Title: A conserved regulatory program drives emergence of the lateral plate mesoderm

Authors: Christopher Hess[1,†], Karin D. Prummel[1,†], Susan Nieuwenhuize[1], Hugo J. Parker[2,3], Katherine W. Rogers[4], Iryna Kozmikova[5], Claudia Racioppi[6], Sibylle Burger[1], Eline C. Brombacher[1], Alexa Burger[1], Anastasia Felker[1], Elena Chiavacci[1], Gopi Shah[7], Jan Huisken[7,8], Zbynek Kozmik[5], Lionel Christiaen[6], Patrick Müller[4], Marianne Bronner[2], Robb Krumlauf[3], Christian Mosimann[1,*].

Affiliations:

[1] Institute of Molecular Life Sciences, University of Zürich, Zürich, Switzerland
[2] California Institute of Technology, Pasadena, United States
[3] Stowers Institute for Medical Research, Kansas City, United States
[4] Friedrich-Miescher-Laboratory of the Max Planck Society, Tübingen, Germany
[5] Institute of Molecular Genetics of the ASCR, Prague, Czech Republic
[6] New York University College of Arts and Science, New York, United States
[7] Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany
[8] Morgridge Institute for Research, Madison, Wisconsin, United States

[†] contributed equally to this work

* correspondence to: christian.mosimann@imls.uzh.ch

Abstract:

Cardiovascular cell lineages emerge with kidney, smooth muscle, and limb skeleton progenitors from the lateral plate mesoderm (LPM). How the LPM emerges during development and how it has evolved to form key lineages of the vertebrate body plan remain unknown. Here, we captured LPM formation by transgenic *in toto* imaging and lineage tracing using the first pan-LPM enhancer element from the zebrafish gene *draculin* (*drl*). *drl* LPM enhancer-based reporters are specifically active in LPM-corresponding territories of several chordate species, uncovering a universal LPM-specific gene program. Distinct from other mesoderm, we identified EomesA, FoxH1, and MixL1 with BMP/Nodal-controlled Smad activity as minimally required factors to drive *drl*-marked LPM formation. Altogether, our work provides a developmental and mechanistic framework for LPM emergence and the *in vitro* differentiation of cardiovascular cell types. Our findings suggest that the LPM may represent an ancient cell fate domain that predates ancestral vertebrates.

Main Text:

Following gastrulation, mesoderm in vertebrates is classically described to partition into axial, paraxial, and lateral domains(Gurdon, 1995). The ventrally forming lateral plate mesoderm (LPM) is composed of highly mobile cells and is mainly defined by its position adjacent to the somite-forming paraxial mesoderm (Hatada and Stern, 1994). Transplantation and lineage tracing experiments in several species have shown that the LPM contains progenitor cells of the circulatory system, smooth muscle lineages, the kidneys (often demarcated as dedicated intermediate mesoderm), and the limb anlagen (Chal and Pourquié, 2017; Gurdon, 1995; Takasato and Little, 2015). Several transcription factors including Hand1/2, Tbx5, Osr1, FoxF1, Prrx1, Mesp1, and Etv2 are expressed in LPM territories and play overlapping roles in cell fate determination (Chal and Pourquié, 2017; Davidson and Zon, 2004; Takasato and Little, 2015), albeit not always with conserved function (Yabe et al., 2016). Anterior-to-posterior expression domains of select transcription factors including Tbx1/10 and Hand are conserved in lampreys and even in the basal cephalochordate amphioxus that lacks major LPM-derived organ systems (Onimaru et al., 2011). These observations raise the possibility of the existence of an ancient upstream program that delineates the prospective LPM domain.

In zebrafish, the LPM emerges from the ventral half of the embryo and becomes readily detectable as a patchwork of bilateral gene expression domains at the end of gastrulation (Davidson and Zon, 2004). *hand2* expression marks in the anterior LPM (ALPM) the heart field and pectoral fin precursors, and a lateral-most domain within the posterior LPM (PLPM) (Perens et al., 2016; Yelon et al., 2000), while *pax2.1* outlines the prospective kidney progenitors that in zebrafish become positioned more laterally than in chick and mouse (Perens et al., 2016; Takasato and Little, 2015). Hematopoietic and endothelial progenitors in distinct ALPM and PLPM domains express a number of lineage-determining transcription factor genes including *etv2*, *scl*/*tal1*, *lmo2*, and later *gata1*, *pu.1*, and *fli1* (Davidson and Zon, 2004). While several studies have uncovered mechanisms that control the lateral-to-medial or anterior-posterior specification of LPM-descendant organ precursors (Chal and Pourquié, 2017; Helker et al., 2015; Tonegawa et al., 1997), how the LPM arises among other mesendoderm remains incompletely understood.

Several *cis*-regulatory elements have been reported with specific activity in the mouse LPM; these include a small upstream enhancer of mouse and human *HoxB6* (Becker et al., 1996), an upstream enhancer of mouse *Gata4* (Rojas et al., 2005), and a downstream enhancer of mouse *Bmp4* (Chandler et al., 2009). In line with the ventral origin of the LPM, the *Gata4* LPM enhancer responds to Smads downstream of BMP signaling (Rojas et al., 2005). Nonetheless, the activities driven by these enhancer elements in mice are confined to the PLPM. In contrast, a transgenic reporter from the 6.35 kb *cis*-regulatory region of the zebrafish-specific gene *draculin* (*drl*) specifically outlines seemingly all LPM derivatives and labels the forming and differentiating LPM from its

emergence in late gastrulation (Mosimann et al., 2015), as confirmed by Cre/*lox*-mediated genetic lineage analysis (Felker et al., 2016; Gays et al., 2017; Henninger et al., 2016; Mosimann et al., 2015). These observations suggest that *-6.35drl* harbors one or several *cis*-regulatory elements active throughout the prospective LPM from gastrulation onwards, raising the intriguing possibility that it may provide a read-out of an LPM-specific program.

To resolve the dynamics of LPM emergence marked by the *-6.35drl cis*-regulatory elements, we performed 3D single-plane illumination microscopy (SPIM) (Fig. 1A-D) and panoramic projections (Schmid et al., 2013) (Fig. 1E-H) of *-6.35drl:EGFP* reporter transgenic zebrafish embryos. *-6.35drl*:*EGFP*-expressing LPM precursors became detectable by early gastrula stages (50% epiboly) and slowly condensed along the embryo margin through the end of gastrulation (tailbud stage) (Fig. 1A,B,F; Movie S1, S2). From tailbud stage onward, *-6.35drl:EGFP*-marked LPM formed a continuous band of cells with condensing anterior and posterior segments interspersed by a scattered medial segment (Fig. 1C,D,G,H). We demonstrated that this EGFP-positive cell band encompasses stripes of previously characterized bilateral LPM sub-domain markers (Fig. 1I-L): *lmo2:dsRED2* labels embryonic hematopoietic and vascular tissues (Zhu et al., 2005), and its expression overlapped with medial *-6.35drl:EGFP*-expressing cells in the ALPM and PLPM (Fig. 1L); *scl:EGFP* (Zhang and Rodaway, 2007) also co-expressed with *-6.35drl:mCherry* in the most medial PLPM domain and in a small ALPM population (Fig. 1J); *pax2.1:EGFP* in early somitogenesis marks pronephric epithelial precursors in zebrafish (Picker et al., 2002), and its expression overlaps with *-6.35drl:mCherry* in the PLPM, while the *pax2.1:EGFP*-expressing midbrain-hindbrain precursor cells of ectodermal origin were negative for *-6.35drl:mCherry* (Fig. 1K). Moreover, *hand2:EGFP* expression, which demarcates the lateral-most PLPM domain and parts of the ALPM-derived heart field and pectoral fin precursors (Perens et al., 2016; Yelon et al., 2000), is also fully situated within the pan-LPM expression domain of *-6.35drl:mCherry* (Fig. 1L). Taken together, these data reveal that transgenic reporters based on *-6.35drl* demarcate the entire emerging zebrafish LPM (Fig. 1M).

To clarify how the *drl*-expressing LPM emerges relative to the endoderm during gastrulation, we performed time lapse imaging of the forming LPM *in toto*. Panoramic SPIM for the *sox17:EGFP*-positive endoderm together with *-6.35drl:mCherry* reporter expression revealed *sox17*;*drl* reporter double-positive cells from the onset of reporter detection through late gastrulation (Fig. 1N-O; Fig. S1A-D). After gastrulation, we detected a continuous band of *-6.35drl*-positive cells around the developing embryo that was clearly separated from the more medial endodermal *sox17* expression domain (Fig. 1N; Fig. S1D; Movie S3, S4). The temporal overlap in expression suggests that the *-6.35drl*-marked territory initially emerges among mesendoderm. Consequently, we performed *-6.35drl:creERT2*-based genetic lineage tracing with *hsp70l:switch* (*hsp70l:loxP-STOP-loxP-EGFP*) with 4-OHT-

induction at shield and 5-6 somite stages (ss) to explore whether endoderm derivatives are marked by the *-6.35drl* reporter at 72 hpf (Fig. 1P). 4-OHT induction at shield stage (Fig. 1Q,T) broadly marked i) LPM lineages as the only mesodermal fates (Fig. 1Q), and ii) broad territories within endodermal organs including pancreas, liver, and pharynx/gut epithelium (Fig. 1Q-T; Fig. S2B-E). In contrast, *sox17:creERT2* exclusively marked endoderm lineages at all induction time points tested (Fig. S2I-K), in line with *sox17* expression demarcating specified endoderm downstream of the key zebrafish endoderm regulator *sox32* (Alexander and Stainier, 1999). Supporting our SPIM observations of the *-6.35drl* reporter and consistent with previous studies (Felker et al., 2016; Gays et al., 2017; Henninger et al., 2016; Mosimann et al., 2015), *-6.35drl:creERT2*-based lineage tracing with 4-OHT induction at 5-6 ss efficiently mapped to LPM lineages with only minimal endoderm contribution (Fig. 1U; Fig. S2B,E). These results are not caused by lineage-biased *loxP* reporter sensitivity, as controlled for by recombining *hsp70l:switch* with the ubiquitous *ubi:creERT2* (Fig. S2B,F-H). Embryos that fail to form endoderm upon perturbation of *sox32* (Alexander and Stainier, 1999) still properly generated *-6.35drl*-traced LPM that patterns into its major descendants including heart, blood, endothelium, and pronephros (Fig. S3). These data establish that during gastrulation, *-6.35drl*-marked LPM gradually refines from a ventral mesendoderm territory to a bilateral LPM domain along the entire anterior-posterior axis of the embryo. The concomitant *drl*;*sox17*-expressing progenitors contribute to seemingly all endodermal organs. The rather rare lineage labeling of somitic muscle by *-6.35drl:creERT2* (Fig. 1Q) (Gays et al., 2017; Mosimann et al., 2015) further suggests that, in zebrafish, the paraxial mesoderm and the LPM develop as distinct mesoderm lineages with minimal overlap (Warga and Nüsslein-Volhard, 1999).

To identify the *cis*-regulatory components in the *-6.35drl* genomic region responsible for pan-LPM expression, we sub-divided *-6.35drl* into smaller regions and tested them in reporter assays (Fig. 2A). These efforts uncovered *drl cis*-regulatory elements surrounding exon 1 that recapitulated strongest *drl* expression during somitogenesis in the differentiating cardiovascular/hematopoietic ALPM and PLPM ($drl^{proximal}$) and selectively in the PLPM (drl^{PLPM}), respectively (Fig. 2B,C). In addition, we found a 968 bp region in the distal first intron, referred to as drl^{LPM}, that as transgenic reporter fully recapitulated early *-6.35drl* and endogenous *drl* expression in zebrafish embryos before fading during somitogenesis (approx. 13 hpf) (Fig. 2D). Corresponding to gastrulation-stage lineage tracing with *-6.35drl:creERT2* (Fig. 1P-U), genetic lineage tracing with drl^{LPM}:*creERT2* and *hsp70l:switch* (Fig. 2E) induced at 50% epiboly specifically labels LPM-derived organs including heart, iSMCs, blood, endothelium, and kidney, and additionally marked endodermal cells (Fig. 2F-H). Trimming of drl^{LPM} resulted a minimal enhancer region of 432 bp (drl^{minLPM}) that still behaved as a pan-LPM enhancer and lost its specificity and activity upon further trimming (Fig. S4). Hence, the *-6.35drl* reporter pattern combines an early mesendoderm/pan-LPM phase and a later ALPM/PLPM-restricted phase

controlled by dedicated *cis*-regulatory elements, consistent with *drl*-based lineage tracing (Fig. 1P-U) (Felker et al., 2016; Gays et al., 2017; Henninger et al., 2016; Mosimann et al., 2015).

We next investigated whether the *drl*^{LPM} enhancer reads out a conserved pan-LPM program by performing reporter assays in other chordate species. Electroporation of *drl*^{LPM} reporter into the primitive streak of *ex ovo*-cultured chicken embryos at HH3+/4 resulted in specific reporter activity in the forming LPM. Depending on the stage and region of electroporation, we observed specific reporter activity in medial LPM stripes or the posterior end of the LPM, including outward migrating cells reminiscent of endothelial/blood progenitors at later stages (Fig. 3A; Fig. S5B-J). In contrast, the previously characterized posterior LPM enhancers of mouse *Gata4* (Rojas et al., 2005) and mouse *Bmp4a* (Chandler et al., 2009) showed specific activity only in the outward migrating endothelial/blood progenitors in chicken embryos (Fig. S6A,B), while the LPM enhancer of mouse *HoxB6* (Becker et al., 1996) showed no specific activity (Fig. S6c). When injected in zebrafish embryos, these constructs showed expression mainly in the notochord and lacked specific LPM activity (Fig. S6D-G). Furthermore, there are species-specific differences in endogenous genes, as FoxF1 in mice is expressed in the early developing LPM (Mahlapuu et al., 2001), while zebrafish *foxf1* mRNA is not detectable by *in situ* hybridization (ISH) before late somitogenesis (Fig. S6H-K). Taken together, these results indicate that activities of previously used LPM reporters do not universally extrapolate to other vertebrates, but that the *drl*^{LPM} enhancer does react to an LPM-specific program in zebrafish and chicken.

We next explored the ability of the *drl*^{LPM} enhancer to read out a pan-LPM program in distinct chordate species. Lampreys are jawless fish that can provide unique insights into vertebrate evolution due to their position at the base of the vertebrate phylogeny. Microinjection of *drl*^{LPM} and *drl*^{minLPM} EGFP-based reporters into sea lamprey embryos (*Petromyzon marinus*) consistently resulted in strong EGFP expression in the mesendoderm starting during neurulation (stage (st) 18) (n = 145/231), as well as in the developing pharynx at st 22-24 (Fig. 3B-F). Transverse embryo sections revealed that this early expression domain partially includes the anterior-most expression of lamprey *pmHandA* that expresses in LPM (Onimaru et al., 2011), with the later pharyngeal expression of EGFP being restricted to the endoderm and mesoderm (Fig. S7A-I).

Electroporation of *drl*^{LPM} reporters into embryos of the tunicate *Ciona robusta*, a chordate species belonging to a sister clade of the vertebrates, revealed specific reporter activity in emerging cardiac and pharyngeal muscle lineages at larval stage (st 26; 18 hpf) (Fig. 3G) (Hotta et al., 2007). We observed *drl*^{LPM}- and *drl*^{minLPM}- driven EGFP reporter activity specifically in the atrial siphon muscle precursors (ASMPs) and in both first and second heart precursors (FHPs and SHPs) as confirmed by co-localization of *Mesp:H2B-mCherry* expression that labels the cardiopharyngeal cell lineage (n = 15/92 and 23/110 respectively; Fig. S7J-L). We rarely observed

EGFP expression in the anterior tail muscles (ATMs) of the *Ciona* larval tail, suggesting that *drl* reporters drive expression specifically in the emerging multipotent cardiopharyngeal progenitors. Moreover, using both *drl* reporters, we observed ectopic reporter activity in the mesenchyme, which often activates transgenes non-specifically (Stolfi and Christiaen, 2012).

Lastly, in the cephalochordate amphioxus (*Branchiostoma lanceolatum*), which belongs to the most basally divergent lineage of chordates (Bertrand et al., 2011), ventral mesoderm that is considered to be equivalent to the vertebrate LPM originates from the ventral half of the somite territory. At mid-neurula stage, the ventral wall of the somite is evaginating as nascent ventral mesoderm, which by late neurula stage fuses at the midline under the gut endoderm (Holland and Holland, 2007; Kozmikova and Kozmik, 2015). Indeed, amphioxus *Hand*, *Csx*, *Vent1*, and *AmphiNk2-tin* are expressed in the ventral half of the somite territory at mid-neurula stage (Holland and Holland, 2007; Kozmikova and Kozmik, 2015; Onimaru et al., 2011). Correspondingly, our injection-based reporter assays in amphioxus showed drl^{LPM} reporter activity at mid-neurula stage (6/7 ss) in the ventral half of somites and evaginating somites (n = 61/65) (Fig. 3H, Fig. S7M,N).

Taken together, these observations establish that the drl^{LPM} regulatory module in the zebrafish *drl* locus is capable of reading out an evolutionarily conserved pan-LPM program that is active from chordates to higher vertebrates.

