© 2021. Published by The Company of Biologists Ltd | Development (2021) 148, dev193789. doi:10.1242/dev.193789

RESEARCH ARTICLE

Spatiotemporal sequence of mesoderm and endoderm lineage segregation during mouse gastrulation

Simone Probst[1,2,*], Sagar[3], Jelena Tosic[1,4,5], Carsten Schwan[1], Dominic Grün[2,3] and Sebastian J. Arnold[1,2,*]

ABSTRACT

Anterior mesoderm (AM) and definitive endoderm (DE) progenitors represent the earliest embryonic cell types that are specified during germ layer formation at the primitive streak (PS) of the mouse embryo. Genetic experiments indicate that both lineages segregate from *Eomes*-expressing progenitors in response to different Nodal signaling levels. However, the precise spatiotemporal pattern of the emergence of these cell types and molecular details of lineage segregation remain unexplored. We combined genetic fate labeling and imaging approaches with single-cell RNA sequencing (scRNA-seq) to follow the transcriptional identities and define lineage trajectories of *Eomes*-dependent cell types. Accordingly, all cells moving through the PS during the first day of gastrulation express *Eomes*. AM and DE specification occurs before cells leave the PS from *Eomes*-positive progenitors in a distinct spatiotemporal pattern. ScRNA-seq analysis further suggested the immediate and complete separation of AM and DE lineages from *Eomes*-expressing cells as last common bipotential progenitor.

KEY WORDS: Gastrulation, Mouse embryo, *Eomes*, Definitive endoderm, Mesoderm, Lineage specification

INTRODUCTION

During mammalian gastrulation, the pluripotent cells of the epiblast become lineage specified and form the three primary germ layers definitive endoderm (DE), mesoderm and (neuro-) ectoderm. Mesoderm and DE are generated at the posterior side of the embryo under the influence of elevated levels of the instructive signals of Tgfß/Nodal, Wnt and Fgf. These signals induce an epithelial-to-mesenchymal transition (EMT) of epiblast cells at the primitive streak (PS), leading to their delamination and the formation of the mesoderm and DE cell layer. The nascent mesoderm layer rapidly extends towards the anterior embryonic pole by cell migration between the epiblast and the visceral endoderm (VE) (reviewed by Arnold and Robertson, 2009; Rivera-Pérez et al., 2003). DE progenitors migrate from the epiblast together with mesoderm cells, before they eventually egress into the VE layer to constitute the DE (reviewed by Rivera-Pérez and Hadjantonakis, 2014; Viotti et al., 2014a).

Current concepts suggest that different cell fates are specified according to the time and position of cell ingression through the PS, reflecting different instructive signaling environments (Rivera-Pérez and Hadjantonakis, 2014). However, the precise morphogenetic mechanisms guiding the emergence of various cell types along the PS still remain uncertain. This is at least in part because of the lack of detailed knowledge about the precise timing and location of individual cells becoming lineage specified, and the challenge to exactly determine the signaling pathway activities during embryonic fate commitment. For example, it remains unclear what is the embryonic correlate of a suggested common mesendoderm progenitor, as described during embryonic stem cell (ESC) differentiation *in vitro*.

Clonal cell labeling and transplantation experiments have proposed the gross patterns and dynamics of cell specification during gastrulation, which have been represented in fate maps of the epiblast and the early germ layers (Tam and Behringer, 1997; Lawson, 1999). Accordingly, first mesoderm cells delaminate from the newly formed PS at the proximal posterior pole of the embryo and give rise to extra-embryonic mesoderm cells (ExM). These migrate proximally and anteriorly to contribute to the mesodermal components of the amnion, chorion and the yolk sac (Parameswaran and Tam, 1995; Kinder et al., 1999). Embryonic anterior mesoderm (AM) giving rise to cardiac and cranial mesoderm follows shortly after ExM (Kinder et al., 1999). As the PS elongates towards the distal embryonic pole, other mesoderm subtypes and DE are generated. The distal domain of the PS (referred to as anterior PS, APS) generates DE and axial mesoderm progenitors, giving rise to the node, notochord and prechordal plate mesoderm (Kinder et al., 2001; Lawson et al., 1991). Additional mesoderm subtypes, such as lateral plate, paraxial and intermediate mesoderm, are generated between the APS and the proximal PS (Lawson et al., 1991; Kinder et al., 1999; Tam et al., 1997; Parameswaran and Tam, 1995).

Tgfß/Nodal and Wnt signals are indispensable for gastrulation onset (Brennan et al., 2001; Conlon et al., 1994; Liu et al., 1999), and genetic experiments revealed that graded levels of Nodal and Wnt signaling instruct distinct lineage identities during gastrulation (Vincent et al., 2003; Dunn et al., 2004; reviewed by Robertson, 2014; Arkell et al., 2013). The T-box transcription factor *Eomes* is a transcriptional target of NODAL/SMAD2/3 signaling (Brennan et al., 2001; Teo et al., 2011; Kartikasari et al., 2013) and is crucial for the specification of all DE and AM progenitors (Arnold et al., 2008; Costello et al., 2011; Probst and Arnold, 2017). Another T-box transcription factor, *Brachyury* (*T*), is essential for the formation of posterior mesoderm starting from embryonic day (E)7.5. Thus, the specification of all types of mesoderm and endoderm relies on either of the two T-box factors *Eomes* or *Brachyury* (Tosic et al., 2019). Experiments using differentiating human ESCs showed that EOMES directly binds and regulates the expression of DE genes together with SMAD2/3 (Teo et al., 2011). Similarly, in the mouse embryo, DE specification relies

[1]Institute of Experimental and Clinical Pharmacology and Toxicology, Faculty of Medicine, University of Freiburg, Albertstrasse 25, D-79104 Freiburg, Germany. [2]Signaling Research Centers BIOSS and CIBSS, University of Freiburg, Schänzlestrasse18, D-79104 Freiburg, Germany. [3]Max Planck Institute of Immunobiology and Epigenetics, Stübeweg 51, D-79108 Freiburg, Germany. [4]Spemann Graduate School of Biology and Medicine (SGBM), University of Freiburg, Albertstrasse 19a, D-79104 Freiburg, Germany. [5]Faculty of Biology, University of Freiburg, Schänzlestrasse 1, D-79104 Freiburg, Germany.

*Authors for correspondence (simone.probst@uniklinik-freiburg.de; sebastian.arnold@pharmakol.uni-freiburg.de)

S.P., 0000-0002-0220-5400; S.J.A., 0000-0002-2688-9210

Handling Editor: James Briscoe
Received 9 June 2020; Accepted 6 November 2020

Development (2021) 148, dev193789. doi:10.1242/dev.193789

on high NODAL/SMAD2/3 signaling levels (Dunn et al., 2004; Vincent et al., 2003). In contrast, in the presence of low or even absent NODAL/SMAD2/3 signals, EOMES activates transcription of key determinants for AM, including *Mesp1* (Saga et al., 1999; Lescroart et al., 2014; Kitajima et al., 2000; Costello et al., 2011; van den Ameele et al., 2012).

Recently, single-cell RNA sequencing (scRNA-seq) analyses allowed for a more detailed view on the cellular composition of embryos during gastrulation stages, including the identification of previously unknown rare and transient cell types (Scialdone et al., 2016; Mohammed et al., 2017; Wen et al., 2017; Lescroart et al., 2018; Pijuan-Sala et al., 2019). Despite the insights into the molecular mechanisms of cell lineage specification, questions about the emergence of the two *Eomes*-dependent cell lineages, AM and DE, remain unresolved. It is still unclear whether both cell populations are generated simultaneously from a common progenitor, and when and where lineage separation occurs. Answers to these questions are required for a comprehensive view on how suggested differences in the signaling environment impact on lineage specification of mesoderm and DE identities that are generated in close proximity within the epiblast of early gastrulation stage embryos that consist of only a few hundred cells (Snow, 1977).

In this study, we used embryo imaging and genetic fate mapping approaches by novel reporter alleles, in combination with molecular characterization by scRNA-seq, to delineate the spatiotemporal patterns of *Eomes*-dependent lineage specification. We show that AM and DE progenitors segregate within the PS into distinct cell lineages. AM progenitors leave the PS earlier and at more proximal regions than DE, demonstrating a clear spatial and temporal separation of lineage specification. The analysis of scRNA-seq experiments suggests that AM and DE progenitors are immediately fate segregated and *Eomes*-positive progenitors that co-express DE and AM markers were not found. This suggests that bipotential *Eomes*-expressing progenitors rapidly progress into either AM or DE lineage-specified cell types preceding cell ingression at the PS.

RESULTS

Eomes marks all cells leaving the PS during the first day of gastrulation

We used our previously described $Eomes^{mTnG}$ fluorescent reporter allele to observe the emergence of *Eomes*-dependent cell lineages during gastrulation. This reporter allele labels *Eomes*-expressing cells with membrane-bound Tomato (mT) and nuclear GFP (nG) (Probst et al., 2017; Fig. 1A-F). Embryos at stages shortly preceding gastrulation onset (E6.25) showed labeling within the cells of the posterior epiblast (Epi) before the formation of the PS (Fig. 1A), and reporter expression in the epiblast persisted until E7.5 (Fig. 1B-D). Importantly, all cells leaving the PS were *Eomes*-positive during these early gastrulation stages, as also seen by the complete reporter staining in the mesoderm layer (Fig. 1B-E,G,H; Movies 1, 2) that contains both mesoderm and DE progenitors (Viotti et al., 2014b). The maximum intensity projection (MIP) of z-stacks at E7.5 showed that the endoderm layer, which at this stage mainly consists of epiblast-derived DE cells, is composed of *Eomes* reporter positive cells (Fig. 1F), and only a few reporter negative cells could be detected. These most likely represent embryonic VE (EmVE) cells that express *Eomes* until E6.5 when it is downregulated (Fig. 1C,F; Movies 1, 2) (Nowotschin et al., 2013). As the fluorescent reporter proteins are more stable than the endogenous protein (Probst et al., 2017), we additionally performed immunofluorescence (IF) staining for EOMES at E7.25 and E7.5, showing the presence of EOMES protein in all cells of the posterior epiblast and in the mesoderm and endoderm layers (Fig. 1G,H). In conclusion, mesoderm and endoderm progenitors generated during the first day of gastrulation from E6.5 to E7.5 are exclusively descendants of *Eomes*-expressing cells (Fig. 1P). These constitute the progenitors of AM and DE, as also demonstrated previously by $Eomes^{Cre}$-mediated fate labeling (Costello et al., 2011).

To molecularly characterize *Eomes*-dependent cell types during early gastrulation, we performed scRNA-seq of cells collected from E6.75 and E7.5 embryos (Fig. 1I-O). A total of 289 handpicked cells from 14 E6.75 embryos, and 371 cells isolated by automated cell sorting from E7.5 pooled litters, were included in the scRNA-seq analysis. To identify transient progenitor populations, we clustered the cells using RaceID3 (Herman et al., 2018), an algorithm specifically developed for the identification of rare cell types within scRNA-seq data (Grün et al., 2015) (Fig. S1A,B). The tissue identities were assigned by the presence of differentially upregulated marker genes in each cluster compared with all other cells (Fig. 1I,J,L,M; Fig. S1A,B; Tables S1, S2). The heatmap representations indicate specifically expressed marker genes in different assigned cell types (Fig. 1J,M). At E7.5, RaceID identified rare cells, such as one single E7.5 primordial germ cell (PGC) (Fig. 1L,M; Table S2). The comparison of t-distributed stochastic neighbour embedding (t-SNE) maps at E6.75 and E7.5 (Fig. 1I,L) showed that, at E6.75, epiblast and PS/native mesoderm (NM) cells clustered closely to each other and only extra-embryonic tissues [extra-embryonic ectoderm (ExE) and VE] are clearly separated on the t-SNE maps (Fig. 1I). The epiblast subclusters at E6.75 (Fig. S1A) do not represent distinct trajectories towards mesoderm or endoderm progenitors. In contrast, at E7.5, separable clusters can be detected within the embryonic cell clusters, demonstrating the increase in transcriptome diversity of embryonic cell types between E6.75 and E7.5 (Fig. 1L).

At E6.75, *Eomes* is expressed in 209 of 289 analyzed cells (72% of all cells), showing highest expression in the PS/NM cluster (Fig. 1K; Fig. S1C) and weaker expression in the epiblast cluster, which is in agreement with the immunofluorescence staining (Fig. 1B; Movie 1). In addition, *Eomes* expression was found in the extra-embryonic tissues ExE and VE (Fig. 1K; Fig. S1C). At E7.5, *Eomes* transcripts were still present in a subset of epiblast cells, the PS, the NM, the node and the mesoderm (AM and ExM) and DE clusters (Fig. 1N; Fig. S1D). However, only 35% of cells showed RNA expression, whereas EOMES protein was still broadly detected (Fig. 1H,N). Thus, at E6.75 *Eomes* mRNA was expressed in relatively more cells and at higher levels than at E7.5 (Fig. 1O). In summary, scRNA-seq analysis showed that the embryonic *Eomes*-expressing cells at E6.75 cluster closely to each other, indicating that they are molecularly similar (Fig. 1I), and we could not identify separate clusters of lineage progenitors for AM and DE.

