A temporal coordination between Nodal and Wnt signalling governs the emergence of the mammalian body plan

André Dias[1,*], Pau Pascual-Mas[1], Gabriel Torregrosa-Cortés[1,2], Harold M. McNamara[3, 4], Alexandra E. Wehmeyer[5], Sebastian J. Arnold[5,6] and Alfonso Martinez Arias[1,7,*]

1. Department of Medicine and Life Sciences, Universitat Pompeu Fabra, Spain

2. EMBL Barcelona, Spain (present address)

3. Lewis-Sigler Institute, Princeton University, USA

4. Wu Tsai Institute and Department of Molecular, Cellular, and Developmental Biology, Yale University, USA (present address)

5. Institute of Experimental and Clinical Pharmacology and Toxicology, Faculty of Medicine, University of Freiburg, Germany

6. Signalling Research Centers BIOSS and CIBSS, University of Freiburg, Germany

7. ICREA, Spain

*Correspondence to: andre.dias@upf.edu and alfonso.martineza@upf.edu

Summary

Nodal and Wnt signalling play an important role in the emergence of the mammalian body plan, primarily by orchestrating gastrulation and primitive streak formation. While the literature suggests that these pathways cooperate to build the primitive streak, their individual contributions remain poorly understood, largely due to pleiotropic effects observed when key pathway components are knocked out in mouse embryos. In this study, we used gastruloids - a non-integrated stem cell model of mammalian gastrulation - to dissect the roles of Wnt and Nodal signalling. We found that Wnt/β-catenin signalling drives a genetic program characteristic of the late primitive streak, promoting the development of posterior body structures in a time and dose-dependent manner. Conversely, Nodal signalling activates a distinct developmental module resembling the early primitive streak. By engineering gastruloids with elevated Nodal signalling, we demonstrate that the decreasing temporal gradient of Nodal activity in the mouse primitive streak likely plays a pivotal role in establishing the anterior body plan, with higher Nodal levels producing more anterior structures in a concentration-dependent manner. Our findings suggest that rather than cooperating, Nodal and Wnt signalling act antagonistically, initiating distinct developmental modules within the primitive streak. This antagonism is likely the core mechanism driving the early body plan in mammals. Additionally, our results also lead to the suggestion that the gastruloid system models the primitive streak and reveals its self-organizing potential.

Introduction

The Primitive Streak (PS) is a transient structure that emerges at the midline of the posterior epiblast in amniotic vertebrate embryos during gastrulation and gives rise to all the main organs of the embryo, except for the brain (Arnold and Robertson, 2009; Sheng et al., 2021). Fate mapping and lineage tracing studies in the mouse embryo have shown that the early PS is configured by cells located initially in the proximal-posterior part of the epiblast, that undergo an epithelial to mesenchymal transition (EMT) (Yang et al., 2020) and give rise, in an ordered sequence, to extraembryonic mesoderm, anterior mesoderm (cranial and cardiac) and endoderm (Kinder et al., 1999; Kinder et al., 2001a; Probst et al., 2021; Scheibner et al., 2021; Tam and Behringer, 1997; Tam et al., 1997). As gastrulation proceeds, some cells of the early PS persist in the anterior/distal-most part of the streak while a new pool from the anterior half of the epiblast emerges at the posterior compartment forming what is referred to as the late posterior PS (Kinder et al., 1999; Lawson et al., 1991; Tam and Behringer, 1997). In addition, while early PS cells also form the node/notochord, cells from the late PS will be incorporated into the caudal epiblast and play a role in generating post-occipital neural and mesodermal derivatives during axial extension (Aires et al., 2018; Binagui-Casas et al., 2021; Lawson et al., 1991; Tam and Behringer, 1997; Wymeersch et al., 2021). Understanding gastrulation requires understanding the regulation of these two cell fate assignment events and their interactions.

It has been known for a long time that mutations of the T-box transcription factor *Eomesodermin* (*Eomes*) result in the loss of derivatives of the early PS (Arnold et al., 2008) with *Brachyury* (*Tbxt*) mutants revealing a complementary phenotype for the late PS streak, with a loss of axial progenitors and post-occipital tissue (Herrmann, 1995; Wymeersch et al., 2021). Recent work has shown that, in addition to their role in fate specification in the PS, *Eomes* and *Tbxt* act synergistically to promote exit from pluripotency by suppressing the pro-neural ground state (Tosic et al., 2019). Interestingly, both transcription factors are co-expressed in the early PS, even though the loss of *Tbxt* shows no phenotype in this region (Herrmann, 1995; Schüle et al., 2023). Recent genetic studies have shown the reason for this: Eomes suppresses the expression of Tbxt targets, and it is the downregulation of *Eomes* at E7.5 that enables the activation of the late PS program (Schüle et al., 2023).

In the mouse, gastrulation is under the control of Nodal and Wnt/β-catenin signalling, with both being active in the posterior part of the epiblast (Arnold and Robertson, 2009; Bardot and Hadjantonakis, 2020; Ben-Haim et al., 2006; Conlon et al., 1994; Ferrer-vaquer et al., 2010; Liu et al., 1999; Lu and Robertson, 2004). However, they display opposing activity dynamics over time, with Nodal signalling becoming confined to the distal region of the epiblast within the territory that will give rise to the node (Ben-Haim et al., 2006; Ferrer-vaquer et al., 2010). While it is known that these signals interact with *Eomes* and *Tbxt* during the laying down of the body plan (Bardot and Hadjantonakis, 2020; Tam and Loebel, 2007), the nature of this relationship and its impact on the fate map that emerges during gastrulation remain open questions.

Here we use 'gastruloids', a stem cell-based model of mammalian gastrulation and early axial elongation (Beccari et al., 2018; Martinez Arias et al., 2022; Martinez Arias et al., 2024; Turner et al., 2017; van den Brink et al., 2014), to explore the role of Wnt and Nodal signalling in mammalian gastrulation. Standard, CHIR (CHIR99021) treated mouse gastruloids exhibit patterns of gene expression and organisation characteristic of

the late PS (Beccari et al., 2018; Mayran et al., 2023; Rosen et al., 2022). However, it has been reported that some of these gastruloids exhibit low and variable levels of expression of genes associated with the early PS reflected, in particular, in varying levels of endoderm (Rosen et al., 2022; Stelloo et al., 2024; van den Brink et al., 2020; Vianello and Lutolf, 2021). However, replacing the CHIR treatment for long-term exposure to CHIR and the Nodal surrogate Activin A results in a large amount of endoderm (Hashmi et al., 2022). This led us to explore the outcome of substituting the Wnt agonist CHIR for Activin A between 48 and 72 hours of gastruloid development.

We find that, in contrast to CHIR, increasing the levels of Nodal signalling using Activin A elicits early PS fates with little of the posterior PS-associated ones. This is reflected in the robust expression of *Eomes*, *Foxa2* and *Lhx1*, the absence of significant *Hox* gene expression, and the emergence of anterior mesoderm and pharyngeal endoderm. The observation that Nodal and Wnt/β-catenin drive two alternative developmental modules in the PS, provides a rationale for the distinct activity dynamics of these signalling in the mouse gastrula and suggests the existence of two functional modules: a Nodal/*Eomes*, governing anterior fates and a Wnt/*Tbxt*, driving posterior ones. Furthermore, we noticed that while high Nodal and low Wnt levels initiate the anterior mesendodermal program, the transition to low Nodal and high Wnt activates posterior fates. In this transition and interactions, we uncover a potential role for Wnt signalling in pacing the process of gastrulation and the orderly establishment of the mammalian embryonic body plan.

Results

Wnt signalling regulates the cell state transition driving the emergence of multipotency in mammals

During the first 48h of the standard mouse gastruloid protocol (Baillie-Johnson et al., 2015; van den Brink et al., 2014), cells undergo major transcriptomic changes at a bulk level as they transition from naïve to formative/primed pluripotency. These changes are associated with the downregulation of genes associated with naïve pluripotency (*Zfp42/Rex1*, *Nanog*, *Essrb* and *Tfcp2l1*) (Ghimire et al., 2018; Kim et al., 2020; Kinoshita et al., 2021; Nichols and Smith, 2009) and the acquisition of an early epiblast-like signature (Arnold and Robertson, 2009; Bardot and Hadjantonakis, 2020; Pfister et al., 2007; Tam and Behringer, 1997; Tam and Loebel, 2007) characterized by the expression of *Fgf5*, *Otx2*, *Pou5f1/Oct4* and *Utf1* (**Figure 1** and **S1**). At this stage, gastruloids resemble epiblast-like stem cells (EpiLSCs) and model the pre-streak epiblast at around embryonic day (E) 5.5/E6.0 (Arnold and Robertson, 2009; Lawson and Wilson, 2016; Tam and Behringer, 1997; Tam and Loebel, 2007), displaying significant levels of *Nodal* and limited *Eomes*, *Wnt3* and *Tbxt* expression (**Figure 1** and **S1**). Upon treatment with the standard CHIR concentration (3µM) (Baillie-Johnson et al., 2015; van den Brink et al., 2014), 72h gastruloids display, at a bulk level, an upregulation of *Tbxt*, *Cdx2*, *Fst*, *Wnt3a*, *Wnt8a* and *Epha1* (**Figure 1C**), which are genes normally associated with the emergence of the caudal epiblast at about E8.0 in the embryo (Amin et al., 2016; de Lemos et al., 2022; Wymeersch et al., 2021). At the same time, we observed a downregulation of early epiblast-related genes and low expression of *Eomes* (**Figure 1C**), suggesting that the early PS program is being suppressed and that these gastruloids represent the late PS.

To better understand the role of Wnt signalling during mammalian gastrulation, we treated mouse gastruloids with different levels of CHIR (**Figure S2**) and compared their transcriptional signatures through bulk RNA sequencing. Using principal component analysis (PCA), we observed that gastruloids treated with different CHIR concentrations cluster separately along PC1, in a concentration-dependent manner (**Figure 1B**). Importantly, we noticed that low concentrations of CHIR (0,5µM) resulted in a gene expression profile more akin to the early PS (Arnold and Robertson, 2009; Pfister et al., 2007; Pijuan-Sala et al., 2019; Tam and Behringer, 1997; Tam and Loebel, 2007) with low levels of *Tbxt*, *Evx1* and *Cdx2*, and high levels of *Eomes*, *Fgf5* and Utf1 , while a higher concentration of CHIR (20µM) led to a gene signature characteristic of the caudal lateral epiblast (CLE) (Wymeersch et al., 2021) with high *Cdx1/2*, *Nkx1-2*, *Wnt3a* and *Greb1*, and low *Nodal*, *Eomes*, *Wnt3* and *Fgf5* (**Figure 1C**). This resemblance to the caudal epiblast was confirmed through bioinformatic comparison against mouse embryo data (Pijuan-Sala et al., 2019) (**Figure 1C and S3**) and can also be observed in the expression of *Hox* genes, like *Hoxb9*, that are prematurely expressed relative to standard gastruloids (**Figure 1C**). As controls for the variations in CHIR concentrations, we analysed gastruloids treated only with the correspondent levels of the CHIR vehicle dimethyl sulfoxide (DMSO). These gastruloids clustered together in the PCA and displayed an epiblast-like signature more similar to that of 48h gastruloids, independent of the DMSO concentration (**Figure 1**).

These results indicate that Wnt signalling drives the emergence of a late PS/caudal epiblast genetic program in a dose-dependent manner and suggest two possible interpretations. One that, in agreement with standard views of Wnt signalling and gene expression, different levels of Wnt/β-catenin elicit distinct transcriptomic profiles. A second one would suggest that the CLE emerges within a schedule in gene expression whose tempo is modulated by Wnt signalling: the higher the level of Wnt signalling, the faster the progress through the schedule. To distinguish between these possibilities, we treated gastruloids with varying levels of CHIR and collected them at 60h for bulk RNAseq. PCA analysis indicated that, at this time, standard gastruloids (3µM CHIR) cluster closer to 0,5µM CHIR-treated gastruloids at 72h, whereas 20µM CHIR gastruloids cluster near standard gastruloids at 72h (**Figure 1**). This observation favours the second interpretation. Further evidence for this is provided by Hybridization Chain Reaction (HCR) of different epiblast markers showing how different concentrations of CHIR control their expression in mouse gastruloids and how that matches with what happens in the embryo (**Figure S4**).

An observation from the transcriptomic comparison of gastruloids treated with different CHIR levels is the possible relationship between increasing levels of Wnt signalling and early signs of EMT (Yang et al., 2020): downregulation of *Cdh1* and increased *Cdh2* and *Snai1* expression (**Figure S5**). However, this does not correlate with gradual changes in the expression of early mesodermal markers, such as *Tbx6*, *Mesp1* or *Tcf15* (**Figure S5**) which might be expected from a classical cell fate decision model. Given the substantial reduction in pluripotency-associated genes in gastruloids exposed to high Wnt levels (e.g. *Oct4* and *Utf1*) (**Figure 1**), our data suggest that Wnt signalling works primarily to control the pace of a state transition leading to multipotency rather than directly acting on fate specification. Importantly, this finding also implies that additional factors likely regulate the early PS and the specification of anterior body structures.

Gastruloids engineered with Activin A model the early primitive streak

Having established that Wnt/β-catenin triggers the emergence of a late PS program and given the key roles that Nodal signalling plays in gastrulation (Brennan J. et al., 2001; Conlon et al., 1994; Lu and Robertson, 2004; Schier, 2003; Shen, 2007), we wondered what would happen if we substituted Wnt signalling for a pulse of Nodal signalling in gastruloids. To do this, we added Activin A instead of CHI between 48 and 72h (**Figure 2A**). An initial comparison of the gene expression patterns in these gastruloids through HCR revealed a robust expression of both *Eomes* and *Tbxt* at 72h, which is consistent with the activation of the early PS program. By 96h, we could detect in the periphery of these gastruloids, the expression of genes related to endoderm (*Foxa2* and *Sox17*) and anterior mesoderm (*Mesp1* and *Tbx6*), which are fates known to arise from the early PS. In addition, we detected node/notochord-like cells (*Noto* expression) in these gastruloids (**Figure 2F and S6**). These patterns of gene expression contrasts with those in CHIR-treated gastruloids that only express *Tbxt* at 72h and do not contain signs of notochord or endodermal-like cells at 96h (**Figure 2F**).

These results suggest that, in gastruloids, Nodal signalling might trigger the early PS program. To further investigate the role of Nodal signalling in the PS and understand how it drives divergence in gastruloid development compared to the pulse of Wnt signalling, we treated gastruloids with Activin A or CHIR and conducted a temporal comparative bulk transcriptomic analysis. PCA analysis indicates that Activin A treated gastruloids at 72h are more similar to gastruloids made with low CHIR (0.5µM) than to those treated with the standard CHIR dose (3µM) (**Figure 2B**). This similarity was evident when examining key epiblast marker genes: Activin A treated gastruloids express high levels of both *Fgf5* and *Otx2*, low CHIR treated gastruloids exhibit intermediate levels of these genes and 3µM CHIR treated gastruloids express only *Fgf5* (**Figure 2C**). Interestingly, these changes in the patterns of gene expression from high Nodal activity to increasing levels of Wnt signalling in gastruloids mirror the temporal progression of PS development during gastrulation. Gastruloids treated with Activin A display an early PS-like signature (*Eomes*, *Foxa2* and *Tbxt* but no *Cdx2* or *Wnt3a*), whereas standard CHIR gastruloids mimic the late PS (no *Eomes* and *Foxa2*, but expression of *Tbxt*, *Cdx2* and *Wnt3a*) and low CHIR elicits an intermediate profile (low levels of *Eomes* and *Cdx2*, moderate *Tbxt* and absence of *Foxa2* and *Wnt3a)* (**Figure 2C**). These results indicate that the Nodal pulse mediated by Activin A triggers an early PS-like program in gastruloids that, although distinct from the standard CHIR one, shares similarities with that triggered by a low CHIR treatment.

The temporal differences between the two PS programs, as represented in Activin A and CHIR gastruloids were further confirmed through Euclidean distance analysis: 72h Activin A treated gastruloids are more similar to E6.75 than 72h standard CHIR gastruloids, which closely resemble E7.25 mouse embryos (**Figure 2E**). At later stages, PCA indicates a divergence in the development of the two types of gastruloids (**Figure 2B**). At 96 and 120h, Activin A treated gastruloids express high levels of genes related to the specification and development of endoderm e.g. *Foxa2*, *Cdh1*, *Epcam*, *Sox17* and *Cpm* (Arnold and Robertson, 2009; Bardot and Hadjantonakis, 2020; Han et al., 2020; Scheibner et al., 2021; Tam and Loebel, 2007), node/notochord e.g. *Noto*, *Gsc*, *Foxa2* and *Shh* (Arnold and Robertson, 2009; Kinder et al., 2001b; Tam and Loebel, 2007; Wymeersch et al., 2019), haematoendothelium e.g. *Kdr*, *Cd44*, *Flt1* and *Etv2* (Kinder et al., 2001a; Lotto et al., 2020; Rossi et al., 2022) and anterior mesoderm e.g. *Tbx1*, *Gata6*

and *Isl1* (Arnold and Robertson, 2009; Nandkishore et al., 2018; Probst et al., 2021; Tam and Loebel, 2007; Vyas et al., 2020) (**Figure 2** and **S7**). This anterior bias of Activin A treated gastruloids follows a highly anterior pattern of *Hox* gene expression (**Figure S8**) and contrasts with the previously reported posterior character of gastruloids engineered with CHIR (Beccari et al., 2018; Lopez-Delisle et al., 2024; Rekaik et al., 2023), which develop essentially posterior neural and mesodermal-like structures (**Figure S7**). Concomitantly, at a phenotypical level, we observed that 120h Activin A treated gastruloids exhibit limited elongation characteristic of the gastruloids developed using CHIR (**Figure S7**) which would be expected in a representation of the anterior domain of the body plan. Together, these results indicate that distinct signalling activities play a key role in regulating the early and late PS and that a temporal progression from Nodal to Wnt signalling activity likely coordinates the emergence of the different mouse body structures.

A temporal gradient of Nodal signalling activity controls the formation of anterior body structures in mammals

In our experimental conditions, while CHIR gastruloids stop to progress around 120h, Activin A treated gastruloids develop at least until 144h. At this stage, these gastruloids continue to exhibit a round shape, with only about 50% showing some signs of elongation (**Figure 3**). Interestingly, reducing the amount of Activin A (from 100ng/ml to 25ng/ml) resulted in gastruloids displaying a significant elongation and a less pronounced anterior compartment (**Figure 4**). Gene expression analysis revealed that genes highly expressed throughout the gastruloids treated with high Activin A were localised to a small, anterior region in gastruloids treated with low Activin A (e.g. *Kdr* and *Otx2*) (**Figure 4D**). Also, genes related to some anterior fates e.g. *Nkx2-5* – cardiac mesoderm, were absent in low Activin A gastruloids (**Figure 4D**). In particular, we observed that in contrast to gastruloids treated with high levels of Activin A, those treated with low levels developed a notochord-like structure with a tip that is positive for both *Tbxt* and *Noto*, indicating the presence of notochord progenitors (**Figure 4D**). These results suggest that a low Activin A treatment in gastruloids possibly results in the formation of more posterior body structures in comparison to high Activin A.

To investigate the importance of the levels of Nodal signalling in PS and further clarify the anterior-posterior (AP) differences in body axis formation between gastruloids treated with high and low levels of Activin A, we performed bulk RNAseq at multiple time points. Unlike gastruloids exposed to different levels of CHIR, gastruloids treated with high or low concentrations of Activin A did not exhibit many significant differences at 72h (**Figure S9**). While we observed slight changes in the levels of *Tbxt*, *Nodal*, *Eomes*, *Pitx2*, *Fgf8* and *Tdgf1*/*Cripto*, there were no significant changes in the expression of Nodal targets like *Lefty1/2* (**Figure S9**). The main difference we observed concerns *Foxa2*, a critical mesendodermal transcription factor (Ang and Rossant, 1994; Burtscher and Lickert, 2009), whose expression was upregulated in gastruloids with high levels of Activin A (**Figure S9**). At 96h, we noticed differences in both the activity of Nodal signalling (higher levels of Activin A resulted in more expression of *Cer1* and *Lefty1/2*) and the expression of key PS-related genes between both types of Activin A treated gastruloids (**Figure S9**). On the one hand, gastruloids treated with high levels of Activin A exhibit stronger expression of genes associated with the early PS e.g. *Lhx1*, *Cripto*,

Foxa2 and *Gsc*. On the other hand, gastruloids treated with lower levels display a significant expression of genes associated with the late PS to CLE transition e.g. *Wnt3a*, *Wnt8a*, *Cdx1/2* and *Nkx1-2* (**Figure S9**). Importantly, these differences seem to impact fate specification as we noticed differences in the expression of genes associated with, for instance, extra-embryonic mesoderm(Fong et al., 1996) e.g. *Flt1* and endoderm e.g. *Sox17* (**Figure S9**).

