Stem Cell Reports, Volume *17*

Supplemental Information

Reconstruction of dynamic regulatory networks reveals signaling-induced topology changes associated with germ layer specification

Emily Y. Su, Abby Spangler, Qin Bian, Jessica Y. Kasamoto, and Patrick Cahan

Supplemental Items

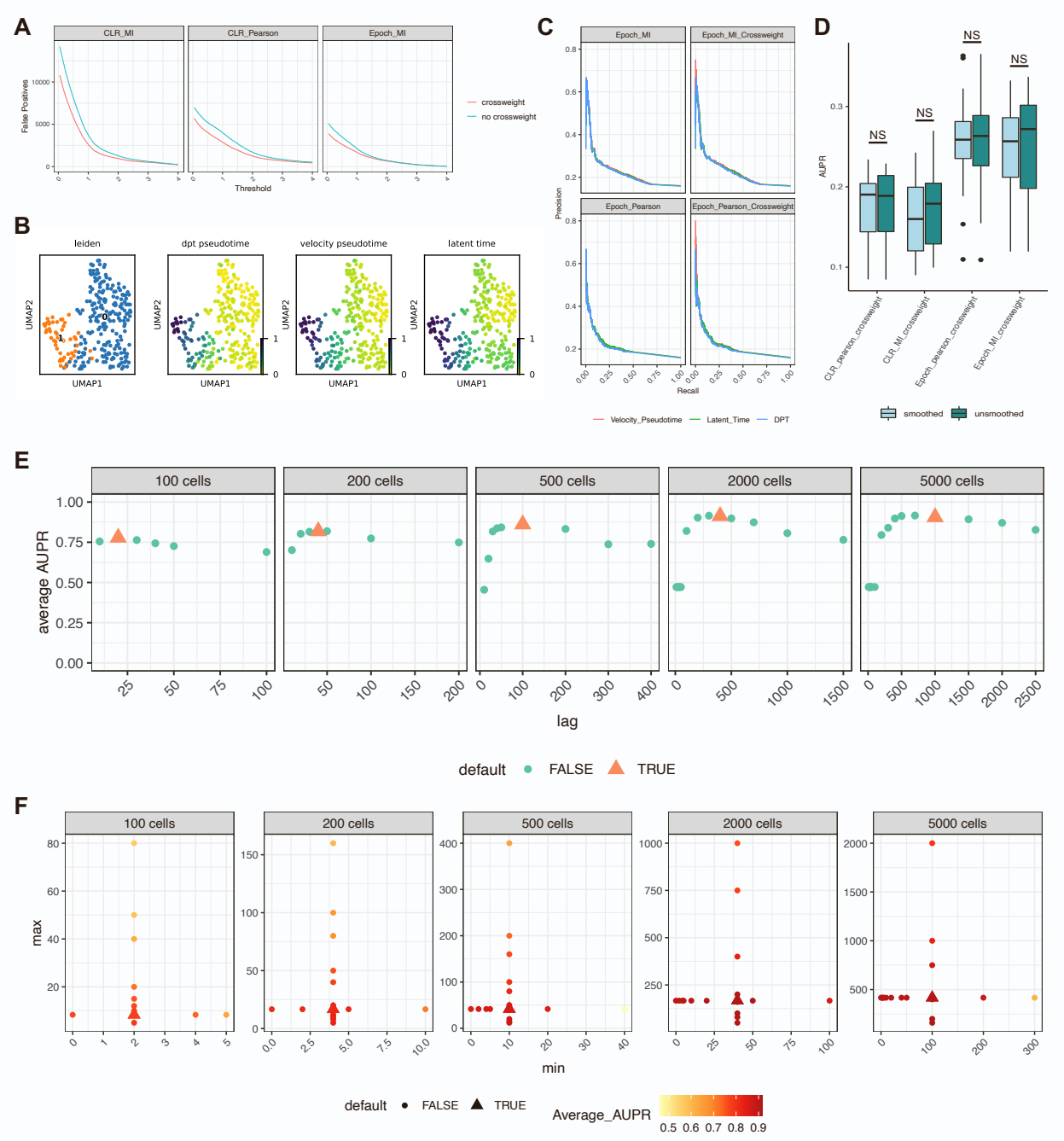

Figure S1. The effect of cross-weighting, TI method, smoothing, and parameter choices on performance.
(A) An example from GRN reconstruction on a synthetic dataset showing the difference in number of false positives based on whether or not cross-weighting is performed. See Note S1 for information on synthetic dataset. (B) Clustering, diffusion pseudotime, velocity pseudotime, and latent time of the E12.5 muscle development data. See Note S1 for dataset details. (C) Precision recall curves of four variations of Epoch using three different pseudotime annotations on the E12.5 muscle development data. There was no clear difference in performance based on pseudotime method used. (D) Effect of smoothing vs. no smoothing on AUPR of reconstruction using various methods. There was no significant difference in performance between using kernel-smoothed data and unsmoothed data. Smoothing functions remain in Epoch's code

to aid in cleaner visualizations of gene dynamics. (E) Average AUPR of reconstruction using different cross-weighting lag times across datasets of varying sizes, taken from BEELINE linear datasets. Each dataset size contains 10 distinct datasets, Default lag time is plotted as the orange triangle, and is computed as one-fifth of the number of cells in the data. (F) Average AUPR of reconstruction using different minimum and maximum windows in cross-weighting. Data shown is reconstruction from BEELINE linear datasets. Each dataset size contains 10 distinct datasets. Default minimum and maximum is indicated by the triangle, and is defined as 1/50 and 1/12 of total cells respectively. High average AUPR is indicated by the darker red colors.

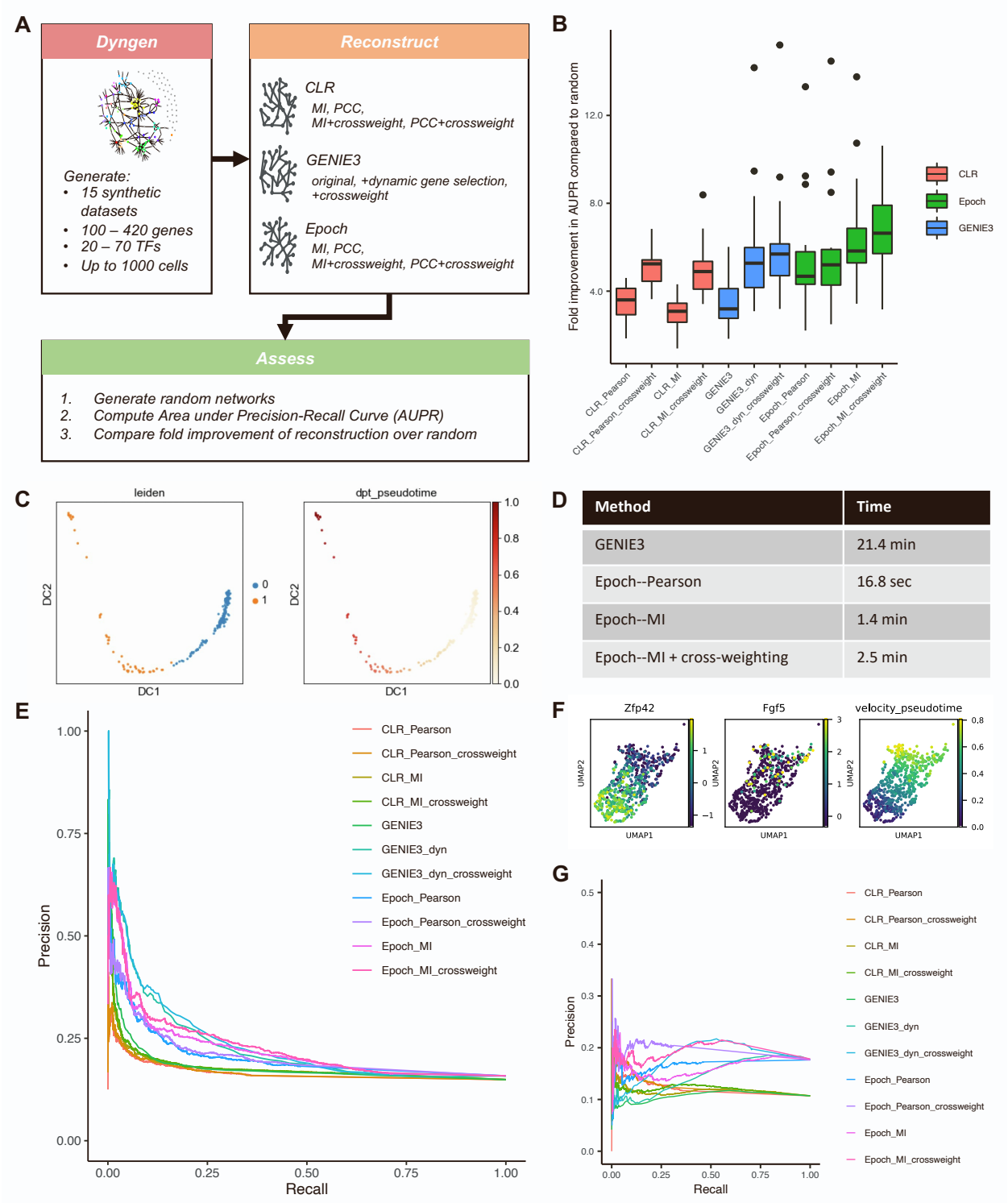

Figure S2. Epoch benchmarking, related to Note S1.
(A) Synthetic dataset generation via Dyngen, reconstruction, and assessment. (B) Full results of benchmarking reconstruction of synthetic data. Four versions of CLR (Pearson vs. MI, cross-weighting vs. no cross-weighting), four versions of Epoch (same as CLR), and three versions of GENIE3 (original, limited to dynamic genes, and with cross-weighting) are compared. Kruskal-Wallis $p=1.75 \times 10^{-11}$. (C) Clustering and diffusion pseudotime for E12.5 muscle development data used in benchmarking. (D) Runtimes for reconstruction on muscle development data. (Runtimes were measured on a MacBook Pro

laptop with 2.9 GHz Intel Core i7 processor, 16GB 2133 MHz DDR3 of RAM). (E) Full results of benchmarking reconstruction of the muscle development data. (F) Early mESC directed differentiation data (subsetted from Spangler et al. 2018). Zfp42 expression, Fgf5 expression, and velocity pseudotime are shown. (G) Full results of benchmarking reconstruction of the early Spangler et al. data.

