Additional file 1

Supplementary text, tables and figures for "In silico perturbation as an adversarial attack: combinatorial search over cell states in single-cell foundation models"

Supplementary text

Implementation details and reproducibility

Two requirements of the overexpression operators are stated, because neither is evident from the reference implementations. First, overexpression in Geneformer must remove a gene from its existing position before reinserting it: concatenating the gene at the front without removal duplicates any gene already expressed in the cell, which occurred in 16.7% of insertions in our data. Second, when the sequence of a cell has reached the maximum length of the model, appending k genes places all of them in the same terminal slot, so that only the last survives and the search collapses to $k = 1$ without any error being raised; we therefore reserved k positions by truncating the base sequence to the maximum length minus k.

The realized mean costs of the attack were 1298 (Geneformer) and 1299 (scGPT) forward passes against the specified 1290 at $k = 2$, 2585 and 2589 against 2580 at $k = 4$, and 2568 and 2574 against 2580 at $k = 8$, because a step evaluates fewer than 128 candidates when duplicates are removed and the last step may overrun.

Searches were run with seed 42, with seeding per (cell, restart) and deterministic `cuBLAS` workspaces. The two acceptance rules (Results, components of the attack) give identical trajectories up to 512 forward passes. For the ranking evaluation, a condition was rerun under identical settings and the per-pathway AUCs of the ranking evaluation (Results) were compared: the largest discrepancy was 1.3×10^{-4}, more than two orders of magnitude below the AUC differences discussed there.

Order sensitivity and the size of the search space

To measure order sensitivity we evaluated the same set of k genes under many permutations and recorded the spread of the objective, expressed as a percentage of the improvement achieved by the search. For Geneformer the spread is 3.60%, 8.25% and 9.97% at $k = 2$, 4 and 8; for scGPT it is 0.00% at every k (between 1.5×10^{-8} and 9.6×10^{-8} in absolute terms, at the level of floating-point noise; Figure S1).

For Geneformer the spread grows with k and reaches a tenth of the total improvement, so order is a genuine dimension of the problem rather than a rounding effect. The search spaces therefore differ by more than an order of magnitude for the same k: at $k = 4$, $P(n, k) = 4.79 \times 10^{13}$ ordered tuples for Geneformer ($n = 2632$) against $C(n, k) = 2.07 \times 10^{12}$ unordered sets for scGPT ($n = 2657$). The attack searches either space at a budget of 2580 forward passes.

Protocol selection

The 160 protocol entries were compiled from the direct reprogramming literature and include haematopoietic, vascular, mesenchymal, epithelial and glial conversions. Collapsing entries that describe the same conversion, including those from differently named starting material, gives 61 distinct pathways, each carrying the union of the factor sets of its protocols.

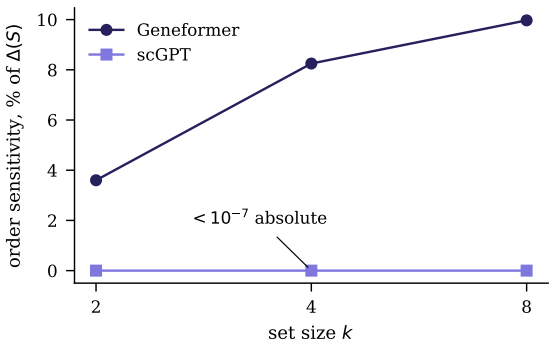


Figure S1: **Order sensitivity.** Change of the objective when the same k genes are permuted, as a percentage of the improvement achieved by search, for Geneformer and scGPT. For scGPT the change is below 10^{-7} in absolute terms.

Three criteria were applied in order: a pathway was retained only if (i) both its source and its target cell type appear in the Tabula Sapiens annotation, (ii) enough cells of each are available to define a source sample and a target centroid, and (iii) the reported factors are identifiable in the transcription factor vocabulary and traceable to a source publication. The first criterion removes most of the 61: conversions to neural targets (27 of the 61) do not survive, because Tabula Sapiens is built from living donors, who do not contribute brain tissue. Of the remaining 34, sixteen were excluded: three describe the same conversion at a finer granularity than the annotation supports (fibroblast to myoblast, hepatocyte to islet cells, pre-B cell to macrophage), two target states that an adult atlas does not contain (cardiac progenitor, induced pluripotent stem cell), and eleven have target cell types that are absent from the annotation or present only in an unrelated tissue (osteoblast, chondrocyte, oligodendrocyte precursor and megakaryocyte have no cells at all; the two adipocyte conversions fail because the 39 cells annotated as adipocytes come from lacrimal gland and ocular tissue). Eighteen pathways were retained (Table S1).

