PNAS
www.pnas.org

Supplementary Information for

Hemogenic and aortic endothelium arise from a common hemogenic angioblast precursor and are specified by the Etv2 dosage

Shizheng ZHAO, Shachuan FENG, Ye TIAN and Zilong WEN

Zilong WEN
E-mail: zilong@ust.hk

This PDF file includes:

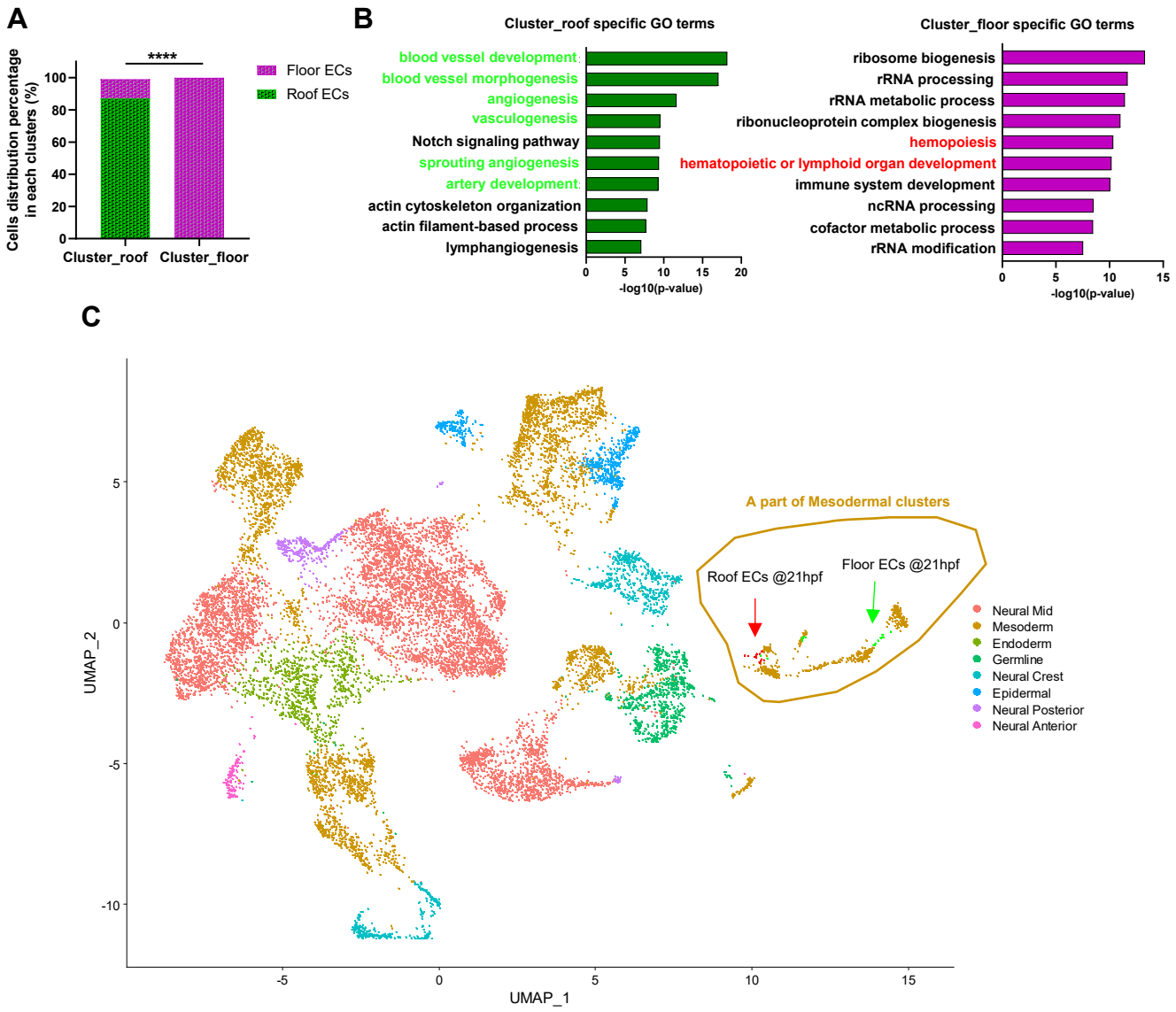

Fig. S1. Clustering and GO analysis of Cluster_roof and Cluster_floor. (A) Barplot graph shows the distribution percentage (%) of roof or floor cells in each cluster shown in Figure. 3A. Cluster_roof was composed of 87% roof cells and 13% floor cells derived from 21 hpf and 28 hpf, while Cluster_floor was composed of 100% floor cells derived from 21 hpf and 28 hpf. (B) GO term analysis on Cluster_roof-specific and Cluster_floor-specific gene-sets, in which the top 10 enriched terms were listed. Informative GO terms were highlighted in green or magenta, respectively. (C) UMAP-plot for cell clustering analysis of total integrated dataset. Dataset in this study (21 hpf) and the total Wagner's dataset (10 hpf, 4280 cells; 14 hpf, 4001 cells; 18 hpf, 6962 cells) are clustered together in (C). Cells in our dataset are predominantly (33/36; 92%) clustered with mesoderm cells defined in Wagner's study (ref.37). Fisher's exact test used (A), ****$p \leq 0.0001$.

