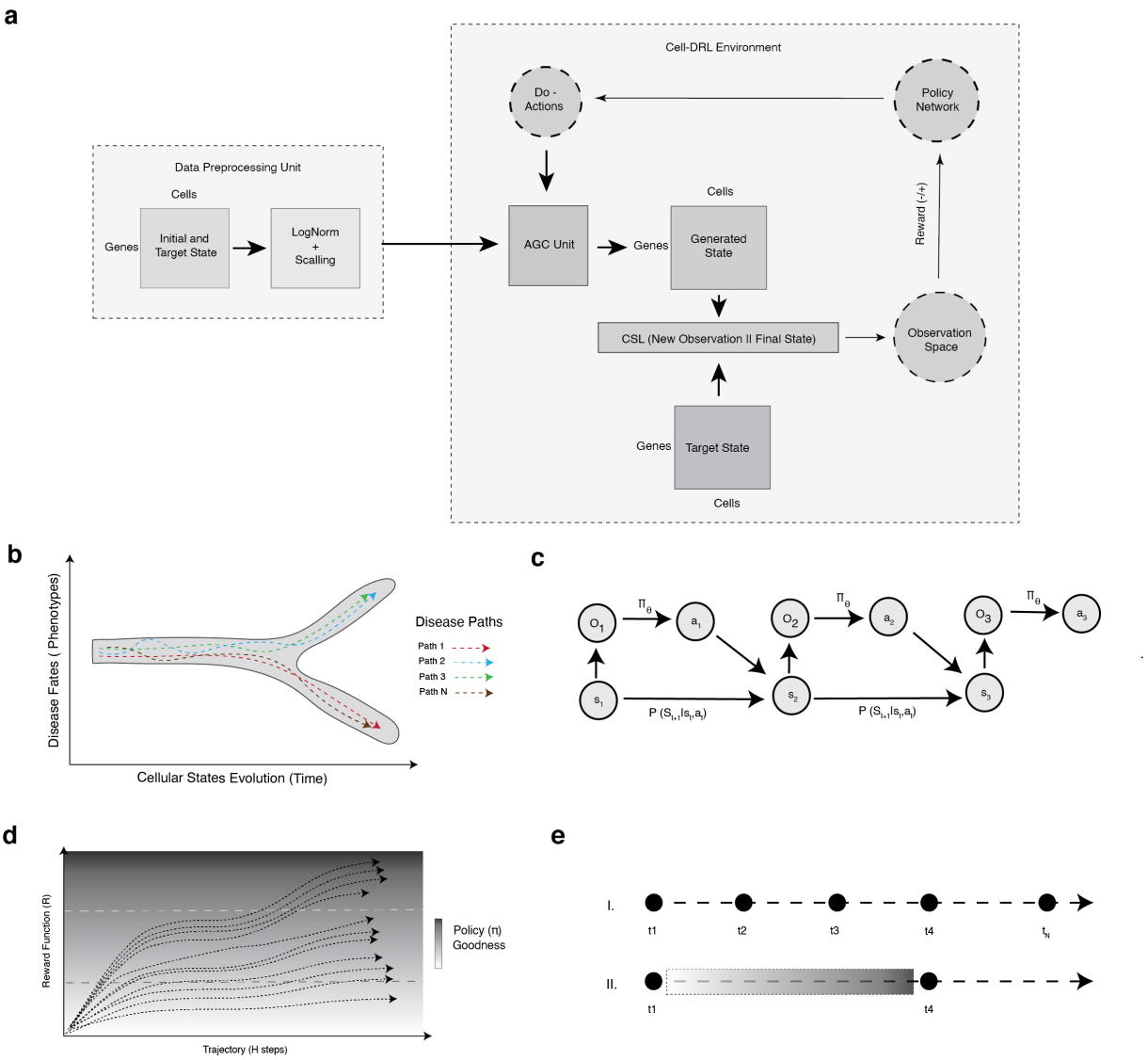
Cell-DRL Reconstructs Unseen Cellular Paths in Health and Disease

Mahmoud et al.

Extended Data Figures 1-17



Extended Data Fig. 1: Cell-DRL environment architecture.

A. Schematic representation of the single-cell RNA-seq data preprocessing unit and Cell-DRL environment architecture. The left side represents the data preprocessing unit where we introduce the initial and target single-cell RNA-seq data and perform normalization and scaling steps. The preprocessed data is then introduced to the Cell-DRL environment. The Cell-DRL environment is composed of an AGC unit where the agent can generate novel cellular states, interacting with the action unit (Do-Unit) that governs the generation process (Gen-Unit). The novel generated state is introduced to the CSL unit to calculate estimated rewards, which are then pushed to the policy network, aiming to learn optimal policies to find the right sequence of actions to maximize overall rewards.

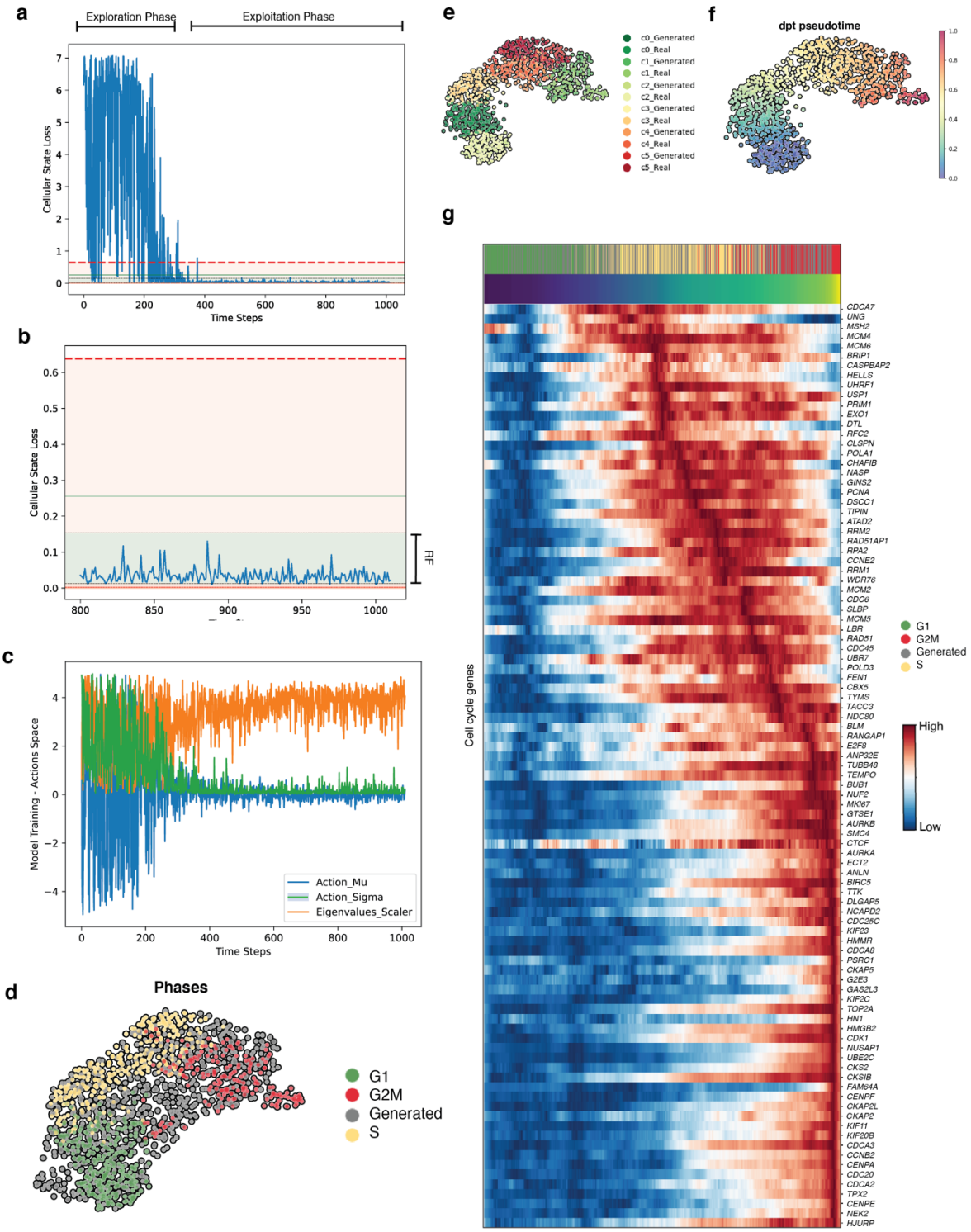
B. Schematic representation of the disease paths that regulate underlying disease state trajectories over our lifetime.

C. Schematic representation of the Markov Decision Process where s_t represents the Cell-DRL environment state, O_t represents the observation space state, and a_t represents the action taken at timepoint t. Π_θ represents the policy network. P denotes the transition probability from the current state s_t to the next state s_{t+1} after the Cell-DRL takes a certain action.

D. Schematic representation depicting the relationship between trajectory steps (H) (x-axis) and the reward (R) function (y-axis). The gray color gradient represents the policy (Π_θ) optimality scale. The grey and white broken lines indicate low and high rewards, respectively.

E. Schematic representation of the two-state challenge, where we introduce data for two states of an evolving biological process data (at t_1 and t_4) to our agent and attempt to reconstruct the hidden

intermediate states. In the real data, we have ground truth for the intermediate states at time points t_2 and t_3, which are used to estimate the reconstruction accuracy of the Cell-DRL agent.



Extended Data Fig. 2: Reconstruction of the cell cycle cellular states trajectory.

A. Diagram illustrating the agent alternating between exploration and exploitation phases during the learning process in the cell cycle cellular state scenario. The x-axis represents the time steps, and the y-axis represents the cellular state loss (CSL) values. The lines indicate the Learning bounds scores (LBS).

B. A zoomed-in diagram displaying the CSL after 800 steps, where we observed that the agent successfully generated novel cellular states in the Reconstruction Field (RF) of interest. The lines

indicate the LBS, with the broken red lines representing the minimum and maximum trajectory bounds, and the broken black lines indicating the minimum and maximum reward bounds. The green-colored area represents the reconstruction field of interest (RF).

