Dissection of anterior mesendoderm segregation at single cell level in zebrafish

Tao Cheng[1,2#], Xiang Liu[2#], Yang Dong[1,2], Yan-Yi Xing[1,2], Cong Liu[2], Yun-Fei Li[2], Ying Huang[2], Jing-Yun Luan[2], Hong-Qing Liang[2], Peng-Fei Xu[1,2]*

[1] Women's Hospital, Zhejiang University School of Medicine, Hangzhou, China.

[2] Institute of Genetics and Department of Human Genetics, Zhejiang University School of Medicine, Hangzhou, China.

*Corresponding author: pengfei_xu@zju.edu.cn

[#]Tao Cheng and Xiang Liu contributed equally to this work and should be regarded as joint first authors.

Keywords

Mesendoderm, cell fate segregation, chromatin organization, multi-omics, zebrafish

Abstract:

During gastrulation, the mesendoderm is firstly specified by morphogens such as Nodal, and then segregates into endoderm and mesoderm in a Nodal concentration-dependent manner. However, so far, the underlying mechanism of this segregation remains unclear. Here, taking zebrafish prechordal plate (PP) and endoderm (Endo) as research model, using single cell multi-omics and live imaging analyses, we show that anterior endodermal progenitors originate directly from prechordal plate. Deconvolution analysis for bulk RNA-seq datasets of Nodal-injected explants reveals that the specification of anterior endoderm from PP was determined by a relatively lower Nodal signaling. And a single-cell transcriptomic trajectory analysis of wild-type, *ndr1* knockdown and *lefty1* knockout Nodal explants confirms the diversification of Endo fate from PP progenitors. Genstoe Ontology (GO) enrichment analysis indicates that chromatin organization potentially underlies the segregation of endodermal cell fate from PP. A further single-cell ATAC and RNA sequencing analysis suggests a positive correlation between Nodal activity and chromatin openness. We then identify two transcriptional regulators, *gsc* and *ripply1*, which are differentially activated in PP and Endo, and manipulation of their

expression levels tilts the cell fate decision between these two lineages. Collectively, our study suggests that different levels of Nodal activity promote transcriptional diversification between PP and Endo potentially through modulation of chromatin states, which eventually leads to diversification in cell fate decisions.

Introduction:

During gastrulation, instructed by genetic and epigenetic signals, naïve cells from blastula progressively acquire distinct cell fates and are allocated to specific domains within the embryo[1,2]. Morphogens, a cluster of diffusible molecules, play vital roles during those processes[1,3-5]. One of the key morphogen molecules, Nodal, which belongs to the transforming growth factor β (TGF-β) superfamily, plays highly conserved roles in mesendoderm induction and patterning during vertebrate gastrulation[6-8]. Two Nodal ligands named Nodal-related 1 and 2 (Ndr1/2) are highly expressed and function in the early zebrafish embryo. Ndr1/2 bind to a receptor complex comprised two type I receptors, one type II receptor and a co-receptor Tdgf1/Oep, and they collectively phosphorylate Smad2, which forms a complex with cofactor Smad4. Together with additional transcription factors such as FoxH1 and Mixer, Nodal signaling initiates the genetic program of mesendoderm specification[8-10].

It is widely accepted that Nodal induces endodermal and mesodermal cell fates via concentration gradient, such that a higher Nodal signaling activity bias cells toward endoderm, while lower levels are thought to promote mesoderm induction[11,12]. Mesodermal and endodermal common progenitors are induced by Nodal and mixed together at the onset of gastrulation[1]; while, how the common progenitors of mesendoderm took on distinct lineages remains elusive. Several recent studies reported that multiple important signaling pathways interacted with Nodal to drive mesendoderm separation[2,13-15]. Notably, lateral endoderm specification may comply with a stochastic cell fate switching model in zebrafish[16]. Specification of dorsal mesendoderm requires the highest Nodal signaling level evidenced by the fact that they are the first to be depleted when applying gradual Nodal inhibition[11]. Despite all these findings so far, how anterior mesendoderm cell fate separation is achieved is still obscure. From anterior mesendoderm, the resulting PP and Endo cell fates are both under the

control of very similar levels of Nodal signaling. What factors downstream of Nodal pathway play major roles in segregating these two fates, when and how PP and Endo cells can be distinguished transcriptionally are the questions remaining to be investigated.

In this study, we constructed a single-cell transcriptional trajectory for cell state segregation between prechordal plate and anterior endoderm using zebrafish embryo and Nodal injected explant. Combined with live imaging analyses, we found that anterior endoderm cells originated and were specified from prechordal plate progenitors, which depended on a relatively lower activity of Nodal compared to that in prechordal plate. Interestingly, we found that the separation of these two cell lineages can be traced back to as early as shield stage. More detailed single-cell trajectory analysis revealed that epigenetic regulation may be involved in the segration of prechordal plate and anterior endoderm. More interestingly, we found Nodal signaling activity was positively correlated with chromatin openness. The differential chromatin states of PP and Endo, especially for master regulators such as *gsc* and *ripply1*, further promoted their differences at transcriptional level, which facilitated the separation of these two cell lineages.

Results:

Anterior endoderm originates from prechordal plate progenitors in zebrafish.

Both anterior endoderm and prechordal plate are derived from embryonic shield and require high Nodal signaling to be specified[11,17,18]. To investigate how these two cell lineages segregate, we collected the single-cell RNA-seq datasets from published studies[19,20], and took subpopulations representing the anterior endoderm and prechordal plate cells to reconstruct a lineage branching tree (Figures 1A-1C and S1A). We found these two linages started to be separated at shield stage, and their common progenitors highly expressed markers of prechordal plate, for example, *gsc* (Figures 1C and S1A). Then, we conducted multiple analyses to further explore the dynamic molecular signatures of these two lineages during their fate determination. Pseudotime analysis indicated that prechordal plate was more mature than endoderm (Figure S1B). Of note, compared to *sox32*, *gsc* displayed earlier transcriptional differences in these two lineages (Figure S1C). The Uniform Manifold Approximation and Projection (UMAP)

analysis[21] showed that the common progenitors of PP and Endo clustered closely to prechordal plate and away from endoderm (Figures S1D and S1E), which indicated that their common progenitors were transcriptionally more similar to prechordal plate cells. The cell lineage prediction analysis[22] also revealed that the progenitor cells resembled more of the prechordal plate lineage (Figures 1D, 1E and S1F), and similar result was obtained using another single-cell RNA-seq dataset (Figure S1G -S1I). Therefore, both UMAP and the cell lineage predication analysis suggest anterior endoderm originates from prechordal palate progenitors and is transcriptionally distinguishable as early as at shield stage.

Cells from hatching gland (derived from prechordal plate) and pharyngeal pouch (derived from anterior endoderm) are distant on the transcriptome-based phylogenetic tree of zebrafish embryo[20]. Thus, it is likely that key molecules that can regulate these two lineages separation are missed when using this *in vivo* single cell datasets for analysis. Our previous study demonstrated that a single Nodal signaling gradient can induce dorsal axial mesendoderm in zebrafish embryonic explant system, including prechordal plate and anterior endoderm[23]. Notably, the single-cell trajectory tree of Nodal explant showed a closer transcriptomic relationship between the prechordal plate and anterior endoderm (Figures 1F). Thus, we reconciled the single-cell transcriptome dynamics to look for the branching point of prechordal plate and anterior endoderm in the Nodal explants (Figures 1G-H, S3A and S3B). Consistent with the findings observed in wild-type embryos, the common progenitors of these two lineages highly expressed prechordal plate marker gene (*gsc*), and the segreagation between PP and Endo had already happened at 6 hpf (hours post fertilization) (Figures 1G, 1H and S3B).

To investigate the molecular signatures during the specification of prechordal plate and anterior endoderm, we performed differential expression analysis for these two lineages in both wild-type embryos and Nodal explants. 18 genes were highly enriched along the trajectory of endoderm specification in both wild-type embryos and Nodal explants (Figure S3C). These genes enriched GO terms such as “gland development”, “negative regulation of transcription, DNA-templated” and “negative regulation of RNA biosynthetic process” (Figure S3D). While for prechordal plate, 49 highly expressed genes were identified (Figure S3C), and GO enrichment analysis showed that those genes were involved in “proteolysis involved in cellular

protein catabolic process", "cellular protein catabolic process" , "protein catabolic process", etc (Figure S3D). Cell fate specification requires the activation of a series of genes, which are regarded as gene modules probably involved in specific lineage formation[24]. We identified two gene modules respectivley enriched with genes activated in endoderm and prechordal plate lineages (Figure S3E). Gene Ontology enrichment analyses[25] showed that terms like "endoderm development" and "cell motility" were enriched for gene module 1 (endoderm), while "embryo development ending in birth or egg hatching" was enriched for gene module 2 (prechordal plate) (Figures S3F-S3H), which were highly correlated with the developmental processes and function of these two cell types. Cell lineage prediction analysis on the Nodal explant datasets also revealed that the common progenitor cells were transcriptionally similar to prechordal plate (Figures 1I and S3I), consistent to our previous observations in wild-type embryo. (Figures 1D, 1E, S1H and S1I).

To further validate this observation and investigate how these two lineages were specified during development *in vivo*, we injected *sox17*: RFP constructs into Tg(*gsc*:EGFP) zebrafish embryos at 1 cell stage, and performed live imaging from the onset to the end of gastrulation (6hpf to 10hpf). Consistent with the bioinformatic analysis, we did find that some RFP positive cells (endoderm) emerged from EGFP positive cells (prechordal plate) during development (Figures 1J and 1K, MovieS1), and this phenomenon could also be observed in zebrafish animal pole upon Nodal injection (Figures S2A and S2B, MovieS2). In short, using single-cell RNA-seq trajectory and live imaging analyses, we discovered that anterior endoderm directly originated from prechordal plate progenitor cells.

