

ARTICLE

<https://doi.org/10.1038/s41467-019-11561-7>

OPEN

# A conserved regulatory program initiates lateral plate mesoderm emergence across chordates

Karin D. Prummel<sup>1,11</sup>, Christopher Hess<sup>1,11</sup>, Susan Nieuwenhuize<sup>1</sup>, Hugo J. Parker<sup>2,3</sup>, Katherine W. Rogers<sup>4</sup>, Iryna Kozmikova<sup>5</sup>, Claudia Racioppi<sup>6</sup>, Eline C. Brombacher<sup>1</sup>, Anna Czarkwiani<sup>7</sup>, Dunja Knapp<sup>7</sup>, Sibylle Burger<sup>1</sup>, Elena Chiavacci<sup>1</sup>, Gopi Shah<sup>8</sup>, Alexa Burger<sup>1</sup>, Jan Huisken<sup>8,9</sup>, Maximina H. Yun<sup>7,8</sup>, Lionel Christiaen<sup>6</sup>, Zbynek Kozmik<sup>5</sup>, Patrick Müller<sup>4</sup>, Marianne Bronner<sup>10</sup>, Robb Krumlauf<sup>2,3</sup> & Christian Mosimann<sup>1</sup>

Cardiovascular lineages develop together with kidney, smooth muscle, and limb connective tissue progenitors from the lateral plate mesoderm (LPM). How the LPM initially emerges and how its downstream fates are molecularly interconnected remain unknown. Here, we isolate a pan-LPM enhancer in the zebrafish-specific *draculin* (*drl*) gene that provides specific LPM reporter activity from early gastrulation. *In toto* live imaging and lineage tracing of *drl*-based reporters captures the dynamic LPM emergence as lineage-restricted mesendoderm field. The *drl* pan-LPM enhancer responds to the transcription factors EomesoderminA, FoxH1, and MixL1 that combined with Smad activity drive LPM emergence. We uncover specific activity of zebrafish-derived *drl* reporters in LPM-corresponding territories of several chordates including chicken, axolotl, lamprey, *Ciona*, and amphioxus, revealing a universal upstream LPM program. Altogether, our work provides a mechanistic framework for LPM emergence as defined progenitor field, possibly representing an ancient mesodermal cell state that predates the primordial vertebrate embryo.

<sup>1</sup>Institute of Molecular Life Sciences, University of Zurich, Zürich 8057, Switzerland. <sup>2</sup>Department of Anatomy and Cell Biology, Kansas University Medical Center, Kansas City, KS 66160, USA. <sup>3</sup>Stowers Institute for Medical Research, Kansas City, MO 64110, USA. <sup>4</sup>Friedrich Miescher Laboratory of the Max Planck Society, Tübingen 72076, Germany. <sup>5</sup>Institute of Molecular Genetics of the ASCR, Prague 142 20, Czech Republic. <sup>6</sup>Center for Developmental Genetics, Department of Biology, New York University, New York, NY 10003, USA. <sup>7</sup>TUD-CRTD Center for Regenerative Therapies Dresden, Dresden 01307, Germany. <sup>8</sup>Max Planck Institute of Molecular Cell Biology and Genetics, Dresden 01307, Germany. <sup>9</sup>Morggridge Institute for Research, Madison, WI 53715, USA. <sup>10</sup>Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA. <sup>11</sup>These authors contributed equally: Karin D. Prummel, Christopher Hess. Correspondence and requests for materials should be addressed to C.M. (email: [christian.mosimann@imls.uzh.ch](mailto:christian.mosimann@imls.uzh.ch))

**K**ey cell fates and organ systems in vertebrates emerge from multipotent progenitors within the embryonic mesoderm. Following gastrulation, the vertebrate mesoderm has been classically described to partition into axial, paraxial, and ventrolateral domains<sup>1</sup>. The latter, referred to as lateral plate mesoderm (LPM), is composed of highly motile cells and is mainly defined by its position adjacent to the somite-forming paraxial mesoderm. Transplantation and lineage tracing experiments in several species have established that the LPM contains progenitors of the circulatory system, smooth muscles, the kidneys (in amniotes often demarcated as intermediate mesoderm), and the limb connective tissue anlagen<sup>2–4</sup>. During segmentation, the LPM principally segregates into the anterior LPM (ALPM) and posterior LPM (PLPM), which further divides into dorsal and ventral domains (somatopleure and splanchnopleure, respectively). 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<sup>2,3,5</sup>, albeit not always with an evolutionarily conserved function<sup>6</sup>. It remains incompletely understood how the LPM arises from an initial mesodermal population that goes on to form distinct endodermal and mesodermal progenitors. This is partly due to the lack of tools and markers to track LPM emergence genetically during development. Further, whether the LPM initially emerges as morphogenetic field in a molecularly coherent unit or as a loosely connected assembly of progenitor cells remains unclear.

Assessing the evolutionary context by which the LPM emerged as a developmental entity also remains challenging, in particular in extant jawless vertebrates such as lamprey or chordate models that do not form the full spectrum of LPM derivatives. Ancestral gene-regulatory repertoires that control higher-order structures in vertebrates previously have been indicated for somatic muscle in lamprey<sup>7</sup> or for the putative equivalents of cardiac and hematopoietic progenitors in amphioxus<sup>8</sup>. Anterior-to-posterior expression domains of key LPM transcription factors including *Tbx1/10* and *Hand* are conserved in lampreys and amphioxus<sup>9</sup>. Furthermore, the tunicate *Ciona* forms cardiac lineages that display genetic regulatory circuits homologous to the cardiac LPM progenitors found in vertebrates<sup>10</sup>. These observations suggest the existence of an ancient regulatory program that delineated prospective LPM progenitors in a common chordate ancestor, dating back to the Cambrian explosion 520–540 million years ago.

Several mammalian *cis*-regulatory elements with broad LPM activity have been reported; these include an upstream enhancer of mouse and human *HoxB6*<sup>11</sup>, an upstream enhancer of mouse *Gata4*<sup>12</sup>, and a downstream enhancer of mouse *Bmp4*<sup>13</sup>. In line with a ventral LPM origin, the *Gata4* LPM enhancer responds to Smads downstream of BMP signaling<sup>12</sup>. Nonetheless, the activities driven by these enhancer elements in mice confine to the PLPM and are seemingly not pan-LPM readouts. In zebrafish, the ventrally and marginally emerging LPM forms during somitogenesis into a patchwork of bilateral gene expression domains, including of the conserved LPM genes *hand2*, *pax2.1*, *scl*, *lmo2*, *etv2*, and *tbx5*<sup>5</sup>. In contrast, transgenic reporters based on the 6.35 kb *cis*-regulatory region of the zebrafish-specific gene *draculin* (*drl*) selectively label the entire LPM from its emergence during gastrulation through initial differentiation<sup>14</sup>. *Cre/lox*-mediated genetic lineage analysis has established that early *drl* reporter expression in zebrafish labels the LPM progenitors forming cardiovascular, blood, kidney, intestinal smooth muscles (iSMCs), and pectoral fin mesenchyme fates<sup>14–16</sup>. While *drl* as putative multimer zinc-finger gene has no obvious ortholog in other vertebrates<sup>14,17,18</sup>, these observations suggest that the 6.35 kb *drl* region harbors *cis*-regulatory elements active throughout the prospective LPM starting from gastrulation, raising the

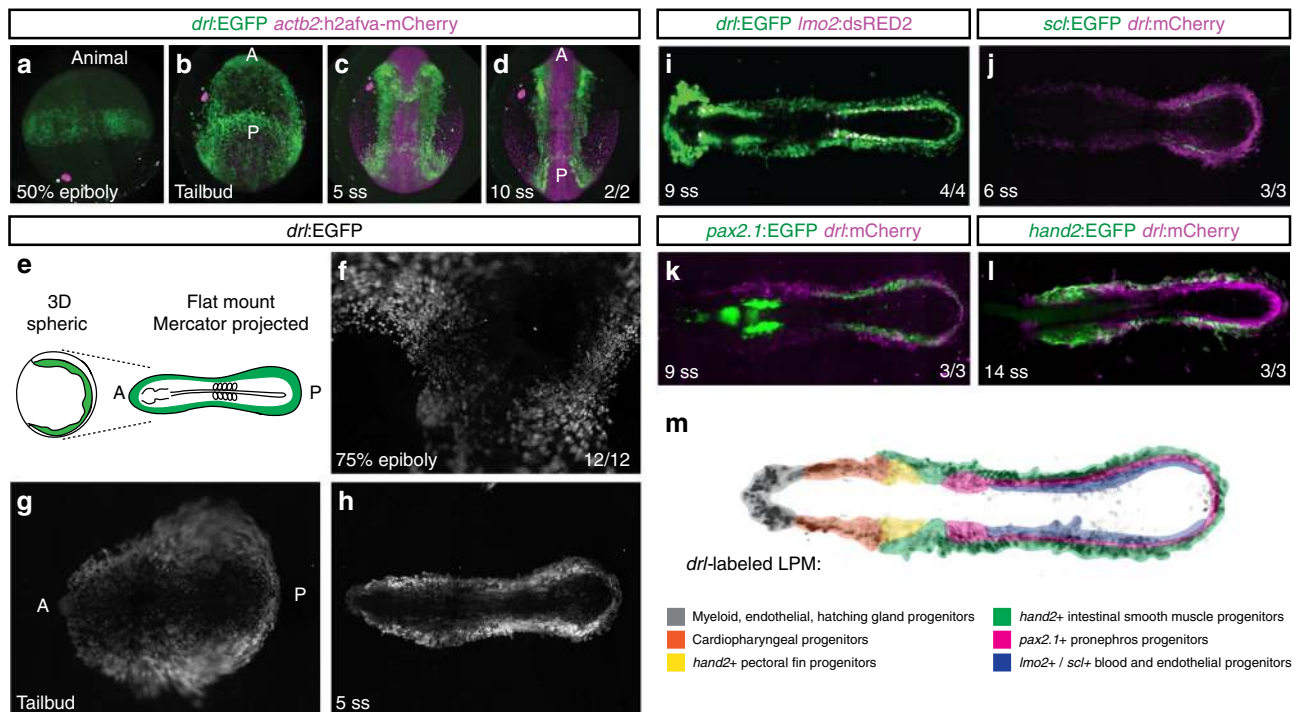
possibility that these regulatory elements read out a pan-LPM program.

Here, we dissect the 6.35 kb *drl cis*-regulatory elements and uncover an intronic enhancer, +2.0*drl*, that in zebrafish is necessary and sufficient for driving LPM-specific expression in all presumptive LPM progenitors from gastrulation to early somitogenesis. Panoramic SPIM and *Cre/lox*-mediated genetic lineage tracing of *drl* reporters demonstrate that the zebrafish LPM forms from a restricted mesoderm territory during gastrulation. As upstream regulatory program read out by the +2.0*drl* pan-LPM enhancer, we identify the combination of mesoderm transcription factors *EomesA*, *FoxH1*, and *MixL1* as sufficient to drive pan-LPM activity. In cross-species assays, we observe specific activity of the zebrafish +2.0*drl* pan-LPM enhancer in LPM-corresponding territories in chicken, axolotl, lamprey, *Ciona*, and amphioxus embryos. These results demonstrate that the zebrafish +2.0*drl* enhancer reads out a universal LPM progenitor program that is conserved across chordates, defining a core transcription factor code for LPM formation. Our data provide a developmental framework for charting the earliest emergence of LPM progenitors across chordates.

