Endocardial identity is established during early somitogenesis by Bmp signalling acting upstream of *npas4l* and *etv2/etsrp*

Samuel J. Capon[1] *, Veronica Uribe Sokolov[2] *, Nicole Dominado[2], Ophelia Ehrlich[2] and Kelly A. Smith[1,2,*,‡]

[1]Division of Genomics of Development and Disease, Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland 4072, Australia.

[2]Department of Anatomy & Physiology, The University of Melbourne, Melbourne, Victoria 3010, Australia.

*These authors contributed equally to this work

[‡]Author for correspondence, email: kelly.smith1@unimelb.edu.au

Keywords: Endocardium, *etv2/etsrp*, *npas4l*/*cloche*, Bmp signalling, zebrafish, cardiac

Summary Statement:

A zebrafish transgenic reporter of the endocardium is identified, permitting transcriptomic analysis and identification of new endocardial markers. Epistasis experiments demonstrate *npas4l* and *etv2* act downstream of Bmp signalling to regulate endocardial differentiation.

Abstract:

The endocardium plays important roles in the development and function of the vertebrate heart however few molecular markers of this tissue have been identified and little is known about what regulates its differentiation. We describe here the *Gt(SAGFF27C); Tg(4xUAS:egfp)* line as a marker of endocardial development in zebrafish. Transcriptomic comparison between endocardium and pan-endothelium confirms molecular distinction between these populations and time-course analysis suggests differentiation as early as 8 somites. To investigate what regulates endocardial identity, we employed *npas4l*/*cloche*, *etv2/etsrp* and *scl* loss-of-function models. Endocardial expression is lost in *cloche* mutants, significantly reduced in *etv2* mutants and only modestly effected upon *scl* loss-of-function. Bmp signalling was also examined: Overactivation of Bmp signalling increased endocardial expression, whilst Bmp inhibition decreased expression. Finally, epistasis experiments showed that overactivation of Bmp signalling was incapable of restoring endocardial expression in *etv2* mutants. By contrast, overexpression of either *npas4l* or *etv2* was sufficient to rescue endocardial expression upon Bmp inhibition. Together, these results describe the differentiation of the endocardium, distinct from vasculature, and place *npas4l* and *etv2* downstream of Bmp signalling in regulating its differentiation.

Introduction:

During early embryonic development, the heart comprises two major tissues: the myocardium and endocardium. To date, much research concerning heart development has focused on the myocardium, the beating muscle of the heart, while the endocardium has received relatively little attention. The endocardium is a specialised subset of endothelium that forms the inner lining of the heart. It plays important roles in the development of the trabecular myocardium, contributes to the coronary vasculature, septa and cardiac cushions, and connects the heart with the adjacent vascular endothelial network (Harris and Black, 2010; Kisanuki et al., 2001; Markwald et al., 1977; Markwald et al., 1975; Meyer and Birchmeier, 1995; Red-Horse et al., 2010; Stankunas et al., 2008; Tian et al., 2017). Despite these important

and unique roles, the developmental cues that distinguish the endocardium from the vascular endothelium remain to be determined.

Molecularly, the earliest specific marker of endocardium is *Nfatc1* (de la Pompa et al., 1998; Wong et al., 2012).The endocardial-specific expression of this marker, combined with the functionally distinct roles of the endocardium, argue that the endocardium is a unique subset of endothelium with its own molecular signature. When and how this distinction is first established, however, remains unclear. In mice, *Nfatc1* expression is already detectable in the medial aspect of the cardiac crescent at E7.5 (de la Pompa et al., 1998), the earliest recognised structure of the murine heart. In contrast, the zebrafish orthologue of this gene, *nfatc1*, is not reported to initiate expression until 22 hpf (Wong et al., 2012), corresponding with the formation of the linear heart tube. In zebrafish, endocardial progenitors bud from bilateral, $kdrl^{+}/fli1a^{+}$ populations at the 12-somite (12s) stage (15 hpf) and migrate to the midline to form the endocardial core of the cardiac disc (Bussmann et al., 2007; Proulx et al., 2010) before forming the linear heart tube. This observation suggests that the endocardial and vascular endothelial progenitors, which remain in the lateral margins of the anterior lateral plate mesoderm (ALPM), are distinct populations as early as 12s. However, molecular evidence that corroborates this early functional distinction in zebrafish has yet to be reported.

Separation of lineages at this early stage is not unprecedented. In fact, distinct expression domains for myeloid (*spi1b*, *runx1*, *c-myb*), (Kalev-Zylinska et al., 2002; Lieschke et al., 2002; Thompson et al., 1998) vascular endothelial (*fli1a*, *kdrl*) (Thompson et al., 1998) and myocardial (*nkx2.5*, *myl7*) progenitors are well described by the 14s stage (16 hpf) (Chen and Fishman, 1996; Yelon et al., 1999). These fields sit adjacent and mutually exclusive to one another, suggesting they are already specified. Importantly, they also respond differently to genetic cues: For example, the anterior myeloid population is dependent on Bmp signalling, whereas the adjacent vascular field at the same developmental stage is not (as determined by *in situ* for *pu.1*/*spi1b*, *scl* and *fli1*) (Hogan et al., 2006). Apparently, location is everything in this context because vascular fate is sensitive to Bmp signalling in the posterior lateral plate mesoderm at the same developmental stage (Liu et al., 2008).

Here, we report the identification and characterisation of a previously reported marker of the zebrafish lymphatic system (Bussmann et al., 2010; Okuda et al., 2012) as a novel marker of early endocardial progenitors. Characterisation of this line identified distinct endocardial expression domains as early as 8s (13 hpf). Similarly, a re-analysis of *nfatc1* expression found overlapping *nfatc1* expression domains. RNA-sequencing of endocardial and endothelial progenitors identified a distinct transcriptional signature in the endocardium, confirming an early divergence in these tissues. Finally, examination of the signals governing endocardial lineage development identified *npas4l*, *etsrp*/*etv2* (hereafter referred to as *etv2*) and Bmp signalling as important regulators of endocardial development and epistasis experiments place Bmp signalling upstream of *npas4l* and *etv2* in the ALPM for induction of endocardial marker expression.

Results:

The *Gt(SAGFF27C); Tg(4xUAS:EGFP)* line labels the endocardium during early-to-mid somitogenesis stages

The *Gt(SAGFF27C); Tg(4xUAS:EGFP)* line, originating from a large gene-trap screen (Asakawa et al., 2008), was observed to have enriched expression in the endocardium at 48 hpf, relative to the previously described vascular expression in Bussmann et al., 2010. To investigate this, we performed time-course analysis of the *Gt(SAGFF27C); Tg(4xuas:egfp)* line (henceforth referred to as *Gt(endocard:egfp*)), and compared it with the endothelial transgenic lines, *Tg(kdrl:Hsa.HRAS-mCherry)* and *Tg(fli1a:egfp)* (Fig 1). eGFP expression in the *Gt(endocard:egfp*) line is first visible as bilateral populations in the anterior lateral plate mesoderm at the 10s stage (14 hpf). It colocalises with *Tg(kdrl:Hsa.HRAS-mCherry)* and localises similar to *Tg(fli1a:egfp)*, confirming that its expression is endothelial, and it is consistent with regions of endocardial progenitors (Fig 1).

Consistent with previous reports (Bussmann et al., 2007; Proulx et al., 2010), the bilateral populations migrate to the midline by 15s, while vascular progenitors remained at the lateral margins of the ALPM (Fig 1). *Gt(endocard:egfp*) expression is largely restricted to these migrating endocardial progenitors. By 20s, endocardial

progenitors have fused to form the endocardial core of the cardiac disc (Fig 1). Whilst some cells appear to be $GFP^{+}/mCherry^{-}$, quantification shows the majority (93%) to be double positive (Fig S1). Analysis of timelapse movies shows that this line also likely labels a few myeloid cells at this time, based on their morphology and migratory movements (Movie 1 and 2). At 24 hpf, endocardial expression is observed in the linear heart tube of *Gt(endocard:egfp)* embryos (Fig 1). By 48 hpf, *Gt(endocard:egfp)* expression has expanded and is also observed in the developing vascular endothelium (Fig 1). At 24 and 48 hpf, some weak GFP expression can also be observed in the myocardium (Fig S2). Together, this data suggests that *Gt(endocard:egfp)* is enriched in the endocardium at early-to-mid somitogenesis stages, making it a unique tool for studying endocardial development in a living embryo.

The endocardium is a molecularly distinct subset of endothelium

Given that at early stages *Gt(endocard:egfp)* expression is enriched in the endocardium, we used this line to compare the transcriptome of the endocardium with *fli1a*:eGFP-positive cells (primarily endothelium and blood cells). At 15s, GFP cells were isolated from *Gt(endocard:egfp)* and *Tg(fli1a:egfp)* embryos by FACS (Fig 2A-D) and mRNA sequencing performed. Principle component analysis showed that endocardial samples clustered separately from endothelial samples (Fig S2A). To confirm the integration site of the *Gt(endocard:egfp)* line, RNA sequencing reads were mapped to the gene-trap cassette. These reads extended into exon 1 of the *map3k22* gene, confirming the reported integration site (Fig S3B) (Bussmann et al., 2010). Despite mapping to this locus, ISH analysis of *map3k22* showed no detectable expression in the endocardial region (S.J. Capon, unpublished observation). Furthermore, no *map3k22* transcripts were detected in the *Tg(fli1a:egfp)* transcriptome. This suggests that rather than the *Gt(SAGFF27C)* gene-trap line reporting *bona fide map3k22* expression, the endocardial expression observed is an artefact either of chromosomal rearrangement from the particular gene-trap insertion event and/or is reporting expression from a nearby locus. Whichever is the case, we have not determined it here.

Hierarchical clustering of the top 200 most variable genes showed that endocardial and endothelial samples separate according to cell type (Fig 2E). Differential expression analysis found a total of 122 statistically significant genes enriched in the endocardium and 165 enriched in the endothelium (Fig 2F). To validate these differences, a subset of genes representing blood, endothelium and endocardium were validated by qPCR (Fig 2F and 2G). Genes related to red blood cell development, *gata1a*, *hbbe3* and *hbae5* were enriched in the *fli1a:egfp*$^{+}$ population as expected due to these cells arising from the intermediate cell mass in the trunk of the embryo (Davidson and Zon, 2004) (Fig 2F), although we were unable to validate the *gata1a* expression by qPCR (Fig 2G). Surprisingly, of the expected pan-endothelial markers analysed, only *kdrl* and *cdh5* were found to show no statistically significant fold change, while *flt4*, *scl*, and *etv2* were all significantly enriched in the endothelium (Fig 2F). Reassuringly, *fn1a* and *nfatc1*, markers previously reported to be expressed in the endocardium (Palencia-Desai et al., 2011; Trinh and Stainier, 2004; Wong et al., 2012), were enriched in the *Gt(endocard:egfp)* population (Fig 2F and G), while GO term analysis on the endocardially enriched gene list retrieved terms relating to heart development and morphogenesis (Fig 2H and Table S1). To identify novel genes expressed in the endocardium, several were selected for ISH-based screening to validate endocardial expression (Fig S4). Of these, a subset were found to have restricted expression in regions consistent with the endocardium, including *cbfa2t3*, *gfi1ab*, *hapln1b* and *nrp1a* (Fig 2 and S4).

To investigate the earliest time point at which endocardial progenitors begin to express these markers, expression of the *Gt(endocard:egfp)* line was examined by *in situ* hybridisation over a developmental time-course from 5s (11.5 hpf) through to 48 hpf (Fig 3). *endocard:egfp* expression was first detected as early as 8s (13 hpf), this expression was observed in faint semilunar domains in the lateral regions at the anterior of the embryo (Fig 3) and were lost in a *npas4l/cloche* mutant (Fig S5). These bilateral expression domains were observed more caudal at 10s (Fig 3). By 15s (16.5 hpf), expression is found at the midline in the region of the endocardial core of the cardiac disc (Fig 3), consistent with the observations made by examining fluorescence (Fig 1). As *nfatc1* expression was also found to be enriched in the endocardial population at 15s (Fig 2F and 2G), its expression was also analysed over the same time course. *nfatc1* expression was similarly found to initiate as early as 8s and was

observed in domains comparable to that of *endocard:egfp* expression, although much weaker (Fig 3). Finally, *cbfa2t3* (identified from transcriptome analysis; Fig 2 and S4) was analysed over the same time-course. *cbfa2t3* expression was also found to initiate at 8s, although at very weak levels, and was observed in similar domains up until at least 15s (Fig 3). By 20s, *cbfa2t3* expression was no longer apparent in the endocardium, although expression continued in the otic placode, olfactory placode, nervous system and head (Fig 3). Together, these results show that the endocardium expresses multiple unique markers as early as 8s, much earlier than previously appreciated in zebrafish development.

npas4l/cloche **and** ***etv2*****, but not** ***scl*****, are required for endocardial progenitor commitment**

Previous studies have elegantly demonstrated that perturbation of fate-determining transcription factors can impact the expression domains of myocardial, endothelial and myeloid populations across the ALPM (Keegan et al., 2005; Lieschke et al., 2002; Palencia-Desai et al., 2011; Schoenebeck et al., 2007; Simoes et al., 2011; Sumanas et al., 2008). However, without a unique marker of endocardial identity, this population is yet to be studied. Having established the *Gt(endocard:egfp)* line as a marker of the endocardium, we first performed expression analysis to confirm its location in relation to these other well-studied cell populations (Fig S6). Double fluorescent ISH analysis of the myocardial marker, *myl7*, combined with *egfp* expression at 15s showed endocardial progenitors located rostral and medial to myocardial progenitors (Fig S6). Double fluorescent ISH analysis of the myeloid marker, *spi1b*, and *egfp* showed endocardium located caudal and medial to myeloid cells (Fig S6). Combining this location with expression of *Gt(endocard:rfp)* on the *Tg(kdrl:egfp)* background at 20s allowed us make a map of the endocardial expression domain in relation to these other developing primordia (Fig S6). Importantly, expression of each of these markers (with the exception of *Gt(endocard:rfp)* and *Tg(kdrl:egfp)*) showed mutually exclusive expression at 15s.

Next, we generated loss-of-function (LOF) models for key regulators of endothelial identity and investigated their effect on endocardial development. CRISPR/Cas9 genome-editing was used to create LOF mutants for *npas4l* and *etv2* genes (Fig S7) and a previously validated morpholino against *tal1/scl* was used to deplete *scl*. Each of these models was phenotypically characterised and found to phenocopy published LOF models (Fig 4 and S7). These models were examined for endocardial (*endocard:egfp*), myeloid (*spi1b*) and myocardial (*myl7*) expression at 14- and 16s, respectively.

As expected, endocardial expression was completely absent in *npas4l* mutants at both 14- and 16s (Fig 4). The myeloid expression domain was also completely absent in mutants, consistent with previous reports (Lieschke et al., 2002). The myocardial expression domains were expanded (Fig 4), as described previously (Schoenebeck et al., 2007).

We next examined *etv2* mutants and found a trend toward decreased staining area for endocardial expression at 14s (Fig 4). Most notably, the bilateral *endocard:egfp* expression in these mutants was extended along the anterior-posterior axis and localised to the lateral margins of the embryo (consistent with vascular endothelial localisation), suggesting presumptive endocardium is failing to migrate to the midline as in wildtype siblings (Fig 4). By 16s, *endocard:egfp* expression is significantly reduced and cell counts show a dramatic 98% reduction in cell numbers, compared with sibling controls (Fig S7). In addition, *spi1b* expression is lost at 14s and *myl7* expression expanded in length at 16s (Fig 4).

Surprisingly, *scl* knockdown had little effect on the expression domain of *endocard:egfp* (Fig 4). This contrasts with cell counts for *scl* morphants, which showed a small (25%) reduction in cell number (Fig S7). Interestingly, these data signify a clear distinction between *etv2* and *scl* function in this endothelial population. At 16s, the spatial organisation of endocardial expression was altered in *scl* morphants, in accordance with a previously described role for *scl* in endocardial cell migration and morphology (Bussmann et al., 2007; Schumacher et al., 2013). Also consistent with previous reports (Gering et al., 1998; Patterson et al., 2005;

Shivdasani et al., 1995), a significant reduction of the myeloid domain was observed (Fig 4) as well as an increased myocardial expression domain.