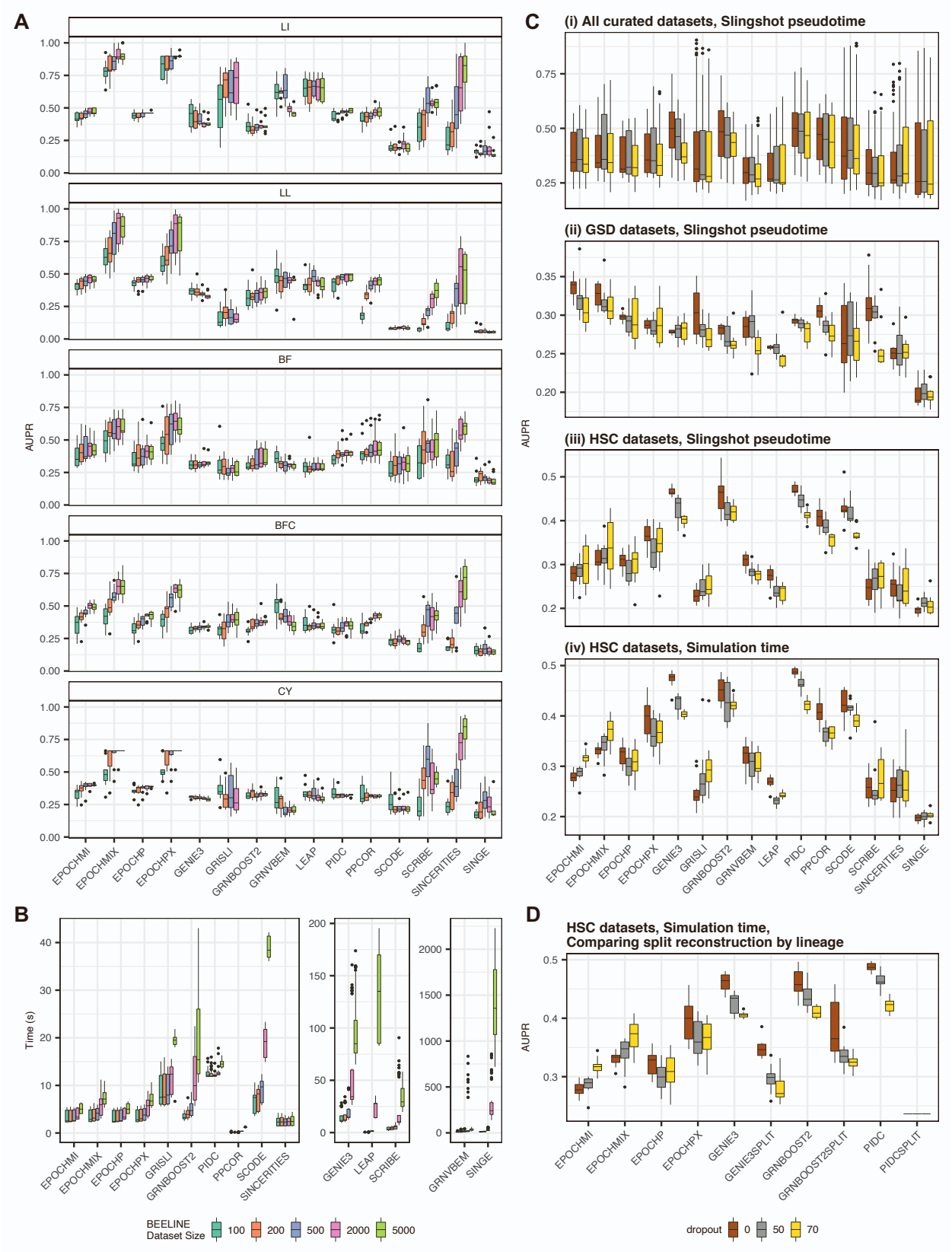

Figure S3. BEELINE data benchmarking, related to Figure 2 and Note S1.
(A) AUPR of reconstruction on synthetic data using the four versions of Epoch (MI = mutual information, P

= Pearson correlation, X = with cross-weighting) and 11 other single-cell GRN reconstruction tools. Results are faceted by trajectory type (LI = linear, LL = long linear, BF = bifurcating, BFC = bifurcating converging, CY = cycle) and colored by number of cells in the dataset. (B) BEELINE synthetic data runtimes. Runtimes (in seconds) are plotted for each method, faceted by magnitude of runtime. Four versions of Epoch are shown (MI = mutual information, P = Pearson correlation, X = with cross-weighting), along with 11 other single-cell GRN reconstruction tools. Data is colored based on number of cells in the dataset. (C) AUPR of reconstruction on curated data (as described in Pratapa et al. 2020), using the four versions of Epoch (MI = mutual information, P = Pearson correlation, X = with cross-weighting) and 11 other single-cell GRN reconstruction tools. Panel (i) summarizes the results from all curated datasets. Panel (ii) shows results from reconstruction on the GSD datasets using BEELINE-provided Slingshot pseudotime. Panel (iii) shows results from reconstruction on the HSC datasets using BEELINE-provided Slingshot pseudotime. Panel (iv) shows results from reconstruction on re-simulated HSC datasets using true simulation time from BoolODE (Pratapa et al., 2020). (D) The results from reconstruction on re-simulated HSC datasets using true simulation time, comparing the results of reconstructing using GENIE3, GRNBoost2, and PIDC with no splitting vs. splitting along each lineage before compiling (SPLIT = with splitting and compiling).

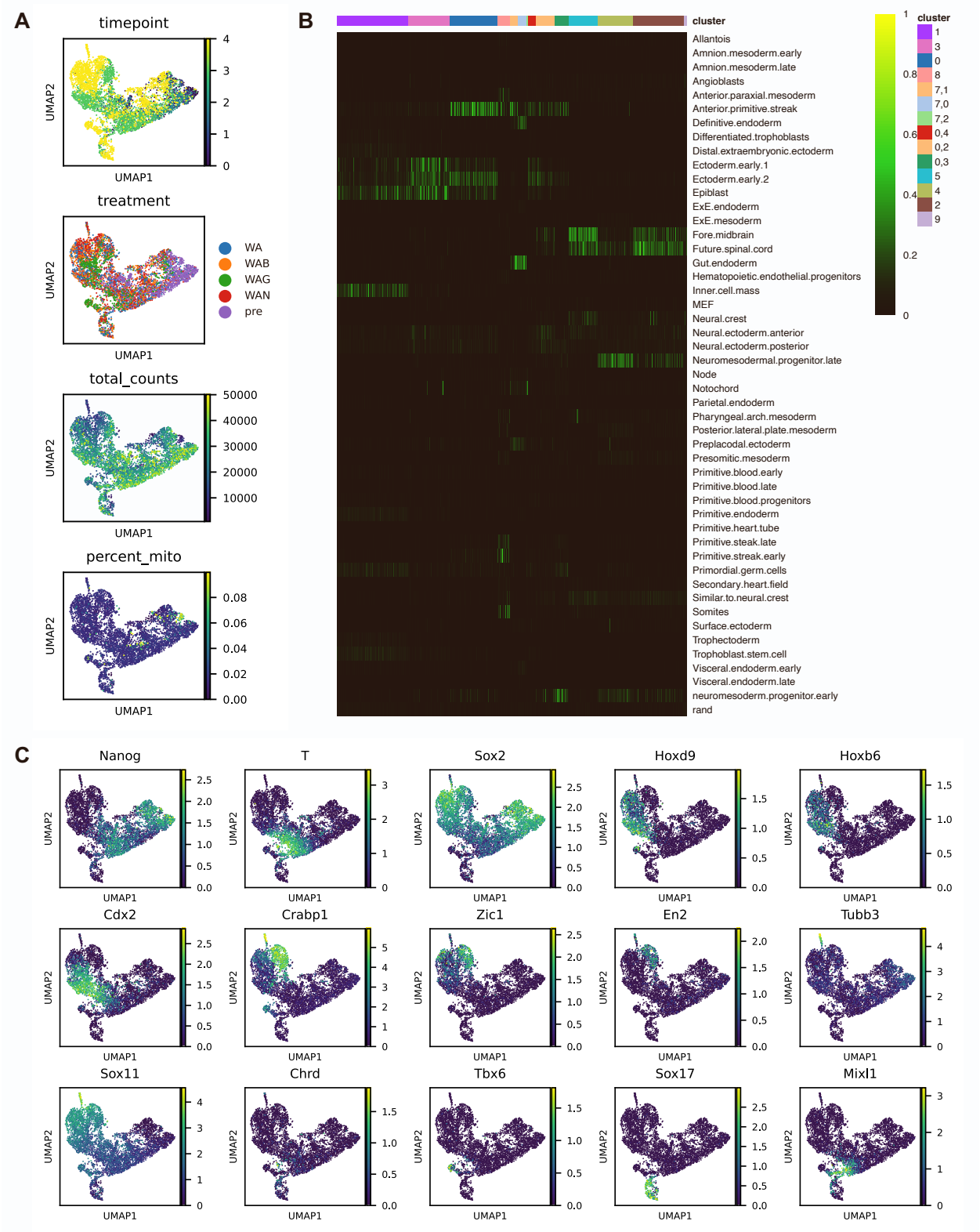

Figure S4. mESC directed differentiation cell annotations, related to Figure 3, Note S2, Table S1. (A) timepoint, treatment, total counts, and percent mitochondrial transcripts content of the MULTI-seq data. (B) SingleCellNet classification results. The classifier was trained from mouse embryo data found in literature, curated by members in the lab. (C) Select marker gene expression of the MULTI-seq data.

