CRISPRsc: Pooled CRISPR screening with single-cell transcriptome resolution.

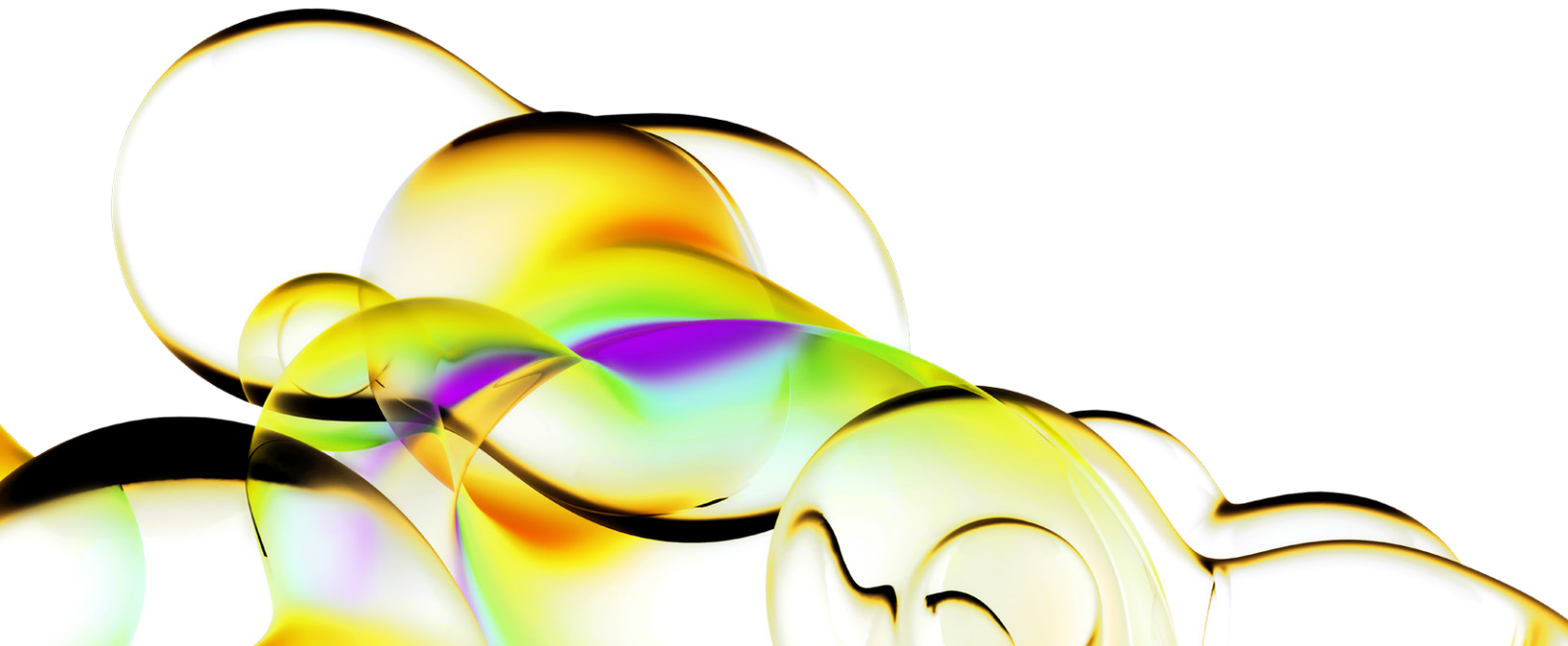
Authors

Maximilian Blanck
Matthew Perkett
John McGonigle
Simon Scrace
Carlos le Sage

Introduction

Pooled knockout screening with CRISPR-Cas9[1] has revolutionized drug discovery and has evolved to become the backbone of many drug discovery pipelines. Pooled CRISPR knockout screens offer a fast, cost effective and precise toolbox to determine drug mechanism of action and to identify genetic resistance and sensitizing factors that enable patient stratification[2,3,4]. Though powerful in its own right, pooled CRISPR screening is often restricted to the measurement of a single phenotype such as proliferation or cell survival, or simple phenotypic changes, such as, measurement of changes in activity of a single gene through the generation of a reporter cell line. Recent publications have demonstrated the feasibility of coupling pooled CRISPR screening to highly complex readouts, such as single cell gene expression profiling[5,6].

