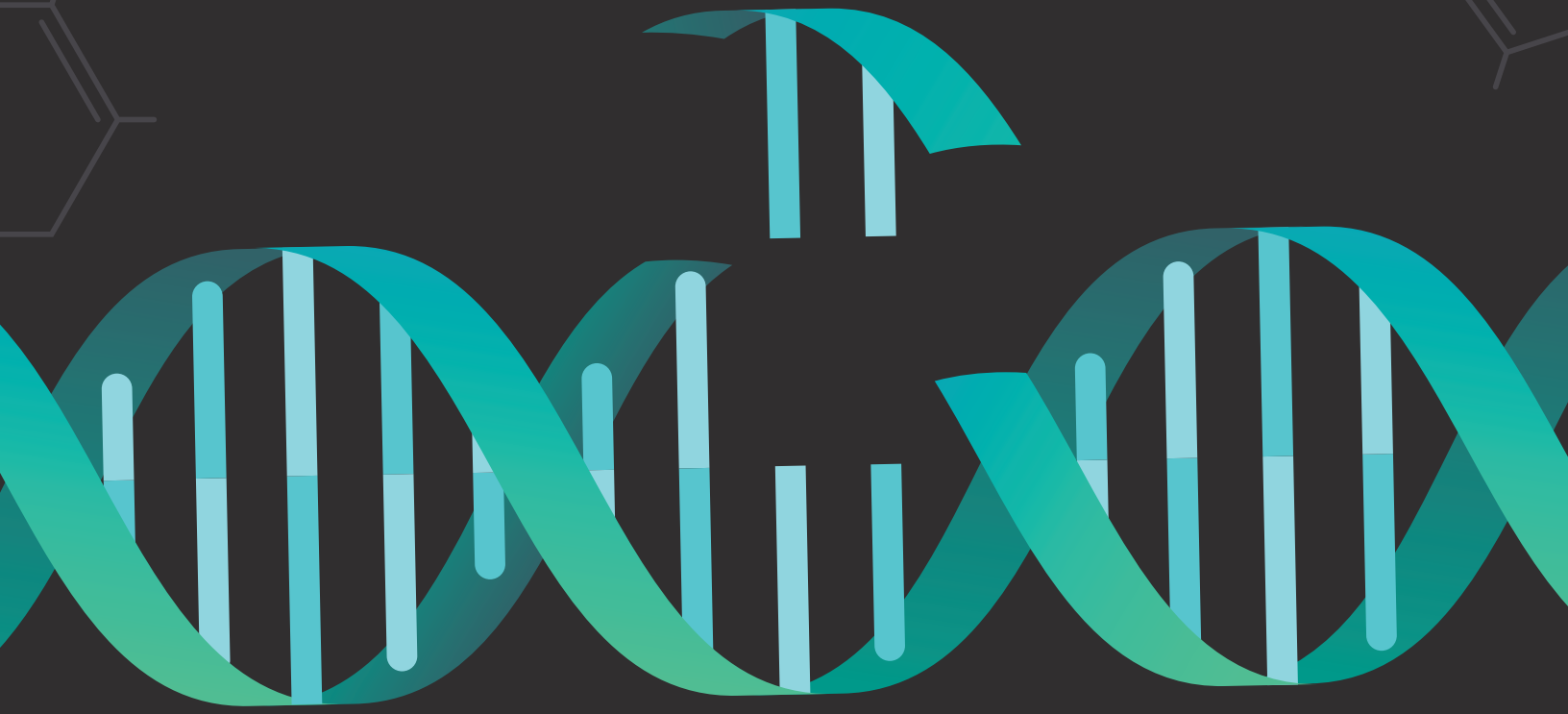
The Changing Landscape of CRISPR Screening

A guidebook to the latest CRISPR screening methodologies and technologies

A SPONSORED PUBLICATION FROM

GEN Genetic Engineering & Biotechnology News

TWIST BIOSCIENCE



Experience Next-Level CRISPR Performance

BE CONFIDENT IN YOUR RESULTS

Highly uniform Oligo Pools eliminate over- and under-represented guides.

High signal to noise enables confident hit calling.

Precise synthesis ensures maximal sequence representation.