## Results

**The LPM emerges as a dedicated mesoderm population.** To resolve the dynamics of LPM emergence *in toto*, we performed time course experiments using single-plane illumination microscopy (SPIM) (Fig. 1a–d) and panoramic projections (Fig. 1e–h) of reporter-transgenic zebrafish embryos based on the full-length 6.35 kb *drl cis*-regulatory region. *drl:EGFP*-expressing LPM precursors became detectable by early gastrula stages (50% epiboly) and continuously condensed along the embryo margin through the end of gastrulation (tailbud stage) (Fig. 1a, b, f, g; Supplementary Movies 1,2). From tailbud stage onward, *drl:EGFP*-marked LPM formed a continuous band of cells with condensing anterior and posterior segments (Fig. 1c, d, g, h). We confirmed that this EGFP-positive cell band encompasses the bilateral stripes of several established LPM sub-domain markers by comparing a series of overlapping expression domains from distinct reporter lines (Fig. 1i–l). First, *lmo2:dsRED2* labels embryonic hematopoietic and vascular tissues, and its expression overlaps with medial *drl:EGFP*-expressing cells in the ALPM and PLPM (Fig. 1i). *scl:EGFP* also co-expressed with *drl:mCherry* in the most medial PLPM domain and in a small ALPM population (Fig. 1j). We find that the *pax2.1:EGFP*-expressing PLPM-derived pronephric epithelial precursors also express *drl:mCherry* (Fig. 1k). Moreover, *hand2:EGFP* expression, which demarcates the lateral-most PLPM domain plus parts of the ALPM-derived heart field and pectoral fin precursors, was also fully situated within the pan-LPM expression domain of *drl:mCherry* (Fig. 1l). Taken together, these data provide a continuous view of the emerging LPM stripes from gastrulation in zebrafish and document that the LPM emerges around the entire circumference of the zebrafish embryo (Fig. 1m).

We next sought to capture how the *drl*-expressing LPM emerges relative to the endoderm. Panoramic SPIM of the *sox17:EGFP*-positive endoderm reporter together with *drl:mCherry* revealed a population of double-positive cells from the onset of reporter detection through late gastrulation (Fig. 2a–d). After gastrulation, we detected a continuous band of *drl* reporter-positive cells around the developing embryo that was separated from the more medial endodermal *sox17* expression domain (Fig. 2d; Supplementary Movies 3,4). To confirm whether endoderm progenitors are also marked by the *drl* reporter during gastrulation, we performed *drl:creERT2*-based genetic lineage tracing with the ubiquitous *hsp70l:loxP-STOP-loxP-EGFP*



**Fig. 1** The LPM forms as a continuous field around the circumference of the developing zebrafish embryo. **a–d** Panoramic SPIM imaging of 50% epiboly to 10 ss embryos transgenic for *drl:EGFP* (green) and *actb2:h2afva-mCherry* (magenta); maximum-intensity-projected, lateral view (**a**), dorso-ventral views (anterior (A) to the top, posterior (P) bottom) ( $n = 2/2$ ) (**b–d**). **e–h** Representative panoramic SPIM imaged *drl:EGFP* zebrafish embryo shown as 2D Mercator projection ( $n = 12/12$ ). **e** Schematic of Mercator projection of spherical embryo, anterior to the left; **f–h** Mercator projections at 75% epiboly, tailbud, and 5 ss stages. **i–l** Single time point projections (anterior to the left) of representative double-transgenic embryos for *drl* reporters and (**i**) *lmo2* ( $n = 4/4$ ), (**j**) *scl* ( $n = 3/3$ ), (**k**) *pax2.1* ( $n = 3/3$ ), or (**l**) *hand2* ( $n = 3/3$ ) reporters co-expressed in dedicated LPM territories. **m** Summary schematic of the LPM fate territories partitioning during early somitogenesis in zebrafish

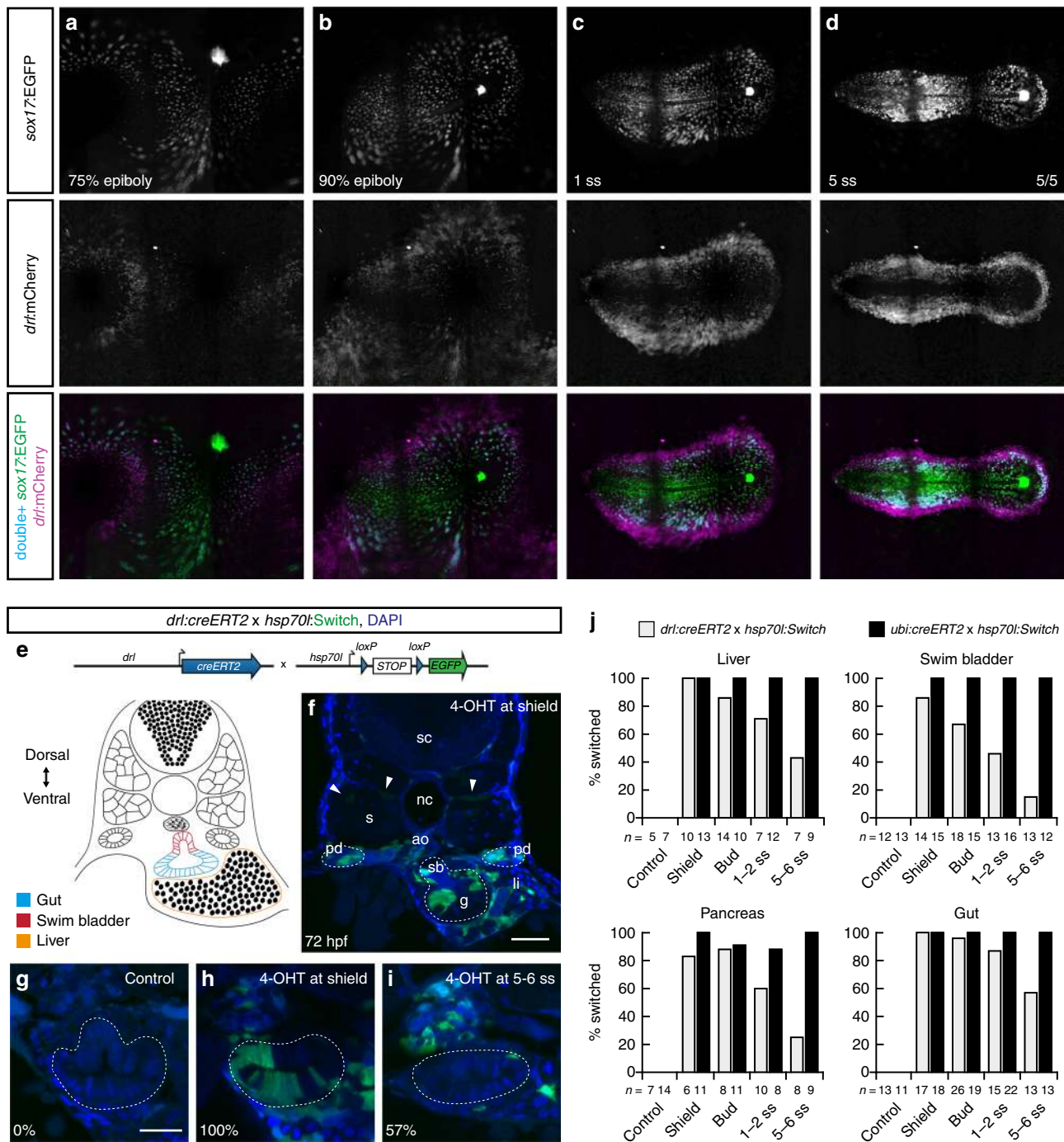
(*hsp70l:Switch*) and 4-OHT-based CreERT2 induction at discrete time points ranging from shield to 5–6 somite stages (ss) followed by analysis of labeling patterns at 72 hpf (Fig. 2e–j). 4-OHT induction at shield stage marked LPM lineages including blood, endothelium, kidney, and iSMCs as the only mesodermal fates (Fig. 2f–h)<sup>14,15</sup>, while lineage labeling also marked broad territories within endodermal organs, including pancreas, liver, and pharynx/gut epithelium (Fig. 2h; Supplementary Fig. 1a,b). 4-OHT induction at later time points gradually decreased endoderm labeling, with minimal to no endodermal lineage signals following 4-OHT induction at 5–6 ss (Fig. 2i, j, Supplementary Figs. 1b and 2b, c). In contrast, LPM structures remained robustly labeled as the exclusive mesoderm fate, consistent with previous work<sup>14–16</sup>. In addition, we observed that the spatio-temporal contribution of *drl* reporter-expressing progenitors to endoderm differs along the anterior-posterior axis. We divided the embryo into four non-overlapping regions along the anterior-to-posterior axis (region I–IV) (Supplementary Fig. 2a) and quantified the switching efficiency. The amount of lineage-labeled gut endothelium increased within individual embryos from the pharynx (region I) towards the caudal gut (region IV), independent of the stage of 4-OHT administration (Supplementary Fig. 2b, c). These results indicate that progenitors expressing the *drl* reporter with ongoing development become progressively restricted to an LPM fate from anterior to posterior, until by early somitogenesis *drl* reporter expression labels only LPM.

In contrast, *sox17:creERT2* exclusively marked endoderm lineages (Supplementary Fig. 1c), supporting that *sox17* expression demarcates zebrafish endoderm progenitors downstream of the key endoderm regulator *sox32*<sup>19</sup>. Supporting the faithfulness

of our lineage tracing experiments, labeling does not appear skewed by any lineage-bias of the *hsp70l:Switch* reporter, as illustrated using the ubiquitous *ubi:creERT2* (Supplementary Fig. 1d). Remarkably, embryos that fail to form endoderm upon *sox32* perturbation still generated *drl*-traced LPM that partitioned into recognizable heart, blood, endothelium, and pronephros (Supplementary Fig. 3).

Altogether, these data establish that during gastrulation the *drl*-marked LPM gradually refines from a ventral-marginal mesoderm territory to a bilateral LPM domain as the sole mesodermal fate along the entire anterior-posterior axis of the embryo. The rare lineage labeling of somitic muscle by *drl:creERT2* (Fig. 2f)<sup>14,15</sup> further underlines that, in zebrafish, the paraxial mesoderm and the LPM develop as distinct mesoderm lineages with only minimal overlap.

**A pan-LPM enhancer in the zebrafish *drl* locus.** The zebrafish *drl* gene encodes a putative 13-mer zinc-finger protein<sup>20</sup> of unknown function and is as genomic locus seemingly zebrafish-specific. First and characteristic for multimeric zinc-finger genes<sup>21</sup>, the putative *drl* ORF lacks a clear ortholog in any other fishes or in vertebrates<sup>17,18</sup>. Second, *drl* is tandem-multiplied, with three *drl*-like genes featuring nearly identical ORFs and introns and a fourth more diverged putative zinc-finger gene, suggesting a recent multiplication event (Fig. 3a)<sup>14,17</sup>. Lastly, the *drl* locus lies in between *ap2b1* and *pex12* in the zebrafish genome, yet these two flanking genes are deeply conserved neighboring loci with no additional genes between them across vertebrates or in other fishes (Fig. 3a). Reminiscent of the also zebrafish-specific *crestin* locus that harbors a neural crest-specific