Together, these results show that transcription factors *npas4l* and *etv2* are required for endocardial development. Furthermore, *scl* is dispensable for early endocardial differentiation, despite its described role in endocardial morphogenesis (Bussmann et al., 2007; Schumacher et al., 2013) and maintenance of endocardial identity during later development (Clay and Ramakrishnan, 2005; Schumacher et al., 2013; Van Handel et al., 2012).

Bmp signalling is required for endocardial differentiation

Next, we examined downstream effectors of two signalling pathways known to be important for cardiovascular development. In zebrafish, VEGF signalling has previously been shown to act through Erk during early stages of angiogenesis (Shin et al., 2016a; Shin et al., 2016b) and Bmp signalling through Smad1/5/8 (Derynck and Zhang, 2003). Analysing phosphorylated Erk 1/2 (pErk) and phosphorylated Smad1/5/8 (pSmad) in *Gt(endocard:egfp*) or *Tg(fli1a:EGFP)* embryos at 15s showed considerable co-labelling of pSmad with endocardial progenitors whereas pErk labelling was observed predominantly in adjacent vascular endothelium (Fig 5A). Quantification confirmed this, showing increased pErk intensity in endothelium compared with endocardial progenitors (Fig 5 and S8). Conversely, an increase in the number of pSmad-positive cells was observed in endocardial progenitors compared with adjacent vascular endothelium. This suggests that Bmp signalling is more active in endocardium than developing vasculature, whereas Erk is more active in vasculature than endocardium.

To investigate whether Bmp signalling plays a functional role in endocardial development, the *Gt(endocard:egfp*) line was crossed with the heat-shock-inducible transgenic lines, *Tg(hsp70l:bmp2b)* and *Tg(hsp70l:nog3)* (Chocron et al., 2007) either activating or inhibiting the Bmp pathway, respectively. Embryos were heat-shocked at tailbud stage (10 hpf) and fixed at the 14- or 16s for analysis of the myeloid, endocardial and myocardial domains. Inhibition of the Bmp signalling pathway by global induction of *noggin3* expression resulted in a reduction of endocardial (Fig 6),

myeloid and myocardial expression domains at the 14s and 16s (Fig S9). Reciprocally, activation of the Bmp signalling pathway by global induction of *bmp2b* expression caused a significant expansion of the endocardial (Fig 6) and myeloid (Fig S9) expression domains but did not increase in the myocardial domain area (Fig S9). Interestingly, activating Bmp signalling not only increased the endocardial expression domain but also resulted in stronger staining compared with siblings lacking the heat-shock transgene (Fig 6). Quantification of cell number showed a trend but no significant increase in cell number following increased Bmp signalling, implying the expansion of the endocardial domain is through increased transcription of *egfp* within endocardial cells. These results show that Bmp signalling is necessary for early endocardial development.

Bmp signals through *npas4l* and *etv2* to instruct endocardial differentiation

Given the dependency of the endocardium on *npas4l* and *etv2* as well as Bmp signalling (Fig 4, 6 and S9), we examined whether *npas4l* and *etv2* expression were perturbed by altered Bmp signalling. Inhibition of Bmp signalling by heat-shock at the tailbud stage did not abolish either *npas4l* or *etv2* expression but significantly reduced the expression of *npas4l* and altered patterning of *etv2* (Fig S10). In addition, weak *etv2* expression could be detected by ISH in the endocardium at 20s. This expression was reduced or loss upon inhibition of Bmp signalling (Fig S10), suggesting that BMP signalling is required for the *etv2* expression.

We next examined the epistatic relationship between Bmp signalling and *etv2* in endocardial development. Heat-shock of *Tg(hsp70l:bmp2b)* at tailbud stage had no effect on endocardial expression in the absence of *etv2*, in that overactivation of Bmp signalling could not rescue the lost endocardial expression observed in *etv2* mutants (Fig 7). This suggests that *etv2* functions downstream of Bmp signalling to induce endocardial differentiation. To prove this, the reciprocal experiment was performed, overexpressing *etv2* in *Tg(hsp70l:nog3)* embryos (Fig S11). Whilst endocardial expression is reduced upon inhibition of Bmp signalling compared with wildtype (Fig 6), overexpression of *etv2* upon inhibition of Bmp signalling was capable of restoring *endocard:egfp* expression (Fig 7). A similar phenomenon was observed for *npas4l*

overexpression in Bmp inhibited embryos, suggesting that *etv2* and *npas4l* are sufficient for endocardial differentiation in the absence of BMP signalling. Together, these results demonstrate that *npas4l* and *etv2* function downstream of Bmp signalling to induce endocardial differentiation.

Discussion:

The endocardium is a unique invention of the vertebrate taxa (Perez-Pomares et al., 2009). Although the endocardium has many molecular similarities to the adjacent vascular endothelium, there are distinct physiological functions of these tissues (Harris and Black, 2010). What distinguishes these lineages on a molecular level remains an important and underexplored question in cardiovascular biology.

Characterisation of the *Gt(endocard:egfp)* line during somitogenesis stages shows expression of this marker enriched in endocardial progenitors as early as 8s. Importantly, we found that *nfatc1* is expressed in similar spatiotemporal domains and a newly identified marker, *cbfa2t3*, also has similar expression. These results show that the endocardium is specified earlier than previously appreciated in zebrafish development and that a unique transcriptional profile is established prior to the formation of the linear heart tube, bringing it in line with what has been reported in mice (de la Pompa et al., 1998). The observed onset of endocardial expression at 8s is suggestive that endocardial identity is assumed close to this time in development. This marries well with observations of *npas4l* and *etv2* (*etv2* is a direct target of *npas4l*) expression – upstream regulators of endothelial identity. *npas4l* expression is reported to peak at tailbud stages and remain high until 15s, whereas *etv2* initiates weakly at tailbud and is highly expressed by 15s (Reischauer et al., 2016). The expression of these two transcription factors that are required for endocardial development is consistent with the onset of endocardial identity from 8s.

The 8s stage represents an intriguing time point for endothelial fate decisions. In *scl* morphants, the mesodermal markers *hhex* and *draculin* are substantially reduced from 7s, despite initiating expression normally at earlier stages (Patterson et al., 2005). These results suggest that this stage marks a crucial time point in development when endothelial progenitors undergo significant changes, perhaps committing to

specialised lineages, such as the endocardium. Adding to this intrigue, both *snai1b* and *myod1* expression in the forming somites undergo unique changes at 7s (Hammerschmidt and Nusslein-Volhard, 1993; Weinberg et al., 1996) suggesting that such fate decisions, may occur more broadly across the mesoderm around this stage.

Surprisingly, in our hands *scl* morphants were found to have only minor defects in early endocardial development despite its important role in blood development, a role in endocardial morphogenesis (Bussmann et al., 2007; Schumacher et al., 2013) and a later maintenance role in endocardial identity (Schumacher et al., 2013; Van Handel et al., 2012). We therefore propose that whilst *scl* is an important transcription factor in endocardial development, it is in fact dispensable for the early differentiation of the endocardial lineage. Unlike *scl*, we found *etv2* is required for endocardial differentiation. Similar observations have been made in mouse, with *Etv2* knock-out resulting in an absence of endocardium (Ferdous et al., 2009). This suggests a conserved role for this *Etv2* in endocardial development.

In addition to our analysis of transcription factors, we also report the requirement for Bmp signalling in endocardial differentiation and find that it is necessary in this regard. Bmp signalling has previously been implicated in endocardial development, particularly in the formation of the cardiac valves (Dietrich et al., 2014; Misfeldt et al., 2009; Saxon et al., 2017; Snider et al., 2014; Wang et al., 2005) and in paracrine signalling between the endocardium and myocardium (Palencia-Desai et al., 2015; Saint-Jean et al., 2019). More broadly in cardiac development, it has also been shown to regulate cardiomyocyte differentiation in the cardiac disc (de Pater et al., 2012) and to instruct cardiac progenitor cell migration during linear heart tube formation (Smith et al., 2008). The successive and yet diverse roles for the Bmp pathway in cardiac development makes it challenging to investigate. For example, somewhat counterintuitive in this study, we observe enriched pSmad in the endocardium after it has differentiated (Fig 5), consistent with previous studies indicating an ongoing requirement for Bmp signalling after differentiation has occurred (Palencia-Desai et al., 2015; Snider et al., 2014). It is feasible and perhaps likely that Bmp signalling plays additional roles in endocardial development, and possibly also endothelial development, that have yet to be described. Whether this is the case or not, we

provide strong evidence here, through expression studies and epistasis experimentation that *npas4l* and *etv2* function downstream of Bmp signalling in endocardial differentiation. Consistent with this, *Etv2* has been reported to function downstream of Bmp signalling in the generation of *flk1*$^{+}$ mesoderm *vitro* (Lee et al., 2008) and in the posterior lateral plate mesoderm in the zebrafish (Row et al., 2018). Furthermore, characterisation of the murine *Etv2* promoter has identified multiple SMAD binding sites (Shi et al., 2015) suggesting direct regulation of *Etv2* by Bmp may occur. Intriguingly, *npas4l/cloche* was recently shown to bind to the *etv2* promoter and regulate its expression (Marass et al., 2019) and we find here that *npas4l* overexpression is capable of inducing endocardial differentiation despite Bmp inhibition. This suggests that SMAD cooperativity is not required for *npas4l* activity in endocardial development. Together, these findings suggest that Bmp signalling, upstream of *npas4l* and *etv2*, plays a crucial role in regulating endocardial development.

Finally, further characterisation of the endocardium by RNA sequencing demonstrates a unique molecular signature distinguishing the endocardium from vascular endothelium. These results are in agreement with an increasing number of single cell RNA sequencing (scRNAseq) studies from Human, Mouse, Chick and Zebrafish, showing the separation of endocardial cells from other endothelial populations. Importantly, many of these studies include genes such as *nfatc1, cbfat2t3 and itga9* (and others identified and confirmed via ISH [Fig S4]) as expressed in the endocardial cell population, supporting the notion of a conserved regulatory program for endocardial development (Chestnut et al., 2020; Hu et al., 2018; Ma et al., 2021; Mantri et al., 2021; Spanjaard et al., 2018; Tucker et al., 2020). This suggests that further analysis of the endocardial transcriptome in model organisms from equivalent stages may identify novel endocardial regulators. Given that the genetic aetiology of many congenital heart disease patients remains unclear (Morton et al., 2022), functional analysis of such regulators is required to improve clinical outcomes. Further, a number of studies have reported the differentiation of endocardial-like cells (Hofbauer et al., 2021; Mikryukov et al., 2021; Neri et al., 2019; Palpant et al., 2015) yet how closely these mimic *in vivo*-derived endocardium remains unclear due in part to a lack of molecular markers for such a comparison. We believe that the

identification of a core subset of conserved endocardial regulators will further aid attempts to generate endocardial tissue and model cardiac development *in vitro*.

Materials and Methods:

Zebrafish lines

All zebrafish strains were maintained and animal work performed in accordance with the guidelines of the animal ethics committee at The University of Queensland and The University of Melbourne, Australia. The previously published transgenic lines used in this study are *Gt(SAGFF27C); Tg(4xUAS:GFP),* referred to as *Gt(endocard:egfp*) in this text (Bussmann et al., 2010), *Tg(kdrl:Hsa.HRAS-mCherry)*^{s916} *(Hogan et al., 2009)*, *Tg(fli1a:EGFP)*^{y1} (Lawson and Weinstein, 2002), *Tg(hsp70l:nog3)*^{fr14} and *Tg(hsp70l:Bmp2b)*^{fr13} (Chocron et al., 2007), *Tg(kdrl:EGFP)*^{s843} (Jin et al., 2005). The construct for the previously published *Tg(myl7:mCherry-CaaX)*^{bns7} line (Uribe et al., 2018) was injected to establish the *uqks35* allele (Koopman et al., 2021).

Antibody staining

Phospho-protein staining was performed as previously described (De Angelis et al., 2017; Okuda et al., 2018) using the following antibodies: GFP (abcam cat. no. ab13970), pErk1/2 (cell signaling technologies cat. no. 4370), pSmad1/5/8 (cell signaling technologies cat. no. 13820), DsRed (Living Colors DsRed Polyclonal Antibody Takara 623496) as previously described (Grassini et al., 2019). Two rounds of independent staining were performed for each pErk1/2 and pSmad1/5/8 stain and n is depicted in the images and accompanying dot-plots.

Fluorescent *in situ* hybridisation

Fluorescent *in situ* hybridisation was performed as previously described (Baek et al., 2019; Brend and Holley, 2009; Clay and Ramakrishnan, 2005) with minor alterations. In brief, embryos were dechorionated and fixed at the desired stage. RNA probes were synthesised with DIG or FLU RNA labelling mix (Roche). Fixed embryos were permeabilised with proteinase K (Invitrogen) and hybridised with 1ng/µL RNA probe in hybridisation buffer overnight at 70°C. After hybridisation, embryos were washed, blocked with western blocking reagent (Roche) and incubated with anti-dig or anti-flu, POD antibodies (Roche) in western blocking reagent overnight at 4°C. After further washes, staining was performed using the tyramide signalling amplification kit (Perkin Elmer) for 2 hours at 37°C. Following staining embryos were fixed at 4°C overnight and mounted for imaging.

FACS and RNA sequencing

To prepare for FACS, an outcross of either homozygous *Gt(endocard:egfp)* or heterozygous *Tg(fli1a:egfp)* fish to WT fish were set-up and embryos collected 30 minute after dividers were removed. Fertilised embryos were kept at 28.5°C until they reached gastrulation stages, at which point they were placed at 23°C to slow down development. The following morning, embryos were examined to determine their developmental stage according to the number of somites present before being returned to 28.5°C. Immediately prior to the 15-somite stage, embryos were dechorionated by adding pronase. At the 15-somite stage, embryos were collected and dissociated for FACS. Briefly, embryos were de-yolked by pipetting embryos in calcium free ringer's solution. Embryos were then centrifuged at 2000rpm for 5 minutes at 4°C, the supernatant removed and replaced with protease solution (a 1:35 dilution of 2.5mg/mL Liberase TM (Roche) in PBS). Embryos were then dissociated by incubating at 28.5°C for 10 minutes with intermittent homogenisation by pipette. Dissociation was stopped by placing samples on ice and adding $CaCl_2$ and FCS to final concentrations of 2mM and 10% respectively. Samples were then centrifuged at 2000rpm for 5 minutes at 4°C, the supernatant removed and the resulting cell pellet resuspended in PBS with 2mM EDTA and sorted at Queensland Brain Institute's flow

cytometry facility. Only high expressing cells were collected and used for further analysis, to limit the contribution of cardiomyocytes that express eGFP weakly. GFP+ cells were collected in TRIzol LS reagent (Thermo Fisher Scientific) and RNA extracted using Direct-zol RNA MiniPrep columns (Zymo Research). As the total amount of RNA recovered from these samples was low, RNA was amplified using the Smart-seq2 protocol (Picelli et al., 2014). Briefly, 2ng of RNA was reverse transcribed using SuperScript II reverse transcriptase (Thermo Fisher Scientific) with oligo-dT and template switching oligo primers. The resulting cDNA was then amplified using KAPA HiFi HotStart ReadyMix (Roche) and the ISPCR primer. The resulting amplified cDNA was then purified using the Axygen AxyPrep Mag Fragment Select Kit (Thermo Fisher Scientific). Illumina library preparation, sequencing, and analysis was performed by the IMB Sequencing Facility.

Read mapping, counting and differential expression analysis

All analysis was performed using RStudio (RStudio Team, 2016). Reads were mapped to the GRCz11 version of the zebrafish genome downloaded from Ensembl using the Rsubread package (Liao et al., 2019). Counting of reads was also performed using Rsubread. Differential expression analysis was performed using the limma package (Law et al., 2014; Ritchie et al., 2015). Figures were produced with the ggplot2 package (Wickham, 2009).

qPCR

qPCR was performed on amplified RNA samples prepared as described above using the Smart-seq2 protocol (Picelli et al., 2014). Amplified cDNA was diluted 1:80 and used in a reaction with SYBR green (ThermoFisher) and primers specific to the gene of interest. Each gene was analysed in triplicate for each sample. Amplification efficiencies for each primer pair were calculated using LineRegPCR (Ramakers et al., 2003). Four housekeeping genes were examined to identify the most stable gene using RefFinder (Xie et al., 2012) and the most stable gene used to normalise samples with Q-Gene (Muller et al., 2002; Simon, 2003).