Additionally, we found some Nodal signaling downstream targets, such as *gsc*, *ndr2* and *lft2*, had already exhibited differential expression before the onset of lineage specification (as *sox17* expression) (Figure S3J). This result suggested varied Nodal signaling activity might already exist before the separation of these two lineages. To test this hypothesis, we performed deconvolution analysis[26] for bulk RNA-seq data of Nodal explants injected with different dosages of Nodal to estimate the resulting proportions of prechordal plate, endoderm, EVL and other cell types (Figure 1L). We found moderate concentration of Nodal was in favour of the endoderm specification, while higher concentration of Nodal promoted prechordal plate

specification (Figure 1M). This observation was supported by the findings demonstrated by *Keisuke Sako et al*., that extended Nodal signaling induces prechordal plate and suppresses endoderm specification[17].

Nodal-lefty regulatory loop is involved in prechordal plate and anterior endoderm separation.

Prechordal plate and anterior endoderm lineages are physically in close proximity and transcription-wise similar[27]. Repulsing effects through cell-cell interactions between two neighbor cell types often promote the robust cell fate segregation during development[27,28]. Prechordal plate cells and anterior endoderm cells may communicate with each other to influence each other's cell fate specification. We utilized LIANA[29], a ligand-receptor analysis framework, to explore the molecular signals involved in cell-cell interactions of prechordal plate and endoderm using single-cell RNA-seq datasets. Firstly, we found prechordal plate interacted strongly with endoderm both in wild-type embryos and Nodal explants (Figures S4A-S4D). Interestingly, when we set prechordal plate as source cells and endoderm as target cells using wild-type embryo dataset[19] (Figure 2A-C), we found that key molecular factors in Nodal-lefty regulatory loops were significantly enriched, including *ndr1*, *lefty1* and *lefty2* (Figure 2A). When the source and target cell types were set conversely (endoderm as source and prechordal plate as target), the number of enriched interactions (ligand-receptor pairs) were reduced, and Nodal-lefty interactions and ligands *lft1* were absent (Figure 2B). In Nodal explant system[23] (Figures 2D-2F), we also found more Nodal-lefty interaction pairs when setting PP as source cell types, especially for both *lefty1 and lefty2* (Figures 2E and 2F). Compared to the conditions of wild-type embryo, *ndr1* was not found in the interactions between prechordal plate and endoderm in Nodal explant (Figures 2E and 2F). Except Nodal-lefty signaling, Fgf and Wnt signalings were also significantly enriched in cell-cell communications between these two cell types both in embryos and Nodal explants (Figures 2A, 2B, 2E, 2F and S4E-S4H). It has been noted that Fgf signaling were previously found to interact with Nodal signaling to regulate mesoderm and endoderm separation[13,30].

Loss of *lft1* or *ndr1* leads to alteration of cell compositions in Nodal explants.

Above results suggested that Nodal-lefty signaling may be invovled in prechordal plate and anterior endoderm fate separation. In principle, it is wildly accepted that Nodal-lefty functions as a reaction-diffusion system to facilitate Nodal gradient formation and tissue patterning[31-35] (Figure 2G). To validate the effect of single Nodal or lefty ligand on Nodal activity and its gradient shape, we generated an ectopic Nodal gradient in zebrafish animal pole cells by injection of *ndr2* mRNA at 128-cell stage in wild-type embryos, *lft1* mutants or *ndr1* morphants (Figures 2H, 2I, S5A-S5C). We didn't observe significant differences of the Nodal activity in *lft1* mutants or *ndr1* morphants at sphere stage (4 hpf), as shown by pSmad2 immunostaining (Figures S5A and S5B). However, at shield stage (6 hpf), Nodal activity was dramatically enhanced in *lft1* mutants, and was decreased upon *ndr1* knockdown (Figures 2H and 2I).

To further investigate if Nodal-lefty system is involved in the regulation of prechordal plate and anterior endoderm separation, we performed single-cell transcriptomic sequencing for Nodal explants constructed from *lft1* mutants and *ndr1* morphants at shield stage (Figure 3A). We integrated these two single-cell RNA-seq datasets with that from the wild-type Nodal explant. After unsupervised clustering, we obtained 9 different cell clusters in this combined single-cell RNA-seq dataset (Figure 3B). Most cells were clustered into epiblast with high expression of *apoeb* and *her3*. For mesodermal lineage, we further identified two cell types, which were named as mesoderm_PP and mesoderm_NT. Mesoderm_PP cells were the progenitors of prechordal plate evidenced by the high expression of *gsc* and *chrd*, while mesoderm_NT cells were the progenitors of notochord charaterized by the high expression of *noto* and *tbxta* (Figures S6A and S6B). We found a small cluster of endodermal cells (marked by *sox32* and *sox17* expression), close to the prechordal plate cells on the UMAP (Figure 3B), which further suggested transcriptional relevance between anterior endoderm and prechordal plate.

We then calculated the proportions of prechordal plate and endoderm cells in these three datasets. It was found that the proportions of these two cell types both increased in *lft1* mutant explants, while decreased in *ndr1* morphants compared to wild-type explants (Figures S6C and S6D).

A meticulous single-cell trajectory tree reveals the involvement of epigenetic regulation in prechordal plate and anterior endoderm specification.

To further estimate Nodal activity during PP and Endo cell fate segregation, we defined a Nodal score by regressing out the expression level of all the Nodal downstream targets which were differentially expressed upon Nodal treatment while also had pSmad2 binding peaks nearby (**See Methods**). We overlapped Nodal targets identified by pSmad2 ChIP-seq experiments in wild-type embryos[36] and 105 Nodal downstream targets identified by our previous bulk RNA-seq analysis in Nodal explants[23]. 29 Nodal downstream targets were identified and considered to be highly-correlated with Nodal activity (Figure 3C), and then these genes were used to calculate Nodal score based on their expression. We found Nodal score was dramatically decreased in *ndr1* knockdown samples and slightly increased in *lft1* mutant samples compared to that in the wild-type samples (Figures 3D and 3E), which was consistent with the results of pSmad2 immunostaining showing enhanced Nodal signaling in *lft1* mutant embryos and the decreased Nodal activity in *ndr1* knockdown embryos (Figures 2H and 2I). Thus, by comparing Nodal scores and the corresponding cell fate identities, we may be able to reveal how quantitative variation in Nodal activity reflects on the cell fate determination.

We assessed the levels of Nodal score across all the cell types in Nodal explants at 6 hpf, and found that axial mesendodermal cells, including prechordal plate, notochord and anterior endoderm, displayed the highest Nodal scores (Figure S6E), which was consistent with the prior knowledge that Nodal activity was the highest in the dorsal margin of zebrafish embryos at the onset of gastrulation[1,37,38]. A further comparsion of Nodel scores between PP and Endo revealed that PP had a sightly higher Nodal score than Endo (Figure S6E). Moreover, to identify the transcriptome dynamic of prechordal plate and anterior endoderm separation, we reconstructed the single-cell trajectory tree of these two cell types based on the three Nodal explant datasets (Figure 3F). A continuous cell-state-transition process from prechordal plate progenitors to anterior endoderm were captured (Figure S6F), and a key branching point in this process was identified (Figure 3F). To obtain the molecular dynamics during this branching process, we performed differential expression analysis for these two lineages around the branching point

(Figure 3G). We identified three clusters of gene sets enriched in pre-branch (progenitors), left-branch (endoderm) and right-branch (prechordal plate) respectively. GO enrichment analysis using these three gene sets identified terms related to mesendodermal cell fate specification. For instance, “regulation of endodermal cell fate specification” was found in in endoderm lineage, while“negative regulation of transcription by RNA polymerase II” was found in prechordal plate lineage. Besides, “chromatin organization” was significantly enriched in the pre-branch cells, and it also remained highly enriched during the prechordal plate specification, but less enriched upon endodermal lineage formation (Figure 3G), suggesting that epigenetic distinctions may underlie the prechordal plate and anterior endoderm separation. Additionally, GO term of “chromatin organization” was also enriched for the DEGs of the prechordal plate cells when comparing *lft1* mutant or *ndr1* knockdown samples to wild-type (Figures 3H-3L and S6G).

Integrative single-nucleus RNA-Seq and ATAC-Seq analysis reveals the impact of chromatin accessibility on prechordal plate and endoderm separation.

To further investigate the role of chromatin organization in prechordal plate and endoderm separation, we employed single-cell assays for transposase-accessible chromatin (ATAC) and nuclear RNA sequencing in zebrafish embryos at 6 hpf (Figure 4A). Cells cannot be properly clustered only based on chromatin accessibility datasets. Thus, we integrated chromatin accessibility data with RNA expression data, and clustered the cells into 10 different groups on the UMAP (Figures 4B, S7A and S7B). Similar to previous single-cell RNA atlas of Nodal explants, endoderm was found to locate closely to the axis mesoderm (prechordal plate and notochord progenitors) on the joint UMAP (Figure 4B).

We used the prechordal plate progenitors marker gene *gsc* to identify PP progenitor cells in axis mesoderm, and then prechordal plate and endoderm cells were selected for downstream analysis (Figure 4C). A small cluster of cells expressing *sox32*, *sox17* and *foxj1a*, which were located far away from prechordal plate and endoderm, were classified as dorsal forerunner cells (DFCs) and excluded in this analysis (Figures 4C and S7C).

Correlation analysis revealed that the chromatin accessibility highly corresponded to RNA

261 expression levels for the marker genes in these two cell types (Figure S7D). To characterize
262 master regulators of prechordal plate and endoderm separation, we performed differential
263 analysis between prechordal plate and endoderm for both chromatin accessibility and gene
264 expression (Figures 4D and S7E). 170 up-regulated genes including *gsc* and *ripply1* in
265 prechordal plate and 261 up-regulated genes such as *sox32* and *sox17* in endoderm were
266 identified (Figures 4E, 4F, S7F, S7G and S7H).