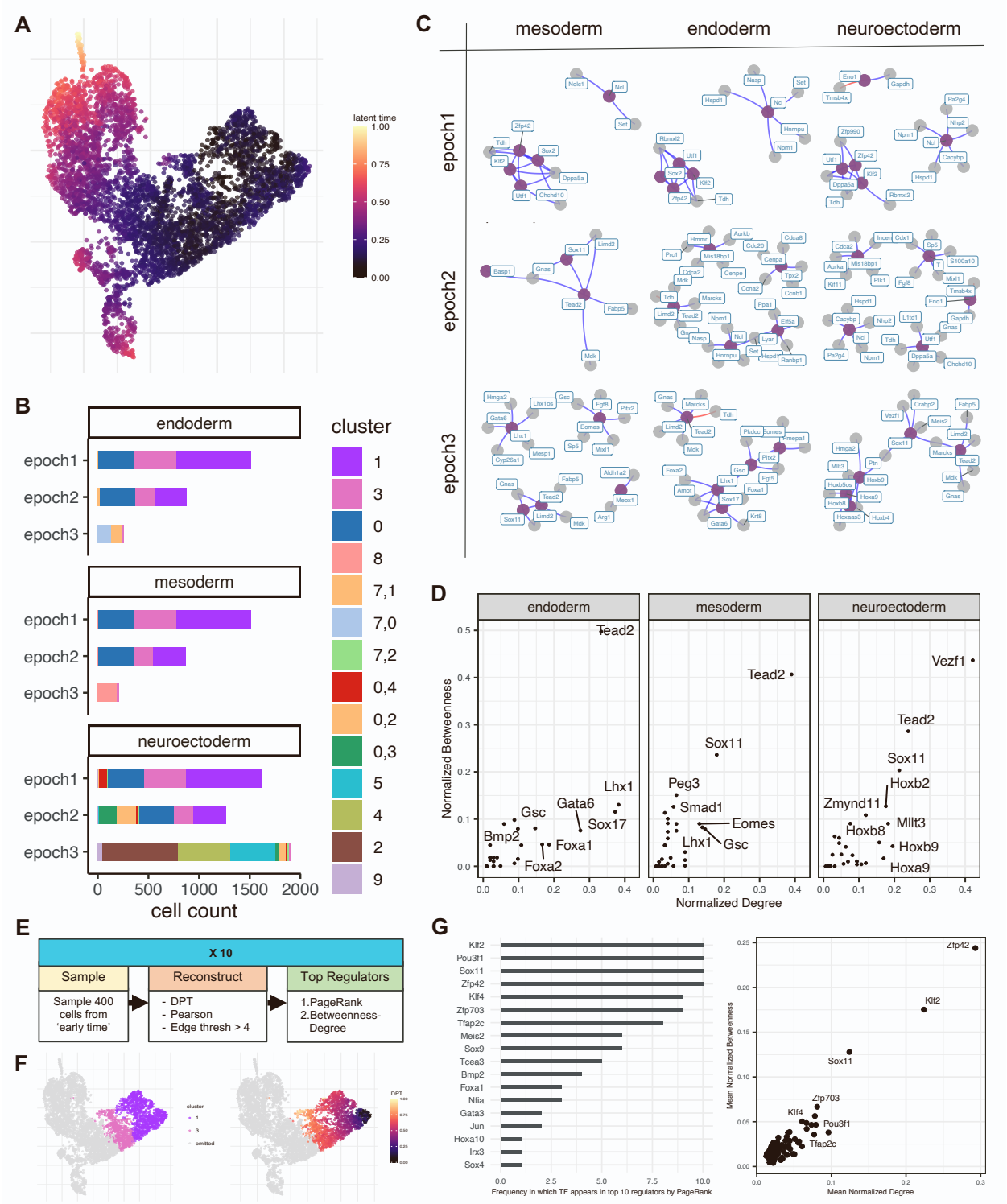

Figure S5. *in vitro* directed differentiation network reconstruction, related to Note S3.
(A) Latent time annotation for MULTI-seq data. (B) Population breakdown by cluster of epoch assignment along each germ layer path. (C) Reconstructed dynamic networks along mesoderm, endoderm, and neuroectoderm paths, with top 5 regulators colored purple and their top targets in gray as determined by PageRank. (D) Top regulators as predicted by Epoch via betweenness and degree for the third epochs along each germ layer path. (E) Workflow for bootstrapped top regulator prediction for inner cell mass to

epiblast-like transition. (F) Only cells from early time were used in this early bootstrap analysis. Clusters used and corresponding diffusion pseudotime shown here. (G) Top regulators for this early fate transition as predicted using frequency a TF scores as a top regulator by PageRank and using mean Betweenness-Degree.

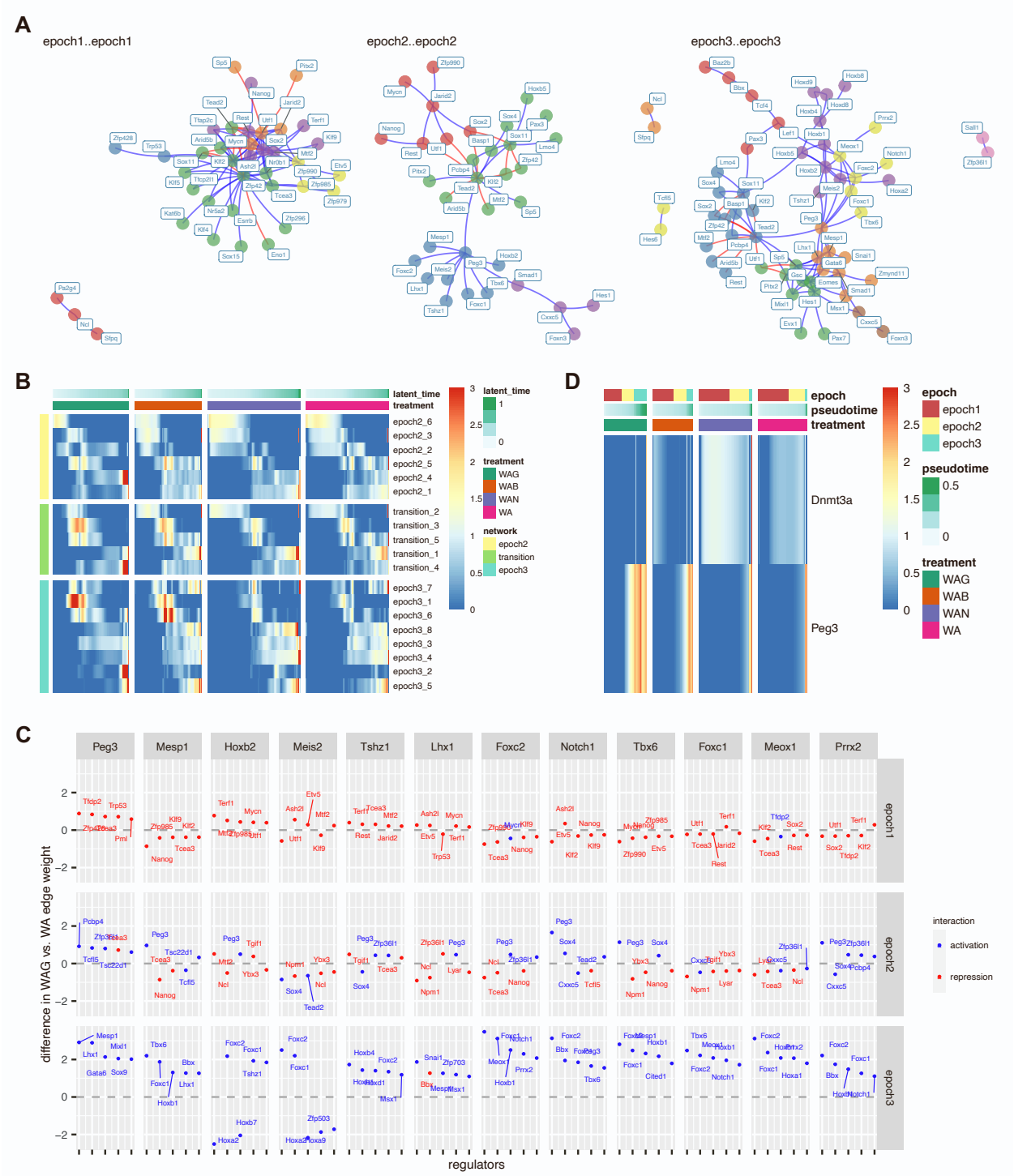

Figure S6. Mesodermal network analysis continued, related to Figure 4 and results section "*Peg3* is a central regulator in mesodermal WAG networks".

(A) Full dynamic mesodermal network. TFs are colored by community. Blue and red edges represent activating and repressive edges respectively. (B) Average community expression over time by treatment along the full dataset. Communities shown (each row) are from the epoch 2 subnetwork (yellow), transition (green), and epoch 3 subnetworks (aqua) of the mesodermal network. (C) Top differential regulators of mesodermal genes. Blue and red represent activators and repressors respectively. Y-axis is plotting difference in edge weights between the WAG and WA networks. Thus, TFs appearing above the

0-line represent interactions that are more unique to the WAG network, and those below are more unique to the WA network. (D) *Peg3* and *Dnmt3a* expression across time along the mesodermal path of the MULTI-seq data.