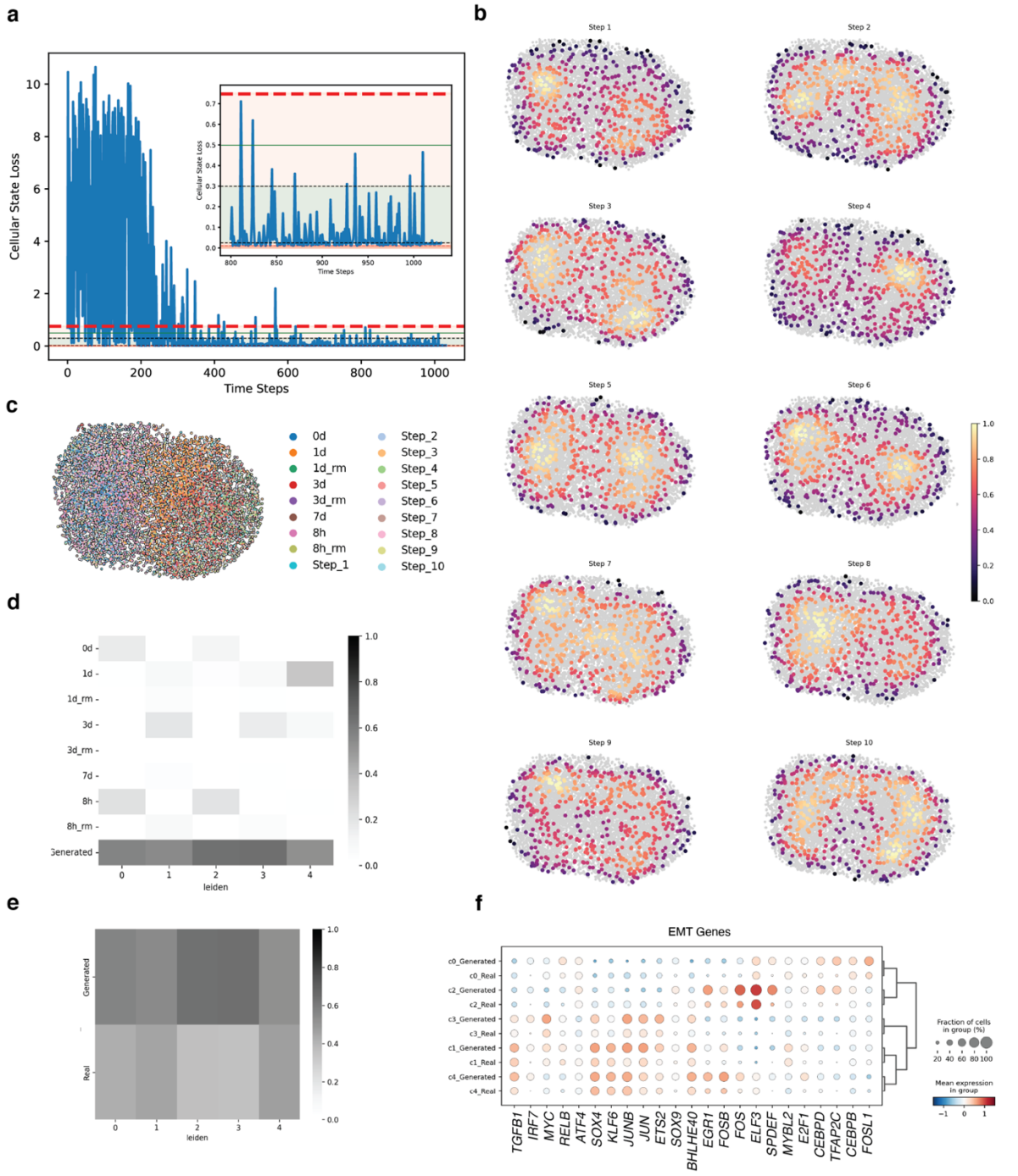
C. Diagram depicting the Do-Unit parameter space values during the Cell-DRL agent training process.

D. UMAP representation displaying the co-embedding of real and generated cell cycle data.

E. UMAP representation displaying the sub-clustering of the co-embedded real and generated cell cycle data.

F. UMAP representation of intermediate cellular states with a color code indicating estimated pseudotime values.

G. Heatmap displaying gene expression of cell cycle genes and the learned latent pseudotime of both real and generated data. The color bars at the top and bottom represent the cellular state groups and the pseudotime scale, respectively. The red heat colors represent high gene expression, and the blue colors represent low gene expression.



Extended Data Fig. 3: Reconstruction of the epithelial-to-mesenchymal transition cellular states trajectory.

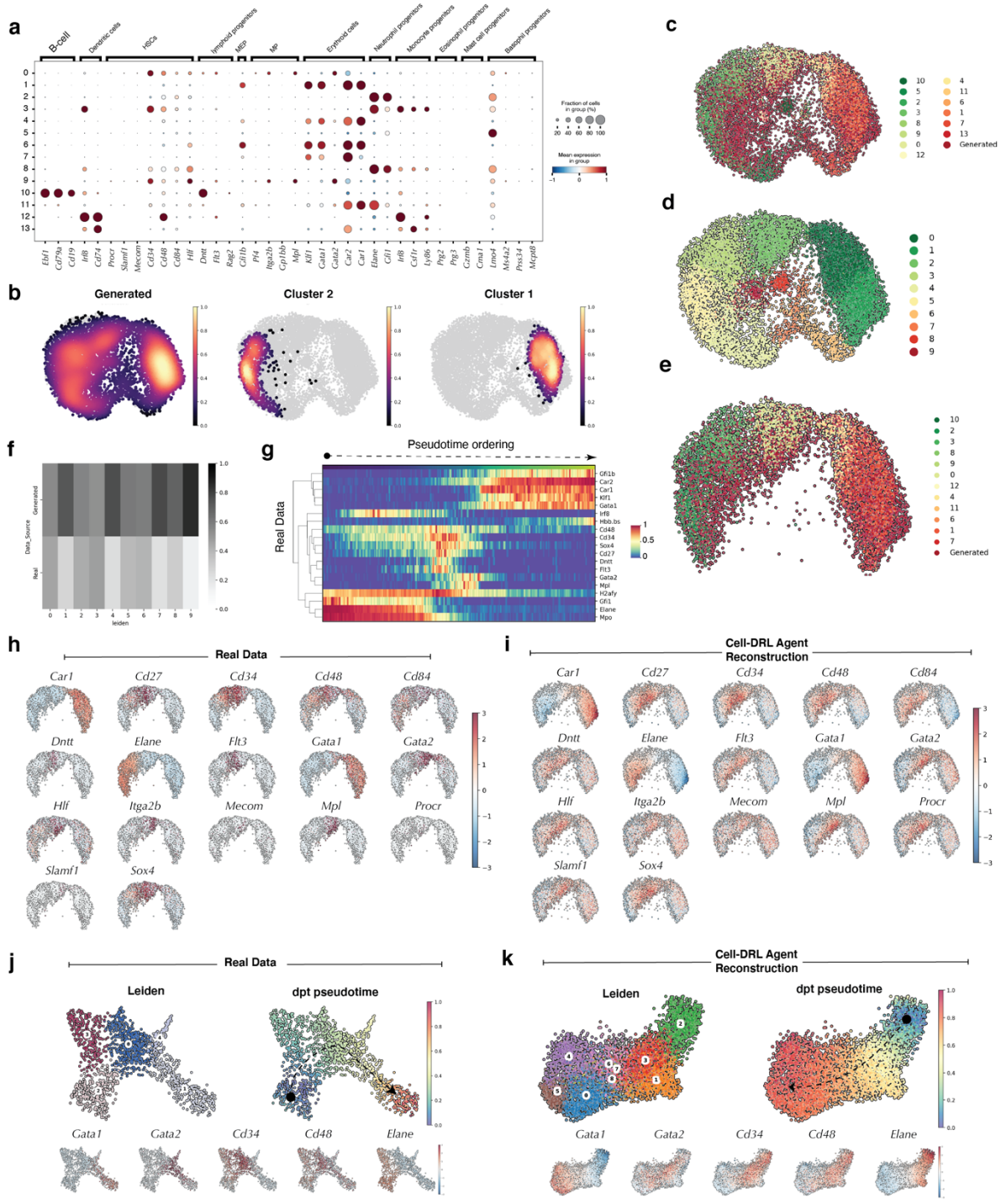
A. Diagram showing the agent's transition between exploration and exploitation phases during the learning process in the EMT state scenario. The x-axis represents the time steps, and the y-axis represents the cellular state loss (CSL) values. The lines indicate the LBS, with the broken red lines representing the minimum and maximum trajectory bounds, and the broken black lines indicating the minimum and maximum reward bounds. The green-colored area represents the reconstruction field of interest (RF).

B. UMAP representation of the generated cellular states at each step of the global generated trajectory.

C. UMAP representation displaying the co-embedding of real and generated data (at each trajectory step of (H)) with real timepoint-specific data.

D. Heatmap displaying the proportions of generated and real data across different time points within each of the Leiden clusters.

E. Heatmap displaying the proportions of generated and real data within each of the Leiden clusters.
F. Heatmap displaying gene expression of EMT genes from both real and generated data for each of the Leiden clusters.