267 Previous studies reported that both Gsc and Ripply1, can act as transcriptional
268 repressors[17,39], and they can induce a secondary axis formation when overexpressed in the
269 ventral side of zebrafish embryos[23,40]. It also has been reported that endoderm differentiation
270 could be suppressed by upregulation of *gsc*[17]. As Ripply1 and Gsc have a similar function in
271 induction of a secondary axis; additionly, *ripply1* was highly expressed in pechordal plate of
272 Nodal explants at 6hpf (Figure S7I), we hypothesized that Ripply1 may also suppress endoderm
273 specification. To test this hypothesis, we injected *ripply1* mRNA at 1-cell stage in zebrafish
274 embryo and found that endoderm cell number was dramatically decreased as shown by whole-
275 mount in situ hybridization (WISH) of *sox32* (Figure 4L). To further investigate the cell
276 autonomous or non-autonomous function of Ripply1 in endoderm repression, we constructed a
277 plasmid that contained *ripply1* cDNA downstream of the *sox17* promoter. Injection of this
278 plasmid at 1 cell stage also resulted in a decrease in endoderm specification shown by EGFP
279 expression (Figure 4K). Taken together, these results demonstrated that Ripply1 acted as an
280 endoderm repressor in a cell autonomous manner. In addition to Gsc and Ripply1, another
281 transcriptional repressor, *osr1* was enriched in prechordal plate (Figure S7J) which had been
282 identified as a negative feedback regulator of Nodal-induced endoderm development[41]. Thus,
283 our data suggested several endoderm repressors may potentially function in prechordal plate to
284 prevent cells from differentiating into endoderm.

285 Our data revealed differential chromatin accessibility existed between prechordal plate and
286 endoderm cells. Previous study reported that Nodal signaling can boost chromatin accessibility
287 both *in vivo* and *ex vivo*[42,43]. Thus, we speculated that the differences in Nodal activity may lead
288 to differential chromatin accessibility levels in PP and Endo. To verify this hypothesis, we
289 defined TSS enrichment score[44] indicating chromatin accessibility levels, and calculated both

Nodal score and TSS enrichment score in prechordal plate and endoderm cells. We found TSS enrichment score was positively correlated with Nodal score across all cells (Figure 4J), and both scores were higher in prechordal plate than in endoderm (Figures 4G, 4H and 4I). All these results suggested that a higher Nodal activity may promote the increased chromatin accessibility level of PP cells in zebrafish.

In summary, our study identified that a variation of Nodal signaling pathway promoted a differential chromatin accessibility between prechordal plate and endoderm, which further led to differential expression of those key regulators such as *gsc* and *ripply1* in these two cell lineages and the segregation of mesendoderm cell fates (Figure 5).

Discussion:

In this work, we studied the specification and segregation of prechordal plate and anterior endoderm lineages in zebrafish by single-cell multi-omics and live imaging analyses. We found that the anterior endoderm originated from prechordal plate progenitors, and identified that differential Nodal signaling levels were correlated with the lineage commitment, which potentially activated cell-fate-specific regulators at epigenetic and transcriptional levels.

The induction of mesendodermal cell fate has been extensively studied over the last decades, and it is well known that Nodal signaling is sufficient and necessary for this process[9,12,45,46]. While, how meso- and endo- lineages separate from common progenitors is still under extensive research. So far, several mechanisms have been reported to support mesendoderm cell fate separation. The most widely accepted mechanism is that high Nodal signaling levels promote endoderm specification, while lower levels are thought to induce mesoderm[11-13]. These quantitative effects of Nodal signaling lead to differential activation of transcription targets which have differential sensitivity to Nodal activation[2,36]. However, recent study proposes that Nodal signal level does not determine endodermal or mesodermal cell fate; instead, Nodal signal provides competency for mesendodermal progenitors to stochastically switch and commit to lineage-specific fates in zebrafish[16]. Another study unveils that the duration of Nodal signaling can specify different cell fates of mesendoderm[47]. Especially, prolonged Nodal signaling promotes prechordal plate specification and suppresses endoderm

319 induction[17]. Besides, other important signaling pathways in early embryonic development also
320 interact with Nodal signaling to regulate cell fate separation[13,30]. Noticeably, the mechanical
321 interaction between neigbouring cells may also contribute to cell fate decision by interpreting
322 morphogen signaling[27].

323 Our study provided a new layer of mechanisms in regulating mesendodermal cell fate
324 segregation by Nodal signaling. We proposed that Nodal signaling variance determined
325 prechordal plate and endoderm cell fate separation potentially by establishing differential
326 chromatin states. When a morphogen gradient is set up, the first task of the signaling-receiving
327 cells is to response and interpret the signaling. It has been reported that many factors could
328 affect morphogen interpretation, including ligand concentration[2,3,48], exposure time[31,49] and the
329 kinetics of target gene induction[36]. In our study, we suggested chromatin openness may function
330 as a key factor that can influence morphogen interpretation. This observation informs us that
331 the influence of chromatin structure should be taken into account when studying how
332 morphogen gradient regulates patterning formation *in vivo*, *ex vivo* and in silico. Morphogen
333 interpretation mediated by chromatin accessibility during development is a widespread
334 mechanism across different species. In Drosophila embryos, bicoid morphogen gradient
335 organizes chromatin state in a concentration dependent manner to facilitate patterning[50,51]. In
336 mouse embryos, the chromatin remodeler Brg1, which is activated by Shh, interacts with the
337 Shh to control bulge stem cells during tissue regeneration[52]. Additionally, in human embryonic
338 stem cells, chromatin structure remolding induced by specific morphogens, such as retinoic
339 acid and BMP4, is essential for p63 to specify diverse transcriptional programs during surface
340 epithelial commitment[53]. In our work, we demonstrated that chromatin remolding together with
341 Nodal morphogen promoted the prechordal plate and endoderm separation in zebrafish.

342 There are also several limitations in our current study. Firstly, we found that the chromatin
343 structures of prechordal plate and endoderm were obviously different, especially for the
344 genomic regions of those master regulators *gsc*, *ripply* and *osr1*. However, it remains unknown
345 which epigenetic regulators function parallel with or can response sensitively to small changes
346 of Nodal signaling and differentially regulate chromatin landscape of prechordal plate and
347 endoderm. Secondly, the anterior endoderm and ventral-lateral endoderm were mixed together

in our single-cell ATAC-seq dataset, which brought in interferences when we interpreted the mechanisms of the separation between prechordal plate and anterior endoderm. More replicates of embryonic single-cell multi-omics or using Nodal explants may help to address this issue in the future studies. Lastly, we used overexpression assay to determine whether *gsc* and *ripply1* can suppress endoderm specification, and this result elicited the conclusion that these two genes cooperatively regulated cell separation of prechordal plate and endoderm. As pointed out in the work of Keisuke Sako et al., "*It is thus conceivable that gsc has a partially redundant function in this process, and other as yet unidentified transcriptional repressors might cooperate with gsc in suppressing endoderm specification*"[17], our current work indicated *ripply1* might be one of the "unidentified transcriptional repressors". Thus, it would be worth studying how *gsc* and *ripply1* cooperatively regulate prechordal and endoderm specification in the future studies.

Methods:

Animals ethics

Zebrafish strains were conducted following standard procedures and experimental procedures were approved by the Institutional Review Board of Zhejiang University. The protocol number is ZJU20220375.

Cell-cell communication analysis

R package, LIANA[29] v0.1.5 (https://github.com/saezlab/liana/), was used for the investigation of ligand-receptor interactions among different cell types. The ligand-receptor analysis was performed against the ligand-receptor reference achieved from Omnipath resource using 5 different methods implemented in LIANA, including Natmi, Connectome, SingleCellSingalR, iTALK and CellPhoneDBv2. The consensus rank aggregated from the results of these methods was used for the identification of enriched ligand-receptor interactions.

Differential expression analysis of prechordal plate cells between wild-type, *lft1*-mutant and *ndr1*-morphant Nodal explant

The three sing-cell RNA-seq datasets of Nodal explants constructed from wild-type, *lft1*-mutant and *ndr1*-morphant were intergrated by Seurat[54]. To investigate the functional and regulatory effects of

Nodal signals on cell differentiation, we compared the wild type mesoderm PP cells to the mesoderm PP cells in *lft1* mutant and *ndr1* MO respectively, and differentially-expressed (DE) genes were identified through Seurat v4.0.2 (https://github.com/satijalab/seurat) with a Bonferroni adjusted p-value < 0.05. As it is known that EVL is not affected by Nodal signaling[23,55], DE genes identified in EVL were regarded as background noises and were removed from the mesoderm PP DE genes. The mesoderm PP DE genes positively correlated with Nodal concentration were then selected, constructing two gene sets: upregulated DE genes in *lft1* mutant and downregulated DE genes in *ndr1* MO. GO enrichment analysis was performed on these two gene sets respectively using clusterProfilter v4.2.2 (https://github.com/YuLab-SMU/clusterProfiler). The enriched GO terms in "biological process" subontology were identified using a Benjamini-Hochberg adjusted p-value < 0.05.

Gene set enrichment analysis

The wild type mesoderm PP gene expression was compared to the mesoderm PP gene expression of *lft1* mutant and *ndr1* MO respectively using Seurat v4.0.2 (https://github.com/satijalab/seurat), and the genes expressed more than 10% cells in either of the two compared populations were retained. A ranked list was formed on the retained genes using sign($\log_2$FC) * (-$\log_{10}$PValue) as the ranking statistic. The mapping linking GO terms to zebrafish genes were achieved through org.Dr.eg.db v3.14.0 (https://bioconductor.org/packages/org.Dr.eg.db/). GSEA was performed on the selected GO terms by clusterProfiler v4.2.2 (https://github.com/YuLab-SMU/clusterProfiler), using fgsea (https://github.com/ctlab/fgsea) with 10^5 interactions. R package, enrichplot v1.14.2 (https://github.com/YuLab-SMU/enrichplot), was then used to visualize the GSEA results.

