

## ***Hapln1b* organizes the ECM to modulate *kit* signaling and control developmental hematopoiesis in zebrafish**

Christopher B. Mahony<sup>1,4,5</sup>, Corentin Pasche<sup>1</sup>, Vincent Braunersreuther<sup>2</sup>, Savvas N. Savvides<sup>3</sup>, Ariane de Agostini<sup>1</sup> and Julien Y. Bertrand<sup>1,5</sup>.

*Running title: the extracellular matrix controls HSC emergence.*

Addresses:

<sup>1</sup> University of Geneva, Faculty of Medicine, Department of Pathology and Immunology, Rue Michel-Servet 1, 1211 Geneva 4, Switzerland.

<sup>2</sup> Service of Clinical Pathology, Geneva University Hospital, Rue Gabrielle-Perret-Gentil 4, 1205 Geneva, Switzerland.

<sup>3</sup> VIB-UGent Center for Inflammation Research, Department of Biochemistry and Microbiology, Technologiepark-Zwijnaarde 71, 9052 Ghent, Belgium.

<sup>4</sup> current address: Institute of Cancer and Genomic Sciences, College of Medical and Dental Sciences, University of Birmingham (UK), B15 2TT.

<sup>5</sup> correspondence to C.B.M. ([C.Mahony@bham.ac.uk](mailto:C.Mahony@bham.ac.uk)) or J.Y.B. ([Julien.Bertrand@unige.ch](mailto:Julien.Bertrand@unige.ch))

Abstract: 207 words

Text: 48844 characters (main text + references + figure legends)

## Key Points

- 1- *hapln1b* encodes a structural protein that is absolutely necessary to HSC emergence
- 2- the extracellular matrix is not only necessary for normal vasculature development, but also for *kitb/kitlgb* signaling in the major vessels, to promote definitive hematopoiesis

## Abstract

During early vertebrate development, hematopoietic stem cells (HSCs) are produced from a hemogenic endothelium located in the dorsal aorta, before they migrate to a transient niche where they expand, the fetal liver and the caudal hematopoietic tissue (CHT), in mammals and zebrafish, respectively. In zebrafish, previous studies have shown that the extracellular matrix (ECM) around the aorta needs to be degraded to allow HSCs to leave the aortic floor and reach blood circulation. However, the role of the ECM components in HSC specification has never been addressed. We show here that *hapln1b*, a key component of the ECM is specifically expressed in hematopoietic sites in the zebrafish embryo. Gain- and loss-of-function experiments all resulted in the absence of HSCs in the early embryo, showing that *hapln1b* is required, at the correct level, to specify HSCs in the hemogenic endothelium. Furthermore, we show that the expression of *hapln1b* is necessary to maintain the integrity of the ECM through its link domain. In addition, by combining functional analyses and computer modelling, we show that *kitlgb* interacts with the ECM, to specify HSCs. Overall, we have demonstrated that the ECM is an integral component of the microenvironment as it mediates specific cytokine signaling that is required for normal HSC specification.

## Introduction

The emergence of blood cells is a highly conserved process that consists of many different waves in distinct anatomical locations. The first emerging cells in zebrafish and mammals consist of primitive myeloid and erythroid cells <sup>1-4</sup>. Definitive hematopoiesis is then initiated by the emergence of the transient erythro-myeloid precursors (EMPs), arising in the yolk sac in mice and humans <sup>1,2</sup>, and in the posterior blood island in zebrafish embryos <sup>5</sup>. Hematopoietic stem cells (HSCs) are then specified in the aorta-gonads-mesonephros (AGM) region where they form small intra-aortic clusters <sup>6-11</sup>. HSCs are derived directly from the aortic hemogenic endothelium (HE), which is specified by the correct balance of several extrinsic factors, such as VEGF, Hedgehog, Notch, BMP and TGF $\beta$  signaling <sup>12-16</sup>. Zebrafish HSC specification occurs between 32-60 hours post fertilization (hpf), from the HE in the dorsal aorta <sup>17,18</sup>. Early HE specification in zebrafish is initiated by the expression of *gata2b* <sup>19</sup> followed by *runx1* <sup>12</sup>. Zebrafish HSC specification is dependent upon inflammatory cytokines produced by neutrophils <sup>20-22</sup>. Macrophage- and vascular-mediated extracellular matrix (ECM) degradation is also necessary to allow HSCs to migrate into the vein <sup>23,24</sup>. HSCs will then migrate to the caudal hematopoietic tissue (CHT), where they interact with endothelial cells and significantly expand their initial number <sup>25,26</sup>.

However, the exact mechanisms controlling HSC emergence from the HE and their expansion in the CHT remain to be fully characterized. We, and others, have previously published that these two processes are highly dependent upon cytokine signalling. In particular, in zebrafish, we showed the important and non-redundant roles of *oncostatin M* and *kit-ligand b(kitlgb)* in this process <sup>25,27</sup>. Proteoglycans, a major component of the ECM, have been previously shown to interact with several hematopoietic cytokines (GM-

CSF, IL-3 and Kitlg) and to maintain a close proximity between stromal cells, HSCs and cytokines in the niche <sup>28-30</sup>. Here, we study the role of *hapln1b*, an ECM-associated protein, in contributing to this process. We focus on *hapln1b*, as it was previously shown to be expressed in the vasculature and in hematopoietic tissues <sup>31</sup>. Furthermore, this gene was shown to be important for correct vascular development of the tail vasculature <sup>31</sup>. In mammals there are *HAPLN1, 2, 3* and *4* genes <sup>32</sup>, whereas in zebrafish there are *hapln1a, 1b, 2, 3* and *4* <sup>31</sup>. *Hapln1* codes for a link protein, required to attach several chondroitin sulfate proteoglycans to the hyaluronic acid (HA) backbone (a ubiquitous glycosaminoglycan (GAG)) to make large, negatively charged, ECM structures <sup>33</sup>. In mammals, *Hapln1* is a secreted ECM protein that stabilizes aggrecan-hyaluronan complexes and is required for correct craniofacial <sup>34</sup> and neocortex development <sup>35</sup>. Loss of *HAPLN1* expression promotes melanoma metastasis but is also required to maintain immune cell motility <sup>36</sup>. *HAPLN1* is also required to maintain lymphatic vessel integrity and reduce endothelial cell permeability preventing visceral metastasis <sup>37</sup>. Knock-Out mouse studies have also revealed a role in maintaining perineural nets (PNNs), a hyaluronan backbone mesh-like network of proteins that surround neurons and regulate neuronal plasticity <sup>38</sup> as well as neural differentiation and development <sup>39-41</sup>. Furthermore, PNNs are responsible for binding chemorepulsive molecules (such as *Semaphorin3a*) <sup>42</sup> as well as transcription factors, such as *Otx2* that is exchanged between different neural cells to enhance cortical plasticity <sup>43</sup>.

Here we show that *hapln1b* is necessary for mediating *kitlgb-kitb* interactions, probably by regulating the ECM stiffness, to induce *runx1* expression in the HE during HSC specification in the zebrafish embryo. Gain- and loss-of-function of *hapln1b* resulted in defective hematopoiesis. Therefore, we conclude that this gene is required, at the correct

dosage, for the specification of HSCs from the HE and their development after their emergence.

## Methods

### Zebrafish strains and husbandry

AB\* (WT), along with transgenic and mutant strains were kept in a 14/10h light/dark cycle at 28°C<sup>44</sup>. We used the following transgenic animals: *lmo2:eGFP<sub>zf71</sub>*<sup>45</sup>, *gata1:DsRed<sub>sd2</sub>*<sup>46</sup>, *kdrl:eGFP<sub>s843</sub>*<sup>47</sup>, *kdrl:Hsa.HRAS-mCherry<sub>s896</sub>*<sup>48</sup>, *cmyb:GFP<sub>zf169</sub>*<sup>49</sup>, *globin:eGFP*<sup>50</sup>, *sox10:mRFP<sub>vu234</sub>*<sup>51</sup>, *mpx:GFP*<sup>52</sup>, *mpeg1:mcherry<sub>gl23</sub>*<sup>53</sup>.

### Whole-mount In Situ hybridization (WISH) staining and analysis

WISH was performed on 4%PFA-fixed embryos at the developmental time points indicated. Digoxigenin labeled probes were synthesized using a RNA Labeling kit (SP6/T7; Roche). RNA probes were generated by linearization of TOPO-TA or ZeroBlunt vectors (Invitrogen) containing the PCR-amplified cDNA sequence. WISH was performed as previously described<sup>54</sup>. Phenotypes were scored by comparing expression with siblings. All injections were repeated three independent times. Analysis was performed using R or GraphPad Prism software. Embryos were imaged in 100% glycerol using an Olympus MVX10 microscope. Oligonucleotide primers used for the production of ISH probes are listed in Table S2.

### Cell sorting and flow cytometry

Zebrafish transgenic embryos (fifteen to twenty per biological replicate) were incubated in 0.5mg/mL Liberase (Roche) solution and shaken for 90min at 33°C, then dissociated, filtered and resuspended in 0.9x PBS and 1% FCS. Dead cells were labeled and excluded

by staining with 5nM SYTOX red (Life Technologies) or 300nM DRAQ7 (Biostatus). Cell sorting was performed using an Aria II (BD Biosciences) or a Biorad S3 (BioRad). Data was analyzed using FlowJo and GraphPad Prism.

### **Quantitative real-time PCR (qPCR) and qPCR analysis**

RNA was extracted using Qiagen RNeasy minikit (Qiagen) and reverse transcribed into cDNA using a Superscript III kit (Invitrogen) or qScript (Quanta Biosciences). qPCR was performed using KAPA SYBR FAST Universal qPCR Kit (KAPA BIOSYSTEMS) and run on a CFX connect Real time system (Bio Rad). All primers are listed in Table S1. Analysis was performed using Microsoft Excel or GraphPad Prism.

### **Synthesis of full-length mRNA and microinjection**

PCR primers to amplify cDNA of interest are listed in Table S2. *Kitlgb* and *kitlga* mRNA was synthesized and injected as previously described<sup>25,27</sup>. mRNA was reverse transcribed using mMessage mMachine kit SP6 (Ambion) from a linearized pCS2+ vector containing PCR products. Following transcription, RNA was purified by phenol-chloroform extraction. *Hapln1b* mRNA was injected at 200pg, unless otherwise stated.

### **Imaging**

WISH were imaged using an Olympus MVX10 microscope in 100% glycerol. Fluorescent images were taken with an Olympus IX83 microscope and processed using cell sense dimension software. All images were processed using Adobe Photoshop. Time-lapse imaging was obtained using a Nikon inverted A1r spectral confocal microscope and processed using Fiji and analysed using GraphPad Prism.