It is widely appreciated that tumour heterogeneity drives metastasis[7], invasion and chemotherapy resistance[8]. Consequently, high levels of cell heterogeneity are also observed in cancer cell lines[9] and results in altered drug resistance[10] as well as immune response to neoantigens that facilitate tumour rejection *in vivo*[11]. To better understand drug responses and the occurrence of treatment resistance it is paramount to investigate the underlying mechanisms at single cell resolution. By probing the transcriptome of each individual cell for its response to a treatment or a stimulus, we can link gene expression signature with drug response which enables the detection of biomarkers and the development of personalized medicines.



CRISPRsc (CRISPR single cell) combines the flexibility of pooled screening with the power of transcriptomics at single cell resolution. By linking a specific guide RNA to single cell whole transcriptome readout, CRISPRsc offers the opportunity to interrogate the phenotype of any given genetic perturbation in unprecedented detail. This allows identification of nuanced cellular responses that are integral to solving complex biological questions. Revvity's CRISPRsc screening platform (Fig. 1) offers a streamlined approach to elucidating valuable and intriguing biological information.

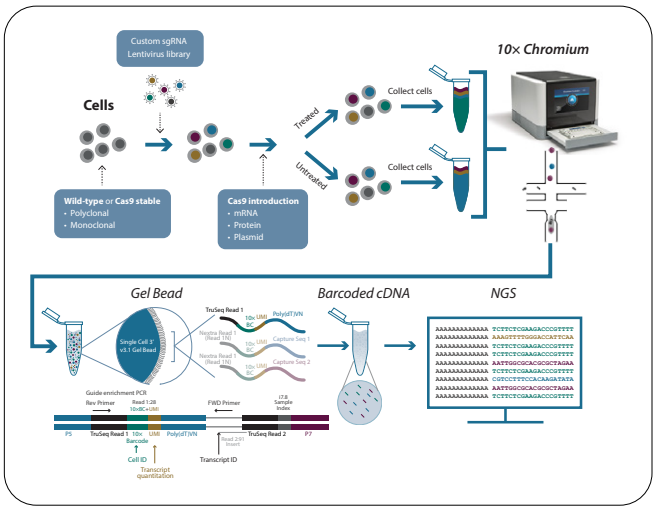


Figure 1: CRISPRsc screen setup. A clonal Cas9 expressing cell line was transduced with pooled lentivirus library. Treated and untreated cell were processed through the 10x chromium controller. Guide sequences were amplified from full length cDNA, indexed and sequenced separately from cellular transcriptomes.

Proof of concept CRISPRsc screen

To test our CRISPRsc screening platform we screened a clonal Cas9 expressing cell line with a total of 200 guide RNAs targeting 50 genes and examined the impact these guides had on drug response. Following guide transduction and antibiotic selection, cells were incubated in the presence or absence of a drug. Cells were then processed using the 10x Chromium controller that facilitates the barcoding of each cell and each transcript. After next-generation sequencing (NGS) the barcodes are used to assign every captured transcript to its cell of origin, enabling the interrogation of the whole transcriptome of each cell for its response to the drug.

The experiment was set up to capture ~10,000 treated and ~10,000 untreated cells, and captured cells were sequenced to a depth of 35,000-40,000 reads per individual cell. Guides were identified as part of the cellular transcriptome by targeted amplification[6]. Amplified guide sequences were indexed and sequenced separately from cellular transcriptomes. Guide RNA detection without separate amplification and sequencing of guide sequences is possible but reduces accurate guide assignment to less than 50% of all captured cells. By contrast, uncoupling of guide RNA detection from whole transcriptome sequencing strongly increases the accuracy and provides excellent guide detection and assignment to individual cells (Fig. 2).

