© 2023. Published by The Company of Biologists Ltd | Development (2023) 150, dev202300. doi:10.1242/dev.202300

SPOTLIGHT

Programming human cell fate: overcoming challenges and unlocking potential through technological breakthroughs

Hsiu-Chuan Lin[1,*,‡], Aly Makhlouf[2,*,‡], Camila Vazquez Echegaray[3,*,‡], Dorota Zawada[4,5,6,*,‡] and Filipa Simões[7,*,‡]

ABSTRACT

In recent years, there have been notable advancements in the ability to programme human cell identity, enabling us to design and manipulate cell function in a Petri dish. However, current protocols for generating target cell types often lack efficiency and precision, resulting in engineered cells that do not fully replicate the desired identity or functional output. This applies to different methods of cell programming, which face similar challenges that hinder progress and delay the achievement of a more favourable outcome. However, recent technological and analytical breakthroughs have provided us with unprecedented opportunities to advance the way we programme cell fate. The Company of Biologists' 2023 workshop on 'Novel Technologies for Programming Human Cell Fate' brought together experts in human cell fate engineering and experts in single-cell genomics, manipulation and characterisation of cells on a single (sub)cellular level. Here, we summarise the main points that emerged during the workshop's themed discussions. Furthermore, we provide specific examples highlighting the current state of the field as well as its trajectory, offering insights into the potential outcomes resulting from the application of these breakthrough technologies in precisely engineering the identity and function of clinically valuable human cells.

KEY WORDS: Cell programming, Genomic engineering, Human cell fate, Reprogramming, Synthetic biology

Introduction

In recent years, we have witnessed significant advancements in the field of programming human cell identity, allowing for the precise design and manipulation of cellular function within controlled environments. One aim of cell programming is to achieve specific fate outcomes tailored to the intended application, such as the generation of disease-specific cell models to study disease mechanisms (Avior et al., 2016; Trounson and DeWitt, 2016), personalised therapies (Barker et al., 2017; Zimmermannova et al., 2023), the production of cells for regenerative medicine (Vierbuchen et al., 2010; Song et al., 2012) or a deeper understanding of the underlying mechanisms governing cell fate decisions (VanHorn and Morris, 2021; Jindal et al., 2023). However, protocols for generating such desired cell types are often limited in efficiency and precision, resulting in engineered cells that fall short of fully replicating the intended identity or functional output. Such challenges are encountered across various methods of cell programming.

Cell programming approaches can be broadly categorised: (1) reprogramming, in which mature, specialised cells undergo a reversal of their developmental state, regaining pluripotency to then adopt a new cellular identity and function; (2) programming from pluripotent cells or progenitors, where less-committed cells are guided towards specific cell fates; (3) direct reprogramming or transdifferentiation, the conversion of one cell type directly into another without the need for an intermediate precursor stage; and (4) synthetic biology, which provides tools to precisely manipulate and programme cell behaviour by engineering synthetic gene circuits, modifying signalling pathways or regulating gene expression at the molecular level with unprecedented resolution.

Novel single-cell technologies hold immense potential for advancing cell and tissue engineering. They enable in-depth molecular characterisation of engineered cell states, allowing for an accurate assessment of the efficiency, precision and harmonisation of existing protocols. In addition, the availability of single-cell atlases for primary developing and adult tissues are valuable resources for guiding cell engineering efforts and predicting the necessary requirements and design strategies for generating selected cell fates. Further progress has been facilitated by CRISPR engineering and synthetic biology, enabling meticulous regulation of gene expression, therapeutic activity, and control over dosage, timing and localisation of programming factors.

The Company of Biologists organised the 'Novel Technologies for Programming Human Cell Fate' workshop in 2023, bringing together experts in human cell fate engineering, single-cell genomics, cell state manipulation and characterisation. The workshop provided a platform for a deeper understanding of the challenges, advancements and potential applications in programming human cell fate.

Here, we summarise the current field of human cell fate programming (Fig. 1). Drawing on insights emerging from thought-provoking discussions at the workshop, we offer a deeper

[1]Department of Biosystems Science and Engineering, ETH Zürich, 4057 Basel, Switzerland. [2]MRC Laboratory of Molecular Biology, University of Cambridge, Cambridge CB2 0QH, UK. [3]Molecular Medicine and Gene Therapy, Lund Stem Cell Centre, Wallenberg Centre for Molecular Medicine, Lund University, 221 84 Lund, Sweden. [4]First Department of Medicine, Cardiology, Klinikum rechts der Isar, Technical University of Munich, School of Medicine and Health, 81675 Munich, Germany. [5]German Center for Cardiovascular Research (DZHK), Munich Heart Alliance, 80636 Munich, Germany. [6]Regenerative Medicine in Cardiovascular Diseases, First Department of Medicine, Klinikum rechts der Isar, Technical University of Munich, School of Medicine and Health, 81675 Munich, Germany. [7]Department of Physiology, Anatomy and Genetics, Institute of Developmental and Regenerative Medicine, University of Oxford, Oxford OX3 7TY, UK.
*These authors contributed equally to this work

‡Authors for correspondence (hsiuchuan.lin@bsse.ethz.ch; amakhlouf@mrc-lmb.cam.ac.uk; camila.vazquez_echegaray@med.lu.se; dm.zawada@tum.de; filipa.simoes@idrm.ox.ac.uk)

H.-C.L., 0000-0001-9587-6483; A.M., 0000-0003-4025-0576; C.V.E., 0000-0001-8849-6853; D.Z., 0000-0003-1127-7980; F.S., 0000-0003-2956-0311

Development (2023) 150, dev202300. doi:10.1242/dev.202300

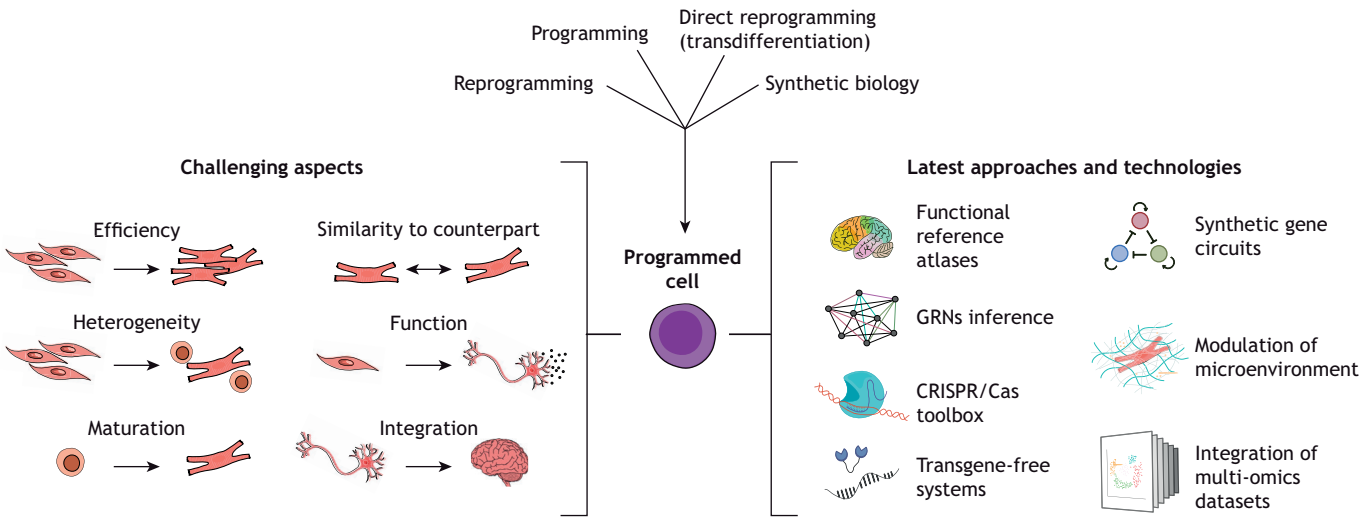

Fig. 1. Technological breakthroughs offer solutions to address challenges in achieving the desired cell output. Programming human cell fate can be achieved through different approaches, including reprogramming, programming, direct reprogramming (transdifferentiation) and synthetic biology. Advances in creating comprehensive reference cell atlases and multi-omics databases, inferring gene regulatory networks (GRNs), designing precise and tunable CRISPR/Cas9-based tools, developing efficient delivery systems, refining synthetic gene circuits and recreating *in vivo* niches can help overcome barriers associated with obtaining the desired cell identity and function. These challenges include issues such as low efficiency, heterogeneity in generated cell products, limited maturation, dissimilarity to counterparts, lack of desired cell function and *in vivo* integration.

understanding of the field's landscape. In addition, we provide a glimpse into the exciting prospects that lie ahead, pushing the boundaries of what can be achieved in this rapidly evolving field of research.