## Morpholinos

All morpholinos oligonucleotides (MOs) were purchased from Gene Tools and listed in Table S3. MO efficiency was tested using PCR from total RNA extracted from a pool of 8-10 embryos at 24hpf. *Hapln1b* full-length primers were used to test MO efficiency and are listed in Table S2. *Hapln1b* morpholino was injected at 3ng, unless otherwise stated. All MOs used in this study are splice blocking MOs.

## Structural modelling and electrostatic properties of *kitlga* and *kitlgb*

Structural models for zebrafish *kitlga* and *kitlgb* were constructed based on a consensus approach employing homology-modeling, template-based structure modelling, and *ab initio* structure prediction as implemented in i-TASSER, Phyre2, and RaptorX<sup>55-57</sup>. Vacuum electrostatic potential calculations were performed and displayed with the built-in module in the program PyMOL v. 2.3 (<https://pymol.org/2/>). Isoelectric point (pI) calculations were performed via the prot-pI server (<https://www.protpi.ch>).

## Glycosaminoglycan staining

Embryos were fixed in 4% formol (Biosystems) and embedded in paraffin (Haslab). 3µm sections were stained with 10mg/ml Alcian blue (Sigma) and counterstained with Mayer's hemalun (Merck) to mark glycosaminoglycans.

## Results

### *hapln1b* is specifically expressed in early embryonic hematopoietic tissues

Previous studies have shown that *hapln1b* was expressed in the vasculature at 24-28hpf, and that its loss-of-function resulted in abnormal angiogenesis<sup>31</sup>. However, its role during embryonic haematopoiesis has never been investigated. In zebrafish, *hapln1b* is the only *hapln* family member to display a vascular and hematopoietic expression pattern (data not shown), despite retaining high sequence identity and conserved peptides capable of forming disulphide bonds (Figure S1A, B). Synteny, phylogeny and sequence identity analyses revealed that *hapln1a* and *hapln1b* originate from a duplication of the *HAPLN1* gene in mammals (Figure S2A, B). Even across these species, peptides capable of forming disulphide bonds within the link domain have been conserved (Figure S2C). We therefore focused our study upon *hapln1b* and first examined its expression pattern. We first established that *hapln1b* is initially expressed from around 12hpf and would therefore not be derived from maternal RNA<sup>58</sup> (Figure S3A). *hapln1b* was expressed between 20hpf and 26hpf along the aorta, in the developing CHT and in the hypochord, as previously described<sup>31</sup>. Further analyses at 26hpf revealed that *hapln1b* was also expressed in the aorta and vein region, ventral to the notochord (Figure 1Ai). Between 30 and 36hpf the expression begins to decrease, becoming more localized to the CHT, before being completely restricted to the marginal fold at 48hpf (Figure 1A). Expression was also scored in the cardiac precursors at 36hpf and in possible cranial cartilaginous structures (Figure 1Aii, Aiii). By 4 and 5dpf expression was restricted to cranial structures and was absent from hematopoietic tissue (Figure 1B).

We then further analyzed the expression of *hapln1b* in different cell populations. We sorted endothelial cells from dissected *kdrl:eGFP* embryos at 26hpf and 48hpf (Figure 1C). As we found by wholemount in situ hybridisation (WISH), *hapln1b* was enriched in endothelial cells extracted from whole embryos at 26hpf (Figure 1D) but no enrichment



was noted at 48hpf (Figure 1E). However, *hapln1b* was highly enriched in early EMPs (*lmo2:GFP;gata1:DsRed* double-positive cells at 28hpf <sup>5</sup>) and nascent HSCs (*kdrl:mCherry;cmyb:GFP* double-positive cells at 36hpf <sup>17</sup>), compared to whole embryos at 28hpf (Figure 1F), which showed a potential link between *hapln1b* and embryonic definitive haematopoiesis. Amino acid structural analysis (using SMART, <http://smart.embl.de/>) revealed that *hapln1b* contains an IG domain and two hyaluronic acid (HA) link domains (Figure 1F). We next investigated how this ECM protein could interfere with developmental hematopoiesis.

### ***hapln1b* is required to specify HSCs from the Hemogenic Endothelium**

To further investigate the role of this gene in hematopoiesis, we injected a splice blocking morpholino (MO) for *hapln1b* which efficiently reduced mRNA levels (Figure S3B, C). Inhibiting *hapln1b* expression did not affect *gata2b* expression, the earliest known marker of HE specification <sup>19</sup> (Figure 2A). However, *runx1* and *cmyb* expression was robustly decreased in the aortic region (Figure 2A). Accordingly, additional downstream expression of *cmyb* at 4dpf in the CHT and *rag1* at 4.5dpf in the thymus were also decreased (Figure 2A, B, and C). To further validate the specificity of our MO, we then rescued the loss of *runx1* in *hapln1b* morphants by co-injecting with *hapln1b* mRNA (Figure S3D). *hapln1b*-morphants displayed normal HSC specification, although in a number of cases the formation of the CHT was perturbed as seen in the “mild” cases, as previously described <sup>31</sup> (Figure 2D). In a small number of embryos, the vasculature was severely affected as represented in the “severe” phenotype (Figure 2D). We observed no change in the emergence of primitive macrophages or red blood cells as marked by *mfap4* and *gata1* at 24hpf (Figure 2E), respectively. We did however note a decrease in primitive

circulating neutrophils, as marked by *mpx* (Figure 2E), although neutrophils were still present on the yolk sac. This change in distribution may be due to a loss of function of the vasculature that we observed in Figure 2D. Loss of *hapln1b* thus results into a specific loss of HSCs but does not impact primitive haematopoiesis. We then investigated the effect of *hapln1b* overexpression on embryonic haematopoiesis.

### ***hapln1b* overexpression is sufficient to reduce HSC emergence and downstream survival**

We next overexpressed *hapln1b* by injecting the full-length mRNA at the one-cell stage and analysed the effect on developmental haematopoiesis. *hapln1b* overexpression did not change vessel development or early HSC specification as marked by *kdrl* at 24hpf and *runx1* at 28hpf, respectively (Figure 3A). However, we noted a decrease in *cmyb* signal at 36hpf in the AGM region (Figure 3A). *hapln1b* overexpression also decreased *cmyb* at 48hpf in the CHT, suggesting that newly formed HSCs did not colonize this tissue. However, in the AGM, the *cmyb* signal was also absent, as in non-injected controls, indicating that HSCs were not lodged and were not unable to migrate from the aorta (Figure 3B). Accordingly, *rag1* staining in the thymus at 4.5dpf and *cmyb* in the CHT at 4dpf were also decreased (Figure 3C). To further analyse this loss of *cmyb* signal, we then imaged double-positive *kdrl*:mCherry;*cmyb*:eGFP embryos to examine HSC emergence from the HE<sup>17</sup>. *hapln1b* overexpression significantly reduced the number of double positive nascent HSCs from the AGM region at 36 and 48hpf (Figure 3D, D'). As expected, the number of HSCs present in the CHT niche at 3 and 4dpf was also significantly reduced (Figure 3E, E'). To further examine this loss of *cmyb* at 36hpf from the HE we used time-lapse confocal microscopy to image double-positive *kdrl*:mCherry; *cmyb*:eGFP embryos

to examine the endothelial-to-hematopoietic transition in more detail in the aorta. We imaged from 34hpf to 42hpf to observe HSC budding from the HE. We observed clusters of cells in the AGM region of control embryos, preparing to bud and enter circulation (Movie S1). However, in *hapln1b*-overexpressing embryos, almost no cells seemed to initiate budding. Occasionally, some *cmyb:GFP+* cells were detected, but none of them ever underwent endothelial-to-hematopoietic transition (Movie S2). We therefore conclude that *hapln1b* overexpression is sufficient to prevent the correct HSC budding from the HE. We next investigated how this could be occurring by examining how *hapln1b* could alter the ECM.

### ***hapln1b* modulates the density of the ECM around vasculature**

The degradation of ECM by primitive macrophages is an important step for HSC emigration out of the aortic region<sup>23</sup>. Therefore, we sought to confirm that *hapln1b* expression levels are capable of modulating the ECM directly. To achieve this, we examined glycosaminoglycans in the AGM and CHT region by staining embryo sections with alcian blue. Non-injected controls at 30hpf showed a high concentration of ECM surrounding the notochord, but little or no ECM around the artery and vein in the AGM (Figure 4A). *hapln1b*-morphants showed less ECM staining, resulting in a change of the shape of the notochord, aorta and vein in the AGM, suggesting that the integrity of the ECM was altered (Figure 4A). In contrast, *hapln1b* mRNA overexpression resulted in increased ECM staining between the artery and vein and around vessels in the AGM (Figure 4A). Concerning the CHT, the artery and the venous plexus were clearly visible in controls morphants (Figure 4B). However, in *hapln1b*-morphants, the venous plexus was highly disrupted with no clear boundary between the artery and vein (Figure 4B).

*hapln1b* mRNA overexpression resulted in a slightly distorted CHT with increased alcian blue staining and densely packed venous plexus (Figure 4B). All these results correlate with the importance of *hapln1b* in coordinating angiogenesis, as described previously<sup>31</sup> and could explain part of our results concerning developing haematopoiesis. We next examined the role that *hapln1b* plays in cytokine signalling to control this process.

### ***hapln1b* mediates *kitlgb-kitb* interactions that are required for proper HSC specification from the HE**

We next attempted to decipher the mechanism by which *hapln1b* can affect haematopoiesis. Our previous studies indicated that HSCs express *kitb* and that *kiltgb* and *kitb* are required for *runx1* expression in the HE at 28hpf<sup>27</sup>. As we found a similar decrease in *runx1* in *hapln1b*-morphants, we investigated a possible link between *kiltgb* signalling and *hapln1b*. Previous studies have indicated that *kitlgb* is expressed in the CHT region by ISH at 24hpf and is likely to exist more commonly as the membrane bound form, due to a loss of one of the two cleavage sites<sup>59</sup>. Whereas *kitlga* would represent the more soluble form of this ligand as it retains the two cleavage sites<sup>59</sup>. We also investigated the expression of *kiltgb* in many different FACS sorted cells and found no enrichment in ICM cells sorted from *globin:GFP* cells at 20hpf, primitive macrophages sorted from *mpeg1:mcherry* embryos at 24hpf, primitive neutrophils sorted from *mpx:eGFP* embryos at 24hpf and neural crest cells sorted from *sox10:mRFP* embryos at 26hpf (Figure S4 A-D). We also found no enrichment of *kitlgb* in endothelial cells sorted from dissected embryos at 19-20hpf (Figure S4E). However, we found a significant enrichment in tail endothelial cells at 26hpf (Figure S4F), corroborating with the previously published ISH expression pattern of *kitlgb*<sup>59</sup>. We therefore concluded that

*kiltgb* would be present in the AGM at low concentrations and would require an additional element to mediate effective interaction with its receptor to initiate *runx1* expression. We tested this hypothesis by attempting to rescue the loss of *runx1* observed in *hapln1b*-morphants and the loss of *cmyb* observed in *hapln1b*-overexpressing embryos by injecting *kitlgb* mRNA.

