nature portfolio

https://doi.org/10.1038/s41586-021-04158-y

Supplementary information

Single-cell transcriptomic characterization of a gastrulating human embryo

In the format provided by the authors and unedited

Supplementary Notes

Supplementary Note 1 - Annotation of gastrula cell types

The Epiblast could be detected by the expression of *SOX2, OTX2, CDH1* and was represented in both the caudal and rostral regions of the embryo (55% caudal, 45% rostral, 0% yolk sac; Figure 1c, Extended Data Figure 2a and SI Table 2). In contrast, the Ectoderm (Amniotic/Embryonic) came predominantly from the rostral portion of the embryo and did not express pluripotency markers but was characterised by high expression of key markers such as *DLX5*, *TFAP2A* and *GATA3,* representative of both the extra-embryonic ectoderm of the amnion as well as the embryonic ectoderm at the rostral boundary of the neural plate[1,2] ((Figure 1d, and Extended Data Figure 2b).

The Primitive Streak was identified by the archetypal marker *TBXT* (*Brachyury*) in combination with *CDH1* and *FST* (Figure 1c and Extended Data Figure 2a). As expected, these cells originated almost exclusively from the caudal portion of the embryo (Figure 1d, and Extended Data Figure 2b). While the majority of Nascent Mesoderm cells were also located in the caudal region and expressed *TBXT*, they could be distinguished from PS cells by the expression of key mesodermal markers *MESP1 and PDGFRA* (Figure 1d, Extended Data Figure 2a and 10a). This co-expression of both PS and mesoderm markers led us to define this mesoderm as 'nascent', representing the forming mesoderm cells in the process of delaminating from the PS (Figure 1c). Axial Mesoderm could be detected by the expression of *TBXT*, *CHRD* and *NOTO* (Figure 1c, and Extended Data Figure 2a).

Two other clusters of embryonic mesoderm could be distinguished by their relative degree of maturation and location within the embryo. We annotated the first as Emergent Mesoderm since it expressed the highest levels of *MESP1* but was negative for *TBXT*, thereby representing a transition from the Nascent Mesoderm towards the more mature Advanced Mesoderm (Figure 1c, Extended Data Figure 2a and 10b). It also expressed *LHX1* and *OTX2* as well as the highest levels of *LEFTY2*, which in the mouse is expressed in mesoderm arising from the mid-distal region of the PS[3] (Extended Data Figure 4). The Advanced Mesoderm cluster was relatively more mature based on the decreased expression of *MESP1* and the highest expression of mesoderm markers *PDGFRA* and GATA6 (Extended Data Figure 2a and 4c). The Advanced Mesoderm also expressed *HAND1*, *BMP4*, *FOXF1* and *SNAI2,* all markers of relatively more mature mesoderm (Extended Data Figure 2a and 4c). The cells that we annotated as Nascent, Emergent and Advanced Mesoderm did not correspond directly to any of the established embryonic mesoderm sub-types such as paraxial or lateral plate mesoderm. Combinatorial marker signatures of such mesodermal sub-types spanned multiple clusters or were seen only within a subset of a cluster (Extended Data Figure 4a). For example, co-expression of *TBX6* and *MSGN1*, which marks presomitic mesoderm [4], was only detected in a subset of the Nascent Mesoderm cluster. In contrast, co-expression of *HAND1* and *GATA6*, which marks precardiac lateral-plate mesoderm, could be detected in multiple clusters including Advanced Mesoderm, Emergent Mesoderm, Extraembryonic Mesoderm as well as the Ectoderm (Amniotic/Embryonic). This suggests that at this stage, the embryonic mesodermal clusters identified do not represent specified mesodermal subtypes and instead correspond to transitional mesodermal states.

