveal the heterogeneity between individual



cells. The recently developed, single-cell transcriptomics has greatly improved the study resolution (Tang et al., 2009; Klein et al., 2015) and can provide more detailed information, which is particularly useful in identifying new cell types (Davie et al., 2018), reconstructing developmental trajectories (Farrell et al., 2018), and revealing interactional networks (Zepp et al., 2017). However, all the aforementioned transcriptome analyses scarcely provide the cellular spatial information, which is essential to decipher how a specific spatial pattern of tissue or organ can regulate the behaviors of resident cells. Currently, the combination of region-specific sample collection and low-input, or even single-cell, transcriptomics makes it feasible to uncover cellular transcriptome state and spatial information simultaneously (Fan et al., 2018; Peng et al., 2016; Junker et al., 2014). With the new methodological advances, it is possible to achieve a complete spatiotemporal map of an intact hematopoietic organ and, therefore, comprehensively dissect HSPC biology.

In this study, we performed a multi-dimensional transcriptome analysis to provide deep insights into zebrafish CHT, a hematopoietic organ in which HSPCs expand and differentiate. Employing transgenic embryos, we first characterized the temporal features of HSPCs at various developmental stages by bulk RNA sequencing (RNA-seq). Then, we performed GEO sequencing (GEO-seq) and single-cell RNA-seq (scRNA-seq) of zebrafish CHT to identify the cellular composition as well as the characteristics of CHT niche and resident HSPCs. We demonstrate that the sophisticated crosstalk between niches and HSPCs is essential for the formation of this hematopoietic organ. Finally, reconstruction of high-resolution transcriptome of CHT organ provides a wealth of resources for HSPC development at the molecular level. The strategy of multi-dimensional transcriptome analysis described in this study also provides a universal pipeline for studies of other tissues or organs and even at a whole-organism level.

RESULTS

Temporal Transcriptome Features of HSPCs at Different Stages

During development, HSPCs at different stages exhibit temporal-specific features of cell state and transcriptome. To characterize the distinct features of HSPCs and the corresponding niche cells, we performed bulk RNA-seq with sorted non-endothelial and non-hematopoietic cells (NCs), endothelial cells (ECs), HE cells, and HSPCs from the zebrafish transgenic lines (*kdr1:mCherry/runx1:en-GFP*) and (*kdr1:mCherry/CD41:GFP*) in distinct regions at six relatively discrete stages, which included HSPC emergence in AGM (28 h postfertilization [hpf]), migration (36 hpf), expansion in CHT (52 hpf; 3 days postfertilization [dpf] and 4 dpf), and maintenance in kidney (3 months postfertilization [mpf]) (Figure 1A). Through principal-component analysis (PCA) after quality control, the results showed that NCs differed from all other cells, possibly because of the mixture of different types of cells, but ECs and HSPCs at the same stage tended to cluster together, indicating their close relationship during development (Figures 1B and S1A). To identify the cell type-specific genes, we performed Bayesian information criterion super *k*-means

(BICSKmeans) clustering (Zhang et al., 2013), based on differentially expressed genes (DEGs; false discovery rate [FDR] < 0.001) between any two stages or cell types. We found that representative marker genes were enriched in the corresponding clustering groups (Figure S1B), indicating that our collected samples can well represent the characteristics of different cell types. To reveal the dynamic change of HSPCs across different stages, we performed PCA and showed a continuous developmental process from HSPC emergence (28 hpf) to maintenance (3 mpf) (Figure 1C). We also used Stouffer's method (Petropoulos et al., 2016) to identify the stage-specifically expressed genes (FDR < 0.001) and explored the functional enrichment of different gene groups (Figure 1D; Table S1). Given the high similarity of expanding HSPCs at different stages (52 hpf and 3 and 4 dpf), we combined those cells together for the subsequent analysis. Gene Ontology (GO) analysis showed that different groups were enriched in functions that correspond to features of HSPCs at different stages (Figure 1D). For example, genes specifically expressed in HSPCs at 36 hpf were mostly enriched in the GO terms "cell adhesion" and "cell migration," which is consistent with the cell behavior at that stage when emerging HSPCs initially bud off and then migrate into CHT from AGM. Genes specifically expressed from 52-hpf to 4-dpf HSPCs were enriched for the terms "rRNA processing," "mRNA splicing," and "ribosome biogenesis," indicating that HSPCs undergo rapid expansion during those stages. In addition, genes specifically expressed in HSPCs at 3 mpf were mainly involved in the terms "immune response" and "antigen processing," suggesting the distinct immunoreaction between embryonic and adult HSPCs. Taken together, our data systematically reveal the dynamic features of HSPCs during development.

Spatial Transcriptome Analysis for Zebrafish Caudal Tissue

Although the states of HSPCs are tightly regulated by a cell-intrinsic transcriptional mechanism, the resident microenvironment is also required to maintain HSPC function. We and others have previously demonstrated that, during the formation of the caudal vein plexus (CVP) in zebrafish, the newly arriving HSPCs remodel the neighboring ECs and stromal cells to form EC "pockets," in which the secreted chemokines (such as Ccl25b) facilitate HSPC lodgment and expansion (Xue et al., 2017; Tamplin et al., 2015). To decode the complex components of CHT, we performed live imaging using the transgenic line (*kdr1:mCherry/cmyb:GFP*), and snapshot live-imaging analysis displayed the mCherry-labeled vasculature structure and GFP-labeled HSPC localization in CHT (Figures S2A and S2B). To further analyze the mechanisms underlying HSPC expansion from the viewpoint of systems biology, we used spatial transcriptome analysis, GEO-seq (Peng et al., 2016), to acquire spatial transcriptome information and generate a three-dimension (3D) atlas of zebrafish CHT as illustrated (Figure 2A). We focused on the early stage (52 hpf) of initiation of HSPC expansion when CHT microenvironment was orchestrated by arriving HSPCs. The tail region of 55-hpf zebrafish embryo was cryo-sectioned, and six sets of samples, including neuro (N), left muscle (L), right muscle (R), caudal artery (CA), caudal vein (CV), and CVP, were obtained from each section by laser capture microdissection

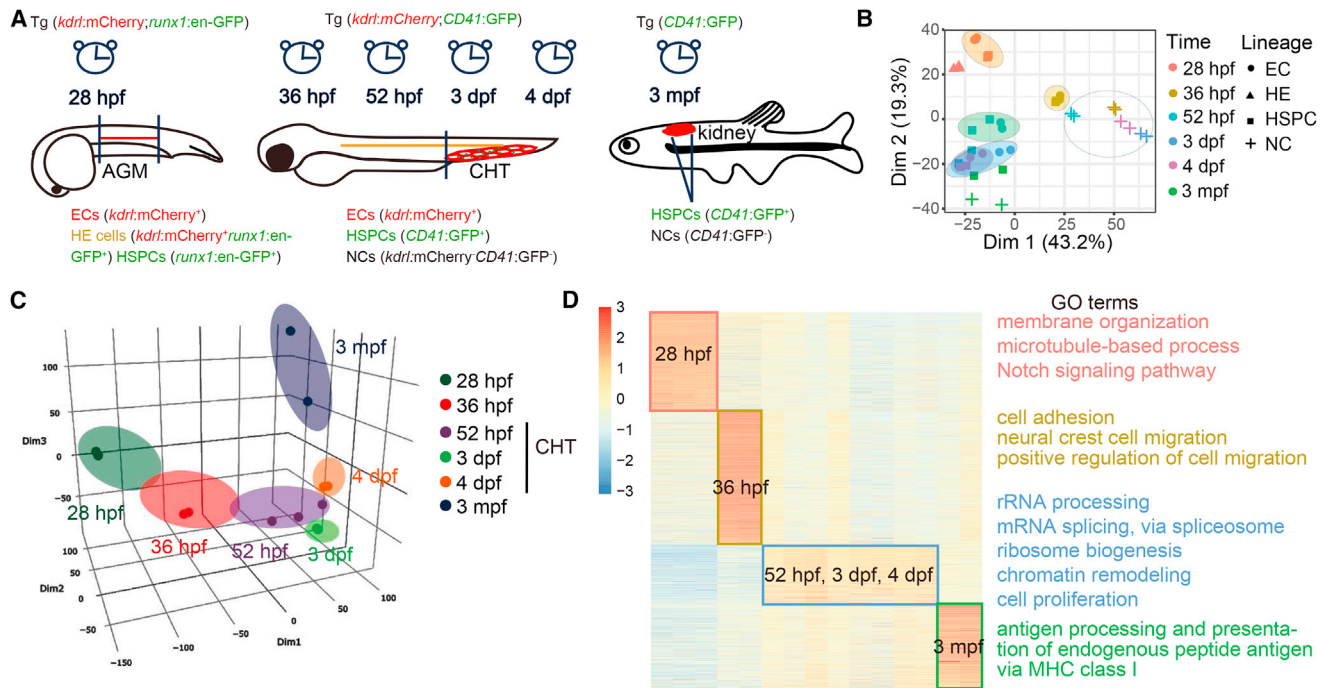


Figure 1. Temporal Transcriptome Features of Discrete HSPCs

(A) Outline of bulk RNA-seq analysis: the non-endothelial and non-hematopoietic cells (NCs), endothelial cells (ECs), hemogenic endothelium (HE), and hematopoietic stem and progenitor cells (HSPCs) from aorta-gonad mesonephros (AGM), caudal hematopoietic tissue (CHT), and kidney at distinct stages were obtained by fluorescence activated cell sorting (FACS).

(B) Principal component analysis (PCA) of all bulk RNA-seq samples. hpf, hours post fertilization; dpf, days post fertilization; mpf, months post fertilization.

(C) Three-dimensional (3D) PCA of HSPC populations at distinct stages shows the progressive development of HSPCs. Each cell population at the corresponding stage includes over two replicates.

(D) Heat map showing the expression of HSPC stage-specific genes (FDR < 0.001) and gene ontology (GO) analysis showing the functional enrichment of HSPC stage-specific genes.

See also Figure S1 and Table S1.

(LCM) (Figure S2C). After quality control, 36 samples were retained for further analysis, and six cryo-sections for LCM are shown (Figure S2D). PCA revealed that samples derived from the same region were generally grouped together (Figure 2B). We used two embryos for comparison, and they showed a high correlation (Figure S2E), indicating a high reproducibility of our GEO-seq data. PCA displayed the separation of CA, CV, and CVP samples (Figure S2F). Hierarchical clustering analysis of the DEGs between L and R showed that there was no significant difference of expression profiling between the left-muscle and right-muscle samples (Figure S2G). BICSKmeans clustering analysis of DEGs (FDR < 0.001) between samples in different regions revealed the presence of seven gene clusters with region-specific expression, which likely correspond to the neuro (G7), muscle (G2 and G5), and CA/CVP/CV (G1, G3, and G4) (Figure S3A). Genes associated with neural crest cell migration, myeloid cell development, and hematopoiesis were enriched in both neuro and CVP. These data suggest that the neural and CVP tissues may share a subset of common genes or that there is a potential regulatory relationship between these two tissues. Genes highly expressed in L/R samples were involved in skeletal muscle tissue development and erythrocyte differentiation. A high degree of concordance was found between the whole

mount *in situ* hybridization (WISH) pattern and a corn plot (which displays the spatial gene expression pattern) (Figure S3B), indicating the high quality of our sample collection.