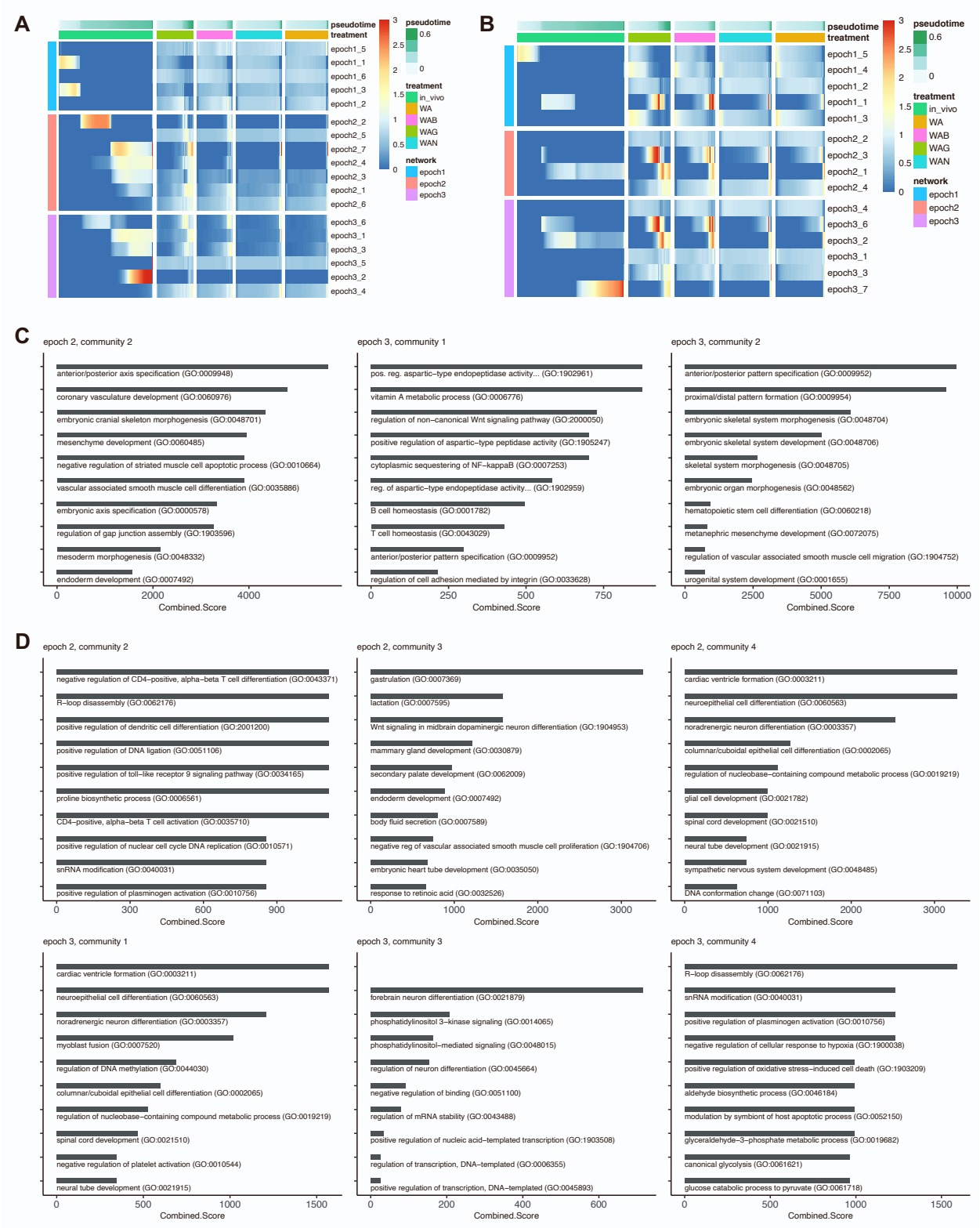

Figure S7. *In vivo* vs. *in vitro* mesodermal network comparison continued, related to Figure 7. (A) Average *in vivo* community expression over time along the mesodermal path for the *in vivo*, *in vitro* WAG, *in vitro* WAB, *in vitro* WAN, and *in vitro* WA data. (B) Average *in vitro* community expression over time along the mesodermal path for the *in vivo*, *in vitro* WAG, *in vitro* WAB, *in vitro* WAN, and *in vitro* WA

data. (C) Top ten enriched terms for *in vivo*-specific modules based on the Combined Score from Enrichr GSEA analysis. The three modules included are activated in the *in vivo* data but not (or weakly activated) in the *in vitro* data (see panel A). (D) Top ten enriched terms for *in vitro*-specific modules based on the Combined Score from Enrichr GSEA analysis. The six modules included are activated in the *in vitro* data but not (or weakly activated) in the *in vivo* data (see panel B).

Table S1. Cell type annotation by cluster of mESC directed differentiation data, related to Figure 3, S4, and Note S2.

Cluster	Cell Type Label	Evidence
1	Undifferentiated ESCs	This cluster classifies strongly as inner cell mass and epiblast. It expresses a number of pluripotency genes including *Zfp42*.
3	Epiblast	This cluster classifies as epiblast with some early ectodermal signature. *Fgf5* expression begins here.
0	Primitive Streak	This cluster classifies strongly as primitive streak with some early ectodermal signature. It expresses *Brachyury*.
8	(mesoderm) Presomitic Mesoderm/Somites	This cluster classifies as presomitic mesoderm and somites, with some residual primitive streak. It expresses a number of mesodermal markers including *Mesp1*.
7,1	(endoderm)	7,0 exhibits strong classification for definitive endoderm and gut endoderm. It expresses definitive endodermal markers including *Foxa2* and *Sox17*. RNA velocity suggests 7,1 cells transition into 7,0, and largely classifies as anterior primitive streak and endoderm.
7,0		
7,2	Node	This cluster strongly classifies as node, and expresses *Chordin*.
0,4	(neuroectoderm) Early Ectoderm	This cluster classifies weakly as a number of ectodermal cell types, but based on RNA velocity seems to be most similar to early ectodermal cells.
0,2	(neuroectoderm) Early Ectoderm, Forebrain/ Midbrain fated	This cluster classifies weakly as a number of ectodermal cell types including early ectoderm. Based on RNA velocity, this cluster transitions into cluster 5 (forebrain/midbrain fated).
5	(neuroectoderm) Future Brain, Forebrain/ Midbrain	This cluster classifies strongly as forebrain/midbrain with some weaker future spinal cord classification. It expresses *Zic1* and *En2*.
9	(neuroectoderm) Future brain	This cluster classifies weakly as future brain, future spinal cord, and neural crest. It expresses *Crabp1* and *Tubb3*, and RNA velocity suggests it transitions from clusters 2 or 5.
0,3	Early Neuromesodermal Progenitor	This cluster classifies weakly as a number of cell types including primitive streak and early ectoderm, but more strongly as early neuromesodermal progenitor. RNA velocity shows cells in this cluster transition into a neuromesodermal progenitor. Cells in this cluster express *T* and some *Sox2*.
4	Neuromesodermal Progenitor	This cluster classifies strongly as neuromesodermal progenitor (late and early). It has some *T* and *Sox2* expression.
2	(neuroectoderm) Future Spinal Cord	This cluster classifies mostly as future spinal cord, and weakly as fore/midbrain. It also expresses many patterning genes from the Hox family indicative of spinal cord fated cells such as *Hoxb9*.

Table S2. Top regulators by PageRank in the mESC directed differentiation data, related to Note S3.
Top 10 regulators in each epoch network along each germ layer path are shown.

mesoderm		endoderm		neuroectoderm	
Gene	PageRank	Gene	PageRank	Gene	PageRank
epoch1		epoch1		epoch1	
Zfp42	0.150	Zfp42	0.105	Zfp42	0.087
Klf2	0.095	Klf2	0.072	Klf2	0.072
Sox2	0.049	Sox2	0.043	Utf1	0.070
Ncl	0.047	Ncl	0.039	Zfp990	0.039
Utf1	0.045	Utf1	0.037	Zfp985	0.026
Mtf2	0.022	Lyar	0.037	Ncl	0.024
Tcea3	0.011	Pa2g4	0.028	Lyar	0.024
Trp53	0.011	Mtf2	0.022	Rest	0.022
Jarid2	0.011	Trp53	0.018	Mtf2	0.019
Eno1	0.011	Tcea3	0.016	Pa2g4	0.018
epoch2		epoch2		epoch2	
Tead2	0.209	Tead2	0.103	Utf1	0.151
Sox11	0.072	Mis18bp1	0.084	Mis18bp1	0.048
Peg3	0.044	Cenpa	0.070	T	0.036
Cenpa	0.037	Ncl	0.039	Rest	0.030
Basp1	0.032	Lyar	0.036	Mtf2	0.028
Pcbp4	0.020	Pa2g4	0.026	Lyar	0.025
Cxxc5	0.019	Phf5a	0.020	Ncl	0.024
Sfpq	0.019	Sap30	0.019	Terf1	0.019
Jarid2	0.017	Jarid2	0.014	Pa2g4	0.017
Tcfl5	0.017	Sox11	0.013	Nanog	0.017
epoch3		epoch3		epoch3	
Tead2	0.134	Tead2	0.139	Vezf1	0.082
Sox11	0.042	Lhx1	0.047	Tead2	0.070
Bbx	0.039	Sox17	0.046	Sox11	0.041
Lhx1	0.027	Gata6	0.027	Hoxb9	0.029
Foxn3	0.026	Pitx2	0.024	Zmynd11	0.027
Basp1	0.020	Gsc	0.021	Hoxb2	0.026
Ndn	0.018	Ndn	0.017	Hoxa9	0.025
Gsc	0.017	Foxa1	0.017	Mllt3	0.024
Peg3	0.016	Foxa2	0.016	Hoxb8	0.022
Eomes	0.015	Basp1	0.015	Meis2	0.019

Table S3. Early functional differences in *in vitro* vs. *in vivo* networks, related to Figure 7 and S7. Enriched terms for *in vitro*-specific modules based on the Combined Score from Enrichr GSEA analysis, specific to the first epoch. These three communities belong to the first epoch and are activated in the *in vitro* data but not (or weakly activated) in the *in vivo* data (see **Fig S7**).

Term	Combined.Score
epoch 1, community 2	
positive regulation of stem cell proliferation (GO:2000648)	773.48
regulation of mRNA splicing, via spliceosome (GO:0048024)	203.75
epoch 1, community 3	
glutathione metabolic process (GO:0006749)	405.99
epoch 1, community 4	
cellular response to laminar fluid shear stress (GO:0071499)	1,027.54
myoblast fusion (GO:0007520)	859.88
response to laminar fluid shear stress (GO:0034616)	584.21
regulation of DNA methylation (GO:0044030)	584.21
cellular response to fluid shear stress (GO:0071498)	444.11
male gonad development (GO:0008584)	182.23
development of primary male sexual characteristics (GO:0046546)	182.23
regulation of protein metabolic process (GO:0051246)	155.19
gonad development (GO:0008406)	142.01

Note S1. Benchmarking Epoch against existing GRN reconstruction tools

To assess the performance of Epoch in reconstructing static networks, we benchmarked it on synthetically generated data. We generated fifteen synthetic datasets using the Dyngen R package (Cannoodt et al., 2021). Networks were designed to encompass varying regulatory motifs and spanned 100-420 genes including 20-70 transcription factors. Synthetic single-cell datasets generated from these networks varied in size, from as low as 20 to as high as 1000 cells. We compared Epoch to the original versions of GENIE3 (Huynh-Thu et al., 2010), which is the GRN reconstruction engine of SCENIC (Aibar et al., 2017), and CLR (Faith et al., 2007) as well as variations on these methods that incorporated them within the Epoch framework. Thus, we assessed a total of 11 methods: CLR using mutual information (original CLR), CLR using Pearson correlation, CLR using mutual information with cross-weighting, CLR using Pearson correlation with cross-weighting, GENIE3 (original GENIE3), GENIE3 limited to dynamically expressed genes, GENIE3 limited to dynamically expressed genes with cross-weighting, Epoch with Pearson correlation (with and without cross-weighting), and Epoch with mutual information (with and without cross-weighting).