Immunostaining and quantification of pSmad2

Samples were fixed at 4% paraformaldehyde overnight for pSmad2/3 immunostaining. The experiment was performed as previously described[23]. Confocal laser scanning microscopy (OLYMPUS CSU-W1) was used to scan the immunostained samples. Amira 6.8 software was used (Thermo Fisher Scientific) to quantify the signaling activity of pSmad2 (488 nm), RFP (561 nm) and DAPI (405 nm). Fistly, we performed maximum Intensity Z-projection to construct a 2D image. In order to make all images to be comparable, we used DAPI signal as an internal control to

408 normalized the signal intensity in other channels. All cells were selected manually by Amira. When
409 we plot a pSmad2 signal intensity gradient along a distance, the cell with the highest P-Smad2 signal
410 was assigned as start point, and then the distance of other cells was calculated by comparing them
411 with start point cell. The fitting curve was plotted by ggplot2 in R (https://www.r-project.org).
412
413 **Production of Nodal explants**
414 Nodal explant was generated as our previous work described[23]. 10 pg *ndr2* mRNA was injected into
415 one blastomere of zebrafish embryonic animal pole at the 128-cell stage. About half of the animal
416 pole region of the embryo was cut off from the blastula at the 1k-cell stage, and was cultured in a
417 Petri dish coated with 1.5% agarose gel filled with Dulbecco's modified Eagle's medium/nutrient
418 mixture F-12 (DMEM/F-12) with 10 mM HEPES, 1x minimum essential medium (MEM)
419 containing non-essential amino acids and supplemented with 7 mM CaCl2, 1 mM sodium pyruvate,
420 50 μg/ml gentamycin, 100 μM 2-mercaptoethanol, 1× antibiotic-antimycotic (15240062, Gibco) and
421 10% serum replacement. For generating *ndr1*-morphant Nodal explants, *ndr1* morpholino was
422 injected into yolk of wild-type zebrafish embryos at 1 cell stage; while for generating *lft1*-mutant
423 Nodal explants, the embryos of *lft1* mutant were used to construct Nodal explants. The sequence of
424 *ndr1* morpholino was 5' - ATGTCAAATCAAGGTAATAATCCAC - 3'.
425
426 **Whole-mount in situ hybridization (WISH)**
427 The processes of WISH and the construction of *sox32* were described in our previous work[23]. For
428 WISH experiments, probes were labeled with digoxigenin-11-UTP (Roche Diagnostics), and the
429 substrate was NBT/BCIP.
430
431 **Processing and analyzing single-cell RNA-seq data**
432 Illumina sequencing reads were aligned to the reference genome (GRCz11) of zebrafish through
433 10× Genomics CellRanger pipeline (version 3.0.2) with default parameters. Nodal explant generated
434 from *lft1* mutant and *ndr1* morphant yielded 5,419 and 5,195 cells respectively. The expression
435 matrix of each sample was obtained via CellRanger pipeline analysis and used for further analysis.
436
437 **Processing and analyzing single-cell multiomics data**
438 Illumina sequencing reads were aligned to the reference genome (GRCz11) of zebrafish using 10×
439 Genomics CellRanger-arc pipeline (version 2.0.0) with default parameters. Expression matrix and

fragments files were obtained after running the pipeline. Signac[44] was used for downstream analyses with default parameters. Filtering out low quality cells with nCount_ATAC < 100000 & nCount_RNA < 50000 & nCount_ATAC > 1000 & nCount_RNA > 1000 & nucleosome_signal < 2 & TSS.enrichment > 1 & percent.mt < 4. And then, gene expression data and DNA accessibility data were processed respectively by Signac with default parameters. The RNA expression level was utilized to identify cluster markers which were used for cell annotation. Gene activity matrix was obtained by assessing the chromatin accessibility associated with each gene through GeneActivity() function.

Single-cell RNA-seq trajectory analysis

Three different methods were used to perform single-cell RNA-seq trajectory analysis. The developmental trajectory tree of zebrafish embryos and Nodal explants was constructed by URD[20] with default parameters as described in previous studies. The branchpoint preference plot of prechordal plate and endoderm was constructed by URD with default parameters. Firstly, branchpointPreferenceLayout() function was used to define the preference layout for the branchpoint. And then stage, pseudotime information and gene expression were plotted on the branchpoint using plotBranchpoint() function. The expression matrix of prechordal plate and endoderm cells were also used to construct trajectory tree by monocle2 and monocle3 with default parameters. And deferential expression analysis was performed at the branching point of prechordal plate and endoderm separation on trajectory tree constructed by BEAM() function in monocle2. Modules of co-regulated genes along prechordal plate or endoderm differentiation lineage were identified by find_gene_modules() function in monocle3 with default parameters. The cells contributed to the preference layout for the endoderm and prechordal plate brachpoint from URD analysis were also selected as inputs for Palantir. We used Harmony to calculate the augmented affinity matrix for data of all time points. Data was visualized by using force directed layouts. Palantir was used for downstream analysis and using a nanog^{high} as the start cell. A frzb^{high} cell (prechordal plate) and a sox17^{high} cell (endoderm) were selected as the terminal cells. Endoderm and prechordal plate cell differential trajectory was detected by palantir.core.run_palantir function from Palantir. And finally, gene expression trend along Palantir inferred pseudotime was calculated by palantir.presults.compute_gene_trends function.

470

471 **Deconvolution of bulk RNA-seq data**

472 The bulk RNA-seq datasets of 6 hpf-Nodal explants injected with different dosages of *ndr2* mRNA
473 (0 pg, 2 pg, 6 pg and 10 pg) at 128 cell stage were used to perform deconvolution analyses. Two
474 different single-cell RNA-seq datasets (zebrafish embryos and Nodal explants single-cell RNA-seq
475 at 6 hpf) were used as references. MuSiC[26] was used to perform deconvolution analysis with default
476 parameters. The estimated cell proportions of EVL, endoderm and prechordal plate were used for
477 statistic analysis between bulk RNA-seq samples with different dosages of Nodal injection.

478

479 **Overexpression of *gsc* and *ripply1* in zebrafish embryos**

480 The mRNA of *gsc* and *ripply1* (synthesized *in vitro*), and *sox17* promotor drived *ripply1* plasmid
481 were co-injected into zebrafish embryos at 1 cell stage. The embryos were imaged or fixed at 8 hpf.
482 The *sox32* probe was used for WISH.

483

484 **Inhibition of heatshock genes in zebrafish embryos**

485 Wild-type embryos, *lft1* mutants and *ndr1* morphants were treated with DMSO (30 μM), hsp70
486 inhibitor (10 μg/ml) and hsp90 inhibitor (10 μg/ml) at 1K stage respectively. The morphology of
487 embryos in each group was photographed by bright filed microscopy at 6 hpf, 10 hpf and 24 hpf.

488

489 **Data availability**

490 The accession number for the single-cell RNA sequencing and single-cell multiomics data reported
491 in this paper is GEO: GSE223636.

492

493 **Code availability**

494 Any custom code and data are available from the authors upon request. All analyses are based on
495 previously published code and software.

496

497 **Additional information**

498 Supplemental Information Supplementary Figures 1-8, Supplementary Data 1-8, Supplementary
499 and Supplementary Movie 1-2.

500

501 **Refference:**

502 1. Schier, A.F., and Talbot, W.S. (2005). Molecular genetics of axis formation in zebrafish.

Annual review of genetics *39*, 561-613. 10.1146/annurev.genet.37.110801.143752.

2. Jones, W.D., and Mullins, M.C. (2022). Cell signaling pathways controlling an axis organizing center in the zebrafish. Current topics in developmental biology *150*, 149-209. 10.1016/bs.ctdb.2022.03.005.

3. Gurdon, J.B., and Bourillot, P.Y. (2001). Morphogen gradient interpretation. Nature *413*, 797-803. 10.1038/35101500.

4. Xu, P.F., Houssin, N., Ferri-Lagneau, K.F., Thisse, B., and Thisse, C. (2014). Construction of a vertebrate embryo from two opposing morphogen gradients. Science (New York, N.Y.) *344*, 87-89. 10.1126/science.1248252.

5. Xu, P.F., Borges, R.M., Fillatre, J., de Oliveira-Melo, M., Cheng, T., Thisse, B., and Thisse, C. (2021). Construction of a mammalian embryo model from stem cells organized by a morphogen signalling centre. Nature communications *12*, 3277. 10.1038/s41467-021-23653-4.

6. Chen, Y., and Schier, A.F. (2001). The zebrafish Nodal signal Squint functions as a morphogen. Nature *411*, 607-610. 10.1038/35079121.

7. Schier, A.F. (2009). Nodal morphogens. Cold Spring Harbor perspectives in biology *1*, a003459. 10.1101/cshperspect.a003459.

8. Hill, C.S. (2018). Spatial and temporal control of NODAL signaling. Current opinion in cell biology *51*, 50-57. 10.1016/j.ceb.2017.10.005.

9. Schier, A.F. (2003). Nodal signaling in vertebrate development. Annual review of cell and developmental biology *19*, 589-621. 10.1146/annurev.cellbio.19.041603.094522.

10. Feldman, B., Gates, M.A., Egan, E.S., Dougan, S.T., Rennebeck, G., Sirotkin, H.I., Schier, A.F., and Talbot, W.S. (1998). Zebrafish organizer development and germ-layer formation require nodal-related signals. Nature *395*, 181-185. 10.1038/26013.

11. Thisse, B., Wright, C.V., and Thisse, C. (2000). Activin- and Nodal-related factors control antero-posterior patterning of the zebrafish embryo. Nature *403*, 425-428. 10.1038/35000200.

12. Dougan, S.T., Warga, R.M., Kane, D.A., Schier, A.F., and Talbot, W.S. (2003). The role of the zebrafish nodal-related genes squint and cyclops in patterning of mesendoderm. Development (Cambridge, England) *130*, 1837-1851. 10.1242/dev.00400.