Challenges

Progress in recent years makes cellular programming an exciting field. Much effort has been invested into developing engineered cells to resemble their *in vivo* counterparts, and this is already reaching clinical potential (Kirkeby et al., 2023). As exciting as this scenario appears, there are challenges and concerns that need to be addressed. Cell programming can cause unexpected outcomes (Treutlein et al., 2016) and many current approaches generate cells that recapitulate only certain aspects of their bona fide *in vivo* counterparts in terms of transcriptome, cellular properties and function. Many reasons might account for this, such as an incomplete understanding of developmental signalling cues, using a minimal set of factors to obtain desired cell types, lack of a detailed reference to guide *in vitro* work or the general artificiality of culture systems. To guarantee the safety and efficacy of cell-based interventions, it is vital to conduct comprehensive analysis and monitoring of programmed cells, through genomic profiling, functional assays and long-term observation to effectively identify and mitigate risks.

Efficiency and heterogeneity

Efficiency in cell programming is crucial because it determines the success of programming processes while minimising off-target effects. The development of highly efficient approaches for cell conversion represents an everyday challenge, and focuses on assessing the time, scalability, reproductivity and robustness for producing specific cell types. These parameters are tested through functional *in vitro* or *in vivo* assays (e.g. transplantation experiments; Balboa et al., 2022; Kirkeby et al., 2017; Tan et al., 2018) with the goal of generating adequate numbers of programmed cells for advancing practical applications. Another constraint that influences efficiency is the presence of programming roadblocks, which involve transcriptional and chromatin features that prevent cell fate changes (Brumbaugh et al., 2019; Arabacı et al., 2021). These barriers are not covered here but are recognised as important factors that determine success when programming cells.

Heterogeneous populations in *in vitro* cell cultures can be viewed in two different ways: high cell type diversity can introduce too much 'noise' in the system and divert the efforts put into programming specific cell types. Therefore, controlling and minimising this effect is crucial for reliable and predictable cell reprogramming. Conversely, some scenarios require the presence of a variety of cell types for proper maturation, as is the case for complex 3D systems. Here, cellular heterogeneity is desired and advantageous (Meier et al., 2023; Drakhlis et al., 2021; Cooke et al., 2023).

Heterogeneity occurs in the source cells used for programming and in the programmed population. For the former, a major challenge is whether the choice of the source cell type or state affects the outcome. Research groups have been optimising protocols tested on different source cell types, which can possibly cause discrepancies in the results obtained and conclusions. Efforts have been invested into standardising these protocols, as well as expanding the diversity of the source cells used for programming (Karow et al., 2018; Lentini et al., 2021; Zhou et al., 2008). The extent to which diversity can be controlled is limited because cells inevitably carry pre-existing mutations. Also, molecular dynamics shape the intrinsic context of cells, which can result in certain states that are not amenable to conversion. For example, how variability in expression levels of endogenous transcription factors can determine whether cells will fail or efficiently convert cell fate (Francesconi et al., 2019).

Maturation

A major challenge is recapitulating maturation processes for the acquisition of functional properties of mature cells *in vitro*, which is crucial for disease modelling and therapeutic translation. For this, understanding developmental maturation is essential but time, in

Development (2023) 150, dev202300. doi:10.1242/dev.202300

many cases, is the limiting factor. It is necessary to speed up maturation processes that can take years *in vivo*. The artificiality of culture systems, which mainly rely on plastic dishes and daily media changes, does not recapitulate the formation and nutrition of cells and tissues in a body. Also, mature phenotype acquisition of many cells does not happen in a vacuum but is dependent on many signalling cues coming from other tissues (Cooke et al., 2023; Arterbery and Bogue, 2014). For example, in direct lineage reprogramming, the generation of a more mature phenotype could be aided by choosing an aged donor as the cell source, but that also brings other challenges regarding the overall fitness of older cells (Oh et al., 2022).

Box 1. Glossary

Attractor state. The stable and well-defined endpoints of the cell engineering process, regardless of the starting conditions or the transitions that occur during the process.

Autoencoders. Neural network architectures used for unsupervised learning tasks that aim to encode and then decode input data with minimal error, often used in dimensionality reduction or feature learning.

Bridge dataset (single cell). An intermediate dataset used to connect or integrate data from different sources or experiments, often applied in single-cell analysis to infer cell states and link modalities.

Cryo-EM. Cryogenic electron microscopy, a technique that enables high-resolution imaging of biological molecules by freezing them in vitreous ice and visualising them using electron microscopy.

Directed evolution. A technique for artificially evolving proteins or other biomolecules in the laboratory and selecting desired properties or functions.

Large language model. A multi-layered computational model, pre-trained on very large datasets of text, that learns to recognise, predict and generate text in a given language.

Perturbation. The introduction of controlled disturbances or changes into a biological system in order to study its response and behaviour.

Microfluidic controlled stem cell regionalization. The manipulation of stem cell differentiation and positioning within microfluidic devices to control morphogen gradients, allowing the study of tissue development and organisation.

Minimal descriptor. A concise representation or characteristic that captures the essential information of a complex system or object.

Molecular recording. The process of capturing and storing molecular-level events or information within cells, often used for tracking cell lineages, cellular behaviour or environmental cues.

Neural network. A computational model composed of layered interconnected processing units (nodes or artificial neurons) that simulate the information processing observed in biological brains. The node layers include input, output and one or more hidden layers. Each node has an associated weight and threshold and is only activated when the output passes the threshold, passing data to the next layer.

Neuromorphic computing. A computing paradigm inspired by the architecture and functioning of the human brain, aimed at developing energy-efficient and parallel processing systems.

Optimal transport (single cell). A mathematical framework to quantify and analyse the transportation of resources from one distribution to another with minimal cost. In single-cell analysis, it is used to compare and match distributions of molecular features such as gene expression profiles.

Organ-on-chip. Microfluidic cell culture platforms that mimic the structure and function of human organs to study physiological responses and drug effects.

Sketching techniques. Methods used to approximate complex datasets with simplified representations, such as subsampling cells while maintaining rare populations, to speed up computation.

Transformer-based models. Advanced neural network architectures that are used in tasks involving sequential data, language understanding and generation.

Similarity to the primary counterpart

Nevertheless, despite the limitations of cell culture systems, we now have several cell types generated *in vitro* that resemble certain features of their *in vivo* counterparts. Conventionally, the comparison between engineered cells and *in vivo* target cells has relied on evaluating general morphologies, biomarker expressions, bulk-omics data and functional assays (Nolbrant et al., 2017; Kroon et al., 2008; Kamao et al., 2014). Recently, it has become common practice to use the transcriptome obtained *in vivo* as the blueprint to engineer the perfect cellular identity and function (Loh et al., 2014). The advent of single-cell RNA-sequencing (scRNA-seq) has revolutionised our ability to systematically and quantitatively characterise cell identities. Efforts including the Human Cell Atlas (https://www.humancellatlas.org/) and Tabula Sapiens Consortium (https://tabula-sapiens-portal.ds.czbiohub.org/) have played pivotal roles in generating comprehensive transcriptome reference atlases at single-cell resolution for diverse primary tissues, providing valuable resources for understanding cell types and states. Leveraging scRNA-seq readout, it becomes possible to quantify the transcriptome similarity between reference atlases and engineered cells, thereby facilitating the estimation of off-target lineages within the culture (La Manno et al., 2016). Though this approach is useful, it does not provide a dynamically resolved view of tissue development and it is challenging to guide *in vitro* work on snapshots that might be obtained with confounding artefacts of sample processing or with different single-cell transcriptomics methods. Although computational approaches have been developed to integrate diverse datasets, the mapping of engineered cell identities remains challenging due to differences in sequencing depth, cell clustering resolution and the absence of universally standardised cell annotations. To generate *in vivo*-like cells, we must start with a well-established *in vivo* reference to assess the validity and relevance of generated cell products and our approaches, preferably with approaches that link *in vitro* obtained properties of cells (transcriptome, proteome, function, etc.) with dynamically recorded data *in vivo* through computational modelling.