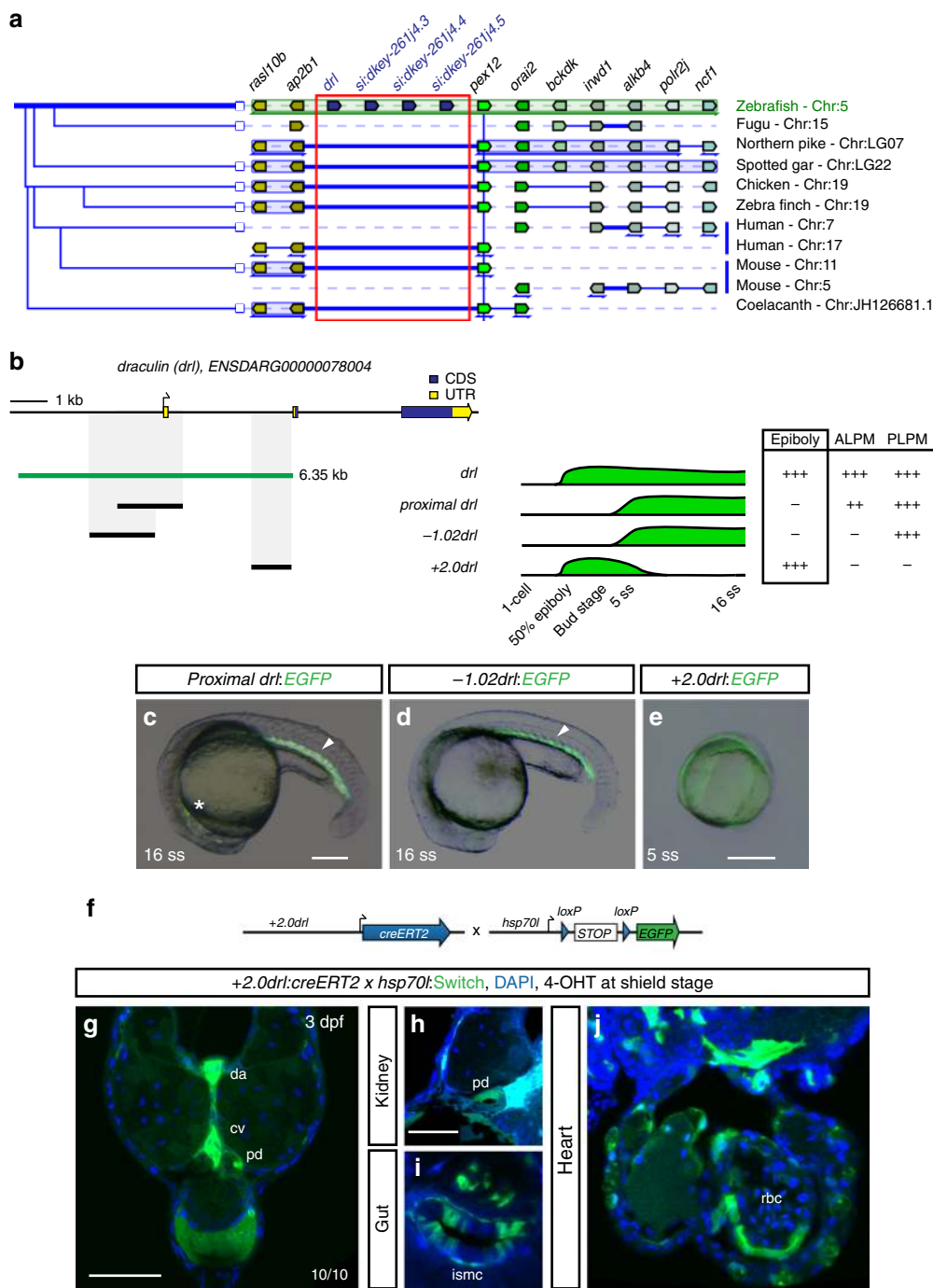


**Fig. 2** The early *drl* reporter-expressing cells comprise endoderm- and LPM-fated progenitor. **a-d** Single time point projections of *sox17:EGFP* marking endoderm progenitors (**a-d**) and *drl:mCherry* marking LPM progenitors from gastrulation until 5 ss; anterior to the left. Double-positive cells for EGFP and mCherry shown in blue (**a-d**) ( $n = 5/5$ ). **e** Schematic of *drl:creERT2* to *hsp70l:Switch* cross for genetic lineage tracing and schematic of transverse section (trunk region) with endoderm- and LPM-derived organs at 3 dpf. **f** Representative 72 hpf transverse section of *drl* lineage-traced embryo 4-OHT-induced at shield stage; arrowheads depict rare trunk muscle labeling. **g-i** Representative transverse sections of *drl* lineage tracing at 72 hpf, control (**g**) versus 4-OHT-induced at shield stage (**h**), and 5-6 ss (**i**); note gradual loss of endoderm labeling (intestinal cells, dashed region). Numbers indicate percentage of embryos with intestinal lineage labeling. **j** Quantification of endoderm lineage labeling in representative organs following 4-OHT induction at indicated time points, comparing *drl:creERT2* versus ubiquitous *ubi:creERT2* control as reference for the *hsp70l:Switch* lineage reporter. N-numbers are mentioned in the graph legends. Notochord (nc), somite (s), spinal cord (sc), pronephric duct (pd), liver (li), swim bladder (sb), gut (g), nuclei in blue (DAPI; **f-i**). Scale bar (**f**) 50  $\mu$ m and (**g-i**) 25  $\mu$ m

regulatory region<sup>22</sup> and despite its peculiar nature, the 6.35 kb *drl* regulatory region provides a means to read out an upstream LPM input.

To identify *cis*-regulatory element(s) in the zebrafish *drl* locus responsible for pan-LPM progenitor expression, we divided the

6.35 kb *drl* regulatory region into smaller fragments and assayed their activity using Tol2-based *EGFP* reporters in F0 zebrafish and in stable transgenics (Fig. 3b). We found that the promoter-proximal region surrounding exon 1 recapitulated *drl* reporter



**Fig. 3** The 6.35 kb *drl* cis-regulatory region contains an early pan-LPM enhancer. **a** Adapted Genomicus-based PhyloView representation of the zebrafish *drl* locus (top row) compared to other vertebrate species. The view is centered on *pex12* (light green, blue vertical line) as anchor, with its orthologous copies across species and corresponding neighboring loci shown by their relative positions. Genes with identical coloration are homologs/orthologs. Blue arrows below gene loci indicate switched local orientation. The phylogenetic relationship tree is depicted on the left. The red box marks the *drl* locus with its neighboring three *drl*-related genes (blue) in between *ap2b1* and *pex12* in zebrafish as unique feature. Note the absence of any genes between *ap2b1* and *pex12* orthologs across fishes and other vertebrates, and conservation of suggested ancestral synteny of *ap2b1* and *pex12*. **b** Schematic of the *drl* locus depicting the 6.35 kb cis-regulatory region (green), and smaller isolated candidate fragments *proximal drl* (region surrounding first exon), *-1.02drl* (upstream region only), and *+2.0drl* (distal first intron) with specific reporter activity. Time line and table indicates expression dynamics (50% epiboly to 16 ss) of stable transgenes for the individual regulatory elements and expression domains (pan-LPM early or somite-stage ALPM, PLPM) with absent expression (-) to strong expression (+ +). **c-e** Representative stable transgenic zebrafish embryos harboring EGFP reporters for *proximal drl*, *-1.02drl*, and *+2.0drl*; at 5 ss and 16 ss *proximal drl* and *-1.02drl* express in PLPM (arrowheads in **c, d**), *proximal drl* additionally in ALPM (asterisk in **c**); note pan-LPM activity of *+2.0drl* (**e**). **f** Schematic of *+2.0drl:creERT2* to *hsp70l:Switch* cross for genetic lineage tracing. **g-j** Transverse sections at 3 dpf of *+2.0drl:creERT2* lineage tracing after 4-OHT induction at shield stage results in specific labeling of LPM-derived tissue including red blood cells (rbc), dorsal aorta (da) and cardinal vein (cv) endothelium, pronephric duct (pd), and intestinal smooth muscle cells (ismc) ( $n = 10/10$ ). *+2.0drl:creERT2* also traces endoderm-derived tissue, shown for gut epithelium. Nuclei counterstained with DAPI (blue). Scale bars (**c, e, g**) 250  $\mu$ m and (**h**) 50  $\mu$ m

expression in ALPM and PLPM from 5–7 ss onwards (*proximal drl*), while the promoter region alone remained active in the posterior endothelial and blood precursors ( $-1.02dr1$ ) (Fig. 3b–d). In addition, we identified a small (968 bp) region in the first intron ( $+2.0dr1$ ) that initially recapitulated early *drl* reporter expression in zebrafish embryos before fading between 5–10 ss (Fig. 3b, e). Genetic lineage tracing with  $+2.0dr1:creERT2$  and *hsp70l:Switch* (Fig. 3f) specifically labeled LPM-derived organs including heart, blood, endothelium, kidney, pectoral fin mesenchyme, and iSMCs, and additionally marked endoderm lineages when induced with 4-OHT at shield stage (Fig. 3g–j). These results correlate well with our lineage tracing using full-length *drl:creERT2* (Fig. 2e–j)<sup>14–16</sup>. Further, deletion of elements within the  $+2.0dr1$  enhancer defined a minimal enhancer region of 432 bp ( $+2.4dr1$ ) that functioned as a pan-LPM enhancer, albeit with higher variability in stable transgenics (Supplementary Fig. 4). These regulatory analyses indicate that the entire *drl* expression pattern in zebrafish derives from distinct *cis*-regulatory elements that control *drl* expression in separable early mesendoderm/pan-LPM and later ALPM versus PLPM domains. The latter pattern is analogous to the hematopoietic lineages that arise during somitogenesis and that are commonly marked with *drl* mRNA ISH<sup>5,20</sup>. These data imply that the  $+2.0dr1$  enhancer contains the key regulatory modules that respond to and integrate an early LPM-defining input.

**Combined EomesA, FoxH1, and MixL1 can drive LPM formation.** We next investigated the upstream input that controls the zebrafish  $+2.0dr1$  pan-LPM enhancer. BMP and Activin/Nodal ligands of the TGF- $\beta$  superfamily trigger key pathways in early vertebrate axis determination and mesendoderm induction (Fig. 4a)<sup>23</sup>. During early gastrulation in zebrafish, BMP ligands are principally secreted from the ventral side, while Nodal ligands are expressed along the margin and the dorsal side<sup>23</sup>. In line with BMP dependence of the LPM, endogenous *drl* expression was virtually absent in embryos maternally mutant for dominant-negative Smad5 (*MZsbn*), which lack BMP activity (Fig. 4b, c)<sup>24</sup>. Similarly, treatment with the BMP inhibitor Dorsomorphin resulted in a pronounced decrease of endogenous *drl* expression (Fig. 4e, f). We also found decreased *drl* expression in embryos with perturbed Nodal signaling: (i) in maternal-zygotic mutant embryos lacking the key Nodal co-receptor Cryptol/Oep (*MZoep*) that cannot transmit Nodal signaling around the embryo margin<sup>25</sup>, and (ii) in embryos treated with the Nodal signaling inhibitor SB-505124 (Fig. 4d, g, h). These results indicate that the *drl* expression domain is sensitive to both BMP and Nodal input. Consistent with our lineage tracing results (Supplementary Fig. 3), embryos devoid of endoderm upon *sox32* knockdown still expressed endogenous *drl*, albeit with overall thinned-out expression and a marked decrease of dorsal *drl* activity (Fig. 4i).

We mined existing whole-embryo ChIP-seq data from zebrafish gastrulation stages<sup>26–28</sup> and identified candidates for transcription factors binding to the  $+2.0dr1$  enhancer. These include the T-box transcription factor EomesoderminA (EomesA), its interaction partner FoxH1, and BMP/Nodal-mediating Smads. Published evidence has uncovered that these factors participate in controlling mesendoderm genes<sup>29–32</sup> and affect *drl* expression during early somitogenesis<sup>29</sup> (Fig. 5a). We found that mRNA injection- or *ubi* promoter-driven expression of constitutive-active forms of EomesA or FoxH1 strongly augmented and prolonged  $+2.0dr1$  reporter and endogenous *drl* expression in their native LPM domain compared to controls (Fig. 5b–e, Supplementary Fig. 5a–d). Ubiquitous expression of wildtype *eomesA* or *foxh1* mRNA was sufficient to increase endogenous *drl* expression (Fig. 5f–h). Addition of a

constitutively-active Smad2 to EomesA and FoxH1 resulted in dorsal widening of the  $+2.0dr1$  reporter expression pattern (Supplementary Fig. 5e, f).