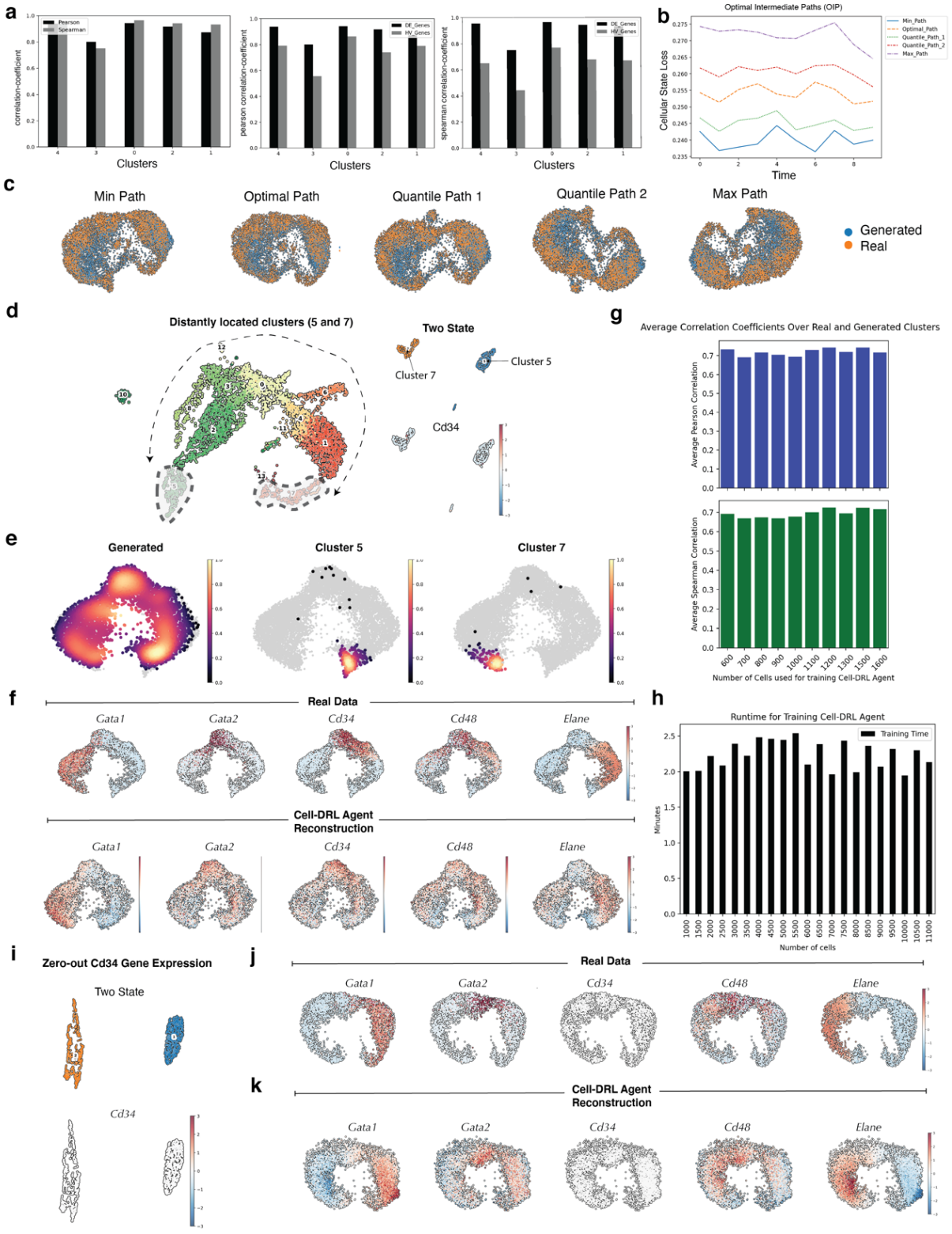
Extended Data Fig. 4: Reconstruction of murine hematopoietic stem and progenitor cellular states.

A. Heatmap showing the gene expression pattern of key stem and progenitor markers across the real single-cell RNA-seq data for each Leiden cluster.

B. Density plot of UMAP representation demonstrating co-localization patterns of real states (clusters 1 and 2) and generated states overlaying the cellular state manifold.

C. UMAP representation displaying the co-embedded real hematopoietic single-cell RNA-seq data clusters and Cell-DRL generated data.

D. UMAP representation of the real and generated data highlighting Leiden clusters obtained after joint re-clustering.
E. UMAP representation of the intermediate cellular state's sub-clusters of real and generated data 2D co-embeddings.
F. Heatmap displaying the proportions of generated and real data within each of the Leiden clusters.
G. Heatmap displaying gene expression of key marker genes and transcription factors regulating HSCs and multipotent progenitor states across latent pseudotime of the real data cellular states.
H. UMAP representation of key hematopoietic stem and multipotent progenitor cell marker gene expression (normalized) of the real gene expression data.
I. UMAP representation of key hematopoietic stem cell (HSC) marker gene expression (normalized) of the Cell-DRL generated cellular states data.
J. UMAP representation displaying the separated embedding and gene expression of the real intermediate hematopoietic cellular states data.
K. UMAP representation displaying the separated embedding and gene expression of the generated intermediate hematopoietic cellular states data.



Extended Data Fig. 5: Cell-DRL reconstruction of hematopoietic stem and progenitor cellular states from distant anchoring states.

A. Barplot indicating Pearson's and Spearman's correlation coefficients between the top 150 differentially expressed (DE) genes for real and generated data, and between highly variable genes (HVGs) within Leiden clusters for real and generated data.

B. Diagram showing the cellular state loss of the top 5 optimal causal paths (OCP) sampled realizations (y-axis) of the Cell-DRL generated states over the sampled trajectory steps (x-axis). The optimal paths correspond to various quantile values “Min_Path = 0.1”, "Quantile_Path_1 = 0.2”, "Optimal_Path = 0.5", "Quantile_Path_2 = 0.7 “, and "Max_Path = 0.9”.

C. UMAP representation displaying the co-embedding of the top 5 optimal causal paths (OCP) generated data with real hematopoietic single-cell RNA-seq data.

D. UMAP representation for the two-state challenge with more distant progenitor clusters (left) and UMAP showing sparse expression of *Cd34* in the anchoring clusters (5 and 7).

E. Density plot on UMAP representation highlighting co-localization of real anchoring state (clusters 5 and 7) data and generated states on the cellular state manifold.

F. UMAP representation of key hematopoietic marker gene expression (normalized) of the real and generated data.

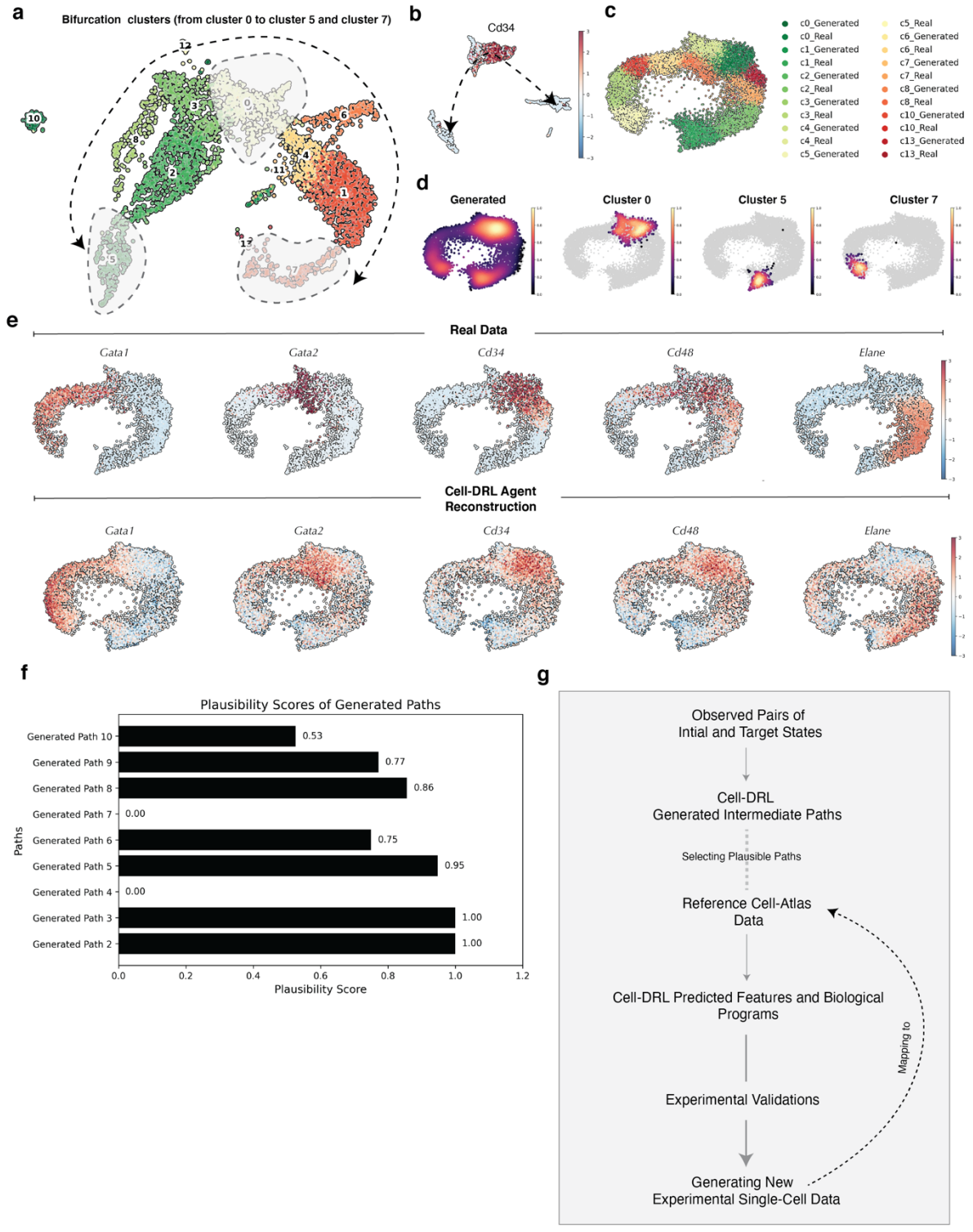
G. Barplot showing the average Pearson’s and Spearman’s correlation coefficients between real and generated clusters after training the Cell-DRL agent with input cell numbers ranging from 600 to 1,600 cells in the anchoring clusters 1 and 2.

H. The barplot displays the number of cells (from the two cellular states used for training) on the x-axis, with the corresponding training runtime in minutes on the y-axis.

I. UMAP representation for the two-state challenge with anchoring clusters 1 and 2 after setting *Cd34* gene expression to zero.

J. UMAP representation of key hematopoietic marker gene expression (normalized) of the real gene expression data for reconstruction after setting *Cd34* expression to zero.

K. UMAP representation of key hematopoietic marker gene expression (normalized) of the Cell-DRL generated cellular states data for reconstruction after setting *Cd34* expression to zero.



Extended Data Fig. 6: Cell-DRL reconstruction of bifurcation events.

A. UMAP representation for the bifurcation challenge with HSCs as the initial cellular state and more distant progenitors' clusters 5 and 7 as anchoring cellular states.