We next sought to gain insight into what upstream input controls the drl^{LPM} enhancer. Mesendoderm induction is coordinated by marginal Nodal and ventral BMP signaling by controlling a range of downstream target genes. In line with ventral LPM emergence, endogenous *drl* expression was virtually absent in embryos maternally mutant for dominant Smad5 (*sbn*) that are devoid of ventral BMP activity (Hild et al., 1999) (Fig. 4A,B). In contrast, *MZoep*-mutant embryos lost the dorsal *drl* expression domain but still weakly expressed *drl* throughout the marginal zone, despite their inability to relay Nodal signaling in the gastrulating embryo (Fig. 4C). Consistent with our lineage tracing results (Fig. S3), embryos devoid of endoderm upon *sox32* knockdown also still expressed *drl* (Fig. 4D). These results underscore the importance of ventral BMP and marginal Nodal signaling in modulating early *drl* expression.

To uncover transcriptionally responsive regions within the drl^{LPM} enhancer, we targeted drl^{LPM} with a small library of sgRNAs to introduce maximally efficient Cas9-mediated deletions in zebrafish embryos transgenic for *-6.35drl:EGFP* or drl^{LPM}*:EGFP* (Fig. S8A) (Burger et al., 2016; Lindsay et al., 2016). This analysis revealed two mutagenesis-sensitive regions that caused mosaic loss of drl^{LPM} and *-6.35drl* reporter activity at mid-gastrulation/bud stage (Fig. S8B-E); our mutagenesis of the early-acting drl^{LPM} did not

significantly perturb overall transgene functionality, as EGFP reporter activity driven by the later-acting regulatory elements in *-6.35drl* remained intact (Fig. S8B``,C``).

The *drl*^{LPM} *3'* region affected in the crispants coincides with established ChIP-seq peaks for the maternally contributed T-box transcription factor Eomesodermin (Eomes) and its interaction partner FoxH1 that act with Nodal/BMP-mediating Smads; these factors together influence endodermal and mesodermal target genes (Charney et al., 2017; Nelson et al., 2014; Nelson et al., 2017; Slagle et al., 2011) including endogenous *drl* expression (Slagle et al., 2011) and bind the *drl*^{LPM} region (Fig. 4E) (Dubrulle et al., 2015; Nelson et al., 2014; Nelson et al., 2017). Ubiquitous expression of constitutively-active EomesA-VP16 strongly augmented and prolonged *drl*^{LPM} reporter expression in the native LPM domain up to mid-/late-somitogenesis compared to uninjected controls (Fig. 4F,G). Similarly, individual ubiquitous expression of native *eomesA* or *foxh1* mRNA strongly increased *drl*^{LPM} reporter expression within its native expression domain during mid-gastrulation (Fig. 4H-J).

From these observations, and since maternal *eomesA* and *foxh1* transcripts are seemingly ubiquitous during gastrulation (Bruce et al., 2003; Slagle et al., 2011), we reasoned that BMP/Nodal-based position information must have additional input on *drl*^{LPM} enhancer activity. To find this additional input, we performed individual and combinatorial mRNA-mediated misexpression of *eomesA* and *foxh1* with a variety of transcription factors implicated in early mesendoderm development (Fig. S9). These experiments identified the homeodomain transcription factor MixL1 as a third factor to act together with Eomes and FoxH1 in controlling *drl*^{LPM} enhancer activity (Fig. S9). MixL1 is a downstream target of BMP and Nodal signaling implicated in controlling endoderm and mesoderm fates (Kunwar et al., 2003; Stainier et al., 1996), retains ventral expression in *MZoep* mutants (Slagle et al., 2011), and triggers extensive blood formation at the expense of dorsal fates upon mRNA injection in *Xenopus* (Mead et al., 1996). Further, MixL1 can physically interact with Eomes (Bjornson et al., 2005) and MixL1, EomesA, FoxH1, and Smads have been detected by ChIP-Seq at the intronic *drl*^{LPM} enhancer region (Dubrulle et al., 2015; Nelson et al., 2017) (Fig. 4e). Mutating *mixl1* by Cas9 resulted in mosaic loss of *drl*^{LPM} reporter activity in F0 crispants (Fig. S9). Mutating the redundant *mixl1* paralog *mezzo* (Poulain and Lepage, 2002) alone or together with *mixl1* did not influence *drl*^{LPM} reporter activity (Fig. S9), suggesting that predominantly MixL1 is acting on *drl*^{LPM}.

Akin to *eomesA* and *foxh1*, native *mixl1* mRNA alone increased *drl*^{LPM} reporter expression in the native LPM domain (Fig. 4K). In contrast, combining the triplet of *eomesA*, *foxH1*, and *mixL1* (*E/F/M*) expressed from native ORF-encoding mRNA or from Tol2-based DNA vectors led to ubiquitous *drl*^{LPM} reporter activation (Fig. 4L,M). Combining native *E/F/M* in endoderm-perturbed

embryos also resulted in ubiquitous *drl*^{LPM} reporter activation (Fig. 4N), suggesting that *drl*^{LPM} reporter-induced cells have an LPM identity. Similarly, we found *E/F/M*-triggered induction of *tmem88a*, expression of which is highly enriched in the early native LPM (Mosimann et al., 2015) (Fig. S10). In contrast, the expression domains of various mesodermal genes remained broadly unaffected (Fig. S10), as for *hand2* expression that initiates in the lateral-most LPM only after gastrulation (Fig. S10; Fig. 1L). Even in *sbn*-mutant embryos without ventral BMP signaling, the constitutive-active mRNA or Tol2-based DNA vector combination of *E/F/M* induced *drl* expression dorsally, an activity that persisted even upon additional endoderm disruption (Fig. 4O; Fig. S11A,B). These observations suggest that *E/F/M* overexpression requires additional Smad activity, which in *sbn* is available in the dorsal Nodal-receiving domain. Conversely, loss of Nodal signaling in *MZoep* mutants led to a ventral upregulation of endogenous *drl* expression upon *E/F/M* overexpression (Fig. S11C,D). Together, these data demonstrate that in zebrafish the combination of EomesA, FoxH1, and MixL1 potentiates ventral Smad-relayed signals to demarcate a mesendoderm territory to become prospective LPM, as read out by the *drl*^{LPM} enhancer.

Its dynamic structure and the lack of a means to genetically mark the ventrally forming LPM in a specific manner has hindered concise comparative studies and *in toto* charting of the emergence and morphogenesis of the LPM. Here, we have shown that the *drl*^{LPM} enhancer reads out LPM-demarcating transcriptional activity in several chordate species, ranging from cephalochordates to amniotes, suggesting that the underlying positional program is of ancient evolutionary origin. In zebrafish, our data reveals that the transcription factors Eomes, FoxH1, and MixL1 are sufficient to trigger a basic LPM program. These factors have been previously implicated in mesendoderm development in various contexts, including blood formation (Mead et al., 1996); our observations now imply that, among their roles in mesendoderm regulation, they cooperate in inducing LPM together with ventral BMP-controlled and marginal Nodal-controlled Smad activity. These findings provide a developmental regulatory framework for the contribution of Eomes and MixL1 in early steps during programming naïve pluripotent stem cells into cardiac lineages (Pereira et al., 2012; Rao et al., 2015), suggesting that these factors in concert with the correct dose of BMP and Nodal establish the correct mesendodermal precursor lineage for forming cardiovascular-competent LPM.

References

Alexander, J. and Stainier, D. Y. (1999). A molecular pathway leading to endoderm formation in zebrafish. *Curr. Biol.* **9**, 1147–57.

Becker, D., Eid, R. and Schughart, K. (1996). The limb/LPM enhancer of the murine Hoxb6 gene: reporter gene analysis in

transgenic embryos and studies of DNA-protein interactions. *Pharm. Acta Helv.* **71**, 29–35.

Bertrand, S., Escriva, H., Williams, N. A., Holland, N. D. and Holland, L. Z. (2011). Evolutionary crossroads in developmental biology: amphioxus. *Development* **138**, 4819–30.

Bjornson, C. R. R., Griffin, K. J. P., Farr, G. H., Terashima, A., Himeda, C., Kikuchi, Y. and Kimelman, D. (2005). Eomesodermin Is a Localized Maternal Determinant Required for Endoderm Induction in Zebrafish. *Dev. Cell* **9**, 523–533.

Bruce, A. E. E., Howley, C., Zhou, Y., Vickers, S. L., Silver, L. M., King, M. Lou and Ho, R. K. (2003). The maternally expressed zebrafish T-box gene eomesodermin regulates organizer formation. *Development* **130**, 5503–5517.

Burger, A., Lindsay, H., Felker, A., Hess, C., Anders, C., Chiavacci, E., Zaugg, J., Weber, L. M. L. M., Catena, R., Jinek, M., et al. (2016). Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes. *Development* **143**, 2025–37.

Chal, J. and Pourquié, O. (2017). Making muscle: skeletal myogenesis in vivo and in vitro. *Development* **144**,.

Chandler, K. J., Chandler, R. L. and Mortlock, D. P. (2009). Identification of an ancient Bmp4 mesoderm enhancer located 46 kb from the promoter. *Dev. Biol.* **327**, 590–602.

Charney, R. M., Forouzmand, E., Cho, J. S., Cheung, J., Paraiso, K. D., Yasuoka, Y., Takahashi, S., Taira, M., Blitz, I. L., Xie, X., et al. (2017). Foxh1 Occupies cis -Regulatory Modules Prior to Dynamic Transcription Factor Interactions Controlling the Mesendoderm Gene Program. *Dev. Cell* **40**, 595–607.e4.

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

Dubrulle, J., Jordan, B. M., Akhmetova, L., Farrell, J. A., Kim, S.-H., Solnica-Krezel, L. and Schier, A. F. (2015). Response to Nodal morphogen gradient is determined by the kinetics of target gene induction. *Elife* **4**,.

Felker, A., Nieuwenhuize, S., Dolbois, A., Blazkova, K., Hess, C., Low, L. W. L. W. L., Burger, S., Samson, N., Carney, T. J. T. J., Bartunek, P., et al. (2016). In Vivo Performance and Properties of Tamoxifen Metabolites for CreERT2 Control. *PLoS One* **11**, e0152989.

Gays, D., Hess, C., Camporeale, A., Ala, U., Provero, P., Mosimann, C. and Santoro, M. M. (2017). An exclusive cellular and molecular network governs intestinal smooth muscle cell differentiation in vertebrates. *Development* **144**, dev.133926.

Gurdon, J. B. (1995). The Formation of Mesoderm and Muscle in Xenopus. In *Organization of the Early Vertebrate Embryo*, pp. 51–59. Boston, MA: Springer US.

Hatada, Y. and Stern, C. (1994). A fate map of the epiblast of the early chick embryo. *Development* **120**, 2879–2889.

Helker, C. S. M., Schuermann, A., Pollmann, C., Chng, S. C., Kiefer, F., Reversade, B. and Herzog, W. (2015). The

hormonal peptide Elabela guides angioblasts to the midline during vasculogenesis. *Elife* **4**, e06726.

Henninger, J., Santoso, B., Hans, S., Durand, E., Moore, J., Mosimann, C., Brand, M., Traver, D. and Zon, L. (2016). Clonal fate mapping quantifies the number of haematopoietic stem cells that arise during development. *Nat. Cell Biol.* **19**, 17–27.

Hild, M., Dick, A., Rauch, G. J., Meier, A., Bouwmeester, T., Haffter, P. and Hammerschmidt, M. (1999). The smad5 mutation somitabun blocks Bmp2b signaling during early dorsoventral patterning of the zebrafish embryo. *Development* **126**, 2149–59.

Holland, L. Z. and Holland, N. D. (2007). A revised fate map for amphioxus and the evolution of axial patterning in chordates. *Integr. Comp. Biol.* **47**, 360–72.

Hotta, K., Mitsuhara, K., Takahashi, H., Inaba, K., Oka, K., Gojobori, T. and Ikeo, K. (2007). A web-based interactive developmental table for the ascidian *Ciona intestinalis* , including 3D real-image embryo reconstructions: I. From fertilized egg to hatching larva. *Dev. Dyn.* **236**, 1790–1805.

Kozmikova, I. and Kozmik, Z. (2015). Gene regulation in amphioxus: An insight from transgenic studies in amphioxus and vertebrates. *Mar. Genomics* **24**, 159–166.

Kunwar, P. S., Zimmerman, S., Bennett, J. T., Chen, Y., Whitman, M. and Schier, A. F. (2003). Mixer/Bon and FoxH1/Sur have overlapping and divergent roles in Nodal signaling and mesendoderm induction. *Development* **130**, 5589–99.

Lindsay, H., Burger, A., Biyong, B., Felker, A., Hess, C., Zaugg, J., Chiavacci, E., Anders, C., Jinek, M., Mosimann, C., et al. (2016). CrispRVariants charts the mutation spectrum of genome engineering experiments. *Nat. Biotechnol.* **34**, 701–702.

Mahlapuu, M., Ormestad, M., Enerbäck, S. and Carlsson, P. (2001). The forkhead transcription factor Foxf1 is required for differentiation of extra-embryonic and lateral plate mesoderm. *Development* **128**, 155–66.

Mead, P. E., Brivanlou, I. H., Kelley, C. M. and Zon, L. I. (1996). BMP-4-responsive regulation of dorsal–ventral patterning by the homeobox protein Mix.1. *Nature* **382**, 357–360.

Mosimann, C., Panáková, D., Werdich, A. A. A., Musso, G., Burger, A., Lawson, K. L. K. L., Carr, L. A. L. A., Nevis, K. R. K. R., Sabeh, M. K. K., Zhou, Y., et al. (2015). Chamber identity programs drive early functional partitioning of the heart. *Nat. Commun.* **6**, 8146.

Nelson, A. C., Cutty, S. J., Niini, M., Stemple, D. L., Flicek, P., Houart, C., Bruce, A. and Wardle, F. C. (2014). Global identification of Smad2 and Eomesodermin targets in zebrafish identifies a conserved transcriptional network in mesendoderm and a novel role for Eomesodermin in repression of ectodermal gene expression. *BMC Biol.* **12**, 81.

Nelson, A. C., Cutty, S. J., Gasiunas, S. N., Deplae, I., Stemple, D. L. and Wardle, F. C. (2017). In Vivo Regulation of the Zebrafish Endoderm Progenitor Niche by T-Box Transcription Factors. *Cell Rep.* **19**, 2782–2795.

Onimaru, K., Shoguchi, E., Kuratani, S. and Tanaka, M. (2011). Development and evolution of the lateral plate mesoderm: comparative analysis of amphioxus and lamprey with implications for the acquisition of paired fins. *Dev. Biol.* **359**, 124–36.

Pereira, L. A., Wong, M. S., Mossman, A. K., Sourris, K., Janes, M. E., Knezevic, K., Hirst, C. E., Lim, S. M., Pimanda, J. E., Stanley, E. G., et al. (2012). Pdgfrα and Flk1 are direct target genes of Mixl1 in differentiating embryonic stem cells. *Stem Cell Res.* **8**, 165–179.

Perens, E. A., Garavito-Aguilar, Z. V, Guio-Vega, G. P., Peña, K. T., Schindler, Y. L. and Yelon, D. (2016). Hand2 inhibits kidney specification while promoting vein formation within the posterior mesoderm. *Elife* **5**, e19941.

Picker, A., Scholpp, S., Böhli, H., Takeda, H. and Brand, M. (2002). A novel positive transcriptional feedback loop in midbrain-hindbrain boundary development is revealed through analysis of the zebrafish pax2.1 promoter in transgenic lines. *Development* **129**, 3227–39.

Poulain, M. and Lepage, T. (2002). Mezzo, a paired-like homeobox protein is an immediate target of Nodal signalling and regulates endoderm specification in zebrafish. *Development* **129**,.

Rao, J., Pfeiffer, M. J., Frank, S., Adachi, K., Piccini, I., Quaranta, R., Araúzo-Bravo, M., Schwarz, J., Schade, D., Leidel, S., et al. (2015). Stepwise Clearance of Repressive Roadblocks Drives Cardiac Induction in Human ESCs. *Cell Stem Cell*.

Rojas, A., De Val, S., Heidt, A. B., Xu, S.-M., Bristow, J. and Black, B. L. (2005). Gata4 expression in lateral mesoderm is downstream of BMP4 and is activated directly by Forkhead and GATA transcription factors through a distal enhancer element. *Development* **132**, 3405–17.

Schmid, B., Shah, G., Scherf, N., Weber, M., Thierbach, K., Campos, C. P., Roeder, I., Aanstad, P. and Huisken, J. (2013). High-speed panoramic light-sheet microscopy reveals global endodermal cell dynamics. *Nat Commun* **4**, 2207.

Slagle, C. E., Aoki, T. and Burdine, R. D. (2011). Nodal-dependent mesendoderm specification requires the combinatorial activities of FoxH1 and Eomesodermin. *PLoS Genet.* **7**, e1002072.

Stainier, D. Y., Fouquet, B., Chen, J. N., Warren, K. S., Weinstein, B. M., Meiler, S. E., Mohideen, M. A., Neuhauss, S. C., Solnica-Krezel, L., Schier, A. F., et al. (1996). Mutations affecting the formation and function of the cardiovascular system in the zebrafish embryo. *Development* **123**, 285–92.

Stolfi, A. and Christiaen, L. (2012). Genetic and genomic toolbox of the chordate Ciona intestinalis. *Genetics* **192**, 55–66.

Takasato, M. and Little, M. H. (2015). The origin of the mammalian kidney: implications for recreating the kidney in vitro. *Development* **142**, 1937–1947.

