nature reviews genetics

https://doi.org/10.1038/s41576-025-00899-y

Supplementary information

Redefining cellular reprogramming with advanced genomic technologies

In the format provided by the authors and unedited

Supplementary Table 1. Key single-cell genomic technology applications to cell engineering

Approach	Overview	Input	Reprogramming applications	Example insights
Single-cell profiling				
scRNA-seq[1]	Profiling transcriptomes at the single-cell level to study cellular heterogeneity and gene regulation	Dissociated cells or nuclei	iPSC reprogramming[2–4]; iN reprogramming[5]; iCM reprogramming [6,7]; iEP reprogramming [8]; Myogenic reprogramming[9,10]	Treutlein *et al.* used scRNA-seq to dissect fibroblast-to-neuron conversion, revealing a continuous reprogramming trajectory with a transient intermediate state distinct from both starting fibroblasts and resulting neurons[5].
scATAC-seq[11,12]	Profiling chromatin accessibility at the single-cell level to study regulatory elements and gene regulation	Dissociated nuclei	iPSC reprogramming [13,14]; iEP reprogramming[8];; Bidirectional XEN-iPSC reprogramming[15]	Chronis *et al.* applied scATAC-seq in early iPSC induction and found that OSK factors cooperatively bind active somatic enhancers and initiate their inactivation,

				thereby reshaping the chromatin landscape to facilitate pluripotency establishment[13].
Multiomics/Joint profiling methods[16–21]	Joint measurement of chromatin accessibility and gene expression in single cells, often extended to include DNA methylation and histone modifications	Dissociated nuclei	Astrocyte-to-neuron reprogramming[22]; Fibroblast-to-pancreatic reprogramming[23]; iCM reprogramming[24]	Joint scRNA-seq and scATAC-seq revealed a two-step fibroblast-to-pancreatic trajectory via endoderm-like intermediates, identifying GATA4-driven maturation as essential for efficient lineage conversion[23].
Spatial methods[25–27]	Technology measuring gene expression within intact tissue sections, preserving spatial context to map cellular	Sectioned tissue; cells fixed on slides	iPSC reprogramming; ESC differentiation[27]; Fibroblast-to-cardiomyocyte reprogramming[28]	Yamada *et al*. applied spatial transcriptomics to reveal localized transcriptional signatures and distinct fibroblast subpopulations during direct

	heterogeneity and gene activity across tissue architecture			fibroblast-to-cardiomyocyte reprogramming in mouse hearts[28]
Cell type classification				
Garnett[29]; ScPred[30]; SingleCellNet[31]; CellTypist[32]; GPTCelltype[33]	Computational tools that use supervised learning, marker genes, and gene expression profiles to classify and annotate cell identities	scRNA-seq; scATAC-seq data	Fibroblast-to-dendritic cell reprogramming[34]	Rosa *et al*. applied CellTypist to annotate single-cell identities, uncovering dynamic transitions and intermediate populations in human fibroblast-to-dendritic cell reprogramming[34].
Quadratic Programming[5]; Capybara[35]	Mathematical optimization to assign cell states by minimizing deviations from predefined gene expression profiles of reference cell types	scRNA-seq	iN reprogramming[5]; iEP reprogramming; iCM reprogramming, iMN programming[35]	Kong *et al.* introduced Capybara, a quadratic programming-based tool that classifies discrete and “hybrid” cell states from scRNA-seq data, enabling quantitative tracking of intermediate cell identities and fate

				transitions during direct lineage reprogramming[35]
scBERT [36]	A transformer-based model for scRNA-seq cell type annotation	scRNA-seq	n/d	n/d
Cell signaling inference				
CellPhoneDB[37], NicheNet[38], and CellChat[39]	Inference of autocrine, paracrine, and juxtacrine signaling via integration of scRNA-seq data with protein-protein interaction databases	scRNA-seq	n/d	n/d
Spectra[40]	Decomposes gene expression into functional programs, maps ligand-receptor interactions and quantifies cell-cell signaling networks to	scRNA-seq	iEP reprogramming[8]	Spectra inferred TGF-β signals promoting off-target mesenchymal fates in fibroblast-to-iEP reprogramming[8].