A thorough transcriptomic evaluation between the two types of Activin A gastruloids at 144h indicates differences not only in the amount of anterior tissues but also in its nature. For instance, higher levels of *Pax9* and *Sox17*, together with *Foxa2*, suggest the presence of pharyngeal endoderm (Phillips et al., 2019; Stothard et al., 2020) in gastruloids treated with high levels of Activin A but not in those treated with low levels of this Nodal surrogate. The endothelial gene expression signature in high Activin A gastruloids is more complete, being positive for *Kdr*, *Esam*, *Igf1*, *Cdh5* and *Etv2*, in contrast to only significant levels of *Kdr* being detected in low Activin A treated gastruloids (**Figure 4A**). Also, higher levels of *Tbx1*, *Pitx2* and *Isl1* are consistent with the development of more anterior mesoderm-like tissues in gastruloids treated with a higher dose of Activin A. The higher expression of *Hesx1*, *Dmbx1* and *Cyp26a1* in these gastruloids, in contrast to higher levels of *Raldh2*, *Pax6* and *Foxb1* in low Activin A gastruloids (**Figure 4A**), further suggests the development of distinct AP body structures in a concentration-dependent manner, with low Activin A treated gastruloids developing slightly more posterior ones. This is supported by a more posterior Hox gene expression (e.g. positive for *Hoxb9*) in gastruloids treated with low levels of Activin A (**Figure 4A** and **S8**).

While experiments at single-cell resolution are still needed to characterise the full diversity of cell types, particularly at later stages, two complementary bioinformatic analyses between the two types of Activin A gastruloids and embryonic data obtained from different regions of E8.5 and E8.75 mouse embryos (Imaz-Rosshandler et al., 2024) further supported our conclusions. Activin A treated gastruloids more closely resemble anterior regions of the body axis while low Activin A-treated gastruloids contain tissues spanning broader AP body axis regions (**Figure 4**). These results suggest two key points: 1) a temporal decreasing gradient of Nodal signalling in the PS is essential for establishing the full AP diversity of anterior mesendodermal tissues and 2) there is an inverse relationship between Nodal and Wnt signalling in gastruloids. To better understand the last one, we evaluated both Wnt and Nodal reporter activity in gastruloids treated with different doses of Activin A (**Figure S10**). At 72h, high levels of Activin A elicit a robust Nodal signalling response with low Wnt signalling that, in most cases, disappears with time. In contrast, low levels of Activin A resulted in higher levels of Wnt signalling at 72h. By 120h the majority of these gastruloids treated with low Activin A exhibited high levels of Wnt signalling at the posterior part and some Nodal activity in the anterior compartment.

These observations provide evidence for how different levels of Nodal signalling control Wnt activity and indeed suggest the existence of an antagonistic interaction between Nodal and Wnt signalling in the PS. Overall, this study provides mechanistic insights into how PS activity is modulated by the temporal dynamics of Nodal and Wnt signalling (**Figure 5**), delineating how these pathways coordinate early body plan formation, and offers a strategy to engineer gastruloids modelling distinct developmental modules of the mammalian PS.

Discussion

Current views of gastrulation in mammals suggest that it is a continuous process driven by a cooperation between Nodal and Wnt signalling, triggered by BMP signalling (Bardot and Hadjantonakis, 2020). However, recent work has revealed the existence of two distinct domains that depend on two T-box transcription factors that divide the PS into defined spatiotemporal compartments: one, dependent on *Eomes* that establishes fates in the anterior PS and a later on, governed by *Tbxt* that promotes posterior fates (Gogolou et al., 2022; Schüle et al., 2023; Wehmeyer et al., 2022). Together these two modules span the length of the body plan, except for the brain, and can be separated from each other suggesting that they are two independent modules (Dias and Martinez Arias, 2023; Schüle et al., 2023; Wehmeyer et al., 2022). Here using gastruloids as a model system, we have shown that each of these transcriptional modules is mainly controlled by a specific signal: Nodal for *Eomes* and Wnt for *Tbxt*.

There are two important features of gastrulation. One is the spatiotemporally ordered cell fate assignments that proceed from anterior to posterior. The second one is the need for the transitions between the distinct fates to occur in a smooth, coordinated manner. While reciprocal transcriptional regulations between the *Eomes* and *Tbxt* modules may contribute to these features, we surmise that it is changes in the Nodal/Wnt signalling landscape that provide the major driver in the order and tempo of the transitions. This could be implemented either through the temporal integration of different relative levels of these signal, as has been shown in other systems (e.g. neural tube (Sagner and Briscoe, 2019)) or by interactions with other signalling systems active during gastrulation (e.g. FGF and/or BMP).

Extrapolating our results in gastruloids to embryos suggests that, during gastrulation the PS progresses from a landscape of high Nodal/low Wnt signalling to one with high Wnt/β-catenin and low or no Nodal signalling (**Figure 5**). These changes are likely to be implemented through repressive interactions between the two signalling pathways that control the sequence and tempo of the fate assignments. Consistent with this model, it has been shown that early spontaneous Nodal signalling predicts future anterior cell fates, and that increasing Nodal activity anteriorizes gastruloids by suppressing Wnt signalling (McNamara et al., 2024).

These molecular interactions complement the cross-regulatory interactions that have been described between *Eomes* and *Tbxt*. Recent work has started to uncover the logic of these interactions revealing that Eomes represses Tbxt activity by restricting chromatin binding and chromatin accessibility at Tbxt-bound sites (Schüle et al., 2023). In our work, we observed that while low levels of Wnt signalling amplify the early PS program involving *Eomes*, *Pitx2* and *Cripto*, these genes are downregulated in a Wnt dose and time-dependent manner. This suggests that such downregulation is likely to be indirect and possibly achieved by a combination of Wnt targets like the *Cdx* genes. Understanding these interactions will require more work at both the transcriptomic and epigenetic levels but that will be important to investigate how the transition between the two modules occurs in the embryo. In that regard, another possibility is that Nodal activity is suppressed by a negative feedback regulation via *Lefty* genes, and that a brief cooperative phase between Nodal and Wnt around E7.0-E7.5 occurs to facilitate a smooth and gradual transition to the posterior module. Although the spatiotemporal dynamics of Nodal and Wnt activity support this hypothesis, further work will benefit

from a systems biology approach and the incorporation of computational modelling (Torregrosa and Garcia-Ojalvo, 2021) would be critical to deciphering the integration of these two modules and their complex GRNs in a manner that explains how the PS is formed in the embryo.

Wnt signalling controls the pace of gastrulation

The role of Wnt signalling during gastrulation has sparked various ideas about the relationship between time and space in cell fate assignments. The two prevailing hypotheses suggest that Wnt signalling drives mesoderm specification or acts as a posteriorizing agent (Niehrs, 2021; Yamaguchi, 2001). Our findings support the latter hypothesis. However, in addition to a spatial gradient as proposed from work in *Xenopus* (Niehrs, 2022) our results suggest that it also works as a timer.

Although Wnt/β-catenin is known to contribute to mesoderm specification, especially in combination with Nodal signalling (Arnold and Robertson, 2009; Bardot and Hadjantonakis, 2020; Yamaguchi, 2001; Yoney et al., 2018), our work suggests that its primary role in mammals lies in controlling the timing of the emergence of axial progenitors. We observed that activating Wnt signalling in an aggregate of mouse formative-to-primed epiblast cells in vitro (mimicking the pre-streak epiblast around E5.5/6.0) triggered an EMT programme of gene expression but did not induce anterior mesoderm formation. Instead, it initiated a developmental program resembling the late PS, with cells adopting an early caudal epiblast-like signature - characterised by the absence of *Nanog* and low *Oct4*, and elevated *Cdx1/2*, along with *Greb1* and *Hox* genes - in a concentration/time-dependent manner, providing a mechanism for how the Wnt posteriorizing function is achieved in mammals. In agreement, other studies also reported a similar EMT associated event (marked by a gradual downregulation of *Cdh1* and concomitant upregulation of *Cdh2*) associated with the emergence of the Caudal Lateral Epiblast and the processes of axial elongation (Dias et al., 2020; Guibentif et al., 2021; Guillot et al., 2021). This EMT programme is functionally different from that required for mesoderm specification and is needed for tailbud formation (Dias et al., 2020).

From a holistic perspective, our results show that, in gastruloids, the activation of the posterior PS module by Wnt signalling can bypass the two previous sequential modules that generate brain and anterior endodermal and mesodermal gene expression programmes. It does so by accelerating the transition between the two. For example, an early/strong Wnt activation could limit the unfolding of the first two modules, by accelerating the posteriorization of the epiblast. Given the activation of some posterior *Hox* genes in gastruloids treated with high levels of CHIR, such posteriorization of the epiblast may be achieved through an acceleration of the Hox timer, whose activation depends on Wnt signalling (Deschamps and Duboule, 2017; Neijts et al., 2016). It may be that, in addition to its role as an initiator of the Hox timer, Wnt also controls its progress, perhaps by acting on the intrinsic CTCF-dependent mechanism (Rekaik et al., 2023). In agreement with this possibility, recent work has shown that Wnt signalling influences the binding patterns of CTCF, altering global chromatin structure (Nordin et al., 2024). Further work is needed to explore these epigenetic processes and how changes in Wnt signalling activity dynamics could be related to differences, for instance, in cranial mesoderm and neck somite numbers across vertebrates (Mallo, 2021; Narita and Kuratani, 2005; Vyas et al., 2020). Understanding the influence of Wnt signalling on

these variations could also shed light on the evolutionary significance of the Wnt-mediated posteriorization in shaping the animal body plan. A general role for Wnt signalling in the modulation and coordination of cell fate decisions across cell populations through a β-catenin effect on chromatin structure has been proposed before (Arias and Hayward, 2006).

Nodal, a mesendodermal specification factor

In contrast to the response of cells in the gastruloid to varying concentrations of Wnt/β-catenin, changing the levels of Activin A, as a surrogate for Nodal signalling, affects fates within the anterior modules rather than the timing of events. While high levels of Activin A elicit the most anterior fates (endoderm, cardiac and cranial mesoderm), low levels displayed a broader anterior-posterior distribution of fates including some posterior ones (e.g. posterior part of the notochord, with notochord progenitors) and the expression of some posterior *Hox* genes. This more posterior *Hox* gene signature in gastruloids treated with low levels of Activin A is probably due to the higher levels of intrinsic Wnt signalling in comparison to gastruloids treated with a high dose of Activin A. In contrast to gastruloids treated with different CHIR concentrations, varying doses of Activin A between 48 and 72h did not result in *Hox* gene expression, even when higher levels were used. This supports the idea that Wnt/β-catenin (and not Nodal signalling) is the activator of *Hox* genes (Rekaik et al., 2023).

Our work raises two possibilities for how Nodal signalling may modulate anterior fate decisions in mammals. On the one hand, Foxa2 - a key mesendodermal factor - was upregulated at 72h when higher levels of Activin A were used. On the other hand, higher levels of Nodal activity were detected in 96 and 120h gastruloids treated with high levels of Activin A. The later suggests that sustained Nodal activity is important for the differentiation of anterior body structures, particularly endoderm, as highlighted in (Hashmi et al., 2022) and work on amphibian animal cap explants treated with Activin A (reviewed in (Asashima and Satou-Kobayashi, 2024)). Another possibility could be epigenetic priming, with higher Activin A doses inducing greater chromatin accessibility at anterior gene loci, enhancing anterior fate specification. Regardless, the findings presented in this work highlight that while Wnt signalling governs the temporal activation of the posterior developmental module, Nodal signalling plays a different role and drives endodermal and mesodermal specification in a concentration and possibly time-dependent manner. A review on some dose-dependent effects of Nodal/Smad signalling during mouse gastrulation is provided in (Robertson, 2014).Understanding the molecular mechanism underlying mesendodermal specification will require more work at both the transcriptome and epigenome levels.

Gastruloids represent the primitive streak

Gastruloids are a versatile tool to explore various features of mammalian development (Turner and Martinez Arias, 2024). The robustness of the system and their experimental ease of use and reproducibility (Merle et al., 2024; Turner et al., 2017; van den Brink et al., 2014) allow for their application in various topics: bioengineering (Girgin et al., 2021a), symmetry breaking and self-organisation (Anlaş et al., 2024; Cermola et al., 2022; Gsell et al., 2024; Mayran et al., 2023; McNamara et al., 2024; Sagy et al., 2019; Suppinger et al., 2023; Turner et al., 2017; van den Brink et al., 2014), lineage (Bolondi

et al., 2024), epigenetics (Zeller et al., 2024), signalling (McNamara et al., 2024; Turner et al., 2017; Underhill and Toettcher, 2023), proteomics (Garge et al., 2024; Stelloo et al., 2024), metabolism (Dingare et al., 2024; Luque et al., 2023; Stapornwongkul et al., 2023), toxicology screening (Mantziou et al., 2021) and even do more theoretical work (Brückner and Tkačik, 2024). Significantly, gastruloids are also being used to study specific developmental events like regulation of *Hox* gene expression (Lopez-Delisle et al., 2024; Rekaik et al., 2023), somitogenesis (Hamazaki et al., 2024; van den Brink et al., 2020; Veenvliet et al., 2020; Yamanaka et al., 2023) and endoderm specification (Farag et al., 2024; Hashmi et al., 2022; Vianello and Lutolf, 2021). However, an important element of these studies is the interpretation of the model (Turner and Martinez Arias, 2024). In the case of the gastruloids an important question is their relationship to the embryo.

It is now clear that while CHIR driven standard gastruloids represent the posterior part of the embryo, with some elements from the anterior region (e.g. endoderm) emerging sporadically in different experiments (Argiro et al., 2024; Beccari et al., 2018; Farag et al., 2024; Hashmi et al., 2022; Mayran et al., 2023; Rosen et al., 2022; Rossi et al., 2021; Rossi et al., 2022; van den Brink et al., 2014; Vianello and Lutolf, 2021). The results presented here support the assignment of the classic gastruloids to the posterior region of the embryo and, together with available transcriptional data, suggests that they correspond to the CLE. Our experiments also suggest that the appearance of anterior fates depends on the levels of Nodal signalling at the time of CHIR addition which will depend on a number of variables such as cell line, growth of the starting cell population (Serum/LIF vs 2i/LIF), culture conditions of the aggregate and quality of the reagents (Turner and Martinez Arias, 2024). We suggest that there is an intrinsic program of gene expression that cells will undergo at the exit of pluripotency, that is initiated by Nodal and that occurs in gastruloids around 48 to 72h. The degree to which this program is implemented will depend on the levels of Nodal signalling accrued by the cells around 48h.

Following this suggestion, our results show that at 48h after seeding, cells in the aggregate have exited pluripotency and begin implementing the gastrulation program with the expression of *Nodal*, *Wnt3*, *Eomes* and associated early PS genes. This happens in a variable and unstable manner. Exposure to 3µM CHIR suppresses this early PS program, launches the aggregate into a state akin to E7.25/7.5 and promotes axial specification and elongation. This pattern of gene expression, together with analysis of the morphogenetic movements (Fiuza et al., 2024) leads us to surmise that the elongation observed in CHIR treated gastruloids between 96h and 120h corresponds to the convergence-extension events happening in the embryo between E7.75 and E8.5 during the PS regression. Lowering the CHIR levels partially releases the suppression of the anterior program and leads to the expression of some early PS-related genes. However, this program is clearly driven by Nodal as shown in the Activin A gastruloids and requires low levels of Wnt signalling ((Barrow et al., 2007) and unpublished data from the Martinez Arias' and Arnold's laboratories). The observation that low levels of Wnt signalling are able to maintain an anterior PS programme of gene expression suggests that the transition from Wnt3 to Wnt3a that is observed during gastrulation (Takemoto et al., 2016) might reflect different levels of Wnt/β-catenin signalling elicited by these ligands.

Overall, at 48h mouse gastruloids resemble the early epiblast of the mouse gastrula. In the embryo, exposure of the posterior epiblast to Wnt and Nodal signalling triggers the emergence of the PS as the multicellular engine of cell fate assignments during gastrulation (Tam et al., 2006). In the PS, cells undergo an EMT before becoming assigned different fates. Although gastruloids represent cell fate assignments during gastrulation, they lack a proper PS or related structure. We surmise that this is because gastruloids represent/model the PS, reveal a self-organized ability of its component - cells - and support the notion of two transcriptional networks (*Eomes* and *Tbxt*) that gastruloids reveal as independent genetic elements. Our work (and see also (Wehmeyer and Arnold)) suggests that an important role of Nodal and Wnt signalling during gastrulation is to govern order, sequence and pace the programmes of cell fate assignments. In this process, the levels of Wnt signalling appear to act as a pacemaker for the process. Consistent with our view, in the absence of both *Eomes* and *Tbxt* (Wehmeyer and Arnold) or of Wnt signalling (Girgin et al., 2021b) gastruloids will develop anterior neural tissues.

Our proposal for the nature of gastruloids suggests that a key element of the emergence of the body plan is the orderly temporal transition through a landscape of varying relative levels of Nodal and Wnt signalling that modulate the activity of *Eomes* and *Tbxt*. Gastruloids might be a useful tool to unravel the way these modules interact to lay down the body plan. Furthermore, given the ability to grow human gastruloids (Moris et al., 2020), it will be of interest to explore whether these modules and their interactions are conserved during human development.

Acknowledgements

We would like to thank Joshua Frenster for providing Epiblast-like stem cells, Anna Bigas (IMIM, Spain) for the E14Tg2A cell line, Tina Balayo for management and cell culture assistance, Raquel Flores Peirats for management assistance, Veronika Mantziou and Tina Balayo for optimization of the HCR protocol, the laboratories of Jordi Garcia Ojalvo (UPF, Spain) and Jared Toettcher (Princeton, USA) for discussions and past and present members of the Martinez Arias' lab, particularly Ulla-Maj Fiuza and Sara Bonavia, for useful discussions during the course of this project. We would also like to thank the IGC Genomics Unit and the CRG Microscope facility. This work has been supported by an EMBO Postdoctoral Fellowship (ALTF 948-2022) to AD, a Generalitat de Catalunya AGAUR Grant (2021 SGR 00175) to AD and GTC, a PhD grant (FPU18/05091) from the Spanish Ministry of Science, Innovation and Universities to GTC, the Lewis-Sigler Scholars program to HMM, the German Research Foundation (DFG) project grants AR 732/3-1, AR 732/2-1, CRC 992 (project ID 192904750) and Germany's Excellence Strategy grant (CIBSS – EXC-2189 – project ID 390939984) and the Ministry of Science and Education of the State of Baden Württemberg to SJA, and an ERC Advanced Grant (834580) to AMA.

Contributions

AD and AMA conceived the project, designed the experiments and analysed the results; AD performed the gastruloid work and RNA extractions; PPM, GTC and AD did the bioinformatics; HMM provided the Wnt/Nodal double reporter results. AD wrote the original draft; AMA reviewed and edited the manuscript; HMM, PPM and GTC commented on the manuscript; SJA and AW shared unpublished data and provided essential input for the interpretation of the results.

Material and Methods

Cell lines and culture

In this work we used wildtype *E14Tg2A* mouse ESCs obtained from the lab of Anna Bigas at the Hospital del Mar Research Institute (Spain), the TCF/Lef reporter cell line "*TLC-mCherry*"(Faunes et al., 2013; Ferrer-vaquer et al., 2010) and a Wnt and Nodal double reporter cell line (McNamara et al., 2024). The Epiblast-like stem cells originate from the *E14Tg2A* mouse ESCs.

The *E14Tg2A* and *TLC-mCherry* cell lines were cultured in ESLIF media (10% fetal bovine serum; ThermoFisher, 10270106), on 0.1% gelatine in 1X Phosphate Buffered Saline (PBS; Biowest, X0520) coated 6-well plates (Greiner, 657160) as described in (Baillie-Johnson et al., 2015). L-Glutamine (ThermoFisher, 25030024) was used instead of Glutamax. These cells were cultured in a humidified incubator at 37°C with 5% CO2 and passaged every three days, with a seeding density of approximately $6{,}5 \times 10^4$ cells per well. The double reporter cell line was cultured in agreement with (McNamara et al., 2024) but ESLIF was used instead of 2iLIF. All cell lines were maintained in culture for at least three passages post-thawing before experimental use and were routinely tested for mycoplasma.

Gastruloid engineering

Standard (CHIR) gastruloids were generated in accordance to previously published work (Baillie-Johnson et al., 2015; Turner et al., 2017; van den Brink et al., 2014) using non-adherent U-bottom 96-well plates (Greiner, 650185) and NDiff 227 (N2B27) media from Takara (Y40002; batch references: 'AM30020S', 'AM10016S', 'AM90020S', 'AN30020S' and 'AO30013S'). Activin A treated gastruloids were made similarly to CHIR treated gastruloids but Activin A (High - 100ng/ml, Low - 25ng/ml; Recombinant human/mouse/rat Activin A PLUS protein, Qkine, Qk005) was used instead of CHIR (Chi99021, Sigma, SML1046) between 48 and 72h. AF-treated gastruloids were cultured with Activin A (25 ng/ml) and bFGF (12.5 ng/ml; Recombinant Murine FGF-basic, Peprotech, 450-33) between 0 and 48h.

Wnt and Nodal reporter analysis

TLC-mCherry gastruloids were imaged in a Zeiss Cell Observer fluorescence microscope, using 5x, 10x and 20x lenses depending on the size of the gastruloid. Images were processed using Fiji (Schindelin et al., 2012). Wnt and Nodal double reporter gastruloids were imaged and processed as in (McNamara et al., 2024).

