Supplementary Material for the manuscript

Detecting critical transition signals from single-cell transcriptomes to infer lineage-determining transcription factors

Running title: **Critical Transition Signal**

Xinan H Yang[1,*], Andrew Goldstein[2], Yuxi Sun[1], Zhezhen Wang[1], Megan Wei[3], Ivan P Moskowitz[1], John M Cunningham[1]

[1] Department of Pediatrics, [2] Department of Statistics, The University of Chicago, Chicago, [3] Johns Hopkins University, Baltimore, Maryland, USA

*: xyang2@uchicago.edu

Contents

Supplementary Tables

Table S1 Data-driven parameter settings for BioTIP application in 6 datasets

Object to count \ dataset		hESC Bargaje 2017	lung Treutlein 2014	E8.25 2019 GSE87038	E8.25 2018 Ibarra-Soria	EB GSE130146	Simulated EMT
Input	Number of analyzed cells	929	131	7240*	11,039	1,531	5,362
Input	Number of detected/expressed genes	96	10,251	10.9k	12,703	15,200	18
Input	Number of global HVGs	Not applicable	3,198	3,073	4,000	4,000	Not applicable
Parameter 1	Cutoff to select variable gene per cluster (b%)	80%	10%	10%	10%	10%	100%
Output 1	Number of the pool of cluster-specific HVG	96	754	1.9k	2.3k	961	18
Parameter 2	FDR for PCC before constructing gene modules	0.2	0.2	0.2	0.05	0.2	0.05
Output 2	Number of genes in a cluster-specific network	70-76	49-340	64-294	131-387	213-336	17-18
Parameter 3	Minimum module size (Number of genes)	10	30[&]	60[&,**]	30[&]	30[&]	6
Parameter 3	Minimum DNB score to select CTS candidate	4[&]	2	2	2	2	0[#]
Output 3	Resultant number of genes in the identified CTSs	18-43	32-180	60-90	33-127	58-64	11 and 12

*: We apply BioTIP to 12 clusters of 7240 cells but discuss the robustness and stability with only 6 clusters of 1362 cells for two reasons -- 1) to speed up the calculation, 2) to compare with QuanTC that has been applied to the same 1362 cells.

**: When discussing robustness against different clustering inputs, we set this parameter to 30 to scan more modules.

#: No signature is higher than 1 nor empirical significant because the expression matrix contains only 18 highly interactive genes.

&: We show consistent BioTIP predictions when tuning this parameter.

Table S2 The classic developing mesoderm cell-lineage markers

Cluster ID, E8.25 2019 data	Cell Identity	Up-regulated marker	Lineage	Reference
C16	Muscle mesenchyme	*Dlk1*	Mesenchymal fibroblast	(Chen et al., 2021)
		Hand1	Mesoderm / cardiac	(Mittnenzweig et al., 2021) (Bargaje et al., 2017)
		Slc2a1	Muscle	(Coudert et al., 2018)
C13	early HEP	*Etv2*	Mesoderm	(Mittnenzweig et al., 2021)
		Dlk1, Mest	Mesenchym fibroblast	(Chen et al., 2021)
		Tal1, Lmo2	Hematopoietic	(Chan et al., 2007)
		Kdr	Hematoendothelial/ cardiac	(Evseenko et al., 2010)
C15	later HEP	*CD34*	Hematoendothelial	(Evseenko et al., 2010)
		Cdh5	Endothelial	(Pijuan-Sala et al., 2019)
		Spi1 (Pu.1), Itga2b	Hematopoietic	(Pijuan-Sala et al., 2019)
C6	Endothelium	*Mest*	Mesenchymal fibroblast	(Chen et al., 2021)
		Esam	Endothelial	(Pijuan-Sala et al., 2019)

Reference:

Bargaje, R., et al. (2017). Cell population structure prior to bifurcation predicts efficiency of directed differentiation in human induced pluripotent cells. Proc Natl Acad Sci U S A *114*, 2271-2276.

Chan, W.Y., *et al.* (2007). The paralogous hematopoietic regulators Lyl1 and Scl are coregulated by Ets and GATA factors, but Lyl1 cannot rescue the early Scl-/- phenotype. Blood *109*, 1908-1916.

Chen, B., et al., and Dun, X.P. (2021). Single Cell Transcriptome Data Analysis Defines the Heterogeneity of Peripheral Nerve Cells in Homeostasis and Regeneration. Front Cell Neurosci *15*, 624826.

Coudert, E., *et al.* (2018). Expression of glucose transporters SLC2A1, SLC2A8, and SLC2A12 in different chicken muscles during ontogenesis. J Anim Sci *96*, 498-509.

Evseenko, D., *et al.* (2010). Mapping the first stages of mesoderm commitment during differentiation of human embryonic stem cells. Proc Natl Acad Sci U S A *107*, 13742-13747.

Mittnenzweig, M., *et al.* (2021). A single-embryo, single-cell time-resolved model for mouse gastrulation. Cell *184*, 2825-2842 e2822.

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

Supplementary Figures with Legends

Figure S1

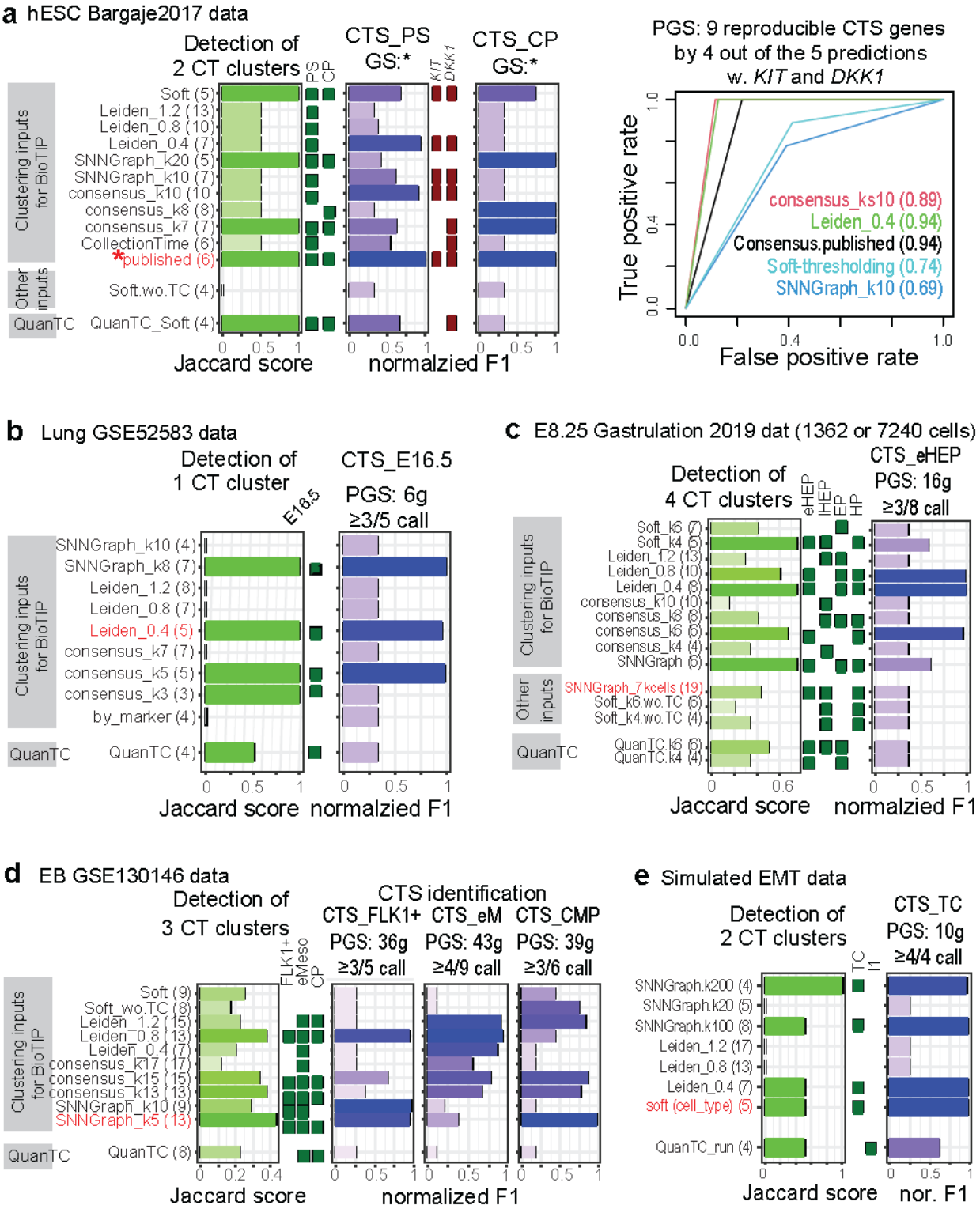


Fig S1. Robustness of BioTIP, related to Figs 2, 3, 5, S7d, S13, S14

a, Results of the hESC data running on different clustering methods and variable key parameters (y-labels). In parathesis is the defined number of cell clusters. Red star indicates the published clusters for this dataset, based on which we detail the BioTIP analysis in **Fig 2**.

Left: Green bars showing the Jaccard scores quantifying a CT detection against the gold standard (GS) bifurcations – the established CT at primitive streak (PS) and the repeatedly detected CT at cardiomyocyte progenitor (CP) (by three tools -- BioTIP, MuTrans, and QuanTC). Green squares check when a prediction includes a transition state. Blue bars show the normalized F1 scores indicating each CT state. Red squares check if a CTS contains the GS markers – a previously evaluated transition markers at day 2.5 (around the PS state), for each run.

Right: ROC plot comparing five clustering methods (with optimal parameters) that detected both GS markers as CTS members at PS, using nine consistently identified CTS member genes as a proxy gold standard (PGS). AUC scores are given in paratheses.

b-e, Like panel a but using proxy gold standard (PGS), showing the results in four independent datasets. There are no evaluated transition marker genes for any of the datasets. Therefore, we infer two types of PGSs for each dataset.

(1) For CT detection, the GS bifurcation(s) are the known transition state in the system and/or the one predicted by both BioTIP and QuanTC. An exception is panel c, in which the HP state is additionally considered because it has significant and stable CTS detected by BioTIP from down-sampled data (**Fig S7d**). These CT clusters serving as proxy GS are listed on the top-right of the green bars.

(2) For CTS identification, the proxy GS markers are consistently predicted genes indicating the best-known transition state in a system, specified atop the blue bars.

The BioTIP results demonstrated in main Figures are highlighted in red. $xg\ y \geq z$ call: x reproducibly identified CTS genes by at least y out of z predictions; SNNGraph: nearest-neighbor graph clustering; Soft.wo.TC: the stable states defined by soft clustering approach using QuanTC pipeline then excluding the transition cells.

Other abbreviations: hESC: human embryonic stem cells; EB: embryoid body; EMT: epithelial-to-mesenchymal transition; SNNGraph: nearest-neighbor graph; TC: transition cell; QuanTC: a model-free method to detect transition cells; E16.5: embryonic day 16.5; eHEP: early haemato-endothelial progenitor; lHRP: later haemato-endothelial progenitor; EP: endothelial progenitor; HP: haematopoietic progenitor; FLK1+: FLK1-expressing mesoderm; eMeso: early mesoderm; CP: cardiomyocyte progenitor.

Figure S2

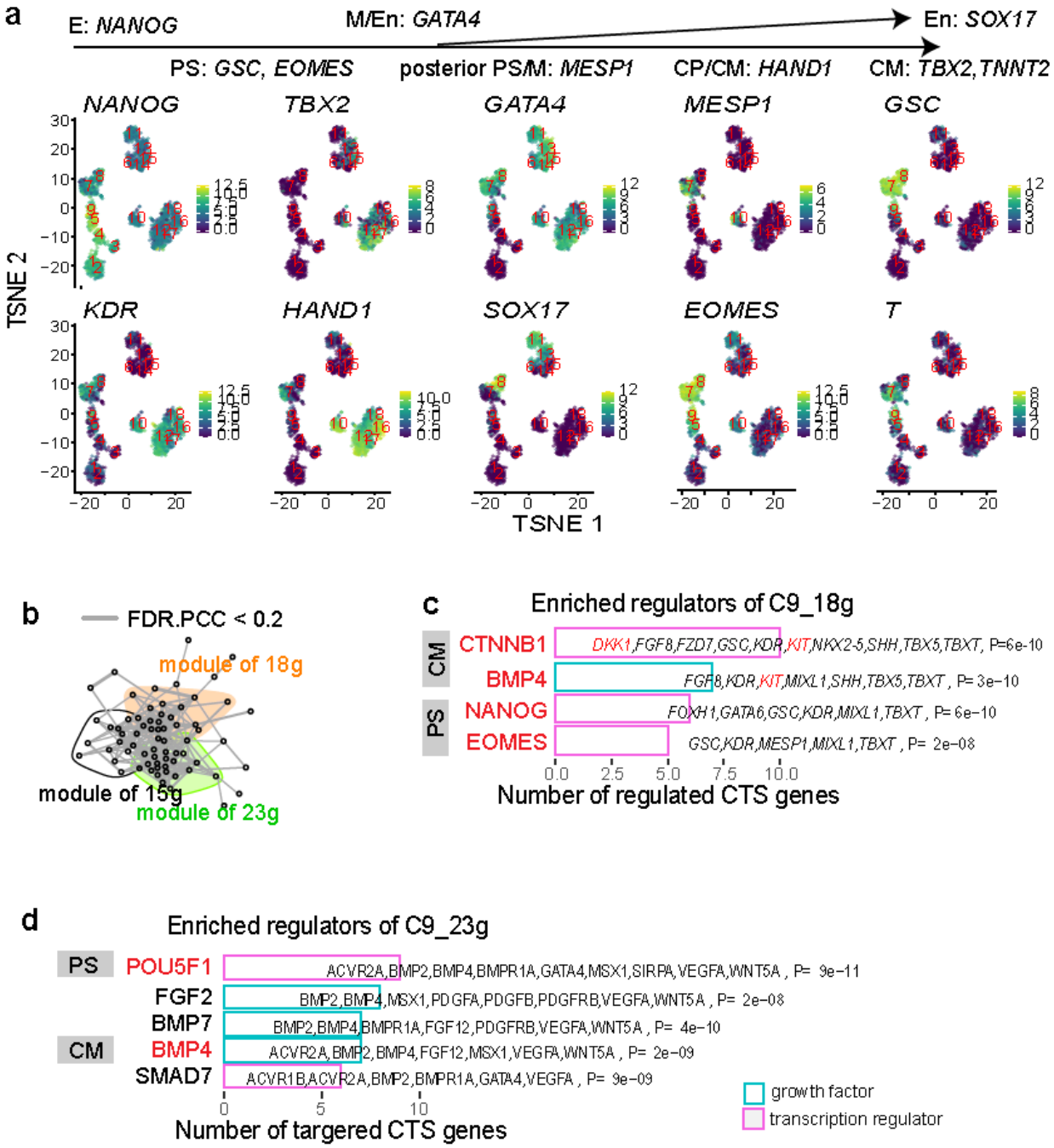

a
E: NANOG
M/En: GATA4
En: SOX17
PS: GSC, EOMES
posterior PS/M: MESP1
CP/CM: HAND1
CM: TBX2,TNNT2
NANOG
TBX2
GATA4
MESP1
GSC
KDR
HAND1
SOX17
EOMES
T
TSNE 2
TSNE 1
b
FDR.PCC < 0.2
module of 18g
module of 15g
module of 23g
c
Enriched regulators of C9_18g
CM
PS
CTNNB1
BMP4
NANOG
EOMES
DKK1,FGF8,FZD7,GSC,KDR,KIT,NKX2-5,SHH,TBX5,TBXT, P=6e-10
FGF8,KDR,KIT,MIXL1,SHH,TBX5,TBXT , P= 3e-10
FOXH1,GATA6,GSC,KDR,MIXL1,TBXT , P= 6e-10
GSC,KDR,MESP1,MIXL1,TBXT , P= 2e-08
Number of regulated CTS genes
d
Enriched regulators of C9_23g
PS
CM
POU5F1
FGF2
BMP7
BMP4
SMAD7
ACVR2A,BMP2,BMP4,BMPR1A,GATA4,MSX1,SIRPA,VEGFA,WNT5A , P= 9e-11
BMP2,BMP4,MSX1,PDGFA,PDGFB,PDGFRB,VEGFA,WNT5A , P= 2e-08
BMP2,BMP4,BMPR1A,FGF12,PDGFRB,VEGFA,WNT5A , P= 4e-10
ACVR2A,BMP2,BMP4,FGF12,MSX1,VEGFA,WNT5A , P= 2e-09
ACVR1B,ACVR2A,BMP2,BMPR1A,GATA4,VEGFA , P= 9e-09
Number of targered CTS genes
growth factor
transcription regulator