B. UMAP representation of clusters 0, 5, and 7, highlighting normalized expression of *Cd34*.

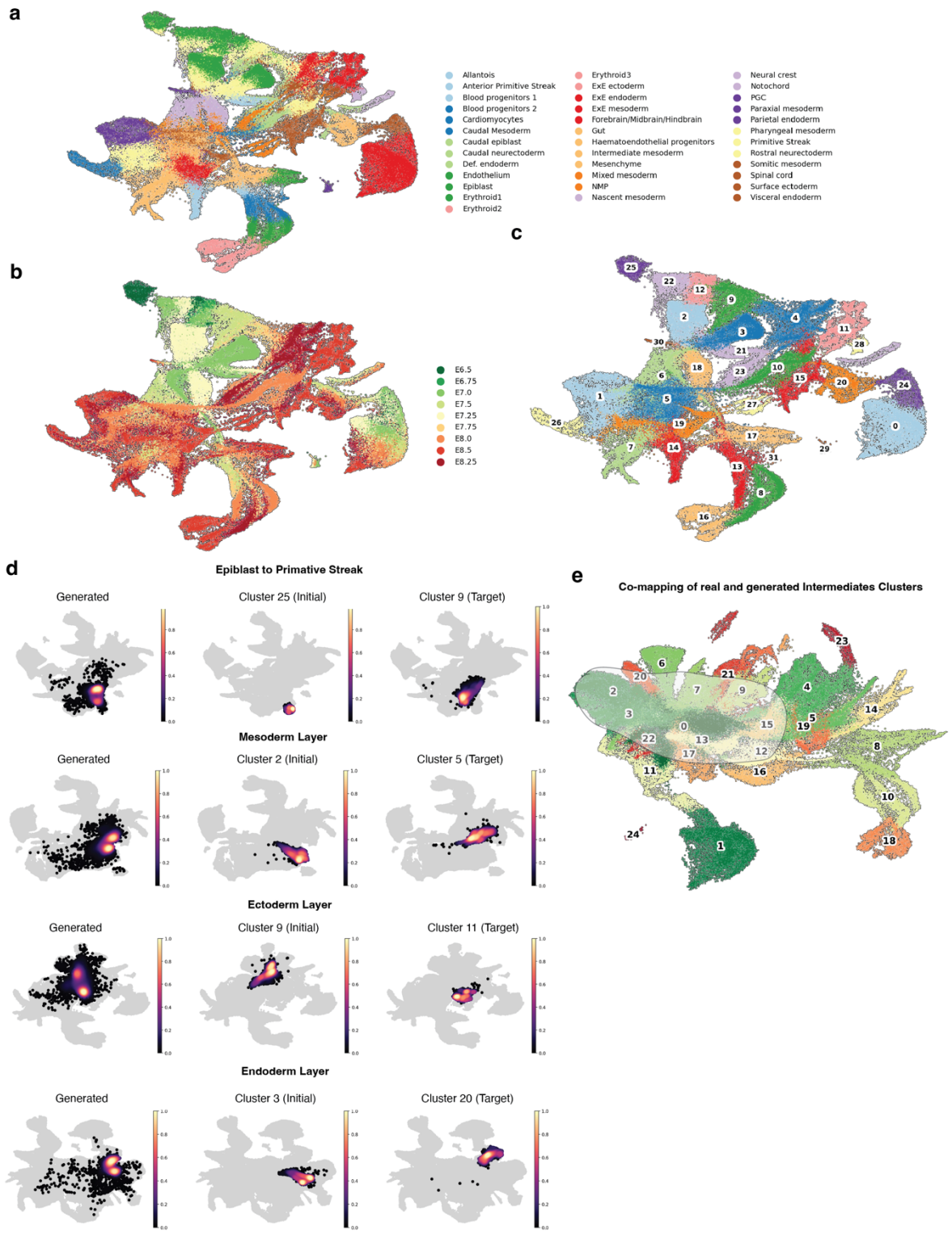
C. UMAP representation displaying the co-embedded real hematopoietic single-cell RNA-seq data and Cell-DRL generated data.

D. Density plot of UMAP representation demonstrating co-localization patterns of real states (clusters 0, 5, and 7) and generated states overlaying the cellular state manifold.

E. UMAP representation of key hematopoietic marker gene expression (normalized) of the real and generated gene expression data.

F. Plausibility Scores of Generated Paths Relative to the reference scRNA-seq dataset (hematopoiesis). The horizontal bar plot illustrates the plausibility scores of different generated paths compared to a reference hematopoiesis dataset. Higher Plausibility Scores indicate higher similarity of the generated paths to the reference single-cell atlas.

G. The figure illustrates the workflow of the Cell-DRL paradigm, where observed pairs of initial and target states are used to generate intermediate paths. These paths are then evaluated against reference data from cell atlases to determine their plausibility.



Extended Data Fig. 7: Cell-DRL reconstruction of embryonic layers development.

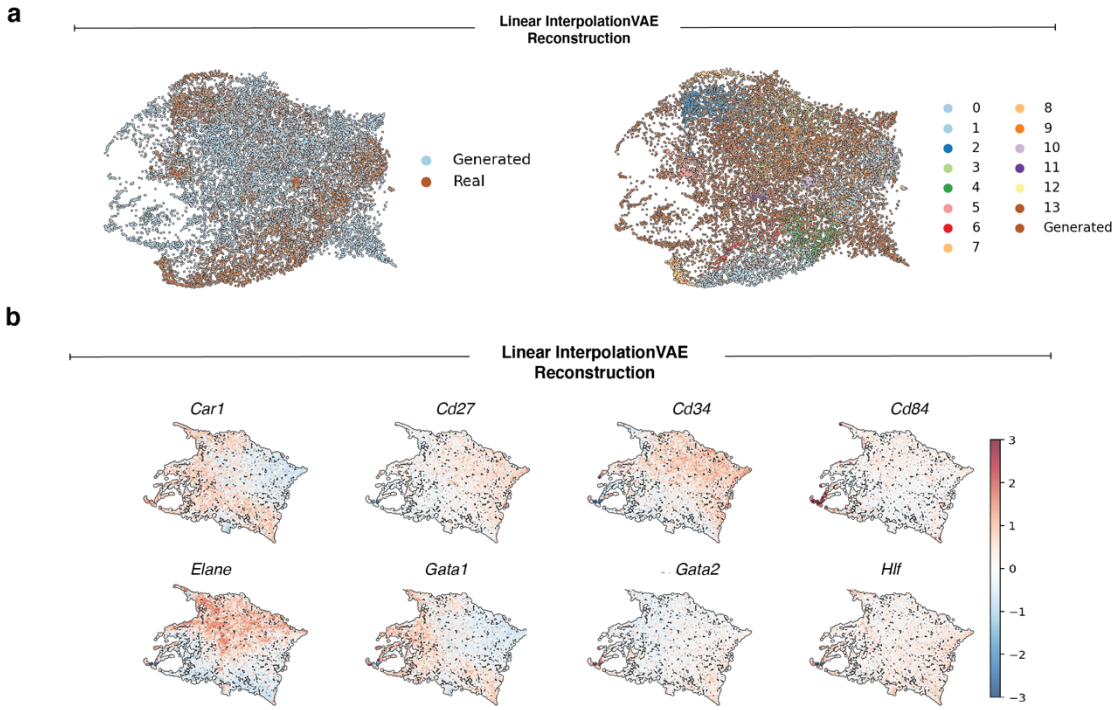
A. UMAP representation of the gastrulation mouse cell atlas. The cells are colored according to the cell type annotations provided by the authors.

B. UMAP representation of the gastrulation mouse cell atlas. The cells are colored according to the embryonic time points (E6.5 to E8.5).

C. UMAP representation of the gastrulation mouse cell atlas. The cells are colored according to Leiden clustering.

D. Density plot of UMAP representation showing the Cell-DRL agents' reconstructed sequence of cellular states from the primitive streak to the three embryonic germ layers: ectoderm, mesoderm, and endoderm. The co-embedding highlights the co-localization patterns with the whole embryogenesis mouse atlas.

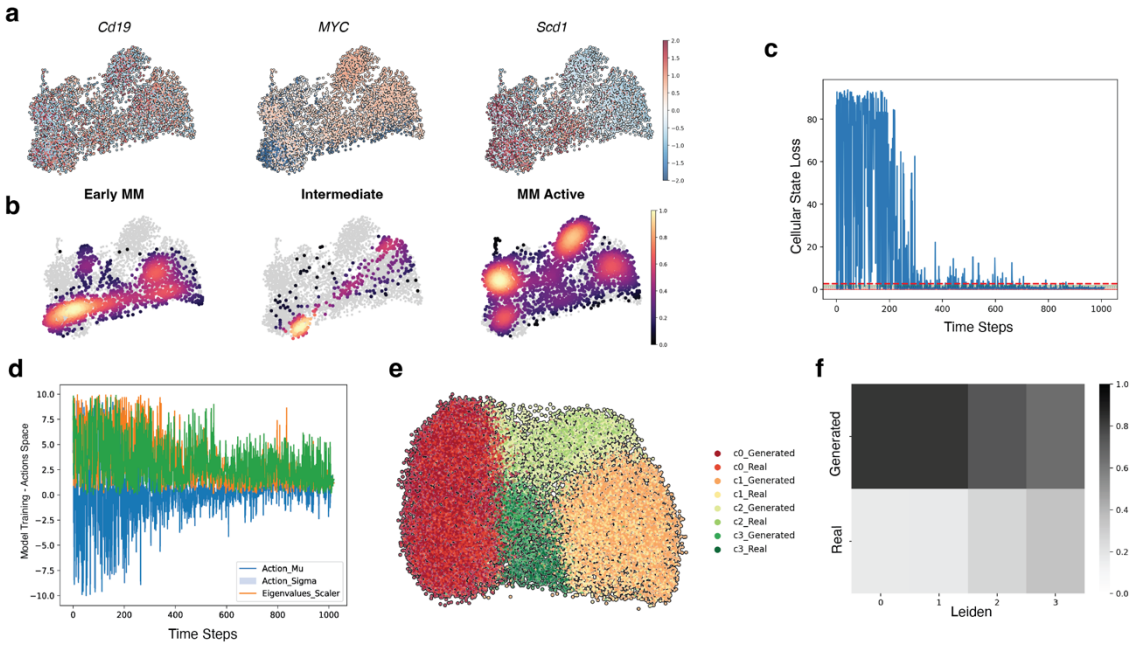
E. UMAP representing the real and generated intermediate cellular states. The cells are colored according to Leiden clustering.



