Title: Time-Aligned Hourglass Gastrulation Models in Rabbit and Mouse

Mayshar Y[1,3], Raz O[1,2,3], Cheng S[1], Ben-Yair R[1], Hadas R[1], Reines N[1], Mittnenzweig M[1], Ben-Kiki O[1], Lifshitz A[1], Tanay A[1,2#], and Stelzer Y[1#]

1 Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel

2 Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel

3 Equally contributing authors

Correspondence:

Yonatan Stelzer Yonatan.stelzer@weizmann.ac.il

Amos Tanay Amos.tanay@weizmann.ac.il

ABSTRACT

The hourglass model describes the convergence of species within the same phylum to a similar body plan during development, yet the molecular mechanisms underlying this phenomenon in mammals remain poorly described. Here, we compare rabbit and mouse time-resolved differentiation trajectories to revisit this model at single cell resolution. We modeled gastrulation dynamics using hundreds of embryos sampled between gestation days 6.0-8.5, and compare the species using a new framework for time-resolved single-cell differentiation-flows analysis. We find convergence toward similar cell state compositions at E7.5, underlied by quantitatively conserved expression of 76 transcription factors, despite divergence in surrounding trophoblast and hypoblast signaling. However, we observed noticeable changes in specification timing of some lineages, and divergence of primordial germ cells programs, which in the rabbit do not activate mesoderm genes. Comparative analysis of temporal differentiation models provides a new basis for studying the evolution of gastrulation dynamics across mammals.

INTRODUCTION

Early mammalian development follows a generally conserved sequence of events resulting in extraordinarily similar embryos at the time of organogenesis, representing a so-called evolutionary hourglass effect[1-3]. The process of gastrulation involves the formation of the embryonic germ layers from the pluripotent epiblast, and laying out of the basic embryonic axes. This critical stage of development has been mainly characterized in the mouse model, in which the developing blastocyst takes on the form of a cup shape, or egg cylinder. The mouse gastrula is however, highly distinct among vertebrates, as most mammals begin gastrulation as a planar embryonic disc. Such gross structural disparity is expected to have dramatic effects on gastrulation by shaping cellular mechanics and spatiotemporal interactions. Moreover, wide variation has been observed between species in relation to implantation strategies, and the development and orientation of the extraembryonic tissues[4,5]. The rabbit stands out among possible alternative mammalian model organisms by presenting many of the advantages of the mouse, namely, relatively short gestation, and large litters that can be accurately timed. Importantly, the rabbit was shown to more closely resemble human development, in particular with respect to the specification of primordial germ cells (PGCs)[6,7]. Interestingly, implantation of rabbit embryos is preceded by a period of rapid growth as an expanded blastocyst[8,9], and in this form gastrulation is initiated[10]. Further, while anti-mesometrial implantation occurs at embryonic day (E)7, placental (mesometrial) implantation occurs only at E8[11], close to the time of somitogenesis. This is in contrast to the mouse, in which implantation occurs early after hatching of the blastocyst at around E4.5, long before the onset of gastrulation.

Single cell transcriptomics and multi-omics are recently being used to map transcriptional and chromatin states in early embryonic development[7,12-22]. This resulted in comprehensive atlases that greatly enrich and refine previous imaging-based data by characterizing precisely transcriptional programs at high cellular resolution. Merging inferred cell states into a manifold model facilitates the inference of cellular differentiation dynamics, at first using computational tools searching for parsimonious differentiation trajectories[23,24]. But the highly parallel and complex nature of embryonic development, and gastrulation in particular, calls for refined strategies to map precisely multi-lineage dynamics of the ensemble of single cells constituting the developing embryo. We recently developed such an approach by sampling single cells from individual embryos, and by inferring a differentiation flow model that describes how single cells coordinately change their manifold state over time[14]. This approach can also be instrumental to elucidate the intrinsic and extrinsic function of genes in embryonic cell specification[25]. Together

with technologies for mapping cell lineage trees (Reviewed in[26]), and spatial transcriptomics[27-29], this can be even further refined to obtain a true holistic view of embryonic development.

Importantly, embryonic atlases were recently described from precious and limited primate[30,31] and human[20,32] samples, but the temporal resolution in these systems is limited. Here, we adapt the rabbit as a second high resolution single cell model for mammalian gastrulation. We generated a time resolved single-embryo, single-cell model from over 100 embryos, and used it for comprehensive synthesis of morphological and molecular classification of rabbit embryonic stages. We define the manifold alignment problem as the evolutionary task of matching molecular states (e.g. cell types) and differentiation dynamics between related species, and describe how to approach this problem when comparing mouse and rabbit models. This leads to characterization of striking resemblance in gastrulation programs between the two species, but also uncovers key differences in lineage coordination and differentiation timing, setting the stage for understanding how conservation of the mammalian intrinsic cell state repertoire can be compatible with changes in morphology and extrinsic signaling.

RESULTS

Charting the rabbit single cell gastrulation manifold using 120 individually profiled E6.0–E8.5 embryos

While overall gestation of the rabbit is significantly longer than that of the mouse (~30 vs ~19 days), major early gestational milestones such as somitogenesis, limb, and eye development occur at highly similar times[33]. Also, both species begin gastrulation shortly after E6.0. However, despite the similarity in timing of gastrulation, and in form during organogenesis, several key differences in the process of gastrulation exist between the species: (i) The rabbit embryo is much larger than the mouse during all corresponding developmental stages. (ii) Uterine implantation, which in the mouse occurs prior to gastrulation, only occurs gradually in the rabbit and at much later stages between embryonic days E7-8. (iii) The egg cylinder shape of the mouse facilities asymmetric interaction between the epiblast and extraembryonic tissues (distal-proximal axis), as opposed to the disc shape of the rabbit embryo which in practice lacks such an axis. (iv) Establishment of the mouse anterior-posterior axis involves coordinated signals from the anterior visceral endoderm (AVE, corresponding to rabbit's anterior hypoblast) and the extraembryonic ectoderm (ExEc). However, in the rabbit, the part of the ExEc that is in contact with the embryo proper (polar trophectoderm, also known as Rauber's layer) deteriorates and disappears rapidly at the onset of gastrulation[34] and is therefore unlikely to play a major signaling role at this stage.

To better understand the conservation of the gastrulation program and its archetypical outcome given such a diverse context, we sought to generate a time-resolved atlas of rabbit gastrulation using a similar approach we previously implemented in the mouse[14]. To this end, we performed scRNA-seq[35] on 160K cells of 120 individual embryos including their immediate surrounding extraembryonic tissue. Each embryo was imaged and assigned a morphological stage[36], and profiles of single cells were associated with their corresponding embryo. Litters were timed to ensure adequate representation of the structural diversity from pre-gastrulation up to the 12 somite stage (**Figures 1A, S1, S2A-D**). Following quality control, we retained 130K profiles of sufficient coverage (**Figure S2A,** median UMI count of 7,142) and constructed a transcriptional manifold model using the metacells2 algorithm[37]. We have extended the rabbit gene annotation to the yet unannotated oryCun3.0 genome assembly using the previous state-of-the-art oryCun2.0 gene annotation as well as the mouse mm9 and Human hg19 gene annotations. Furthermore, we account for any unannotated transcriptional activity that can be mapped to oryCun3.0, adding 120K putative 3'-end loci to the OryCun3.0 25K annotated TSSs (**Figure S2G**). In total we defined 3,584 metacells (MCs), facilitating precise quantitative and comparative analysis of the represented transcriptional states (**Figure 1B**). We annotated ExEc and extraembryonic endoderm (ExEnd) cells and analyzed them separately as elaborated below. We then annotated the embryo manifold using 41 different cell types/states. The entire model is accessible interactively at tanaylab.weizmann.ac.il/rabemb_wex. For subsequent analysis of embryo temporal dynamics, embryos were excluded for either insufficient embryonic cells (<20 cells, n=6), or obvious outliers in cell type composition (indicating a technical error, e.g. embryos with only posterior cell types, n=6).

Refining rabbit embryonic stages by integrating morphology and transcriptional analysis

To describe rabbit gastrulation on an absolute temporal axis, we performed morphology-based ranking of the embryos and ranking by k-NN similarities of single-cell profiles (**Figures S1 and S2H**). Both measurements showed high concordance and generally matched the respective litter's absolute gestation time (**Figure S2E-F**). We then defined twelve temporal bins comprising groups of embryos with similar cell state compositions (see **Methods**), which were used in subsequent modeling and refinement of the current morphology-based nomenclature (**Figure 1C**). Under the assumption that all embryos represent discreet times along a shared process, each bin was manually assigned a representative time considering the absolute gestation time of embryos comprising it, defined as E_t (See **Methods**, and **Table S1**). Each embryo was also

assigned to a specific morphological developmental stage, resulting in a multi-tiered annotation of the rabbit gastrulation continuum (**Figure 1D**). We find that cell type composition changes in many cases reflect visible morphological events, such as; a) the elongation of the embryo from its initial circular form that coincides with the appearance of the mesoderm and anterior primitive streak; b) the appearance of a narrowed primitive streak structure (E_t7.2, stage 4) together with definitive endoderm and ectoderm, hematoendothelial cells, and expansion of the extraembryonic nascent mesoderm, and, c) the brief window of time between the first observation of the notochordal process anterior to the node (E_t 7.5, stage 4c), and its full anterior extension (E_t7.7, stage 5), that is accompanied by an explosion of advanced cell types and disappearance of the nascent mesoderm. Cell type composition analysis also allows direct comparison with similar mouse analysis, providing an empirical measure of corresponding developmental stages across species. In summary, single cell analysis of single rabbit embryos provided us with high resolution description of the rabbit gastrulation stages. It thereby set the stage for quantitative modeling of the differentiation process, transforming the pluripotent epiblast into the rich repertoire of embryonic lineages during 2.5 days from E6.0 to E8.6.

A network flow model tracks rabbit gastrulation dynamics in absolute time

To infer a differentiation model from the single-cell and single embryo rabbit dataset, we used an improved version of our network flows algorithm that was initially demonstrated in the mouse[14] (see **Methods**). The algorithm resolved differentiation flows for metacells distributed over the twelve time-bins, balancing similarities between expression states and estimation of cell proliferation rates (**Figures 2, S3A, and S3B**). The latter was performed by computing co-expression of S-phase and M-phase related genes, and quantifying a distinctive non-proliferating cell subpopulation (**Methods**, **Figure S3C**). Consistent with the extremely high proliferation rate in gastrulating rabbit embryos (estimated doubling every 6-7 hours[9]) we observe that almost all cells express the cell cycle programs. Nevertheless, some metacells of the endoderm lineage, notochord, and cardiomyocytes were linked with considerably lower expression of S- and M-phase related genes, suggesting a slow-down in proliferation rates in these cell states (**Figure S3D**). Estimations of the fraction of non-dividing cells per cell state and time were used to calibrate differential growth rates and growth loss estimation for the flows 'mass conservation' constraint (See **Methods**). Yet, for some extraembryonic lineages, we could not assume such conservation due to incomplete or biased sampling. Namely, the rabbit choriovitelline placenta tissue is extensive and lacks visible borders by standard light microscopy. Therefore, rabbit embryo resection entails arbitrary removal of much of the yolk sac endoderm and related extraembryonic mesoderm, which extends radially from the embryo and thus progressively undersampled. Hence,

analysis of cellular hierarchy was excluded for hematopoietic, ExEnd, and trophectoderm lineages (even though the timing of their sequential appearance remains highly informative, as discussed below). Apart from these lineages, the rabbit flow model uncovers the exceptionally coordinated process of differentiation in the ectoderm, mesoderm, and endoderm with high temporal and quantitative precision. Importantly, this model is based on similar assumptions to those used to create the analogous mouse model and therefore allows direct quantitative alignment between the two species.

Manifold alignment uncovers highly conserved gastrulation states in rabbit and mouse
Given fully time resolved manifold and flow models for rabbit and mouse gastrulation, we wished to define a framework for their principled comparison. Previous comparative single cell atlas work analyzed either highly diverged species with non-alignable genomes[38,39], or in the case of mammals, using lower resolution data[19,30,40,41]. These analyses therefore were restricted to qualitative and coarse grain comparisons. To facilitate a high-resolution comparative analysis, we developed a manifold alignment algorithm that simultaneously computes matching states between mouse and rabbit while adjusting functional gene orthology. Our approach begins with a naïve assumption of complete functional conservation of 9,748 orthologous gene pairs for which significant transcription is detected in both species. This pairing was used for initial approximation of the similarity between all metacells in the two manifolds (**Figure S4A**), identifying 79 reciprocally best orthologous metacell (RBOM) pairs (**Figure S4B**). Notably, the identification of such states is a highly non-trivial result, showing that the two manifolds are indeed alignable over a very rich collection of transcriptional states. Using the 79 RBOMs as a new common reference, we next recomputed gene pairing between the species. This involved; a) identifying sequence orthologs that are transcriptionally diverged (**Figure S4C**); b) adding constant fold factor correction for orthologs that show species-specific constant bias (**Figure S4D**); c) discovering new putative ortholog associations for un-annotated (**Figure S4E**), or ambiguously aligned (**Figure S4F**) rabbit genes that are linked uniquely to mouse genes. Following functional orthology adjustments, we recomputed rabbit-mouse similarity over all manifold states. This procedure uncovered a highly specific positive correlation between related cell type pairs and anti-correlation for unrelated pairs (**Figure 3A**). We note that while overall alignment between the two manifolds is specific and inclusive, it can be of variable precision with significant gene divergence observed even for highly correlated states. Notably, the matching of PGC states between the species was lower than observed for other states (**Figure 3B**). Manifold state alignment can also represent broad similarities between related groups of states rather than a one-to-one state

correspondence. For example, the rabbit 'median TFAP' state is not directly matched by a corresponding mouse annotation, but still shows some similarity to the mouse epiblast, definitive ectoderm, primitive streak, and PGCs (**Figure 3A**).

A conserved gene regulatory core of transcription factors underlies gastrulation

Aligned manifold models at metacell resolution facilitate a fully quantitative comparison of the gastrulation transcriptional states. As exemplified in **Figure 3C**, the expression of gastrulation genes is, in most cases, within two-fold difference between the aligned species states. This analysis also allows the detection of genes that escape such conservation, such as SNAI1 for mouse allantois, NR2F1 for rabbit somitic mesoderm, and more (**Figure 3C** and **Table S3**). Next, we sought to systematically characterize the core gene regulatory conservation between the two gastrulation manifolds. We computed a cross-species gene by gene correlation matrix and searched for pairs of genes (not necessarily sequence orthologs) that were reciprocally best correlated. Of these, 147 gene pairs were found to be best reciprocal sequence orthologs, the majority of which (n=76) code for transcription factors (TFs) (**Figures 3D-E**. See complete list in **Table S2**). These TFs constitute a massive core of conserved regulatory logic that drives gastrulation. This logic is rapidly emerging over time, where 72% of the core TFs are significantly expressed (log2(UMI frequency)>-14) in at least one metacell by E_t7.5. In almost all cases, the identity of these early and concurrently induced TFs is still unmatched with specialized downstream programs (i.e. characterized by structural and functional genes). The deep conservation of the gastrulation TF core is particularly surprising given the high degree of regulatory redundancy in the system, with multiple TFs active and regulating each lineage[14]. This points toward a finely-tuned quantitative balance that goes beyond redundancy, suggesting such balance is crucial for coordinating the complex cellular multifurcation that installs the blueprint of cellular lineages during gastrulation.