A novel $Mesp1^{mVenus}$ allele identifies *Eomes*-dependent AM progenitors

Mesp1 represents one of the earliest markers of mesoderm within the *Eomes*-positive cell population, and is a direct transcriptional target gene of EOMES (Costello et al., 2011). Lineage tracing with a *Mesp1-Cre* allele shows that it faithfully labels the ExM and the AM (Saga et al., 1999; Lescroart et al., 2014, 2018; Chan et al., 2013). To distinguish mesoderm from DE progenitors during the first day of germ layer formation, we generated a fluorescent $Mesp1^{mVenus}$ reporter allele by inserting the sequence of membrane-bound Venus (mV) into the start codon of the *Mesp1* locus followed by the *Mesp1* coding sequence (Fig. 2A-C). Homozygous $Mesp1^{mVenus}$

Development (2021) 148, dev193789. doi:10.1242/dev.193789

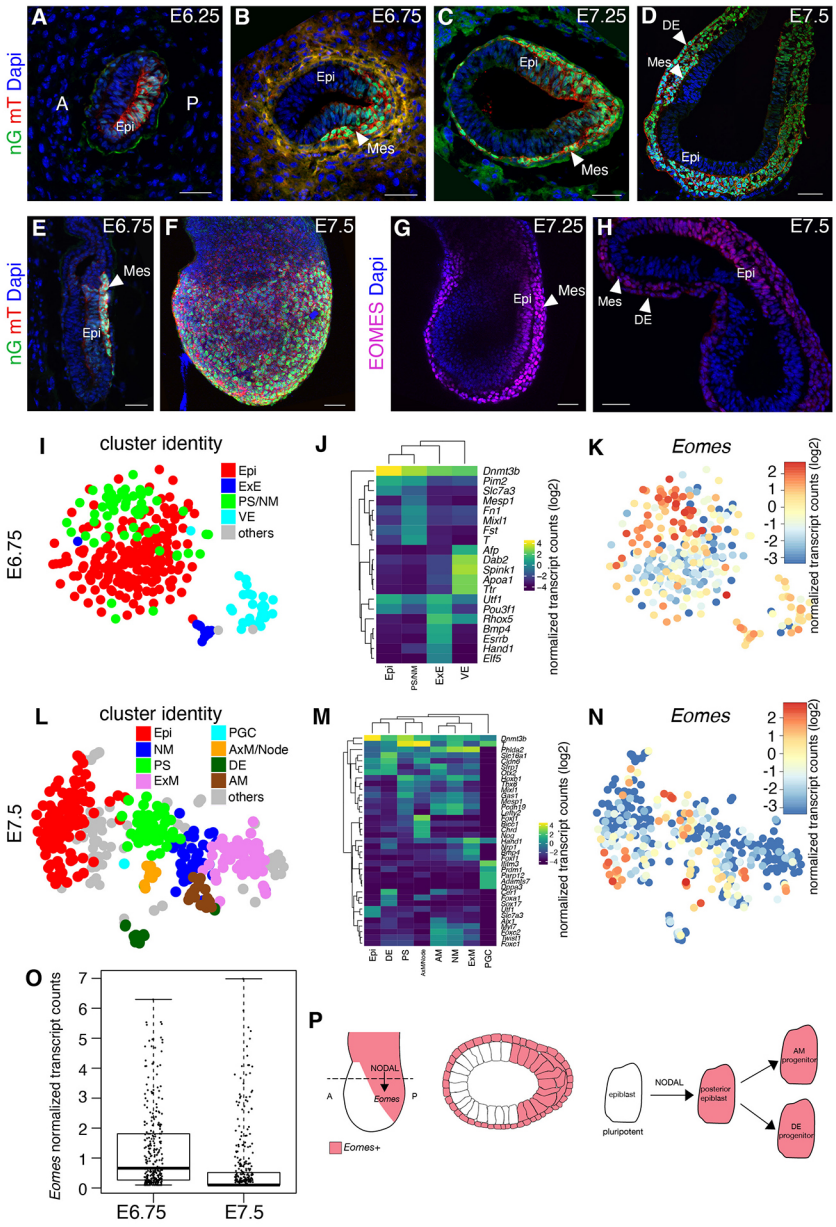


Fig. 1. All cells of the posterior epiblast and AM and DE progenitors between E6.25 and E7.5 express *Eomes*. (A-H) Immunofluorescence staining of *Eomes*^{mTnG} and wild-type embryos ($n \geq 3$ embryos). In this and all following figures, embryos are oriented with anterior (A) to the left and posterior (P) to the right. (A-D) Transverse sections of *Eomes*^{mTnG} embryos showing nuclear GFP (nG) and membrane-bound Tomato (mT) in *Eomes*-expressing cells. (A) Before gastrulation onset at E6.25, *Eomes*^{mTnG}-positive cells mark the posterior half of the proximal epiblast (Epi). (B) At E6.75 and (C) E7.25, the posterior epiblast and all cells in the mesoderm layer (Mes) are positive for *Eomes*^{mTnG}. (D) At E7.5, the *Eomes*^{mTnG}-positive mesoderm wings have migrated to the anterior, and *Eomes*^{mTnG}-positive DE cells integrated into the outer endoderm layer. (E) Sagittal section of an E6.75 embryo showing *Eomes*^{mTnG} expression in the nascent mesoderm layer. (F) MIP of an E7.5 *Eomes*^{mTnG} embryo. (G) Sagittal section of an E7.25 wild-type embryo showing EOMES expression in the epiblast, mesoderm and endoderm. (H) Transverse section of an E7.5 wild-type embryo. Endogenous EOMES protein remains present in the posterior epiblast and in the mesoderm and DE layers; protein levels are reduced in the more anterior mesoderm and DE. Scale bars: 50 µm. (I-O) scRNA-seq of wild-type embryos at E6.75 and E7.5. (I,L) t-SNE plots with assigned identities to different clusters at (I) E6.75 and (L) E7.5. (J,M) Heat maps of selected marker genes for the cluster identities indicated in I and L at E6.75 (J) and at E7.5 (M). (K,N) t-SNE plots showing the expression of *Eomes* in single cells at E6.75 (K) and E7.5 (N). The scale bars represent log2 normalized transcript counts. (O) Boxplot showing the expression levels of *Eomes* by normalized transcript counts in single cells at both time points, indicating a higher proportion of *Eomes*-expressing cells at E6.75. The box shows the interquartile range and the whiskers show the minimum and maximum values of the data. (P) Schematic illustrating the generation of *Eomes*-dependent cell lineages in the posterior embryo at E6.75.

(*Mesp1*^{mV}) mice are viable and fertile, demonstrating sufficient *Mesp1* expression from the reporter allele.

We analyzed the emergence of *Eomes*-dependent mesoderm progenitors in *Mesp1*^{mV} embryos and found *Mesp1*^{mV} reporter-expressing cells as early as E6.5 in the proximal epiblast during early PS formation (Fig. 2D,E). At this stage, the PS had not yet extended towards the distal part of the embryo and no *Mesp1*^{mV}-positive cells were present in distal portions of the epiblast (Fig. 2F). Notably, most cells leaving the early proximal PS showed *Mesp1* reporter expression, identifying them as mesoderm progenitors (Fig. 2E, inset). Once *Mesp1*^{mV}-positive cells leave the PS they rapidly migrate proximally and anteriorly to their destinations of ExM and AM (Fig. 2G). Importantly, *Mesp1*^{mV}-reporter positive cells were detected in the epithelial portion of the PS (Fig. 2H inset, arrowheads), indicating that mesoderm fate specification takes place before cells delaminate from the epiblast. At E7.25, the mesodermal wings had migrated far anteriorly (Fig. 2J). *Mesp1*^{mV}-positive cells constituted the major population within the EOMES-positive mesodermal layer (Fig. 2I). *Mesp1*^{mV}-negative cells were found intermingled between *Mesp1*^{mV}-reporter expressing cells mostly towards distal regions (Fig. 2I,L, inset, arrowheads). At E7.25, nascent *Mesp1*^{mV} cells were still emerging from the proximal PS (Fig. 2K), whereas, more distally, no *Mesp1*^{mV}-expressing cells were present in the PS (Fig. 2L). *Mesp1* therefore marks the earliest population of mesoderm progenitors that are continuously produced between E6.5 and E7.5 from *Eomes*-expressing cells. *Mesp1*-positive progenitors are present throughout the mesoderm layer but they are preferentially generated in the proximal domains of the PS.

DE and AM progenitors become fate specified in different regions of the epiblast

Next, we investigated the spatial distribution of the *Eomes*-dependent cell lineages (Costello et al., 2011; Arnold et al., 2008). To simultaneously detect DE and AM progenitors, we used FOXA2 IF staining of embryos carrying the *Mesp*^{mV} reporter allele. Previous reports and our data show that *Foxa2* is expressed in the VE and during gastrulation from E6.5 onwards in the epiblast, the APS/node, and its derivatives DE and axial mesoderm (AxM)

Development (2021) 148, dev193789. doi:10.1242/dev.193789

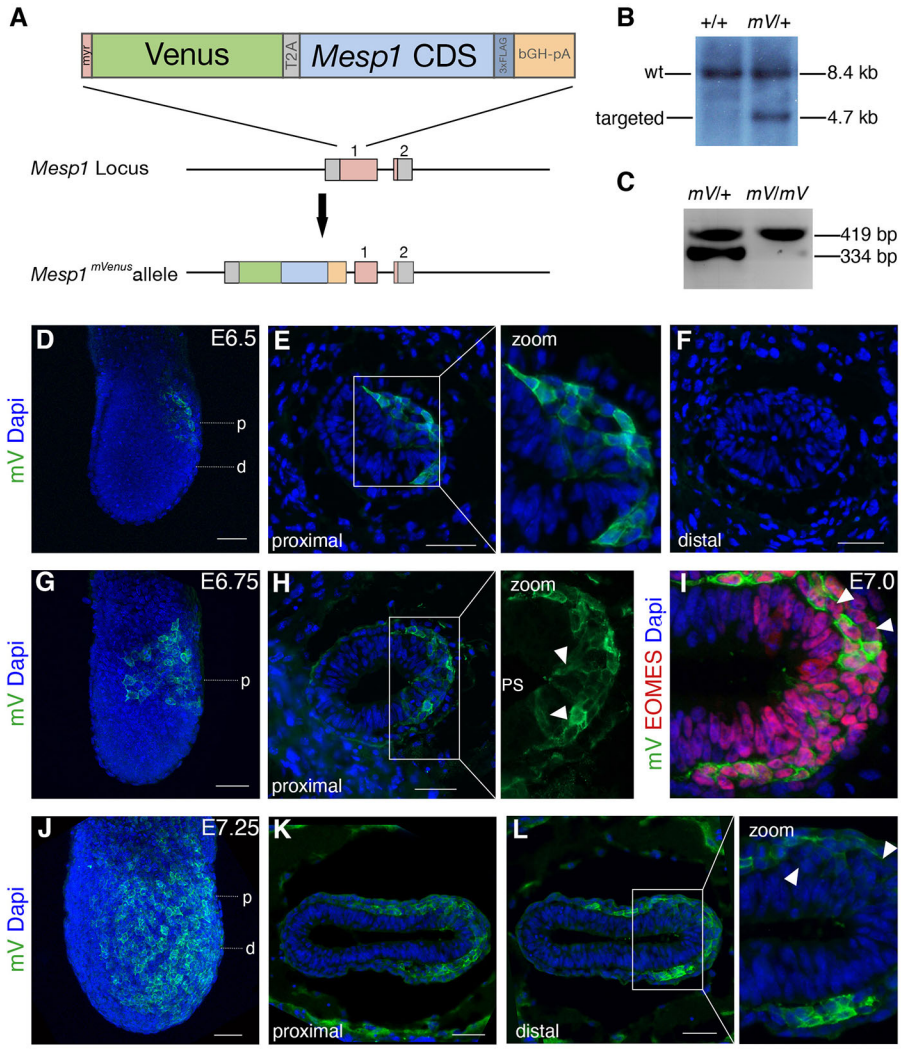


Fig. 2. Generation of a novel *Mesp1*^{mVenus} allele to identify the *Eomes*-dependent mesoderm progenitors. (A) Schematic of the *Mesp1*^{mVenus} allele. The sequence for a membrane-targeted (myr) Venus protein (mV) and the *Mesp1* CDS were inserted into the ATG start codon of the *Mesp1* gene to generate the *Mesp1*^{mVenus} (*Mesp1*^{mV}) allele. (B) Southern blot analysis of targeted ESC clones showing a wild-type (wt) (+/+) and a correctly targeted (*mV*/+) clone. The wild-type band was detected at 8.4 kb and the targeted band was detected at 4.7 kb. (C) Genotyping PCR of a heterozygous (*mV*/+) and a homozygous (*mV*/*mV*) mouse showing the wild-type band at 334 bp and the *mV* band at 419 bp. (D-L) Immunofluorescence stainings with anti-GFP antibody to enhance mV protein in *Mesp1*^{mV} embryos ($n \geq 3$ embryos). (D-F) *Mesp1*^{mV}-positive cells appear during the initiation of gastrulation at E6.5. (E,F) Transverse sections at E6.5 show that early gastrulating cells in the proximal embryo are positive for *Mesp1*^{mV} expression (E), whereas in more distal regions the PS has not yet formed (F). (G) By E6.75, *Mesp1*^{mV}-positive cells rapidly migrate proximally towards the extra-embryonic domain and anteriorly. (H) *Mesp1*^{mV}reporter-positive cells were detected in the epithelial PS (arrowheads in the zoomed inset). (I-L) By E7.0 and E7.25, the *Mesp1*^{mV}-positive cells constitute a large population within the mesoderm layer. (I) Co-staining with anti-EOMES antibody shows that *Mesp1*^{mV}-expressing cells represent a subpopulation of EOMES-positive^{+} cells (arrowheads indicate a few *Mesp1*^{mV}-negative cells). (J-L) At E7.25, the mesoderm layer contains many *Mesp1*^{mV}-positive cells, and the proximal PS is also positive for *Mesp1*^{mV} (K). More distally, the PS contains no *Mesp1*^{mV} cells (arrowheads in zoomed inset indicate *Mesp1*$^{mV-}$ cells) (L). D, G and J, show MIPs. The approximate levels of the transverse sections are indicated in the MIPs. d, distal; p, proximal. Scale bars: 50 µm.

(Fig. 3; Ang et al., 1993; Sasaki and Hogan, 1993; Monaghan et al., 1993; Viotti et al., 2014b). Previous lineage tracing by Cre-induced recombination and imaging by fluorescent reporters showed that *Foxa2* expression faithfully labels DE progenitors (Park et al., 2008; Frank et al., 2007; Imuta et al., 2013).