Furthermore, we performed a self-organizing map (SOM) of GEO-seq data to visualize the spatial distinctions of gene modules. The anterior and posterior samples showed a similar gene profile in L, R, and N groups, whereas diverse patterns among samples were observed in CV, CVP, and CA groups (Figure 2C), suggesting that distinct positional CA, CVP, and CV may possess different features. Based on gene set enrichment analysis (GSEA), we identified the activated signaling pathways in each sample (see Methods). The corn plots showed the distribution of various signaling pathways in each region (Figure 2D; Table S2). Transforming growth factor β (TGF- β), Notch, and Wnt signaling pathways were highly enriched in neural tissues. Fibroblast growth factor receptor (FGFR) signaling was enriched in L and R muscle samples, whereas Janus kinase/signal transducers and activators of transcription (JAK/STAT), integrin, and inflammatory signaling pathways were enriched in CA, CVP, and CV regions. The gene expression heat map displayed the region-restricted patterns of signaling pathway-associated genes (Figure S3C). Altogether, the regional distribution of signaling pathways in zebrafish caudal tissues can be summarized as

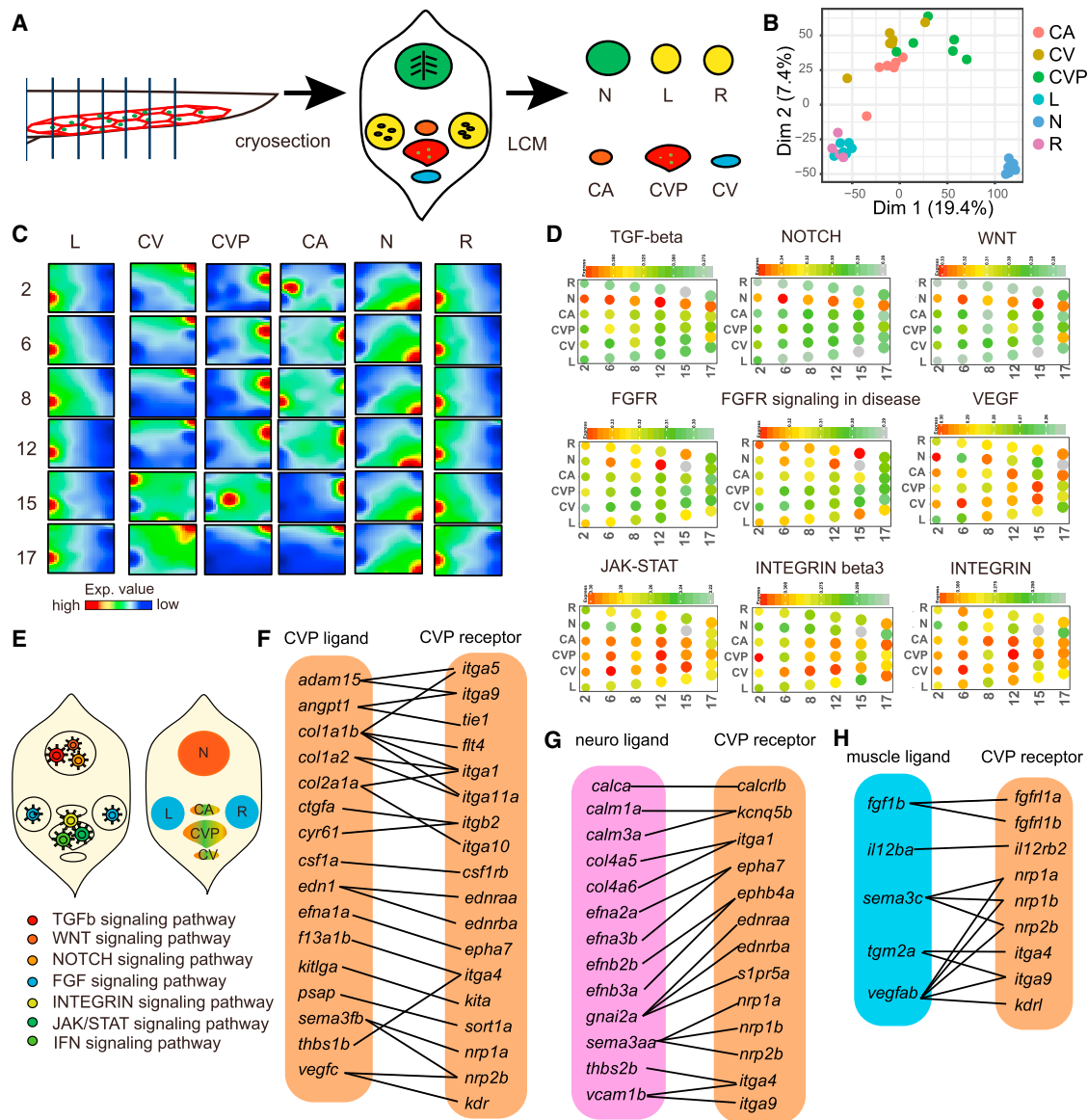


Figure 2. Spatial Transcriptome for Zebrafish Caudal Tissue

(A) Workflow of GEO-seq analysis: the CHT region at 55 hpf was embedded for cryo-section, and subsequently six regions, including neuro (N), left muscle (L), right muscle (R), caudal artery (CA), caudal vein (CV), and CVP were collected from each section by laser capture microdissection (LCM).

(B) PCA of different regions from six sections.

(C) Self-organizing map (SOM) analysis for all 36 GEO-seq samples. The numbers on the left represent the ordinal codes of cryo-sections.

(D) Corn plots of signaling pathways showing the specific distribution of signals in the CHT region.

(E) The sketch model showing the regional-specific distribution of signaling pathway in CHT.

(F) The ligand-receptor pairs of internal caudal vein plexus (CVP).

(G) The ligand-receptor pairs between CVP and neural cells.

(H) The ligand-receptor pairs between CVP and muscle cells.

See also [Figures S2](#) and [S3](#) and [Tables S2](#) and [S3](#).

shown in a sketch model (Figure 2E). To investigate how CHT tissue-to-tissue crosstalk occurs, we analyzed the spatial interactions of signaling pathways and *in silico* secretomics (based on genes encoding annotated secretory proteins and receptors) between different GEO-seq spatial domains (see Method Details). Secretomics analysis uncovered a ligand-receptor regulatory

model in the zebrafish caudal region, including both proximal and distal regulation (Figure S3D). Integrin, Notch and fibroblast growth factor (FGF) signaling pathways mediated CVP-to-CVP, neuro-to-CVP, and muscle-to-CVP crosstalk, respectively (Figures 2F–2H; Table S3). These interactions suggest that the diverse niche cells and sophisticated gene regulatory networks

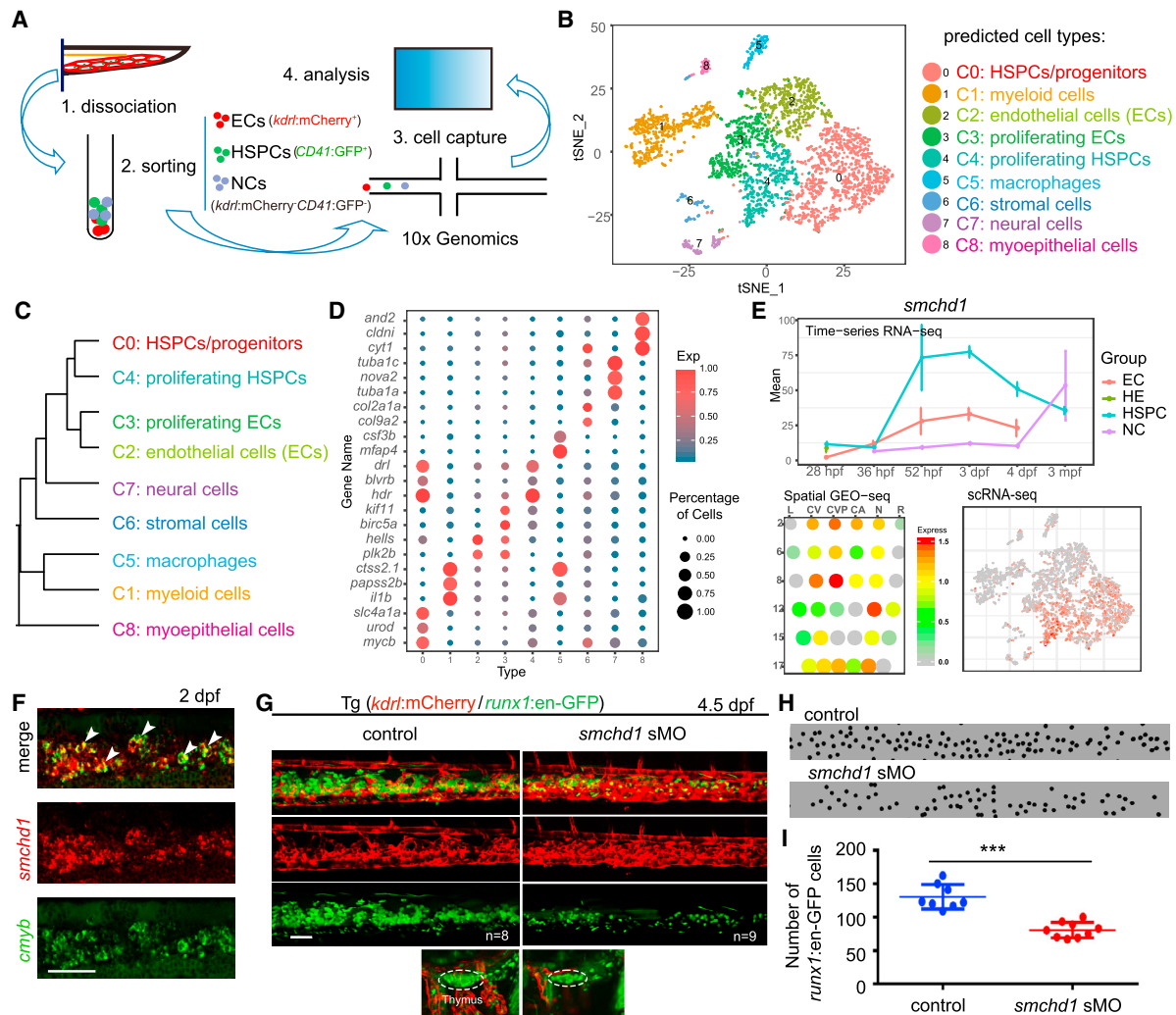


Figure 3. A Single-Cell Resolution Map of Zebrafish CHT

(A) The illustration of CHT single-cell RNA-seq (scRNA-seq) at 55 hpf.