No fixed cell-count threshold was applied: six of the retained conversions have fewer than 400 target cells (the smallest has 144), and the number of reported factors ranges from 1 to 18. The constraint is the atlas rather than single-cell data in general; the same census contains 72488 adipocytes outside Tabula Sapiens.

Representativeness of the five pathways

The 2×2 decomposition and the anytime analysis used five of the 18 pathways (P01, P03, P07, P12 and P14). These span the range of the full set: ranked by the improvement that the attack achieves at $k = 2$, they occupy positions 1, 4, 10, 17 and 18, so that they include both the largest and the smallest, and their mean is within 3% of the mean over all 18. The matched-budget comparison agrees between the five and the full set: the attack-to-greedy ratio over the five pathways was 1.036, 1.180 and 1.361 at $k = 2$, 4 and 8 in Geneformer, against 1.043, 1.181 and 1.399 over all 18.

Distance to the optimum

At $k = 2$ the space is small enough to enumerate: 6924792 ordered pairs per cell (Figure S2). Measured against the true per-cell optimum, gradient-guided hill climbing attains 85.3% of the optimal improvement on average (median 93.2%), the full attack 84.3%, greedy 82.5%, uniform hill climbing 65.6% and the best of an equal number of random sets 66.8%. The searches operate at 1/5368 of the exhaustive cost.

The gap between the two gradient-guided methods and the two uniform ones, roughly twenty

Table S1: **The eighteen reprogramming pathways.**

	source	target	factors	cells	references
P01	fibroblast	regular atrial cardiac myocyte	12	2000	[1–12]
P02	fibroblast	hepatocyte	18	2000	[13–27]
P03	B cell	macrophage	2	2000	[28, 29]
P04	fibroblast	skeletal muscle satellite stem cell	6	2000	[30–32]
P05	T cell ($CD4^+$, $CD8^+$)	macrophage	2	2000	[33]†
P06	fibroblast	endothelial cell	1	2000	[34, 35]
P07	fibroblast	smooth muscle cell	1	2000	[36]
P08	fibroblast	myeloid dendritic cell	3	153	[37]
P09	fibroblast	erythroid progenitor cell	4	2000	[38]
P10	fibroblast	hematopoietic stem cell	5	2000	[39]
P11	fibroblast	enterocyte of epithelium proper of duodenum	4	2000	[40]
P12	fibroblast	type B pancreatic cell	5	306	[41]
P13	hepatocyte	type B pancreatic cell	7	306	[42–49]
P14	pancreatic acinar cell	type B pancreatic cell	3	306	[50]
P15	fibroblast	melanocyte	3	306	[51]
P16	fibroblast	Schwann cell	2	144	[52]
P17	fibroblast	kidney epithelial cell	4	2000	[53]
P18	fibroblast	pulmonary alveolar type 2 cell	4	2000	[54]

Cell type names are the Tabula Sapiens annotations. "factors" is the size of the union of the factor sets reported for that pathway across its protocols, 86 in total across the eighteen and 62 distinct genes, so some factors are curated for more than one pathway. † the source cells in this publication are committed pre-T progenitors rather than the mature T cells of the atlas annotation. "references" lists the publications of the protocols merged for the pathway. "cells" is the number available to form the target centroid; where it exceeds 200 a random sample of 200 is used.

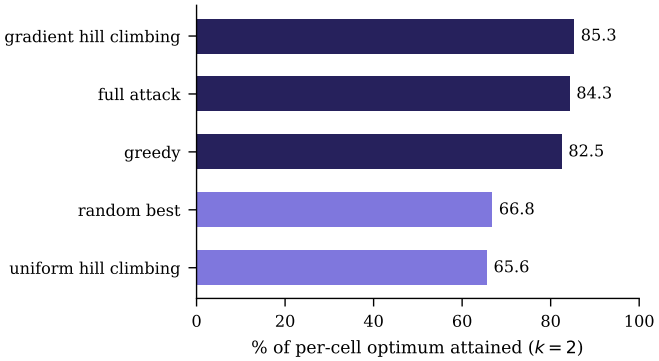


Figure S2: **Distance to the optimum.** At $k = 2$, where the space can be enumerated exhaustively, the fraction of the per-cell optimum attained by each method at 1/5368 of the exhaustive cost.