The simultaneous analysis of *Mesp1*^{mV} and FOXA2 showed that AM and DE progenitors were generated at distinct levels along the PS (Fig. 3A-J). At E6.5, FOXA2-positive DE progenitors still remained within the epithelial epiblast and were located more distally in relation to proximally located *Mesp1*^{mV}-expressing AM cells, of which some have already delaminated from the PS, as observed in sagittal (Fig. 3A) or in consecutive transverse sections (Fig. 3C-F). Additionally, FOXA2 broadly marks VE cells (Fig. 3A-J). The distribution of proximally located AM and distal DE progenitors was also found at E6.75 (Fig. 3B) and E7.0 (Fig. 3G-J), at which point the most proximal sections showed *Mesp1*^{mV} expression in the PS and in cells of the mesoderm layer (Fig. 3G). More distal regions of the PS contained a mix of *Mesp1*^{mV} and FOXA2 single-positive cells (Fig. 3H,I). At the distal tip of the PS, only FOXA2-positive cells were found within the streak or the delaminated cells (Fig. 3J). Only rarely *Mesp1*^{mV}-positive cells that also showed a FOXA2 signal were found (Fig. 3H,I, arrowheads). Careful analysis of stained embryos showed that *Mesp1*^{mV} and FOXA2 double-positive cells were located mostly in the mesoderm layer and only few were found in the PS. These double-positive cells were found at an intermediate level between the proximal domain of *Mesp1*^{mV}-reporter positive cells and the distal domain of FOXA2-positive cells (Fig. 3H,I; Movies 3, 4). These cells most likely represent recently described FOXA2-positive progenitors that contribute to the cardiac ventricles and the outflow tract (Bardot et al., 2017; Ivanovitch et al., 2020 preprint). These double-positive cells could also represent a transient bipotential progenitor population for DE and ME. At both E6.5 and E7.0, we found *Mesp1*^{mV} single-positive cells in the proximal PS and mesoderm layer (Fig. 3C,G). Similarly, in distal PS only FOXA2 single-positive cells were found (Fig. 3E,J), suggesting that double-positive cells in the mid PS region represent a specific cell population that does not contribute to the majority of AM or DE cells. In summary, AM and DE progenitors are generated in mostly non-overlapping domains, and cells are already lineage separated when they are still located within the epithelial epiblast at the level of the PS.

Next, we employed scRNA-seq to analyze the segregation of *Foxa2* expressing DE and *Mesp1* expressing AM progenitors by their RNA expression profiles (Fig. 3K-N). At E6.75 and E7.5, *Mesp1* and *Foxa2* expression was found in distinct cells in the t-SNE maps (Fig. 3K-M; Fig. S2A-C). As our scRNA-seq analysis contained a limited amount of cells, we also employed a published scRNA-seq dataset that contains higher cell numbers (Pijuan-Sala et al., 2019). Here, we included and combined time points from E6.5

Development (2021) 148, dev193789. doi:10.1242/dev.193789

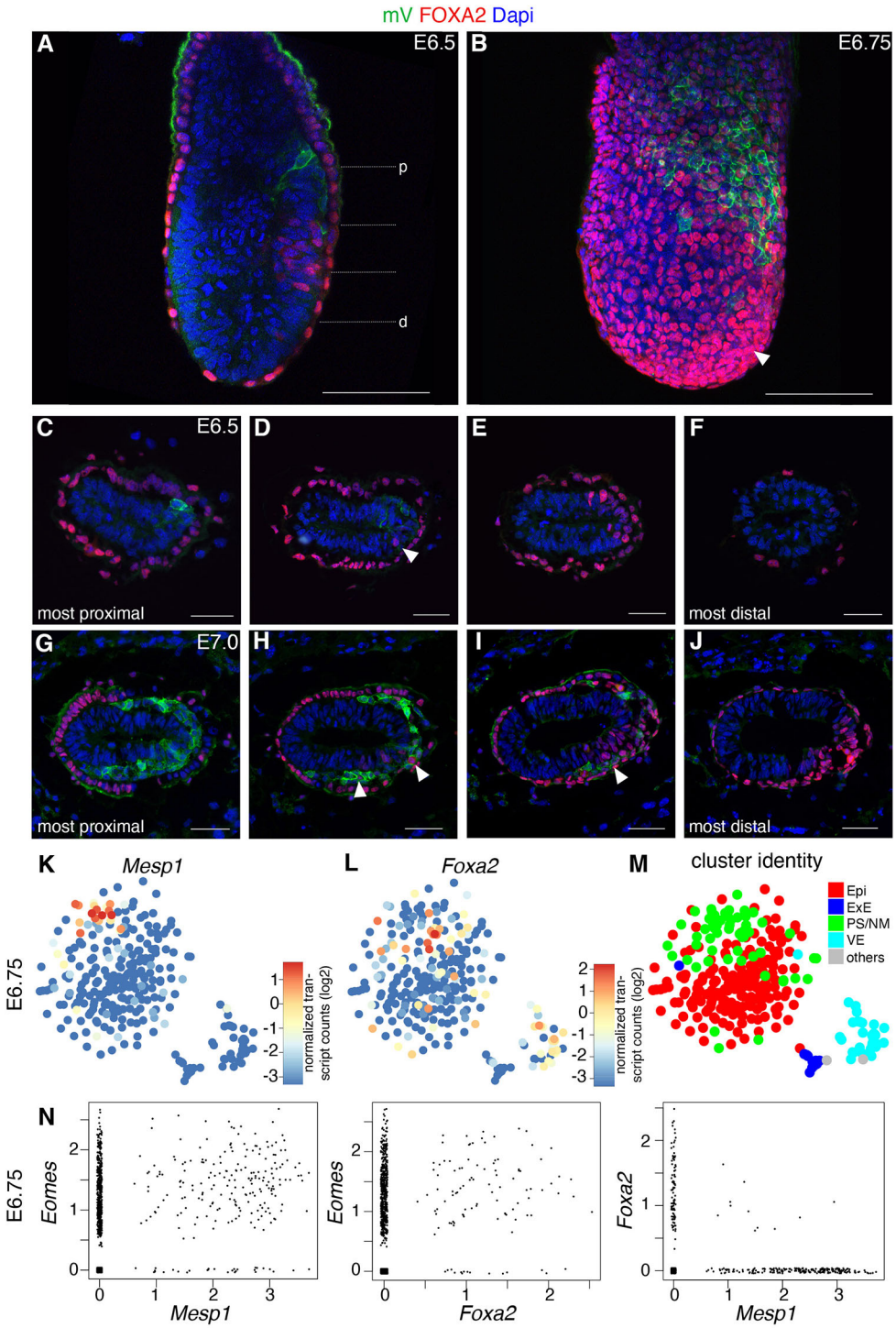


Fig. 3. Spatial separation of *Eomes*-dependent *Mesp1*^{mV} labeled AM and FOXA2^{+} DE progenitors in the posterior epiblast and PS.
(A-J) Immunofluorescence staining in *Mesp1*^{mV} embryos using anti-GFP (green) and anti-FOXA2 (red) antibodies ($n \geq 3$ embryos). FOXA2 is present in the cells of the VE. (A) Sagittal section of an E6.5 *Mesp1*^{mV} embryo showing proximal *Mesp1*^{mV} reporter-expressing cells and distal FOXA2^{+} cells in the posterior epiblast (n=1 embryo).
(B) MIP of an E6.75 *Mesp1*^{mV} embryo (n=2 embryos). *Mesp1*^{mV}-positive cells are present in the proximal region of the embryo. FOXA2^{+} cells in the VE are covering the whole embryo. In the posterior distal embryo FOXA2-positive cells are especially dense, probably corresponding to newly generated FOXA2-positive DE progenitors (arrowhead). (C-J) Transverse sections at different levels of E6.5 (C-F) and E7.0 (G-J) *Mesp1*^{mV} embryos. C and G show the most proximal sections, and F and J show the most distal sections. The approximate levels of the transverse sections are indicated in A, along the proximal (p) to distal (d) axis. Proximal sections contain *Mesp1*^{mV} reporter-expressing cells and the more distal sections contain FOXA2-positive cells. A single FOXA2-positive cell in the *Mesp1*^{mV} positive domain of the epiblast in D is indicated with an arrowhead. At E7.0, there is an intermediate zone of mixed *Mesp1*^{mV}-positive and FOXA2-positive cells (H and I). Few *Mesp1*^{mV}/FOXA2 double-positive cells are present (H and I, arrowheads). Scale bars: 100 µm (A,B); 50 µm (C-J). (K,L) t-SNE representation of E6.75 scRNA-seq data showing the expression of *Mesp1* (K) and *Foxa2* (L). The scale represents log2 normalized transcript counts. (M) t-SNE plot with assigned identities to different clusters at E6.75. (N) Scatter plots of single cells at E6.75 indicating *Eomes*/*Mesp1*, *Eomes*/*Foxa2* and *Foxa2*/*Mesp1* expression (data from Pijuan-Sala et al., 2019). Correlation scores are *Eomes*/*Mesp1*: 0.3991671, *Eomes*/*Foxa2*: 0.2262012, and *Foxa2*/*Mesp1*: −0.0782863. *x*- and *y*-axes indicate normalized transcript counts.

to E7.5 (E6.5, E6.75, E7.0, E7.25 and E7.5), and performed cell clustering using the Seurat package (Stuart et al., 2019). Similar clusters were identified between both datasets (Fig. S3A), and also in the larger dataset *Mesp1* and *Foxa2* expressing cells were largely non-overlapping on the Uniform Manifold Approximation and Projection (UMAP) representations (Fig. S3C,D). Interestingly, at E6.75, the *Mesp1*-positive cells cluster closely together on the t-SNE map, whereas *Foxa2*-expressing cells were found more scattered within the clusters of Epi/PS/NM (Fig. 3K-M; Fig. S2E). This suggests that *Foxa2*-positive cells have less homogenous expression profiles that are more similar to unspecified epiblast cells at these early time points of analysis, whereas at E7.5, *Foxa2*-positive cells form discrete clusters of node and DE cells (Fig. S2B,C,F).

Plotting cells for their expression of *Eomes*, *Mesp1* and *Foxa2* at E6.75 shows co-expression of *Mesp1* or *Foxa2* with *Eomes* in most cells (Fig. 3N, first and second plot). *Mesp1* and *Foxa2* expression was mostly exclusive, with the exception of 11 observed *Mesp1*/*Foxa2* double-positive cells out of 324 *Mesp1*$^{-}$ and/or *Foxa2*-positive cells (Fig. 3N, third plot; Fig. S3E). At E7.5, *Eomes* was rapidly downregulated and, consequently, increasing numbers of *Mesp1* or *Foxa2* single-positive *Eomes*-negative cells were found (Fig. S2D, Tables S3, S4). *Mesp1*-positive and *Foxa2*-positive cell

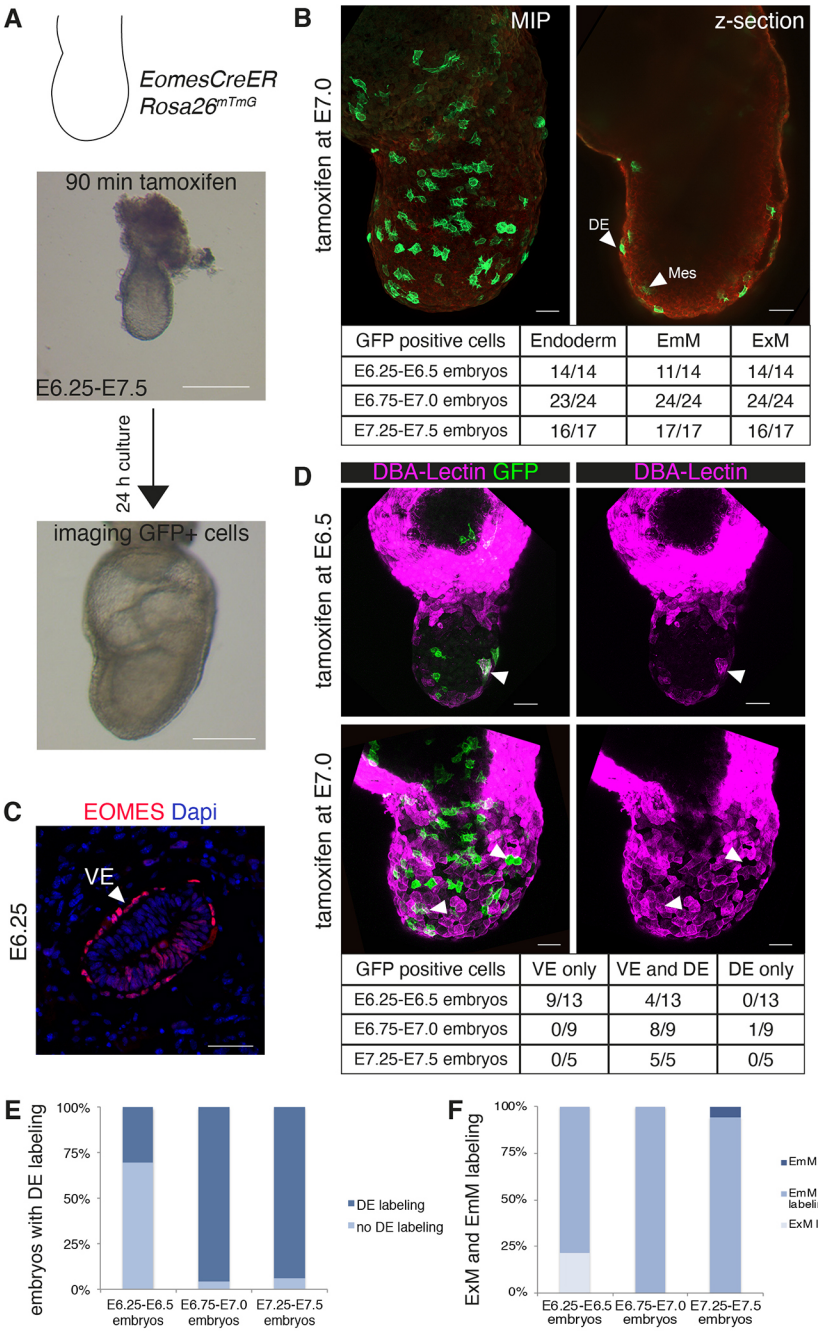
Development (2021) 148, dev193789. doi:10.1242/dev.193789

populations remained mostly exclusive (Fig. S2D). Quantification of *Mesp1*/*Foxa2* double-positive cells within both datasets showed that more than 95% of *Mesp1*$^{-}$ or *Foxa2*-expressing cells were single-positive, and only between 1.7% to 5% of cells were double-positive at different time points (Fig. S3E). This analysis was confirmed by quantification of *Mesp1*^{mV} reporter-expressing and FOXA2-immunostained cells in whole sectioned embryos, which showed similar results (Fig. S3E).