Fig S2. Analysis of the hESC dataset, related to Fig 2

a, Top: Established cell lineage markers. Bottom: TSNE plot showing marker gene expression of individual cells, numbered by 18 unique cell cluster IDs. Each dot represents a single cell. Dot color decodes expression levels on log-2 scale. E: epiblast, PS: primitive streak; M: mesoderm, En: endoderm, CP: cardiomyocyte progenitor, CM: cardiomyocyte.

b, A graph view of network modulation determined by the random-walk approach for preselected HVGs (dots) for C9 cells. Background colors represent different gene modules. The module size is listed in parentheses.

c-d, Bar plot of significant upstream regulators for the 18 genes (panel c) or 23 genes (panel d) charactering C9. Bar color decodes the molecular types of these upstream regulators. Also shown are the target genes and enrichment p-values (IPA analysis). Red font highlights the established fate-determining TFs (for either PS or CM state) and early-warning genes for the bifurcation.

Figure S3

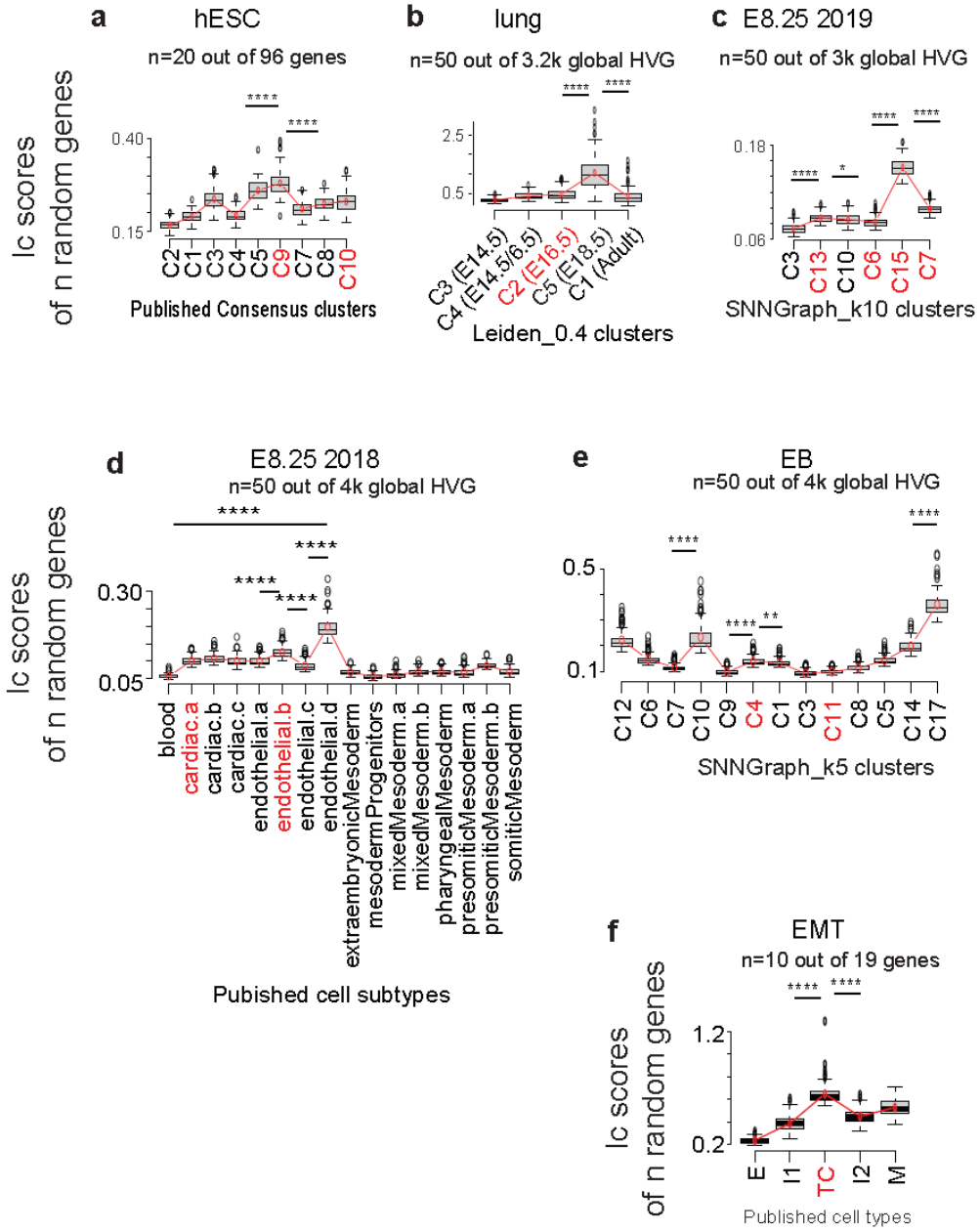


Fig S3. Applying existing Ic approach to six datasets, related to Fig 5.

a, Ic scores in hESC after sampling n=20 background (measured) genes. Boxplot shows the scores per cluster after 100 runs. The red line connects the average scores. Clusters are ordered along an inferred trajectory in **Fig 5c**. The red label highlights the gold-standard TC clusters, each is compared to adjacent clusters using Wilcox test. *: P<0.05, **: P<0.01, ***: P<0.001, ****:P<0.0001.

b-f, Similar to panel a, one panel showing the results of one dataset. For big datasets, the global HVG is the background genes. The value of n is adjusted due to the size of background genes, although Ic scores has been shown to be robust to the value of n (Bargaje et al., 2017).

Figure S4

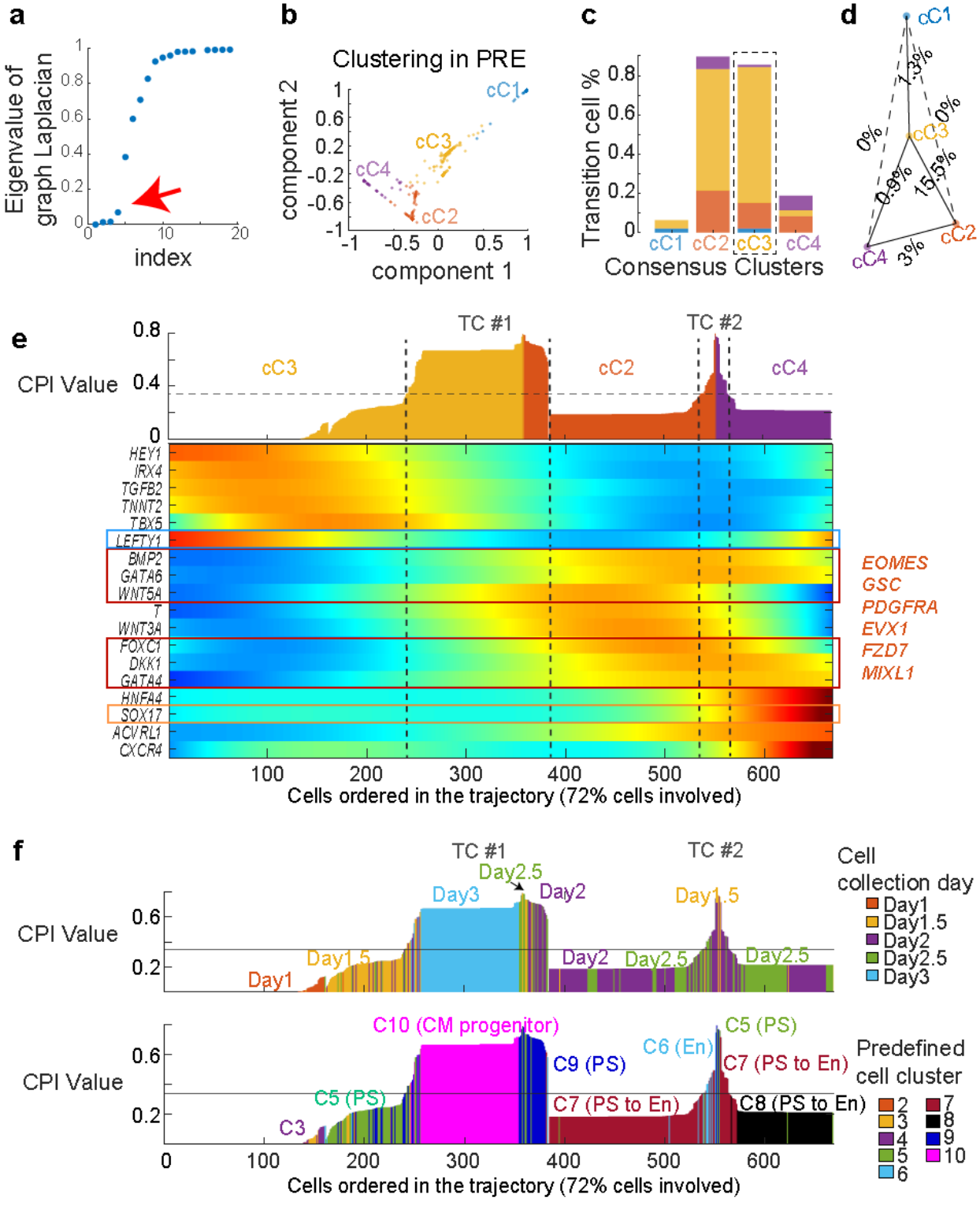

a
Eigenvalue of graph Laplacian
index
b
Clustering in PRE
component 2
component 1
cC1
cC2
cC3
cC4
c
Transition cell %
Consensus Clusters
cC1 cC2 cC3 cC4
d
cC1
cC2
cC3
cC4
7.3%
0%
0%
0.9%
15.5%
3%
e
CPI Value
TC #1
TC #2
cC3
cC2
cC4
HEY1
IRX4
TGFB2
TNNT2
TBX5
LEFTY1
BMP2
GATA6
WNT5A
T
WNT3A
FOXC1
DKK1
GATA4
HNF4A
SOX17
ACVRL1
CXCR4
EOMES
GSC
PDGFRA
EVX1
FZD7
MIXL1
Cells ordered in the trajectory (72% cells involved)
f
CPI Value
TC #1
TC #2
Day1
Day1.5
Day2
Day2.5
Day3
Cell collection day
Day1
Day1.5
Day2
Day2.5
Day3
C10 (CM progenitor)
C5 (PS)
C6 (En)
C9 (PS)
C7 (PS to En)
C3
C8 (PS to En)
Predefined cell cluster
2 7
3 8
4 9
5 10
6
Cells ordered in the trajectory (72% cells involved)

Fig S4. Applying QuanTC to the Bargaje dataset, related to Figs 2 and 5.

a, The first 20 sorted eigenvalues (x-axis) of the graph Laplacian of the constructed consensus matrix. The number of consensus clusters (cC) (k=4, red arrow) was predicted, where the largest gap along the y-axis is observed.

b, Two-dimensional probabilistic regularized embedding (PRE) visualization for four consensus clusters (left) or cell collection time points (right).

c, Percentage of identified transition cells (TC) in each cluster relative to the total number of TC. Font color in the x-axis is consistent with panel b. Dashed box: the relatively stable cluster where the most transition cells are similar to the true cell label chosen as the starting cluster to infer potential transition trajectories.

d, Transition trajectories with node colors consistent with panel b. Percentages of TC between clusters are shown.

e, Heatmap of normalized expression of 5 top marker genes and 6 top transition genes for the chosen trajectory with the highest proportion of cells involved. Columns represent cells ordered along the transition trajectory and rows represent genes. Color indicates the normalized expression of each gene: red represents high expression and blue represents low expression. The identified transition genes are marked in two boxes. The expression of genes either decrease during the transition from cC3 to cC2 (blue box, TC #1) or increase (maroon box, TC #2). Top: Histogram of cell plasticity index (CPI) values of each cell along the transition trajectory. Dashed horizontal line marks the cutoff of CPI. Other identified transition genes for the transition cells (TC #1) are listed on the right. Only one transition gene (*Sox17*, orange box) was detected for the TC #2.

f, Histogram of cell plasticity indexes (CPI) values of each cell along the transition trajectory, colored by collection day (top) and the cell identity published by the Bargaje (bottom). TC #1 describes the transition from PS (C9) to cardiomyocyte progenitor (C10, see **Figure 2a**). TC #2 at the PS state describes possibly gaining endothelial fate. E: epiblast, PS: primitive streak; En: endoderm.