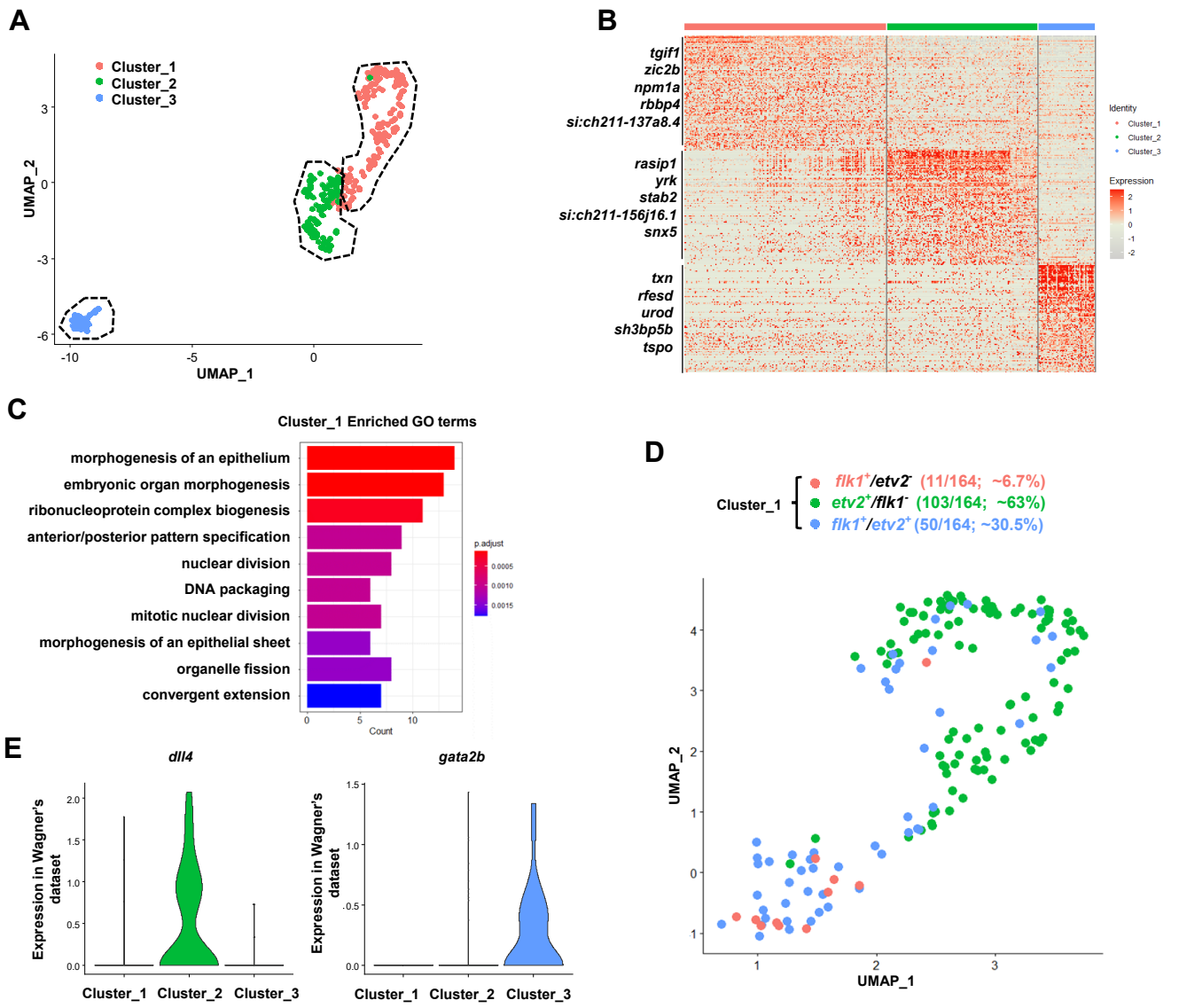

Fig. S2. Clustering and GO analysis of integrated dataset. (A) UMAP-plot for cell clustering analysis of integrated dataset after selecting *etv2*$^{+}$ and *flk1*$^{+}$ cells. Three clusters, Cluster_1, Cluster_2 and Cluster_3, were identified. (B) Expression heatmap of the top 100 feature genes of three clusters in (A). Left panel shows the top 5 feature genes in each cluster. (C) GO analysis on Cluster_1 top 100 gene-sets, in which the top 10 enriched terms were listed. (D) Enlarged plot of Cluster_1 in (A) showing distribution and percentage of *etv2* or *flk1* single positive and *etv2/flk1* double positive cells. (E) Violin-plot for HEC marker *gata2b* and cEC marker *dll4* gene expression in Cluster_1, Cluster_2, and Cluster_3 of cells from Wagner's dataset.