In summary, the simultaneous analysis of early emerging AM and DE progenitors at E6.5 reveals the spatial separation of their sites of origin. *Mesp1*^{mV} mesoderm progenitors are generated from the proximal PS and *Foxa2*-expressing DE progenitors from the distal PS. ScRNA-seq analysis shows mostly exclusive lineage marker expression, suggesting that AM and DE progenitors are separated. Only a few *Mesp1*^{mV}/*Foxa2* double-positive cells were detected, most likely representing a separate progenitor population.

Eomes-dependent AM progenitors are specified at earlier time points than DE progenitors

Our analyses and published literature show that the generation of mesoderm and DE progenitors is spatially separated along the forming PS (Fig. 3; Lawson and Pedersen, 1987; Tam and Beddington, 1987; Lawson et al., 1991; Tam and Behringer, 1997). The fact that the PS elongates over time in a proximal to distal fashion suggests that mesoderm and DE progenitor specification is also temporally separated. To test the temporal sequence of lineage specification downstream of *Eomes*, we performed time-dependent genetic lineage tracing using a tamoxifen-inducible *Eomes*^{CreER} mouse line expressing CreER from the *Eomes* locus (Pimeisl et al., 2013) in



GFP positive cells	Endoderm	EmM	ExM
E6.25-E6.5 embryos	14/14	11/14	14/14
E6.75-E7.0 embryos	23/24	24/24	24/24
E7.25-E7.5 embryos	16/17	17/17	16/17

GFP positive cells	VE only	VE and DE	DE only
E6.25-E6.5 embryos	9/13	4/13	0/13
E6.75-E7.0 embryos	0/9	8/9	1/9
E7.25-E7.5 embryos	0/5	5/5	0/5

Fig. 4. AM and DE are specified from *Eomes*-positive cells following a sequential temporal order. (A) Schematic of time-dependent lineage tracing in E6.25 to E7.5 embryos carrying the *Eomes*^{CreER} and the *Rosa26*^{mTmG} reporter alleles (Muzumdar et al., 2007). Embryos were dissected, staged and treated for 90 min with tamoxifen to induce CreER activity, followed by culture for 24 h without tamoxifen and three-dimensional imaging. A total of 55 embryos were analyzed. (B) MIP and an optical section of an exemplary embryo treated with tamoxifen at E7.0. The contribution of *Eomes*-expressing cells from labeling at different time points to different cell types is summarized in the table below the images. (C) Transverse section of an E6.25 embryo shows EOMES in the VE and in the epiblast. (D) MIPs of embryos stained for DBA-lectin to identify VE cells. The upper panels show an embryo treated with tamoxifen at E6.5. Lectin staining (arrowhead) of GFP-positive cells indicates VE cells. The other GPF-positive cells are in the mesoderm layer. The lower panels show an embryo that was treated with tamoxifen at E7.0, and GFP-expressing cells in the endoderm layer are of both VE (lectin-positive) and DE (lectin-negative) origin, as shown by arrowheads. The table summarizes the amounts of embryos with the contribution of GFP-positive cells to VE or DE, or both. A total of 27 embryos were re-analyzed for DBA-lectin staining. (E) Bar graph representing the percentage of embryos with GFP labeling in the DE. (F) Bar graph representing the percentage of embryos with ExM or EmM labeling, or with labeling in both ExM and EmM. Embryonic time points indicated in tables in B and D, and in the graphs in E and F, are time points of tamoxifen treatment. Scale bars: 200 µm (A); 50 µm (B-D).

Development (2021) 148, dev193789. doi:10.1242/dev.193789

combination with a Cre-inducible fluorescent reporter (Muzumdar et al., 2007; Fig. 4A). This *Rosa26*^{mTmG} reporter strain ubiquitously expresses membrane-bound Tomato that switches to membrane-bound GFP following Cre recombination. Short-term administration of tamoxifen (90 min) to dissected and morphologically staged embryos in culture was used to label *Eomes*-expressing cells at different developmental time points (Fig. 4A). Embryos were sorted into three groups according to the stage at the time of dissection (E6.25-E6.5, E6.75-E7.0 and E7.25-7.5), and were cultured for an additional 24 h. Whole embryos were imaged as *z*-stacks to evaluate whether the presence of GFP-labeled cells within the mesoderm and endoderm layers depends on the time point of Cre induction (Fig. 4A, B). Of note, in addition to the labeling of epiblast-derived cell types, this approach also marks *Eomes*-expressing VE cells (Fig. 4C).

In a first analysis, 53 of 55 embryos showed labeling of both endoderm and mesoderm, including embryonic mesoderm (EmM) and ExM (Fig. 4B). Interestingly, three E6.25-labeled embryos expressed GFP only within the ExM, supporting the notion that ExM is the first cell population generated in the PS (Fig. 4F) (Parameswaran and Tam, 1995; Kinder et al., 1999). As we were interested in the DE population within the labeled cells of the endoderm layer originating from *Eomes*-expressing cells in the posterior epiblast/PS, we needed to discriminate VE from DE cells. Therefore, we additionally stained embryos with the lectin dolichos biflorus agglutinin (DBA-lectin) that specifically labels VE cells but not DE (Fig. 4D) (Kimber, 1986). This revealed that the GFP-positive cells in the endoderm layer of 9 out of 13 (69%) E6.25-E6.5-labeled embryos were exclusively of VE origin, indicating that no DE was formed yet in most of the E6.25-E6.5 embryos (Fig. 4D). All E6.25-E6.5-labeled embryos showed GFP-positive cells in mesoderm cells (EmM 10/13) (Fig. 4B,F). The GFP positive cells in the endoderm layer of the remaining four E6.25-E6.5-labeled embryos were of mixed DE and VE origin. Thus, we confirmed the existence of a short time window before E6.5 during which *Eomes*-expressing cells in the posterior epiblast give rise to mesoderm (Fig. 4D,E,F). Starting from E6.5, progenitors of mesoderm and DE are both present (Fig. 3) and therefore embryos that were tamoxifen treated at E6.5-E6.75 or later showed GFP labeling both in mesoderm and DE cells (Fig. 4D,E,F). These experiments thus confirm that mesoderm and DE specification is also temporally separated so that mesoderm progenitor specification slightly precedes DE formation.

DE progenitors are transcriptionally closer to the epiblast and show a less complete EMT signature

As the RaceID algorithm did not identify distinct progenitor populations for AM and DE within the *Eomes*-positive cells of the epiblast (Fig. 1I,L), we wanted to investigate scRNA-seq expression profiles during this lineage segregation in more detail. Thus, we analyzed the transcriptomes of *Eomes*-expressing cells (expression cutoff, 0.3 normalized transcript counts) from the published dataset (Pijuan-Sala et al., 2019) at time points from E6.5 to E7.5. According to our analysis (Figs 1, 3), the *Eomes*-positive population should include the unspecified progenitors, as well as early AM and DE progenitors. VE and ExE cells were excluded from the analysis. VarID (Grün, 2020) identified cell clusters representing the posterior epiblast and two branches consisting of the proximal PS, NM, AM and ExM, and of the distal PS, AxM, node and DE (Fig. 5A; Fig. S4A).

Eomes-positive cells were categorized into three groups of *Eomes/Mesp1* double-positive cells (blue), *Eomes/Foxa2* double positive cells (red), and *Eomes* single-positive cells (grey) (expression cut-off 0.3 normalized transcript counts for all three genes) (Fig. 5B). Differential gene expression analysis between the *Eomes*/*Mesp1* or *Eomes*/*Foxa2* double positive cells and *Eomes* single-positive cells showed that *Foxa2*-positive cells expressed higher endoderm and axial mesoderm marker genes (e.g. *Sox17*, *Cer1* and *Gsc*), and *Mesp1*-positive cells showed increased expression of mesodermal/mesenchymal/EMT genes (e.g. *Fn1*, *Lefty2*, *Myl7* and *Snai1*) (Fig. S4B; Table S5). Both *Mesp1*-expressing and *Foxa2*-expressing cells showed a downregulation of anterior epiblast markers (e.g. *Pou3f1*, *Utf1* and *Slc7a3*) (Fig. S4B; Table S5), indicating the differentiation of these two cell populations towards their respective fates. Overall, more genes were differentially regulated in *Mesp1*-expressing cells than in *Foxa2*-expressing cells (126 versus 43 genes were more than twofold changed, respectively), suggesting that DE progenitors are transcriptionally more similar to cells of the epithelial epiblast or that AM progenitors are further differentiated. *Foxa2*-positive DE progenitors delaminate from the PS and migrate anteriorly together with the mesoderm cells before they intercalate into the VE layer to form the DE layer (Viotti et al., 2014b). To analyze whether EMT regulation of DE progenitors differs from EMT in mesoderm progenitors, we compared the expression of EMT- and migration-associated genes during gastrulation. Several EMT (including *Zeb2*, *Twist1* and *Snai1*) and migration genes (e.g. *Itga5* and *Rasgrp3*) were expressed at lower levels in *Eomes* single-positive and *Eomes/Foxa2* double positive cells compared with *Eomes/Mesp1* double-positive cells (Fig. 5D,E). This indicates that even though DE progenitors delaminate from the PS, the regulation of EMT is different from mesoderm cells at the transcriptional level.

To further analyze the rare *Mesp1/Foxa2* double-positive cells found at the intermediate level between *Mesp1*^{mV} and FOXA2^{+} PS regions, we plotted *Mesp1/Foxa2* double-positive cells onto the *Eomes*-positive population. These cells are distributed across different clusters, but are enriched in the distal PS cluster close to the branch point of DE and AM (Fig. 5C). Differential gene expression analysis shows that *Eomes/Mesp1/Foxa2* triple-positive cells express both mesoderm and DE gene markers at low levels, which would fit with a role as very transient bipotential progenitors (Fig. S4C; Table S6). However, clustering of *Eomes/Mesp1/Foxa2* triple-positive cells and visualization of gene signatures of AM, DE and epiblast clusters in these cells suggests that they are not a particularly short-lived population. They rather represent a continuum of states from the epiblast to lineage-specified progenitors (Fig. S4D-F). Finally, lineage tracing with a *Mesp1*^{Cre} mouse line (Saga et al., 1999) indicated that *Mesp1*-expressing cells generally do not give rise to DE progenitors (Fig. S4G) (Yoshida et al., 2008; MacDonald et al., 2008; Park et al., 2006; Zhang et al., 2005; Saga et al., 2000). Thus, the rare *Foxa2/Mesp1* double-positive cells are unlikely bipotential progenitors for the majority of AM and DE progenitors.

Eomes-expressing epiblast cells directly differentiate to either AM or DE lineages

Next, we investigated whether *Eomes* single-positive cells are differentiation biased towards either AM or DE progenitors by using FateID, which uses a random forests-based approach to assign a fate bias (on a scale of 0 to 1) to all the cells included in the analysis. It requires the input of target cells, i.e. the differentiated cells, which are used to train an iterative random forests classifier for the inference of the fate probabilities of the remaining cells included in the analysis based on the characterized transcriptome (Herman et al., 2018). To avoid artefacts originating from different developmental stages of cells, we analyzed each time point separately, with the exception of E6.5 and E6.75 cells that were combined to increase cell numbers (Fig. 6A,E; Fig. S5B). We

Development (2021) 148, dev193789. doi:10.1242/dev.193789

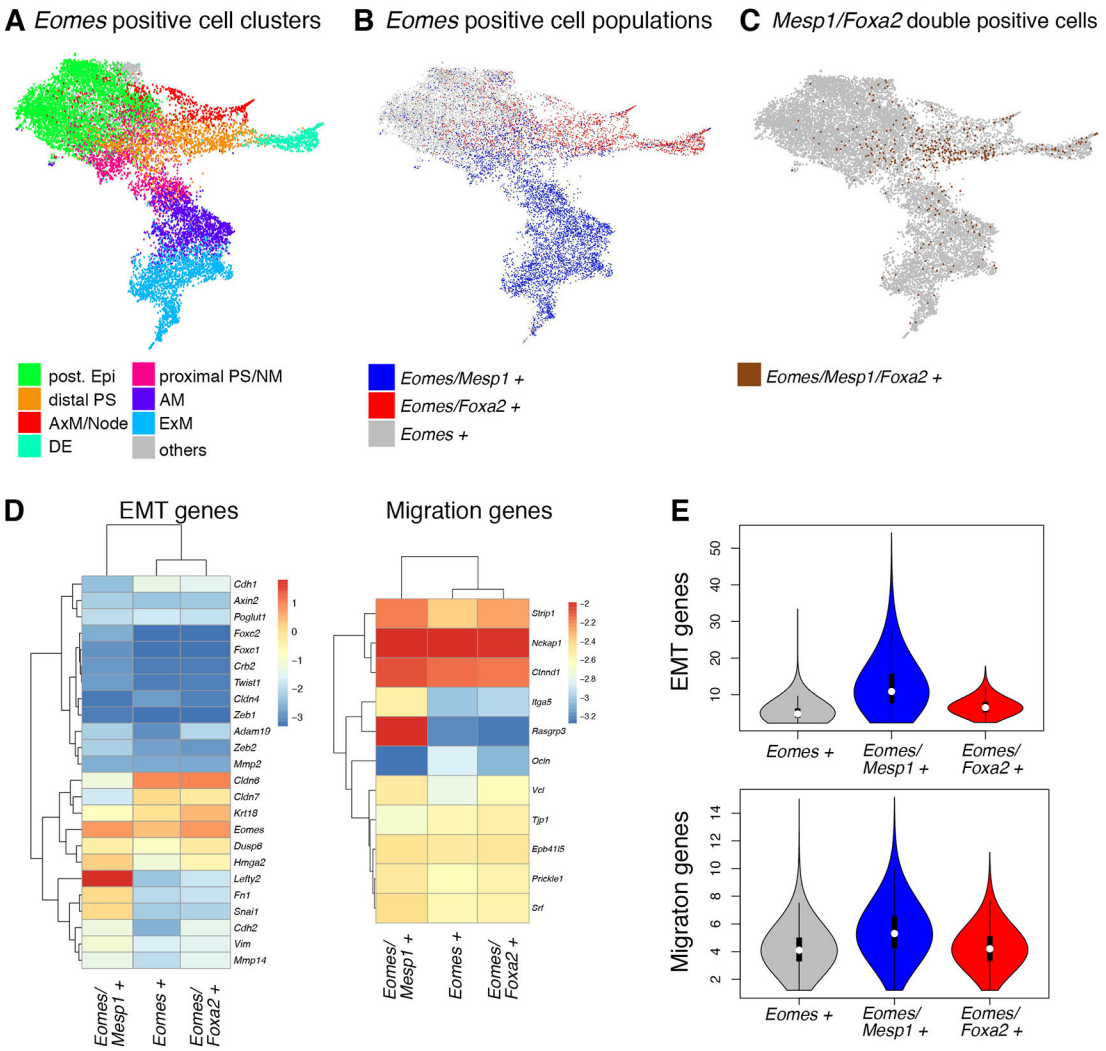


Fig. 5. The DE progenitor signature is more epithelial than that of AM progenitors. (A) UMAP representation of all *Eomes*-positive cells from E6.5 to E7.5. Expression cutoff was set at 0.3 counts (data from Pijuan-Sala et al., 2019). Assigned cluster identities are indicated (posterior epiblast, post. Epi). (B) UMAP representation of *Eomes*/*Mesp1* (blue) and *Eomes*/*Foxa2* double-positive cells (red), and *Eomes* single-positive cells (grey). The cutoff for *Mesp1* and *Foxa2* expression was set to 0.3 counts. (C) Mesp1/Foxa2-double-positive cells within the Eomes-positive population. (D,E) Heat map representation (D) and violin plots (E) of selected genes involved in EMT and migration during gastrulation, comparing *Eomes* single-positive cells with *Eomes*/*Mesp1* and *Eomes*/*Foxa2* double-positive cells. Heat map scale bar represents log2 normalized transcript counts. *y*-axis in violin plots indicates aggregated gene expression of all genes included in the analysis. Migration genes are expressed at very low levels. The following genes were not included in the violin plots: EMT – *Cdh1*, *Cldn4*, *Cldn6*, *Cldn7* and *Krt18*; migration – *Ocln*.