(B) t-Distributed Stochastic Neighbor Embedding (t-SNE) clustering map of 3,776 single cells from zebrafish CHT.

(C) The hierarchy structure of 9 cell clusters. Single cells of the same group were combined together for hierarchical clustering.

(D) Dot plot showing representative marker genes of the defined cell groups. Color indicates the ratio between the expression level of each group and the maximum expression level of all groups. Size of the dot is proportional to the percentage of cells expressing this gene (>1) in each group.

(E) The expression pattern of *smchd1* generated from bulk RNA-seq, GEO-seq, and scRNA-seq, respectively.

(F) Double fluorescence *in situ* hybridization (FISH) analysis showing co-expression of *smchd1* and *cmyb* in the CHT region at 3 dpf. White arrowheads mark the *smchd1*⁺*cmyb*⁺ double positive cells. Scale bar, 50 μ m.

(G) Confocal imaging of the transgenic line (*kdr*:mCherry/*runx1*:en-GFP) showing the number of GFP cells (*runx1*⁺) in the CHT and thymus regions in control and *smchd1* morphants at 4.5 dpf. Dashed line circles mark the thymus region. $n_{\text{control}} = 8$; $n_{\text{smchd1 sMO}} = 9$. Scale bar, 50 μ m.

(H) The black spots mimicked the *runx1*⁺ cells in CHT region in (G).

(I) The quantification data of *runx1*⁺ cells in CHT in control and *smchd1* morphants, showing as means \pm SEM. ***p < 0.001. The HSPCs were counted in about 6-somite regions of CHT in control embryos and morphants.

See also Figures S4 and S5 and Table S4.

are involved in rendering zebrafish CHT to facilitate HSPC expansion.

A Single-Cell Resolution Map of Zebrafish CHT

The development of scRNA-seq approaches makes it possible to better understand the fundamental biological processes at an unprecedented resolution. To investigate the characteristics

of CHT HSPCs, as well as the interaction between CHT niche and HSPCs at a higher resolution, we performed scRNA-seq using 10 \times genomics for sorted ECs, HSPCs, and other neighboring cells from zebrafish CHT (Figure 3A). After quality control, 3,776 single cells were retained for further analysis, and nine major clusters were identified after dimension reduction by Seurat (Macosko et al., 2015) (Figure 3B). Based on their specific signature

genes (Figure S4A; Table S4), these clusters were assigned as HSPCs and progenitors (C0), myeloid cells (C1), ECs (C2), proliferating ECs (C3), proliferating HSPCs (C4), macrophages (C5), stromal cells (C6), neural cells (C7), and myoepithelial cells (C8). The cluster of proliferating HSPCs was segregated with the cluster of HSPCs and progenitors by a set of cell cycle- and cell division-related genes (Table S4). Expression similarity or distance between HSPCs and various niche cells was further revealed using hierarchical clustering analysis (Figure 3C). Consistently, dot plots of top marker genes confirmed the assignment of individual clusters (Figure 3D).

To confirm the findings from scRNA-seq data, we next analyzed the data from bulk RNA-seq and scRNA-seq together. Consequently, we identified 40 HSPC signature genes, including *atf7ip* and *smchd1* (Figure S4B; Table S4). The expression of candidate HSPC signature genes identified from the combinatorial screening was further validated using WISH analysis (Figure S4C). Given that *smchd1* encodes structural maintenance of chromosomes flexible hinge-domain containing 1, which is an epigenetic regulator, and may modulate a set of hematopoietic or cell cycle-related genes (Leong et al., 2013; Wang et al., 2018), we focused on *smchd1* in the following analysis. Combinatorial profiling using the three-sequencing methods (Figure 3E) showed that *smchd1* is preferentially expressed in HSPCs, suggesting its potential function during hematopoiesis. Furthermore, fluorescence *in situ* hybridization (FISH) and WISH analyses showed that *smchd1* was specifically expressed in HSPCs (Figures 3F and S4C). To examine the function of *smchd1*, we performed morpholino (MO) knockdown with the efficacy of *smchd1* splice MO (sMO) validated by RT-PCR (Figures S5A and S5B). Confocal imaging of transgenic embryos indicated that, compared with the control samples, *smchd1* sMO-injected embryos had fewer *runx1*⁺ GFP cells and *CD41*⁺ GFP cells in the CHT and thymus regions at 4.5 dpf (Figures 3G–3I and S5C). WISH analysis indicated that definitive hematopoiesis was impaired in *smchd1* sMO-injected embryos, whereas primitive hematopoiesis and thymic endothelium were not affected (Figures S5D–S5F). To confirm that hematopoietic phenotype, the second MO (*smchd1* ATG MO [aMO]) was used. Normal HSPC emergence in AGM at 28 hpf and impaired hematopoiesis in CHT and thymus at 4 dpf were also validated in *smchd1* aMO-injected embryos (Figures S5G and S5H). Additionally, to assess HSPC proliferation, we performed bromodeoxyuridine (BrdU) assay and FISH analysis and found that the number of *draculin*⁺BrdU⁺ double-positive cells was reduced in *smchd1* morphants at 2.5 dpf, in comparison with control (Figures S5I and S5J). These data demonstrate that *smchd1* is required for the expansion of HSPCs.

The Cell Cycle Status of CHT HSPCs

To systematically investigate the underlying mechanism of HSPC expansion, we first explored the expression pattern of cell division genes across all nine clusters identified from scRNA-seq data, which showed that these genes were preferentially expressed in C3 (proliferating ECs) and C4 (proliferating HSPCs), indicating that partial ECs and HSPCs in CHT synchronously underwent rapid expansion. Because C0 and C4 are highly correlated with HSPCs and to further analyze HSPC devel-

opment in CHT, we conducted sub-clustering of C0 and C4 into 5 sub-groups (Figure 4A). Based on top marker genes of each sub-cluster (subC), we reasoned that subC0s were lymphoid progenitors with high expression of *bcl11ba*; subC1s were erythroid and megakaryocyte progenitors enriched with *mycb* and *tcima*; subC2s were myeloid progenitors with expression of *coro1a*, *lcp*, and *cebpb*; subC3s were nascent HSPCs and progenitors that still retained the expression of some endothelial genes; and subC4s were erythrocytes with expression of a set of hemoglobin genes (Figure 4B; Table S5). We revealed that HSPCs in the CHT region undergo a lineage-restricted process, a phenomenon that can be observed in adult HSPCs of zebrafish kidney (Athanasiadis et al., 2017). In addition, to reconstruct the lineage differentiation path of HSPCs, we performed Monocle analysis to clarify the trajectory of HSPC differentiation in CHT. Based on our prior knowledge, we set the HSPCs and progenitors as the root state and the following progression of the differentiation trajectory, myeloid progenitors, lymphoid progenitors, erythroid and megakaryocyte progenitors, and erythrocytes were derived (Figure 4C). The expression of lineage-specific marker genes in the trajectory map confirmed the differentiation process from nascent HSPC to lineage-restricted progenitors in CHT (Figure 4D). Furthermore, the cell cycle status was mapped into the HSPC trajectory (Figure 4E), suggesting that most HSPCs and progenitors, erythroid and megakaryocyte progenitors, and erythrocytes are in G2-M phase, whereas other cell types, lymphoid progenitors, and myeloid progenitors are mainly in G1 phase (Figure 4E). Collectively, these data indicate that cell cycle features are associated with HSPC differentiation.

Tissue and Molecular Crosstalk Underlie the Organization of CHT

An efficient organ formation requires a complex of tissue-tissue crosstalk and cell-cell interaction. To address how the CHT niche facilitates HSPC expansion, we analyzed comprehensive transcriptomic data to identify a complex of ligand-receptor pairs between HSPCs and niche cells using an SOM, a CSI (connection specificity index), and an *in silico* secretomics analysis. First, to unravel extrinsic factors involved in dynamic development process of HSPC expansion and differentiation, we explored the global profile of ECs and HSPCs from 2 to 4 dpf by SOM analysis of bulk RNA-seq (Figure 5A). We found that from 52 hpf to 4 dpf, functional enrichments of stage-specific gene sets were changed from “regulation of transcription” and “cell cycle” to “calcium and ion transport” and “cell adhesion.” A coordinated change of transcriptome indicates that HSPC expansion has a high association with ECs. Furthermore, the co-expressed genes were enriched in the GO terms “JAK/STAT signaling” and “inflammatory response,” which may mediate interaction between HSPCs and EC niche. Second, we performed secretomics analysis of bulk RNA-seq and found potential pathways regulating HSPC expansion in the CHT niche, such as integrin, chemokine, Notch, and Wnt pathways (Figure 5B). Third, secretomics analysis based on scRNA-seq also revealed a series of signaling pathways that mediate niche-HSPC interaction, such as Notch, BMP, and integrin signaling pathways (Figure 5C). Fourth, GEO-seq data also revealed the enrichment of integrin