We reconstructed individual networks with each method for each dataset, using normalized simulation time as the pseudotemporal annotation for Epoch. To assess performance, we compared each resulting network to a set of five random networks, unique to each reconstructed network, that were generated by permuting edge weights of targets for each transcription factor in the reconstructed network (**Fig S2A**). As a gold standard, all networks were compared against the original network used to generate the synthetic single-cell data. Based on fold improvement in area under the precision-recall curve over the random networks, all variations of Epoch outperformed original versions of CLR and GENIE3 (**Fig S2B**). Additionally, we found that Epoch using mutual information with or without cross-weighting resulted in higher fold improvement over random compared to any other method. The addition of cross-weighting improved median AUPR fold change over random for all methods. Further, limitation to dynamic genes (e.g. versions of CLR vs. corresponding versions of Epoch as well as GENIE3 vs. GENIE3 limited to dynamic genes) also improved performance. In all cases, it was apparent that key steps in the Epoch framework improved overall static reconstruction, as limiting reconstruction to dynamically expressed genes and applying the cross-weighting scheme both led to increased fold improvement in AUPR over random, demonstrating the utility of Epoch's framework.

We next compared Epoch to recently developed GRN reconstruction methods designed for single-cell data. We used the benchmarking platform, BEELINE (Pratapa et al., 2020), to assess Epoch's performance against 11 other single-cell methods across the synthetic datasets available within the tool (**Fig 2B, S3A**). Importantly, these covered a range of cell state trajectories (linear, long linear, bifurcating, bifurcating converging, cycle) as well as dataset sizes (100 to 5000 cells). Our results demonstrate that Epoch with cross-weighting (using either mutual information or Pearson correlation) performed best based on AUPR. Additionally, while SINCERITIES almost approaches Epoch's performance at the 5000-cell dataset size, it performs markedly worse with smaller datasets, unlike Epoch. Finally, additional comparison of computational runtimes demonstrated that Epoch has the additional benefit of running magnitudes more efficiently than other methods including GRNBOOST2, PIDC, and GENIE3 (**Fig S3B**).

We then compared Epoch against the same 11 single-cell GRN reconstruction methods across the curated datasets available in BEELINE (**Fig S3C**). Importantly these datasets, though synthetic, were simulated from literature-derived networks and pseudotime was computed using Slingshot (Street et al., 2018). In this case performance based on AUPR was more equal across the methods, with GENIE3, GRNBoost2, and PIDC slightly out-performing other methods (**Fig S3C i**). On closer inspection, it was clear that performance was dataset dependent. For example, Epoch and other pseudotime-based methods performed relatively better in reconstructing the GSD network vs the HSC network (**Fig S3C ii, iii**). The same phenomenon was present in the BEELINE publication where it was initially attributed to imprecision of trajectory inference in estimating pseudotime. To test this, we used the underlying literature derived-networks and re-simulated the datasets. We then used the true simulation time, rather than Slingshot-derived pseudotime, for reconstruction and benchmarking. Our results demonstrated that despite using the true simulation time, AUPR of pseudotime-based reconstruction methods was largely unchanged, suggesting that poor pseudotime computation was not at fault (**Fig S3C iv**).

Finally, we hypothesized that decreased performance in the pseudotime-based methods was largely driven by reconstruction on split trajectories. Namely, for pseudotime-based methods, reconstruction of branched data was performed along each lineage separately and resulting networks

were aggregated. For example, HSC datasets represented four distinct lineages (branches), and so the 2000 cell dataset was broken into roughly four 500 cell datasets for reconstruction by pseudotime-based methods before aggregation. For non-pseudotime-based methods, including GENIE3, GRNBoost2, and PIDC, reconstruction was done on the full 2000 cell dataset. To test the extent to which this impacted performance, we re-benchmarked GENIE3, GRNBoost2, and PIDC splitting the reconstruction along each lineage before aggregating, mirroring the pseudotime-based methods. The resulting performance was substantially decreased, suggesting that splitting reconstruction is a factor contributing to decreased performance (**Fig S3D**). Ultimately, other factors, such as network density and the presence of specific connectivity motifs, warrant further exploration, but was unfortunately infeasible with the limited data.

Next, while the results from assessment of networks reconstructed from synthetic data demonstrated Epoch's favorable performance as a network inference tool, we acknowledge that there are caveats associated with synthetic data, including limited stochasticity and reliance on simulations that may not encompass the full extent of potential cell state responses seen in real biological systems. Thus, we next turned our attention to benchmarking on real world data. To assess Epoch's performance, we began by assessing the networks reconstructed by the variations of Epoch, CLR and GENIE3 listed above from data of E12.5 mouse muscle development (**Fig S2C-E**). For this dataset, we used Diffusion based pseudotime (DPT) to compute pseudotime for Epoch reconstruction (Haghverdi et al., 2016). We used Chip-X data from the Enrichr database, which includes ChIP-Seq, ChIP-ChIP, ChIP-PET, DamID used to profile transcription factor binding to DNA, covering 220 TFs and >35,000 genes, as a gold standard network (Lachmann et al., 2010). Based on precision recall, we found that all variations of Epoch once again outperformed original versions of CLR and GENIE3. Interestingly, we found that the two variations of GENIE3 embedded within the Epoch framework performed best overall in this assessment, followed by the two variations of Epoch using mutual information (with and without cross-weighting). It was again apparent that limiting reconstruction to dynamically expressed genes and cross-weighting significantly improved overall reconstruction (**Fig S2E**).

To determine if these results were consistent across other real world datasets, we turned to single-cell transcriptomic data collected from day 4 *in vitro* mESC directed differentiation (Spangler et al., 2018). We limited our reconstruction to cells in earlier stages of differentiation that had not yet transitioned into any committed germ layer-like state, as determined by pseudotemporal analysis, RNA velocity analysis, expression of *Zfp42*, marking naive pluripotent cells, and expression of *Fgf5*, marking cells in a primed state (**Fig S2F**). The filtering ensured the trajectory remained linear, which, while absolutely not a requirement for Epoch reconstruction, was employed to minimize performance variability of trajectory inference methods in which branched structures may be more difficult to order. This allowed us to narrow our focus on reconstruction performance while minimizing impact from potentially inaccurate pseudotemporal ordering. Based on precision-recall, we found that Epoch variations using cross-weighting performed best, followed by GENIE3 embedded within the Epoch framework (**Fig S2G**). Overall, these benchmarking results demonstrate Epoch's utility as a reconstruction tool, and further demonstrate the usefulness of Epoch as a flexible framework that can improve reconstruction of other methods.

Note S2. Early *in vitro* mouse ESC directed differentiation cluster annotations

We annotated the *in vitro* mESC directed differentiation clusters based on gene marker expression, differential expression analysis, RNA velocity analysis, and singleCellNet classification (**Fig 3B-D, S4**). Undifferentiated cells populated cluster 1, which was positive for *Zfp42* expression, a pluripotency marker (Rogers et al., 1991). This population additionally classified as inner cell mass, further supporting their pluripotent annotation. RNA velocity indicated these cells transitioned into cluster 3, which marked the beginning of *Fgf5* expression, an epiblast maker (Haub and Goldfarb, 1991; Pelton et al., 2002). SCN revealed early ectoderm and epiblast signature for cluster 3, indicating to us that it was representative of a primed cell state. Cells in this cluster transitioned to cluster 0 and 0,4. Cells in cluster 0 were largely representative of primitive streak, as determined by *Brachyury* expression, with some early ectoderm and residual epiblast signature, as determined by SCN. Interestingly, cluster 0,4, which expressed a lower level of *Brachyury* and exhibited some primitive streak and early ectoderm SCN signature, was comprised of day 0 cells. It is possible that these cells, which had yet to undergo the

differentiation protocol, had escaped LIF-controlled naïve pluripotency and were already in a primed early ectodermal state.

Cluster 0 cells expressing *Mixl1*, a regulator of mesoderm and endoderm (Hart et al., 2002) further branched into two distinct lineages according to RNA velocity: mesoderm (cluster 8) and endoderm (clusters 7,1 and 7,0). Cluster 8 was largely enriched for *Mesp1*, and some cells at later time (determined by RNA velocity) in this cluster also expressed *Tbx6*. SCN analysis revealed that this cluster exhibited multiple mesoderm-related signatures including primitive streak, anterior paraxial mesoderm, presomitic mesoderm, and somites, ultimately supporting its mesoderm annotation. Along the endoderm path, cluster 7,1 transitioned from cluster 0 and further transitioned into cluster 7,0. Both 7,1 and 7,0 are enriched for *Foxa2* and *Sox17* expression. SCN analysis suggested an increasing endodermal signature down this path, with cluster 7,1 classifying as anterior primitive streak and weaker as definitive endoderm and gut endoderm, while cluster 7,0 more strongly classified as definitive and gut endoderm.

The neuroectodermal path begins from both cluster 0 and 0,4 according to RNA velocity results. We annotated clusters 0,3 and 4 as neuromesodermal progenitor (NMP) populations. SCN further verified cluster 0,3, which expresses *T* and *Sox2*, as early NMP and cluster 4, which expresses *Sox2* and exhibits variable *T* expression, as late NMP. Cluster 7,2, which expresses *Chordin*, classified strongly as node. Interestingly, RNA velocity suggested the existence of two parallel paths along the neuroectodermal lineage, with cluster 0,2 transitioning into cluster 5, and cluster 4 transitioning into cluster 2. Along the first path, cluster 0,2, classified weakly as a number of ectodermal cell types including early ectoderm. Because RNA velocity indicated this cluster transitioned into the later neuroectoderm cluster 5, we annotated it as early ectoderm. Cluster 5 classified strongly as forebrain/midbrain, with some weaker future spinal cord classification. Given this, and it's *Zic1* and *En2* expression, we annotated this cluster as future brain. Cluster 9, which expresses *Crabp1* and differentially expresses *Tubb3* (suggesting a later neuronal identity compared to cluster 2 or 5), seems to transition from this cluster, and is thus annotated as future brain. Along the second neuroectodermal path, cells in cluster 2 classified more strongly as future spinal cord with a weaker forebrain/midbrain classification. Importantly this cluster expresses many patterning genes from the Hox family indicative of spinal cord fated cells, including *Hoxd9* and *Hoxb6* (Philippidou and Dasen, 2013). For this reason, we annotated this cluster as future spinal cord. A summary of our cluster labels can be found in the supplemental tables (**Table S1**).