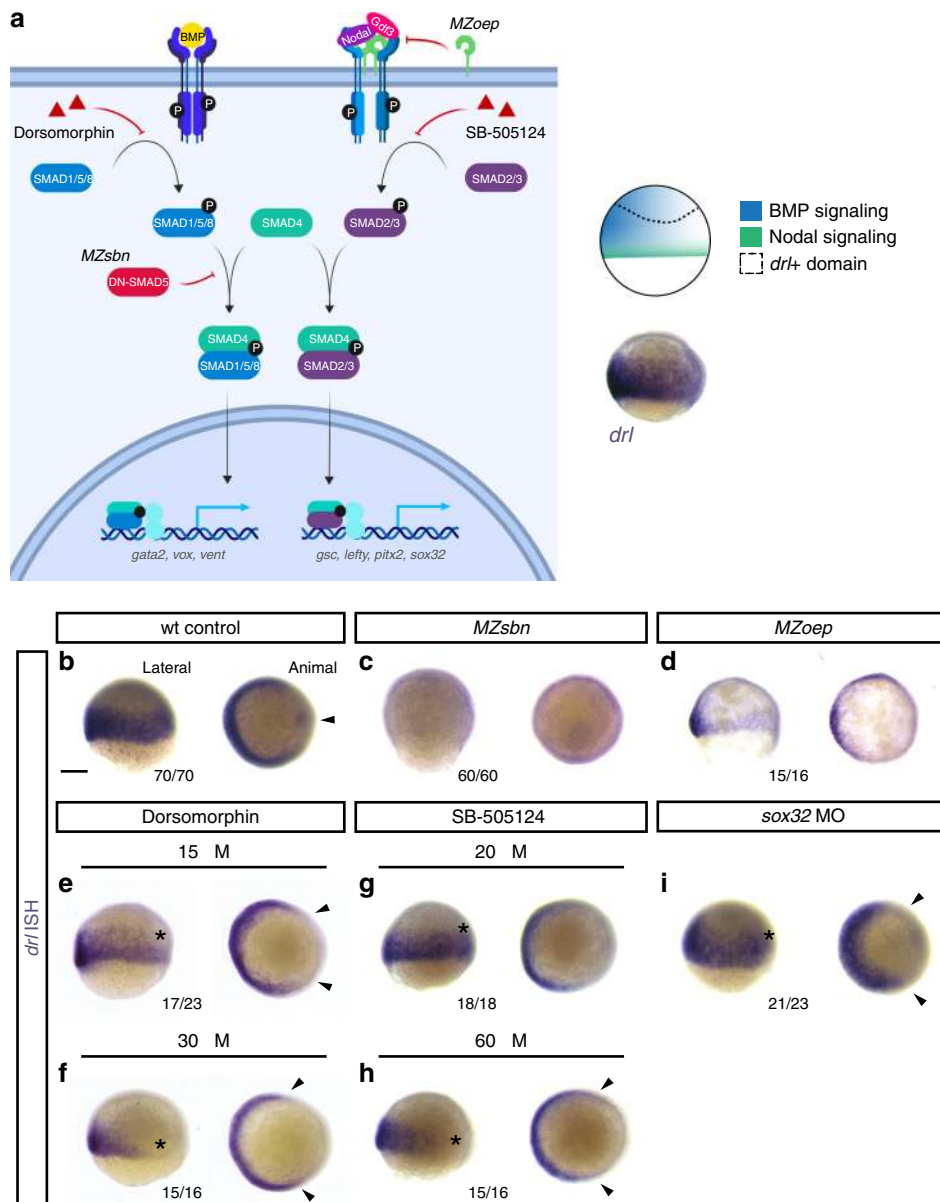
EomesA and FoxH1 are maternally contributed and ubiquitously distributed during gastrulation<sup>29,33</sup>, during which the LPM emerges in the BMP and Smad activity domain (Figs. 1, 4). We therefore hypothesized that at least one additional, ventrally and marginally expressed transcription factor might be required for LPM formation. In published ChIP-seq data<sup>26</sup>, we identified the homeodomain protein MixL1 as a third possible transcription factor that acts together with EomesA and FoxH1 in controlling  $+2.0dr1$  enhancer activity (Fig. 5a). MixL1 is a downstream target of BMP and Nodal signaling implicated in controlling endoderm and mesoderm fates<sup>34</sup> and it retains ventral expression in *MZoep* mutants<sup>29</sup>. Furthermore, MixL1 can form a complex with EomesA<sup>35</sup>. Reminiscent of *eomesA* or *foxh1* mRNA injections (Fig. 5g, h), microinjected *mixl1* mRNA also resulted in increased  $+2.0dr1$  reporter expression within in the native LPM domain (Fig. 5i).

Combining the triplet of wildtype mRNAs or Tol2-based DNA constructs encoding full-length EomesA, FoxH1, and MixL1 (shortened as *eff/m*) led to ubiquitous  $+2.0dr1$  reporter activation in embryos (Fig. 5j, Supplementary Fig. 5g). In *MZsbn*-mutant embryos without BMP signaling, *eff/m* misexpression induced *drl* expression dorsally (Supplementary Fig. 5h, i). These observations suggest that in *eff/m* overexpression conditions there is still a requirement for additional Smad activity, which in *MZsbn* embryos is only available in the dorsal Nodal-positive domain. Conversely, loss of Nodal signaling in *MZoep* mutants led to a ventral upregulation of *drl* expression upon mRNA-based *eff/m* overexpression (Supplementary Fig. 5k, l). Combining native *eff/m* in wildtype and *MZsbn* embryos devoid of endoderm following *sox32* knockdown also resulted in patchy, ubiquitous  $+2.0dr1$  reporter activation (Fig. 5k, Supplementary Fig. 5j). This suggests that most, if not all, of the  $+2.0dr1$  reporter-positive cells have an LPM identity. Mutating *mixl1* by CRISPR-Cas9 resulted in mosaic loss of  $+2.0dr1$  reporter activity in F0 crispants (Fig. 5l–n), while mutating the *mixl1* paralog *mezzo*<sup>36</sup> alone or together with *mixl1* did not influence  $+2.0dr1$  reporter activity (Fig. 5l, o, p). This indicates that MixL1 is the predominant Mix paralog acting on the  $+2.0dr1$  enhancer in zebrafish. Furthermore, CRISPR-Cas9-mediated mutagenesis of the  $+2.0dr1$  enhancer in the region of predicted FoxH1 and MixL1 sites in the context of the full-length *drl:EGFP* transgene resulted in specific perturbation of early LPM reporter expression, without affecting the later ALPM and PLPM patterns (Supplementary Fig. 6).

EomesA, FoxH1, and MixL1 misexpression also induced weak yet detectable expression of *tmem88a*, which is highly enriched in the early native LPM<sup>14</sup> (Supplementary Fig. 7a, b). In contrast, the expression domains of other early expressed mesodermal genes either showed a slight broadening of their native domains or appeared unaffected (Supplementary Fig. 7). This was illustrated further by the lack of changes in *hand2* expression, which normally initiates in the lateral-most LPM after gastrulation (Supplementary Fig. 7k–m). Together, these data suggest a regulatory model in zebrafish whereby the combination of EomesA, FoxH1, and MixL1 potentiates Smad-relayed BMP signals to demarcate a mesendoderm territory that becomes prospective LPM.

### The $+2.0dr1$ enhancer detects a LPM program across chordates.

We next explored whether the zebrafish-derived  $+2.0dr1$  enhancer could read out a putative pan-LPM program in diverse chordate species. First, we revisited several previously characterized enhancers with activity in the posterior LPM of mice:

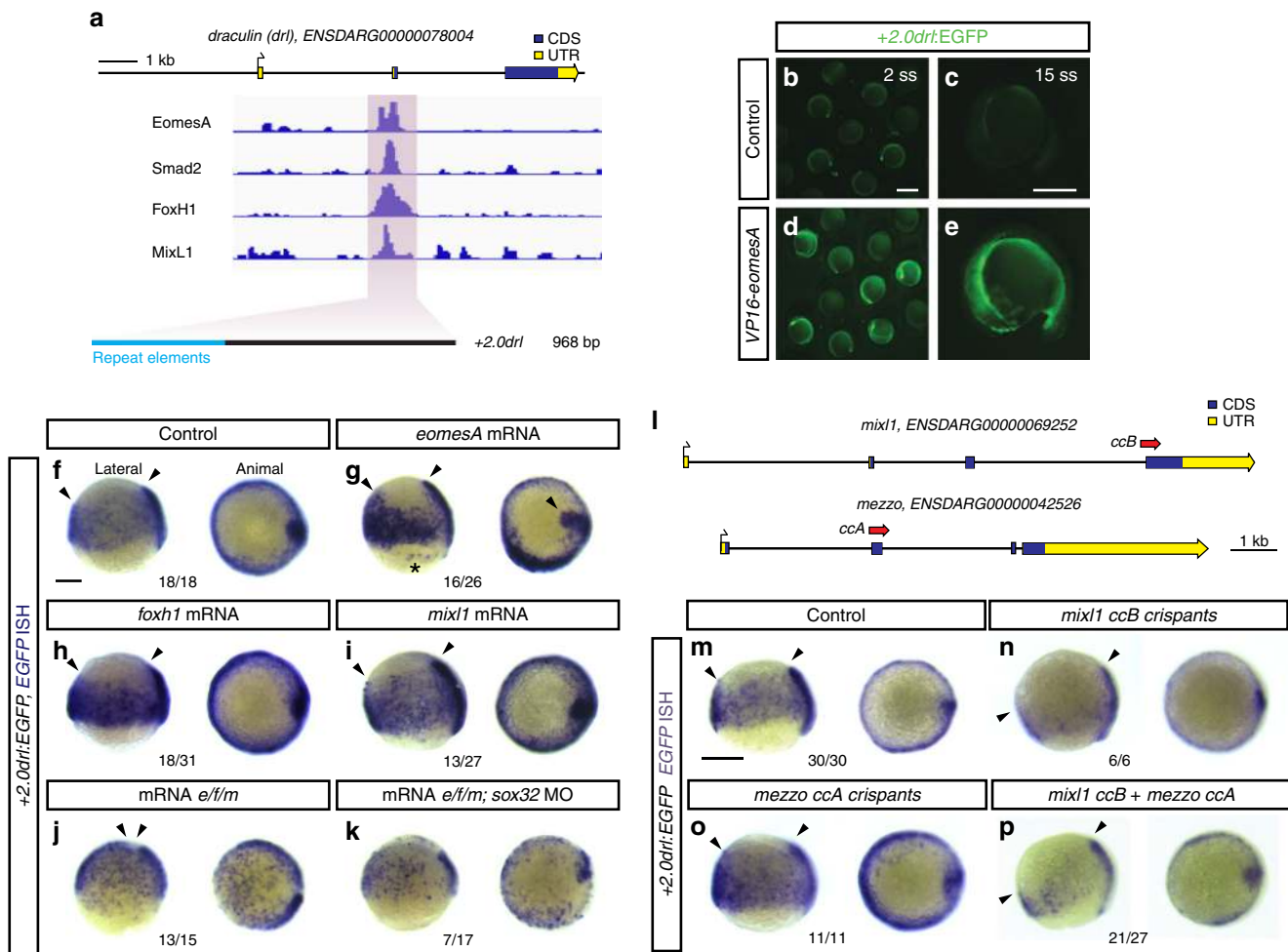


**Fig. 4** Early *drl*-expressing LPM progenitors respond to ventral BMP. **a** Schematic of the BMP and Nodal signaling pathways and the corresponding signaling territories in the gastrulation-stage zebrafish embryo, compared to endogenous *drl* expression marked by mRNA ISH. Pathway schematic generated with BioRender. **b–i** mRNA ISH for endogenous *drl* expression during early gastrulation (shield stage to 75% epiboly), lateral view on the left and animal view on the right. **b–d** Wildtype (wt) controls ( $n = 70/70$ ), BMP-perturbed (*MZsbn*, a dominant-negative *Smad5* allele,  $n = 60/60$ ) and Nodal-perturbed (*MZoep*, mutant for the required Nodal co-receptor *Oep*,  $n = 15/16$ ) showing that BMP perturbation abolishes endogenous *drl* expression. **e–h** Chemical inhibition of BMP via Dorsomorphin (15  $\mu\text{M}$  ( $n = 17/23$ ) and 30  $\mu\text{M}$  ( $n = 15/16$ ), **e, f**) decreases ventral *drl* expression, while chemical Nodal inhibition with SB-505124 (20  $\mu\text{M}$  ( $n = 18/18$ ) and 60  $\mu\text{M}$  ( $n = 15/16$ ), **g, h**) decreases dorsal *drl* expression, as indicated with the asterisks in the lateral views and the arrowheads in the animal views. **i** *sox32* morphants that fail to form endoderm also lose dorsal *drl* expression, as indicated with asterisk in the lateral view and with arrowheads in the animal view ( $n = 21/23$ ). Scale bar (**b**) 250  $\mu\text{m}$

*Gata4*<sup>12</sup>, *Bmp4*<sup>13</sup>, and *HoxB6*<sup>11</sup>. Reporter transgenes based on mouse *Gata4* and *Bmp4* showed restricted activity in the outward migrating endothelial/blood progenitors and in the PLPM when electroporated into the primitive streak of *ex-ovo*-cultured chicken embryos after the onset of gastrulation (HH3+/4) (Supplementary Fig. 8a–c). The mouse *HoxB6* LPM enhancer showed no specific activity in this assay (Supplementary Fig. 8d). In contrast, when microinjected into zebrafish embryos, reporters based on these three mouse enhancers all resulted in expression mainly in the notochord without specific LPM activity (Supplementary Fig. 8e–h). This indicates that while some of the previously isolated LPM enhancers from mice express faithfully in

the PLPM of chick embryos, their activity does not recapitulate an LPM pattern in zebrafish. These results suggest that these mammalian LPM enhancers may have specialized during amniote evolution.