Bulk RNA sequencing

RNA for bulk RNA sequencing was extracted from gastruloids using the Qiagen RNeasy Mini Kit (74104). The number of pooled gastruloids varied according to their developmental stage: full 96-well plate if at 48 or 72h, half a 96-well plate if at 96 and one-third of a 96-well plate if at 120 or 144h (at least 3 replicate samples were obtained per condition, from 3 independent experiments). RNA was also extracted from 2D cultures of ESCs and EPI-like stem cells (T=0h) following the same procedure. RNA concentration and purity were determined using a Fragment Analyzer (Agilent). cDNA libraries (non-stranded) were prepared from total RNA using the SMART-Seq2 protocol (Picelli et al., 2013). Sequencing was performed on an Illumina NextSeq 2000 at the IGC Genomics Facility (Instituto Gulbenkian de Ciência, Portugal), generating around 25 million single-end 100 base pair reads per sample.

Bulk RNA sequencing analysis

Data Preprocessing:

The quality of the raw sequencing files was initially assessed using *FastQC* (version 0.11.9)(Andrews, 2010). Adapter trimming and low-quality filtering were performed using *TrimGalore!* (version 0.6.7)(Krueger, 2015) where reads of Phred score below 20 were removed. Following trimming, the quality of the cleaned reads was re-evaluated with *FastQC* as a quality checkpoint to ensure the effectiveness of the trimming and filtering steps.

The processed sequencing files were mapped to the mouse genome (GRCm38.p6) using the GENCODE annotation (version 25) with the *STAR* alignment algorithm (version 2.7.10a)(Dobin et al., 2013). The mapping was performed with the default settings and the aligned read counts were generated using the *quantMode* function within *STAR*. The resulting count matrix was created by concatenating the assembled reads using the *Python*3 environment.

Downstream analysis and visualisation:

The downstream statistical and visualization analysis was performed in a *Python* 3.10.12 environment. Gene expression counts were normalised using the *pyDESeq2* package (version 0.4.0). Variance-stabilising transformation (VST) was used to transform raw counts across samples(Love et al., 2014). Feature selection and dimensionality reduction using PCA were performed with *sklearn* (version 1.2.2), while *seaborn* (version 0.13.2) and *matplotlib* (version 3.7.2) were used for plotting.

Gene differential expression analysis was also conducted using *pyDESeq2*. Genes with an adjusted p-value < 0.1, as determined by the Benjamini-Hochberg method for multiple testing correction, and an absolute log2 fold change > 0.5 were determined as significantly differentially expressed (DEGs).

Mixed populations of embryo sections from (Imaz-Rosshandler et al., 2024; Pijuan-Sala et al., 2019) were computed by creating random pseudo-samples using *KFold* from *sklearn*. A 5% step was set to create a gradient of mixed conditions.

The complete code, including all functions, parameters, and the data used for the analysis and visualisation, will be made available on GitHub upon publication.

In situ hybridization chain reaction

The HCR protocol for gastruloids was adapted from (Choi et al., 2018) and the HCR probes and amplifier hairpins were purchased from Molecular Instruments. First, gastruloids were fixed in 4% paraformaldehyde (PFA; Electron Microscopy Sciences, 15710) in PBS without calcium chloride and magnesium chloride (Sigma, D8537), supplemented with 0.1% Tween-20 (Sigma, P9416; PBST) for 20 to 25 minutes (depending on the developmental stage of the gastruloids) at room temperature (RT). Following two washes in PBST, the gastruloids were dehydrated in methanol (series of 25% increase of methanol in PBST) and stored at -20°C. Before starting the HCR protocol per se, gastruloids were rehydrated through a reverse methanol/PBST series. After two washes in PBST, gastruloids were incubated in hybridization buffer at 37°C for 30 minutes, and then overnight (ON) with the different gene probes. On the following day, gastruloids were washed four times every 15 minutes in 30% probe wash buffer and three times for 5 minutes each in 5X SSC buffer (Sigma, S6639) with 0.1% Tween-20 (SSCT). For the pre-amplification step, gastruloids were incubated in amplification buffer for 30 minutes at RT. The hairpins were snap-cooled (90 seconds at 95°C, followed by 30 minutes at RT) before being added to the amplification buffer solution where the gastruloids were incubated ON (RT). On the next day, the gastruloids were washed four times for 5 minutes each, plus three additional times every 20 minutes in SSCT (RT) before being stained with 4',6-Diamidino-2-Phenylindole, Dihydrochloride (DAPI; ThermoFisher, D1306) ON at 4°C. Finally, the gastruloids were washed in PBST and mounted in depression slides (Sigma, BR475505) with RapiClear 1.47 (SunJinLab, RC147001) as described in (Dias et al., 2021). Imaging was done on a Zeiss LSM980 Airyscan 2 microscope (objective lens: 10x/0.45, 20x/0.8) and image analysis was done using Fiji/ImageJ.

References

Aires, R., Dias, A. and Mallo, M. (2018). Deconstructing the molecular mechanisms shaping the vertebrate body plan. *Curr. Opin. Cell Biol.* **55**, 81–86.

Amin, S., Neijts, R., Simmini, S., van Rooijen, C., Tan, S. C., Kester, L., van Oudenaarden, A., Creyghton, M. P. and Deschamps, J. (2016). Cdx and T Brachyury Co-activate Growth Signaling in the Embryonic Axial Progenitor Niche. *Cell Rep.* **17**, 3165–3177.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. *Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc*.

Ang, S. and Rossant, J. (1994). HNF-3β Is Essential for Node and Notochord Formation in Mouse Development. *Cell* **78**, 561–574.

Anlaş, K., Gritti, N., Nakaki, F., Salamó Palau, L., Tlili, S. L., Oriola, D., Arató, K., Le Lim, J., Sharpe, J. and Trivedi, V. (2024). Early autonomous patterning of the anteroposterior axis in gastruloids. *Development* **151**, dev202171.

Argiro, L., Chevalier, C., Choquet, C., Nandkishore, N., Ghata, A., Baudot, A., Zaffran, S. and Lescroart, F. (2024). Gastruloids are competent to specify both cardiac and skeletal muscle lineages. *Nat. Commun.* **15**, 10172.

Arias, A. M. and Hayward, P. (2006). Filtering transcriptional noise during development: concepts and mechanisms. *Nat. Rev. Genet.* **7**, 34–44.

Arnold, S. J. and Robertson, E. J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. *Nat. Rev. Mol. Cell Biol.* **10**, 91–103.

Arnold, S. J., Hofmann, U. K., Bikoff, E. K. and Robertson, E. J. (2008). Pivotal roles for eomesodermin during axis formation, epithelium-to-mesenchyme transition and endoderm specification in the mouse. *Development* **135**, 501–511.

Asashima, M. and Satou-Kobayashi, Y. (2024). Spemann-Mangold organizer and mesoderm induction. *Cells Dev.* **178**, 203903.

Baillie-Johnson, P., van den Brink, S. C., Balayo, T., Turner, D. A. and Arias, A. M. (2015). Generation of aggregates of mouse embryonic stem cells that show symmetry breaking, polarization and emergent collective behaviour in vitro. *J. Vis. Exp.* e53252.

Bardot, E. S. and Hadjantonakis, A. K. (2020). Mouse gastrulation: Coordination of tissue patterning, specification and diversification of cell fate. *Mech. Dev.* **163**, 103617.

Barrow, J. R., Howell, W. D., Rule, M., Hayashi, S., Thomas, K. R., Capecchi, M. R. and McMahon, A. P. (2007). Wnt3 signaling in the epiblast is required for proper orientation of the anteroposterior axis. *Dev. Biol.* **312**, 312–20.

Beccari, L., Moris, N., Girgin, M., Turner, D. A., Baillie-Johnson, P., Cossy, A. C., Lutolf, M. P., Duboule, D. and Arias, A. M. (2018). Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids. *Nature* **562**, 272–276.

Ben-Haim, N., Lu, C., Guzman-Ayala, M., Pescatore, L., Mesnard, D., Bischofberger, M., Naef, F., Robertson, E. J. J. and Constam, D. B. (2006). The Nodal Precursor Acting via Activin Receptors Induces Mesoderm by Maintaining a Source of Its Convertases and BMP4. *Dev. Cell* **11**, 313–323.

Binagui-Casas, A., Dias, A., Guillot, C., Metzis, V. and Saunders, D. (2021). Building consensus in neuromesodermal research: Current advances and future biomedical perspectives. *Curr. Opin. Cell Biol.* **73**, 133–140.

Bolondi, A., Law, B. K., Kretzmer, H., Gassaloglu, S. I., Buschow, R., Riemenschneider, C., Yang, D., Walther, M., Veenvliet, J. V., Meissner, A., et al. (2024). Reconstructing axial progenitor field dynamics in mouse stem cell-derived embryoids. *Dev. Cell* **59**, 1489–1505.

Brennan J., Lu C.C., Norris D. P., Rodriguez T.A., Beddington R.S.P. and Robertson E.J. (2001). Nodal signalling in the epiblast patterns the early mouse embryo. *Nature* **411**, 965–969.

Brückner, D. B. and Tkačik, G. (2024). Information content and optimization of self-organized developmental systems. *PNAS* **121**, e2322326121.

Burtscher, I. and Lickert, H. (2009). Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. *Development* **136**, 1029–1038.

Cermola, F., Amoroso, F., Saracino, F., Ibello, E., De Cesare, D., Fico, A., Cobellis, G., Scalera, E., Casiraghi, C., D'Aniello, C., et al. (2022). Stabilization of cell-cell adhesions prevents symmetry breaking and locks in pluripotency in 3D gastruloids. *Stem Cell Reports* **17**, 2548–2564.

Choi, H. M. T. T., Schwarzkopf, M., Fornace, M. E., Acharya, A., Artavanis, G., Stegmaier, J., Cunha, A. and Pierce, N. A. (2018). Third-generation in situ hybridization chain reaction: Multiplexed, quantitative, sensitive, versatile, robust. *Development* **145**, dev165753.

Conlon, F. L., Lyons, K. M., Takaesu, N., Barth, K. S., Kispert, A., Herrmann, B. and Robertson, E. J. (1994). A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. *Development* **120**, 1919–1928.

de Lemos, L., Dias, A., Nóvoa, A. and Mallo, M. (2022). Epha1 is a cell-surface marker for the neuromesodermal competent population. *Development* **149**, dev198812.

Deschamps, J. and Duboule, D. (2017). Embryonic timing, axial stem cells, chromatin dynamics, and the Hox clock. *Genes Dev.* **31**, 1406–1416.

Dias, A. and Martinez Arias, A. (2023). Laying out the mammalian body plan to a T. *Dev. Cell* **58**, 1625–1626.

Dias, A., Lozovska, A., Wymeersch, F. J., Nóvoa, A., Binagui-Casas, A., Sobral, D., Martins, G. G., Wilson, V. and Mallo, M. (2020). A Tgfbr1/Snai1-dependent developmental module at the core of vertebrate axial elongation. *Elife* **9**, e56615.

Dias, A., Martins, G. G., Lopes, A. and Mallo, M. (2021). Three and four-dimensional visualization and analysis approaches to study vertebrate axial elongation and segmentation. *J. Vis. Exp.* e62086.

Dingare, C., Cao, D., Yang, J. J., Sozen, B. and Steventon, B. (2024). Mannose controls mesoderm specification and symmetry breaking in mouse gastruloids. *Dev. Cell* **59**, 1523–1537.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21.

Farag, N., Sacharen, C., Avni, L. and Nachman, I. (2024). Coordination between endoderm progression and mouse gastruloid elongation controls endodermal morphotype choice. *Dev. Cell* **59**, 2364–2374.

Faunes, F., Hayward, P., Descalzo, S. M., Chatterjee, S. S., Balayo, T., Trott, J., Christoforou, A., Ferrer-Vaquer, A., Hadjantonakis, A. K., Dasgupta, R., et al. (2013). A membrane-associated β-catenin/Oct4 complex correlates with ground-state pluripotency in mouse embryonic stem cells. *Development* **140**, 1171–1183.

Ferrer-vaquer, A., Piliszek, A., Tian, G., Aho, R. J., Dufort, D. and Hadjantonakis, A.-K. (2010). A sensitive and bright single-cell resolution live imaging reporter of Wnt/ß-catenin signaling in the mouse. *BMC Dev. Biol.* **10**, 1471–213X.

Fiuza, U.-M., Bonavia, S., Pascual-Mas, P., Torregrosa-Cortés, G., Casani-Galdon, P., Robertson, G., Dias, A. and Martinez Arias, A. (2024). Morphogenetic constrains in the development of gastruloids : implications for mouse gastrulation. *bioRxiv*.

Fong, G. H., Klingensmith, J., Wood, C. R., Rossant, J. and Breitman, M. L. (1996). Regulation of flt-1 expression during mouse embryogenesis suggests a role in the establishment of vascular endothelium. *Dev. Dyn.* **207**, 1–10.

Garge, R. K., Lynch, V., Fields, R., Casadei, S., Best, S., Stone, J., Mcgann, C. D., Shendure, J., Starita, L. M., Hamazaki, N., et al. (2024). The proteomic landscape and temporal dynamics of mammalian gastruloid development. *bioRxiv*.

Ghimire, S., Van Der Jeught, M., Neupane, J., Roost, M. S., Anckaert, J., Popovic, M., Van Nieuwerburgh, F., Mestdagh, P., Vandesompele, J., Deforce, D., et al. (2018). Comparative analysis of naive, primed and ground state pluripotency in mouse embryonic stem cells originating from the same genetic background. *Sci. Rep.* **8**, 5884.

Girgin, M. U., Broguiere, N., Hoehnel, S., Brandenberg, N., Mercier, B., Arias, A. M. and Lutolf, M. P. (2021a). Bioengineered embryoids mimic post-implantation development in vitro. *Nat. Commun.* **12**, 5140.

Girgin, M. U., Broguiere, N., Mattolini, L. and Lutolf, M. P. (2021b). Gastruloids generated without exogenous Wnt activation develop anterior neural tissues. *Stem Cell Reports* **16**, 1143–1155.

Gogolou, A., Souilhol, C., Granata, I., Wymeersch, F. J., Manipur, I., Wind, M., Frith, T. J. R., Guarini, M., Bertero, A., Bock, C., et al. (2022). Early anteroposterior regionalisation of human neural crest is shaped by a pro-mesodermal factor. *Elife* **11**, e74263.

Gsell, S., Tlili, S., Merkel, M. and Lenne, P.-F. (2024). Marangoni-like tissue flows enhance symmetry breaking of embryonic organoids. *bioRxiv*.

Guibentif, C., Griffiths, J. A., Imaz-Rosshandler, I., Ghazanfar, S., Nichols, J., Wilson, V., Göttgens, B. and Marioni, J. C. (2021). Diverse Routes toward Early Somites in the Mouse Embryo. *Dev. Cell* **56**, 141–153.

Guillot, C., Djeffal, Y., Michaut, A., Rabe, B. and Pourquié, O. (2021). Dynamics of primitive streak regression controls the fate of neuromesodermal progenitors in the chicken embryo. *Elife* **10**, e64819.

Hamazaki, N., Yang, W., Kubo, C. A., Qiu, C., Martin, B. K., Garge, R. K., Regalado, S. G., Nichols, E. K., Pendyala, S., Bradley, N., et al. (2024). Retinoic acid induces human gastruloids with posterior embryo-like structures. *Nat. Cell Biol.* **26**, 1790–1803.

Han, L., Chaturvedi, P., Kishimoto, K., Koike, H., Nasr, T., Iwasawa, K., Giesbrecht, K., Witcher, P. C., Eicher, A., Haines, L., et al. (2020). Single cell transcriptomics identifies a signaling network coordinating endoderm and mesoderm diversification during foregut organogenesis. *Nat. Commun.* **11**, 4158.

Hashmi, A., Tlili, S., Perrin, P., Lowndes, M., Peradziryi, H., Brickman, J. M., Arias, A. M. and Lenne, P. F. (2022). Cell-state transitions and collective cell movement generate an endoderm-like region in gastruloids. *Elife* **11**, e59371.

Herrmann, B. G. (1995). The mouse Brachyury (T) gene. *Semin. Dev. Biol.* **6**, 385–394.

Imaz-Rosshandler, I., Rode, C., Guibentif, C., Harland, L. T. G. G., Ton, M. L. N., Dhapola, P., Keitley, D., Argelaguet, R., Calero-Nieto, F. J., Nichols, J., et al. (2024). Tracking early mammalian organogenesis - prediction and validation of differentiation trajectories at whole organism scale. *Development* **151**, dev201867.

Kim, H. J., Osteil, P., Humphrey, S. J., Cinghu, S., Oldfield, A. J., Patrick, E., Wilkie, E. E., Peng, G., Suo, S., Jothi, R., et al. (2020). Transcriptional network dynamics during the progression of pluripotency revealed by integrative statistical learning. *Nucleic Acids Res.* **48**, 1828–1842.

Kinder, S. J., Tsang, T. E., Quinlan, G. A., Hadjantonakis, A.-K., Nagy, A. and Tam, P. P. L. (1999). The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo. *Development* **126**, 4691–701.

Kinder, S. J., Loebel, D. A. F. and Tam, P. P. L. (2001a). Allocation and early differentiation of cardiovascular progenitors in the mouse embryo. *Trends Cardiovasc. Med.* **11**, 177–184.

Kinder, S. J., Tsang, T. E., Wakamiya, M., Sasaki, H., Behringer, R. R., Nagy, A. and Tam, P. P. L. (2001b). The organizer of the mouse gastrula is composed of a dynamic population of progenitor cells for the axial mesoderm. *Development* **128**, 3623–3634.

Kinoshita, M., Barber, M., Mansfield, W., Cui, Y., Spindlow, D., Stirparo, G. G., Dietmann, S., Nichols, J. and Smith, A. (2021). Capture of Mouse and Human Stem Cells with Features of Formative Pluripotency. *Cell Stem Cell* **28**, 453–471.

Krueger, F. (2015). A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. *Available online at: https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/*.

Lawson, K. A. and Wilson, V. (2016). Chapter 3: A Revised Staging of Mouse Development Before Organogenesis. In *Kaufman's Atlas of Mouse Development Supplement* (ed. Richard Baldock), Bard, J.), Davidson, D. R.), and Morriss-Kay, G.), pp. 51–64. Academic Press.

Lawson, K. A., Meneses, J. J. and Pedersen, R. A. (1991). Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development* **113**, 891–911.

Liu, P., Wakamiya, M., Shea, M. J., Albrecht, U., Behringer, R. R. and Bradley, A. (1999). Requirement for Wnt3 in vertebrate axis formation. *Nat. Genet.* **22**, 361–365.

Lopez-Delisle, L., Zakany, J., Bochaton, C., Osteil, P., Mayran, A., Darbellay, F., Mascrez, B., Rekaik, H. and Duboule, D. (2024). CTCF- dependent insulation of Hoxb13 and the heterochronic control of tail length. *PNAS* **121**, e2414865121.

Lotto, J., Drissler, S., Cullum, R., Wei, W., Setty, M., Bell, E. M., Boutet, S. C., Nowotschin, S., Kuo, Y. Y., Garg, V., et al. (2020). Single-Cell Transcriptomics Reveals Early Emergence of Liver Parenchymal and Non-parenchymal Cell Lineages. *Cell* **183**, 702–716.

Love, M. I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550.

Lu, C. C. and Robertson, E. J. (2004). Multiple roles for Nodal in the epiblast of the mouse embryo in the establishment of anterior-posterior patterning. *Dev. Biol.*

273, 149–159.
Luque, A. V., Savill, R., López-Anguita, N., Bolondi, A., Garai, S., Gassaloglu, S. I., Poddar, A., Bulut-Karslioglu, A. and Veenvliet, J. V (2023). Integrated Molecular-Phenotypic Profiling Reveals Metabolic Control of Morphological Variation in Stembryos. *bioRxiv*.
Mallo, M. (2021). Of necks, trunks and tails: Axial skeletal diversity among vertebrates. *Diversity* **13**, 289.
Mantziou, V., Baillie-Benson, P., Jaklin, M., Kustermann, S., Arias, A. M. and Moris, N. (2021). In vitro teratogenicity testing using a 3D, embryo-like gastruloid system. *Reprod. Toxicol.* **105**, 72–90.
Martinez Arias, A., Marikawa, Y. and Moris, N. (2022). Gastruloids: Pluripotent stem cell models of mammalian gastrulation and embryo engineering. *Dev. Biol.* **488**, 35–46.
Martinez Arias, A., Rivron, N., Moris, N., Tam, P., Alev, C., Fu, J., Hadjantonakis, A. K., Hanna, J. H., Minchiotti, G., Pourquie, O., et al. (2024). Criteria for the standardization of stem-cell-based embryo models. *Nat. Cell Biol.* **26**, 1625–1628.
Mayran, A., Kolly, D., Lopez-Delisle, L., Romaniuk, Y., Leonardi, M., Cossy, A. C., Lacroix, T., Amândio, R., Osteil, P. and Duboule, D. (2023). Cadherins modulate the self-organizing potential of gastruloids. *bioRxiv*.
McNamara, H. M., Solley, S. C., Adamson, B., Chan, M. M. and Toettcher, J. E. (2024). Recording morphogen signals reveals mechanisms underlying gastruloid symmetry breaking. *Nat. Cell Biol.* **26**, 1832–1844.
Merle, M., Friedman, L., Chureau, C., Shoushtarizadeh, A. and Gregor, T. (2024). Precise and scalable self-organization in mammalian pseudo-embryos. *Nat. Struct. Mol. Biol.* **31**, 896–902.
Moris, N., Anlas, K., van den Brink, S. C., Alemany, A., Schröder, J., Ghimire, S., Balayo, T., van Oudenaarden, A. and Martinez Arias, A. (2020). An in vitro model of early anteroposterior organization during human development. *Nature* **582**, 410–415.
Nandkishore, N., Vyas, B., Javali, A., Ghosh, S. and Sambasivan, R. (2018). Divergent early mesoderm specification underlies distinct head and trunk muscle programmes in vertebrates. *Development* **145**, dev173187.
Narita, Y. and Kuratani, S. (2005). Evolution of the Vertebral Formulae in Mammals: A Perspective on Developmental Constraints. *J. Exp. Zool.* **304B**, 91–106.
Neijts, R., Amin, S., Van Rooijen, C., Tan, S., Creyghton, M. P., De Laat, W. and Deschamps, J. (2016). Polarized regulatory landscape and Wnt responsiveness underlie Hox activation in embryos. *Genes Dev.* **30**, 1937–1942.
Nichols, J. and Smith, A. (2009). Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487–92.
Niehrs, C. (2021). The role of Xenopus developmental biology in unraveling Wnt signalling and antero-posterior axis formation. *Dev. Biol.* **482**, 1–6.
Pfister, S., Steiner, K. A. and Tam, P. P. L. (2007). Gene expression pattern and progression of embryogenesis in the immediate post-implantation period of mouse development. *Gene Expr. Patterns* **7**, 558–573.
Phillips, H. M., Stothard, C. A., Shaikh Qureshi, W. M., Kousa, A. I., Alberto Briones-Leon, J., Khasawneh, R. R., O'Loughlin, C., Sanders, R., Mazzotta, S., Dodds, R., et al. (2019). Pax9 is required for cardiovascular development and interacts with

Tbx1 in the pharyngeal endoderm to control 4th pharyngeal arch artery morphogenesis. *Development* **146**, dev177618.