	predict pathway activity			
Trajectory inference				
*Monocle[9,10]; Slingshot[41]; Wanderlust[42]; Wishbone [43]	Arranging cells along a trajectory that reflects their transcriptional progression rather than the time of collection	scRNA-seq; scATAC-seq	Myogenic reprogramming [9,10]; Fibroblast-to-cardiomyocyte reprogramming[44]	Cacchiarelli *et al.* applied pseudotemporal ordering (Monocle) to MyoD-induced myogenic reprogramming, revealing that aberrant insulin and BMP signaling underlie unsuccessful conversions to muscle fate[9].
WaddingtonOT[3]	An optimal transport-based method for inferring probabilistic cell fate trajectories from single-cell transcriptomic data	scRNA-seq	iPSC reprogramming[3]	Schiebinger *et al.* leveraged an optimal transport framework (Waddington-OT) on time-series scRNA-seq data of OSKM-induced reprogramming, reconstructing probabilistic cell-fate trajectories and successfully predicting molecular

				interventions that improved reprogramming efficiency[3].
SPRING[45]	Embedding single-cell data into force-directed graphs, preserving the continuum of transcriptional states for interactive exploration	scRNA-seq	MN differentiation and programming[46]	Briggs *et al.* used the SPRING single-cell graph visualization to show that pluripotent stem cells can differentiate via multiple distinct paths to the same terminal state, demonstrating convergence of divergent lineage trajectories during differentiation[46].
Lineage tracing				
CellTagging[8,47]; Rewind[48]	Lentiviral integration of random barcodes into cells. Barcodes are expressed and captured as part of the single-cell transcriptome, allowing	scRNA-seq and scATAC-seq of barcoded cells	iEP reprogramming[8,47]; iPSC reprogramming[48]	Biddy *et al.* developed high-throughput *CellTagging* barcodes to trace clonal lineages in fibroblast-to-iEP reprogramming, uncovering two early-diverging cell trajectories

	clonally related cells to be identified through shared barcodes.			(productive vs. dead-end) and identifying *Mettl7a1* as a marker and enhancer of successful reprogramming when added to the TF cocktail[47].
GRN Inference				
Single-cell GRN inference: Pando[49], SCENIC[50], SCENIC+[51]	Incorporation of single-cell expression information to infer GRNs and nominate TF regulators	scRNA-seq and scATAC-seq	iPSC reprogramming[14,52]; iEP reprogramming[53]	Tran *et al.* combined single-cell transcriptomics with network analysis to define checkpoints in OSKM-driven reprogramming, showing that continued mesenchymal gene expression hinders pluripotency and that 2i plus epigenetic modulators synergistically repress somatic programs while activating

				pluripotency networks [52].
Experimental perturbation				
Reprogram-seq[54]; Human TFome[55]	Systematic assessments of transcriptional responses to combinatorial TF perturbations in reprogramming and differentiation	scRNA-seq of cells overexpressing TFs	Epicardial-like reprogramming[54]; Human ESC differentiation[55]	Duan *et al.* used *Reprogram-seq* – single-cell RNA-seq coupled with pooled TF overexpression – to systematically evaluate transcription factor combinations, identifying an optimal cocktail that efficiently reprograms fibroblasts into epicardial-like cells[54].
Single-cell CRISPR-based perturbation[56,57]	A high-throughput technique that perturbs multiple genes simultaneously in a population of cells	Single-cell profiling of CRISPR-perturbed cells	ESC and fibroblast-to-neuron[58]; iPSC reprogramming[59]; iEP reprogramming[53]	Nourreddine et al. (preprint) performed genome-scale single-cell CRISPRi in human iPSCs, identifying novel regulators that modulate pluripotency and cell identity

				transitions during reprogramming[59].
***In silico* perturbation analysis**				
CellOracle[60]; SCENIC+[51]	Computational tools that infer GRNs from single-cell data and predict cell identity changes following TF perturbations	scRNA-seq and ATAC-seq	iEP reprogramming[53]	Kamimoto *et al.* developed *CellOracle*, which infers single-cell gene regulatory networks and simulates TF perturbations in silico, allowing the prediction of factors that drive cell identity shifts (as validated in iEP reprogramming to identify key pro- and anti-reprogramming TFs)[53].
GEARS[61]	A graph-based computational framework that predicts transcriptional outcomes of novel multi-gene perturbations using gene regulatory interactions	scRNA-seq	n/d	n/d