Note S3. Epoch elucidates regulatory networks driving lineage specification in early *in vitro* mouse ESC differentiation

To begin our network analysis exploring lineage specification, we applied Epoch to the day 0 through day 4 directed differentiation data, using latent time (Bergen et al., 2020) to order cells, and reconstructing the networks underlying each of the three lineages. Latent time was divided into three epochs, and a dynamic network was extracted for each lineage (**Fig S5A-C**). In examining the Epoch-predicted top regulators and their top targets (**Fig S5D, Table S2**), we observed noteworthy regulatory programs emerging over time along the three lineages. For example, a module including the genes *Gsc*, *Lhx1*, *Sox17*, *Foxa1*, *Foxa2*, and *Gata6* emerges in the third epoch along the endoderm path. *Sox17* and *Foxa2* serve as endoderm markers, as knockout of *Sox17* has been shown to deplete definitive gut endoderm and lack of *Foxa2* results in failure to develop fore- and midgut endoderm (Dufort et al., 1998; Kanai-Azuma et al., 2002). *Foxa1* has also been implicated in the initiation and maintenance of the endodermal lineage (Ang et al., 1993), while inactivation of *Lhx1*, which has been shown to bind to the enhancer region of *Foxa2*, disrupts the development of definitive endoderm (Costello et al., 2015). The regulator *Sox11* and a collection of Hox genes emerge in the third epoch of the neuroectoderm lineage, suggesting the activation of programs that pattern brain and spinal cord (Philippidou and Dasen, 2013). Ultimately, these networks served as a foundation for the remainder of the analyses in this study, representing TF-target regulation present along each path of lineage specification.

In addition to directly extracting top regulators via PageRank and Betweenness-Degree from the reconstructed lineage networks, we highlight a more nuanced methodology to resolve influential TFs based on bagging, since the consensus of many reconstructed networks would be less susceptible to noise. Specifically, we limited our scope to early time, corresponding to a naïve to primed pluripotency transition, and bootstrapped the network reconstruction (**Fig S5E-G**). For each of these sampled and reconstructed networks, we predicted the top 10 regulators involved in the transition by both PageRank

and Betweenness-Degree. We then applied a consensus approach in which regulators were ranked by their frequency as a top 10 regulator (based on PageRank) or average importance (based on Betweenness-Degree) amongst the bootstrapped networks. The top regulators recovered by both PageRank and Betweenness-Degree were largely overlapping, and included TFs such as *Zfp42*, *Klf2*, *Sox11*, *Pou3f1*, and *Klf4*.

Of the top 10 regulators predicted through this consensus approach, *Klf2* and *Klf4* are critical in maintaining ground state pluripotency (Yeo et al., 2014). For example, it has been demonstrated that nuclear export of *Klf4* causes rapid decline in *Klf4* transcription, a trigger that leads to the exit of naïve pluripotency, and the blocking of which prevents ESC differentiation (Dhaliwal et al., 2018). Thus, it is unsurprising that they are predicted to have an influential role in dictating network topology. Interestingly, a number of predicted top regulators, such as *Pou3f1* and *Sox11*, are known to promote neural fate (Barral et al., 2019; Wang et al., 2013; Zhu et al., 2014). Knockdown of another predicted top regulator, *Tcea3*, has also been shown to bias differentiation toward mesendodermal fates (Park et al., 2013). Taken together, this suggests that the underlying network is primed for the neuroectodermal specification fate choice, aligning with the default model of neural induction in mESCs (Muñoz-Sanjuán and Brivanlou, 2002; Tropepe et al., 2001) and the apparent preference for neuroectoderm fate amongst cells in this dataset.

Note S4. *In vivo* dynamic network top regulator prediction

Of the 22 unique top regulators in the second and third epochs of the *in vivo* mesoderm network (**Fig 7A**), at least 17 have known roles in guiding exit from pluripotency, mesoderm specification, or somitogenesis. In particular, a number of top regulators in the second epoch are required for guiding the exit of pluripotency, such as *Nr6a1* and *Snai1* (Galvagni et al., 2015; Gu et al., 2005). Others in this epoch are components of the Wnt signaling pathway including *Lef1*, a Wnt effector, and *Sp5*, which itself is directly activated by Wnt signals and further aids in activating Wnt target genes to promote differentiation into multipotent mesodermal progenitors (Kennedy et al., 2016). Further, TFs like *Brachyury* and *Lhx1* are necessary for primitive streak formation and normal development of primitive streak-derived tissues (Shawlot et al., 1999; Wilson et al., 1995). Spread across both the second and third epochs are a number TFs responsible for patterning and promoting mesoderm fate. These include TFs like *Hoxb1*, *Hoxb2*, *Hoxb9*, *Cdx1*, *Cdx2*, *Mesp1*, *Meis2*, and *Id1* (Bernardo et al., 2011; Cecconi et al., 1997; Foley et al., 2019; Gouveia et al., 2015; Malaguti et al., 2013; Roux et al., 2015; Takahashi et al., 2005, 2007). Of these, many are also responsible for inhibiting alternative fates, such as *Id1*'s role in inhibiting neural induction and *Cdx2*'s role in blocking cardiac differentiation (Mendjan et al., 2014; Zhang et al., 2010). Finally, at least two TFs in the third epoch, *Tbx6* and *Hes7*, are essential for specifying and coordinating somitogenesis (Bessho et al., 2001; Chapman et al., 2003). Ultimately, the predicted top regulators match well with prior literature describing the ESC to presomitic mesoderm lineage trajectory from which the network was reconstructed. Additionally, this further corroborates Epoch's usefulness and ability in isolating important genes driving dynamic processes.

Supplemental Experimental Procedures

mESC maintenance and differentiation

GFP-Brychyury mESC cells (Gadue et al., 2006) were maintained and differentiated through day 2 of the directed differentiation protocol according to Spangler et al. (Spangler et al., 2018). At day 2, four different primitive-streak induction treatments were established by adding growth factors Wnt3a (25ng/mL) and Activin A (WA; 9ng/mL) along with one of the following additional growth factors: Noggin (150ng/mL), Gsk inhibitor (10mM), or BMP4 (0.5ng/mL). On the day of sequencing, EBs and monolayers were dissociated to a single-cell suspension through the use of TrypLE and 40uM cell strainers.

Multi-Seq protocol

We utilized the MULTI-seq sample barcoding and library preparation protocol (McGinnis et al., 2019) in order to sequence all samples in one 10x capture run. In short, the protocol involved tagging cell membranes with sample specific barcodes using a lipid-modified oligonucleotide (LMO). The LMOs (reagents obtained from the McGinnis lab) anchor into the cell membrane and allow for the attachment of a normal ssDNA oligonucleotide (MULTI-Seq barcode). A unique MULTI-seq barcode was added to each sample after which, all samples were pooled and sequenced as if they were one sample. An equal number of cells from each sample were combined to make up the pooled samples and ensure equal representation of each sample after down-stream sequencing. MULTI-Seq barcodes were used down-stream to identify which cells were from which samples.

Library preparation and sequencing

After cells were successfully tagged and pooled, they were submitted to a sequencing core facility for 10x capture and library preparation. A Truseq library preparation was performed with the necessary adjustments made to accommodate the MULTI-Seq platform. Libraries were sequenced on Illumina NovaSeq.

Single-cell data processing

Sequencing alignment to the mm10 reference genome was performed using CellRanger (version 3.1), and spliced, unspliced, and ambiguous read counts were extracted using Velocyto. Sample barcode classification was performed using the deMULTIplex R package (https://github.com/chris-mcginnis-ucsf/MULTI-seq). Subsequent quality control filtering, normalization, clustering, and differential gene expression analysis was performed using Scanpy (version 1.4.5) (Wolf et al., 2018). Specifically, doublets and negatives from the sample barcode classification were removed, along with 271 cells we identified as residual fibroblasts (Thy1+). Next, genes were excluded if they were detected in less than 5 cells; cells were excluded if their mitochondrial gene content exceeded 10% of their total reads or if they had fewer than 500 unique genes. This left us with 5530 cells and 17528 genes. The data was then normalized and log transformed before highly variable genes were identified (1938 genes). The data was scaled and PCA was performed. Leiden clustering was performed and visualized on a UMAP embedding. We further removed a cluster of possible dying cells expressing relatively high mitochondrial gene content and low total read count, leaving a dataset with 5228 cells and 1938 variable genes. Leiden clustering was refined, and RNA velocity was computed using scvelo (version 0.2.2) (Bergen et al., 2020). Latent time was computed with scvelo, using root and end states that were computed through CellRank (version 1.0.0-rc.0). Cell type classification was performed via SingleCellNet (Tan and Cahan, 2019), using complied embryo data for training the classifier. For the E12.5 muscle development data and day 4 early directed differentiation data, diffusion maps and diffusion pseudotime were computed through Scanpy.

CLR

The Context Likelihood of Relatedness (CLR) algorithm was first proposed in 2007 by Faith et al. as a method for inferring gene regulatory networks from expression data. In short, CLR computes the mutual

information (MI) between expression levels of every potential transcription factor and target gene. From here the method scores the MI values in the context of the network by computing a z-score relying on (1) the distribution of MI values for all potential regulators of the target gene, z_i, and on (2) the distribution of MI values for all potential targets of the transcription factor, z_j:

$$z_{ij} = \sqrt{z_i^2 + z_j^2}$$

The result is a weighted adjacency matrix describing the likelihood of TF-target gene interactions that can be thresholded to derive a network structure. For Epoch, a version of CLR is applied to dynamically expressed genes to reconstruct an initial static network via the 'minet' package in R. Users can choose between implementing the MI version or a version using Pearson correlation in place of MI. In either case, Epoch at this step will return a GRN table. Each row outlines an interaction in the initial network identifying the transcription factor, target gene, correlation, and CLR-based z-score.

Synthetic benchmarking

15 synthetic datasets were generated using the Dyngen package in R (Cannoodt et al., 2021). Specifically, networks were generated with 100-420 genes (including 20-70 TFs), containing varied topological motifs, and cell trajectories were simulated over time. Simulated single-cell sequencing was performed, resulting in datasets that included up to 1000 cells. CLR was implemented with the minet package in R (Meyer et al., 2008), and GENIE3 was implemented with the GENIE3 package in R (Aibar et al., 2017; Huynh-Thu et al., 2010). Variations of these two methods were carried out using the Epoch framework. Performance (AUPR) was compared against sets of 5 random networks generated for each reconstructed network via the random permutation of edge weights across targets.

