CRISPR-Cas9 Knockout Screening Using Primary Human Cells in Drug Discovery

While significant progress has been made carrying out CRISPR screens in immortalised cell lines, a more physiological and clinically relevant alternative are human primary cells, such as T cells, regulatory B cells, or natural killer (NK) cells. The use of phenotypically relevant primary cells, though, is not without biological and technical challenges, especially when their intended use includes gene editing and scale-up of these specialised cell types for screening. However, overcoming these challenges and utilising new detection technologies for assay readouts herald a new era of cellular screening. Thus, functional genomic screening (FGS) with gene-edited human primary cells offer unique opportunities to accurately identify relevant drug targets and more reliably introduce validated therapeutics to the clinic.

CRISPR-Cas9 Knockout Screening to Understand Gene Function

A logical method to understand the role of a factor within a system is to remove it and assess the result. In drug discovery, this loss-of-function approach is often employed to identify new drug targets by knocking genes down or out and investigating the resulting biological characteristics or phenotype. If the phenotype is unaltered, the targeted gene likely does not contribute to the phenotype under investigation. However, if the phenotype is altered, then the gene, and its protein, might be a potential drug target. Ideally, these screens involve systematic knockdown or knockout of genes from small scale up to whole genome level and are conducted in a cellular system that resembles in vivo conditions. Harnessing innovative technologies that include gene editing, cell culture, automation, end-point acquisition, and robust analysis is no longer science fiction, but a fully established reality.

CRISPR-Cas9 Knockout Technology

Until recently, FGS was performed using RNA interference (RNAi). However, RNAi does not completely abrogate gene expression and can potentially contribute to off-target effects. Permanent silencing of genes at the DNA level with minimal off-target effects is achieved with the CRISPR-Cas system allowing knockout of gene function using both Cas9 nuclease and guide RNA (gRNA). The genomic target can be any ~20 nucleotide DNA sequence provided the sequence is unique compared to the rest of the genome and is immediately adjacent to a protospacer adjacent motif (PAM). The gRNA guides the Cas9 protein to the target sequence while the PAM serves as a binding signal. Then, the Cas9 protein and the gRNA form a complex enabling the Cas9 to cleave the DNA creating a DNA double-strand break (DSB) within the target site. To preserve the integrity of the DNA, DSBs are swiftly repaired. The most active and efficient DNA repair mechanism is the non-homologous end joining pathway that ligates the broken ends together, frequently giving rise to small nucleic acid insertions and deletions (indels) at the break site. This imperfect repair causes frame shift mutations, where the messenger RNA no longer encodes a functional protein effectively producing gene knockout.

CRISPR-Cas9 FGS

There are two screening formats: pooled screening (Figure 1) and arrayed screening (Figure 2) that often utilise CRISPR knockout (CRISPRko) technology. Pooled screening involves introducing a pool of single guide RNAs (sgRNAs) or a sgRNA library into a population of cells. This is achieved by transducing cells with lentiviral particles packaged with sgRNA and Cas9-containing vectors (one vector per particle). Expression of sgRNA and Cas9 lead to target gene knockout. As these knockouts occur in a one cell population, pooled screening is restricted to readouts that physically separate edited cells (e.g., fitness endpoint or