Picelli, S., Björklund, Å. K., Faridani, O. R., Sagasser, S., Winberg, G. and Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10**, 1096–1100.

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.

Probst, S., Sagar, Tosic, J., Schwan, C., Grün, D. and Arnold, S. J. (2021). Spatiotemporal sequence of mesoderm and endoderm lineage segregation during mouse gastrulation. *Development* **148**, dev193789.

Rekaik, H., Lopez-Delisle, L., Hintermann, A., Mascrez, B., Bochaton, C., Mayran, A. and Duboule, D. (2023). Sequential and directional insulation by conserved CTCF sites underlies the Hox timer in stembryos. *Nat. Genet.* **55**, 1164–1175.

Robertson, E. J. (2014). Dose-dependent Nodal/Smad signals pattern the early mouse embryo. *Semin. Cell Dev. Biol.* **32**, 73–79.

Rosen, L. U., Stapel, L. C., Argelaguet, R., Barker, C. G., Yang, A., Reik, W. and Marioni, J. C. (2022). Inter-gastruloid heterogeneity revealed by single cell transcriptomics time course: implications for organoid based perturbation studies. *bioRxiv*.

Rossi, G., Broguiere, N., Miyamoto, M., Boni, A., Guiet, R., Girgin, M., Kelly, R. G., Kwon, C. and Lutolf, M. P. (2021). Capturing Cardiogenesis in Gastruloids. *Cell Stem Cell* **28**, 230–240.

Rossi, G., Giger, S., Hübscher, T. and Lutolf, M. P. (2022). Gastruloids as in vitro models of embryonic blood development with spatial and temporal resolution. *Sci. Rep.* **12**, 13380.

Sagner, A. and Briscoe, J. (2019). Establishing neuronal diversity in the spinal cord: a time and a place. *Development* **146**, dev182154.

Sagy, N., Slovin, S., Allalouf, M., Pour, M., Savyon, G., Boxman, J. and Nachman, I. (2019). Prediction and control of symmetry breaking in embryoid bodies by environment and signal integration. *Development* **146**, dev181917.

Scheibner, K., Schirge, S., Burtscher, I., Büttner, M., Sterr, M., Yang, D., Böttcher, A., Ansarullah, Irmler, M., Beckers, J., et al. (2021). Epithelial cell plasticity drives endoderm formation during gastrulation. *Nat. Cell Biol.* **23**, 692–703.

Schier, A. F. (2003). Nodal Signaling in Vertebrate Development. *Annu. Rev. Cell Dev. Biol.* **19**, 589–621.

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

Schüle, K. M., Weckerle, J., Probst, S., Wehmeyer, A. E., Zissel, L., Schröder, C. M., Tekman, M., Kim, G.-J. J., Schlägl, I.-M. M., Sagar, S., et al. (2023). Eomes restricts Brachyury functions at the onset of mouse gastrulation. *Dev. Cell* **58**, 1627–1642.

Shen, M. M. (2007). Nodal signaling: developmental roles and regulation. *Development* **134**, 1023–1034.

Sheng, G., Martinez Arias, A. and Sutherland, A. (2021). The primitive streak and

cellular principles of building an amniote body through gastrulation. *Science* **374**, 1213.

Stapornwongkul, K. S., Hahn, E., Palau, L. S., Arato, K., Gritti, N., Anlas, K., Polinski, P., Lopez, M. O., Eibisuya, M. and Trivedi, V. (2023). Metabolic control of germ layer proportions through regulation of Nodal and Wnt signalling. *bioRxiv*.

Stelloo, S., Alejo-Vinogradova, M. T., van Gelder, C. A. G. H., Zijlmans, D. W., van Oostrom, M. J., Valverde, J. M., Lamers, L. A., Rus, T., Sobrevals Alcaraz, P., Schäfers, T., et al. (2024). Deciphering lineage specification during early embryogenesis in mouse gastruloids using multilayered proteomics. *Cell Stem Cell* **31**, 1072–1090.

Stothard, C. A., Mazzotta, S., Vyas, A., Schneider, J. E., Mohun, T. J., Henderson, D. J., Phillips, H. M. and Bamforth, S. D. (2020). Pax9 and Gbx2 interact in the pharyngeal endoderm to control cardiovascular development. *J. Cardiovasc. Dev. Dis.* **7**, 20.

Suppinger, S., Zinner, M., Aizarani, N., Lukonin, I., Ortiz, R., Azzi, C., Stadler, M. B., Vianello, S., Palla, G., Kohler, H., et al. (2023). Multimodal characterization of murine gastruloid development. *Cell Stem Cell* **30**, 867–884.

Takemoto, T., Abre, T., Kiyonari, H., Nakao, K., Furuta, Y., Suzuki, H., Takada, S., Fujimori, T. and Kondoh, H. (2016). R26-WntVis reporter mice showing graded response to Wnt signal levels. *Genes to Cells* **21**, 661–669.

Tam, P. P. and Behringer, R. R. (1997). Mouse gastrulation: the formation of a mammalian body plan. *Mech. Dev.* **68**, 3–25.

Tam, P. P. L. and Loebel, D. A. F. (2007). Gene function in mouse embryogenesis: get set for gastrulation. *Nat. Rev. Genet.* **8**, 368–381.

Tam, P. P. L., Parameswaran, M., Kinder, S. J. and Weinberger, R. P. (1997). The allocation of epiblast cells to the embryonic heart and other mesodermal lineages: the role of ingression and tissue movement during gastrulation. *Development* **124**, 1631–1642.

Tam, P. P., Loebel, D. A. and Tanaka, S. S. (2006). Building the mouse gastrula: signals, asymmetry and lineages. *Curr. Opin. Genet. Dev.* **16**, 419–425.

Torregrosa, G. and Garcia-Ojalvo, J. (2021). Mechanistic models of cell-fate transitions from single-cell data. *Curr. Opin. Syst. Biol.* **26**, 79–86.

Tosic, J., Kim, G. J., Pavlovic, M., Schröder, C. M., Mersiowsky, S. L., Barg, M., Hofherr, A., Probst, S., Köttgen, M., Hein, L., et al. (2019). Eomes and Brachyury control pluripotency exit and germ-layer segregation by changing the chromatin state. *Nat. Cell Biol.* **21**, 1518–1531.

Turner, D. A. and Martinez Arias, A. (2024). Three-dimensional stem cell models of mammalian gastrulation. *BioEssays* **46**, e2400123.

Turner, D. A., Girgin, M., Alonso-Crisostomo, L., Trivedi, V., Baillie-Johnson, P., Glodowski, C. R., Hayward, P. C., Collignon, J., Gustavsen, C., Serup, P., et al. (2017). *Anteroposterior polarity and elongation in the absence of extraembryonic tissues and of spatially localised signalling in gastruloids: mammalian embryonic organoids*.

Underhill, E. J. and Toettcher, J. E. (2023). Control of gastruloid patterning and morphogenesis by the Erk and Akt signaling pathways. *Development* **150**, dev201663.

van den Brink, S. C., Baillie-Johnson, P., Balayo, T., Hadjantonakis, A. K., Nowotschin,

S., Turner, D. A. and Arias, A. M. (2014). Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells. *Development* **141**, 4231–4242.

van den Brink, S. C., Alemany, A., van Batenburg, V., Moris, N., Blotenburg, M., Vivié, J., Baillie-Johnson, P., Nichols, J., Sonnen, K. F., Martinez Arias, A., et al. (2020). Single-cell and spatial transcriptomics reveal somitogenesis in gastruloids. *Nature* **582**, 405–409.

Veenvliet, J. V., Bolondi, A., Kretzmer, H., Haut, L., Scholze-Wittler, M., Schifferl, D., Koch, F., Guignard, L., Kumar, A. S., Pustet, M., et al. (2020). Mouse embryonic stem cells self-organize into trunk-like structures with neural tube and somites. *Science* **370**, 1291.

Vianello, S. and Lutolf, M. P. (2021). In vitro endoderm emergence and self-organisation in the absence of extraembryonic tissues and embryonic architecture. *bioRxiv*.

Vyas, B., Nandkishore, N. and Sambasivan, R. (2020). Vertebrate cranial mesoderm: developmental trajectory and evolutionary origin. *Cell. Mol. Life Sci.* **77**, 1933–1945.

Wehmeyer, A. E. and Arnold, S. J. in preparation.

Wehmeyer, A. E., Schüle, K. M., Conrad, A., Schröder, C. M., Probst, S. and Arnold, S. J. (2022). Chimeric 3D gastruloids – a versatile tool for studies of mammalian peri-gastrulation development. *Development* **149**, dev200812.

Wymeersch, F. J., Skylaki, S., Huang, Y., Watson, J. A., Economou, C., Marek-Johnston, C., Tomlinson, S. R. and Wilson, V. (2019). Transcriptionally dynamic progenitor populations organised around a stable niche drive axial patterning. *Development* **146**, dev168161.

Wymeersch, F. J., Wilson, V. and Tsakiridis, A. (2021). Understanding axial progenitor biology in vivo and in vitro. *Development* **148**, dev180612.

Yamaguchi, T. P. (2001). Heads or tails: Wnts and anterior-posterior patterning. *Curr. Biol.* **11**, 713–724.

Yamanaka, Y., Hamidi, S., Yoshioka-Kobayashi, K., Munira, S., Sunadome, K., Zhang, Y., Kurokawa, Y., Ericsson, R., Mieda, A., Thompson, J. L., et al. (2023). Reconstituting human somitogenesis in vitro. *Nature* **614**, 509–520.

Yang, J., Antin, P., Berx, G., Blanpain, C., Brabletz, T., Bronner, M., Campbell, K., Cano, A., Casanova, J., Christofori, G., et al. (2020). Guidelines and definitions for research on epithelial–mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* **21**, 341–352.

Yoney, A., Etoc, F., Ruzo, A., Carroll, T., Metzger, J. J., Martyn, I., Li, S., Kirst, C., Siggia, E. D. and Brivanlou, A. H. (2018). WNT signaling memory is required for ACTIVIN to function as a morphogen in human gastruloids. *Elife* **7**, e38279.

Zeller, P., Blotenburg, M., Bhardwaj, V., Barbanson, B. A. de, Salmén, F. and Oudenaarden, A. van (2024). T-ChIC: multi-omic detection of histone modifications and full-length transcriptomes in the same single cell. *bioRxiv*.

Figures

Figure 1 - Wnt levels modulate the temporal activation of the late posterior primitive streak.

A - Schematic diagram of the gastruloid protocol using different CHIR concentrations and their corresponding DMSO controls. The protocol starts with the seeding of approximately 300 mouse embryonic stem cells (mESCs) cultured in ESLIF media (see **Material and Methods** section), as single cells, per well of a 96-well plate in N2B27 media (0h). Following aggregation, CHIR or DMSO is introduced at 48h in fresh N2B27. Gastruloids were then collected for bulk RNAseq at various time points. **B** - PCA of distinct temporal gastruloid bulk RNA sequencing samples, with some of the top 60 genes derived from the loadings of components 1 and 2. Vectors representing genes associated with naive pluripotency (e.g. *Tfcp2l1*) are directed towards the 0h replicates, which cluster far apart from the other conditions. Early epiblast markers such as *Nodal*, *Eomes*, *Fgf5*, and *Pitx2*, are represented in loading vectors pointing towards the 48h and DMSO conditions, indicating a change in cell state from 0 to 48h (see **Figure S1**). DMSO replicates/conditions cluster together, suggesting that varying DMSO concentrations do not significantly alter the gastruloid transcriptome at 72h. Genes related to the late posterior PS (e.g. *Tbxt*, *Wnt8a* and *Cdx2*) are in loading vectors pointing along the CHIR conditions, which capture its concentration dependence. The 72h gastruloids treated with 3µM CHIR cluster between those treated with 0.5 and 20µM CHIR along this axis. Gastruloids at 60h treated with 3µM CHIR cluster near the 72h gastruloids treated with 0.5µM CHIR, whereas those at 60h that were treated with 20µM CHIR cluster close to 72h gastruloids treated with 3µM CHIR. Together, this highlights the temporal and concentration-dependent effects of Wnt signalling on gastruloids progression along PC1, which represents developmental time. **C** - Gene expression heatmap (mean variance-stabilized read counts of replicates) showing that the early epiblast-like signature observed at 48h and in DMSO conditions (e.g. *Utf1*, *Fgf5* and *Otx2*) gradually disappears with the increase of CHIR levels. In contrast, expression of genes related to the late PS (e.g. *Tbxt*, *Evx1* and *Fst*) increases with CHIR exposure. This transition is more pronounced in 72h gastruloids treated with 20µM CHIR, which display a strong downregulation of pluripotent and early epiblast/PS markers (e.g. Nanog and *Eomes*, respectively) and a concomitant upregulation of genes related to the caudal epiblast (e.g. *Cdx2*, *Nkx1-2* and *Greb1*). **D** – Pseudo-bulk, time-resolved points of scRNA-seq mouse embryo gastrulation atlas(Pijuan-Sala et al., 2019) (blue dots, see **Material and Methods** for the pseudo-bulk construction) projected onto PCA of various CHIR-treated gastruloids suggests that Wnt signalling controls the pace of gastrulation. The projection is robust to the analysis parameters (see PCA robustness analysis in **Figure S3**). Gastruloids at 48h cluster near E6.5 mouse embryos, whereas 72h gastruloids treated with 20µM CHIR cluster towards the opposite end, closer to E8.0-8.5 mouse embryos. 0,5 and 3µM CHIR treated gastruloids cluster halfway along dimension 1 in a concentration-dependent manner, reflecting the temporal progression of mouse embryos. Among the top 50 genes derived from the loadings of components 1 and 2, the presence of early epiblast and late/caudal epiblast markers (e.g. *Eomes* and *Cdx2*, respectively) pointing to opposite ends of dimension 1 further validate the temporal

developmental changes occurring in the mouse epiblast during gastrulation. **E** - Euclidean distance between the gastruloid and embryo (pseudo-bulk atlas (Pijuan-Sala et al., 2019)) datasets suggests that the transcriptome of 48h gastruloids closely resembles that of E6.5 mouse embryos, whereas 72h gastruloids more closely resemble later embryonic stages in a concentration-dependent manner. The increase in distance for later time points is likely due to the increase of complexity in mouse cell diversity in comparison to gastruloids.

Figure 2 – The gastruloid as a stem cell-based embryo model of the primitive streak.

A - Schematic diagram of the gastruloid engineering process with either CHIR or Activin A (ActA). CHIR treated gastruloids exhibit an elongated phenotype at 120h, their endpoint. Conversely, gastruloids made with Activin A develop until 144h and remain mostly spherical. **B** - PCA of temporal Activin A and CHIR-treated gastruloid bulk RNAseq samples, with some of the top 100 genes derived from the loadings of components 1 and 2. The organisation of these genes and their vectors along the PCA mimics PS development; epiblast markers in the left quadrant (e.g. *Otx2* and *Fgf5*), early streak-related genes in the upper one (e.g. *Eomes*, *Mixl1* and *Foxa2*) and late streak markers further to the top-right (*Tbxt*, *Cdx2* and *Wnt3a*). This temporal sequence along dimension 1 ends with markers characteristic of epiblast neural and mesodermal derivatives during early axial elongation (e.g. *Meox1*, *Pax3* and *Nr2f2*), reflecting the positioning of the distinct gastruloid samples in the PCA. The 48h and 72h DMSO replicates/conditions cluster in the lower left quadrant, followed by the 72h Activin A and 72h low (0,5µM) CHIR samples. The 72h standard CHIR (3µM) replicates/conditions cluster further along dimension 1, suggesting they resemble a later PS developmental timepoint. Beyond this point, there is a bifurcation, with Activin A gastruloids following a different path in the PCA than that of CHIR treated gastruloids. **C** – Heatmap of gene expression (mean of the replicates variance-stabilized read counts) comparing 72h Activin A and CHIR treated gastruloids, with 48h and 72h DMSO gastruloids as a reference. Activin A treated gastruloids exhibit a gene expression characteristic of the early PS, with higher levels of *Nodal*, *Cripto* (*Tdgf1*), *Eomes* and *Foxa2* and lower levels of *Tbxt*, *Cdx2*, *Wnt3a* in comparison to CHIR treated gastruloids. Also, Activin A treated gastruloids express higher levels of genes associated with the early epiblast (e.g. *Otx2* and *Fgf5*). The epiblast state of Activin A gastruloids at 72h is more mature/advanced than that of 48h and 72h DMSO gastruloids because of the higher levels of *Nanog* and *Fgf5*, together with lower levels of *Utf1*. The 72h 0.5µM CHIR gastruloids display a gene profile more similar to Activin A treated gastruloids than 3µM CHIR gastruloids (e.g. *Eomes*, *Otx2*, *Myc* and *Cdx2*). The activation of the posterior PS module in CHIR treated gastruloids occurs in a concentration-dependent manner (e.g. *Cdx2*, *Wnt3a* and *Hoxa1*). **D** - Gene expression heatmap (mean of the replicates variance-stabilized read counts) of 96h CHIR (3µM) and Activin A treated gastruloids. Gastruloids made with CHIR express genes that are associated with the early caudal epiblast (e.g. *Gdf11*, *Cdx2* and *Wnt3a*, with lack of *Wnt3*) and its early mesoderm derivatives (e.g. *Tbx6*, *Cited1* and *Msgn1*). In contrast, Activin A treated gastruloids are delayed and express genes associated with the early to mid-streak transition (e.g. *Eomes*, *Mixl1*, *Wnt3* and *Cdx2*, and the lack of *Gdf11*). Activin A treated gastruloids also show signs of having derivatives of the early PS as

highlighted by the expression of *Noto* and *Gsc* (suggestive of node/notochord), *Foxa2* and *Sox17* (endoderm), *Kdr* and *Gata6* (haematoendothelium) and *Tbx6* and *Mesp1* (mesoderm). **E** - Euclidean distance analysis between gastruloid and embryo (pseudo-bulk atlas (Pijuan-Sala et al., 2019)) datasets shows that both types of gastruloids follow a similar developmental progression. However, Activin A treated gastruloids seem to resemble earlier developmental time points compared to CHIR gastruloids. This difference starts at 72h, with Activin A treated gastruloids being more similar to E6.75 mouse embryos while CHIR-treated gastruloids resemble E7.25 embryos. **F** – HCR (MIP) on 72h Activin A and CHIR treated gastruloids showing the expression of *Tbxt* and *Eomes* (in green and magenta, respectively). CHIR-treated gastruloids express only *Tbxt*, resembling the late posterior PS. In contrast, Activin A treated gastruloids mimic the early PS, displaying both *Eomes* and *Tbxt* expression (white implies co-localization). The scale bar is equivalent to 50µm. **G** - HCR (MIP) highlighting the expression of PS and early differentiation marker genes in 96h gastruloids. Activin A treated gastruloids display a mid-streak-like signature with high levels of *Tbxt* and expression of *Eomes* and *Wnt3*. On the contrary, gastruloids made with CHIR show weak expression of *Eomes* and *Wnt3* (brown and red arrowheads) and a reduced, polarized *Tbxt* expression. Similarly, genes related to endoderm, node/notochord and haematoendothelium are expressed in Activin A gastruloids but show limited (yellow and orange arrowhead) or no expression in CHIR treated gastruloids. Early mesoderm is present in both types of gastruloids as indicated by the expression of *Mesp1* and *Tbx6*. The scale bar is equivalent to 50µm.

Figure 3 – Activin A-treated gastruloids develop several anterior body features.