Partial convergence of rabbit and mouse extraembryonic endoderm programs

The conservation of gastrulation transcriptional states is compatible with the known morphological convergence of early mammalian embryos termed the phylotypic stage, and the related hourglass model. In contrast to this conservation, the rabbit and mouse extraembryonic tissues highly diverge in their structure, with the rabbit displaying typical planar mammalian hypoblast and trophoblast, largely distinct from the mouse (**Figure 4A**). To understand how such morphological divergence may potentially impact signaling between extraembryonic and embryonic tissues and

hence their role in axial patterning[42], we created a rabbit ExEnd manifold (**Figure 4B**) and compared its states to the mouse counterpart (**Figures 4C-D**). This identified a program characterized by expression of Laminin (e.g. LAMA1, LAMB1), which characterized mouse parietal endoderm (PE), and in the rabbit included both PE, and hypoblast (Hyp) states (the latter showing higher expression of the pluripotent factor ESRRB). A second matching behavior was defined by strong expression of Apolipoproteins (APO-A/B/C), involving rabbit yolk sac endoderm (YSE) states[43], and in mouse embryo-associated visceral endoderm (VE) and extraembryonic VE. We also detected a subpopulation of rabbit hypoblast cells expressing the TFs EOMES, LHX1, HHEX, and the BMP antagonist CER1, together with relatively lower WNT3. This subpopulation therefore exhibits a transcriptional repertoire closely reminiscent of the mouse anterior visceral endoderm (AVE), and was termed anterior hypoblast[44,45]. As all rabbit epiblast cells are initially in direct contact with the hypoblast cell layer, specialized hypoblast programs were previously suggested to underlie the anterior-posterior polarization of the embryo[46]. Experimentally, by resection of the rabbit hypoblast, it was demonstrated to play an inhibitory role, confining the domain of the primitive streak to its stereotypic form[47]. Importantly, the convergence toward common functional programs in the endoderm of both species occurs even though some of the regulatory machinery is diverged, as is the case for EPAS1 and MYC, two key TFs that follow reversed patterns (**Figure 4D**).

Delayed and diverged rabbit trophoblast differentiation

In contrast to the overall transcriptional convergence of ExEnd states, the rabbit extraembryonic ectoderm (ExEc) states (**Figures 4E**) show little resemblance to those of the mouse (**Figures S5A-B**). The early bifurcating and diverging mouse chorion and ectoplacental cone (EPC) populations[48,49] could not be observed in rabbit ExEc. Instead, we found highly homogeneous states that gradually reduced DPPA3, RBP2 and increase IRX2, LDHA, among many other genes through E_t7.5 and until E_t8.0 (stages 4b-6) (**Figures 4F** - top, and **S5C-D**). Only following E_t8.0 (somite stages), coinciding with time of mesometrial implantation[11], diversification of the ExEc programs become evident, along a gradient spanned by diversifying expression of genes including GATA2, TFAP2A, MITF, and CCNE1 (**Figure 4F** - bottom) and inverse regulation of DNMT3B, LIFR, ALDOC, and CD34 (**Figure 4F** - middle). Comparison of this spectrum of late rabbit ExEc states to the mouse chorion and EPC states (**Figure 4G**) reveals important common features, including several TFs (TFAP2C, IRX2, ELF3, GATA3) and structural genes such as PLET1 and LMNA. Nevertheless, the differentiation gradient of late rabbit ExEc states did not associate strongly with either mouse chorion or derivatives of the ectoplacental cone lineage. This

may either suggest dramatically altered placental transcriptional states, or that rabbit differentiation toward functional placenta precursors initiates beyond the 8-12 somite stage (E_t8.6).

Differential expression of growth factors in early extraembryonic tissues

In the mouse, reciprocal interactions between the epiblast and ExEc that involve BMP4, NODAL, and WNT3, regulate the induction of the primitive streak (PS)[50-52]. Similarly, inhibitory signals produced by the mouse AVE, namely, the BMP and WNT inhibitors CER1 and DKK1, are critical for proper anterior-posterior axis formation[45]. Analysis of the rabbit hypoblast and ExEc states at the relevant time, prior to E_t6.5, showed conservation of BMP2 and WNT3 expression in the hypoblast. Notably, while CER1 is enriched in the anterior hypoblast, DKK1 appears to be expressed throughout the hypoblast (**Figure 4H**). BMP4 however, is not expressed in early ExEc cells, and only at low levels in the hypoblast. This, despite its robust expression in the embryo proper at later time stages, suggests the predominance of the more robustly expressed BMP2 for early patterning (**Figures S5A, F**). Finally, we find IGF1 to be highly expressed in the rabbit hypoblast, while in the mouse IGF1 expression is confined to the much later endothelial cells (also shared in the rabbit) (**Figure S5G**). In line with this finding is the demonstrated role of IGF1 in activation of embryoblast specific signal transduction[53], and maturation of ex utero cultured rabbit blastocysts[54]. Together, this analysis suggests a minimal signaling role for the rabbit ExEc in embryonic patterning (consistent with the deterioration of the Rauber's layer), and supports previous findings regarding the prominent role of hypoblast polarity. The detailed temporal expression of signaling genes (**Table S3**) complements previous efforts using traditional molecular techniques[46,54-56], and provides a basis for further understanding intercellular signaling developmental regulation in non-rodent gastrulating embryos.

Aligned rabbit and mouse gastrulation time axis highlights an hourglass-like bottleneck

Following alignment of the rabbit and mouse gastrulation manifolds, we next sought a principled strategy for comparing the two differentiation *processes*, as represented by our models. We therefore tested whether the independently determined rabbit and mouse gastrulation "clocks" could be aligned over a common time axis. To eliminate potential cell type annotation bias we generated a unified representation of rabbit/mouse cell state distributions (See **Methods**), and in this manner computed cross-species embryo cell-state frequency similarities over time (**Figure 5A**). The resulting similarity matrix revealed a stereotypical structure in which pre-gastrulation states are aligned but not synchronized, leading toward a bottleneck at approximately E_t7.5 in

both species, followed by a more synchronized gastrulation process with potential gradual loss of coherence. This analysis showed that overall we can use the absolute time axes of the two species for comparing their gastrulation processes. It also clearly demonstrates divergence and compensation in some key stages, particularly highlighting the earlier and more gradual emergence of PS populations in the mouse (**Figures 5B, C**). Remarkably, although the rabbit PS emerges later, it then converges to embryonic frequencies that closely match those observed in mouse. We next performed detailed comparison of the conserved and diverged gene expression signatures in the epiblast and PS of both species (**Figure S6A, B**). This highlighted the conservation of most of the previously characterized PS TFs and signaling genes, but also detected intriguing divergence of several developmental genes such as GATA4, FGF4, and POU3F1. This data further demonstrates conservation in the establishment of the core PS program, despite sizing, contextual and morphological differences between rabbit and mouse pre-gastrulation embryos.

Interestingly, initiation of gastrulation between the species is also characterized by clear differences in the induction of the anterior primitive streak (APS). In contrast to the gradual acquisition of APS through an intermediate PS state in the mouse, rabbit APS rapidly accumulates and appears to be induced directly from the epiblast (**Figure 5D**). **Figure 5E** shows that mouse PS states co-expressing T/Brachyuri (key TF of the PS) and FOXA2 (early endoderm TF) are not matched in the rabbit manifold. Instead, rabbit epiblast differentiation toward the PS quickly induces MSX1 (mesodermal TF) and CDX1 (caudal homeobox), distinguishing itself from high FOXA2 states linked with the APS. Together, these results indicate that rabbit endoderm may be established without transiently activating a meso-endo like program. This also serves as an example for conserved states that can be developmentally reached, even though the trajectories leading to them have changed during evolution.

Conserved rostral-caudal signatures in morphologically diverged gastrulas

Following the observation of species-specific timing of PS bifurcation, we next tested systematically whether this might further propagate to affect the later rostral-caudal program. To this end, we searched for conserved caudal and rostral gene regulatory modules (**Figure 5F**), and computed for each metacell a rostral/caudal score by calculating the expression ratio between the two. We then used the flow models to traceback the differentiation trajectories of metacells, grouped by type, and charted the rostral/caudal scores of each trajectory over time (**Figures 5G and S7A-H**). This showed remarkable conservation in differentiation toward most

fates, not restricted to fates defined specifically by their rostral/caudal transcription (e.g. caudal mesoderm and caudal neurectoderm), as well as in lineages creating significant caudal-rostral diversity (e.g. neural tube). Yet, rostral/caudal identities of some lineages are inconsistent, possibly reflecting morphological differences. For example, the data indicates strong separation between the large fraction of rabbit extraembryonic mesodermal amnion cells, that develop a strong transient caudal expression signature in E_t~7.7, and a minority showing a transient rostral signature, converging after E_t8 (**Figure S7A**). Similar separation is also seen in the surface ectoderm, though in this case the separation is maintained throughout. The strong expression of the axial gene signature indicating a more polarized development of the amnion in the rabbit vs the mouse, suggests that this module may be used by evolution at different stages and lineages, to compensate for the morphological divergence (size and shape) between the species.

Screening for evolutionarily diverged differentiation kinetics toward conserved fates

To further characterize the conservation or divergence of gene expression kinetics, we compared entire differentiation trajectories between mouse and rabbit. Specifically, we wished to examine if temporal kinetics of key cell-type-specific genes would be tightly conserved, shifted in phase, or compensated by other genes, and whether the kinetics of genes that are transiently active during state differentiation are also conserved. For example, when studying the differentiation toward somitic mesoderm, we observed high conservation between rabbit and mouse trajectories and expression kinetics (**Figures 6A** and **S7B)**. In-depth analysis of TF kinetics along this trajectory showed that while some key factors are tightly conserved between the species (**Figure 6B**, top), other TFs have altered timing (**Figure 6B,** second row). Regulation of the somitic mesoderm program is also correlated with conserved transient expression kinetics, though such behavior seems to span a more extended period in the mouse (**Figure 6B**, lower rows). Thus, this coordinated network of TFs (such as T, TBX6, and IRX5) not only leads toward the ultimate expression of a conserved somitic mesoderm program (including for example, retinoic acid signaling genes, collagens, etc.), but also to highly conserved timing of the entire somitic mesoderm specification process.

As a contrary example, the differentiation kinetics of the allantois greatly diverges, and rabbit embryos develop an allantois cell state later and more synchronously than the mouse (**Figures 6C** and **S7C**). Comparative kinetics analysis (**Figure 6D**) outlines the potential mechanistic basis for such divergence. Notably, some TFs that characterize the extraembryonic mesoderm program define a conserved basis for allantois differentiation and emerge at similar time points (**Figure**

6D, top), but other more specific allantois factors are delayed (**Figure 6D** second and third rows). Perhaps most significant is the allantois-specific transcription factor TBX4 which shows a very significant delay in rabbit compared to the mouse. The delayed commitment of the allantois (visible from E9, **Figure S7I**) is possibly related to the aforementioned overall delay in maturation of the rabbit placenta, perhaps relaxing the requirement for its early maturation. In summary, comparing differentiation processes using kinetics of orthologous genes can provide a powerful tool for identifying conserved regulatory links and may suggest how similar regulation is used within diverging contexts. It is however only a basis for even deeper exploration of the evolution of gastrulation logic which must also consider regulatory rewiring events, for example those involving compensation by paralogous genes as shown for SNAI1/2 (**Figure 6D,** bottom**)**, or repurposing factors to perform diverged tasks.

Early specification of rabbit primordial germ cells without transient mesoderm program activation

From all embryonic gastrulation states, primordial germ cells (PGCs) were associated with much lower conservation scores between the species (**Figures 3A, B**). The flow model indicates that rabbit PGCs appear early at E_t7.0 (stage3) (**Figure 1C**), and are specified either directly from the epiblast state or from a variant of the epiblast state distinctly marked by TFAP2A and TFAP2C expression (hence denoted “median TFAP”) (**Figure 2A**). In a previous study[7], numerous TFAP2C/POU5F1 positive cells were observed in the posterior end of gastrulating rabbit embryos. First, broadly in the posterior epiblast layer of stage2 embryos (E6.5-6.75), and later also in the mesoderm layer, becoming more spatially restricted by stage3-4 (E6.5-7.0). Further, in a recent human in vitro study, a TFAP2A/TFAP2C hierarchy was shown to regulate PGC specification from embryonic stem cells (ESCs)[57]. Intriguingly, we find rabbit PGCs rapidly downregulate TFAP2A while TFAP2C levels remain high (**Figures 7A and S8A**), yet a similar cell state does not exist in the mouse (**Figure S8A**). To further examine transcriptional dynamics underlying rabbit PGC specification, we screened for TFs that were specifically enriched in rabbit PGC metacells and compared their time-dependent expression between epiblast, median TFAP, PS, Nascent mesoderm, and PGC. The PGC states are associated with high expression levels of TFAP2C, SOX15, SOX17, CITED1, FOXA3, and NANOS3 (**Figure 7A**). Interestingly, SOX15 has been shown to be a primate and rabbit naïve pluripotency TF, though not in mouse[13,40]. Analysis of PS and nascent mesoderm TFs, clearly showed that PGCs from all embryonic stages do not upregulate most canonical TFs, such as T, EOMES, MIXL1, SNAI1 and HOXA1. This is in contrast to the transient but significant upregulation of the PS genes T and EOMES observed in vitro[7]. In

addition, DPPA3 (Stella) that is gradually silenced with time in the epiblast and all mesodermal states (**Figure S8B**) remains expressed throughout in PGCs – further supporting a direct differentiation trajectory. Comparing expression of PGC-specific genes between mouse and rabbit (using epiblast as a common baseline) showed conserved features, including the upregulation of PRDM1, TFAP2C, SOX17, and KIT (**Figure 7B**). Consistent with the previously described downregulation of SOX2 in non-rodent PGCs[7,58], we find rapid downregulation of this gene in the PGC trajectory starting at the median TFAP progenitor state (**Figure S8C**). The expression of Fragilis (IFITM3), ALPL, and additional mouse-specific genes were completely missing in rabbit PGC (some possibly due to technical lack of orthology). At the same time, a number of the strongest rabbit markers were either not expressed in mouse (CD34, RELN, FOXA3), or did not have a mouse orthologue, such as the rabbit orthologue for human germ-cell-specific gene PASD1[59]. The early specification of rabbit PGCs is also supported by cell cycle analysis (**Figure 7C).** While it is impossible to accurately deduce the actual reduction in PGC proliferation, the data suggest it to be mild, and stable up to E_t8.6. Together with flow analysis (**Figure S8D**), these results are most consistent with an early specified self-sustaining PGC population with declining frequency but slowly increasing absolute numbers, within a more rapidly growing embryo.