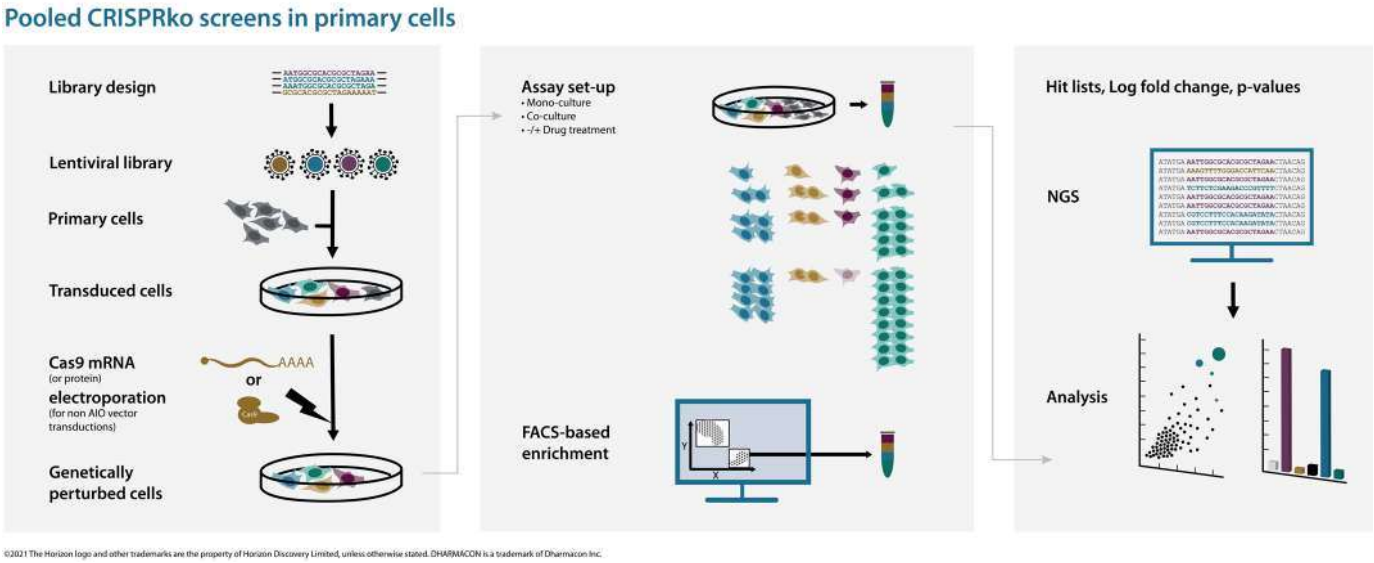


Figure 1. Workflow of pooled CRISPRko screening in primary cells. CRISPR-Cas9 editing of one cell population with either an all-in-one vector or a two-step approach (sgRNA lentiviral library, Cas9 mRNA/RNP electroporation), followed by assay set up and data analysis.

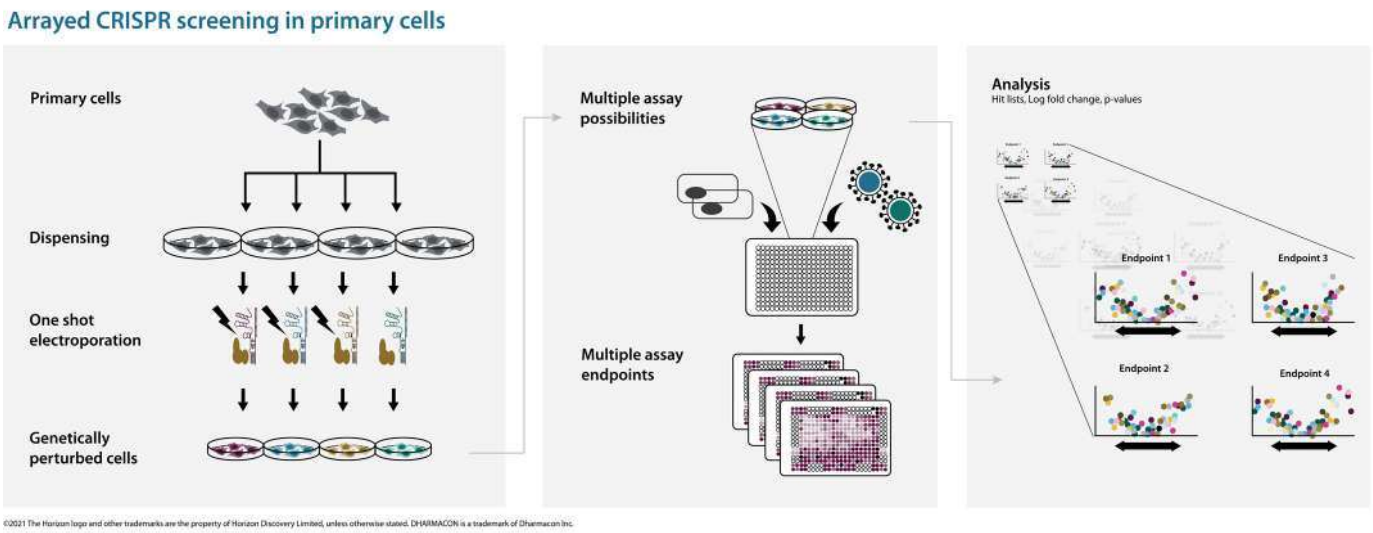


Figure 2. Workflow of arrayed CRISPRko screening in primary cells. CRISPR-Cas9 editing of cells dispensed into individual wells by electroporation, followed by assay set up (mono-cultures or multi-cultures) with multiplexed assay read-out and data analysis.