Extraembryonic Mesoderm was identified based both on its anatomical origin (69% yolk sac, 29% rostral, 2% caudal) as well as the expression of extraembryonic mesoderm markers such as *POSTN* and *ANXA1* (Extended Data Figure 2a, 2b and SI Table 16). In mice, *POSTN* is a marker of extraembryonic mesoderm overlying not only the amnion but also the yolk sac [5]. This cluster is likely to represent the yolk sac mesoderm given the spatial origin of the

majority of cells. However, it may also include some amniotic mesoderm cells, given the commonalities in marker expression between these two cell types.

Consistent with the possibility that some cells of the Advanced Mesoderm might contribute later in development to the Extraembryonic Mesoderm, there was a degree of overlap in marker expression between the two (e.g. *HAND1* and *SNAI2*), and proximity between them on the diffusion map. However, we noted that while other mesoderm populations formed a well-connected trajectory, there was a clear separation between Advanced and Extraembryonic Mesoderm populations (Extended Data Figure 3d and SI Table 16), suggesting the Extraembryonic mesoderm may not be as closely related as other mesoderm populations. This might be because, unlike in rodents, in humans the extra-embryonic mesoderm is already present pre-gastrulation at CS 5[6,7]. This early extraembryonic mesoderm is thought to arise from the hypoblast or parietal endoderm of the bilaminar disk stage embryo[8]. The high DC1 value of Extraembryonic Mesoderm is consistent with the notion of an early origin, with cells arising both prior to and during gastrulation[9].

The information we retained on the spatial origin of these cells also allowed us to track the progression of mesoderm maturity (Extended Data Figure 3d). Cells that had more time (since they emerged from the PS) to mature might also, in that time, be expected to have migrated further from the PS. Consistent with this, Nascent Mesoderm was almost entirely collected from the caudal portion of the embryo that encompassed tissue immediately adjacent to the PS (99% caudal, 0% rostral, 1% yolk sac; Figure 1d, Extended Data Figure 2b and SI Table 2). Similarly, Axial Mesoderm was only located in the caudal region, consistent with it having just emerged from the PS when the embryo was collected. In contrast, Emergent Mesoderm and Advanced Mesoderm were collected from both the rostral and caudal regions of the embryo (Emergent: 70% caudal, 30% rostral; Advanced: 58% caudal, 42% rostral), highlighting that they had migrated rostrally away from the PS. These two mesoderm clusters also showed evidence of sub-structure based on Rostral-Caudal differences in origin (See below)

Other mesoderm-derived clusters included primitive erythroblasts, characterised by the expression of globin genes including the embryonic globins *HBZ* and *HBE1* as well as the eryhroid-related transcription factor *GATA1*. The majority of erythroblasts were collected from the yolk sac (81% yolk sac, 19% caudal, 0% rostral). We annotated a Hemato-Endothelial Progenitor population based on the expression of both endothelial makers (*PECAM1* and *MEF2C*) as well as hematopoietic markers (*RUNX1* and *GATA1*). Hemato-Endothelial Progenitor cells were also located predominantly in the yolk sac, although some cells also came from caudal and rostral regions (72% yolk sac, 15% caudal, 12% rostral).

Both blood-related populations expressed erythroid marker genes[10]; however the HEP population had a mixed expression profile of endothelial, myeloid and erythroid markers, suggesting a higher order substructure (Extended Data Figure 10b and c). Unsupervised clustering of the HEP population revealed four different subpopulations with distinct transcriptional and isoform signatures (Figure 4d, Extended Data Figure 10c and d). One subpopulation represented Endothelium (Endoth) based on the high expression of *PECAM1*, *CDH5*, *KDR* and *TEK*. Another expressed both megakaryocyte (*GP1BB*, *ITGA2B* (CD41), *NFE2*) and erythroid (*GATA1*, *KLF1*, *GYPB*, *HBE1*) markers, which we annotated as Megakaryocyte-Erythroid Progenitors (MEP). A Myeloid Progenitor sub-population could be identified on the basis of high levels of monocyte/macrophage markers *CD36*, *CSF1R*, and *LYVE1*. The final subcluster had an unusual transcriptional profile given the early stage of the sample, expressing a range of myeloid and erythroid markers including *KIT*, *CSF1R*, *MYB*, *SPI1* (PU.1), *CD34*, *PTPRC* (CD45), CD52 and *NFE2*[11]. Based on these markers and the