Tonegawa, A., Funayama, N., Ueno, N. and Takahashi, Y. (1997). Mesodermal subdivision along the mediolateral axis in chicken controlled by different concentrations of BMP-4. *Development* **124**,.

Warga, R. M. and Nüsslein-Volhard, C. (1999). Origin and development of the zebrafish endoderm. *Development* **126**, 827–38.

Yabe, T., Hoshijima, K., Yamamoto, T. and Takada, S. (2016). Quadruple zebrafish mutant reveals different roles of Mesp genes in somite segmentation between mouse and zebrafish. *Development* **143**,.

Yelon, D., Ticho, B., Halpern, M. E., Ruvinsky, I., Ho, R. K., Silver, L. M. and Stainier, D. Y. R. (2000). The bHLH transcription factor hand2 plays parallel roles in zebrafish heart and pectoral fin development. *Development* **127**, 2573–82.

Zhang, X. Y. and Rodaway, A. R. F. (2007). SCL-GFP transgenic zebrafish: In vivo imaging of blood and endothelial development and identification of the initial site of definitive hematopoiesis. *Dev. Biol.* **307**, 179–194.

Zhu, H., Traver, D., Davidson, A. J., Dibiase, A., Thisse, C., Thisse, B., Nimer, S. and Zon, L. I. (2005). Regulation of the lmo2 promoter during hematopoietic and vascular development in zebrafish. *Dev Biol* **281**, 256–269.

Author Contributions

C.H, K.D.P., S.N., S.B, C.M. designed, performed, and interpreted zebrafish experiments; S.N., E.C., A.B. performed chicken experiments; K. P. performed lightsheet imaging with technical and equipment support by G.S., J.H; E.C.B, K.P., A.F. performed section-based zebrafish lineage tracing; K.W.R, P.M. provided and generated mutants and maternal-zygotic mutant zebrafish; H.J.P, M.B., R.K. designed, performed, and interpreted lamprey experiments; C.R., L.C. designed, performed, and interpreted *Ciona* experiments; I.K, Z.K designed, performed, and interpreted amphioxus experiments; C.H., K.P., and C.M. assembled and wrote the manuscript with contributions from all co-authors.

Acknowledgements

We thank Sibylle Burger and Seraina Bötschi for technical and husbandry support; the lab of Dr. Stephan Neuhauss for zebrafish husbandry support; the labs of Dr. Esther Stöckli and Dr. Jerome Gros for chicken experimentation support; the ZBM at UZH for imaging support; Dr. Fiona Wardle for input and support on the ChIP-seq panel in Fig. 4E; the lab of Dr. Magdalini Polymenidou for vibratome access; Karolína Ditrychová for cloning the *pKD001* construct; Dr. Ashley Bruce, Dr. Rebecca Burdine, and Dr. Michael Tsang for sharing transcription factor constructs; the Stowers Institute histology facility for assistance with lamprey embryo sectioning; and all members of the Mosimann lab for constructive input. This work has been supported by a Swiss National Science Foundation (SNSF) professorship [PP00P3_139093] and SNSF R'Equip grant 150838 (Lightsheet Fluorescence Microscopy), a Marie Curie Career Integration Grant from the European Commission [CIG PCIG14-GA-2013-631984], the Canton of Zürich, the

UZH Foundation for Research in Science and the Humanities, the Swiss Heart Foundation, and the ZUNIV FAN/UZH Alumni to C.M; a UZH CanDoc to C.H.; EuFishBioMed and Company of Biologists travel fellowships to K.D.P.; the Stowers Institute (grant #1001) to H.J.P. and R.K.; NIH/NHLBI R01 award HL108643, trans-Atlantic network of excellence award 15CVD01 from the Leducq Foundation to L.C.; a long-term fellowship ALTF 1608-2014 from EMBO to C.R.

Figures

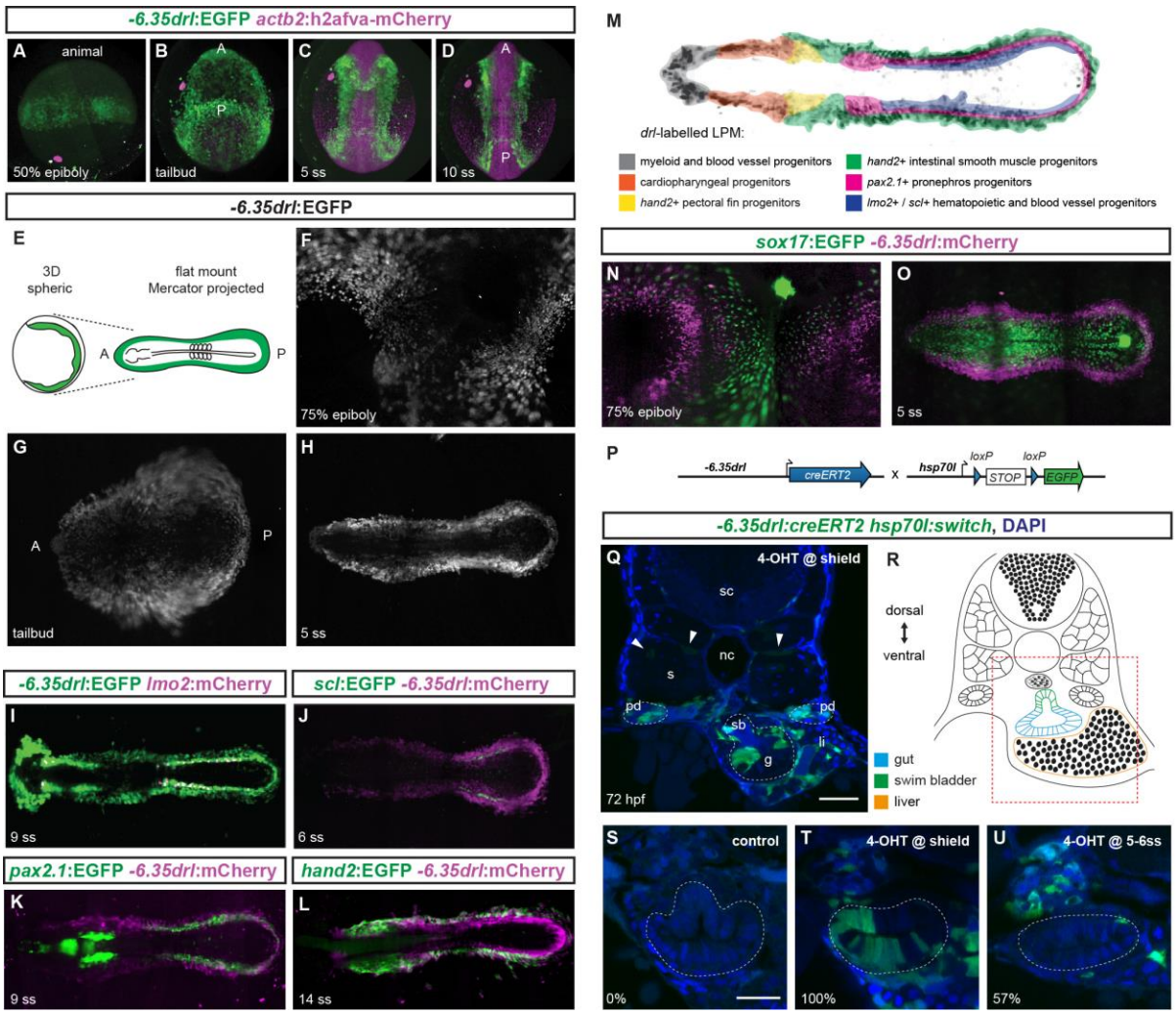

Figure 1: The LPM forms around the entire circumference of the developing zebrafish embryo. (**A-D**) Panoramic SPIM imaging of 50% epiboly to 10 ss embryos transgenic for *-6.35drl*:*EGFP* (green) and *actb2:h2afva-mCherry* (nuclei marker, magenta); maximum-intensity-projected, lateral (**A**), dorsal (**B-D**) views. (**E-O**) Panoramic SPIM imaged zebrafish embryos shown as Mercator projections. (**E**) Schematic of Mercator projection to visualize entire embryo. (**F-l**) Single time point projections (dorsal views) of double-transgenics for *-6.35drl* reporters and *lmo2*, *scl*, *pax2.1*, or *hand2* reporters co-expressed in dedicated LPM territories. (**M**) *-6.35drl* reporter embryo schematic, depicting LPM sub-territories at 10 ss. (**N,O**) Time lapse projections of *-6.35drl:mCherry* (magenta) with *sox17:EGFP*, showing LPM versus *sox17*-marked endoderm from gastrulation until 10 ss. (**P**) Schematic of *-6.35drl:creERT2 x hsp70l:switch* cross for genetic lineage tracing. (**Q**) Representative whole-body transverse section of 3 dpf *-6.35drl* lineage-traced embryo 4-OHT-induced at shield stage. Arrowheads depict rare trunk muscle labeling. (**R**) Schematic of transverse section with endoderm-derived organs at 3 dpf. (**S-U**) Transverse sections of *-6.35drl* lineage tracing, 4-OHT induction shield stage to 5-6 ss; note gradual loss of endoderm labeling (intestinal cells, dashed region). Numbers indicate percentage of embryos with intestinal lineage-labeling. Notochord (nc), somite (s), spinal cord (sc), pronephric duct (pd), liver (li), swim bladder (sb), and gut (g) labeled in (**q**), nuclei in blue (DAPI; **Q**,**S-U**). (**F-O**) Anterior to the left. Scale bar (**Q**) 50 µm and (**S-U**) 25 µm.

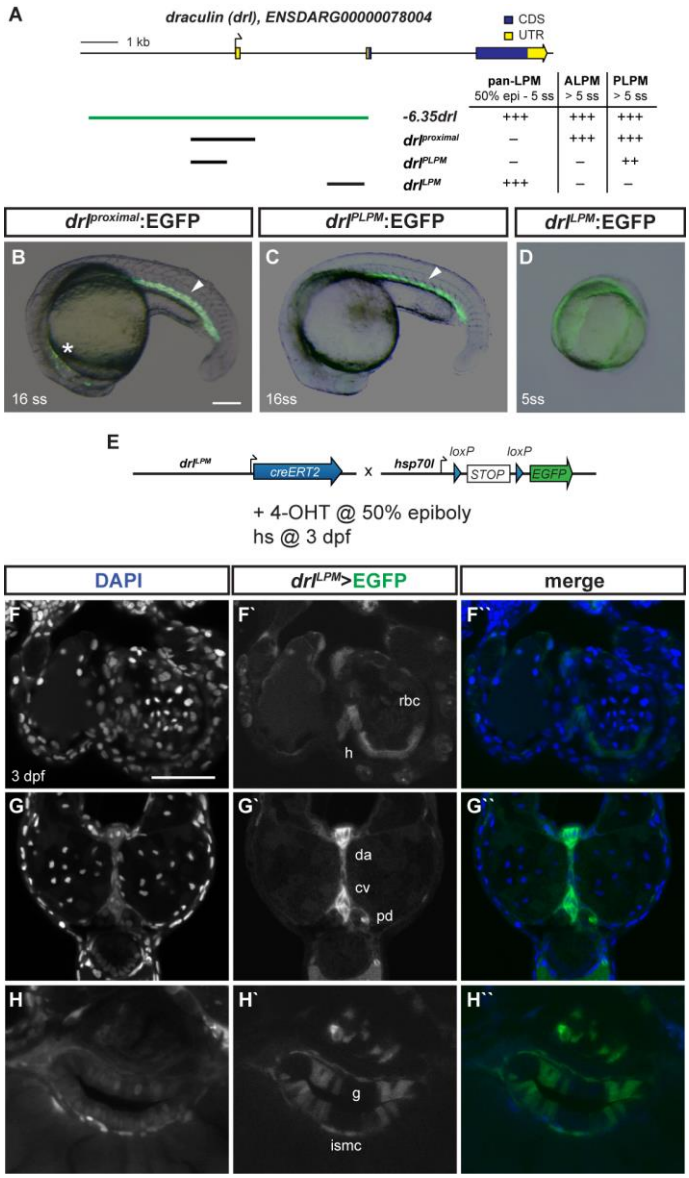

Figure 2: The *-6.35drl cis*-regulatory region contains an early pan-LPM enhancer. (**A**) Schematic of *drl* locus, *-6.35drl cis*-regulatory region (green), smaller isolated fragments $drl^{proximal}$, drl^{PLPM}, and drl^{LPM} with specific reporter activity. Table indicates pan-LPM expression (50% to 5 ss) or ALPM/PLPM from 5 ss onwards, absent expression (-) to strong expression (+++). (**B-D**) Spatio-temporal expression of drl^{PLPM}, $drl^{proximal}$, and drl^{LPM} reporter transgenics; at 16 ss, drl^{PLPM} and $drl^{proximal}$ express in PLPM (arrowheads in **B**,**C**), drl^{PLPM} additionally in ALPM (asterisk in **B**); note pan-LPM activity of drl^{LPM} (**D**). (**E**) Schematic representation of drl^{LPM}*:creERT2 x hsp70l:switch* cross for genetic lineage tracing. (**F-H**) Transverse sections at 3 dpf of drl^{LPM}*:creERT2* lineage tracing after 4-OHT induction at shield stage results in specific labeling of LPM-derived tissue including heart (h), red blood cells (rbc), dorsal aorta (da) and cardinal vein (cv) endothelium, pronephric duct (pd), and intestinal smooth muscle cells (ismc). drl^{LPM}*:creERT2* also traces endoderm-derived tissue, shown for gut (g). Nuclei counterstained with DAPI (blue). Scale bars (**B**) 250 µm and (**F**) 50 µm.

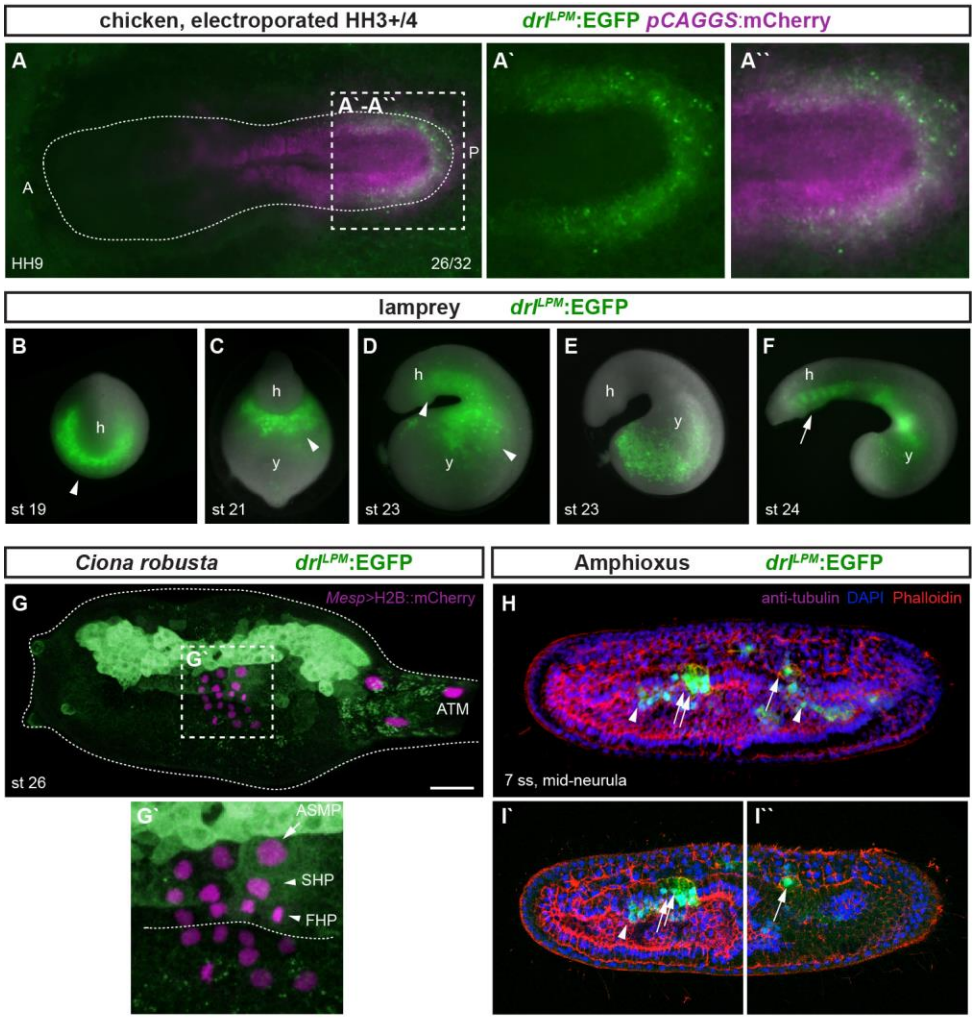

Figure 3: The zebrafish *drl^{LPM}* *cis*-regulatory element reads out a LPM program in various chordates. (**A**) HH9-staged *ex ovo*-cultured chicken embryos electroporated at HH3+/4 with *drl^{LPM}:EGFP* and ubiquitous *pCAGGS:mCherry* control (magenta), showing specific *drl^{LPM}* reporter expression in the electroporated LPM. The dashed line depicts outline of chicken embryo, anterior to the left; zoom-in depicts boxed region in (**A**), single and merged channel (**A`, A``**). (**B-F**) Transient transgenic lamprey embryos with *drl^{LPM}*:EGFP expression in the anterior mesendoderm (arrowheads) and overlying the yolk at neurula stages (st 19-21) (**B-C**), and in the developing pharynx (arrow) during head protrusion (st 23-24) (**C-F**); views anterior (**B**), ventral (**C**), lateral (**D-F**), head (h) and yolk (y). (**G**) *Ciona robusta* embryos at 18 hpf (st 26) expressing *drl^{LPM}*:EGFP, with *Mesp*>H2B::mCherry to track the B7.5 cardiopharyngeal cell lineage (magenta). EGFP is specifically expressed in mesenchymal lineages and partly overlaps with B7.5 derivatives including atrial siphon muscles precursors (ASMP) (white arrows) and both first and second heart precursors (FHPs and SHPs) (white arrowheads). Representative larva shown, zoom-in of boxed region in (**G`**); dashed line indicates midline, anterior to the left, scale bar (**G**) 25 µm. (**H**) Confocal Z-stack of mid-neurula stage (7 ss) amphioxus embryo injected with *drl^{LPM}:EGFP*, showing specific reporter activity in lateral endoderm (arrowhead), ventral half of somites (arrows), and evaginated somites (double arrows). Embryo counterstained with Phalloidin (red), DAPI (blue); lateral view as 3D-rendering (**H**) and Z-stack sagittal sections (**H`- H``**), anterior to the left, dorsal to the top.