defined early *Eomes*/*Mesp1* and *Eomes*/*Foxa2* double-positive cells as target cells and excluded more differentiated clusters (target cells, shown in red cells in Fig. 6C,D,G,H; Fig. S5D,E). On the respective UMAP representations of E6.5/E6.75 cells, early *Mesp1*-expressing cells were grouped and *Foxa2*-expressing cells were more scattered (as described in Fig. 3K,L; Fig. 6A), whereas at later time points *Eomes/Mesp1* and *Eomes/Foxa2* double-positive cells formed two distinct branches (Fig. 6E; Fig. S5B). *Utf1* expression is shown to indicate the undifferentiated epiblast population (Fig. 6B,F; Fig. S5C) (Tosic et al., 2019; Galonska et al., 2014). We then calculated the fate bias probabilities (between 0 and 1) of *Eomes* single-positive cells for each target group, i.e. *Eomes/Mesp1* or *Eomes/Foxa2* double-positive cells in red (fate bias probability of 1). The target cells of the respective other fate appear blue (fate bias probability of 0). This analysis revealed that at E6.5/E6.75 *Eomes* single-positive cells have a similar fate bias probability towards both lineages of AM and DE (yellow cells, Fig. 6C,D), and thus are not fate biased towards either lineage. Accordingly, only very few differentially expressed genes could be found when we compared expression values in mesoderm- and endoderm-biased cells (fate bias probability cutoff was set to 0.6 for each of the respective lineages) (12 genes were more than twofold changed, Fig. S5A; Table S7). At E7.25, the *Eomes* single-positive cells of the undifferentiated epiblast (Fig. 6F, *Utf1*$^{+}$) were biased towards *Eomes/Foxa2* double-positive target cells (orange cells, Fig. 6H). The cells closer to the branching point mostly did not display a clear fate bias showing an intermediate probability for both lineages (yellow cells, Fig. 6G,H). This indicates that at E7.25 most mesoderm downstream of *Eomes* has already been generated and the majority of the remaining *Eomes* single-positive epiblast cells will give rise to DE/axial mesoderm. Analysis of differentially expressed genes between endoderm- and mesoderm-biased cells at E7.25 revealed that epiblast markers were more strongly expressed in endoderm-fated cells, whereas mesoderm-fated cells already

Development (2021) 148, dev193789. doi:10.1242/dev.193789

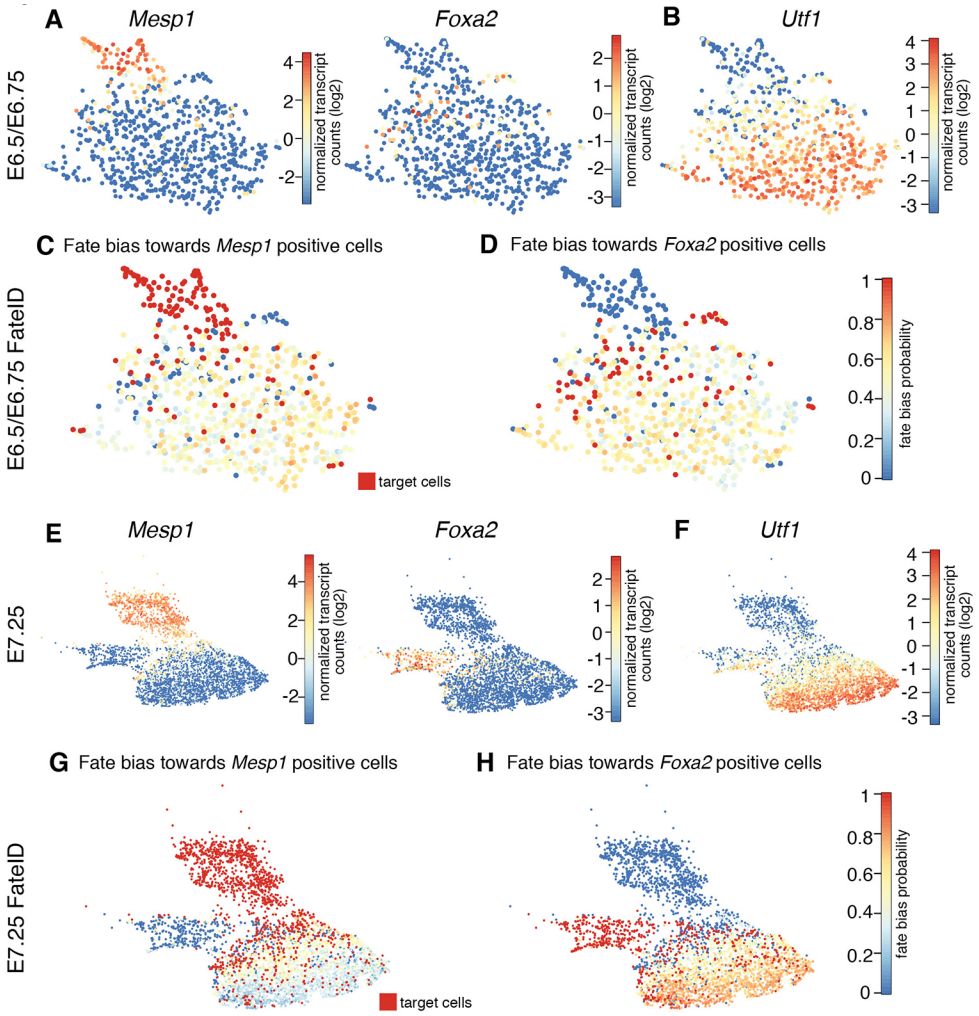


Fig. 6. *Eomes*-positive posterior epiblast cells directly differentiate to either AM or DE. (A,B,E,F) *Eomes*-positive cells from E6.5 and E6.75 (A,B) and E7.25 (E,F) were clustered, and *Mesp1* and *Foxa2* (A,E) and *Utf1* (B,F) expression was plotted onto the UMAP representation. (C,D,G,H) FateID analysis of embryonic *Eomes*-expressing cells from time points E6.5 and E6.75 (C,D), and E7.25 (G,H). The fate bias probability is indicated in single *Eomes*-positive cells towards *Mesp1*-positive cells (C,G) and *Foxa2*-positive cells (D,H) (red, target cells). Color scale represents fate bias probabilities on the scale from 0 to 1. Scale bar for gene expression on UMAP maps represents log2 normalized transcript counts. Data from Pijuan-Sala et al. (2019).

show a pronounced mesodermal expression profile (52 genes were more than twofold changed, Fig. S5A; Table S7). At E7.0, FateID analysis showed the progressively increasing fate bias towards the *Foxa2*-expressing population in *Eomes*-expressing epiblast cells (Fig. S5D,E).

In conclusion, until E7.0, the *Eomes* single-positive posterior epiblast cells are not fate biased towards either lineage before the onset of *Mesp1* or *Foxa2* expression, and *Mesp1* and *Foxa2* are mostly expressed in distinct cell populations. We therefore propose that cells differentiate directly from a posterior *Eomes*-positive epiblast state to either mesoderm or endoderm lineages, without passing through an intermediate mesendoderm progenitor state.

DISCUSSION

To date, the understanding of lineage specification on the level of individual cells within the gastrulation stage embryo remains limited. It is still unclear how cells in close proximity acquire different fates according to local signaling environments and how these specification events are regulated in a temporal manner. In this study, we have analyzed the emergence of AM and DE populations that are both dependent on the transcription factor *Eomes* (Arnold et al., 2008; Costello et al., 2011; Teo et al., 2011). *Eomes*-expressing cells give rise to the mesoderm derivatives of the anterior embryo and the entire DE (Costello et al., 2011; Arnold et al., 2008). The *Eomes*-expressing population in the early PS was thought to be one of several populations leaving the PS between E6.5 and E7.5 (Robertson, 2014). However, our data indicate that early posterior epiblast cells uniformly express *Eomes* and all cells passing through the PS during the first day of gastrulation (E6.5 to E7.5) are positive for EOMES. Thus, between E6.5 and E7.5, only cells that will contribute to the mesoderm of the anterior embryo and the DE progenitors leave the PS, and posterior mesodermal tissues are generated after E7.5. Accordingly, spatial gene regulatory network analysis of gastrulation stage embryos indicates that separate anterior and posterior mesoderm populations exist at E7.5 (Peng et al., 2019). The FateID analysis further indicates that AM downstream of *Eomes* is mainly generated until E7.25. During the following stages, mesoderm formation is most likely regulated by other factors, such as the related T-box factor *Brachyury* and Wnt signaling (Koch et al., 2017; Wymeersch et al., 2016). The existence of distinct anterior and posterior mesoderm populations downstream of different T-box factors has been suggested previously, such as in the zebrafish (Kimelman and Griffin, 2000). However, the molecular details of this transition in the regulation of gastrulation are currently incompletely understood.

The first lineage decision following *Eomes* expression in the epiblast segregates AM and DE. Here, we show that the AM and DE marker genes, *Mesp1* and *Foxa2*, are already expressed in epithelial epiblast cells at the PS. Therefore, we can place the event of lineage specification within the PS before cells migrate to form the mesoderm layer. Earlier cell tracing experiments have shown that cells are restricted in their potency after their passage through the PS (Tam et al., 1997). Our simultaneous marker analysis shows that

Development (2021) 148, dev193789. doi:10.1242/dev.193789

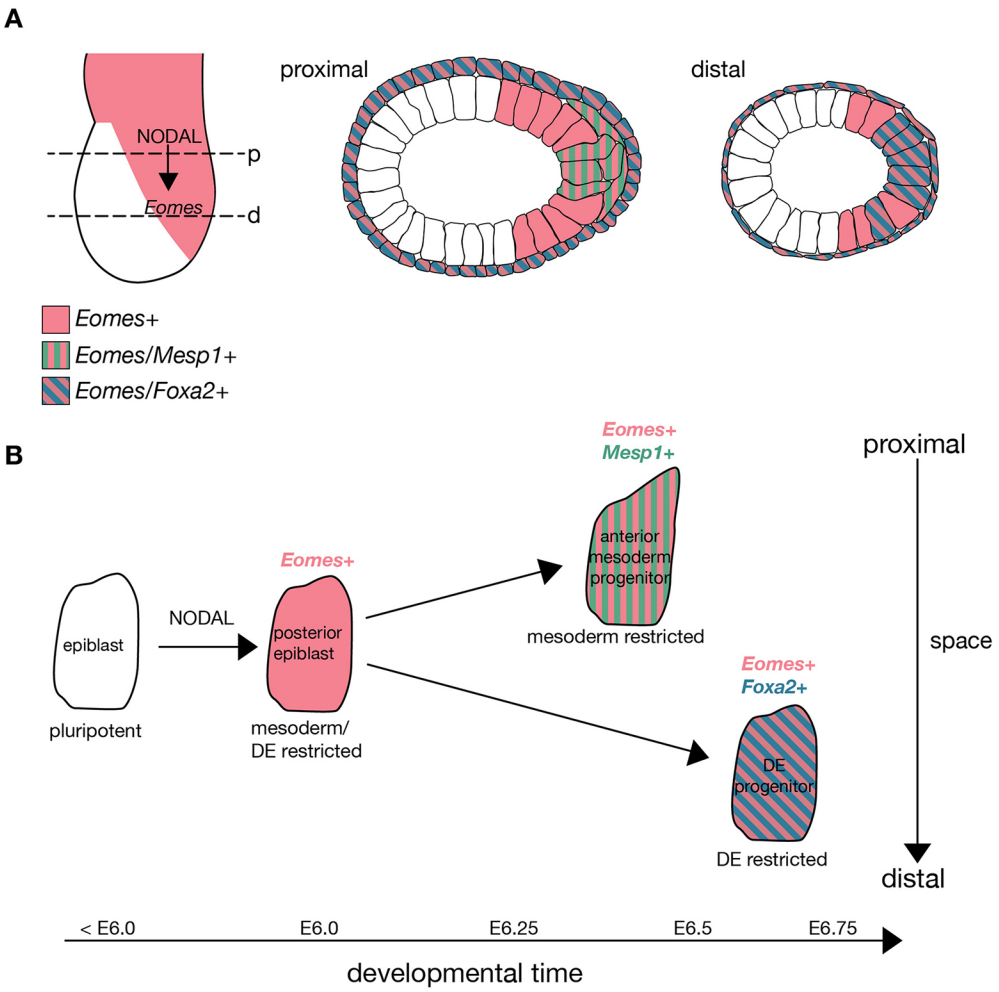


Fig. 7. Model of spatial and temporal separation of DE and AM lineage specification downstream of *Eomes*. (A) Induction of *Eomes* by NODAL/SMAD2/3 signals leads to the specification of AM and DE lineages. AM marked by *Mesp1* expression is generated in the proximal PS and *Foxa2*-positive cells give rise to DE in the distal PS. At E6.5, few AM progenitor cells have already delaminated from the PS, whereas DE progenitor cells are still entirely located in the epiblast. The VE is positive for *Foxa2*. p, proximal; d, distal. (B) AM and DE progenitor cells are specified from an unbiased *Eomes*-expressing progenitor cell at different localizations along the proximo-distal axis and at different time points. *Mesp1* and *Foxa2* indicate the first fully specified AM and DE cells, respectively.

mesoderm and DE are produced at distinct places along the PS (Fig. 7A,B). The proximal domain of the PS generates only mesoderm from the initiation of gastrulation. With a slight temporal delay, the most distal tip of the PS produces only *Foxa2*-positive DE and axial mesoderm progenitors. However, this study did not address the generation and distribution of the axial mesoderm that is also derived from *Foxa2*-expressing progenitors. At intermediate levels, the PS generates both mesoderm and endoderm progenitors, and here we also find rare *Mesp1/Foxa2* double-positive cells. The separation of BRACHYURY-positive and FOXA2-positive domains in the PS of E6.5 embryos (Burtscher and Lickert, 2009) further suggests the existence of distinct areas of progenitor specification within the PS.