DISCUSSION

Gastrulation is the hallmark developmental process that translates “raw” intracellular pluripotent regulatory logic into the stereotypical blueprint of the vertebrate body plan. It requires single cells to acquire increasingly more specialized internal states in a precise and stable fashion. At the same time, and perhaps more importantly, gastrulation is defined by the capacity of cell ensembles to acquire such identity in a coordinated fashion. All the while, the embryo proliferates massively, and the initially unlimited potential of the epiblast is restricted and shaped by a combination of changes in cells’ internal state, cellular interactions, and physical forces - toward defined germ layers, cell types, and organ precursors. In this work, we used extensive sampling of the ensemble of single cells from single gastrulating rabbit embryos to combine quantitative characterization of transcriptional cell states with a model tracking the coordinated dynamics of the cell ensembles over time. Using comparative analysis of the new rabbit model and our recently published mouse gastrulation model[14], we could begin to approach fundamental questions on the universal and conserved properties of mammalian gastrulation and characterize diverged mechanisms ultimately leading toward the diversity across mammals in general, and between rabbit and mouse in particular.

An initial “static” comparative strategy of mouse and rabbit gastrulation manifolds shows remarkable conservation between the transcriptional states of both species. Such conservation is both specific and quantitative. Its specificity is indicated by the identification of 79 distinct RBOMs - transcriptional states that are reciprocally matched between the species. The discovery of a large number of distinct RBOMs reflects a truly rich conserved transcriptional space going far beyond a broad conservation of the germ layers states or specialized cell types like the primitive erythrocytes. Furthermore, conservation is highly quantitative, as shown by comparison of gene expression frequencies of thousands of genes in matched states, indicating almost all genes are conserved quantitatively up to a two-fold ratio. This precision of conservation is surprising since the gastrulation manifold does not consist of a large number of highly distinct states, but is instead a continuum of multifurcating programs[14,23]. Such plasticity could in theory support significant regulatory redundancy and subsequent evolutionary flexibility. But instead, at the core of the gastrulation programs, which we define by reciprocal best orthologous genes, we can identify 76 (and possibly more) transcription factors that retain a distinct and rigid regulatory identity between mouse and rabbit. Such conservation suggests that these factors coordinate the gastrulation process in a highly quantitative and balanced fashion, which limits the evolutionary plasticity of this process. Notably, 7 of these TFs are of the HOX family, supporting their suggested central role in establishing the phylotypic state[1]. Our model therefore calls for a quantitative and multi-factorial approach toward understanding the regulation of gastrulation state transitions in ways that were so far difficult to practice.

As noted above, understanding gastrulation at the transcriptional level entails more than mapping static cell states and types at single cell resolution, as the process inherently involves ensembles of interactive and proliferating cells. We therefore developed a framework for comparative analysis of the gastrulation process as it occurs over a carefully quantified time scale. The data coming out of this analysis unequivocally showed that gastrulation is conserved as a process, and highlighted a specific point in time, coinciding with the onset of massive differentiation (stages 4c-5; E_t7.5-7.7), where both mouse and rabbit gastrulation are aligned with maximum specificity (**Figure 5A**). This striking conservation, which provides fresh support and a possible mechanistic basis to the hourglass model for vertebrate evolution should be considered in light of the highly diverged extraembryonic tissues, overall difference in size, as well as dramatic morphological differences between the rabbit bilaminar disc and mouse cup-shaped structures[4]. Still, the structural divergence of the rabbit and mouse extraembryonic endoderm remains compatible with

conserved signaling. Namely, the canonical organizer function of the anterior hypoblast and the corresponding mouse AVE. One key difference however, is the apparent compensation by the rabbit hypoblast for the lack of functional signaling from the trophoblast. In general, rabbit trophoblast is markedly delayed in its development compared to the mouse extraembryonic ectoderm, in line with similar retardation of the allantois. The divergence of extraembryonic structures are contradictory to the strong conservation of the gastrulation process. This could suggest that the crucial conserved steps underlying embryonic diversification are more intrinsically driven, for example by the rich core of conserved TFs discussed above. Such a process is unlikely to rely on complex maternal or extraembryonic interactions that would strongly depend on conserved structure.

Our detailed rabbit gastrulation flow model and the new comparative methodology hereby provide a scaffold for further comprehensive and robust interpretation of other mammalian models, for which similar resolution and accuracy is yet to be attained. Human (and primate) gastrulation is already known to share many states and characteristics with the mouse. However, the temporal process of differentiation and proliferation for human cell ensembles during gastrulation would have to be modeled more carefully, especially considering the sparsity of such data. As flow models become integrated with new data and models for primates and other species, it can be expected that we will come closer toward truer understanding of human developmental process. This, will be useful for designing the control of embryonic stem cell differentiation, tissue models, and analysis of ex-utero and synthetic embryos. Lastly, such models will place us in an unprecedented position to synthesize the molecular phenotypes of gastrulation states between and within phyla[60] with an underlying genome and epigenome evolutionary theory.

FIGURE LEGENDS

Figure 1. **Generation of a temporal rabbit single cell transcription manifold**

(A) Morphological characterization of rabbit embryos analyzed in this study. Shown is a schematic representation of the main morphological features observed in each embryonic stage, alongside an image of a representative embryo. Gestation time reflects a single idealized time for each embryo time group. Images are uniformly scaled. Scale bar = 500um.
(B) Color coded 2-D UMAP projection of the metacells (MCs) graph representing the entire transcriptional manifold. MCs (nodes) are connected by edges, indicating the most similar neighbors for each MC.
(C) Cell type distribution per embryo, arranged according to intrinsic transcriptional rank and separated into 12 characteristic age groups.
(D) Morphological staging of individual embryos, indicated by a black square. Embryos that could not be assigned a stage (lacking proper image) are left empty. Color bar indicates the actual gestation time of each embryo, calculated from the time of mating. Some of the most distinctive morphological and cell state events are noted in the boxes below.

Figure 2. **Rabbit gastrulation network flow model**

(A) The gastrulation network flow model consists of MCs (nodes in rows) distributed in time (x-axis), and flows (edges) that link MCs between adjacent time points. The first time-point represents a common source for all MCs. Annotation (color coded) relies on marker expression and flow-based fate mapping.
(B) Heatmap showing relative expression (log2) of key lineage-specific genes. Extraembryonic endoderm and ectoderm lineages are presented separately, see Figure 4.

Figure 3. **Rabbit/Mouse cross-species manifold alignment**

(A) Comparison of independently annotated rabbit and mouse transcriptional states according to similarity of global gene expression profiles (here, based on 955 genes that are the intersection of each manifolds set of feature genes).
(B) Accumulated maximal correlation between rabbit MCs of each type, to any MC in the mouse manifold.
(C) Comparable gene expression in selected cell types. Each point indicates absolute expression levels in rabbit and mouse for each gene (log2 of UMI frequency, x- and y- axes, respectively, averaged by cell type).
(D) Gene-gene correlations between reciprocally best orthologous metacells (RBOMs). Shown are most highly correlated transcription factors between species.
(E) MC-MC correlations between RBOMs for a number of representative genes (following linear correction, see Methods).

Figure 4. **Similarities and differences in extraembryonic endoderm and ectoderm**

(A) Schematic of structural organization of late streak rabbit and mouse embryos. Midsagittal sections of mouse and rabbit are shown alongside an en-face view of the rabbit embryo. The planar form of the hypoblast and trophoblast observed in rabbits diverges from the cup-shaped mouse embryo.
(B, E) Color coded 2-D UMAP projection of the MC graph representing the extraembryonic endoderm and ectoderm manifolds, respectively. MCs (nodes) are connected by edges, indicating the most similar neighbors for each MC.
(C) MC-MC correlation matrix of rabbit and mouse extraembryonic endoderm.
(D) Side by side expression heat map (log2 of UMI frequency) of highly variable genes in rabbit and mouse extraembryonic endoderm.
(F) Temporal expression of key genes during ExEc development. Each dot represents a single MC.
(G) Heatmap showing relative expression of most highly variable genes within the ExEc (log2 fold-change relative to the embryonic average).
(H) Expression of key signaling factors during early gastrulation. Each dot represents a single MC. Only MCs with calculated time of E6-6.5 are shown.
AMC – anterior marginal crescent; PGE – posterior gastrula extension; PS – primitive streak; Epi – epiblast; Hyp – hypoblast; ExEc – extraembryonic ectoderm; VE – visceral endoderm; AVE – anterior visceral endoderm; ExVE – extraembryonic visceral endoderm; YSE – yolk sac endoderm; AHyp – anterior hypoblast; PE – parietal endoderm; Chor – chorion; Chor.p – chorion progenitors.

Figure 5. **Interspecies temporal alignment**

(A) Heatmap of interspecies correlation of cell type composition over time. For comparability, rabbit MC identities were derived de novo by the most similar mouse annotation according to gene expression (see Methods). Dynamic time warping (DTW) is plotted as a solid line along the diagonal (dashed), explaining differences as temporal compression/stretching. Plotted below, the DTW derivative, highlighting the temporal segments of major changes.
(B) Temporal expression of key genes during epiblast to PS transition. Each dot represents a single MC.
(C) Temporal alignment of PS frequency in mouse and rabbit. Lines represent Loess smoothing, standard error bounds indicated by shading.
(D) Vein plots describing the continuous transition of cell types from the epiblast (represented by diagonal flows spanning time points), and the dynamic relative frequencies of these cell types in the embryo over time (vein width on the y- axis). Vein connection width represents flow flux at each time point.

Figure 6. **Conservation and divergence in expression kinetics over differentiation tracebacks**

(A, B) Tracebacks were performed from (A) somitic-mesoderm and (C) allantois, in flow models of both Rabbit (left) and Mouse (right, mirrored). The latest rabbit time bin, and the first mouse time bin, for which there are no corresponding data from the other species, were trimmed.

Network plots depict traced-back flows, colored by cell types (top). Heatmaps show average expression kinetics for genes with conserved gene expression profiles (see Table S3), displaying the highest variance over the trajectories. Below, colored polygons represent the cell type composition of the traceback at each time point for the full flow maps.
(B, D) Absolute expression level (log2 of UMI frequency, y-axis) for select marker genes along each of the trajectories shown in the respective network plot. Dots represent individual tracebacks for each terminal MC, lines indicate respective medians.

Figure 7. **PGC differentiation**

(A) Absolute expression level (log2 of UMI frequency, y axis) for select marker genes.
(B) Relative expression in PGC vs epiblast (log2 of UMI frequency in PGC - log2 of UMI frequency in epiblast) in rabbit (Dark grey) and mouse (light grey). Mouse genes with no known rabbit ortholog are annotated with "*" (top) and rabbit genes lacking mouse orthology with "**" (bottom).
(C) Cell cycle score. A cycling score combining the expression of S- and M- phase characteristic genes (see Methods) was assigned to each single cell (y- axis). Due to the low number of PGCs in the manifold, cells were binned to four groups (6-6.6, 7-7.4, 7.5-7.9, 8.1-8.6). Partial cycling slowdown in the PGC population is observed, when compared to the notochord (substantially halting) or other early populations (fully cycling).

Figure S1. **Rabbit embryo compendium**

Pictures of embryos sampled in the study. Embryos shown in sequence according to intrinsic k-NN transcriptional rank. Color bars indicates separation into 12 age groups. Scale bar = 500um.

Figure S2. **Embryo collection and scRNA-seq data preparation**

(A) Distribution of UMIs per single cell, exclusion cells criteria denoted by red and green vertical lines.
(B) Distribution of fraction of excluded gene UMIs per single cell (see Metacells vignette). Exclusion cells criteria denoted by red vertical line.
(C) Number of embryonic cells per embryo (after QC, ranked from “youngest” to “oldest”).
(D) Number of embryos participating in the flow analysis, collected over gestation time.
(E) Embryo stages by morphological assessment vs gestation time.
(F) Calculated time (Et) vs gestation time for each time bin.
(G) Genome annotation enhancement pipeline.
(H) Temporal ordering and binning of embryos. Left, k-NN embryo/embryo similarity (horizontal lines mark separation of time bins, numbered in bold). Right, cell-type frequency per embryo (see Figure 1).

Figure S3. **Flows support**

(A) Embryonic cells per time bin, plotted over the 2-D UMAP projection, color coded by cell type.
(B) MC/MC distance metric used as constraint for the flow problem formulation, indicating transcriptional logistic distance.
(C) Cell cycle S-phase score vs M-phase score based on UMI count fractions. Cutoff line (bottom left diagonal) represents cells with low cycling score.
(D) Distribution of MCs cycling score, by type. MC cycling score indicates fraction of cells below cutoff (as in C).

Figure S4. **MC alignment support**

(A) Correlation across all MCs of the rabbit and mouse manifolds, ordered by cell type and internally by time.
(B) Correlation across the 79 RBOMs of the rabbit and mouse manifolds (ordered as in (A)).
(C) Example of transcriptionally diverged sequence orthologs of APLN.
(D) Example of sequence orthologs showing species-specific constant bias. Shown before and after linear correction (factor of 0.41) according to the manifold alignment.
(E) Example of new putative ortholog association for un-annotated rabbit MESP1 gene, matched with the transcriptional hostpot33433.
(F) Example of annotation assignment for the rabbit TWIST1 gene based on alignment to the mouse genome (mm9). Table lists explicit OryCun3.0 genome coordinates for genes mentioned in (C-F).

Figure S5. **Extraembryonic ectoderm (ExEc) interspecies comparison**

(A) Side by side gene expression heatmap in mouse and rabbit ExEc.
(B) MC-MC correlation matrix of rabbit and mouse ExEc.
(C) Relative expression of notable ExEc and ExEnd genes (rabbit only).
(D) Temporal distributions of rabbit ExEc.
(E) Temporal expression (log2 of UMI frequency) of key genes during ExEc development. Each dot represents a single MC.
(F, G) Temporal expression of rabbit BMP4 and mouse IGF1, respectively. Each dot represents a single MC.

Figure S6. **Conserved and diverged gene expression signatures in the epiblast and PS**

(A) Comparison of PS - Epiblast differential gene expression (log2) between rabbit (x- axis) and mouse (y- axis). Canonical TFs such as MIXL1 and T are upregulated in the epiblast to PS transition in both species (top right corner).
(B) Heatmap of genes with extreme profiles in their PS-Epiblast differential expression, grouped by the quadrants of (A), each column represents a single MC.

Figure S7. **Gene expression trends over trajectories**

(A-H) Rostral-Caudal score of individual MC trajectories of terminal types, as indicated.
(I) Representative E9.0 embryo (16 somite pairs, not sequenced).

Figure S8. **PGC specification**

(A) Expression over time of TFAP2A for mouse and rabbit.
(B) Expression of DPPA3 (Stella) over time. While gradually silenced with time in the epiblast and mesodermal states, DPPA3 remains expressed in PGCs.
(C) Expression over time of SOX2 for mouse and rabbit.
(D) Internal PGC flow over time. After an initial differentiation of non-PGC cell types to PGC, beyond Et7, the number of PGC cells can be explained solely by internal proliferation.