Table S2: **Paired t-tests and Wilcoxon signed-rank tests for every contrast.**

	Geneformer			scGPT		
contrast	effect	p_t	p_W	effect	p_t	p_W
restart mean						
gradient, under improve-only	-0.016	0.846	0.88	-0.031	0.0223	0.731
gradient, under unconditional	$+0.572$	1.76×10^{-9}	3.00×10^{-10}	$+0.033$	4.69×10^{-6}	2.88×10^{-12}
unconditional, under uniform	-0.108	0.0907	0.00579	-0.025	2.94×10^{-7}	1.17×10^{-11}
unconditional, under gradient	$+0.479$	6.36×10^{-12}	2.35×10^{-14}	$+0.039$	3.57×10^{-4}	3.73×10^{-9}
main effect of candidate source	$+0.278$	1.22×10^{-4}	5.27×10^{-5}	$+0.001$	0.884	1.75×10^{-5}
main effect of acceptance rule	$+0.185$	2.19×10^{-5}	6.07×10^{-7}	$+0.007$	0.201	0.048
interaction	$+0.588$	7.91×10^{-9}	1.23×10^{-11}	$+0.064$	7.42×10^{-7}	1.71×10^{-14}
best of three restarts						
gradient, under improve-only	$+0.045$	0.624	0.575	-0.018	0.0794	0.624
gradient, under unconditional	$+0.472$	5.75×10^{-4}	6.45×10^{-11}	$+0.019$	4.08×10^{-7}	1.05×10^{-12}
unconditional, under uniform	-0.105	0.401	0.00437	-0.018	4.48×10^{-7}	3.83×10^{-9}
unconditional, under gradient	$+0.322$	2.08×10^{-5}	8.10×10^{-10}	$+0.019$	0.0257	7.70×10^{-7}
main effect of candidate source	$+0.259$	0.00628	1.55×10^{-7}	$+0.000$	0.965	1.33×10^{-6}
main effect of acceptance rule	$+0.108$	0.162	0.0065	$+0.001$	0.892	5.73×10^{-4}
interaction	$+0.427$	0.0019	5.24×10^{-8}	$+0.037$	2.02×10^{-4}	1.66×10^{-11}

$n = 100$ source cells per model. Effects are in units of 10^{-3}. The interaction is significant under both tests, in both models, and under both aggregations. The two tests disagree in eight of the 28 cells, none of them the interaction: the effect of unconditional acceptance under uniform candidates in Geneformer under both aggregations, the effect of gradient candidates under improve-only acceptance in scGPT at the restart mean, both main effects in scGPT under both aggregations, and the main effect of the acceptance rule in Geneformer under best-of-restarts.

percentage points, is much larger here than at $k = 8$, consistent with the transient gradient advantage described in the Results (components of the attack): at this budget the search is still in the regime where proposal quality governs ascent rate.

Simple effects and main effects of the 2×2

Introducing gradient candidates while retaining improve-only acceptance changes nothing in Geneformer ($p = 0.85$) and does not help in scGPT, where the mean is negative and significant by a t-test ($p = 0.022$) but the signed-rank test is far from significance ($p = 0.73$); we therefore claim only the absence of benefit. Adopting unconditional acceptance while retaining uniform candidates is harmful in both models. In scGPT it degrades the objective in 88 of 100 cells ($p = 2.9\times10^{-7}$). In Geneformer the mean is diluted by a few cells with large positive differences, so the t-test reaches only $p = 0.091$ while the signed-rank test, which reflects the 65 of 100 cells made worse, gives $p = 0.0058$.