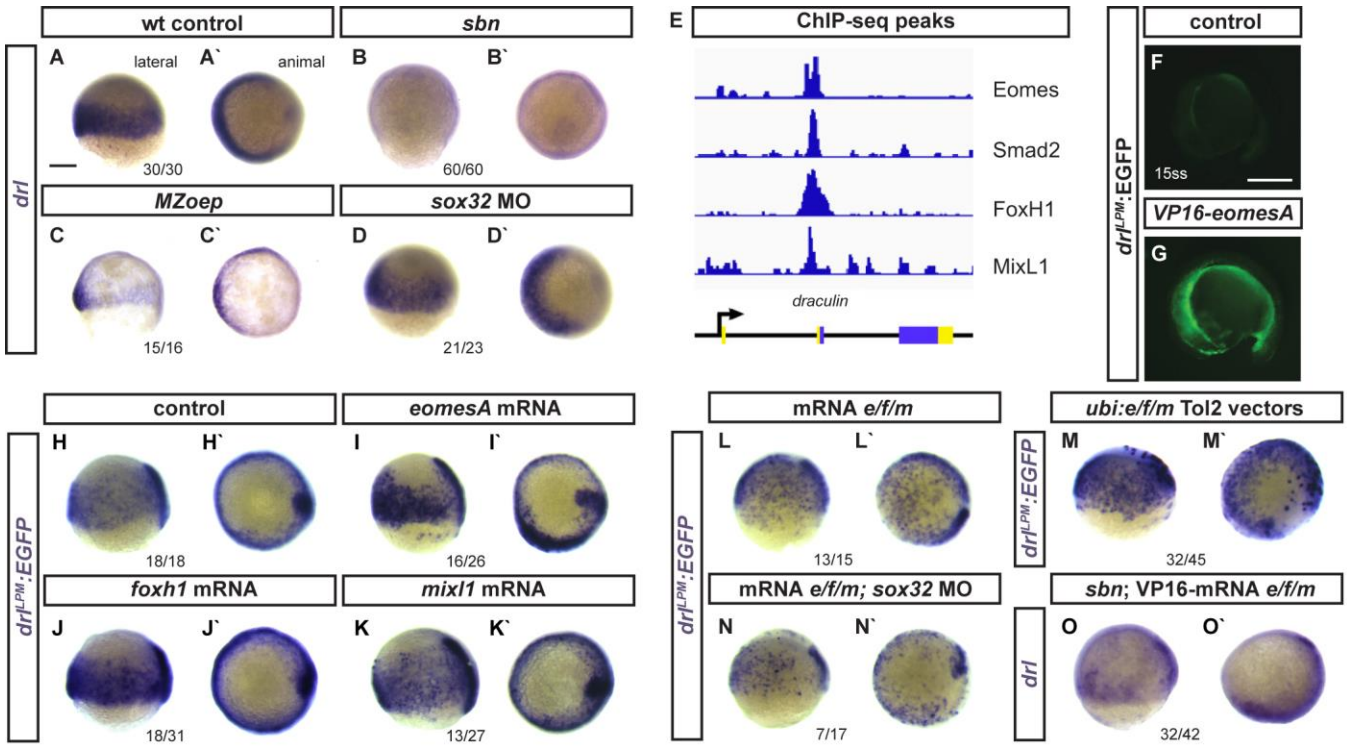

Figure 4: Eomes, FoxH1, and MixL1 induce *drl*^{LPM} activity. (**A-D**) mRNA ISH for endogenous *drl* expression in wildtype (wt), BMP-perturbed (*sbn*), Nodal-perturbed (*MZoep*), and endoderm-perturbed (*sox32* morphant) backgrounds, shield stage; *drl* expression is strongly suppressed in *sbn*, dorsally lost in *MZoep*, and unaffected upon *sox32* perturbation. (**E**) ChIP-seq peaks of Eomes, Smad2 (Nelson et al., 2014), FoxH1 (Charney et al., 2017), and MixL1 (Nelson et al., 2017) in the endogenous *drl* locus. (**F**,**G**) Microinjection of *VP16-eomesA* in *drl*^{LPM}*:EGFP* reporter transgenics enhances and prolongs EGFP expression in native *drl* expression domain. (**H-K**) *EGFP* ISH expression on *drl*^{LPM}*:EGFP* embryos injected with full-length *eomesA*, *foxh1*, or *mixl1* shows enhanced reporter activity in the native LPM territory. (**L-O**) *EGFP* ISH in *drl*^{LPM}*:EGFP* embryos injected with *eomesA (e)*, *foxh1 (f)*, and *mixl1 (m)* (*e/f/m*) as mRNA (**L**) or as *ubi*-driven DNA vectors (**M**) show ectopic reporter expression, also when endoderm is perturbed in *sox32* morphants (**N**). (**O**) ISH for endogenous *drl* in ventral BMP mutants *(sbn)* injected with *e/f/m* as constitutively active VP16 fusion ORFs, note a partial restoration of *drl* expression at dorsal Nodal-active margin. (**A-D**,**H-O**) Lateral and animal views shown at shield stage as indicated. Scale bar (**A**) 250 μm and (**F**) 400 μm.

Material and Methods

Animal experiments and husbandry

Zebrafish and chick experiments were carried out in accordance with the recommendations of the national authorities of Switzerland (Animal Protection Ordinance). The protocols and the experiments were approved by the cantonal veterinary office of the Canton Zurich (Kantonales Veterinäramt). Zebrafish care and all experimental procedures were carried out in accordance with the European Communities Council Directive (86/609/EEC), according to which all embryo experiments performed before 120 hours post fertilization are not considered animal experimentation and do not require ethics approval. Adult zebrafish for breeding were kept and handled according to animal care regulation of the Kantonales Veterinäramt Zürich (TV4209). All zebrafish (*Danio rerio*) were raised, kept, and handled as described (Westerfield, 2007). To assure same-staged embryos, adults were separated before mating by dividers. Embryos were collected and mixed to ensure a homologous diversity of clutches, and subsequently raised in E3 medium in temperature-controlled incubators at 21, 28, and 32°C respectively. Lamprey studies were conducted in accordance with the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health, and protocols were approved by the Institutional Animal Care and Use Committees of the California Institute of Technology (Protocol # 1436-11).

Transgenic constructs and transgenic zebrafish lines

The upstream *cis*-regulatory elements of the zebrafish *drl* gene (*ENSDARG00000078004*; *ZDB-GENE-991213-3*) were amplified from zebrafish wildtype genomic DNA (Extended Data Table 1; regulatory elements) and TOPO-cloned into the *pENTR™ 5'-TOPO® TA* Cloning® plasmid (Invitrogen) according to the manufacturer's instructions.

Subsequent cloning reactions were performed with the Multisite Gateway system with LR Clonase II Plus (Life Technologies) according to the manufacturer's instructions. *drl^{PLPM}:EGFP (pDestTol2pA2_ drl^{PLPM}:EGFP)* and *drl^{proximal}:EGFP (pDestTol2pA2_ drl^{proximal}:EGFP)* were assembled from pENTR/5'_*drl^{PLPM}* or pENTR/5'_*drl^{proximal}* together with Tol2kit vectors *#383* (*pME-EGFP*), *#302* (*p3E_SV40polyA*), and *#394* (*pDestTol2A2*) as backbone. *drl^{LPM}:EGFP (pDestTol2pA2_drl^{LPM}:EGFP)* and *drl^{minLPM}:EGFP (pDestTol2pA2_ drl^{minLPM}:EGFP)* were assembled from pENTR/5'_*drl^{LPM}* or pENTR/5'_*drl^{minLPM}* together with *pME-β-globin_minpromoter_EGFP* (Tamplin et al., 2015), Tol2kit vectors *#302* (*p3E_SV40polyA*), and *#394* (*pDestTol2A2*) as backbone (Kwan et al., 2007). *drl^{LPM}:creERT2 (pDestTol2CY_drl^{LPM}:creERT2,alpha-crystallin:YFP)* and *-5.0sox17:creERT2 (pDestTol2CY_-5.0sox17:creERT2,alpha-crystallin:YFP)* were assembled from *pENTR/5'_drl^{LPM}* or *pENTR/5'_-5.0sox17* together with *pCM293* (*pENTR/D_creERT2*)(Mosimann et al., 2011), Tol2kit vector *#302* (*p3E_SV40polyA*), and *pCM326 (pDestTol2CY)* as backbone (Mosimann et al., 2015).

The regulatory elements of mouse-specific LPM enhancers were PCR-amplified (Extended Data Table 1; regulatory elements), TOPO-cloned into the *pENTR5'* plasmid, and assembled together with *pKD001* (*pME-β-globin_minpromoter_EGFP* with improved Kozak sequence), Tol2kit vectors *#302* (*p3E_SV40polyA*), and *#394* (*pDestTol2A2*) as backbone.

Assembled reporter constructs were injected at a concentration of 25 ng/µL together with 25 ng/µL *Tol2* mRNA for Tol2-mediated zebrafish transgenesis (Felker and Mosimann, 2016). Injected F0 founders were screened for specific EGFP or *alpha-crystallin:*YFP expression. Fish were raised to adulthood and screened in F1 for germline transmission. Single-insertion transgenic strains were generated, and microscopy images were taken on a Leica M205FA with a Leica DFC450C digital camera or a Leica SP8 confocal microscope with Plan-Apochromat 20x/0.5 objective. Images were processed using Leica LAS and Fiji.

Established transgenic and mutant lines used in this study included *-6.35drl:EGFP* (Mosimann et al., 2015), *-6.35drl:mCherry^{zh705}* (Sánchez-Iranzo et al., 2018), *-6.35drl:creERT2* (Mosimann et al., 2015), *ubi:creERT2* (Mosimann et al., 2011), *ubi:switch* (Mosimann et al., 2011), *hsp70l:switch^{zh701}*, *lmo2:dsRED2* (Zhu et al., 2005), *scl/tal1:EGFP* (Jin et al., 2006), *pax2.1:EGFP* (Picker et al., 2002), *hand2:EGFP* (Yin et al., 2010),

actb2:h2afva-mCherry (Krens et al., 2011), maternal and zygotic EGF-CFC co-receptor *oep* gene mutants (*MZoep*) (Gritsman et al., 1999), and *somitabun*^{dtc24} mutants (*sbn, smad5*) (Hild et al., 1999; Mullins et al., 1996).

Zebrafish CreERT2-based lineage tracing

Lineage tracing experiments were performed by crossing female *hsp70l:switch* or *ubi:switch* (Mosimann et al., 2011) reporter carriers with the male *creERT2* drivers *-6.35drl:creERT2* (Mosimann et al., 2015), *sox17:creERT2*, and *drl*^{LPM}*:creERT2*. Embryos were induced using 4-Hydroxytamoxifen (4-OHT) (Cat#H7904; Sigma H7904) from fresh or pre-heated (65°C for 10 min) stock solutions in DMSO at a final concentration of 10 µM in E3 embryo medium at indicated time points (Felker et al., 2016). Embryos were washed in untreated E3 medium at 24 hpf. To induce EGFP transcription in *hsp70l:switch* carrying embryos, the embryos were incubated at 37°C for 1 hour, 3 hours before fixation. Embryos were fixed with 4% Paraformaldehyde (PFA) overnight at 4°C at 3 dpf and processed for confocal analysis.

Zebrafish transverse vibratome sections

Transverse sections were generated as previously described (Gays et al., 2017). Fixed embryos were washed in PBS, embedded in 6% low-melting agarose (Sigma-Aldrich) in PBS/0.1% Tween-20 (Sigma-Aldrich), and cut into 130 µm thick sections using a vibratome (Leica VT 1000S). Sections were mounted in DAPI-containing Vectashield (Cat#H-1200; Vector Laboratories). Sections were analyzed with a Zeiss LSM710 confocal microscope with a Plan-Apochromat 40x/1.3 oil DIC M27 objective. Images were cropped and adjusted for brightness using ImageJ/Fiji (Schindelin et al., 2012). Graphs were generated in Graphpad Prism 5.

Zebrafish gene knockdown experiments

The previously established *sox32* morpholino (Sakaguchi et al., 2001) was synthesized by GeneTools (sequence: *5'-TGCTTTGTCATGCTGATGTAGTGGG-3'*) and dissolved in nuclease-free water to a stock concentration of 10 µg/µl. The *sox32* morpholino injection mix of 4 µg/µl was incubated for 10 min at 65°C, and 8 ng were microinjected into one-cell stage zebrafish embryos.

Zebrafish CRISPR-Cas9 site-directed mutagenesis

sgRNAs were obtained in an oligo-based approach as previously described (Burger et al., 2016). Briefly, primer extension was performed using Phusion polymerase (NEB) of forward primer 5'-*GAAATTAATACGACTCACTATAGG-N*$_{20}$*-GTTTTAGAGCTAGAAATAGC*-3' (including 20 nt target site) and invariant reverse primer 5'-*AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC*-3' (PAGE-purified) (Bassett et al., 2013). The following target sites are shown in this manuscript: 5'-*AGATTTGTTTAGTCAGTGTC*-3' (*drl ccG*), 5'-*GTCTGGAAACAGTCTGAATC*-3' (*drl ccH*), 5'-*GAGACTTCGCCCTTCGGTTC*-3' (*mixl1 ccB*), and 5'-*GACAGAACAGGCCACGTTGA*-3' (*mezzo ccA*).

In vitro transcription of sgRNAs was performed as previously described (Burger et al., 2016) using MAXIscript T7 (Ambion). Afterwards, RNA was precipitated with ammonium acetate, washed with 75% ethanol, and dissolved in DEPC water. RNA quality was checked on a MOPS gel. RNPs of Cas9 protein (Cas9-mCherry, pMJ293 (available from Addgene)) (Burger et al., 2016) and sgRNA were *in vitro*-assembled for 10 min at 37°C in 300 mM KCl to ensure maximum cleavage efficiency and microinjected into the cell of one-cell stage embryos (Burger et al., 2016). The CRISPR target regions in the transgenic locus were amplified using specific target region amplification primers (Table S1; target region amplifications).

Sequencing analysis was performed with the R software package CrispRVariants. CrispRVariants enables variant counting and visualization (Lindsay et al., 2015).

Zebrafish overexpression experiments

For early developmental transcription factor genes, ORFs were PCR-amplified from mixed-stage zebrafish cDNA using ORF-specific primers (Table S1; coding sequence primers); full-length zebrafish *eomesA* was derived from I.M.A.G.E clone *IRBOp991A0739D* (Source BioScience LifeSciences, UK). All CDS were TOPO-cloned into the *pENTR/D*™ Directional TOPO® plasmid (Invitrogen) according to the manufacturer's instructions. Subsequent cloning reactions were performed with the Multisite Gateway system with LR Clonase II Plus (Invitrogen) according to the manufacturer's instructions and as previously described (Felker and Mosimann, 2016). The *CDS* as pENTR/D vectors were assembled with *pENTR5'_ubi* (Mosimann et al., 2011), *pENTR5'_T7-VP16*, *pENTR5'_T7-eng* or *pENTR5'_T7* together with Tol2kit vectors *#302* (*p3E_SV40polyA*), and *#394* (*pDestTol2A2*) as backbone (Kwan et al., 2007). Afterwards, plasmids were linearized (in case of *T7-VP16*, *T7-eng* or *T7*) and *in vitro* transcribed (Roche). The *EomesA-VP16* plasmid (containing 153aa-431aa of the zebrafish EomesA ORF) (Bruce et al., 2003) was kindly provided by Dr. Rebecca Burdine. The plasmid was linearized via NotI followed by SP6 *in vitro* transcription. The *Ntl-VP16* plasmid was kindly provided by Dr. Ashley Bruce. The *etv5a-VP16* and *etv5a-eng* plasmids were kindly provided by Dr. Michael Tsang. All mRNAs were precipitated with ammonium acetate, washed with 75% ethanol, and dissolved in DEPC water. mRNA quality was checked on MOPS gels as described (Burger et al., 2016).