Geneformer[62]; scGPT	Transformer-based deep learning models trained on large-scale single-cell data to predict gene expression patterns and transcriptional responses to perturbations	scRNA-seq	iPSC reprogramming[62]	Theodoris *et al.* trained a large-scale transformer model (Geneformer) on ~30 million single-cell profiles to enable context-aware predictions of gene expression and perturbation responses, which they showed can forecast transcriptional changes during human iPSC reprogramming and identify influential regulators[62].
Molecular recording				
CAMERA[63]; DOMINO[64]	CRISPR-based genome editing to record biological signals	CRISPR-edited cells (signal records)	n/d	n/d
DNA Typewriter[65]; ENGRAM[66]	CRISPR-Cas9-based genome editing to sequentially record events	CRISPR-edited cells (signal records)	Mouse embryonic stem cell differentiation into gastruloids[66].	Chen *et al.* developed *ENGRAM*, a CRISPR-mediated “DNA

	into cellular DNA			typewriter" that sequentially writes cellular signaling events into the genome, which they used to reconstruct the timing of differentiation cues during embryoid body formation from stem cells[66].
Calling Cards[67]	PiggyBac Transposase-based system to record TF binding	PiggyBac Transposase-TF fusion (Calling Card) marked cells	n/d	n/d
uLIPSTIC[68]	Sortase A-mediated substrate transfer to mark cell-cell interactions	Sortase A-mediated-labeled cells	n/d	n/d
Synthetic biology				
Synthetic gene circuits[69]	Cooperative TF assemblies for highly tunable gene regulation	Engineered cells	n/d	n/d
CRISPR-based synthetic transcription[70]	Modular and combinatorial control of TF activity with	CRISPR-edited cells	n/d	n/d

	engineered dCas9-fusion proteins			
Engineered synthetic matrices and cell-mimetic scaffolds[71,72]	Engineering strategies to recapitulate the niche	Cells and supporting materials	Organoid culture[71,72]	Cruz-Acuña et al. showed synthetic hydrogels enhance human intestinal organoid differentiation and colonic regeneration[71]
Synthetic signaling systems[73]	Optogenetic and chemical-inducible pathways to allow real-time control of extrinsic signals		Human embryonic stem cells differentiation to mesendoderm[73].	Repina et al. demonstrated that optogenetic activation of Wnt signaling precisely directs human pluripotent stem cell fate[73]

n/d: not yet documented in a reprogramming context. This table is not exhaustive and is primarily intended to highlight the examples discussed in this Review. *Trajectory inference has been broadly applied across many reprogramming strategies.

References

1. Tang, F. *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* **6**, 377–382 (2009).
2. Buganim, Y. *et al.* Single-Cell Expression Analyses during Cellular Reprogramming Reveal an Early Stochastic and a Late Hierarchic Phase. *Cell* **150**, 1209–22 (2012).
3. Schiebinger, G. *et al.* Optimal-Transport Analysis of Single-Cell Gene Expression Identifies Developmental Trajectories in Reprogramming. *Cell* **176**, 928-943.e22 (2019).
4. Panariello, F. *et al.* Cellular population dynamics shape the route to human pluripotency. *Nature Communications 2023 14:1* **14**, 1–15 (2023).
5. Treutlein, B. *et al.* Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. *Nature* (2016) doi:10.1038/nature18323.

