RESEARCH ARTICLE SUMMARY

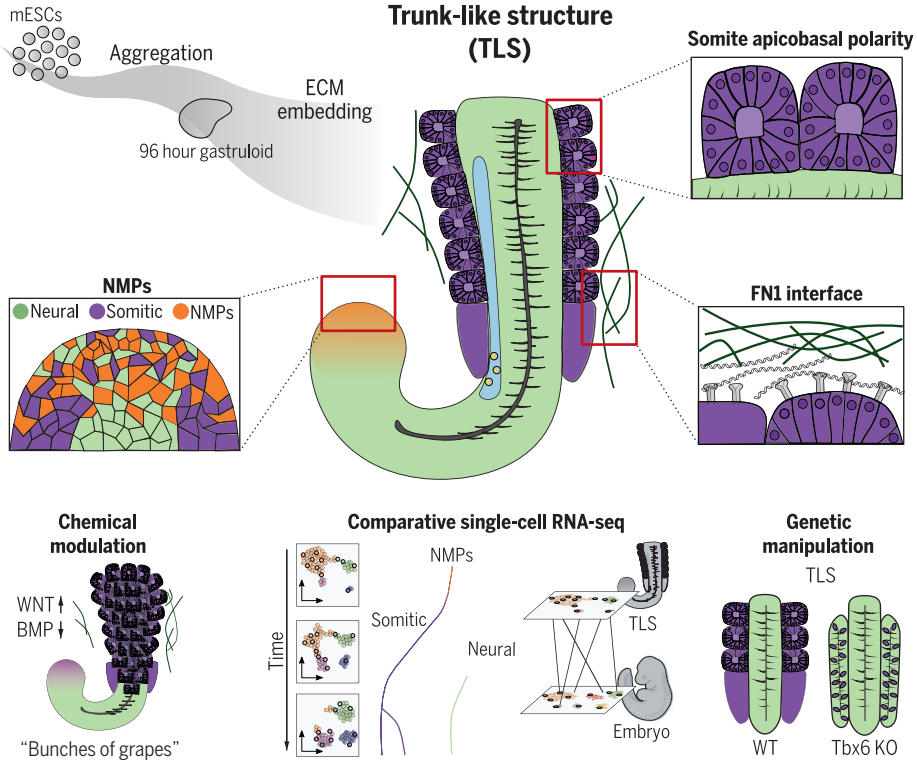
DEVELOPMENTAL BIOLOGY

Mouse embryonic stem cells self-organize into trunk-like structures with neural tube and somites

Jesse V. Veenvliet*†, Adriano Bolondi†, Helene Kretzmer‡, Leah Haut‡, Manuela Scholze-Wittler, Dennis Schifferl, Frederic Koch, Léo Guignard, Abhishek Sampath Kumar, Milena Pustet, Simon Heimann, René Buschow, Lars Wittler, Bernd Timmermann, Alexander Meissner*, Bernhard G. Herrmann*

INTRODUCTION: Vertebrate development comprises multiple complex morphogenetic processes that shape the embryonic body plan through self-organization of pluripotent stem cells and their descendants. Because mammalian embryogenesis proceeds in utero, it is difficult to study the dynamics of these processes, including much-needed analysis at the cellular and molecular level. Various three-dimensional stem cell systems ("embryoids") have been developed to circumvent this impediment. The most advanced models of post-implantation development achieved so far are gastruloids, mouse embryonic stem cell (mESC)–derived aggregates with organized gene expression domains but lacking proper morphogenesis.

RATIONALE: To advance the current models, we explored the usage of Matrigel, an extracellular matrix (ECM) surrogate. During embryonic development, the ECM provides essential chemical and mechanical cues. In vitro, lower percentages of Matrigel can drive complex tissue morphogenesis in organoids, which led us to use Matrigel embedding in various media conditions to achieve higher-order embryo-like architecture in mESC-derived aggregates.



Engineering the embryonic trunk in a dish. Embedding of mouse embryonic stem cell (mESC) aggregates in an extracellular matrix (ECM) enables generation of trunk-like structures (TLSs) with an in vivo–like architecture including gut, and neuromesodermal progenitor (NMP)–derived neural tube and somites. Comparative single-cell RNA sequencing revealed that TLS cell states and differentiation dynamics match those of the embryo. Chemical modulation and genetic manipulation highlight the utility of the TLS as a scalable, tractable, and accessible model for investigating mid-gestational embryogenesis. FN1, fibronectin.

RESULTS: We found that embedding of 96-hour gastruloids in 5% Matrigel is sufficient to induce the formation of highly organized "trunk-like structures" (TLSs), comprising the neural tube and bilateral somites with embryo-like polarity. This high level of self-organization was accompanied by accumulation of the matrix protein fibronectin at the Matrigel-TLS interface and the transcriptional up-regulation of fibronectin-binding integrins and other cell adhesion molecules. Chemical modulation of signaling pathways active in the developing mouse embryo [WNT and bone morphogenetic protein (BMP)] resulted in an excess of somites arranged like a "bunch of grapes." Comparative time-resolved single-cell RNA sequencing of TLSs and embryos revealed that TLSs follow the same stepwise gene regulatory programs as the mouse embryo, comprising expression of critical developmental regulators at the right place and time. In particular, trunk precursors known as neuromesodermal progenitors displayed the highest differentiation potential and continuously contributed to neural and mesodermal tissue during TLS formation. In addition, live imaging demonstrated that the segmentation clock, required for rhythmic deposition of somites in vivo, ticks at an embryo-like pace in TLSs. Finally, a proof-of-principle experiment showed that Tbx6-knockout TLSs generate ectopic neural tubes at the expense of somite formation, mirroring the embryonic phenotype.

CONCLUSION: We showed that embedding of embryonic stem cell–derived aggregates in an ECM surrogate generates more advanced in vitro models that are formed in a process highly analogous to embryonic development. Trunk-like structures represent a powerful tool that is easily amenable to genetic, mechanical, chemical, or other modulations. As such, we expect them to facilitate in-depth analysis of molecular mechanisms and signaling networks that orchestrate embryonic development as well as studies of the ontogeny of mutant phenotypes in the culture dish. The scalable, tractable, and highly accessible nature of the TLS makes it a complementary in vitro platform for deciphering the dynamics of the molecular, cellular, and morphogenetic processes that shape the post-implantation embryo, at an unprecedented spatiotemporal resolution. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: veenvlie@molgen.mpg.de (J.V.V.); meissner@molgen.mpg.de (A.M.); herrmann@molgen.mpg.de (B.G.H.)
†These authors contributed equally to this work.
‡These authors contributed equally to this work.
Cite this article as J. V. Veenvliet *et al.*, *Science* **370**, eaba4937 (2020). DOI: 10.1126/science.aba4937

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.aba4937

RESEARCH ARTICLE

DEVELOPMENTAL BIOLOGY

Mouse embryonic stem cells self-organize into trunk-like structures with neural tube and somites

Jesse V. Veenvliet[1]*†, Adriano Bolondi[2]†, Helene Kretzmer[2]‡, Leah Haut[1,2]‡, Manuela Scholze-Wittler[1], Dennis Schifferl[1], Frederic Koch[1], Léo Guignard[3], Abhishek Sampath Kumar[2], Milena Pustet[1], Simon Heimann[1], René Buschow[4], Lars Wittler[1], Bernd Timmermann[5], Alexander Meissner[2,6,7,8]*, Bernhard G. Herrmann[1,9]*

Post-implantation embryogenesis is a highly dynamic process comprising multiple lineage decisions and morphogenetic changes that are inaccessible to deep analysis in vivo. We found that pluripotent mouse embryonic stem cells (mESCs) form aggregates that upon embedding in an extracellular matrix compound induce the formation of highly organized "trunk-like structures" (TLSs) comprising the neural tube and somites. Comparative single-cell RNA sequencing analysis confirmed that this process is highly analogous to mouse development and follows the same stepwise gene-regulatory program. *Tbx6* knockout TLSs developed additional neural tubes mirroring the embryonic mutant phenotype, and chemical modulation could induce excess somite formation. TLSs thus reveal an advanced level of self-organization and provide a powerful platform for investigating post-implantation embryogenesis in a dish.