Figure S5

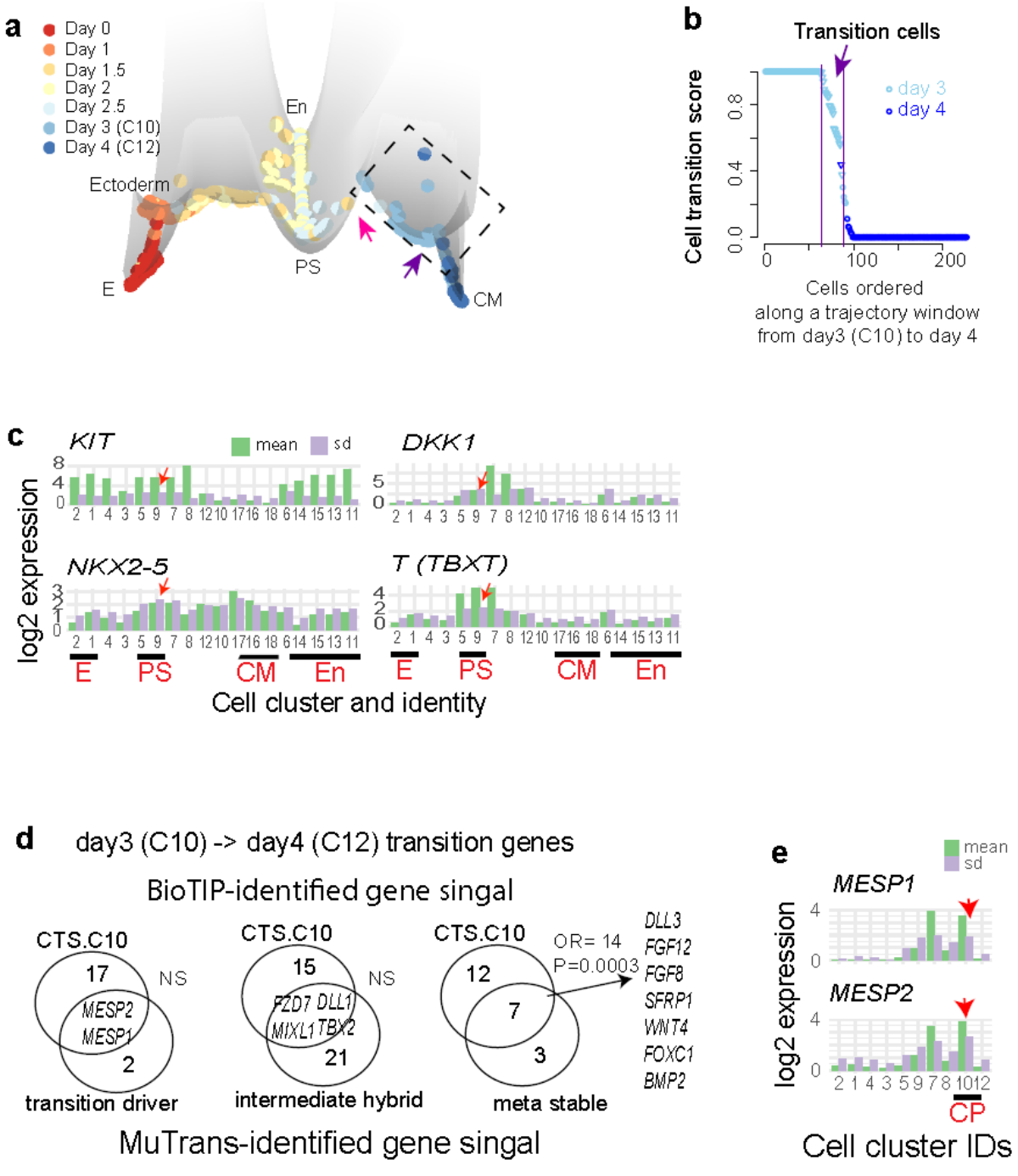

a
Day 0
Day 1
Day 1.5
Day 2
Day 2.5
Day 3 (C10)
Day 4 (C12)
Ectoderm
En
PS
E
CM
b
Transition cells
Cell transition score
day 3
day 4
Cells ordered
along a trajectory window
from day3 (C10) to day 4
c
KIT
DKK1
NKX2-5
T (TBXT)
mean
sd
log2 expression
E
PS
CM
En
Cell cluster and identity
d
day3 (C10) -> day4 (C12) transition genes
BioTIP-identified gene singal
CTS.C10
17
NS
MESP2
MESP1
2
transition driver
CTS.C10
15
NS
FZD7 DLL1
MIXL1 TBX2
21
intermediate hybrid
CTS.C10
12
7
3
OR= 14
P=0.0003
DLL3
FGF12
FGF8
SFRP1
WNT4
FOXC1
BMP2
meta stable
MuTrans-identified gene singal
e
mean
sd
MESP1
MESP2
log2 expression
CP
Cell cluster IDs

Fig S5. Comparing BioTIP to MuTrans using the hESC dataset, related to Fig 2.

a, The constructed dynamical manifold of cells collected from day 0-2.5 and day 3-4 along the cardiomyocyte lineage using MuTrans. Colored arrow distinguishes the verified tipping point at day 2-2.5 (red) or the newly identified one at day 3 (purple). The color of each individual cell represents the cell-collection day. The dashed square indicates the transition event (panel b).

b, Transition cell scores (y-axis) estimated for 227 cells of day 3 and day 4. Purple lines isolate 24 MuTrans-predicted transition cells, among which 23 are day 3 cells (subset of C10).

c, Bar plot displaying the log2-scaled expression patterns of four CTS member genes for C9. Red arrows point to the highest standard deviation (sd) at C9 for each gene. E: epiblast, PS: primitive streak; CM: cardiomyocyte, En: endoderm.

d, Venn diagrams comparing BioTIP's precision for C10 with three types of MuTrans-predicted transition genes. Transition driver genes vary during transition. Intermediate hybrid genes are expressed in both stable and the transition cells; meta-stable genes are expressed in the stable states (Zhou et al., 2021). The Fisher's exact test was performed using 96 measured genes as background. NS: non-significant.

e,. Same as panel c, except with CP marker *MESP1/2* and C10. CP: cardiomyocyte progenitor.

Figure S6

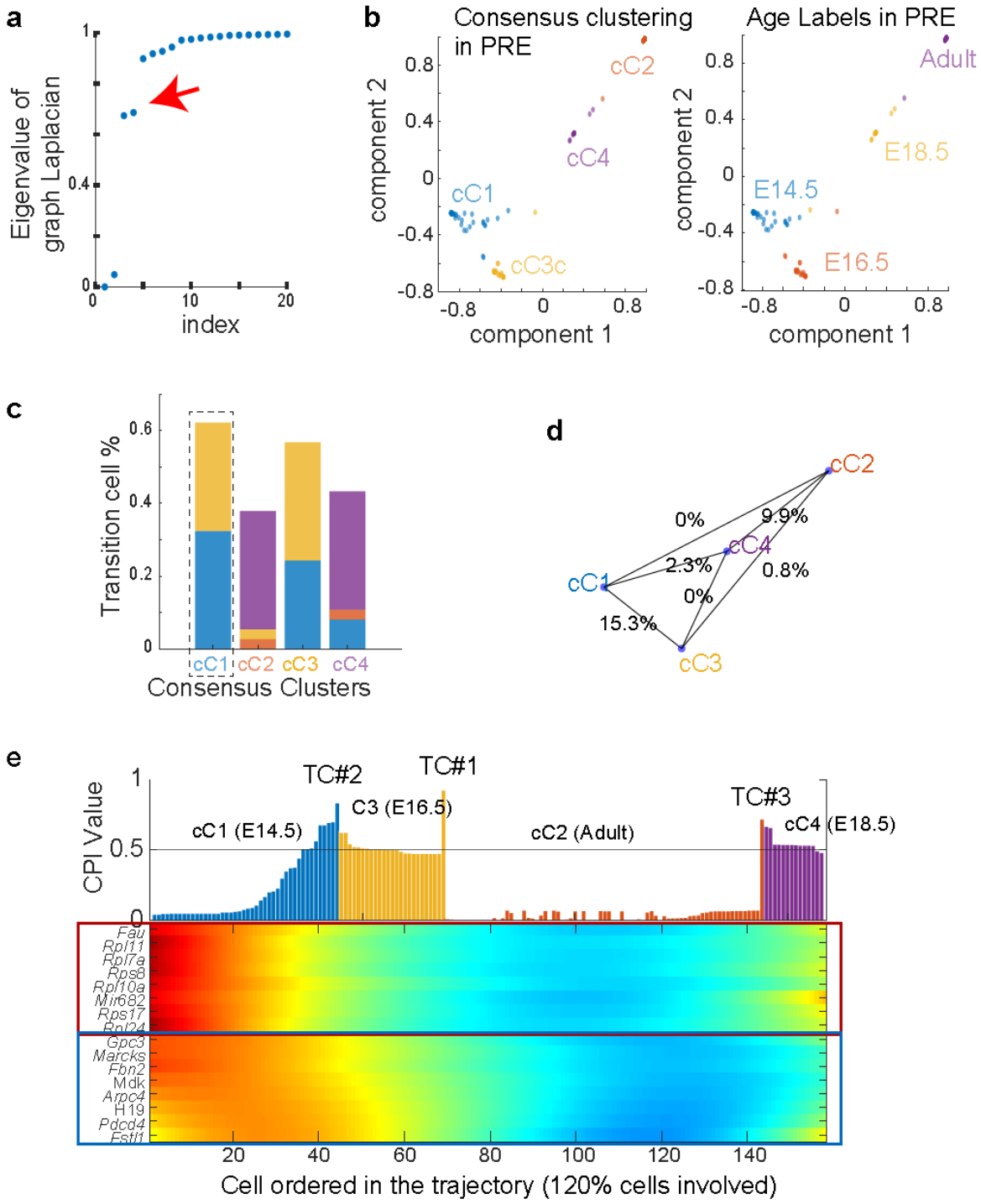
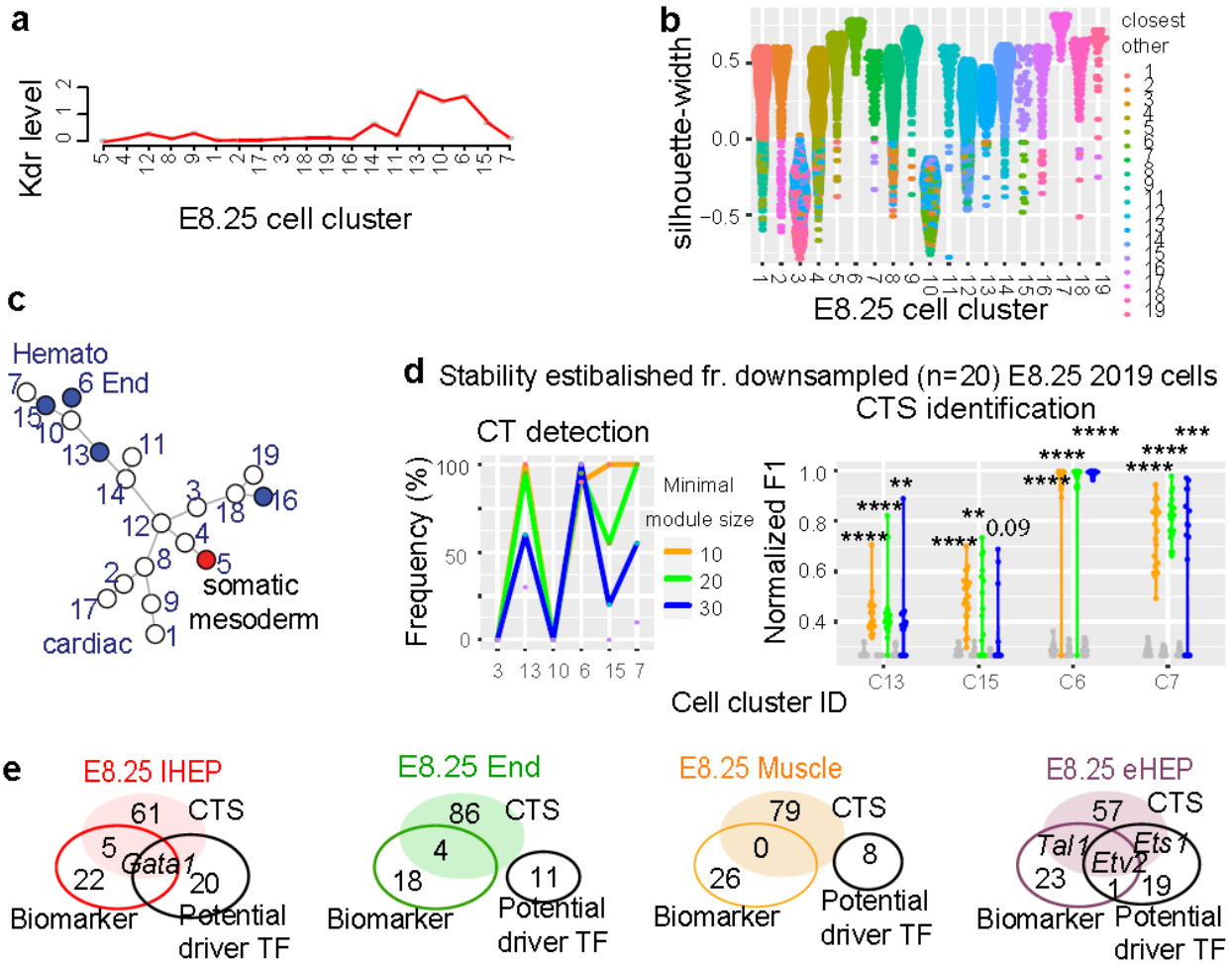


Fig S6. Applying QuanTC to the lung dataset, related to Figs 3 and 5.

a-e, Same as **Figure S4**, except applied to the lung dataset. The identified transition genes are in two boxes – decreasing in E16.5 cells (TC #1) and decreasing in the transition from E14.4 cells to E16.5 cells (TC #2). CPI: cell plasticity indexes.

Figure S7



f Applying to scRNA-seq data (BioTIP: 7240 cells; QuanTC: 1362 cells)

QuanTC identifications

BioTIP identifications	4 stable states	6 stable states							
		trajectory starting at C6				trajectory starting at C3			
	M&HEP->HEP (42g)	eHEP->EP (33g)	HEP->EP (8g)	#3 (6g)	#4 (5g)	eHEP->EP (11g)	HEP->EP (6g)	#3 (5g)	#4 (4g)
IHEP_67g	OR = 8* *Rhoj, Vamp5*	OR=5 *Tal1*	OR=23* *Oit3*			OR=11.5 *Gngt2*	OR=33* *Oit3*		
End_90g			OR=41*** *Oit3, Lyve1*			OR=27*** *Plvap, Ctla2a*	OR=61*** *Oit3, Lyve1*		
eHEP_60g	OR = 15*** *Rhoj, Elk3, Rasip1*	OR=12* *Tal1, Etv2*	OR=26* *Oit3*			OR=71*** *Plvap, Gngt2, Ctla2a*	OR=37* *Oit3*		
Muscle_79g									

(10,938 gene background) * P ≤ 0.05; **P < 0.01; *** P < 0.005

Fig S7. Analysis of the E8.25, 2019 dataset, related to Figs 4-6.

a, Lines show the average expression levels of *Kdr* in each of the 19 cell states.

b, Distribution of the approximate silhouette width (y-axis) across 19 cell states (x-axis) of the dataset. Each point represents a cell and is colored with the identity of its own cluster if its silhouette width is positive, or that of the closest other cluster if the width is negative. The higher a silhouette score is, the more consistent the cluster.

c, Trajectory reconstruction of the 19 clusters using the Minimum-spanning-tree algorithm. The red dot is the knowledge-based root in the trajectory. Blue points are the BioTIP-predicted CT states.

d, Stability of BioTIP estimated from 1362 cells and 3 k HVGs after down-sampling 95% genes and 95% cells (20 runs). Left: Frequency of identifying each cluster as a significant CT state. Right: The normalized F1 score for CTS identification at each of the four states serving as proxy gold standard (Detailed in **Fig S1c**). In both subpanels, color encodes the minimal gene-module size to be detected. **:P<0.001; ***:P<0.001 in t-test.

e, Venn-diagram comparing each CTS gene members (filled circle) with their potential regulatory TFs (black circle), and the up-regulated biomarkers of the representative critical transition state (colored circle). lHEP: later haemato-endothelial progenitor; End: endothelial; eHEP: early HEP.

f, Comparing QuanTC-identified transition genes with BioTIP-identified CTSs. When there is common identification, the odds ratio (OR) and p-value of the Fisher's exact test between the two identifications are shown. QuanTC's results with 4 or 6 stable states are shown (Detailed in **Figs S9, S10**). QuanTC was run on the 1362 cells spanning three BioTIP-predicted transition states and their neighboring clusters in the UMAP space but not the muscle mesenchyme state. Grey box: the negative control where no common identification was expected. QuanTC run with 6 states predicted two transition trajectories with different starting points; each detected four sets of transition genes (Detailed in **Fig S10**). HEP: haemato-endothelial progenitor; HP: hematopoietic progenitor; EP: endothelial progenitor; M: muscle mesenchyme.