Function

Regardless of the approach used to generate a certain cell type, it is essential that obtained cells exhibit a certain set of functions and behaviours that are characteristic for equivalent cells within the body. This is especially true for drug testing or disease modelling platforms. It seems fairly straightforward to assess cell functionality *in vitro* – whether by the measurement of action potentials for cells that are electrically active or by the measurement of proteins secreted upon stimuli for cells that are metabolically active. However, results of these measurements might be difficult to interpret: should measurements obtained on isolated cells *in vitro* be the same as cells *in vivo*? Is it possible that cells that respond to stimuli *in vitro* would fail to respond to similar stimuli *in vivo* in a timely and dose-dependent manner?

One of the best approaches to achieve the full functionality of cell products is to expose them to the *in vivo* niche through transplantation. The signalling cues coming from the microenvironment can provide the necessary stimuli, enhancing terminal differentiation or maturation (Balboa et al., 2022; Kirkeby et al., 2017). Many challenges and questions remain: is it sufficient to bring cells to a certain state of functionality and let the *in vivo* niche do the rest? Can we predict how cells will behave and mature following transplantation? Can we recreate the niche *in vitro*? Depending on the goal of the study, we might need approaches to mature cells *ex vivo*.

Development (2023) 150, dev202300. doi:10.1242/dev.202300

Integration

Obtaining cellular products that are suitable for cell replacement is crucial not only to ensure their functional maturation but also functional integration with the surrounding microenvironment upon transplantation. The cellular behaviours and interactions within the native tissue must be considered to create an environment that supports the functionality and longevity of engineered cells. Strategies to enhance cell competitiveness including cell survival, optimising delivery methods and promoting tissue-specific interactions, can enhance functional integration.

Ultimately, ensuring the similarity of generated cells in terms of morphology, gene expression and functionality is essential to safeguard the validity and relevance of research findings, as well as the safety and efficacy of cell therapy.

Overcoming current challenges

Working with bona-fide reference networks for data integration

To gain a comprehensive understanding of engineered cell fates, establishing a unified framework for describing cell identities across diverse human cell atlases is crucial. The reference cell tree is a valuable approach that combines molecular states and lineage histories to address this need (Domcke and Shendure, 2023). Although single-cell genomics has transformed our understanding of cell identities, debates arise regarding whether cell identity should be solely defined by function, as cells with similar gene expression patterns may exhibit different functional behaviours (Scala et al., 2021). However, in scenarios where the *in vivo* function of cell subtypes is undefined, we contend that gene expression can serve as a bridge for extrapolating and comprehending these uncharted functions. In addition to solely relying on observed feature similarities, we can potentially disentangle cell identity emphasising functional attributes by considering cellular responses to environmental cues or perturbations (see Glossary, Box 1). Perturbation experiments and the construction of functional reference atlases can provide deeper insights into cellular functions, facilitating engineering of functional cells (Rauscher et al., 2017; http://genomecrispr.dkfz.de; https://orcs.thebiogrid.org).

It is also crucial to unravel the underlying mechanisms driving cell fate specification to effectively guide cell engineering efforts. Gene regulatory networks (GRNs) play a fundamental role in orchestrating cell fate determination, which can be inferred from gene expression patterns and relevant regulatory elements (Fleck et al., 2022; Kamimoto et al., 2023; Aibar et al., 2017; Davidson and Erwin, 2006). By comparing the GRNs of engineered cells with reference networks, we can explore whether identical fates require similar regulatory routes or if alternative expedited pathways lead to the same desired attractor states (see Glossary, Box 1). However, inferred causal relationships within GRNs often represent indirect predictions or associations. Is it possible to approach a more accurate representation of the molecular events that drive cell state transitions, thereby leveraging this knowledge to enhance cell engineering outcomes? Recent advancements in molecular recording technologies (see Glossary, Box 1) offer promising avenues for uncovering transcriptional histories and the dynamic processes underlying cell fate transitions (Farzadfard and Lu, 2018; Schmidt et al., 2018; Choi et al., 2022; Chen et al., 2022; Horns et al., 2023).

Tools for enhancing the accuracy and precision of cell fate programming

Transcription factor-mediated cell programming remains the most widespread approach for manipulating cell fate (Takahashi and Yamanaka, 2006; Barretto et al., 2020; Missinato et al., 2023; Pierson Smela et al., 2023). Directed evolution (see Glossary, Box 1) has been used to generate new transcription factor variants with improved reprogramming speeds and efficiencies, compared with wild-type counterparts (Tan et al., 2021). However, traditional transcription factor-mediated approaches continue to raise concerns over the need to introduce exogenous genes in human cells. Alternative tools, such as CRISPR activation (CRISPRa) and CRISPR interference (CRISPRi), rely on guide RNA (gRNA)-mediated targeting of DNA, in conjunction with a deactivated Cas9 nuclease, to modulate the activity of endogenous genes (Perez-Pinera et al., 2013; Sokka et al., 2022). CRISPR-Cas13 is also a potentially versatile tool for the direct targeting and post-transcriptional regulation of RNA (Adler et al., 2022). In addition, the CRISPR-Cas12a system has the ability to process multiple CRISPR-RNAs (crRNAs) in a single CRISPR array, potentially useful for multi-gene regulation (Magnusson et al., 2021). Most cellular phenotypes, including cell states, are polygenically regulated and rely on the coordinated dosage control of multiple gene products, making multi-gene regulation a challenging yet crucial bottleneck for efficient cell fate programming. Developing robust transcriptional control systems by designing programmable gene regulatory elements with reproducible effects on the transcriptional outputs of multiple genes has remained somewhat elusive, mainly due to the cellular and genetic context-dependencies of how regulatory elements behave. Together, these tools might facilitate cell fate engineering without over-expressing exogenous genes, but they still rely on generating transgenic cells and, depending on the approach, do not necessarily confer heritable cell identities.

Transgene-free systems have emerged as promising alternatives to bypass the burden, and potential risks, of genetically modified cells. One example is the use of an antibody-based reprogramming approach, amenable to high-throughput, combinatorial library screening, which can identify cell surface-targeting antibodies with the same downstream signalling effects in the cell for reprogramming as the respective transcription factors (Blanchard et al., 2017). There is also the added benefit that antibodies have higher selectivity compared with small molecules, another popular approach for manipulating signalling pathways during cell reprogramming. Therefore, the ability to design fit-for-purpose antibodies could dramatically expand the potential of this approach. Indeed, recent advances in large language models (see Glossary, Box 1) are already seeing growing applications in generative protein design and evolution (Hie et al., 2023; Watson et al., 2023). As such, the protein design space is becoming increasingly more attainable and can serve as a ‘playground’ for generating new and functional protein-based tools for cell reprogramming. Nanobodies, for example, are increasingly applied to reprogramming immune cells for cancer immunotherapy and CAR-T cell engineering (Hie et al., 2023; Ma et al., 2020). With the potential benefits of cost, size, delivery and binding specificity compared with traditional antibodies, nanobodies may serve as a promising tool for future innovations in cell reprogramming. In addition, there are RNA-based approaches, including miRNA and mRNA (Yakubov et al., 2010; Kogut et al., 2018; Warren et al., 2010). Active areas of research aim to improve targeted cell delivery approaches (Paunovska et al., 2022), RNA encapsulation methods (Hou et al., 2021; Tanaka et al., 2023), RNA modifications that confer higher stability and lower immunogenicity (Paunovska et al., 2022; Kim et al., 2022; Liu and Wang, 2022), as well as reduce variability associated with RNA transfection efficiency (Shin and Min, 2023).