METHODS

Embryo Collection and Documentation

Embryos were collected from pregnant New Zealand White female rabbits (mated by Envigo, and shipped on the morning of each experiment), between E6.0-8.5. All animal experiments were approved by the Institutional Animal Care and Use Committee (07100820-2). Animals were sacrificed by intravenous injection of lethal dose of pentobarbital (100mg/kg), and embryos removed as described by[61]. In brief, uteri were placed in ice-cold PBS, and embryos removed individually from dissected uterus segments containing a single site using fine iridectomy scissors on darkened (Activated charcoal C9157, Sigma-Aldrich) Sylgard184 (Dow) coated 10cm plates. Embryos were resected from most of the extraembryonic tissues, washed in fresh PBS, and transferred to chilled DMEM (Phenol-red free, Gibco) supplemented with 10% FBS (Biological Industries) for imaging prior to dissociation. Phase contrast images were taken with an Eclipse Ti2 inverted microscope (Nikon) and Zyla sCMOS camera (Andor). Embryo staging was performed according to[36,62].

Flow Cytometry

Prior to dissociation, to focus on tissues that contribute to the embryo proper, most of the extraembryonic ectoderm and yolk sac were removed using fine forceps and fine iridectomy scissors. For isolation of single cells for scRNA-seq, embryos were dissociated with 0.25% Trypsin-A, 0.02% EDTA (Biological Industries) solution for 5' at 37°C, and resuspended in DMEM w/o phenol red (Gibco) supplemented with 10% FBS (Biological Industries). Samples were sorted to 384well lysis plates using a FACSAria-III or Symphony S6 flow cytometer (BD Biosciences), excluding doublets and cell debris by FSC and SSC metrics.

Single-cell RNA-sequencing and Metacell analysis

Single cell cDNA libraries were prepared using the MARS-seq method, as described[35,63], with the following modifications: The final concentration of the RT1 primers was 2nM, and pooling was done via centrifugation to VBLOCK200 reservoir (Clickbio). Klenow reaction was not followed by heat inactivation. The volume of the first RT and Exonuclease I reactions mix were scaled down to 1 and 0.5ul, respectively, and dispensed by MANTIS liquid handler (FORMULATRIX). MARS-seq reads were processed following the MARS-seq2.0 pipeline[63]. Briefly, demultiplexing plate barcodes (4 base pairs) from Read 1, the reads were mapped to the OryCun3.0 genome using STAR. Demultiplexing of the reads and construction of the single-cell UMI matrix was based on well barcodes (7 base pairs) and UMI barcodes (8 base pairs) from Read 2. Embryo identities were recorded per well during sorting and recovered as part of the MARS-seq post-processing stage. Overall we processed 165,504 wells, out of which we retained 117,660 well-covered cells for further analysis. In particular, cells with less than 2,600 or over 50,000 UMIs were filtered out (median 7,142 UMIs). Following the Metacells vignette[37], mitochondrial, stress, cycle-related, and batchy transcripts, were excluded. Feature genes for MC analysis were identified and scrutinized using manual evaluation of gene modules. The derived reference model included 3,584 MCs on 117,660 cells.

Reference genome and annotation

At the time this work took place, the state-of-the-art Rabbit reference genome was OryCun2.0 [GCA_000003625.1], covered at 7.5X and annotated to some extent. We've extended the former annotation with additional RefSeq gene orthologies. Recently OryCun3.0 was sequenced and assembled [GCA_013371645.1], we've converted the extended OryCun2.0 using LiftOver mapping to OryCun3.0 and appended additional gene annotations from direct OryCun3 LiftOver mapping of mm9 and hg19 reference genomes. For completeness, we further extended the annotations with mapping hotspots, such that any expressed transcript is accounted for. See illustrated annotation pipeline in Figure S2G.

Embryo k-NN ordering

Following an initial coarse-grained embryo ordering, we applied the k-NN order refinement heuristic described by[14]. Briefly, a cell-cell transcriptional similarity k-NN graph is composed of all the embryos' cells, excluding extraembryonic tissues. Embryo similarity is defined by the number of edges crossing between the embryos' cells, normalized for the number of cells per embryo. A greedy pair-wise order switching heuristic minimizes the differences between adjacent embryos, improving upon the initial order until convergence (Figure S2H).

Defining absolute developmental time

Following k-NN guided binning of embryos, each bin was assigned a time stamp based on the absolute time of embryos making it up (considering mating as time 0, and assigning all littermates the same time. See Table S1). For the latest somite stages, an idealized time was given based on a 2hr cycle for each somite production and multiplied by the median difference in somites between consecutive bins.

Cell cycle modules and metacell proliferation rate

For each cell, we computed an M- and S- phase score as the fraction of UMIs for these genes, out of the total number of UMIs per cell (Figure S3C). The group of M-phase genes was guided by MKI67, TOP2A, UBE2C and expanded to ARHGAP11A, ARHGAP27P1-BPTFP1-KPNA2P3, ARL6IP1, ASPM, AURKA, AURKB, BUB3, CCNB1, CCNB2, CCNF, CDCA2, CDCA3, CENPE, CENPF, CKAP2, CKAP2L, DEPDC1B, G2E3, GTSE1, HYLS1, INCENP, KIF11, KIF14, KIF18A, KIF20A, KIF23, KIFC1, MIS18BP1, MKI67, NCAPG, NUSAP1, PLK1, PRC1, PRR11, PSRC1, RACGAP1, SGO2, SPAG5, TOP2A, TPX2, TTK, UBE2C, hotspot44564. The group of S-phase genes was guided by PCNA, RRM2 and expanded to include DTL, FEN1, MCM2, MCM4, MCM6, MCM7, PCNA, RFC3, RRM2, hotspot60823. Cells presenting reduced S- and M- phase scores were considered non-cycling, reflecting on the MCs overall cycling score. We detected the highest cycling reduction in the extraembryonic tissues as well as the notochord, corroborated by the exponential decay of the fraction of cells per embryo as a function of inferred embryo time.

Network flow inference and modeling

Temporally dense whole-embryo sampling of a highly reproducible process allows the inference of a network flow model, first presented by[14]. In short, a "mass conservation" constraint is formulated around cell numbers and applied on each MC's representation across time groups (Figure S3A). Differential MC's proliferation rate (as calculated above in "Cell cycle modules metacell proliferation rate"), is incorporated into the mass conservation constraints in the form of

dilution and renormalization. These constraints, together with distances (Figure S3B) and a cost function that accounts for deviations from the allowed capacities, compose a classical network min-cost max-flow problem with convex costs that can be solved in polynomial time using a classic simplex solver.

Manifold alignment

Given two gastrulation manifolds of highly diverged species with non-alignable genomes and poorly comparable transcriptional annotation, we suggest a manifold alignment algorithm to facilitate a high-resolution comparative analysis. Initially, we naively assume complete functional conservation of named-orthologous gene-pairs with minimal technical filtering. Based on this initial gene matching, we identify reciprocally best orthologous metacells (RBOMs, Figure S4B) pairs. That is, for each metacell in one manifold, we pick the best correlating metacell in the other manifold only if the correlation is also maximal for that second manifold metacell across all the metacells of the first. Using the RBOMs as a new common anchor, the initial gene pairing is re-evaluated and adjusted with limited linear correction (Figure S4D), following which, the rabbit-mouse similarity is recomputed over all manifold states.

Projected annotation

Given two manifolds, suggested annotation can be projected from one annotated manifold onto the other by “voting” among a small set of top correlating metacells. Following the previous manifold-alignment gene-orthology assertion and correction, MC/MC correlations are calculated between the manifolds. Each MC in the projected atlas “votes” its annotations such that any MC in the other atlas is annotated by the majority of its top correlating annotated MCs.

Temporal alignment

Dynamic time warping (DTW) is a common technique to align two given (time-dependent) sequences, applied here on the sequence of cell type compositions as the embryos go through gastrulation. Given annotated cell type composition of our characteristic embryo timepoints, intermediate compositions are linearly interpolated to produce a continuous series. Comparable cell type frequencies follow the mouse annotation (via “Projected annotation“, see above).

In Figure 5A, an overlapping gastrulation timeframe E6-8.3 is selected for alignment, showing a later differentiation onset in the Rabbit followed by a delayed progression which then goes back into remarkable sync at E7.5 and later diverges slightly.

ACKNOWLEDGMENTS

We thank the Tanay and Stelzer group members for discussion and advice. Special thanks to Hernan Rubinstein for help in graphic design and embryo illustrations. Y.S. is the incumbent of the Louis and Ida Rich Career Development Chair and a member of the European Molecular Biology Organization (EMBO) Young Investigator Program. Research in the Stelzer lab is supported by European Research Council (ERC_StG 852865), Moross Integrated Cancer Center, Helen and Martin Kimmel Stem Cell Institute, the Schwartz/Reisman Collaborative Science Program, and Abisch Frenkel Foundation. This research was also generously supported by Barry and Janet Lang, Hadar Impact Fund, Lord Sieff of Brimpton Memorial Fund, Janet and Steven Anixter, JoAnne Silva, Maurice and Vivienne Wohl Biology Endowment, and Lester and Edward Anixter Family. M.M. was a postdoctoral fellow of the Minerva Stiftung and is supported by the Walter Benjamin program of the German Research Foundation (DFG). A.T. is supported by the European Research Council (ERC CoG scAssembly), the EU BRAINTIME project, the Israel Science Foundation, and the Chen-Zuckerberg Foundation. This research was further supported Israeli Council for Higher Education (CHE) Data Science program, and by a grant from Madame Olga Klein-Astracha.

AUTHOR CONTRIBUTIONS

Y.M., O.R., A.T., and Y.S. conceived and designed the experiment, and performed data analysis and its interpretation. S.C., R.B.Y, and R.H. assisted in the collection of single-cells from individual embryos. N.R. prepared scRNA-seq libraries. A.L., M.M., and O.B.K. assisted with scRNA-seq data analyses. Y.M., O.R., A.T., and Y.S. wrote the manuscript with input from all the authors.

REFERENCES

1. Duboule, D. (1994). Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. Dev Suppl, 135-142.
2. Raff, R.A. (1996). The Shape of Life: Genes, Development, and the Evolution of Animal Form (University of Chicago press).
3. Domazet-Loso, T., and Tautz, D. (2010). A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns. Nature *468*, 815-818. 10.1038/nature09632.
4. Eakin, G.S., and Behringer, R.R. (2004). Diversity of germ layer and axis formation among mammals. Semin Cell Dev Biol *15*, 619-629. 10.1016/j.semcdb.2004.04.008.
5. King, B.F., and Enders, A.C. (1993). Comparative Development of the Mammalian Yolk Sac. In The Human Yolk Sac and Yolk Sac Tumors, F.F. Nogales, ed. (Springer Berlin, Heidelberg), pp. 1-32. https://doi.org/10.1007/978-3-642-77852-0.
6. Irie, N., Tang, W.W., and Azim Surani, M. (2014). Germ cell specification and pluripotency in mammals: a perspective from early embryogenesis. Reprod Med Biol *13*, 203-215. 10.1007/s12522-014-0184-2.
7. Kobayashi, T., Castillo-Venzor, A., Penfold, C.A., Morgan, M., Mizuno, N., Tang, W.W.C., Osada, Y., Hirao, M., Yoshida, F., Sato, H., et al. (2021). Tracing the emergence of primordial germ cells from bilaminar disc rabbit embryos and pluripotent stem cells. Cell Rep *37*, 109812. 10.1016/j.celrep.2021.109812.
8. Denker, H.W. (2000). Structural dynamics and function of early embryonic coats. Cells Tissues Organs *166*, 180-207. 10.1159/000016732.
9. Fischer, B., Chavatte-Palmer, P., Viebahn, C., Navarrete Santos, A., and Duranthon, V. (2012). Rabbit as a reproductive model for human health. Reproduction *144*, 1-10. 10.1530/REP-12-0091.
10. Viebahn, C., Mayer, B., and Hrabe de Angelis, M. (1995). Signs of the principle body axes prior to primitive streak formation in the rabbit embryo. Anat Embryol (Berl) *192*, 159-169. 10.1007/BF00186004.
11. Nishimura, M. (2001). Timing of implantation in New Zealand White rabbits. Congenital Anomalies, 198-203.
12. Argelaguet, R., Clark, S.J., Mohammed, H., Stapel, L.C., Krueger, C., Kapourani, C.A., Imaz-Rosshandler, I., Lohoff, T., Xiang, Y., Hanna, C.W., et al. (2019). Multi-omics profiling of mouse gastrulation at single-cell resolution. Nature *576*, 487-491. 10.1038/s41586-019-1825-8.
13. Bouchereau, W., Jouneau, L., Archilla, C., Aksoy, I., Moulin, A., Daniel, N., Peynot, N., Calderari, S., Joly, T., Godet, M., et al. (2022). Major transcriptomic, epigenetic and metabolic changes underlie the pluripotency continuum in rabbit preimplantation embryos. Development *149*. 10.1242/dev.200538.
14. Mittnenzweig, M., Mayshar, Y., Cheng, S., Ben-Yair, R., Hadas, R., Rais, Y., Chomsky, E., Reines, N., Uzonyi, A., Lumerman, L., et al. (2021). A single-embryo, single-cell time-resolved model for mouse gastrulation. Cell *184*, 2825-2842 e2822. 10.1016/j.cell.2021.04.004.