Pools of any size to meet your needs from whole genome to highly focused screens.
Expand your idea of what's possible, design without limits!

LEARN MORE AT TWISTBIOSCIENCE.COM

TWIST BIOSCIENCE

The Changing Landscape of CRISPR Screening

A guidebook to the latest CRISPR screening methodologies and technologies

Contents

Cover: Getty Images / AliseFox

Introduction

We find ourselves at an inflection point in functional genomics, one characterized by precision, expanding functionality, and high-throughput discovery. With this new era comes a need for high performance tools and the expansion of resources that help researchers navigate the changing technological landscape.

This eBook offers a contemporary synthesis of CRISPR screening technologies for functional geneticists. In the chapters that follow, we'll dive deeper into some of the most recent advances in the field, we'll see how the technologies are being used today, and we'll highlight CRISPR experiment design tips from leading experts in the field to help scientists see success with the next generation of functional genomic screens.

CRISPR Catalyzing a Revolution

In the nine months spanning August 2012 to May 2013, a series of papers published by multiple labs revealed the genome editing potential of a little-known bacterial immune system called Clustered Regularly Interspersed Short Palindromic Repeats (CRISPR). In less than a decade, these landmark studies have already inspired the publication of more than 25,000 papers, the granting of the 2020 Nobel Prize in Chemistry, and innumerable technological innovations. Even so, the transformative impact of CRISPR has yet to be fully realized.

CRISPR provides a simple and customizable approach to genome editing wherein a CRISPR-associated (Cas) nuclease induces DNA strand breaks at precise locations within the genome. The remarkable precision of CRISPR is gained through the use of single-guide RNA (sgRNA) molecules that hold complementarity to the genomic target. Imperfect repair of Cas-induced breaks leads to frameshift mutations and loss of gene function.

The simplicity and accuracy of CRISPR means the technology has quickly overwhelmed previous techniques used for genomic manipulation, such as RNA interference (RNAi), which relied on the destruction of target gene transcripts mediated by the introduction of short interfering RNA.

Functional genomic screens using CRISPR allow researchers to methodically perturb gene function in a high-throughput, large-scale fashion. Such studies can reveal previously unknown gene-phenotype

> "One way to think about it is that we picked up the ball where RNAi screening technology left off. The difference is that when you knock down a gene with RNAi, sometimes you get it right, but with CRISPR you're almost always right."
>
> — John Doench, PhD,
> director of research and development
> Genetic Perturbation Platform of the
> Broad Institute of MIT and Harvard

associations, highlight potential therapeutic targets, and elucidate the complex genetic interactions that underlie basic biology.