To verify the specificity of our *kitlgb* injection, we also injected *kitlga* mRNA, which does not play any role in HSC specification and development<sup>27</sup>. Injection of either *kitlga* or *kitlgb* alone, did not alter expression of *runx1* at 28hpf, as anticipated (Figure 5A, A'). Injection of the *hapln1b*-MO gave a similar loss of *runx1* at 28hpf (Figure 5A, A'). We then co-injected the *hapln1b*-MO along with *kitlgb* mRNA, which in 16/34 embryos resulted in a rescue of the *runx1* signal (Figure 5A, A'). As expected, *kitlga* could not rescue the loss of *runx1* in *hapln1b*-morphants (Figure 5A, A'). We then attempted to rescue the decrease of *cmyb* signal in the *hapln1b*-overexpressing embryos by injecting *kitlgb* mRNA. Compared to non-injected controls, *hapln1b* mRNA reduced *cmyb* expression at 36hpf as observed earlier (Figure 5B, B'). Whereas *kitlgb* injection alone, did not alter the *cmyb* expression at 36hpf (Figure 5B, B'), co-injection of both *hapln1b* and *kitlgb* mRNAs increased *cmyb* expression to control levels (Figure 5B, B'). Altogether, this data suggests that the dose of *hapln1b* controls the stiffness of the ECM, therefore controlling the accessibility of *kitlgb* to the HE.

As proteoglycans are negatively charged, we sought to examine the possible electrostatic compatibility of *kitlgb* with such a biological environment. In the absence of an experimentally determined structure for zebrafish *kitlgb* and *kitlga*, we leveraged available structural information at high resolution on human KITL (SCF)<sup>60,61</sup> and the

related hematopoietic cytokines Flt3 ligand and colony stimulating factor 1<sup>62-64</sup>, to derive reliable models based on homology-based approaches and *ab initio* structure prediction. Our analysis reveals that *Kitlgb* would be expected to display an overall positive electrostatic potential as manifested by a Z-potential value of +1.2 at physiological pH and as illustrated by extensive patches of positively charged amino acids at the protein surface (Figure 5C). Such physicochemical properties are in sharp contrast to zebrafish *Kitlga* and human *KITL*, which are pronouncedly acidic with very similar electrostatic properties characterized by strongly negative electrostatic Z-potentials at physiological pH. As proteoglycans are negatively charged, the overall positively charged *Kitlgb*, but not the negatively charged *Kitlga*, would mediate favourable interactions with the ECM in hemogenic and hematopoietic tissues, therefore activating the *Kitb* receptor, consistent with our findings.

### **The link domain is necessary for *hapln1b* functions**

Finally, we wanted to verify that *hapln1b* was interacting with HA by its link domain, to maintain ECM integrity. We cloned a truncated version of the *hapln1b* mRNA lacking exon 4 and 5 (*hapln1b* $\Delta$ 4 $\Delta$ 5) that encode the HA link domain (Figure 6A). Compared to non-injected controls, *hapln1b* $\Delta$ 4 $\Delta$ 5 mRNA injected did not reduce *cmyb* levels at 36hpf whereas wild-type *hapln1b* mRNA did (Figure 6B, B'), as shown earlier (Figure 3A). We therefore conclude that the link domain is necessary for *hapln1b* to exert its function, which is to assemble the ECM. We therefore propose the following model, where *kitlgb* is probably distributed along the aorta, by interacting with the ECM. The ECM structure is maintained by proteoglycans forming links with the HA backbone (mediated by the link domain of *hapln1b*), allowing *kiltgb* to reach the AGM. A loss of *hapln1b* will prevent HA

links to be formed and will prevent *kitlgb* distribution to the AGM, reducing *runx1* expression (Figure 6C). When *hapln1b* is overexpressed, HSC budding is reduced as the HA links are increased, resulting in a tight ECM that would also reduce *kitlgb* distribution to the AGM. Therefore, a lower concentration of *kitlgb* is present in the AGM would result in fewer HSCs budding and developing (Figure 6C).

## Discussion

We have shown that *hapln1b* is expressed along the embryonic endothelium prior to HSC emergence, before it becomes restricted to non-hematopoietic tissue at later stages (Figure 1). *hapln1b* is required, in the correct concentration, to specify HSCs from the HE and to maintain HSC budding and release into circulation. We found that loss of *hapln1b* results in a loss of *runx1* and loss of downstream HSC specification markers (Figure 2A-C). As proteoglycans can bind cytokines, it is possible that *hapln1b* is responsible for maintaining HA linking to proteoglycans to favor cytokine-receptor interaction. Our previous data has shown that *kitlgb-kitb* signalling is required for maintaining *runx1* expression and HSC specification in the HE <sup>27</sup>. However, we, and others <sup>59</sup>, have demonstrated that *kitlgb* is highly expressed in the posterior blood island but not the aortic HE region. When we modulate *hapln1b* expression, we are able to change the structure of the ECM as the loss of *hapln1b* affects the ECM/GAG deposition around axial vasculature, as well as in the CHT vessels and affects the overall CHT structure (Figure 4). The disruption of the CHT structure following *hapln1b* knockdown was also noted in a previous study <sup>31</sup>. This, coupled with the fact that we are able to rescue *hapln1b*

morphants with *kitlgb* mRNA injections (Figure 5A), lead us to conclude that *hapln1b* mediates HA linking with proteoglycans and is responsible for stabilising the ECM scaffold and distributing *kitlgb* to the HE (Figure 6C). Our data shows that the ECM likely plays a crucial role in maintaining the signalling environment in close proximity to the AGM to mediate HSC specification from the HE. Furthermore, we show an involvement of the ECM in the CHT, further suggesting that the ECM is a key player in maintaining the embryonic niche to help expanding HSCs. Although we have highlighted *kiltgb* as a key cytokine that interacts with the ECM, there are likely many others that contribute to the extrinsic signalling microenvironment. Fully characterizing and understanding these molecules will be an important step to improve the currently challenging derivation of HSCs from iPSCs.

In contrast to our loss-of-function assays, overexpression of *hapln1b* resulted in normal HSC specification as indicated by normal *runx1* expression (Figure 3A), but a reduction in HSC emergence and later expansion within hematopoietic tissue, as indicated by a reduction of *cmyb* expression (Figure 3B). As we found that HSC emergence was reduced, but never completely ablated, we reasoned that overexpression of *hapln1b* is likely to result in a denser ECM, reducing the accessibility of *kitlgb* to the hemogenic microenvironment. We confirmed this by rescuing HSC expansion in *hapln1b*-overexpressing embryos by overexpressing *kiltgb* mRNA (Figure 5B). Recent mouse studies have indicated that *Kitlg* is required throughout HSC development in the AGM region <sup>65</sup>. This was further corroborated by additional studies showing that *c-kit* is expressed by proliferating HSCs in the AGM <sup>66</sup> and that *Kitlg* maintains HSCs in mouse AGM cultures <sup>67</sup>. It is therefore likely that reducing the concentration of *kitlgb* in our *hapln1b*-overexpressing embryos restricts HSC emergence and decreases downstream survival and HSC maturation (Figure 6C). It is also possible that the ECM is too dense to



allow HSC to bud and migrate to the CHT, as previous studies have indicated that ECM degradation is essential to HSC to enter blood circulation <sup>23,24</sup>.

A similar role for *Hapln1* was also demonstrated in the PNNs in the mouse central nervous system, as it is required to maintain the HA mesh, surrounding neurons <sup>38</sup>. These PNNs bind molecules to maintain neuronal plasticity and development, further implicating this gene in maintaining the ECM. Recent studies have identified that HSCs migrate to the fetal niche, the CHT, and expand their initial number in response to a number of cytokines <sup>24-26</sup>. It is possible that many cytokine gradients and concentrations are maintained in the fetal niche by ECM components, such as *hapln1b*. Understanding these in more details will allow to fully characterizing HSC expansion providing alternative methods to improve HSC expansion *ex vivo*.

## **Acknowledgements**

J.Y.B. was endorsed by a Chair in Life Sciences funded by the Gabriella Giorgi-Cavaglieri Foundation and is also funded by the Swiss National Fund (310030\_184814) and the “Fondation Privée des HUG”.

## **Author contributions**

SNS performed structural analysis of *KITL*, *kitlga* and *kitlgb*. VB and AA performed Alcian blue stainings on embryo sections. CBM and CP performed all other experiments. CBM and JYB designed the research and wrote the manuscript. SNS and AA edited the manuscript.

## Conflict of interests

The authors declare no conflict of interest

## References

1. Bertrand JY, Jalil A, Klaine M, Jung S, Cumano A, Godin I. Three pathways to mature macrophages in the early mouse yolk sac. *Blood*. 2005;106(9):3004-3011.
2. McGrath KE, Frame JM, Fegan KH, et al. Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo. *Cell Rep*. 2015;11(12):1892-1904.
3. Palis J. Primitive and definitive erythropoiesis in mammals. *Front Physiol*. 2014;5:3.
4. Palis J, Chan RJ, Koniski A, Patel R, Starr M, Yoder MC. Spatial and temporal emergence of high proliferative potential hematopoietic precursors during murine embryogenesis. *Proc Natl Acad Sci U S A*. 2001;98(8):4528-4533.
5. Bertrand JY, Kim AD, Violette EP, Stachura DL, Cisson JL, Traver D. Definitive hematopoiesis initiates through a committed erythromyeloid progenitor in the zebrafish embryo. *Development*. 2007;134(23):4147-4156.
6. Zovein AC, Hofmann JJ, Lynch M, et al. Fate tracing reveals the endothelial origin of hematopoietic stem cells. *Cell Stem Cell*. 2008;3(6):625-636.
7. Chen MJ, Yokomizo T, Zeigler BM, Dzierzak E, Speck NA. Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. *Nature*. 2009;457(7231):887-891.
8. Boisset JC, van Cappellen W, Andrieu-Soler C, Galjart N, Dzierzak E, Robin C. In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. *Nature*. 2010;464(7285):116-120.
9. Bollerot K, Pouget C, Jaffredo T. The embryonic origins of hematopoietic stem cells: a tale of hemangioblast and hemogenic endothelium. *Apmis*. 2005;113(11-12):790-803.