Electroporation of the zebrafish +2.0*drl* reporter into the primitive streak of HH3+/4 chicken embryos resulted in reporter activity specifically in the forming LPM: depending on the exact stage and region of electroporation, we observed specific reporter activity in several LPM territories. Most frequently observed expression patterns included medial and posterior LPM domains (Fig. 6a–c, Supplementary Fig. 9a–f), and we noted ALPM reaching the head fold in individual embryos (Supplementary



**Fig. 5** EomesA, FoxH1, and MixL1 together can induce the +2.0drl enhancer. **a** ChIP-seq tracks for EomesA, FoxH1, MixL1, and Smad2 in the *drl* locus. See text for details. Bottom depicts the +2.0drl intronic enhancer and the smaller, minimally specific region +2.4drl after removal of repeat sequences. **b–e** Constitutively-active VP16-EomesA boosts +2.0drl:EGFP reporter expression in its native territory. Compared to injection controls (**b**, **c**), microinjection of VP16-eomesA mRNA in +2.0drl:EGFP reporter transgenics enhances and prolongs EGFP expression in the native reporter expression domain (**d**, **e**). **f–k** Gastrulation-stage (shield to 75% epiboly) zebrafish embryos, lateral view left and animal view right, probed with EGFP ISH to detect expression of +2.0drl:EGFP. Compared to controls ( $n = 18/18$ ) (**f**), embryos injected with mRNA encoding full-length *eomesA* ( $n = 16/26$ ), *foxh1* ( $n = 18/31$ ), or *mixl1* ( $n = 13/27$ ) show enhanced +2.0drl:EGFP reporter activity (**g–i**), as indicated by arrowheads. The asterisk in **g** points out ectopic expression in the enveloping layer cells. **j** Combining mRNAs encoding full-length *eomesA* (*e*), *foxh1* (*f*), and *mixl1* (*m*) (*e/f/m*) triggers ectopic reporter expression also in dorsal blastomeres ( $n = 13/15$ , compared to native reporter expression pattern in **f**), an activity that also remained in embryos devoid of endoderm after *sox32* morpholino injection ( $n = 7/17$ ) (**k**). **l–p** Mixl1 acts on the +2.0drl enhancer as analyzed in Cas9 RNP-mediated cisprants. **l** Schematic representation of the *mixl1* and *mezzo* loci, with the individual sgRNAs for mutagenesis annotated (cc for CRISPR cutting, followed by sgRNA index). **m–p** mRNA in situ hybridization of EGFP expression in +2.0drl:EGFP embryos as cisprant control ( $n = 30/30$ ) (**m**), injected with Cas9 RNPs of (**n**) *mixl1 ccB* ( $n = 6/6$ ), (**o**) *mezzo ccA* ( $n = 11/11$ ), and (**p**) *mixl1 ccB* together with *mezzo ccA* sgRNA ( $n = 21/27$ ); the resulting mosaic *mixl1 ccB* cisprants show diminished +2.0drl reporter expression at late gastrulation, as pointed out by arrowheads. Lateral and animal views as indicated. Scale bar in (**b**) 500  $\mu$ m, (**c**) 250  $\mu$ m

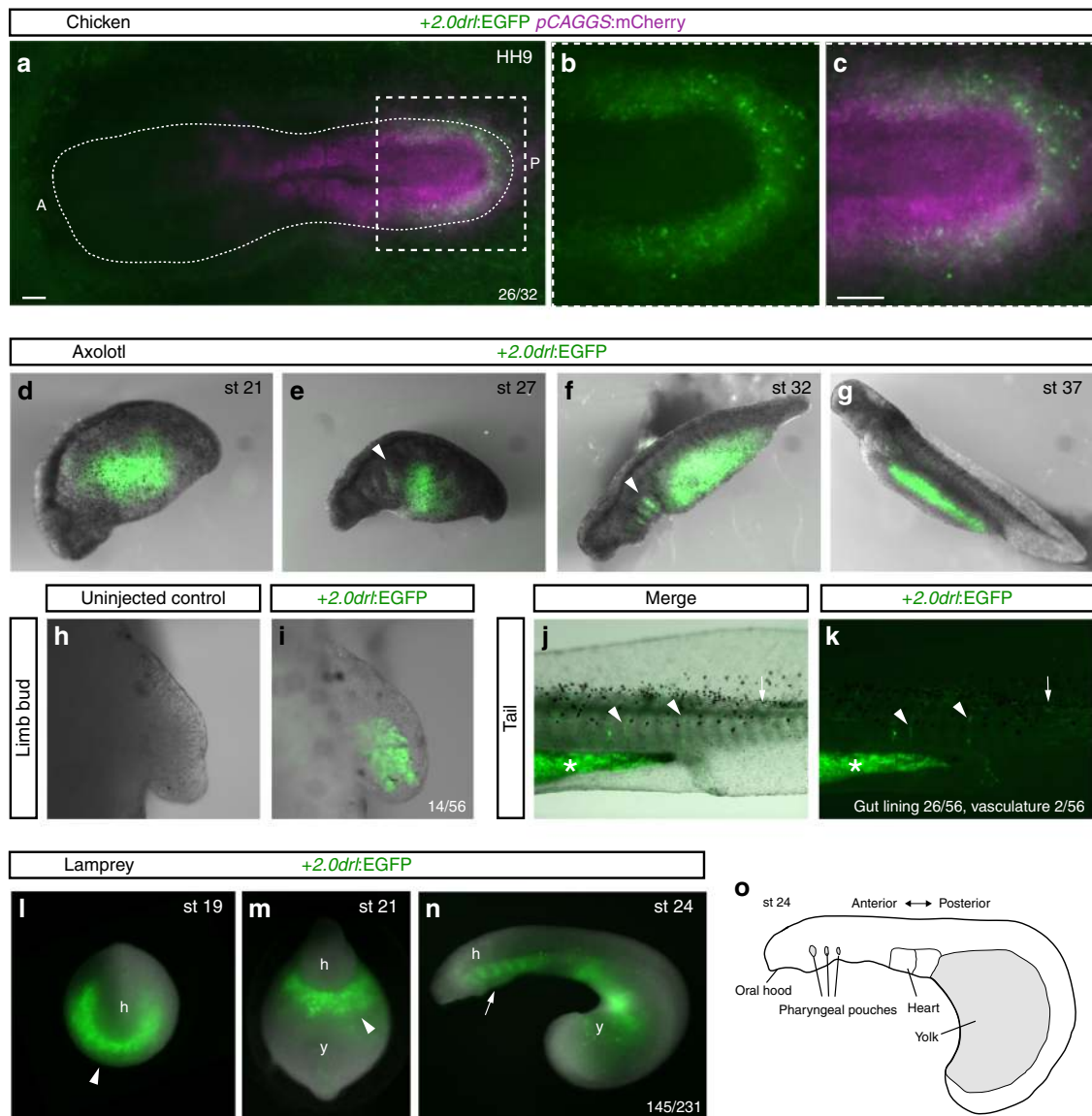
Fig. 9f). These observations suggest that a basic upstream program underlying LPM formation, as read out by the +2.0drl reporter, continues to function in birds as representative amniotes.

We then tested whether zebrafish +2.0drl is responding to LPM-inducing cues in other tetrapods. Axolotl embryos microinjected with the +2.0drl EGFP reporter marked putative endodermal and LPM territories beginning from early somite-stages (st 21), additionally marking the pharyngeal regions at tailbud stages (st 27, 32; Fig. 6d–g). Transversal sections of stage 32 embryos confirmed the presence of EGFP-positive cells in the endoderm and lateral mesendoderm (Supplementary Fig. 9g–j). Notably, EGFP fluorescence was present throughout axolotl development and could be readily detected in the gut, as well as in LPM-derived tissues including the limb bud ( $n = 14/56$ ; Fig. 6h,

i), heart ( $n = 15/56$ ), blood vessels ( $n = 2/56$ ; Fig. 6j, k), and gut lining ( $n = 26/56$ ; Supplementary Fig. 9k–n). These results support the notion that the zebrafish-derived +2.0drl enhancer also interprets an LPM program active in amphibians.

Next, we asked if the +2.0drl enhancer reads out a pan-LPM program in more distantly related vertebrates. Lampreys are jawless vertebrates (cyclostomata) that can provide unique insights into vertebrate evolution due to the early divergence of their lineage from jawed vertebrates. Microinjection of the +2.0drl EGFP-based reporter into sea lamprey embryos (*Petromyzon marinus*) consistently resulted in robust EGFP expression in the lateral mesendoderm starting during neurulation (st 18–21) ( $n = 145/231$ ; Fig. 6l, m), as well as in the developing pharynx at st 22–24 (Fig. 6n, o). Transverse embryo sections revealed that this early expression domain includes the anterior-most,