Vertebrate post-implantation development comprises a multitude of complex morphogenetic processes resulting from self-organization of stem cells and their descendants shaping the embryonic body plan (*1*). Recently developed stem cell models represent powerful platforms for deconstructing the dynamics of these processes in vitro (*1*, *2*). The most advanced models in terms of developmental stage are gastruloids, self-organizing mESC aggregates that form elongating structures that comprise postoccipital embryo derivatives of all three germ layers but lack proper morphogenesis (*1*–*3*). In vivo, the extracellular matrix (ECM) provides chemical signals and exerts mechanical constraints via the basement membrane, which has a critical role in tissue morphogenesis (*4*, *5*). In vitro, Matrigel can serve as an ECM surrogate, and embedding of gastruloids in 10% Matrigel allowed the formation of a string of somite-like structures with anterior-posterior polarity (*6*). Lower percentages of Matrigel facilitated complex morphogenesis in organoids (*7*).

[1]Department of Developmental Genetics, Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany. [2]Department of Genome Regulation, Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany. [3]Max Delbrück Center for Molecular Medicine and Berlin Institute of Health, 10115 Berlin, Germany. [4]Microscopy and Cryo-Electron Microscopy, Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany. [5]Sequencing Core Facility, Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany. [6]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA. [7]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [8]Institute of Chemistry and Biochemistry, Freie Universität Berlin, 14195 Berlin, Germany. [9]Institute for Medical Genetics, Charité–University Medicine Berlin, Campus Benjamin Franklin, 12203 Berlin, Germany. ***Corresponding author. Email: veenvlie@molgen.mpg.de (J.V.V.); meissner@molgen.mpg.de (A.M.); herrmann@molgen.mpg.de (B.G.H.)** †These authors contributed equally to this work. ‡These authors contributed equally to this work.

Matrigel embedding drives trunk-like morphogenesis

To achieve more advanced embryo-like morphogenetic features in gastruloids, we used Matrigel with various media conditions (Fig. 1A). To facilitate high-throughput characterization and quantification, we generated mESCs with *T::H2B-mCherry* (hereafter T^{mCH}) and *Sox2::H2B-Venus* (hereafter $Sox2^{VE}$) reporters, marking the mesodermal (ME) or neural (NE) lineage, respectively (fig. S1A). Embedding of 96-hour aggregates in 5% Matrigel resulted in segmentation in the T^{mCH+} ME domain and formation of a $Sox2^{VE+}$ neural tube–like structure (Fig. 1, A to C, fig. S1, B and C, and movies S1 and S2). The vast majority of structures [hereafter referred to as trunk-like structures (TLSs)] elongated and formed a T^{mCH+} pole at the posterior end, with segmentation occurring in about half of the TLSs (Fig. 1D). Whole-mount in situ hybridization for Tcf15 and Uncx confirmed somite identity and revealed embryo-like anterior-posterior polarity (Fig. 1E) (*8*). In 61% of the segmented TLSs, bilateral somites were observed (fig. S1D). Additional WNT activation using CHIR99021 (hereafter TLS^{C}) or combined with bone morphogenetic protein (BMP) inhibition by LDN193189 (hereafter TLS^{CL}) improved the physical separation of neighboring segments without affecting T^{mCH+} pole formation or elongation (Fig. 1, A to D, and fig. S1, E to H) and resulted in an excess of segments at the anterior end, arranged like a "bunch of grapes" (Fig. 1, B, C, and F) (*9*). Moreover, the ME domain expanded at the expense of the NE compartment, with apparent disorganization of the posterior end and neural tissue (Fig. 1B and fig. S1, G and H). This phenotype has not been observed in vivo and may be explained by shifting the lineage choice of neuromesodermal progenitors (NMPs)—bipotent cells giving rise to both postoccipital NE and ME (*10*)—toward ME as a result of dominant WNT signaling. In addition, $T^{mCH+}/Sox2^{VE+}$ putative NMPs were reduced, as confirmed by flow cytometry (fig. S1I).

Phalloidin and N-cadherin staining demonstrated that the cells of the neural tube and somites of TLS, TLS^{C}, and TLS^{CL} show proper apicobasal polarity, a characteristic of epithelial tissues, with F-actin and N-cadherin accumulating at the apical side (Fig. 1, G and H, and fig. S2, A and B). In gastruloids, by contrast, F-actin distribution appeared random and organized epithelial structures were not detected (fig. S2C).

Whole-mount immunofluorescence analysis of FOXA2 and SOX17, transcription factors characteristic for endoderm, revealed gut formation in a subset of TLSs (fig. S3, A to D). Cells at the posterior base were SOX17-negative but coexpressed FOXA2 with high levels of T^{mCH} (fig. S3B). Our data show that embedding in Matrigel is both necessary and sufficient to drive complex, embryo-like tissue morphogenesis of the three embryonic germ layers.

We next performed a detailed morphometric analysis of TLSs and their substructures. The data demonstrate reproducibility of the three protocols with respect to size and shape of the whole structure, somites, and neural tube, whereas the gut-like structure shows more variation (fig. S3E and fig. S4, A to D). Time-resolved whole-structure morphometry showed similar morphogenetic changes over time (fig. S4B). TLSs in general were larger than gastruloids (fig. S4A). Comparisons of the different TLS protocols revealed that, relative to TLS, TLS^{C} and TLS^{CL} were slightly larger and formed more somites, whereas their neural domains were reduced in length and narrowed toward the anterior end (fig. S4, A to D). In all protocols, somites were similar in shape but smaller than their embryonic counterparts (fig. S4D).

To assess whether the segmentation clock, an oscillator involved in somitogenesis, is active in TLSs, we performed live imaging (*11*). In line with recent observations, we found that segmentation occurs in a rhythmic fashion at an embryo-like pace in all three TLS conditions (fig. S5, A to C, and movies S3 to S5) (*6*). TLS^{C} and TLS^{CL} showed consecutive formation of multiple somites (movies S6 to S9).

Transcriptional characterization highlights selective responses to chemical modulation

To characterize the structures in detail at the molecular level, we performed RNA sequencing