Zebrafish whole-mount *in situ* hybridization

Total RNA was extracted from zebrafish embryos from various stages during development. This RNA was used as template for generation of first-strand complementary DNA (cDNA) by Superscript III First-Strand Synthesis kit (Cat#18080051; Invitrogen). *In situ* hybridization (ISH) probes were designed with an oligonucleotide-based method (including T7 promoter added to the reverse primers) using zebrafish cDNA (Table S1; *in situ* hybridization probes). The following oligonucleotide pairs (including T7 promoter added to the reverse primers) were used to amplify the DNA template from zebrafish cDNA. The ISH probe for *admp* was obtained from pCS2_ADMP plasmid, and *sizzled* from pCS2_Sizzled. *Admp* and *sizzled* were linearized by ClaI. The *gata2a* probe was obtained from the middle entry vector *pCM238.*. For *in vitro* transcription, T7 RNA polymerase (Roche) and digoxygenin (DIG)-labeled NTPs (Roche) were used. Afterwards, RNA was precipitated with lithium chloride, washed with 75% ethanol, and dissolved in DEPC water. RNA quality was checked on a MOPS gel. ISH on whole-mount zebrafish embryos was executed as described before (Thisse and Thisse, 2008). After ISH, embryos were transferred to 80-95% glycerol (Sigma-Aldrich), and microscopy images were taken on a Leica M205FA with a Leica DFC450C digital camera. Images were cropped and adjusted for brightness using ImageJ/Fiji.

Zebrafish selective plan illumination microscopy (SPIM)

At 30-50% epiboly, embryos in the chorion were embedded into 1% low-melting agarose with optional 0.016% Ethyl 3-aminobenzoate methanesulfonate salt (Tricaine, Cat#A5040; Sigma) in E3 embryo medium, and sucked into a FEP tube (inner diameter: 2.0 mm, wall thickness: 0.5 mm). 6-7 embryos were positioned on top of each other. The FEP tube was mounted in the microscope imaging chamber filled with E3 medium. Panoramic (3D) SPIM/lightsheet microscopy and subsequent image processing (Mercator projections) were performed as previously described (Schmid et al., 2013). A z-stack of 402 planes was obtained from every embryo with an interval of 2 min for a period of 14-17 hours. Images were processed using Leica LAS, ImageJ/Fiji, and Photoshop CS6.

Chicken embryo incubation and *ex evo* culturing

Ex ovo culturing was adapted from previously established protocols (New, 1955). Fertilized chicken eggs were obtained from a local hatchery and stored at 12°C up to maximum of 14 days. Prior to use, eggs were incubated horizontally for 17 hours until Hamburger-Hamilton (HH) 3+/4 in a 39°C incubator with 55-65% humidity. After incubation, the eggs were kept for at least 30 min at RT before opening. Eggs were opened in

a petri dish and a layer of thick albumin together with the chalaziferous layer was removed using a plastic Pasteur pipette. A paper ring was placed around the embryo on the yolk and dissection scissors were used to cut the yolk membrane around the ring. The paper ring with the embryo was cleaned from remaining yolk and transferred and placed upside down on a semisolid albumin/agarose (43.5 ml thin albumin incubated for 2 hours at 55ºC, 5 ml 2% agarose, 1.5 ml 10% glucose in 30 mm petri dishes) culturing plate. Embryos were recovered for at least 2 hours at RT before electroporation.

Chicken embryo injection and electroporation

For electroporations, a customized electroporation chamber was used containing an electrode with a positive pole on the bottom of the chamber and separate negative electrode on a holder (kindly provided by the lab of Dr. Jerome Gros). Both electrodes were mounted and connected to a square wave electroporator (BTX ECM 830). The electroporation chamber was filled with HBSS (Gibco Life Technologies), and the embryo-containing paper ring was placed in the chamber with the dorsal side up. The DNA mixture was injected by a mouth injector along the primitive streak beneath the pellucid membrane. The positive electrode holder was placed on top of the streak to allow electricity pulses flow through the embryo (3 pulses, 8V x 50 milliseconds, 500 milliseconds interval). The embryos were placed back on the albumin culturing plates with the ventral side up and placed back at 39°C until HH8-9. Microscopy images of the embryos were taken at HH8-9 on a Leica M205FA with a Leica DFC450C digital camera. Images were cropped and adjusted for brightness using ImageJ/Fiji.

All injection mixtures for electroporations contained 0.1% fast green dye, 0.1% methyl-cellulose, 300 ng/µl control plasmid *pCAGGs* (*pCMV:H2B-CAGG-RFP*, abbreviated for *chicken β-actin promoter CAGG-mCherry)* and 1 µg/µl of the plasmid of interest.

Lamprey experiments

The drl^{LPM} and drl^{minLPM} regulatory elements were amplified from the zebrafish vectors drl^{LPM}*:EGFP* and drl^{minLPM}*:EGFP* by PCR using KOD Hot Start Master Mix (Novagen) (Table S1; regulatory elements). The amplified enhancers were cloned into the HLC vector for lamprey transgenesis (Parker et al., 2014b), containing the mouse *c-Fos* minimal promoter, by Gibson assembly using the Gibson Assembly Master Mix (NEB).

Injections for I-SceI meganuclease-mediated lamprey transient transgenesis were performed using *P. marinus* embryos at the one-cell stage with injection mixtures containing 0.5 U/µl I-SceI enzyme and 20 ng/µl reporter construct as described previously (Parker et al., 2014a). Selected EGFP-expressing embryos were fixed in MEM-FA and dehydrated in methanol for *in situ* hybridization. EGFP-expressing embryos were imaged using a Zeiss SteREO Discovery V12 microscope with variable zoom and a Zeiss Axiocam MRm camera with AxioVision Rel 4.6 software. Images were cropped and adjusted for brightness using Adobe Photoshop CS5.1.

For mRNA ISH, total RNA was extracted from st 21-26 *P. marinus* embryos using the RNAqueous Total RNA Isolation Kit (Ambion). This was used as a template for 3' rapid amplification of cDNA ends (RACE) with the GeneRacer Kit and SuperScript III RT (Invitrogen). A 339bp-long *pmHandA in situ* probe was designed based on a characterized cDNA sequence from the closely related Arctic lamprey (*Lethenteron camtschaticum*) (Kuraku et al., 2010), and this sequence was amplified by PCR from 3' RACE cDNA using KOD Hot Start Master Mix (Novagen) with the following primers: 5'-*GCGGAGGACATTGAGCATC*-3' (forward) and 5'-*TGGAATTCGAGTGCCCACA*-3' (reverse). The cDNA fragment was cloned into the *pCR4-TOPO* vector (Invitrogen). The 709bp-long *eGFP* probe was described previously (Parker et al., 2014b).

DIG-labelled probes were generated and used in lamprey whole-mount ISH as described previously(Sauka-Spengler et al., 2007). Embryos were cleared in 75% glycerol and imaged using a Leica MZ APO microscope with variable zoom and Lumenera Infinity 3 camera with Lumenera Infinity Capture v6.5.3 software. Images were cropped and adjusted for brightness using Adobe Photoshop CS5.1.

After ISH, selected embryos were transferred into 30% sucrose in PBS, embedded in O.C.T. Compound (Tissue-Tek), and cut into 10 µm-thick cryosections using a CryoStar NX70 cryostat (Thermo Scientific). Images were taken using a Zeiss Axiovert 200 microscope with AxioCam HRc camera and AxioVision Rel 4.8.2 software.

Ciona experiments

The regulatory elements *-6.35drl* and *drl*^{LPM} were amplified from the zebrafish vectors *drl*^{LPM}*:EGFP* and *drl*^{minLPM}*:EGFP* and subcloned in a vector upstream of *unc76:GFP* (Table S1; regulatory elements).

Gravid *Ciona robusta* adults were obtained from M-REP (San Diego CA, USA). To test the activity of the zebrafish enhancers in *Ciona robusta*, 80 µg of *drl*^{LPM}*:EGFP* or *drl*^{minLPM}*:EGFP* were injected in a mixture with the reporter plasmid *Mesp* (Davidson et al., 2005) to mark the B7.5 cardiopharyngeal lineage with *H2B:mCherry* (10 µg).

For antibody staining, embryos were fixed in 4% MEM-PFA for 30 min, rinsed several times in PBT (PBS/0.1% Tween-20), and incubated with anti-GFP (1:500, Mouse mAb, Roche) with 2% normal goat serum in PBT at 4°C overnight. Embryos were washed in PBT and then incubated with donkey anti-mouse secondary antibody (1:1000) coupled to Alexa Fluor 488 (Life Technologies) in PBT with 2% normal goat serum for 2 hours at RT, then washed in PBT(Racioppi et al., 2014).

Amphioxus experiments

The regulatory elements *-6.35drl* and *drl*^{LPM} were amplified from the zebrafish vectors *-6.35drl:EGFP* and *drl*^{LPM}*:EGFP* and subcloned upstream of a EGFP reporter in the pPB vector carrying PiggyBac transposon terminal repeats (Kozmikova and Kozmik, 2015). Adults of *Branchiostoma lanceolatum* were collected in Banyuls-sur-mer, France, prior to the summer breeding season and raised in the laboratory until spawning. The spawning of amphioxus male and females was induced by shifting of the temperature as described (Fuentes et al., 2007). For microinjection of amphioxus eggs, a mixture of *pPB-6.35drl:EGFP* or *pPB-drl*^{LPM}*:EGFP* (200 ng/µl) with PiggyBac transposase mRNA (100 ng/µl) in 15% glycerol was used. Transgenic embryos were allowed to develop until neurula stage, fixed in 4% PFA overnight at 4°C, mounted with Vectashield with DAPI (Vector Laboratories), and analyzed by Leica SP5 confocal microscope. The confocal images were adjusted for brightness and contrast with ImageJ/Fiji.

Table S1: Primer sequences

regulatory elements	
drl^{PLPM} zebrafish	*5'-TTTCAATTTCAAAGGAGC-3'* *5'-TGTTTTGAGTGAGTCGGTTAGG-3'*
drl^{proximal} zebrafish	*5'-TTTCAATTTCAAAGGAGC-3'* *5'-CAACTGGGACATATTTAAACC-3'*
drl^{LPM} zebrafish	*5'-CACTTCCTAATCTGCTATTGCC-3'* *5'-CCAAGTGTGAATTGGGATCG-3'*
for lamprey expression	*5'-CCCTCGAGGTCGACGCACTTCCTAATCTGCTATTG-3'* *5'- GAGGATATCGAGCTCGCCAAGTGTGAATTGGGATC-3'*
for ciona expression	*5'-CTCGAGCACTTCCTAATCTGCTATTGCC-3' (w/ Xho1)* *5'-TCTAGACCAAGTGTGAATTGGGATCG-3' (w/ Xba1)*
drl^{minLPM}	*5'-GTCAAATCAATCAATGCGGTGC-3'*

	5'-CCAAAGACAATTAAACCACC-3'
for lamprey expression	*5'-AGGGTAATGAGGGCCGTCAAATCAATCAATGCGGTGCTC-3'* *5'GAGGATATCGAGCTCGCCAAAGACAATTAAACCACCAGATTCAG-3'*
for ciona expression	*5'-CTCGAGGTCAAATCAATCAATGCGG-3' (w/ Xho1)* *5'-TCTAGACCAAAGACAATTAAACCACC-3' (w/ Xba1)*
-5.0sox17	*5'-TGACGACCCCCAGGTCAACC-3'* *5'-CTCAAACACGCACCGGGC-3'*
BMP4	*5'-GGGGATGAAAGTAGCATCCTG-3'* *5'-TTCCACTTTGCTTCCCAAACTGG-3'*
HOXB6	*5'-AATGGAGAACTGGCTGGTG-3* *5'-ATCCTTTAGAGCAGCAACCC-3'*
GATA4	*5'-CCTTCCTATACACGCTTGC-3'* *5'-TCCTTCCCAAGCCCTGAGG-3'*

target region amplification	
drl	*5'-AATCACAAAGTCTGGGTTAACAGAG-3'* *5'-CGACCAGGATGGGCACCAC-3' (EGFP region)*
mixl1	*5'-GGCTGCTGTTATCAGTCTC-3'* *5'-GGGATGGTGGTTGGGTCG-3'*
mezzo	*5'-CGCTGTGGCGACCTCTGA-3'* *5'-CTTGCTCTCTGGAAGCAGCG-3'*

coding sequence	
foxh1	*5'-CACCATGACAAAGCACTGGGGG-3'* *5'-TTAAAGAGAATATTTGCAAAGG-3'*
mixl1	*5'-CACCATGGCAGTCGTGCACGGAAACC-3'* *5'-CTAAATGAAGCCATTAAAACC-3'*
spr2	*5'-CACCATGGCTGCATTAACCTTATCC-3'* *5'-TTACTCTCTCTTGATGTTGCTC-3'*
osr1	*5'-CACCATGGGTAGTAAGACGCTCC -3'* *5'-TTACTTTATCTTGGCTGGC-3'*
mespaa	*5'-CACCATGGACGCCTCCACG-3'* *5'-TCAGCACTGTCCATGGAA-3'*
scl/tal1	*5`-CACCATGATGGAAAAACTGAAATCC-3`* *5`-TCACCGCTGGGCATTTC-3'*

***in situ* hybridization**

EGFP	*5'-ATGGTGAGCAAGGGCGAGGAGC-3'* *5'-TAATACGACTCACTATAGGG-3'*
drl	*5'-ATGAAGAATACAACAAAACCC-3'* *5'-TAATACGACTCACTATAGGTGAGAAGCTCTGGCCGC-3'*
hand2	*5'-CACCATGAGTTTAGTTGGAGGGTTTCC-3'* *5'-TAATACGACTCACTATAGGGTCATTGCTTCAGTTCCAATGCC-3'*
foxf1	*5'-ATACCGCGCTATCATACC-3'* *5'-GAATTCTAATACGACTCACTATAGGGAAAGCAGTCTTAACTCTGG-3'*
ved	*5'-CACCATGAAGGGTCAGTTCTCC-3'* *5'-TAATACGACTCACTATAGGGCACAATAAAGCCAGTCC-3'*
tmem88a	*5'-ATGAGTCTTCCACGAAACGG-3'* *5'-TAATACGACTCACTATAGGGATGAAGCACAGGCCC-3'*

References

Bassett, A. R., Tibbit, C., Ponting, C. P. and Liu, J.-L. (2013). Highly efficient targeted mutagenesis of Drosophila with the CRISPR/Cas9 system. *Cell Rep.* **4**, 220–8.

Bruce, A. E. E., Howley, C., Zhou, Y., Vickers, S. L., Silver, L. M., King, M. Lou and Ho, R. K. (2003). The maternally expressed zebrafish T-box gene eomesodermin regulates organizer formation. *Development* **130**, 5503–5517.

Burger, A., Lindsay, H., Felker, A., Hess, C., Anders, C., Chiavacci, E., Zaugg, J., Weber, L. M. L. M., Catena, R., Jinek, M., et al. (2016). Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes. *Development* **143**, 2025–37.

Davidson, B., Shi, W. and Levine, M. (2005). Uncoupling heart cell specification and migration in the simple chordate Ciona intestinalis. *Development* **132**,.

Felker, A. and Mosimann, C. (2016). Contemporary zebrafish transgenesis with Tol2 and application for Cre/lox recombination experiments. *Methods Cell Biol.* **135**, 219–44.

Felker, A., Nieuwenhuize, S., Dolbois, A., Blazkova, K., Hess, C., Low, L. W. L. W. L., Burger, S., Samson, N., Carney, T. J. T. J., Bartunek, P., et al. (2016). In Vivo Performance and Properties of Tamoxifen Metabolites for CreERT2 Control. *PLoS One* **11**, e0152989.

Fuentes, M., Benito, E., Bertrand, S., Paris, M., Mignardot, A., Godoy, L., Jimenez-Delgado, S., Oliveri, D., Candiani, S., Hirsinger, E., et al. (2007). Insights into spawning behavior and development of the european amphioxus (Branchiostoma lanceolatum). *J. Exp. Zool. Part B Mol. Dev. Evol.* **308B**, 484–493.

Gays, D., Hess, C., Camporeale, A., Ala, U., Provero, P., Mosimann, C. and Santoro, M. M. (2017). An exclusive cellular and molecular network governs intestinal smooth muscle cell differentiation in vertebrates. *Development* **144**, dev.133926.

Gritsman, K., Zhang, J., Cheng, S., Heckscher, E., Talbot, W. S. and Schier, A. F. (1999). The EGF-CFC Protein One-Eyed Pinhead Is Essential for Nodal Signaling. **97**, 121–132.

Hild, M., Dick, A., Rauch, G. J., Meier, A., Bouwmeester, T., Haffter, P. and Hammerschmidt, M. (1999). The smad5 mutation somitabun blocks Bmp2b signaling during early dorsoventral patterning of the zebrafish embryo. *Development* **126**, 2149–59.

Jin, H., Xu, J., Qian, F., Du, L., Tan, C. Y., Lin, Z., Peng, J. and Wen, Z. (2006). The 5' zebrafish scl promoter targets transcription to the brain, spinal cord, and hematopoietic and endothelial progenitors. *Dev Dyn* **235**, 60–67.

Kozmikova, I. and Kozmik, Z. (2015). Gene regulation in amphioxus: An insight from transgenic studies in amphioxus and vertebrates. *Mar. Genomics* **24**, 159–166.

Krens, S. F. G., Möllmert, S. and Heisenberg, C.-P. (2011). Enveloping cell-layer differentiation at the surface of zebrafish germ-layer tissue explants. *Proc. Natl. Acad. Sci. U. S. A.* **108**, E9–10; author reply E11.