The main effects are not informative as summaries, because each averages two simple effects of opposite sign. The main effect of candidate source is $+0.278$ in Geneformer, the mean of -0.016 under improve-only acceptance and $+0.572$ under unconditional acceptance, and $+0.001$ in scGPT, the mean of -0.031 and $+0.033$. In scGPT the two halves almost cancel, and the resulting figure describes neither of them. The acceptance rule behaves in the same way: $+0.185$ in Geneformer, the mean of -0.108 and $+0.479$. The two tests disagree on both main effects in scGPT under either aggregation, and in Geneformer they agree at the restart mean but disagree on the acceptance rule under best-of-restarts (Table S2). The interaction, by contrast, is significant under both tests, in both models and under both aggregations, with signed-rank p values smaller than the corresponding t-test values throughout.

Spread and sign structure of the single-gene ranking evaluation

Differential expression had the higher pathway-specific AUC in 14 of the 18 pathways and the two tied in one, but the per-pathway difference between Geneformer and differential expression ranged from -0.194 to $+0.705$, and the pathway-specific AUC of Geneformer was bimodal, at or below 0.60 in eight pathways and at or above 0.98 in three.

With a single positive the pathway-specific AUC reduces to the rank percentile of that one gene among the 61 curated factors of the other pathways: it takes 62 distinct values and moves in steps of 0.016. Two of the eighteen pathways have one curated factor each, and both sit at an extreme of the distribution, as is expected of a measure with this granularity. In the fibroblast to endothelial pathway the factor is ETV2, which Geneformer places below 54 of the 61 others, giving 0.115. In the fibroblast to smooth muscle pathway it is MYOCD, which Geneformer places above all 61, giving 1.000. Dropping the fibroblast to endothelial pathway moves the means over the remaining 17 pathways to 0.904 and 0.735 for Geneformer against 0.887 and 0.904 for differential expression, which is 0.017 apart conventionally and 0.1687 apart pathway-specifically. Weighting pathways by their number of positives gives a gap of 0.1694 rather than 0.198.

The direction of the two measures also differs by method. For Geneformer the conventional AUC exceeds the pathway-specific one in 17 of 18 pathways, a mean gap of $+0.175$; the ordering is reversed for the target-expression baseline in all 18 pathways (mean -0.105) and for differential expression in 12 of 18 (mean -0.024). Geneformer therefore ranks curated factors above the rest of the vocabulary but cannot separate a pathway's own factors from those of another pathway, which is what a general master-regulator signal would produce. The expression baselines can.

Complementarity of the search with the expression baselines

The searched objective is complementary to the expression baselines rather than superior to them. A factor is counted as lifted when the baseline places it below the midpoint of the vocabulary and the method places it in the top 10%. Against the target-expression baseline the two rankings of curated factors correlate only weakly (Spearman 0.249). That baseline places 13 of the 86 curated factor-pathway pairs below the midpoint of the vocabulary; this method ranks every one of them higher, moving the median rank from 1969 to 231, and lifts seven into the top 10% of the vocabulary: NEUROG3 in three separate pathways and PAX4 in two, both transiently expressed endocrine lineage-specifying factors, and SOX2 and GATA3 in the fibroblast to hepatocyte pathway. Most of this recovery is not an artefact of embedding norm: dividing every score by the norm of its gene embedding leaves six of the seven. Two pairs move the other way, SRF and ZFPM2, both in the fibroblast to cardiomyocyte pathway. Against the differential-expression baseline the two rankings correlate more strongly (Spearman 0.406) and the complementarity is weaker: that baseline places nine pairs below the midpoint, this method ranks all nine higher and lifts five into the top 10% (four after the norm correction), and three pairs move the other way (SRF, ZFPM2 and ETV2), though that baseline is the stronger of the two on its own. A rank average of this method and that baseline reaches a conventional AUC of 0.922, higher than either alone, though its pathway-specific AUC of 0.880 does not reach the 0.899 of the baseline by itself (0.701 for this method alone).

References

[1] Ieda M, Fu JD, Delgado-Olguin P, Vedantham V, Hayashi Y, Bruneau BG, et al. Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. Cell. 2010;142(3):375–386.

[2] Song K, Nam YJ, Luo X, Qi X, Tan W, Huang GN, et al. Heart repair by reprogramming non-myocytes with cardiac transcription factors. Nature. 2012;485(7400):599–604.

[3] Protze S, Khattak S, Poulet C, Lindemann D, Tanaka EM, Ravens U. A new approach to transcription factor screening for reprogramming of fibroblasts to cardiomyocyte-like cells. J Mol Cell Cardiol. 2012;53:323–332.