GO term analysis

To examine GO terms exclusively associated with the endocardium and vascular endothelium, an overrepresentation test was performed using the PANTHER database available at http://pantherdb.org/ (Mi et al., 2019). Upregulated and downregulated gene lists were filtered from differentially expressed genes with a $\log_2$ Fold Change cut-off of ≥1 and ≤-1 and an adjusted P-value of ≤0.05, applied using TREAT (McCarthy and Smyth, 2009) in RStudio (RStudio Team, 2016).

In situ hybridisation

In situ hybridisation was performed as previously described (Grassini et al., 2018). All RNA probes but one were transcribed from plasmids. Plasmids containing the *spi1b* and *gfp* coding sequences were a kind gift from the laboratory of Ben Hogan. To generate the *etv2* plasmid used in this study, pCS2+ vector was prepared by digesting empty vector with BamHI and XbaI. The *etv2* cDNA sequence was amplified by PCR using the primer sequences below, the PCR product was purified and inserted into the digested pCS2+ vector using circular polymerase extension cloning (Quan and Tian, 2011).

*etv2*_cDNA_F:

AAGCTACTTGTTCTTTTTGCAGGATCTGTCAAAACCCCTGATATAGTG

*etv2*_cDNA_R:

TGGATCTACGTAATACGACTCACTATAGTTCTAGCAATCTGCTGCAAAGTCC

For the *npas4l* probe, primers from Reischauer et al., 2016 were used, with the Sp6 promoter on the R primer to amplify the template from cDNA.

For all *in situ* data, at least two experimental replicates were performed. Biological replicates are indicated in images, depicting the *n* for each condition.

CRISPR/Cas9 mutagenesis

To generate the mutant lines used in this study CRISPR/Cas9 mutagenesis was performed as described previously (Capon et al., 2017; Gagnon et al., 2014). Embryos and fish were screened for indels by HRMA (Dahlem et al., 2012) and carriers sequenced to identify frame-shift mutations that truncate the protein (Fig S7). Specific details of each of the mutants generated for this study are below.

The *npas4l* mutant allele carries a 7bp deletion in the second exon of the *npas4l* gene and is referred to as the *npas4l*^{uq14ks} allele. This deletion creates a frame-shift at amino acid 47 (E47D) followed by 7 missense amino acids before a premature stop codon. The predicted protein product of this truncated allele partially retains the DNA-binding bHLH domain but lacks the two PAS domains as well as the three transcriptional activation domains (Fig S7). An incross of fish carrying the *npas4l*^{uq14ks} allele produces a mutant phenotype in 25% of the progeny, consisting of a bell-shaped heart and lacking the endocardial layer at 48 hpf, as previously described for the *cloche* mutant (Stainier et al., 1995).

The *etv2* mutant allele has a 23bp deletion in the exon five of the *etv2* gene and is referred to as the *etv2*^{uq13ks} allele. This deletion creates a frame-shift at amino acid 137 (P137R) resulting in the incorporation of 10 missense amino acids before a stop codon. This truncated allele completely lacks the DNA-binding ETS domain (Fig S7). An incross of fish heterozygous for the *etv2*^{uq13ks} allele produces a mutant phenotype in 25% of progeny. This phenotype is characterised by a lack of blood circulation, as previously reported (Pham et al., 2007), collapsed heart at 48 hpf, impaired intersegmental vessel sprouting and extopic expression of vascular markers in skeletal muscle, as previously reported (Chestnut et al., 2020).

Morpholino oligonucleotide (MO) reagents

All MOs were ordered from Genetools, LLC. The MO sequences and concentrations used are as follows: 0.8 pmol *scl* – AATGCTCTTACCATCGTTGATTTCA (Dooley et al., 2005), 1 pmol *gata1a* – CTGCAAGTGTAGTATTGAAGATGTC (Galloway et

al., 2005), 1.77 pmol *spi1b* – GATATACTGATACTCCATTGGTGGT (Rhodes et al., 2005).

Heat-shock treatment

All heat-shock experiments were performed by outcrossing heterozygous carriers of either *Tg(hsp70l:nog3)* or *Tg(hsp70l:bmp2b)* to non-heat-shock transgenic lines, generating clutches of embryos with an expected Mendelian ration of 50% transgenic-to-50% wildtype. Heat-shock was performed at 39°C for 30 minutes in pre-warmed medium on mixed-genotype clutches of embryos. Following heat-shock, embryos were returned to a 28.5°C incubator until fixation.

***In vitro* transcription**

To produce the *etv2* and *npas4l* RNA for overexpression experiments, plasmids containing the *etv2* or *npas4l* coding sequences were linearised by restriction enzyme digest. Purified linearised plasmid was transcribed with the mMessage mMachine SP6 kit (Thermo Fisher Scientific). RNA was purified with the RNA clean and concentrator kit (Zymo Research). The *npas4l* plasmid was a kind gift from the laboratory of Didier Stainier.

Imaging

Embryos from *in situ* hybridisation were dehydrated, cleared in murray's solution (a 2:1 ratio of benzyl benzoate:benzyl alcohol) and imaged using an Olympus BX51 Microscope with Olympus DP70 CCD camera on glass slides with bridged coverslips.

For wide-field bright-field and fluorescence imaging, embyros were mounted in 3% methyl cellulose (Sigma, cat. no. M0387) and imaged using a Leica M165 FC stereo microscope with a DFC425 C camera.

Confocal imaging was performed on Zeiss LSM710 FCS, LSM880 with 40x immersive objective with Airyscan detector and Zeiss LSM900 confocal microscope. Live embryos were mounted for confocal imaging using 0.5-1% low-melting agarose (Sigma, cat. no. A9414) in glass-bottom petri dishes. Fixed embryos for confocal imaging were de-yolked using lash tools and mounted in vectashield with DAPI (Vector laboratories, cat. no. H-1200) on glass-slides with coverslips. Alternatively, the fixed embryos were mounted in 1% agarose and imaged with an immersive 40x objective.

All images were processed using FIJI (Schindelin et al., 2012) and/or Imaris software. For all imaging involving embryos involving crosses of *etv2* or *npas4l* mutants or *Tg(hsp70l:nog3)* or *Tg(hsp70l:bmp2b)* lines, imaging was performed blind and samples genotyped post-imaging to assign genotypes.

Quantification of pErk and pSmad staining (shown in Fig 5) was generated using a mask in Imaris software. pErk and pSmad signal was quantified within the mask regions. Example masks are represented in Fig S8.

Statistical testing

A student's t-test with Welch's correction was performed using Prism software. For data-sets containing 3 or more conditions, one-way ANOVA testing with Tukey's post hoc test for multiple comparisons was performed using Prism software.

Acknowledgements:

We acknowledge members of the Smith laboratory for useful discussions and Ben Hogan and lab members for discussions and sharing of reagents. We thank the UQBR aquatics team and University of Melbourne DrUM facility for animal husbandry; the QBI Flow Cytometry Facility for performing FACS; the IMB Sequencing Facility for advice, library preparation and sequencing; and the Australian Cancer Research Foundation's Cancer Ultrastructure and Function Facility at the IMB and University

of Melbourne Biological Optimal Microscopy Platform (BOMP) for assistance and maintenance of microscopes.

Funding:

SJC was supported by an Australian Postgraduate Award (APA). The research was funded with external support from ARC Discovery Project grants (DP170101217 and DP200103642) and internal support from the University of Queensland and University of Melbourne.

Competing interests:

The authors declare that no competing interests exist.

References

Asakawa, K., Suster, M. L., Mizusawa, K., Nagayoshi, S., Kotani, T., Urasaki, A., Kishimoto, Y., Hibi, M. and Kawakami, K. (2008). Genetic dissection of neural circuits by Tol2 transposon-mediated Gal4 gene and enhancer trapping in zebrafish. *Proc Natl Acad Sci U S A* **105**, 1255-1260.

Baek, S., Oh, T. G., Secker, G., Sutton, D. L., Okuda, K. S., Paterson, S., Bower, N. I., Toubia, J., Koltowska, K., Capon, S. J., et al. (2019). The Alternative Splicing Regulator Nova2 Constrains Vascular Erk Signaling to Limit Specification of the Lymphatic Lineage. *Dev Cell* **49**, 279-292 e275.

Brend, T. and Holley, S. A. (2009). Zebrafish whole mount high-resolution double fluorescent in situ hybridization. *J Vis Exp*.

Bussmann, J., Bakkers, J. and Schulte-Merker, S. (2007). Early endocardial morphogenesis requires Scl/Tal1. *PLoS Genet* **3**, e140.

Bussmann, J., Bos, F. L., Urasaki, A., Kawakami, K., Duckers, H. J. and Schulte-Merker, S. (2010). Arteries provide essential guidance cues for lymphatic endothelial cells in the zebrafish trunk. *Development* **137**, 2653-2657.

Capon, S. J., Baillie, G. J., Bower, N. I., da Silva, J. A., Paterson, S., Hogan, B. M., Simons, C. and Smith, K. A. (2017). Utilising polymorphisms to achieve allele-specific genome editing in zebrafish. *Biol Open* **6**, 125-131.

Chen, J. N. and Fishman, M. C. (1996). Zebrafish tinman homolog demarcates the heart field and initiates myocardial differentiation. *Development* **122**, 3809-3816.

Chestnut, B., Casie Chetty, S., Koenig, A. L. and Sumanas, S. (2020). Single-cell transcriptomic analysis identifies the conversion of zebrafish Etv2-deficient vascular progenitors into skeletal muscle. *Nat Commun* **11**, 2796.

Chocron, S., Verhoeven, M. C., Rentzsch, F., Hammerschmidt, M. and Bakkers, J. (2007). Zebrafish Bmp4 regulates left-right asymmetry at two distinct developmental time points. *Dev Biol* **305**, 577-588.

Clay, H. and Ramakrishnan, L. (2005). Multiplex fluorescent in situ hybridization in zebrafish embryos using tyramide signal amplification. *Zebrafish* **2**, 105-111.

Dahlem, T. J., Hoshijima, K., Jurynec, M. J., Gunther, D., Starker, C. G., Locke, A. S., Weis, A. M., Voytas, D. F. and Grunwald, D. J. (2012). Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. *PLoS Genet* **8**, e1002861.

Davidson, A. J. and Zon, L. I. (2004). The 'definitive' (and 'primitive') guide to zebrafish hematopoiesis. *Oncogene* **23**, 7233-7246.

De Angelis, J. E., Lagendijk, A. K., Chen, H., Tromp, A., Bower, N. I., Tunny, K. A., Brooks, A. J., Bakkers, J., Francois, M., Yap, A. S., et al. (2017). Tmem2 Regulates Embryonic Vegf Signaling by Controlling Hyaluronic Acid Turnover. *Dev Cell* **40**, 123-136.

de la Pompa, J. L., Timmerman, L. A., Takimoto, H., Yoshida, H., Elia, A. J., Samper, E., Potter, J., Wakeham, A., Marengere, L., Langille, B. L., et al. (1998). Role of the NF-ATc transcription factor in morphogenesis of cardiac valves and septum. *Nature* **392**, 182-186.

de Pater, E., Ciampricotti, M., Priller, F., Veerkamp, J., Strate, I., Smith, K., Lagendijk, A. K., Schilling, T. F., Herzog, W., Abdelilah-Seyfried, S., et al. (2012). Bmp signaling exerts opposite effects on cardiac differentiation. *Circ Res* **110**, 578-587.

Derynck, R. and Zhang, Y. E. (2003). Smad-dependent and Smad-independent pathways in TGF-beta family signalling. *Nature* **425**, 577-584.

Dietrich, A. C., Lombardo, V. A., Veerkamp, J., Priller, F. and Abdelilah-Seyfried, S. (2014). Blood flow and Bmp signaling control endocardial chamber morphogenesis. *Dev Cell* **30**, 367-377.

Dooley, K. A., Davidson, A. J. and Zon, L. I. (2005). Zebrafish scl functions independently in hematopoietic and endothelial development. *Dev Biol* **277**, 522-536.

Ferdous, A., Caprioli, A., Iacovino, M., Martin, C. M., Morris, J., Richardson, J. A., Latif, S., Hammer, R. E., Harvey, R. P., Olson, E. N., et al. (2009). Nkx2-5 transactivates the Ets-related protein 71 gene and specifies an endothelial/endocardial fate in the developing embryo. *Proc Natl Acad Sci U S A* **106**, 814-819.

Gagnon, J. A., Valen, E., Thyme, S. B., Huang, P., Akhmetova, L., Pauli, A., Montague, T. G., Zimmerman, S., Richter, C. and Schier, A. F. (2014). Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS One* **9**, e98186.

Galloway, J. L., Wingert, R. A., Thisse, C., Thisse, B. and Zon, L. I. (2005). Loss of gata1 but not gata2 converts erythropoiesis to myelopoiesis in zebrafish embryos. *Dev Cell* **8**, 109-116.

Gering, M., Rodaway, A. R., Gottgens, B., Patient, R. K. and Green, A. R. (1998). The SCL gene specifies haemangioblast development from early mesoderm. *EMBO J* **17**, 4029-4045.

Grassini, D. R., da Silva, J., Hall, T. E., Baillie, G. J., Simons, C., Parton, R. G., Hogan, B. M. and Smith, K. A. (2019). Myosin Vb is required for correct trafficking of N-cadherin and cardiac chamber ballooning. *Dev Dyn* **248**, 284-295.

Grassini, D. R., Lagendijk, A. K., De Angelis, J. E., Da Silva, J., Jeanes, A., Zettler, N., Bower, N. I., Hogan, B. M. and Smith, K. A. (2018). Nppa and Nppb act redundantly during zebrafish cardiac development to confine AVC marker expression and reduce cardiac jelly volume. *Development* **145**.

Hammerschmidt, M. and Nusslein-Volhard, C. (1993). The expression of a zebrafish gene homologous to Drosophila snail suggests a conserved function in invertebrate and vertebrate gastrulation. *Development* **119**, 1107-1118.

Harris, I. S. and Black, B. L. (2010). Development of the endocardium. *Pediatr Cardiol* **31**, 391-399.

Hofbauer, P., Jahnel, S. M., Papai, N., Giesshammer, M., Deyett, A., Schmidt, C., Penc, M., Tavernini, K., Grdseloff, N., Meledeth, C., et al. (2021). Cardioids reveal self-organizing principles of human cardiogenesis. *Cell* **184**, 3299-3317 e3222.

Hogan, B. M., Bos, F. L., Bussmann, J., Witte, M., Chi, N. C., Duckers, H. J. and Schulte-Merker, S. (2009). Ccbe1 is required for embryonic lymphangiogenesis and venous sprouting. *Nat Genet* **41**, 396-398.

Hogan, B. M., Layton, J. E., Pyati, U. J., Nutt, S. L., Hayman, J. W., Varma, S., Heath, J. K., Kimelman, D. and Lieschke, G. J. (2006). Specification of the primitive myeloid precursor pool requires signaling through Alk8 in zebrafish. *Curr Biol* **16**, 506-511.

Hu, P., Liu, J., Zhao, J., Wilkins, B. J., Lupino, K., Wu, H. and Pei, L. (2018). Single-nucleus transcriptomic survey of cell diversity and functional maturation in postnatal mammalian hearts. *Genes Dev* **32**, 1344-1357.

Jin, S. W., Beis, D., Mitchell, T., Chen, J. N. and Stainier, D. Y. (2005). Cellular and molecular analyses of vascular tube and lumen formation in zebrafish. *Development* **132**, 5199-5209.

Kalev-Zylinska, M. L., Horsfield, J. A., Flores, M. V., Postlethwait, J. H., Vitas, M. R., Baas, A. M., Crosier, P. S. and Crosier, K. E. (2002). Runx1 is required for zebrafish blood and vessel development and expression of a human RUNX1-CBF2T1 transgene advances a model for studies of leukemogenesis. *Development* **129**, 2015-2030.

Keegan, B. R., Feldman, J. L., Begemann, G., Ingham, P. W. and Yelon, D. (2005). Retinoic acid signaling restricts the cardiac progenitor pool. *Science* **307**, 247-249.

Kisanuki, Y. Y., Hammer, R. E., Miyazaki, J., Williams, S. C., Richardson, J. A. and Yanagisawa, M. (2001). Tie2-Cre transgenic mice: a new model for endothelial cell-lineage analysis in vivo. *Dev Biol* **230**, 230-242.

Koopman, C. D., De Angelis, J., Iyer, S. P., Verkerk, A. O., Da Silva, J., Berecki, G., Jeanes, A., Baillie, G. J., Paterson, S., Uribe, V., et al. (2021). The zebrafish grime mutant uncovers an evolutionarily conserved role for Tmem161b in the control of cardiac rhythm. *Proc Natl Acad Sci U S A* **118**.