Kuraku, S., Takio, Y., Sugahara, F., Takechi, M. and Kuratani, S. (2010). Evolution of oropharyngeal patterning mechanisms involving Dlx and endothelins in vertebrates. *Dev. Biol.* **341**, 315–23.

Kwan, K. M., Fujimoto, E., Grabher, C., Mangum, B. D., Hardy, M. E., Campbell, D. S., Parant, J. M., Yost, H. J., Kanki, J. P. and Chien, C. Bin (2007). The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev Dyn* **236**, 3088–3099.

Lindsay, H., Burger, A., Felker, A., Hess, C., Zaugg, J., Chiavacci, E., Anders, C., Jinek, M., Mosimann, C. and Robinson, M. D. (2015). *CrispRVariants: precisely charting the mutation spectrum in genome engineering experiments*. Cold Spring Harbor Labs Journals.

Mosimann, C., Kaufman, C. K. C. K., Li, P., Pugach, E. K. E. K., Tamplin, O. J. O. J. and Zon, L. I. L. I. (2011). Ubiquitous transgene expression and Cre-based recombination driven by the ubiquitin promoter in zebrafish. *Development* **138**, 169–177.

Mosimann, C., Panáková, D., Werdich, A. A. A., Musso, G., Burger, A., Lawson, K. L. K. L., Carr, L. A. L. A., Nevis, K. R. K. R., Sabeh, M. K. K., Zhou, Y., et al. (2015). Chamber identity programs drive early functional partitioning of the heart. *Nat. Commun.* **6**, 8146.

Mullins, M. C., Hammerschmidt, M., Kane, D. A., Odenthal, J., Brand, M., van Eeden, F. J., Furutani-Seiki, M., Granato, M., Haffter, P., Heisenberg, C. P., et al. (1996). Genes establishing dorsoventral pattern formation in the zebrafish embryo: the ventral specifying genes. *Development* **123**, 81–93.

New, D. A. T. (1955). A New Technique for the Cultivation of the Chick Embryo in vitro. *Development* **3**,.

Parker, H. J., Sauka-Spengler, T., Bronner, M. and Elgar, G. (2014a). A reporter assay in lamprey embryos reveals both functional conservation and elaboration of vertebrate enhancers. *PLoS One* **9**, e85492.

Parker, H. J., Bronner, M. E. and Krumlauf, R. (2014b). A Hox regulatory network of hindbrain segmentation is conserved to the base of vertebrates. *Nature* **advance on**,.

Picker, A., Scholpp, S., Böhli, H., Takeda, H. and Brand, M. (2002). A novel positive transcriptional feedback loop in midbrain-hindbrain boundary development is revealed through analysis of the zebrafish pax2.1 promoter in transgenic lines. *Development* **129**, 3227–39.

Racioppi, C., Kamal, A. K., Razy-Krajka, F., Gambardella, G., Zanetti, L., di Bernardo, D., Sanges, R., Christiaen, L. A. and Ristoratore, F. (2014). Fibroblast growth factor signalling controls nervous system patterning and pigment cell formation in Ciona intestinalis. *Nat. Commun.* **5**, 4830.

Sakaguchi, T., Kuroiwa, A. and Takeda, H. (2001). A novel sox gene, 226D7, acts downstream of Nodal signaling to specify endoderm precursors in zebrafish. *Mech. Dev.* **107**, 25–38.

Sánchez-Iranzo, H., Galardi-Castilla, M., Minguillón, C., Sanz-Morejón, A., González-Rosa, J. M., Felker, A., Ernst, A., Guzmán-Martínez, G., Mosimann, C. and Mercader, N. (2018). Tbx5a lineage tracing shows cardiomyocyte plasticity during zebrafish heart regeneration. *Nat. Commun.* **9**, 428.

Sauka-Spengler, T., Meulemans, D., Jones, M. and Bronner-Fraser, M. (2007). Ancient Evolutionary Origin of the Neural Crest Gene Regulatory Network. *Dev. Cell* **13**, 405–420.

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**, 676–82.

Schmid, B., Shah, G., Scherf, N., Weber, M., Thierbach, K., Campos, C. P., Roeder, I., Aanstad, P. and Huisken, J. (2013). High-speed panoramic light-sheet microscopy reveals global endodermal cell dynamics. *Nat Commun* **4**, 2207.

Tamplin, O. J., Durand, E. M., Carr, L. A., Childs, S. J., Hagedorn, E. J., Li, P., Yzaguirre, A. D., Speck, N. A. and Zon, L. I. (2015). Hematopoietic stem cell arrival triggers dynamic remodeling of the perivascular niche. *Cell* **160**, 241–52.

Thisse, C. and Thisse, B. (2008). High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat. Protoc.* **3**, 59–69.

Westerfield, M. (2007). *The Zebrafish Book: a guide for the laboratory use of zebrafish (Danio rerio)*. 5th ed. Eugene: University of Oregon Press.

Yin, C., Kikuchi, K., Hochgreb, T., Poss, K. D. and Stainier, D. Y. R. (2010). Hand2 regulates extracellular matrix remodeling essential for gut-looping morphogenesis in zebrafish. *Dev. Cell* **18**, 973–984.

Zhu, H., Traver, D., Davidson, A. J., Dibiase, A., Thisse, C., Thisse, B., Nimer, S. and Zon, L. I. (2005). Regulation of the lmo2 promoter during hematopoietic and vascular development in zebrafish. *Dev Biol* **281**, 256–269.

Supplementary Figures

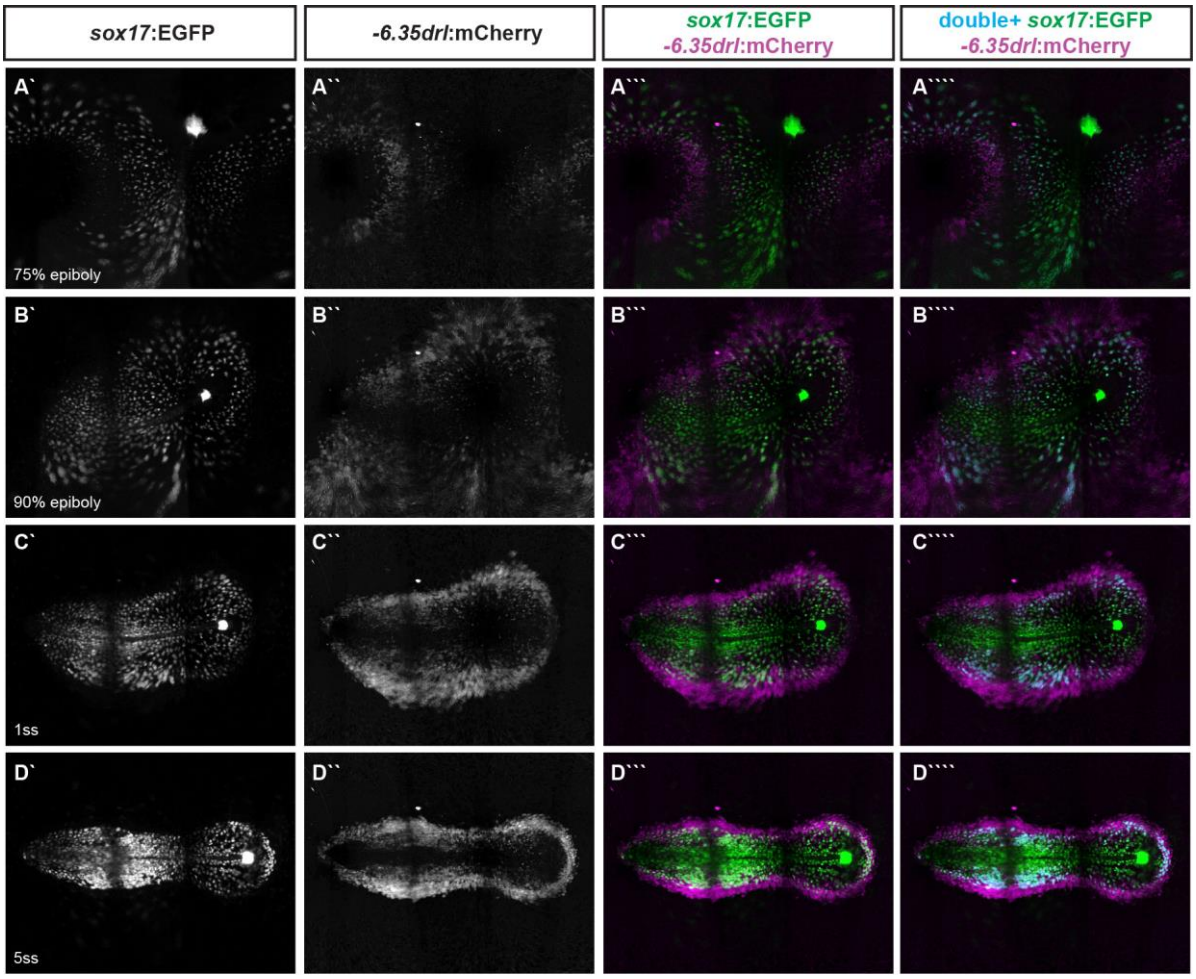

Fig. S1

***-6.35drl* reporter expression overlaps with *sox17*-marked endoderm during gastrulation.** (**A-D**) Supporting Mercator projections of the SPIM time-lapse imaging of *-6.35drl:mCherry* (magenta) combined with *sox17:EGFP* to visualize the *drl*-expressing emerging cell population versus endoderm from 70% epiboly until 5 ss (see also Movies S3 and S4). (**A`-D`**,**A``-D``**) Single and (**A```-D```**) merged channels are shown. Double-positive cells are labelled in blue in (**A``''-D``''**). Of note, the blue labeling in the 5 ss embryo in panel (**D``''**) represents not *per se* double-positive cells, but the overlap between the endodermal and *drl*-positive cell layer.

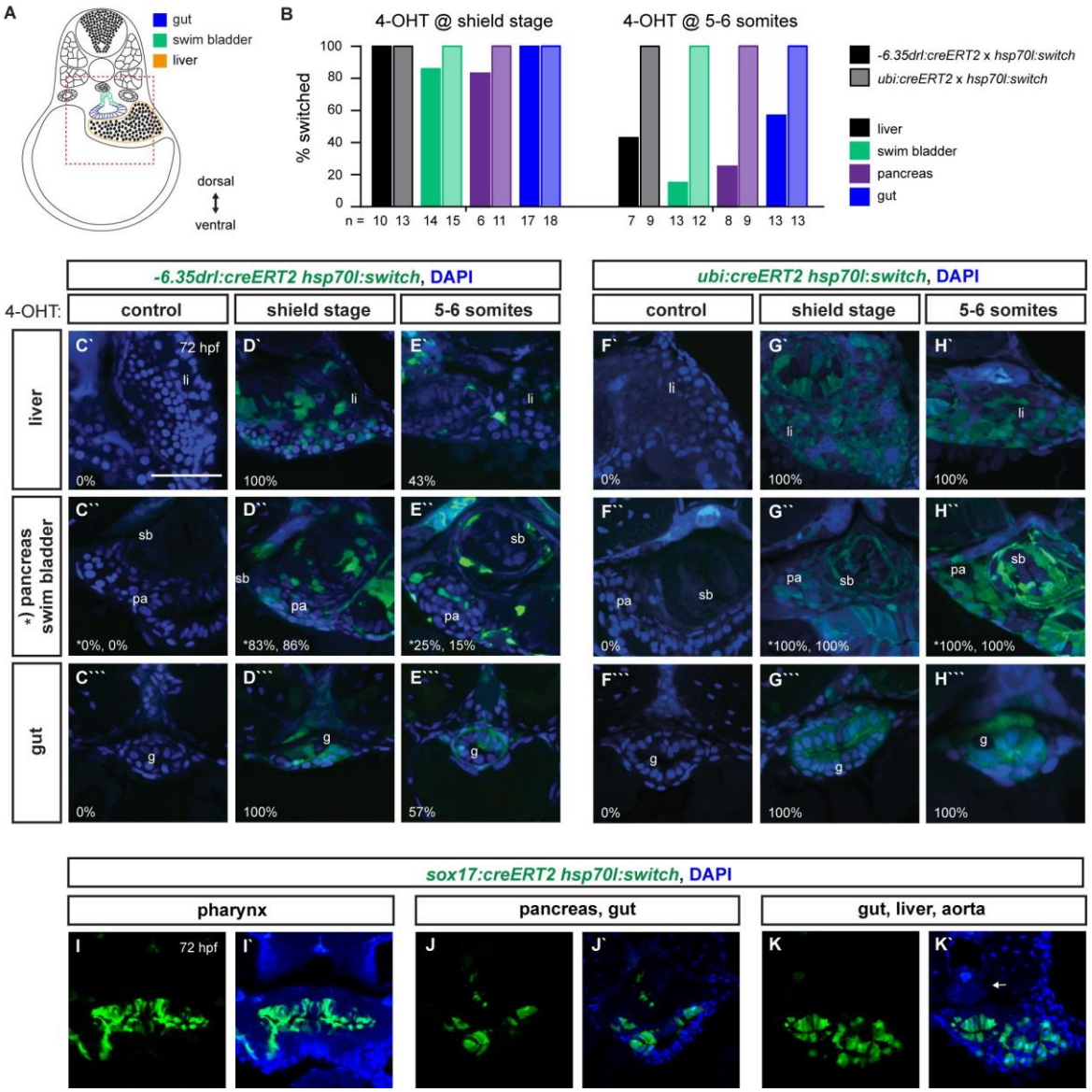

Fig. S2

Genetic lineage tracing time course shows *-6.35drl*-mediated endoderm tracing during gastrulation. *sox17*-based lineage tracing exclusively labels endoderm-derived structures.

(**A**) Schematic representation of a transverse section indicating endoderm-derived organs. (**B-H**) Whole-body transverse sections of *-6.35drl:creERT2*- (**C-E**) or *ubi:creERT2*-lineage traced cells (**F-H**) at different 4-OHT induction time points (shield and 5-6 ss) to trigger *hsp70l:switch* recombination (green). (**B**) Quantification of observed lineage contribution at individual time points; 4-OHT induction at shield stage traces, besides of LPM-derived tissue, a high percentage of endoderm-derived tissue, as depicted for liver (li), swim bladder (sb), pancreas (pa), and gut (g), while 4-OHT induction at 5-6 ss traces shows a strong reduction in percentage of endoderm-derived tissue. No tissue bias in lineage reporter labeling could be detected for the *hsp70l:switch* line, as shown using the ubiquitous *ubi:creERT2* (**F-H**). (**I-K**) Confocal transverse sections of *sox17:creERT2* lineage-traced endoderm cells after 4-OHT induction at shield stage and heat shock and fixation at 3 dpf. Lineage labeling is confined to endodermal lineages and absent from mesodermal and specifically LPM lineages, as depicted for the pharynx, pancreas, gut, and liver, in contrast to LPM-derived tissue such as the aorta (arrow in **K`**). Nuclei were counterstained with DAPI (blue). Numbers indicate the percentage of embryos showing lineage-labeling in depicted organs. Scale bar 50 µm.

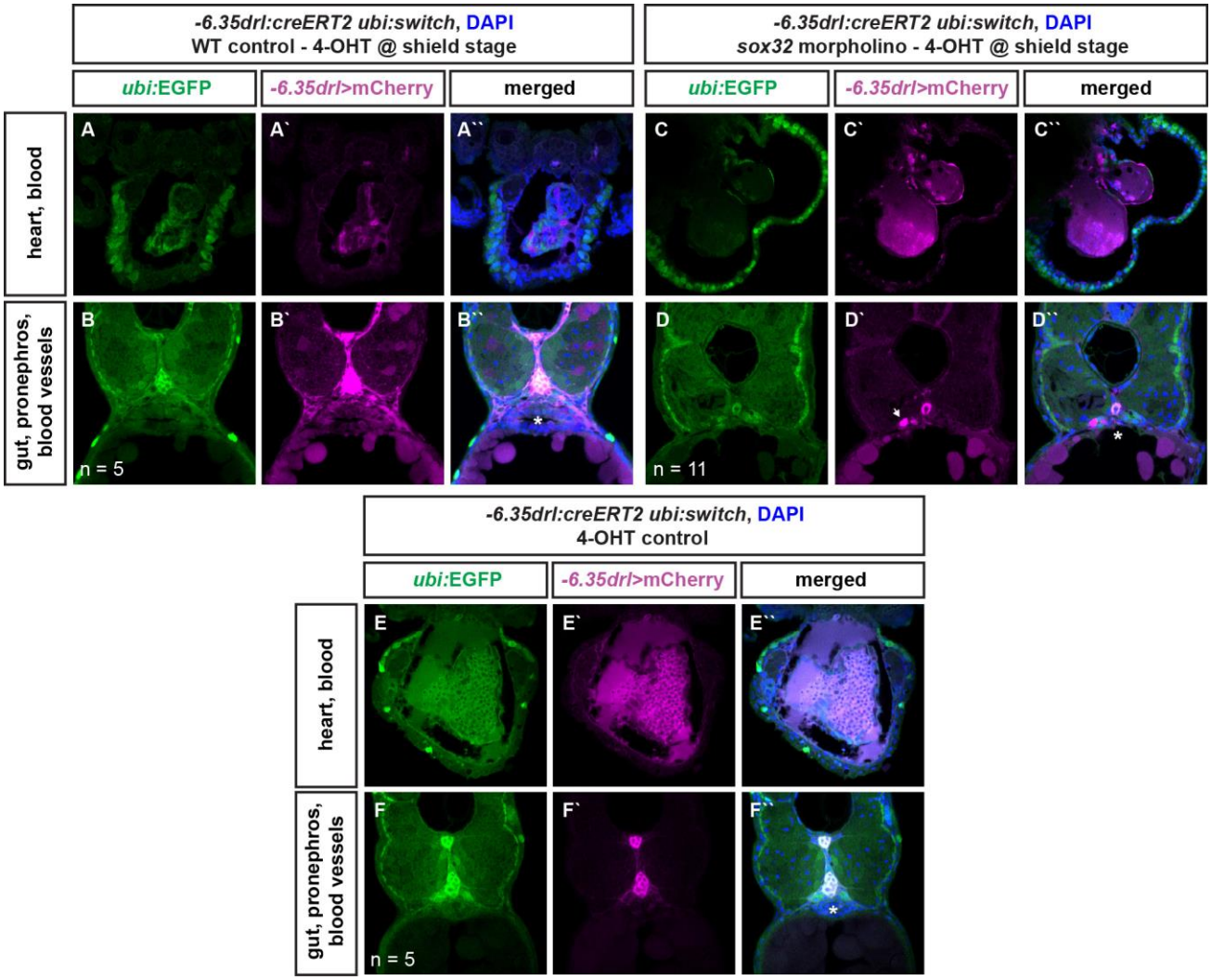

Fig. S3

LPM lineages and organs still arise in embryos devoid of endoderm. Confocal transverse sections of *-6.35drl:creERT2* x *ubi:switch* embryos fixed at 3 dpf. (**A-B**) Embryos were induced with 4-OHT at shield stage, bringing *mCherry* under control of the *ubi* promoter in cells with active CreERT2 at shield stage (magenta). LPM-derived organs, as depicted for heart, blood, pectoral fins, pronephros, and endothelial cells, and endoderm-derived lineages, including swim bladder, gut epithelium, and liver, are lineage-traced. (**C-D**) In *sox32* morphants that are devoid of endoderm, 4-OHT induction at shield stage traced LPM-derived organs, as depicted for heart, blood, pectoral fins, pronephros (arrow), and endothelial cells. (**E-F**) Control without 4-OHT admission reveals absence of any background activity of the used reporter (note autofluoresence of blood), confirming absence of leakiness of the *creERT* and *loxP* lines used. Asterisks indicate endoderm-derived gut epithelium (**B``**, **F``**) or absence thereof (**D``**). Nuclei counterstained with DAPI (blue).