cell sorting) with a phenotype of interest from those that do not. To link a phenotype to an individual gene, the integrated sgRNAs are deep sequenced and data are deconvoluted. Recent studies have shown that CRISPR-Cas9 based screens using whole-genome sgRNA-libraries can quickly and efficiently identify genes involved in different biological processes.[1,2] The other screening paradigm is arrayed screening and involves targeting one gene per well in a multi-well plate format. Delivery of Cas9 and sgRNA is accomplished by transient transfection (e.g., electroporation) or lentiviral transduction. Because targeted genes are physically separated, unlike pooled screening, arrayed screens are compatible with a wide variety of complex assays and multiparametric endpoint readouts. The choice between pooled or arrayed format is highly dependent on the biological question, cell type and laboratory equipment.

CRISPR-knockout Screening in Primary Human Cells: Challenges and Solutions

Traditionally, immortalised cell lines have been used in FGS because they are stable, predictable, and easy to handle. However, their ability to proliferate indefinitely (due to accumulated or induced genetic mutations) may affect their original physiological properties and cause them to differentiate from the human tissue they model. Therefore, immortalised cell lines may not provide the most biologically predictive results to determine therapeutic effect. By contrast, human primary cells are more physiologically relevant and therefore, screens using such cells can identify more clinically relevant targets.

In the last few years publications on CRISPR-Cas9 editing in primary human cells, such as fibroblasts, pancreatic islet cells, megakaryocytic precursor, monocytes, NK cells, B cells and T cells have emerged (Table 1). Most of these studies are proof-of-concept experiments targeting a very limited set of genes because primary cells used for CRISPR-Cas9 FGS encounter multiple challenges (Table 2).

Primary Cells in CRISPR-Cas9 FGS: Challenges

Low cell yield due to limited biological abundance

Cell lines have been immortalised to allow continuous propagation, while primary cells derived from human tissue have a limited capacity to propagate in vitro. In addition, some primary cell populations are rarely present in body fluids and tissues. This poses an issue as the number of cells required for screening increases with the size of the sgRNA library. A sufficient cell yield to conduct whole genome screens can be achieved by sourcing the appropriate blood product (e.g., 10^9 peripheral blood mononuclear cells per leukopak), expanding/polarising cells (e.g., naïve T cells to induced regulatory T cells), pooling cells from various donors (e.g., myeloid cells) and assay miniaturisation (e.g., optimising cell numbers).

Viability issues due to limited lifespan

Primary cells are inherently more difficult to culture than immortalised cell lines. As primary cells are isolated from human tissue or blood, they are sensitive to changes in their environment and often require special nutrient medium for their survival and growth. While cell lines can be kept alive in vitro for numerous passages, the in vitro lifespan of primary cells highly depends on its cell type. Some immune cell types, such as B cells and unstimulated T cells, have a very limited lifespan, but can be kept alive and functional by using specialist cell culture consumables and reagents.

Editing efficiency due to limited susceptibility to CRISPR/Cas9 gene editing

Pooled and arrayed screening both involve the introduction of sgRNA and Cas9 into cells. In pooled screens using cell lines, sgRNA libraries and Cas9 are delivered using viral transduction; in arrayed screens using cell lines, CRISPR components are delivered in the form of DNA plasmid, lentivirus and synthetic sgRNA. However, primary immune system cells, such as T cells and NK cells, have an innate mechanism to resist foreign genetic material as they are perceived as signs of infection resulting in low editing efficiencies of T cells[3,4] and NK cells.[5]

To circumvent this mechanism, high editing efficiency is needed for a penetrating phenotype and to select for edited cells. While edited cells in pooled screening could be selected by cell sorting and antibiotic selection, this is not feasible for arrayed screening at large scale. In addition, unstimulated lymphocytes are not efficiently transduced by lentiviral vectors due to low levels of low-density lipid receptor[6] which serves as a major

entry port for lentiviral vectors in humans.[7] The requirement to stimulate lymphocytes for lentiviral transduction excludes pooled screening as an approach to study the polarisation of T cells or B cells into subsets as stimulation of these cells will drive them towards a default phenotype before editing has occurred.