Development (2023) 150, dev202300. doi:10.1242/dev.202300

In future, advances in cell fate engineering will rely on the creation of cellular systems that can respond to spatiotemporal cues with high precision and appropriate sensitivity in a cell type-specific manner. The field of synthetic biology offers the opportunity to engineer programmable, multi-gene and multicellular systems that can potentially be constructed from modularised components, such as synthetic receptors (Morsut et al., 2016), synthetic cell-cell signalling networks (Toda et al., 2018) or synthetic, multi-stable gene circuits (Zhu et al., 2022). Such complex synthetic biological systems rely on the ability to encode complex cellular logic through genetic circuits, which have more traditionally taken inspiration from digital circuit design. More recent work has been inspired by principles of neuromorphic computing (see Glossary, Box 1; Rizik et al., 2022), by creating a neural network (see Glossary, Box 1) architecture in *Escherichia coli* to perform perceptron-based cellular computations, which can go beyond the capabilities of basic digital and analogue circuits. Together, synthetic biology tools can potentially be used for both *in vitro* applications, where precise, automated, multi-step cell fate acquisition may be needed, or *in vivo* applications, where engineered cell states that can process *in vivo* signalling cues and respond accordingly are required.

Finally, techniques that facilitate cell fate engineering by direct manipulation of the cell microenvironment, rather than the cell itself, are growing. Techniques, such as StemBond hydrogels, modulate active cell signalling pathways through the selective control of the mechanical and/or biochemical properties of the substrate. Recent work using alginate hydrogels (Elosegui-Artola et al., 2023) has shown that the viscoelastic properties of the extracellular matrix can control cell signalling, symmetry-breaking, proliferation and morphology, particularly in the context of multicellular aggregates and organoids. Complex organoid systems also rely on the co-culture of a target cell type with one or more auxiliary cell types to promote target cell fate acquisition by mechanical constraints, cell-cell signalling and growth factor secretion, which combine to create a supportive cell niche. This is exemplified by recent advances in stem cell-based models of human embryos, which precisely co-culture several defined cell types to generate higher-order complexity *in vitro* (Weatherbee et al., 2023; Ai et al., 2023; Yuan et al., 2023 preprint; Hislop et al., 2023 preprint; Pedroza et al., 2023; Oldak et al., 2023). Microfluidic approaches can also precisely control spatiotemporal signalling cues. For example, the microfluidic-controlled stem cell regionalisation (MiSTR; see Glossary, Box 1) system generates spatially patterned WNT gradients for studying neural tube development (Rifes et al., 2020). More broadly, the advent of 'Organ-on-Chip' systems (see Glossary, Box 1), although hampered by issues of complexity, scalability, cost and standardisation, could provide opportunities for achieving reproducible cell fate reprogramming in more chemically-defined and physically constrained microenvironments.

Comprehensive multi-modal data integration

We previously highlighted the power of single-cell transcriptomes to describe engineered cell identities. However, to understand the mechanisms behind cell fate transitions and predict outcomes, relying solely on transcriptomics may be insufficient. Integration of diverse omics datasets could further our understanding of cell states and transitions. By combining single-cell transcriptomics with other modalities, such as chromatin accessibility, specific chromatin modifications, DNA methylation and protein measurements, we can obtain a detailed and holistic view of cellular identities (Stuart and Satija, 2019; Zhu et al., 2020; Baysoy et al., 2023). This multi-layered approach goes beyond gene expression, unravelling the intricate regulatory mechanisms that govern cell fate decisions. Incorporating protein measurements is particularly important because protein levels often deviate from gene expression levels (Vogel and Marcotte, 2012; Reimegård et al., 2021). Although direct measurement of all these modalities in single cells is not currently possible, emerging computational approaches, including integrating modalities based on a bridge dataset or optimal transport (see Glossary, Box 1), offer promising solutions (Hao et al., 2023; Klein et al., 2023 preprint). Computational efficiency has also been enhanced through sketching techniques (see Glossary, Box 1) and graphical processing unit (GPU) acceleration, enabling effective processing of atlas-scale datasets containing up to millions of cells (Hao et al., 2023; Nolet et al., 2022 preprint). Scalable single-cell multi-omics empowers us to monitor and predict outcomes following perturbations, facilitating the design of precise and effective interventions in cell engineering. This comprehensive understanding and integration of multiple omics dimensions significantly advances our capability to engineer cells with desired identities.

Single-cell spatial omics is a powerful complement to single-cell molecular multi-omics, providing valuable spatial context to molecular profiles. Technologies such as spatial transcriptomics, multiplexed protein staining, imaging mass cytometry and 3D spatial mass cytometry allow us to capture cell morphology, intracellular organisation and cellular polarity that emerge during cell-fate manipulation (Bhatia et al., 2022; Kuett et al., 2022; Rodriques et al., 2019; Gut et al., 2018; Moffitt et al., 2016; Giesen et al., 2014). Cryo-electron microscopy (Cryo-EM; see Glossary, Box 1) offers high-resolution structural information of cellular components and molecular complexes and the potential to decipher architectural changes during cell fate transitions (Pfeffer and Mahamid, 2018). Single-cell spatial omics also contribute to our understanding of cell-cell interactions and tissue microenvironments (Kanemaru et al., 2023). It enables the mapping of molecular gradients, cell-cell signalling pathways and spatially restricted niche factors that influence cell behaviour and fate decisions, which is crucial for recreating complex tissue microenvironments in engineered systems or for engineering cellular therapies that can integrate seamlessly within native tissues. For example, CellPhoneDB (Efremova et al., 2020) leverages data on the combined expression of multi-subunit ligand–receptor complexes to infer intercellular communication, and NicheNet models cell-cell communication by linking ligands to target genes (Browaeys et al., 2020). Importantly, there is now opportunity to combine intercellular communication inference tools with existing single-cell spatial omics data, integrating both molecular and spatial data to build unified, higher confidence models of functional cell-cell interactions. The common thread among these tools reflects a move towards trying to understand cell fate specification from a more holistic, multicellular context, acknowledging that cells are not solely the products of their own intrinsic molecular programmes, but also respond to their surroundings and to neighbouring cells in a cell-extrinsic manner.

There are tools to infer how single-cell GRNs respond to perturbations, with clear applications for *in silico* cell reprogramming (Kamimoto et al., 2023; Jung et al., 2021; Lotfollahi et al., 2019). Autoencoders (see Glossary, Box 1) are a deep neural network framework seeing increasing utility in the field of network biology (Theodoris et al., 2023), with the compositional perturbation autoencoder (CPA) learning to predict transcriptional perturbation responses at the single-cell level *in silico* for unseen drug dosages, cell types, time points and species (Lotfollahi et al., 2023). Of

course, the question remains whether the prediction of perturbation responses on a handful of genes is sufficiently predictive of changes in cell identity. Capybara is a tool that focuses more explicitly on cell identity, exploring the continuous space of intermediate or 'hybrid' cell states, which are not necessarily captured *in vivo* but could still be functionally relevant in engineering cell fate transitions (Kong et al., 2022). In addition, there have been recent attempts at using transformer-based models (see Glossary, Box 1), such as the 'single-cell bidirectional encoder representations from transformers' (scBERT) model, which can decode and annotate both large transcriptomic and multi-omic datasets (Yang et al., 2022). As omics technologies continue to emerge, we can expect to improve our capacity to capture molecular information at cellular resolution, pertaining to the genome, epigenome, transcriptome, proteome, metabolome and molecular interactomes, among others. Importantly, computational tools will be crucial in understanding how to achieve a maximally informative but minimal descriptor (see Glossary, Box 1) of cell state, and to reduce the design space for cell reprogramming.

Conclusions

There have been significant advancements in programming human cell identity, allowing for greater manipulation and design of cell function. Yet, challenges remain, hindering the full replication of desired cell identity and function. The Company of Biologists' 2023 workshop on 'Novel Technologies for Programming Human Cell Fate' highlighted the limitations and opportunities in this field, emphasising the potential impact of recent technological breakthroughs on precisely engineering clinically valuable human cells.