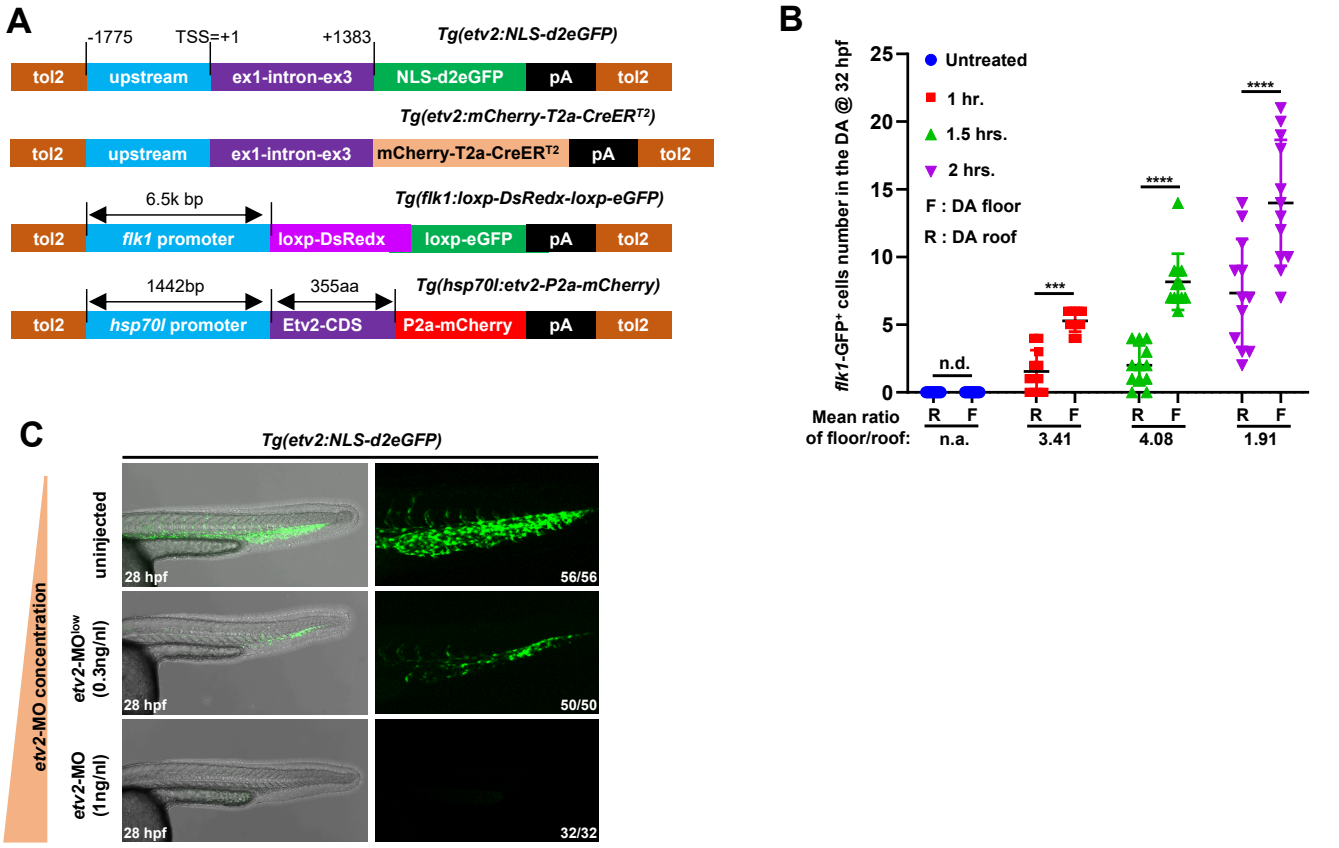

Fig. S3. Transgenic zebrafish construction and Etv2-dosage dependent lineage tracing. (A) Schematic diagram of *Tg(etv2:NLS-d2eGFP)*, *Tg(etv2:mCherry-T2a-CreER^{T2})*, *Tg(hsp70l:etv2-P2a-mCherry)* and *Tg(flk1:loxp-DsRedx-loxp-eGFP)* construction. (B) Lineage tracing using double transgenic line *Tg(etv2:mCherry-T2a-CreER^{T2};flk1:loxp-DsRedx-loxp-eGFP)*. Fish were untreated or treated with median dose of 4-OHT (5μM) for different time lengths starting from 15 hpf and GFP^{+} ECs in the DA roof or floor were quantified at 32 hpf. Mean ratio was calculated by the mean number of floor GFP^{+}/roof GFP^{+} cells in each conditions. (C) *etv2* MO dose testing. MO concentration $\geq$1 ng/nl with 2 nl injection volume per embryo was sufficient to completely block *etv2-d2eGFP* expression (referred as *etv2-MO* or *etv2* morphants), while MO concentration $\leq$0.3ng/nl only partially blocked *etv2-d2eGFP* expression (referred as *etv2-MO^{low}*). TSS, transcription starting site; pA, SV40 polyA sequence; n.d., not detectable; n.a., not available. Data are acquired by two indpendent experiment in (B) or two different clutches of embryos in (C). n/N reports the number of embryos with pattern in image/total embryos in (C). Embryos number for quantification in Untreated/0.5/1/1.5 hrs. groups are 10/9/11/12 in (B). Data are represented as mean ± SD and exact p values were labelled in (B). Two-way ANOVA test used in (B). ***p $\leq$ 0.001, ****p $\leq$ 0.0001, n.d. not detectable, n.a. not available.