[4] Wada R, Muraoka N, Inagawa K, Yamakawa H, Miyamoto K, Sadahiro T, et al. Induction of human cardiomyocyte-like cells from fibroblasts by defined factors. Proc Natl Acad Sci USA. 2013;110:12667–12672.

[5] Addis RC, Ifkovits JL, Pinto F, et al. Optimization of direct fibroblast reprogramming to cardiomyocytes using calcium activity as a functional measure of success. J Mol Cell Cardiol. 2013;60:97–106. https://doi.org/10.1016/j.yjmcc.2013.04.004.

[6] Fu JD, Stone NR, Liu L, Spencer CI, Qian L, Hayashi Y, et al. Direct reprogramming of human fibroblasts toward a cardiomyocyte-like state. Stem Cell Rep. 2013;1:235–247.

[7] Christoforou N, Chellappan M, Adler AF, et al. Transcription factors MYOCD, SRF, Mesp1 and SMARCD3 enhance the cardio-inducing effect of GATA4, TBX5, and MEF2C during direct cellular reprogramming. PLoS One. 2013;8(5):e63577. https://doi.org/10.1371/journal.pone.0063577.

[8] Nam YJ, Song K, Luo X, Daniel E, Lambeth K, West K, et al. Reprogramming of human fibroblasts toward a cardiac fate. Proc Natl Acad Sci USA. 2013;110(14):5588–5593.

[9] Muraoka N, Yamakawa H, Miyamoto K, Sadahiro T, Umei T, Isomi M, et al. MiR-133 promotes cardiac reprogramming by directly repressing Snai1 and silencing fibroblast signatures. EMBO J. 2014;33(14):1565–1581. https://doi.org/10.15252/embj.201387605.

[10] Zhou Y, Wang L, Vaseghi HR, Liu Z, Lu R, Alimohamadi S, et al. Bmi1 is a key epigenetic barrier to direct cardiac reprogramming. Cell Stem Cell. 2016;18:382–395.

[11] Singh VP, Mathison M, Patel V, Sanagasetti D, Gibson BW, Yang J, et al. MiR-590 promotes transdifferentiation of porcine and human fibroblasts toward a cardiomyocyte-like fate by directly repressing specificity protein 1. J Am Heart Assoc. 2016;5(11):e003922. https://doi.org/10.1161/JAHA.116.003922.

[12] Zhou H, Morales MG, Hashimoto H, Dickson ME, Song K, Ye W, et al. ZNF281 enhances cardiac reprogramming by modulating cardiac and inflammatory gene expression. Genes Dev. 2017;31:1770–1783.

[13] Sekiya S, Suzuki A. Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. Nature. 2011;475(7356):390–393.

[14] Huang P, He Z, Ji S, Sun H, Xiang D, Liu C, et al. Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. Nature. 2011;475(7356):386–389.

[15] Yu B, He ZY, You P, Han QW, Xiang D, Chen F, et al. Reprogramming fibroblasts into bipotential hepatic stem cells by defined factors. Cell Stem Cell. 2013;13(3):328–340.

[16] Du Y, Wang J, Jia J, Song N, Xiang C, Xu J, et al. Human hepatocytes with drug metabolic function induced from fibroblasts by lineage reprogramming. Cell Stem Cell. 2014;14:394–403.

[17] Huang P, Zhang L, Gao Y, He Z, Yao D, Wu Z, et al. Direct reprogramming of human fibroblasts to functional and expandable hepatocytes. Cell Stem Cell. 2014;14:370–384.

[18] Simeonov KP, Uppal H. Direct reprogramming of human fibroblasts to hepatocyte-like cells by synthetic modified mRNAs. PLoS One. 2014;9(6):e100134.

[19] Pournasr B, Asghari-Vostikolaee MH, Baharvand H. Transcription factor-mediated reprograming of fibroblasts to hepatocyte-like cells. Eur J Cell Biol. 2015;94(12):603–610.

[20] Song G, Pacher M, Balakrishnan A, Yuan Q, Tsay HC, Yang D, et al. Direct reprogramming of hepatic myofibroblasts into hepatocytes in vivo attenuates liver fibrosis. Cell Stem Cell. 2016;18:797–808.