15. Niu, Y., Sun, N., Li, C., Lei, Y., Huang, Z., Wu, J., Si, C., Dai, X., Liu, C., Wei, J., et al. (2019). Dissecting primate early post-implantation development using long-term in vitro embryo culture. Science *366*. 10.1126/science.aaw5754.
16. Nowotschin, S., Setty, M., Kuo, Y.Y., Liu, V., Garg, V., Sharma, R., Simon, C.S., Saiz, N., Gardner, R., Boutet, S.C., et al. (2019). The emergent landscape of the mouse gut endoderm at single-cell resolution. Nature *569*, 361-367. 10.1038/s41586-019-1127-1.
17. Pijuan-Sala, B., Griffiths, J.A., Guibentif, C., Hiscock, T.W., Jawaid, W., Calero-Nieto, F.J., Mulas, C., Ibarra-Soria, X., Tyser, R.C.V., Ho, D.L.L., et al. (2019). A single-cell molecular map of mouse gastrulation and early organogenesis. Nature *566*, 490-495. 10.1038/s41586-019-0933-9.
18. Rayon, T., Maizels, R.J., Barrington, C., and Briscoe, J. (2021). Single-cell transcriptome profiling of the human developing spinal cord reveals a conserved genetic programme with human-specific features. Development *148*. 10.1242/dev.199711.
19. Ton, M.-L.N., Keitley, D., Theeuwes, B., Guibentif, C., Ahnfelt-Rønne, J., Andreassen, T.K., Calero-Nieto, F.J., Imaz-Rosshandler, I., Pijuan-Sala, B., Nichols, J., et al. (2022). Rabbit Development as a Model for Single Cell Comparative Genomics. bioRxiv.
20. Tyser, R.C.V., Mahammadov, E., Nakanoh, S., Vallier, L., Scialdone, A., and Srinivas, S. (2021). Single-cell transcriptomic characterization of a gastrulating human embryo. Nature *600*, 285-289. 10.1038/s41586-021-04158-y.
21. Wagner, D.E., Weinreb, C., Collins, Z.M., Briggs, J.A., Megason, S.G., and Klein, A.M. (2018). Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. Science *360*, 981-987. 10.1126/science.aar4362.
22. Qiu, C., Cao, J., Martin, B.K., Li, T., Welsh, I.C., Srivatsan, S., Huang, X., Calderon, D., Noble, W.S., Disteche, C.M., et al. (2022). Systematic reconstruction of cellular trajectories across mouse embryogenesis. Nat Genet *54*, 328-341. 10.1038/s41588-022-01018-x.
23. Moris, N., Pina, C., and Arias, A.M. (2016). Transition states and cell fate decisions in epigenetic landscapes. Nat Rev Genet *17*, 693-703. 10.1038/nrg.2016.98.
24. Tritschler, S., Buttner, M., Fischer, D.S., Lange, M., Bergen, V., Lickert, H., and Theis, F.J. (2019). Concepts and limitations for learning developmental trajectories from single cell genomics. Development *146*. 10.1242/dev.170506.
25. Cheng, S., Mittnenzweig, M., Mayshar, Y., Lifshitz, A., Dunjic, M., Rais, Y., Ben-Yair, R., Gehrs, S., Chomsky, E., Mukamel, Z., et al. (2022). The intrinsic and extrinsic effects of TET proteins during gastrulation. Cell *185*, 3169-3185 e3120. 10.1016/j.cell.2022.06.049.
26. Sankaran, V.G., Weissman, J.S., and Zon, L.I. (2022). Cellular barcoding to decipher clonal dynamics in disease. Science *378*, eabm5874. 10.1126/science.abm5874.
27. Chen, A., Liao, S., Cheng, M., Ma, K., Wu, L., Lai, Y., Qiu, X., Yang, J., Xu, J., Hao, S., et al. (2022). Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. Cell *185*, 1777-1792 e1721. 10.1016/j.cell.2022.04.003.
28. Liu, C., Li, R., Li, Y., Lin, X., Zhao, K., Liu, Q., Wang, S., Yang, X., Shi, X., Ma, Y., et al. (2022). Spatiotemporal mapping of gene expression landscapes and developmental trajectories during zebrafish embryogenesis. Dev Cell *57*, 1284-1298 e1285. 10.1016/j.devcel.2022.04.009.
29. 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. 10.1038/s41586-019-1469-8.
30. Nakamura, T., Okamoto, I., Sasaki, K., Yabuta, Y., Iwatani, C., Tsuchiya, H., Seita, Y., Nakamura, S., Yamamoto, T., and Saitou, M. (2016). A developmental coordinate of pluripotency among mice, monkeys and humans. Nature *537*, 57-62. 10.1038/nature19096.

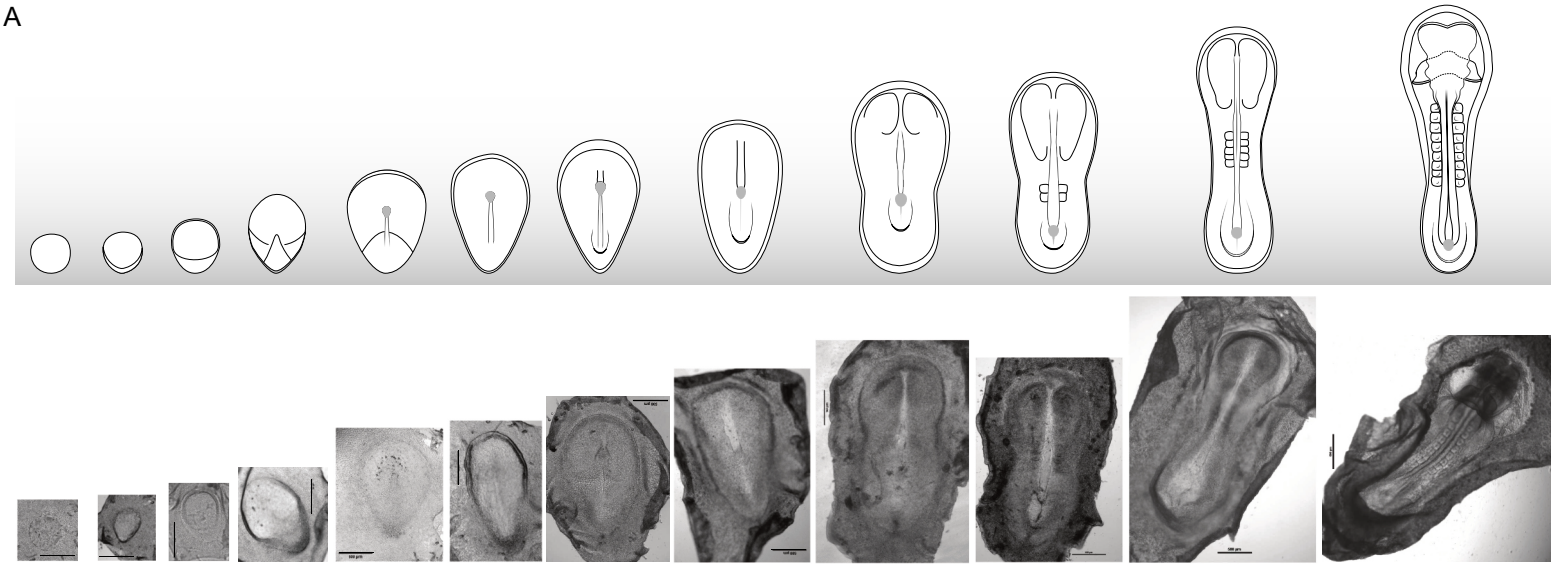
31. Nakamura, T., Yabuta, Y., Okamoto, I., Sasaki, K., Iwatani, C., Tsuchiya, H., and Saitou, M. (2017). Single-cell transcriptome of early embryos and cultured embryonic stem cells of cynomolgus monkeys. Sci Data *4*, 170067. 10.1038/sdata.2017.67.
32. Xu, Y., Zhang, T., Zhou, Q., Hu, M., Qi, Y., Xue, Y., Wang, L., Nie, Y., Bao, Z., and Shi, W. (2022). A single-cell transcriptome atlas of human early embryogenesis. bioRxiv.
33. Ihara, T. (1970). Comparative study of developmental progress in the mouse, rat and rabbit in their stages of organogenesis. Congenital Anomalies *10 (2)*, 67-81.
34. Williams, B.S., and Biggers, J.D. (1990). Polar trophoblast (Rauber's layer) of the rabbit blastocyst. Anat Rec *227*, 211-222. 10.1002/ar.1092270210.
35. Jaitin, D.A., Kenigsberg, E., Keren-Shaul, H., Elefant, N., Paul, F., Zaretsky, I., Mildner, A., Cohen, N., Jung, S., Tanay, A., and Amit, I. (2014). Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. Science *343*, 776-779. 10.1126/science.1247651.
36. Blum, M., Andre, P., Muders, K., Schweickert, A., Fischer, A., Bitzer, E., Bogusch, S., Beyer, T., van Straaten, H.W., and Viebahn, C. (2007). Ciliation and gene expression distinguish between node and posterior notochord in the mammalian embryo. Differentiation *75*, 133-146. 10.1111/j.1432-0436.2006.00124.x.
37. Ben-Kiki, O., Bercovich, A., Lifshitz, A., and Tanay, A. (2022). Metacell-2: a divide-and-conquer metacell algorithm for scalable scRNA-seq analysis. Genome Biol *23*, 100. 10.1186/s13059-022-02667-1.
38. Sebe-Pedros, A., Saudemont, B., Chomsky, E., Plessier, F., Mailhe, M.P., Renno, J., Loe-Mie, Y., Lifshitz, A., Mukamel, Z., Schmutz, S., et al. (2018). Cnidarian Cell Type Diversity and Regulation Revealed by Whole-Organism Single-Cell RNA-Seq. Cell *173*, 1520-1534 e1520. 10.1016/j.cell.2018.05.019.
39. Tanay, A., and Sebe-Pedros, A. (2021). Evolutionary cell type mapping with single-cell genomics. Trends Genet *37*, 919-932. 10.1016/j.tig.2021.04.008.
40. Boroviak, T., Stirparo, G.G., Dietmann, S., Hernando-Herraez, I., Mohammed, H., Reik, W., Smith, A., Sasaki, E., Nichols, J., and Bertone, P. (2018). Single cell transcriptome analysis of human, marmoset and mouse embryos reveals common and divergent features of preimplantation development. Development *145*. 10.1242/dev.167833.
41. Hodge, R.D., Bakken, T.E., Miller, J.A., Smith, K.A., Barkan, E.R., Graybuck, L.T., Close, J.L., Long, B., Johansen, N., Penn, O., et al. (2019). Conserved cell types with divergent features in human versus mouse cortex. Nature *573*, 61-68. 10.1038/s41586-019-1506-7.
42. 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. 10.1038/nrm2618.
43. Cindrova-Davies, T., Jauniaux, E., Elliot, M.G., Gong, S., Burton, G.J., and Charnock-Jones, D.S. (2017). RNA-seq reveals conservation of function among the yolk sacs of human, mouse, and chicken. Proc Natl Acad Sci U S A *114*, E4753-E4761. 10.1073/pnas.1702560114.
44. Hoshino, H., Shioi, G., and Aizawa, S. (2015). AVE protein expression and visceral endoderm cell behavior during anterior-posterior axis formation in mouse embryos: Asymmetry in OTX2 and DKK1 expression. Dev Biol *402*, 175-191. 10.1016/j.ydbio.2015.03.023.
45. Robb, L., and Tam, P.P. (2004). Gastrula organiser and embryonic patterning in the mouse. Semin Cell Dev Biol *15*, 543-554. 10.1016/j.semcdb.2004.04.005.
46. Yoshida, M., Kajikawa, E., Kurokawa, D., Tokunaga, T., Onishi, A., Yonemura, S., Kobayashi, K., Kiyonari, H., and Aizawa, S. (2016). Conserved and divergent expression patterns of markers of axial development in eutherian mammals. Dev Dyn *245*, 67-86. 10.1002/dvdy.24352.

47. Idkowiak, J., Weisheit, G., Plitzner, J., and Viebahn, C. (2004). Hypoblast controls mesoderm generation and axial patterning in the gastrulating rabbit embryo. Dev Genes Evol *214*, 591-605. 10.1007/s00427-004-0436-y.
48. Latos, P.A., and Hemberger, M. (2016). From the stem of the placental tree: trophoblast stem cells and their progeny. Development *143*, 3650-3660. 10.1242/dev.133462.
49. Lau, K.Y.C., Rubinstein, H., Gantner, C.W., Hadas, R., Amadei, G., Stelzer, Y., and Zernicka-Goetz, M. (2022). Mouse embryo model derived exclusively from embryonic stem cells undergoes neurulation and heart development. Cell Stem Cell *29*, 1445-1458 e1448. 10.1016/j.stem.2022.08.013.
50. Ben-Haim, N., Lu, C., Guzman-Ayala, M., Pescatore, L., Mesnard, D., Bischofberger, M., Naef, F., Robertson, E.J., and Constam, D.B. (2006). The nodal precursor acting via activin receptors induces mesoderm by maintaining a source of its convertases and BMP4. Dev Cell *11*, 313-323. 10.1016/j.devcel.2006.07.005.
51. Donnison, M., Beaton, A., Davey, H.W., Broadhurst, R., L'Huillier, P., and Pfeffer, P.L. (2005). Loss of the extraembryonic ectoderm in Elf5 mutants leads to defects in embryonic patterning. Development *132*, 2299-2308. 10.1242/dev.01819.
52. Georgiades, P., and Rossant, J. (2006). Ets2 is necessary in trophoblast for normal embryonic anteroposterior axis development. Development *133*, 1059-1068. 10.1242/dev.02277.
53. Navarrete Santos, A., Ramin, N., Tonack, S., and Fischer, B. (2008). Cell lineage-specific signaling of insulin and insulin-like growth factor I in rabbit blastocysts. Endocrinology *149*, 515-524. 10.1210/en.2007-0821.
54. Thieme, R., Ramin, N., Fischer, S., Puschel, B., Fischer, B., and Santos, A.N. (2012). Gastrulation in rabbit blastocysts depends on insulin and insulin-like-growth-factor 1. Mol Cell Endocrinol *348*, 112-119. 10.1016/j.mce.2011.07.044.
55. Hopf, C., Viebahn, C., and Puschel, B. (2011). BMP signals and the transcriptional repressor BLIMP1 during germline segregation in the mammalian embryo. Dev Genes Evol *221*, 209-223. 10.1007/s00427-011-0373-5.
56. Schroder, S.S., Tsikolia, N., Weizbauer, A., Hue, I., and Viebahn, C. (2016). Paraxial Nodal Expression Reveals a Novel Conserved Structure of the Left-Right Organizer in Four Mammalian Species. Cells Tissues Organs *201*, 77-87. 10.1159/000440951.
57. Chen, D., Sun, N., Hou, L., Kim, R., Faith, J., Aslanyan, M., Tao, Y., Zheng, Y., Fu, J., Liu, W., et al. (2019). Human Primordial Germ Cells Are Specified from Lineage-Primed Progenitors. Cell Rep *29*, 4568-4582 e4565. 10.1016/j.celrep.2019.11.083.
58. Kobayashi, T., Zhang, H., Tang, W.W.C., Irie, N., Withey, S., Klisch, D., Sybirna, A., Dietmann, S., Contreras, D.A., Webb, R., et al. (2017). Principles of early human development and germ cell program from conserved model systems. Nature *546*, 416-420. 10.1038/nature22812.
59. Michael, A.K., Harvey, S.L., Sammons, P.J., Anderson, A.P., Kopalle, H.M., Banham, A.H., and Partch, C.L. (2015). Cancer/Testis Antigen PASD1 Silences the Circadian Clock. Mol Cell *58*, 743-754. 10.1016/j.molcel.2015.03.031.
60. Levin, M., Anavy, L., Cole, A.G., Winter, E., Mostov, N., Khair, S., Senderovich, N., Kovalev, E., Silver, D.H., Feder, M., et al. (2016). The mid-developmental transition and the evolution of animal body plans. Nature *531*, 637-641. 10.1038/nature16994.
61. Puschel, B., and Viebahn, C. (2010). Rabbit mating and embryo isolation. Cold Spring Harb Protoc *2010*, pdb prot5350. 10.1101/pdb.prot5350.
62. Viebahn, C., Stortz, C., Mitchell, S.A., and Blum, M. (2002). Low proliferative and high migratory activity in the area of Brachyury expressing mesoderm progenitor cells in the gastrulating rabbit embryo. Development *129*, 2355-2365. 10.1242/dev.129.10.2355.
63. Keren-Shaul, H., Kenigsberg, E., Jaitin, D.A., David, E., Paul, F., Tanay, A., and Amit, I. (2019). MARS-seq2.0: an experimental and analytical pipeline for indexed sorting

combined with single-cell RNA sequencing. Nat Protoc *14*, 1841-1862. 10.1038/s41596-019-0164-4.