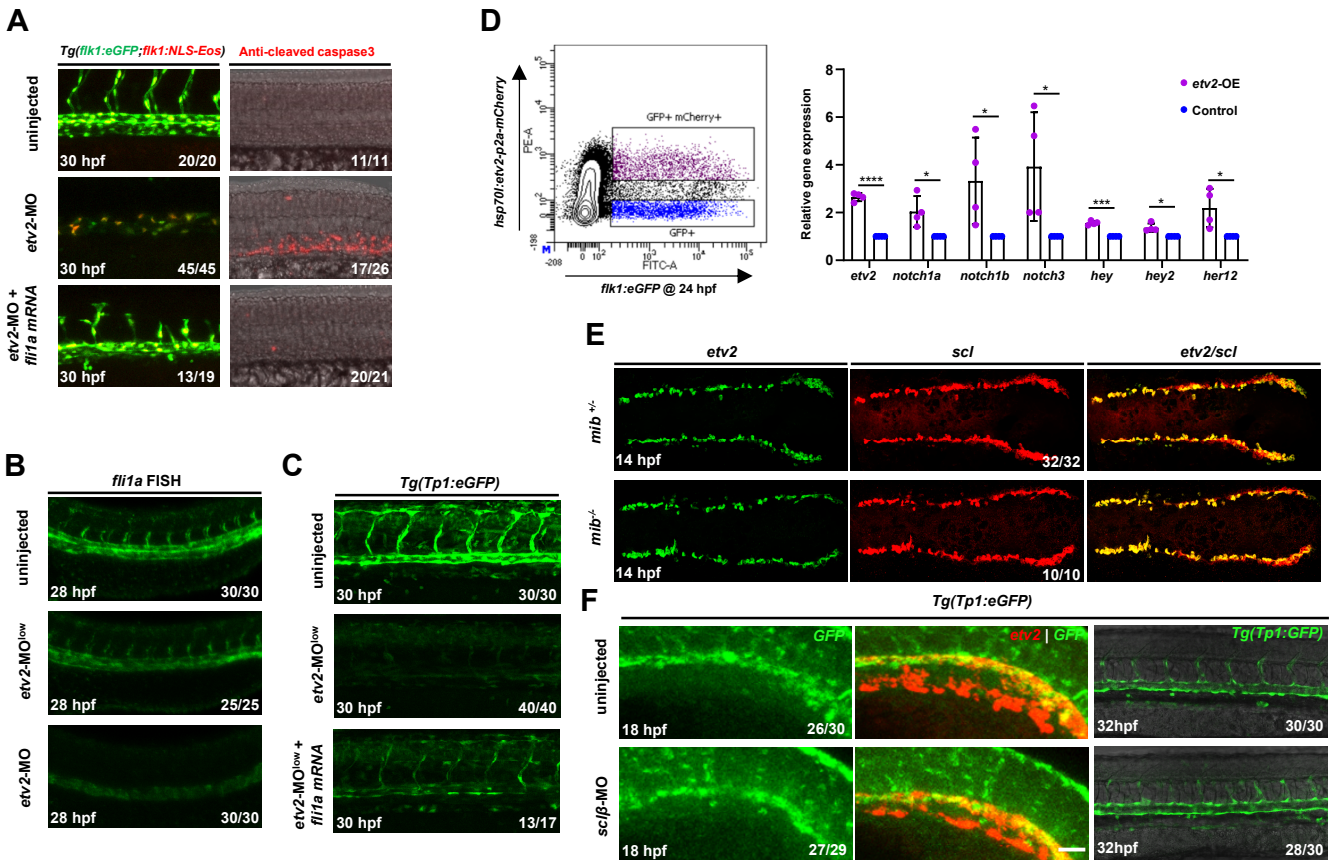

Fig. S4. Interplay among Etv2, Fli1a, Sclβ and Notch signaling. (A) *fli1a* mRNA injection partially rescued the DA structure and largely rescued EC apoptosis in *etv2* morphants. (B) FISH of *fli1a* in *etv2* morphants and *etv2-MO*^{low} embryos. The embryos were injected with or without low or high dose of *etv2* MO. (C) *fli1a* mRNA injection partially rescued *Tp1-eGFP* reporter in low dose *etv2* MO knockdown embryos. (D) Flow cytometric and gene expression analysis in *flk1:eGFP*$^{+}$ endothelial cells with or without *etv2* overexpression. Etv2-mCherry was induced by waterbath heatshock at 14 hpf for 45min. (E) dFISH of *etv2* and *scl* (sclα/β) at 14 hpf in PLPM in *mib* sibling and *mib* mutant embryos. (F) dFISH of *eGFP* and *etv2* and representative single slice confocal image in the *scl*β MO-injected *Tg(Tp1:eGFP)* and uninjected embryos. Data are representative of two independent experiments with four biological replicates (D) or two different clutches of embryos for dFISH or WISH assay (A-C,E-F). n/N reports the number of embryos with pattern in image/total embryos (A-C,E-F). Scale bars, 30 μm; TSS, transcription starting site; pA, SV40 polyA sequence; PLPM, posterior lateral plate mesoderm. Paired student's t tests used in (D). Data are represented as mean ± SD, *p $\leq$ 0.05, ***p $\leq$ 0.001, ****p $\leq$ 0.0001.