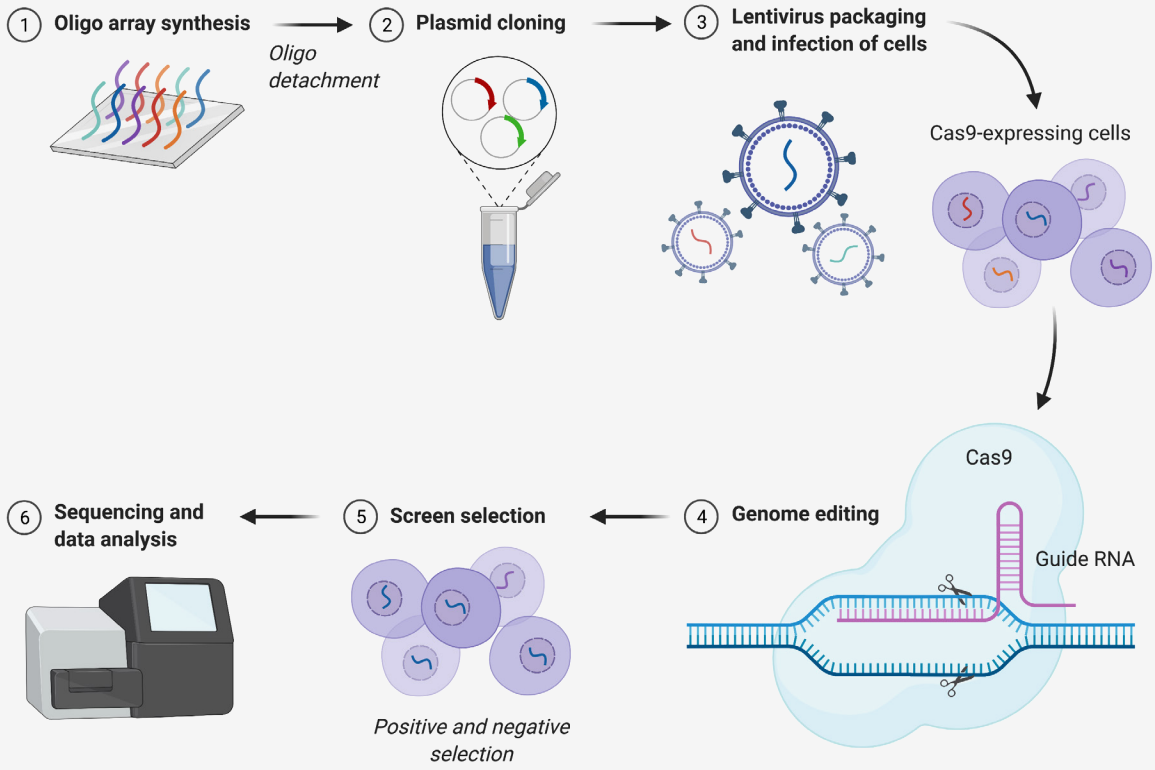
In its initial design, CRISPR provided a powerful tool for these purposes. However, recent innovations have greatly expanded CRISPR's functionality and discovery power. As a result, CRISPR screening is evolving and developing faster now than it ever has before. Early screening efforts cast a broad net, using libraries of guide RNAs that targeted genes spanning the genome. Most of the perturbations generated might have little relevance to the researchers' interests and required low-throughput, resource-intensive followup studies. Now, the increased availability of highly customized sgRNA libraries enables highly focused CRISPR screens that can drill down into a particular pathway or set of targets.

For instance, a genome-wide association study may uncover a number of genes that impact a disease or phenotype of interest. Creating a custom sgRNA library targeting just those genes enables a far more efficient and informative set of experiments that would not be feasible for a genome-wide screen, such as single-cell RNA sequencing or spatial transcriptomics. The targeted focus of these custom libraries allows for more in-depth study of possible phenotypes and interactions between the genes of interest.

"Having the ability to synthesize whatever [guide RNAs] you want on demand, rapidly and inexpensively, means that you don't need to rely on what is already available," says Professor John Doench, director of research and development in the Genetic Perturbation Platform of the Broad Institute and Harvard. "Maybe you're working in primary cells, so you can't screen all 20,000 genes, and you certainly can't screen 20,000 genes squared to study a combinatorial phenotype."

Combinatorial phenotypes are increasingly relevant as researchers investigate complex diseases.

Example CRISPR Screen workflow



Adapted from "CRISPR Screening Protocol", by BioRender.com (2021). Retrieved from https://app.biorender.com/biorender-templates.

Shown here is an example workflow for CRISPR screening. The first step in any screen is to design the sgRNA library and sequencing strategy. Once designed, oligonucleotide synthesis platforms can be leveraged to generate sgRNA pools. [*Considerations for sgRNA design and generation are provided in Chapter 3.*] Oligo pools are then cloned and transduced into experimental cell lines (some researchers choose to transduce vectors containing just sgRNA templates into cell lines stably expressing Cas9). Following transduction with the sgRNA library, editing takes place and cells are often allowed to clonally expand in the intervening time. DNA sequencing is then used to determine which sgRNAs are present and which have dropped out.

The search for anti-cancer drugs, for instance, has recently coalesced around the idea of synthetic lethality, or cell death, caused by the disruption of two genes at the same time. Increasingly, researchers seek combination therapies that kill cancer cells with lower overall toxicity. To find suitable targets, CRISPR screens that perturb genes in pairs, rather than individually, are proving invaluable. Pairwise knockouts can reveal genetic targets that don't especially harm the cancer cell when drugged separately, but deliver

a fatal blow when both are targeted. Dual-sgRNA libraries make these combinatorial screens more accessible than ever before by delivering two sgRNAs in a single vector, allowing researchers to test precise combinations of gene knockouts on a large-scale.