Figure S8

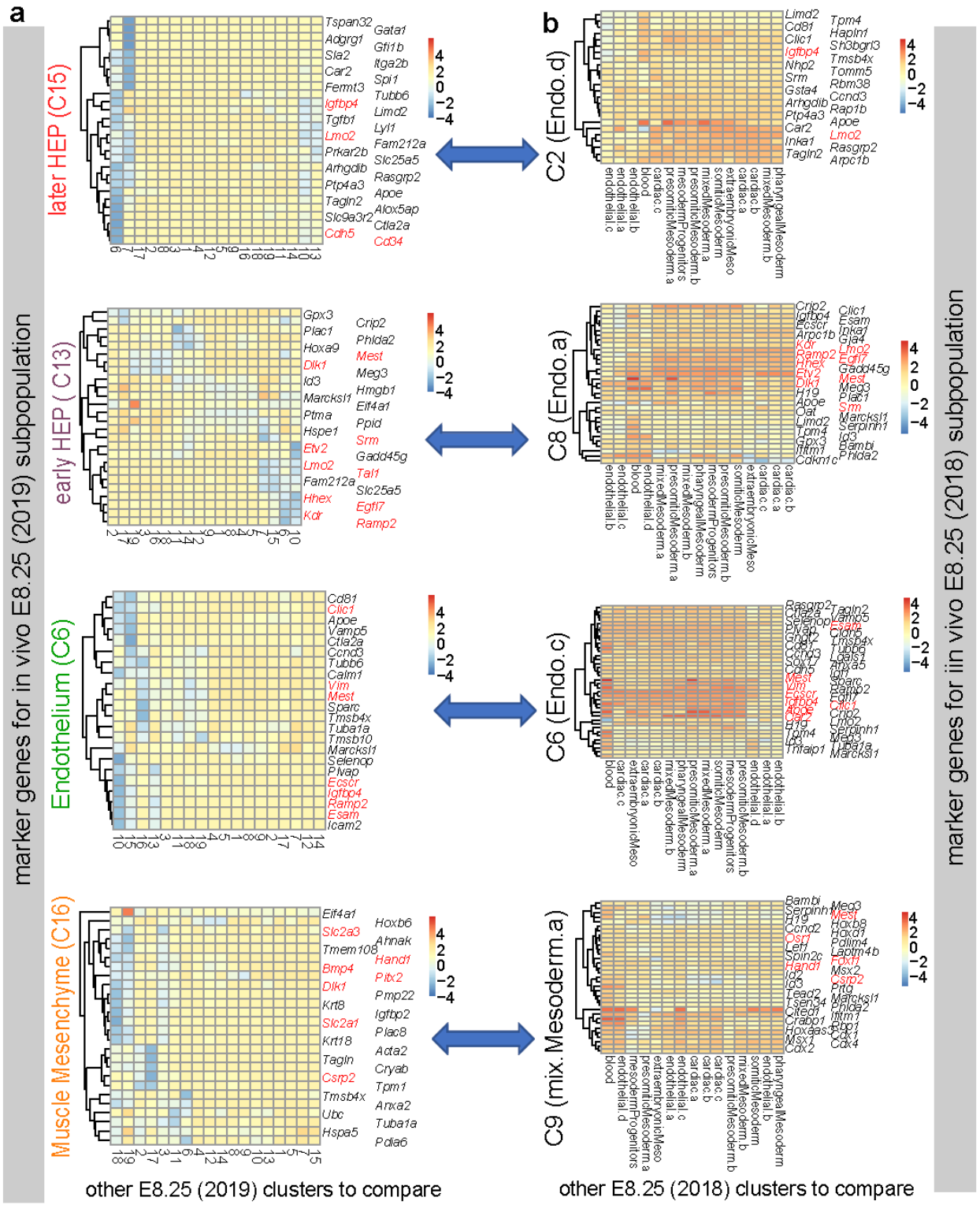

a
b
marker genes for in vivo E8.25 (2019) subpopulation
marker genes for iin vivo E8.25 (2018) subpopulation
later HEP (C15)
early HEP (C13)
Endothelium (C6)
Muscle Mesenchyme (C16)
C2 (Endo.d)
C8 (Endo.a)
C6 (Endo.c)
C9 (mix.Mesoderm.a)
other E8.25 (2019) clusters to compare
other E8.25 (2018) clusters to compare

Fig S8. Comparing the classic up-regulated markers between two E8.25 datasets, related to Fig 4

a, Heatmap showing the top-5 ranked regulated genes each identified E8.25 bifurcation state over other E8.25 states (2019). Euclidean distance was measured, and normalized log counts were centered and scaled along the x-axis. Pairwise comparisons between cell states were run using the Wilcoxon rank-sum test.

b, Like panel a but for the top 10 up-regulated markers detected from the E8.25 (2018) dataset. Each blue arrow connects clusters of different datasets predicted to be the same bifurcation state (See **Fig 4g**). Pairwise comparisons between cell states were run using a t-test. The up-regulated markers were identified as a summary logFC > 1, FDR<0.01, and rank ≤ 10, using the R package scran.

In both panels, classic cell identification markers annotating the cluster are highlighted in red.

Figure S9

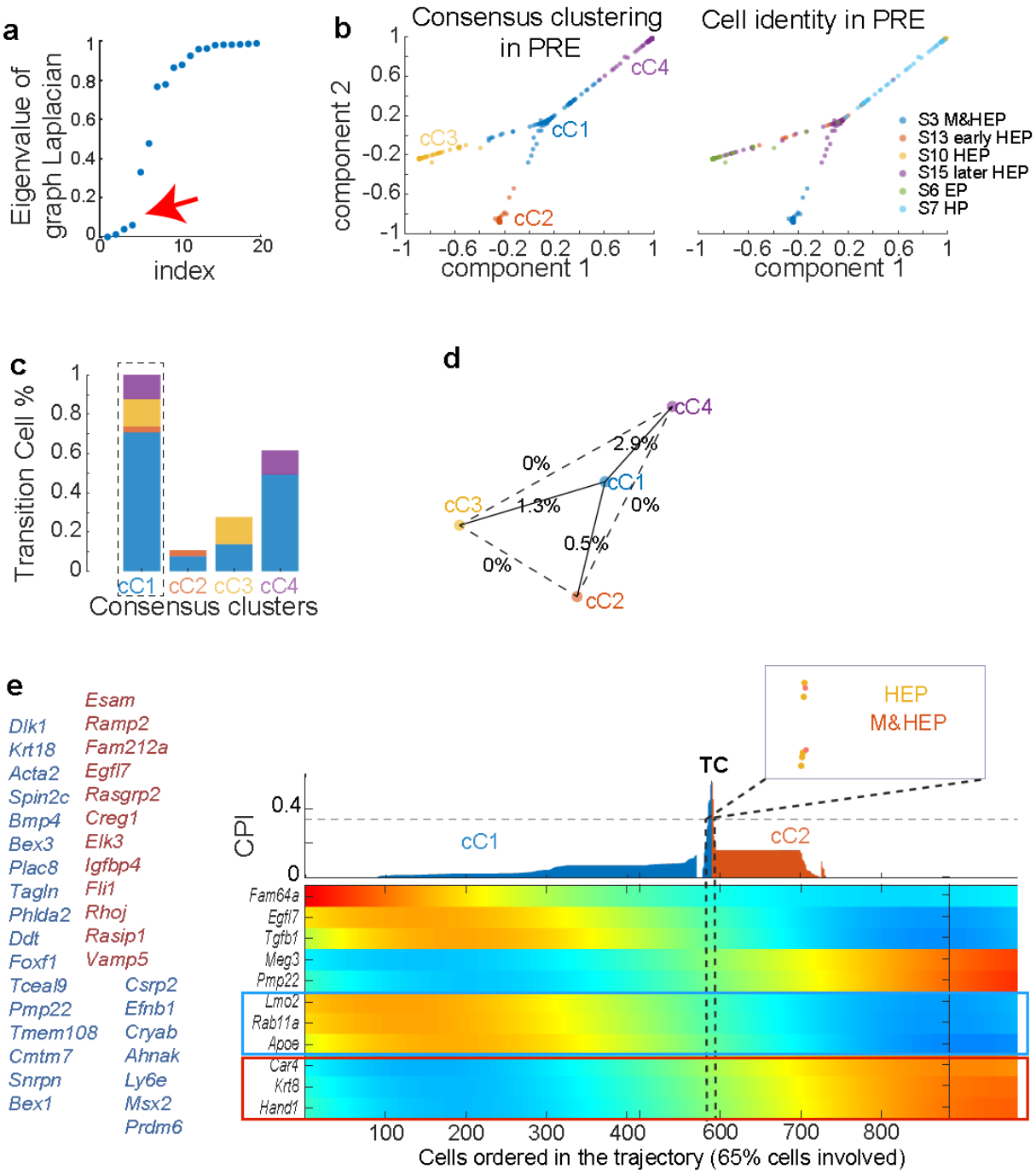


Fig S9. Applying QuanTC (k=4) to the E8.25 2019 dataset, related to Fig 4.

a-e, Same as **Fig S4**, except applied to the E8.25 2019 dataset and with the number of consensus clusters (k=4, red arrow) was predicted, where the largest gap along the y-axis is observed. M: mesenchyme; HEP: haemato-endothelial progenitor; HP: hematopoietic progenitor; EP: endothelial progenitor. CPI: cell plasticity indexes.

Figure S10

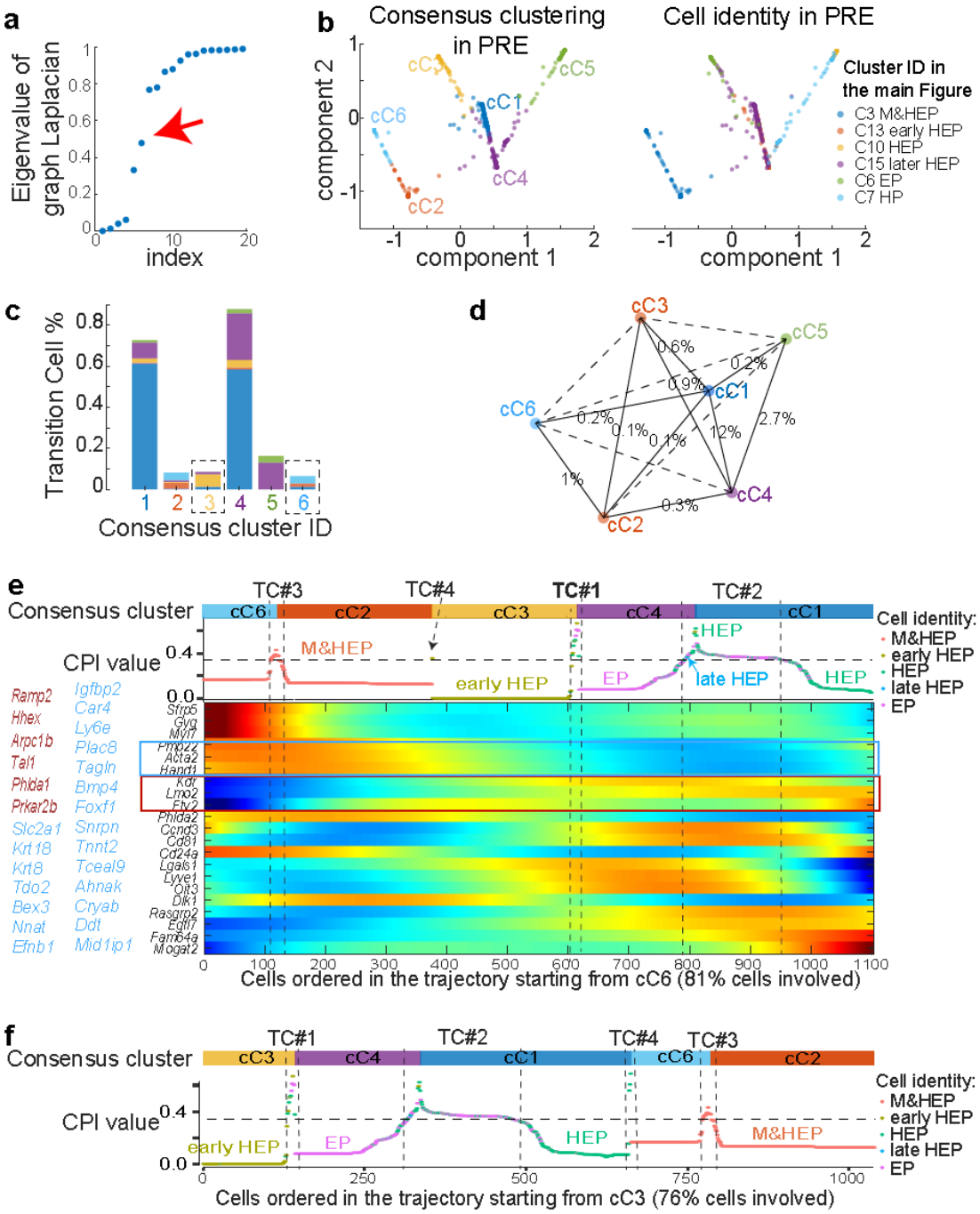


Fig S10. Applying QuanTC (k=6) to the E8.25 2019 dataset, related to Figs 4-5.

a-e, Same as **Fig S9**, except k=6 (panel a, red arrow) was tried, where the 2nd largest gap along the y-axis is observed. Panel e has cC6 chosen to be at the start of the transition trajectory.

f, Histogram of CPI along the trajectory starting from cC3. Each dot is a cell colored with cell identity, same colors as in panel b 'consensus clustering in PRE'. Three predicted transition-cell (TC) populations (#1, #2, #3) are the same as in panel e.

