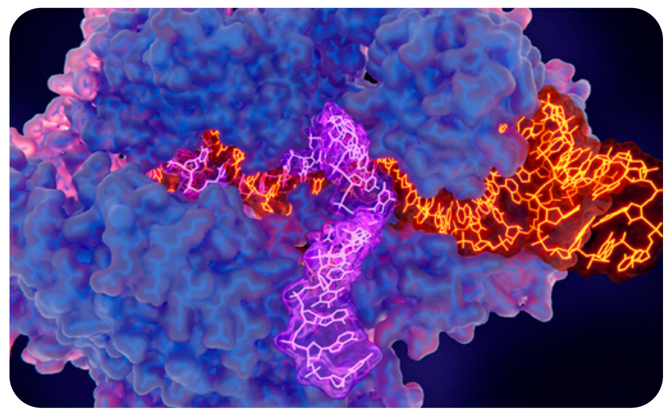
Linking genes to phenotypes with CRISPR screening

A novel approach to understand the genetic underpinnings of complex biological pathways

Introduction

The discovery of the CRISPR-Cas9 system has led to the rapid adoption of gene editing techniques within pharmaceutical and biotechnology research. Gene editing is now being used at an industrial scale to identify and validate new biological targets for precision medicines. Numerous gene editing approaches are available including systematic gene knockdown, knockout, or over-expression. These experimental approaches are collectively known as functional genomics and provide a crucial link between observed biological phenomena and the genes that influence and drive those phenomena.

While RNA interference (RNAi)-based gene silencing has been exploited as a powerful functional genomics tool over recent years, there remain challenges associated with the variability and off-target effects inherent to this technology. CRISPR-Cas9 approaches appear not to suffer from the variability evident with RNAi, and as such this technology is increasingly providing novel and compelling opportunities for drug discovery. This article explores the various approaches researchers are taking to understand the genetic underpinnings of complex biological pathways using CRISPR screening.



The evolution of CRISPR screening

Functional genomic screening with CRISPR is a powerful tool for exploring gene-regulated responses and correlating cellular phenotypes with genome profiles in high-throughput screens. The technique was first demonstrated using the RNA-guided endonuclease Cas9, which introduces a double-strand break at the chromosomal site specified by a guide RNA. Cells repair the double-strand break by the endogenous DNA repair machinery, which often involves the error-prone process of nonhomologous end joining. This can introduce insertions or deletions (indels) at the repaired locus, causing frameshifts or premature stop codons that ablate gene function.

Compared to RNAi, which can suppress gene expression but rarely eliminates it entirely, CRISPR-Cas9 can drive gene deletion to homozygosity at a high frequency, maximizing the phenotypic impact of the perturbation. There are, however, some nuances of the CRISPR knockout (CRISPRko) approach. For example, essential genes when targeted by CRISPRko are lost from the population, making the discovery of any additional phenotypic contributions of these genes challenging. This has led to the development of alternative systems such as CRISPR interference (CRISPRi) and CRISPR activation (CRISPRa), which enable the repression or overexpression of target genes, respectively. Both CRISPRi and CRISPRa rely on a nuclease-deficient version of Cas9 and can potentially identify gene targets where modulation of the gene is important for the desired phenotypic effect rather than complete loss of gene transcription.

Compared to CRISPRko, a CRISPRi approach might provide a better screening window by reducing rather than ablating gene transcription. Meanwhile, CRISPRa provides something entirely new to genomic-based screening in the form of a gain-of-function tool that drives gene expression.

CRISPR screening is typically conducted in two formats: pooled or arrayed. The decision on which screening format to use is based on a number of considerations, including the assay and cell model being used as well as the read-out and process considerations like automation, high-content data capture, and powerful biostatistical analysis.

Pooled CRISPR screens

Pooled CRISPR screens offer the opportunity to efficiently explore cell systems at the whole-genome level with high data resolution. This approach is used to address multiple biological questions, including insights into the action of drugs, toxins, and pathogens, as well as gene essentiality. Pooled CRISPR screening workflows begin with the selection of targets and the design of guides associated with those genes. Once synthesized, these guides are then cloned into a suitable library vector and introduced into cells by a lentiviral cassette. Screening assays often involve the addition of a drug (drug–gene interaction screening) or the comparative analysis of variable genotype cells (genetic interaction screening), but in all cases samples are analyzed by deep sequencing the small guide RNAs (sgRNAs) to quantitatively determine barcode/genotype abundance in each sample or condition.

A low or decreased barcode abundance indicates that the genotype marked by that barcode has been negatively selected for. An example of a negative selection screen is in the measurement of gene essentiality and the contribution of each gene in the human genome to cell fitness. Conversely, an increase in the abundance of the sgRNA barcode suggests that these genes when lost provide some ability of the cell to avoid death, known as positive selection. The most common application of positive selection is in the discovery of genes that provide resistance to a given cytotoxic agent.