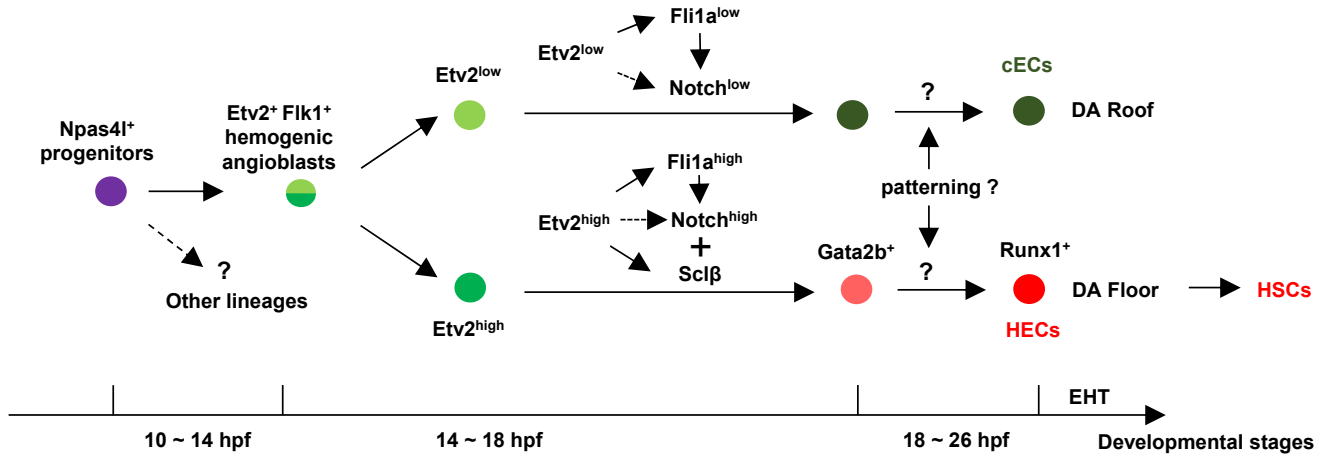

Fig. S5. Proposed working model of HEC and cEC fate specification. During 10-14 hpf, Etv2^{+}Flk1^{+} hemogenic angioblasts are formed from Npas4l^{+} precursors in the posterior lateral plate mesoderm. Whether other lineages (e.g., primitive erythrocytes) and Etv2^{+} angioblasts arise from a common Npas4l^{+} precursor remain unknown. During 14-18 hpf, the first wave of Etv2^{+}Flk1^{+} hemogenic angioblasts migrate towards midline to form the DA, during which HECs and cECs are specified as a result of Etv2 dosage together with the differential activation of Etv2 downstream factors, Fli1a, Notch and Sclβ. While Etv2^{low}-Fli1a^{low}-Notch^{low} axis is sufficient to maintain EC lineage survival, differentiation and DA identity, HEC lineage requires high level of Etv2, Fli1a, Notch, and Sclβ. During 18-26 hpf, HEC and cEC precursors are patterned to the DA roof and floor respectively by an unknown mechanism.