Figure S11

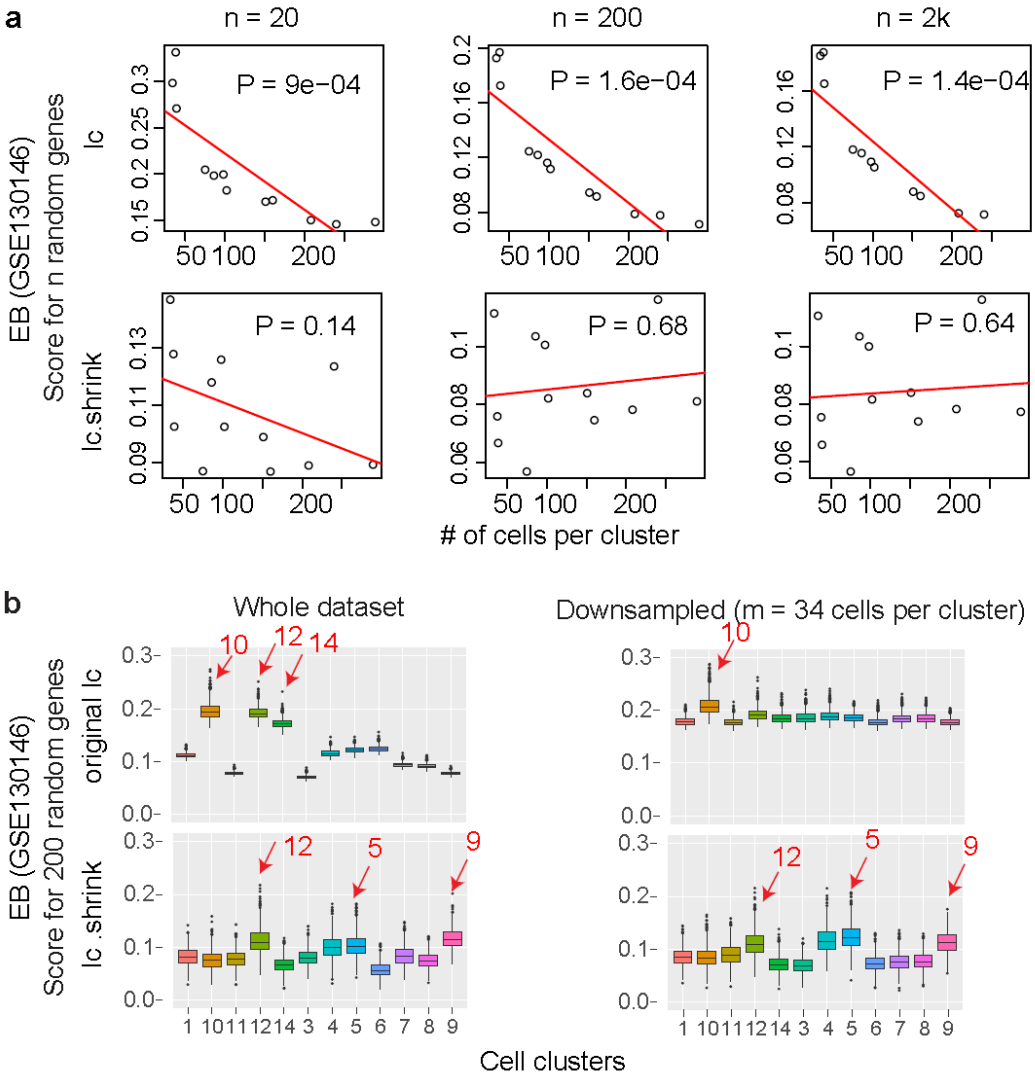


Fig S11. Comparing Ic.shrink with the existing Ic methods, related to Fig 5

a-b, Same as **Fig 5 a-b**, except applied to the EB dataset.

Figure S12

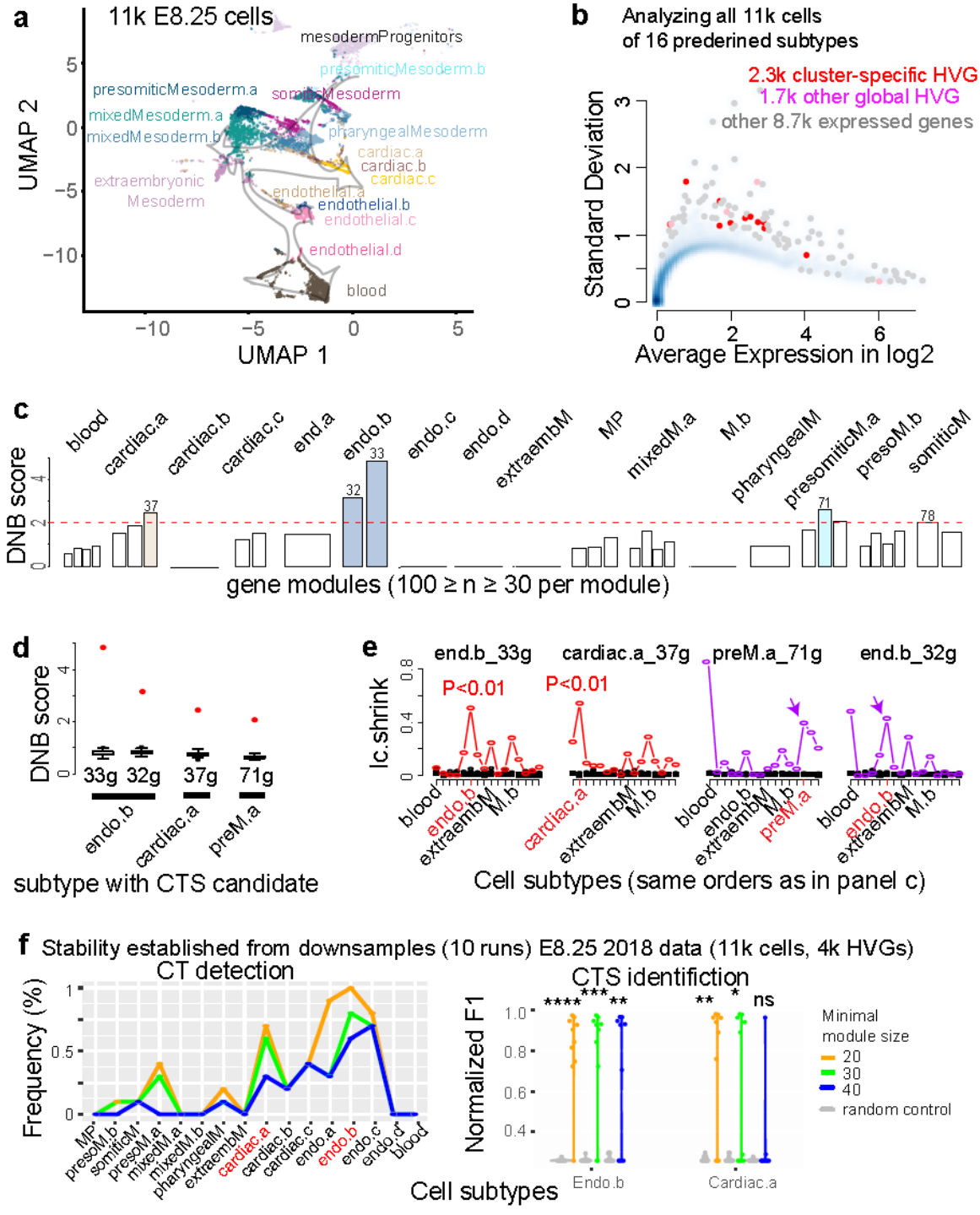


Fig S12. Applying BioTIP to the E8.25 2018 dataset, predefined subcell types, related to Fig 5.

a-f subpanels are presented in the same way as **Fig 3**. One exception is that in panel e, purple arrows point to where the observed score at the intended cluster (labeled red in axis) failed to be the highest score in the system and is rejected.

Figure S13

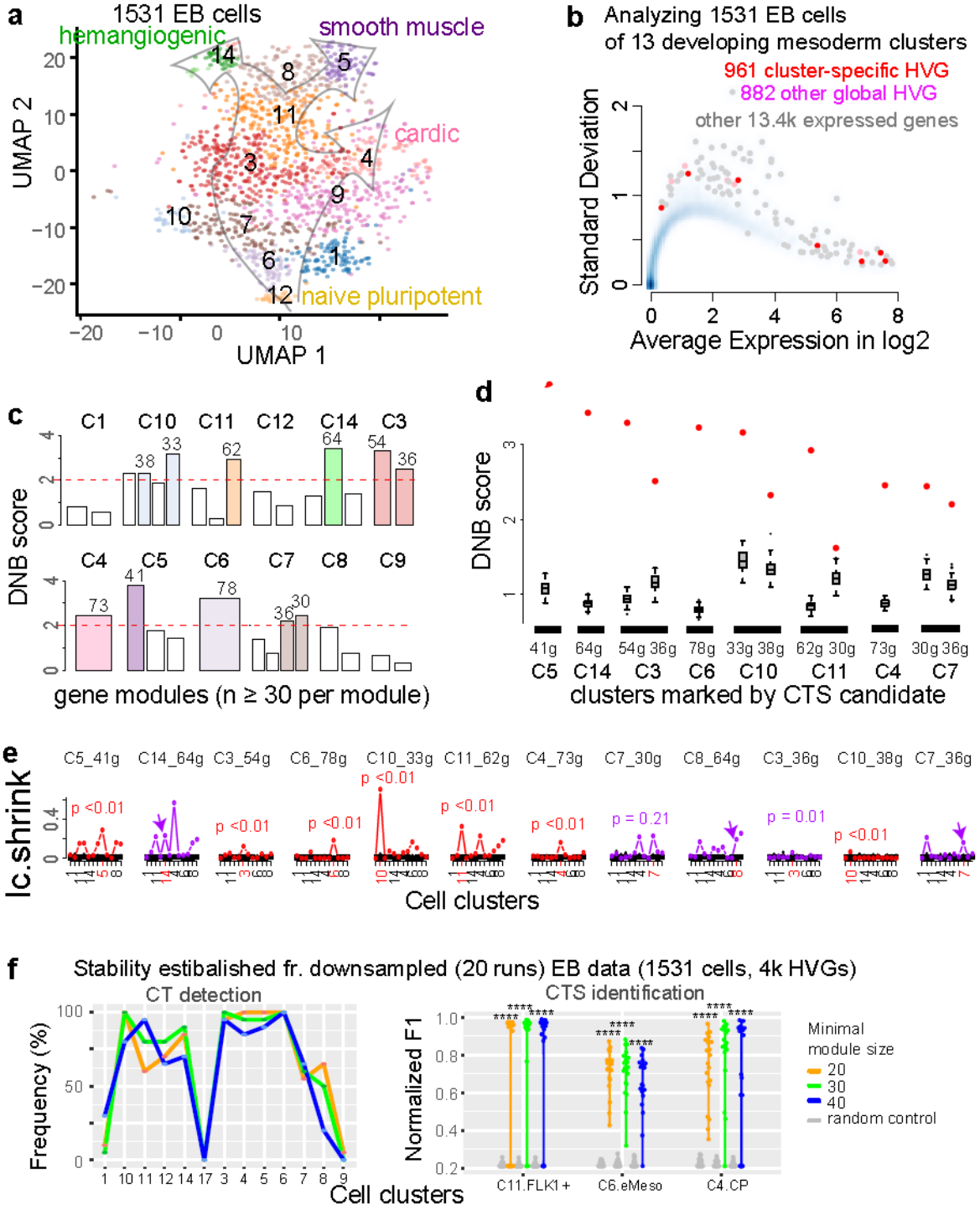


Fig S13. Applying BioTIP to the EB dataset, related to Fig 5.

a-f subpanels are presented similar to **Fig 3**. One exception is that in panel e, purple arrows point to where the observed score at the intended cluster (labeled red in axis) failed to be the highest. FLK1+: FLK1-expressing mesoderm; eMeso: early mesoderm; CP: cardiomyocyte progenitor.

Figure S14

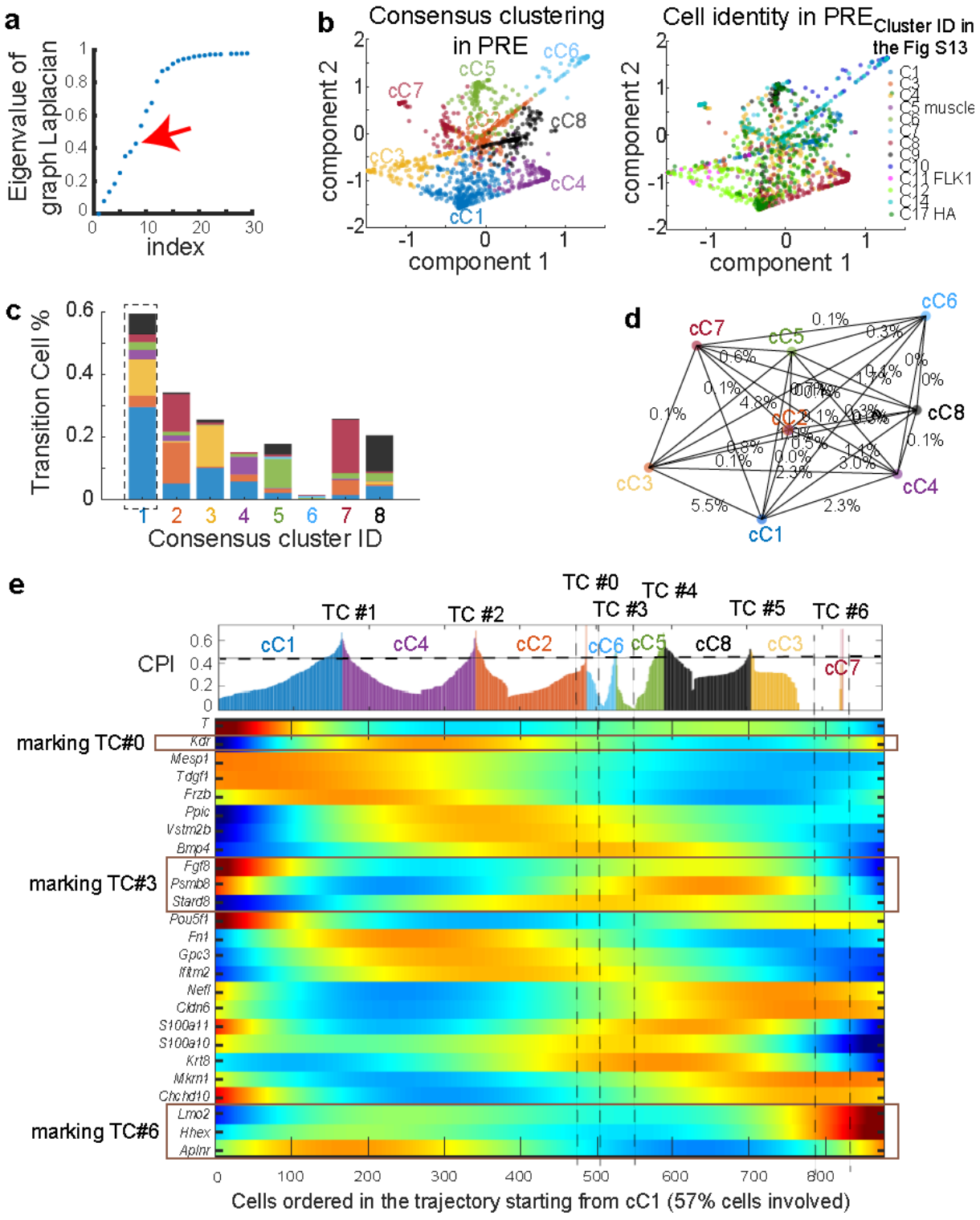


Fig S14. Applying QuanTC (k=8) to the simulated EMT dataset, related to Fig 5.

a-e subpanels are presented the same as **Fig S4** except applied to the simulated EMT dataset.