HCR images (MIP or optical section) highlighting the expression of key genes involved in anterior body formation on Activin A treated gastruloids at 144h. **A** – The localised *Pou5f1/Oct4* expression pattern is suggestive of PGC-like cells or ectopic pluripotency. **B-C** – *Tbxt* and *Foxa2* co-localization indicates the presence of notochord-like cells, resembling the anterior-most part of the notochord in mouse embryos. This is inferred from the absence of *Noto*, a marker of notochord progenitor cells. The *Foxa2* expression that does not overlap with *Tbxt* suggests the presence of endoderm-like cells. These cells show an organisation similar to the embryonic gut, including an internal cavity (white asterisk). Additionally, the colocalization of *Foxa2* and *Sox17* at one end of this gut-like structure suggests its resemblance to the foregut of mouse embryos. **D** – Colocalization of *Foxa2* and *Sox2* (red arrowhead) also suggests the presence of endoderm-like cells. However, *Sox2* expression alone (orange arrowhead) indicates the presence of neural tube-like cells, which can be observed going from one tip to the other, align to the gut and notochord as in the embryo. **E-F** – Endothelial-like cells are also present in Activin A treated gastruloids, as shown by *Kdr* expression. *Isl1* and *Pitx2* expression patterns would be consistent with the development of pharyngeal endoderm and cardiac-like cells. **G-H** – *Tbx1* expression suggests the presence of anterior mesoderm. The co-localization of *Tbx1* and *Tcf15* (grey arrowhead) suggests the presence of cranial mesoderm-like cells, while *Tcf15* expression adjacent to *Tbx1* is supportive of the existence of paraxial mesoderm-like cells. *Nkx2-5* expression near *Tbx1* suggests the presence of cardiac-like cells.

Figure 4 – Nodal signalling activity drives anterior body formation in a dose-dependent manner.

A - Heatmap showing gene expression (mean of the replicates variance-stabilized read counts) in 144-hour gastruloids treated with varying levels of Activin A (ActA). Gastruloids treated with higher doses of Activin A exhibit increased expression of *Cyp26a1*, *Sox17*, *Pitx2*, *Pax9*, *Isl1*, and *Tbx1*, indicative of more anterior endoderm and mesoderm-like structures compared to gastruloids exposed to lower levels of Nodal signalling. In the latter gastruloids, *Aldh1a2/Raldh2*, typically found in trunk regions, is highly expressed, along with a more posterior *Hox* gene signature (e.g. *Hoxa5* and *Hoxb8*). *Noto* expression, marking notochord progenitors, is limited to gastruloids treated with low Activin A levels. Additionally, *Pax6* and *Olig2* are expressed at higher levels in low Activin A gastruloids, suggesting the presence of more posterior neural tube-like cells compared to high Activin A gastruloids. Increased expression of *Kdr*, *Esam* and *Cdh5* in gastruloids treated with a high dose of Activin A suggests the presence of more endothelial-like cells compared to those with lower levels of treatment. **B** - Comparative transcriptome analysis between bulk RNA-seq from different gastruloid types and pseudo-bulk mixtures of E8.5 mouse embryos (see **Material and Methods**). Embryonic data from Imaz-Rosshandler et al.(Imaz-Rosshandler et al., 2024) is divided into "A" (anterior) and "P" (posterior) regions, in accordance to the dissections indicated in their work and highlighted in this scheme. B1 – PCA of bulk RNAseq data from gastruloid samples/replicates, projected onto computationally reconstructed anterior and posterior embryonic cell populations. CHIR-treated 120h gastruloids cluster near the posterior region, while 144h gastruloids treated with high Activin A cluster near the anterior region, though not fully resembling 100% anterior cells ("Anterior"). Low Activin A treated gastruloids cluster closer to mixed anterior-posterior populations, suggesting they mimic more posterior regions of the body axis compared to high Activin A treated gastruloids. B2 – Euclidean distance analysis shows that CHIR-treated gastruloids differ from the anterior part of E8.5 embryos, while Activin A treated gastruloids more closely resemble this region. **C** - Transcriptome comparison between 144h gastruloids treated with different levels of Activin A and E8.75 mouse embryos, divided into "A" (anterior), "M" (medial) and "P" (posterior) sections (Imaz-Rosshandler et al., 2024). C1 – PCA of 144h gastruloid RNA-seq data, projected onto reconstructed anterior and medial embryonic cell populations. Activin A treated gastruloids cluster between these two populations, though those treated with high Activin A resemble more anterior populations. C2 – Euclidean distance analysis suggests that 144h gastruloids treated with a higher Activin A concentration mimic more anterior embryonic regions (anterior and medial) than those treated with lower doses, which display a broader body axis regional diversity. **D** - HCR (MIP) images highlighting differences between gastruloids treated with low or high Activin A. In high Activin A treated gastruloids, genes like *Kdr*, *Cyp26a1*, *Cdh5* and *Otx2* are widely expressed, whereas their expression is restricted to small regions in low Activin A treated gastruloids. *Tbx1* expression is higher and present in more cells in high Activin A treated gastruloids. *Nkx2-5* is also expressed in gastruloids developed with a high dose of Activin A (red arrow) but absent in low Activin A-treated gastruloids. *Cdx2* expression is higher in low Activin A treated gastruloids, with an enlarged expression domain (red arrowhead). *Hoxb9* expression is also broader in low Activin A-treated gastruloids. Co-expression of *Noto* and *Tbxt*, marking notochord progenitors, is observed

only at the tip of low Activin A-treated gastruloids. Inlets highlight individual gene expression. The scale bar is equivalent to 100µm.

Figure 5 – A temporal coordination between Nodal and Wnt signalling governs the emergence of the mammalian body plan.

A – Schematic representation of the signalling-dependent modularity in the mouse PS during gastrulation. The early PS module (yellow) is mainly controlled by the activity of Nodal signalling, which decreases over time and is responsible for setting out the various mesendodermal fates. Conversely, Wnt/β-catenin activity increases over time and regulates the emergence of the late PS module (red), thus pacing gastrulation. Extra-embryonic tissues are shown in light and dark grey. Other embryonic tissues that are not the PS are shown in light blue. **B** – Schematic representation of an E8.5 mouse embryo highlighting the two compartments that arise from the early and late PS modules, regulated by the two opposing gradients of Nodal and Wnt signalling activity (yellow and red, respectively). Brain tissues (blue) relate to a distinct gastrulation module, which requires the absence of Wnt and Nodal signalling activity. Schemes of Activin A (Nodal) and CHIR (Wnt) treated gastruloids are also provided with the respective colours, in a gradient manner.

Supplementary Figures

Figure S1 - Transition from naive to primed pluripotency occurs between 0 and 48h of the gastruloid protocol.

A - PCA of gastruloid and EpiLSCs (Epi) bulk RNA sequencing samples. mESCs at the beginning of the gastruloid protocol (0h) form a distinct cluster, far away from all the other samples. EpiLSCs cluster between the 48 and 72h DMSO gastruloids, suggesting that a transition from naive to formative/primed pluripotency occurs between 0 and 48h in the standard gastruloid protocol. The addition of Activin A and Fgf (AF; see **Material and Methods**) during this time window (0-48h) does not result in significant transcriptomic differences as these gastruloids at 48h cluster between the 48h control gastruloids and the EpiLSCs samples. **B** - Gene expression heatmap (mean of the replicates variance-stabilized read counts) illustrating different states of pluripotency. Contrary to all 48h gastruloids and EpiLSCs samples, mESCs (0h) display high expression of genes associated with naive pluripotency (e.g. *Zfp42*, *Nanog* and *Essrb*) and low levels of epiblast-related genes (e.g. *Otx2* and *Fgf5*). EpiLSCs exhibit slightly higher expression of PS-related genes (e.g. *Tbxt*) compared to 48h gastruloids, further distinguishing their transcriptomic profile.

Figure S2 - Wnt signalling activity in distinct CHIR and DMSO-treated gastruloids.

Maximum intensity projection (MIP) of the TCF/Lef reporter at 48h (**A**) and 72h gastruloids (**B**-**G**), upon different CHIR and DMSO treatments provided at 48h. Contrary to DMSO controls (B-D), which do not show a significant change in Wnt activity, CHIR-

treated gastruloids display higher reporter expression in a concentration-dependent manner. The scale bar is equivalent to 100µm.

Figure S3 - Robustness analysis of mouse gastruloids and embryos transcriptomic distance.

Correlation between pseudo-bulk data from Pijuan-Sala et al.(Pijuan-Sala et al., 2019) and 48h gastruloids (**A**),72h gastruloids treated with 0.5 (**B**), 3 (**C**) and 20µM CHIR. Different choices (**D**) of three analysis parameters were also tested: 1) variance threshold, which selected the number of genes to use for the PCA (**Figure 1D**), 2) number of PCA components and 3) metric to compute the pairwise distance between the embryo and gastruloid samples. All the parameter choices lead to similar results, with 48h gastruloids being closer to early-stage embryos and 72h gastruloids close to late-stage embryos, in a CHIR concentration manner. The main outlier is the parameter set using a different metric (cosine metric), which displays a similar trend but shifts the expected minimum at each CHIR concentration to earlier stages in the 0.5 and 3µM CHIR samples.

Figure S4 - Expression of gastrulation marker genes in mouse embryos and gastruloids.

A - HCR (MIP) for early and late PS-related genes in gastruloids made with different levels of Wnt signalling. At 72h, gastruloids treated with DMSO do not express any of the late PS marker genes (e.g. *Cdx2*), with only *Nodal* being detected. In contrast, CHIR gastruloids show an upregulation of *Tbxt* and *Cdx2* and a downregulation of genes associated with early developmental stages (e.g. *Nodal* and *Wnt3*) in a concentration-dependent manner. These changes promoted by Wnt signalling seem to occur also in a temporal manner as 20µM CHIR treated gastruloids at 60h display a gene expression resembling that of 3µM CHIR treated gastruloids at 72h (e.g. low *Nodal* and *Cdx2,* high *Tbxt*). The scale bar is equivalent to 50µm. **A** – Pseudo-bulk temporal expression of specific gastrulation-related genes in mouse embryos (containing extra-embryonic tissues). The expression of the early PS marker genes *Nodal* and *Wnt3* significantly decreases after E7.0, with the upregulation of *Tbxt*, *Cdx2* and *Hoxa1*. The drop in the levels of *Tbxt* after E7.5 is consistent with the regression of the PS. High levels of *Cdx2* at early stages are consistent with its expression in extra-embryonic tissues. The gene expression values displayed in the figure were derived from the scRNA-seq mouse atlas (Pijuan-Sala et al., 2019) and are the mean expression calculated over all the cells assigned to each timepoint, with the data normalised and logp1-fold scaled. Dashed lines between timepoints represent linear interpolations to aid visualization.

Figure S5 – Wnt signalling is not primarily involved in mesoderm specification.

A - Gene expression heatmap (mean variance-stabilized read counts of replicates) of gastruloid bulk RNAseq samples at 48h and 72h. Mesenchymal markers, such as *Cdh2*, *Vim*, and *Snai1*, are upregulated upon CHIR treatment. Notably, *Cdh2* expression increases in a concentration-dependent manner, with the highest expression observed

with 20µM CHIR. In contrast, the expression of epithelial-related genes (*Cdh1* and *Epcam*) decreased when CHIR was used. Despite this classical EMT gene expression profile, the levels of key mesodermal-related genes remain low (e.g. *Tbx6*, *Msgn1, Mesp1* and *Tcf15)*, even when higher concentrations of CHIR. **B** – HCR (MIP) showing the lack of significant *Mesp1* and *Tbx6* expression in 72h gastruloids, including in 20µM CHIR treated gastruloids. The scale bar is equivalent to 50µm.

Figure S6 – Radial cell fate specification in Activin A treated gastruloids.

HCR medial optical sections unless stated otherwise (MIP). At 72 hours, Activin A induces a radial organisation in gastruloids as indicated by the expression of *Eomes* and *Tbxt* across most of the gastruloid except for its central region, which contains some *Sox2*-positive cells. *Oct4* is uniformly expressed throughout the gastruloid at this stage. The radial asymmetry becomes more pronounced by 96h, with a distinct core of *Sox2*-positive cells emerging in the centre, surrounded by *Tbxt*-expressing cells. *Eomes* expression is predominantly localised in the outer regions, while *Oct4* continues to be expressed throughout the gastruloid. At this stage, derivatives of the early PS are observed in the periphery of the Activin A treated gastruloids. These include *Noto* and *Foxa2*, marking node/notochord-like cells, *Foxa2* and *Sox17* for endodermal lineage, *Kdr* for haematoendothelial progenitors, and *Tbx6* and *Mesp1* for anterior mesoderm. Inlets provide magnified views of these expression patterns. "G" denotes gastruloid. The scale bar represents 50 µm.

Figure S7 – Gene expression comparison between Activin A and CHIR-treated gastruloids at 120h.

A – Heatmap showing gene expression (mean of variance-stabilized read counts across replicates) from bulk RNAseq of 120h gastruloids treated with (3 µM) CHIR or (high) Activin A. Unlike Activin A, the CHIR treatment leads to a gene expression signature characteristic of derivatives of the late PS. For example, genes associated with the caudal epiblast, such as *Cdx4*, *Fgf17*, and *Gdf11*, are exclusively expressed in CHIR treated gastruloids. Additionally, *Hoxc10* – a tailbud marker - is expressed only in gastruloids made with CHIR, suggesting they are more developmentally advanced and posteriorized compared to Activin A treated gastruloids. The expression of several Hox genes, including *Hoxc4*, *Hoxb6*, *Hoxd8*, and *Hoxa9*, in CHIR-treated gastruloids further supports this posteriorization. Conversely, the presence of *Wnt8a*, a late gastrulation marker, in Activin A but not CHIR treated gastruloids suggests a temporal delay in the first ones. Moreover, the expression of *Sox2*/*Pax6*, and *Uncx*/*Meox1*/*Lfng*/*Tbx6* in CHIR-treated gastruloids indicates the development of trunk neural and mesodermal-like cells, respectively. In contrast, Activin A treated gastruloids lack *Pax6* expression, which suggests an absence of trunk neural tissue. The mesodermal gene expression signature in these gastruloids includes *Tbx1*, indicative of its anterior identity. The expression of *Sox17*, *Foxa2*, *Noto*, *Gsc* and *Chrd* further emphasizes the anteriorization of Activin A treated gastruloids compared to those treated with CHIR. Additionally, *Kdr* and *Flt1* expression in Activin A treated gastruloids suggests the presence of endothelial-like cells. **B** - HCR (MIP) images illustrate gene expression in 120h gastruloids treated with Activin A or CHIR. *Pax1* is

expressed in the anterior compartment of CHIR treated gastruloids but absent in Activin A gastruloids. *Tbx6* is expressed in the posterior regions of both gastruloid types (indicated by red arrowheads), indicating mesoderm formation. *Mesp1* is expressed in the determination front of gastruloids made with CHIR (yellow arrowhead) but shows a distinct pattern in Activin A treated gastruloids, next to *Tbx1* expression. *Sox17*, *Foxa2* and *Noto* expression in Activin A treated gastruloids highlights the spatial distribution of endoderm and node/notochord-like cells, respectively. Expression of *Otx2* exclusively in gastruloids made with Activin A is also in accordance with the bulk RNAseq data. Scale bar is equivalent to 50µm.

Figure S8 – Comparative temporal *Hox* gene expression in CHIR and Activin A treated gastruloids.

No *Hox* gene expression is detected at 48h. By 72h, *Hoxa1* and *Hoxb1* begin to be expressed in CHIR treated gastruloids but not in gastruloids made with Activin A, even when a high dose is used. At 96 hours, Activin A treated gastruloids begin expressing *Hoxa1* and *Hoxb1*, though expression levels vary inversely with the dose of Activin. CHIR treated gastruloids exhibit at this stage a broader *Hox* gene profile, including the expression of some posterior ones like *Hoxb9*. At 120 hours, CHIR treated gastruloids show a highly posteriorized *Hox* gene signature, with significant levels of *Hoxc10 (*tailbud marker). In contrast, gastruloids treated with a low dose of Activin A express a *Hox* gene signature similar to that of 96h CHIR treated gastruloids, with notable expression of posterior *Hox* genes like *Hoxa5* and *Hoxb9*, but not *Hoxc5*, *Hoxc6*, or *Hoxa/c/d9*. Gastruloids exposed to a high dose of Activin A exhibit a more limited *Hox* gene expression, predominantly featuring anterior *Hox* genes such as *Hoxb1*, *Hoxa3* and *Hoxa5*. By 144 hours, gastruloids made with a low dose of Activin A show expression of several posterior *Hox* genes, though not *Hoxc10*. Meanwhile, gastruloids treated with high levels of Activin A continue to display a strong anterior *Hox* gene signature, with low levels of posterior *Hox* genes.

Figure S9 – Comparative gene expression analysis of gastruloids treated with low or high levels of Activin A.

A – Volcano plot highlighting the small difference between gastruloids treated with high versus low levels of Activin A (ActA) at 72 hours. Red dots represent differentially expressed genes (DEG) meeting the criteria of log fold change (LFC) greater than 1.5 and a p-value below 0.05; non-significant genes (NS; do not meet the previous conditions) are shown in grey. Out of 15,291 genes analysed, only 39 were considered significant based on these thresholds. **B** – Heatmap illustrating gene expression (mean of variance-stabilized read counts across replicates) from bulk RNAseq of 48h gastruloids and 72-144h gastruloids treated with either low or high levels of Activin A. At 72h, a slight upregulation of *Otx2*, *Nodal*, *Tdgf1*, *Tbxt*, *Fgf8*, *Evx1*, *Eomes*, and *Pitx2* expression is observed in gastruloids treated with a higher amount Activin A, with the most notable increase seen in the expression *Foxa2*. By 96h, the differences become more pronounced, with high Activin A treated gastruloids showing elevated expression of *Otx2*, *Lefty1/2*, *Tdgf1*, *Eomes*, *Lhx1*, *Dkk1*, *Msx1*, *Cer1*, *Gsc*, *Foxa2*, *Sox17*, *Gata4*

indicative of a more anterior patterning. In contrast, gastruloids treated with low Activin A upregulate genes associated with the late PS, like *Wnt3a*, *Cdx1*, *Cdx2* and *Meis2*. Of note, Activin A treated gastruloids express higher levels of *Cyp26a1*, which is involved in the degradation of retinoic acid (RA). At 120h, the divergence between the two conditions becomes even more distinct. Gastruloids exposed to higher levels of Nodal signalling continue to express higher levels of genes associated with the development of anterior body structures like *Dkk1* and *Lhx1*, as well as endothelial-related genes like *Kdr* and *Flt1*. The higher levels of *Foxa2*, *Sox17*, *Shh* and *Nog* expression in gastruloids developed with a high dose of Activin A suggests the presence of more endoderm and notochord-like cells in these gastruloids. Conversely, low Activin A treated gastruloids show increased expression of *Wnt3a*, *Wnt8a*, *Cdx1/2/4*, *Fgf17*, and *Nkx1-2*, markers linked to the late PS. Gastruloids treated with higher Activin A continue to maintain significantly higher levels of *Cyp26a1*, suggesting greater protection against RA activity in comparison to those treated with a lower Activin A dose. High levels of *Lefty1* in 120h gastruloids made with a high Activin A suggest again that Nodal signalling activity might be more elevated in comparison to low Activin A treated gastruloids.

Figure S10 – Nodal and Wnt signalling activity in gastruloids treated with different levels of Activin A.

A – At 72h, higher levels of Nodal activity (orange) are observed when a higher dose of Activin A was used to make gastruloids. In contrast, Wnt signalling (cyan) is more active in low Activin A treated gastruloids. **B** – Both types of Activin A treated gastruloids display a polarized domain of Wnt activity at 96h. Nodal activity is increased in gastruloids treated with high levels of Activin A and is more prominent in the outer part of the gastruloids. **C** – At 120h, we observed a mixture of elongated and round gastruloids. Elongated gastruloids were more frequently observed in the low Activin A condition and they showed strong Wnt levels posteriorly and some Nodal activity in the anterior compartment. Non-elongated gastruloids were mostly observed in the high Activin A condition and they exhibited significant Nodal activity in the outer shell, more pronounced towards one side of the gastruloid, and low levels of Wnt signalling. Scale bar is equivalent to 100µm.