Extended Data Fig. 8: Linear VAE interpolation for the hematopoietic cellular states.

A. The UMAP shows a co-embedding of real data and data generated by linear interpolation in the latent space of a variational autoencoder (VAE) for the hematopoiesis data (with the anchoring clusters 1 and 2).

B. UMAP representation of key hematopoietic marker gene expression in the unseen cellular states generated by linear interpolation in the latent space of a VAE.



Extended Data Fig. 9: Reconstruction of murine multiple myeloma (MM) tumor evolution and patient-specific relapse state trajectory.

A. UMAP representation of *Cd19, MYC,* and *Scd1* gene expressions (normalized) in the B cell compartment of the MM mouse single-cell RNA-seq data.

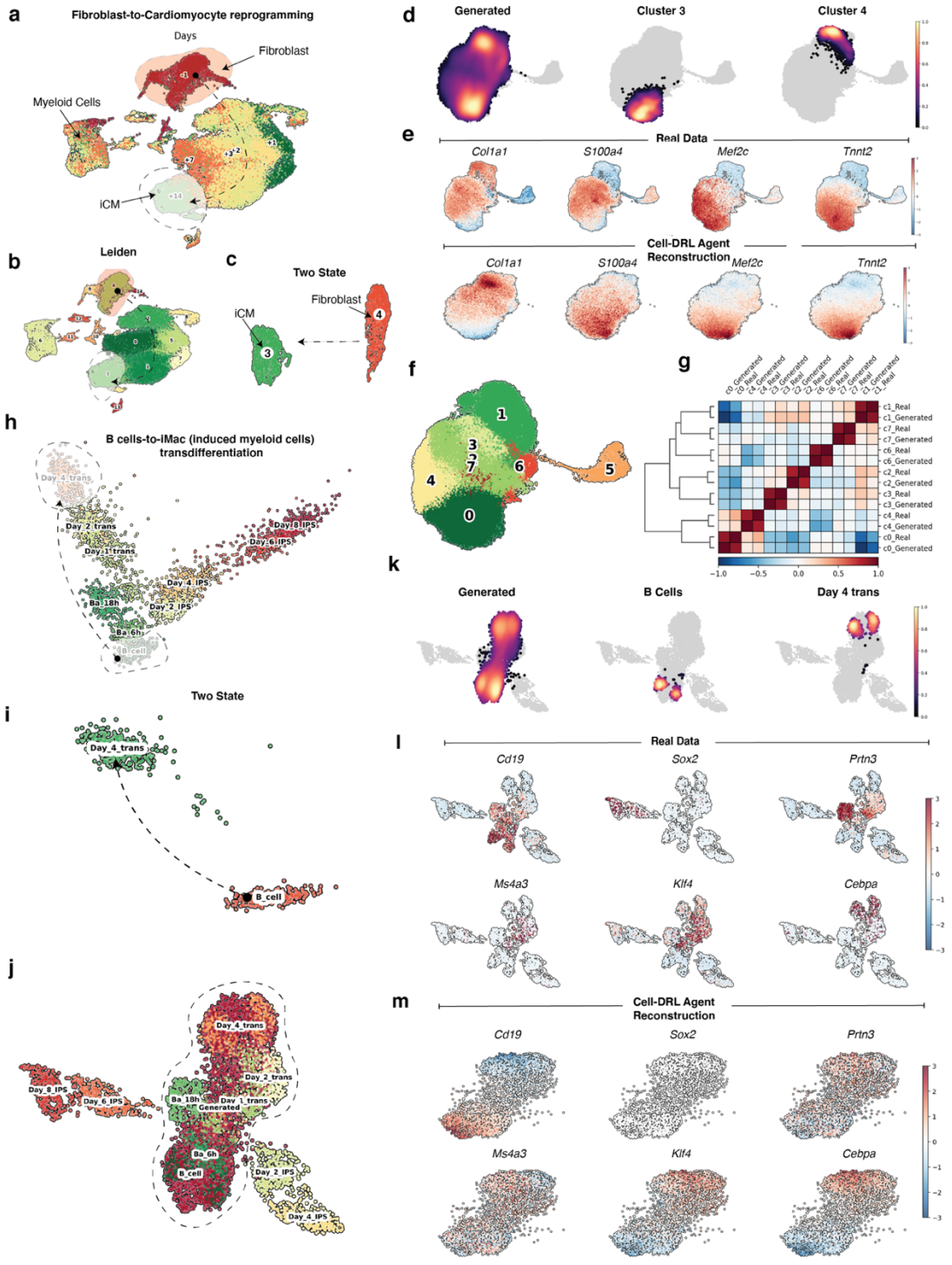
B. Density plot on UMAP representation highlighting localization patterns of the MM disease stages.

C. Diagram showing the agent's transition between exploration and exploitation phases during the learning process in the MM mouse model state scenario. The x-axis represents the time steps, and the y-axis represents the cellular state loss (CSL) values. The lines indicate the LBS, with the broken red lines representing the minimum and maximum trajectory bounds, and the broken black lines indicating the minimum and maximum reward bounds. The green-colored area represents the reconstruction field of interest (RF).

D. Diagram depicting the Do-Unit parameter space values during the Cell-DRL agent training process.

E. UMAP representation displaying the co-clustered and co-embedded real and generated MM mouse model data.

F. Heatmap displaying the proportions of generated and real data within each Leiden cluster.



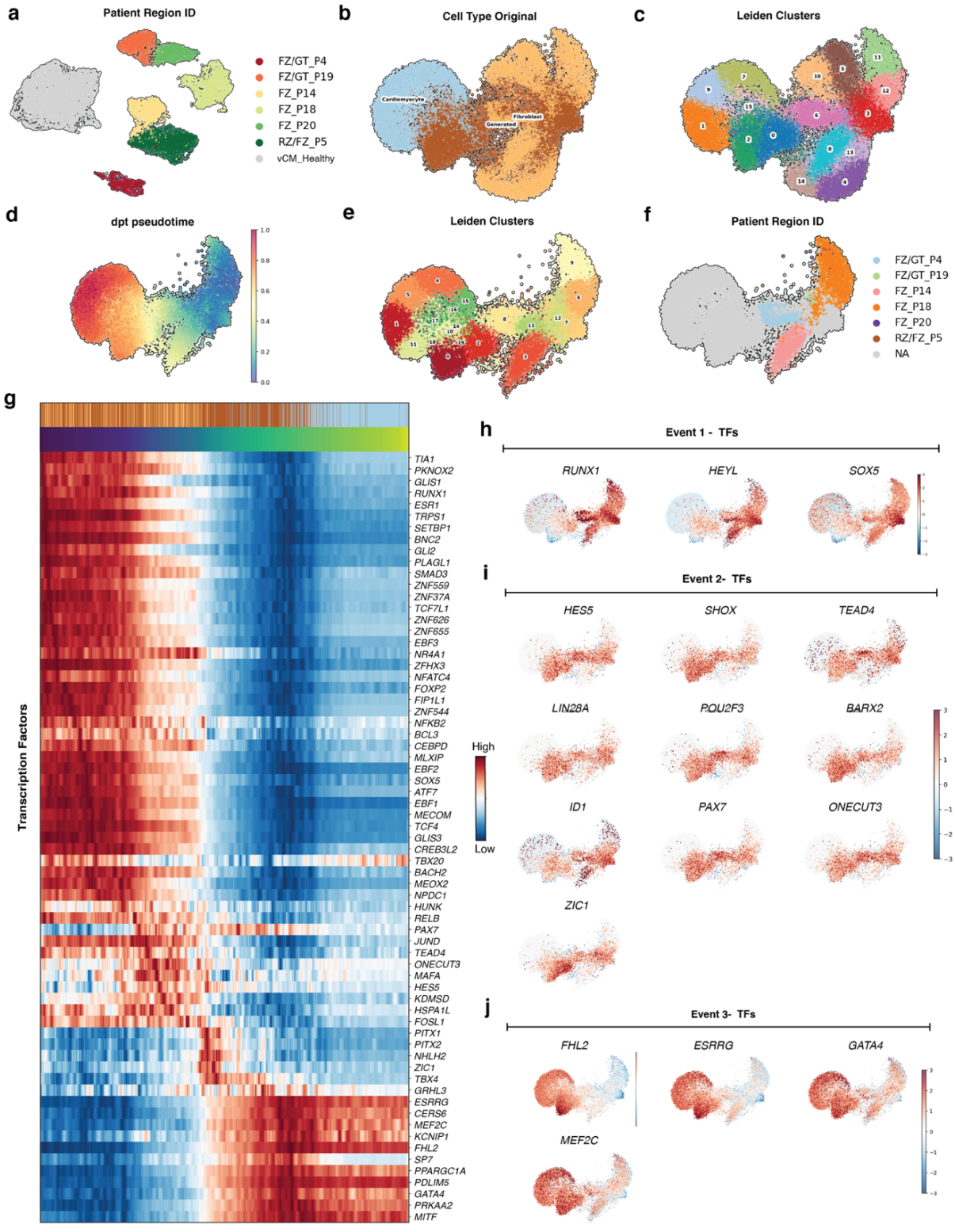
Extended Data Fig. 10: Reconstruction of murine fibroblasts-to-cardiomyocytes and B-cell-to-induced macrophages (iMac) trans-differentiation trajectories.

A. UMAP representation of murine fibroblasts-to-cardiomyocytes reprogramming single-cell RNA-seq data. Colors indicate the sampled cells over the reprogramming days. iCM: induced cardiomyocytes.

B. UMAP representation of murine fibroblasts-to-cardiomyocytes reprogramming single-cell RNA-seq data with cell type labels based on marker gene expression. Colors indicate Leiden clusters.

C. UMAP representation for the two-state challenge presenting murine fibroblast and cardiomyocyte clusters (3 and 4) to the Cell-DRL agent.