6. Stone, N. R. *et al.* Context-Specific Transcription Factor Functions Regulate Epigenomic and Transcriptional Dynamics during Cardiac Reprogramming. *Cell Stem Cell* **25**, 87-102.e9 (2019).
7. Liu, Z. *et al.* Single-cell transcriptomics reconstructs fate conversion from fibroblast to cardiomyocyte. *Nature* **551**, 100–104 (2017).
8. Jindal, K. *et al.* Single-cell lineage capture across genomic modalities with CellTag-multi reveals fate-specific gene regulatory changes. *Nat Biotechnol* **42**, 946–959 (2024).
9. Cacchiarelli, D. *et al.* Aligning Single-Cell Developmental and Reprogramming Trajectories Identifies Molecular Determinants of Myogenic Reprogramming Outcome. *Cell Syst* **7**, 258-268.e3 (2018).
10. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* **32**, 381–6 (2014).
11. Buenrostro, J. D. *et al.* Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* **523**, 486–490 (2015).
12. Cusanovich, D. A. *et al.* Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science (1979)* **348**, 910–914 (2015).
13. Chronis, C. *et al.* Cooperative Binding of Transcription Factors Orchestrates Reprogramming. *Cell* **168**, 442-459.e20 (2017).
14. Xing, Q. R. *et al.* Diversification of reprogramming trajectories revealed by parallel single-cell transcriptome and chromatin accessibility sequencing. *Sci Adv* **6**, 18 (2020).
15. Garg, V. *et al.* Single-cell analysis of bidirectional reprogramming between early embryonic states identify mechanisms of differential lineage plasticities in mice. *Dev Cell* (2024) doi:10.1016/J.DEVCEL.2024.11.022.
16. Ma, S. *et al.* Chromatin Potential Identified by Shared Single-Cell Profiling of RNA and Chromatin. *Cell* **183**, 1103-1116.e20 (2020).
17. Cao, J. *et al.* Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science (1979)* **361**, 1380–1385 (2018).
18. Chen, S., Lake, B. B. & Zhang, K. High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell. *Nature Biotechnology 2019 37:12* **37**, 1452–1457 (2019).
19. Zhu, C. *et al.* An ultra high-throughput method for single-cell joint analysis of open chromatin and transcriptome. *Nature Structural & Molecular Biology 2019 26:11* **26**, 1063–1070 (2019).
20. Clark, S. J. *et al.* scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nature Communications 2018 9:1* **9**, 1–9 (2018).
21. Zhu, C. *et al.* Joint profiling of histone modifications and transcriptome in single cells from mouse brain. *Nat Methods* **18**, 283–292 (2021).
22. Pereira, A. *et al.* Direct neuronal reprogramming of mouse astrocytes is associated with multiscale epigenome remodeling and requires Yy1. *Nature Neuroscience 2024 27:7* **27**, 1260–1273 (2024).
23. Fei, L., Zhang, K., Poddar, N., Hautaniemi, S. & Sahu, B. Single-cell epigenome analysis identifies molecular events controlling direct conversion of human fibroblasts to pancreatic ductal-like cells. *Dev Cell* **58**, 1701-1715.e8 (2023).