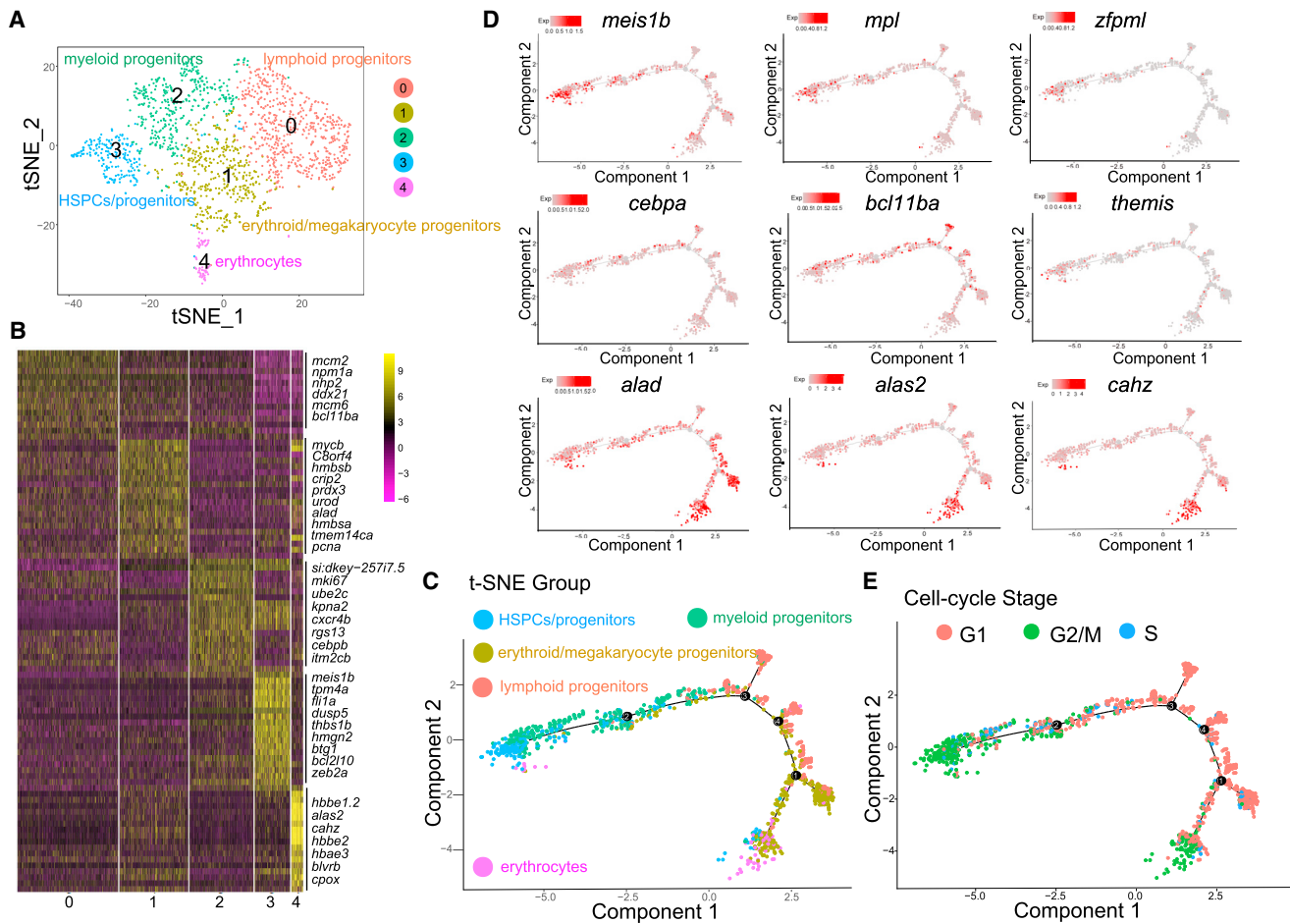


Figure 4. The Cell Cycle Status of CHT HSPCs

(A) t-SNE map further clusters the original C0 (HSPCs and progenitors) and C4 (proliferating HSPCs) groups into five subgroups. (B) Heat map showing the expression of genes in each sub-cluster defined in (A). Based on these marker genes in each group, we assigned the five groups into lymphoid progenitors, erythroid and megakaryocyte progenitors, myeloid progenitors, HSPCs and progenitors, and erythrocytes. (C) Trajectory analysis by Monocle 2 reconstructed the lineage separation path from HSPCs to multiple progenitors. (D) The expression of HSPC and lineage marker genes along with the trajectory. (E) The cell cycle status along with the trajectory of speculated differentiation of HSPCs in CHT. See also Table S5.

and interferon (IFN) signaling pathways in the CVP region (Figures 2D and 2E).

The comparison of ligands and receptors detected using the three-sequencing methods is shown in Venn diagram (Figures S6A and S6B). The combinatorial transcriptome analysis, especially between bulk RNA-seq and GEO-seq, classified integrin signaling as one of the most common signaling pathways during HSPC expansion in CHT. Bulk RNA-seq data analysis showed the distribution of integrin signaling ligands and receptors in three cell types, including ECs, HSPCs, and NCs (Figure 5D). *ctgfa* and *itgb2* were detected in the CHT region by WISH (Figure S6C). Furthermore, FISH analysis indicated that *itgb2* and *cmyb* were partially co-localized in CHT, whereas *ctgfa* and *fli1a* were co-expressed in CHT (Figure 5E). To investigate the function of the *ctgfa-itgb2* pathway, we also used the validated antisense MOs (Figure S6D) to knock down the endogenous

expression of those two genes. WISH data showed that HSPC generation in the AGM region was normal in *ctgfa* and *itgb2* morphants (Figure S6E). Interestingly, we observed an increase in the numbers of *runx1*⁺ and *CD41*⁺ cells (HSPCs), *gata1*⁺ cells (erythrocytes), and *coro1a*⁺, *mfap4*⁺ cells (myeloid cells) in *ctgfa* and *itgb2* morphants, and qPCR results confirmed the upregulation of the expression of hematopoietic genes (*cmyb*, *pu.1*, and *gata1*) in *ctgfa* and *itgb2* morphants (Figures 5F, 5G, and S6F–S6I), indicating that the *ctgfa-itgb2* pathway has a negative role during HSPC expansion in CHT. Further, we employed Tg (*hsp70l:ccn2a-2A-EGFP*) (Mokalled et al., 2016) and heat shocked the embryos at 36 hpf and 2.5 dpf. The expression of *ctgfa* was enhanced in GFP⁺ embryos with overexpressing *ctgfa*, compared with GFP⁺ embryos at 3.5 dpf (Figure S6J). WISH results displayed that the expression of *cmyb* was increased in *ctgfa* morphants and decreased in *ctgfa* overexpressed

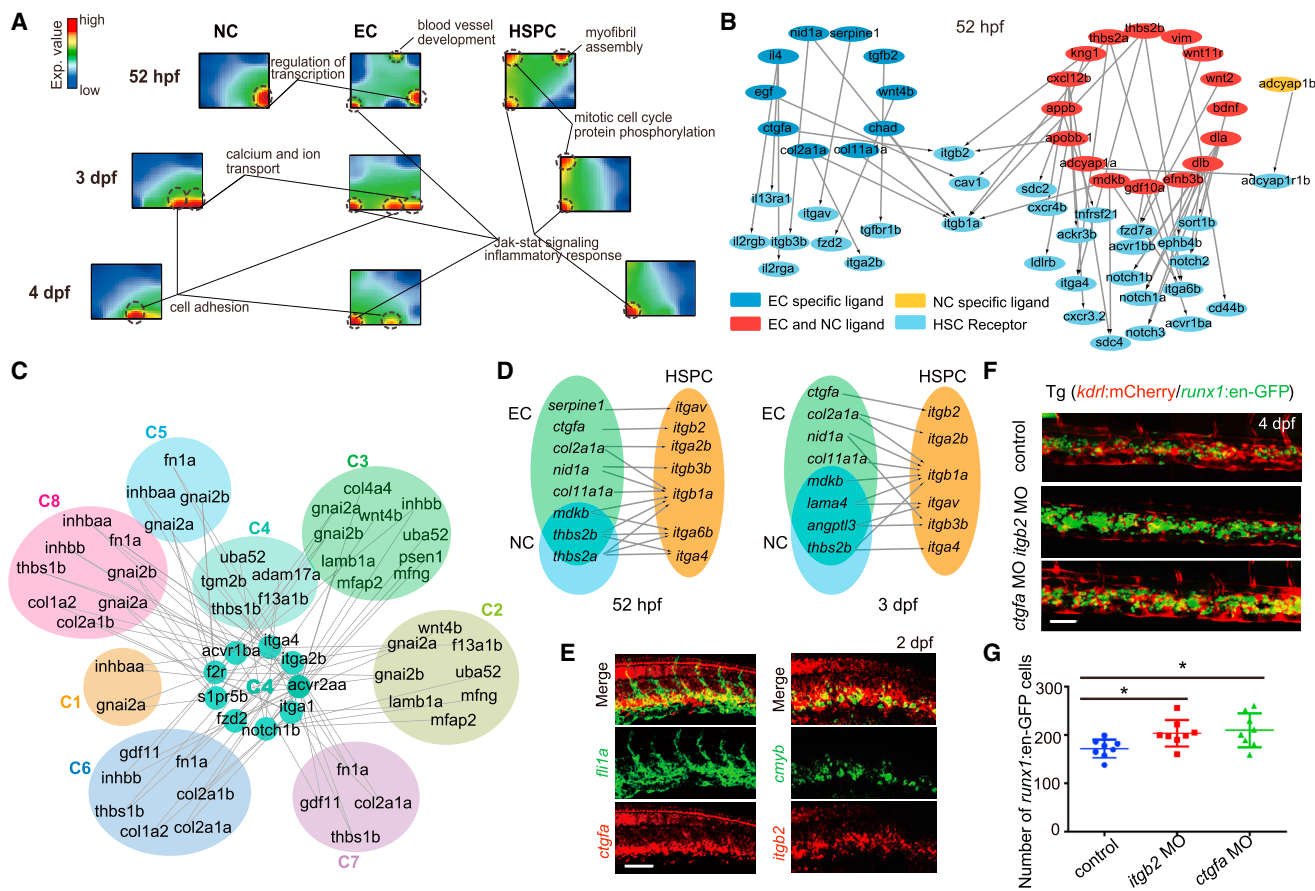


Figure 5. The Crosstalk between CHT Niche and HSPCs

(A) SOM showing the distinct gene regulatory structure among NCs, ECs, and HSPCs at 52 hpf and 3 and 4 dpf.
 (B) Secretomics analysis showing the ligand-receptor regulatory relationship between HSPCs and CHT niche at 52 hpf.
 (C) The secretomics analysis for scRNA-seq data revealing the interaction pairs in distinct cell types, receptor-expressing C4 (proliferating HSPC) and ligand-expressing clusters.
 (D) The distribution of ligands and receptors in ECs (*kdr1*⁺ cells), HSPCs (*CD41*⁺ cells), and NCs (*kdr1*⁻*CD41*⁻ cells) at 52 hpf and 3 dpf based on bulk RNA-seq analysis.
 (E) Double FISH showing the co-expression of *ctgfa* with *fli1a*, and *itgb2* with *cmyb* at 2 dpf. Scale bar, 50 μ m.
 (F) The confocal imaging of the transgenic line (*kdr1:mCherry/runx1:en-GFP*) showing the number of *runx1*⁺ cells in control, *ctgfa* morphants, and *itgb2* morphants at 4 dpf. Scale bar, 50 μ m.
 (G) The quantification of confocal imaging displayed in (F) are shown as means \pm SEM. **p* < 0.05, *n* = 8. The HSPCs were counted in about the eight-somite region of the CHT in control embryos and morphants.
 See also Figure S6.

embryos at 4 dpf, whereas the dysregulation of *cmyb* expression level in *ctgfa* morphants was rescued by the overexpression of *ctgfa* (Figure S6K). To further confirm that data, we generated an *itgb2* mutant using CRISPR/Cas9 (Figures S6L and S6M). Consistently, similar HSPC phenotypes were observed in *itgb2*-loss embryos (Figure S6N). Taken together, our data suggest that both intrinsic factors and niche-HSPC interplay can contribute to rapid expansion of HSPCs in zebrafish CHT.