expression of *MYB,* which in the mouse marks Erythro-Myeloid Progenitors (EMP), we annotated this cluster as EMP. Supporting this annotation is their co-expression of *CD34*, *CD45* and *CD44,* which have recently been used to define a yolk sac-derived myeloid-biased progenitor in CS11 human embryos[12]. These CS11 cells have been shown to have multi-lineage potential and are thought to correspond with murine EMPs. Our results therefore indicate that such progenitors might already start to emerge as early as CS7.

Endoderm could be identified by the expression of *SOX17*, *GATA6*, *FOXA2* and *TTR*. Endoderm was collected from all three anatomical regions (64% rostral, 19% yolk sac, 17% caudal; Extended Data Figure 9b and e). YS endoderm was identified based on spatial location (47% yolk-sac, 45% rostral, 8% caudal) and expression of the established marker genes *AFP* and *TTR* (Extended Data Figure 9e). In addition, we identified *GJB1* as a new marker of the Yolk Sac Endoderm and validated this by Hybridization Chain Reaction in the mouse (Extended Data Figure 9e and 9).

The Smart-Seq2 protocol also allowed us to differentiate between transcript isoforms (SI Table 3). An analysis of this was able to detect the cluster specific expression of particular isoforms of genes such as *MEST* and *GCNT2* (Extended Data Figure 2e). *GCNT2* is known to be differentially spliced during mouse embryonic development, validating this approach [13]. We observed that *GCNT2* isoform *201* was expressed in most clusters except Axial mesoderm and Erythroblasts, *202* was more epiblast and PS specific, whilst *225* was expressed most strongly in the Advanced Mesoderm and Endoderm populations. This approach also identified *MEST* as a gene which had cell type specific isoform variation. Whilst MEST 202 was expressed in all clusters, 201 expression was low in the Endoderm and Erythroblasts clusters and 203 was more strongly expressed in the Nascent Mesoderm and Extraembryonic Mesoderm. This analysis further validates our clustering and highlights the depth of this data set.

Supplementary Note 2 - Comparison with Gastruloids

A recent advance in developing *in vitro* models of human gastrulation are hESC derived gastruloids[14]. We used the human gastrula to benchmark the rostral-caudal patterning of gastruloids, by comparing the transcriptional patterns of sections collected along the rostral-caudal axis of human gastruloids[14] with the cell type and the spatial information from our human gastrula dataset (Extended Data Figure 3c). This verified an overall agreement between the rostral-caudal gene expression patterns in gastruloids compared to the gastrula. In future, once single-cell data from gastruloids become available, in order to further refine protocols for producing gastruloids, more detailed analyses can be performed, for example, by comparing cell types produce *in vitro* with their natural counterparts (see Methods and Extended Data Figure 3a-c).

Supplementary Note 3 - Rostral and Caudal differences in diversification of mesodermal subtypes

The earliest mesoderm to emerge from the PS gives rise to the *extra-embryonic mesoderm*, that form the early blood islands[8,15]. Later cells to stream through the PS remain embryonic and give rise to the: *lateral plate mesoderm* (LPM), *intermediate mesoderm*, *paraxial mesoderm*, and *axial mesoderm*. On the basis of just marker gene expression, the mesodermal clusters at this stage did not correspond specifically to any of these established mesodermal

sub-types, but rather had mixed expression of markers. This suggests that the clusters capture transient *states* of maturation of mesoderm as they transitioning towards more clearly specified sub-types.