Figure S15

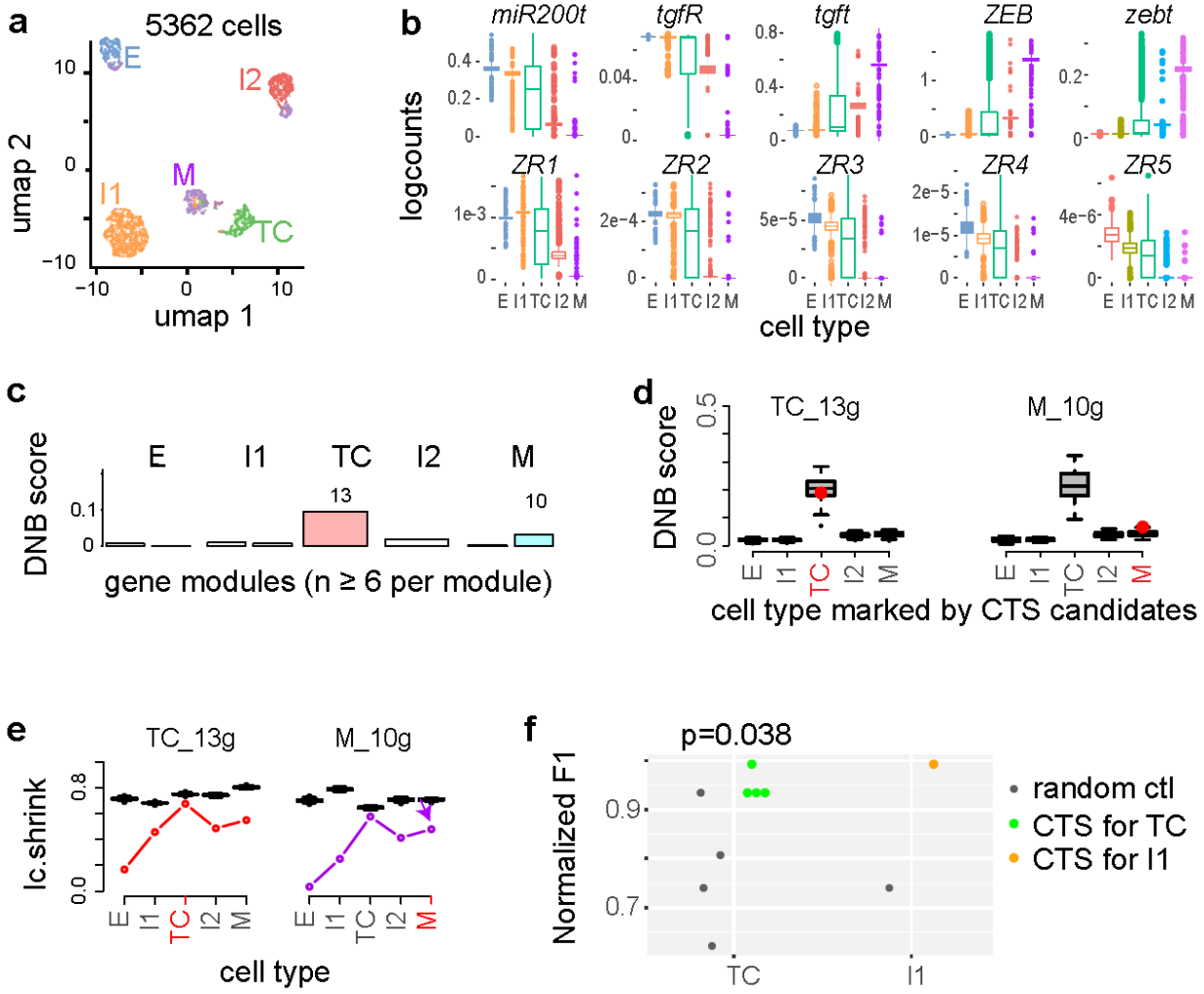


Fig S15. Applying BioTIP to the simulated EMT dataset, related to Fig 5.

a-e subpanels are presented in the same way as **Figure 2**. In panel **e**, purple arrows points to where the observed score at the intended cluster (labeled red in axis) failed to be the highest score in the system and is rejected. Note the observed DNB and Ic.shrink scores fall into the range of their 'random' controls (c-d). This is because there are only 18 selected genes measured in this data, we cannot simulate true random control. Therefore, we simplify call the cluster with the highest score as the identification in this system, resulting in one identification of 13 genes at the cluster of transition cell (TC).

f, Stability of BioTIP estimated from 5362 cells of 4 predefined clusters and 18 genes after down-sampling 95% genes and 95% cells (20 runs). The normalized F1 score for CTS identification at two CT clusters (TC and I1) is graphed. F1 compares each run and the identified CTSs from the whole dataset. Shown are those positive F1 scores (outputs with commonly identified genes) and the t-test statistics.

Figure S16

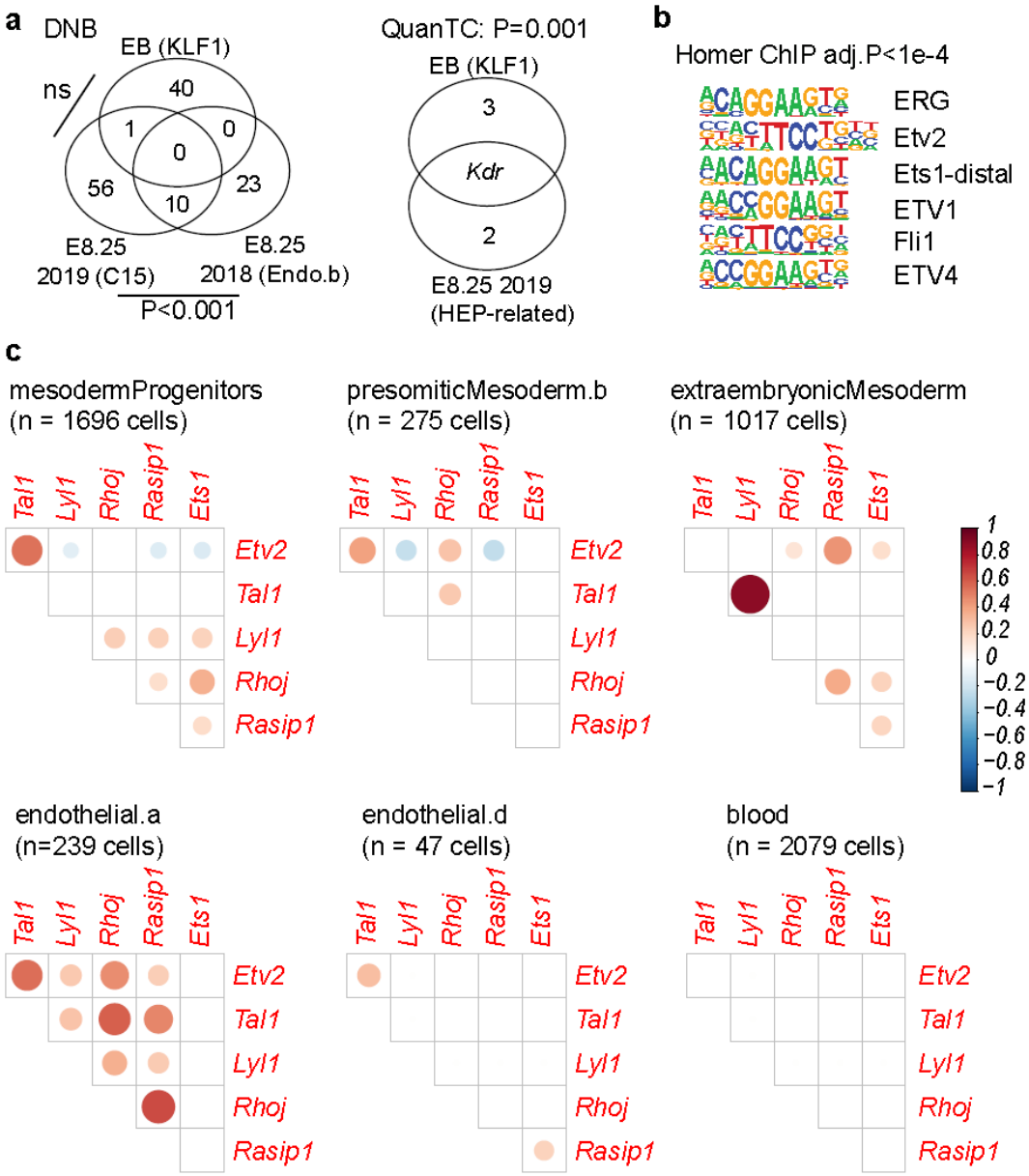


Fig S16. Computational evaluation of Etv2 targets, related to Figs 5, 6

a, Venn diagram comparing the CTS identification by DNB and the transition markers identified by QuanTC from three datasets. Among QuanTC's many predictions, the transition related to haemato-endothelial bifurcation are considered. Also shown is the empirical p-value to observe the overlap among three global HVG.

b, ETS-binding motifs are enriched at the proximal promoters ([-200, 100] around TSS of the 60 early HEP CTS genes.

c, Pairwise correlation between Etv2 and its 5 direct targets in the independent E8.25 2018 dataset. We observe distinct patterns for each gene pair. Pearson coefficients P<0.05 are shown.

Supplementary methods

1. Predicting upstream regulatory transcription factors

We exam four pieces of evidence – IPA (Kramer et al., 2014), TF-binding motif, and ChIP-seq, and literature – to predict the upstream regulator of identified CTS genes. IPA is a curated repository (QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis) (Kramer et al., 2014). IPA explains the expression changes of CTS genes with regulators whose changes in expression are relevant to what is expected from the literature. The cutoff settings were FDR<0.005, at least 10% of target genes, and molecular type='transcription regulator' or 'growth factor'. Because transcriptional regulation requires the binding of transcription factors (TFs), we also search for enriched 'known' TF-binding motifs from curated repositories which are mostly based on the analysis of public ChIP-Seq data sets, using Homer software (Heinz et al., 2010). Significance settings were Benjamini-adjusted p<0.005 and at least 20% of target promoters ([-200,100] around TSS) with a known motif. We further analyze TF-promoter interactions derived from ChIP-seq data for knowledge discovery, using Homer software. Promoters were directly extracted from the assembled transcripts of RNA-seq but were retrieved from the Ensembl (GRCm38_release97) annotations for the mouse gene symbols. Promoters overlapping with the blacklist were removed (Amemiya et al., 2019). We considered significance at a level of Benjamini-adjusted p<0.005 and at least 20% of target sequences with a known motif.

2. Network partition

We calculate co-expression between all pairs of cluster-specific HVG across cells in that cluster and build a cluster-specific network. A parameter FDR $< x$ of between-gene correlation impacts the node (gene) size of this network, for which we use a value from 0.05 to 0.2 to focus on 50~200 genes per cell cluster (**Table S1**). Then, we construct communities in this network via random walks. The ideal is that short random walks tend to stay in the same community (Pons and Latapy, 2005). This step partitions co-expressed, cluster-specific HVG into modules.

3. Evaluating stability and robustness

a. Inferring proxy gold standard (GS)

A GS per dataset is needed for a quantitative comparison. For CT detection, the established transition state or the consistently identified state(s) by BioTIP and QuanTC serve as a proxy gold standard. For CTS identification per transition state, the consistently identified CTS-member genes by BioTIP running on variable clusters or by QuanTC serve as a proxy GS (**Fig S1**). An exception is the estimation of the largest E8.25 2018 data (11k cells), for which we simply use the BioTIP's prediction from the whole dataset as a proxy GS.

*b. Evaluating stability (**Figs 2f, 3f, S7d, S12f, S13f, S15f**)*

We modify the methodology described in a previous benchmark (Saelens, Cannoodt et al. 2019) to evaluate the stability of BioTIP. The stability is evaluated by the similarity between each output using down-sampled data and the (proxy) gold standard inferred from the whole dataset. We sample 95% of the cells and 95% of the genes iteratively and then apply BioTIP, doing *n* (20 or 10) iterations.

Regarding CT detection, we calculate the frequency to predict each cluster using down-sampled data. We ask whether among *n* runs, the CT cluster(s) serving as a GS present the highest frequency (or consistency). We first input our original clustering results and parameter settings into BioTIP to calculate this frequency. Additionally, we calculate the frequency when tuning the parameters: minimum DNB cutoff and minimum gene module size.

Regarding CTS identification at the best-known transition state in a system, we calculate the F1 score that quantifies the similarities of two sets of genes (Saelens, Cannoodt et al. 2019). Given two gene sets, one is a CTS predicted in run *i* and another is the GS genes per dataset, the agreement is quantified by counting their shared members based on the Jaccard similarity. Let $m \ni \{g\}$ denote a prediction (*i.e.*, the significant module we named CTS), $m' \ni \{g_{GS}\}$ the set of the GS genes (*i.e.*, consistently identified CTS member genes in this dataset), and |.| the cardinality operator which counts how many elements are in a set. This gives:

$$Jaccard_{gene}(m, m') = \frac{|m \cap m'|}{|m \cup m'|}$$

For each run with significant CTSs using down-sampled data, there will always be a GS-best-matched CTS prediction even this run fails to predict the best-known transition. Therefore, we introduce a numerical weight to indicate whether the best-matched prediction represents the GS transition state. Then, we define the 'Recovery' as the maximal similarity between all predicted CTSs and this GS gene set. We also defined the 'Relevance' as the average maximal similarity to this GS set for every predicted CTSs. Since BioTIP can predict multiple CTSs, this Relevance estimates the precision of CTS identification. A harmonic mean between Recovery and Relevance gives an F1 score for the best-known transition per dataset. Let M_i be the set of predicted CTSs in run *i* from a dataset. This gives the $F1$ score:

$$x = argmax_{m \in M_i} Jaccard_{gene}(m, m')$$

$$w(m, m') = \begin{cases} 1, & x \text{ and } m' \text{ represent the same state} \\ -1, & x \text{ and } m' \text{ represent different states} \end{cases}$$

$$Recovery = max_{m \in M_i} Jaccard_{gene}(m, m') * w(m, m')$$

$$Relevance = \frac{1}{|M_i|} \sum_{m \in M_i} \{Jaccard_{gene}(m, m') * w(m, m')\}$$

$$F1 = 2 * \frac{Recovery * Relevance}{Recovery + Relevance} = \frac{2}{\frac{1}{Recovery} + \frac{1}{Relevance}}$$

*c. Evaluating robustness against clustering methods (**Fig S1**)*

Similarly, we modify the methodology described in a previous benchmark (Saelens, Cannoodt et al. 2019) to evaluate the robustness of BioTIP. The robustness is discussed when iteratively applying BioTIP to a dataset, after the same cells have been clustered by different clustering methods with variable parameters. Each clustering procedure *i* generates a new set of clusters, and BioTIP subsequently has a new prediction.