Transcriptome Reconstruction of Zebrafish Hematopoietic Organ

Molecular reconstruction of CHT will provide clues for understanding the underlying mechanisms of HSPC expansion and

differentiation *in vitro*. The global regulatory network during HSPC development in CHT region is shown in Figure 6A. Given that the original spatial information is lost in scRNA-seq, we used the zip-code mapping method (Peng et al., 2016) to map the scRNA-seq samples to the corresponding regions of GEO-seq samples (Figure 6B). The results showed that single-cell samples of neural cells (C7) were mainly mapped to the neuro section (N), whereas ECs (C2 and C3), HSPCs (C0 and C4), and hematopoietic cells (C1 and C5) were mapped to the CA section and the CVP. From the mapping result, CA and CVP regions also contain abundant stromal cells (C6), whereas CV section (CV) includes many myoepithelial cells (C8). The mapping results implied a high coordination of spatial

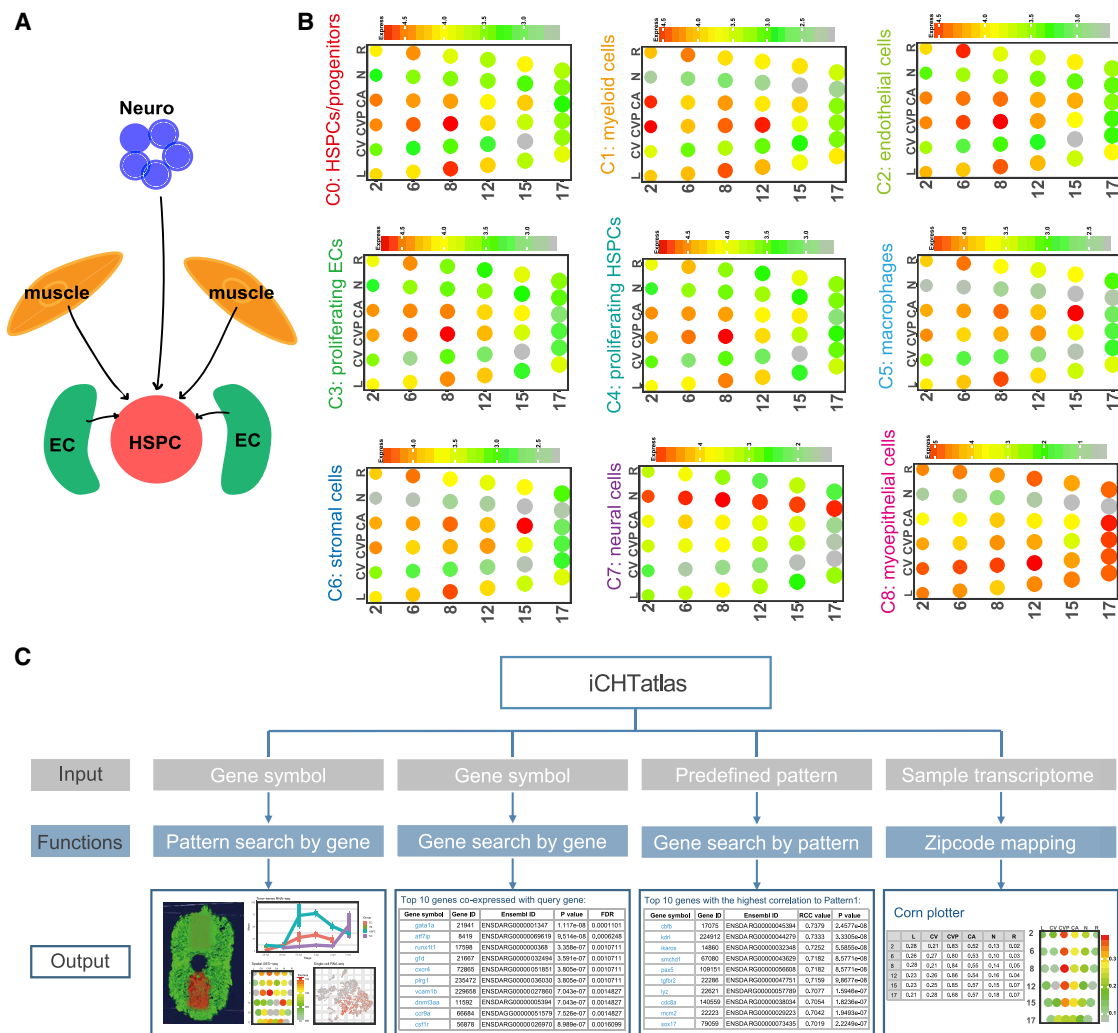


Figure 6. The Transcriptome Reconstruction of Zebrafish CHT

(A) The illustration showing the regulatory network among HSPCs and niche cells in CHT.

(B) The corn plot showing the probability of each single cell cluster in each GEO region by mapping single-cell nine clusters to GEO-seq data.

(C) Web server for query of the expression pattern of each gene (including pattern search by gene, gene search by gene, gene search by pattern, and zip code mapping).

distribution between ECs and hematopoietic cells. A comprehensive web server (<http://www.picb.ac.cn/hanlab/ichtatlas/Home/>) was generated to help the community to further investigate this hematopoietic organ (Figure 6C). The web server provides spatiotemporal, high-resolution transcriptomic information covering 21,588 genes that are expressed in the CHT region in zebrafish during HSPC expansion. These data are collated into a comprehensive and interactive 3D transcriptome profiles-the iCHT atlas. Our web server thus provides the following service: (1) visualizing the 3D expression pattern, the 2D corn plot with or without single-cell transcriptome mapped, and the time-series expression pattern for all of genes; (2) mapping individual query samples into the 3D position; and (3) querying a gene of interest with a predefined expression pattern (Figure 6C).

DISCUSSION

In this study, we used multi-dimensional transcriptome analysis, combined with functional validation, to dissect an intact hematopoietic organ for HSPC expansion, from a perspective of systems biology. Application of multiple sequencing methods allowed the generation of a panoramic (temporal and spatial) and high-resolution (single-cell level) transcriptome landscape for zebrafish CHT.

To systemically study a crucial organ, the top priority is to dissect the major components and understand the internal regulatory network. Previous studies on HSPCs and their niche have mainly focused on the cell-cell contact (Tamplin et al., 2015) and mutual signaling transduction (Himburg et al., 2012; Winkler et al., 2012). However, the regulatory mechanism of the global

microenvironment within a hematopoietic organ is still ambiguous. To address this concern, spatial transcriptome analyses, such as Tomo-seq (Junker et al., 2014) and GEO-seq (Peng et al., 2016), have been developed. Here, in our studies of zebrafish CHT, the secretomics analysis based on GEO-seq reveals not only proximal tissue (CA and CV) communications, such as EC-HSPC in the CVP region, which is in line with previous studies (Mahony et al., 2016; Blaser et al., 2017; Xue et al., 2017), but also distal tissue (muscle and neuro) regulations, such as neuro-HSPC interaction. Considering that these complementary regulation machineries are prevalent, any component within a niche is essential. To fully interrogate the dynamic developmental process of organ formation, time-series transcriptomics was performed, which demonstrated that the transcriptome of HSPCs and niche cells tends to show a synchronized transition from 52 hpf to 4 dpf. Similarly, a previous study has also reported that rapidly proliferating ECs adjust vascular structure to enlarge lumen volume for HSPC expansion (Khan et al., 2016). At each stage, HSPCs and niche cells possess a specific group of co-expressed genes. For instance, CHT ECs and HSPCs at 3 dpf shared a series of genes that are associated with RNA and DNA synthesis and cell cycle, suggesting that HSPCs accompanied by their niche cells assume an actively proliferating state. Furthermore, scRNA-seq was used to create high-resolution transcriptome map of an intact organ. Using this approach, we classified nine cell clusters and unveiled additional complex regulation networks among different niche cell types and HSPCs in zebrafish CHT, such as stromal cell (C6)-HSPC and myoepithelial cell (C8)-HSPC interaction. Taken together, HSPC development in CHT is modulated by various molecular mechanisms, thereby inducing a steady expansion of HSPCs in the resident niche.

We observed that HSPCs in C0 and C4 can be sub-clustered into five subpopulations, based on their transcriptome states. Importantly, we also found that these transcriptionally different subpopulations possess distinct cell cycle status, which points to a possible link between cell cycle status and lineage differentiation outputs. Consistently, previous studies indicate that the cell cycle regulation of lineage-specific gene expression is responsible for stem cell differentiation (Hao et al., 2016; Laurenti et al., 2015; Pauklin and Vallier, 2013; Sun et al., 2017). Therefore, manipulation of the cell cycle might be considered in rapid expansion and biased differentiation of HSPCs. Alternatively, the alterations in cell cycle status might be the concomitant phenomenon in which stem cells are gradually committed to lineage-restricted progenitors (Nygren et al., 2006; Rocco et al., 2013). Thus, future investigation is needed to further illustrate the exact relationship between cell cycle status and HSPC differentiation.

In summary, our data demonstrate the sophistication of niche-HSPC interactions and reconstructs a 3D transcriptome of an intact hematopoietic organ, providing a molecular framework for hematopoiesis study *in vivo*. Importantly, we elucidate that the integrity of the hematopoietic organ components is essential to establish a functional organoid for HSPC *in vitro* culture.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
 - MOs Microinjection, Overexpression and CRISPR/Cas9 Mutant
 - Whole-mount and Fluorescence *in Situ* Hybridization (WISH and FISH)
 - BrdU Assay
 - Quantitative Real-Time PCR and Plasmids Construction
 - Cell Sorting, RNA Extraction and mRNA Sequencing
 - Confocal Microscopy
 - Tissue Preparation for GEO-seq
 - Cell Preparation for single cell RNA-seq
 - Preprocessing of the Bulk and GEO RNA-seq Data
 - Bulk RNA-seq Analysis
 - GEO-seq Data Analysis
 - *In silico* Secretomics Analysis
 - Preprocessing of the Single Cell RNA-seq Data
 - Single Cell RNA-seq Analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
 - Image Analysis
 - Statistical Analysis
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.celrep.2019.04.030>.

ACKNOWLEDGMENTS

We thank members of the Liu lab for critical reading of, and insightful comments on, the manuscript. We thank Mayssa Mokalled for giving us the Tg (*hsp70l:ccn2a-2A-EGFP*) line as a gift. This work was supported by grants from the Ministry of Science and Technology of China (2016YFA0100500 to F.L. and 2015CB964803 and 2016YFE0108700 to J.-D.J.H.), the Strategic Priority Research program of the Chinese Academy of Sciences, China (XDA16010207 to F.L. and XDA01010303 to J.-D.J.H.), and the National Natural Science Foundation of China (31425016, 81530004, and 31830061 to F.L. and 91749205, 91329302, and 31210103916 to J.-D.J.H.). This study was supported by the State Key Laboratory of Membrane Biology, China.