10. Bollerot K, Romero S, Dunon D, Jaffredo T. Core binding factor in the early avian embryo: cloning of Cbfbeta and combinatorial expression patterns with Runx1. *Gene Expr Patterns*. 2005;6(1):29-39.
11. Jaffredo T, Gautier R, Brajeul V, Dieterlen-Lievre F. Tracing the progeny of the aortic hemangioblast in the avian embryo. *Dev Biol*. 2000;224(2):204-214.
12. Burns CE, Traver D, Mayhall E, Shepard JL, Zon LI. Hematopoietic stem cell fate is established by the Notch-Runx pathway. *Genes Dev*. 2005;19(19):2331-2342.
13. Wilkinson RN, Pouget C, Gering M, et al. Hedgehog and Bmp polarize hematopoietic stem cell emergence in the zebrafish dorsal aorta. *Dev Cell*. 2009;16(6):909-916.
14. Monteiro R, Pinheiro P, Joseph N, et al. Transforming Growth Factor beta Drives Hemogenic Endothelium Programming and the Transition to Hematopoietic Stem Cells. *Dev Cell*. 2016;38(4):358-370.
15. Kim AD, Melick CH, Clements WK, et al. Discrete Notch signaling requirements in the specification of hematopoietic stem cells. *Embo j*. 2014;33(20):2363-2373.
16. Gering M, Patient R. Hedgehog signaling is required for adult blood stem cell formation in zebrafish embryos. *Dev Cell*. 2005;8(3):389-400.
17. Bertrand JY, Chi NC, Santoso B, Teng S, Stainier DY, Traver D. Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature*. 2010;464(7285):108-111.
18. Kissa K, Herbomel P. Blood stem cells emerge from aortic endothelium by a novel type of cell transition. *Nature*. 2010;464(7285):112-115.
19. Butko E, Distel M, Pouget C, et al. Gata2b is a restricted early regulator of hemogenic endothelium in the zebrafish embryo. *Development*. 2015;142(6):1050-1061.
20. Espin-Palazon R, Stachura DL, Campbell CA, et al. Proinflammatory signaling regulates hematopoietic stem cell emergence. *Cell*. 2014;159(5):1070-1085.

21. Sawamiphak S, Kontarakis Z, Stainier DY. Interferon gamma signaling positively regulates hematopoietic stem cell emergence. *Dev Cell*. 2014;31(5):640-653.
22. Li Y, Esain V, Teng L, et al. Inflammatory signaling regulates embryonic hematopoietic stem and progenitor cell production. *Genes Dev*. 2014;28(23):2597-2612.
23. Travnickova J, Tran Chau V, Julien E, et al. Primitive macrophages control HSPC mobilization and definitive haematopoiesis. *Nat Commun*. 2015;6:6227.
24. Theodore LN, Hagedorn EJ, Cortes M, et al. Distinct Roles for Matrix Metalloproteinases 2 and 9 in Embryonic Hematopoietic Stem Cell Emergence, Migration, and Niche Colonization. *Stem Cell Reports*. 2017;8(5):1226-1241.
25. Mahony CB, Fish RJ, Pasche C, Bertrand JY. tfec controls the hematopoietic stem cell vascular niche during zebrafish embryogenesis. *Blood*. 2016;128(10):1336-1345.
26. Tamplin OJ, Durand EM, Carr LA, et al. Hematopoietic stem cell arrival triggers dynamic remodeling of the perivascular niche. *Cell*. 2015;160(1-2):241-252.
27. Mahony CB, Pasche C, Bertrand JY. Oncostatin M and Kit-Ligand Control Hematopoietic Stem Cell Fate during Zebrafish Embryogenesis. *Stem Cell Reports*. 2018.
28. Roberts R, Gallagher J, Spooncer E, Allen TD, Bloomfield F, Dexter TM. Heparan sulphate bound growth factors: a mechanism for stromal cell mediated haemopoiesis. *Nature*. 1988;332(6162):376-378.
29. Bruno E, Luikart SD, Long MW, Hoffman R. Marrow-derived heparan sulfate proteoglycan mediates the adhesion of hematopoietic progenitor cells to cytokines. *Exp Hematol*. 1995;23(11):1212-1217.
30. Gupta P, Oegema TR, Jr., Brazil JJ, Dudek AZ, Slungaard A, Verfaillie CM. Structurally specific heparan sulfates support primitive human hematopoiesis by formation of a multimolecular stem cell niche. *Blood*. 1998;92(12):4641-4651.
31. Gomez GA, Veldman MB, Zhao Y, Burgess S, Lin S. Discovery and characterization of novel vascular and hematopoietic genes downstream of etsrp in zebrafish. *PLoS One*. 2009;4(3):e4994.

32. Spicer AP, Joo A, Bowling RA, Jr. A hyaluronan binding link protein gene family whose members are physically linked adjacent to chondroitin sulfate proteoglycan core protein genes: the missing links. *J Biol Chem*. 2003;278(23):21083-21091.
33. Day AJ, Prestwich GD. Hyaluronan-binding proteins: tying up the giant. *J Biol Chem*. 2002;277(7):4585-4588.
34. Watanabe H, Yamada Y. Mice lacking link protein develop dwarfism and craniofacial abnormalities. *Nat Genet*. 1999;21(2):225-229.
35. Long KR, Newland B, Florio M, et al. Extracellular Matrix Components HAPLN1, Lumican, and Collagen I Cause Hyaluronic Acid-Dependent Folding of the Developing Human Neocortex. *Neuron*. 2018;99(4):702-719.e707.
36. Kaur A, Ecker BL, Douglass SM, et al. Remodeling of the collagen matrix in aging skin promotes melanoma metastasis and affects immune cell motility. *Cancer Discov*. 2018.
37. Ecker BL, Kaur A, Douglass SM, et al. Age-Related Changes in HAPLN1 Increase Lymphatic Permeability and Affect Routes of Melanoma Metastasis. *Cancer Discov*. 2018.
38. Foscarin S, Ponchione D, Pajaj E, et al. Experience-dependent plasticity and modulation of growth regulatory molecules at central synapses. *PLoS One*. 2011;6(1):e16666.
39. Eskici NF, Erdem-Ozdamar S, Dayangac-Erden D. The altered expression of perineuronal net elements during neural differentiation. *Cell Mol Biol Lett*. 2018;23:5.
40. Kecskes S, Gaal B, Racz E, Birinyi A, Hunyadi A, Matesz C. Extracellular matrix molecules exhibit unique expression pattern in the climbing fiber-generating precerebellar nucleus, the inferior olive. *Neuroscience*. 2015;284:412-421.
41. Gaal B, Kecskes S, Matesz C, Birinyi A, Hunyadi A, Racz E. Molecular composition and expression pattern of the extracellular matrix in a mossy fiber-generating precerebellar nucleus of rat, the prepositus hypoglossi. *Neurosci Lett*. 2015;594:122-126.
42. Dick G, Tan CL, Alves JN, et al. Semaphorin 3A binds to the perineuronal nets via chondroitin sulfate type E motifs in rodent brains. *J Biol Chem*. 2013;288(38):27384-27395.

43. Beurdeley M, Spatazza J, Lee HH, et al. Otx2 binding to perineuronal nets persistently regulates plasticity in the mature visual cortex. *J Neurosci*. 2012;32(27):9429-9437.
44. Westerfield M. The zebrafish book: A guide for the laboratory use of zebrafish (*Brachydanio rerio*). *University of Oregon Press*. 1994(2.1).
45. Zhu H, Traver D, Davidson AJ, et al. Regulation of the *lmo2* promoter during hematopoietic and vascular development in zebrafish. *Dev Biol*. 2005;281(2):256-269.
46. Traver D, Paw BH, Poss KD, Penberthy WT, Lin S, Zon LI. Transplantation and in vivo imaging of multilineage engraftment in zebrafish bloodless mutants. *Nat Immunol*. 2003;4(12):1238-1246.
47. Jin SW, Beis D, Mitchell T, Chen JN, Stainier DY. Cellular and molecular analyses of vascular tube and lumen formation in zebrafish. *Development*. 2005;132(23):5199-5209.
48. Chi NC, Shaw RM, De Val S, et al. *Foxn4* directly regulates *tbx2b* expression and atrioventricular canal formation. *Genes Dev*. 2008;22(6):734-739.
49. North TE, Goessling W, Walkley CR, et al. Prostaglandin E2 regulates vertebrate haematopoietic stem cell homeostasis. *Nature*. 2007;447(7147):1007-1011.
50. Ganis JJ, Hsia N, Trompouki E, et al. Zebrafish globin switching occurs in two developmental stages and is controlled by the LCR. *Dev Biol*. 2012;366(2):185-194.
51. Kirby BB, Takada N, Latimer AJ, et al. In vivo time-lapse imaging shows dynamic oligodendrocyte progenitor behavior during zebrafish development. *Nat Neurosci*. 2006;9(12):1506-1511.
52. Renshaw SA, Loynes CA, Trushell DM, Elworthy S, Ingham PW, Whyte MK. A transgenic zebrafish model of neutrophilic inflammation. *Blood*. 2006;108(13):3976-3978.
53. Ellett F, Pase L, Hayman JW, Andrianopoulos A, Lieschke GJ. *mpeg1* promoter transgenes direct macrophage-lineage expression in zebrafish. *Blood*. 2011;117(4):e49-56.
54. Thisse C, Thisse B, Schilling TF, Postlethwait JH. Structure of the zebrafish *snail1* gene and its expression in wild-type, spadetail and no tail mutant embryos. *Development*. 1993;119(4):1203-1215.