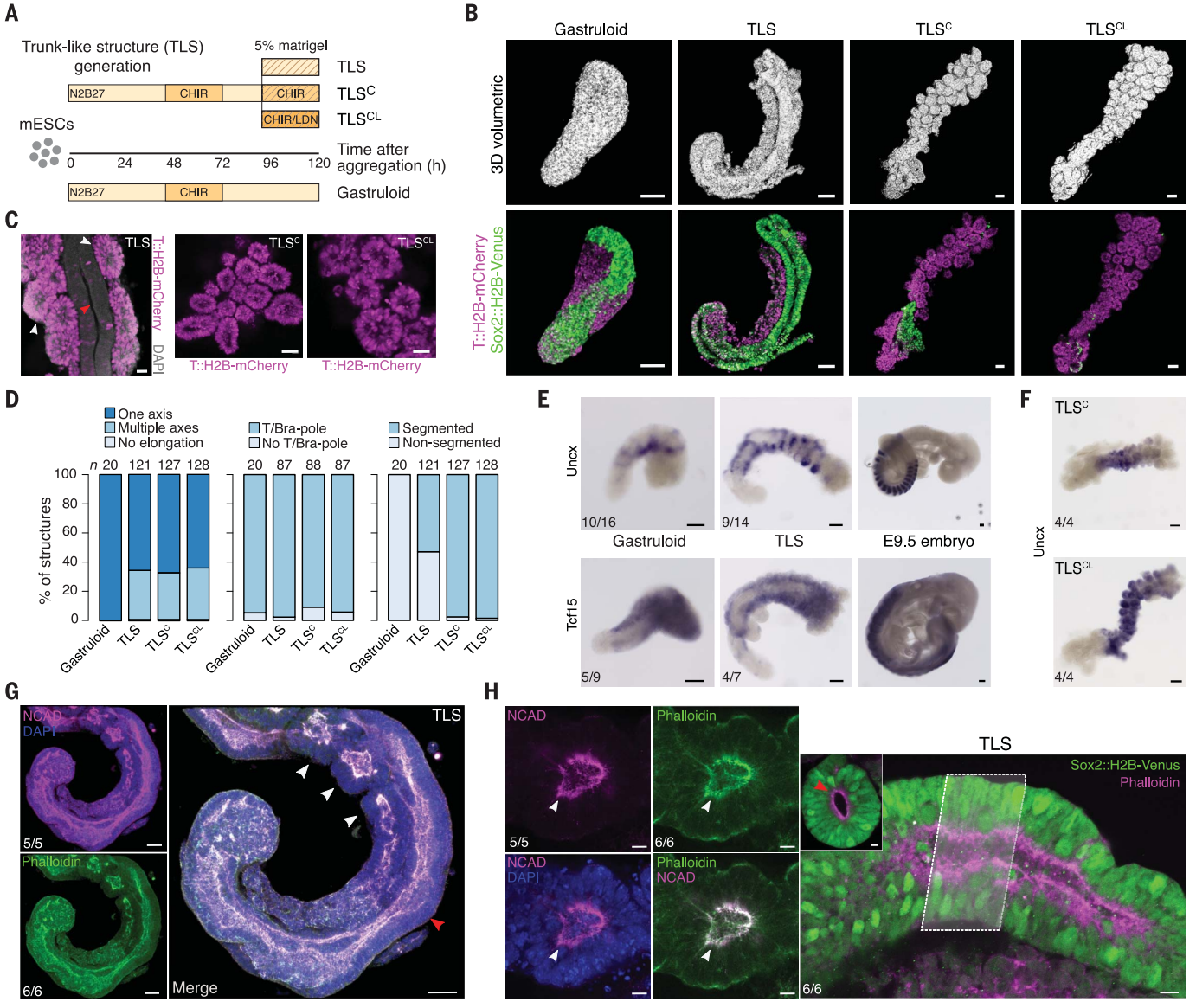


Fig. 1. Generation of trunk-like structures (TLSs) with somites and a neural tube. (**A**) Schematic overview: 200 to 250 mESCs were aggregated in ultralow-attachment plates; *Wnt* agonist CHIR99021 (CHIR) was added between 48 and 72 hours (*3*). At 96 hours, aggregates were (i) cultured as gastruloids (*3*), (ii) embedded in 5% Matrigel (TLS), and (iii) treated with WNT signaling activator CHIR99021 (CHIR) alone (TLS^C), or (iv) treated with CHIR and the BMP signaling inhibitor LDN193189 (TLS^{CL}). The two compounds have been reported to induce a (pre-)somitic mesoderm fate in 2D and 3D differentiation protocols (*8*). (**B**) 3D volumetric renderings (top) and confocal sections (bottom) of a gastruloid, TLS, TLS^C, and TLS^{CL}. Scale bars, 100 μm. Each image is representative of at least 10 biological replicates with similar morphology and expression patterns. (**C**) Segments in TLS are T^{mCH+} and are positioned adjacent to the neural tube. In TLS^C and TLS^{CL} the segments are arranged in "bunches of grapes." Scale bars, 25 μm. Red arrowhead, neural tube; white arrowheads, somites. (**D**) Quantification of morphogenetic features in gastruloids, TLSs, TLS^C, and TLS^{CL} (see supplementary materials for scoring criteria). (**E**) Segments express somitic markers Uncx and Tcf15 as shown by whole-mount in situ hybridization. Note the characteristic stripe-like expression pattern of Uncx in TLS due to posterior restriction, whereas Tcf15 is expressed throughout the segments (as in the embryo). Scale bars, 100 μm. (**F**) In TLS^C and TLS^{CL}, Uncx is detected throughout the segments, indicating loss of anterior-posterior polarity. Scale bars, 100 μm. (**G** and **H**) Confocal sections showing that cells of somites and neural tube display apical-basal polarity with NCAD and F-actin (phalloidin) accumulating at the apical surface. Inset represents an optical section of the neural tube shown in the main panel. Scale bars, 50 μm (G), 10 μm (H). White arrowheads, somites; red arrowheads, neural tubes.

(RNA-seq) analysis (fig. S6A) and found that the TLS models the postoccipital embryo, similar to gastruloids (Fig. 2A) (*3*). Relative to TLS, both TLS^C and TLS^{CL} showed significant up-regulation of genes involved in (pre)somitic development [e.g., *Tbx6*, *Msgn1*, *Hes7* (*8*, *10*, *12*)] at the expense of NE marker genes [e.g., *Sox1*, *Pax6*, *Irx3* (*13*)], corroborating the flow cytometry and imaging results (Fig. 2A and fig. S6, B and C). The analysis of marker gene sets for NMPs, for their direct descendants undergoing lineage choice (NMP ME and NMP NE), and for committed NE and ME cells substantiated this finding. Moreover, TLS^C and TLS^{CL} displayed reduced expression of markers in all clusters, including ME, relative to TLS (Fig. 2B). In contrast, on average (pre)somitic mesoderm [(P)SM]–specific markers were up-regulated, whereas intermediate ME (IM) and lateral plate ME (LPM) markers were down-regulated in TLS^C and further reduced