Instead of using an all-in-one vector with sgRNA and Cas9 in one plasmid, researchers developed a dual system, using both lentivirus (for the sgRNA library) and electroporation (for Cas9 ribonucleoprotein (RNP) or mRNA) to overcome these challenges.[3,4] High knockout efficiencies are achieved using algorithm-designed gRNA[8] and screening libraries that employ a multi-guide design strategy, in which several sgRNAs target one gene, inducing one or more fragment deletions. Because these deletions remove several amino acids, they are highly likely to cause a complete knockout. This multi-guide strategy results in higher and more consistent knockout efficiencies compared to editing with just one sgRNA.

For electroporation, a thorough optimisation phase in which parameters such as different electroporators, electroporation buffers, CRISPR-Cas9 amounts are tested to find the best condition in terms of cell morphology, cell viability, genome editing efficiency. To monitor any effects introduced by the electroporation pulse alone, mock and non-electroporated controls are necessary.

The inclusion of appropriate controls is crucial for any screen. As negative controls, non-targeting sgRNA or sgRNAs targeting genes not associated with the downstream phenotypic readout may be advisable in case the DSB, itself, or the DNA repair pathways influence the readout. As positive controls, sgRNAs could be added that have a predictable effect on the phenotype being assayed. Other controls that provide information on editing efficiency in a binary fitness readout (live/dead) include lethal sgRNA mixes or target genes essential for cell health.

Assay variability due to heterogenous cell populations and donor-to-donor variability

Higher variability of results should be expected in screens with primary cells than those using cell lines due to intra-donor (heterogenicity of cell populations) and inter-donor variations (donor demography). Intra- and inter-donor variability can be controlled by increasing technical and biological replicates, respectively. Even though increased variability might make data interpretation challenging, it also provides an insight into the spectrum of patient responses that could be expected in clinic and should not be viewed as a major drawback.

Scale up due to restricted technology for automation, data acquisition and analysis

While the scale of a screen with cell lines is largely determined by the size of the sgRNA library and technical replicates, screens with primary human immune cells can easily double or triple in scope depending on the number of donors. Harnessing state-of-the art automation and robotics, liquid handling systems enable handling and dispensing of numerous assay plates and flasks for replicable assay set up.

Similar to assay set up, readouts should not only be customised to address the biological question at hand but also be scalable depending on the size of the gene library. Hence, there are limitations to assess knockout efficiency depending on the size of the screen. On a genomic level, analysis such as sequencing, tracking of indels by decomposition (TIDE) or T7 endonuclease I mismatch assay (T7EI) are not feasible for large gRNA library screens. On a phenotypic level, analysis such as flow cytometry, qPCR and western blotting might not reflect the genomic knockout as CRISPR guides target small regions of protein that render them non-functional. As such, unless there is direct overlap between the guide and the epitope to which detection antibodies bind, or protein transcription/translation is impacted, it is conceivable that the antibodies will bind to a protein that is otherwise non-functional. Being appreciative of these limitations for large-scale screens, it is advisable to plan for a validation screen to narrow down a gene list for thorough gene efficiency analysis.

The timing for assay set up and data acquisition is crucial for FGS with human primary cells. A fine balance between lifespan/viability and absence of target protein must be identified. While immortalised cell lines have a rapid protein turnover as they are metabolically very active due to continuous propagation, primary cells, especially unstimulated or non-dividing ones, have a very slow protein turnover. So, even when a target gene has been knocked out successfully its functional protein may persist because it is not yet degraded. For the gene knockout to have a functional effect, the protein must be absent or non-functional by the time of assay set-up.

FGS relies on readout approaches that allow qualitative or quantitative evaluation at large scale. While screens utilising cell lines are often monocultures with a binary readout deploying high-throughput-friendly methods using plate readers (*e.g.,* ATPlite) or high-content imaging, these are not always suitable options for co-cultured primary cells with a multiplexed readout. For example, to assess the effect of gene knockouts in primary T cells in response to cell stimulation various surface expressed markers, cytokines, proliferation, and mRNA expression is of interest. All could be answered by classical protein or mRNA detection techniques such as tube-based flow cytometry, ELISA or western blot. However, these traditional readout methods have their limitations when it comes to scale up. To acquire data from a screen conducted with multiple donors and a 20+ sgRNA library, technology and approaches designed for microwell sampling and miniaturisation, such as plate-based flow cytometry screener or homogenous time resolved fluorescent assays, are harnessed. Ideally, these technologies are supported by auto-fill stations and robotic arms.