AUTHOR CONTRIBUTIONS

Y.X. performed zebrafish experiments with help from Y.D., S.G., Y.Z., X.W., and P.L.; D.L. performed the bioinformatics analysis with help from D.A., S.S., C.Z., Y.L., and X.C., under the guidance of J.-D.J.H.; G.C., G.P., and N.J. assisted with GEO-seq; and Y.X., D.L., J.-D.J.H., and F.L. conceived the project, analyzed the data, and wrote the paper. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 13, 2018

Revised: January 30, 2019

Accepted: April 3, 2019

Published: April 30, 2019

REFERENCES

- Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169.
- Athanasiadis, E.I., Botthof, J.G., Andres, H., Ferreira, L., Lio, P., and Cvejic, A. (2017). Single-cell RNA-sequencing uncovers transcriptional states and fate decisions in haematopoiesis. *Nat. Commun.* 8, 2045.
- Bertrand, J.Y., Chi, N.C., Santoso, B., Teng, S., Stainier, D.Y., and Traver, D. (2010). Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature* 464, 108–111.
- Blaser, B.W., Moore, J.L., Hagedorn, E.J., Li, B., Riquelme, R., Lichtig, A., Yang, S., Zhou, Y., Tamplin, O.J., Binder, V., and Zon, L.I. (2017). CXCR1 remodels the vascular niche to promote hematopoietic stem and progenitor cell engraftment. *J. Exp. Med.* 214, 1011–1027.
- Boisset, J.C., van Cappellen, W., Andrieu-Soler, C., Galjart, N., Dzierzak, E., and Robin, C. (2010). In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. *Nature* 464, 116–120.
- Buettner, F., Natarajan, K.N., Casale, F.P., Proserpio, V., Scialdone, A., Theis, F.J., Teichmann, S.A., Marioni, J.C., and Stegle, O. (2015). Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat. Biotechnol.* 33, 155–160.
- Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* 36, 411–420.
- Chen, J., Suo, S., Tam, P.P., Han, J.J., Peng, G., and Jing, N. (2017). Spatial transcriptomic analysis of cryosectioned tissue samples with Geo-seq. *Nat. Protoc.* 12, 566–580.
- Covassin, L., Amigo, J.D., Suzuki, K., Teplyuk, V., Straubhaar, J., and Lawson, N.D. (2006). Global analysis of hematopoietic and vascular endothelial gene expression by tissue specific microarray profiling in zebrafish. *Dev. Biol.* 299, 551–562.
- Crane, G.M., Jeffery, E., and Morrison, S.J. (2017). Adult haematopoietic stem cell niches. *Nat. Rev. Immunol.* 17, 573–590.
- Davie, K., Janssens, J., Koldere, D., De Waegeneer, M., Pech, U., Kreft, L., Aibar, S., Makhzami, S., Christiaens, V., Bravo Gonzalez-Blas, C., et al. (2018). A single-cell transcriptome atlas of the aging *Drosophila* brain. *Cell* 174, 982–998.e20.
- Dobin, A., and Gingeras, T.R. (2015). Mapping RNA-seq reads with STAR. *Curr. Protoc. Bioinformatics* 51, 11.14.1–19.
- Fan, X., Dong, J., Zhong, S., Wei, Y., Wu, Q., Yan, L., Yong, J., Sun, L., Wang, X., Zhao, Y., et al. (2018). Spatial transcriptomic survey of human embryonic cerebral cortex by single-cell RNA-seq analysis. *Cell Res.* 28, 730–745.
- Farrell, J.A., Wang, Y., Riesenfeld, S.J., Shekhar, K., Regev, A., and Schier, A.F. (2018). Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science* 360, eaar3131.
- Fukui, H., Terai, K., Nakajima, H., Chiba, A., Fukuhara, S., and Mochizuki, N. (2014). S1P-Yap1 signaling regulates endoderm formation required for cardiac precursor cell migration in zebrafish. *Dev. Cell* 31, 128–136.
- Gao, S., and Liu, F. (2018). Fetal liver: an ideal niche for hematopoietic stem cell expansion. *Sci. China Life Sci.* 61, 885–892.
- Gao, X., Xu, C., Asada, N., and Frenette, P.S. (2018). The hematopoietic stem cell niche: from embryo to adult. *Development* 145, dev139691.
- Hall, C., Flores, M.V., Storm, T., Crosier, K., and Crosier, P. (2007). The zebrafish lysozyme C promoter drives myeloid-specific expression in transgenic fish. *BMC Dev. Biol.* 7, 42.
- Hao, S., Chen, C., and Cheng, T. (2016). Cell cycle regulation of hematopoietic stem or progenitor cells. *Int. J. Hematol.* 103, 487–497.
- Himburg, H.A., Harris, J.R., Ito, T., Daher, P., Russell, J.L., Quarmyne, M., Doan, P.L., Helms, K., Nakamura, M., Fixsen, E., et al. (2012). Pleiotrophin regulates the retention and self-renewal of hematopoietic stem cells in the bone marrow vascular niche. *Cell Rep.* 2, 964–975.
- Huang, W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37, 1–13.
- Huang, W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57.
- Junker, J.P., Noël, E.S., Guryev, V., Peterson, K.A., Shah, G., Huiskens, J., McMahon, A.P., Berezikov, E., Bakkers, J., and van Oudenaarden, A. (2014). Genome-wide RNA Tomography in the zebrafish embryo. *Cell* 159, 662–675.
- Khan, J.A., Mendelson, A., Kunisaki, Y., Birbrair, A., Kou, Y., Arnal-Estapé, A., Pinho, S., Ciero, P., Nakahara, F., Ma'ayan, A., et al. (2016). Fetal liver hematopoietic stem cell niches associate with portal vessels. *Science* 351, 176–180.
- Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360.
- Kissa, K., and Herbomel, P. (2010). Blood stem cells emerge from aortic endothelium by a novel type of cell transition. *Nature* 464, 112–115.
- Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A., and Kirschner, M.W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* 161, 1187–1201.
- Laurenti, E., and Göttgens, B. (2018). From haematopoietic stem cells to complex differentiation landscapes. *Nature* 553, 418–426.
- Laurenti, E., Frelin, C., Xie, S., Ferrari, R., Dunant, C.F., Zandi, S., Neumann, A., Plumb, I., Doulatov, S., Chen, J., et al. (2015). CDK6 levels regulate quiescence exit in human hematopoietic stem cells. *Cell Stem Cell* 16, 302–313.
- Leong, H.S., Chen, K., Hu, Y., Lee, S., Corbin, J., Pakusch, M., Murphy, J.M., Majewski, I.J., Smyth, G.K., Alexander, W.S., et al. (2013). Epigenetic regulator Smc4d1 functions as a tumor suppressor. *Cancer Res.* 73, 1591–1599.
- Lin, H.F., Traver, D., Zhu, H., Dooley, K., Paw, B.H., Zon, L.I., and Handin, R.I. (2005). Analysis of thrombocyte development in CD41-GFP transgenic zebrafish. *Blood* 106, 3803–3810.
- Löffler-Wirth, H., Kalcher, M., and Binder, H. (2015). oposSOM: R-package for high-dimensional portraying of genome-wide expression landscapes on bioconductor. *Bioinformatics* 31, 3225–3227.
- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550.
- Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 161, 1202–1214.
- Mahony, C.B., Fish, R.J., Pasche, C., and Bertrand, J.Y. (2016). tfec controls the hematopoietic stem cell vascular niche during zebrafish embryogenesis. *Blood* 128, 1336–1345.
- Mokalled, M.H., Patra, C., Dickson, A.L., Endo, T., Stainier, D.Y., and Poss, K.D. (2016). Injury-induced ctgfa directs glial bridging and spinal cord regeneration in zebrafish. *Science* 354, 630–634.
- Mosimann, C., Panáková, D., Werdich, A.A., Musso, G., Burger, A., Lawson, K.L., Carr, L.A., Nevis, K.R., Sabeh, M.K., Zhou, Y., et al. (2015). Chamber identity programs drive early functional partitioning of the heart. *Nat. Commun.* 6, 8146.
- Murayama, E., Kissa, K., Zapata, A., Mordet, E., Briolat, V., Lin, H.F., Handin, R.I., and Herbomel, P. (2006). Tracing hematopoietic precursor migration to successive hematopoietic organs during zebrafish development. *Immunity* 25, 963–975.
- Nygren, J.M., Bryder, D., and Jacobsen, S.E.W. (2006). Prolonged cell cycle transit is a defining and developmentally conserved hemopoietic stem cell property. *J. Immunol.* 177, 201–208.
- Pauklin, S., and Vallier, L. (2013). The cell-cycle state of stem cells determines cell fate propensity. *Cell* 155, 135–147.
- Peng, G., Suo, S., Chen, J., Chen, W., Liu, C., Yu, F., Wang, R., Chen, S., Sun, N., Cui, G., et al. (2016). Spatial transcriptome for the molecular annotation of lineage fates and cell identity in mid-gastrula mouse embryo. *Dev. Cell* 36, 681–697.