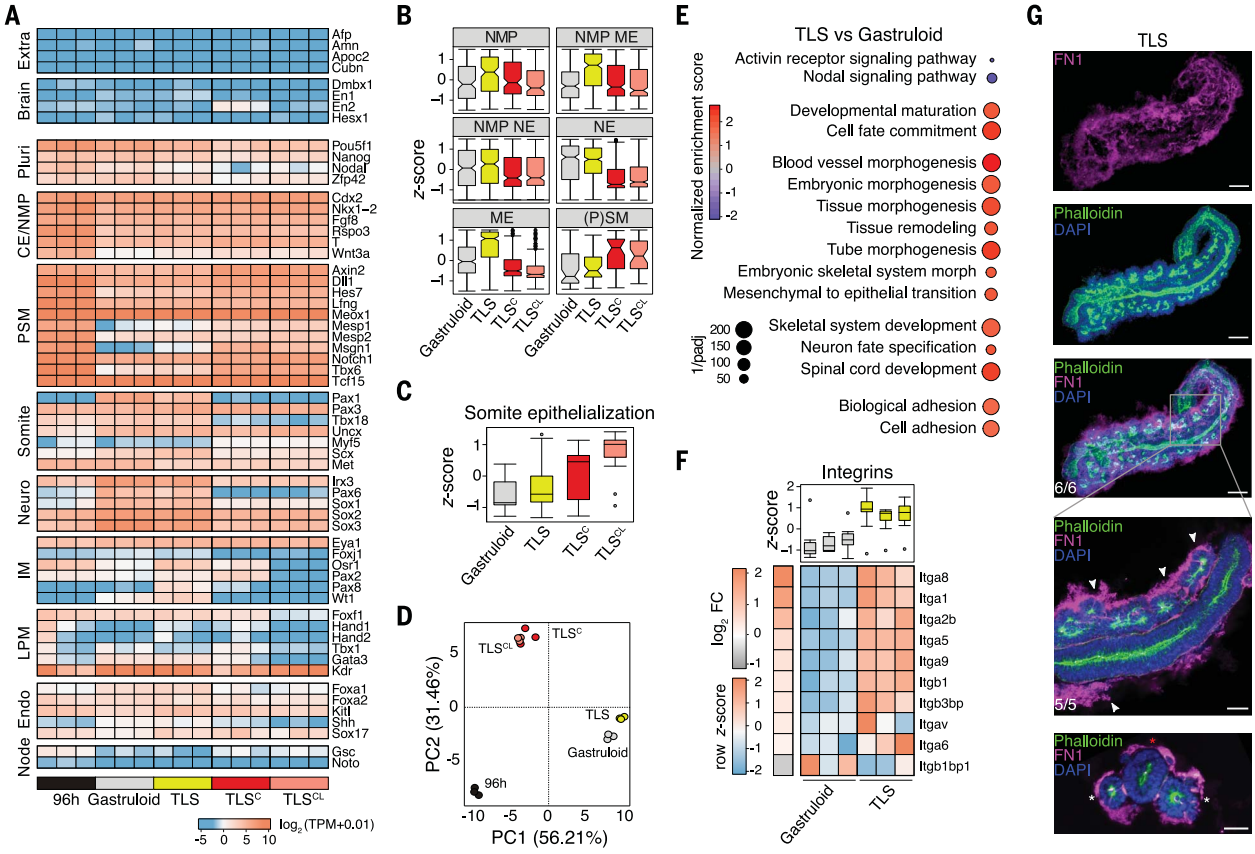


Fig. 2. Transcriptional characterization of TLSs. (**A**) Heatmap of $\log_2$(TPM + 0.01) expression (TPM, transcripts per million) of selected genes associated with development of indicated embryonic structures in 96-hour aggregates and 120-hour gastruloids, 120-hour TLSs, 120-hour TLS^{C}, and 120-hour TLS^{CL}, as measured by RNA-seq. Replicates were derived from pools of independent biological samples (see fig. S6A for experimental setup). CE, caudal end; NMP, neuromesodermal progenitors; PSM, presomitic mesoderm; LPM, lateral plate mesoderm; IM, intermediate mesoderm. (**B**) Box plots showing distribution of marker genes for indicated cell types (average *z*-score). Boxes indicate interquartile range (IQR); whiskers extend to 1.5 × IQR from the hinge. Dots indicate outliers. Notches are centered on the median. See data S1 for lists of genes used for each category and statistical analysis. (**C**) Box plot representing average *z*-score per column (pool of three replicates) for somite epithelialization factors (see fig. S6E for individual genes). (**D**) PCA analysis of samples from (A) with color coding of individual samples (dots) as in (A). PC1 and PC2 represent the two components with highest percentage of explained variance. (**E**) Selected significant terms of gene set enrichment analysis (GSEA) enriched in TLSs as compared to gastruloids at 120 hours. See data S2 for full list of significant terms [false discovery rate (FDR) < 0.05]. (**F**) Heatmap of scaled expression (row *z*-score) and $\log_2$ FC of integrins with significantly different expression ($P_{adj(FDR)}$ < 0.05) in 120-hour TLSs versus gastruloids. Box plot represents *z*-score per column (sample), with boxes indicating interquartile range, whiskers extending to 1.5 × IQR from the hinge, dots showing outliers, and central line representing median. Each column represents one of three biological replicates. See data S1 for statistical analysis. (**G**) 3D maximum-intensity projection (top three images) and confocal section showing FN1 accumulation around TLS somites and neural tube (zoomed-in image, white arrowheads). Phalloidin staining shows apical-basal polarity. Scale bars, 100 μm, 50 μm for magnification. Bottom: Light-sheet optical transversal section showing FN1 accumulation around the somites (white asterisk) and neural tube (red asterisk) in TLSs. Scale bar, 50 μm.

in TLS^{CL} (Fig. 2B and fig. S6D) (*3*, *12*), in line with the known role of WNT versus BMP signaling in PSM versus IM and LPM specification (Fig. 2A and fig. S6D) (*8*).

We next searched for gene expression differences possibly underlying improved physical separation of somites observed in TLS^{C} and TLS^{CL}. Among the most strongly up-regulated genes relative to TLSs was *Wnt6*, which acts as a somite epithelialization factor in vivo (fig. S6E) (*14*). In addition, multiple ephrins and their receptors, and other factors involved in somite epithelialization, were up-regulated (Fig. 2C and fig. S6E) (*15*, *16*). We observed expression changes of selected somite polarity markers and their inducers, in line with the role played by WNTs, SHH, BMPs, and their antagonists in somite compartmentalization in vivo (fig. S6, F and G) (*8*, *17*). Our data show that exposure to CHIR or CHIR/LDN improved segment boundary formation but affected somite cellular composition.

Tissue morphogenesis and remodeling genes are up-regulated in TLSs

Principal components analysis (PCA) indicated a high transcriptional similarity between gastruloids and TLSs despite profound morphological differences (Fig. 2D). The latter are better highlighted by gene set enrichment analysis (GSEA), which showed that Matrigel embedding promotes tissue morphogenesis and remodeling (Fig. 2E and fig. S6H). Zooming in on embryonic and tissue morphogenesis gene sets revealed markers of blood vessel development among up-regulated genes, which suggests the induction of capillary morphogenesis in TLSs (fig. S6I). GSEA also showed an enrichment of cell adhesion terms and overall a significant up-regulation of corresponding marker genes in TLSs (Fig. 2E and fig. S7, A and B). The most pronounced increase was observed for integrins, transmembrane receptors mediating cell adhesion to the ECM (important for, e.g., neural tube formation, blood vessel development, and segmentation) (Fig. 2F and fig. S7, A and B) (*18*–*21*). Because binding of integrin to the glycoprotein fibronectin (FN1) and matrix

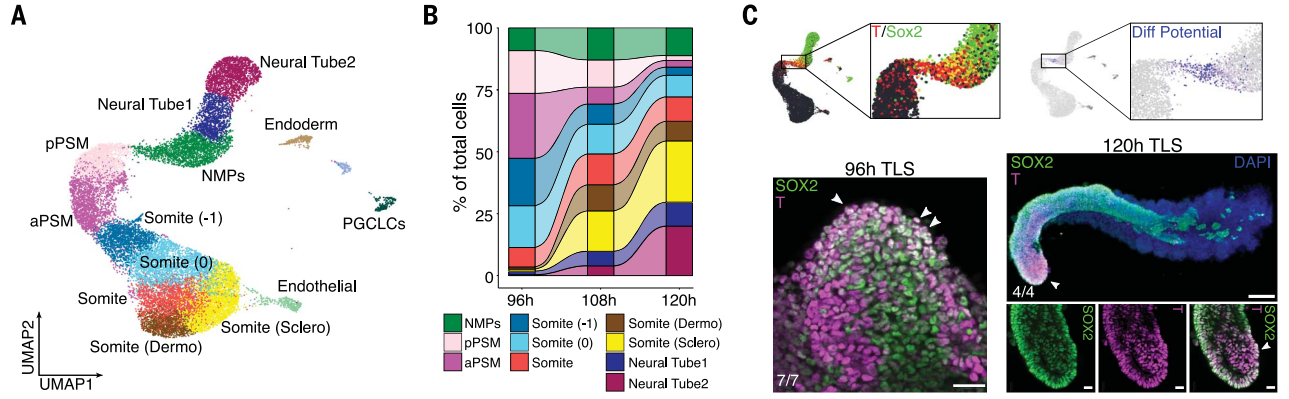


Fig. 3. Time-resolved single-cell RNA-seq of TLSs. In total, 20,294 cells were sampled from TLSs at 96, 108, and 120 hours (see fig. S9A for experimental setup). (**A**) UMAP (uniform manifold approximation and projection) colored by the 14 clusters identified. (**B**) Alluvial plot of percentage of neuromesodermal progenitors (NMPs), posterior presomitic mesoderm (pPSM), anterior PSM, somitic, and neural tube cells over time. (**C**) NMPs coexpress Sox2 and T (top left; blending with normalized expression threshold = 0.25) and are characterized by the highest differentiation potential (top right; see supplementary materials for differentiation potential calculation). NMPs coexpress T and SOX2 at 96 and 120 hours and reside at the posterior end of the TLS (confocal sections, bottom left and magnifications, 3D maximum-intensity projection, whole structure); white arrowheads, NMPs. Scale bars, 50 μm for 96-hour TLS, 100 μm for 3D maximum-intensity projection, 20 μm for magnifications.