Experiments with ESCs indicate that cells go through a transient state in which they co-express mesoderm and DE lineage markers (often referred to as mesendoderm cells) (Tada et al., 2005; Kubo et al., 2004). This might represent a transient bipotential state that all AM and DE progenitors pass through. However, in the embryo, this co-expression of lineage markers is found in only a few cells within a restricted domain of the PS and mesoderm layer. Their scRNA expression profiles and localized presence of these cells in the PS suggest that they do not represent a bipotential progenitor for all the AM and DE. Cells within the PS are lineage specified but not lineage determined (Tam et al., 1997). In culture conditions, cell state changes might take place that will not occur in the embryo. Similarly, changes in signals delivered by the culture conditions (medium composition and/or other signals) might contribute to the appearance of cell signatures during ESC differentiation that are only rarely found in an undisturbed embryo or in limited cells. For example, DE differentiation was described to generally pass through a *Brachyury*-positive state during ESC differentiation (Kubo et al., 2004; D'Amour et al., 2005), whereas lineage tracing in the embryo suggests that only the hindgut is generated from *Brachyury*-expressing cells (Imuta et al., 2013; Perantoni et al., 2005; MacDonald et al., 2008). Therefore, ESC differentiation through a *Brachyury*-positive state might produce mostly hindgut DE and does not represent the specification events for all DE progenitors in the embryo.

Live embryo imaging analysis has shown that DE and mesoderm progenitors leave the PS and migrate together within the mesodermal wings before DE cells insert into the outer VE layer (Viotti et al., 2014b). Our data show that DE cells are already specified as they leave the PS. Single-cell transcriptome analysis of EMT and migration genes indicates that they are differentially regulated between *Eomes*-dependent AM and DE progenitors. It will be interesting to further investigate how DE cells behave within the mesoderm population and which mechanisms are used to separate them during the anterior-ward migration.

Analysis of the fate bias of *Eomes*-positive cells, which are not yet expressing *Mesp1* or *Foxa2* markers, indicates that an unbiased posterior epiblast state directly progresses to either AM or DE, arguing for fast-acting control mechanisms that independently promote AM or DE programs (Fig. 7B). However, we cannot rule out the existence of different already lineage-restricted progenitors for AM and DE within the *Eomes* single-positive population, as we

Development (2021) 148, dev193789. doi:10.1242/dev.193789

demonstrate that these are spatially separated cell populations. Embryonic clonal lineage analyses had suggested that the common posterior epiblast progenitor for mesoderm and DE represents only a very transient cell population as clones containing both AM and DE cells were only very rarely detected by genetic or labeled lineage tracing (Tzouanacou et al., 2009; Lawson et al., 1991). Novel approaches using a combination of scRNA-seq and molecular recording of cell lineage might be able to provide information about the lineage segregation and relationship of AM and DE in more detail, which was not explored in existing datasets to date (Chan et al., 2019).

In conclusion, this study demonstrates that the generation of the *Eomes*-dependent lineages of AM and DE is spatiotemporally separated during early gastrulation. These cells are molecularly separated early during the differentiation process and share as last common progenitor the *Eomes*-expressing posterior epiblast cells.

MATERIALS AND METHODS

Generation of the *Mesp1*^{mVenus} allele

To generate a fluorescent allele to follow *Mesp1*-expressing cells during gastrulation, we targeted the *Mesp1* locus by homologous recombination to insert a membrane-bound Venus fluorescent protein reporter (mVenus) into the locus. Following the mVenus coding sequence, the *Mesp1* coding sequence, including a 3xFLAG C-terminal tag, was inserted. These two coding sequences are linked by a T2A peptide, which leads to co-translational cleavage of the two proteins, resulting in independent mVenus and Mesp1-3xFLAG proteins (Fig. 2A). The start site of mVenusT2AMesp1-3xFLAG was inserted at the translational start site of the *Mesp1* gene. The 5′ homologous arm spans from an AfeI site upstream of *Mesp1* exon 1 to the *Mesp1* translational start site. The mVenusT2AMesp1-3xFLAG sequence followed by a bGH PolyA signal was then inserted via the 5′ untranslated region EcoRI site and a FspA1 site within *Mesp1* exon 1, thereby deleting 337 bp of the *Mesp1* coding sequence (CDS). A PGK-neomycin resistance (neoR) cassette flanked by loxP sites was inserted downstream of mVenusT2AMesp1-3xFLAG. Between the mVenusT2AMesp1-3xFLAG insert and the neoR cassette, a NdeI site was introduced for screening by Southern blot analysis. The 3′ homology region spanned to an NsiI site downstream of *Mesp1* exon 2 and was flanked by a pMCI-TK negative selection cassette.

Linearized targeting vector was electroporated into CCE mouse ESCs (derived from 129/Sv), and neomycin-resistant and fialuridine-insensitive ESC clones were screened by genomic Southern blot. Genomic DNA was digested with NdeI and probed with an external 3′ probe (wild type allele, 8.4 kb; mutant allele, 4.7 kb; Fig. 2B). Two independent positive clones were injected at the morula stage for chimera generation. *Mesp1*^{mVenus} mice were genotyped by PCR at 62°C annealing temperature to detect the wild-type allele (334 bp) and the knock-in allele (419 bp) using the following primers: wt forward primer 5′-CGCTTCACACCTAGGGCTCA-3′; wt reverse primer 5′-TGTGCGCATACGTAGCTTCTCC-3′; ki forward primer 5′-GCCAATGCAATCCCGAAGTCTC-3′; and ki reverse primer 5′-GCCCTTGGACACCATTGTCTTG-3′ (Fig. 2C). The neomycin cassette was removed by crossing *Mesp1*^{Venus}-positive males to females carrying the *Sox2::Cre* transgene (Vincent and Robertson, 2003).

Mice

Mesp1^{Venus} mice were backcrossed to the NMRI strain and otherwise kept as homozygotes, as they were viable and fertile, and showed no obvious phenotypic differences to wild type. *Eomes*^{mTnG} mice (Probst et al., 2017) were also kept on a NMRI background and were kept as heterozygotes. *Mesp1*^{Cre} mice (Saga et al., 1999) and *R26*^{R}-reporter mice (Soriano, 1999) were kept on a mixed background. Mice were maintained as approved by the Regierungspräsidium Freiburg (license numbers G11/31 and X19/O2F).

Whole-mount immunofluorescence of embryos

Embryos were dissected in PBST (PBS with 0.1% Tween 20), fixed for 1 h in 4% paraformaldehyde (PFA) in PBS at 4°C or on ice, and washed twice in PBST. At this point embryos could be kept at 4°C for at least 1 month. To perform the staining, embryos were permeabilized in 0.3% Triton X-100 in PBST at room temperature for 30 min. Embryos were blocked for 2 h at room temperature in blocking solution [1% bovine serum albumin (BSA) in PBST]. The primary antibodies (antibodies are listed in the next section) were incubated in blocking solution at 4°C overnight. Embryos were washed four times for 5 min each time in PBST at room temperature and then incubated with the secondary antibodies in blocking solutions for 3 h at room temperature. Embryos were washed two times for 5 min each time in PBST at room temperature, stained with DAPI for 30 min at room temperature and washed with PBST. Embryos were stored and imaged in PBST. Imaging was performed using a Zeiss inverted laser-scanning microscope or a Zeiss spinning disk inverted microscope in glass bottom dishes.

Immunofluorescence and LacZ staining on embryo sections

Embryos were fixed in 4% PFA for 1 h at 4°C in the deciduae that were opened to expose the embryo. Deciduae were washed with PBST and then processed through 15% and 30% sucrose/PBS at 4°C, and incubated for at least 1 h in embedding medium (15% sucrose/7.5% gelatin in PBS) at room temperature before embedding. Sections (6-7 µm) were cut with a Leica cryotome. To perform the immunofluorescence staining, sections were washed three times for 5 min each time in PBS and permeabilized in PBST containing 0.2% Triton X-100. Sections were blocked in blocking solution (1% BSA in PBST) for 1 h at room temperature. The primary antibodies were added in blocking solution at 4°C overnight. The slides were washed three times for 5 min each time with PBS and then incubated with the secondary antibody in blocking solution for 1 h at room temperature. After washing the antibody away with PBS, the sections were stained with DAPI in PBST for 5 min and then mounted with ProLong Diamond Antifade Mountant (Life Technologies, P36970), and imaged using an inverted Zeiss Axio Observer Z1 microscope. The following primary antibodies were used: GFP (1:1000, Abcam, ab13970), RFP (1:500, Rockland, 600-401-379), EOMES (1:300, Abcam, ab23345) and FOXA2 (1:500, Cell Signaling Technology 8186). Secondary Alexa Fluor-conjugated antibodies (Life Technologies) were used at a dilution of 1:1000. LacZ staining was performed as described previously (Nagy et al., 2003).

Counting of *Mesp1*$^{-}$ and *Foxa2*-expressing cells in sections of immunofluorescence-stained embryos

For E6.75 and E7.25, all *Mesp1*^{mV}, FOXA2, and *Mesp1*^{mV}/FOXA2 double-positive cells were counted in two embryos each. Entire embryos were imaged from 7 µm transversal sections, and cells were counted from every second section to avoid double counting of same cells. The outer layer (VE and DE) was not counted to avoid the inclusion of VE cells in the analysis. This might lead to a slight underrepresentation of FOXA2-positive cells within the data. In addition, the *Mesp1*^{mV} reporter expression might not entirely reflect endogenous MESP1 protein.

Time-dependent lineage tracing using the *Eomes*^{CreER} allele

Embryos were isolated at E6 or E7 in prewarmed dissection medium [10% fetal calf serum (FCS) in Dulbecco's modified Eagle's Medium (DMEM)/F12 containing Glutamax) and were then placed in embryo culture medium (50% DMEM/F-12 containing GlutaMAX and 50% rat serum) containing 10 µM of 4-OH-tamoxifen (Sigma-Aldrich, H7904; dissolved at 10 mM in DMSO) for 90 min. Embryos were washed three times in dissection medium and placed individually in ibidi eight-well slides in embryo culture medium without 4-OH-tamoxifen. Embryos were cultured for 24 h in regular tissue culture incubators at 37°C with 5% CO_2. A picture was taken of each individual embryo before and after this 24 h period. After the incubation, embryos were fixed and stained with GFP and RFP antibodies as described above, and imaged using a Zeiss spinning disk inverted microscope in glass bottom dishes.

Fluorescent DBA-Lectin staining on whole-mount embryos

Embryos from the lineage-tracing experiments were re-stained with the biotinylated DBA-lectin (Sigma-Aldrich, L6533). Because embryos were

Development (2021) 148, dev193789. doi:10.1242/dev.193789

already stained with GFP and RPF antibodies, no extra blocking step was performed. Embryos were washed in PBST and then the DBA-lectin was added at a dilution of 1:1000 in PBS with 1% BSA at 4°C overnight. The next day embryos were washed three times for 10 min each time with PBST and then incubated with Alexa Fluor-647-streptavidin (Molecular Probes, S21374; dissolved in PBS at 1 mg/ml) in PBS containing 1% BSA at a dilution of 1:500 for 1 h at room temperature. Before adding the streptavidin, the tube was briefly centrifuged. Finally, embryos were washed three times in PBST and imaged using a Zeiss inverted laser-scanning microscope in glass bottom dishes.

Collection of embryo cells for single-cell RNA sequencing

Embryos were dissected in pre-warmed dissection medium (10% FCS in DMEM/F12 containing GlutaMAX) and washed in pre-warmed PBS. For the E6.75 time point, the extra-embryonic part was cut off and a picture of each embryo was taken, and single embryos were transferred into the wells of a pre-warmed non-adhesive 96-well plate containing 40 µl of TrypLE Express (Gibco, 12604013). The wells were coated with FCS before adding the TrypLE. Embryos were incubated at 37°C for 10 min with pipetting up and down once during incubation and at the end to make a single-cell solution. Dissociation was stopped with 120 µl of dissection medium, and cells were centrifuged for 2 min at 200 ***g*** in the 96-well plate. The supernatant was removed and cells from one embryo were resuspended in 200 µl ice-cold PBS. For handpicking, the drop containing the cells was placed in a plastic Petri dish. Cells were picked under a Leica M165 FC binocular using ES-blastocyst injection pipettes (BioMedical Instruments, blunt, bent ID 15 µm, BA=35°) and placed into 1.2 µl lysis buffer containing polyT primer with unique cell barcode. Embryos from the E7.5 time point were cut under the chorion to include the extra-embryonic mesoderm in the analysis. The embryos were imaged and the embryos of one or two litters were pooled and processed in an FCS-coated Eppendorf tube in the same way as the E6.75 embryos. After centrifugation the cells were resuspended in 200 µl PBS and kept on ice until flow sorting.