Beyond supporting high throughput combinatorial knockout screens, dual-sgRNA libraries expand the types of genetic elements that are available for CRISPR targeting. As researchers have begun to recognize the biological relevance of long noncoding RNAs (lncRNAs), the demand for a scalable technique for screening them has increased. While CRISPR with single sgRNAs can easily disrupt protein coding regions by creating a frameshift, lncRNAs don't have open reading frames, and a small insertion or deletion wouldn't necessarily be catastrophic. In contrast, dual-guide vectors can be designed to introduce large deletions by targeting two sgRNAs to adjacent loci, increasing the likelihood of a loss-of-function mutation in both protein coding and non-protein coding genes.

New techniques are also allowing investigation of a broader range of phenotypes using CRISPR screens. First-generation screens were built around the idea of perturbing a single gene in each cell, then testing these cells for survival or growth under specific conditions, such as in the presence of a drug or toxin. But sgRNAs can have myriad effects on a cell that don't affect its viability, such as altering gene expression and cell identity. Drawing out what these effects are can help uncover nuanced gene functions and build a better understanding of genetic interactions.

To this end, several methods have been developed that combine CRISPR screening with single-cell RNA sequencing technology.

A method called Perturb-seq combines high-throughput single-cell RNA sequencing (RNA-seq) with barcoded sgRNAs. This allows for in-depth interrogation of complex biological pathways by linking specific genetic perturbations with detailed transcriptome readouts, rather than relying on simple selectable markers. A recent evolution of this method, known as direct-capture Perturb-seq, leverages an sgRNA target enrichment strategy to improve sgRNA sequencing sensitivity. With better sensitivity and the ability to deliver predefined pairs of sgRNAs to individual cells, direct capture Perturb-seq opens the door to new possibilities using combinatorial or dual guide libraries.

"We have made tremendous leaps in understanding the genome by doing CRISPR-based perturbation screens," says Seth Shipman, PhD, synthetic biology researcher at UCSF. "But those are all elimination of genes or changing of the abundance of different proteins and RNAs. The next era of functional genomics is going to go one layer more nuanced into making modifications to the nucleotide sequence at the individual base level."

Several technologies have been developed that allow researchers to make precise, single-base

modifications throughout the genome to understand the influence of specific polymorphisms on a phenotype. Enzymes that are capable of converting a single DNA base to another can be fused to Cas9 and directed to precise locations throughout the genome. The effects of resulting single base edits can be measured and maps of mutational burden can be built.

Prime editing offers another approach to single base editing, combining the precision of Cas9 with targeted homology directed repair. A reverse transcriptase is guided to precise genomic regions alongside a template sequence, allowing for the precise incorporation of up to around 30 base pairs of new genetic information. The homology-based nature of prime editing offers increased flexibility over base editors and is increasingly being deployed in large-scale screens to improve the efficacy of CRISPR based gene therapy prototypes.

Although base editors and prime editors are both cutting-edge technologies, Cas9 fusion proteins have been a mainstay in the CRISPR toolbox since its inception. Other modern Cas9 fusions expand the capabilities of functional genomic screens to genetic control. CRISPRi and CRISPRa guide proteins that influence gene activation or repression to specific regulatory elements in the genome. CRISPRon and CRISPRoff guide methyltransferases for precise epigenetic control over gene expression. When paired with large-scale custom sgRNA library synthesis and improved guide design algorithms, these new CRISPR screening tools are helping researchers gain an unprecedented window into the genome's inner complexities through high throughput functional genomic screens.

A conversation with the experts:

Designing sgRNA Libraries for High-throughput CRISPR Screening

Even for seasoned researchers, designing an sgRNA library can be a daunting task. Each guide in a library of thousands requires careful consideration to ensure fidelity for its target gene. And, accuracy doesn't always equate to effectiveness—approximately 20% of CRISPR mutagenesis leads to in-frame mutations, many of which are unlikely to affect protein function[1].

In the decade since CRISPR was first applied to gene editing in mammalian cells, guidelines for designing sgRNA libraries have continually evolved in response to new insights and new technologies. With this evolution has been a growing need for custom CRISPR sgRNA libraries, tailored to maximize efficiency while taking full advantage of CRISPR screening's discovery power.