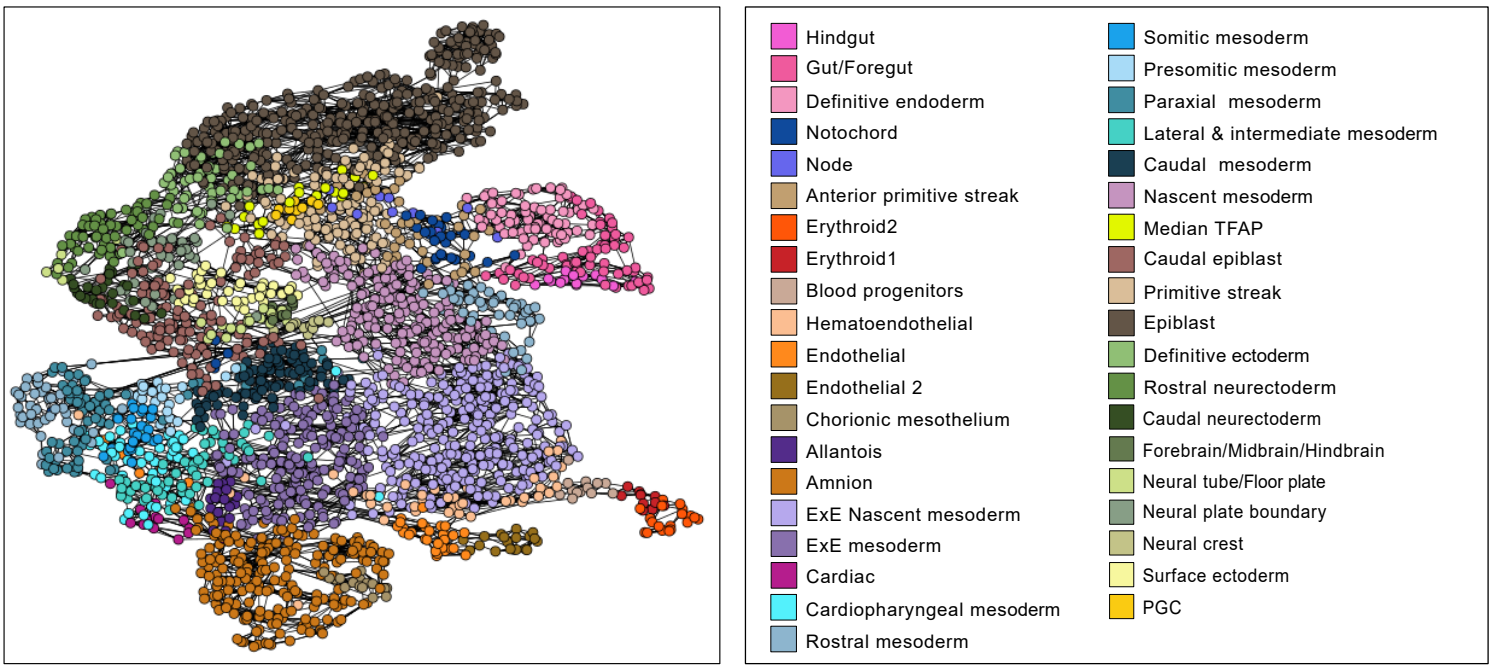
Figure 1

A

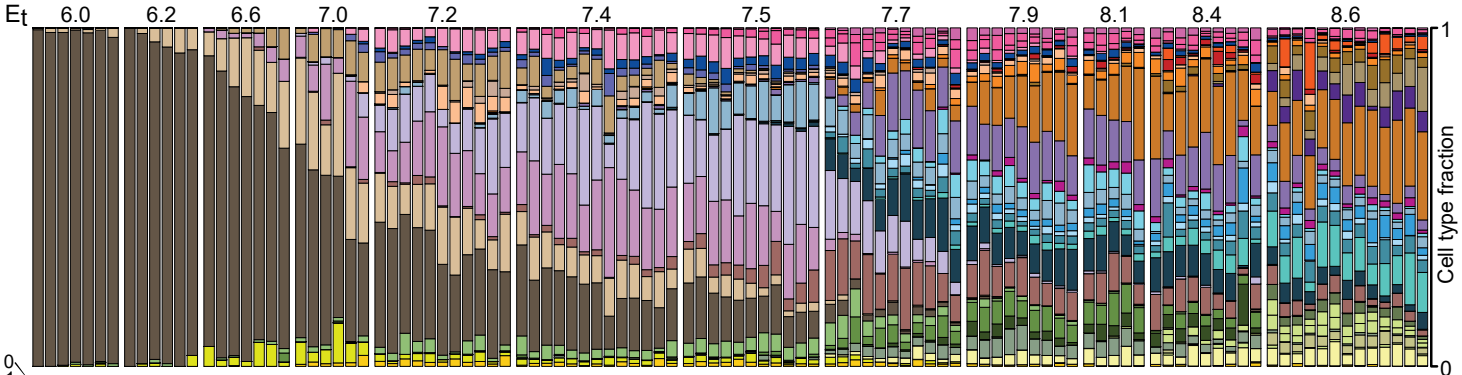


Gestation time	6.0	6.2	6.6	7.0	7.2	7.4	7.5	7.7	7.9	8.1	8.4	8.6
Developmental stage	st. 0	st. 1	st. 2	st. 3	st. 4	st. 4b	st. 4c	st. 5	st. 6	1-3 somites	4-7 somites	8-12 somites
Corresponding mouse stage	Pre-S	ES	ES/MS	MS	LS	LS/0B	LS/0B	EB/LB	EHF-LHF	1-3 somites	4-7 somites	8-12 somites

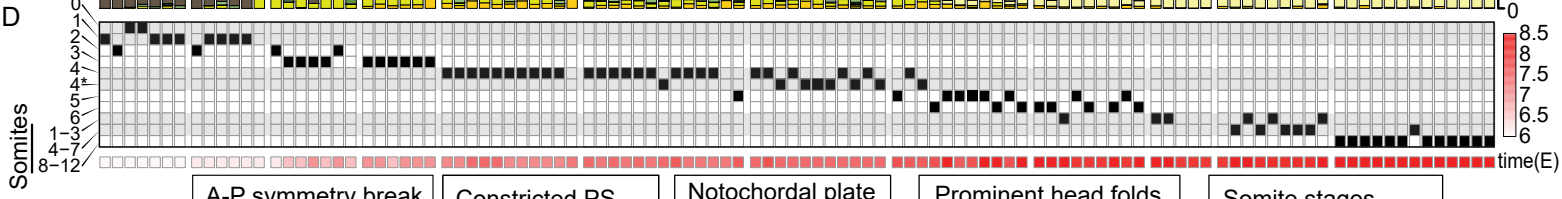
B



C



D



A-P symmetry break posterior gastrula extension (PGE)

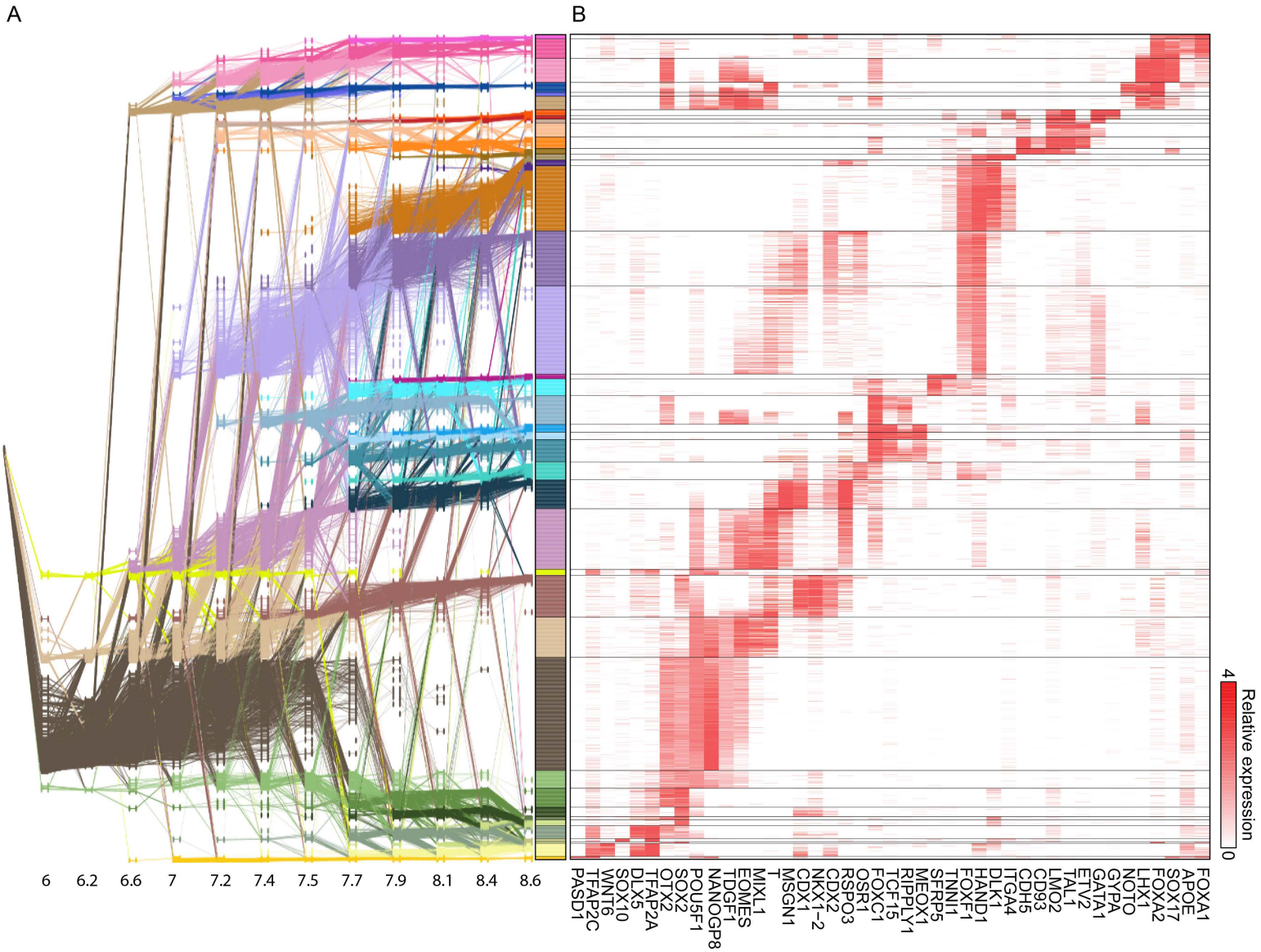
Constricted PS, defined node, robust endoderm and ectoderm

Notochordal plate

Prominent head folds, caudal epiblast, cellular diversification

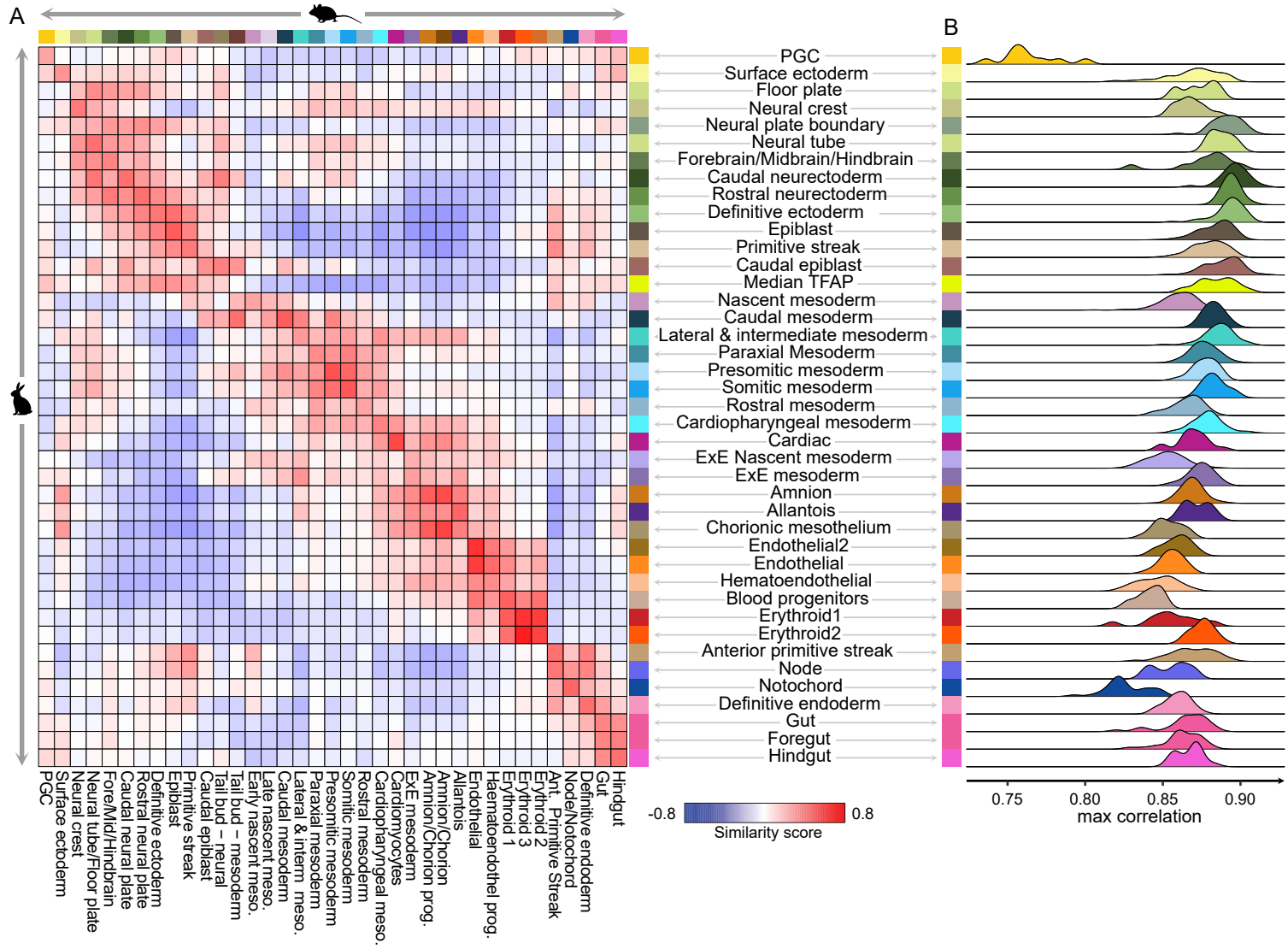
Somite stages, embryo elongation, cardiac crescent, neural tube