Single-cell RNA amplification and library preparation

Single-cell RNA sequencing of 576 handpicked cells (E6.75) was performed using the CEL-Seq2 protocol, whereas sequencing of 1152 flow-sorted cells (E7.5) was performed using the mCEL-Seq2 protocol (Hashimshony et al., 2016; Herman et al., 2018). Eighteen libraries with 96 cells each were sequenced per lane on Illumina HiSeq 2500 or 3000 sequencing systems (pair-end multiplexing run) at a depth of ∼200,000-250,000 reads per cell.

Quantification of transcript abundance

Paired-end reads were aligned to the transcriptome using BWA (version 0.6.2-r126) with default parameters (Li and Durbin, 2010). The transcriptome contained all gene models based on the mouse ENCODE VM9 release downloaded from the University of California Santa Cruz genome browser comprising 57,207 isoforms, with 57,114 isoforms mapping to fully annotated chromosomes (1 to 19, X, Y, M). All isoforms of the same gene were merged to a single gene locus. Furthermore, gene loci overlapping by more than 75% were merged to larger gene groups. This procedure resulted in 34,111 gene groups. The right mate of each read pair was mapped to the ensemble of all gene loci and to the set of 92 External RNA Controls Consortium spike-ins in sense direction (Baker et al., 2005). Reads mapping to multiple loci were discarded. The left read contained the barcode information: the first six bases corresponded to the unique molecular identifier (UMI), followed by six bases representing the cell specific barcode. The remainder of the left read contained a polyT stretch. For each cell barcode, the number of UMIs per transcript was counted and aggregated across all transcripts derived from the same gene locus. Based on binomial statistics, the number of observed UMIs was converted into transcript counts (Grün et al., 2014).

Clustering and visualization of mCEL-Seq2 data

Clustering analysis and visualization of the data generated in this study were performed using the RaceID3 algorithm (Herman et al., 2018). The numbers of genes quantified were 19,574 and 20,108 in the E6.75 and E7.5 datasets, respectively. Cells with a total number of transcripts of less than 3000 were discarded, and count data of the remaining cells were normalized by downscaling. Cells expressing more than 2% of *Kcnq1ot1*, a potential marker for low-quality cells (Grün et al., 2016), were not considered for analysis. Additionally, transcripts correlating to *Kcnq1ot1* with a Pearson's correlation coefficient of more than 0.65 were removed. The following parameters were used for RaceID3 analysis: mintotal=3000, minexpr=5, outminc=5 and probthr=10^{-4}. Mitochondrial genes, ribosomal genes, as well as genes starting with 'Gm', were excluded from the analysis. We observed strong batch effects in the E6.75 dataset based on the day of the handpicking. Batch effects were corrected by matching mutual nearest neighbors (MNNs) as described previously (Haghverdi et al., 2018). mnnCorrect function from the scran package was used for the batch correction (Lun et al., 2016). MNN-based batch correction was also performed on the combined *Eomes*-positive dataset used for FateID analysis.

Clustering and visualization of mouse gastrulation atlas data

Processed atlas data on mouse organogenesis from Pijuan-Sala et al. (2019) were downloaded from ArrayExpress (accession number: E-MTAB-6967). The following time points and sequencing batches were analyzed: E6.5 (sequencing batch 1), E6.75 (sequencing batch 1), E7 (sequencing batches 1, 2 and 3), E7.25 (sequencing batch 2) and E7.5 (sequencing batches 1 and 2). Cells defined as doublets in the study were removed from the analysis. Integration of datasets from different time points and sequencing batches was performed using Seurat version 3 with default settings (Stuart et al., 2019). Ribosomal genes (small and large subunits), as well as genes with *Gm*-identifiers were excluded from the data before integration. The integrated dataset contained 45,196 cells. A focused analysis of *Eomes*-expressing cells was performed using VarID (Grün, 2020). From the complete dataset containing 45,196 cells, cells with a total number of transcripts of less than 6000 were discarded, and count data of the remaining cells were normalized by downscaling. Cells having normalized *Eomes* transcript counts of more than 0.3 were considered as *Eomes*-positive (14,329 cells) and further clustered and visualized using VarID with the following parameters: large=TRUE, pcaComp=100, regNB=TRUE, batch=batch, knn=50, alpha=10 and no_cores=20. Each batch contained cells from different time points and sequencing libraries. Dimensionality reduction of the datasets was performed using UMAP.

FateID analysis

In order to investigate the transcriptional priming of single *Eomes*-expressing cells towards the mesodermal and DE fates, FateID (Herman et al., 2018) was run on the mouse gastrulation data (Pijuan-Sala et al., 2019) separately at the following different time points: E6.5/E6.75, E7.0 and E7.25 with cells having normalized *Eomes* transcript counts of more than 0.3 using default parameters. *Mesp1*-positive (mesoderm specified, normalized transcript count of more than 0.3) and *Foxa2*-positive (DE specified, normalized transcript count more than 0.3) cells were used as target cells. Extra-embryonic cells were excluded from the FateID analysis and t-distributed stochastic neighbor embedding was used for dimensional reduction and visualization of the results. Differential gene expression analysis was performed between cells biased towards one of the lineages with a fate bias probability of more than 0.5 using the diffexpnb function. UMAP coordinates from the VarID analysis were used for the visualization of the results.

Differential gene expression analysis

Differential gene expression analysis was performed using the diffexpnb function of the RaceID3 algorithm. Differentially expressed genes between two subgroups of cells were identified in a similar way to a previously published method (Anders and Huber, 2010). First, negative binomial distributions reflecting the gene expression variability within each subgroup were inferred based on the background model for the expected transcript count variability computed by RaceID3. Using these distributions, a *P*-value for the observed difference in transcript counts between the two subgroups was calculated and multiple testing corrected by the Benjamini-Hochberg method.

Acknowledgements
We thank Yumiko Saga for the *Mesp1*^{Cre} mouse line; Thilo Bass for technical assistance; the staff of the Life Imaging Center of the Albert-Ludwigs-Universität Freiburg for help with microscopy; Shankar Srinivas for teaching Simone Probst mouse imaging techniques; Matthias Weiß and the CEMT team for support with animal care; and Katrin Schüle for critical reading of the manuscript.

Competing interests
The authors declare no competing or financial interests.

Author contributions
Conceptualization: S.P., S.J.A.; Methodology: S.P., S., S.J.A.; Formal analysis: S.P., S., J.T., S.J.A.; Investigation: S., S.; Data curation: S.; Writing - original draft: S.P., S.J.A.; Writing - review & editing: S.P., S., J.T., D.G., S.J.A.; Visualization: S.P., S.; Supervision: C.S., D.G., S.J.A.; Funding acquisition: C.S., D.G., S.J.A.

Funding
This work was supported by the Deutsche Forschungsgemeinschaft through the Emmy Noether- and Heisenberg-Programs (AR 732/1-1/2/3 and AR 732/3-1), project grant AR 732/2-1, project B07 of SFB 1140 (project ID 246781735), project A03 of SFB 850 (project ID 89986987 to S.J.A.) and by Germany's Excellence Strategy (CIBSS – EXC-2189 – Project ID 390939984 to D.G. and S.J.A.). D.G. was supported by the Max-Planck-Gesellschaft, the Deutsche Forschungsgemeinschaft (SPP1937 GR4980/1-1, GR4980/3-1 and GRK2344 MeInBio) the European Research Council (818846 – ImmuNiche – ERC-2018-COG) and the Behrens-Weise-Foundation. C.S was supported by the Deutsche Forschungsgemeinschaft (SCHW 1708/2-1)

Data availability
The primary read files, as well as expression count files for the single-cell RNA-sequencing datasets generated in this study, have been deposited in GEO under accession number GSE151824.

Supplementary information
Supplementary information available online at
https://dev.biologists.org/lookup/doi/10.1242/dev.193789.supplemental

Peer review history
The peer review history is available online at
https://dev.biologists.org/lookup/doi/10.1242/dev.193789.reviewer-comments.pdf

References

Anders, S. and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106. doi:10.1186/gb-2010-11-10-r106

Ang, S. L., Wierda, A., Wong, D., Stevens, K. A., Cascio, S., Rossant, J. and Zaret, S. (1993). The formation and maintenance of the definitive endoderm lineage in the mouse: involvement of HNF3/forkhead proteins. *Development* **119**, 1301-1315.

Arkell, R. M., Fossat, N. and Tam, P. P. L. (2013). Wnt signalling in mouse gastrulation and anterior development: new players in the pathway and signal output. *Curr. Opin. Genet. Dev.* **23**, 454-460. doi:10.1016/j.gde.2013.03.001

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. doi:10.1038/nrm2618

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. doi:10.1242/dev.014357

Baker, S. C., Bauer, S. R., Beyer, R. P., Brenton, J. D., Bromley, B., Burrill, J., Causton, H., Conley, M. P., Elespuru, R., Fero, M. et al. (2005). The External RNA Controls Consortium: a progress report. *Nat. Methods* **2**, 731-734. doi:10.1038/nmeth1005-731

Bardot, E., Calderon, D., Santoriello, F., Han, S., Cheung, K., Jadhav, B., Burtscher, I., Artap, S., Jain, R., Epstein, J. et al. (2017). Foxa2 identifies a cardiac progenitor population with ventricular differentiation potential. *Nat. Commun.* **8**, 14428. doi:10.1038/ncomms14428

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. doi:10.1038/35082103

Burtscher, I. and Lickert, H. (2009). Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. *Development* **136**, 1029-1038. doi:10.1242/dev.028415

Chan, S. S.-K., Shi, X., Toyama, A., Arpke, R. W., Dandapat, A., Iacovino, M., Kang, J., Le, G., Hagen, H. R., Garry, D. J. et al. (2013). Mesp1 patterns mesoderm into cardiac, hematopoietic, or skeletal myogenic progenitors in a context-dependent manner. *Cell Stem Cell* **12**, 587-601. doi:10.1016/j.stem.2013.03.004

Chan, M. M., Smith, Z. D., Grosswendt, S., Kretzmer, H., Norman, T. M., Adamson, B., Jost, M., Quinn, J. J., Yang, D., Jones, M. G. et al. (2019). Molecular recording of mammalian embryogenesis. *Nature* **570**, 77-82. doi:10.1038/s41586-019-1184-5

Conlon, F. L., Lyons, K. M., Takaesu, N., Barth, K. S., Kispert, A., Hermann 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.

Costello, I., Pimeisl, I.-M., Dräger, S., Bikoff, E. K., Robertson, E. J. and Arnold, S. J. (2011). The T-box transcription factor Eomesodermin acts upstream of Mesp1 to specify cardiac mesoderm during mouse gastrulation. *Nat. Cell Biol.* **13**, 1084-1091. doi:10.1038/ncb2304

D'Amour, K. A., Agulnick, A. D., Eliazer, S., Kelly, O. G., Kroon, E. and Baetge, E. E. (2005). Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nat. Biotechnol.* **23**, 1534-1541. doi:10.1038/nbt1163

Dunn, N. R., Vincent, S. D., Oxburgh, L., Robertson, E. J. and Bikoff, E. K. (2004). Combinatorial activities of Smad2 and Smad3 regulate mesoderm formation and patterning in the mouse embryo. *Development* **131**, 1717-1728. doi:10.1242/dev.01072

Frank, D. U., Elliott, S. A., Park, E. J., Hammond, J., Saijoh, Y. and Moon, A. M. (2007). System for inducible expression of cre-recombinase from the Foxa2 locus in endoderm, notochord, and floor plate. *Dev. Dyn.* **236**, 1085-1092. doi:10.1002/dvdy.21093

Galonska, C., Smith, Z. D. and Meissner, A. (2014). In Vivo and in vitro dynamics of undifferentiated embryonic cell transcription factor 1. *Stem Cell Rep.* **2**, 245-252. doi:10.1016/j.stemcr.2014.01.007

Grün, D. (2020). Revealing dynamics of gene expression variability in cell state space. *Nat. Methods* **17**, 45-49. doi:10.1038/s41592-019-0632-3

Grün, D., Kester, L. and Van Oudenaarden, A. (2014). Validation of noise models for single-cell transcriptomics. *Nat. Methods* **11**, 637-640. doi:10.1038/nmeth.2930

Grün, D., Lyubimova, A., Kester, L., Wiebrands, K., Basak, O., Sasaki, N., Clevers, H. and Van Oudenaarden, A. (2015). Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* **525**, 251-255. doi:10.1038/nature14966

Grün, D., Muraro, M. J., Boisset, J.-C., Wiebrands, K., Lyubimova, A., Dharmadhikari, G., van den Born, M., van Es, J., Jansen, E., Clevers, H. et al. (2016). De novo prediction of stem cell identity using single-cell transcriptome data. *Cell Stem Cell* **19**, 266-277. doi:10.1016/j.stem.2016.05.010

Haghverdi, L., Lun, A. T. L., Morgan, M. D. and Marioni, J. C. (2018). Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat. Biotechnol.* **36**, 421-427. doi:10.1038/nbt.4091

Hashimshony, T., Senderovich, N., Avital, G., Klochendler, A., de Leeuw, Y., Anavy, L., Gennert, D., Li, S., Livak, K. J., Rozenblatt-Rosen, O. et al. (2016). CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol.* **17**, 77. doi:10.1186/s13059-016-0938-8

Herman, J. S., Sagar, and Grün, D., (2018). FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. *Nat. Methods*, **15**, 379-386. doi:10.1038/nmeth.4662

Imuta, Y., Kiyonari, H., Jang, C.-W., Behringer, R. R. and Sasaki, H. (2013). Generation of knock-in mice that express nuclear enhanced green fluorescent protein and tamoxifen-inducible Cre recombinase in the notochord from Foxa2 and T loci. *Genesis* **51**, 210-218. doi:10.1002/dvg.22376

Ivanovitch, K., Soro-Barrio, P., Chakravarty, P., Jones, R. A., Neda, S., Gharavy, M., Stamataki, D., Delile, J., Smith, J. C. and Briscoe, J. (2020). Ventricular, atrial and outflow tract heart progenitors arise from spatially and molecularly distinct regions of the primitive streak. *bioRxiv*, **3**, 2020.07.12.198994.