JULIAN JUDE, PH.D., one of Twist Bioscience's CRISPR experts and co-author of the Vienna Bioactivity CRISPR (VBC) score algorithm for scoring sgRNA design[1], has been helping researchers improve CRISPR screening for the better part of a decade. In the following interview, Twist Bioscience's Chief Technology Officer, **SIYUAN CHEN, PH.D.,** joins Julian to discuss some of the lessons they've learned about designing sgRNA libraries for CRISPR screens, and provide some simple tips for how researchers might improve the quality of their CRISPR screens.

Where should researchers start when designing a sgRNA library?

SIYUAN: Designing the best sgRNA library for your experiment will depend in part on your screening goals. The perfect sgRNA for a CRISPRi screen will be different from the perfect sgRNA for a classic knockout screen, and should be inserted into a backbone that is designed to match your method of delivery. Therefore it's important to start from a position that builds on generalized rules, but ultimately considers your specific needs.

JULIAN: As you said, a good sgRNA library is carefully designed to address specific needs. But there are some generalities that should be considered. Many excellent resources are available to help interested researchers catch up on the base rules for sgRNA design. Briefly, sgRNAs should have minimal complementarity with non-target DNA sequences, have between 40% and 80% GC content, and the spacer RNA portion of the sgRNA should extend approximately 17 to 24 nucleotides in length (the exact length will depend on the associated Cas protein you choose to use).

Identifying the appropriate CRISPR-associated (Cas) protein to use is a critical step. At this stage, there are many different types of nucleases that have been developed for genomic screens. To pick the one that's right for you requires an analysis of available PAM sequences near your target genes and consideration of what type of perturbation you need the Cas protein to perform. If you need guidance on this, I'd suggest reviewing John Doench's 2019 review of CRISPR technologies[2], including a nice review of the various Cas proteins, their functions, and their associated PAM sequences.

What are the latest and greatest tools available for sgRNA design?

JULIAN: If your goal is to perform a knockout screen, there are ways to optimize sgRNAs to increase the likelihood of achieving on-target knockout or, at least, to induce loss-of-function in-frame mutations. I was recently part of a team of researchers from the Vienna BioCenter, Austria, developed Vienna Bioactivity CRISPR score (VBC), a high-performance sgRNA prediction tool that helps researchers select sgRNAs to reliably generate loss-of-function alleles in mammalian cells[1]. In developing this tool, we assessed the qualities of sgRNAs that were most likely to lead to successful loss-of-function mutations. Our analysis revealed that sgRNAs designed to target stretches of highly conserved amino acids were more likely to result in loss of protein function—with the best results observed for stretches of seven amino acids (21 base pairs). Similarly, the sgRNAs that targeted DNA sequences corresponding to hydrophobic domains

General considerations for designing optimal sgRNAs for CRISPR experiments*

GENERAL CONSIDERATIONS

- GC content between 40% and 80% to ensure strong binding between sgRNA and the target DNA.
- Ensure that secondary structures like hairpins and polymerase termination sequences are limited in your sgRNA design. These structures can affect cloning efficiency and the transcription of guides.
- Remove polyA sites, these can disrupt packaging if you're using viral delivery.
- Minimize mismatches, particularly within the first 10 bases upstream of the PAM. Mismatches can be tolerated in the latter half of the spacer sequence, but mismatches nearer to the PAM site can disrupt binding and subsequent editing.
- Match the spacer sequence length in the sgRNA to the Cas protein being used. Discordance between these two can severely reduce editing efficiency.
- Research and select the best Cas protein for your needs. Cas proteins differ in the PAM sequences they recognize, the type of cut they perform, and multiple other factors that affect editing performance.
- Ensure that sgRNA promoters aren't placed in overlapping conflict with other promoters, such as the EF-1a promoter that's frequently used to transcribe selection resistance genes. Overlap with this promoter has resulted in less efficient transcription of the sgRNA.