BEELINE benchmarking

The BEELINE benchmarking tool (Pratapa et al., 2020) was used to compare Epoch against the following single-cell reconstruction methods (using the pre-built Docker images from BEELINE): GENIE3, GRISLI, GRNBoost2, GRNVBEM, LEAP, PIDC, PPCOR, SCODE, SCRIBE, SINCERITIES, and SINGE. We added the four versions of Epoch into the BEELINE framework for benchmarking. The synthetic and curated datasets (including simulation time and pseudotime) on which we did the benchmarking were the same as those used in the BEELINE publication, and downloaded from https://doi.org/10.5281/zenodo.3378975. Briefly, for synthetic datasets, each trajectory type included 5 dataset sizes (100 cells, 200 cells, 500 cells, 2000 cells, and 5000 cells), and each dataset size included 10 distinct datasets (for a total of 50 datasets for each trajectory type). Curated datasets were simulated from four literature derived networks (GSD: gonadal sex determination, HSC: hematopoietic stem cell differentiation, VSC: ventral spinal cord, mCAD: mammalian cortical area development). 10 datasets were simulated for each network at 3 different dropout levels (0%, 50%, 70%), for a total of 30 datasets per network. AUPR, early precision, and runtimes were evaluated. BoolODE (Pratapa et al., 2020) was used to re-simulate the curated datasets for the purposes of extracting the true simulation time.

Bootstrapped network reconstruction of early MULTI-seq data

We isolated early latent time clusters 1 and 3 and uniformly sampled 400 cells, 10 times to reconstruct 10 networks. For each reconstructed network, we extracted top regulators by PageRank and identified TFs that were most frequently ranked in the top 10 regulators. We further computed the average normalized betweenness and average normalized degree across the 10 reconstructed networks to identify top regulators.

Differential network analysis and assessment

As one method of network comparison, we used a naïve edge weight-based approach to extract the most differentially weighted edges between networks. Briefly, this involves computing edge weight differences for each possible edge in the two networks. Edges are then considered "differential" and specific to a given network if the edge weight difference is greater than a given threshold (in the case of both the treatment difference analysis and the *in vivo* vs. *in vitro* comparison, this threshold was set to 2). Such an

approach resulted in “differential networks” which are representative of unique topological features existing in each treatment-specific network. To assess the feasibility of this simple method in extracting real and important topological differences between networks, we first applied the same method to compare the three lineage networks in the dataset. The ground truth on which we assessed these differential networks was a set of lineage-unique “pseudo-ChIP-seq” networks we compiled similar to TF-promoter binding predictions generated by (Lu and Mar, 2020). In short, TF-target interactions were acquired by mapping the PWM of a TF (Weirauch et al., 2014) to the promoter region sequence of potential target genes using the Find Individual Motif Occurrences (FIMO) software (described in more detail below). These ground truth networks were filtered for dynamically expressed genes along given paths (for lineage-specificity), and further filtered for unique edges to mimic extraction of unique topological features (lineage-unique). Using this set of networks, we compared AUPR of our path-differential networks against random networks (generated by shuffling edge weights of targets of each TF in the differential networks). On average the differential networks had a 1.7 fold improvement over random AUPR, with many comparisons falling in the 1.8-3.0 range (data not shown).

Construction of path-unique TF-target pseudo-gold standard

Transcription star site (TSS) annotations were downloaded from the USCS Genome Browser (mouse GRCm38/mm10 genome assembly) (Haeussler et al., 2019). Promoter regions were defined as -750 to +250 around the TSSs. We then acquired position weight matrices (PWMs) of transcription factors from the Catalog of Inferred Sequence Binding Preferences (CIS-BP Database build version 2.00) (Weirauch et al., 2014). We used the Find Individual Motif Occurrences tool (Grant et al., 2011) to obtain potential TF-promoter binding pairs for 648 TFs. We kept all interactions for which p-value < 5×10^{-5}. To generate path-unique gold standards, for each lineage (mesoderm, endoderm, neuroectoderm) we filtered interactions for dynamically expressed genes along each path as computed by Epoch. Further, when assessing the “differential network” pairwise comparisons, we filtered out interactions that were redundant between the gold standards, and only kept interactions unique to each path.

Comprehensive signaling effector TF targets list

We acquired binding score (MACS2) data for 18 signaling effector TFs from the ChIP-Atlas (Oki et al., 2018), and processed each accordingly: target genes were ranked by maximum binding score and the top 2000 targets were retained (or all retained, if less than 2000 targets). To compute average effector activity, target lists were filtered for those appearing in the MULTI-seq directed differentiation data, and not considered in the analysis if the effector had less than 10 targets after filtering.

Targets of TFs by treatment comparison

To quantify the extent of topology differences between treatments along the mesodermal path, we performed pairwise comparison on the four treatment-specific networks and computed Jaccard similarity between predicted targets of TFs. The comparison was carried out on 72 TFs that were selected based on their active expression on day 3 or day 4 (i.e. after induction treatments). We computed baseline similarities for each TF in each treatment by bootstrapping the network reconstruction (10 network reconstructions per treatment, 400 sampled cells per network reconstruction) and averaging their pairwise Jaccard similarities. We selected the 13 TFs exhibiting the greatest differences in targets amongst the treatments and performed gene set enrichment analysis using the Enrichr R package (Chen et al., 2013; Kuleshov et al., 2016) and the GO Consortium Biological Processes database (Ashburner et al., 2000; Gene Ontology Consortium, 2021) on the targets of each TF. We further filtered out terms that did not meet the criteria of adjusted p-value<0.05. Finally, we summarized the results between the 13 TFs by counting the frequency (out of 13) that a term was considered enriched and ranked the results accordingly.

In vivo comparison

To create a comparable in vivo dataset, 250 cells were randomly sampled from each of the following annotated populations in the Grosswendt et al. gastrulation dataset: “Primitive streak anterior”,

“Endoderm”, “Ectoderm early 1”, “Ectoderm early 2”, “Epiblast”, “Fore_midbrain”, “Future spinal cord”, “Endoderm gut”, “Neural crest”, “Ectoderm neural anterior”, “Ectoderm neural posterior”, “Neuromesodermal progenitor late”, “Node”, “Notochord”, “Mesoderm pharyngeal arch”, “Mesoderm presomitic”, “Primitive streak late”, “Primitive streak early”, “Somites”, “Neuromesodermal progenitor early.” The resulting dataset spanned E6.5 to E8.5. Quality control filtering, normalization, and Diffusion Pseudotime trajectory inference was performed using Scanpy (version 1.4.5) (Wolf et al., 2018), analogous to the processing of the MULTI-Seq data described above. Further comparison was limited to the mesodermal lineage – this included populations labeled “Epiblast”, “Primitive streak early”, “Primitive streak late”, and “Mesoderm presomitic”.

Epoch was employed to reconstruct the in vivo dynamic network corresponding to this data. To compare the *in vivo* network to the *in vitro* network, both networks were thresholded such that the top 2% of non-zero edges were kept. Network comparison including differential network extraction and community detection were performed using Epoch. Average module expression was analyzed by computing mean expression of predicted activated members over time if the module included at least 8 such members (members predicted to be repressed were not included in the average so as not to improperly depress this measure of activity). GSEA was performed using the Enrichr R package (Chen et al., 2013; Kuleshov et al., 2016) and the GO Consortium Biological Processes database (Ashburner et al., 2000; Gene Ontology Consortium, 2021). Both the 2021 and 2018 versions of this database were used for enrichment analyses. Figures shown here show results using the 2021 version of this database for all enrichment analyses except for epoch3, community 3. This community was not enriched for any terms in the 2021 database, but was enriched for multiple terms in the 2018 pathway.

Supplemental References

Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.-C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nat. Methods *14*, 1083–1086.

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

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene Ontology: tool for the unification of biology. Nat. Genet. *25*, 25–29.

Barral, A., Rollan, I., Sanchez-Iranzo, H., Jawaid, W., Badia-Careaga, C., Menchero, S., Gomez, M.J., Torroja, C., Sanchez-Cabo, F., Göttgens, B., et al. (2019). Nanog regulates Pou3f1 expression at the exit from pluripotency during gastrulation. Biol. Open *8*.

Bergen, V., Lange, M., Peidli, S., Wolf, F.A., and Theis, F.J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. Nat. Biotechnol. *38*, 1408–1414.

Bernardo, A.S., Faial, T., Gardner, L., Niakan, K.K., Ortmann, D., Senner, C.E., Callery, E.M., Trotter, M.W., Hemberger, M., Smith, J.C., et al. (2011). BRACHYURY and CDX2 mediate BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and extraembryonic lineages. Cell Stem Cell *9*, 144–155.

Bessho, Y., Sakata, R., Komatsu, S., Shiota, K., Yamada, S., and Kageyama, R. (2001). Dynamic expression and essential functions of Hes7 in somite segmentation. Genes Dev. *15*, 2642–2647.

Cannoodt, R., Saelens, W., Deconinck, L., and Saeys, Y. (2021). Spearheading future omics analyses using dyngen, a multi-modal simulator of single cells. Nat. Commun. *12*, 3942.

Cecconi, F., Proetzel, G., Alvarez-Bolado, G., Jay, D., and Gruss, P. (1997). Expression of Meis2, a Knotted-related murine homeobox gene, indicates a role in the differentiation of the forebrain and the somitic mesoderm. Dev. Dyn. *210*, 184–190.

Chapman, D.L., Cooper-Morgan, A., Harrelson, Z., and Papaioannou, V.E. (2003). Critical role for Tbx6 in mesoderm specification in the mouse embryo. Mech. Dev. *120*, 837–847.

Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., Clark, N.R., and Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics *14*, 128.

Costello, I., Nowotschin, S., Sun, X., Mould, A.W., Hadjantonakis, A.-K., Bikoff, E.K., and Robertson, E.J. (2015). Lhx1 functions together with Otx2, Foxa2, and Ldb1 to govern anterior mesendoderm, node, and midline development. Genes Dev. *29*, 2108–2122.

Dhaliwal, N.K., Miri, K., Davidson, S., Tamim El Jarkass, H., and Mitchell, J.A. (2018). KLF4 Nuclear Export Requires ERK Activation and Initiates Exit from Naive Pluripotency. Stem Cell Rep. *10*, 1308–1323.

Dufort, D., Schwartz, L., Harpal, K., and Rossant, J. (1998). The transcription factor HNF3beta is required in visceral endoderm for normal primitive streak morphogenesis. Development *125*, 3015–3025.

Faith, J.J., Hayete, B., Thaden, J.T., Mogno, I., Wierzbowski, J., Cottarel, G., Kasif, S., Collins, J.J., and Gardner, T.S. (2007). Large-scale mapping and validation of Escherichia coli transcriptional regulation from a compendium of expression profiles. PLoS Biol. *5*, e8.