Kartikasari, A. E. R., Zhou, J. X., Kanji, M. S., Chan, D. N., Sinha, A., Grapin-Botton, A., Magnuson, M. A., Lowry, W. E. and Bhushan, A. (2013). The histone demethylase Jmjd3 sequentially associates with the transcription factors Tbx3 and Eomes to drive endoderm differentiation. *EMBO J.* **32**, 1393-1408. doi:10.1038/emboj.2013.78

Kimber, S. J. (1986). Distribution of lectin receptors in postimplantation mouse embryos at 6-8 days gestation. *Am. J. Anat.* **177**, 203-219. doi:10.1002/aja.1001770207

Kimelman, D. and Griffin, K. J. P. (2000). Vertebrate mesendoderm induction and patterning. *Curr. Opin. Genet. Dev.* **10**, 350-356. doi:10.1016/S0959-437X(00)00095-2

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**, 4691-4701.

Kinder, S. J., Loebel, D. A. F. and Tam, P. P. L. (2001). Allocation and early differentiation of cardiovascular progenitors in the mouse embryo. *Trends Cardiovasc. Med.* **11**, 177-184. doi:10.1016/S1050-1738(01)00091-3

Kitajima, S., Takagi, A., Inoue, T. and Saga, Y. (2000). MesP1 and MesP2 are essential for the development of cardiac mesoderm. *Development* **127**, 3215-3226.

Koch, F., Scholze, M., Wittler, L., Schifferl, D., Sudheer, S., Grote, P., Timmermann, B., Macura, K. and Herrmann, B. G. (2017). Antagonistic

activities of Sox2 and brachyury control the fate choice of neuro-mesodermal progenitors. *Dev. Cell* **42**, 514-526.e7. doi:10.1016/j.devcel.2017.07.021

Kubo, A., Shinozaki, K., Shannon, J. M., Kouskoff, V., Kennedy, M., Woo, S., Fehling, H. J. and Keller, G. (2004). Development of definitive endoderm from embryonic stem cells in culture. *Development* **131**, 1651-1662. doi:10.1242/dev.01044

Lawson, K. A. (1999). Fate mapping the mouse embryo. *Int. J. Dev. Biol.* **43**, 773-775.

Lawson, K. A. and Pedersen, R. A. (1987). Cell fate, morphogenetic movement and population kinetics of embryonic endoderm at the time of germ layer formation in the mouse. *Development* **101**, 627-652.

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.

Lescroart, F., Chabab, S., Lin, X., Rulands, S., Paulissen, C., Rodolosse, A., Auer, H., Achouri, Y., Dubois, C., Bondue, A. et al. (2014). Early lineage restriction in temporally distinct populations of Mesp1 progenitors during mammalian heart development. *Nat. Cell Biol.* **16**, 829-840. doi:10.1038/ncb3024

Lescroart, F., Wang, X., Lin, X., Swedlund, B., Gargouri, S., Sànchez-Dànes, A., Moignard, V., Dubois, C., Paulissen, C., Kinston, S. et al. (2018). Defining the earliest step of cardiovascular lineage segregation by single-cell RNA-seq. *Science* **359**, 1177-1181. doi:10.1126/science.aao4174

Li, H. and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589-595. doi:10.1093/bioinformatics/btp698

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. doi:10.1038/11932

Lun, A. T. L., McCarthy, D. J. and Marioni, J. C. (2016). A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Research* **5**, 2122. doi:10.12688/f1000research.9501.2

Macdonald, S. T., Bamforth, S. D., Chen, C.-M., Farthing, C. R., Franklyn, A., Broadbent, C., Schneider, J. E., Saga, Y., Lewandoski, M. and Bhattacharya, S. (2008). Epiblastic Cited2 deficiency results in cardiac phenotypic heterogeneity and provides a mechanism for haploinsufficiency. *Cardiovasc. Res.* **79**, 448-457. doi:10.1093/cvr/cvn101

Mohammed, H., Hernando-Herraez, I., Savino, A., Scialdone, A., Macaulay, I., Mulas, C., Chandra, T., Voet, T., Dean, W., Nichols, J. et al. (2017). Single-cell landscape of transcriptional heterogeneity and cell fate decisions during mouse early gastrulation. *CellReports* **20**, 1215-1228. doi:10.1016/j.celrep.2017.07.009

Monaghan, A. P., Kaestner, K. H., Grau, E. and Schütz, G. (1993). Postimplantation expression patterns indicate a role for the mouse forkhead/HNF-3 alpha, beta and gamma genes in determination of the definitive endoderm, chordamesoderm and neuroectoderm. *Development* **119**, 567-578.

Muzumdar, M. D., Tasic, B., Miyamichi, K., Li, L. and Luo, L. (2007). A global double-fluorescent Cre reporter mouse. *Genesis* **45**, 593-605. doi:10.1002/dvg.20335

Nagy, A., Gerstenstein, M., Behringer, R. R. and Vintersten, K. (2003). *Manipulating the Mouse Embryo*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Nowotschin, S., Costello, I., Piliszek, A., Kwon, G. S., Mao, C.-, Klein, W. H., Robertson, E. J. and Hadjantonakis, A.-K. (2013). The T-box transcription factor Eomesodermin is essential for AVE induction in the mouse embryo. *Genes Dev.* **27**, 997-1002. doi:10.1101/gad.215152.113

Parameswaran, M. and Tam, P. P. L. (1995). Regionalisation of cell fate and morphogenetic movement of the mesoderm during mouse gastrulation. *Dev. Genet.* **17**, 16-28. doi:10.1002/dvg.1020170104

Park, E. J., Ogden, L. A., Talbot, A., Evans, S., Cai, C.-L., Black, B. L., Frank, D. U. and Moon, A. M. (2006). Required, tissue-specific roles for Fgf8 in outflow tract formation and remodeling. *Development* **133**, 2419-2433. doi:10.1242/dev.02367

Park, E. J., Sun, X., Nichol, P., Saijoh, Y., Martin, J. F. and Moon, A. M. (2008). System for tamoxifen-inducible expression of cre-recombinase from the Foxa2 locus in mice. *Dev. Dyn.* **237**, 447-453. doi:10.1002/dvdy.21415

Peng, G., Suo, S., Cui, G., Yu, F., Wang, R., Chen, J., Chen, S., Liu, Z., Chen, G., Qian, Y. et al. (2019). Molecular architecture of lineage allocation and tissue organization in early mouse embryo. *Nature* **572**, 528-532. doi:10.1038/s41586-019-1469-8

Perantoni, A. O., Timofeeva, O., Naillat, F., Richman, C., Pajni-Underwood, S., Wilson, C., Vainio, S., Dove, L. F. and Lewandoski, M. (2005). Inactivation of FGF8 in early mesoderm reveals an essential role in kidney development. *Development* **132**, 3859-3871. doi:10.1242/dev.01945

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. doi:10.1038/s41586-019-0933-9

Pimeisl, I.-M., Tanriver, Y., Daza, R. A., Vauti, F., Hevner, R. F., Arnold, H.-H. and Arnold, S. J. (2013). Generation and characterization of a tamoxifen-inducible Eomes CreER mouse line. *Genesis* **51**, 725-733. doi:10.1002/dvg.22417

Probst, S. and Arnold, S. J. (2017). Eomesodermin-at dawn of cell fate decisions during early embryogenesis. *Curr. Top. Dev. Biol.* **122**, 93-115. doi:10.1016/bs.ctdb.2016.09.001

Probst, S., Daza, R. A., Bader, N., Hummel, J. F., Weiß, M., Tanriver, Y., Hevner, R. F. and Arnold, S. J. (2017). A dual-fluorescence reporter in the Eomes locus for live imaging and medium-term lineage tracing. *Genesis* **55**, e23043. doi:10.1002/dvg.23043

Rivera-Pérez, J. A. and Hadjantonakis, A.-K. (2014). The dynamics of morphogenesis in the early mouse embryo. *Cold Spring Harbor Perspect. Biol.* **7**, a015867. doi:10.1101/cshperspect.a015867

Rivera-Pérez, J. A., Mager, J. and Magnuson, T. (2003). Dynamic morphogenetic events characterize the mouse visceral endoderm. *Dev. Biol.* **261**, 470-487. doi:10.1016/S0012-1606(03)00302-6

Robertson, E. J. (2014). Dose-dependent Nodal/Smad signals pattern the early mouse embryo. *Semin. Cell Dev. Biol.* **32**, 73-79. doi:10.1016/j.semcdb.2014.03.028

Saga, Y., Miyagawa-Tomita, S., Takagi, A., Kitajima, S., Miyazaki, J. and Inoue, T. (1999). MesP1 is expressed in the heart precursor cells and required for the formation of a single heart tube. *Development* **126**, 3437-3447.

Saga, Y., Kitajima, S. and Miyagawa-Tomita, S. (2000). Mesp1 expression is the earliest sign of cardiovascular development. *Trends Cardiovasc. Med.* **10**, 345-352. doi:10.1016/S1050-1738(01)00069-X

Sasaki, H. and Hogan, B. L. (1993). Differential expression of multiple fork head related genes during gastrulation and axial pattern formation in the mouse embryo. *Development* **118**, 47-59.

Scialdone, A., Tanaka, Y., Jawaid, Y., Moignard, V., Wilson, N. K., Macaulay, I. C., Marioni, J. C. and Göttgens, B. (2016). Resolving early mesoderm diversification through single-cell expression profiling. *Nature* **535**, 289-293. doi:10.1038/nature18633

Snow, M. H. L. (1977). Gastrulation in the mouse: growth and regionalization of the epiblast. *Development* **42**, 293-303.

Soriano, P. (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat. Genet.* **21**, 70-71. doi:10.1038/5007

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., III, Hao, Y., Stoeckius, M., Smibert, P. and Satija, R. (2019). Comprehensive integration of single-cell data. *Cell* **177**, 1888-1902.e21. doi:10.1016/j.cell.2019.05.031

Tada, S., Era, T., Furusawa, C., Sakurai, H., Nishikawa, S., Kinoshita, M., Nakao, K., Chiba, T. and Nishikawa, S.-I. (2005). Characterization of mesendoderm: a diverging point of the definitive endoderm and mesoderm in embryonic stem cell differentiation culture. *Development* **132**, 4363-4374. doi:10.1242/dev.02005

Tam, P. P. and Beddington, R. S. (1987). The formation of mesodermal tissues in the mouse embryo during gastrulation and early organogenesis. *Development* **99**, 109-126.

Tam, P. P. L. and Behringer, R. R. (1997). Mouse gastrulation: the formation of a mammalian body plan. *Mech. Dev.* **68**, 3-25. doi:10.1016/S0925-4773(97)00123-8

Tam, P. P., 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.

Teo, A. K. K., Arnold, S. J., Trotter, M. W. B., Brown, S., Ang, L. T., Chng, Z., Robertson, E. J., Dunn, N. R. and Vallier, L. (2011). Pluripotency factors regulate definitive endoderm specification through eomesodermin. *Genes Dev.* **25**, 238-250. doi:10.1101/gad.607311

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. doi:10.1038/s41556-019-0423-1

Tzouanacou, E., Wegener, A., Wymeersch, F. J., Wilson, V. and Nicolas, J.-F. (2009). Redefining the progression of lineage segregations during mammalian embryogenesis by clonal analysis. *Dev. Cell* **17**, 365-376. doi:10.1016/j.devcel.2009.08.002

Van Den Ameele, J., Tiberi, L., Bondue, A., Paulissen, C., Herpoel, A., Iacovino, M., Kyba, M., Blanpain, C. and Vanderhaeghen, P. (2012). Eomesodermin induces Mesp1 expression and cardiac differentiation from embryonic stem cells in the absence of Activin. *EMBO Rep.* **13**, 355-362. doi:10.1038/embor.2012.23

Vincent, S. D. and Robertson, E. J. (2003). Highly efficient transgene-independent recombination directed by a maternally derived SOX2CRE transgene. *Genesis* **37**, 54-56. doi:10.1002/gene.10226

Vincent, S. D., Dunn, N. R., Hayashi, S., Norris, D. P. and Robertson, E. J. (2003). Cell fate decisions within the mouse organizer are governed by graded Nodal signals. *Genes Dev.* **17**, 1646-1662. doi:10.1101/gad.1100503

Viotti, M., Foley, A. C. and Hadjantonakis, A.-K. (2014a). Gutsy moves in mice: cellular and molecular dynamics of endoderm morphogenesis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369**, 20130547. doi:10.1098/rstb.2013.0547

Viotti, M., Nowotschin, S. and Hadjantonakis, A.-K. (2014b). SOX17 links gut endoderm morphogenesis and germ layer segregation. *Nat. Cell Biol.* **16**, 1146-1156. doi:10.1038/ncb3070

Wen, J., Zeng, Y., Fang, Z., Gu, J., Ge, L., Tang, F., Qu, Z., Hu, J., Cui, Y., Zhang, K. et al. (2017). Single-cell analysis reveals lineage segregation in early post-implantation mouse embryos. *J. Biol. Chem.* **292**, 9840-9854. doi:10.1074/jbc.M117.780585

Wymeersch, F. J., Huang, Y., Blin, G., Cambray, N., Wilkie, R., Wong, F. C. K. and Wilson, V. (2016). Position-dependent plasticity of distinct progenitor types in the primitive streak. *eLife* **5**, e10042. doi:10.7554/eLife.10042

Yoshida, T., Vivatbutsiri, P., Morriss-Kay, G., Saga, Y. and Iseki, S. (2008). Cell lineage in mammalian craniofacial mesenchyme. *Mech. Dev.* **125**, 797-808. doi:10.1016/j.mod.2008.06.007

Zhang, Z., Cerrato, F., Xu, H., Vitelli, F., Morishima, M., Vincentz, J., Furuta, Y., Ma, L., Martin, J. F., Baldini, A. et al. (2005). Tbx1 expression in pharyngeal epithelia is necessary for pharyngeal arch artery development. *Development* **132**, 5307-5315. doi:10.1242/dev.02086