Law, C. W., Chen, Y., Shi, W. and Smyth, G. K. (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* **15**, R29.

Lawson, N. D. and Weinstein, B. M. (2002). In vivo imaging of embryonic vascular development using transgenic zebrafish. *Dev Biol* **248**, 307-318.

Lee, D., Park, C., Lee, H., Lugus, J. J., Kim, S. H., Arentson, E., Chung, Y. S., Gomez, G., Kyba, M., Lin, S., et al. (2008). ER71 acts downstream of BMP, Notch, and Wnt signaling in blood and vessel progenitor specification. *Cell Stem Cell* **2**, 497-507.

Liao, Y., Smyth, G. K. and Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res* **47**, e47.

Lieschke, G. J., Oates, A. C., Paw, B. H., Thompson, M. A., Hall, N. E., Ward, A. C., Ho, R. K., Zon, L. I. and Layton, J. E. (2002). Zebrafish SPI-1 (PU.1) marks a site of myeloid development independent of primitive erythropoiesis: implications for axial patterning. *Dev Biol* **246**, 274-295.

Liu, F., Walmsley, M., Rodaway, A. and Patient, R. (2008). Fli1 acts at the top of the transcriptional network driving blood and endothelial development. *Curr Biol* **18**, 1234-1240.

Ma, H., Liu, Z., Yang, Y., Feng, D., Dong, Y., Garbutt, T. A., Hu, Z., Wang, L., Luan, C., Cooper, C. D., et al. (2021). Functional coordination of non-myocytes plays a key role in adult zebrafish heart regeneration. *EMBO Rep* **22**, e52901.

Mantri, M., Scuderi, G. J., Abedini-Nassab, R., Wang, M. F. Z., McKellar, D., Shi, H., Grodner, B., Butcher, J. T. and De Vlaminck, I. (2021). Spatiotemporal single-cell RNA sequencing of developing chicken hearts identifies interplay between cellular differentiation and morphogenesis. *Nat Commun* **12**, 1771.

Marass, M., Beisaw, A., Gerri, C., Luzzani, F., Fukuda, N., Gunther, S., Kuenne, C., Reischauer, S. and Stainier, D. Y. R. (2019). Genome-wide strategies reveal target genes of Npas4l associated with vascular development in zebrafish. *Development* **146**.

Markwald, R. R., Fitzharris, T. P. and Manasek, F. J. (1977). Structural development of endocardial cushions. *Am J Anat* **148**, 85-119.

Markwald, R. R., Fitzharris, T. P. and Smith, W. N. (1975). Sturctural analysis of endocardial cytodifferentiation. *Dev Biol* **42**, 160-180.

McCarthy, D. J. and Smyth, G. K. (2009). Testing significance relative to a fold-change threshold is a TREAT. *Bioinformatics* **25**, 765-771.

Meyer, D. and Birchmeier, C. (1995). Multiple essential functions of neuregulin in development. *Nature* **378**, 386-390.

Mi, H., Muruganujan, A., Ebert, D., Huang, X. and Thomas, P. D. (2019). PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res* **47**, D419-D426.

Mikryukov, A. A., Mazine, A., Wei, B., Yang, D., Miao, Y., Gu, M. and Keller, G. M. (2021). BMP10 Signaling Promotes the Development of Endocardial Cells from Human Pluripotent Stem Cell-Derived Cardiovascular Progenitors. *Cell Stem Cell* **28**, 96-111 e117.

Misfeldt, A. M., Boyle, S. C., Tompkins, K. L., Bautch, V. L., Labosky, P. A. and Baldwin, H. S. (2009). Endocardial cells are a distinct endothelial lineage derived from Flk1+ multipotent cardiovascular progenitors. *Dev Biol* **333**, 78-89.

Morton, S. U., Quiat, D., Seidman, J. G. and Seidman, C. E. (2022). Genomic frontiers in congenital heart disease. *Nat Rev Cardiol* **19**, 26-42.

Muller, P. Y., Janovjak, H., Miserez, A. R. and Dobbie, Z. (2002). Processing of gene expression data generated by quantitative real-time RT-PCR. *Biotechniques* **32**, 1372-1374, 1376, 1378-1379.

Neri, T., Hiriart, E., van Vliet, P. P., Faure, E., Norris, R. A., Farhat, B., Jagla, B., Lefrancois, J., Sugi, Y., Moore-Morris, T., et al. (2019). Human pre-valvular endocardial cells derived from pluripotent stem cells recapitulate cardiac pathophysiological valvulogenesis. *Nat Commun* **10**, 1929.

Okuda, K. S., Astin, J. W., Misa, J. P., Flores, M. V., Crosier, K. E. and Crosier, P. S. (2012). lyve1 expression reveals novel lymphatic vessels and new mechanisms for lymphatic vessel development in zebrafish. *Development* **139**, 2381-2391.

Okuda, K. S., Baek, S. and Hogan, B. M. (2018). Visualization and Tools for Analysis of Zebrafish Lymphatic Development. *Methods Mol Biol* **1846**, 55-70.

Palencia-Desai, S., Kohli, V., Kang, J., Chi, N. C., Black, B. L. and Sumanas, S. (2011). Vascular endothelial and endocardial progenitors differentiate as cardiomyocytes in the absence of Etsrp/Etv2 function. *Development* **138**, 4721-4732.

Palencia-Desai, S., Rost, M. S., Schumacher, J. A., Ton, Q. V., Craig, M. P., Baltrunaite, K., Koenig, A. L., Wang, J., Poss, K. D., Chi, N. C., et al. (2015). Myocardium and BMP signaling are required for endocardial differentiation. *Development* **142**, 2304-2315.

Palpant, N. J., Pabon, L., Roberts, M., Hadland, B., Jones, D., Jones, C., Moon, R. T., Ruzzo, W. L., Bernstein, I., Zheng, Y., et al. (2015). Inhibition of beta-catenin signaling respecifies anterior-like endothelium into beating human cardiomyocytes. *Development* **142**, 3198-3209.

Patterson, L. J., Gering, M. and Patient, R. (2005). Scl is required for dorsal aorta as well as blood formation in zebrafish embryos. *Blood* **105**, 3502-3511.

Perez-Pomares, J. M., Gonzalez-Rosa, J. M. and Munoz-Chapuli, R. (2009). Building the vertebrate heart - an evolutionary approach to cardiac development. *Int J Dev Biol* **53**, 1427-1443.

Pham, V. N., Lawson, N. D., Mugford, J. W., Dye, L., Castranova, D., Lo, B. and Weinstein, B. M. (2007). Combinatorial function of ETS transcription factors in the developing vasculature. *Dev Biol* **303**, 772-783.

Picelli, S., Faridani, O. R., Bjorklund, A. K., Winberg, G., Sagasser, S. and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc* **9**, 171-181.

Proulx, K., Lu, A. and Sumanas, S. (2010). Cranial vasculature in zebrafish forms by angioblast cluster-derived angiogenesis. *Dev Biol* **348**, 34-46.

Quan, J. and Tian, J. (2011). Circular polymerase extension cloning for high-throughput cloning of complex and combinatorial DNA libraries. *Nat Protoc* **6**, 242-251.

Ramakers, C., Ruijter, J. M., Deprez, R. H. and Moorman, A. F. (2003). Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci Lett* **339**, 62-66.

Red-Horse, K., Ueno, H., Weissman, I. L. and Krasnow, M. A. (2010). Coronary arteries form by developmental reprogramming of venous cells. *Nature* **464**, 549-553.

Reischauer, S., Stone, O. A., Villasenor, A., Chi, N., Jin, S. W., Martin, M., Lee, M. T., Fukuda, N., Marass, M., Witty, A., et al. (2016). Cloche is a bHLH-PAS transcription factor that drives haemato-vascular specification. *Nature* **535**, 294-298.

Rhodes, J., Hagen, A., Hsu, K., Deng, M., Liu, T. X., Look, A. T. and Kanki, J. P. (2005). Interplay of pu.1 and gata1 determines myelo-erythroid progenitor cell fate in zebrafish. *Dev Cell* **8**, 97-108.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W. and Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47.

Row, R. H., Pegg, A., Kinney, B. A., Farr, G. H., 3rd, Maves, L., Lowell, S., Wilson, V. and Martin, B. L. (2018). BMP and FGF signaling interact to pattern mesoderm by controlling basic helix-loop-helix transcription factor activity. *Elife* **7**.

Saint-Jean, L., Barkas, N., Harmelink, C., Tompkins, K. L., Oakey, R. J. and Baldwin, H. S. (2019). Myocardial differentiation is dependent upon endocardial signaling during early cardiogenesis in vitro. *Development* **146**.

Saxon, J. G., Baer, D. R., Barton, J. A., Hawkins, T., Wu, B., Trusk, T. C., Harris, S. E., Zhou, B., Mishina, Y. and Sugi, Y. (2017). BMP2 expression in the endocardial lineage is required for AV endocardial cushion maturation and remodeling. *Dev Biol* **430**, 113-128.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682.

Schoenebeck, J. J., Keegan, B. R. and Yelon, D. (2007). Vessel and blood specification override cardiac potential in anterior mesoderm. *Dev Cell* **13**, 254-267.

Schumacher, J. A., Bloomekatz, J., Garavito-Aguilar, Z. V. and Yelon, D. (2013). tal1 Regulates the formation of intercellular junctions and the maintenance of identity in the endocardium. *Dev Biol* **383**, 214-226.

Shi, X., Zirbes, K. M., Rasmussen, T. L., Ferdous, A., Garry, M. G., Koyano-Nakagawa, N. and Garry, D. J. (2015). The transcription factor Mesp1 interacts with cAMP-responsive element binding protein 1 (Creb1) and coactivates Ets variant 2 (Etv2) gene expression. *J Biol Chem* **290**, 9614-9625.

Shin, M., Beane, T. J., Quillien, A., Male, I., Zhu, L. J. and Lawson, N. D. (2016a). Vegfa signals through ERK to promote angiogenesis, but not artery differentiation. *Development* **143**, 3796-3805.

Shin, M., Male, I., Beane, T. J., Villefranc, J. A., Kok, F. O., Zhu, L. J. and Lawson, N. D. (2016b). Vegfc acts through ERK to induce sprouting and differentiation of trunk lymphatic progenitors. *Development* **143**, 3785-3795.

Shivdasani, R. A., Mayer, E. L. and Orkin, S. H. (1995). Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. *Nature* **373**, 432-434.

Simoes, F. C., Peterkin, T. and Patient, R. (2011). Fgf differentially controls cross-antagonism between cardiac and haemangioblast regulators. *Development* **138**, 3235-3245.

Simon, P. (2003). Q-Gene: processing quantitative real-time RT-PCR data. *Bioinformatics* **19**, 1439-1440.

Smith, K. A., Chocron, S., von der Hardt, S., de Pater, E., Soufan, A., Bussmann, J., Schulte-Merker, S., Hammerschmidt, M. and Bakkers, J. (2008). Rotation and asymmetric development of the zebrafish heart requires directed migration of cardiac progenitor cells. *Dev Cell* **14**, 287-297.

Snider, P., Simmons, O., Wang, J., Hoang, C. Q. and Conway, S. J. (2014). Ectopic Noggin in a Population of Nfatc1 Lineage Endocardial Progenitors Induces Embryonic Lethality. *J Cardiovasc Dev Dis* **1**, 214-236.

Spanjaard, B., Hu, B., Mitic, N., Olivares-Chauvet, P., Janjuha, S., Ninov, N. and Junker, J. P. (2018). Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. *Nat Biotechnol* **36**, 469-473.

Stainier, D. Y., Weinstein, B. M., Detrich, H. W., 3rd, Zon, L. I. and Fishman, M. C. (1995). Cloche, an early acting zebrafish gene, is required by both the endothelial and hematopoietic lineages. *Development* **121**, 3141-3150.

Stankunas, K., Hang, C. T., Tsun, Z. Y., Chen, H., Lee, N. V., Wu, J. I., Shang, C., Bayle, J. H., Shou, W., Iruela-Arispe, M. L., et al. (2008). Endocardial Brg1 represses ADAMTS1 to maintain the microenvironment for myocardial morphogenesis. *Dev Cell* **14**, 298-311.

Sumanas, S., Gomez, G., Zhao, Y., Park, C., Choi, K. and Lin, S. (2008). Interplay among Etsrp/ER71, Scl, and Alk8 signaling controls endothelial and myeloid cell formation. *Blood* **111**, 4500-4510.

Thompson, M. A., Ransom, D. G., Pratt, S. J., MacLennan, H., Kieran, M. W., Detrich, H. W., 3rd, Vail, B., Huber, T. L., Paw, B., Brownlie, A. J., et al. (1998). The cloche and spadetail genes differentially affect hematopoiesis and vasculogenesis. *Dev Biol* **197**, 248-269.

Tian, X., Li, Y., He, L., Zhang, H., Huang, X., Liu, Q., Pu, W., Zhang, L., Li, Y., Zhao, H., et al. (2017). Identification of a hybrid myocardial zone in the mammalian heart after birth. *Nat Commun* **8**, 87.

Trinh, L. A. and Stainier, D. Y. (2004). Fibronectin regulates epithelial organization during myocardial migration in zebrafish. *Dev Cell* **6**, 371-382.

Tucker, N. R., Chaffin, M., Fleming, S. J., Hall, A. W., Parsons, V. A., Bedi, K. C., Jr., Akkad, A. D., Herndon, C. N., Arduini, A., Papangeli, I., et al. (2020). Transcriptional and Cellular Diversity of the Human Heart. *Circulation* **142**, 466-482.

Uribe, V., Ramadass, R., Dogra, D., Rasouli, S. J., Gunawan, F., Nakajima, H., Chiba, A., Reischauer, S., Mochizuki, N. and Stainier, D. Y. R. (2018). In vivo analysis of cardiomyocyte proliferation during trabeculation. *Development* **145**.

Van Handel, B., Montel-Hagen, A., Sasidharan, R., Nakano, H., Ferrari, R., Boogerd, C. J., Schredelseker, J., Wang, Y., Hunter, S., Org, T., et al. (2012). Scl represses cardiomyogenesis in prospective hemogenic endothelium and endocardium. *Cell* **150**, 590-605.

Wang, J., Sridurongrit, S., Dudas, M., Thomas, P., Nagy, A., Schneider, M. D., Epstein, J. A. and Kaartinen, V. (2005). Atrioventricular cushion transformation is mediated by ALK2 in the developing mouse heart. *Dev Biol* **286**, 299-310.

Weinberg, E. S., Allende, M. L., Kelly, C. S., Abdelhamid, A., Murakami, T., Andermann, P., Doerre, O. G., Grunwald, D. J. and Riggleman, B. (1996). Developmental regulation of zebrafish MyoD in wild-type, no tail and spadetail embryos. *Development* **122**, 271-280.

Wickham, H. (2009). ggplot2. Elegant Graphics for Data Analysis. *Springer Science & Business Media.*

Wong, K. S., Rehn, K., Palencia-Desai, S., Kohli, V., Hunter, W., Uhl, J. D., Rost, M. S. and Sumanas, S. (2012). Hedgehog signaling is required for differentiation of endocardial progenitors in zebrafish. *Dev Biol* **361**, 377-391.

Xie, F., Xiao, P., Chen, D., Xu, L. and Zhang, B. (2012). miRDeepFinder: a miRNA analysis tool for deep sequencing of plant small RNAs. *Plant Mol Biol*.

Yelon, D., Horne, S. A. and Stainier, D. Y. (1999). Restricted expression of cardiac myosin genes reveals regulated aspects of heart tube assembly in zebrafish. *Dev Biol* **214**, 23-37.