The latest addition to readout approaches is single-cell RNA sequencing (RNA-seq), which allows insight into mechanisms by which each gene perturbation mediates its effect. Single-cell RNA-seq analyses transcriptomes on a cell-by-cell basis through microfluidic partitioning by capturing and barcoding single cells to obtain a next-generation sequencing (NGS) cDNA library. Although powerful, these applications of single-cell CRISPR screening have been limited by low throughput and complex analytical methods. However, continued innovations in single-cell sequencing technologies can increase the utility of single-cell CRISPR screening for scale up.

The larger the scope of the screen (number of target genes, donors, technical replicates, and readouts), the larger the volume

Primary cell type	Healthy Donors	Editing approach	Size of library	Functional assay	Readout	Authors
Lung Fibroblasts	2	crRNA:tracrRNA, Cas9 RNP electroporation	SMAD2, SMAD3, PI4K4	Scar in a Jar assay	Western blot, immunofluorescence	Martufi et al., CRISPR J 2019
Pancreatic islet cells	3	All-in-one lentivirus	PDX1, KCNJ11	Insulin and glucagon measurement, immune staining, transplantation into mice, patch clamp studies	Insulin and glucagon measurement, TIDE analysis, ddPCR, qRT-PCR, ELISA, immunostaining, Na^+, Ca^{2+} and K_{ATP} currents Insulin and glucagon measurement, TIDE analysis, ddPCR, qRT-PCR, ELISA, immunostaining, Na^+, Ca^{2+} and K_{ATP} currents	Bevacqua et al., Nat Commun 2021
Mega-karyocytic precursor (MkPs)	3	All-in-one lentivirus	60 genes highly expressed in MkPs	Colony forming unit-megakaryocyte functional assay	qRT-PCR, FACS	Zhu et al., PNAS 2018
Monocytes	3	crRNA:tracrRNA or sgRNA, Cas9 RNP electroporation	B2M, CD14, CD81	Differentiation to DC and macrophages, phagocytosis assay	FACS, ELISA, imaging, phagocytosis, Sanger sequencing	Freud et al., J Exp Med 2020
Monocytes	3	crRNA:tracrRNA, Cas9 RNP electroporation	SAMHD1, CXCR4, b2M, PPIA, AAVS1	Differentiation to DC and macrophages, in vitro infection of MDMs by Mycobacterium tuberculosis, HIV infection Differentiation to DC and macrophages, in vitro infection of MDMs by Mycobacterium tuberculosis, HIV infection	FACS, immunofluorescent imaging, Western blot, TIDE, Luminescence, RNA-seq	Hiatt et al., Cell Reports 2021
NK cells	1	crRNA:tracrRNA, Cas9 RNP electroporation	CXCR4	Transwell migration assay	Plate reader-based fluorescence-based method for quantifying cells, Flow cytometry, qRT-PCR, ddPCR, PCR, Sequencing, TIDEPlate reader-based fluorescence-based method for quantifying cells, Flow cytometry, qRT-PCR, ddPCR, PCR, Sequencing, TIDE	Lambert et al., Methods in Molecular Biology 2020
NK cells	3	crRNA:tracrRNA, Cas9 RNP electroporation, all-in-one vector	CD45, PTPRC, NCR1, CISH	Cytotoxicity and proliferation assay	FACS, Western blot	Rautela et al., JLB 2020
NK cells	3	crRNA:tracrRNA, Cas9 RNP electroporation	TGFBR2, HPRT1	Cytotoxicity assay	RT-PCR, plate reader	Kararoudi et al., J Vis Exp 2018
B cell	3	crRNA:tracrRNA, Cas9 RNP electroporation	CD46 and CDKN2A	EBV infection, B cell proliferation, plasma cell differentiation	ELISA, FACS, NGS, metabolically labeled with 4sU	Akidil et al., PLOS Pathogens 2021
B cells	3	crRNA:tracrRNA, Cas9 RNP electroporation	CCR5, PRDM1	Naïve cell to plasma cell differentiation	T7EI assay, Illumina sequencing for percentages of on-target indels, viability, ELISA, NSG mouse transplant	Hung et al., Mol Ther 2018
B cells	2	crRNA:tracrRNA, Cas9 RNP electroporation	CD19, B2M	-	FACS	Laoharawee et al., Methods Mol Biol 2020
CD4+ and CD8 + T cells	3	crRNA:tracrRNA, Cas9 RNP electroporation	Classical T cell markers	Various classical T cell assays including anti-tumour efficiency of CAR T cells, HIV infection	e.g. FACS, in vivo tumour clearance, deep sequencing	Schuhmann et al., PNAS 2015; Ren et al., Clin Cancer Res 2017; Rupp et al., Scientific Reports 2017; Hultquist et al., Nat Protoc 2019
CD3+ T cells	-	gRNA-expressing Cas9-GFP plasmid	TCR, CD52	-	IDAA, FACS	Kamali et al., BMC Biotechnology 2021
CD4+ T cells & CD34+ cells	multiple	gRNA lentiviral, Cas9 in complex with NTC gRNA	2,585 genes and 13,243 genes	Proliferation	FACS-based enrichment, NGS	Ting et al., Nature Methods 2018
CD4+ T cells	1-2	crRNA, fluorescently labelled tracrRNA, Cas9 RNP, electroporation enhancer	TRAC, ZC3H12D, TBX21	Proliferation	FACS, T7EI, intracellular cytokine staining	Leoni et al., PLOS One 2021
CD8+ T cells	2	crRNA lentivirus, Cas9 RNP	Whole genome (19,114 genes, 77,441 sgRNA) & 6 genes (2 guides per gene)	Proliferation, T cell cloning	FACS-enriched NGS analysis for whole genome library & RNA-seq for custom library of six genes	Shifrut et al., Cell 2018
CD8+ T cells	4	crRNA:tracrRNA, Cas9 RNP electroporation	12 genes identified in a WG pooled screen to impact to enhance T cell proliferation	Proliferation	FACS	Shifrut et al., Cell 2018