Sample	Cells Found	Mean Reads Per Cell	Median Genes Per Cell	Number of reads	Q30 in RNA Reads
Untreated cells	10,815	35,563	4,093	3.85e8	89.1%
Treated cells	10,009	40,745	4,324	4.04e8	89.1%

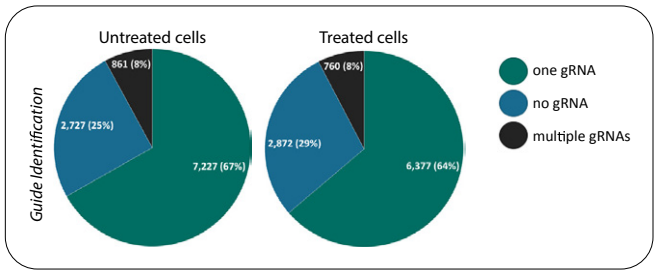


Figure 2: NGS of single cell CRISPR screen shows excellent quality metrics with high number of genes detected per individual cell and highly accurate guide RNA detection and assignment.

Screen quality was assessed by the performance of different control groups as measured by the abundance of cells expressing control guides. Control groups are composed of non-essential genes (NEG_CTRL), essential genes (POS_CTRL) and non-targeting controls that do not bind anywhere in the genome. As expected, NEG_CTRL and non-targeting guides had no effect on cell viability as shown by the normalized cell counts in untreated cells. Importantly, the presence of cells with guides targeting POS_CTRL genes was strongly reduced in the screen population as a result of decreased cell viability in the absence of essential genes (Fig. 3). Overall, control group behavior was excellent and indicated good screen performance, which is a prerequisite for high confidence hit calling.

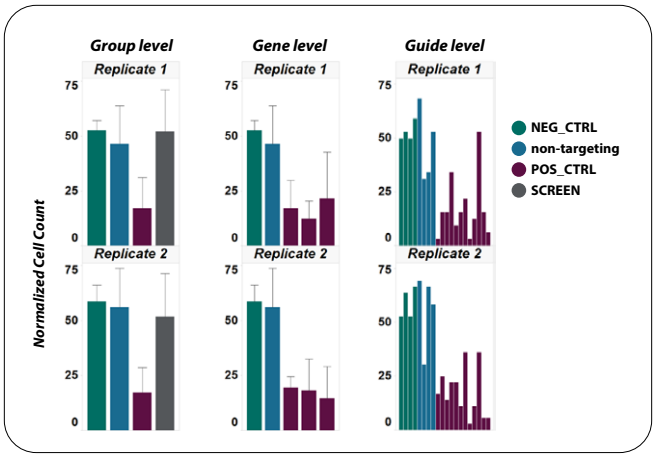


Figure 3: Control group behaviour. Excellent control group performance with reduced detection of cells that express positive control guides targeting essential genes that result in cellular demise compared to negative and non-targeting control guides that do not impact cell viability.

Hit identification

One of the great advantages of CRISPRsc screening is the possibility to analyse datasets in the more traditional pooled screening and in the single cell screening paradigms for data comparison. Pooled analysis of whole transcriptome sequencing data can identify hits that are up- or down-regulated in response to a treatment and can help to verify screen performance if the treatment is known to exert a certain response (e.g. change in gene expression) (Fig. 4).

Population level data can then be further probed to better understand the biology that governs these responses by investigating single cell behaviour that underpin this phenotype. Dimensional reduction of data was conducted by t-distributed Stochastic Neighbour Embedding (t-SNE) to cluster individual cells in line with their gene expression profile. T-SNE projection of the data showed that untreated and treated cells clustered into distinct populations (Fig. 5A). Importantly, single cell data confirmed hits from the pooled data analysis and demonstrated that *Gene A* had reduced expression levels (Fig. 5B) while *Gene B* had increased expression levels in the treated cell population (Fig. 5C). Excitingly, these data show that expression of *Gene A* and *Gene B* are highly variable within treated and untreated cell populations, which demonstrates that the cellular response to the treatment is not unidirectional but instead results in a plethora of different gene expression profiles and responses that cannot be deduced from pooled screening analyses (Fig. 4). Thus, CRISPRsc screening enables nearly limitless resolution of individual cell phenotypes to facilitate significantly greater understanding of a biological question.