There are a variety of purposes for engineering cells, from gaining a deeper developmental understanding to the clinical application of functional cell products. In each case, there will be specific limitations and outcomes, all-in-all depending on what questions we are trying to answer by programming cells.

The field of cell fate programming is rapidly advancing, there is therefore a need to achieve a unified framework to describe cell identities across diverse cell atlases. Although single-cell genomics has been transformative in understanding cell identities, there is an ongoing debate about defining cell identity solely based on gene expression profiles. It is essential to consider functional attributes and cellular responses to environmental cues. A limitation of all the technologies discussed above is the integration of approaches and datasets to uncover findings that cannot be described from a single-sided perspective. Overall, overcoming this constraint will allow us to engineer cells with desired identities and improve our ability to read and manipulate molecular information at a more holistic level. Continued advancements in multi-omics technologies and computational tools combined with the latest cell culture techniques will be crucial in shaping the future of the field.

Acknowledgements

We thank Marisa Karow, Samantha Morris and Barbara Treutlein for guidance during conceptualization and revision of the manuscript. We also thank everyone that participated in the discussions and The Company of Biologists for making this workshop possible.

Competing interests

The authors declare no competing or financial interests.

Funding

H.-C.L. is supported by the Human Frontier Science Program (LT000399/2020-L). A.M. is supported by a Laboratory of Molecular Biology Cambridge Scholarship funded by the Cambridge Trust. C.V.E. is supported by a H2020 Marie Skłodowska-Curie Actions Postdoctoral Fellowship (REA grant no. 101067501). D.Z. is financed by the European Research Council (grant 788381) and the Deutsche Forschungsgemeinschaft (Transregio Research Unit 267). F.S. is supported by a British Heart Foundation Intermediate Basic Science Research Fellowship (FS/IBSRF/21/25088). Open Access funding provided by University of Oxford. Deposited in PMC for immediate release.

References

Adler, B. A., Hessler, T., Cress, B. F., Lahiri, A., Mutalik, V. K., Barrangou, R., Banfield, J. and Doudna, J. A. (2022). Broad-spectrum CRISPR-Cas13a enables efficient phage genome editing. *Nat. Microbiol.* **7**, 1967-1979. doi:10.1038/s41564-022-01258-x

Ai, Z., Niu, B., Yin, Y., Xiang, L., Shi, G., Duan, K., Wang, S., Hu, Y., Zhang, C., Zhang, C. et al. (2023). Dissecting peri-implantation development using cultured human embryos and embryo-like assembloids. *Cell Res.* **33**, 661-678. doi:10.1038/s41422-023-00846-8

Aibar, S., González-Blas, C. B., Moerman, T., Huynh-Thu, V. A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.-C., Geurts, P., Aerts, J. et al. (2017). SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083-1086. doi:10.1038/nmeth.4463

Arabacı, D. H., Terzioğlu, G., Bayırbaşı, B. and Önder, T. T. (2021). Going up the hill: chromatin-based barriers to epigenetic reprogramming. *FEBS J.* **288**, 4798-4811. doi:10.1111/febs.15628

Arterbery, A. S. and Bogue, C. W. (2014). Endodermal and mesenchymal cross talk: a crossroad for the maturation of foregut organs. *Pediatr. Res.* **75**, 120-126. doi:10.1038/pr.2013.201

Avior, Y., Sagi, I. and Benvenisty, N. (2016). Pluripotent stem cells in disease modelling and drug discovery. *Nat. Rev. Mol. Cell Biol.* **17**, 170-182. doi:10.1038/nrm.2015.27

Balboa, D., Barsby, T., Lithovius, V., Saarimäi-Vire, J., Omar-Hmeadi, M., Dyachok, O., Montaser, H., Lund, P.-E., Yang, M., Ibrahim, H. et al. (2022). Functional, metabolic and transcriptional maturation of human pancreatic islets derived from stem cells. *Nat. Biotechnol.* **40**, 1042-1055. doi:10.1038/s41587-022-01219-z

Barker, R. A., Parmar, M., Studer, L. and Takahashi, J. (2017). Human trials of stem cell-derived dopamine neurons for Parkinson's disease: dawn of a new era. *Cell Stem Cell* **21**, 569-573. doi:10.1016/j.stem.2017.09.014

Barretto, N., Zhang, H., Powell, S. K., Fernando, M. B., Zhang, S., Flaherty, E. K., Ho, S.-M., Slesinger, P. A., Duan, J. and Brennand, K. J. (2020). ASCL1- and DLX2-induced GABAergic neurons from hiPSC-derived NPCs. *J. Neurosci. Methods* **334**, 108548. doi:10.1016/j.jneumeth.2019.108548

Baysoy, A., Bai, Z., Satija, R. and Fan, R. (2023). The technological landscape and applications of single-cell multi-omics. *Nat. Rev. Mol. Cell Biol.* **24**, 695-713. doi:10.1038/s41580-023-00615-w

Bhatia, H. S., Brunner, A.-D., Öztürk, F., Kapoor, S., Rong, Z., Mai, H., Thielert, M., Ali, M., Al-Maskari, R., Paetzold, J. C. et al. (2022). Spatial proteomics in three-dimensional intact specimens. *Cell* **185**, 5040-5058.e19. doi:10.1016/j.cell.2022.11.021

Blanchard, J. W., Xie, J., El-Mecharrafie, N., Gross, S., Lee, S., Lerner, R. A. and Baldwin, K. K. (2017). Replacing reprogramming factors with antibodies selected from combinatorial antibody libraries. *Nat. Biotechnol.* **35**, 960-968. doi:10.1038/nbt.3963

Browaeys, R., Saelens, W. and Saeys, Y. (2020). NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat. Methods* **17**, 159-162. doi:10.1038/s41592-019-0667-5

Brumbaugh, J., Di Stefano, B. and Hochedlinger, K. (2019). Reprogramming: identifying the mechanisms that safeguard cell identity. *Development* **146**, dev182170. doi:10.1242/dev.182170

Chen, W., Guillaume-Gentil, O., Rainer, P. Y., Gäbelein, C. G., Saelens, W., Gardeux, V., Klaeger, A., Dainese, R., Zachara, M., Zambelli, T. et al. (2022). Live-seq enables temporal transcriptomic recording of single cells. *Nature* **608**, 733-740. doi:10.1038/s41586-022-05046-9

Choi, J., Chen, W., Minkina, A., Chardon, F. M., Suiter, C. C., Regalado, S. G., Domcke, S., Hamazaki, N., Lee, C., Martin, B. et al. (2022). A time-resolved, multi-symbol molecular recorder via sequential genome editing. *Nature* **608**, 98-107. doi:10.1038/s41586-022-04922-8

Cooke, C. B., Barrington, C., Baillie-Benson, P., Nichols, J. and Moris, N. (2023). Gastruloid-derived primordial germ cell-like cells develop dynamically within integrated tissues. *Development* **150**, dev201790. doi:10.1242/dev.201790

Davidson, E. H. and Erwin, D. H. (2006). Gene regulatory networks and the evolution of animal body plans. *Science* **311**, 796-800. doi:10.1126/science.1113832

Domcke, S. and Shendure, J. (2023). A reference cell tree will serve science better than a reference cell atlas. *Cell* **186**, 1103-1114. doi:10.1016/j.cell.2023.02.016

Drakhlis, L., Biswanath, S., Farr, C.-M., Lupanow, V., Teske, J., Ritzenhoff, K., Franke, A., Manstein, F., Bolesani, E., Kempf, H. et al. (2021). Human heart-forming organoids recapitulate early heart and foregut development. *Nat. Biotechnol.* **39**, 737-746. doi:10.1038/s41587-021-00815-9

Efremova, M., Vento-Tormo, M., Teichmann, S. A. and Vento-Tormo, R. (2020). CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat. Protoc.* **15**, 1484-1506. doi:10.1038/s41596-020-0292-x

Elosegui-Artola, A., Gupta, A., Najibi, A. J., Seo, B. R., Garry, R., Tringides, C. M., de Lázaro, I., Darnell, M., Gu, W., Zhou, Q. et al. (2023). Matrix viscoelasticity controls spatiotemporal tissue organization. *Nat. Mater.* **22**, 117-127. doi:10.1038/s41563-022-01400-4