D. Density plot on UMAP representation demonstrating co-localization patterns of real states (clusters 3 and 4) and generated states overlaying the cellular state manifold.
E. UMAP representation of key fibroblast and cardiomyocyte marker gene expression (normalized) of the real and generated gene expression data.
F. UMAP representation displaying clustering of co-embedded real and generated data, colored by Leiden cluster IDs.
G. Heatmap illustrating correlation patterns of gene expression between real and generated data within each Leiden cluster.
H. PCA representation of B cells-to-induced macrophages (iMac) trans-differentiation single-cell RNA-seq data. Colors indicate the sampled cells over the trans-differentiation time points.
I. UMAP representation for the two-state challenge, presenting B cells and day 4 transdifferentiated cells (iMac) to the Cell-DRL agent.
J. UMAP representation displaying the co-clustered and co-embedded real and generated data for B cells-to-induced macrophage (iMac) trans-differentiation.
K. Density plot on UMAP representation demonstrating co-localization patterns of real states (B cells, and day 4 trans-differentiation data) and generated states overlaying the cellular state manifold.
L. UMAP representation of key B cell and trans-differentiation marker gene expression (normalized) of the real gene expression data.
M. UMAP representation of key B cell and trans-differentiation marker gene expression (normalized) of the generated gene expression data.

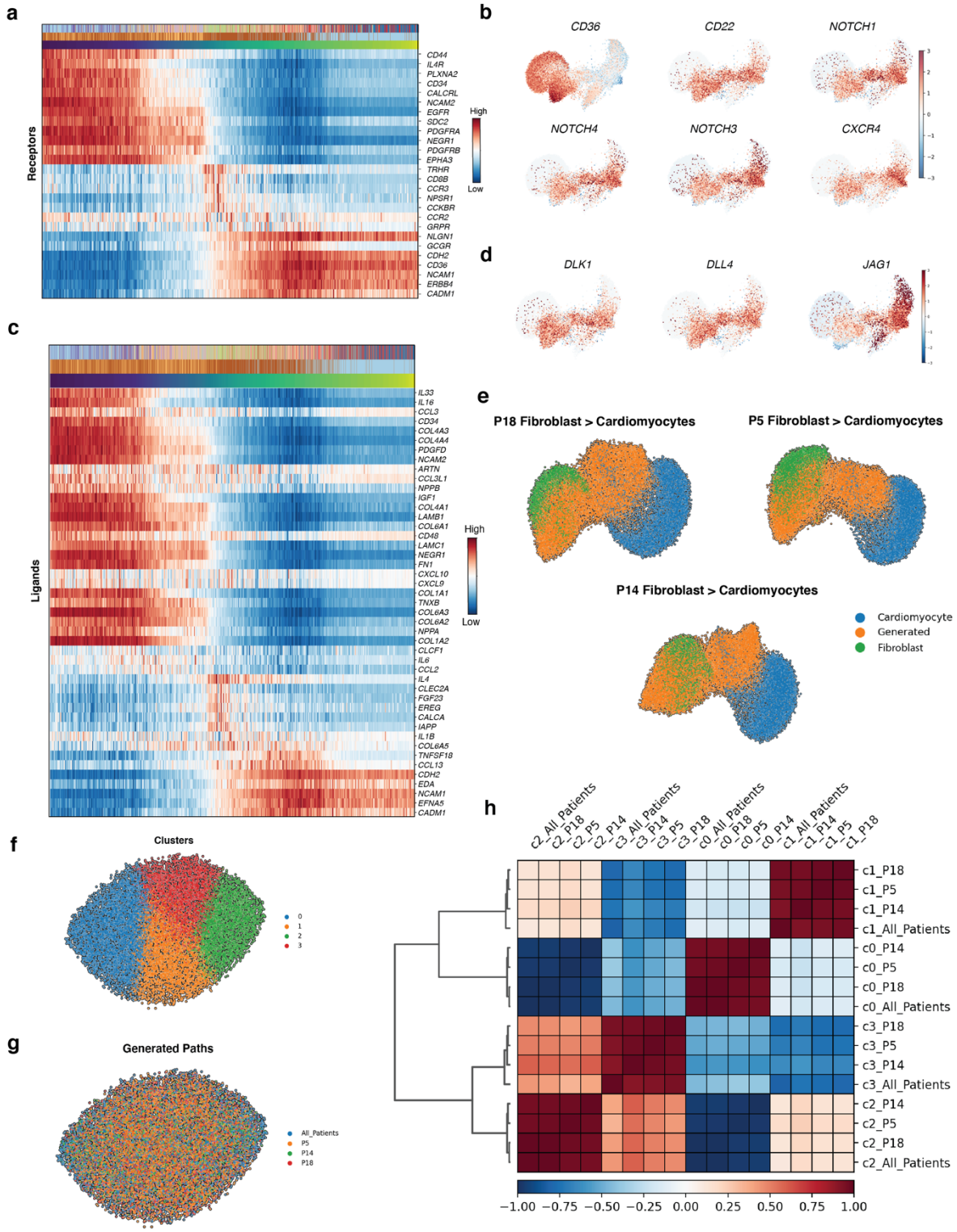


Extended Data Fig. 11: Predicting a novel trans-differentiation path (direct reprogramming) from fibroblasts to cardiomyocytes.

A. UMAP representation of fibroblast and cardiomyocyte populations extracted from the Fibrotic Zone (FZ) of Infarcted Heart patients. Colors indicate patients and anatomical region IDs.

B. UMAP representation illustrating the novel Cell-DRL reconstructed trans-differentiation Path between fibroblasts and cardiomyocytes.

C. UMAP representation highlighting the Leiden clusters of the reconstructed trans-differentiation Path merged with real fibroblast and cardiomyocyte data.
D. UMAP representation of the cellular states with a color code indicating estimated pseudotime values.
E. UMAP representation displaying the clustering of the co-embedded real and generated data. Colors indicate Leiden cluster IDs.
F. UMAP representation displaying the co-embedded real and generated data. Colors indicate patients and anatomical region IDs.
G. Heatmap displaying gene expression of the differentially expressed (DE) transcription factors over the reconstructed trans-differentiation Path sorted by the estimated pseudotime. The color bars at the top and bottom represent the cellular state groups, clusters, and the pseudotime scale, respectively. The red heat colors represent high gene expression, and the blue colors represent low gene expression.
H, I, and J. UMAP representations of key DE transcription factors (normalized) in Event 1, Event 2, and Event 3, respectively.



Extended Data Fig. 12: Robustness of the Cell-DRL-predicted reprogramming path from fibroblasts to cardiomyocytes.

A. Heatmap displaying gene expression of the DE receptors over the reconstructed trans-differentiation path sorted by the estimated pseudotime. The color bars at the top and bottom represent the cellular state groups, clusters, and the pseudotime scale, respectively. The red heat colors represent high gene expression, and the blue colors represent low gene expression.

B. UMAP representations of normalized gene expression for the predicted receptors.

C. Heatmap displaying gene expression of the DE ligands over the reconstructed trans-differentiation path sorted by the estimated pseudotime order. The color bars at the top and bottom represent the cellular state groups, clusters, and the pseudotime scale, respectively. The red heat colors represent high gene expression, and the blue colors represent low gene expression.

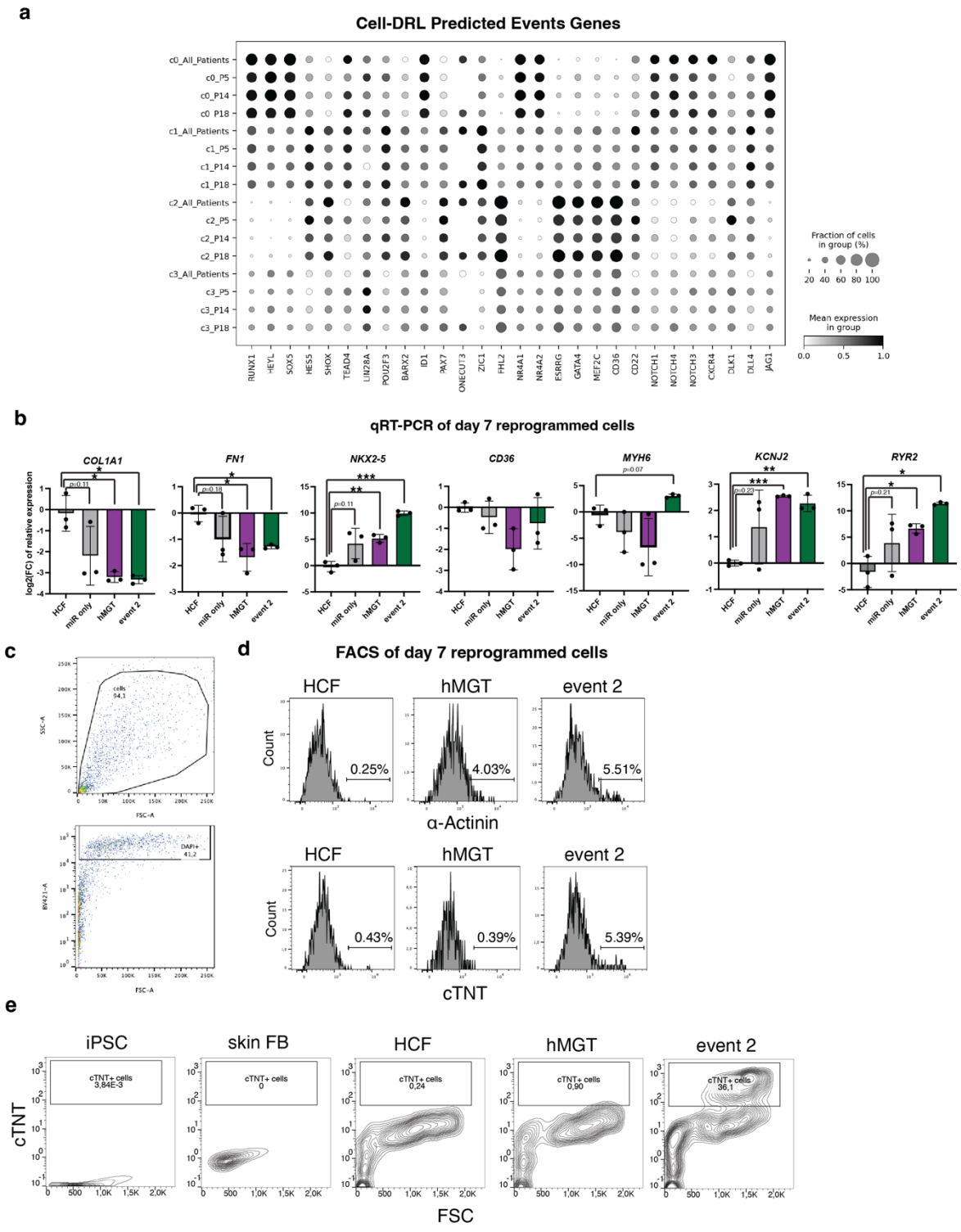
D. UMAP representations of normalized gene expression for the predicted ligands.