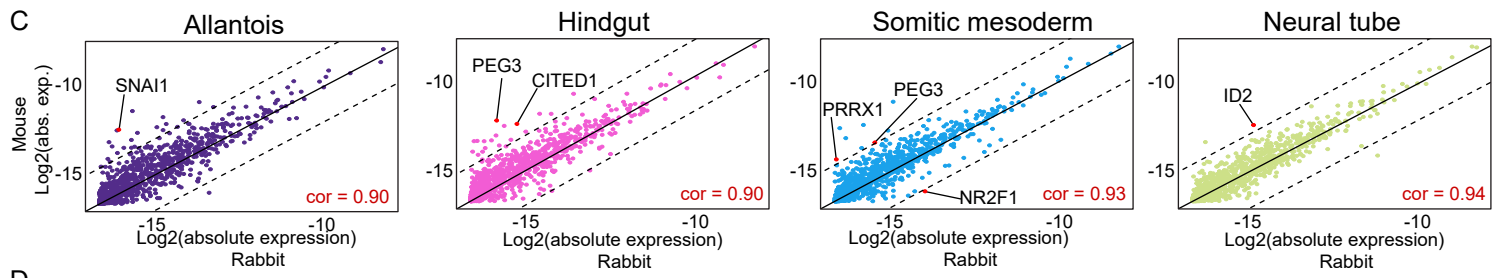
Figure 2



Hindgut	Blood progenitors	ExE mesoderm	Caudal mesoderm	Caudal neurectoderm
Gut/Foregut	Hematoendothelial	Cardiac	Nascent mesoderm	Forebrain/Midbrain/Hindbrain
Definitive endoderm	Endothelial	Cardiopharyngeal mesoderm	Median TFAP	Neural tube/Floor plate
Notochord	Endothelial 2	Rostral mesoderm	Caudal epiblast	Neural plate boundary
Node	Chorionic mesothelium	Somitic mesoderm	Primitive streak	Neural crest
Anterior primitive streak	Allantois	Presomitic mesoderm	Epiblast	Surface ectoderm
Erythroid2	Amnion	Paraxial mesoderm	Definitive ectoderm	PGC
Erythroid1	ExE Nascent mesoderm	Lateral & intermediate meso	Rostral neurectoderm	

Figure 3


A
B
PGC
Surface ectoderm
Floor plate
Neural crest
Neural plate boundary
Neural tube
Forebrain/Midbrain/Hindbrain
Caudal neurectoderm
Rostral neurectoderm
Definitive ectoderm
Epiblast
Primitive streak
Caudal epiblast
Median TFAP
Nascent mesoderm
Caudal mesoderm
Lateral & intermediate mesoderm
Paraxial Mesoderm
Presomitic mesoderm
Somitic mesoderm
Rostral mesoderm
Cardiopharyngeal mesoderm
Cardiac
ExE Nascent mesoderm
ExE mesoderm
Amnion
Allantois
Chorionic mesothelium
Endothelial2
Endothelial
Hematoendothelial
Blood progenitors
Erythroid1
Erythroid2
Anterior primitive streak
Node
Notochord
Definitive endoderm
Gut
Foregut
Hindgut
-0.8
0.8
Similarity score
0.75
0.80
0.85
0.90
max correlation


C
Allantois
Hindgut
Somitic mesoderm
Neural tube
SNAI1
PEG3
CITED1
PRRX1
NR2F1
ID2
cor = 0.90
cor = 0.90
cor = 0.93
cor = 0.94
Mouse
Log2(abs. exp.)
Log2(absolute expression)
Rabbit

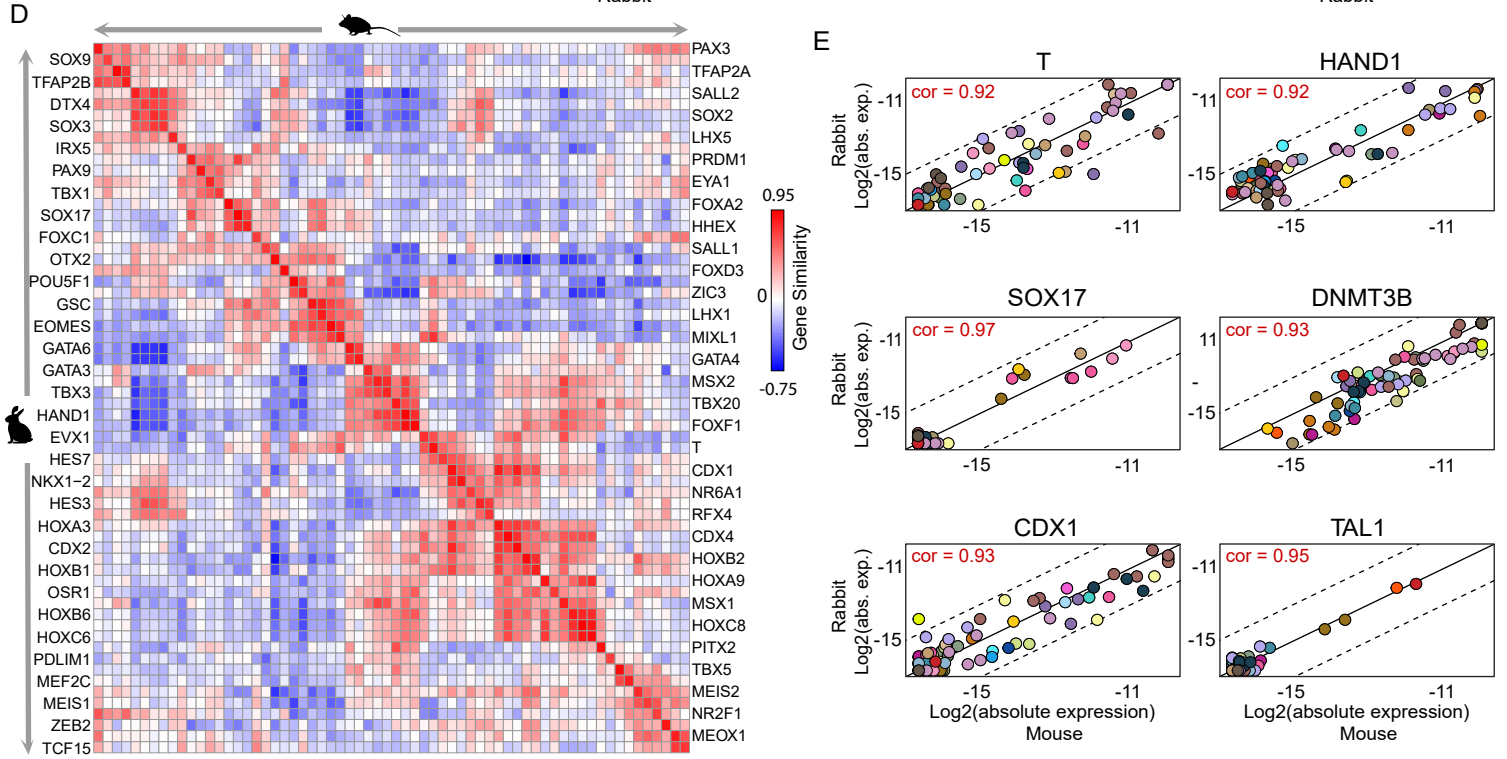

D
SOX9
TFAP2B
DTX4
SOX3
IRX5
PAX9
TBX1
SOX17
FOXC1
OTX2
POU5F1
GSC
EOMES
GATA6
GATA3
TBX3
HAND1
EVX1
HES7
NKX1-2
HES3
HOXA3
CDX2
HOXB1
OSR1
HOXB6
HOXC6
PDLIM1
MEF2C
MEIS1
ZEB2
TCF15
PAX3
TFAP2A
SALL2
SOX2
LHX5
PRDM1
EYA1
FOXA2
HHEX
SALL1
FOXD3
ZIC3
LHX1
MIXL1
GATA4
MSX2
TBX20
FOXF1
T
CDX1
NR6A1
RFX4
CDX4
HOXB2
HOXA9
MSX1
HOXC8
PITX2
TBX5
MEIS2
NR2F1
MEOX1
0.95
0
-0.75
Gene Similarity
E
T
HAND1
SOX17
DNMT3B
CDX1
TAL1
cor = 0.92
cor = 0.92
cor = 0.97
cor = 0.93
cor = 0.93
cor = 0.95
Rabbit
Log2(abs. exp.)
Log2(absolute expression)
Mouse

Figure 4

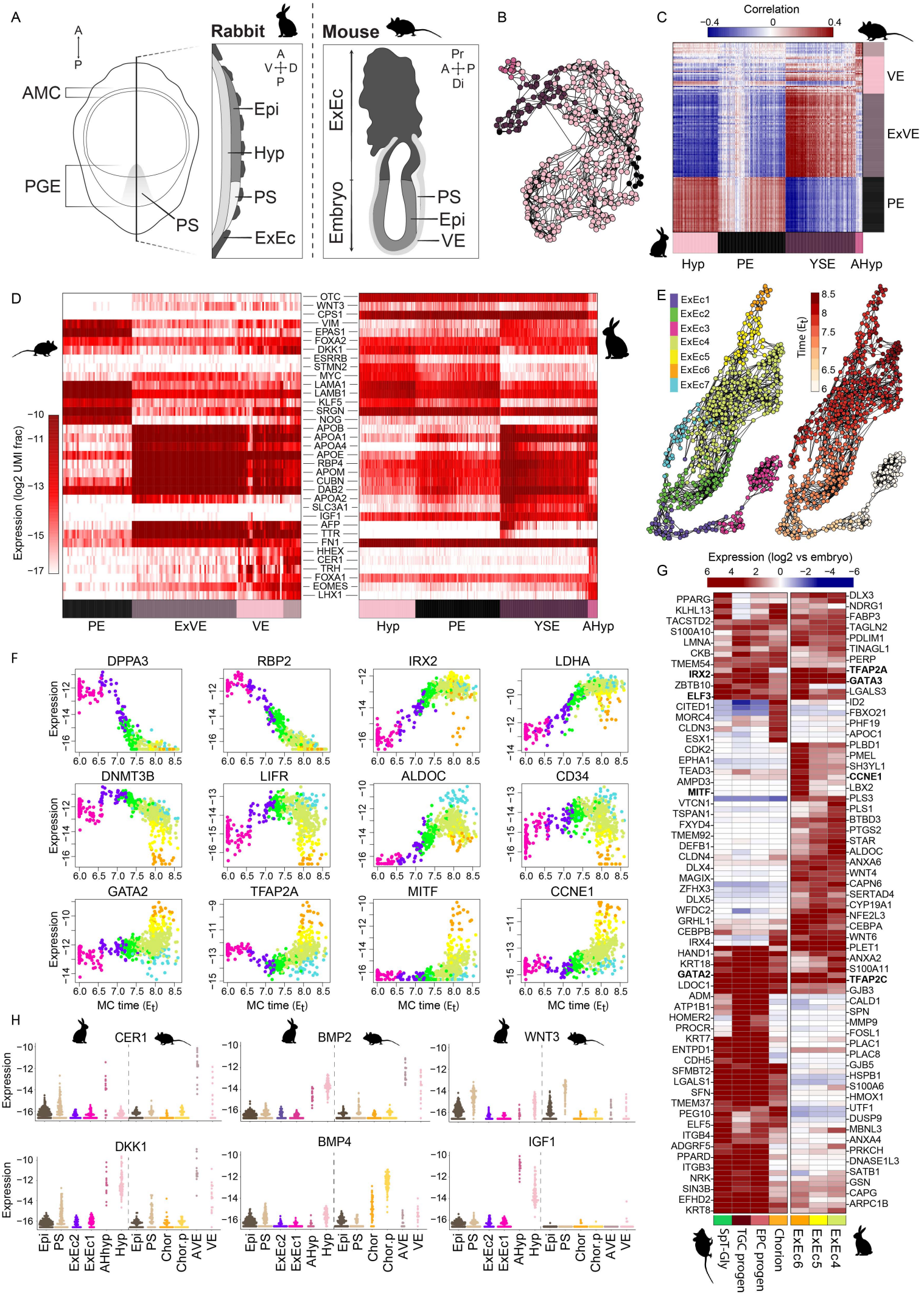

A
Rabbit
Mouse
AMC
PGE
PS
Epi
Hyp
ExEc
ExEc
Embryo
VE
B
C
Correlation
−0.4
0
0.4
VE
ExVE
PE
Hyp
PE
YSE
AHyp
D
Expression (log2 UMI frac)
OTC
WNT3
CPS1
VIM
EPAS1
FOXA2
DKK1
ESRRB
STMN2
MYC
LAMA1
LAMB1
KLF5
SRGN
NOG
APOB
APOA1
APOA4
APOE
RBP4
APOM
CUBN
DAB2
APOA2
SLC3A1
IGF1
AFP
TTR
FN1
HHEX
CER1
TRH
FOXA1
EOMES
LHX1
PE
ExVE
VE
Hyp
PE
YSE
AHyp
E
ExEc1
ExEc2
ExEc3
ExEc4
ExEc5
ExEc6
ExEc7
Time (Et)
F
DPPA3
RBP2
IRX2
LDHA
DNMT3B
LIFR
ALDOC
CD34
GATA2
TFAP2A
MITF
CCNE1
Expression
MC time (Et)
G
Expression (log2 vs embryo)
PPARG
KLHL13
TACSTD2
S100A10
LMNA
CKB
TMEM54
IRX2
ZBTB10
ELF3
CITED1
MORC4
CLDN3
ESX1
CDK2
EPHA1
TEAD3
AMPD3
MITF
VTCN1
TSPAN1
FXYD4
TMEM92
DEFB1
CLDN4
DLX4
MAGIX
ZFHX3
DLX5
WFDC2
GRHL1
CEBPB
IRX4
HAND1
KRT18
GATA2
LDOC1
ADM
ATP1B1
HOMER2
PROCR
KRT7
ENTPD1
CDH5
SFMBT2
LGALS1
SFN
TMEM37
PEG10
ELF5
ITGB4
ADGRF5
PPARD
ITGB3
NRK
SIN3B
EFHD2
KRT8
DLX3
NDRG1
FABP3
TAGLN2
PDLIM1
TINAGL1
PERP
TFAP2A
GATA3
LGALS3
ID2
FBXO21
PHF19
APOC1
PLBD1
PMEL
SH3YL1
CCNE1
LBX2
PLS3
PLS1
BTBD3
PTGS2
STAR
ALDOC
ANXA6
WNT4
CAPN6
SERTAD4
CYP19A1
NFE2L3
CEBPA
WNT6
PLET1
ANXA2
S100A11
TFAP2C
GJB3
CALD1
SPN
MMP9
FOSL1
PLAC1
PLAC8
GJB5
HSPB1
S100A6
HMOX1
UTF1
DUSP9
MBNL3
ANXA4
PRKCH
DNASE1L3
SATB1
GSN
CAPG
ARPC1B
SpT-Gly
TGC progen
EPC progen
Chorion
ExEc6
ExEc5
ExEc4
H
CER1
BMP2
WNT3
DKK1
BMP4
IGF1
Expression
Epi
PS
ExEc2
ExEc1
AHyp
Hyp
Epi
PS
Chor
Chor.p
AVE
VE