Farzadfard, F. and Lu, T. K. (2018). Emerging applications for DNA writers and molecular recorders. *Science* **361**, 870-875. doi:10.1126/science.aat9249

Fleck, J. S., Jansen, S. M. J., Wollny, D., Zenk, F., Seimiya, M., Jain, A., Okamoto, R., Santel, M., He, Z., Camp, J. G. et al. (2022). Inferring and perturbing cell fate regulomes in human brain organoids. *Nature* **621**, 365-372. doi:10.1038/s41586-022-05279-8

Francesconi, M., Di Stefano, B., Berenguer, C., de Andrés-Aguayo, L., Plana-Carmona, M., Mendez-Lago, M., Guillaumet-Adkins, A., Rodriguez-Esteban, G., Gut, M., Gut, I. G. et al. (2019). Single cell RNA-seq identifies the origins of heterogeneity in efficient cell transdifferentiation and reprogramming. *eLife* **8**, e41627. doi:10.7554/eLife.41627

Giesen, C., Wang, H. A. O., Schapiro, D., Zivanovic, N., Jacobs, A., Hattendorf, B., Schüffler, P. J., Grolimund, D., Buhmann, J. M., Brandt, S. et al. (2014). Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods* **11**, 417-422. doi:10.1038/nmeth.2869

Gut, G., Herrmann, M. D. and Pelkmans, L. (2018). Multiplexed protein maps link subcellular organization to cellular states. *Science* **361**, eaar7042. doi:10.1126/science.aar7042

Hao, Y., Stuart, T., Kowalski, M. H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C. et al. (2023). Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* doi:10.1038/s41587-023-01767-y

Hie, B. L., Shanker, V. R., Xu, D., Bruun, T. U. J., Weidenbacher, P. A., Tang, S., Wu, W., Pak, J. E. and Kim, P. S. (2023). Efficient evolution of human antibodies from general protein language models. *Nat. Biotechnol.* **57**. doi:10.1038/s41587-023-01763-2

Hislop, J., Alavi, A., Song, Q., Schoenberger, R., Keshavarz, F. K., LeGraw, R., Velazquez, J., Mokhtari, T., Taheri, M. N., Rytel, M. et al. (2023). Modelling human post-implantation development via extra-embryonic Niche engineering. *bioRxiv*. doi:10.1101/2023.06.15.545118

Horns, F., Martinez, J. A., Fan, C., Haque, M., Linton, J. M., Tobin, V., Santat, L., Maggiolo, A. O., Bjorkman, P. J., Lois, C. et al. (2023). Engineering RNA export for measurement and manipulation of living cells. *Cell* **186**, 3642-3658.e32. doi:10.1016/j.cell.2023.06.013

Hou, X., Zaks, T., Langer, R. and Dong, Y. (2021). Lipid nanoparticles for mRNA delivery. *Nat. Rev. Mater.* **6**, 1078-1094. doi:10.1038/s41578-021-00358-0

Jindal, K., Adil, M. T., Yamaguchi, N., Yang, X., Wang, H. C., Kamimoto, K., Rivera-Gonzalez, G. C. and Morris, S. A. (2023). Single-cell lineage capture across genomic modalities with CellTag-multi reveals fate-specific gene regulatory changes. *Nat. Biotechnol.* doi:10.1038/s41587-023-01931-4

Jung, S., Appleton, E., Ali, M., Church, G. M. and del Sol, A. (2021). A computer-guided design tool to increase the efficiency of cellular conversions. *Nat. Commun.* 12, 1659. doi:10.1038/s41467-021-21801-4

Kamao, H., Mandai, M., Okamoto, S., Sakai, N., Suga, A., Sugita, S., Kiryu, J. and Takahashi, M. (2014). Characterization of human induced pluripotent stem cell-derived retinal pigment epithelium cell sheets aiming for clinical application. *Stem Cell Rep.* **2**, 205-218. doi:10.1016/j.stemcr.2013.12.007

Kamimoto, K., Stringa, B., Hoffmann, C. M., Jindal, K., Solnica-Krezel, L. and Morris, S. A. (2023). Dissecting cell identity via network inference and in silico gene perturbation. *Nature* **614**, 742-751. doi:10.1038/s41586-022-05688-9

Kanemaru, K., Cranley, J., Muraro, D., Miranda, A. M. A., Ho, S. Y., Wilbrey-Clark, A., Patrick Pett, J., Polanski, K., Richardson, L., Litvinukova, M. et al. (2023). Spatially resolved multiomics of human cardiac niches. *Nature* **619**, 801-810. doi:10.1038/s41586-023-06311-1

Karow, M., Camp, J. G., Falk, S., Gerber, T., Pataskar, A., Gac-Santel, M., Kageyama, J., Brazovskaja, A., Garding, A., Fan, W. et al. (2018). Direct pericyte-to-neuron reprogramming via unfolding of a neural stem cell-like program. *Nat. Neurosci.* **21**, 932-940. doi:10.1038/s41593-018-0168-3

Kim, S. C., Sekhon, S. S., Shin, W.-R., Ahn, G., Cho, B.-K., Ahn, J.-Y. and Kim, Y.-H. (2022). Modifications of mRNA vaccine structural elements for improving mRNA stability and translation efficiency. *Mol. Cell Toxicol.* **18**, 1-8. doi:10.1007/s13273-021-00171-4

Kirkeby, A., Nolbrant, S., Tiklova, K., Heuer, A., Kee, N., Cardoso, T., Ottosson, D. R., Lelos, M. J., Rifes, P., Dunnett, S. B. et al. (2017). Predictive markers guide differentiation to improve graft outcome in clinical translation of hESC-based therapy for Parkinson's disease. *Cell Stem Cell* **20**, 135-148. doi:10.1016/j.stem.2016.09.004

Kirkeby, A., Nelander, J., Hoban, D. B., Rogelius, N., Bjartmarz, H., Storm, P., Fiorenzano, A., Adler, A. F., Vale, S., Mudannayake, J. et al. (2023). Preclinical quality, safety, and efficacy of a human embryonic stem cell-derived product for the treatment of Parkinson's disease, STEM-PD. *Cell Stem Cell* **30**, 1299-1314.e9. doi:10.1016/j.stem.2023.08.014

Klein, D., Palla, G., Lange, M., Klein, M., Piran, Z., Gander, M., Meng-Papaxanthos, L., Sterr, M., Bastidas-Ponce, A., Tarquis-Medina, M. et al. (2023). Mapping cells through time and space with moscot. *bioRxiv*, 2023.05.11.540374. doi:10.1101/2023.05.11.540374

Kogut, I., Mccarthy, S. M., Pavlova, M., Astling, D. P., Chen, X., Jakimenko, A., Jones, K. L., Getahun, A., Cambier, J. C., Pasmooij, A. M. G. et al. (2018). High-efficiency RNA-based reprogramming of human primary fibroblasts. *Nat. Commun.* **9**, 745. doi:10.1038/s41467-018-03190-3

Kong, W., Fu, Y. C., Holloway, E. M., Garipler, G., Yang, X., Mazzoni, E. O. and Morris, S. A. (2022). Capybara: A computational tool to measure cell identity and fate transitions. *Cell Stem Cell* **29**, 635-649.e11. doi:10.1016/j.stem.2022.03.001

Kroon, E., Martinson, L. A., Kadoya, K., Bang, A. G., Kelly, O. G., Eliazer, S., Young, H., Richardson, M., Smart, N. G., Cunningham, J. et al. (2008). Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. *Nat. Biotechnol.* **26**, 443-452. doi:10.1038/nbt1393

Kuett, L., Catena, R., Özcan, A., Plüss, A., Ali, H. R., Sa'd, M. A., Alon, S., Aparicio, S., Battistoni, G., Balasubramanian, S. et al. (2022). Three-dimensional imaging mass cytometry for highly multiplexed molecular and cellular mapping of tissues and the tumor microenvironment. *Nat. Cancer* **3**, 122-133. doi:10.1038/s43018-021-00301-w