Table 1. Overview of publications using CRISPR-Cas9 editing in primary human cells. crRNA:tracrRNA (crRNA and tracrRNA complexed together), ddPCR (droplet digital PCR), ELISA (enzyme-linked immunosorbent assay), FACS (fluorescence-activated cell sorting), GFP (green fluorescent protein), IDAA (indel detection by amplicon analysis), NTC (non-targeting control), and qRT-PCR (quantitative reverse transcription PCR).

of data is generated. In parallel, methodological advances are driving renewed development in statistical modelling, machine learning, and artificial intelligence. Data analysis can only be as good as the data provided. Therefore, before a screen is conducted, statistical modelling must address the number of controls required for a powerful statistical analysis. Data complexity increases further by combining data in pathway analysis, for example. Another key challenge is the graphical presentation, where various software options are available to choose.

A successful CRISPR screen can be measured by the hit list of potential targets generated for a new therapeutic. Next steps require prioritising targets that might have an impact in the clinic.

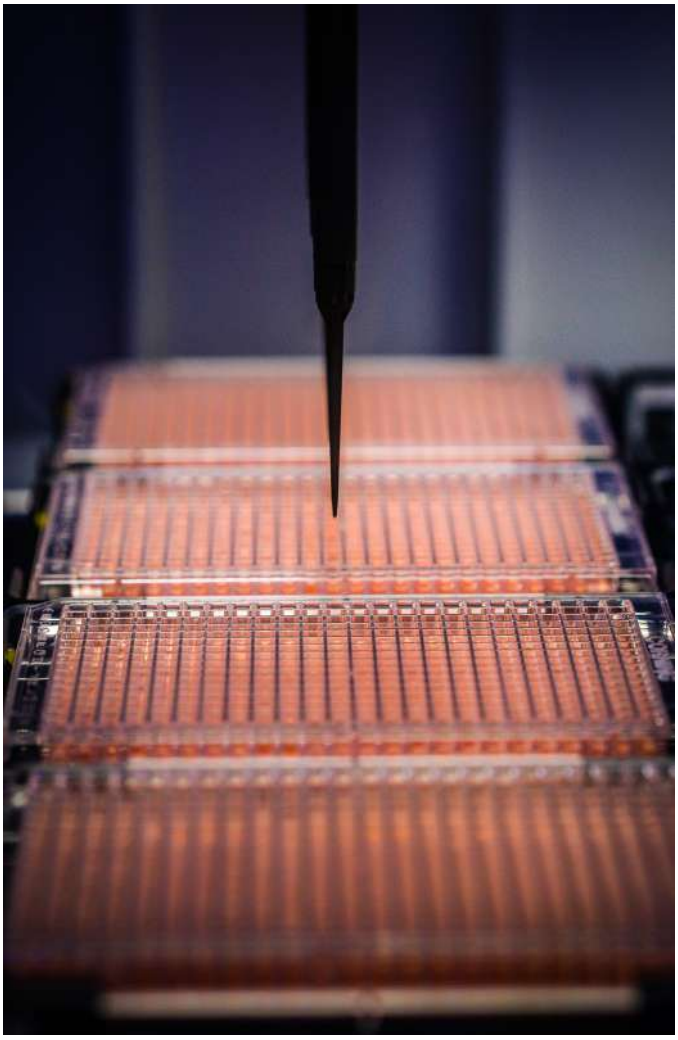
	Challenge	Solution
Cell yield	Limited capacity to propagate; Rare cell subsets	Appropriate blood product; Expansion/polarisation; Pooling donors; Assay miniaturisation
Viability	Sensitivity to change	Specialist cell culture consumables and reagents
Editing efficiency	Susceptibility to CRISPR-Cas9 reagents; Knockout efficiency	Stimulation; Choose the right screening format: Pooled vs Arrayed; Dual lentivirus/electroporation protocol; Algorithm designed; Optimising transfection/transduction conditions; Positive, negative and fitness controls
Variability	Intra-donor variability; Inter-donor variability	Technical replicates; Biological replicates
Scale up	Handling of plates and flask; Readout (editing efficiency and endpoint); Data acquisition and analysis	Automation/robotics, liquid handling systems; Validation studies; HTP-automation friendly methods; Software tools, computer power and scripting languages

Table 2. Challenges and solutions for CRISPR/Cas9-knockout screening with primary human cells

While bioinformatics and data mining help, experiments such as validation screens or compound cell panel screens are integral to move to the next stage.

Conclusion