Figures

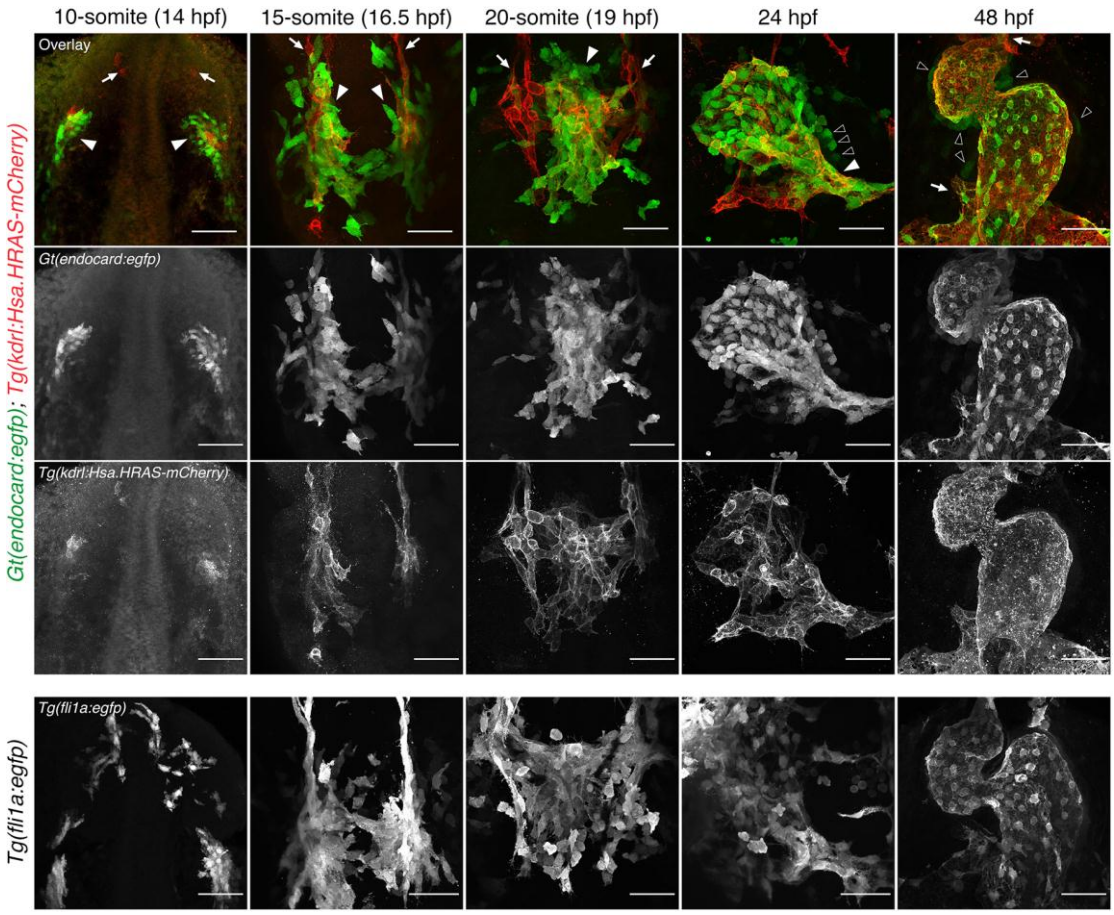


Fig 1. Time-course analysis of *Gt(endocard:egfp)* fluorescence demonstrates enriched endocardial expression at somitogenesis stages in the zebrafish embryo. Immunofluorescence staining of *Gt(endocard:egfp)*; *Tg(kdrl:Hsa.HRAS-mCherry)* and *Tg(fli1a:egfp)* embryos from the 10s through to 48 hpf. Anterior to the top. Anterior views at 48 hpf, all other images show dorsal views. White arrowheads label kdrl^{+}/endocard:egfp^{+} endocardial cells. White arrows label kdrl^{+}/endocard:egfp^{-} vascular endothelial cells. Black arrowheads label presumptive myocardial cells. Scale bars represent 50 μm.

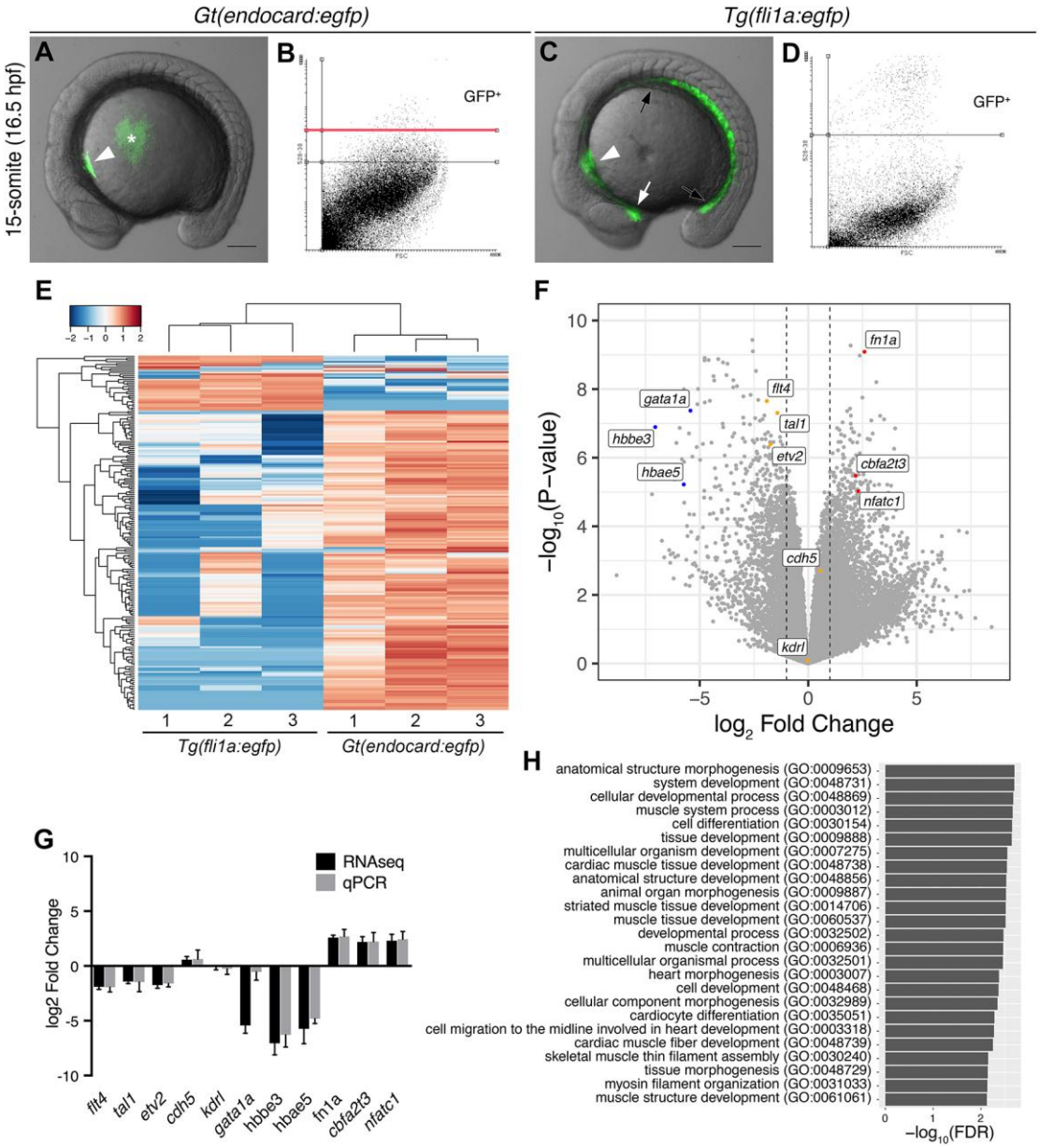


Fig 2. The endocardium is molecularly distinct from the vascular endothelium. Live images of *Gt(endocard:egfp)* and *Tg(fli1a:egfp)* embryos at the 15 somite-stage (16.5 hpf) showing GFP fluorescence (**A** and **C**). GFP fluorescence can be seen in the endocardial progenitors (white arrowhead) of *Gt(endocard:egfp)* embryos (**A**), some auto-fluorescence is seen in the yolk (asterisk). In *Tg(fli1a:egfp)* embryos (**C**), GFP fluorescence can be seen in the entire vascular endothelium of the trunk (black arrows) and head (white arrow) as well as the endocardium (white arrowhead). Scale bars represent 50μm. FACS plots from dissociated *Gt(endocard:egfp)* and *Tg(fli1a:egfp)* embryos show the gating strategy to collect GFP^+ cells, two gates were used to collect high and low GFP^+ populations for *Gt(endocard:egfp)* embryos, and all analysis was performed on cells collected above the red line (**B**). A single

population was collected for *Tg(fli1a:egfp)* embryos (**D**). Hierarchical clustering of the top 200 most variable genes across all samples shows that samples segregate according to transgene (**E**). Volcano plot showing the spread of differentially expressed genes with a subset of genes labelled, the dotted lines show a $\log_2$ Fold Change of 1 or -1 (**F**). Validation of the subset of genes highlighted in **F** by qPCR (**G**), $\log_2$ Fold Change is represented, error bars indicate 95% confidence intervals**.** GO term analysis using a PANTHER overrepresentation test was performed on TREAT lists of genes upregulated in the *Gt(endocard:egfp)*^{High} samples relative to *Tg(fli1a:egfp)* (**H**). The 25 most significant GO terms for genes enriched in the endocardium are shown in (**H**).

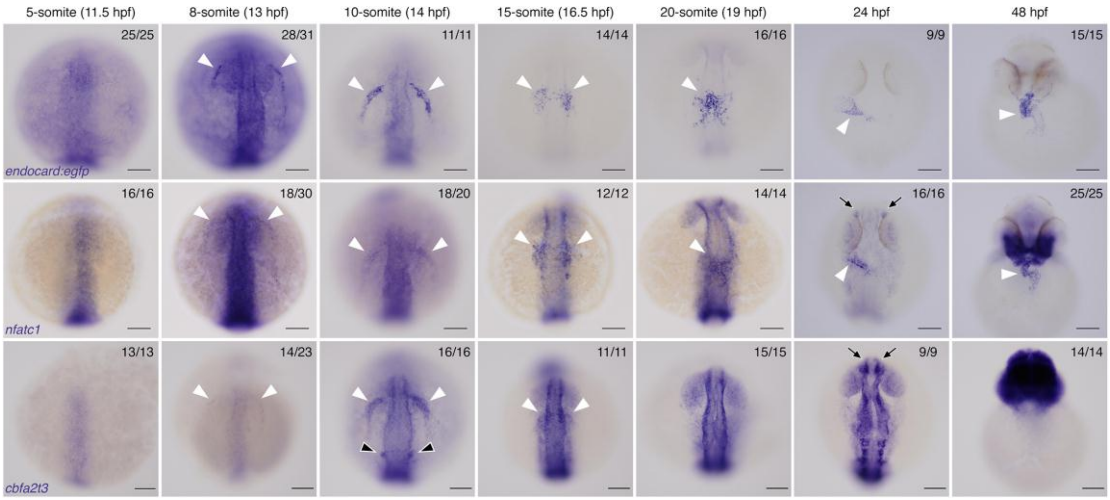


Fig 3. Endocardial markers are first expressed at 8s in zebrafish embryos. *In situ* hybridisation for *endocard:egfp*, *nfatc1* and *cbfa2t3* expression in embryos from 5s (11.5 hpf) through to 48 hpf, showing expression emerging in the endocardium from 8s onwards. White arrowheads show the expression domains corresponding to endocardial cells and their progenitors. *nfatc1* expression is also observed in the olfactory placode (black arrows) as previously described. *cbfa2t3* expression is also seen in otic vesicle progenitors (black arrowheads) and throughout the neural system (black arrows). Anterior to the top in all images. Anterior views at 48 hpf, all other images show dorsal views. Scale bars represent 100 μm. The number of embryos matching the image shown is indicated in the top right of each image.

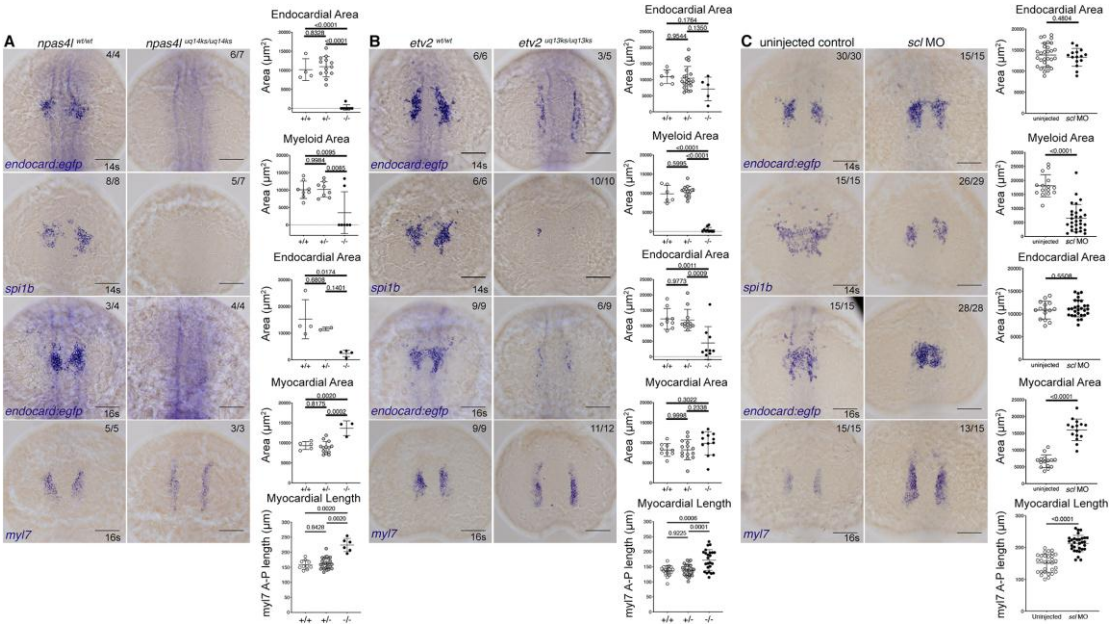


Fig 4. ***Gt(endocard:egfp)*** **expression is reduced in** ***npas4l*** **and** ***etv2*** **mutants.** *In situ* hybridisation for *endocard:egfp* and *spi1b* expression at the 14s and *endocard:egfp* and *myl7* expression at the 16s in (**A)** *npas4l*^{uq14ks} mutants, (**B)** *etv2*^{uq13ks} mutants and (**C)** *scl* morphants. Quantification of the expression domain was measured by calculating the surface area of the stain or the length of the staining domain. These measurements are shown in adjacent graphs, where wildtype (+/+), heterozygous (+/-) and mutant (-/-) quantification is depicted. Dorsal views, with anterior to the top. Scale bars represent 100 µm. P-values are present in graphs.