The challenge with both approaches is that the datasets are not usually supported by obvious or robust expected outcomes, and the technical aspects of accurately measuring a rare or decreasingly abundant event with next-generation sequencing (NGS) renders the approach slightly noisy. Overcoming this hurdle has been a major focus of drug discovery CRISPR screeners and a number of possible adaptations have been suggested, including increasing the size of the experiment or quality of the library. Additionally, it is becoming increasingly popular to use multiple independent screening systems to augment the datasets and provide better pathway validation.

Pooled CRISPR screens in complex physiological systems

The need to investigate diseases such as cancer in an environment that recapitulates human physiology has led to the development of pooled CRISPR screens *in vivo* using either cell line-based approaches or patient-derived xenografts (PDXs). Xenograft approaches involve the transplantation of cancer cells that have been transduced with a pooled lentiviral library *ex vivo* into immunodeficient animals. The resulting primary tumor and/or metastatic sites can then be harvested and analyzed by NGS to identify genes that selectively alter the growth of the tumor, sometimes in the presence of a late-stage preclinical asset. This contrasts with direct *in vivo* screens, which are more complex and require direct injection of pooled lentiviral or adeno-associated viral (AAV) particles at the anatomical site of interest to evaluate the impact of the gene perturbation.

There is also increasing interest in functional genomic approaches in primary immune cells. However, adapting CRISPR screening tools to primary immune cells has proved challenging, in part because the introduction of a nuclease proficient Cas9 is often only feasible when using certain delivery routes. To help facilitate CRISPR screens in primary T cells, low-throughput electroporation methods have been adapted to high-throughput and high-scale experimentation and researchers are now investigating the feasibility of gene editing in other primary tissues.

Adaptations for complex disease analysis

The analysis of screen data by direct NGS quantitation can be limiting when it comes to more complex diseases. It can also be challenging to analyze phenotypes where cellular pathophysiology is uncoupled from cell health. To address this, high-throughput fluorescence-activated cell sorting (FACS) can be used to determine cell phenotypes based on a biomarker signal. This allows the monitoring of protein expression levels, post-translational modification, reporter system response, and even complex protein–protein interactions. FACS-linked screening has also been used to monitor genes impacting cell death or cell death pathways.

Another strategy is to combine CRISPR-based pooled genetic screening with single-cell RNA sequencing (scRNA-seq) to increase the phenotypic throughput of CRISPR screening. Examples include perturb-seq, CRISPseq, and CROP-seq, which are all reverse genetic methods that directly link CRISPR-driven genetic perturbations to gene expression phenotypes at the single-cell level. By providing rich transcriptome-derived datasets, these technologies are particularly powerful in dissecting complex signaling pathways that cannot be studied or are difficult to study through the behavior of an individual marker.

Arrayed phenotypic CRISPR screening

Functional genomic screening in arrayed format enables the direct exploration of phenotypic readouts that are less compatible with pooled screening approaches. Arrayed library screening involves targeting one gene per well across a multiwell plate format. Using this approach, it is possible to explore phenotypes that arise from a vast number of perturbations in parallel. Arrayed screens are also amenable to high-content imaging (HCI) and more sophisticated assays. Using HCI, a wide range of cellular phenotypes can be visualized, including morphological features, protein trafficking, and post-translational protein modifications such as protein phosphorylation.

Arrayed screens are performed using cells grown in 96- or 384-well microwell plates. Library delivery can be achieved through various transfection methods, for example lipid transfection or nucleofection. Once the cells have been perturbed, they are stained either at low or high complexity, and image-based readouts are acquired using automated microscopy. The imaging data is then analyzed to explore any novel relationships between phenotypes and the perturbagens that induce them. Due to the size and complexity of the data generated from arrayed screens, high-performance computing and informatics-based systems are often required.

Dual CRISPR screening

One final tool in the armory of the CRISPR screener is the dual screening approach. This technique uses parallel loss-and gain-of-function analyses, for example CRISPRi and CRISPRa, and can identify genetic interactions through paired gene perturbation analysis. By combining screening platforms, the quality and value of data derived from the screening campaigns is significantly increased. Having both directions of perturbation for each gene also allows for novel hit identification and provides insights that would not be accessible from a single technology. Dual-direction screening is therefore becoming an industry standard for CRISPR-based functional genomic screening.

Conclusions and future outlook

CRISPR has proved to be a powerful tool for functional genomics and the implementation of new adaptations and innovations is continuing to add value to the screening researcher's toolbox. Combining CRISPR screening platforms improves the robustness of screening datasets, and the impact of these tools for drug discovery is only just starting to be understood. In an age of personalized medicine, functional genomics will play a fundamental role in discovery research, with creativity, expertise, and pragmatism being critical if screening tools are to yield benefits for patients.

This article is an abridged version of 'CRISPR: A Screener's Guide' which can be viewed here.

Revvity, Inc.
940 Winter Street
Waltham, MA 02451 USA
www.revvity.com