- Petropoulos, S., Edsgård, D., Reinius, B., Deng, Q., Panula, S.P., Codeluppi, S., Plaza Reyes, A., Linnarsson, S., Sandberg, R., and Lanner, F. (2016). Single-cell RNA-seq reveals lineage and X chromosome dynamics in human pre-implantation embryos. *Cell* 165, 1012–1026.
- Ramilowski, J.A., Goldberg, T., Harshbarger, J., Kloppmann, E., Lizio, M., Satagopam, V.P., Itoh, M., Kawaji, H., Carninci, P., Rost, B., and Forrest, A.R. (2015). A draft network of ligand-receptor-mediated multicellular signalling in human. *Nat. Commun.* 6, 7866.
- Roccio, M., Schmitter, D., Knobloch, M., Okawa, Y., Sage, D., and Lutolf, M.P. (2013). Predicting stem cell fate changes by differential cell cycle progression patterns. *Development* 140, 459–470.
- Sánchez, M.J., Holmes, A., Miles, C., and Dzierzak, E. (1996). Characterization of the first definitive hematopoietic stem cells in the AGM and liver of the mouse embryo. *Immunity* 5, 513–525.
- Sergushichev, A. (2016). An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. *bioRxiv*. <https://doi.org/10.1101/060012>.
- Shaw, N.D., Brand, H., Kupchinsky, Z.A., Bengani, H., Plummer, L., Jones, T.I., Erdin, S., Williamson, K.A., Rainger, J., Stortchevoi, A., et al. (2017). SMCHD1 mutations associated with a rare muscular dystrophy can also cause isolated arhinia and *Bosma arhinia* microphthalmia syndrome. *Nat. Genet.* 49, 238–248.
- Smith, T., Heger, A., and Sudbery, I. (2017). UMI-tools: modeling sequencing errors in unique molecular identifiers to improve quantification accuracy. *Genome Res.* 27, 491–499.
- Sun, N., Yu, X., Li, F., Liu, D., Suo, S., Chen, W., Chen, S., Song, L., Green, C.D., McDermott, J., et al. (2017). Inference of differentiation time for single cell transcriptomes using cell population reference data. *Nat. Commun.* 8, 1856.
- Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P., et al. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43, D447–D452.
- Tamplin, O.J., Durand, E.M., Carr, L.A., Childs, S.J., Hagedorn, E.J., Li, P., Yzaguirre, A.D., Speck, N.A., and Zon, L.I. (2015). Hematopoietic stem cell arrival triggers dynamic remodeling of the perivascular niche. *Cell* 160, 241–252.
- Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* 6, 377–382.
- Teschendorff, A.E., and Enver, T. (2017). Single-cell entropy for accurate estimation of differentiation potency from a cell's transcriptome. *Nat. Commun.* 8, 15599.
- Wang, L., Zhang, P., Wei, Y., Gao, Y., Patient, R., and Liu, F. (2011). A blood flow-dependent klf2a-NO signaling cascade is required for stabilization of hematopoietic stem cell programming in zebrafish embryos. *Blood* 118, 4102–4110.
- Wang, C.Y., Jegu, T., Chu, H.P., Oh, H.J., and Lee, J.T. (2018). SMCHD1 merges chromosome compartments and assists formation of super-structures on the inactive X. *Cell* 174, 406–421.
- Wei, W., Wen, L., Huang, P., Zhang, Z., Chen, Y., Xiao, A., Huang, H., Zhu, Z., Zhang, B., and Lin, S. (2008). Gfi1.1 regulates hematopoietic lineage differentiation during zebrafish embryogenesis. *Cell Res.* 18, 677–685.
- Winkler, I.G., Barbier, V., Nowlan, B., Jacobsen, R.N., Forristal, C.E., Patton, J.T., Magnani, J.L., and Lévesque, J.P. (2012). Vascular niche E-selectin regulates hematopoietic stem cell dormancy, self renewal and chemoresistance. *Nat. Med.* 18, 1651–1657.
- Xue, Y., Gao, S., and Liu, F. (2015). Genome-wide analysis of the zebrafish Klf family identifies two genes important for erythroid maturation. *Dev. Biol.* 403, 115–127.
- Xue, Y., Lv, J., Zhang, C., Wang, L., Ma, D., and Liu, F. (2017). The vascular niche regulates hematopoietic stem and progenitor cell lodgment and expansion via klf6a-ccl25b. *Dev. Cell* 42, 349–362.
- Zepp, J.A., Zacharias, W.J., Frank, D.B., Cavanaugh, C.A., Zhou, S., Morley, M.P., and Morrissey, E.E. (2017). Distinct mesenchymal lineages and niches promote epithelial self-renewal and myofibrogenesis in the lung. *Cell* 170, 1134–1148.
- Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Girón, C.G., et al. (2018). Ensembl 2018. *Nucleic Acids Res.* 46 (D1), D754–D761.
- Zhang, W., Liu, Y., Sun, N., Wang, D., Boyd-Kirkup, J., Dou, X., and Han, J.D. (2013). Integrating genomic, epigenomic, and transcriptomic features reveals modular signatures underlying poor prognosis in ovarian cancer. *Cell Rep.* 4, 542–553.
- Zhang, P., He, Q., Chen, D., Liu, W., Wang, L., Zhang, C., Ma, D., Li, W., Liu, B., and Liu, F. (2015). G protein-coupled receptor 183 facilitates endothelial-to-hematopoietic transition via Notch1 inhibition. *Cell Res.* 25, 1093–1107.
- Zhang, Y., Gao, S., Xia, J., and Liu, F. (2018). Hematopoietic Hierarchy - An Updated Roadmap. *Trends Cell Biol.* 28, 976–986.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-Digoxigenin-AP	Roche	11093274910; AB_2734716
Alexa Fluor 555 Goat Anti-Mouse IgG (H+L)Antibody	Invitrogen	A21422; AB_141822
Chemicals, Peptides, and Recombinant Proteins		
DPBS basic (1x)	GIBCO	14190-250
Fetal Bovine Serum	GIBCO	16000-044
2.5% Trypsin (10 x)	GIBCO	15090046
Tissue freezing medium(OCT)	Leica	020108926
Cresyl violet acetate	Sigma	C5042
Glycogen	Roche	10901393001
RNase inhibitor	TOYOBO	SIN-201
Guanidine isothiocyanate	Invitrogen	15577-018
Magnesium chloride	Sigma	M1028-100ML
KAPA HiFi HotStart ReadyMix (2X)	KAPA Biosystems	KK2601
SuperScript II reverse transcriptase	Invitrogen	18064-014
Betaine solution	Sigma	B0300-1VL
DIG RNA labeling mix	Roche	96714
Flu RNA labeling mix	Roche	60750
BM purple	Roche	11621
Critical Commercial Assays		
5-Bromo-2'-deoxy-uridine Labeling and Detection Kit	Roche	11296736001
MoloneyMurineLeukemiaVirus (M-MLV) ReverseTranscript	Promega	90694
TSA™ plus Cyanine 3/Fluorescein System	PerkinElmer	2516252
mirVana™ miRNA Isolation Kit	Invitrogen	00670905
Deposited Data		
Bulk RNA-seq, GEO-seq, single cell RNA-seq	This paper	https://www.ncbi.nlm.nih.gov/geo/ ; Accession number GEO: GSE120581
Webserver	This paper	http://www.picb.ac.cn/hanlab/ichtatlas/ Home/
Experimental Models: Organisms/Strains		
Zebrafish: <i>kdr1</i> :mCherry	from B. Zhang Lab in Peking University	NA
Zebrafish: <i>CD41</i> :GFP	Lin et al., 2005	NA
Zebrafish: <i>runx1</i> :en-GFP	Zhang et al., 2015	NA
Zebrafish: <i>gata1</i> :DsRed	Hall et al., 2007	NA
Zebrafish: -6.35drl:EGFP	Mosimann et al., 2015	NA
Zebrafish: <i>hsp70l</i> : <i>ccn2a</i> -2A-GFP	Mokalled et al., 2016	NA
Zebrafish: <i>gfi1</i> :GFP	Wei et al., 2008	NA
Oligonucleotides		
See Table S6 for list of primers	This paper	NA
<i>ctgfa</i> ATG MO: 5'-GAGTCATTCAGAAA ACATGATGAC-3'	Fukui et al., 2014	NA
<i>itgb2</i> ATG MO: 5'- CATCTCAGTTGTTGA AATGCAACGC -3'	This paper	NA
<i>smchd1</i> ATG MO: 5'- ATCACTTCCACAA CCGTCACTTTTC -3'	This paper	NA
<i>smchd1</i> splice MO: 5'- AGGTGTGATTTC GACTTACGCAAC -3'	Shaw et al., 2017	NA

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Recombinant DNA		
Plasmid: <i>ctgfa</i> -Pegfpn1	This paper	NA
Plasmid: <i>itgb2</i> -Pegfpn1	This paper	NA
Software and Algorithms		
Data analysis	Graph Pad Prism 6	https://www.graphpad.com/
Self-organizing maps	oposSOM	https://www.ncbi.nlm.nih.gov/pubmed/26063839
Secretomics analysis	FANTOM	http://fantom.gsc.riken.jp/
Single cell data analysis pipeline	Seurat 2.3	https://satijalab.org/seurat/
Single cell lineage construction	Monocle 2.4	http://cole-trapnell-lab.github.io/monocle-release/
Imaging analysis	Bitplane Imaris 7.4.2	http://www.bitplane.com/
Imaging analysis	photoshop CC2018	https://www.adobe.com/cn/
Imaging analysis	Adobe Illustrator CC 2018	https://www.adobe.com/cn/
Imaging analysis	ImageJ	https://imagej.nih.gov/ij/

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Feng Liu (liuf@ioz.ac.cn).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Adult zebrafish including wild-type Tubingen and transgenic lines including *kdr1*:mCherry (from B. Zhang Lab in Peking University), *CD41*:GFP (Lin et al., 2005), *gata1*:DsRed (Hall et al., 2007), *runx1*:en-GFP (Zhang et al., 2015), -6.35*kdr1*:EGFP (Mosimann et al., 2015), *gfi1*:GFP (Wei et al., 2008) and *hsp70l:ccn2a-2A-EGFP* (Mokalled et al., 2016) were raised under a standard condition. This study was approved by the Ethical Review Committee of Institute of Zoology, Chinese Academy of Sciences, China.

METHOD DETAILS

MOs Microinjection, Overexpression and CRISPR/Cas9 Mutant

The antisense MOs (MOs) including *smchd1* splice MO, *itgb2* ATG MO and *ctgfa* ATG MO were purchased from GeneTools. The sequences of gene-specific MOs are described as previously reported, i.e., *ctgfa* ATG MO (5'-GAGTCATTCCAGAAAACATGATGAC-3') (Fukui et al., 2014), *itgb2* ATG MO (5'-CATCTCAGTTGTTGAAATGCAACGC-3'), *smchd1* ATG MO (5'-ATCACTTCCACAACCGTCACCTTTTC-3') and *smchd1* splice MO (5'-AGGTGTGATTTCAGACTTACG

CAAC-3') (Shaw et al., 2017). 1 mM MO stock solutions were prepared by ddH₂O dilution, and 4-8 ng of MOs were injected into 1-cell-stage zebrafish embryos. The Tg (*hsp70l:ccn2a-2A-EGFP*) line used for overexpression of *ctgfa* was a gift from Prof. Mokalled (Washington University School of Medicine). The *itgb2* (ENSART00000149704.2) guide RNA was designed for targeting exon5 (target sequence, GGACCCTCCAGAGGAAGCC).

Whole-mount and Fluorescence *in Situ* Hybridization (WISH and FISH)

The RNA probes targeting the genes of *ccl25b*, *cmyb*, *irf4a*, *be1-globin*, *gata1*, *pu.1*, *mfap4*, *scl*, *myod*, *rag1*, *foxn1*, *egfp*, *smchd1*, *ctgfa*, *itgb2*, *atf7ip*, *apln*, *hdr*, *tcima* and *kiaa0101* were transcribed by SP6 or T7 polymerase labeled by digoxigenin (DIG) or fluorescein (FLU). WISH was performed as previously described (Wang et al., 2011). We performed FISH to detect co-expression of genes. The fixed embryos were treated with proteinase K (10 µg/ml) and then hybridized with DIG/FLU-labeled antisense RNA probe at 65°C for 15 to 18 hours (h). After washing with PBST and blocking with MABT, embryos were incubated with anti-DIG/FLU-POD antibody (1:100) at 4°C overnight. After removing antibody and washing embryos with PBST, embryos were stained with TSA-FITC/Cy3 amplification reagent (1:100). Before the secondary color staining or immunofluorescence staining, the first color reaction was stopped by gradient methanol. After removing the reaction buffer, embryos were washed sequentially by 25%, 50% and 75% methanol/PBST (10 minutes/each buffer); 1% H₂O₂/methanol (30 minutes); 75%, 50% and 25% methanol/PBST (10 minutes/each buffer); PBST (2 × 5 minutes). The secondary antibody, anti-FLU/DIG-POD was used for incubation. The followed steps were similar with first color

staining procedure. The secondary immunofluorescence staining process can refer to BrdU assay. The primers are listed in Table S6.