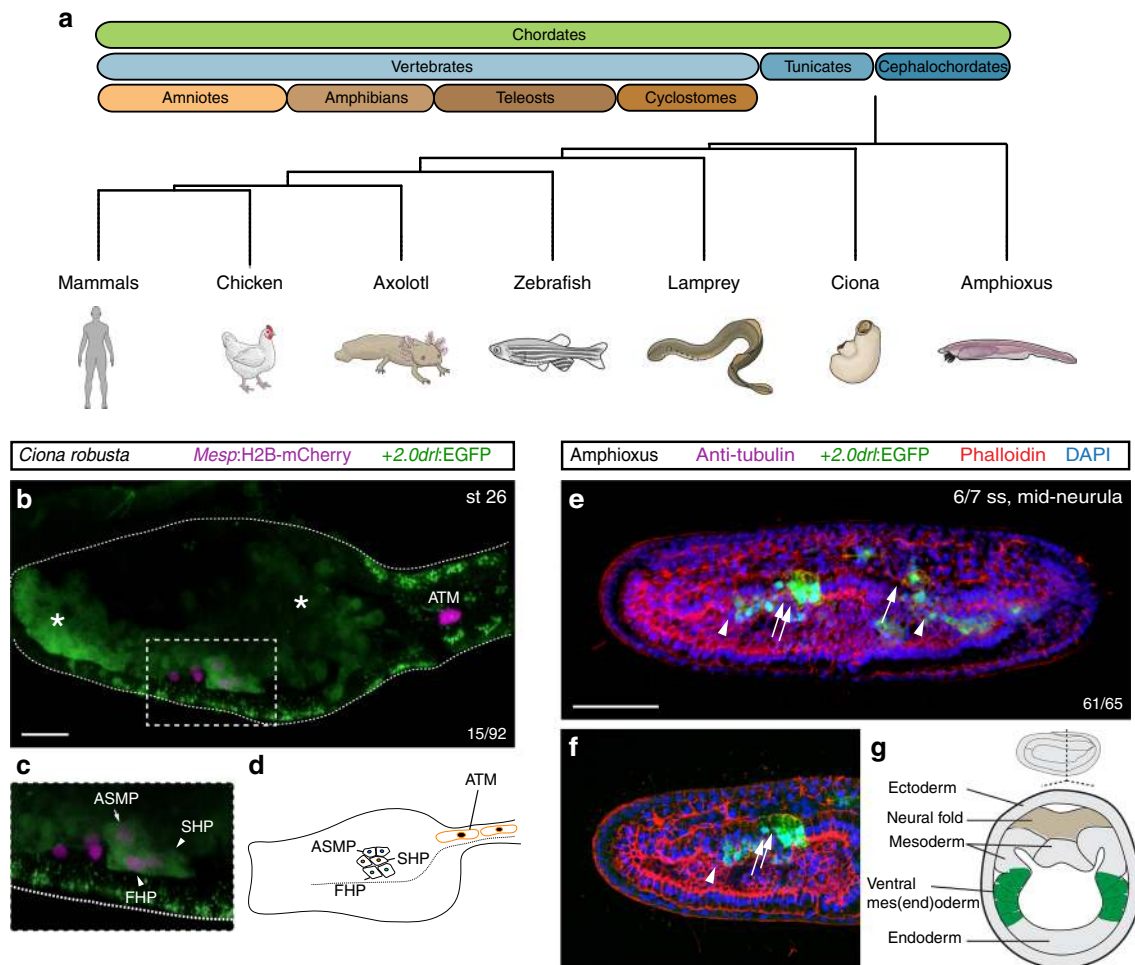


**Fig. 6** The zebrafish +2.0 *drl* enhancer reads out a LPM program across vertebrates. **a–c** Representative HH9 ex-ovo-cultured chicken embryo electroporated at HH3+/4 with +2.0*drl*:EGFP (green, **a–c**) and ubiquitous pCAGGS:mCherry control (magenta, **a, c**), showing specific +2.0*drl* reporter expression in the electroporated LPM ( $n = 26/32$ ). The dashed line depicts the outline of the chicken embryo, anterior (A) to the left, posterior (P) to the right; boxed region (**a**) depicts magnified area shown for single and merged channels (**b, c**). **d–k** Expression of +2.0*drl*:EGFP upon transient injections in axolotl at the indicated embryonic stages. Note EGFP expression in the lateral portion of the embryo, future gut region, and pharyngeal arches (arrowheads in **e, f**). **h, i** EGFP expression in mesenchymal cells of the developing axolotl limb bud, indicative of LPM origin ( $n = 14/56$ ). **j, k** EGFP fluorescence in axolotl st 43 larvae in the gut lining (asterisk,  $n = 26/56$ ) and blood vessels (arrowheads,  $n = 2/56$ ). Expression is also found occasionally in a small fraction of muscle fibers (arrows). **l–o** Transient transgenic lamprey embryos (*Petromyzon marinus*) with +2.0*drl*:EGFP expression in the anterior mesendoderm (arrowheads) and overlying the yolk at neurula stages (st 19–21) (**l, m**), and in the developing pharynx (arrow) during head protrusion (st 23–24) (**n**); views anterior (**l**), ventral (**m**), lateral (**n**), head (h) and yolk (y) ( $n = 145/231$ ). **o** Schematic depiction of st 23 lamprey embryo to outline key features. Scale bar (**a, c**) 250  $\mu\text{m}$

LPM-linked expression of lamprey *pmHandA*<sup>9</sup>, with the later pharyngeal expression of EGFP being restricted to the endoderm and mesoderm (Supplementary Fig. 10a–k). We conclude that the +2.0*drl* enhancer is capable of integrating regulatory outputs from an upstream LPM program that remains conserved across vertebrates.

Next, we asked if the +2.0*drl* enhancer also responds to upstream activity in LPM-linked cell fates dating back to the chordate radiation (Fig. 7a). First, we electroporated the zebrafish-derived +2.0*drl* reporter into embryos of the tunicate *Ciona robusta*, a chordate species belonging to a sister clade of vertebrates (Fig. 7a). While missing the full complement of LPM-

derived organ systems found in vertebrates, the LPM is echoed in the cardiopharyngeal progenitors forming in *Ciona* embryos<sup>10,37</sup>. We detected +2.0*drl*:EGFP reporter activity in emerging cardiac and pharyngeal muscle lineages at *Ciona* larval stage (st 26) (Fig. 7b–d): we observed +2.0*drl*:EGFP reporter activity in the atrial siphon muscle precursors (ASMPs) and in both first and second heart precursors (FHPs and SHPs). This was confirmed by co-localization of *Mesp:H2B-mCherry* expression that labels the cardiopharyngeal cell lineage ( $n = 15/92$ ; Supplementary Fig. 10l, m). In agreement with the *drl*-based LPM lineage tracing in zebrafish, we found minimal to no overlap with paraxial mesoderm progenitors and the anterior tail muscles (ATMs)



**Fig. 7** The zebrafish +2.0drl enhancer reads out an LPM program in tunicates and cephalochordates. **a** Phylogeny of chordates depicting the species used in this study. **b–d** *Ciona robusta* embryo at 18 hpf (st 26), electroporated with +2.0drl:EGFP (green), and with Mesp:H2B-mCherry (magenta) to track the B7.5 cardiopharyngeal cell lineage. **b, c** Representative larva shown with boxed region (**b**) magnified for detail (**c**); dashed line indicates midline, anterior to the left, schematic larva depicted in (**d**) ( $n = 15/92$ ). +2.0drl-driven EGFP 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). EGFP is also detected in mesenchymal lineages (white asterisks). **e–g** Mid-neurula stage (6/7 ss) amphioxus embryo, confocal Z-stack anterior to the left and dorsal to the top; embryo injected with +2.0drl:EGFP (green), counterstained with Phalloidin (red), DAPI (blue), lateral view as 3D rendering (**e**) and Z-stack sagittal section (**f**) ( $n = 61/65$ ). +2.0drl:EGFP showing specific reporter activity in lateral mesendoderm (arrowhead), ventral half of somites (arrows), and elongating somites (double arrows). **g** Schematic depiction of amphioxus embryo; lateral view on top, dotted line represents transverse section shown below with green depicting domain of +2.0drl:EGFP expression. Scale bar (**b**) 25 μm, (**e**) 50 μm

(Fig. 7b–d). These results indicate that the zebrafish +2.0drl enhancer responds to regulatory input in the emerging multipotent cardiopharyngeal progenitors in *Ciona*.

Lastly, we examined the cephalochordate amphioxus (*Branchiostoma lanceolatum*), which belongs to the most basally divergent lineage of chordates. In amphioxus, the LPM forms from a continuous sheet of cells that encompass the dorsally emerging somites, the LPM, and the ventral-most forming endoderm<sup>9,38–40</sup>. At mid-neurula stage (equivalent to early somitogenesis in zebrafish), the ventral wall of the somites evaginates as nascent ventral mesoderm, which by late neurula stage fuses at the midline under the gut endoderm<sup>9,38,39</sup>. Indeed, the amphioxus orthologs of LPM-expressed transcription factors including *Hand*, *Csx*, *Vent1*, and *AmphiNk2-tin* are expressed in the ventral half of the somite territory at mid-neurula stage<sup>8,38,39</sup>. We observed that injection of zebrafish +2.0drl-based reporters into amphioxus embryos showed specific reporter activity in the ventral half of somites and in the elongating somites at mid-neurula stage (6/7 ss) ( $n = 61/65$ ; Fig. 7e–g, Supplementary

Fig. 10n, o). At early larvae stage, the activity of +2.0drl-based reporter was present in the pharyngeal region (Supplementary Fig. 10p) where LPM is located<sup>40</sup>. Hence, also in amphioxus as cephalochordate, the +2.0drl enhancer reads out the positional input active in the LPM-corresponding territory during development.

Taken together, these observations establish that the zebrafish-derived intronic +2.0drl enhancer reads out a position-dependent LPM program that remains active in tunicates, cyclostomes, teleosts, amphibians, and amniotes, and thus across all tested chordates.

## Discussion

The dynamic nature of the LPM has made it challenging to precisely monitor its emergence and morphogenesis and has hindered comparative studies of its properties during chordate evolution. Here, we show that an enhancer from the seemingly zebrafish-specific *drl* locus (+2.0drl) reads out an

LPM-demarcating transcriptional activity in six chordate species, ranging from cephalochordates to amniotes, suggesting that this LPM-underlying transcriptional program is of ancient evolutionary origin. Characterization of the properties of this enhancer in zebrafish revealed that the transcription factors Eomes, FoxH1, and MixL1 are sufficient to trigger this basic LPM program. These observations in zebrafish suggest a regulatory model whereby, among their roles in mesendoderm regulation, Eomes, FoxH1, and MixL1 cooperate in inducing LPM together with position-dependent Smad activity. These factors have been individually implicated in mesendoderm development in several vertebrates<sup>27,29,32,35,41,42</sup>, in diverse LPM-associated contexts such as in blood formation<sup>43</sup>, and in reprogramming towards cardiac and renal fates<sup>44–46</sup>. Based on our series of comparative and mechanistic studies, we postulate that LPM-like origins in ancestral chordates already featured the basic molecular building blocks that enabled the increasing specialization of the LPM into its sophisticated descendant cell fates observed in vertebrates.

In vertebrates, the LPM is readily detectable after gastrulation through its position lateral to the forming somites and by several progenitor markers for hemangioblast, renal, and smooth muscle fates. Genetic tracking of several aspects of LPM emergence and patterning has been achieved previously: in mouse, transgenic strains based on *HoxB6*, *Prrx1*, *Bmp4*, *Gata2*, or *FoxF1* enable labeling of the LPM post-gastrulation<sup>11–13,47,48</sup>. We had previously found and applied the *cis*-regulatory region of the zebrafish *drl* gene to genetically track LPM emergence during both gastrulation and early somitogenesis in zebrafish<sup>14–16</sup>. While *drl* encodes a putative zinc-finger protein without a clear ortholog outside of zebrafish<sup>17,18</sup>, the early LPM-confined expression mediated by its +2.0*drl* enhancer provides a unique tool to investigate LPM origins across chordates. Using *in toto* live imaging of *drl*-based reporters together with lineage-restricted reporter transgenics, we charted LPM formation in zebrafish as a continuous process building around the entire circumference of the forming embryo (Fig. 1, Supplementary Movie 1). This mode of progenitor formation is distinct from paraxial and axial mesoderm, which both form by progressive extension over time<sup>1</sup>.

Despite the close developmental relationship between the endoderm and the LPM, also indicated by *drl* labeling a mesendoderm population that becomes progressively dedicated towards an LPM fate (Fig. 2), LPM progenitors do not seem to require endoderm for their initial morphogenesis in zebrafish. While LPM midline migration is perturbed in *sox32* (*casanova*) mutants or morphants devoid of endoderm progenitors, such embryos still form bilateral, contracting hearts<sup>19</sup> and maintain kidney and iSMC progenitor markers<sup>49</sup>. Our LPM lineage tracing confirmed that the LPM stripes still form even without endoderm and documented how they develop into structures resembling pronephros, iSMC-like structures, endothelium, erythrocytes, and pectoral fin mesenchyme (Supplementary Fig. 3). These data imply that, despite close or even joint origin, the *sox17*-positive endoderm progenitors have minimal influence on initial LPM fate determination and LPM morphogenesis. Our imaging and lineage tracing data further indicates minimal overlap between early paraxial mesoderm progenitors and LPM progenitors, as evident in the rare occurrence of somatic muscle labeling by *drl*-expressing precursors (Fig. 2f). While there is considerable heterogeneity of cell fate domains among the post-gastrulation LPM (Fig. 1), our findings collectively suggest that the LPM initially emerges as a field of cells endowed with common properties.

Guiding the therapeutically relevant differentiation of cultured embryonic or induced pluripotent stem (iPS) cells towards cardiovascular, hematopoietic, or renal cell fates remains challenging<sup>44,50–52</sup>. Initial differentiation leads to broadly defined mesodermal progenitors that, depending on the protocol, show a

preference to early versus late primitive streak regions, mimicking the anterior-to-posterior progression of vertebrate body axis formation<sup>53</sup>. Other protocols combined expression of transcription factor combinations to drive direct differentiation into specific cell fates, such as achieved for cardiomyocytes or kidney cells<sup>2,51,54</sup>. Nonetheless, directed differentiation of uncommitted cells into correct LPM progenitor states would be highly desirable to achieve increased reprogramming efficiency<sup>52,55</sup>. In this regard, our functional analyses in zebrafish showed that EomesA, FoxH1, and MixL1 together with BMP-induced Smads are able to drive cells towards an LPM program.

Eomes and FoxH1 cooperate in controlling BMP/Nodal target genes together with Smads<sup>29,35,41</sup>. Eomes, FoxH1, and MixL1 have been implicated separately or in pairwise combinations in mesendoderm development<sup>31,35,56–58</sup>. Our findings indicate that the combination of Eomes, FoxH1, and MixL1 modulates mesendodermal target genes required for progression towards LPM formation. The requirement for the combined action of all three factors becomes apparent when testing each factor individually, as there was only a marginal increase in +2.0*drl* expression. In contrast, the combination of all three factors was sufficient to ubiquitously induce the +2.0*drl* pan-LPM reporter. The three factors do however not merely boost mesendoderm fate per se, as demonstrated in embryos devoid of endodermal progenitors following *sox32* perturbation (Fig. 5k). The dependency of early *drl* expression on BMP, an less so on Nodal based on comparing genetic mutants, is further in line with the classic definition of a ventral-marginal emergence of the LPM<sup>1,4</sup>. From this data, we propose the following model: maternal Eomes and FoxH1 cooperate with BMP- and Nodal-triggered Smads to prime mesendodermal target genes. Ventral induction of MixL1 provides an instructive signal that cooperates with the previous permissive mesendoderm state to trigger an LPM fate in BMP-receiving blastomeres (Figs. 4e–i, 5f–j). These findings provide a framework for the contribution of Eomes, FoxH1, and MixL1 in programming of naïve pluripotent stem cells into cardiovascular and renal lineages by generating the correct mesendodermal precursor lineage. In-depth analysis of genomic targets of the three transcription factors is warranted to i) establish how this program conveys key LPM properties to these uncommitted progenitor cells, and ii) if or which orthologs of these T-box, Forkhead, and Homeobox factors drive LPM progenitor formation across chordates.