To investigate the differentiation trajectory of human mesoderm, we took all mesoderm-related clusters together with the PS and performed a diffusion map and RNA-velocity analysis (Extended Data Figure 4b). Diffusion component 1 captured the time-course of mesoderm formation from the Primitive Streak through Nascent to Extra-embryonic Mesoderm. This trajectory was further highlighted by the RNA-velocity analysis, which showed vectors leading from PS through the Nascent Mesoderm towards the more mature mesoderm types. This computational analysis was supported by marker expression and the anatomical location from which cells were collected (Extended Data Figure 4b). The differentiation trajectory could also be observed by the transition along DC1 in the expression of transcription factors *TBXT*, *MESP1* and *HAND1* (Extended Data Figure 4c) that are respectively markers of the PS, early mesoderm and later relatively more mature mesoderm[16–18]. These results suggest that the pattern of emergence and maturation of mesoderm in the human is similar to that in chick and mouse.

Advanced Mesoderm predominantly represented LPM based on the expression of markers such as *HAND1*, *BMP4, GATA6* and *PDGFRA* (Extended Data Figure 4b). Within the Advanced Mesoderm cluster we could observe a separation in cells collected from rostral and caudal regions in both the UMAP and the diffusion map (Extended Data Figure 4b). To investigate this difference, we analyzed the differentially expressed genes between Advanced Mesoderm cells derived from these two regions (Extended Data Figure 4d). Rostral cells of the Advanced Mesoderm consistently showed elevated levels of expression for cardiomyocyte related contractile genes such as *TNNT2*, *MYL7*, *TNNI1* and *MYH10* (Extended Data Figure 4d and f). Consistent with this region of the LPM giving rise to the cardiac crescent.

However, several canonical markers of cardiac progenitors in the mouse were expressed at surprisingly low levels, if at all. The archetypal cardiac progenitor marker *NKX2-5* was expressed in very few cells throughout the mesoderm clusters, with no difference in caudal or rostral expression [19](Extended Data Figure 4d and e). *TBX5* was also expressed in very few cells, but in contrast to *NKX2-5*, expression was restricted specifically to rostral cells. *ISL1* was more strongly expressed but had a relatively broad distribution. In contrast *HOPX*, a homeodomain protein required for cardiac development in Zebrafish and Mouse, acting downstream of *NKX2-5* [20] was strongly and specifically expressed in caudal Advanced Mesoderm cells (Extended Data Figure 4d and e), despite the absence of *NKX2-5*. The cardiac-related marker *MAB21L2* [21,22], was strongly expressed in the Advanced Mesoderm, particularly in cells collected from the rostral region (Extended Data Figure 4d and e). Together these data suggest that in the human embryo, early cardiac contractile gene expression is initiated independently of *NKX2-5* and that other factors may play an important role in generating the cardiac progenitor state.

In the caudal cells of the Advanced Mesoderm cluster, there was an increase in PS markers such as *HOXA1* and *CDX1* consistent with their location (Extended Data Figure 4d and g). This region also strongly expressed *CDX2*, a marker of the early allantoic bud [23]. *CDH1*, required for placenta formation, is also expressed at elevated levels in the caudal cells. The combined expression of *CDX2* and *CDH1* suggests the caudal population of Advanced Mesoderm represents mesoderm emerging from the PS, to form allantoic mesoderm, consistent with the presence of an allantoic bud structure (Extended Data Figure 4d and g).

We could also detect the earliest signatures of other mesodermal subtypes. Chordin (*CHRD*) and *NOTO*, which are strongly expressed in the node and notochord could be detected in the axial mesoderm cluster, (Extended Data Figure 4h). The Emergent Mesoderm cluster had a broad and heterogenous marker gene profile, expressing genes reported in a number of different mesoderm types including, paraxial, intermediate, lateral plate mesoderm. The intermediate and axial midline marker, *LHX1*, was also expressed in this cluster along with other axial midline genes such as *OTX2* and *EOMES* [24] (Extended Data Figure 4j). The paraxial mesoderm marker *TBX6* and *SFRP2* was restricted to the Nascent Mesoderm and the caudal cells of the Emergent Mesoderm (Extended Data Figure 4i). The broad mesoderm signature of this cluster highlights its relatively immature state and suggests they emerged from the mid to anterior PS, supported by the specific expression of the anterior PS marker *LEFTY2* [3](Extended Data Figure 4j).