Figures

Figure 1

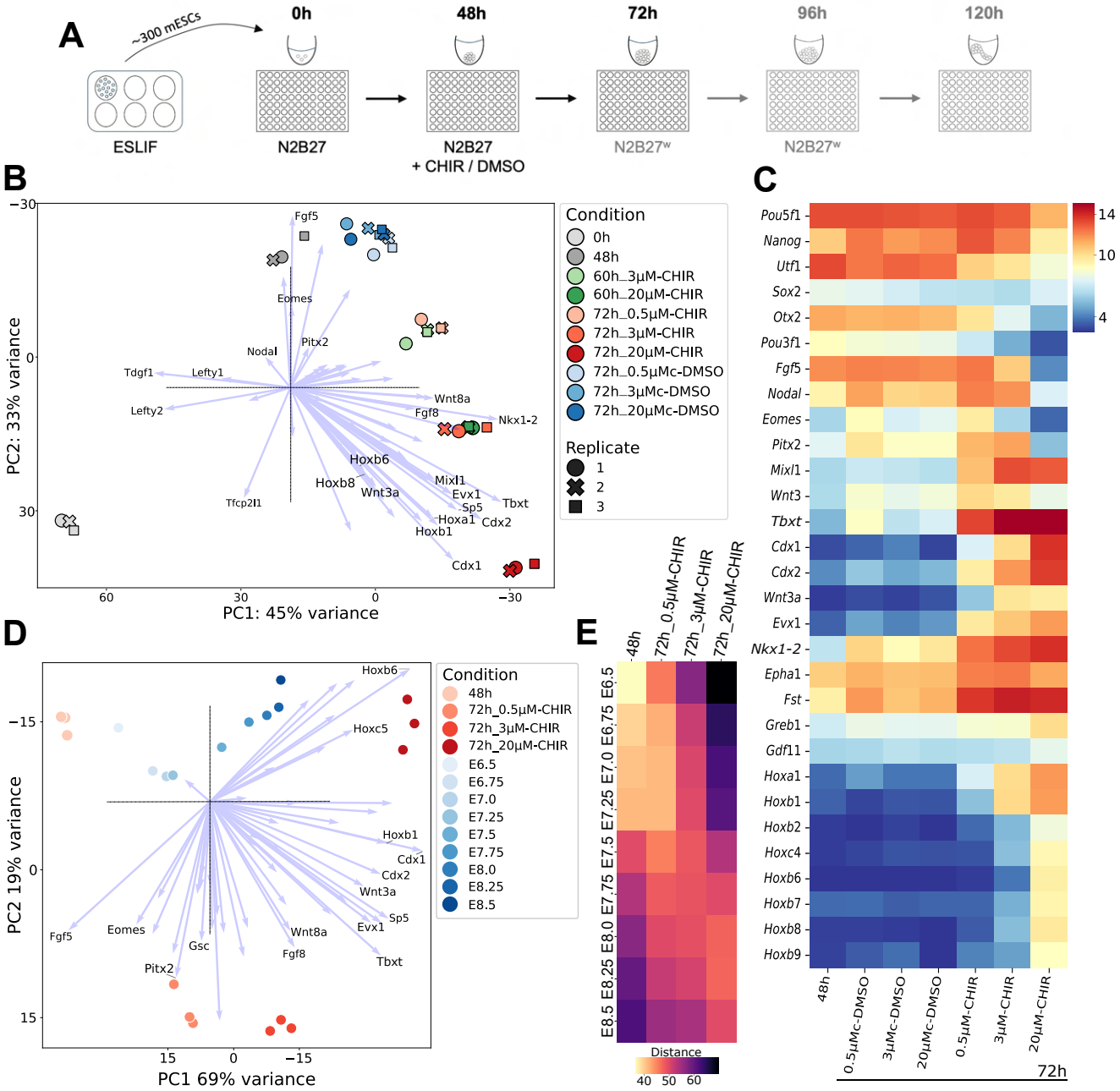


Figure 2

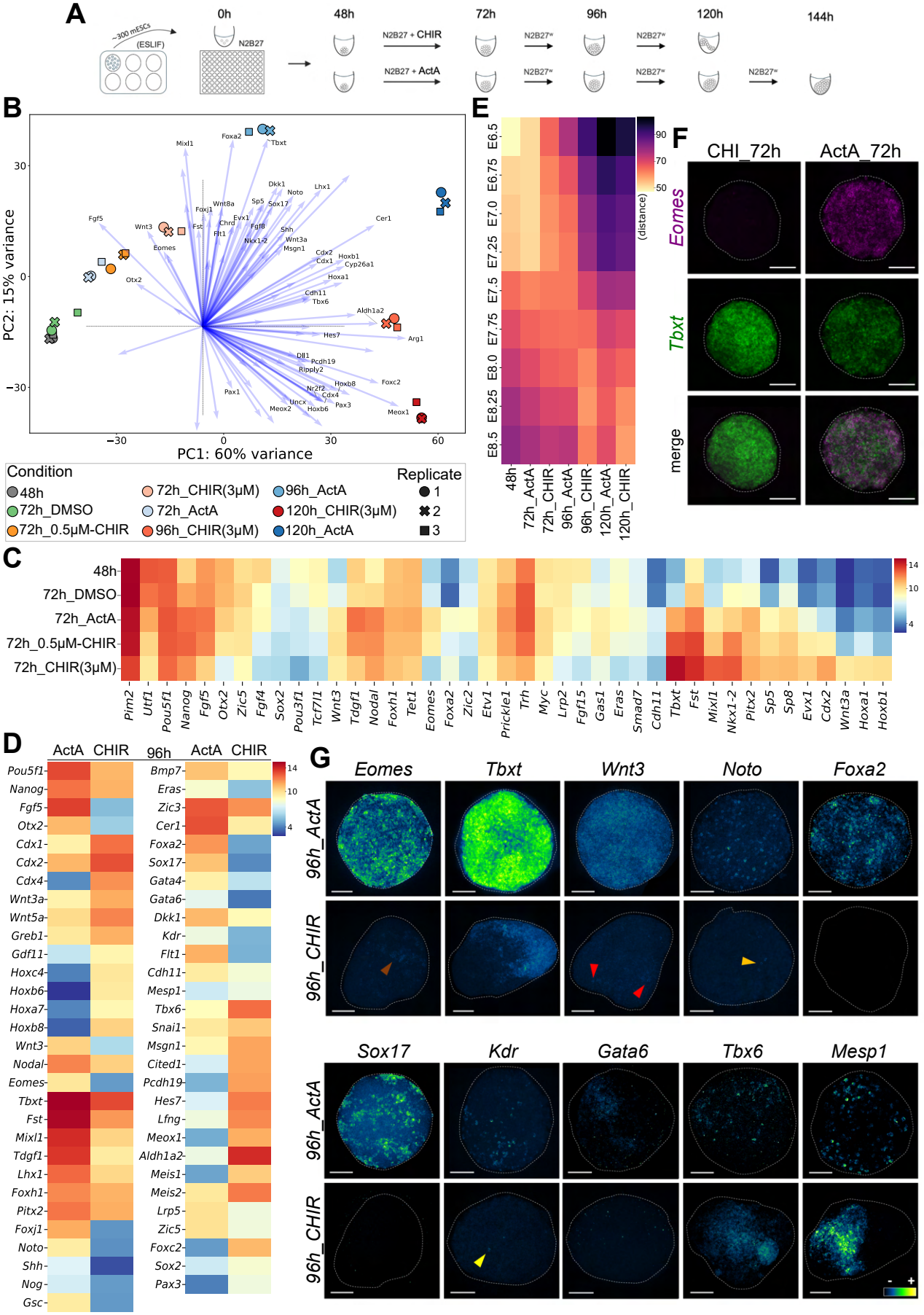

A
~300 mESCs
(ESLIF)
N2B27
0h
48h
72h
96h
120h
144h
N2B27 + CHIR
N2B27 + ActA
B
PC2: 15% variance
PC1: 60% variance
Condition
48h
72h_DMSO
72h_0.5μM-CHIR
72h_CHIR(3μM)
72h_ActA
96h_CHIR(3μM)
96h_ActA
120h_CHIR(3μM)
120h_ActA
Replicate
1
2
3
E
(distance)
48h
72h_ActA
72h_CHIR
96h_ActA
96h_CHIR
120h_ActA
120h_CHIR
E6.5
E6.75
E7.0
E7.25
E7.5
E7.75
E8.0
E8.25
E8.5
F
CHI_72h
ActA_72h
Eomes
Tbxt
merge
C
48h
72h_DMSO
72h_ActA
72h_0.5μM-CHIR
72h_CHIR(3μM)
Pim2 Utf1 Pou5f1 Nanog Fgf5 Otx2 Zic5 Fgf4 Sox2 Pou3f1 Tcf7l1 Wnt3 Tdgf1 Nodal Foxh1 Tet1 Eomes Foxa2 Zic2 Etv1 Prickle1 Trh Myc Lrp2 Fgf15 Gas1 Eras Smad7 Cdh11 Tbxt Fst Mixl1 Nkx1-2 Pitx2 Sp5 Sp8 Evx1 Cdx2 Wnt3a Hoxa1 Hoxb1
D
ActA CHIR
96h ActA CHIR
Pou5f1 Nanog Fgf5 Otx2 Cdx1 Cdx2 Cdx4 Wnt3a Wnt5a Greb1 Gdf11 Hoxc4 Hoxb6 Hoxa7 Hoxb8 Wnt3 Nodal Eomes Tbxt Fst Mixl1 Tdgf1 Lhx1 Foxh1 Pitx2 Foxj1 Noto Shh Nog Gsc
Bmp7 Eras Zic3 Cer1 Foxa2 Sox17 Gata4 Gata6 Dkk1 Kdr Flt1 Cdh11 Mesp1 Tbx6 Snai1 Msgn1 Cited1 Pcdh19 Hes7 Lfng Meox1 Aldh1a2 Meis1 Meis2 Lrp5 Zic5 Foxc2 Sox2 Pax3
G
Eomes
Tbxt
Wnt3
Noto
Foxa2
96h_ActA
96h_CHIR
Sox17
Kdr
Gata6
Tbx6
Mesp1

Figure 3

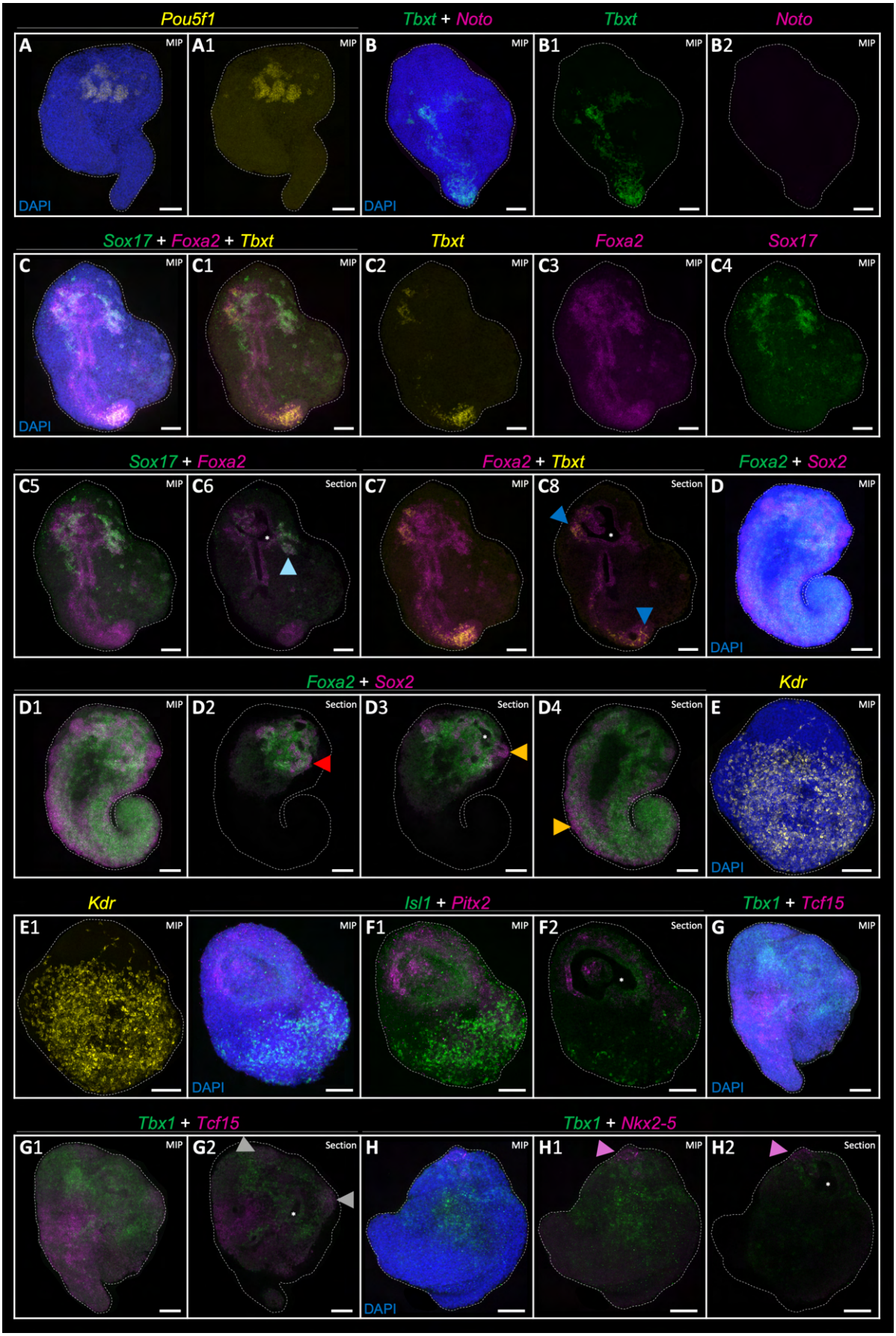

Pou5f1
Tbxt + Noto
Tbxt
Noto
A
MIP
A1
MIP
B
MIP
B1
MIP
B2
MIP
DAPI
DAPI
Sox17 + Foxa2 + Tbxt
Tbxt
Foxa2
Sox17
C
MIP
C1
MIP
C2
MIP
C3
MIP
C4
MIP
DAPI
Sox17 + Foxa2
Foxa2 + Tbxt
Foxa2 + Sox2
C5
MIP
C6
Section
C7
MIP
C8
Section
D
MIP
DAPI
Foxa2 + Sox2
Kdr
D1
MIP
D2
Section
D3
Section
D4
Section
E
MIP
DAPI
Kdr
Isl1 + Pitx2
Tbx1 + Tcf15
E1
MIP
MIP
DAPI
F1
MIP
F2
Section
G
MIP
DAPI
Tbx1 + Tcf15
Tbx1 + Nkx2-5
G1
MIP
G2
Section
H
MIP
DAPI
H1
MIP
H2
Section

Figure 4

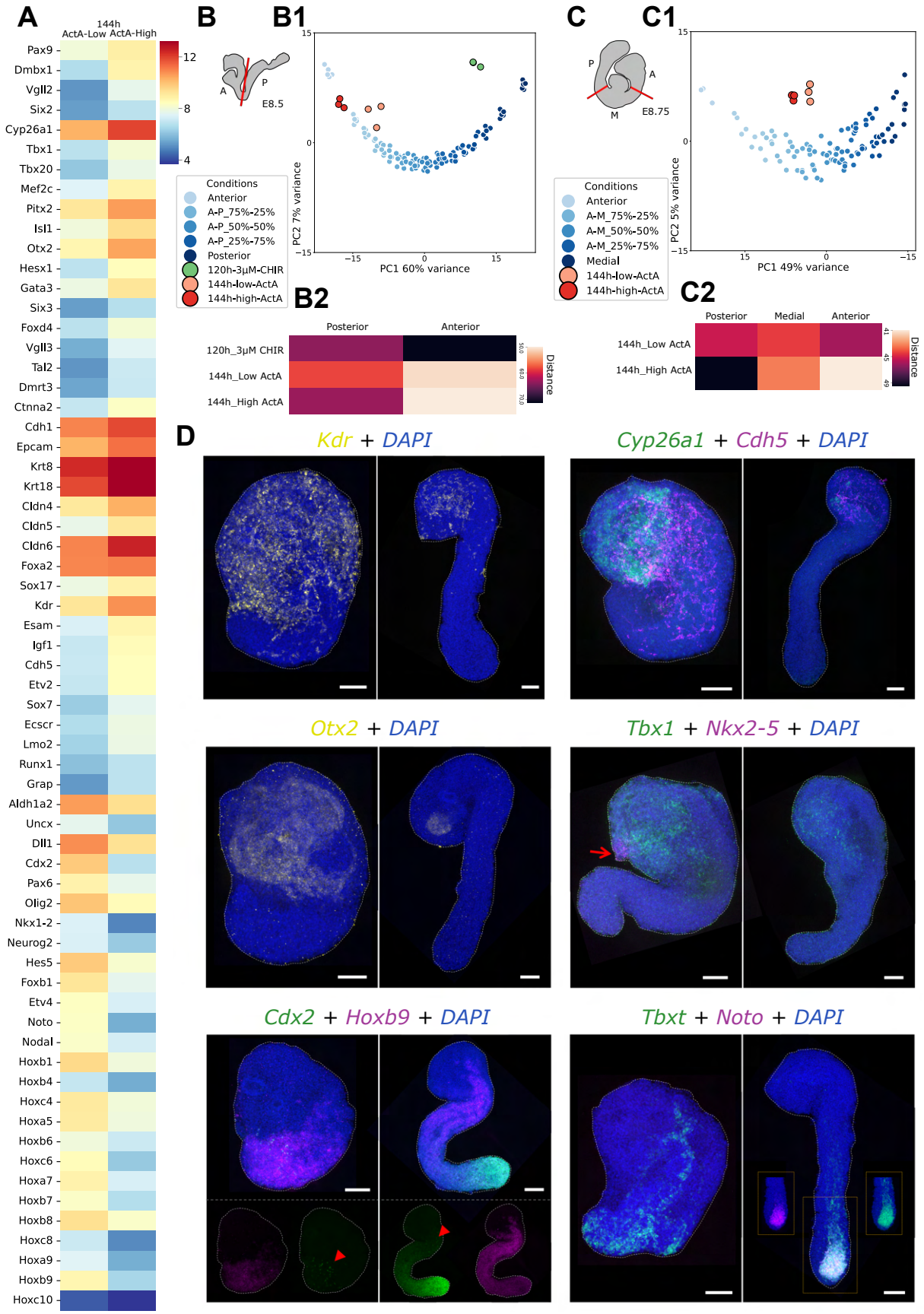
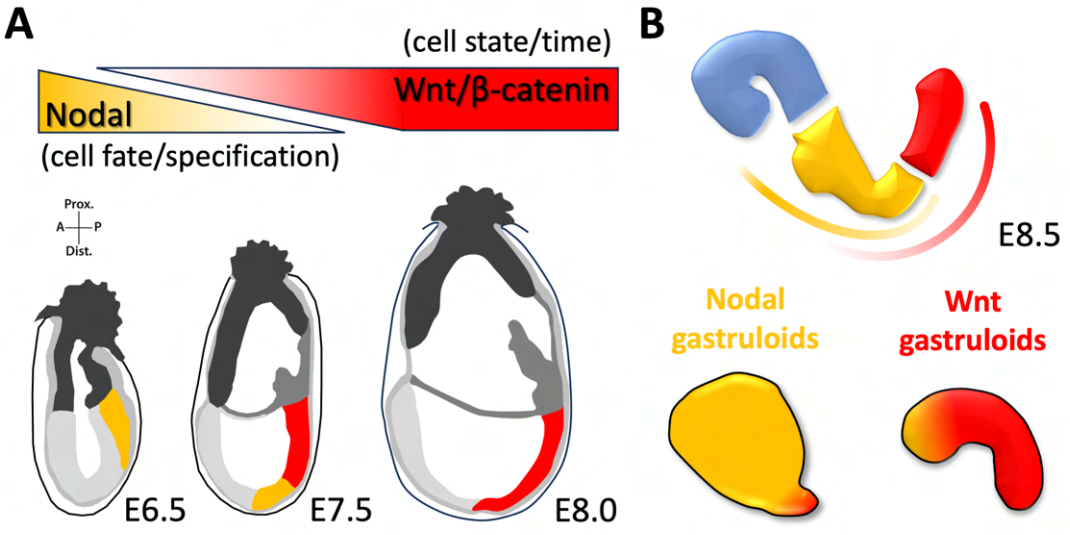

A
144h
ActA-Low ActA-High
Pax9
Dmbx1
Vgll2
Six2
Cyp26a1
Tbx1
Tbx20
Mef2c
Pitx2
Isl1
Otx2
Hesx1
Gata3
Six3
Foxd4
Vgll3
Tal2
Dmrt3
Ctnna2
Cdh1
Epcam
Krt8
Krt18
Cldn4
Cldn5
Cldn6
Foxa2
Sox17
Kdr
Esam
Igf1
Cdh5
Etv2
Sox7
Ecscr
Lmo2
Runx1
Grap
Aldh1a2
Uncx
Dll1
Cdx2
Pax6
Olig2
Nkx1-2
Neurog2
Hes5
Foxb1
Etv4
Noto
Nodal
Hoxb1
Hoxb4
Hoxc4
Hoxa5
Hoxb6
Hoxc6
Hoxa7
Hoxb7
Hoxb8
Hoxc8
Hoxa9
Hoxb9
Hoxc10
12
8
4
B
A
P
E8.5
B1
Conditions
Anterior
A-P_75%-25%
A-P_50%-50%
A-P_25%-75%
Posterior
120h-3μM-CHIR
144h-low-ActA
144h-high-ActA
PC2 7% variance
PC1 60% variance
B2
Posterior
Anterior
120h_3μM CHIR
144h_Low ActA
144h_High ActA
Distance
C
P
A
M
E8.75
C1
Conditions
Anterior
A-M_75%-25%
A-M_50%-50%
A-M_25%-75%
Medial
144h-low-ActA
144h-high-ActA
PC2 5% variance
PC1 49% variance
C2
Posterior
Medial
Anterior
144h_Low ActA
144h_High ActA
Distance
D
Kdr + DAPI
Cyp26a1 + Cdh5 + DAPI
Otx2 + DAPI
Tbx1 + Nkx2-5 + DAPI
Cdx2 + Hoxb9 + DAPI
Tbxt + Noto + DAPI

Figure 5


A
(cell state/time)
Wnt/β-catenin
Nodal
(cell fate/specification)
Prox.
A
P
Dist.
E6.5
E7.5
E8.0
B
E8.5
Nodal gastruloids
Wnt gastruloids