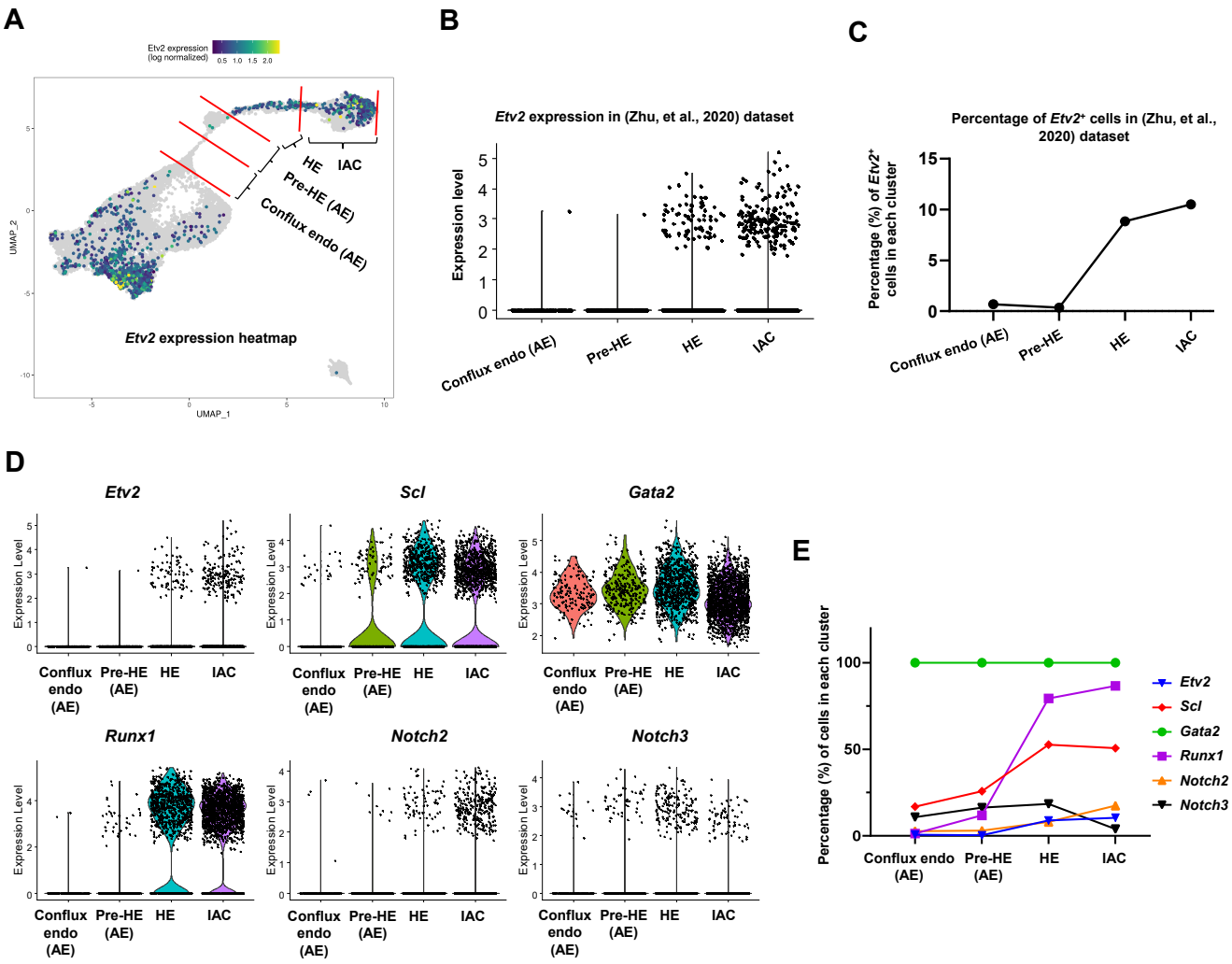

Fig. S6. Cross-species comparison between this study and murine scRNA-seq dataset. (A) Overview of *Etv2* gene expression heatmap in pre-HSC ontogeny trajectory from (Zhu, et al.,2020) study (ref.17). (B,C) *Etv2* gene expression level and percentage of expressing cells in defined clusters by (Zhu, et al.,2020) study. (D,E) *Etv2, Scl, Gata2, Runx1, Notch2* and *Notch3* gene expression levels and percentage of expressing cells in defined clusters. HSC, hematopoietic stem cell; AE, arterial endothelium; HE, hemogenic endothelium; IAC, intra-aortic clusters.

Table S1. qPCR primers used in this study