Supplementary Note 4 - Comparison of EMT pathway member expression during human and mouse gastrulation

During gastrulation, epithelial cells of the epiblast undergo an epithelial to mesenchymal transformation (EMT) by downregulating adherens junction molecules such as E-Cadherin (*CDH1*) so they can delaminate from the epiblast and migrate away as mesenchymal cells. Given this sample contains cells actively undergoing gastrulation, we sought to examine the transcriptional changes which occur during gastrulation in the human and compare them to the mouse, the leading model for studying mammalian gastrulation.

Pseudotime analysis showed that in the CS7 human gastrula, the PS marker *TBXT* (*Brachyury*) increased during the transition from Epiblast to Nascent Mesoderm, peaking in the Primitive Streak, while as expected, the mesoderm marker *MESP1* increased with the formation of Nascent Mesoderm (Figure 4c and Extended Data Figure 5b). A core event during EMT in mouse is a switch in the adherens junction molecules Cadherins, from E- to N-Cadherin (*CDH1* to *CDH2*)[25,26]. In the human, we could detect a similar trend in these genes, with CDH1 decreasing towards Nascent Mesoderm while *CDH2* increased (Extended Data 4b). In total, we found 3,350 genes that were differentially expressed along the developmental trajectory between Epiblast and Nascent Mesoderm (SI Table 5), of which 449 genes correlated with the expression profile of *TBXT* (FDR < 0.01; SI Table 6).

To test in an unbiased manner similarities and differences between human and mouse gastrulation, we used pseudotime analyses to compare the transition from epiblast to early mesoderm in the human cells with the equivalent populations in the Mouse Gastrula Single Cell Atlas[27] (Figure 2d and Extended Data Figure 6a). First, we compared the trends of genes that were differentially expressed in both human and mouse. For this analysis, we only considered genes with strong changes (i.e., with a log2-fold change > 1 along the trajectory, see Methods). By doing this, we identified 662 genes that were differentially expressed along the developmental trajectories from Epiblast to Nascent Mesoderm in both species (Extended Data Figure 6b and SI Table 7). Of these genes, the vast majority (531) shared the same trend across pseudotime, either increasing (117) or decreasing (414). For example, in both mouse and human, *CDH1* decreased during transition from epiblast to nascent mesoderm, *TBXT* was transiently expressed and *SNAI1* continuously increased towards nascent mesoderm (Figure 2d and Extended Data Figure 6c). This verifies that the mouse represents a broadly faithful model of human gastrulation, despite the evolutionary distance between the two and the

differences in gastrula morphology (human trilaminar disc compared to rodent specific egg cylinder).
Additionally, we also found some genes that were differentially expressed in only one of the two species. One example was in the expression of the zinc-finger transcription factor *SNAI2* (Slug), a regulator of EMT. In the human, *SNAI2* levels increased dramatically during Nascent Mesoderm formation, however *Snai2* was not detected during this transition in the mouse or cynomolgus macaque (Figure 2d). The absence of *Snai2* transcript in mouse during this transition was confirmed by additional independent mouse transcriptomic datasets from the same stages[28]. By contrast in the chick, as in the human, *SNAI2* is expressed within the PS and interfering in its expression results in the impaired emergence of mesoderm from the PS[29]. Together this suggests that in contrast to the mouse, in human, *SNAI2* may play a role in regulating EMT during gastrulation.

Another difference was in the expression of the transcription factor *MSGN1*. In the mouse *Msgn1* is expressed only weakly, if at all, in the streak and is expressed most strongly in the paraxial mesoderm[27,28]. Consistent with this, *Msgn1* null mouse embryos do not show a gastrulation phenotype [4,30], but show defects in somitogenesis[31]. Similar to mouse, *Msgn1* was not detected during gastrulation in the cynomolgus macaque. In contrast, the human gastrula showed robust and widespread expression of *MSGN1* in Nascent Mesoderm (Extended Data Figure 6c) raising the possibility that it may be functional more broadly in the human than the mouse.