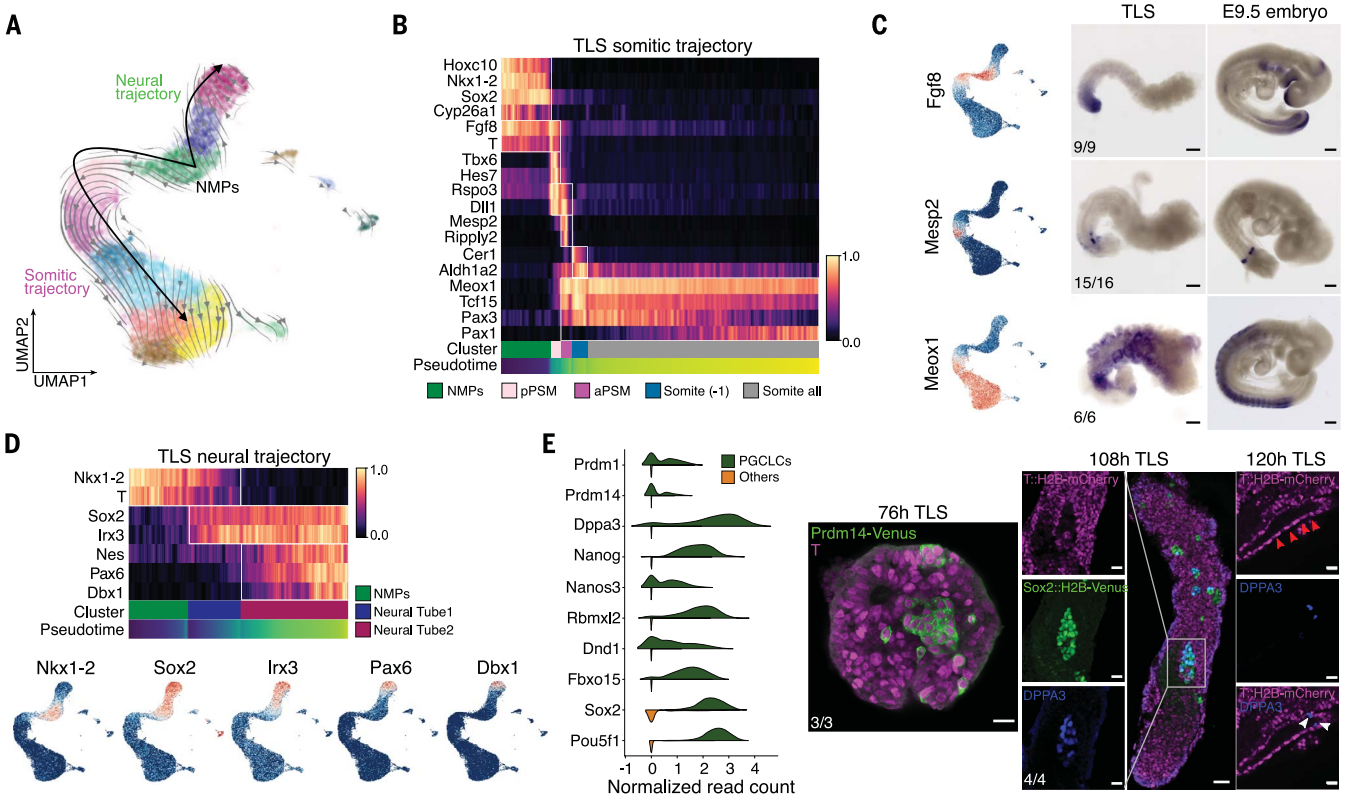


Fig. 4. TLS differentiation trajectories. (**A**) UMAP colored by identified clusters with trajectories inferred from RNA velocity. Gray arrow flows represent calculated velocity trajectories. (**B**) Heatmap with scaled expression of genes involved in somitogenesis measured in 5966 cells from 120-hour TLSs rooted in NMPs and ordered by pseudotime. (**C**) UMAP colored by expression of indicated genes (left) and whole-mount in situ hybridization for the same genes in TLS and E9.5 embryos (right). Numbers indicate the fraction of TLSs with embryo-like expression. Scale bars, TLS, 100 μm; embryo, 200 μm. (**D**) Heatmap with scaled expression of genes involved in neural development measured in 3462 cells from 120-hour TLSs rooted in NMPs and ordered by pseudotime (top) and UMAP colored by expression of indicated genes (bottom). (**E**) Split violin plots showing expression of marker genes for primordial germ cell–like cells (PGCLCs, left) and confocal sections of TLS showing PGCLC specification dynamics: T/PRDM14^{VE} double positive at 76 hours, SOX2$^{VE\text{-}high}$/DPPA3^{+} PGCLCs at 108 hours, and DPPA3^{+} cells in close contact with the T^{mCH+} gut-like structure at 120 hours. At 120 hours, Sox2$^{VE\text{-}high}$ cells in contact with FOXA2^{+} gut-like domain or DPPA3^{+} cells in contact with T^{mCH+} gut-like domain were observed in seven of nine TLSs (see also fig. S14D). Scale bars, 20 μm for 76-hour TLS, 50 μm for 108-hour TLS (20 μm for magnifications), 25 μm for 120-hour TLS. Red arrowheads, gut-like structure; white arrowheads, DPPA3^{+} PGCLCs.