[21] Lim KT, Lee SC, Gao Y, Kim KP, Song G, An SY, et al. Small molecules facilitate single factor-mediated hepatic reprogramming. Cell Rep. 2016;15(4):814–829.

[22] Zakikhan K, Pournasr B, Nassiri-Asl M, Baharvand H. Enhanced direct conversion of fibroblasts into hepatocyte-like cells by Kdm2b. Biochem Biophys Res Commun. 2016;474(1):97–103.

[23] Rezvani M, Español-Suñer R, Malato Y, Dumont L, Grimm AA, Kienle E, et al. In vivo hepatic reprogramming of myofibroblasts with AAV vectors as a therapeutic strategy for liver fibrosis. Cell Stem Cell. 2016;18(6):809–816.

[24] Nakamori D, Akamine H, Takayama K, Sakurai F, Mizuguchi H. Direct conversion of human fibroblasts into hepatocyte-like cells by ATF5, PROX1, FOXA2, FOXA3, and HNF4A transduction. Sci Rep. 2017;7(1):16675.

[25] Wu W, Jin YQ, Gao Z. Directly reprogramming fibroblasts into adipogenic, neurogenic and hepatogenic differentiation lineages by defined factors. Exp Ther Med. 2017;13:2685–2690.

[26] Chen C, Pla-Palacín I, Baptista PM, Shang P, Oosterhoff LA, van Wolferen ME, et al. Hepatocyte-like cells generated by direct reprogramming from murine somatic cells can repopulate decellularized livers. Biotechnol Bioeng. 2018;115:2807–2816.

[27] Park S, Hwang SI, Kim J, Hwang S, Kang S, Yang S, et al. The therapeutic potential of induced hepatocyte-like cells generated by direct reprogramming on hepatic fibrosis. Stem Cell Res Ther. 2019;10(1):21.

[28] Xie H, Ye M, Feng R, Graf T. Stepwise reprogramming of B cells into macrophages. Cell. 2004;117(5):663–676.

[29] Bussmann LH, Schubert A, Vu Manh TP, De Andres L, Desbordes SC, Parra M, et al. A robust and highly efficient immune cell reprogramming system. Cell Stem Cell. 2009;5(5):554–566.

[30] Davis RL, Weintraub H, Lassar AB. Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell. 1987;51:987–1000.

[31] Ito N, Kii I, Shimizu N, Tanaka H, Takeda S. Direct reprogramming of fibroblasts into skeletal muscle progenitor cells by transcription factors enriched in undifferentiated subpopulation of satellite cells. Sci Rep. 2017;7(1):8097.

[32] Wakao J, Kishida T, Fumino S, Kimura K, Yamamoto K, Kotani S, et al. Efficient direct conversion of human fibroblasts into myogenic lineage induced by co-transduction with MYCL and MYOD1. Biochem Biophys Res Commun. 2017;488:368–373.

[33] Laiosa CV, Stadtfeld M, Xie H, de Andres-Aguayo L, Graf T. Reprogramming of committed T cell progenitors to macrophages and dendritic cells by C/EBPα and PU.1 transcription factors. Immunity. 2006;25(5):731–744.

[34] Morita R, Suzuki M, Kasahara H, Shimizu N, Shichita T, Sekiya T, et al. ETS transcription factor ETV2 directly converts human fibroblasts into functional endothelial cells. Proc Natl Acad Sci USA. 2015;112(1):160–165.

[35] Lee S, Park C, Han JW, Kim JY, Cho K, Kim EJ, et al. Direct reprogramming of human dermal fibroblasts into endothelial cells using ER71/ETV2. Circ Res. 2017;120(5):848–861.

[36] Wang Z, Wang DZ, Pipes GT, Olson EN. Myocardin is a master regulator of smooth muscle gene expression. Proc Natl Acad Sci USA. 2003;100(12):7129–7134.

[37] Rosa FF, Pires CF, Kurochkin I, Ferreira AG, Gomes AM, Palma LG, et al. Direct reprogramming of fibroblasts into antigen-presenting dendritic cells. Sci Immunol. 2018;3(30):eaau4292.