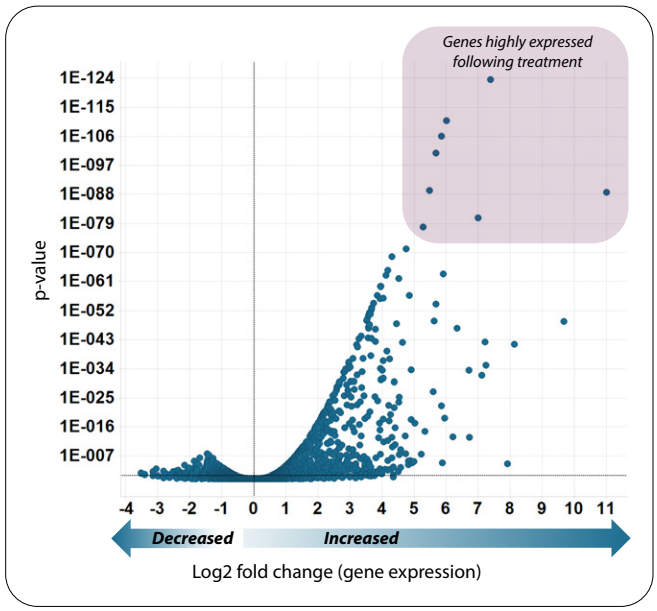


Figure 4: CRISPRsc screening pooled analysis of whole transcriptome sequencing data identifies genes that are strongly up- or down-regulated in response to the treatment on the population level.

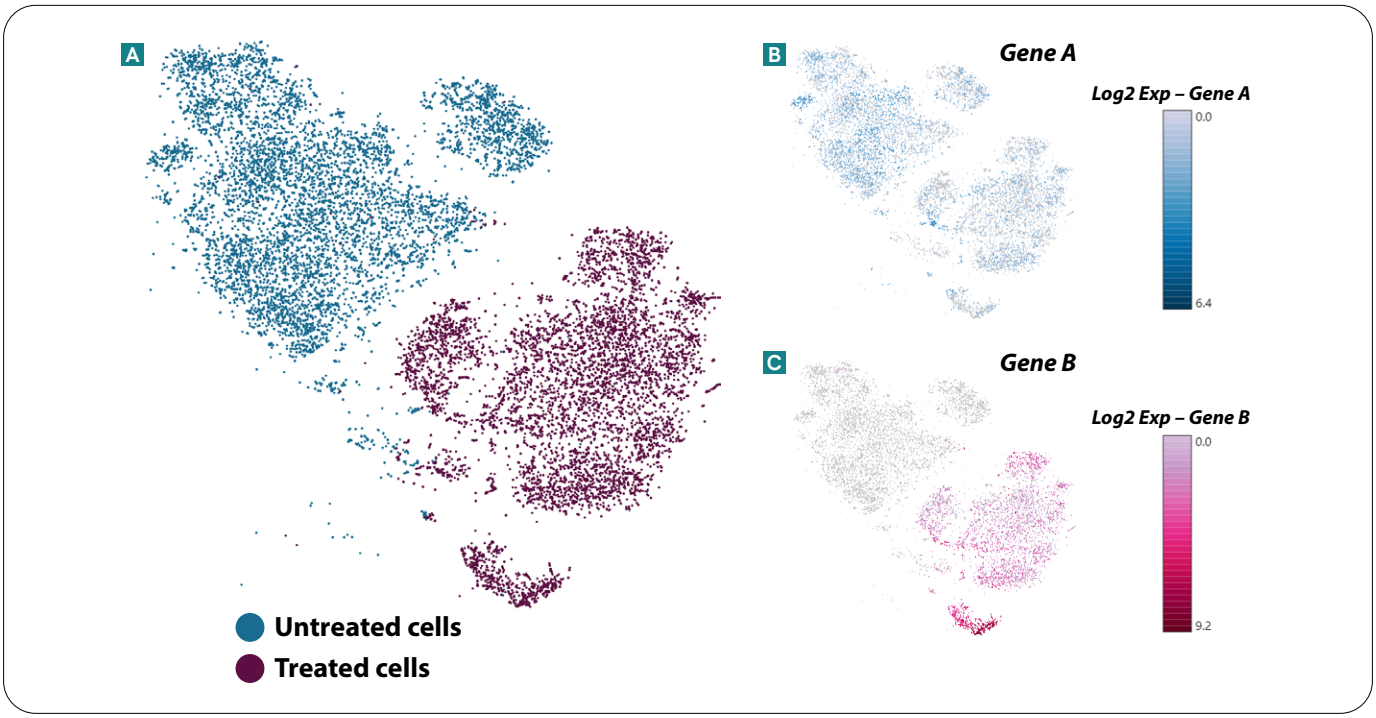


Figure 5: A) t-SNE plotting of treated and untreated cell samples enables clear distinction based on gene expression profiling B) Treatment of cells resulted in downregulation of *Gene A* C) Treatment of cells resulted in upregulation of *Gene B*.