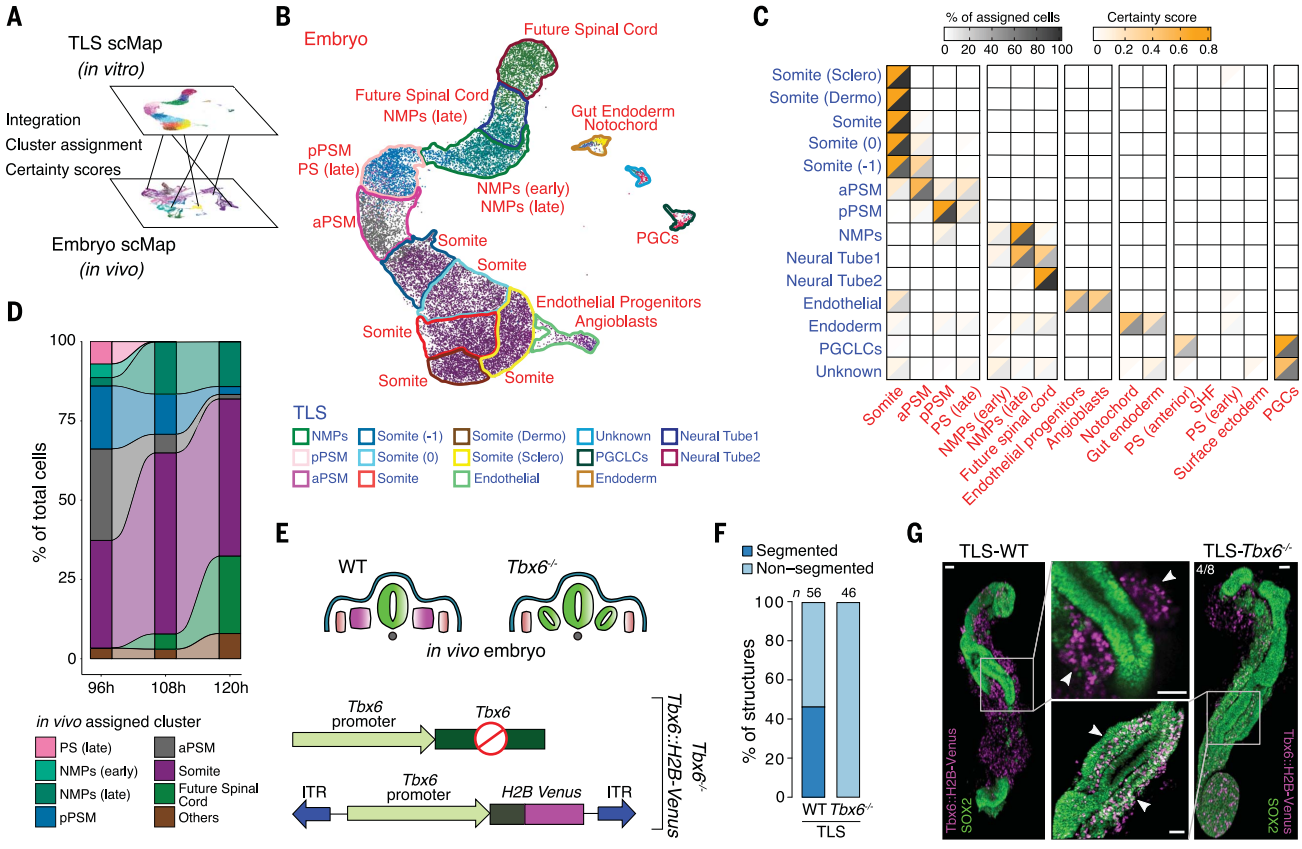


Fig. 5. TLS cell states are embryo-like and *Tbx6*$^{-/-}$ TLSs recapitulate the embryonic knockout phenotype. (**A**) Schematic of our comparative transcriptome analysis of TLSs with postoccipital E7.5 and E8.5 embryos at the single-cell level. (**B**) TLS UMAP colored by assigned embryonic cell states. TLS clusters are projected as corresponding colored contours. Blue font, TLS clusters; red font, embryo clusters. (**C**) Split heatmap with percentage of assigned cells (dark gray) and certainty score (orange) for TLS cells from the indicated cluster upon unbiased mapping to the in vivo counterpart. Font colors as in (B). (**D**) Alluvial plot of percentage of cells assigned to the indicated in vivo clusters in 96-, 108-, and 120-hour TLSs. (**E**) Simplified schematics of *Tbx6*$^{-/-}$ in vivo phenotype and knockout/reporter constructs. (**F**) Quantification of segmentation phenotype in TLS-*Tbx6*$^{-/-}$. Data represent three different experiments performed with two independent mESC lines of each genotype. (**G**) Formation of ectopic neural tubes in TLS-*Tbx6*$^{-/-}$. Ectopic neural tubes are identified as SOX2^{+}/Tbx6^{VE+} tubular structures flanking the main SOX2^{+}-only neural tube. Green, SOX2; magenta, Tbx6^{VE}. White arrowheads indicate Tbx6^{VE+} somites in the wild type (WT) and Tbx6^{VE+}/SOX2^{+} ectopic neural tubes in *Tbx6*$^{-/-}$. Scale bars, 50 μm.

assembly play an important role in somite and neural tube formation (*21*, *22*), we asked whether FN1, which is highly expressed in TLSs, was retained by Matrigel, as recently shown in 2D culture (fig. S7C) (*23*). FN1 protein clearly accumulated at the TLS-Matrigel interface; this was not observed in gastruloids, where FN1 likely diffuses into the medium (Fig. 2G, fig. S7D, and movie S10). Taken together, our data suggest that the activation of morphogenetic programs by Matrigel embedding is driven by ECM assembly involving integrins and fibronectin.

Single-cell RNA-seq demonstrates embryo-like dynamics of cell differentiation

We next focused on the TLS condition because it produced the most in vivo–like structures. After confirming reproducibility at the molecular level (fig. S8, A to C), we performed a time-resolved single-cell RNA-seq (scRNA-seq) analysis on a total of 20,294 postprocessed cells sampled from TLSs at 96, 108, and 120 hours (fig. S9A). Clustering analysis identified 14 different cell states. The larger clusters corresponded to derivatives of the PSM and NE that flank putative NMPs, whereas smaller clusters comprised endoderm, endothelial cells, and primordial germ cell–like cells (PGCLCs) (fig. S9B). The main clusters organized into a continuum of states recapitulating spatiotemporal features of the developing postoccipital embryo (Fig. 3A). Across the three time points sampled, progenitor cells gradually decreased in favor of more mature neural and somitic cells as development progressed (Fig. 3B and fig. S9C). Putative NMPs coexpressing T and SOX2, or T^{mCH}, Sox2^{VE}, and CDX2, were located at the posterior end at 96 and 120 hours (Fig. 3C and fig. S10, A to C) (*10*, *24*). TLS-NMPs thus display an in vivo–like NMP signature and have the highest differentiation potential, as they give rise to differentiating cells of both neural and mesodermal lineages (Fig. 3C and fig. S10, D to F).

RNA velocity analysis revealed neural and somitic trajectories rooted in the NMPs, further suggesting that the TLS recapitulates the developmental dynamics observed in the mid-gestational embryo (Fig. 4A and fig. S11A) (*25*). In vivo, NMPs and their descendants are arranged in order of progressive maturity along the posterior-to-anterior axis (*8*). Accordingly, ordering of TLS-derived cells along a pseudotemporal trajectory showed that the somitic trajectory reflects the genetic cascade observed in the embryo (Fig. 4B and fig. S12A). For example, the trajectory from Fgf8^{+} NMPs and PSM, through the determination front marked by Mesp2, to Meox1^{+} somites, was faithfully recapitulated, and the embryo-like spatial arrangement was confirmed by whole-mount in situ hybridization (Fig. 4C) (*8*). Likewise, the genetic cascade from NMPs to neural progenitors reflected the in vivo differentiation path in space and time (Fig. 4D). Subclustering of the neural cells demonstrated that TLSs generate

both dorsal and ventral neural subtypes, with dorsal subtypes being more prevalent (fig. S12B) (*13*). The analysis of *Hox* gene expression at consecutive time points showed in vivo–like collinear activation, as described for gastruloids (fig. S12C) (*3*). To test whether TLS somites establish cell states segregated along the dorsal-ventral (D-V) and anterior-posterior (A-P) somite axes in vivo, we reclustered all somitic cells. At 96 hours, we detected two main groups corresponding to the Uncx^{+} posterior and Tbx18^{+} anterior somite domains, in line with the A-P polarity established during segmentation (fig. S13, A and B) (*6*, *8*). At 120 hours, we found distinct clusters of Pax3^{+} (putative dorsal dermomyotomal) and Pax1^{+} (putative ventral sclerotomal) cells, as well as a small cluster of Lbx1^{+}/Met^{+} putative migratory limb muscle precursors (fig. S13, C to F) (*8*, *26*). In addition, Scx^{+} syndetome cells were detected (fig. S13G), and Uncx and Tbx18 expression were anticorrelated (fig. S13H).