E. UMAP representation illustrating the novel Cell-DRL reconstructed trans-differentiation path between fibroblasts (starting with patient-specific populations from the myocardial infarction atlas – Kuppe et al dataset) and cardiomyocytes (from healthy adult heart atlas Litviňuková et al dataset).

F. UMAP representation illustrating the co-embedding of the generated paths from different patients with generated data of the initial states of all patients' fibroblast populations. Colors indicate Leiden clusters.

G. UMAP representation illustrating the co-embedding of the generated paths from different patients with generated data of the initial states of all patients' fibroblast populations. Colors indicate the patients' specific generated paths.

H. Heatmap illustrating correlation patterns of gene expression between real and generated data within each Leiden cluster.



Extended Data Fig. 13: Validation of the Cell-DRL-predicted novel reprogramming path from fibroblasts to cardiomyocytes.

A. Heatmap showing the gene expression patterns of the predicted Cell-DRL events across patient-specific generated paths in each Leiden cluster from the myocardial infarction atlas – *Kuppe et al* dataset.

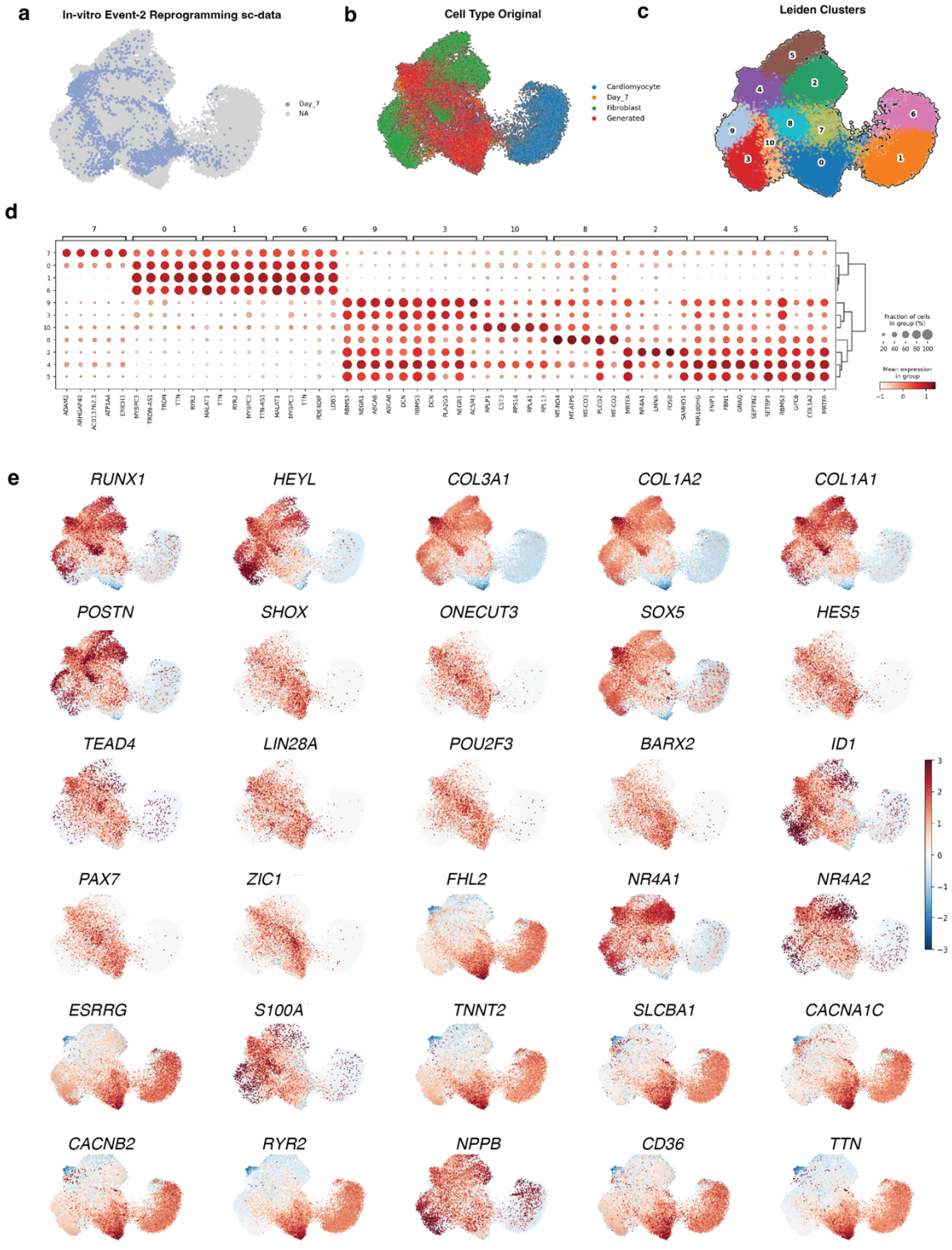
B. qRT-PCR quantification of fibroblast identity genes, CM progenitor/identity genes, and CM functional genes comparing across all reprogramming conditions, on day 7 of reprogramming. Gene expressions

are normalized relative to *GAPDH*. Fold changes of each sample against the averaged control values are plotted. Unpaired t-test was performed against control.

C. Gating strategy for the enrichment of nucleated cells (DAPI+, BV421 channel) isolated from the cell culture.

D. FACS analysis of day 7 reprogrammed cells with antibodies against α-Actinin and cTNT, respectively. Quantification of signals was determined by the log2 fold change (log2FC) of each measured sample against the averaged signal intensity of HCF condition. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

E. FACS analysis of day 7 reprogrammed cells with antibodies against cTNT and other negative controls



Extended Data Fig. 14: Co-mapping of *in vitro* Event 2 direct reprogramming data with the predicted Cell-DRL path.

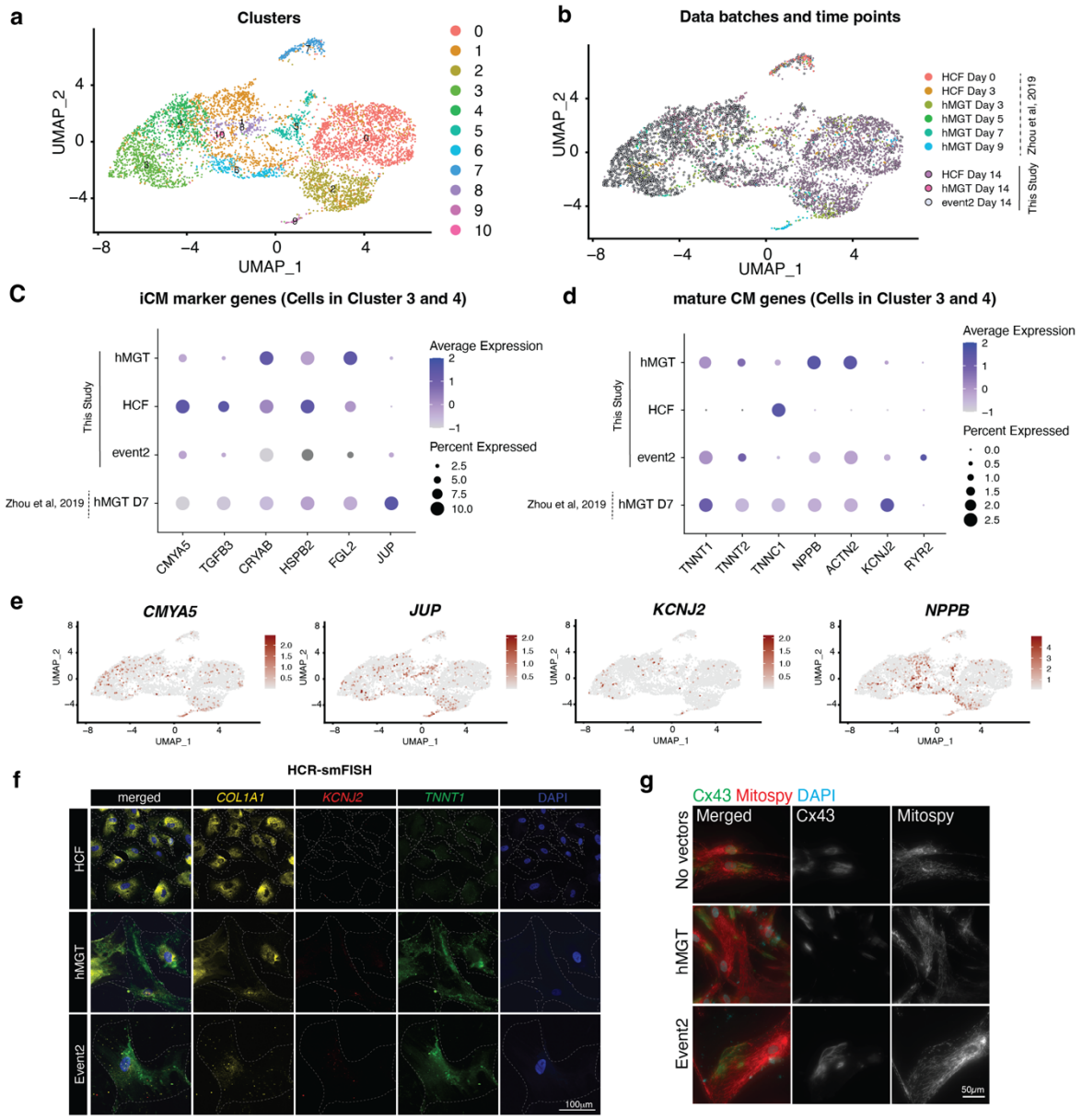
A. UMAP representation of *in vitro* Event 2 snRNA-seq data with the reconstructed trans-differentiation path, merged with real fibroblast and cardiomyocyte data. Colors indicate time point day 7 of *in vitro* reprogramming.

B. UMAP representation of in vitro Event 2 snRNA-seq data with the reconstructed trans-differentiation path, merged with real fibroblast and cardiomyocyte data. Colors indicate cell types and generated states together with Event 2 experimental data.