Supplementary Figures

Figure S1

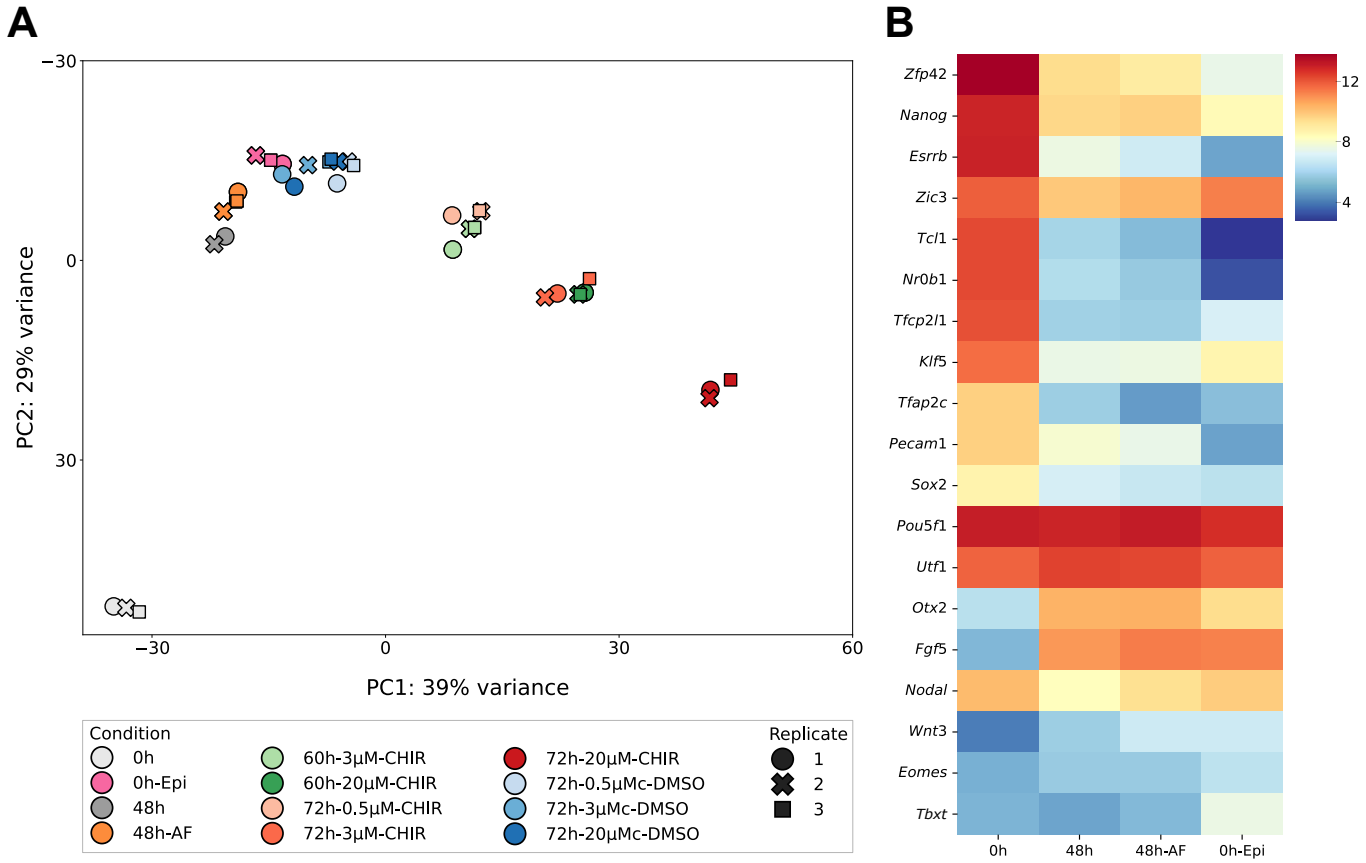


Figure S2

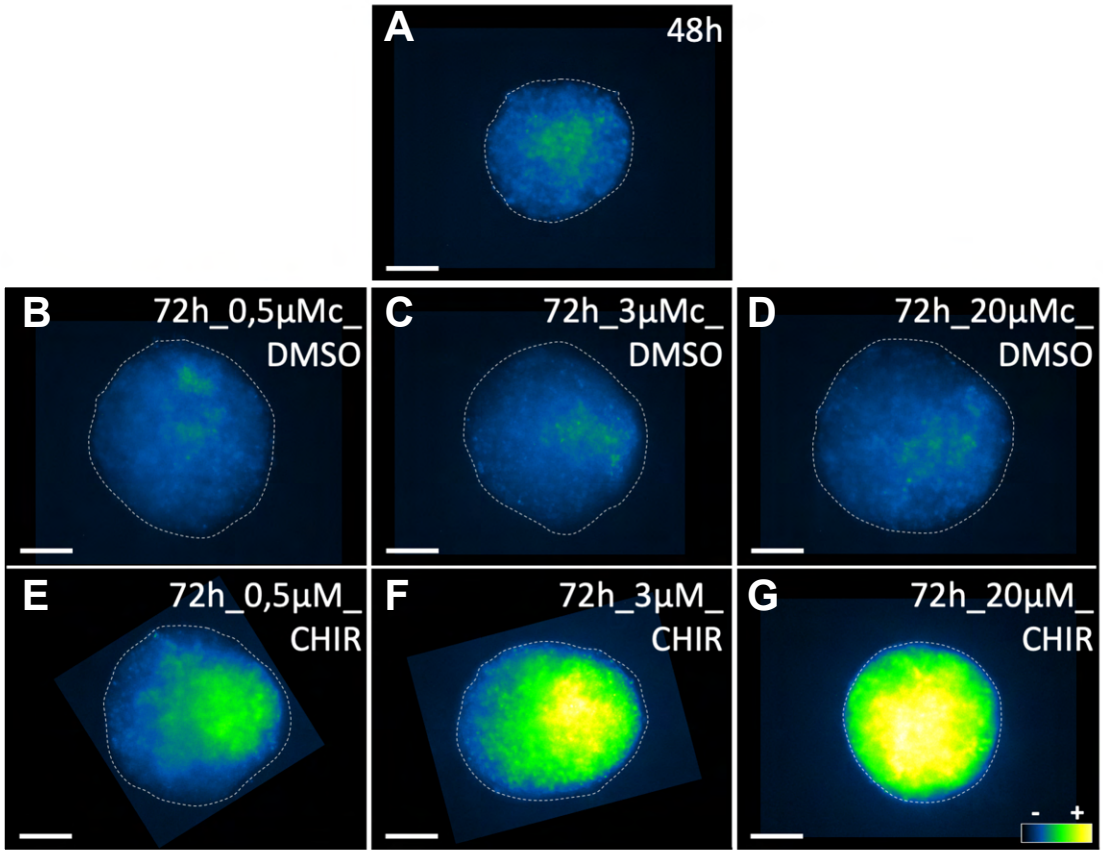

A
48h
B
72h_0,5µMc_
DMSO
C
72h_3µMc_
DMSO
D
72h_20µMc_
DMSO
E
72h_0,5µM_
CHIR
F
72h_3µM_
CHIR
G
72h_20µM_
CHIR
-
+

Figure S3

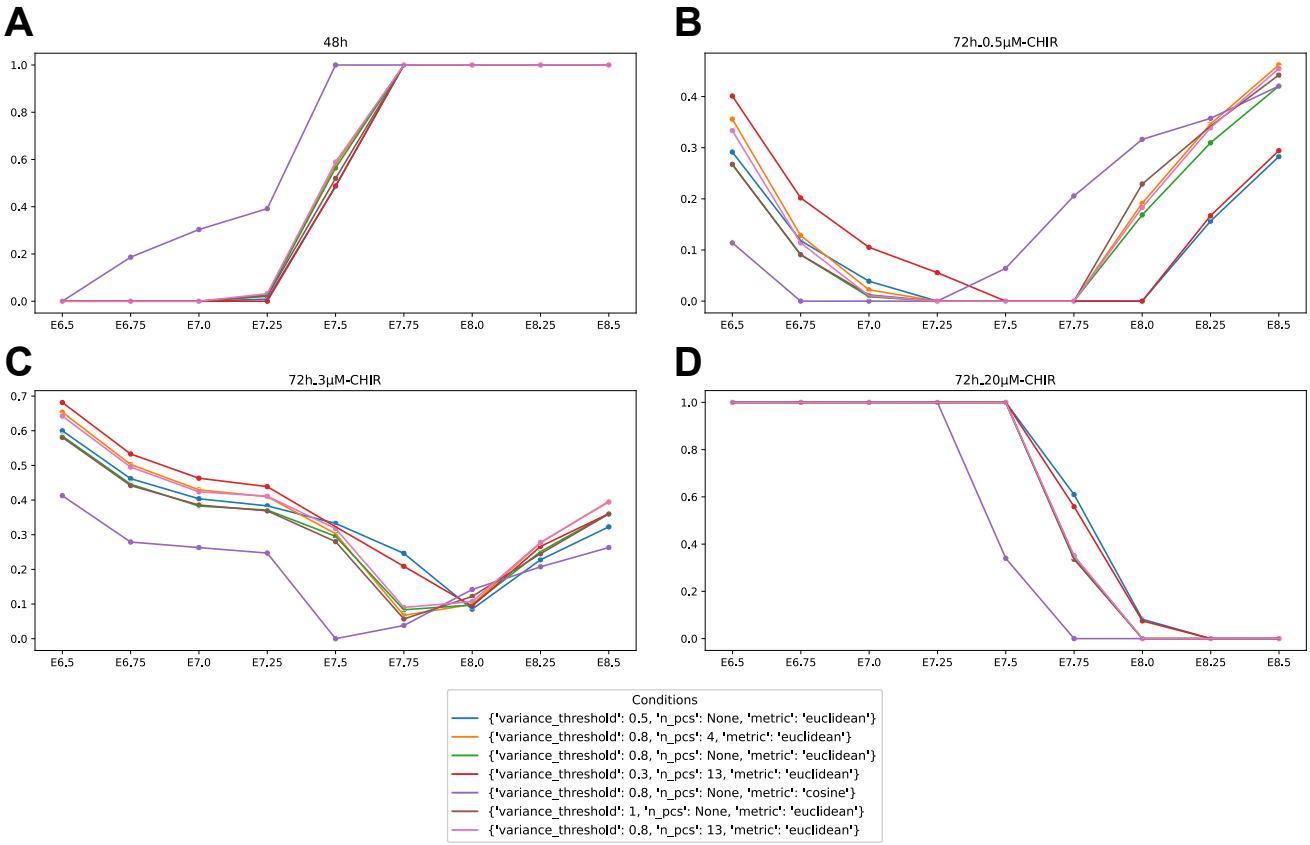

A
48h
B
72h_0.5μM-CHIR
C
72h_3μM-CHIR
D
72h_20μM-CHIR
E6.5
E6.75
E7.0
E7.25
E7.5
E7.75
E8.0
E8.25
E8.5
Conditions
{'variance_threshold': 0.5, 'n_pcs': None, 'metric': 'euclidean'}
{'variance_threshold': 0.8, 'n_pcs': 4, 'metric': 'euclidean'}
{'variance_threshold': 0.8, 'n_pcs': None, 'metric': 'euclidean'}
{'variance_threshold': 0.3, 'n_pcs': 13, 'metric': 'euclidean'}
{'variance_threshold': 0.8, 'n_pcs': None, 'metric': 'cosine'}
{'variance_threshold': 1, 'n_pcs': None, 'metric': 'euclidean'}
{'variance_threshold': 0.8, 'n_pcs': 13, 'metric': 'euclidean'}

Figure S4

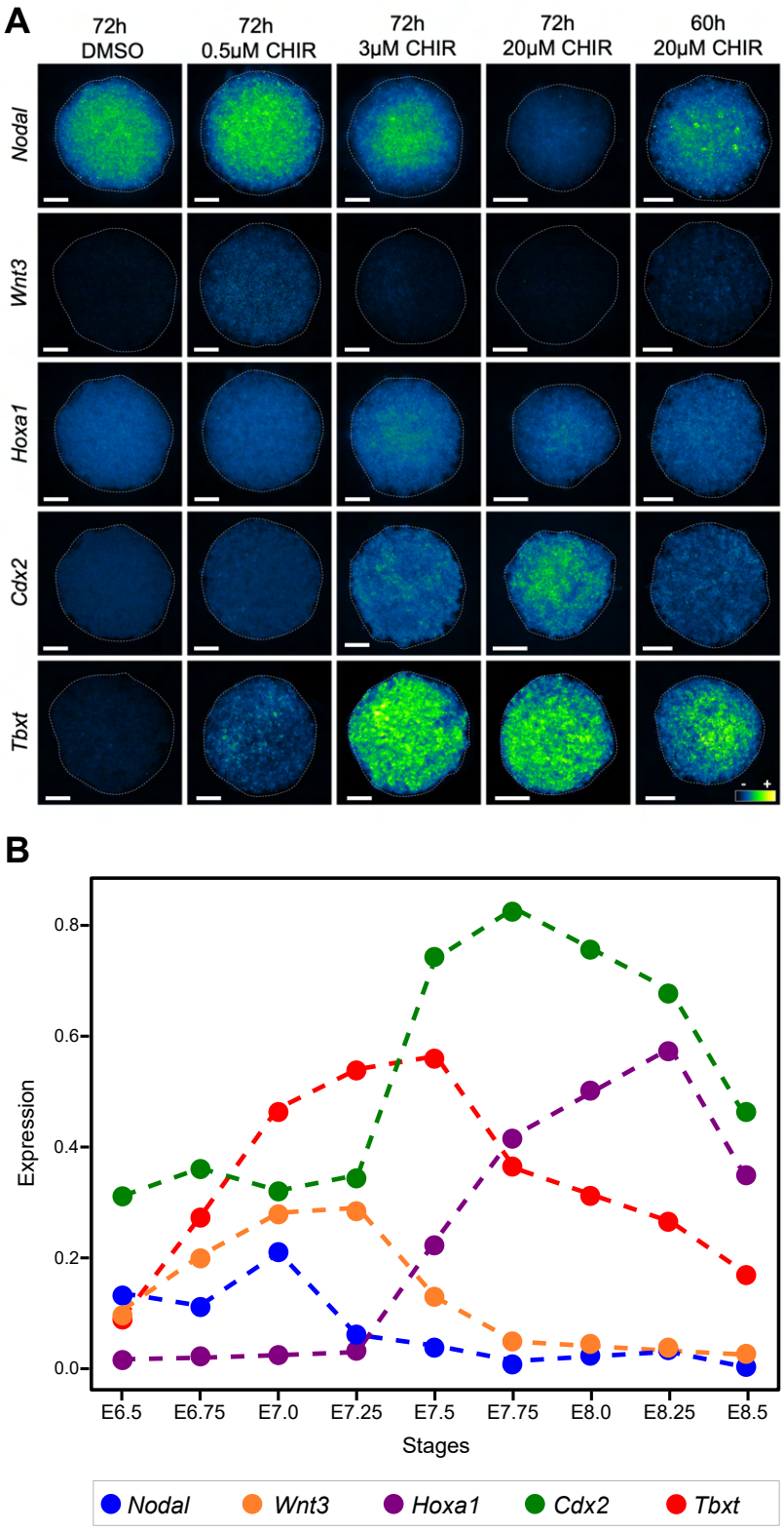

A
72h DMSO
72h 0.5μM CHIR
72h 3μM CHIR
72h 20μM CHIR
60h 20μM CHIR
Nodal
Wnt3
Hoxa1
Cdx2
Tbxt
-
+
B
Expression
0.0
0.2
0.4
0.6
0.8
E6.5
E6.75
E7.0
E7.25
E7.5
E7.75
E8.0
E8.25
E8.5
Stages
Nodal
Wnt3
Hoxa1
Cdx2
Tbxt

Figure S5

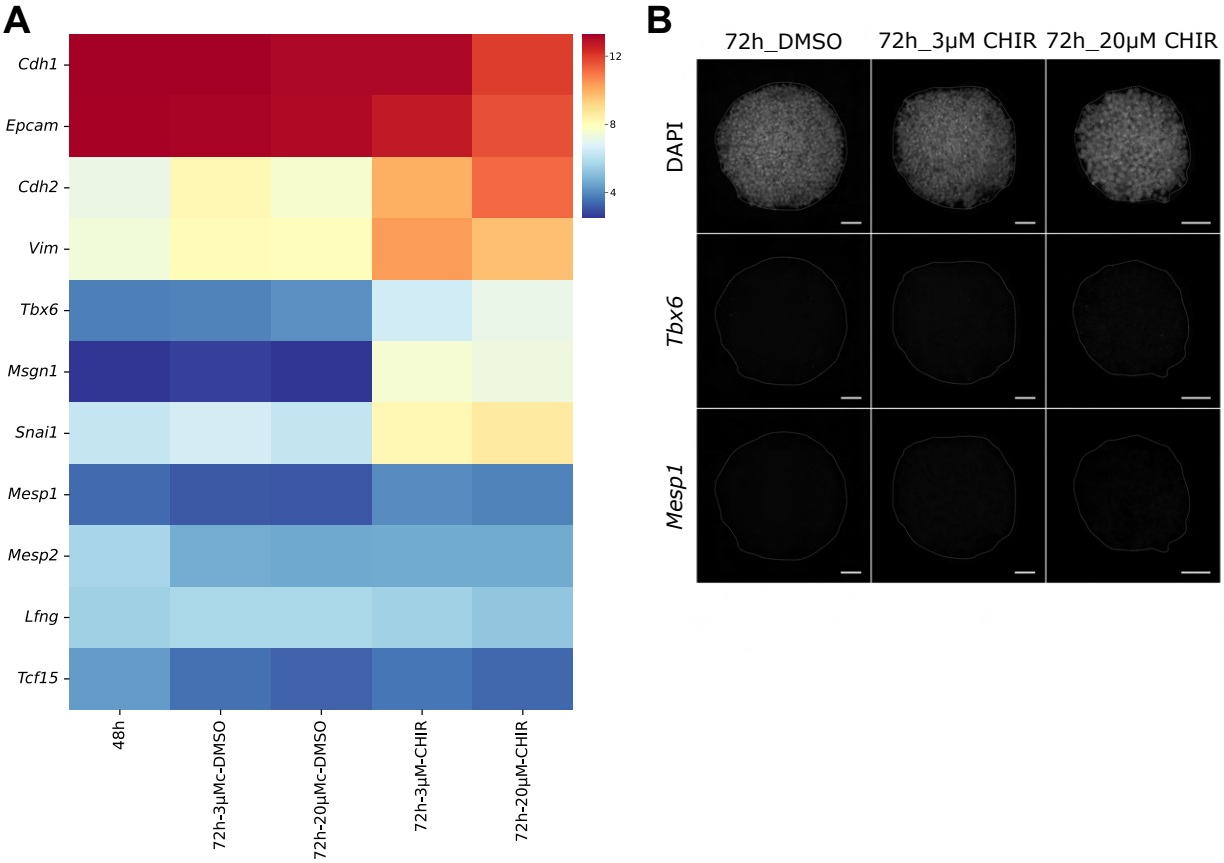

A
Cdh1
Epcam
Cdh2
Vim
Tbx6
Msgn1
Snai1
Mesp1
Mesp2
Lfng
Tcf15
48h
72h-3μMc-DMSO
72h-20μMc-DMSO
72h-3μM-CHIR
72h-20μM-CHIR
12
8
4
B
72h_DMSO
72h_3μM CHIR
72h_20μM CHIR
DAPI
Tbx6
Mesp1

Figure S6

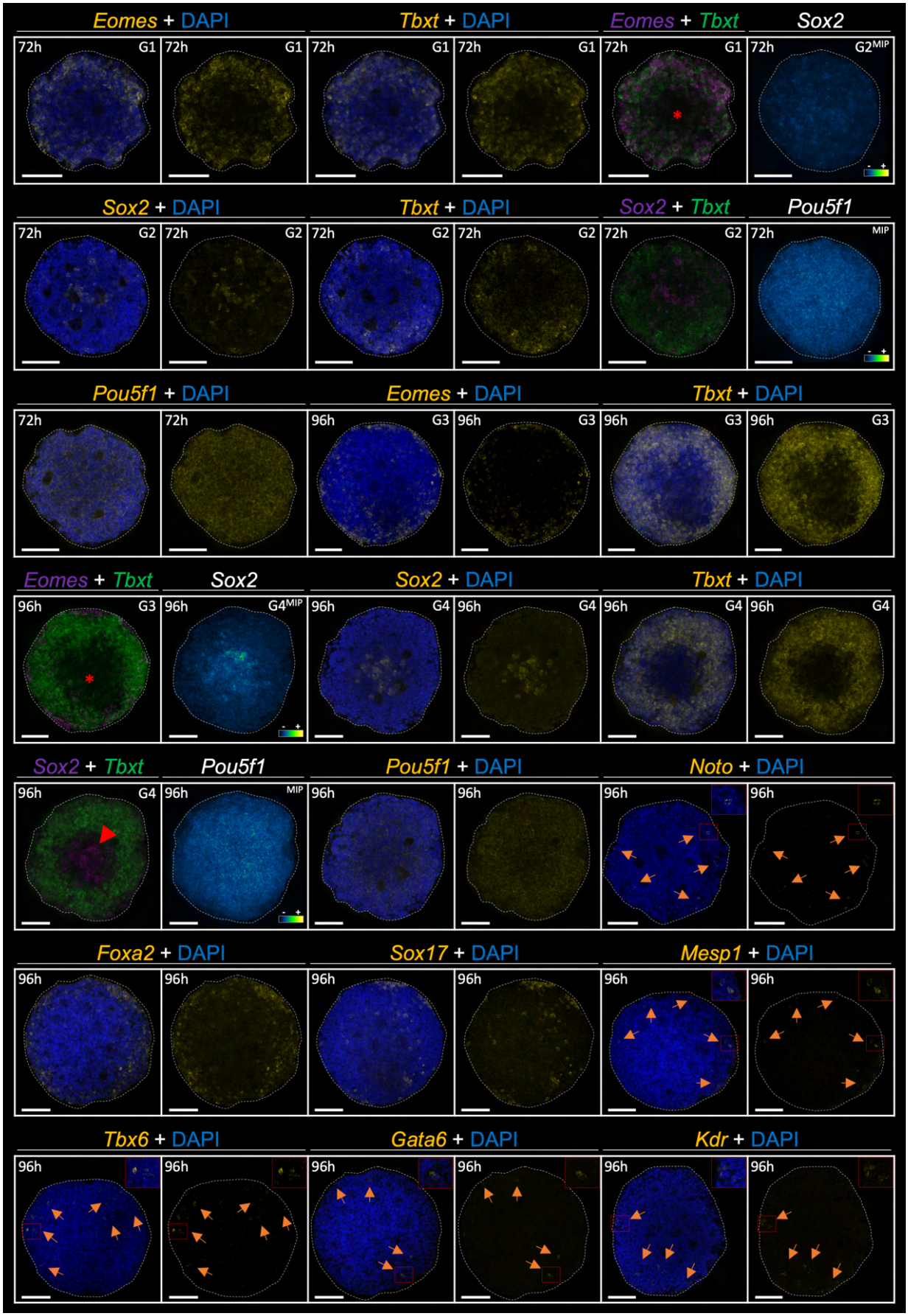

Eomes + DAPI
Tbxt + DAPI
Eomes + Tbxt
Sox2
Sox2 + DAPI
Tbxt + DAPI
Sox2 + Tbxt
Pou5f1
Pou5f1 + DAPI
Eomes + DAPI
Tbxt + DAPI
Eomes + Tbxt
Sox2
Sox2 + DAPI
Tbxt + DAPI
Sox2 + Tbxt
Pou5f1
Pou5f1 + DAPI
Noto + DAPI
Foxa2 + DAPI
Sox17 + DAPI
Mesp1 + DAPI
Tbx6 + DAPI
Gata6 + DAPI
Kdr + DAPI
72h
96h
G1
G2
G3
G4
G2MIP
G4MIP
MIP