In the mouse, the expression of various signaling molecules is crucial for EMT, germ layer specification and migration[32–34]. Hence, we analyzed specifically the expression trends of signaling molecules across the human, mouse and cynomolgus monkey. Our analyses again revealed broad similarities, but we also did observe some striking differences (Extended Data Figure 8). In the mouse, *TDGF1*, a NODAL co-receptor essential for normal mesodermal patterning, shows an increase in expression during Primitive Streak and Nascent Mesoderm formation. In contrast, in the human gastrula *TDGF1* expression showed the opposite trend, decreasing as Nascent Mesoderm formed (Figure 2d). FGF8 is the only known FGF directly required for gastrulation[35], playing a particularly important role in the migration of cells away from the PS[36]. In contrast, FGF8 was completely absent during the transition from Epiblast to Nascent Mesoderm in human (Figure 2d). We noted however that other FGF members were expressed in the human during this transition, raising the possibility that they may be serving the function that FGF8 does in the mouse. For example, we observed expression of *FGF4* (which is also expressed in the mouse), and *FGF2*, which is not required for gastrulation in the mouse[37,38] nor expressed, as confirmed in other datasets[28](Figure 2d). Consistent with the notion of an overlap in FGF function during gastrulation, treatment of *in vitro* cultured mouse epiblast with FGF2 results in altering the fate of these cells from ectoderm to mesoderm[39].

To experimentally validate these human specific transcriptional trends, we used an *in vitro* model of the transition from Epiblast to Nascent Mesoderm. For this, we differentiated pluripotent human ESC (PLU) to mesendoderm progenitors (ME) (Figure 2e and Extended Data Figure 7)[40–42]. ME colonies showed hallmarks of the EMT accompanying gastrulation, such as dispersed morphology and downregulated E-Cadherin (Figure 2e and Extended Data Figure 7a), in addition to the upregulation of key EMT and gastrulation markers (Figure 2e and Extended Data Figure 7b). Serving as a negative control, blockage of FGF/ERK signaling by a MEK inhibitor prevented all of these responses and directed ME+PD cells towards a non-neural ectoderm state (Extended Data Figure 7c). These results indicated that our hESC model recapitulates the transition from Epiblast to Nascent Mesoderm in the diffusion map (Figure 2c).

Using this *in vitro* system, we tested the gene expression changes identified in the human gastrula. Consistent with our earlier findings, *SNAI2* and *MSGN1* increased, while *FGF2* and *TDGF1* decreased significantly during the transition from PLU to ME states. *FGF8* was expressed at very low level throughout the differentiation (Figure 2e and Extended Data Figure 7b). Upon MEK inhibition, all these markers were maintained at levels comparable to that in PLU cells, with the exception of *TDGF1*, possibly because it is regulated by multiple pathways. Together, these results indicate that there is a broad conservation of molecular players in human and mouse gastrulation, while the roles of specific members in these gene modules may vary between humans and mice.

Supplementary References

1. Yang, L. *et al.* An early phase of embryonic Dlx5 expression defines the rostral boundary of the neural plate. *J. Neurosci.* (1998). doi:10.1523/JNEUROSCI.18-20-08322.1998

2. Streit, A. The preplacodal region: An ectodermal domain with multipotential progenitors that contribute to sense organs and cranial sensory ganglia. *International Journal of Developmental Biology* (2007). doi:10.1387/ijdb.072327as

3. Meno, C. *et al.* Mouse lefty2 and zebrafish antivin are feedback inhibitors of nodal signaling during vertebrate gastrulation. *Mol. Cell* (1999). doi:10.1016/S1097-2765(00)80331-7

4. Nowotschin, S., Ferrer-Vaquer, A., Concepcion, D., Papaioannou, V. E. & Hadjantonakis, A. K. Interaction of Wnt3a, Msgn1 and Tbx6 in neural versus paraxial mesoderm lineage commitment and paraxial mesoderm differentiation in the mouse embryo. *Dev. Biol.* (2012). doi:10.1016/j.ydbio.2012.04.012

5. Dobreva, M. P. *et al.* Periostin as a biomarker of the amniotic membrane. *Stem Cells Int.* (2012). doi:10.1155/2012/987185

6. O'Rahilly, R. & Müller, F. Developmental Stages in Human Embryos. *Contrib. Embryol., Carnegie Inst. Wash* **637**, (1987).