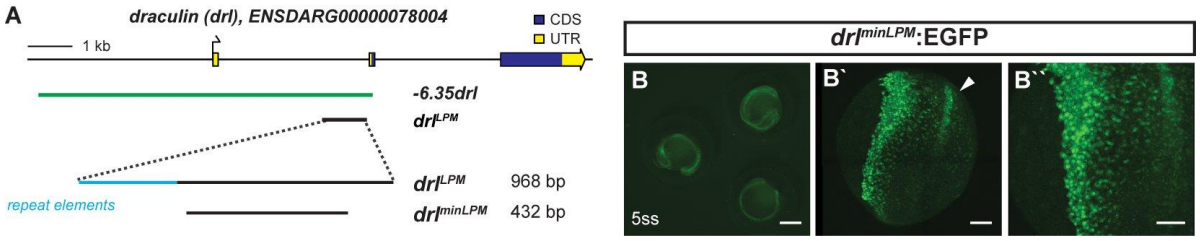

Fig. S4

Trimming of *drl*^{LPM} to the minimal 432bp pan-LPM *drl*^{minLPM} enhancer region. (**A**) Schematic representation of the *-6.35drl* locus including the PCR-based trimming approach to map the smallest LPM specific regulatory enhancer region; repetitive elements highlighted in blue. (**B**) Expression of *drl*^{minLPM}*:EGFP* at 5 ss; arrow indicates axial mesoderm expression. Scale bar (**B**) 400 μm (**B`**) and 80 μm (**B``**).

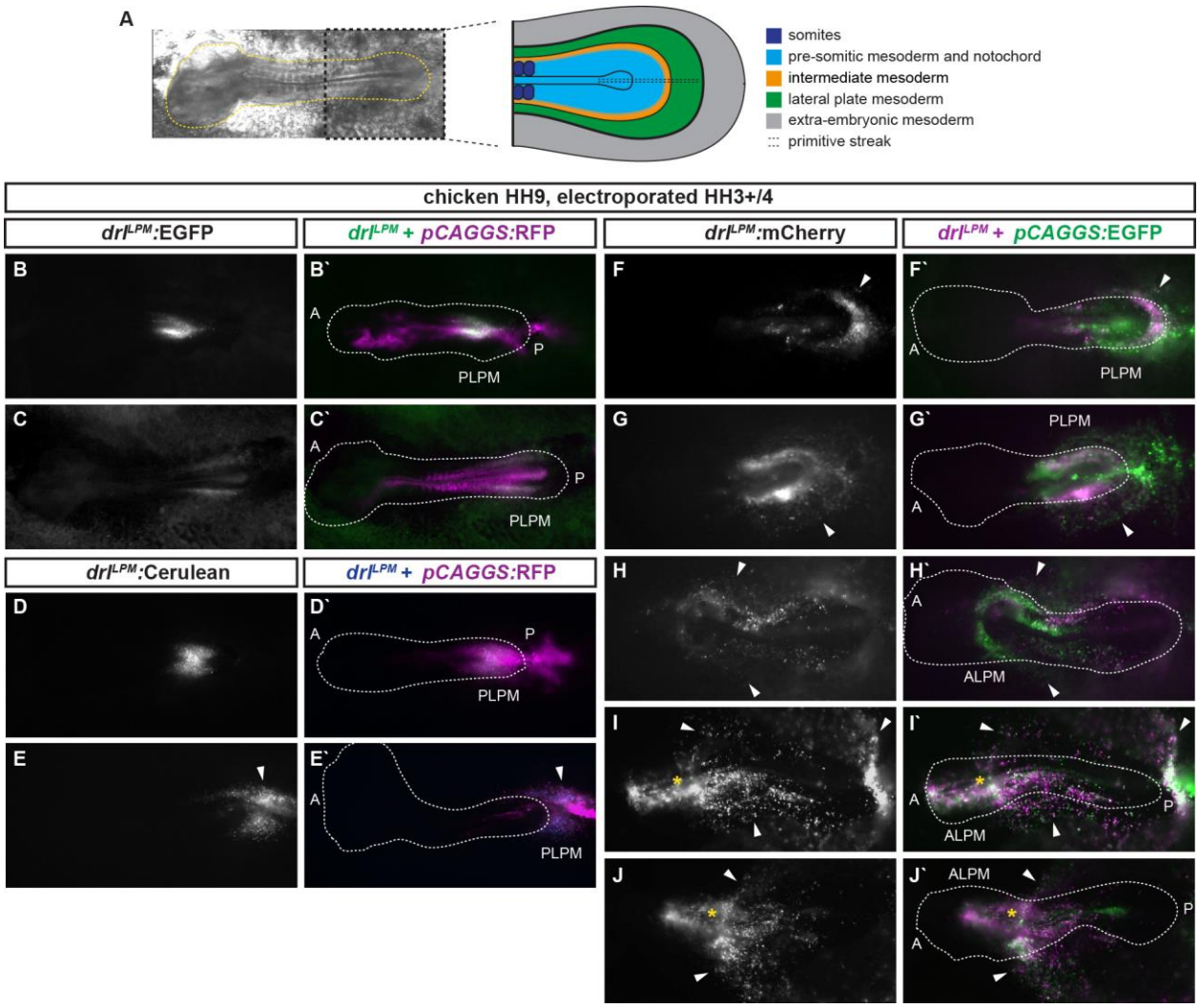

Fig. S5

Examples of electroporated chicken embryos with *drl*^{LPM} reporters driving specific expression in the chicken LPM. (**A**) Brightfield picture of an HH9-staged chicken embryo with the zoomed-in region represented in a schematic, showing the different territories in the posterior. (**B-J**) *Ex ovo* cultured chicken embryos electroporated with *drl*^{LPM} *cis*-regulatory reporters as (**B-C**) *drl*^{LPM}*:EGFP*, (**D-E**) *drl*^{LPM}*:Cerulean*, and (**F-J**) *drl*^{LPM}*:mCherry* (magenta) together with control plasmid (**B-E**) *pCAGGS:RFP* (magenta) or (**F-J**) *pCAGGS:EGFP* at HH3+/HH4. Arrowheads indicate extra-embryonic endothelial/blood progenitors and asterisks the heart field. The dash lines indicate the outline of the chicken embryos, with anterior (a) located to the left.

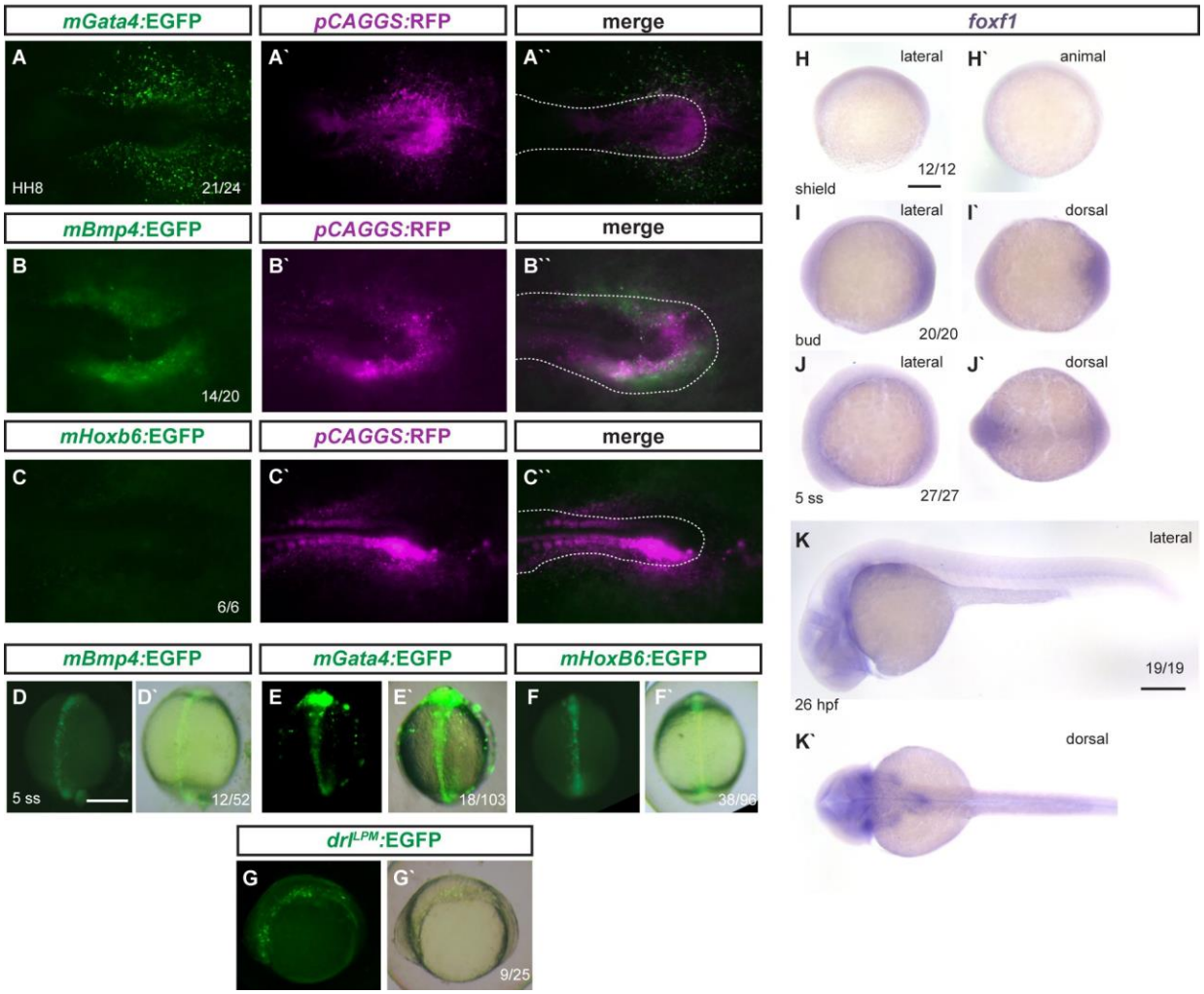

Fig. S6

No universal LPM-specific activity of previously characterized mouse LPM markers in both chicken and zebrafish. (**A**) The mouse *Gata4*, (**B**) *Bmp4*, and (**c**) *Hoxb6* LPM elements as *EGFP*-expressing reporters electroporated together with control plasmid *pCAGGS:RFP* (magenta) in *ex ovo* cultured chicken embryos at HH3+/H4. Widefield imaging reveals expression in (**A**) extraembryonic blood vessels and (**B**) the posterior arch, or (**C**) unspecifically/absent. The dash lines indicate the outline of the chicken embryos. (**D-G**) Mouse *Bmp4* (**D**), *Gata4* (**E**), and *Hoxb6* (**F**) LPM elements as *EGFP*-expressing Tol2-based reporter constructs injected in zebrafish embryos, showing confined axial mesoderm activity. (**G**) *drl*^{LPM}*:EGFP* as injection reference. (**H-K**) *In situ* hybridization (ISH) for *foxf1* from shield stage to 26 hpf in zebrafish does not mark any LPM-derived structures or lineages. Scale bars in (**D**,**H**,**K**) 300 μm.

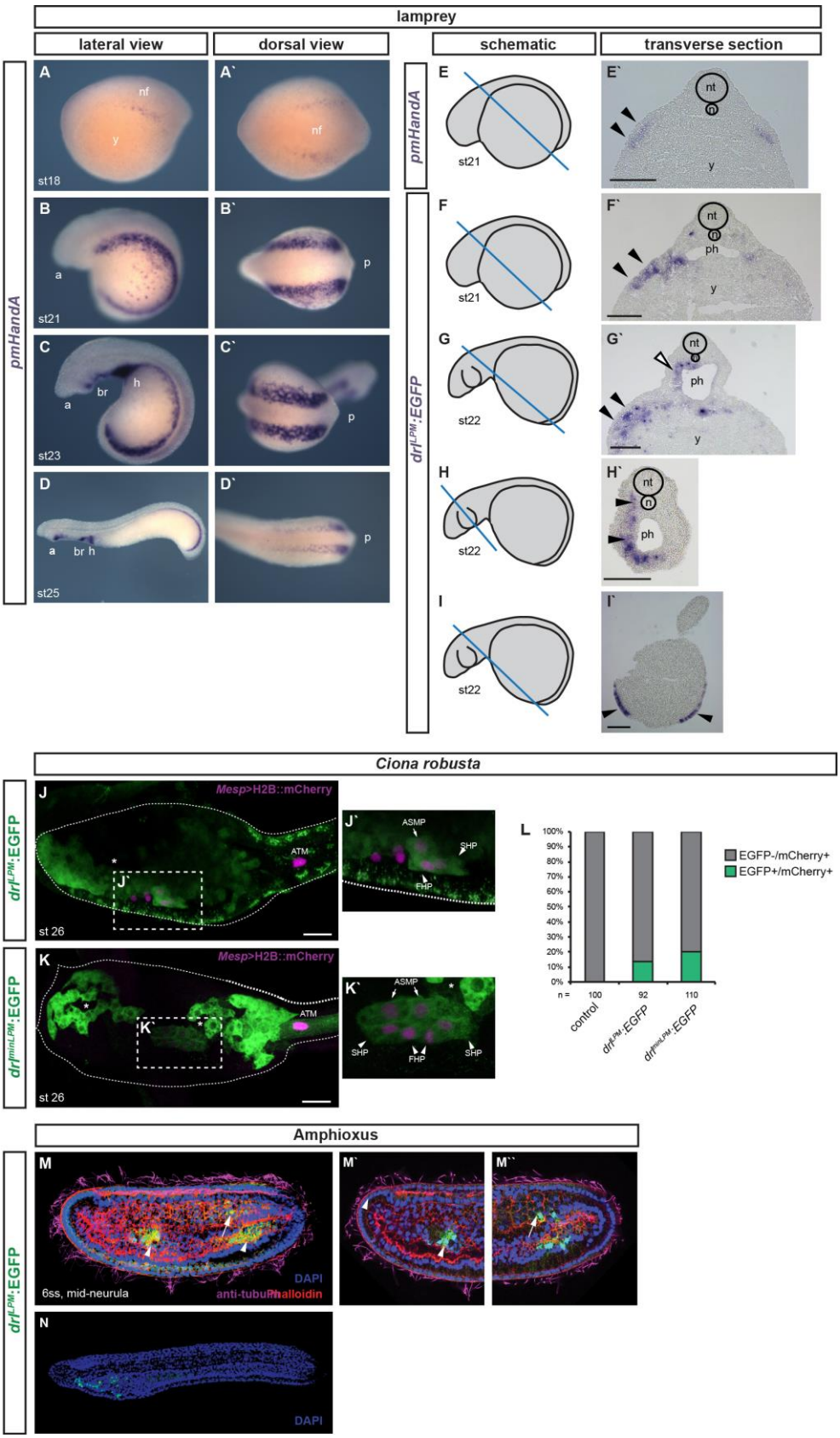
lamprey
lateral view
dorsal view
schematic
transverse section
pmHandA
drfL-PM:EGFP
st18
st21
st23
st25
st22
nf
y
a
br
h
p
nt
n
ph
Ciona robusta
Mesp>H2B::mCherry
ATM
ASMP
SHP
FHP
st 26
drfminLPM:EGFP
EGFP-/mCherry+
EGFP+/mCherry+
control
Amphioxus
6ss, mid-neurula
anti-tubulin
phalloidin
DAPI