qPCR primer name	Sequence(5'-3')
gata2b-fwd	ACCACCACACTCTGGAGAC
gata2b-rev	CTGTTGCGTGTCTGAATACC
etv2-fwd	GAGCTGTTGCACAAAGGTCA
etv2-rev	CAGAGAGGGACGAGGTTCTG
fli1a-fwd	TCGTCCTCAGCCAGATCC
fli1a-rev	TGGTTCCTTCCCAGGTGA
scl-fwd	GGAGATGCGGAACAGTATGG
scl-rev	GAAGGCACCGTTCACATTCT
notch1a-fwd	CGGGCCTGACGGATTCAC
notch1a-rev	GGACTCCAGCAGACGTTTAGC
notch1b-fwd	ACAGAAGGGGGAGACCAGTT
notch1b-rev	CCTCATGATAAGCACAAATTCCT
notch3-fwd	GCATTGACCGACCTAATGGA
notch3-rev	TGCTCTCACACAGTCTTCCTTC
hey-fwd	GCGCCTTTGAGAAACAGGGCTCA
hey-rev	ACGGATGCGGAGGGGATCTGT
hey2-fwd	GTGGCTCACCTACAACGACA
hey2-rev	CTCCAACTTGGCAGATCCCT
her12-fwd	GACATGGCACCCCACTCAGCC
her12-rev	GGTGCTGGGATCGTGACTGTGG