The evolutionary origin of the LPM has remained unaddressed. In part, the discussion of evolutionary origins of key features in the vertebrate body plan is tangled by the deduction of ancestral versus derived features without an existing common chordate ancestor<sup>59</sup>. Jawed vertebrate species share thousands of conserved non-coding regulatory regions<sup>60</sup> and a greatly reduced number can be traced to jawless vertebrates like lamprey<sup>61</sup>. Nonetheless, while some of the previously isolated LPM enhancers of mouse *Gata4*, *Bmp4*, and *Hoxb6* expressed faithfully in the PLPM of chick embryos, their activity did not recapitulate a LPM pattern in zebrafish (Supplementary Fig. 8), suggesting specialization during amniote evolution. In contrast, our cross-species regulatory analyses of the zebrafish-derived +2.0*drl* enhancer, while as endogenous sequence absent in the tested species, uncovered a remarkable degree of regulatory conservation. While cryptic reporter activity in cross-species assays can bias results, we established that the zebrafish +2.0*drl* enhancer drove specific fluorescent reporter expression in LPM or LPM-related structures in six analyzed chordates: chick, axolotl, zebrafish, lamprey, *Ciona*, and amphioxus (Figs. 6 and 7). Of note, in *Ciona*, the +2.0*drl*:EGFP reporter also resulted in EGFP activity in cells of the developing mesenchyme; if this reporter activity is specific or ectopic activity of the used transgene plasmid as previously observed<sup>62</sup>, remains to be determined. Nonetheless, the

zebrafish-derived *+2.0drl* enhancer provides a unique tool to investigate the upstream regulatory networks and the emergence of LPM structures across chordate development. This remarkable conservation of upstream reporter inputs further suggests that *+2.0drl* enhancer activity uncovers a deeply rooted LPM-inducing program, dating back to the last shared chordate ancestor.

Our findings further provide a genetic approach for investigating an LPM-delineating program that sets this mesodermal lineage apart from axial and paraxial mesoderm progenitors. The LPM program responds to ancient upstream regulatory inputs that defines the LPM from its early developmental origins across chordates. The activity of the *+2.0drl* reporter in the prospective LPM of amphioxus is particularly striking (Fig. 7), as cephalochordates form few and only rudimentary equivalents of the vertebrate LPM-derived organ systems. The LPM in amphioxus forms as a non-segmented mesodermal sheet that is continuous with the ventral prospective endoderm and the more dorsally folding somites (Fig. 7)<sup>38–40</sup>. This configuration makes it tempting to speculate that the LPM evolved from mesenchymal mesendoderm that did not integrate into the definitive endoderm or into the paraxial somites, providing ample material for diversification over deep time.

## 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, permit no. 150). 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 h 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 at 25–28 °C if not indicated otherwise. White mutant (*d/d*) axolotls (*Ambystoma mexicanum*) were obtained from the axolotl facility at the TUD-CRTD Center for Regenerative Therapies Dresden, Germany. 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 and TOPO-cloned into the *pENTR*<sup>™</sup> 5'-*TOPO*<sup>®</sup> TA Cloning<sup>®</sup> plasmid (Invitrogen) according to the manufacturer's instructions. Subsequent cloning reactions for all used transgenesis constructs were performed with the Multisite Gateway system with LR Clonase II Plus (Life Technologies) according to the manufacturer's instructions. Cloning details, transgenesis, transgenic zebrafish strains, and applied zebrafish techniques used in this study are outlined in Supplementary Methods. Primer sequences used for cloning and sgRNAs are outlined in Supplementary Table 1.

**Zebrafish transverse vibratome sections.** 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- $\mu$ m-thick sections using a vibratome (Leica VT 1000 S). 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 40 $\times$ /1.3 oil DIC M27 objective. Images were cropped and adjusted for brightness using ImageJ/Fiji<sup>63</sup>. Graphs were generated in GraphPad Prism 5.

**Zebrafish chemical treatments.** Chemicals for performed zebrafish treatments were dissolved in DMSO. Dorsomorphin (10–30  $\mu$ M; Sigma-Aldrich) and SB-505124 (30–60  $\mu$ M; Sigma-Aldrich) were administered at 1-cell stage and embryos kept in the treated E3 until fixation.

**Zebrafish selective plane illumination microscopy.** 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 an FEP tube (inner diameter: 2.0 mm, wall thickness: 0.5 mm). In all, 6–7 embryos were positioned on top of each other. The FEP tube was mounted in the microscope imaging chamber filled with E3 medium.

Time-lapse acquisition was performed by a standardized image acquisition pipeline<sup>64</sup>. The subsequent real-time image processing, registration of time points, and 2D map (Mercator) projections were performed with published Fiji scripts<sup>64</sup>. A Z-stack of 402 planes was obtained from every embryo with an interval of 2 min for a period of 14–17 h. Images were processed using ImageJ/Fiji and Photoshop CS6.

**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 the 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 (Supplementary Table 1; 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 a *pCS2\_ADMP* plasmid, and *sizzled* from *pCS2\_Sizzled*. *admp* and *sizzled* were linearized by *Cl*A. The *gata2a* probe was obtained from the middle entry vector *pCM238*. For in vitro transcription, T7 RNA polymerase (Roche) and digoxigenin (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 using standard zebrafish ISH protocols. 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.

**Chicken embryo incubation and ex-ovo culturing.** 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 h until Hamburger-Hamilton (HH) 3 +1/4 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 h at 55 °C, 5 ml 2% agarose, 1.5 ml 10% glucose in 30 mm petri dishes) culturing plate. Embryos were recovered for at least 2 h 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 Jerome Gros, Institut Pasteur, Paris). 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, 8 V  $\times$  50 ms, 500 ms interval). All injection mixtures for electroporations contained 0.1% fast green dye, 0.1% methyl-cellulose, 300 ng/ $\mu$ l control plasmid *pCAGGS* (*pCMV:H2B-CAGG-RFP*, abbreviated for chicken  $\beta$ -*actin* promoter *CAGG-mCherry*) and 1  $\mu$ g/ $\mu$ l of the plasmid of interest. 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 processed using Leica LAS, and cropped and adjusted for brightness using ImageJ/Fiji.

**Axolotl experiments.** The generation of transgenic animals and determination of developmental stages were performed following standardized protocols<sup>65,66</sup>. Animals at stage 43 were anaesthetized by bathing in 0.01% benzocaine<sup>65</sup>. Live imaging was performed on an Olympus SZX16 fluorescence stereomicroscope. Time lapse movies were acquired using an Axio Zoom.V16 (Zeiss) stereomicroscope. Confocal images were acquired on a Zeiss LSM780-FCS inverted microscope.

For immunostaining, embryos were fixed in MEMFA at 4 °C overnight, washed in PBS, embedded in 2% low-melting temperature agarose and sectioned by vibratome into 200- $\mu$ m-thick sections. Fibronectin was detected using mouse anti-Fibronectin antibody (IST-9, mouse monoclonal, ab6328, Abcam) at 5  $\mu$ g/ml.

**Lamprey experiments.** The *+2.0drl* regulatory element was amplified from the zebrafish vector *+2.0drl:EGFP* by PCR using KOD Hot Start Master Mix (Novagen) (Supplementary Table 1; regulatory elements). The amplified enhancers were cloned into the HLC vector for lamprey transgenesis<sup>67</sup>, 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/ $\mu$ l I-SceI enzyme and 20 ng/ $\mu$ l reporter construct. 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 RNeasy 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*)<sup>68</sup>, 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'-TGGAATTCGAGTGCCCA-3' (reverse). The cDNA fragment was cloned into the *pCR4-TOPO* vector (Invitrogen).

Lamprey whole-mount ISH was performed using DIG-labeled probes including for *eGFP*<sup>67</sup>. 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 an AxioCam HRC camera and AxioVision Rel 4.8.2 software.

**Ciona experiments.** +2.0*drl* was amplified from the zebrafish vector +2.0*drl*:EGFP and sub-cloned upstream of *unc76:GFP* to generate a *Ciona* reporter construct including minimal promoter (*pBuS24*; see Supplementary Table 1, regulatory elements for primer sequences). 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 +2.0*drl*:EGFP was injected in a mixture with the reporter plasmid for *Mesp*<sup>69</sup> 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 h at RT, then washed in PBT<sup>70</sup>.

**Amphioxus experiments.** The regulatory elements *drl* (entire 6.35 kb) and +2.0*drl* were amplified from the zebrafish reporter vector *drl*:EGFP<sup>14</sup> and +2.0*drl*:EGFP and subcloned upstream of a EGFP reporter in the *pPB* vector carrying *PiggyBac* transposon terminal repeats<sup>71</sup>. 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<sup>72</sup>. For microinjection of amphioxus eggs, a mixture of *pPB-drl*:EGFP or *pPB*- +2.0*drl*: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, stained with monoclonal anti-acetylated Tubulin antibody (T6793 Sigma-Aldrich, mouse ascites fluid, clone 6-11B-1, dilution 1:500), mounted with Vectashield with DAPI (Vector Laboratories), and analyzed using a Leica SP5 confocal microscope. The confocal images were adjusted for brightness and contrast with ImageJ/Fiji.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The authors declare that the data supporting the findings of this study are available within the paper and its supplementary information files. Original data underlying the lamprey experiments in this manuscript are accessible from the Stowers Original Data Repository at <http://odr.stowers.org/websimr/>.

The source data underlying Fig. 2j and Supplementary Fig. 2b and 10m are provided as a Source Data file. Reagents are available upon request.