Researchers have made significant strides towards overcoming challenges associated with editing and screening primary human cells. As shown in Table 1, many of these insights are from recent years and are limited to proof-of-concept studies. FGS pipelines are a key tool in drug discovery; by efficiently identifying specific genes that are relevant to a particular phenotype, researchers can more quickly identify drug targets that will be successful in clinical trials. However, the development of a robust, large-scale screening platform, especially for arrayed screening, is costly and currently only feasible for major pharmaceutical companies and specialist CROs. This dynamic is reflected by the emergence of only very few scientific reports sharing data – not due to lack of feasibility but rather linked to intellectual property concerns, proprietary methods, and internal policies.



REFERENCES

1. Wang, T. *et al.* Genetic Screens in Human Cells Using the CRISPR-Cas9 System. Science. 80–84 (2014).
2. Shalem, O. *et al.* Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. Science. 84–87 (2014).
3. Shifrut, E. *et al.* Genome-wide CRISPR Screens in Primary Human T Cells Reveal Key Regulators of Immune Function. Cell. 1958–1971 (2018).
4. https://www.nature.com/articles/d42473-019-00168-7, visited on 20 Jun 2021.
5. Rautela, J. *et al.* Drug target validation in primary human natural killer cells using CRISPR RNP. J. Leukoc. Biol. 1397–1408 (2020).
6. Amirache, F. *et al.* Mystery solved: VSV-G-LVs do not allow efficient gene transfer into unstimulated T cells, B cells, and HSCs because they lack the LDL receptor. Blood. 1422–1424 (2014).
7. Finkelshtein, D. *et al.* LDL receptor and its family members serve as the cellular receptors for vesicular stomatitis virus. PNAS. 7306–7311 (2013).
8. https://horizondiscovery.com/en/resources/featured-articles/top-5-reasons-to-use-algorithm-designed-guide-rna, visited on 20 Jun 2021.

Dr. Verena Brucklacher-Waldert

Verena is a Manager in the Cell-based Screening Department at Horizon Discovery, a PerkinElmer company, Cambridge, UK. She joined Horizon Discovery as Principal Scientist after supporting the Immuno-Oncology drug discovery/development pipeline in the biotechnology industry, as a Senior Scientist. A postdoctoral research fellowship from the German Research Foundation enabled her to intensify her studies in Immunology in Cambridge, UK, after receiving her doctorate in Neuroimmunology from the University of Tübingen, Germany.

Email: Verena.Brucklacher-Waldert@horizondiscovery.com