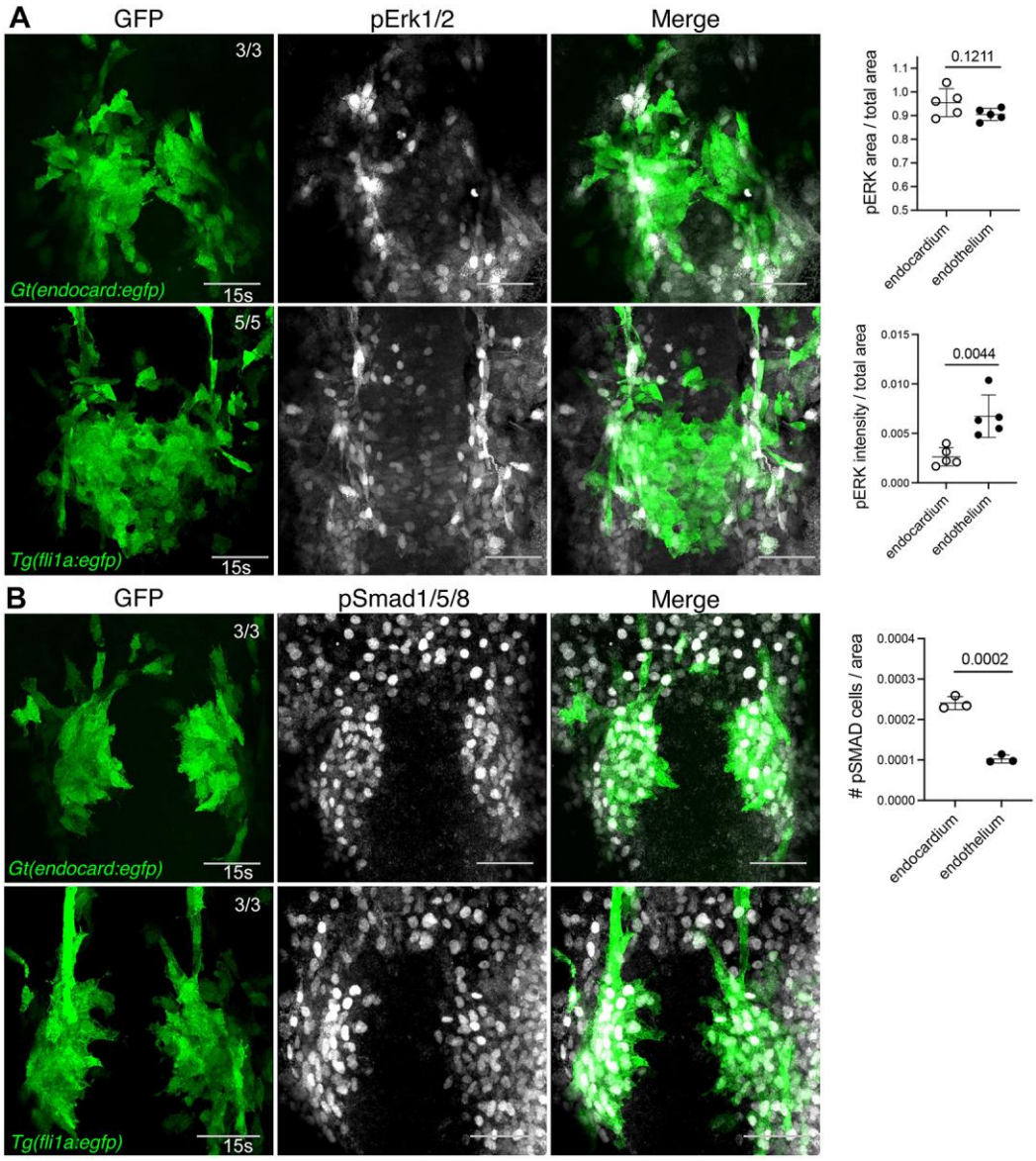


Fig 5. Bmp signalling is active in developing endocardial cells. Immunofluorescence staining for pErk1/2 (**A**) in *Gt(endocard:egfp)* and *Tg(fli1a:egfp)* embryos at 15s shows minimal pErk1/2 signal in endocardial cells and high activity in the adjacent vascular endothelium. By contrast, pSmad1/5/8 (**B**) in *Gt(endocard:egfp)* and *Tg(fli1a:egfp)* embryos at 15s shows high pSmad1/5/8 activity in developing endocardial cell but minimal activity in adjacent vasculature. Quantification of pErk1/2 area and intensity or pSMAD cell number in endocardium or adjacent vasculature is depicted in neighbouring graphs. Dorsal views are shown with anterior to the top in all images. Scale bars represent 50 μm. P-values are present in graphs.

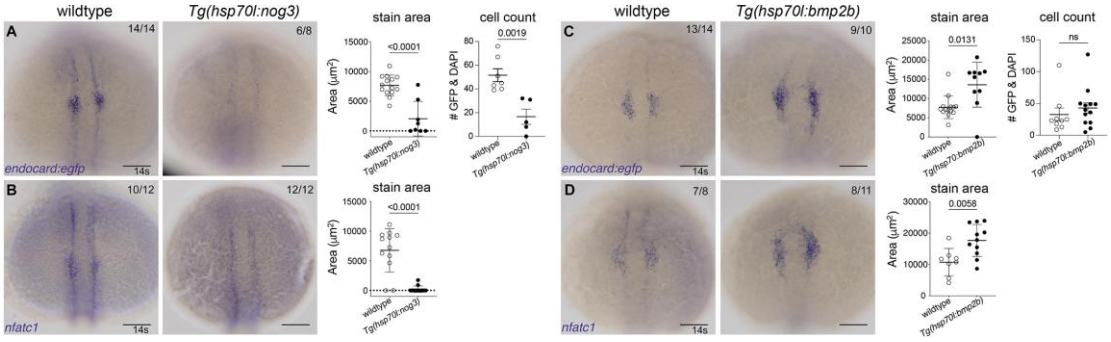


Fig 6. **Bmp signalling is required for endocardial development.** *In situ* hybridisation for endocardial markers *endocard:egfp* (**A**) and *nfatc1* (B) in wildtype (transgenic negative siblings) or *Tg(hsp70l:nog3)* embryos at the 14s (heat-shocked performed at tailbud stage (10 hpf)) show significantly reduced staining of the endocardial domain upon inhibition of Bmp signalling. Quantification of the staining area of expression is shown in adjacent graphs. Quantification of GFP-positive & DAPI-positive cell number is also shown, as determined by confocal imaging of *Gt(endocard:egfp*), DAPI-stained embryos at 14-15s. Reciprocally, wildtype sibling controls or *Tg(hsp70l:bmp2b)* embryos at 14ss (heat-shocked at tailbud) show increased staining area in embryos stained for *endocard:egfp* (**C**) or *nfatc1* (**D**). Quantification of the area of expression shown in adjacent graphs**.** Dorsal views are shown with anterior to the top in all images. Scale bars represent 100 µm. P-values are present in graphs.

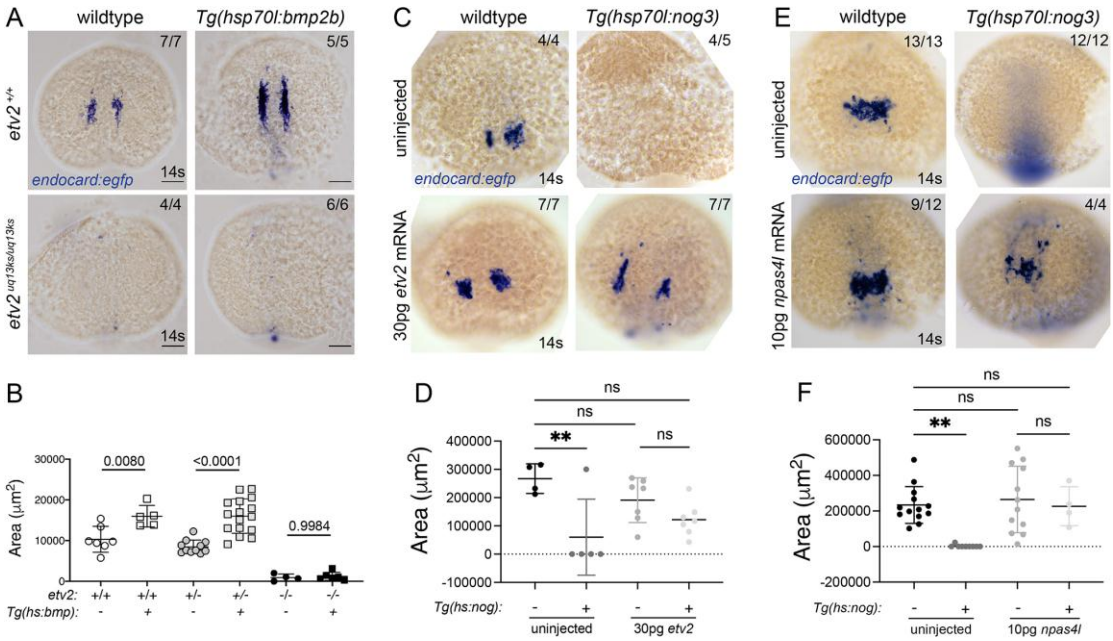


Fig 7. Epistasis experiments show that *npas4l* and *etv2* are downstream of Bmp signalling. *In situ* hybridisation for *endocard:egfp* expression in wildtype and *etv2* mutant with and without the *hsp70l:bmp2b* transgene. (**A**) *etv2* mutants have reduced *endocard:egfp* expression, irrespective of Bmp overexpression. (**B**) Quantification of the staining area shows Bmp overexpression increased the *endocard:egfp* expression domain whereas *endocard:egfp* expression is lost in *etv2* mutants, regardless of Bmp overexpression. **(C-F)** Reciprocal experiments involving *in situ* hybridisation for *endocard:egfp* expression in wildtype (transgenic negative siblings) or *Tg(hsp70l:nog3)* embryos injected with 30pg of *etv2* (**C**) or 10pg of *npas4l* (**E**) mRNA. Quantification of staining area shows that both *etv2* (**D**) and *npas4l* (**F**) mRNA overexpression can restore *endocard:egfp* expression upon Bmp inhibition. Scale bars represent 100 µm. P-values are indicated in graphs. ** $p < 0.01$.

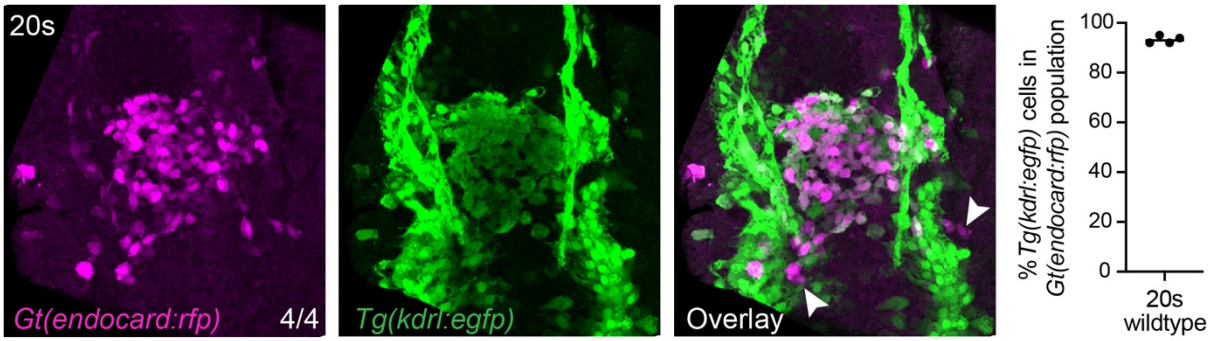


Figure S1

Fig. S1. Quantification of *Tg(kdrl:egfp)* positive cells in the *Gt(endocard:egfp)* population, expressed as a percentage, in 20s stage embryos.

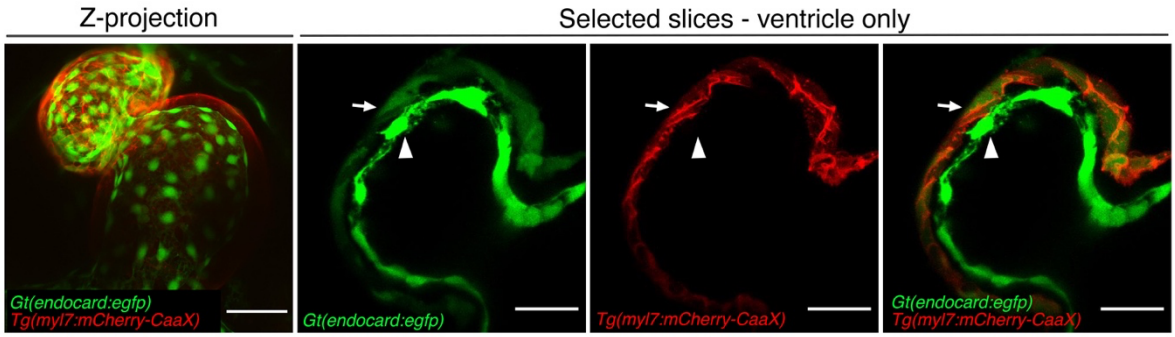


Figure S2

Fig. S2. Live imaging of *Gt(endocard:egfp)*; *Tg(myl7:mCherry-CaaX)* embryos at 48 hpf. A Z-projection of the full heart is shown. Selected slices of the ventricle are also shown demonstrating weak myocardial expression in the *Gt(endocard:egfp)* line. All images are ventral views with anterior to the top. Scale bars represent 50 µm (in Z-projection) and 25 µm (in slices). White arrowheads label endocardium, white arrows label myocardium.

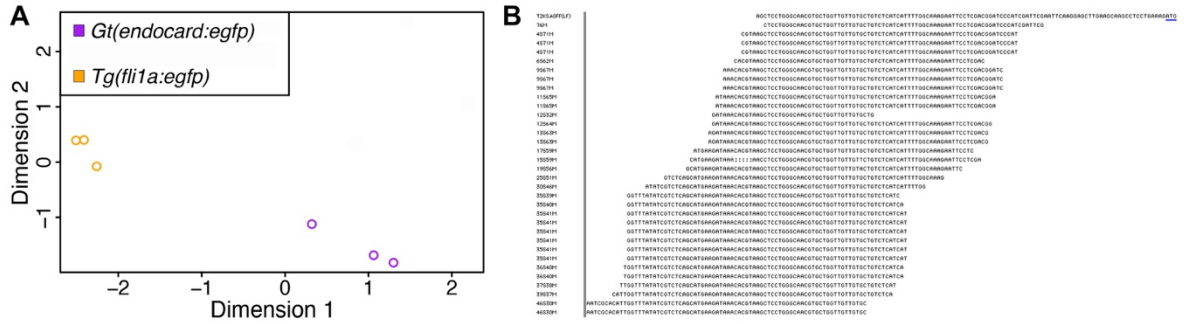


Figure S3

Fig. S3. A. A PCA plot showing the relationship of samples according to dimensions 1 and 2 shows *Gt(endocard:egfp)* and *Tg(fli1a:egfp)* samples form distinct clusters according to both dimensions. **B.** Sequencing reads were mapped to the gene-trap cassette to validate the reported insertion site of the *Gt(endocard:egfp)* line. Soft-clipped reads from *Gt(endocard:egfp)* samples were aligned and the sequence 5' of the gene-trap cassette blasted against the zebrafish genome. This sequence was found to align perfectly with the 3' end of the first exon of the *map3k22* gene confirming the insertion of the cassette in the first intron of this gene. The start codon of the GAL4FF coding sequence in the gene-trap cassette is underlined in blue for reference.

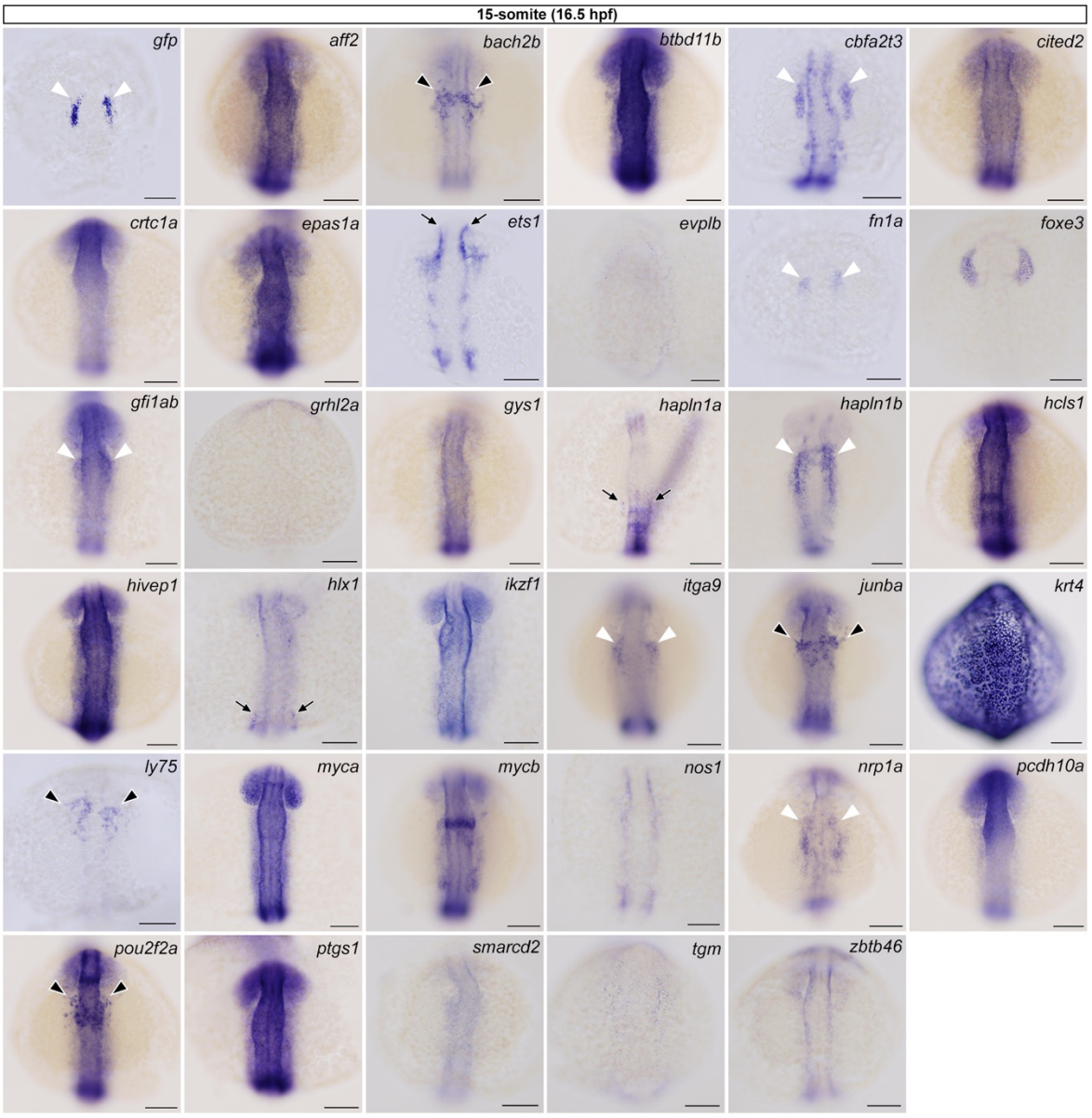


Figure S4

Fig. S4. *In situ* hybridisation for 34 candidate genes showing enrichment in the endocardium in some. All expression patterns were examined at the 15-somite stage (16.5 hpf). As a positive control, a *gfp* probe was examined in *Gt(endocard:egfp)* embryos to label the endocardium. White arrowheads label endocardial domains, black arrowheads label myeloid domains, black arrows label vascular endothelial expression. Ventral views are shown with anterior to the top in all images. Scale bars represent 50µm.