55. Kallberg M, Wang H, Wang S, et al. Template-based protein structure modeling using the RaptorX web server. *Nat Protoc.* 2012;7(8):1511-1522.
56. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. *Nat Methods.* 2015;12(1):7-8.
57. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc.* 2015;10(6):845-858.
58. Tadros W, Lipshitz HD. The maternal-to-zygotic transition: a play in two acts. *Development.* 2009;136(18):3033-3042.
59. Hultman KA, Bahary N, Zon LI, Johnson SL. Gene Duplication of the zebrafish kit ligand and partitioning of melanocyte development functions to kit ligand a. *PLoS Genet.* 2007;3(1):e17.
60. Yuzawa S, Opatowsky Y, Zhang Z, Mandiyan V, Lax I, Schlessinger J. Structural basis for activation of the receptor tyrosine kinase KIT by stem cell factor. *Cell.* 2007;130(2):323-334.
61. Zhang Z, Zhang R, Joachimiak A, Schlessinger J, Kong XP. Crystal structure of human stem cell factor: implication for stem cell factor receptor dimerization and activation. *Proc Natl Acad Sci U S A.* 2000;97(14):7732-7737.
62. Savvides SN, Boone T, Andrew Karplus P. Flt3 ligand structure and unexpected commonalities of helical bundles and cystine knots. *Nat Struct Biol.* 2000;7(6):486-491.
63. Verstraete K, Vandriessche G, Januar M, et al. Structural insights into the extracellular assembly of the hematopoietic Flt3 signaling complex. *Blood.* 2011;118(1):60-68.
64. Elegheert J, Bracke N, Pouliot P, et al. Allosteric competitive inactivation of hematopoietic CSF-1 signaling by the viral decoy receptor BARF1. *Nat Struct Mol Biol.* 2012;19(9):938-947.
65. Azzoni E, Frontera V, McGrath KE, et al. Kit ligand has a critical role in mouse yolk sac and aorta-gonad-mesonephros hematopoiesis. *EMBO Rep.* 2018;19(10).
66. Batsivari A, Rybtsov S, Souilhol C, et al. Understanding Hematopoietic Stem Cell Development through Functional Correlation of Their Proliferative Status with the Intra-aortic Cluster Architecture. *Stem Cell Reports.* 2017;8(6):1549-1562.

67. Rybtsov S, Batsivari A, Bilotkach K, et al. Tracing the origin of the HSC hierarchy reveals an SCF-dependent, IL-3-independent CD43(-) embryonic precursor. *Stem Cell Reports*. 2014;3(3):489-501.

## Figure legends

### Figure 1. *hapln1b* is expressed by vascular cells and haematopoietic progenitors

(A) WMISH of *hapln1b* expression from 20hpf to 48hpf. (Ai) Section at 26hpf, displaying aorta (a) and vein (v) and notochord (n). (Aii and iii) ventral view at 36hpf and 48hpf. (B) WMISH of *hapln1b* expression at 4 and 5dpf. (Biv and v) Dorsal view at 4 and 5dpf. (C-E) Experimental outline and qPCR expression of *hapln1b* from FACS sorted endothelial cells. (F) qPCR expression of *hapln1b* in FACS sorted haematopoietic progenitors (EMPs and HSCs). All qPCR data is from biological triplicates. In D, two-tailed Student's t-test, whole GFP- and GFP+,  $p=0.0035$ , heads and tails  $p=0.75$ . In E and F, analysis was performed by a one way ANOVA with multiple comparisons. In E, whole GFP- and GFP+,  $p=0.99$ , whole GFP- and heads,  $p=0.44$ , whole GFP- and tails,  $p=0.99$ . In F, whole and EMPs:  $p=0.0081$ , whole and HSCs:  $p=0.0449$ . (F) SMART (<http://smart.embl.de/>) prediction of *hapln1b* protein structure. IG: immunoglobulin like-domain, Link: HA Link domain, aa: amino acids.

### Figure 2. *hapln1b* is necessary for the specification of HSC from the hemogenic endothelium

(A) *gata2b*, *runx1* and *cmyb* ISH in control MO, or *hapln1b* MO injected embryos (injected at 3ng throughout, unless stated). (B) ISH expression of *cmyb*, in control MO, or *hapln1b* MO injected embryos. (C) *rag1* expression in the thymus after control MO or



*hapln1b* MO injection. (D) ISH expression of *kdrl* in control MO, or *hapln1b* MO injected embryos. *Hapln1b* morphants displayed three phenotypes: normal, mild and sever. (E) ISH expression of *mfap4*, *mpx* and *gata1* in control MO, or *hapln1b* MO injected embryos. H1b: *hapln1b*.

### **Figure 3. *hapln1b* overexpression is sufficient to inhibit HSC budding and development**

(A) ISH expression of *kdrl*, *runx1* and *cmyb* in non-injected, or *hapln1b* mRNA injected embryos. (B, C) ISH expression of *cmyb* in non-injected, or *hapln1b* mRNA injected embryos. (D, D') Imaging double transgenic *kdrl:mCherry/cmyb:GFP* embryos at 36 and 48hpf. Scale, 50 $\mu$ m. In D', between NI and *+h1b* injected at 36hpf,  $p < 0.0001$  and at 48hpf  $p = 0.0003$ . (E, E') Imaging double transgenic *kdrl:mCherry/cmyb:GFP* embryos at 3dpf and 4dpf. Scale, 50 $\mu$ m. In E', between NI and *+h1b* injected at 3dpf,  $p = 0.0003$  and at 4dpf  $p = 0.004$ . NI: non-injected, *+h1b*: *hapln1b* mRNA injected.

### **Figure 4. *hapln1b* is principally required for maintaining ECM integrity in the CHT**

(A) Alcian blue and hemalun staining of glycosaminoglycans in the AGM region at 30hpf in NI. Black arrows indicate areas of ECM staining. (B) Alcian blue and hemalun staining of glycosaminoglycans in the CHT region at 30hpf. A= aorta, V=vein, N=Notochord. Scale bars, 10 $\mu$ m.

### **Figure 5. *hapln1b* interacts with *kitlgb* to maintain *kitlgb-kitb* interactions in the AGM**

(A) ISH expression pattern of *runx1* in control morphants, *kitlga* mRNA injected embryos, *kitlgb* mRNA injected embryos, *hapln1b* morphants, embryos injected with both *hapln1b* MO and *kitlga* mRNA and embryos injected with both *hapln1b* MO and *kitlgb* mRNA. (A') analysis of *runx1* expression. All analysis is compared to control MO. Data was analysed by fishers exact test using R. ctl MO vs +*kitlga*: p=0.05826, NI vs +*kitlgb*: p=0.7939, NI vs +*hapln1b*: p<0.0001, NI vs +*hapln1b+kitlga*: p<0.0001, NI vs +*hapln1b+kitgb*: p<0.0001.

(B) ISH expression of *cmyb* in NI control embryos, *hapln1b* mRNA injected embryos, *kitlgb* mRNA injected embryos and embryos injected with both *hapln1b* and *kitlgb*. (B') analysis of *cmyb* expression. All analysis is compared to NI control. Data was analysed by fishers exact test using R. NI vs +*kitlga*: p=0.8373, NI vs +*kitlgb*: p=1, NI vs +*hapln1b*: p=0.0051, NI vs +*hapln1b+kitlga*: p=0.036, NI vs +*hapln1b+kitgb*: p=0.8617. h1b, hapln1b. ka, kitlga. Kb, kitlgb. (C) Electrostatic potential of *KITL*, *kiltga* and *kitlgb* at biological pH.

### Figure 6. *hapln1b* maintains ECM integrity through the link domains

(A) Schematic representation of WT *hapln1b* and a truncated version that lacks exon 4 and 5 (*hapln1b* $\Delta$ 4 $\Delta$ 5). (B) ISH expression of *cmyb* in NI control embryos, *hapln1b* mRNA injected embryos and *hapln1b* $\Delta$ 4 $\Delta$ 5 injected embryos at 36hpf. (C) Summary of proposed model. In the WT situation, hapn1b helps *kitlgb* to interact with *kitb* in the AGM to permit *runx1* expression and HSC formation. Loss of *hapln1b* results in a loss of *kitlgb* interacting with *kitb*, no *runx1* expression and no HSCs forming. Over expression of *hapln1b* results in a dense ECM and reduced *kitlgb* interacting with *kitb*, impairing HSC emergence and survival. HE: hemogenic endothelium.