Primordial germ cell (PGC) specification in the embryo occurs between embryonic day (E) 6.0 and E6.5 via T-mediated activation of *Blimp1* and *Prdm14*, and at E7.5, nascent PGCs can be identified as a group of DPPA3^{+} cells in the posterior primitive streak, which later migrate along the hindgut to the gonads (*27*, *28*). We assigned PGCLC identity using marker genes characteristic for PGCs and identified their location in the TLS (Fig. 4E and fig. S14). At 76 hours, roughly corresponding to stage E6.5, we detected T/Prdm14^{VE}–coexpressing cell clusters (Fig. 4E and fig. S14, A and B). At 108 hours, we found a group of Sox2$^{VE\text{-}high}$ cells that coexpressed DPPA3 (Fig. 4E and fig. S14C). At 120 hours, Sox2$^{VE\text{-}high}$ cells were detected in contact with FOXA2^{+} cells, and DPPA3^{+} cells in contact with a T^{mCH+} gut-like epithelial structure (Fig. 4E and fig. S14D). These data show that TLSs contain cells displaying characteristics typical for PGCs.

TLSs display a high complexity of cell states that match their in vivo counterparts

Single-cell comparison of 120-hour gastruloids with 120-hour TLSs identified different proportions of the major cell states (fig. S15, A to C). A more refined analysis revealed a higher complexity of cell states in TLSs (fig. S16, A to E, fig. S17, A to F, and fig. S18, A to D), and expression of later (more posterior) *Hox* genes suggests development into more advanced trunk stages (fig. S18, E and F). The comparison of TLSs with TLS^{CL} showed that in the latter, (i) sclerotomal and more mature neural cells are virtually absent, and (ii) somitic as well as endothelial cell identities are altered (fig. S15, B and C, fig. S16, A and B, fig. S17, A to G, and fig. S18, A and B). Application of RNA velocity confirmed that in TLS^{CL}, NMPs are highly biased toward the mesodermal lineage, whereas contribution to the neural lineage is diminished relative to TLS-NMPs; this is further corroborated by up-regulation of posterior PSM and down-regulation of neural marker genes (fig. S15, D and E) (*8*, *13*).

To investigate how close the cellular states identified in TLSs resemble those in embryos, we mapped TLS single-cell transcriptomes to a scRNA-seq compendium of postoccipital embryonic cellular subtypes (Fig. 5A) (*29*). The data revealed globally high accordance of TLS and embryonic cell states including characteristic marker genes, and pairwise comparison of mapped clusters identified only a small fraction of differentially expressed genes (Fig. 5, B and C, and fig. S19, A to D). Of note, PS- and early NMP–like cells were exclusively present at 96 hours and were replaced by late NMP–like cells at 108 and 120 hours (Fig. 5D). Taken together, our scRNA-seq analyses demonstrate that the TLS executes gene regulatory programs in a spatiotemporal order resembling that of the embryo.

Knockout TLSs display the embryonic mutant phenotype

Finally, to explore the utility of TLSs further, we conducted a proof-of-concept experiment to test whether gene ablation would reproduce the embryonic mutant phenotype. In vivo, loss of *Tbx6* results in transdifferentiation of prospective PSM and subsequent formation of ectopic neural tubes at the expense of PSM and somites (Fig. 5E) (*30*, *31*). We deleted *Tbx6* from *Tbx6::H2B-Venus* (*Tbx6*^{Ve}) mESCs and generated TLSs, which clearly failed to form somites even upon treatment with CHIR or CHIR/LDN (Fig. 5, E and F, and fig. S20, A and B). Quantitative polymerase chain reaction analysis on fluorescence-activated cell sorter–purified Tbx6^{VE+} cells revealed up-regulation of neural markers at the expense of (P)SM markers in *Tbx6*$^{-/-}$ cells, thus recapitulating the in vivo phenotype at the molecular level (fig. S20C). Finally, whole-mount immunofluorescence analysis for SOX2 showed that *Tbx6*$^{-/-}$ TLSs generated ectopic Tbx6^{VE+} neural tubes, whereas gastruloids, TLS^{C}, and TLS^{CL} formed an excess of morphologically indistinct SOX2^{+} tissue (Fig. 5G and fig. S20, D to F).

Discussion

The TLS model provides a scalable, tractable, readily accessible platform for investigating the lineage decisions and morphogenetic processes that shape the mid-gestational embryo with high spatiotemporal resolution. Our results show that the TLS faithfully reproduces key features of postoccipital embryogenesis, including axial elongation with coordinated neural tube, gut, and somite formation as well as PGCLCs. Accordingly, genetic manipulation of the TLS faithfully reproduced the morphogenetic and molecular changes observed in vivo. Thus, the TLSs will enable deeper analysis of the ontogeny of mutant phenotypes and provide an additional tool for investigating morphogenetic mechanisms unavailable in vivo. We also envision that the TLS^{C} and TLS^{CL} models may become important for testing current concepts of somitogenesis—for instance, the hypothesis that somite size and shape are controlled by local cell-cell interactions (*9*).

Mechanistically, our data highlight a crucial role for the ECM surrogate in unlocking the potential of in vitro derived mESC aggregates, although future efforts will have to address the exact functional contribution of individual components and biophysical properties (fig. S21), possibly using modular synthetic 3D matrices (*7*, *32*). Alternatively, the single-cell expression catalog of TLSs and gastruloids can provide initial guidance for further exploration of cell-cell and cell-matrix interactions and their control of embryonic architecture (fig. S7, A and B, and figs. S22 and S23).

REFERENCES AND NOTES

1. M. N. Shahbazi, E. D. Siggia, M. Zernicka-Goetz, Self-organization of stem cells into embryos: A window on early mammalian development. *Science* **364**, 948–951 (2019). doi: 10.1126/science.aax0164; pmid: 31171690
2. M. N. Shahbazi, M. Zernicka-Goetz, Deconstructing and reconstructing the mouse and human early embryo. *Nat. Cell Biol.* **20**, 878–887 (2018). doi: 10.1038/s41556-018-0144-x; pmid: 30038253
3. L. Beccari *et al*., Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids. *Nature* **562**, 272–276 (2018). doi: 10.1038/s41586-018-0578-0; pmid: 30283134
4. I. Matsuo, R. Hiramatsu, Mechanical perspectives on the anterior-posterior axis polarization of mouse implanted embryos. *Mech. Dev.* **144A**, 62–70 (2017). doi: 10.1016/j.mod.2016.09.002; pmid: 27697519
5. H. K. Kleinman, D. Philp, M. P. Hoffman, Role of the extracellular matrix in morphogenesis. *Curr. Opin. Biotechnol.* **14**, 526–532 (2003). doi: 10.1016/j.copbio.2003.08.002; pmid: 14580584
6. S. C. van den Brink *et al*., Single-cell and spatial transcriptomics reveal somitogenesis in gastruloids. *Nature* **582**, 405–409 (2020). doi: 10.1038/s41586-020-2024-3; pmid: 32076263
7. J. A. Brassard, M. P. Lutolf, Engineering Stem Cell Self-organization to Build Better Organoids. *Cell Stem Cell* **24**, 860–876 (2019). doi: 10.1016/j.stem.2019.05.005; pmid: 31173716
8. J. Chal, O. Pourquié, Making muscle: Skeletal myogenesis *in vivo* and *in vitro*. *Development* **144**, 2104–2122 (2017). doi: 10.1242/dev.151035; pmid: 28634270
9. A. S. Dias, I. de Almeida, J. M. Belmonte, J. A. Glazier, C. D. Stern, Somites without a clock. *Science* **343**, 791–795 (2014). doi: 10.1126/science.1247575; pmid: 24407478
10. F. Koch *et al*., Antagonistic Activities of Sox2 and Brachyury Control the Fate Choice of Neuro-Mesodermal Progenitors. *Dev. Cell* **42**, 514–526.e7 (2017). doi: 10.1016/j.devcel.2017.07.021; pmid: 28826820
11. O. Pourquié, The segmentation clock: Converting embryonic time into spatial pattern. *Science* **301**, 328–330 (2003). doi: 10.1126/science.1085887; pmid: 12869750
12. J. Chal *et al*., Differentiation of pluripotent stem cells to muscle fiber to model Duchenne muscular dystrophy. *Nat. Biotechnol.* **33**, 962–969 (2015). doi: 10.1038/nbt.3297; pmid: 26237517
13. A. Sagner, J. Briscoe, Establishing neuronal diversity in the spinal cord: A time and a place. *Development* **146**, dev182154 (2019). doi: 10.1242/dev.182154; pmid: 31767567
14. J. von Maltzahn, N. C. Chang, C. F. Bentzinger, M. A. Rudnicki, Wnt signaling in myogenesis. *Trends Cell Biol.* **22**, 602–609 (2012). doi: 10.1016/j.tcb.2012.07.008; pmid: 22944199
15. M. Rowton *et al*., Regulation of mesenchymal-to-epithelial transition by PARAXIS during somitogenesis. *Dev. Dyn.* **242**, 1332–1344 (2013). doi: 10.1002/dvdy.24033; pmid: 24038871