7. Enders, A. C. & King, B. F. Formation and differentiation of extraembryonic mesoderm in the rhesus monkey. *Am. J. Anat.* (1988). doi:10.1002/aja.1001810402

8. Bianchi, D. W., Wilkins-Haug, L. E., Enders, A. C. & Hay, E. D. Origin of extraembryonic mesoderm in experimental animals: Relevance to chorionic mosaicism in humans. *Am. J. Med. Genet.* (1993). doi:10.1002/ajmg.1320460517

9. Ross, C. & Boroviak, T. E. Origin and function of the yolk sac in primate embryogenesis. *Nat. Commun.* (2020). doi:10.1038/s41467-020-17575-w

10. McGrath, K. E., Frame, J. M. & Palis, J. Early hematopoiesis and macrophage development. *Seminars in Immunology* (2015). doi:10.1016/j.smim.2016.03.013

11. Palis, J. Hematopoietic stem cell-independent hematopoiesis: emergence of erythroid, megakaryocyte, and myeloid potential in the mammalian embryo. *FEBS Letters* (2016). doi:10.1002/1873-3468.12459

12. Bian, Z. *et al.* Deciphering human macrophage development at single-cell resolution. *Nature* (2020). doi:10.1038/s41586-020-2316-7

13. Revil, T., Gaffney, D., Dias, C., Majewski, J. & Jerome-Majewska, L. A. Alternative splicing is frequent during early embryonic development in mouse. *BMC Genomics* (2010). doi:10.1186/1471-2164-11-399

14. Moris, N. *et al.* An in vitro model of early anteroposterior organization during human development. *Nature* (2020). doi:10.1038/s41586-020-2383-9

15. Kinder, S. J. *et al.* The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo. *Development* (1999).

16. Wilkinson, D. G., Bhatt, S. & Herrmann, B. G. Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature* **343**, 657–658 (1990).

17. Saga, Y. *et al.* MesP1: A novel basic helix-loop-helix protein expressed in the nascent mesodermal cells during mouse gastrulation. *Development* (1996).

18. Cserjesi, P., Brown, D., Lyons, G. E. & Olson, E. N. Expression of the Novel Basic Helix-Loop-Helix Gene eHAND in Neural Crest Derivatives and Extraembryonic Membranes during Mouse Development. *Dev. Biol.* (1995). doi:10.1006/dbio.1995.1245

19. Harvey, R. P. NK-2 homeobox genes and heart development. *Developmental Biology* (1996). doi:10.1006/dbio.1996.0212

20. Chen, F. *et al.* Hop is an unusual homeobox gene that modulates cardiac development. *Cell* (2002). doi:10.1016/S0092-8674(02)00932-7

21. Saito, Y., Kojima, T. & Takahashi, N. Mab21l2 is essential for embryonic heart and liver development. *PLoS One* (2012). doi:10.1371/journal.pone.0032991

22. Tyser, R. C. V. *et al.* Characterization of a common progenitor pool of the epicardium and myocardium. *Science (80-.).* (2021). doi:10.1126/science.abb2986

23. Beck, F., Erler, T., Russell, A. & James, R. Expression of Cdx-2 in the mouse embryo and placenta: Possible role in patterning of the extra-embryonic membranes. *Dev. Dyn.* (1995). doi:10.1002/aja.1002040302

24. Costello, I. *et al.* Lhx1 functions together with Otx2, Foxa2, and Ldb1 to govern anterior mesendoderm, node, and midline development. *Genes Dev.* (2015). doi:10.1101/gad.268979.115

25. Smith, D. E., Franco Del Amo, F. & Gridley, T. Isolation of Sna, a mouse gene homologous to the Drosophila genes snail and escargot: Its expression pattern suggests multiple roles during postimplantation development. *Development* (1992).