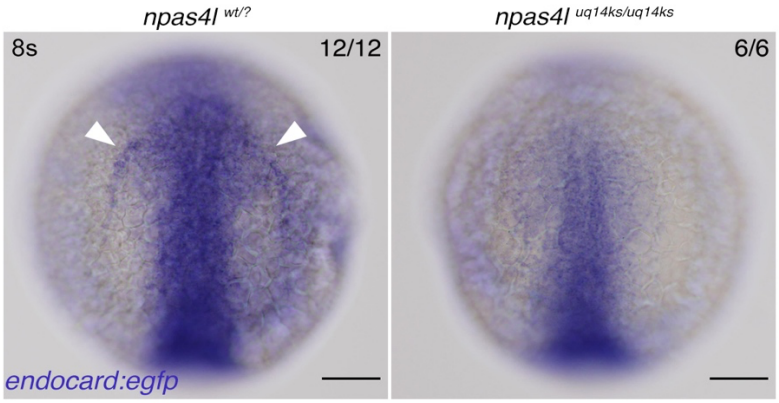


Figure S5

Fig. S5. *In situ* hybridisation for endocardial *gfp* expression in *Gt(endocard:egfp)*; *npas4l* wildtype sibling and homozygous mutant embryos at 8s (13 hpf). Dorsal views with anterior to the top in all images. Scale bars represent 100 μm. The number of embryos matching the image is indicated in the top right of each image.

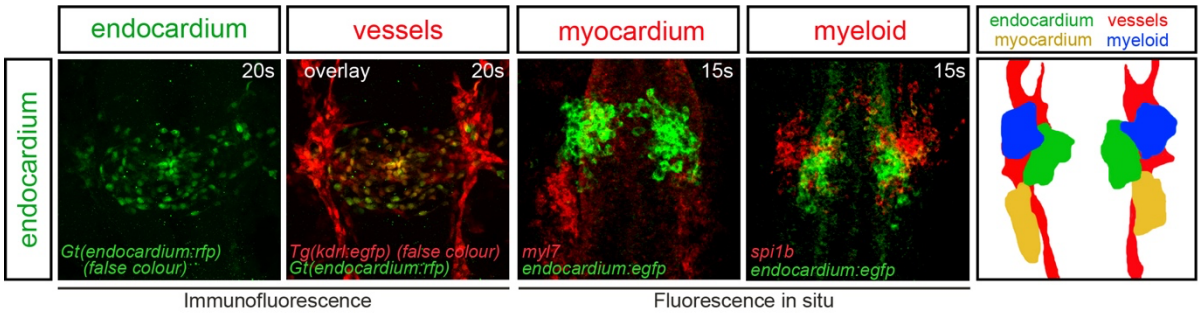


Figure S6

Fig. S6. Expression domains of mesodermal lineages in the mid-somitogenesis stage embryo. Live imaging of *Gt(endocard:rfp)*;*Tg(kdrl:egfp)* embryos at 20s (false coloured). Fluorescent double *in situ* hybridisation *myl7* (red; myocardium) overlaid with *endocardium:egfp* (green; endocardium) and spi1b (red; myeloid) overlaid with *endocardium:egfp* (green; endocardium) at 15s. Schematic representation of spatial arrangement of endocardium, vessels, myocardium and myeloid expression domains in 15-20s.

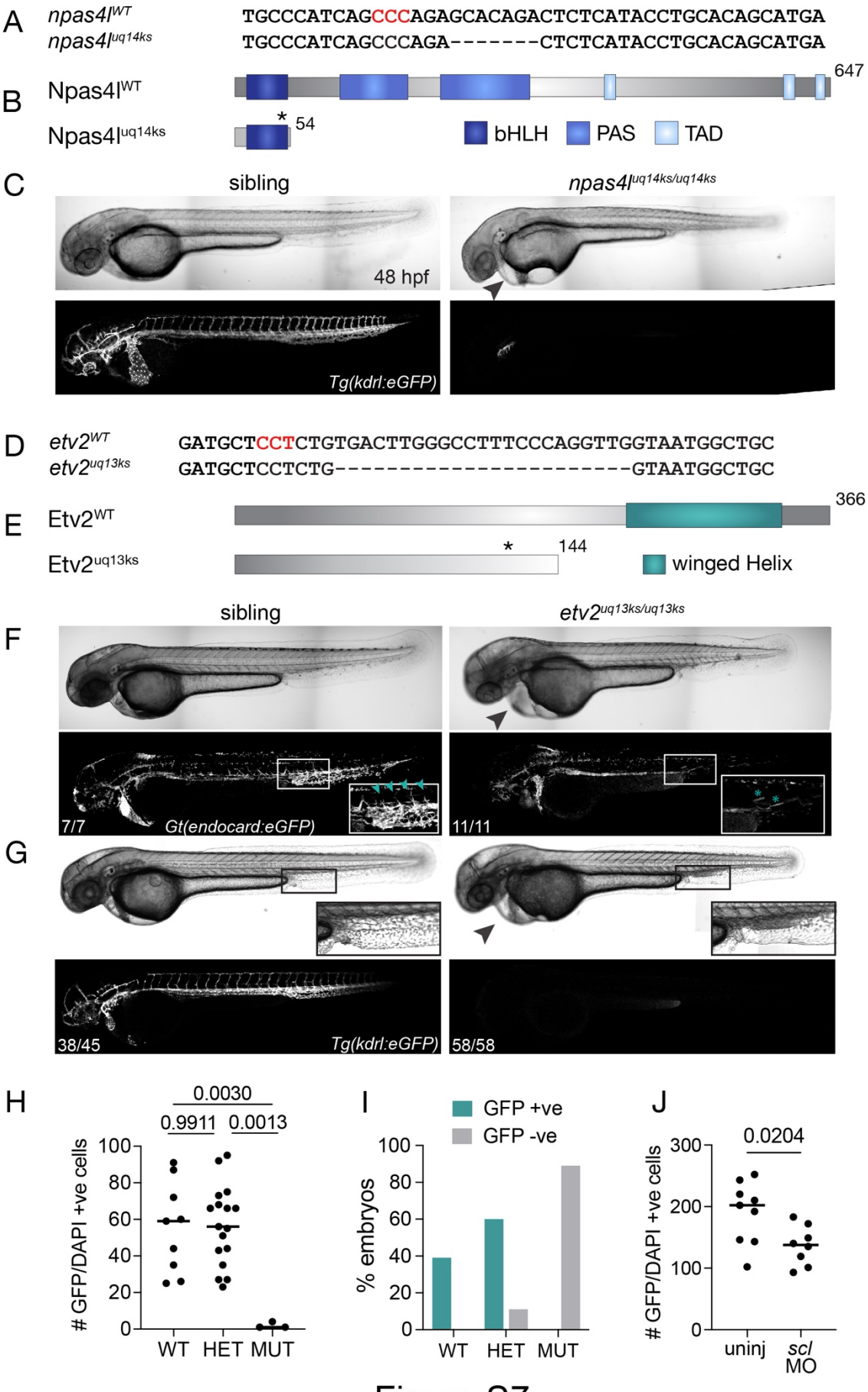


Figure S7

Fig. S7. A. Sequence of CRISPR-Cas9-induced lesion in *npas4l*^{uq14ks} allele. **B.** Schematic representing the predicted protein resulting from the *npas4l*^{uq14ks} mutation. **C.** Bright field (above) and confocal image of *Tg(kdrl:egfp)* expression (below) in wildtype sibling versus *npas4l*^{uq14ks} mutant. Almost all eGFP expression is absent in homozygous mutants, consistent with previously published alleles. Pronounced pericardial oedema is present in mutant embryos (arrowhead). **D.** Sequence of CRISPR-Cas9-induced lesion in *etv2*^{uq13ks} allele. **E.** Schematic representing the predicted protein resulting from the *etv2*^{uq13ks} mutation. **F.** Bright field (above) and confocal image of *Gt(endocard:egfp)* expression (below) in wildtype sibling versus *etv2*^{uq13ks} mutant. Pronounced pericardial oedema is present in mutant embryos (black arrowhead). The *Gt(endocard:egfp)* reporter shows intersegmental venous sprouts (teal arrowheads) present in siblings that are absent from *etv2*^{uq13ks} mutants (inset), accompanied by reduced endothelial expression and ectopic expression in skeletal muscle fibres, as previously reported (inset; teal asterisks) (Chestnut et al., 2020). **G.** Bright field (above) and confocal image of *Tg(kdrl:egfp)* expression (below) in wildtype sibling versus *etv2*^{uq13ks} mutant. As above, pericardial oedema is present in mutant embryos (arrowhead) and the caudal vein plexus is absent. Inset: caudal vein plexus is observable by bright field contrast imaging, whereas accumulated blood is observed in mutant embryos. The *kdrl:egfp* transgene was never observed in *etv2* $^{uq13ks/uq13ks}$ homozygous mutant embryos (n=58), whereas homozygous wildtype embryos were always observed to segregate with the *kdrl:egfp* transgene (n=15), suggesting *etv2* is genetically linked with the *Tg(kdrl:egfp)* locus. **H.** Quantification of *Gt(endocard:egfp)* positive cells (determined by DAPI counterstain) in *etv2*$^{+/+}$(WT), *etv2*$^{+/uq13ks}$ (HET) and *etv2* $^{uq13ks/uq13ks}$ (MUT) at 14s. **I.** Genotyping of GFP-positive and GFP-negative embryos for the *etv2*^{uq13ks} allele, confirms genetic linkage to the the *Tg(kdrl:egfp)* locus (χ^2 test: $p < 0.0001$). **J.** Quantification of *Gt(endocard:egfp)* positive cells (determined by DAPI counterstain) in uninjected control and *scl* morphant embryos at 14s.

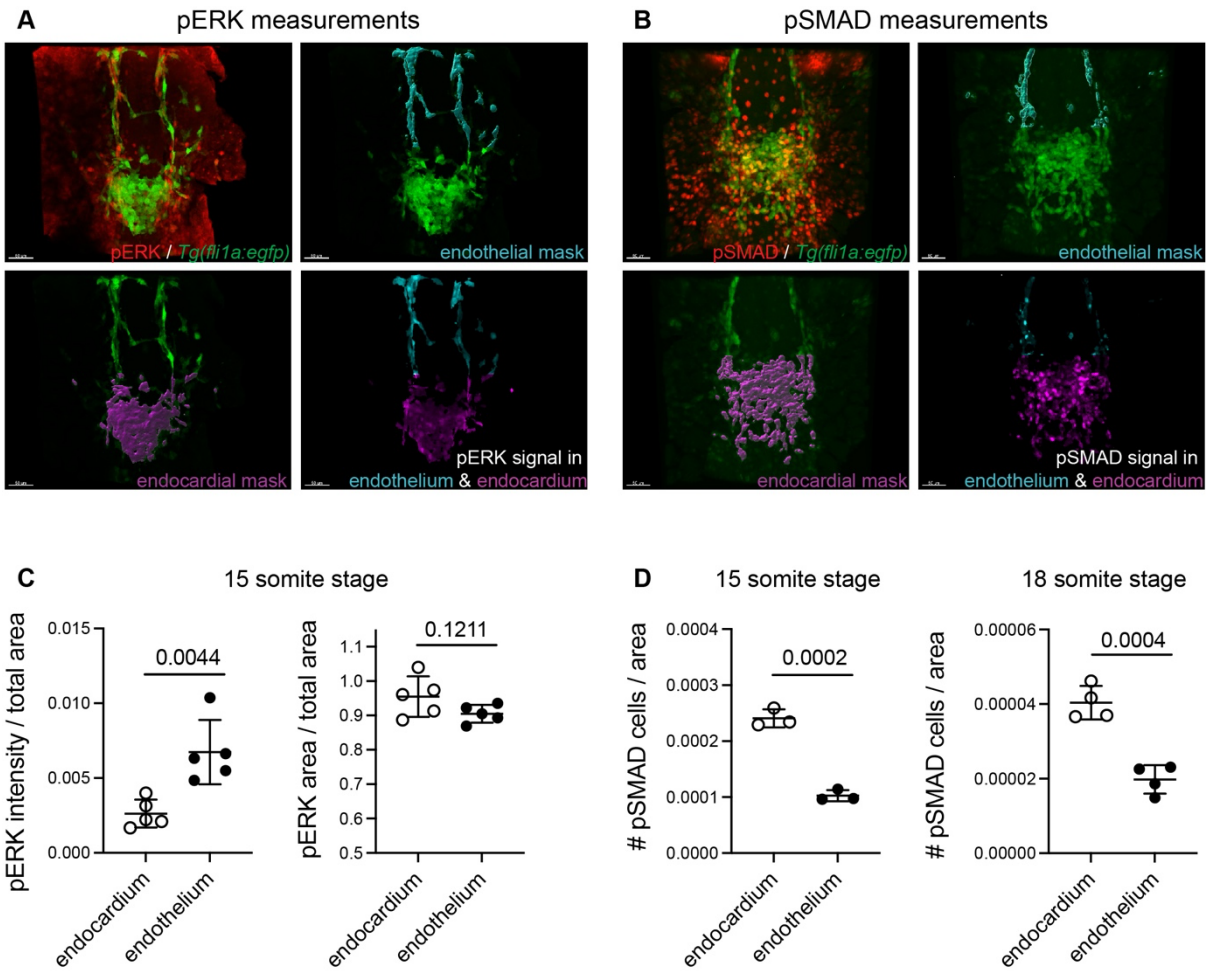


Figure S8

Fig. S8. Representative images of confocal scans for **A.** pERK or **B.** pSMAD and α-GFP immunostaining on *Tg(fli1a:egfp)* background. Images show examples of endocardial or endothelial masks created in Imaris to quantify pERK or pSMAD staining in these two tissues. **C.** and **D.** Dot plots depicting quantification of pERK (**C**) or pSMAD (**D**) in endocardium versus endothelium.