13. van Boxtel, A.L., Economou, A.D., Heliot, C., and Hill, C.S. (2018). Long-Range Signaling Activation and Local Inhibition Separate the Mesoderm and Endoderm Lineages. Developmental cell *44*, 179-191.e175. 10.1016/j.devcel.2017.11.021.

14. Martyn, I., Kanno, T.Y., Ruzo, A., Siggia, E.D., and Brivanlou, A.H. (2018). Self-organization of a human organizer by combined Wnt and Nodal signalling. Nature *558*, 132-135. 10.1038/s41586-018-0150-y.

15. Poulain, M., Furthauer, M., Thisse, B., Thisse, C., and Lepage, T. (2006). Zebrafish endoderm formation is regulated by combinatorial Nodal, FGF and BMP signalling. Development (Cambridge, England) *133*, 2189-2200. 10.1242/dev.02387.

16. Economou, A.D., Guglielmi, L., East, P., and Hill, C.S. (2022). Nodal signaling establishes a competency window for stochastic cell fate switching. Developmental cell *57*, 2604-2622 e2605. 10.1016/j.devcel.2022.11.008.

17. Sako, K., Pradhan, S.J., Barone, V., Inglés-Prieto, Á., Müller, P., Ruprecht, V., Čapek, D., Galande, S., Janovjak, H., and Heisenberg, C.P. (2016). Optogenetic Control of Nodal

Signaling Reveals a Temporal Pattern of Nodal Signaling Regulating Cell Fate Specification during Gastrulation. Cell reports *16*, 866-877. 10.1016/j.celrep.2016.06.036.

18. Warga, R.M., and Nusslein-Volhard, C. (1999). Origin and development of the zebrafish endoderm. Development (Cambridge, England) *126*, 827-838. 10.1242/dev.126.4.827.

19. Wagner, D.E., Weinreb, C., and Collins, Z.M. (2018). Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. Science (New York, N.Y.) *360*, 981-987. 10.1126/science.aar4362.

20. Farrell, J.A., and Wang, Y. (2018). Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. Science (New York, N.Y.) *360*. 10.1126/science.aar3131.

21. Becht, E., and McInnes, L. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nature biotechnology. 10.1038/nbt.4314.

22. Setty, M., Kiseliovas, V., Levine, J., Gayoso, A., Mazutis, L., and Pe'er, D. (2019). Characterization of cell fate probabilities in single-cell data with Palantir. Nature biotechnology *37*, 451-460. 10.1038/s41587-019-0068-4.

23. Cheng, T., Xing, Y.-Y., Li, Y.-F., Liu, C., Huang, Y., Zhang, Y.-J., Megason, S.G., and Xu, P.-F. (2021). Single cell response landscape of graded Nodal signaling in zebrafish explants. bioRxiv. 10.1101/2021.04.25.441305.

24. Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. Nature methods *14*, 979-982. 10.1038/nmeth.4402.

25. Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nature communications *10*, 1523. 10.1038/s41467-019-09234-6.

26. Wang, X., Park, J., Susztak, K., Zhang, N.R., and Li, M. (2019). Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. Nature communications *10*, 380. 10.1038/s41467-018-08023-x.

27. Barone, V., Lang, M., Krens, S.F.G., Pradhan, S.J., Shamipour, S., Sako, K., Sikora, M., Guet, C.C., and Heisenberg, C.P. (2017). An Effective Feedback Loop between Cell-Cell Contact Duration and Morphogen Signaling Determines Cell Fate. Developmental cell *43*, 198-211.e112. 10.1016/j.devcel.2017.09.014.

28. Xia, P., Gutl, D., Zheden, V., and Heisenberg, C.P. (2019). Lateral Inhibition in Cell Specification Mediated by Mechanical Signals Modulating TAZ Activity. Cell *176*, 1379-1392 e1314. 10.1016/j.cell.2019.01.019.

29. Dimitrov, D., Turei, D., Garrido-Rodriguez, M., Burmedi, P.L., Nagai, J.S., Boys, C., Ramirez Flores, R.O., Kim, H., Szalai, B., Costa, I.G., et al. (2022). Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data. Nature communications *13*, 3224. 10.1038/s41467-022-30755-0.

30. Mathieu, J., Griffin, K., Herbomel, P., Dickmeis, T., Strahle, U., Kimelman, D., Rosa, F.M., and Peyrieras, N. (2004). Nodal and Fgf pathways interact through a positive regulatory loop and synergize to maintain mesodermal cell populations. Development (Cambridge, England) *131*, 629-641. 10.1242/dev.00964.

31. Müller, P., Rogers, K.W., Jordan, B.M., Lee, J.S., Robson, D., Ramanathan, S., and Schier, A.F. (2012). Differential diffusivity of Nodal and Lefty underlies a reaction-diffusion patterning

591 system. Science (New York, N.Y.) *336*, 721-724. 10.1126/science.1221920.
592 32. Chen, Y., and Schier, A.F. (2002). Lefty proteins are long-range inhibitors of squint-
593 mediated nodal signaling. Current biology : CB *12*, 2124-2128. 10.1016/s0960-
594 9822(02)01362-3.
595 33. Sekine, R., Shibata, T., and Ebisuya, M. (2018). Synthetic mammalian pattern formation
596 driven by differential diffusivity of Nodal and Lefty. Nature communications *9*, 5456.
597 10.1038/s41467-018-07847-x.
598 34. Kuhn, T., Landge, A.N., Morsdorf, D., Cossmann, J., Gerstenecker, J., Capek, D., Muller, P.,
599 and Gebhardt, J.C.M. (2022). Single-molecule tracking of Nodal and Lefty in live zebrafish
600 embryos supports hindered diffusion model. Nature communications *13*, 6101.
601 10.1038/s41467-022-33704-z.
602 35. Liu, L., Nemashkalo, A., Rezende, L., Jung, J.Y., Chhabra, S., Guerra, M.C., Heemskerk, I.,
603 and Warmflash, A. (2022). Nodal is a short-range morphogen with activity that spreads
604 through a relay mechanism in human gastruloids. Nature communications *13*, 497.
605 10.1038/s41467-022-28149-3.
606 36. Dubrulle, J., Jordan, B.M., Akhmetova, L., Farrell, J.A., Kim, S.H., Solnica-Krezel, L., and
607 Schier, A.F. (2015). Response to Nodal morphogen gradient is determined by the kinetics
608 of target gene induction. eLife *4*. 10.7554/eLife.05042.
609 37. Xing, C., Shen, W., Gong, B., Li, Y., Yan, L., and Meng, A. (2022). Maternal Factors and
610 Nodal Autoregulation Orchestrate Nodal Gene Expression for Embryonic Mesendoderm
611 Induction in the Zebrafish. Front Cell Dev Biol *10*, 887987. 10.3389/fcell.2022.887987.
612 38. Pinheiro, D., Kardos, R., Hannezo, E., and Heisenberg, C.P. (2022). Morphogen gradient
613 orchestrates pattern-preserving tissue morphogenesis via motility-driven unjamming.
614 Nature Physics. 10.1038/s41567-022-01787-6.
615 39. Kawamura, A., Koshida, S., Hijikata, H., Ohbayashi, A., Kondoh, H., and Takada, S. (2005).
616 Groucho-associated transcriptional repressor ripply1 is required for proper transition
617 from the presomitic mesoderm to somites. Developmental cell *9*, 735-744.
618 10.1016/j.devcel.2005.09.021.
619 40. Dixon Fox, M., and Bruce, A.E. (2009). Short- and long-range functions of Goosecoid in
620 zebrafish axis formation are independent of Chordin, Noggin 1 and Follistatin-like 1b.
621 Development (Cambridge, England) *136*, 1675-1685. 10.1242/dev.031161.
622 41. Terashima, A.V., Mudumana, S.P., and Drummond, I.A. (2014). Odd skipped related 1 is a
623 negative feedback regulator of nodal-induced endoderm development. Developmental
624 dynamics : an official publication of the American Association of Anatomists *243*, 1571-
625 1580. 10.1002/dvdy.24191.
626 42. Coda, D.M., Patel, H., Gori, I., Gaarenstroom, T.E., Song, O.R., Howell, M., and Hill, C.S.
627 (2022). A network of transcription factors governs the dynamics of NODAL/Activin
628 transcriptional responses. Journal of cell science *135*. 10.1242/jcs.259972.
629 43. Akhmetova, L. (2016). Transcriptional Regulation of Nodal Target Genes in Early Zebrafish
630 Development. Doctor (Harvard University).
631 44. Stuart, T., Srivastava, A., Madad, S., Lareau, C.A., and Satija, R. (2021). Single-cell chromatin
632 state analysis with Signac. Nature methods *18*, 1333-1341. 10.1038/s41592-021-01282-
633 5.
634 45. Shen, M.M. (2007). Nodal signaling: developmental roles and regulation. Development

(Cambridge, England) *134*, 1023-1034. 10.1242/dev.000166.

46. Takenaga, M., Fukumoto, M., and Hori, Y. (2007). Regulated Nodal signaling promotes differentiation of the definitive endoderm and mesoderm from ES cells. Journal of cell science *120*, 2078-2090. 10.1242/jcs.004127.

47. Hagos, E.G., and Dougan, S.T. (2007). Time-dependent patterning of the mesoderm and endoderm by Nodal signals in zebrafish. BMC developmental biology *7*, 22. 10.1186/1471-213X-7-22.

48. Ashe, H.L., and Briscoe, J. (2006). The interpretation of morphogen gradients. Development (Cambridge, England) *133*, 385-394. 10.1242/dev.02238.

49. Harvey, S.A., and Smith, J.C. (2009). Visualisation and quantification of morphogen gradient formation in the zebrafish. PLoS biology *7*, e1000101. 10.1371/journal.pbio.1000101.