Fig. S7

***drl^{LPM}* drives specific reporter expression in lamprey, *Ciona*, and amphioxus.** (**A-D**) Whole-mount ISH for *pmHandA* marks the LPM in lamprey embryos from neurula to hatching stages (st 18-25). (**A**) Expression is first detectable as bilateral stripes in the PLPM at st 18. (**B**,**C**) Later, the PLPM expression expands covering the yolk, and *pmHandA* is upregulated in the heart tube, branchial region, and caudal to the heart. Lateral and dorsal views are shown with the anterior (a) of the embryo to the left. (**E-I**) Transverse sections of fixed st 21-22 embryos. The schematics indicate the plane of sectioning in blue and sections are shown with dorsal to the top. (**E**) ISH for *pmHandA* at st 21 marking the LPM (black arrowheads). (**F-I**) ISH for *EGFP* in transverse sections of *drl^{LPM}:EGFP* transient transgenic lamprey embryos fixed at st 21-22, showing enhancer activity in the anterior mesendoderm subjacent to the ectoderm (**F`**,**G`**, black arrowheads), in the pharyngeal endoderm (**G`**, white arrowhead), in the pharyngeal mesoderm (**H`**, black arrowheads), and in some embryos in the ectoderm as commonly observed unspecific expression of reporter plasmids (**I`**, black arrowheads). (**J,K**) Immunostaining for EGFP in *Ciona* larvae embryo (st 26) expressing *drl* reporters (green) and *Mesp*>H2B::mCherry to track the B7.5 nuclei cell lineage (red). Expression of *drl^{LPM}*- (**J**) and *drl^{minLPM}*- (**K**) driven EGFP reporter in the larvae is stained in the mesenchymal lineage (white asterisks) and in the B7.5 cell progeny including ASM precursors (ASMP) (white arrows) and both cardiac first and second heart precursors (FHPs and SHPs) (white arrowheads), zoomed in (**J`,K`**). Dashed line indicates midline. Anterior is to the left. (**l**) Proportion of larvae embryos expressing both GFP and mCherry in the B7.5 lineage when co-electroporated *drl* reporter and *Mesp*>H2B::mCherry in comparison to the control. N: number of electroporated larval halves. (**M-M``**) Confocal Z-stack of amphioxus embryo at mid-neurula stage (6 ss), injected with *drl^{LPM}:EGFP*, showing specific reporter activity in the lateral endoderm (arrowhead), and ventral half of the somites (arrows). (**N**) At early larvae stage, the activity of *drl^{LPM}* reporter was observed in the endoderm of the pharynx (n = 15/30) and absent in the ventral mesoderm (n = 0/30). Embryos counterstained with Phalloidin (red), embryos/larvae with DAPI (blue), and anti-acetylated tubulin antibody (magenta). Lateral view shown as 3D-rendering (**M**) or Z-stack sagittal sections (**M`**,**M``**), anterior to the left and dorsal to the top. (**N**) At early larvae stage, the activity of *drl^{LPM}* reporter was observed in the endoderm of the pharynx (n = 15/30) and absent in the ventral mesoderm (n = 0/30). Branchial region (br), heart (h), neural folds (nf), neural tube (nt), notochord (n), pharynx (ph), and yolk (y) labeled in (**E-H**). Scale bars in (**E`-I`**) are 200 µm, and in (**K**) 25 µm.

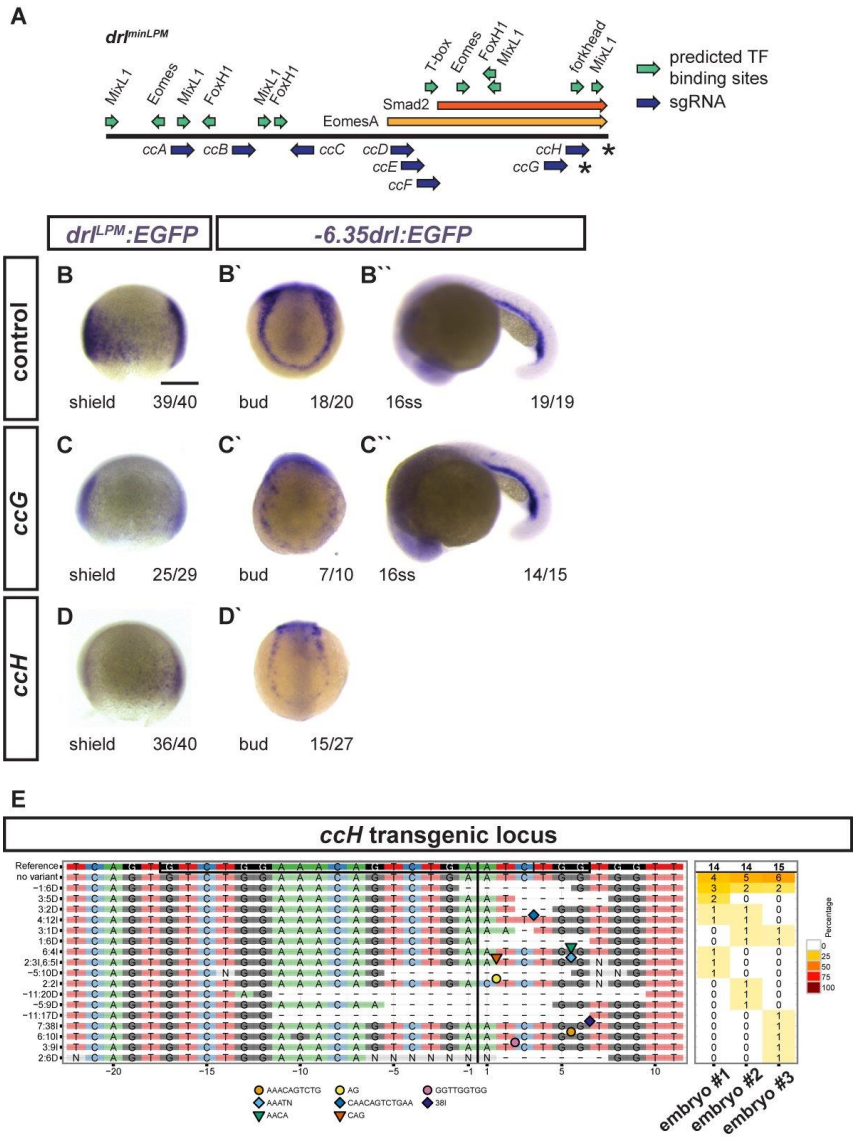

Fig. S8

Crispant analysis of the minimally required *drl*^{minLPM} region. (**A**) Schematic representation of the minimal *drl*^{minLPM} region with transcription factor binding site predictions (JASPAR), Eomes and Smad2 ChIP-seq annotations, and individual sgRNAs annotated. (**B-D**) ISH for *EGFP* in *drl*^{LPM}:*EGFP* embryos individually mutagenized using reconstituted Cas9 protein-sgRNA complexes with either sgRNA (**C**) *ccG* or (**D**) *ccH* (asterisks in **A**), showing diminished reporter expression at bud stage compared to uninjected controls. (**B``**,**C``**) Reporter expression remains functional at later stages, here shown for 16 ss. (**E**) Panel plot shows mutagenesis efficiency and allele spectrum in three representative *ccH* crispants, as created by CrispRVariants. The genomic reference sequence is shown on top, with the 20 bp *ccH* sgRNA followed by a 3 bp PAM indicated by the boxed regions. The Cas9 cleavage site is represented by a black vertical bar. Deletions are indicated by a '-', and inserted sequences by symbols. The right column of the plot shows the allele frequency, showing that some mutant alleles occurred in different crispants. Scale bar 250 μm.

A Individual and combinatorial misexpression of VP16 / eng mRNA construct

	VP16	*eng*
EomesA	++	nc
FoxH1	nc	nc
MixL1	nc	nc
Spr2	nc	nc
Osr1	nc	nc
Scl/Tal1	nc	
Mespaa	nc	
Etv5a	- -	++
Ntl	nc	
E/F/M	+++	nc

++ drl^{LPM} expression increase in native LPM territory

+++ drl^{LPM} ectopic expression

- - drl^{LPM} expression decrease

nc no change in drl^{LPM} expression

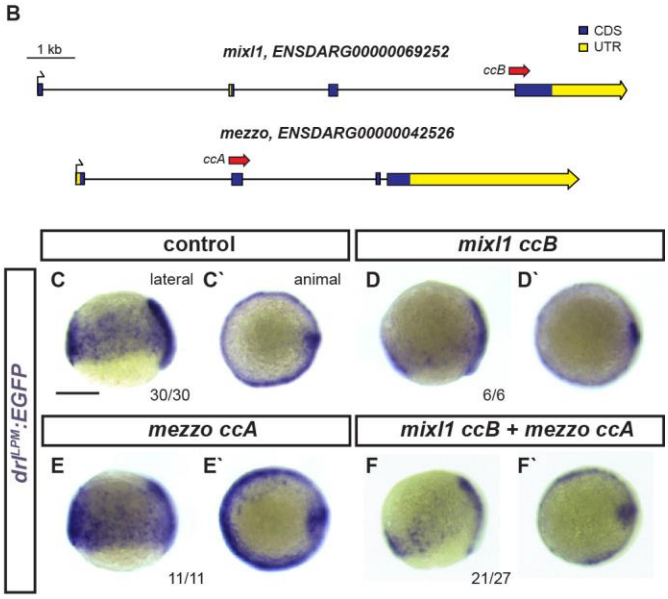

Fig. S9

mRNA-mediated misexpression screen uncovered Mixl1 as a third component acting on *drl^{LPM}*. (**A**) Transcription factors were fused N-terminally with the transactivation domain *VP16* or the *engrailed* (*eng*) repressor domain ORF, and tested for impact on *drl^{LPM}* reporter activity. (**B**) Schematic representation of the *mixl1* and *mezzo* loci with the individual sgRNAs annotated. (**C-F**) Cas9 RNP-mediated crispant analysis by mRNA *in situ* hybridization of *EGFP* expression in *drl^{LPM}*:*EGFP* embryos injected with (**D**) *mixl1 ccB*, (**E**) *mezzo ccA*, and (F) *mixl1 ccB* plus *mezzo ccA* sgRNA; the resulting mosaic *mixl1 ccB* crispants show diminished *drl^{LPM}* reporter expression at late gastrulation. Lateral and animal views are shown as indicated. Scale bar in (**C**) 250 μm.

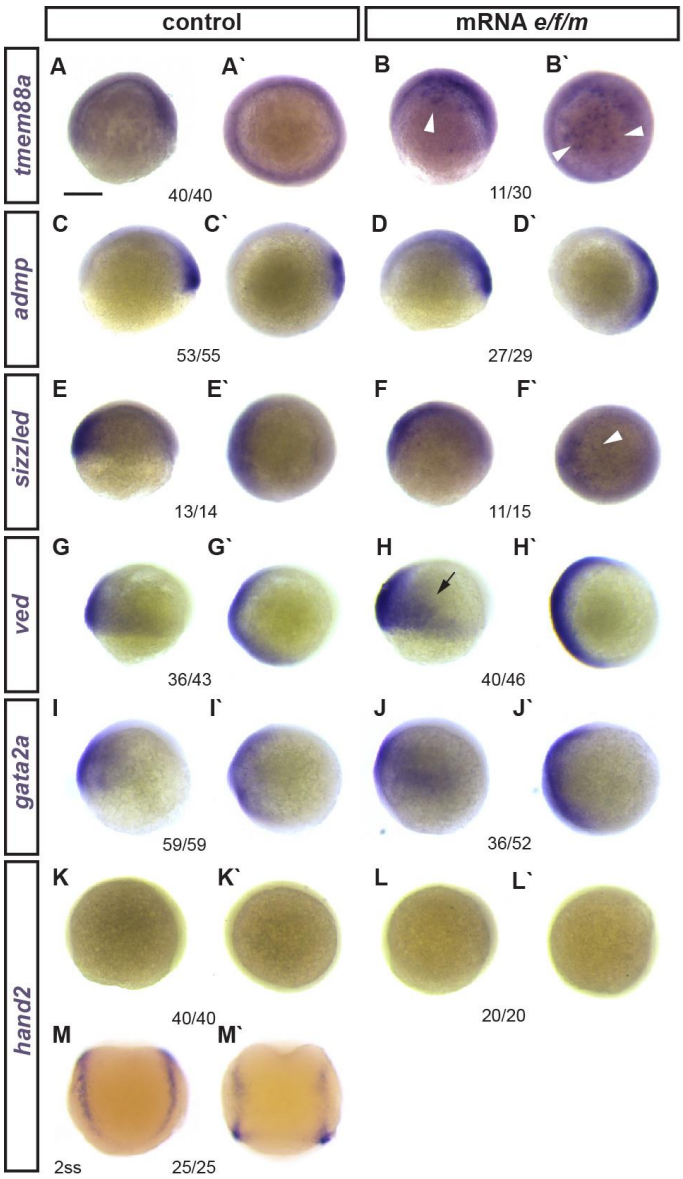

Fig. S10

***e/f/m* misexpression drives ectopic expression of the LPM gene *tmem88a* while the expression domains of BMP target genes and the post-gastrulation LPM maker *hand2* remain broadly intact.** ISH for endogenous expression of *tmem88a* (ectopic expression, indicated by arrowheads) (**A**,**B**), *admp* (**C**,**D**), *sizzled* (slight increased expression, indicated by arrowhead) (**E**,**F**), *ved* (slight increased expression, indicated by arrow) (**G**,**H**), *gata2a* (**I**,**j**), and *hand2* (**K-M**) in embryos injected with *e/f/m* mRNA compared to representative control embryos. Lateral and animal views are shown as indicated. Scale bar (**A**) 250 μm.

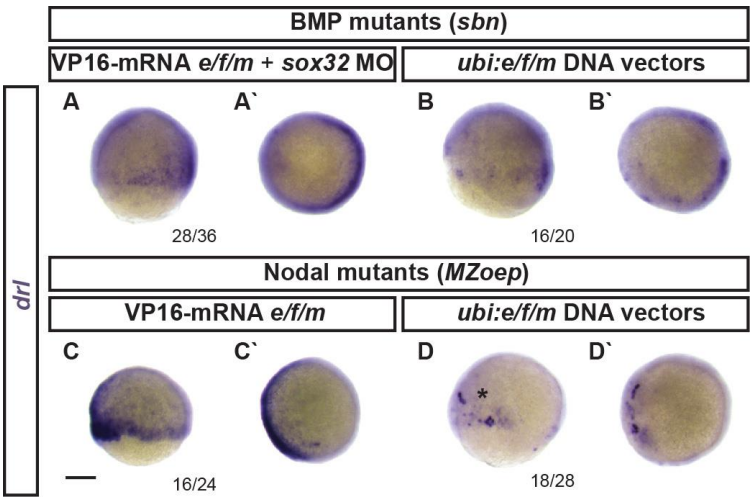

Fig. S11

***VP16-e/f/m* mRNA injections in Nodal and BMP mutants.** (**A**,**B**) mRNA ISH for endogenous *drl* expression in BMP-mutant (*sbn*) embryos injected with either (**A**) *VP16-e/f/m* mRNA in an endoderm-perturbed background (*sox32* morpholino) or (**B**) *ubi:e/f/m* DNA vectors, showing enhanced ventral *drl* expression. (**C-D**) Additional mRNA ISH for endogenous *drl* expression in Nodal-mutant (*MZoep*) embryos injected with (**C**) *VP16-e/f/m* mRNA or (**D**) *ubi:e/f/m* DNA vectors. Scale bar in (**A**) 250 μm.

Movie S1: 3D lightsheet imaging of the arising *-6.35drl* reporter-labeled LPM. SPIM time lapse imaging of zebrafish embryo transgenic for *-6.35drl*:EGFP (green) together with nuclear marker *actb2:*H2afva-mCherry (magenta) to visualize LPM emergence from 50% epiboly to 10 ss.

Movie S2: Panoramic lightsheet imaging of the arising *-6.35drl* reporter-labeled LPM. Mercator projection of SPIM time lapse imaging of zebrafish embryos transgenic for *-6.35drl:EGFP* to visualize LPM emergence from 50% epiboly to 10 ss. From somitogenesis onwards, the embryo is oriented with anterior to the left and posterior to the right.

Movie S3: 3D lightsheet imaging showing overlap between *-6.35drl* reporter expression and *sox17*-marked endoderm during gastrulation. SPIM time lapse imaging of *-6.35drl:mCherry* (red) combined with *sox17:EGFP* (green) to visualize the *drl*-expressing emerging cell population versus endoderm from 50% epiboly until 16 ss.

Movie S4: Panoramic lightsheet imaging showing overlap between *-6.35drl* reporter expression and *sox17*-marked endoderm during gastrulation. Mercator projection of SPIM time lapse imaging of *-6.35drl:mCherry* (magenta) combined with *sox17:EGFP* to visualize the *drl*-expressing emerging cell population versus endoderm from 50% epiboly until 16 ss. From somitogenesis onwards, the embryo is oriented with anterior to the left and posterior to the right. Double-positive cells during gastrulation and early somitogenesis are shown in blue.