24. Wang, H., Yang, Y., Qian, Y., Liu, J. & Qian, L. Delineating Chromatin Accessibility Re-patterning at Single Cell Level during Early Stage of Direct Cardiac Reprogramming. *J Mol Cell Cardiol* **162**, 62 (2022).
25. Rao, A., Barkley, D., França, G. S. & Yanai, I. Exploring tissue architecture using spatial transcriptomics. *Nature 2021 596:7871* **596**, 211–220 (2021).
26. Moffitt, J. R., Lundberg, E. & Heyn, H. The emerging landscape of spatial profiling technologies. *Nat Rev Genet* **23**, 741–759 (2022).
27. Kobayashi-Kirschvink, K. J. *et al.* Prediction of single-cell RNA expression profiles in live cells by Raman microscopy with Raman2RNA. *Nature Biotechnology 2024* 1–9 (2024) doi:10.1038/s41587-023-02082-2.
28. Yamada, Y. *et al.* Cardiac Reprogramming and Gata4 Overexpression Reduce Fibrosis and Improve Diastolic Dysfunction in Heart Failure with Preserved Ejection Fraction. *Circulation* (2024) doi:10.1161/CIRCULATIONAHA.123.067504/ASSET/237043EC-0D3F-4248-BC36-F4FBC7FE9D4B/ASSETS/GRAPHIC/CIRCULATIONAHA.123.067504.FIG08.JPG.
29. Pliner, H. A., Shendure, J. & Trapnell, C. Supervised classification enables rapid annotation of cell atlases. *Nat Methods* **16**, 983–986 (2019).
30. Alquicira-Hernandez, J., Sathe, A., Ji, H. P., Nguyen, Q. & Powell, J. E. scPred: accurate supervised method for cell-type classification from single-cell RNA-seq data. *Genome Biol* **20**, 264 (2019).
31. Tan, Y. & Cahan, P. SingleCellNet: A Computational Tool to Classify Single Cell RNA-Seq Data Across Platforms and Across Species. *Cell Syst* **9**, 207-213.e2 (2019).
32. Domínguez Conde, C. *et al.* Cross-tissue immune cell analysis reveals tissue-specific features in humans. *Science (1979)* **376**, (2022).
33. Hou, W. & Ji, Z. Assessing GPT-4 for cell type annotation in single-cell RNA-seq analysis. *Nature Methods 2024 21:8* **21**, 1462–1465 (2024).
34. Rosa, F. F. *et al.* Single-cell transcriptional profiling informs efficient reprogramming of human somatic cells to cross-presenting dendritic cells. *Sci Immunol* **7**, (2022).
35. Kong, W. *et al.* Capybara: A computational tool to measure cell identity and fate transitions. *Cell Stem Cell* **29**, 635-649.e11 (2022).
36. Yang, F. *et al.* scBERT as a large-scale pretrained deep language model for cell type annotation of single-cell RNA-seq data. *Nature Machine Intelligence 2022 4:10* **4**, 852–866 (2022).
37. Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nature Protocols 2020 15:4* **15**, 1484–1506 (2020).
38. Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nature Methods 2019 17:2* **17**, 159–162 (2019).
39. Jin, S. *et al.* Inference and analysis of cell-cell communication using CellChat. *Nature Communications 2021 12:1* **12**, 1–20 (2021).
40. Kunes, R. Z., Walle, T., Land, M., Nawy, T. & Pe'er, D. Supervised discovery of interpretable gene programs from single-cell data. *Nature Biotechnology 2023 42:7* **42**, 1084–1095 (2023).
41. Street, K. *et al.* Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).