Figure S7

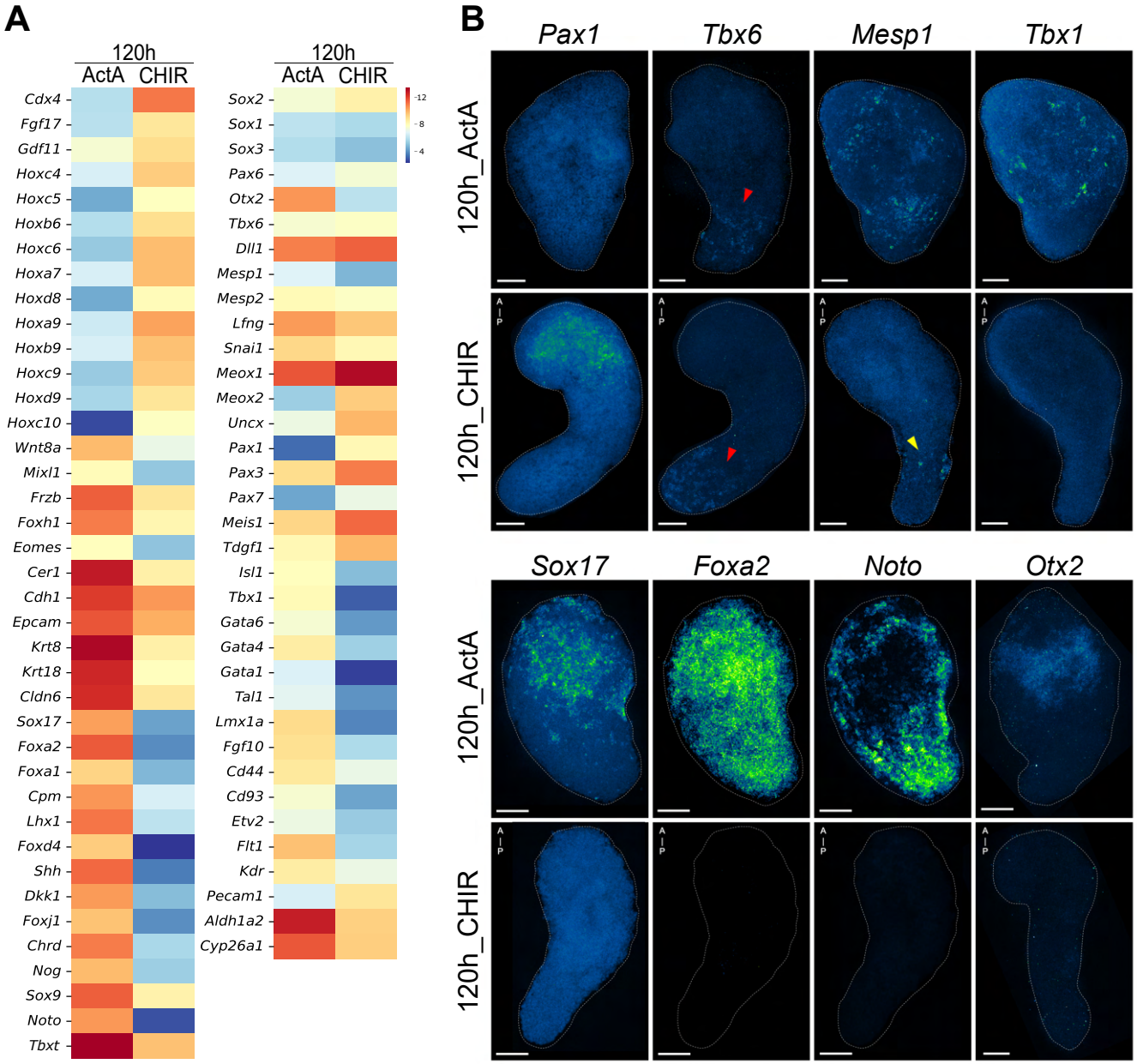

A
120h
ActA
CHIR
Cdx4
Fgf17
Gdf11
Hoxc4
Hoxc5
Hoxb6
Hoxc6
Hoxa7
Hoxd8
Hoxa9
Hoxb9
Hoxc9
Hoxd9
Hoxc10
Wnt8a
Mixl1
Frzb
Foxh1
Eomes
Cer1
Cdh1
Epcam
Krt8
Krt18
Cldn6
Sox17
Foxa2
Foxa1
Cpm
Lhx1
Foxd4
Shh
Dkk1
Foxj1
Chrd
Nog
Sox9
Noto
Tbxt
Sox2
Sox1
Sox3
Pax6
Otx2
Tbx6
Dll1
Mesp1
Mesp2
Lfng
Snai1
Meox1
Meox2
Uncx
Pax1
Pax3
Pax7
Meis1
Tdgf1
Isl1
Tbx1
Gata6
Gata4
Gata1
Tal1
Lmx1a
Fgf10
Cd44
Cd93
Etv2
Flt1
Kdr
Pecam1
Aldh1a2
Cyp26a1
-12
-8
-4
B
Pax1
Tbx6
Mesp1
Tbx1
120h_ActA
120h_CHIR
A
P
Sox17
Foxa2
Noto
Otx2

Figure S8

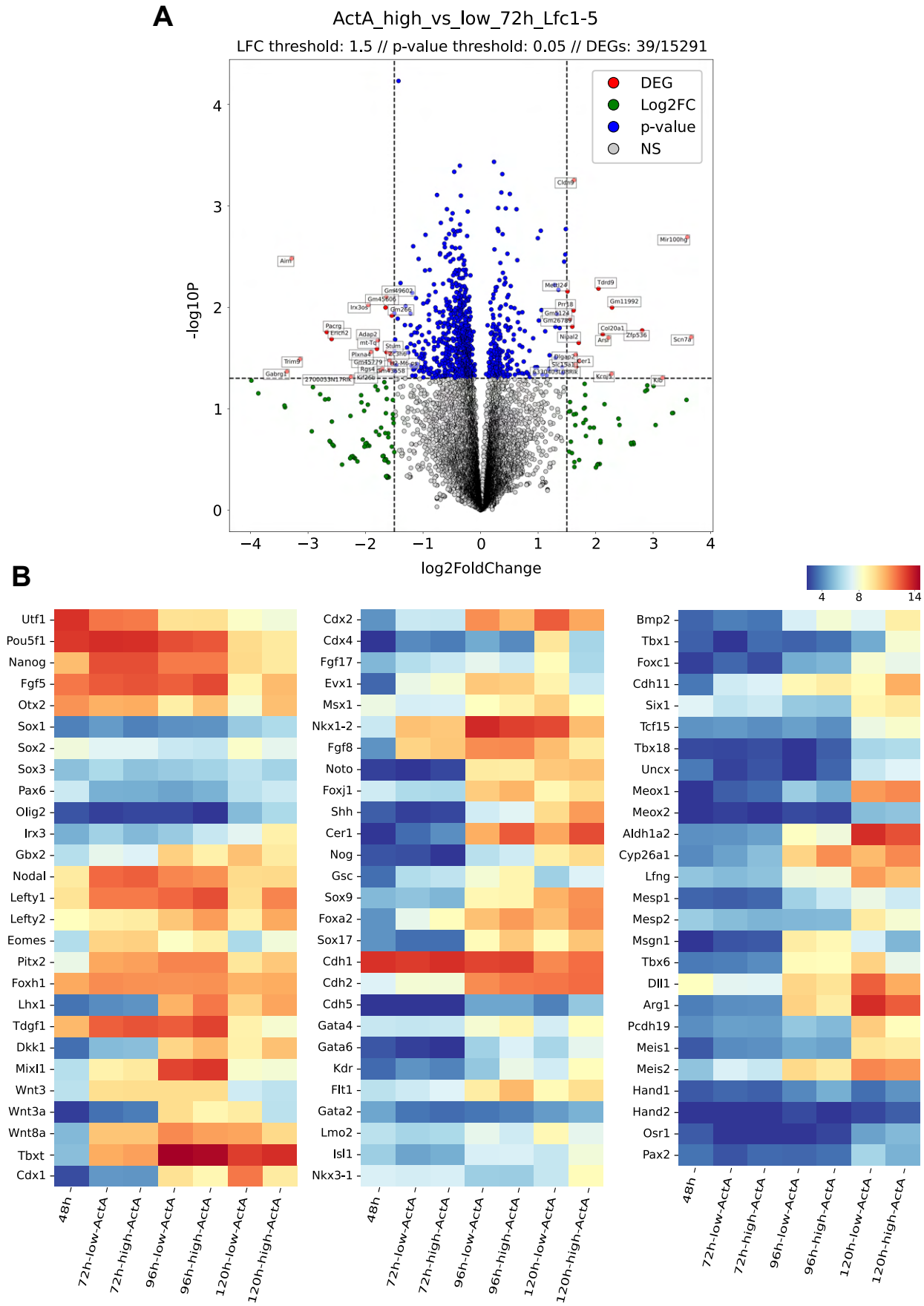

A
ActA_high_vs_low_72h_Lfc1-5
LFC threshold: 1.5 // p-value threshold: 0.05 // DEGs: 39/15291
DEG
Log2FC
p-value
NS
-log10P
log2FoldChange
B
Utf1
Pou5f1
Nanog
Fgf5
Otx2
Sox1
Sox2
Sox3
Pax6
Olig2
Irx3
Gbx2
Nodal
Lefty1
Lefty2
Eomes
Pitx2
Foxh1
Lhx1
Tdgf1
Dkk1
Mixl1
Wnt3
Wnt3a
Wnt8a
Tbxt
Cdx1
Cdx2
Cdx4
Fgf17
Evx1
Msx1
Nkx1-2
Fgf8
Noto
Foxj1
Shh
Cer1
Nog
Gsc
Sox9
Foxa2
Sox17
Cdh1
Cdh2
Cdh5
Gata4
Gata6
Kdr
Flt1
Gata2
Lmo2
Isl1
Nkx3-1
Bmp2
Tbx1
Foxc1
Cdh11
Six1
Tcf15
Tbx18
Uncx
Meox1
Meox2
Aldh1a2
Cyp26a1
Lfng
Mesp1
Mesp2
Msgn1
Tbx6
Dll1
Arg1
Pcdh19
Meis1
Meis2
Hand1
Hand2
Osr1
Pax2
48h
72h-low-ActA
72h-high-ActA
96h-low-ActA
96h-high-ActA
120h-low-ActA
120h-high-ActA
4
8
14

Figure S9

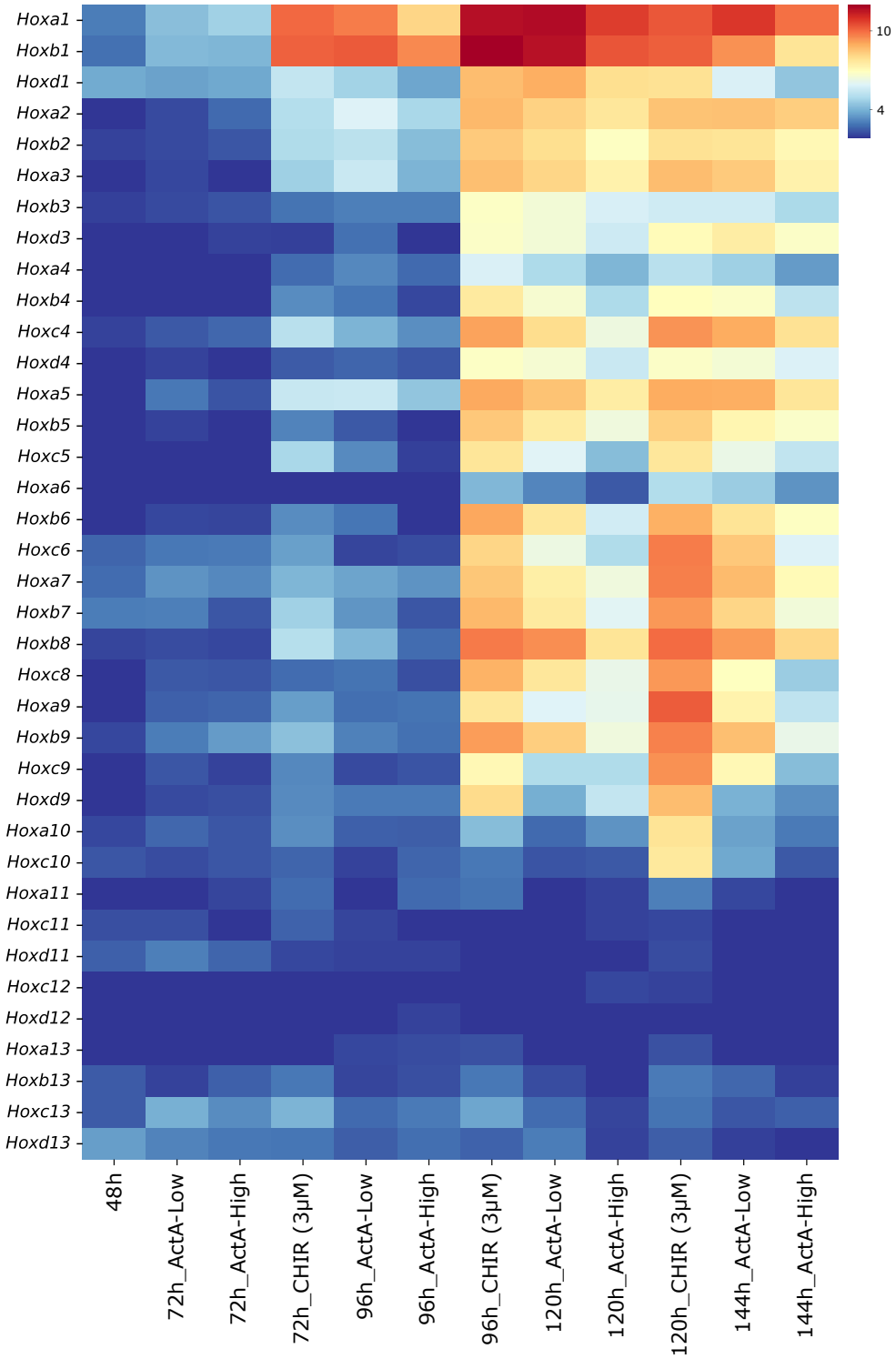
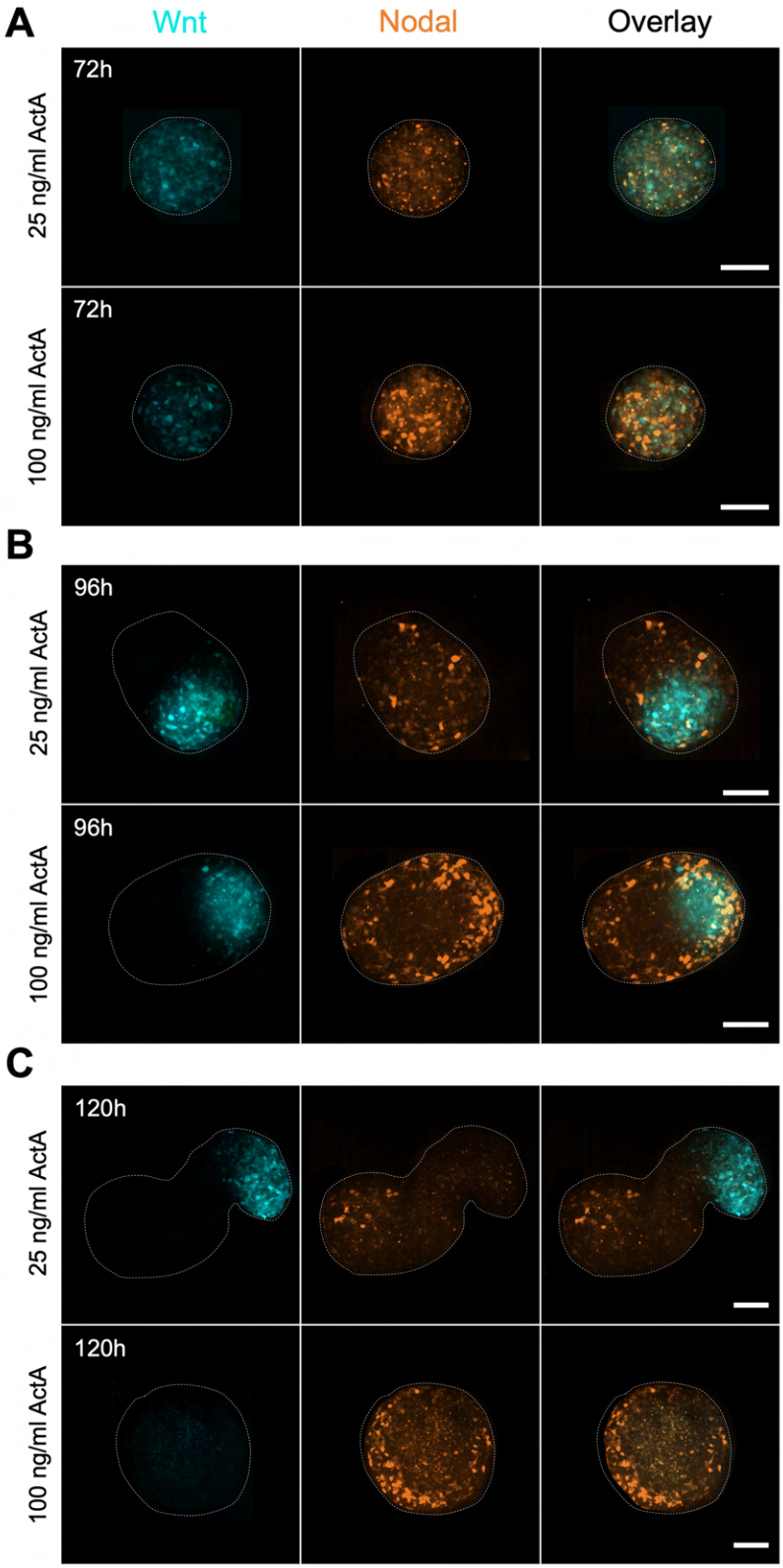

Hoxa1
Hoxb1
Hoxd1
Hoxa2
Hoxb2
Hoxa3
Hoxb3
Hoxd3
Hoxa4
Hoxb4
Hoxc4
Hoxd4
Hoxa5
Hoxb5
Hoxc5
Hoxa6
Hoxb6
Hoxc6
Hoxa7
Hoxb7
Hoxb8
Hoxc8
Hoxa9
Hoxb9
Hoxc9
Hoxd9
Hoxa10
Hoxc10
Hoxa11
Hoxc11
Hoxd11
Hoxc12
Hoxd12
Hoxa13
Hoxb13
Hoxc13
Hoxd13
48h
72h_ActA-Low
72h_ActA-High
72h_CHIR (3μM)
96h_ActA-Low
96h_ActA-High
96h_CHIR (3μM)
120h_ActA-Low
120h_ActA-High
120h_CHIR (3μM)
144h_ActA-Low
144h_ActA-High
10
4

Figure S10


A
Wnt
Nodal
Overlay
72h
25 ng/ml ActA
72h
100 ng/ml ActA
B
96h
25 ng/ml ActA
96h
100 ng/ml ActA
C
120h
25 ng/ml ActA
120h
100 ng/ml ActA