We first map all new cell clusters to every GS cluster to mimic a proxy GS for this new set of clusters. This is because the proxy GS for transition states is induced based on a reliable clustering method (red x-labels in **Fig S1** and presented in **Fig 5c** which are either predefined or our originally chosen clustering method). Given new cell cluster assignments of the same dataset, we find the GS-best-matched cluster based on the Jaccard similarity of cells. We calculate a $Jaccard_{cell}(Cell, Cell')$ metric, where $Cell \ni \{cells\}$ denotes a cell cluster, and $Cell' \ni \{cells_{of\ a\ transiiton\ cluster}\}$ denotes the transition cells that served as GS. Among a new set of clusters in procedure *i*, the one having the highest $Jaccard_{cell}$ score will represent the GS state for this prediction in this dataset.

Regarding CT detection per clustering procedure *i* per dataset, we then compute the $Jaccard_{cluser}$ score between the new set of predicted transition clusters $\{C_i\}$ and its mimicked proxy GS clusters $\{C_{GS}\}$. This gives:

$$Jaccard_{cluster}(C_i, C_{GS}) = \frac{|C_i \cap C_{GS}|}{|C_i \cup C_{GS}|}.$$

Regarding CTS identification, we calculate the $F1_i$ score as above described, where *i* indexes cell-clustering procedures.

*d. Quantitative method comparison (**Fig 5, d-e**)*

For each of the 6 datasets, we apply four methods (DNB, Ic, BioTIP, QuanTC) and get a set of CT clusters per method. Each CT prediction is compared to the GS transition states by the $Jaccard_{cluster}$ score. For the three methods (DNB, BioTIP, QuanTC) that also predict transition markers, we assess their predicted gene sets by the similarity to the GS markers per dataset, using the above-detailed $F1_i$ score, where *i* indexes the studied methods.

e. F1 Normalization

To ensure that easy and difficult datasets have equal influence on method comparisons, we normalize the scores on each dataset as previously described (Saelens, Cannoodt et al. 2019). To normalize, we first scale and center the F1 scores to $\sigma = 1$ and $\mu = 0$, then applied the unit probability function that moved the score values back to the range [0, 1].

4. Presentation of the analyses in six independent datasets

In the main **Figures 2-4** and the supplementary **Figures S12, S13**, and **S15**, we structure figure panels consistently:

a) Describing the data and the system (lineage differentiation and cell clusters),

b) Presenting maker gene expression (for <100 detected genes) or HVG selection (from >10k measured genes),

c) Constructing gene modules,

d) Identifying CTS candidates,

e) Identifying significant CTSs (each characterizing one CT cluster),

f) Evaluating the robustness and stability using down-sampled data (or the reproducibility using an independent dataset in **Fig 4**), and

g) Summarizing the system and pointing the known CT cluster(s), for two benchmark datasets (**Figs 2g, 3g**).

In each figure, we also specify the number of cells measured and the number of analyzed cells. The number of cells measured could be larger than the number of analyzed cells to have a better view of the system.

Identified CTSs are named in the format of its representing cluster ID together with gene numbers. This is because one cluster could have multiple CTS identifications.

BioTIP analysis was conducted using the wrap function in our developed R package. QuanTC analysis was conducted using the optimally-selected number of clusters. Ic analysis was also conducted using the BioTIP R package but with 10-50 randomly selected HVGs.

5. Analysis of the benchmark hESC dataset of early cardiogenesis (Figs 2, S4, S5)

Design. To demonstrate the accuracy of CTS predictions, we reanalyzed the single-cell TR-PCR data with an experimentally validated lineage bifurcation and its marker gene *KIT* (Bargaje et al., 2017). It is a time-course collection of cells when the activin A, BMP4, combined with a Wnt pathway activator were added to induce pluripotent stem (iPS) cells differentiating into cardiomyocytes. Day 2-2.5 was the verified pitchfork bifurcation when multipotent primitive streak (PS)-like progenitor cells branched out into either the mesoderm cardiomyocyte (CM) lineage (marked by *Hand1*) or the competing endoderm (En) lineage (marked by *Sox17*). **Fig S2a** presents the lineage mark gene expression. The expression levels of *DKK1*, *WNT5B*, and *PDGFRA* were highly correlated with the BMP-induced differentiation efficiency towards cardiac cell fate (Bargaje et al., 2017).

Transcriptome. This dataset contains the gene expression profiles of 96 developmental genes for 1,934 cells collected from eight timepoints (Bargaje et al., 2017) (**Fig 2a**). Gene expression matrix, cell collection date, and the cells' consensus cluster IDs were downloaded from the original publication.

Data analysis. We focused on the temporal gene expression profiles of 929 cells at six time points (days 0 - 2.5, and mesoderm-specific day 3) as previously analyzed by Ic (Bargaje et al., 2017). Data for 96 genes and these 929 cells were transformed into a SingleCellExperiment object in R. Data visualization was performed on the reduced dimension of the gene expression space, using the package *scater*.

BioTIP analysis was conducted on the previously defined consensus clusters (Bargaje et al., 2017). When analyzing the author-defined consensus clusters, we excluded the cluster 6 (C6) which is a small population of endoderm-specific cells collected at day 2.5 (n=15, **Fig 2a**). This allows us to focus on the differentiation path from PS towards mesoderm-specific cells. From both analyses, we narrowed down to the top 80% most variable genes of the 96 measured genes, allowing 69-76 genes selected (using the 'optimize.sd_selection' function with default parameters except for cutoff = 0.8). To build gene modules, we used the 'getNetwork' function while controlling the FDR of PCC (≤ 0.2). Finally, we fed to the DNB-scoring system these gene modules and set a minimum model size to be 10 genes for the downstream analyses.

QuanTC analysis was conducted on the same 929 cells that had been analyzed by BioTIP and Ic, using MATLAB version R2020b. The software package was downloaded from https://github.com/yutongo/QuanTC (Sha et al., 2020). A cell-cell similarity matrix was generated using the R package SC3 version 1.18.0 with default parameters except for that gene filter = FALSE. We took the average of the consensus clustering results (k=5:10) to robustly estimate cell-cell similarity. Additional inputs for the QuanTC analysis include the normalized count matrix, the gene symbols, and the cell collection date. We excluded the gene-preselection process, given the small number of measured genes. To identify the transition state and gene signatures, we set the two parameters: number of clusters to be 4 where the largest gap is observed from the sorted eigenvalues of symmetric normalized graph (**Fig S4a**), and the threshold of cell plasticity index (CPI) to select transition cells = 0.34 by default. The transition trajectory from C3 to C2 to C4 (mostly of day 2.5 cells) included 72% of total cells, indicating that this path dominates the cell transitions from PS to cardiomyocytes, at least in this dataset (**Fig S4e**). This analysis detected 13 transition genes from C3 to C2, and one (the endothelial mark *SOX17*) transition gene from C2 to C4.

MuTrans analysis (**Fig S5**) is conducted using MATLAB version R2020b. The software package was downloaded from https://github.com/cliffzhou92/MuTrans-release (Zhou et al., 2021).

6. Analysis of the mouse lung epithelial cells (Figs 3, S6)

Design. This published data (GSE52583) describes how lung-epithelium progenitor cells differentiate into two alveolar types (ATs). Cells were collected from four embryonic days (E): E14.5 (early progenitors), E16.5 (around the critical transition), E18.5 (transitioned to AT1 or AT2 with distinct expression patterns), and mature AT2 lineage cells (Treutlein et al., 2014). Among these four timepoints, a known lineage bifurcation occurs at E18.5, when cells co-

express certain "bipotent progenitor" markers: either *Ager* and *S100a6* for AT1 or *Sftpc* for AT2 (Guo et al., 2019; Treutlein et al., 2014). A previous study using the Ic model found that E16.5 is a CT state (Mojtahedi et al., 2016).

Transcriptome. 198 single-cell transcriptomes from mouse lung epithelium were downloaded from GEO (GSE52583). The downloaded transcript levels (mm10) were already quantified as fragments per kilobase of transcript per million mapped reads (FPKM) followed by a depth-matching process. The phenotypic cell annotations were downloaded from publication (Treutlein et al., 2014).

We analyzed 22,854 mouse RNAs (mm10, including 1.5k annotated non-coding RNAs) after a three-step data pre-processing method. First, we removed the ERCC controls. Then, we annotated these transcripts using the Bioconductor package biomaRt, and finally collapsed the FPKM values for multiple transcripts of the same gene symbol by the mean value. Then, the gene FPKM values were converted to gene counts using the R package Monocle3.

Gene and cell filtering. Genes with a count value below 2e-18 (the left tail of the overall count distribution) were considered 'not expressed.' 10.3k genes that were expressed in at least 10 cells were preserved for downstream analysis. After removing the transcripts having no matched gene symbols in the MM10 genome (biomaRt_2.42.0 or DBI v1.1.0), 15,897 transcripts were preserved for further analysis.

Often, doublets or triplets have roughly twice the mRNA recovered as true single cells. After removing two adult cells with either very low mRNA recovery or far more mRNA than the typical cells, 196 single cells were kept for further analysis.

Data analysis. Cell clustering using with Leiden community detection (k=5), principal component analysis, visualization for marker genes, and pseudo-trajectory construction were performed using the R package Monocle3 [11]. To demonstrate the ability of BioTIP to characterize the tipping-point state along a linear topological trajectory, we focused on 131 cells by excusing the potential AT1-lineage cells.

BioTIP analysis was conducted with the parameters detailed in **Table S1** (**Fig 3).**

QuanTC analysis was performed with a high cut of CPI value (0.5 rather than the default 0.35) to select the most promising transition cells, and other by default parameters (**Fig S6**).

7. Analysis of the mouse E8.25 developing mesoderm 2019 cells (Figs 4, 6, S7-S10)

Design. To demonstrate the application of BioTIP, we analyzed a single-cell map of mouse embryogenesis to study 'tipping points' in development (Pijuan-Sala et al., 2019). We focused on early organogenesis at embryonic day (E) 8.25 when mesodermal layers were connected (Pijuan-Sala et al., 2019). There were 15.9k E8.25 cells measured in this dataset.

Transcriptome. We extracted the single-cell gene expression counts, size factors, normalized values, experimental batches (10x sample IDs), and gene annotations from MouseGastrulationData package in R (Pijuan-Sala et al., 2019). This database also provides cell metadata, batch-corrected principal components, and manually annotated cell types. We focused on 7,240 E8.25 cells after removing cells that were annotated as putative doublets or cytoplasm-stripped nuclei.

Data analysis. We focused on 10.9k expressed genes from the total 29k measured genes that met two criteria: 1) having positive biological components in each dataset after splitting the overall variance of the log-normalized expression into biologically relevant and noise components, and 2) having positive biological components in all combined E8.25 cells while accounting for batch effects and assuming Poisson distribution about the noise.

PCA was performed using these expressed genes, and the first 10 principal components were used for cells clustering. All 7,240 E8.25 cells were clustered into 19 subpopulations (states, **Fig 4a**) after constructing a graph that considers the k=10 nearest neighbors (buildSNNGraph). To detect coherent and 'poorly' separated clusters, we performed silhouette-width analysis on the dimension-reduced expression profiles (top 50 principal components) using the 'approx' function. Silhouette in *bluster* package in R was used to approximate the average distances for faster computation in large datasets. Two clusters (C3 and C10) with mostly negative silhouette-width were 'poorly'-separated cell states (**Fig S7b**).

Marker genes defining each state were identified using the 'findMarkers' function in *scran* (Lun et al., 2016) (Wilcox test, pairwise comparisons, top 5 genes for each comparison) and these were used to annotate clusters based on well-known cell-type specific genes.

C13 and C10 clusters standout, given that they have both common and distinct expression patterns -- both are the sub-clusters of HE progenitors which share up-regulated marker genes *Kdr, Etv2,* and *Lmo2* (**Fig S8a**); other commonly up-regulated genes include both the hematopoietically expressed homeobox gene *Hhex* and the endothelial marker *Slc25a5*, suggesting the multipotentiality. States C13 and C10 also have notable differences in gene expression. C13 has higher average expression of *Kdr* than C10 and is marked by *Etv2* target *Tal1*. C10 is marked by average expression of *Flt1* that can inhibit *Kdr* expression (Koyano-Nakagawa et al., 2015).

Trajectory analysis was performed among all 19 clusters using the minimum spanning tree (MST) algorithm. We first summarize the cells into a smaller set of discrete units and compute cluster centroids by averaging the coordinates of its member cells. We then form the MST across those centroids.

Using BioTIP, we identified four putative CTSs from 19 cell clusters. To this end, we first pre-selected 3,073 highly variable genes (HVG) across all populations, using the 'getTopHVGs' function in *scran* (Lun et al., 2016). From these genes, we narrowed down to the top 10% variable genes within each subpopulation (using the 'optimize.sd_selection' function with

default parameters except for cutoff = 0.1). This second feature selection defined 1.9k state-specific HVG (**Fig 4b**). To build gene modules from these 1.9k genes, we used the 'getNetwork' function while controlling the FDR of PCC (≤ 0.2). Finally, we fed to the DNB-scoring system these gene modules and set a minimum model size to be 60 genes for the downstream analyses (**Fig 4c**). With gene permutation statistics estimated from the 3k HVG, we confirmed the significance of these four CTSs using BioTIP's CTS-scoring plus Ic.shrink-scoring system (**Fig 4, d-e**).

To speed up the computation of the stability, we just ran BioTIP on the same six clusters of 1,362 cells that we have applied QuanTC below.

QuanTC analysis was designed to analyze 13 or fewer cell clusters (Sha et al., 2020). This dataset has 10.9k expressed genes with 19 cell clusters (**Fig 4a**). We focused on six clusters covering the early mesenchyme and haemato-endothelial progenitor (HEP) (C3), three HEP (C13, C10, C15), the hematopoietic progenitor (C7), and the endothelial progenitor (C6). We focus on this subset of clusters to compare QuanTC with BioTIP. Three CTSs for the involving states were used as positive controls and the CTS for excluded state C16 was used as a negative control (**Fig S7f**). There were 1,362 cells in these six clusters.