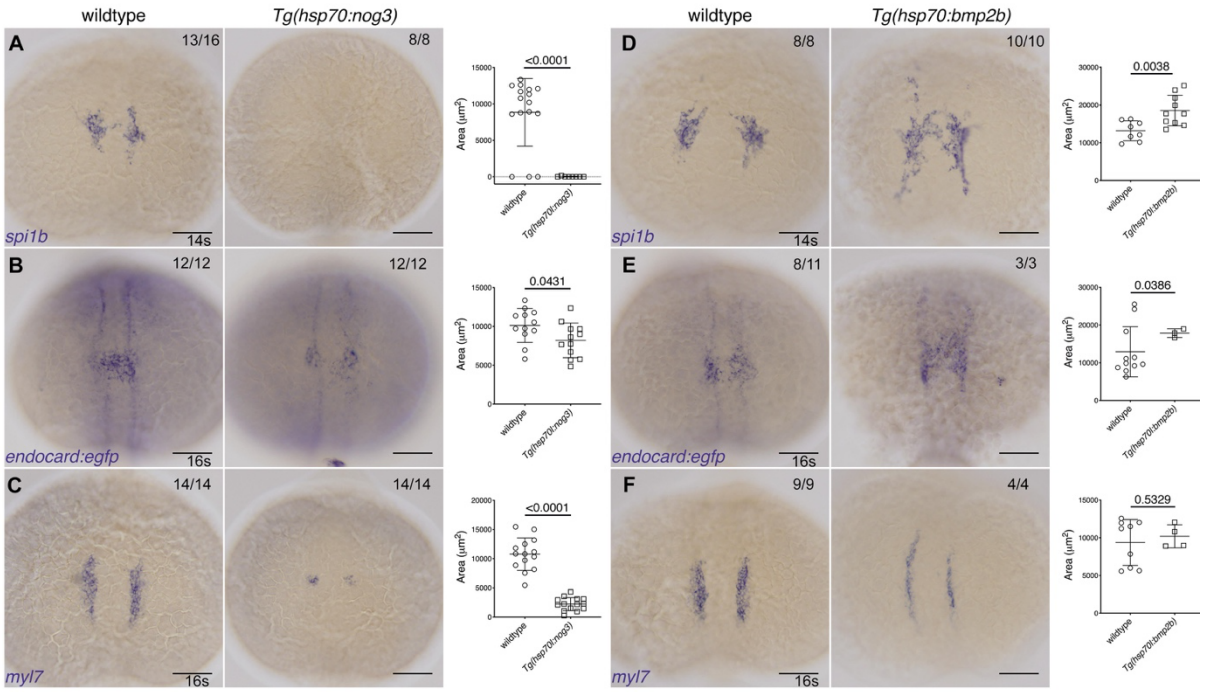


Figure S9

Fig. S9. *In situ* hybridisation for myeloid expression, *spi1b* (**A** and **D**), at the 14s as well as for endocardial, *endocard:egfp* (**B** and **E**), and myocardial, *myl7* (**C** and **F**), expression at the 16s in embryos from crosses of the *Gt(endocard:egfp)* line to either the *Tg(hsp70l:nog3)* or *Tg(hsp70l:bmp2b)* lines. Quantification of the expression is shown in adjacent graphs. Dorsal views are shown with anterior to the top in all images. Scale bars represent 100 µm. P-values are indicated in graphs.

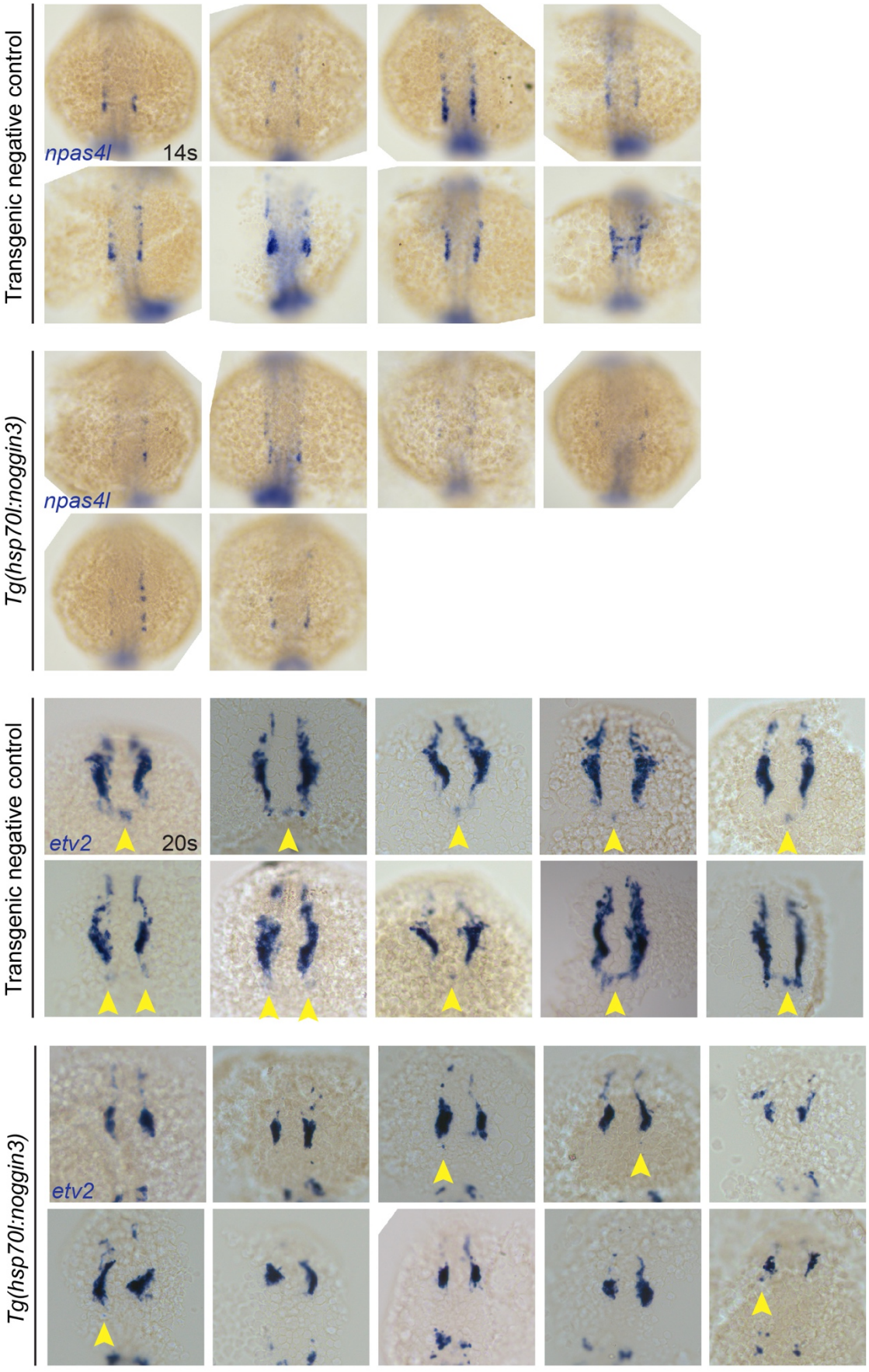


Figure S10

Fig. S10. *In situ* hybridisation for *npas4l* expression in wildtype sibling controls and *Tg(hsp70l:nog3)* embryos at 14s, showing reduced or absent expression of *npas4l* upon Bmp inhibition. Figure shows one of two experiments performed (total embryos analysed n = 10 for transgenic negative controls and n=8 for transgene positive). *In situ* hybridisation for *etv2* expression in wildtype sibling controls and *Tg(hsp70l:nog3)* embryos at 20s. At 20s, the majority of *etv2* expression is enriched in the presumptive vasculature. A small amount of *etv2* expression can be observed in the developing endocardium (yellow arrowheads). Upon Bmp inhibition, the expression domain is smaller and endocardial expression is greatly reduced (yellow arrowheads). Figure shows one of two experiments (total embryos analysed n = 25 for transgenic negative controls and n=11 for transgene positive)

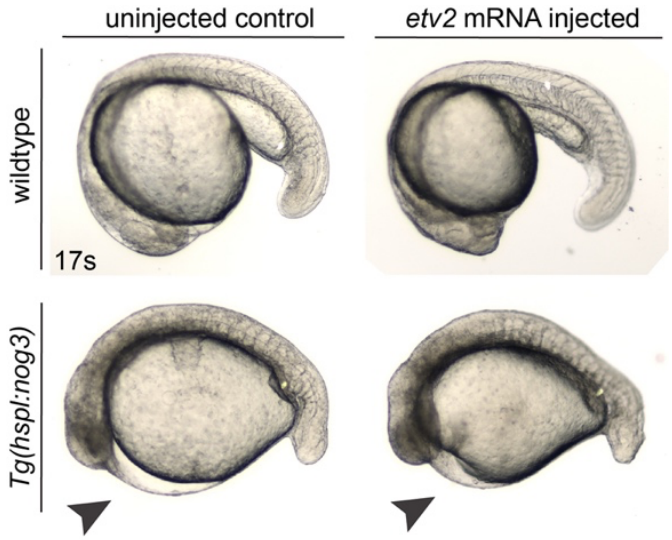
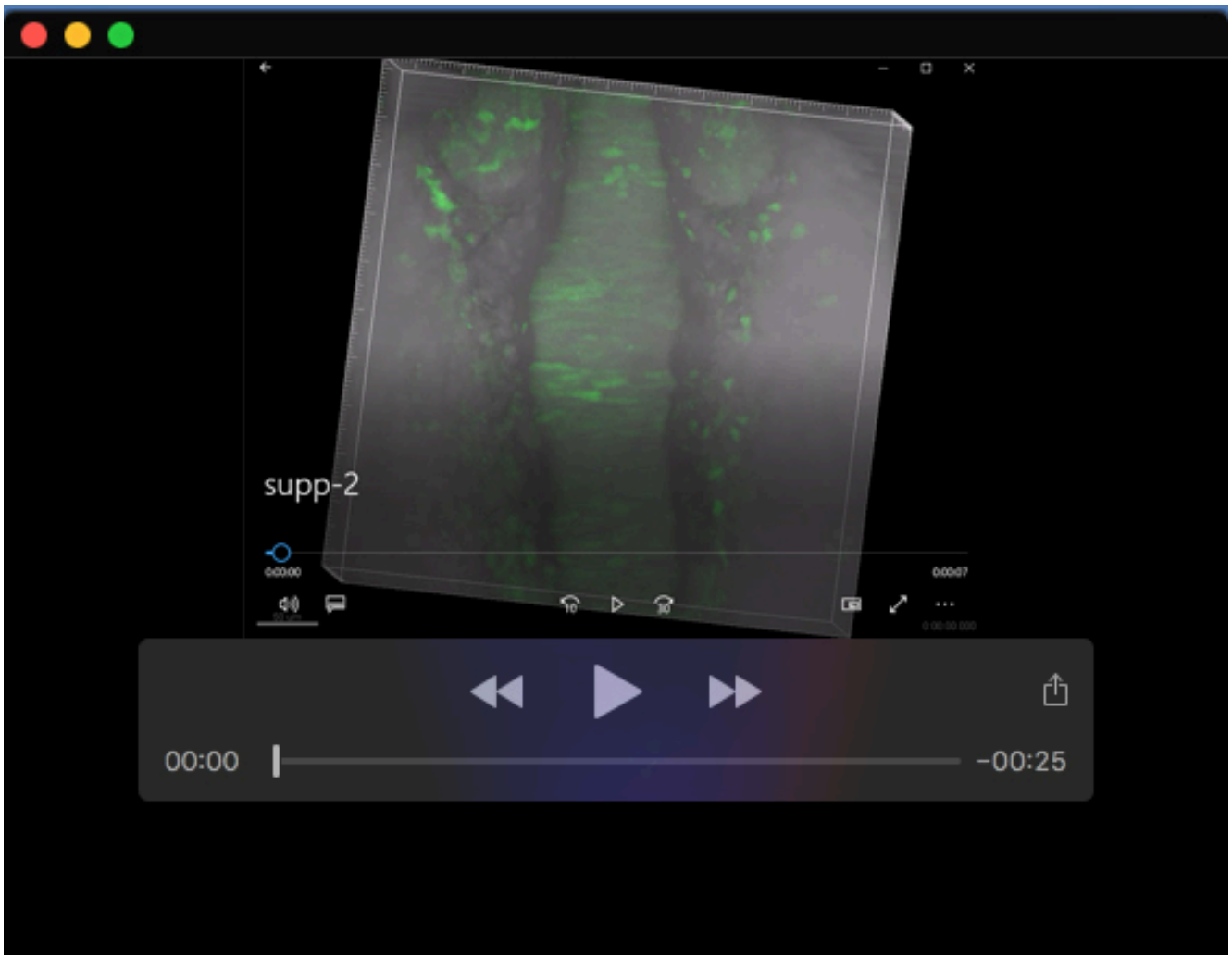


Figure S11

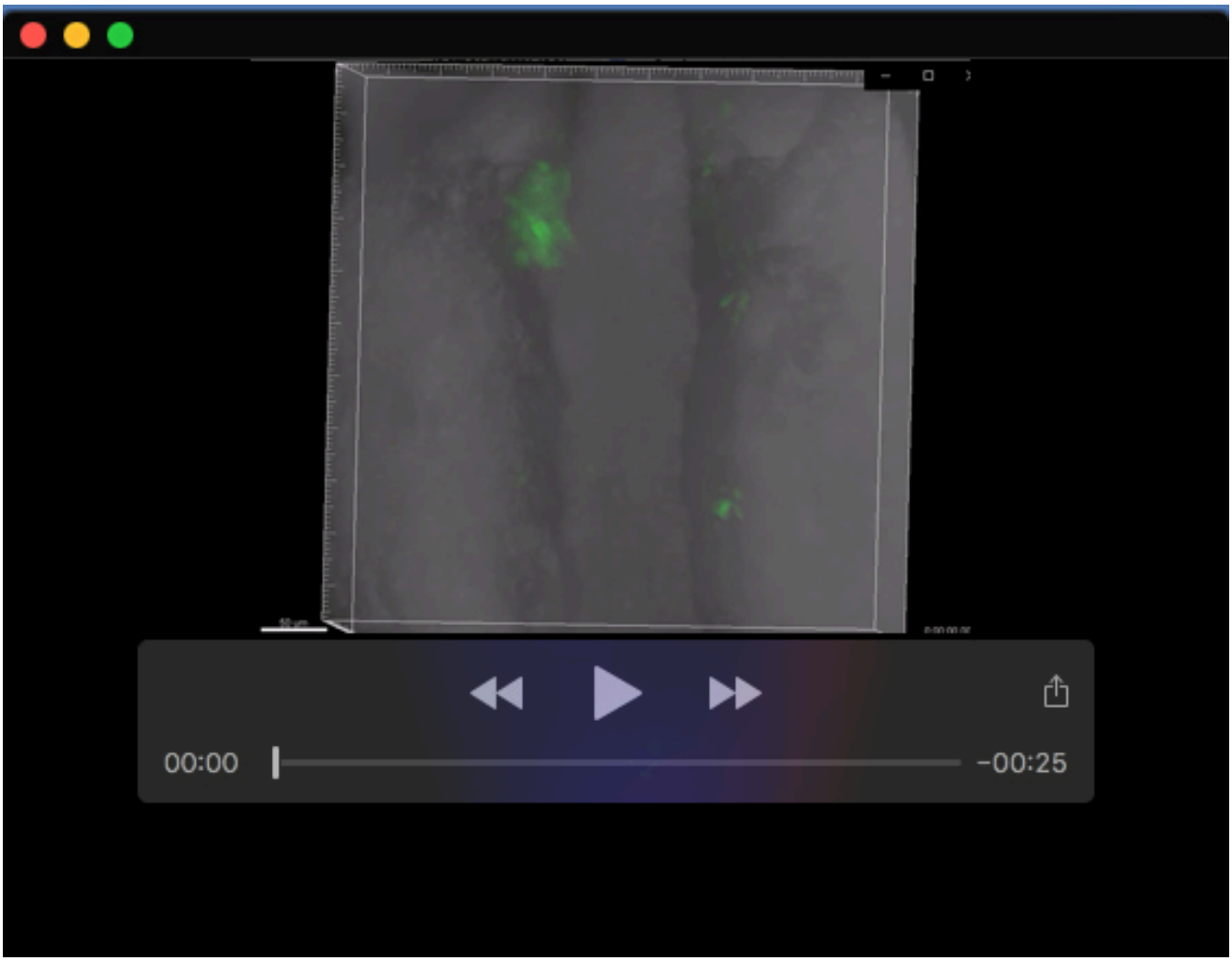
Fig. S11. Bright field lateral view images wildtype of *Tg(hsp70l:nog3)* embryos at 17s, either uninjected or injected with *etv2* mRNA. *Tg(hsp70l:nog3)* embryos heat-shocked at tailbud stage present with a characteristic dorsalisation phenotype. Whilst *etv2* overexpression may rescue the *endocard:egfp* expression in the ALPM, it does not rescue the overall patterning defects of the embryo, nor does it prevent pericardial oedema that forms in *Tg(hsp70l:nog3)* embryos (black arrowhead).

Table S1. Full list of RNAseq analysis and GO terms associated with the differentially expressed data. The lists correspond to the complete list of genes from the differential expression analysis (DE_endocard-egfpVSfliegfp), as well as all significantly up (GO_sig_down_up-list) or down (GO_sig_down_ref-list) gene lists and significantly up or down gene lists with a $\log_2$ Fold Change cut-off of ≥1 (GO_TREAT_up_ref-list) and ≤-1 (GO_TREAT_down_ref-list) applied by TREAT.

Click here to download Table S1



Movie 1.



Movie 2.