C. UMAP representation of in vitro Event 2 snRNA-seq data with the reconstructed trans-differentiation path, merged with real fibroblast and cardiomyocyte data. Colors indicate cell types and generated states together with Event 2 experimental data. Colors indicate Leiden clusters.

D. Dot plot showing differentially expressed gene patterns in each Leiden cluster, co-mapped with single-cell data from *in vitro* Event-2 data and the reconstructed trans-differentiation path, merged with real fibroblast and cardiomyocyte data.

E. UMAP representations of normalized gene expression levels for fibroblast and cardiac marker genes.



Extended Data Fig. 15: Single-cell data for the Event-2 and hMGT reprogramming experiments.

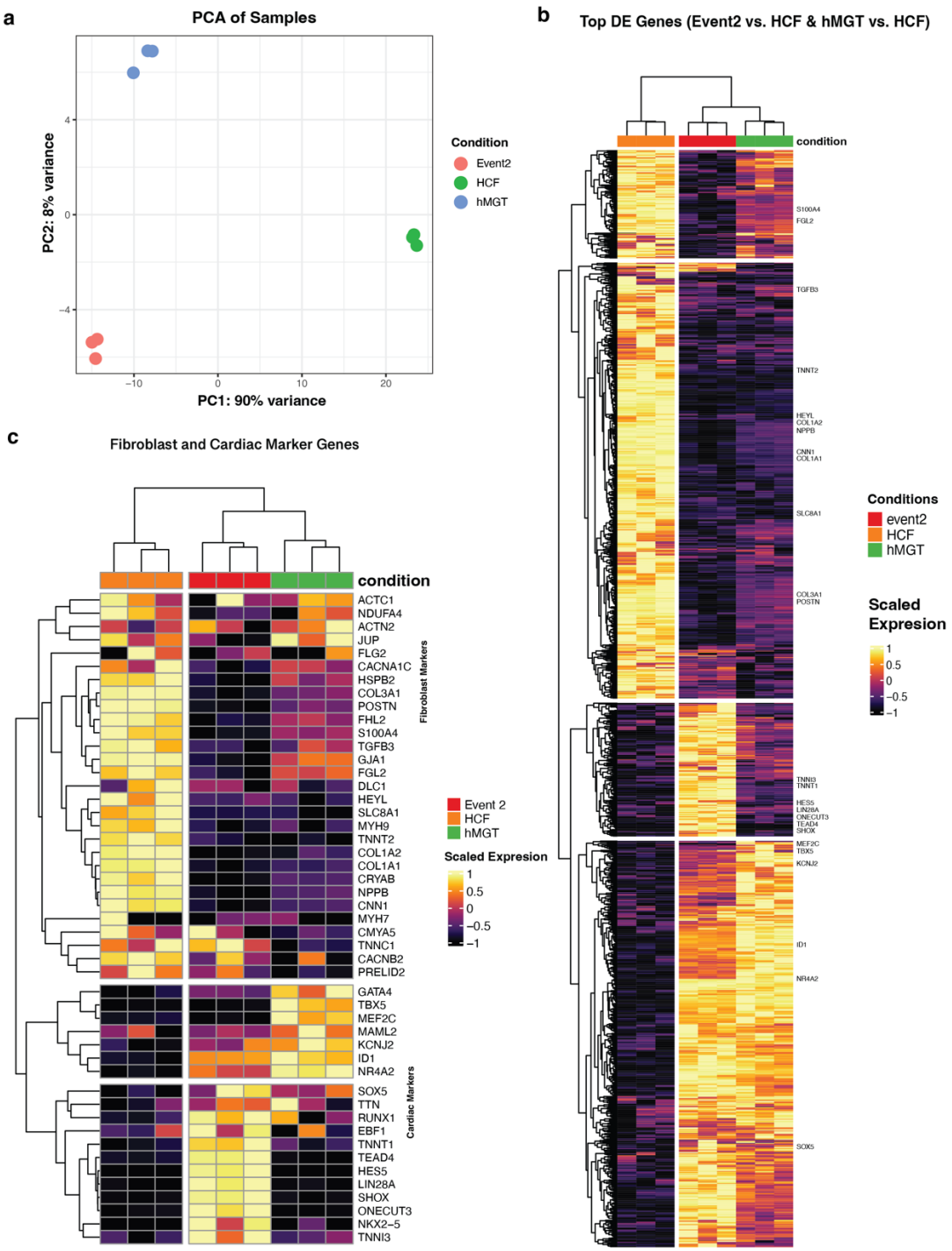
A, B. UMAP co-mapping of reprogrammed cells from Zhou *et al.,* 2019, and this study highlighting Louvain cluster (A) and annotated reprogramming conditions (B).

C, D. Dotplots of normalized gene expression in reprogrammed cells of different conditions in cluster 3 and 4, highlighting expression of induced CM (iCM) marker genes (C) and mature CM genes (D).

E. UMAP representation displaying the gene expression of iCM (*CMYA5 and JUP*) and mature CM marker genes (*KCNJ2* and *NPPB*).

F. Representative HCR-smFISH of day 14 reprogrammed cells, detecting mRNA transcripts of COL1A1 (fibroblasts marker gene), KCNJ2, and TNNT1 (muscle markers genes).

G. Representative immunofluorescence staining of reprogrammed cells at day 15, showing Cx43 (green) and Mitospy (red). Right panels, dashed white line shows the border between the cells. White arrow heads, Cx43 foci.

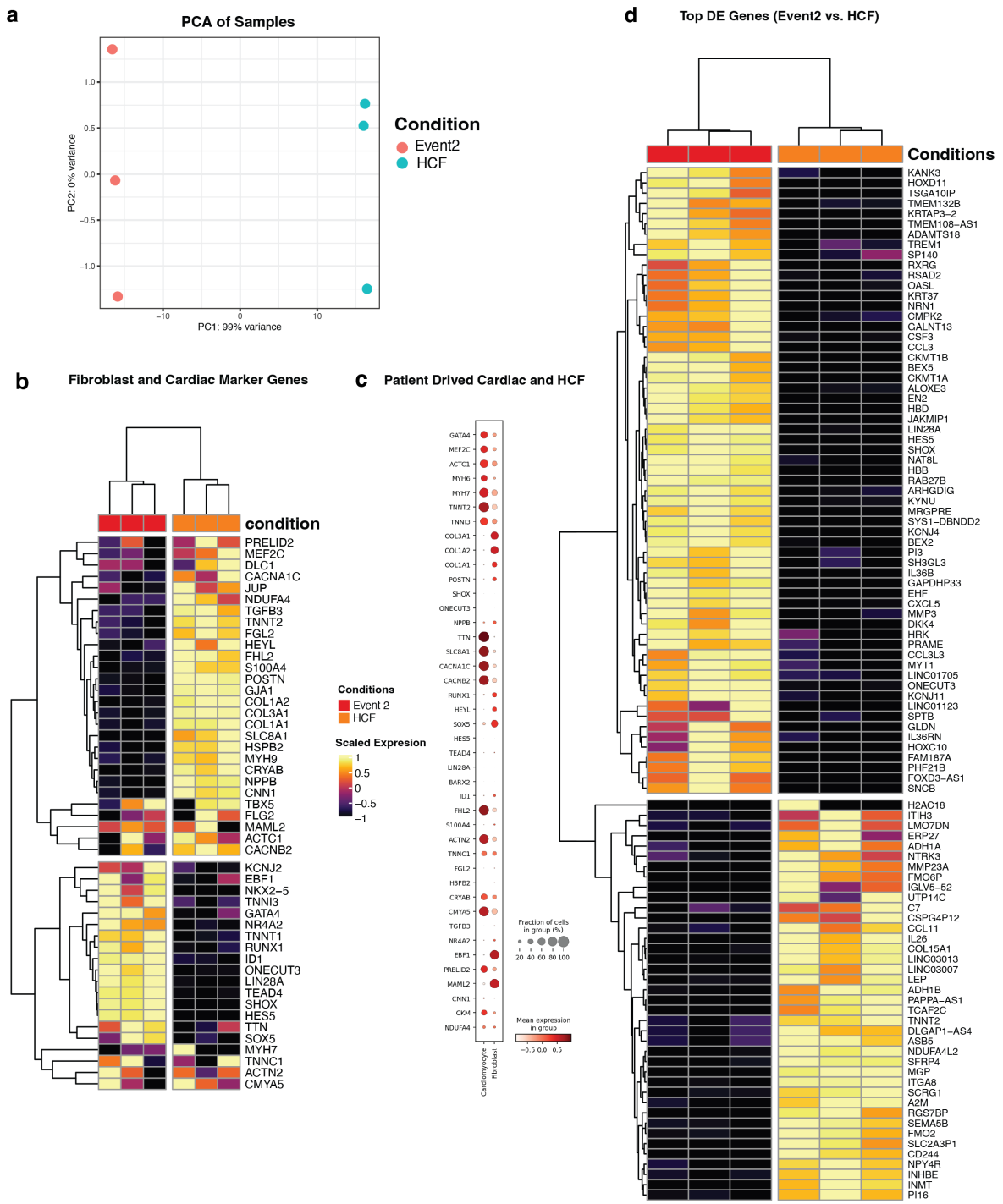


Extended Data Fig. 16: Bulk RNA-seq data for the Event 2 and hMGT reprogramming experiments at Day 7.

A. Principal Component Analysis (PCA) plot displaying the clustering of samples based on gene expression data.

B. Hierarchical clustering heatmap of the top differentially expressed genes between Event2, hMGT, and HCF conditions.

C. Heatmap showing the expression of selected fibroblast and cardiac marker genes across different conditions and time points.



Extended Data Fig. 17: Bulk RNA-seq data for the Event 2 reprogramming experiments at Day 7.

A. Principal Component Analysis (PCA) plot displaying the clustering of samples based on gene expression data.

B. Heatmap showing the expression of selected fibroblast and cardiac marker genes across different conditions and time points

C. Dotplot shows the gene expression pattern of cardiac and fibroblast markers in the Human Cardiac Atlas (Litviňuková *et al,* 2020 Nature), highlighting the partial expression of TNNT2 in the fibroblast population.

D. Hierarchical clustering heatmap of the top differentially expressed genes between Event2, and HCF conditions.