Figure 5

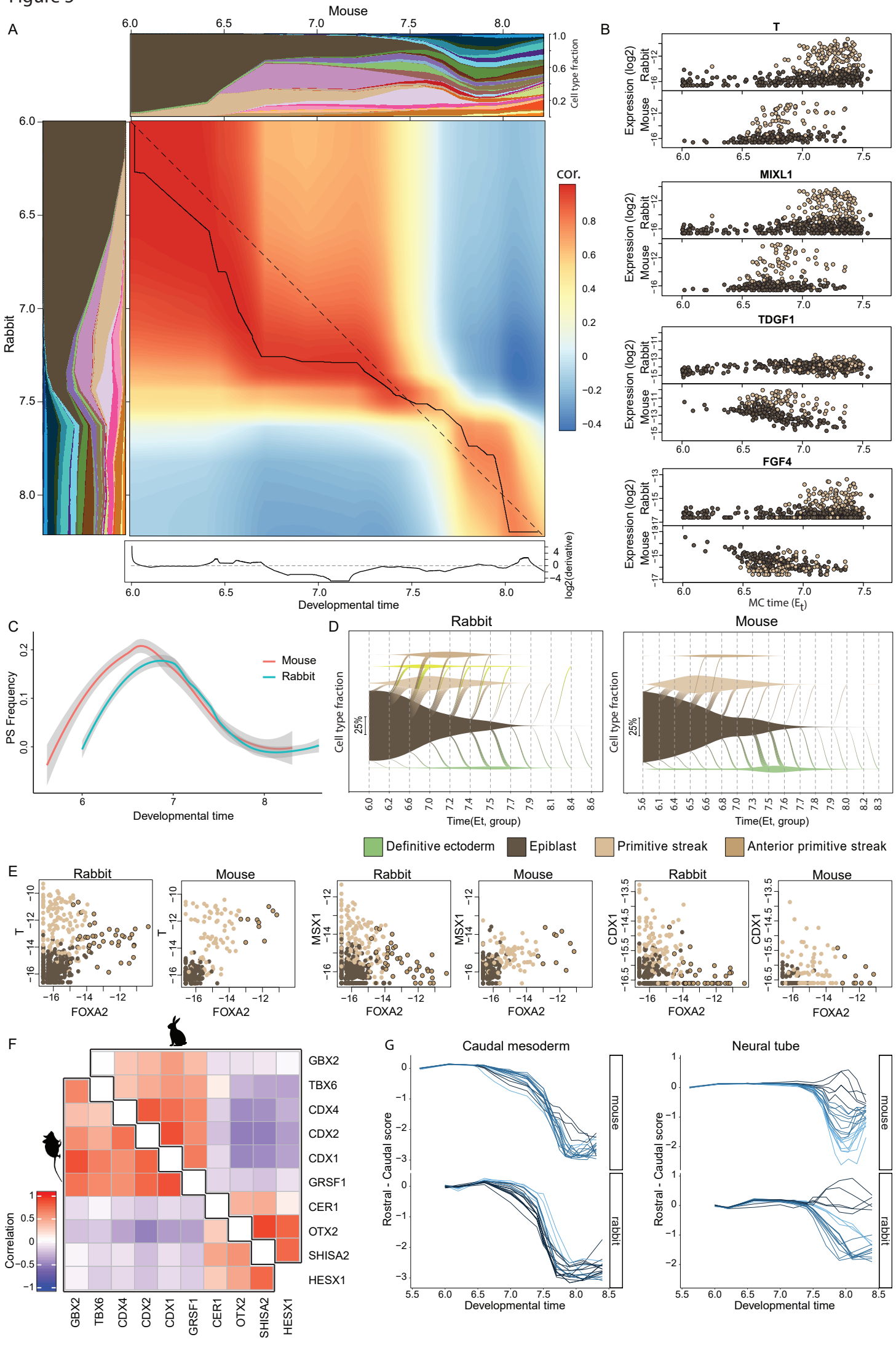

A
Mouse
Cell type fraction
Rabbit
cor.
log2(derivative)
Developmental time
B
T
MIXL1
TDGF1
FGF4
Expression (log2)
Rabbit
Mouse
MC time (Et)
C
PS Frequency
Mouse
Rabbit
Developmental time
D
Rabbit
Mouse
Cell type fraction
25%
Time(Et, group)
Definitive ectoderm
Epiblast
Primitive streak
Anterior primitive streak
E
Rabbit
Mouse
T
MSX1
CDX1
FOXA2
F
GBX2
TBX6
CDX4
CDX2
CDX1
GRSF1
CER1
OTX2
SHISA2
HESX1
Correlation
G
Caudal mesoderm
Neural tube
Rostral - Caudal score
mouse
rabbit
Developmental time

Figure 6

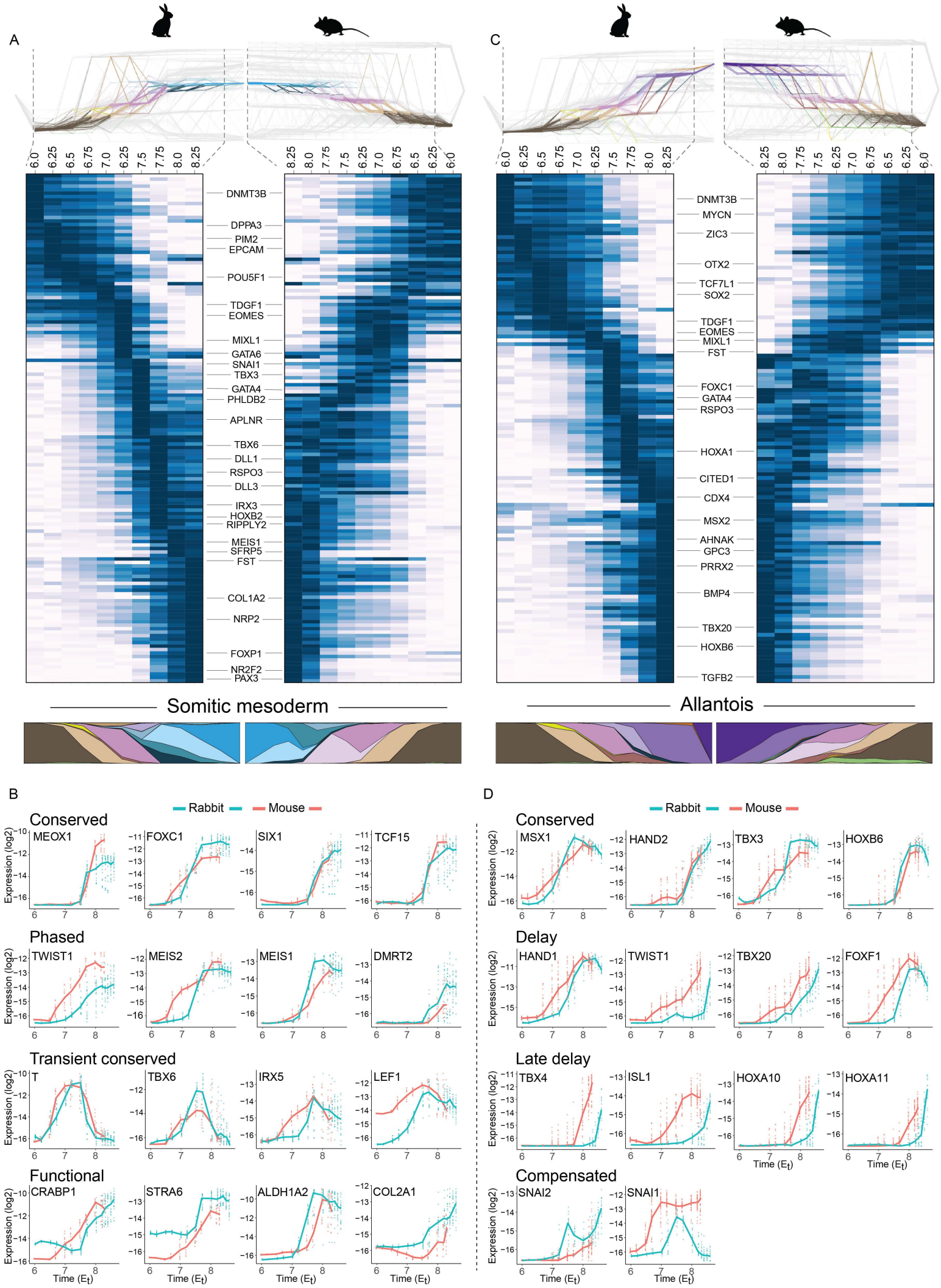
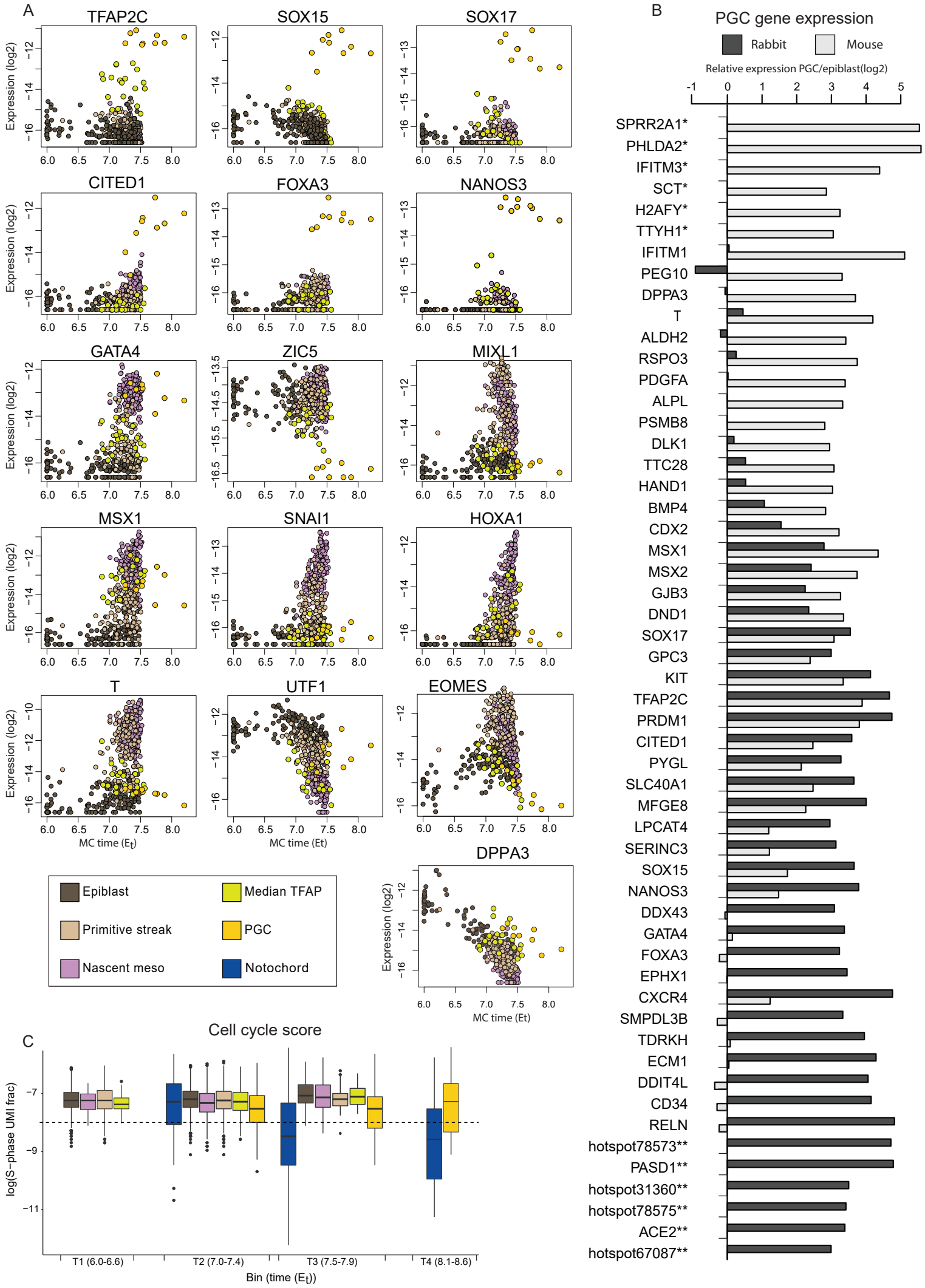

A
C
DNMT3B
DPPA3
PIM2
EPCAM
POU5F1
TDGF1
EOMES
MIXL1
GATA6
SNAI1
TBX3
GATA4
PHLDB2
APLNR
TBX6
DLL1
RSPO3
DLL3
IRX3
HOXB2
RIPPLY2
MEIS1
SFRP5
FST
COL1A2
NRP2
FOXP1
NR2F2
PAX3
MYCN
ZIC3
OTX2
TCF7L1
SOX2
FOXC1
HOXA1
CITED1
CDX4
MSX2
AHNAK
GPC3
PRRX2
BMP4
TBX20
HOXB6
TGFB2
Somitic mesoderm
Allantois
B
D
Rabbit
Mouse
Conserved
MEOX1
FOXC1
SIX1
TCF15
Phased
TWIST1
MEIS2
MEIS1
DMRT2
Transient conserved
T
TBX6
IRX5
LEF1
Functional
CRABP1
STRA6
ALDH1A2
COL2A1
MSX1
HAND2
TBX3
HOXB6
Delay
HAND1
TWIST1
TBX20
FOXF1
Late delay
TBX4
ISL1
HOXA10
HOXA11
Compensated
SNAI2
SNAI1
Expression (log2)
Time (Et)

Figure 7


A
TFAP2C
SOX15
SOX17
CITED1
FOXA3
NANOS3
GATA4
ZIC5
MIXL1
MSX1
SNAI1
HOXA1
T
UTF1
EOMES
DPPA3
Expression (log2)
MC time (Et)
Epiblast
Primitive streak
Nascent meso
Median TFAP
PGC
Notochord
B
PGC gene expression
Rabbit
Mouse
Relative expression PGC/epiblast(log2)
SPRR2A1*
PHLDA2*
IFITM3*
SCT*
H2AFY*
TTYH1*
IFITM1
PEG10
DPPA3
T
ALDH2
RSPO3
PDGFA
ALPL
PSMB8
DLK1
TTC28
HAND1
BMP4
CDX2
MSX1
MSX2
GJB3
DND1
SOX17
GPC3
KIT
TFAP2C
PRDM1
CITED1
PYGL
SLC40A1
MFGE8
LPCAT4
SERINC3
SOX15
NANOS3
DDX43
GATA4
FOXA3
EPHX1
CXCR4
SMPDL3B
TDRKH
ECM1
DDIT4L
CD34
RELN
hotspot78573**
PASD1**
hotspot31360**
hotspot78575**
ACE2**
hotspot67087**
C
Cell cycle score
log(S−phase UMI frac)
T1 (6.0-6.6)
T2 (7.0-7.4)
T3 (7.5-7.9)
T4 (8.1-8.6)
Bin (time (Et))