50. Bozek, M., Cortini, R., Storti, A.E., Unnerstall, U., Gaul, U., and Gompel, N. (2019). ATAC-seq reveals regional differences in enhancer accessibility during the establishment of spatial coordinates in the Drosophila blastoderm. Genome research *29*, 771-783. 10.1101/gr.242362.118.

51. Mir, M., Reimer, A., Haines, J.E., Li, X.Y., Stadler, M., Garcia, H., Eisen, M.B., and Darzacq, X. (2017). Dense Bicoid hubs accentuate binding along the morphogen gradient. Genes & development *31*, 1784-1794. 10.1101/gad.305078.117.

52. Xiong, Y., Li, W., Shang, C., Chen, R.M., Han, P., Yang, J., Stankunas, K., Wu, B., Pan, M., Zhou, B., et al. (2013). Brg1 governs a positive feedback circuit in the hair follicle for tissue regeneration and repair. Developmental cell *25*, 169-181. 10.1016/j.devcel.2013.03.015.

53. Pattison, J.M., Melo, S.P., Piekos, S.N., Torkelson, J.L., Bashkirova, E., Mumbach, M.R., Rajasingh, C., Zhen, H.H., Li, L., Liaw, E., et al. (2018). Retinoic acid and BMP4 cooperate with p63 to alter chromatin dynamics during surface epithelial commitment. Nat Genet *50*, 1658-1665. 10.1038/s41588-018-0263-0.

54. Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nature biotechnology *36*, 411-420. 10.1038/nbt.4096.

55. Sagerstrom, C.G., Gammill, L.S., Veale, R., and Sive, H. (2005). Specification of the enveloping layer and lack of autoneuralization in zebrafish embryonic explants. Developmental dynamics : an official publication of the American Association of Anatomists *232*, 85-97. 10.1002/dvdy.20198.

56. Pijuan-Sala, B., Griffiths, J.A., Guibentif, C., Hiscock, T.W., Jawaid, W., Calero-Nieto, F.J., Mulas, C., Ibarra-Soria, X., Tyser, R.C.V., Ho, D.L.L., et al. (2019). A single-cell molecular map of mouse gastrulation and early organogenesis. Nature *566*, 490-495. 10.1038/s41586-019-0933-9.

Figures:

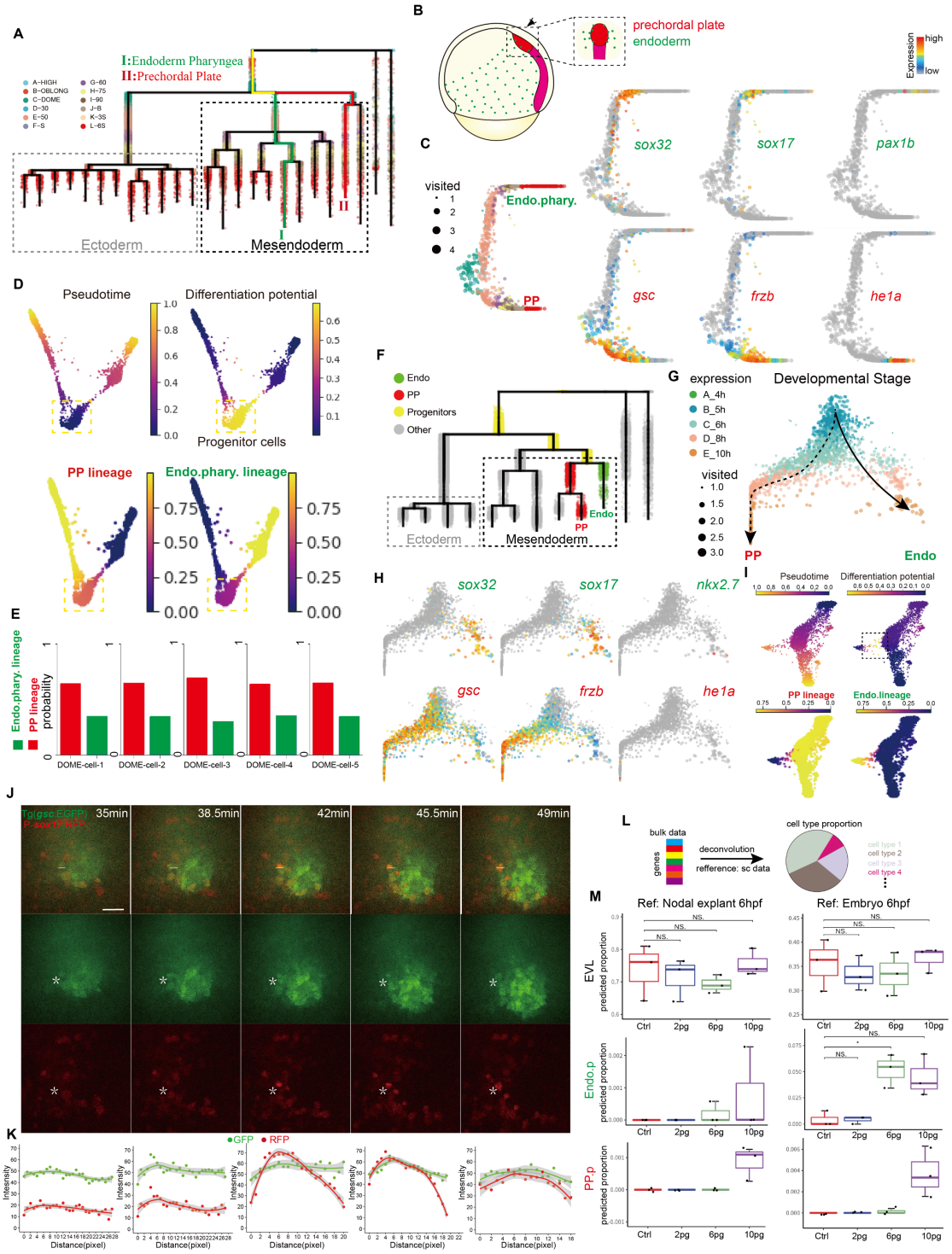

Figure 1. Single-cell trajectory analysis for studying prechordal plate and anterior endoderm specification. (A) URD differentiation tree of zebrafish embryos from blastula stage to somitogenesis stage[56]. Cells are colored according to developmental stages. Gray dotted frame indicates the ectoderm linages and black dotted frame indicates the mesendoderm

686 lineages. Green line shows the endoderm pharyngea lineage (anterior endoderm, roman number
687 I), red line shows the prechordal plate lineage (roman number II) and yellow line indicates their
688 progenitors. (B) Simplified schematics of the patterns of endoderm and axial mesoderm of
689 zebrafish embryos at 75% epiboly. Dotted frame in the right indicates the enlarged region of
690 dorsal anterior. Anterior endoderm sporadically locates near prechordal plate. (C) URD
691 branchpoint plots of the prechordal plate and anterior endoderm development from scRNA-seq
692 data of zebrafish embryo[56], showing pseudotime (x-axis) and random walk visitation preference
693 from endoderm pharyngea to prechordal plate domains (y-axis). Cells are colored by
694 developmental stages (left), expression of endodermal markers (*sox32*, *sox17* and *pax1b*; right
695 top) and prechordal plate markers (*gsc*, *frzb* and *he1a*; right bottom). (D) Force-directed layout
696 of prechordal plate and anterior endoderm cells of zebrafish embryos from blastula stage to
697 somitogenesis stage. Cells are colored by Palantir[22] pseudo-time (top left), differentiation
698 potential (top right) and branch probabilities of prechordal plate (bottom left) and anterior
699 endoderm (bottom right). (E) Branch probabilities of five cells randomly selected from
700 progenitor cells. Bars are colored by cell types as the legend shown on the top of the panel. (E)
701 Expression dynamics of prechordal plate and endoderm markers along pseudotime axis
702 inferred by Monocle[24]. The green frame and the dotted red frame indicate the points that *sox32*
703 (Also as the yellow dotted line showing) and *gsc* start to differentially express in these two
704 lineages respectively. (F) URD differentiation tree of Nodal explants from blastula stage to the
705 end of gastrulation. Gray dotted frame indicates the ectoderm linages and black dotted frame
706 indicates the mesendoderm lineages. The green, red and yellow dots indicate the endodermal
707 cells (anterior endoderm), prechordal plate cells and the progenitor cells repectively, and the
708 gray dots indicate other cells. (G-H) URD branchpoint plots of the prechordal plate and
709 endoderm development from scRNA-seq data of Nodal explants showing pseudotime (y-axis)
710 and random walk visitation preference from prechordal plate to endoderm domains (x-axis).
711 Cells are colored by developmental stages (G), expression of endodermal markers (*sox32*,
712 *sox17* and *nkx2.7*; right top) and prechordal plate markers. (*gsc*, *frzb* and *he1a*; right bottom)
713 (H) (I) Force-directed layout of prechordal plate and anterior endoderm cells of Nodal explants
714 from blastula stage to the end of gastrulation. Cells are colored by the levels of Palantir[22]

pseudo-time (top left), differentiation potential (top right) and branch probabilities of prechordal plate (bottom left) and anterior endoderm (bottom right). (J) Time series of confocal one-stack image of a representative zebrafish embryo. The prechordal plate cells are marked by *gsc* promotor drived expression of EGFP (green), and the endodermal cells are marked by *sox17* promotor drived expression of RFP (red). The white snowflakes indicate the cell transiting from *gsc* positive cell to *sox17* positive cell. (K) The intensity distribution of GFP (*gsc*) and RFP (*sox17*) channel along the line drawn in A (the white lines in the merge images in panel A). The intensity of RFP channel increases during development. Each live imaging experiment was performed on at least 3 independent replicates. Scale bar: 50 μm. (L) Schematic diagram of estimated cell type fractions in Nodal bulk RNA-seq data inferring from single-cell RNA-seq data of Nodal explant by a deconvolution method named MuSiC[26]. (M) The estimated cell proportions of EVL (top), endoderm (middle) and prechordal plate (bottom) in bulk RNA-seq data of Nodal explants injected with 0 pg, 2 pg, 6 pg and 10 pg Nodal mRNA inferring from single-cell RNA-seq data of Nodal explants (left) or embryos (right) at 6 hpf. Statistical differences between two samples were evaluated by Student's t-test. P-value < 0.05 was considered as statistically significant.