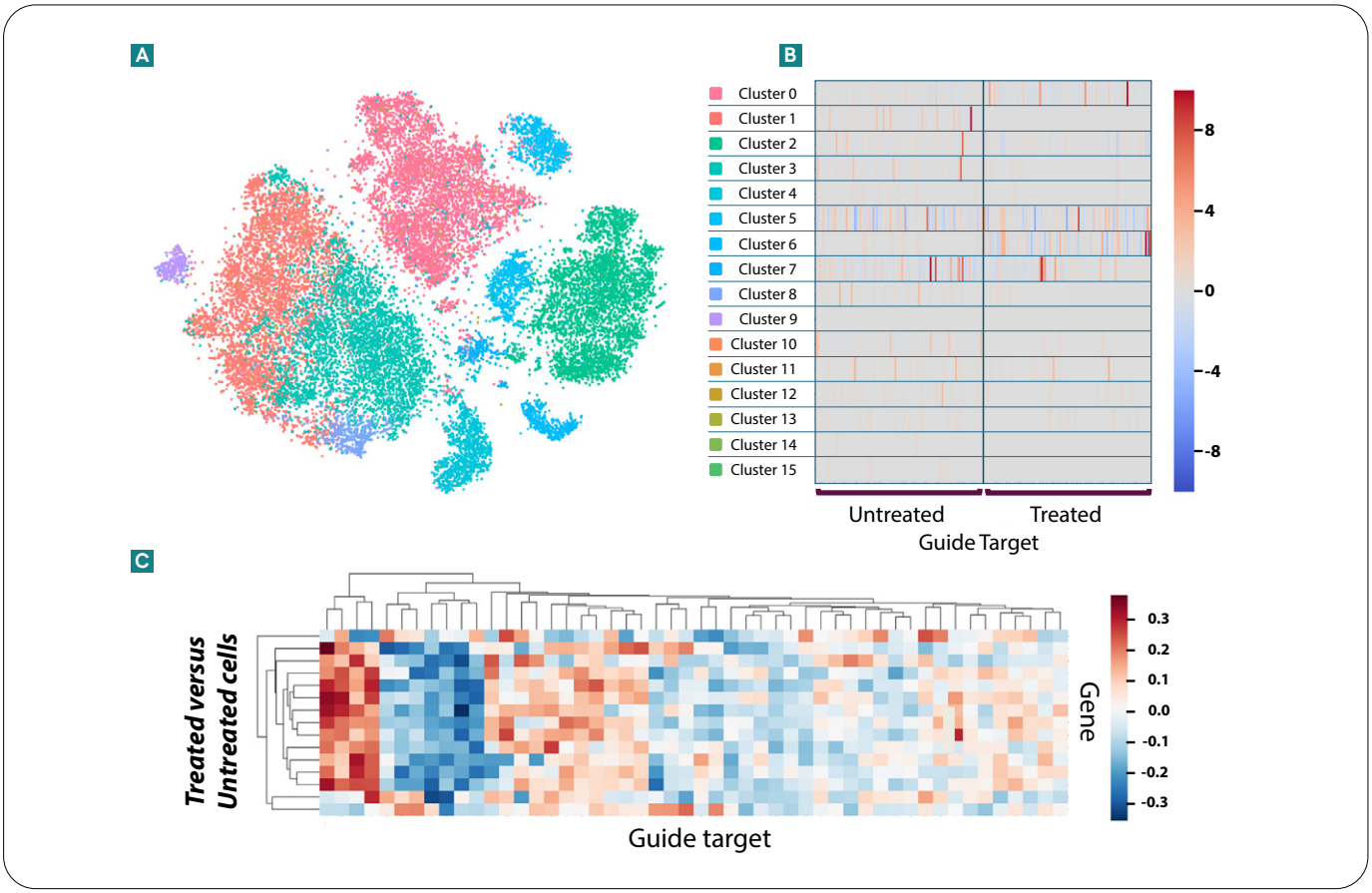


Figure 6: A) Unbiased algorithm-based t-SNE plotting of treated and untreated cell samples identified 16 individual clusters based on gene expression signature B) Chi-squared test identified clusters that deviate from expected guide distribution C) Heatmap with dendrogram pinpointing key regulators of treatment response.

CRISPRsc screening identifies pathways that affect treatment response

Gene expression profiling and unbiased algorithm-based clustering of treated and untreated cell populations identified 16 individual cell clusters (Fig 6A). This demonstrates that cancer cell lines have a dramatic intrinsic level of heterogeneity, impacting biological responses to a treatment and thus resulting in diverse outcomes. Cellular responses are further modulated by Cas9 dependent gene deletion. The impact of the genotype on the drug response driven phenotype and subsequent cluster formation, can be queried by calculating the expected guide distribution for each cluster. Chi-squared testing indicated that clusters 0, 5, 6 and 7 had the most significant deviation from the expected guide RNA distribution (Fig. 6B). Genes differentially expressed in treated cells compared to untreated cells were grouped using hierarchical clustering at the guide level to identify key regulators that drive treatment-based signalling (Fig. 6C).

Discussion

CRISPRsc screening provides a new and powerful approach that enables significantly greater understanding of highly complex biological behaviors and insights into the extent of cell population heterogeneity. However, generation of data by single cell CRISPR screening with whole transcriptome readout is heavily reliant on NGS and as such is currently only feasible as a validation screen with a targeted library. To reduce NGS burden, CRISPRsc can also be coupled with targeted amplification and sequencing of transcripts of interest[12], concomitantly allowing for