La Manno, G., Gyllborg, D., Codeluppi, S., Nishimura, K., Salto, C., Zeisel, A., Borm, L. E., Stott, S. R. W., Toledo, E. M., Villaescusa, J. C. et al. (2016). Molecular diversity of midbrain development in mouse, human, and stem cells. *Cell* **167**, 566-580.e19. doi:10.1016/j.cell.2016.09.027

Lentini, C., D'orange, M., Marichal, N., Trottmann, M.-M., Vignoles, R., Foucault, L., Verrier, C., Massera, C., Raineteau, O., Conzelmann, K.-K. et al. (2021). Reprogramming reactive glia into interneurons reduces chronic seizure activity in a mouse model of mesial temporal lobe epilepsy. *Cell Stem Cell* **28**, 2104-2121.e10. doi:10.1016/j.stem.2021.09.002

Liu, A. and Wang, X. (2022). The pivotal role of chemical modifications in mRNA therapeutics. *Front. Cell Dev. Biol.* **10**, 901510. doi:10.3389/fcell.2022.901510

Loh, K. M., Ang, L. T., Zhang, J., Kumar, V., Ang, J., Auyeong, J. Q., Lee, K. L., Choo, S. H., Lim, C. Y. Y., Nichane, M. et al. (2014). Efficient endoderm induction from human pluripotent stem cells by logically directing signals controlling lineage bifurcations. *Cell Stem Cell* **14**, 237-252. doi:10.1016/j.stem.2013.12.007

Lotfollahi, M., Wolf, F. A. and Theis, F. J. (2019). scGen predicts single-cell perturbation responses. *Nat. Methods* **16**, 715-721. doi:10.1038/s41592-019-0494-8

Lotfollahi, M., Klimovskaia Susmelj, A., De Donno, C., Hetzel, L., Ji, Y., Ibarra, I. L., Srivatsan, S. R., Naghipourfar, M., Daza, R. M., Martin, B. et al. (2023). Predicting cellular responses to complex perturbations in high-throughput screens. *Mol. Syst. Biol.* **19**, e11517. doi:10.15252/msb.202211517

Ma, L., Zhu, M., Gai, J., Li, G., Chang, Q., Qiao, P., Cao, L., Chen, W., Zhang, S. and Wan, Y. (2020). Preclinical development of a novel CD47 nanobody with less toxicity and enhanced anti-cancer therapeutic potential. *J. Nanobiotechnology* **18**, 12. doi:10.1186/s12951-020-0571-2

Magnusson, J. P., Rios, A. R., Wu, L. and Qi, L. S. (2021). Enhanced Cas12a multi-gene regulation using a CRISPR array separator. *eLife* **10**, e66406. doi:10.7554/eLife.66406

Meier, A. B., Zawada, D., De Angelis, M. T., Martens, L. D., Santamaria, G., Zengerle, S., Nowak-Imialek, M., Kornherr, J., Zhang, F., Tian, Q. et al. (2023). Epicardioid single-cell genomics uncovers principles of human epicardium biology in heart development and disease. *Nat. Biotechnol.* doi:10.1038/s41587-023-01718-7

Missinato, M. A., Murphy, S., Lynott, M., Yu, M. S., Kervadec, A., Chang, Y.-L., Kannan, S., Loreti, M., Lee, C., Amatya, P. et al. (2023). Conserved transcription factors promote cell fate stability and restrict reprogramming potential in differentiated cells. *Nat. Commun.* **14**, 1709. doi:10.1038/s41467-023-37256-8

Moffitt, J. R., Hao, J., Bambah-Mukku, D., Lu, T., Dulac, C. and Zhuang, X. (2016). High-performance multiplexed fluorescence in situ hybridization in culture and tissue with matrix imprinting and clearing. *Proc. Natl. Acad. Sci. USA* **113**, 14456-14461. doi:10.1073/pnas.1617699113

Morsut, L., Roybal, K. T., Xiong, X., Gordley, R. M., Coyle, S. M., Thomson, M. and Lim, W. A. (2016). Engineering customized cell sensing and response behaviors using synthetic Notch receptors. *Cell* **164**, 780-791. doi:10.1016/j.cell.2016.01.012

Nolbrant, S., Heuer, A., Parmar, M. and Kirkeby, A. (2017). Generation of high-purity human ventral midbrain dopaminergic progenitors for in vitro maturation and intracerebral transplantation. *Nat. Protoc.* **12**, 1962-1979. doi:10.1038/nprot.2017.078

Nolet, C., Lal, A., Ilango, R., Dyer, T., Movva, R., Zedlewski, J. and Israeli, J. (2022). Accelerating single-cell genomic analysis with GPUs. *bioRxiv*, 2022.05.26.493607. doi:10.1101/2022.05.26.493607

Oh, Y. M., Lee, S. W., Kim, W. K., Chen, S., Church, V. A., Cates, K., Li, T., Zhang, B., Dolle, R. E., Dahiya, S. et al. (2022). Age-related Huntington's disease progression modeled in directly reprogrammed patient-derived striatal neurons highlights impaired autophagy. *Nat. Neurosci.* **25**, 1420-1433. doi:10.1038/s41593-022-01185-4

Oldak, B., Wildschutz, E., Bondarenko, V., Comar, M. Y., Zhao, C., Aguilera-Castrejon, A., Tarazi, S., Viukov, S., Pham, T. X. A., Ashouokhi, S. et al. (2023). Complete human day 14 post-implantation embryo models from naive ES cells. *Nature* **622**, 562-573. doi:10.1038/s41586-023-06604-5

Paunovska, K., Loughrey, D. and Dahlman, J. E. (2022). Drug delivery systems for RNA therapeutics. *Nat. Rev. Genet.* **23**, 265-280. doi:10.1038/s41576-021-00439-4

Pedroza, M., Gassaloglu, S. I., Dias, N., Zhong, L., Hou, T.-C. J., Kretzmer, H., Smith, Z. D. and Sozen, B. (2023). Self-patterning of human stem cells into post-implantation lineages. *Nature* **622**, 574-583. doi:10.1038/s41586-023-06354-4

Perez-Pinera, P., Kocak, D. D., Vockley, C. M., Adler, A. F., Kabadi, A. M., Polstein, L. R., Thakore, P. I., Glass, K. A., Ousterout, D. G., Leong, K. W. et al. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* **10**, 973-976. doi:10.1038/nmeth.2600

Pfeffer, S. and Mahamid, J. (2018). Unravelling molecular complexity in structural cell biology. *Curr. Opin. Struct. Biol.* **52**, 111-118. doi:10.1016/j.sbi.2018.08.009

Pierson Smela, M. D., Kramme, C. C., Fortuna, P. R. J., Adams, J. L., Su, R., Dong, E., Kobayashi, M., Brixi, G., Kavirayuni, V. S., Tysinger, E. et al. (2023). Directed differentiation of human iPSCs to functional ovarian granulosa-like cells via transcription factor overexpression. *eLife* **12**, e83291. doi:10.7554/eLife.83291

Rauscher, B., Heigwer, F., Breinig, M., Winter, J. and Boutros, M. (2017). GenomeCRISPR - a database for high-throughput CRISPR/Cas9 screens. *Nucleic Acids Res.* **45**, D679-D686. doi:10.1093/nar/gkw997

Reimegård, J., Tarbier, M., Danielsson, M., Schuster, J., Baskaran, S., Panagiotou, S., Dahl, N., Friedländer, M. R. and Gallant, C. J. (2021). A combined approach for single-cell mRNA and intracellular protein expression analysis. *Commun. Biol.* **4**, 624. doi:10.1038/s42003-021-02142-w

Rifes, P., Isaksson, M., Rathore, G. S., Aldrin-Kirk, P., Møller, O. K., Barzaghi, G., Lee, J., Egerod, K. L., Rausch, D. M., Parmar, M. et al. (2020). Modeling neural tube development by differentiation of human embryonic stem cells in a microfluidic WNT gradient. *Nat. Biotechnol.* **38**, 1265-1273. doi:10.1038/s41587-020-0525-0