Figure 1

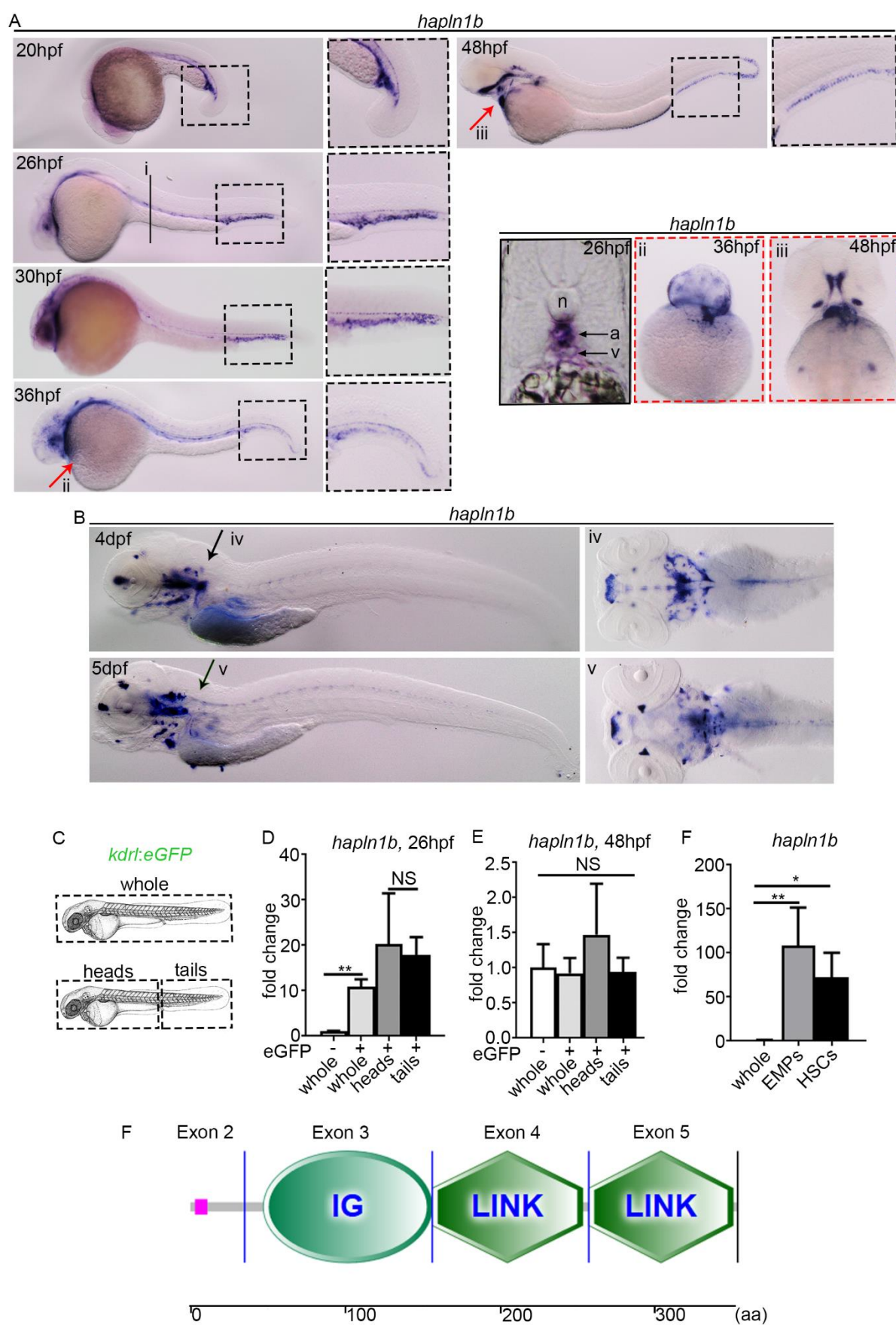


Figure 2

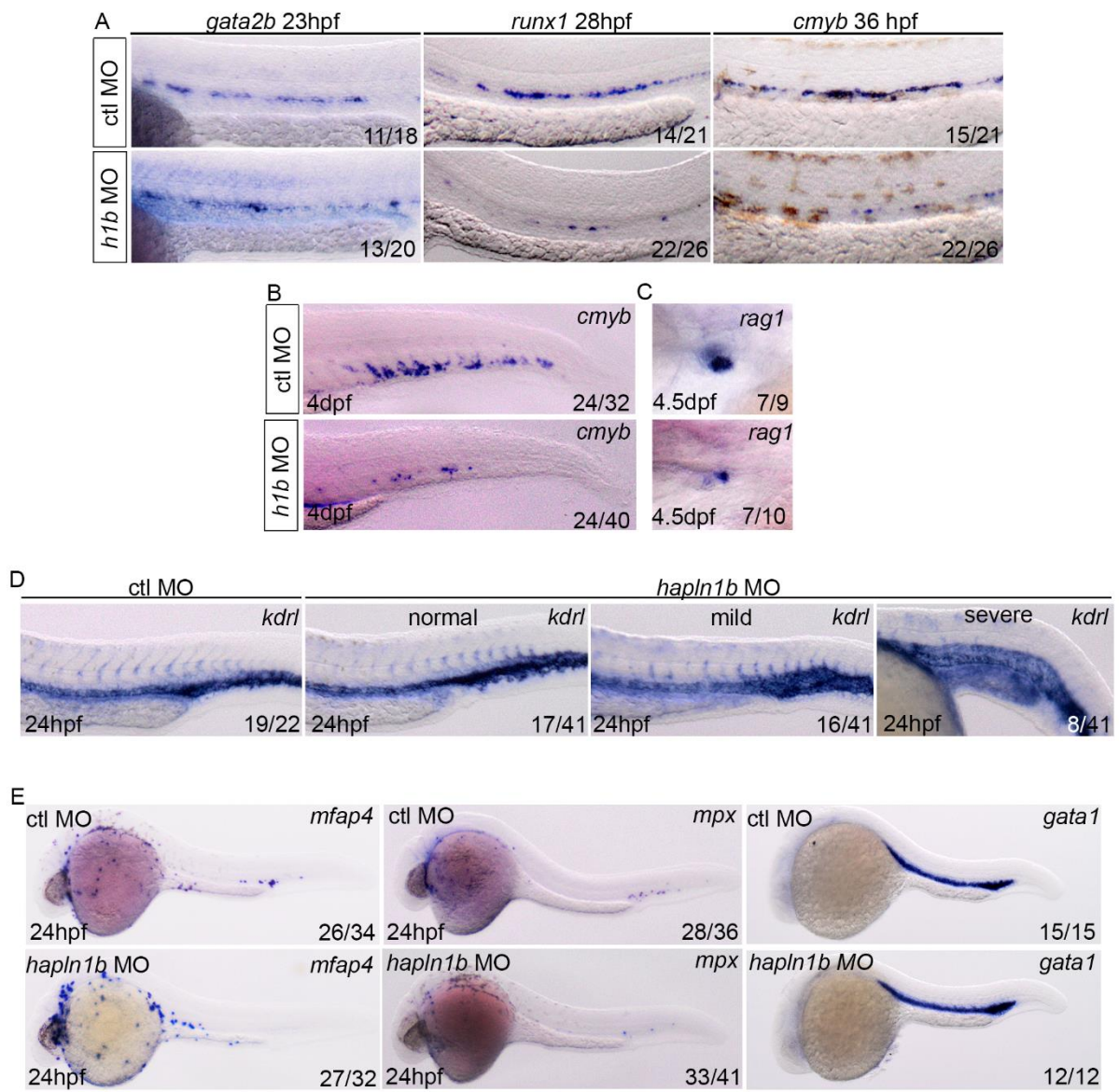


Figure 3

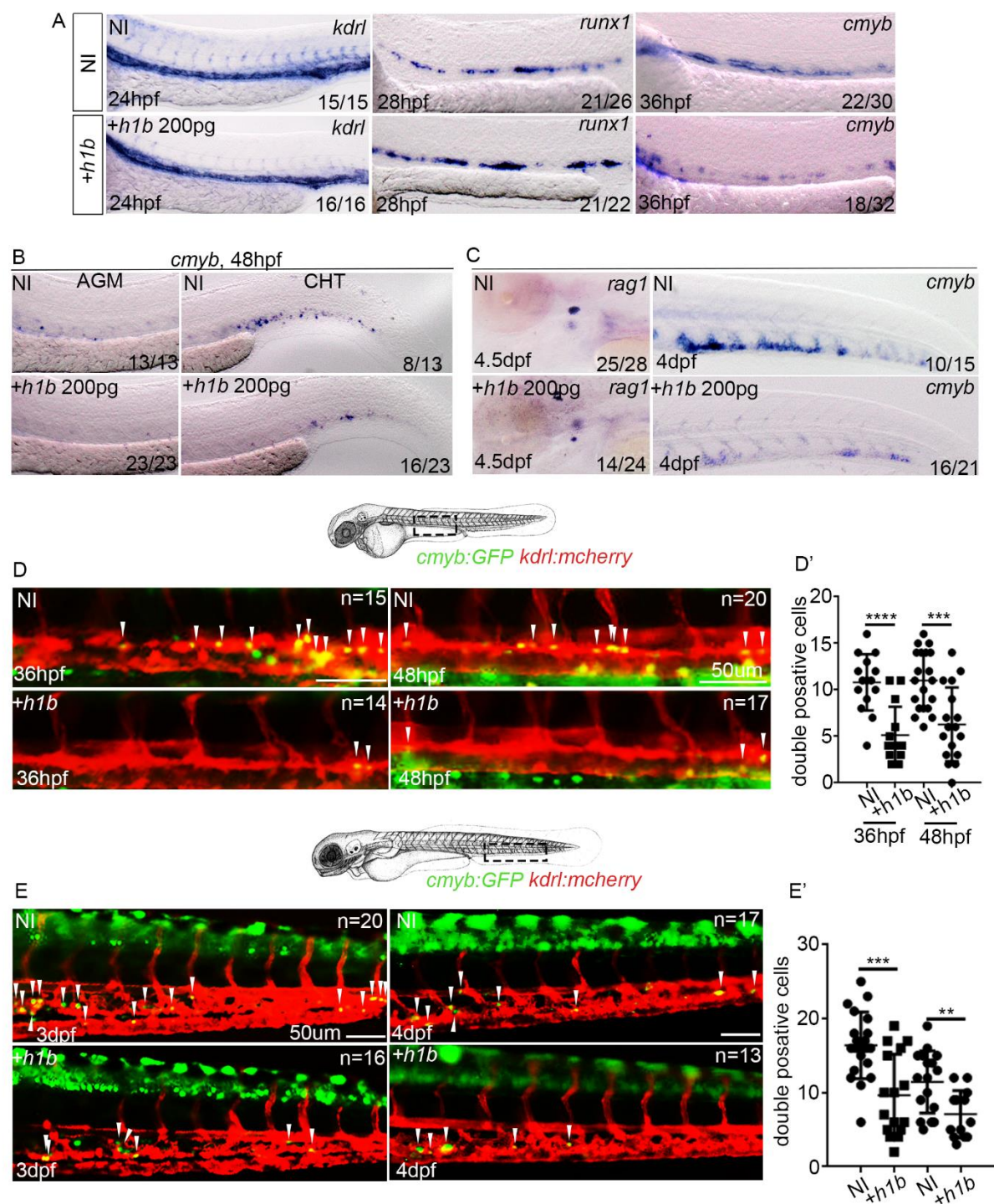


Figure 4

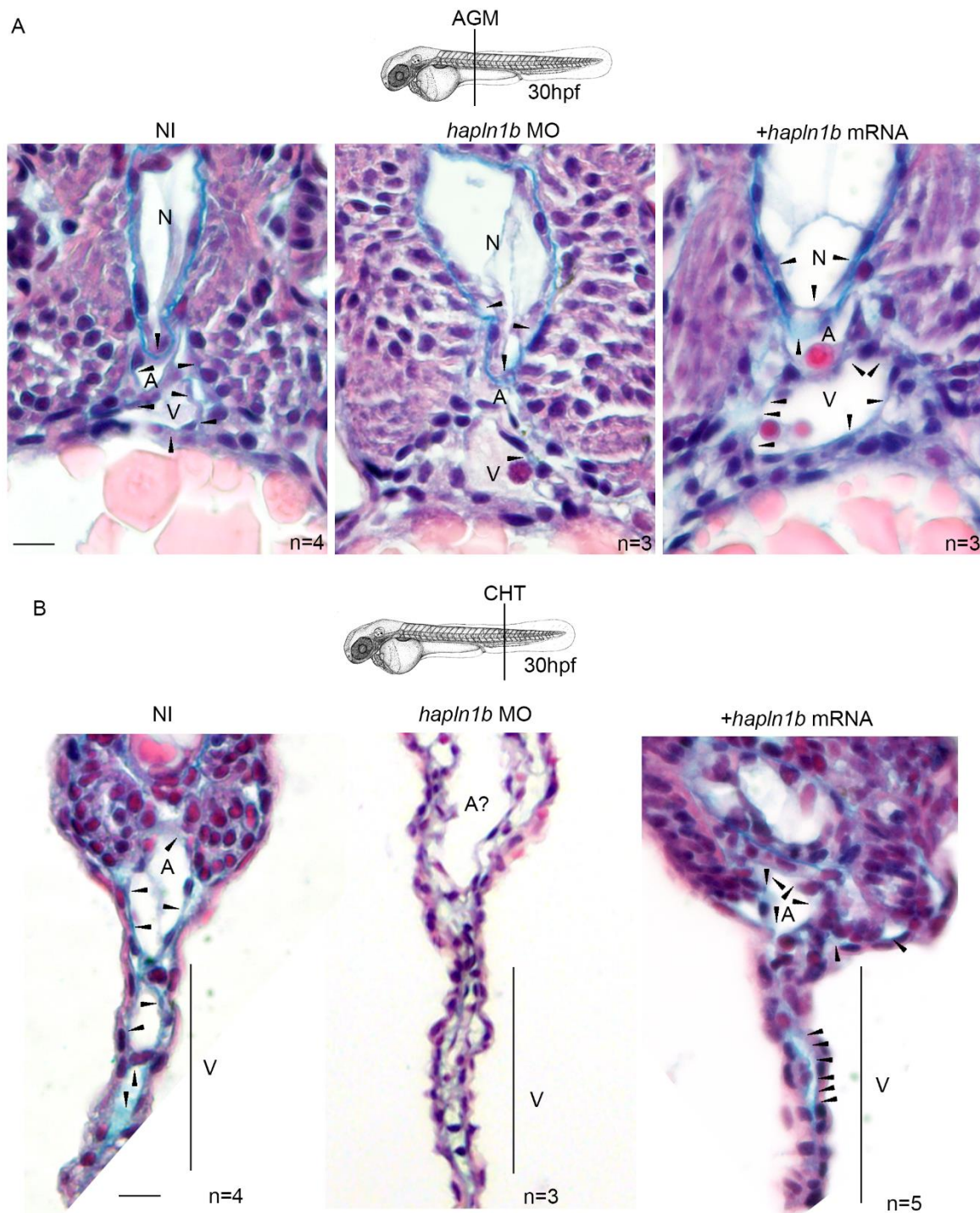


Figure 5

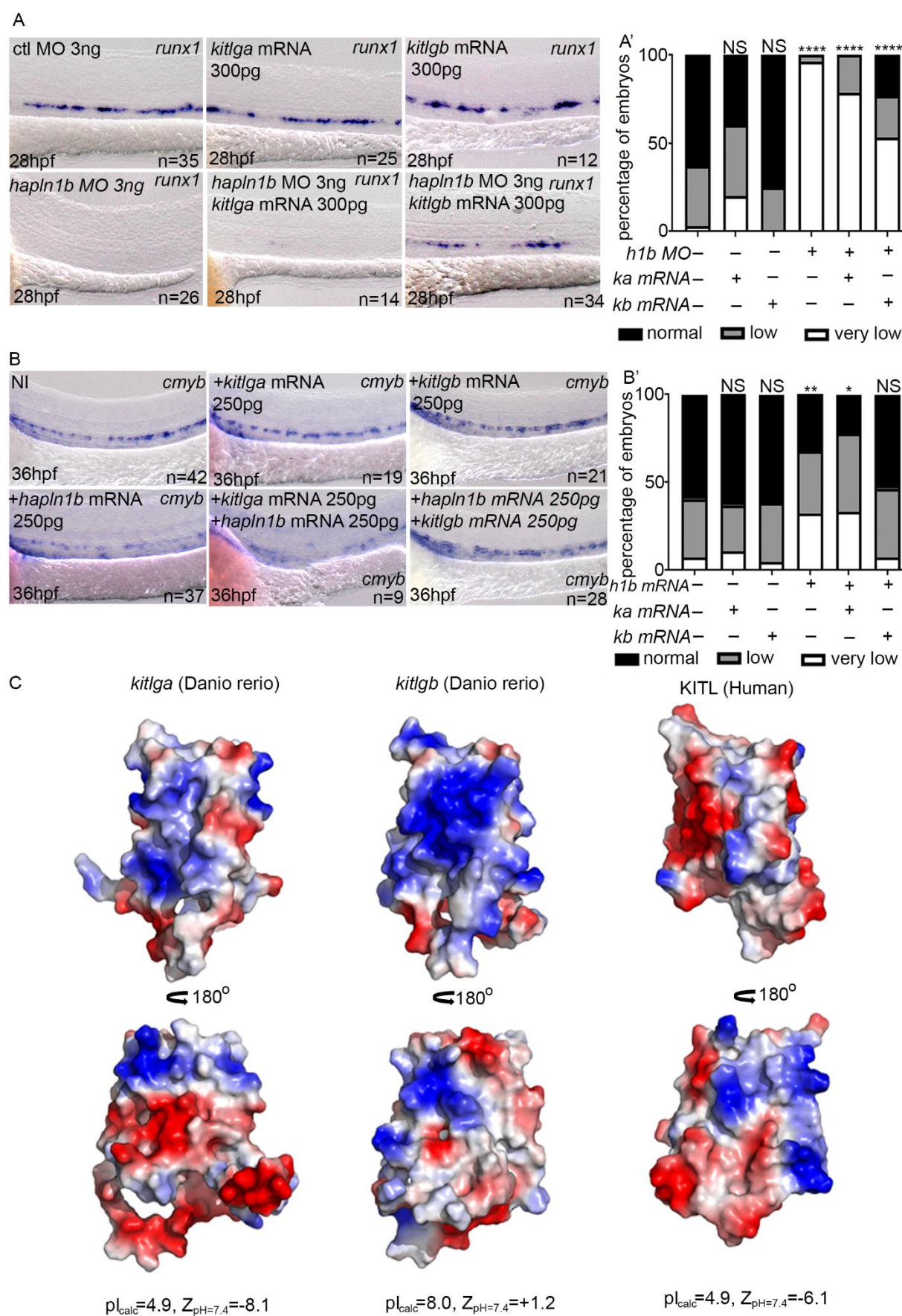
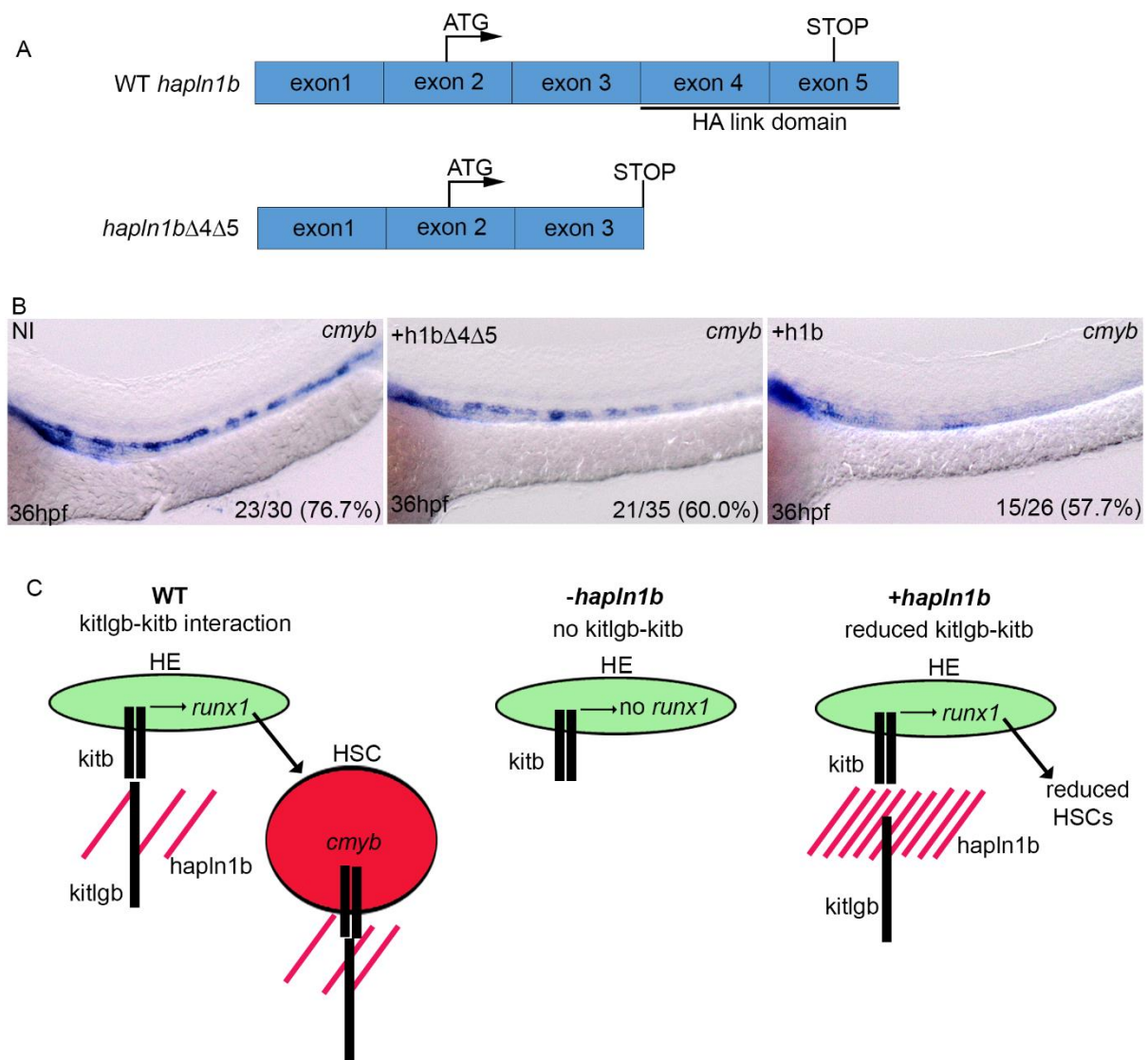


Figure 6





## Supplementary Tables

ef1a F	GAGAAGTTCGAGAAGGAAGC
ef1a R	CGTAGTATTGCTGGTCTCG
hapln1b F	ACAACTTTGAGGAGGCGTACA
hapln1b R	TGTGCTGTGTCGACTGAGAGT
kitlgb F	ATGGCATGAACTCTGCGGTT
kitlgb R	CGCATTCTTGCTCCACAACG

**Table S1. QPCR primers**

hapln1b ISH probe F	CAGAGCTGACAACTTTGAGG
hapln1b ISH probe R	ATCCCATAGTTCCTGACTCC
hapln1b full length F	AAAGAATTCAAGATGACCTTCCTGCTCCT
hapln1b full length R	AAACTCGAGGCGAACTCAGTTGTTGCTCT