BrdU Assay

BrdU labeling was performed as described previously (Xue et al., 2015). In brief, embryos were injected with BrdU (10 mM) at 2.5 days post fertilization (dpf) and were fixed using 4% paraformaldehyde (PFA) 2 h after injection. After washing, embryos were digested with the Proteinase K (10 μ g/ml), then treated with HCl (2 mol/L) for 1 h. After blocking with 1% BSA for 1 h, the embryos were incubated with anti-BrdU antibody (1:800) (5-Bromo-2'-deoxy-uridine Labeling and Detection Kit, Roche) at 4°C overnight. After washing with PBST, the embryos were incubated with Alexa Fluor 555 Goat Anti-Mouse IgG (H+L) Antibody (1:500).

Quantitative Real-Time PCR and Plasmids Construction

Total RNAs used for RT-PCR or qPCR were extracted by TRIzol reagent from the whole embryos or dissected tails. The cDNA templates were reversely transcribed by oligo-dT or random primers from total RNAs and subsequently diluted 4 times. The RT-PCR was used to validate the efficiency of *smchd1* splice MO. The primers were designed to amplify the region from the third exon to the fifth exon of *smchd1* gene sequence. The detailed protocols were followed as described previously (Wang et al., 2011). To validate the efficiency of *ctgfa* and *itgb2* ATG MOs, we generated the *ctgfa*-GFP and *itgb2*-GFP fusion constructs using PEGFPN1 basic vector. The construct contained a partial sequence of *ctgfa* or *itgb2* gene including the ATG MO recognition site and fused GFP in frame.

Cell Sorting, RNA Extraction and mRNA Sequencing

Tg (*kdr1*:mCherry/*CD41*:GFP) embryos at successive stages, including 36 hours post fertilization (hpf), 52 hpf, 3 dpf and 4 dpf, were collected after removing heads or cutting tail regions, respectively. The dissociated tissues were washed by calcium free Ringer solution in 6-well plates and incubated with 2.5% trypsin in PBS for 15 min, then stopped by adding CaCl_2 to a final concentration of 1 mM and fetal calf serum to 10%. Cells were centrifuged for 6 min at 300 g and washed by PBS once. The single cell suspensions for FACS were filtrated through 300 Mesh nylon filter and were subjected to FACS analysis (Covassin et al., 2006). The *CD41*:GFP⁺ cells and *kdr1*:mCherry⁺ cells were sorted by MoFlo XDP (Beckman). The total RNAs of sorted cells were extracted by QIAGEN RNeasy Mini Kit (Cat. No. 74G4). The mRNA sequencing libraries were generated using NEBNext Ultra RNA Library Prep Kit for Illumina. The library preparations were sequenced on an Illumina HiSeq 4000 platform and 150 bp paired-end reads were generated.

Confocal Microscopy

After anesthetized and mounted in 1.2% low melting agarose, transgenic zebrafish embryos were imaged by Nikon confocal A1 laser microscope with 20 \times objective.

Tissue Preparation for GEO-seq

All the reagents and instruments to be used in this experiment were RNase-free. The fresh 55 hpf zebrafish embryonic tail was embedded in OCT with vertical position and rapid frozen in liquid nitrogen. The cryosection was stained with cresyl violet before laser microdissection. Six samples were acquired from each section using specific LCM platforms and lysed with lysis buffer. Then trace amounts of total RNAs were prepared for single cell-based PCR and further amplification. After quality control and qPCR validation, the cDNA libraries were prepared for sequencing. The detailed steps were described previously (Chen et al., 2017).

Cell Preparation for single cell RNA-seq

The tails of 55 hpf transgenic embryos were dissected into single cell suspension for fluorescence activated cell sorting (FACS). The 50,000 *kdr1*:mCherry⁺ cells, 50,000 *CD41*:GFP⁺ cells and 50,000 double negative cells were mixed for 10 \times Genomics analysis (CapitalBio Technology Corporation). Then single cell suspensions (300-1000 living cells per microliter determined by trypan blue staining) were loaded on a Chromium Single Cell Controller (10 \times Genomics) to generate single-cell gel beads in emulsion (GEMs) by using Single Cell 30 Library and Gel Bead Kit V2 (10 \times Genomics, 120237). Cell lysis and RNAs barcoding were accomplished through reverse transcription in individual GEMs. Barcoded cDNAs were pooled and cleaned up by beads (Invitrogen, 37002D). Single-cell RNA-seq libraries were prepared using Single Cell 30 Library Gel Bead Kit V2 (10 \times Genomics, 120237) following the manufacture's introduction. Sequencing was performed on an Illumina HiSeq X Ten with pair end 150 bp (PE150).

Preprocessing of the Bulk and GEO RNA-seq Data

Cell population RNA-seq and GEO-seq data were first subjected to "FastQC" (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) for checking the data quality. And then clean reads were mapped to the reference genome (Version: GRCz10) downloaded from Ensemble (<https://www.ensembl.org/>) by HISAT (Kim et al., 2015) with the parameter "-qc -filter." Uniquely mapped reads were utilized to calculate the counts of each gene by HTSeq (Anders et al., 2015) based on the GTF annotation downloaded from Ensemble (Version: GRCz10.89). Reads counts matrix was normalized to library size by DESeq2 (Love et al., 2014) with default parameters, and normalized reads were then used to calculate the FPKM (Fragments Per Kilobase of transcript per Million fragments mapped).

Bulk RNA-seq Analysis

Principal component analysis (PCA) was performed by FactoMineR package in R, based on the expressed genes (FPKM > 1 in at least two samples). Stage specific genes were identified based on the previous described Stouffer's method (Petropoulos et al., 2016). We first combined the Z scores of each gene obtained from the differential expression analysis of one stage against the other three groups. And selected genes that have FDR < 0.01 as stage specific genes. Differentially expressed genes (DEGs) were identified by DESeq2 (Love et al., 2014) between pairwise groups. Genes were defined as DEGs if the FDR < 0.01 between any two groups. BICSKmeans (Zhang et al., 2013) was utilized to determine the number of clusters of DEGs. And the output of BICSKmeans was imported into the "Java Treeview" for visualization. Functional enrichment analysis was performed by DAVID (Huang et al., 2009a, 2009b) with the default parameters.

GEO-seq Data Analysis

PCA, DEGs and BICSKmeans clustering were performed as previous described. To perform self-organizing maps (SOM), we first identified the most variable 2000 genes by fitting a loess curve between the mean versus the standard deviation of all samples. And these variable genes were then feed into oposSOM (Löffler-Wirth et al., 2015) package in R for constructing the SOM, with the parameter "standard.spot.modules=overexpression."

Activation of signaling pathway was estimated by a GSEA-like method. First, we downloaded the gene list of 122 signaling pathways (Version: c2.cp.v6.1). And then, these gene lists were used for calculated normalized enrichment scores (NES) together with the gene expression pattern by fgsea (Sergushichev, 2016) package in R. And at last, corn plot (Chen et al., 2017) was utilized to visualize the NES value for each signaling pathway.

Construction of the 3D gene expression of CHT region was performed as previous described (Peng et al., 2016). The expression pattern of six layers of CHT region was smoothed into 30 layers. And then, based on the microscope imaging picture of each layer, we colored corresponding region based on the smoothed expression level of this region. Vaa3D (<http://home.penglab.com/proj/vaa3d/Vaa3D/Download.html>) were then used to integrated these pictures together for visualization.

In silico Secretomics Analysis

The annotation between ligand-receptor was downloaded from FANTOM (Ramilowski et al., 2015). Because this database was based on human annotation. So we transformed these human gene symbol into zebrafish gene symbol by Ensembl (Zerbino et al., 2018). And to further filter some false positive regulatory relationship, we utilized STRING (v10) (Szklarczyk et al., 2015) of zebrafish to keep those only annotated in this database. We obtained 1865 regulatory relationship, containing 437 ligands and 415 receptors. At last, we kept those differentially expressed (FDR < 0.01) ligands and receptors among different tissues for visualization.

Preprocessing of the Single Cell RNA-seq Data

We used UMI-tools (Smith et al., 2017) for data quality and gene expression quantification from fastq files. First, in "whitelist" step we set 5000 as cell number and check quality plots. Then in "extract" step we filter some low-quality reads and non-UMI reads with recommend parameters, so that we can use STAR (Dobin and Gingeras, 2015) to map clean reads to the reference genome. Finally, in "count" step we generate a gene expression contains unique molecular identifies (UMIs) matrix for following analysis.

Single Cell RNA-seq Analysis

Seurat (Butler et al., 2018) was utilized for data quality control and group clustering. After filter these cells expressed too low (< 800) or too high (> 4000) gene number, we obtained 3,774 high quality single cells for further analysis. Then, we selected 2,000 most variable genes and fed into Seurat for dimension reduction by PCA. The top 10 PCs were selected for t-SNE clustering. The number of clustering is depended on the parameter resolution. We tried different parameters and finally selected 0.8, by which the number of groups is 9. We propose this number of groups may best present the structure of single cells. DEGs of each group were also identified by Seurat, with the cut off FDR < 0.05.

The cell cycle phase of each single cell were identified by scran (Buettner et al., 2015) package in R. And the signaling entropy was calculated by scent (Teschendorff and Enver, 2017) package in R, with the input of STRING network database (Szklarczyk et al., 2015).

We performed zip-code mapping of single cell into the GEO-seq as previous described (Peng et al., 2016). Zip-code genes were contained in the PC1, PC2 loading genes and DEGs of the single cells, to present both the spatial information and cell type signature. And for each single cell, we calculated the spearman rank correlation between this single cell and samples from GEO-seq. And take the best hit (with p value < 0.01 and the lowest p value) as the mapping position of each single cell.

QUANTIFICATION AND STATISTICAL ANALYSIS

Image Analysis

Raw image data were processed using ImageJ, photoshop CC 2018 and Adobe Illustrator CC 2018. The Imaris 7.4.2 (Bitplane Scientific Software) was used for digital image processing and the HSPCs were counted in about 6-8-somite region of CHT.

Statistical Analysis

GraphPad Prism 6 was used to analyze the data. Values are presented as means \pm SEM. Student's unpaired 2-tailed t test was applied for comparisons unless otherwise indicated.

DATA AND SOFTWARE AVAILABILITY

The accession number for the RNA-seq data reported in this paper is Gene Expression Omnibus database: GSE120581.