Rizik, L., Danial, L., Habib, M., Weiss, R. and Daniel, R. (2022). Synthetic neuromorphic computing in living cells. *Nat. Commun.* **13**, 1-17. doi:10.1038/s41467-022-33288-8

Rodriques, S. G., Stickels, R. R., Goeva, A., Martin, C. A., Murray, E., Vanderburg, C. R., Welch, J., Chen, L. M., Chen, F. and Macosko, E. Z. (2019). Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463-1467. doi:10.1126/science.aaw1219

Scala, F., Kobak, D., Bernabucci, M., Bernaerts, Y., Cadwell, C. R., Castro, J. R., Hartmanis, L., Jiang, X., Laturnus, S., Miranda, E. et al. (2021). Phenotypic variation of transcriptomic cell types in mouse motor cortex. *Nature* **598**, 144-150. doi:10.1038/s41586-020-2907-3

Schmidt, F., Cherepkova, M. Y. and Platt, R. J. (2018). Transcriptional recording by CRISPR spacer acquisition from RNA. *Nature* **562**, 380-385. doi:10.1038/s41586-018-0569-1

Shin, H. and Min, D.-H. (2023). Highly efficient messenger RNA transfection of hard-to-transfect cells using carbon nanodots. *ACS Omega* **8**, 29113-29121. doi:10.1021/acsomega.3c01394

Sokka, J., Yoshihara, M., Kvist, J., Laiho, L., Warren, A., Stadelmann, C., Jouhilahti, E.-M., Kilpinen, H., Balboa, D., Katayama, S. et al. (2022). CRISPR activation enables high-fidelity reprogramming into human pluripotent stem cells. *Stem Cell Rep.* **17**, 413-426. doi:10.1016/j.stemcr.2021.12.017

Song, K., Nam, Y.-J., Luo, X., Qi, X., Tan, W., Huang, G. N., Acharya, A., Smith, C. L., Tallquist, M. D., Neilson, E. G. et al. (2012). Heart repair by reprogramming non-myocytes with cardiac transcription factors. *Nature* **485**, 599-604. doi:10.1038/nature11139

Stuart, T. and Satija, R. (2019). Integrative single-cell analysis. *Nat. Rev. Genet.* **20**, 257-272. doi:10.1038/s41576-019-0093-7

Takahashi, K. and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676. doi:10.1016/j.cell.2006.07.024

Tan, Y.-T., Ye, L., Xie, F., Beyer, A. I., Muench, M. O., Wang, J., Chen, Z., Liu, H., Chen, S.-J. and Kan, Y. W. (2018). Respecifying human iPSC-derived blood cells into highly engraftable hematopoietic stem and progenitor cells with a single factor. *Proc. Natl. Acad. Sci. USA* **115**, 2180-2185. doi:10.1073/pnas.1718446115

Tan, D. S., Chen, Y., Gao, Y., Bednarz, A., Wei, Y., Malik, V., Ho, D. H.-H., Weng, M., Ho, S. Y., Srivastava, Y. et al. (2021). Directed evolution of an enhanced POU reprogramming factor for cell fate engineering. *Mol. Biol. Evol.* **38**, 2854-2868. doi:10.1093/molbev/msab075

Tanaka, H., Hagiwara, S., Shirane, D., Yamakawa, T., Sato, Y., Matsumoto, C., Ishizaki, K., Hishinuma, M., Chida, K., Sasaki, K. et al. (2023). Ready-to-use-type lyophilized lipid nanoparticle formulation for the postencapsulation of messenger RNA. *ACS Nano* **17**, 2588-2601. doi:10.1021/acsnano.2c10501

Theodoris, C. V., Xiao, L., Chopra, A., Chaffin, M. D., Al Sayed, Z. R., Hill, M. C., Mantineo, H., Brydon, E. M., Zeng, Z., Liu, X. S. et al. (2023). Transfer learning enables predictions in network biology. *Nature* **618**, 616-624. doi:10.1038/s41586-023-06139-9

Toda, S., Blauch, L. R., Tang, S. K. Y., Morsut, L. and Lim, W. A. (2018). Programming self-organizing multicellular structures with synthetic cell-cell signaling. *Science* **361**, 156-162. doi:10.1126/science.aat0271

Treutlein, B., Lee, Q. Y., Camp, J. G., Mall, M., Koh, W., Shariati, S. A. M., Sim, S., Neff, N. F., Skotheim, J. M., Wernig, M. et al. (2016). Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. *Nature* **534**, 391-395. doi:10.1038/nature18323

Trounson, A. and Dewitt, N. D. (2016). Pluripotent stem cells progressing to the clinic. *Nat. Rev. Mol. Cell Biol.* **17**, 194-200. doi:10.1038/nrm.2016.10

Vanhorn, S. and Morris, S. A. (2021). Next-generation lineage tracing and fate mapping to interrogate development. *Dev. Cell* **56**, 7-21. doi:10.1016/j.devcel.2020.10.021

Vierbuchen, T., Ostermeier, A., Pang, Z. P., Kokubu, Y., Südhof, T. C. and Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**, 1035-1041. doi:10.1038/nature08797

Vogel, C. and Marcotte, E. M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* **13**, 227-232. doi:10.1038/nrg3185

Warren, L., Manos, P. D., Ahfeldt, T., Loh, Y.-H., Li, H., Lau, F., Ebina, W., Mandal, P. K., Smith, Z. D., Meissner, A. et al. (2010). Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* **7**, 618-630. doi:10.1016/j.stem.2010.08.012

Watson, J. L., Juergens, D., Bennett, N. R., Trippe, B. L., Yim, J., Eisenach, H. E., Ahern, W., Borst, A. J., Ragotte, R. J., Milles, L. F. et al. (2023). De novo design of protein structure and function with RFdiffusion. *Nature* **620**, 1089-1100. doi:10.1038/s41586-023-06415-8

Weatherbee, B. A. T., Gantner, C. W., Iwamoto-Stohl, L. K., Daza, R. M., Hamazaki, N., Shendure, J. and Zernicka-Goetz, M. (2023). Pluripotent stem cell-derived model of the post-implantation human embryo. *Nature* **622**, 584-593. doi:10.1038/s41586-023-06368-y

Yakubov, E., Rechavi, G., Rozenblatt, S. and Givol, D. (2010). Reprogramming of human fibroblasts to pluripotent stem cells using mRNA of four transcription factors. *Biochem. Biophys. Res. Commun.* **394**, 189-193. doi:10.1016/j.bbrc.2010.02.150

Yang, F., Wang, W., Wang, F., Fang, Y., Tang, D., Huang, J., Lu, H. and Yao, J. (2022). scBERT as a large-scale pretrained deep language model for cell type annotation of single-cell RNA-seq data. *Nat. Mach. Intell.* **4**, 852-866. doi:10.1038/s42256-022-00534-z

Yuan, G., Wang, J., Liu, Z., Chen, M., Zhu, P., Zhang, H., Hu, Z., Cui, Y., Yuan, Y. and Sha, J. (2023). Establishment of a novel non-integrated human pluripotent stem cell-based gastruloid model. *bioRxiv*. doi:10.1101/2023.06.28.546720

Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J. and Melton, D. A. (2008). In vivo reprogramming of adult pancreatic exocrine cells to β-cells. *Nature* **455**, 627-632. doi:10.1038/nature07314

Zhu, C., Preissl, S. and Ren, B. (2020). Single-cell multimodal omics: the power of many. *Nat. Methods* **17**, 11-14. doi:10.1038/s41592-019-0691-5

Zhu, R., del Rio-Salgado, J. M., Garcia-Ojalvo, J. and Elowitz, M. B. (2022). Synthetic multistability in mammalian cells. *Science* **375**, eabg9765. doi:10.1126/science.abg9765

Zimmermannova, O., Ferreira, A. G., Ascic, E., Velasco Santiago, M., Kurochkin, I., Hansen, M., Met, Ö., Caiado, I., Shapiro, I. E., Michaux, J. et al. (2023). Restoring tumor immunogenicity with dendritic cell reprogramming. *Sci. Immunol.* **8**, eadd4817. doi:10.1126/sciimmunol.add4817