[38] Capellera-Garcia S, Pulecio J, Dhulipala K, Siva K, Rayon-Estrada V, Singbrant S, et al. Defining the minimal factors required for erythropoiesis through direct lineage conversion. Cell Rep. 2016;15(11):2550–2562.

[39] Batta K, Florkowska M, Kouskoff V, Lacaud G. Direct reprogramming of murine fibroblasts to hematopoietic progenitor cells. Cell Rep. 2014;9(5):1871–1884.

[40] Miura S, Suzuki A. Generation of mouse and human organoid-forming intestinal progenitor cells by direct lineage reprogramming. Cell Stem Cell. 2017;21(4):456–471.

[41] Fontcuberta-PiSunyer M, García-Alamán A, Prades È, Téllez N, Alves-Figueiredo H, Ramos-Rodríguez M, et al. Direct reprogramming of human fibroblasts into insulin-producing cells using transcription factors. Commun Biol. 2023;6(1):256.

[42] Ferber S, Halkin A, Cohen H, Ber I, Einav Y, Goldberg I, et al. Pancreatic and duodenal homeobox gene 1 induces expression of insulin genes in liver and ameliorates streptozotocin-induced hyperglycemia. Nat Med. 2000;6:568–572.

[43] Ber I, Shternhall K, Perl S, Ohanuna Z, Goldberg I, Barshack I, et al. Functional, persistent, and extended liver to pancreas transdifferentiation. J Biol Chem. 2003;278:31950–31957.

[44] Wang AY, Ehrhardt A, Xu H, Kay MA. Adenovirus transduction is required for the correction of diabetes using Pdx-1 or Neurogenin-3 in the liver. Mol Ther. 2007;15(2):255–263.

[45] Meivar-Levy I, Sapir T, Berneman D, Weissbach T, Polak-Charcon S, Ravassard P, et al. Human liver cells expressing albumin and mesenchymal characteristics give rise to insulin-producing cells. J Transplant. 2011;2011(1):252387.

[46] Nagasaki H, Katsumata T, Oishi H, Tai PH, Sekiguchi Y, Koshida R, et al. Generation of insulin-producing cells from the mouse liver using β cell-related gene transfer including Mafa and Mafb. PLoS One. 2014;9:e113022. https://doi.org/10.1371/journal.pone.0113022.

[47] Berneman-Zeitouni D, Molakandov K, Elgart M, Mor E, Fornoni A, Domínguez MR, et al. The temporal and hierarchical control of transcription factors-induced liver to pancreas transdifferentiation. PLoS One. 2014;9(2):e87812.

[48] Lu J, Dong H, Lin L, Wang Q, Huang L, Tan J. miRNA-302 facilitates reprogramming of human adult hepatocytes into pancreatic islets-like cells in combination with a chemical defined media. Biochem Biophys Res Commun. 2014;453(3):405–410.

[49] Cerdá-Esteban N, Naumann H, Ruzittu S, Mah N, Pongrac IM, Cozzitorto C, et al. Stepwise reprogramming of liver cells to a pancreas progenitor state by the transcriptional regulator Tgif2. Nat Commun. 2017;8(1):14127.

[50] Zhou Q, Brown J, Kanarek A, Rajagopal J, Melton DA. In vivo reprogramming of adult pancreatic exocrine cells to β-cells. Nature. 2008;455:627–632.

[51] Yang R, Zheng Y, Li L, Liu S, Burrows M, Wei Z, et al. Direct conversion of mouse and human fibroblasts to functional melanocytes by defined factors. Nat Commun. 2014;5(1):5807.

[52] Mazzara PG, Massimino L, Pellegatta M, Ronchi G, Ricca A, Iannielli A, et al. Two factor-based reprogramming of rodent and human fibroblasts into Schwann cells. Nat Commun. 2017;8(1):14088.

[53] Kaminski MM, Tosic J, Kresbach C, Engel H, Klockenbusch J, Müller AL, et al. Direct reprogramming of fibroblasts into renal tubular epithelial cells by defined transcription factors. Nat Cell Biol. 2016;18(12):1269–1280.

[54] Morita A, Ishii M, Asakura T, Yotsukura M, Hegab AE, Kusumoto T, et al. Direct reprogramming of mouse fibroblasts into self-renewable alveolar epithelial-like cells. npj Regen Med. 2025;10(1):30.