increased library size at reduced cost. For primary screening campaigns we recommend Revvity's whole genome pooled CRISPR screening platform[3,4] that can be seamlessly integrated with our CRISPRsc capacity. Using a targeted 50 gene library with 200 guides, we demonstrate that treatment response is not unidirectional but evokes many different cell states that cannot be captured with pooled analysis. By integrating single cell transcriptomics and CRISPR-Cas9 driven functional genomics we unveil the genetic interactions that underlie a treatment response.

Materials and methods

Cells were transduced with a lentiviral Cas9 construct and a clonal cell line was established. Cells were transduced with a pooled guide RNA library (50 genes, 200 guides). After selection cells were treated with drug or control and 10,000 cells of each condition were loaded in duplicate onto a Chromium single cell Chip B and processed using v3 chemistry. Guides were captured as part of the transcriptome and amplified by PCR of full-length cDNA. Guide amplification PCR products were indexed and sequenced separately at a 1,000-fold coverage. Whole transcriptome sequencing of individual cells was conducted by NGS at a read depth of 50,000 reads per cell.

References

1. Wang, T. et al., *Science* 350, 1096–1101 (2015)
2. Cross, B. C. S. et al., *Sci. Rep*. 6 (2016)
3. Le Sage, C. et al., *Sci. Rep*. 7 (2017)
4. Blanck, M. et al., *CRISPR J*. 3, 211–222 (2020)
5. Datlinger, P. et al., *Nat. Met*. 14, 297–301 (2017)
6. Hill, A. et al., *Nat. Met*. 15, 271–274 (2018)
7. Kim, N. et al., *Nat. Commun*. 11, 2285 (2020)
8. Liu, D. et al., *Nat. Commun*. 8, 2193 (2017)
9. Dexter, D.L. et al., *Am. J. Med*. 71, 949–956 (1981)
10. Ben-David, U. et al, Nature 560, 325–330 (2018)
11. Wolf, Y. et al., *Cell* 179, 219–235 (2019)
12. Mercer, T. et al., *Nat. Protoc*. 9, 989–1009 (2014)

Revvity, Inc.
940 Winter Street
Waltham, MA 02451 USA

(800) 762-4000
www.revvity.com