Foley, T.E., Hess, B., Savory, J.G.A., Ringuette, R., and Lohnes, D. (2019). Role of Cdx factors in early mesodermal fate decisions. Development *146*.

Gadue, P., Huber, T.L., Paddison, P.J., and Keller, G.M. (2006). Wnt and TGF-beta signaling are required for the induction of an in vitro model of primitive streak formation using embryonic stem cells. Proc. Natl. Acad. Sci. USA *103*, 16806–16811.

Galvagni, F., Lentucci, C., Neri, F., Dettori, D., De Clemente, C., Orlandini, M., Anselmi, F., Rapelli, S., Grillo, M., Borghi, S., et al. (2015). Snai1 promotes ESC exit from the pluripotency by direct repression of self-renewal genes. Stem Cells *33*, 742–750.

Gene Ontology Consortium (2021). The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res. *49*, D325–D334.

Gouveia, A., Marcelino, H.M., Gonçalves, L., Palmeirim, I., and Andrade, R.P. (2015). Patterning in time and space: HoxB cluster gene expression in the developing chick embryo. Cell Cycle *14*, 135–145.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif.

Bioinformatics *27*, 1017–1018.
Gu, P., LeMenuet, D., Chung, A.C.-K., Mancini, M., Wheeler, D.A., and Cooney, A.J. (2005). Orphan nuclear receptor GCNF is required for the repression of pluripotency genes during retinoic acid-induced embryonic stem cell differentiation. Mol. Cell. Biol. *25*, 8507–8519.
Haeussler, M., Zweig, A.S., Tyner, C., Speir, M.L., Rosenbloom, K.R., Raney, B.J., Lee, C.M., Lee, B.T., Hinrichs, A.S., Gonzalez, J.N., et al. (2019). The UCSC Genome Browser database: 2019 update. Nucleic Acids Res. *47*, D853–D858.
Haghverdi, L., Büttner, M., Wolf, F.A., Buettner, F., and Theis, F.J. (2016). Diffusion pseudotime robustly reconstructs lineage branching. Nat. Methods *13*, 845–848.
Hart, A.H., Hartley, L., Sourris, K., Stadler, E.S., Li, R., Stanley, E.G., Tam, P.P.L., Elefanty, A.G., and Robb, L. (2002). *Mixl1* is required for axial mesendoderm morphogenesis and patterning in the murine embryo. Development *129*, 3597–3608.
Haub, O., and Goldfarb, M. (1991). Expression of the fibroblast growth factor-5 gene in the mouse embryo. Development *112*, 397–406.
Huynh-Thu, V.A., Irrthum, A., Wehenkel, L., and Geurts, P. (2010). Inferring regulatory networks from expression data using tree-based methods. PLoS One *5*.
Kanai-Azuma, M., Kanai, Y., Gad, J.M., Tajima, Y., Taya, C., Kurohmaru, M., Sanai, Y., Yonekawa, H., Yazaki, K., Tam, P.P.L., et al. (2002). Depletion of definitive gut endoderm in Sox17-null mutant mice. Development *129*, 2367–2379.
Kennedy, M.W., Chalamalasetty, R.B., Thomas, S., Garriock, R.J., Jailwala, P., and Yamaguchi, T.P. (2016). Sp5 and Sp8 recruit β-catenin and Tcf1-Lef1 to select enhancers to activate Wnt target gene transcription. Proc. Natl. Acad. Sci. USA *113*, 3545–3550.
Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. *44*, W90-7.
Lachmann, A., Xu, H., Krishnan, J., Berger, S.I., Mazloom, A.R., and Ma'ayan, A. (2010). ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. Bioinformatics *26*, 2438–2444.
Lu, T., and Mar, J.C. (2020). Investigating transcriptome-wide sex dimorphism by multi-level analysis of single-cell RNA sequencing data in ten mouse cell types. Biol. Sex Differ. *11*, 61.
Malaguti, M., Nistor, P.A., Blin, G., Pegg, A., Zhou, X., and Lowell, S. (2013). Bone morphogenic protein signalling suppresses differentiation of pluripotent cells by maintaining expression of E-Cadherin. Elife *2*, e01197.
McGinnis, C.S., Patterson, D.M., Winkler, J., Conrad, D.N., Hein, M.Y., Srivastava, V., Hu, J.L., Murrow, L.M., Weissman, J.S., Werb, Z., et al. (2019). MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. Nat. Methods *16*, 619–626.
Mendjan, S., Mascetti, V.L., Ortmann, D., Ortiz, M., Karjosukarso, D.W., Ng, Y., Moreau, T., and Pedersen, R.A. (2014). NANOG and CDX2 pattern distinct subtypes of human mesoderm during exit from pluripotency. Cell Stem Cell *15*, 310–325.
Meyer, P.E., Lafitte, F., and Bontempi, G. (2008). minet: A R/Bioconductor package for inferring large transcriptional networks using mutual information. BMC Bioinformatics *9*, 461.
Muñoz-Sanjuán, I., and Brivanlou, A.H. (2002). Neural induction, the default model and embryonic stem cells. Nat. Rev. Neurosci. *3*, 271–280.
Oki, S., Ohta, T., Shioi, G., Hatanaka, H., Ogasawara, O., Okuda, Y., Kawaji, H., Nakaki, R., Sese, J., and Meno, C. (2018). ChIP-Atlas: a data-mining suite powered by full integration of public ChIP-seq data. EMBO Rep. *19*.
Park, K.-S., Cha, Y., Kim, C.-H., Ahn, H.-J., Kim, D., Ko, S., Kim, K.-H., Chang, M.-Y., Ko, J.-H., Noh, Y.-S., et al. (2013). Transcription elongation factor Tcea3 regulates the pluripotent differentiation potential of mouse embryonic stem cells via the Lefty1-Nodal-Smad2 pathway. Stem Cells *31*, 282–292.
Pelton, T.A., Sharma, S., Schulz, T.C., Rathjen, J., and Rathjen, P.D. (2002). Transient pluripotent cell populations during primitive ectoderm formation: correlation of in vivo and in vitro pluripotent cell development. J. Cell Sci. *115*, 329–339.
Philippidou, P., and Dasen, J.S. (2013). Hox genes: choreographers in neural development, architects of circuit organization. Neuron *80*, 12–34.
Pratapa, A., Jalihal, A.P., Law, J.N., Bharadwaj, A., and Murali, T.M. (2020). Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data. Nat. Methods *17*, 147–154.

Rogers, M.B., Hosler, B.A., and Gudas, L.J. (1991). Specific expression of a retinoic acid-regulated, zinc-finger gene, Rex-1, in preimplantation embryos, trophoblast and spermatocytes. Development *113*, 815–824.
Roux, M., Laforest, B., Capecchi, M., Bertrand, N., and Zaffran, S. (2015). Hoxb1 regulates proliferation and differentiation of second heart field progenitors in pharyngeal mesoderm and genetically interacts with Hoxa1 during cardiac outflow tract development. Dev. Biol. *406*, 247–258.
Shawlot, W., Wakamiya, M., Kwan, K.M., Kania, A., Jessell, T.M., and Behringer, R.R. (1999). Lim1 is required in both primitive streak-derived tissues and visceral endoderm for head formation in the mouse. Development *126*, 4925–4932.
Spangler, A., Su, E.Y., Craft, A.M., and Cahan, P. (2018). A single cell transcriptional portrait of embryoid body differentiation and comparison to progenitors of the developing embryo. Stem Cell Res. *31*, 201–215.
Street, K., Risso, D., Fletcher, R.B., Das, D., Ngai, J., Yosef, N., Purdom, E., and Dudoit, S. (2018). Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genomics *19*, 477.
Takahashi, Y., Hiraoka, S., Kitajima, S., Inoue, T., Kanno, J., and Saga, Y. (2005). Differential contributions of Mesp1 and Mesp2 to the epithelialization and rostro-caudal patterning of somites. Development *132*, 787–796.
Takahashi, Y., Yasuhiko, Y., Kitajima, S., Kanno, J., and Saga, Y. (2007). Appropriate suppression of Notch signaling by Mesp factors is essential for stripe pattern formation leading to segment boundary formation. Dev. Biol. *304*, 593–603.
Tan, Y., and Cahan, P. (2019). SingleCellNet: A Computational Tool to Classify Single Cell RNA-Seq Data Across Platforms and Across Species. Cell Syst. *9*, 207–213.e2.
Tropepe, V., Hitoshi, S., Sirard, C., Mak, T.W., Rossant, J., and van der Kooy, D. (2001). Direct neural fate specification from embryonic stem cells: a primitive mammalian neural stem cell stage acquired through a default mechanism. Neuron *30*, 65–78.
Wang, Y., Lin, L., Lai, H., Parada, L.F., and Lei, L. (2013). Transcription factor Sox11 is essential for both embryonic and adult neurogenesis. Dev. Dyn. *242*, 638–653.
Weirauch, M.T., Yang, A., Albu, M., Cote, A.G., Montenegro-Montero, A., Drewe, P., Najafabadi, H.S., Lambert, S.A., Mann, I., Cook, K., et al. (2014). Determination and inference of eukaryotic transcription factor sequence specificity. Cell *158*, 1431–1443.
Wilson, V., Manson, L., Skarnes, W.C., and Beddington, R.S. (1995). The T gene is necessary for normal mesodermal morphogenetic cell movements during gastrulation. Development *121*, 877–886.
Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. *19*, 15.
Yeo, J.-C., Jiang, J., Tan, Z.-Y., Yim, G.-R., Ng, J.-H., Göke, J., Kraus, P., Liang, H., Gonzales, K.A.U., Chong, H.-C., et al. (2014). Klf2 is an essential factor that sustains ground state pluripotency. Cell Stem Cell *14*, 864–872.
Zhang, K., Li, L., Huang, C., Shen, C., Tan, F., Xia, C., Liu, P., Rossant, J., and Jing, N. (2010). Distinct functions of BMP4 during different stages of mouse ES cell neural commitment. Development *137*, 2095–2105.
Zhu, Q., Song, L., Peng, G., Sun, N., Chen, J., Zhang, T., Sheng, N., Tang, W., Qian, C., Qiao, Y., et al. (2014). The transcription factor Pou3f1 promotes neural fate commitment via activation of neural lineage genes and inhibition of external signaling pathways. Elife *3*.