**Table S2. Cloning primers**

Std Ctrl	CCTCCTACCTCAGTTACAATTTATA
hapln1b	ATCCTAAACATCCTCTCACCTGAGA

**Table S3. Mopholinos**



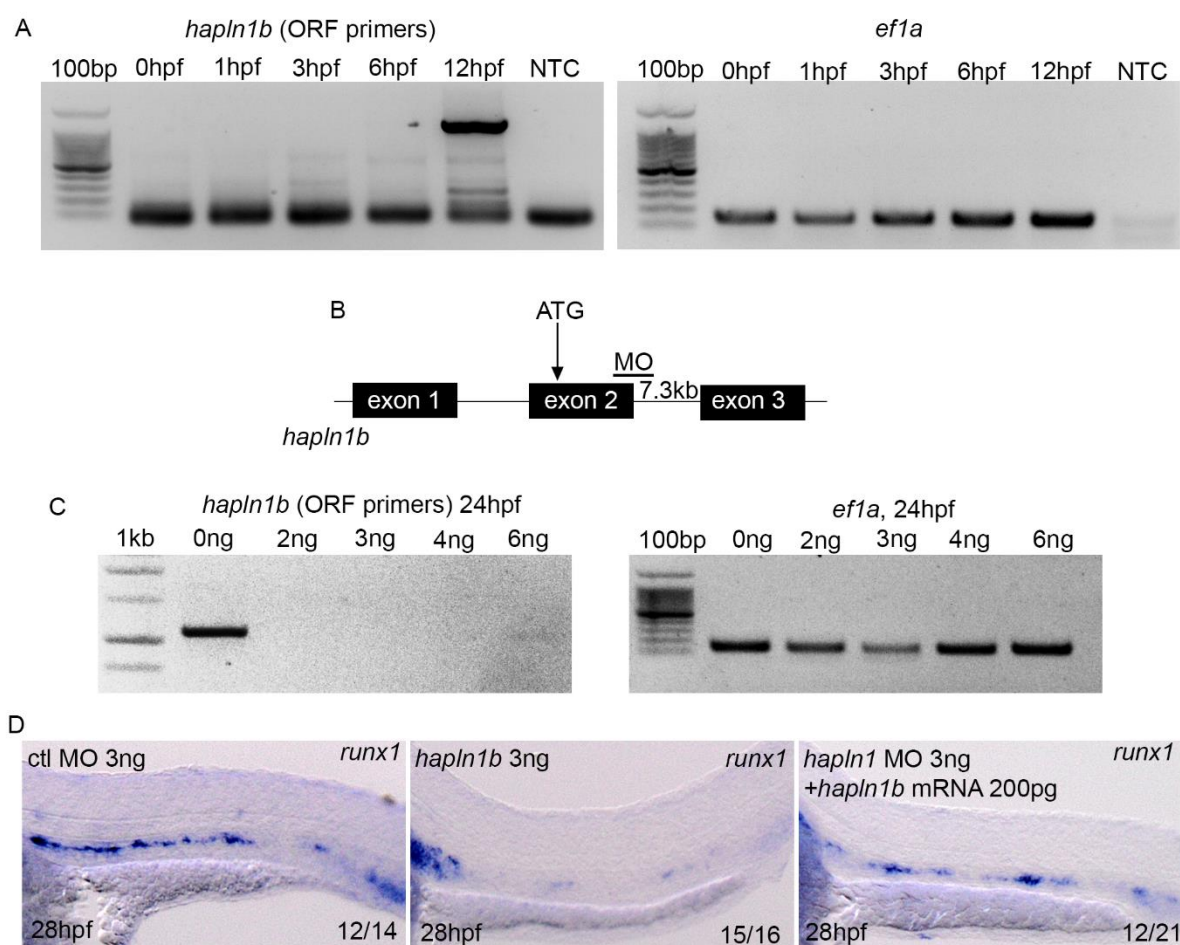
### **Figure S1. Comparison of hapln family member protein structures in zebrafish**

(A) Summary of known expression pattern of other members of the *hapln* family in zebrafish. ZFIN: <https://zfin.org/>. (B) Amino acid sequence alignment (using uniprot) with highlighted signal peptide and disulphide bonds of zebrafish hapln family and percentage identity compared to hapln1b as calculated by clustal 2.1. Q1LXE1= *hapln1a*, E7FBW8= *hapln1b*, A2CF10=*hapln2*, Q6NV41=*hapln3*, Q1LV25=*hapln4*.



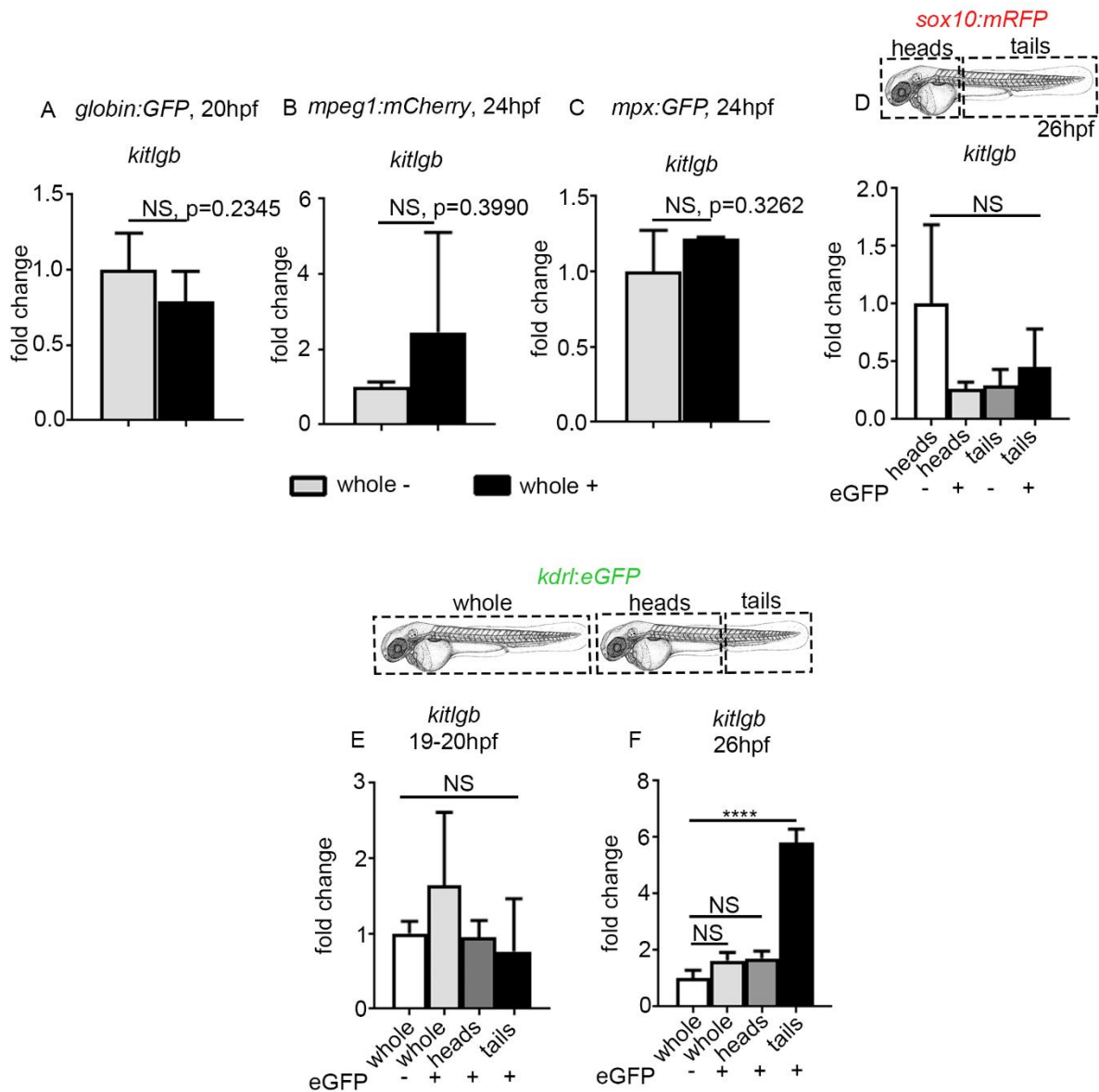
**Figure S2. *hapln1b* and *hapln1a* arise from duplication of the mammalian orthologue**

(A) Synteny analysis, showing duplication of the mammalian *hapln1* to give *hapln1a* and *hapln1b* in zebrafish. (B) phylogenetic analysis of *hapln1b* and its comparison to *hapln1a* in zebrafish (dr) and *hapl1b* in mice (m) and humans (h) as calculated by clustal 2.1. (C) Amino acid sequence alignment of zebrafish *hapl1b*, and *hapl1a*, *Hapl1* (mouse) and *HAPLN1* (human) with highlighted disulphide bond, propeptide, signal peptide and glycosylation site (analysed using uniprot). Percentage identity is compared to *hapln1b* as calculated by clustal 2.1. Q1LXE1= *hapln1a*, E7FBW8= *hapln1b*.



### Figure S3. Validation of the *hapln1b* MO

(A) *hapln1b* expression by PCR from pools of 8-10 embryos extracted at the indicated timepoints. Primers for cloning full length *hapln1b* were used. *Ef1a* was used as a housekeeping control. NTC, no template control. (B) Schematic indicating MO target site, schematic is not to scale. (C) PCR after injection of the *hapln1b* MO at the indicated timepoints. RNA was extracted from pools of 8-10 embryos at 24hpf. *Hapln1b* full length primers used (ORF: open reading frame). *Ef1a* was used as a housekeeping control. (D) ISH expression of *runx1* in control morphants, *hapln1b* morphants and *hapln1b* morphants injected with *hapln1b* mRNA.



**Figure S4. *kitlgb* is specifically expressed by caudal endothelial cells**

(A-D) qPCR expression of *kitlgb* in either negative cells or positive cells from the indicated transgenic lines. In A-C, data is analysed using a two-tailed Student's t-test, A:  $p=0.2345$ , B:  $p=0.3990$ , C:  $p=0.3262$ . In D, cells were sorted from dissection of heads or tails as indicated in the schematic, data was analysed using one-way anova. Heads- vs. heads+:  $p=0.1137$ , heads- vs. tails -:  $p=0.1294$ , heads- vs. tails+:  $p=0.2708$ . (E,F) Experimental outline and qPCR expression of *kitlgb* from FACS sorted endothelial cells from different dissections at either 19-20hpf or 26hpf. Data was analysed using one way anova. In E,

whole- vs. whole+:  $p=0.4839$ , whole- vs. head +:  $p=0.9994$ , whole- vs. tail+:  $p=0.9338$ . In

F, whole- vs. whole+:  $p=0.1300$ , whole- vs. head +:  $p=0.0866$ , whole- vs. tail+:  $p<0.0001$ .