42. Bendall, S. C. *et al.* Single-Cell Trajectory Detection Uncovers Progression and Regulatory Coordination in Human B cell Development. *Cell* **157**, 714 (2014).
43. Setty, M. *et al.* Wishbone identifies bifurcating developmental trajectories from single-cell data. *Nat Biotechnol* **34**, 637–645 (2016).
44. Zhou, Y. *et al.* Single-Cell Transcriptomic Analyses of Cell Fate Transitions during Human Cardiac Reprogramming. *Cell Stem Cell* **25**, 149-164.e9 (2019).
45. Weinreb, C., Wolock, S. & Klein, A. M. SPRING: a kinetic interface for visualizing high dimensional single-cell expression data. *Bioinformatics* **34**, 1246–1248 (2018).
46. Briggs, J. A. *et al.* Mouse embryonic stem cells can differentiate via multiple paths to the same state. *Elife* **6**, (2017).
47. Biddy, B. A. *et al.* Single-cell mapping of lineage and identity in direct reprogramming. *Nature 2018 564:7735* **564**, 219–224 (2018).
48. Jain, N. *et al.* Retrospective identification of cell-intrinsic factors that mark pluripotency potential in rare somatic cells. *Cell Syst* **0**, (2024).
49. Fleck, J. S. *et al.* Inferring and perturbing cell fate regulomes in human brain organoids. *Nature 2022 621:7978* **621**, 365–372 (2022).
50. Aibar, S. *et al.* SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* **14**, 1083–1086 (2017).
51. Bravo González-Blas, C. *et al.* SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks. *Nature Methods 2023 20:9* **20**, 1355–1367 (2023).
52. Tran, K. A. *et al.* Defining Reprogramming Checkpoints from Single-Cell Analyses of Induced Pluripotency. *Cell Rep* **27**, 1726-1741.e5 (2019).
53. Kamimoto, K. *et al.* Gene regulatory network reconfiguration in direct lineage reprogramming. *Stem Cell Reports* **18**, 97–112 (2023).
54. Duan, J. *et al.* Rational Reprogramming of Cellular States by Combinatorial Perturbation. *Cell Rep* **27**, 3486 (2019).
55. Joung, J. *et al.* A transcription factor atlas of directed differentiation. *Cell* **186**, 209-229.e26 (2023).
56. Bock, C. *et al.* High-content CRISPR screening. *Nature Reviews Methods Primers 2022 2:1* **2**, 1–23 (2022).
57. Schraivogel, D., Steinmetz, L. M. & Parts, L. Pooled Genome-Scale CRISPR Screens in Single Cells. *Annu Rev Genet* **57**, 223–244 (2023).
58. Liu, Y. *et al.* CRISPR Activation Screens Systematically Identify Factors that Drive Neuronal Fate and Reprogramming. *Cell Stem Cell* **23**, 758–771 (2018).
59. Nourreddine, S. *et al.* A Perturbation Cell Atlas of Human Induced Pluripotent Stem Cells. *bioRxiv* 2024.11.03.621734 (2024) doi:10.1101/2024.11.03.621734.
60. Kamimoto, K. *et al.* Dissecting cell identity via network inference and in silico gene perturbation. *Nature 2023 614:7949* **614**, 742–751 (2023).
61. Roohani, Y., Huang, K. & Leskovec, J. Predicting transcriptional outcomes of novel multigene perturbations with GEARS. *Nature Biotechnology 2023 42:6* **42**, 927–935 (2023).
62. Theodoris, C. V. *et al.* Transfer learning enables predictions in network biology. *Nature 2023 618:7965* **618**, 616–624 (2023).
63. Tang, W. & Liu, D. R. Rewritable multi-event analog recording in bacterial and mammalian cells. *Science (1979)* **360**, (2018).

64. Farzadfard, F. *et al.* Single-Nucleotide-Resolution Computing and Memory in Living Cells. *Mol Cell* **75**, 769-780.e4 (2019).
65. Choi, J. *et al.* A time-resolved, multi-symbol molecular recorder via sequential genome editing. *Nature 2022 608:7921* **608**, 98–107 (2022).
66. Chen, W. *et al.* Symbolic recording of signalling and cis-regulatory element activity to DNA. *Nature 2024* 1–9 (2024) doi:10.1038/s41586-024-07706-4.
67. Moudgil, A. *et al.* Self-Reporting Transposons Enable Simultaneous Readout of Gene Expression and Transcription Factor Binding in Single Cells. *Cell* **182**, 992-1008.e21 (2020).
68. Nakandakari-Higa, S. *et al.* Universal recording of immune cell interactions in vivo. *Nature 2024 627:8003* **627**, 399–406 (2024).
69. Bashor, C. J. *et al.* Complex signal processing in synthetic gene circuits using cooperative regulatory assemblies. *Science (1979)* **364**, 593–597 (2019).
70. Chen, W. C. W. *et al.* A synthetic transcription platform for programmable gene expression in mammalian cells. *Nature Communications 2022 13:1* **13**, 1–16 (2022).
71. Cruz-Acuña, R. *et al.* Synthetic hydrogels for human intestinal organoid generation and colonic wound repair. *Nature Cell Biology 2017 19:11* **19**, 1326–1335 (2017).
72. Gjorevski, N. *et al.* Designer matrices for intestinal stem cell and organoid culture. *Nature 2016 539:7630* **539**, 560–564 (2016).
73. Repina, N. A. *et al.* Optogenetic control of Wnt signaling models cell-intrinsic embryogenic patterning using 2D human pluripotent stem cell culture. *Development (Cambridge)* **150**, (2023).