Received: 8 April 2019 Accepted: 22 July 2019

Published online: 26 August 2019

## References

- Gurdon, J. B. *Organization of the Early Vertebrate Embryo* 51–59 (Springer, Boston, MA US, 1995).
- Takasato, M. & Little, M. H. The origin of the mammalian kidney: implications for recreating the kidney in vitro. *Development* **142**, 1937–1947 (2015).
- Chal, J. & Pourquié, O. Making muscle: skeletal myogenesis in vivo and in vitro. *Development* **144**, 2104–2122 (2017).
- Lane, M. C. & Smith, W. C. The origins of primitive blood in *Xenopus*: implications for axial patterning. *Development* **126**, 423–434 (1999).
- Davidson, A. J. & Zon, L. I. The 'definitive' (and 'primitive') guide to zebrafish hematopoiesis. *Oncogene* **23**, 7233–7246 (2004).
- Yabe, T., Hoshijima, K., Yamamoto, T. & Takada, S. Quadruple zebrafish mutant reveals different roles of *Mesp* genes in somite segmentation between mouse and zebrafish. *Development* **143**, 2842–2852 (2016).
- Kusakabe, R. & Kuratani, S. Evolutionary perspectives from development of mesodermal components in the lamprey. *Dev. Dyn.* **236**, 2410–2420 (2007).
- Pascual-Anaya, J. et al. The evolutionary origins of chordate hematopoiesis and vertebrate endothelia. *Dev. Biol.* **375**, 182–192 (2013).
- Onimaru, K., Shoguchi, E., Kuratani, S. & Tanaka, M. 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**, 124–136 (2011).
- Kaplan, N., Razy-Krajka, F. & Christiaen, L. Regulation and evolution of cardiopharyngeal cell identity and behavior: insights from simple chordates. *Curr. Opin. Genet. Dev.* **32**, 119–128 (2015).
- Becker, D., Eid, R. & Schughart, K. 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 (1996).
- Rojas, A. et al. *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**, 3405–3417 (2005).
- Chandler, K. J., Chandler, R. L. & Mortlock, D. P. Identification of an ancient *Bmp4* mesoderm enhancer located 46 kb from the promoter. *Dev. Biol.* **327**, 590–602 (2009).
- Mosimann, C. et al. Chamber identity programs drive early functional partitioning of the heart. *Nat. Commun.* **6**, 8146 (2015).
- Gays, D. et al. An exclusive cellular and molecular network governs intestinal smooth muscle cell differentiation in vertebrates. *Development* **144**, 464–478 (2017).
- Felker, A. et al. Continuous addition of progenitors forms the cardiac ventricle in zebrafish. *Nat. Commun.* **9**, 2001 (2018).
- Pimtung, W., Datta, M., Ulrich, A. M., Rhodes, J. & Zon, L. I. *Drl3* governs primitive hematopoiesis in zebrafish. *Sci. Rep.* **4**, 5791 (2015).
- Sumanas, S., Zhang, B., Dai, R. & Lin, S. 15-Zinc finger protein Bloody Fingers is required for zebrafish morphogenetic movements during neurulation. *Dev. Biol.* **283**, 85–96 (2005).
- Alexander, J., Rothenberg, M., Henry, G. L. & Stainier, D. Y. Casanova plays an early and essential role in endoderm formation in zebrafish. *Dev. Biol.* **215**, 343–357 (1999).
- Herbomel, P., Thisse, B. & Thisse, C. Ontogeny and behaviour of early macrophages in the zebrafish embryo. *Development* **126**, 3735–3745 (1999).
- Emerson, R. O. & Thomas, J. H. Adaptive evolution in zinc finger transcription factors. *PLoS Genet.* **5**, e1000325 (2009).
- Kaufman, C. K. et al. A zebrafish melanoma model reveals emergence of neural crest identity during melanoma initiation. *Science* **351**, aad2197 (2016).
- Langdon, Y. G. & Mullins, M. C. Maternal and zygotic control of zebrafish dorsoventral axial patterning. *Annu. Rev. Genet.* **45**, 357–377 (2011).
- Hild, M. et al. The *smad5* mutation somitabun blocks *Bmp2b* signaling during early dorsoventral patterning of the zebrafish embryo. *Development* **126**, 2149–2159 (1999).
- Gritsman, K. et al. The EGF-CFC protein one-eyed pinhead is essential for nodal signaling. *Cell* **97**, 121–132 (1999).
- Nelson, A. C. et al. In vivo regulation of the zebrafish endoderm progenitor niche by T-Box transcription factors. *Cell Rep.* **19**, 2782–2795 (2017).
- Nelson, A. C. et al. 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 (2014).
- Dubrulle, J. et al. Response to Nodal morphogen gradient is determined by the kinetics of target gene induction. *Elife* **4**, e05042 (2015).
- Slagle, C. E., Aoki, T. & Burdine, R. D. Nodal-dependent mesendoderm specification requires the combinatorial activities of *FoxH1* and *Eomesodermin*. *PLoS Genet.* **7**, e1002072 (2011).
- Charney, R. M. et al. *Foxh1* occupies cis-regulatory modules prior to dynamic transcription factor interactions controlling the mesendoderm gene program. *Dev. Cell* **40**, 595–607 (2017).
- Chen, X., Rubock, M. J. & Whitman, M. A transcriptional partner for *MAD* proteins in TGF-β signalling. *Nature* **383**, 691–696 (1996).
- Germain, S., Howell, M., Esslemont, G. M. & Hill, C. S. Homeodomain and winged-helix transcription factors recruit activated *Smads* to distinct promoter elements via a common *Smad* interaction motif. *Genes Dev.* **14**, 435–451 (2000).

33. Bruce, A. E. E. et al. The maternally expressed zebrafish T-box gene *eomesodermin* regulates organizer formation. *Development* **130**, 5503–5517 (2003).
34. Kunwar, P. S. et al. Mixer/Bon and FoxH1/Sur have overlapping and divergent roles in Nodal signaling and mesendoderm induction. *Development* **130**, 5589–5599 (2003).
35. Bjornson, C. R. R. et al. Eomesodermin is a localized maternal determinant required for endoderm induction in Zebrafish. *Dev. Cell* **9**, 523–533 (2005).
36. Poulain, M. & Lepage, T. Mezzo, a paired-like homeobox protein is an immediate target of Nodal signalling and regulates endoderm specification in zebrafish. *Development* **129**, 4901–4914 (2002).
37. Diogo, R. et al. A new heart for a new head in vertebrate cardiopharyngeal evolution. *Nature* **520**, 466–473 (2015).
38. Kozmik, Z. et al. Characterization of amphioxus Amphivent, an evolutionarily conserved marker for chordate ventral mesoderm. *Genesis* **29**, 172–179 (2001).
39. Holland, N. D., Venkatesh, T. V., Holland, L. Z., Jacobs, D. K. & Bodmer, R. AmphixNk2-tin, an amphioxus homeobox gene expressed in myocardial progenitors: insights into evolution of the vertebrate heart. *Dev. Biol.* **255**, 128–137 (2003).
40. Holland, N. D. Formation of the initial kidney and mouth opening in larval amphioxus studied with serial blockface scanning electron microscopy (SBSEM). *EvoDevo* **9**, 16 (2018).
41. Arnold, S. J., Hofmann, U. K., Bikoff, E. K. & Robertson, E. J. Pivotal roles for eomesodermin during axis formation, epithelium-to-mesenchyme transition and endoderm specification in the mouse. *Development* **135**, 501–511 (2008).
42. Zhang, H., Fraser, S. T., Papazoglu, C., Hoatlin, M. E. & Baron, M. H. Transcriptional activation by the Mixl1 homeodomain protein in differentiating mouse embryonic stem cells. *Stem Cells* **27**, 2884–2895 (2009).
43. Mead, P. E., Brivanlou, I. H., Kelley, C. M. & Zon, L. I. BMP-4-responsive regulation of dorsal–ventral patterning by the homeobox protein Mix.1. *Nature* **382**, 357–360 (1996).
44. Takasato, M. et al. Directing human embryonic stem cell differentiation towards a renal lineage generates a self-organizing kidney. *Nat. Cell Biol.* **16**, 118–126 (2014).
45. Costello, I. et al. The T-box transcription factor Eomesodermin acts upstream of *Mesp1* to specify cardiac mesoderm during mouse gastrulation. *Nat. Cell Biol.* **13**, 1084–1091 (2011).
46. Pfeiffer, M. J. et al. Cardiogenic programming of human pluripotent stem cells by dose-controlled activation of EOMES. *Nat. Commun.* **9**, 440 (2018).
47. Ormestad, M., Astorga, J. & Carlsson, P. Differences in the embryonic expression patterns of mouse *Foxf1* and *-2* match their distinct mutant phenotypes. *Dev. Dyn.* **229**, 328–333 (2004).
48. Martin, J. F. & Olson, E. N. Identification of a *prxl* limb enhancer. *Genesis* **26**, 225–229 (2000).
49. Reichenbach, B. et al. Endoderm-derived Sonic hedgehog and mesoderm *Hand2* expression are required for enteric nervous system development in zebrafish. *Dev. Biol.* **318**, 52–64 (2008).
50. De Los Angeles, A. & Daley, G. Q. Stem cells: reprogramming in situ. *Nature* **502**, 309–310 (2013).
51. Song, K. et al. Heart repair by reprogramming non-myocytes with cardiac transcription factors. *Nature* **485**, 599–604 (2012).
52. Lee, J. H., Protze, S. I., Laksman, Z., Backx, P. H. & Keller, G. M. Human pluripotent stem cell-derived atrial and ventricular cardiomyocytes develop from distinct mesoderm populations. *Cell Stem Cell* **21**, 179–194 (2017). e4.
53. Mendjan, S. et al. NANOG and CDX2 pattern distinct subtypes of human mesoderm during exit from pluripotency. *Cell Stem Cell* **15**, 310–325 (2014).
54. Takeuchi, J. K. & Bruneau, B. G. Directed transdifferentiation of mouse mesoderm to heart tissue by defined factors. *Nature* **459**, 708–711 (2009).
55. Murry, C. E. & Keller, G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell* **132**, 661–680 (2008).
56. Zhang, G. et al. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* **346**, 1311–1320 (2014).
57. Henry, G. L. & Melton, D. A. Mixer, a homeobox gene required for endoderm development. *Science* **281**, 91–96 (1998).
58. Kikuchi, Y. et al. The zebrafish *bonnie and clyde* gene encodes a Mix family homeodomain protein that regulates the generation of endodermal precursors. *Genes Dev.* **14**, 1279–1289 (2000).
59. Technau, U. & Scholz, C. B. Origin and evolution of endoderm and mesoderm. *Int. J. Dev. Biol.* **47**, 531–539 (2003).
60. McEwen, G. K. et al. Early evolution of conserved regulatory sequences associated with development in vertebrates. *PLoS Genet.* **5**, e1000762 (2009).
61. Parker, H. J., Piccinelli, P., Sauka-Spengler, T., Bronner, M. & Elgar, G. Ancient Pbx-Hox signatures define hundreds of vertebrate developmental enhancers. *BMC Genomics* **12**, 637 (2011).
62. Stolfi, A. & Christiaen, L. Genetic and genomic toolbox of the chordate *Ciona intestinalis*. *Genetics* **192**, 55–66 (2012).
63. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
64. Schmid, B. et al. High-speed panoramic light-sheet microscopy reveals global endodermal cell dynamics. *Nat. Commun.* **4**, 2207 (2013).
65. Khattak, S. et al. Optimized axolotl (*Ambystoma mexicanum*) husbandry, breeding, metamorphosis, transgenesis and tamoxifen-mediated recombination. *Nat. Protoc.* **9**, 529–540 (2014).
66. Bordzilovskaya, N. P., Dettlaf, T. A., Duhon, S. T. & Malacinski, G. M. in *Developmental Biology of the Axolotl* (eds Armstrong, J. B. & Malacinski, G. M.) 201–219 (Oxford University Press, Oxford, 1989).
67. Parker, H. J., Bronner, M. E. & Krumlauf, R. A Hox regulatory network of hindbrain segmentation is conserved to the base of vertebrates. *Nature* **514**, 490–493 (2014).
68. Kuraku, S., Takio, Y., Sugahara, F., Takechi, M. & Kuratani, S. Evolution of oropharyngeal patterning mechanisms involving *Dlx* and *endohelins* in vertebrates. *Dev. Biol.* **341**, 315–323 (2010).
69. Davidson, B., Shi, W. & Levine, M. Uncoupling heart cell specification and migration in the simple chordate *Ciona intestinalis*. *Development* **132**, 4811–4818 (2005).
70. Racioppi, C. et al. Fibroblast growth factor signalling controls nervous system patterning and pigment cell formation in *Ciona intestinalis*. *Nat. Commun.* **5**, 4830 (2014).
71. Kozmikova, I. & Kozmik, Z. Gene regulation in amphioxus: An insight from transgenic studies in amphioxus and vertebrates. *Mar. Genomics* **24**, 159–166 (2015).
72. Fuentes, M. et al. Insights into spawning behavior and development of the european amphioxus (*Branchiostoma lanceolatum*). *J. Exp. Zool. B Mol. Dev. Evol.* **308B**, 484–493 (2007).

## Acknowledgements

We thank Seraina Böttschi, Lukas Obernosterer, and Vesna Barros for technical and husbandry support; the lab of Stephan Neuhaus for zebrafish husbandry support; the labs of Esther Stöckli and Jerome Gros for chicken experimentation support; the ZMB at UZH for imaging support; Fiona Wardle for input and support on the ChIP-seq panel in Fig. 5a; the lab of Magdalini Polymenidou for vibratome access; Karolína Ditrychová for cloning the *pKD001* construct; Rebecca Burdine for sharing transcription factor constructs; Mark Miller for animal illustrations in Fig. 7a; the Stowers Institute histology facility for assistance with lamprey embryo sectioning; Hans-Henning Epperlein for discussions on salamander embryology; Diego Safian for suggesting Genomicus for synteny analysis; 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.; Alexander von Humboldt fellowship to A.C. and DFG Research Center (DFG FZ 111) and Cluster of Excellence (DFG EXC 168) funds to M.H.Y.; Czech Science Foundation 17-15374 S to Z.K.

## Author contributions

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

**Additional information**

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-019-11561-7>.

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Peer review information:** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019