16. L. Durbin *et al*., Eph signaling is required for segmentation and differentiation of the somites. *Genes Dev.* **12**, 3096–3109 (1998). doi: 10.1101/gad.12.19.3096; pmid: 9765210
17. C. Marcelle, M. R. Stark, M. Bronner-Fraser, Coordinate actions of BMPs, Wnts, Shh and noggin mediate patterning of the dorsal somite. *Development* **124**, 3955–3963 (1997). pmid: 9374393
18. D. Jülich *et al*., Cross-Scale Integrin Regulation Organizes ECM and Tissue Topology. *Dev. Cell* **34**, 33–44 (2015). doi: 10.1016/j.devcel.2015.05.005; pmid: 26096733
19. J. H. Shawky, L. A. Davidson, Tissue mechanics and adhesion during embryo development. *Dev. Biol.* **401**, 152–164 (2015). doi: 10.1016/j.ydbio.2014.12.005; pmid: 25512299
20. S. Koshida *et al*., Integrinalpha5-dependent fibronectin accumulation for maintenance of somite boundaries in zebrafish embryos. *Dev. Cell* **8**, 587–598 (2005). doi: 10.1016/j.devcel.2005.03.006; pmid: 15809040
21. E. Nikolopoulou, G. L. Galea, A. Rolo, N. D. E. Greene, A. J. Copp, Neural tube closure: Cellular, molecular and biomechanical mechanisms. *Development* **144**, 552–566 (2017). doi: 10.1242/dev.145904; pmid: 28196803
22. G. G. Martins *et al*., Dynamic 3D cell rearrangements guided by a fibronectin matrix underlie somitogenesis. *PLOS ONE* **4**, e7429 (2009). doi: 10.1371/journal.pone.0007429; pmid: 19829711
23. J. Lu *et al*., Basement membrane regulates fibronectin organization using sliding focal adhesions driven by a contractile winch. *Dev. Cell* **52**, 631–646.e4 (2020). doi: 10.1016/j.devcel.2020.01.007; pmid: 32004443
24. S. Amin *et al*., Cdx and T Brachyury Co-activate Growth Signaling in the Embryonic Axial Progenitor Niche. *Cell Rep.* **17**, 3165–3177 (2016). doi: 10.1016/j.celrep.2016.11.069; pmid: 28009287
25. G. La Manno *et al*., RNA velocity of single cells. *Nature* **560**, 494–498 (2018). doi: 10.1038/s41586-018-0414-6; pmid: 30089906
26. M. Buckingham *et al*., The formation of skeletal muscle: From somite to limb. *J. Anat.* **202**, 59–68 (2003). doi: 10.1046/j.1469-7580.2003.00139.x; pmid: 12587921
27. S. Aramaki *et al*., A mesodermal factor, T, specifies mouse germ cell fate by directly activating germline determinants. *Dev. Cell* **27**, 516–529 (2013). doi: 10.1016/j.devcel.2013.11.001; pmid: 24331926
28. M. Saitou, M. Yamaji, Primordial germ cells in mice. *Cold Spring Harb. Perspect. Biol.* **4**, a008375 (2012). doi: 10.1101/cshperspect.a008375; pmid: 23125014
29. S. Grosswendt *et al*., Epigenetic regulator function through mouse gastrulation. *Nature* **584**, 102–108 (2020). doi: 10.1038/s41586-020-2552-x; pmid: 32728215
30. D. L. Chapman, V. E. Papaioannou, Three neural tubes in mouse embryos with mutations in the T-box gene Tbx6. *Nature* **391**, 695–697 (1998). doi: 10.1038/35624; pmid: 9490412
31. D. Concepcion *et al*., Cell lineage of timed cohorts of *Tbx6*-expressing cells in wild-type and *Tbx6* mutant embryos. *Biol. Open* **6**, 1065–1073 (2017). doi: 10.1242/bio.026203; pmid: 28606934
32. A. Ranga *et al*., Neural tube morphogenesis in synthetic 3D microenvironments. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E6831–E6839 (2016). doi: 10.1073/pnas.1603529113; pmid: 27742791

ACKNOWLEDGMENTS

We are grateful for the support and feedback received from members of the Herrmann and Meissner laboratories, in particular S. Grosswendt, Z. Smith, and A. Taguchi. We thank D. Micic and J Fiedler for animal care; N. Mages for assistance with (sc)RNA-seq; C. Giesecke-Thiel and U. Marchfelder for assistance with FACS; T. Mielke and B. Fauler for help with microscopy; S. Währisch, F. Tobor, I. Gassaloglu, and P. Burton for technical assistance; and N. Rajewsky (MDC/BIMSB) for providing access to the NanoString. **Funding:** Supported by an Alexander von Humboldt Fellowship (J.V.V.), NIH grant HG006193 (A.M.), and the Max Planck Society. **Author contributions:** B.G.H. initiated the study; J.V.V. and B.G.H. conceived the project; J.V.V., A.M., and B.G.H. supervised the project; J.V.V., A.B., and L.H. designed, performed, and quantified most experiments; H.K. performed bulk and scRNA-seq computational analysis with help from J.V.V. and A.B.; M.S.-W., D.S., F.K., M.P., and A.S.K. generated and derived the mESC reporter lines; L.W. performed tetraploid complementation; S.H. performed pilot experiments to optimize culture media; R.B. helped with image acquisition and performed image analysis with L.G., A.B., and J.V.V.; B.T. supervised next-generation sequencing; J.V.V. drafted the first version of the manuscript; and the final manuscript was written by J.V.V., A.B., A.M., and B.G.H. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. Bulk and scRNA-seq data have been deposited in the Gene Expression Omnibus (GEO) under accession code GSE141175. Computational code is available at https://github.com/HeleneKretzmer/TLS and https://github.com/leoguignard/TLS-morpho.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/370/6522/eaba4937/suppl/DC1
Materials and Methods
Figs. S1 to S24
Movies S1 to S10
Data S1 to S10
References (*33*–*71*)
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

9 December 2019; resubmitted 13 May 2020
Accepted 24 September 2020
10.1126/science.aba4937