The QuanTC package was downloaded from https://github.com/yutongo/QuanTC and run with Matlab version R2020b. A cell-cell similarity matrix was generated using the R package SC3 version 1.18.0, with default parameters. To robustly estimate cell-cell similarity, we took the average of the consensus clustering (cC) results (k=5,6,9,10) that better resembled our six clusters in a range between 3 to 10. Additional inputs for the QuanTC analysis include the normalized count matrix, the gene symbols, and the cell collection date. To identify the transition state and gene signatures, we set the threshold of cell plasticity index (CPI) to select transition cells = 0.34 by default and selected 3000 most informative genes in the preprocessing function as previously described (Sha et al., 2020). Additionally, we tried two different numbers of clusters (k=4 and 6, respectively) where the two largest gaps are observed from the sorted eigenvalues of symmetric normalized graph (**Figs S9a, S10a**). With k=4, the transition trajectory from cC1 to cC2 included 65% of analyzed cells, indicating that this path dominates the HEP transitions (**Fig S9e**). This analysis detected one transition trajectory and 42 transition genes, including the BioTIP-detected *Rhoj*, *Rasip1*, but not *Etv2* (**Fig S7f, the 1st column**). With k=6, two potential starting clusters resulted in distinct transition trajectories. One identified trajectory starting from cC6 to cC1 via cC2, cC3, and cC4 included 81% of analyzed cells, indicating that this path dominates (**Fig S10e**). The other identified trajectory starting from cC3 to cC2 via cC4, cC1, and cC6 included 76% of analyzed cells. Along each trajectory four sets of transition genes were predicted (**Fig S7f**).

Validated Etv2-target **(Fig 6c)**. Additionally, we downloaded 73 Etv2 targets that were validated in 4 populations from an *in-vitro* differentiation model of ES cells (Zhao and Choi, 2017). This model utilized *T/Brachyury*, *Etv2*, and *Scl (Tal1)* expression together with *PDGFRα* and *FLK1+* mesodermal markers to track hemangiogenic cell lineage development. By applying a CRISPR

(clustered regularly interspaced short palindromic repeats) screening to this model, Zhao and Choi reported 73 verified Etv2 targets in hemangiogenic cell lineage specification in their Supplemental Data 1 (GSE85641) (Zhao and Choi, 2017).

Etv2 direct targets derived from ChIP-seq binding data **(Fig 6c)**. First, Etv2-binding sites were defined by ChIP-seq in *in vitro* differentiated mouse ES cells (iER71) at day 3.5 (GSE59402) (Liu et al., 2015). In these ES cells, Etv2-expression was induced from day 2 to 3.5, a time frame when *Etv2* is normally expressed (Liu et al., 2015). To extract Etv2-targets, we first selected 11.2k loci from the reported peaks (GSE59402) that are reproducible in at least two out of three experimental conditions (wild-type and cells with induced Etv2 that measured by two independent antibodies). We then overlapped these loci to gene promoters ([-2500, 1000] around the TSS annotated in the mouse genome NCBIM37) and extracted the unique gene symbols, using the ChIPpeakAnno (Zhu et al., 2010) package in R. This step identified 1079 unique gene symbols. The authors also published 15 evaluated Etv2 target genes, including nine located far [-250k, 81k] away from any TSS. Merging these 15 evaluated targets, we got 1087 unique gene symbols as Etv2 direct targets.

8. Analysis of the mouse E8.25 developing mesoderm 2018 cells (Figs 4, 5, S8, S12, S16c)

Design. We test the reproducibility of the four CTSs identified from one mouse E8.25 dataset (Pijuan-Sala et al., 2019) in this independent dataset of mouse gastrulation at E8.25 (Ibarra-Soria et al., 2018).

Transcriptome. The normalized count matrix of 20,809 genes, cell labeling of 33 previously defined subtypes (19.4k cells) were downloaded from ArrayExpress (Access number E-MTAB-6153). These subtypes span from mesoderm progenitor to blood and endothelium. We log-transformed the normalized counts (y=log2(x+1)).

Data analysis. To annotate all gene ids in the data matrix, we queried the *AnnotationHub* package for Mus musculus genes (Ensembl release 103). We removed genes that have been annotated to abnormal chromosomes, remaining 20.5k genes. PCA analysis was run based on the top 10% HVGs selected with the *scran* package. And Horn's parallel analysis was performed to choose 21 principal components to retain. UMAP analysis was then conducted on the selected principal components. Upregulated marker genes for these subtypes were identified using the t-test with the function *findMarkers* in the R package scran and the criterial: logFC>1, FDR<0.01, and rank≤ 10 (**Fig S8b**).

To validate four identified CTSs, 16 developing mesoderm cells subtypes (11,039 cells) were analyzed (**Fig 4, f-h**). These 16 subtypes were extraembryonicMesoderm, endothelial.a, endothelial.b, endothelial.c, endothelial.d,blood, mesodermProgenitors, presomiticMesoderm.b, presomiticMesoderm.a, somiticMesoderm, mixedMesoderm.a, pharyngealMesoderm, mixedMesoderm.b, cardiac.a, cardiac.b, and cardiac.c. Ic.shrink scores were calculated using

the log-transformed normalized counts (**Fig 4f**). Between-gene correlations were calculated on the log-transformed normalized counts (**Fig S16c**).

Independent BioTIP analysis was conducted with the parameters detailed in **Table S1** (**Fig S12**). Additionally, Ic scores (**Figs 5, S3**), and DNB scores (**Fig 5**) were calculated using the BioTIP R package we developed.

9. Analysis of mouse *in-vivo* embryo body (EB) cells (Figs 5, S13, S14)

Design. In vitro differentiation model of embryonic stem (ES) cells has been extensively used to study lineage development. The ES model overcomes the cell-number limitations for early embryonic studies, thus allowing large-scale transcriptomic snapshots. We reanalyzed the published scRNA-seq data of day-4 EB cells during differentiation when hemangiogenic cells extensively emerge and pluripotent stem cells are still present (Zhao and Choi, 2019).

Transcriptome. We downloaded the 10x genomics matrixes from GEO (GSE130146).

Gene and cell filter. We kept genes expressed in at least one cell for analysis. We removed cells with more than 5% mitochondria reads or fewer than 2000 unique genes detected. By these filtering, we kept 33,456 genes and 1,731 cells for further analysis.

Data analysis. We normalized the gene expression values by library size factors, using the R package scran (v 1.18.0). Based on the 1k HVG, we performed dimension reduction and cell clustering. All 1,731 cells were grouped into 17 clusters based on a nearest-neighbor graph clustering method, considering 5 nearest neighbors. We then annotated cell identities according to the average expressed gene markers per cluster and constructed the developmental trajectory. After dropping off 4 clusters of endoderm or primordial germ cells, we kept 1,531 cells along the path from naive pluripotent cells to either hemangiogenic or smooth muscle lineages for further analysis.

BioTIP analysis was conducted on 13 reliable clusters of 1531 genes (**Fig S13a**). We excluded cluster 17 which is a small proportion of blood progenitor cells (n=11). We used the master function in the BioTIP R package to do the analysis. All parameters are given in **Table S1**.

QuanTC analysis was performed with a high cut of CPI value (0.45 rather than the default 0.35) to select the most promising transition cells, and other by default parameters (**Fig S14**). Given 8 consistent clusters (cC) defined by QuanTC, we had multiple options on choosing a trajectory starting point. **Fig S14e** showed the results of one option that had more cells (57%) involved than other options.

10. Analysis of the simulated EMT dataset (Figs 5, S15)

Design. Reversible epithelial-to-mesenchymal transition (EMT) is a prominent example of saddle-node bifurcations in development. During this transition, the balance between EMT-

promoting and EMT-inhibiting factors is a critical causal of intermediate states. The established promoting factors include Zeb1, Snail, and TGFβ, while the inhibiting factors include Ovol2 and miR200t (Hong et al., 2015). The QuanTC method has previously described model-free simulation dataset of a EMT transition. This is a mimicked single-cell expression profile of four states during epithelial-mesenchymal transition. One intermediate state (I1) has the highest proportion of mimicked transition cells.

Transcriptome. We downloaded the simulation dataset in July 2021 (https://github.com/yutongo/QuanTC). This dataset has 18 genes expressed in 5,363 cells from five distinct states (one epithelial state (E), two intermediate cell states (I1 and I2), one mesenchymal state (M), and mimicked transition cells (TC)) (**Fig S15a**). The highest proportion of TC were found in both E and the intermediate state closer to the E (I1). QuanTC detected an intermediated cell state (ICS) between E and I1 with eight transition signal genes including Ovol2 and miR200t (Sha et al., 2020). These genes fluctuate in I1 and highly expressed in transition-involving states E or I2 (**Fig S15b**).

Data analysis. Downloaded expression counts were inputted into Monocle 3 for normalization and then log2-transformation. With the log-transformed normalized counts, cells were classified into four clusters using Monocle 3 (considering k=200 nearest neighbors). Pearson correlations were calculated between gene pairs in each cluster.

BioTIP analysis was first performed on the normalized and log-transformed matrix. No feature selection was conducted due to the low-dimensionality of the data (18 genes). To build gene modules, we used the 'getNetwork' function while controlling the FDR of PCC (≤ 0.05). We fed these genes into the CTS-scoring system and set a minimum model size to be 6 genes (30% of all detected genes) for the downstream analyses. We failed to observe significance in either DNB-scoring or Ic.shrink-scoring systems. This is because a successful significance assessment required large global HVG, but this data has only 18 genes. Instead, we computationally evaluated the top two DNB-scored modules by observed Ic.shrink scores. We accepted one module whose Ic.shrink score peaked at the intended cell cluster. This module characterized the simulated transition cells which is intended.

To compare with QuanTC for the robustness to different clustering methods, BioTIP analysis was also respectively performed on the six and four consensus clusters that QuanTC had analyzed.

11. Analysis of chromatin accessibility (Fig 6e)

Design. Chromatin accessibility is associated with cell-type-specific transcription factor activity that regulates target gene expression. To understand how the expression threshold of Etv2 functions during HE bifurcation, we asked if the oscillation of Etv2 expression causes epigenetic changes that are responsible for CTS genes' expression changes.

ATAC-seq data. We downloaded the published chromatin accessibility data (GSE92537) in two types of cells: the Etv2-deficient ($Etv2^{KO}$) control cell and the endothelial cell mimicking a normal HEP state by overexpressing Etv2 in Etv2-deficient cell ($Etv2^{KO}$iEtv2) (Duan, 2016).

Data analysis. We visualize the ATAC-seq signal using the R packages profileplyr and Gviz. 60 CTS genes were split into two groups according to the occurrence of the Etv2-binding sites within a window of [-2500, 1000] bp around the TSS (**Fig 6c**). We assessed the differential accessibility between two cell types using the pairwise Wilcox-tested and multiple-testing adjusted p-value (**Fig 6e**, boxes atop).

References for Supplementary Method

Amemiya, H.M., Kundaje, A., and Boyle, A.P. (2019). The ENCODE Blacklist: Identification of Problematic Regions of the Genome. Sci Rep *9*, 9354.

Bargaje, R., Trachana, K., Shelton, M.N., McGinnis, C.S., Zhou, J.X., Chadick, C., Cook, S., Cavanaugh, C., Huang, S., and Hood, L. (2017). Cell population structure prior to bifurcation predicts efficiency of directed differentiation in human induced pluripotent cells. Proc Natl Acad Sci U S A *114*, 2271-2276.

Duan, D. (2016). Transcriptional Regulation of Hemato-vascular Lineage Specification during Embryogenesis. In Molecular, Cell and Developmental Biology (California, United States: University of California, Los Angeles), pp. 176.

Guo, M., Du, Y., Gokey, J.J., Ray, S., Bell, S.M., Adam, M., Sudha, P., Perl, A.K., Deshmukh, H., Potter, S.S., *et al.* (2019). Single cell RNA analysis identifies cellular heterogeneity and adaptive responses of the lung at birth. Nat Commun *10*, 37.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Hong, T., Watanabe, K., Ta, C.H., Villarreal-Ponce, A., Nie, Q., and Dai, X. (2015). An Ovol2-Zeb1 Mutual Inhibitory Circuit Governs Bidirectional and Multi-step Transition between Epithelial and Mesenchymal States. PLoS Comput Biol *11*, e1004569.

Ibarra-Soria, X., Jawaid, W., Pijuan-Sala, B., Ladopoulos, V., Scialdone, A., Jorg, D.J., Tyser, R.C.V., Calero-Nieto, F.J., Mulas, C., Nichols, J., *et al.* (2018). Defining murine organogenesis at single-cell resolution reveals a role for the leukotriene pathway in regulating blood progenitor formation. Nat Cell Biol *20*, 127-134.

Koyano-Nakagawa, N., Shi, X., Rasmussen, T.L., Das, S., Walter, C.A., and Garry, D.J. (2015). Feedback Mechanisms Regulate Ets Variant 2 (Etv2) Gene Expression and Hematoendothelial Lineages. J Biol Chem *290*, 28107-28119.

Kramer, A., Green, J., Pollard, J., Jr., and Tugendreich, S. (2014). Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics *30*, 523-530.

Liu, F., Li, D., Yu, Y.Y., Kang, I., Cha, M.J., Kim, J.Y., Park, C., Watson, D.K., Wang, T., and Choi, K. (2015). Induction of hematopoietic and endothelial cell program orchestrated by ETS transcription factor ER71/ETV2. EMBO Rep *16*, 654-669.

Lun, A.T., McCarthy, D.J., and Marioni, J.C. (2016). A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. F1000Res *5*, 2122.

Mojtahedi, M., Skupin, A., Zhou, J., Castano, I.G., Leong-Quong, R.Y., Chang, H., Trachana, K., Giuliani, A., and Huang, S. (2016). Cell Fate Decision as High-Dimensional Critical State Transition. PLoS Biol *14*, e2000640.

Pijuan-Sala, B., Griffiths, J.A., Guibentif, C., Hiscock, T.W., Jawaid, W., Calero-Nieto, F.J., Mulas, C., Ibarra-Soria, X., Tyser, R.C.V., Ho, D.L.L., *et al.* (2019). A single-cell molecular map of mouse gastrulation and early organogenesis. Nature *566*, 490-495.

Pons, P., and Latapy, M. (2005). Computing communities in large networks using random walks. Lect Notes Comput Sc *3733*, 284-293.

Saelens, W., Cannoodt, R., Todorov, H., and Saeys, Y. (2019). A comparison of single-cell trajectory inference methods. Nat Biotechnol *37*, 547-554.

Sha, Y., Wang, S., Zhou, P., and Nie, Q. (2020). Inference and multiscale model of epithelial-to-mesenchymal transition via single-cell transcriptomic data. Nucleic Acids Res *48*, 9505-9520.

Treutlein, B., Brownfield, D.G., Wu, A.R., Neff, N.F., Mantalas, G.L., Espinoza, F.H., Desai, T.J., Krasnow, M.A., and Quake, S.R. (2014). Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature *509*, 371-375.

Zhao, H., and Choi, K. (2017). A CRISPR screen identifies genes controlling Etv2 threshold expression in murine hemangiogenic fate commitment. Nat Commun *8*, 541.

Zhao, H., and Choi, K. (2019). Single cell transcriptome dynamics from pluripotency to FLK1(+) mesoderm. Development *146*.

Zhou, P., Wang, S., Li, T., and Nie, Q. (2021). Dissecting transition cells from single-cell transcriptome data through multiscale stochastic dynamics. Nat Commun *12*, 5609.

Zhu, L.J., Gazin, C., Lawson, N.D., Pages, H., Lin, S.M., Lapointe, D.S., and Green, M.R. (2010). ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. BMC Bioinformatics *11*, 237.