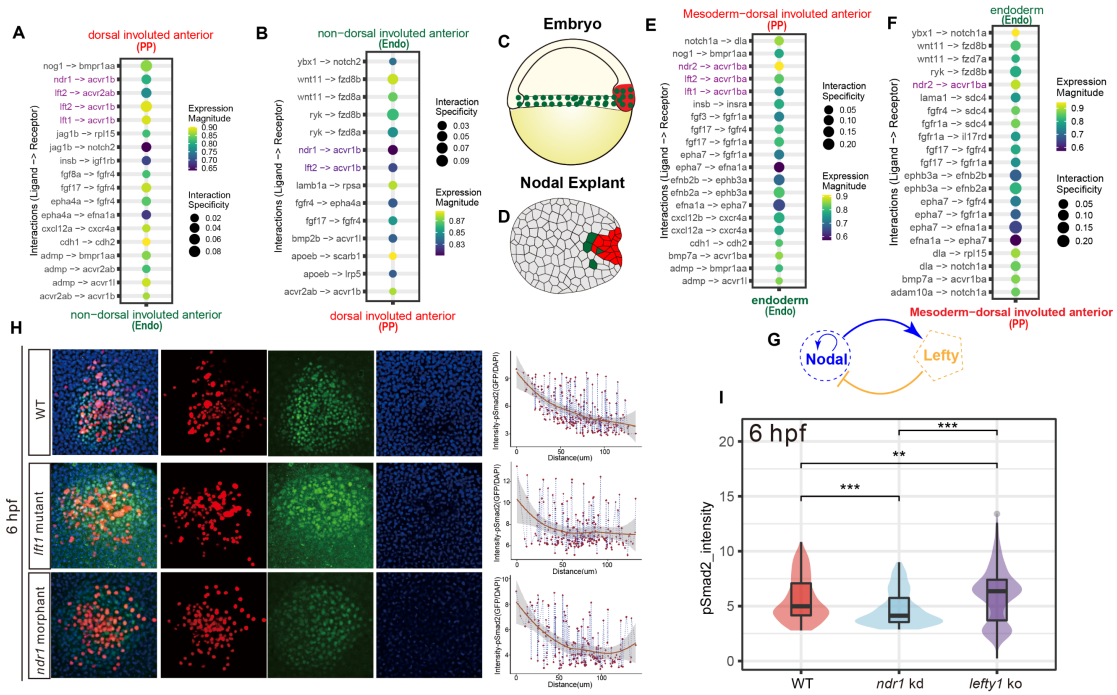

Figure 2. Cell-cell communication analyses reveal that Nodal-Lefty regulatory network is probably involed in prechordal plate and anterior endoderm separation. (A-B) Dotplots visualizing ligand-receptor interactions between prechordal plate (PP) and endoderm (endo). The interactions from prechordal plate (ligand) to endoderm (receptor) and from endoderm (ligand) to prechordal plate (receptor) are plotted respectively in A and B. (C) Schematic diagram of prechordal plate (red) and endoderm (green) in zebrafish embryos at 6 hpf. (D) Schematic diagram of prechordal plate (red) and endoderm (green) in Nodal explants at 6 hpf. (E-F) Dotplots visualizing ligand-receptor interactions between prechordal plate (PP) and endoderm (endo). The interactions from prechordal plate (ligand) to endoderm (receptor) and from endoderm (ligand) to prechordal plate (receptor) are plotted respectively in E and F. Interactions are filtered by an aggregate rank < 0.05 (A, B, E and F). (G) Simplified schematic showing the interaction network between Nodal and Lefty. (H) confocal scanned images showing pSmad2 levels in zebrafish embryonic animal pole stimulated by Nodal injection at 128-cell stage. pSmad2 immunostaining (left) and its quantifications (right) in wild-type embryos, *lft1* mutants and *ndr1* morphants at 6 hpf. (I) Comparisons of pSmad2 levels in wild-type embryos, *lft1* mutants and *ndr1* morphants at 6 hpf (as in H). Statistical differences between two samples were evaluated by Student's t-test. *indicates P-value < 0.05, **indicates P-value < 0.01 and ***indicates P-value < 0.001.

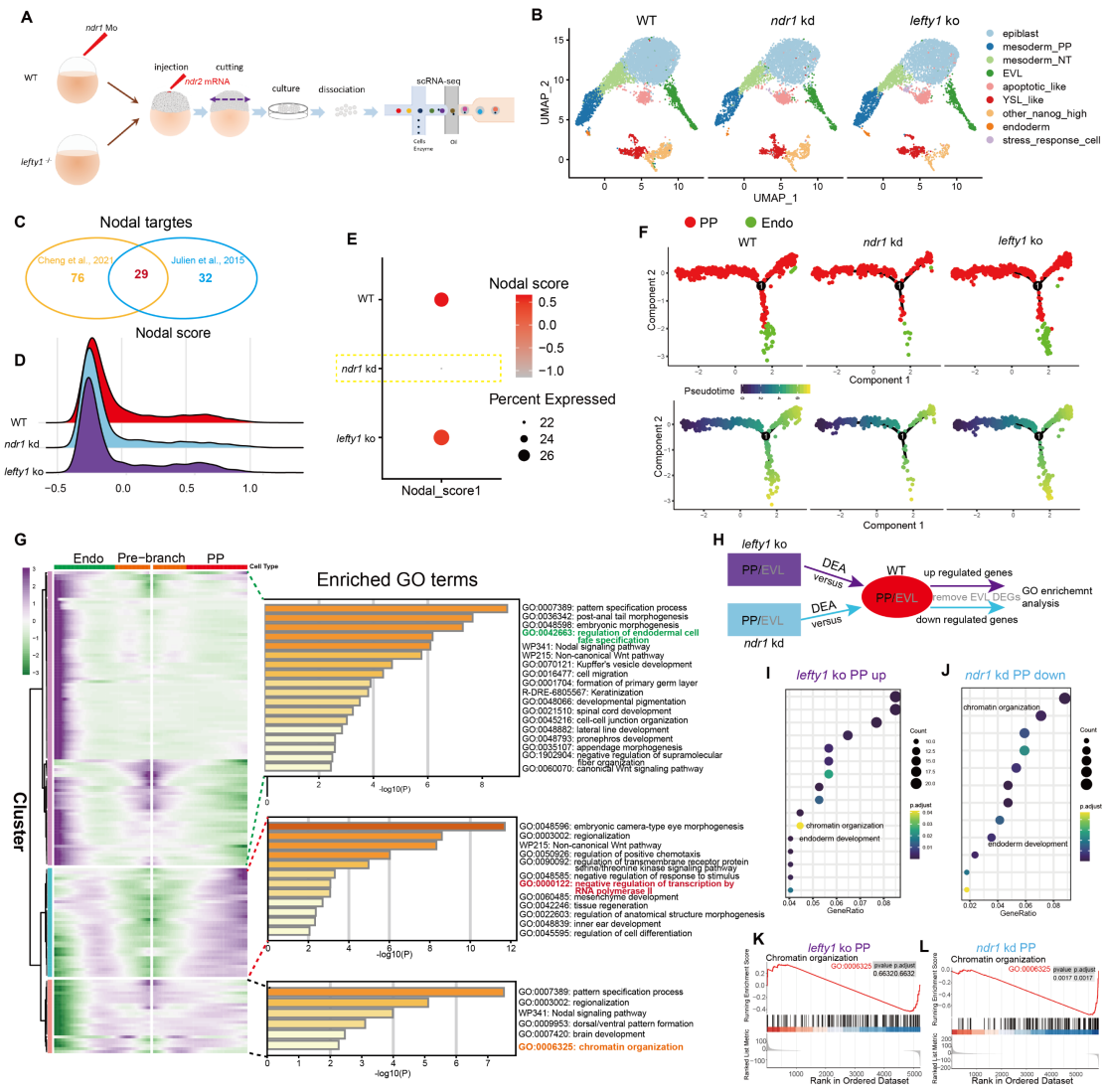

Figure 3. Reconstructing single-cell trajectory tree of prechordal plate and anterior endoderm by integrating the single-cell datasets of Nodal explants generated from wild-type, *lft1* mutants and *ndr1* morphants. (A) Schematic diagram showing the experimental workflow of single-cell RNA sequencing of Nodal explants constructed from *lft1* mutants and *ndr1* morphants at 6 hpf. (B) Overall UMAP plot of integrated single cell datasets of Nodal explants generated from wild-type (left), *ndr1* morphants (middle) and *lft1* mutants (right). (C) Venn plot showing the selection of Nodal direct target genes used for defining Nodal score (**see Methods**). (D) Ridge plot showing the Nodal score in wild-type, *ndr1*-morphant and *lft1*-mutant Nodal explants. (E) Dot plot visualizing Nodal score in wild-type, *ndr1*-morphant and *lft1*-mutant Nodal explants. Dot size indicates the percentage of cells within the cell group; the color indicates the average Nodal score levels. Nodal score is obviously decreased in *ndr1*-