26. Cano, A. *et al.* The transcription factor Snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat. Cell Biol.* (2000). doi:10.1038/35000025

27. Pijuan-Sala, B. *et al.* A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature* **566**, 490–495 (2019).

28. Peng, G. *et al.* Spatial Transcriptome for the Molecular Annotation of Lineage Fates and Cell Identity in Mid-gastrula Mouse Embryo. *Dev. Cell* (2016). doi:10.1016/j.devcel.2016.02.020

29. Nieto, M. A., Sargent, M. G., Wilkinson, D. G. & Cooke, J. Control of cell behavior during vertebrate development by Slug, a zinc finger gene. *Science (80-.).* (1994). doi:10.1126/science.7513443

30. Jeong Kyo Yoon & Wold, B. The bHLH regulator pMesogenin1 is required for maturation and segmentation of paraxial mesoderm. *Genes Dev.* (2000). doi:10.1101/gad.850000

31. Chalamalasetty, R. B. *et al.* Mesogenin 1 is a master regulator of paraxial presomitic mesoderm differentiation. *Dev.* (2014). doi:10.1242/dev.110908

32. Ciruna, B. & Rossant, J. FGF Signaling Regulates Mesoderm Cell Fate Specification and Morphogenetic Movement at the Primitive Streak. *Dev. Cell* (2001). doi:10.1016/S1534-5807(01)00017-X

33. Ding, J. *et al.* Cripto is required for correct orientation of the anterior-posterior axis in the mouse embryo. *Nature* (1998). doi:10.1038/27215

34. Jin, J. Z. & Ding, J. Cripto is required for mesoderm and endoderm cell allocation during mouse gastrulation. *Dev. Biol.* (2013). doi:10.1016/j.ydbio.2013.05.029

35. Ornitz, D. M. & Itoh, N. The fibroblast growth factor signaling pathway. *Wiley Interdiscip. Rev. Dev. Biol.* (2015). doi:10.1002/wdev.176

36. Sun, X., Meyers, E. N., Lewandoski, M. & Martin, G. R. Targeted disruption of Fgf8 causes failure of cell migration in the gastrulating mouse embryo. *Genes Dev.* (1999). doi:10.1101/gad.13.14.1834

37. Zhou, M. *et al.* Fibroblast growth factor 2 control of vascular tone. *Nat. Med.* (1998). doi:10.1038/nm0298-201

38. Ortega, S., Ittmann, M., Tsang, S. H., Ehrlich, M. & Basilico, C. Neuronal defects and delayed wound healing in mice lacking fibroblast growth factor 2. *Proc. Natl. Acad. Sci. U. S. A.* (1998). doi:10.1073/pnas.95.10.5672

39. Burdsal, C. A., Flannery, M. L. & Pedersen, R. A. FGF-2 alters the fate of mouse epiblast from ectoderm to mesoderm in vitro. *Dev. Biol.* (1998). doi:10.1006/dbio.1998.8898

40. Teo, A. K. K. *et al.* Pluripotency factors regulate definitive endoderm specification through eomesodermin. *Genes Dev.* (2011). doi:10.1101/gad.607311

41. Mendjan, S. *et al.* NANOG and CDX2 pattern distinct subtypes of human mesoderm during exit from pluripotency. *Cell Stem Cell* (2014). doi:10.1016/j.stem.2014.06.006

42. Yiangou, L. *et al.* Method to Synchronize Cell Cycle of Human Pluripotent Stem Cells without Affecting Their Fundamental Characteristics. *Stem Cell Reports* (2019). doi:10.1016/j.stemcr.2018.11.020