765 morphant Nodal explants. (F) Pseudotime trajectory analysis of prechordal plate and anterior
766 endoderm using integrated single-cell datasets of wild-type (left), *ndr1*-morphant (middle) and
767 *lft1*-mutant (right) Nodal explants. Cells are colored by cell types (top) and pseudotime levels
768 (bottom). (G) Heatmaps representing modules of genes that co-vary along the pseudotime
769 during prechordal plate and anterior endoderm separating from common progenitors. Top GO
770 terms enriched in each class of genes are listed with their corresponding adjusted P-values.
771 Representative GO terms of each class are highlighted by different colors. (H) Diagram of gene
772 set selection for GO analysis. As a control of batch effects, EVL differential expression (DE)
773 genes are removed from mesoderm prechordal plate (PP) DE genes. (I-J) GO analysis of PP
774 mesoderm PP up-regulated genes in *lefty1* mutant (I) and down-regulated genes in *ndr1*
775 morphant (J). Redundancy of enriched GO terms are removed, and each GO term contains at
776 least 10 transcripts. Note that only the top 16 enriched GO terms identified in *lefty1* mutant are
777 plotted for a visualization purpose. (K-L) GSEA of "chromatin organization" term on PP cells
778 from *lefty1* mutant (K) and *ndr1* morphant (L) respectively. P-adjust indicates Benjamini-
779 Hochberg adjusted p-value.
780

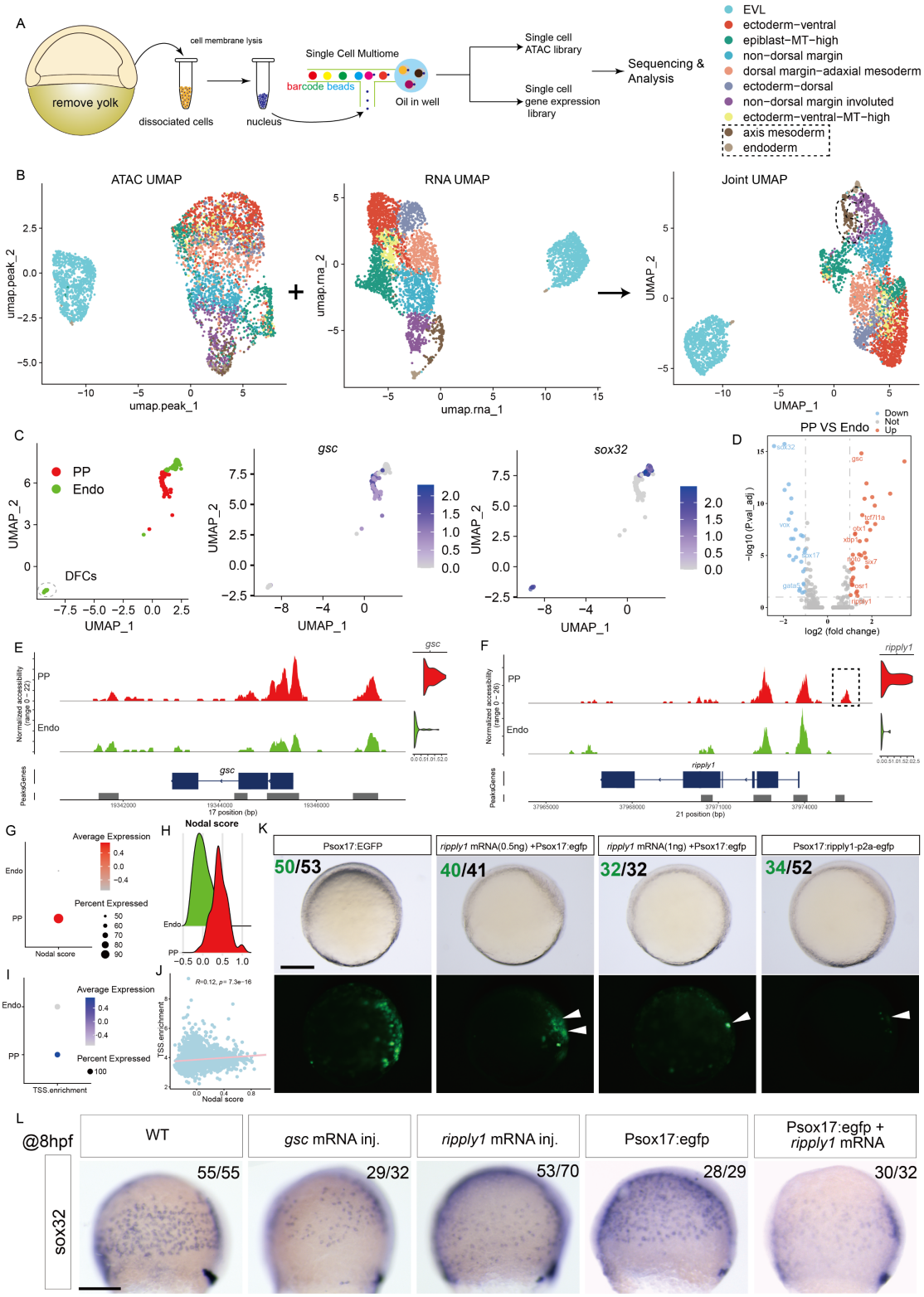

Figure 4. Investigating the mechanisms of prechordal plate and endoderm separation through integrative single-cell RNA-seq and ATAC-seq. (A) Schematic diagram showing the experimental workflow of single-cell multiomics of 6 hpf zebrafish embryos. (B) UMAPs based on scATAC-Seq (left) and scRNA-Seq (middle) datasets, and co-embedding scRNA-seq

and scATAC-seq datasets (right). Cells are colored by different cell types. (C) UMAP plot showing single-cell map of the prechordal plate and endoderm cells. Cells are colored by cell types (left) and the expression of *gsc* (middle) and *sox32* (right). (D) Volcano plot showing the differential expressed genes in prechordal plate and endoderm based on RNA levels of single-cell multiomics datasets. (E-F) Track plot showing chromatin accessibility of each cell-type on the gene locus of *gsc* (E) and *ripply1* (F). Violin plot showing the RNA levels of these two genes in prechordal plate and endoderm (right of each panel). (G-H) Dot plot (G) and ridge plot (H) showing the level of Nodal score in prechordal plate and endoderm. (I) Dot plot showing the level of TSS.enrichment (Transcription Start Site Enrichment Score, **see Methods**) in prechordal plate and endoderm. (J) Scatter plots and linear regression line for the comparison of TSS.enrichment (y-axis) and Nodal score (x-axis) across all cells. (K-L) EGFP drived by *sox17* promotor (K) and WISH of *sox32* (L) are used to indicate the endoderm cells. Both mRNA of *ripply1* and *sox17* promotor drived *ripply1* plasmid lead to the decrease of endoderm cell in zebrafish embryos. *gsc* overexpression is used as a positive control. The statistics of observed/total sampled embryos were shown in top left (K) and top right (L) of the panels. At least 3 independent replicates were used in each overexpression and ISH experiments. Scale bar: 200 μm (K and L).

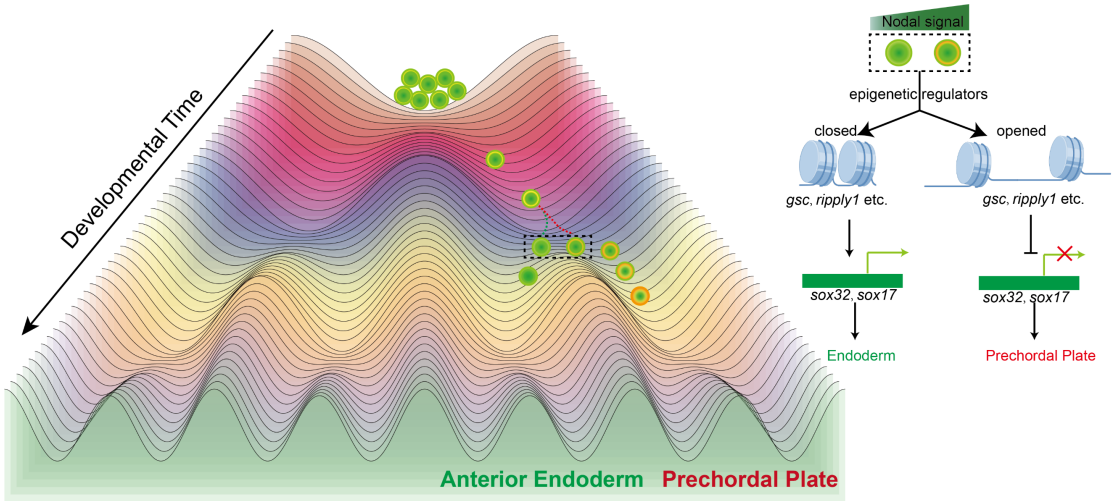

Figure 5. Model of the separation of prechordal plate and anterior endoderm. An outline depicting prechordal plate and anterior endoderm specification according to the Waddington's epigenetic landscape. Sight Nodal signaling differences drive a differential chromatin statue

between prechordal plate and anterior endoderm, and then those key transcriptional factors such as *gsc* and *ripply1* differentially expressed in these two lineages and promote cell fate separation.

Acknowledgments:

We would like to thank Dr. Alex Schier for his kind gift of the *lft1* mutant. We also thank Dr. Nai-He Jing at University of Chinese Academy of Sciences, Dr. Jun Ma, Dr. Xiao-hang Yang, Dr. Min-Xin Guan, Dr. Feng He and the members of Laboratory of Development and Organogenesis (LDO) at Zhejiang University for helpful suggestions and discussions. We thank Mr. Guang-Xu Zhang from Biostar Technology for the technical help on constructing 10x single cell multiomics sequencing library. This work was supported by grants from the Chinese National Key Research and Development Project (2022YFA1103100, 2019YFA0802402), the National Scientific Foundation of China (32050109, 31970757, 31900576).

Conflict of Interest:

The authors declare no competing interests.

Author Contributions

PFX and TC conceived and designed the research; TC, XL, CL, YYX, CL, YH and JYL performed research; TC and YD analyzed the single cell RNA-seq and single-cell multiomics data; TC, YD, HQL and PFX wrote the manuscript; all authors reviewed and approved the manuscript. Supervision and funding for the study was provided by PFX.