GENE KNOCKOUT SPECIFIC CONSIDERATIONS**

- Target guides to conserved amino acid sequences (ideally 7aa in length). It's likely that these sequences have been conserved because they are important to protein function, therefore disruption in these regions is more likely to be effective.
- Target hydrophobic domains in the protein's core. It's not clear why editing core hydrophobic domains is more likely to disrupt protein function, but experimental evidence indicates that this is an effective targeting approach.
- Target larger exons to avoid alternative splicing as this can circumvent gene knockout despite successful editing.
- Don't target regions near the end of the protein-coding sequence because some proteins may still be expressed and functional despite slight truncations.
- Avoid targeting locations near alternative start sites. As with alternative splice sites, alternative start sites can compensate for the loss of the standard transcript to prevent gene knockout.
- Avoid locations that frequently contain polymorphisms. Even if your sgRNA is designed to perfectly match your target sequence, polymorphisms in the target sequence can lead to unexpected mismatches and reduced editing efficiency.

CRISPRI/A/ON/OFF CONSIDERATIONS

- Limited sgRNA options are available as you need to target the transcription start site/ CpG islands
- Ideal target window is +25 to +75 nts downstream of the TSS for CRISPRi sgRNAs, based on tiling arrays that show optimal disruption of gene expression within this window.
- For CRISPRa: 150–75 nucleotides upstream of the TSS provides optimal activation.
- CRISPRoff and CRISPRon sgRNAs are most effective when targeted to a 1kb window centered on the transcription start site. Methylation in this region is likely to disrupt key transcription factor binding and recruitment of polymerases.

* *Insights for this box were derived from several recent studies1-7.*

** *These insights were derived from studies focused on mammalian cells and may not hold true in other systems.*

in the protein's core were more likely to perform well.

What's highlighted in our work, and captured in the VBC score, is that sgRNA libraries are more likely to achieve knockout when we consider more than just sgRNA specificity, but rather select guides based on gene and protein structure. [*See box to the left for a list of general considerations for sgRNAs that came out of this work.*]

SIYUAN: For experiments that aim to modulate gene expression through CRISPRi, CRISPRa, CRISPRon or CRISPRoff, there are some constraints that limit the number of sgRNAs to choose from. For example, studies have shown that CRISPRi sgRNAs are most effective when targeting sequences within a 75 base pair window immediately downstream of the transcription start site. This window is broadened for CRISPRoff, such that sgRNAs targeting sequences within 500 base pairs upstream or downstream of the transcriptional start site appear most effective at transcriptional inhibition[3].

JULIAN: Also, as single-cell CRISPR screening can use knockout or epigenetic perturbation methods, these sgRNA libraries must also be designed to ensure that each sgRNA is faithfully linked to its perturbation and subsequent transcriptomic profile. Single-cell CRISPR screening has been challenging because functional sgRNAs can't be polyadenylated,

yet many RNA sequencing methods rely on capture of polyadenylated transcripts. Methods have been developed to overcome this hurdle by inserting barcodes into the 3' long terminal repeat section downstream of plasmid selection genes. And while this approach is effective, it has some limitations, such as dissociation of the barcode from the sgRNA as a result of template switching during the cloning step (multiple studies suggest this can affect as much as 50% of plasmids.)[4-6]

It's therefore very important to consider where in the backbone your sgRNA is inserted and how you'll detect it when sequencing. Methods like CROP-seq and direct capture Peturb-seq offer alternative approaches. [See chapter 7 of this eBook for an in depth analysis on the latest single cell CRISPR screening technology]

There are several tools available to help researchers design sgRNAs, including the VBC score described above. These tools have been collected in a helpful **github repository**.

There are lots of premade CRISPR sgRNA libraries available. Should researchers use these or consider custom-designed libraries?

SIYUAN: Researchers who are either wary of designing their own sgRNA libraries or are unsure of the right backbone to use for CRISPR screening have two options: use off-the-shelf libraries for a general screen, or get help to design custom sgRNA libraries.

The former may be a good option for researchers who have a limited budget and want to go with a library that others have validated. These libraries are usually ready to go, meaning there's little wait time for them to be made and deployed. However, such libraries have some significant drawbacks. The pace of advancement in the CRISPR field is such that any static sgRNA library is likely to be outdated by the time it is used. Pre-made libraries are also not malleable to a researcher's specific needs—libraries are designed to meet several potential research goals and may provide too much, or too little coverage over areas of specific interest.

Custom sgRNA libraries offer many advantages. Inherent to the concept of a custom sgRNA library is flexibility, and that flexibility can be particularly helpful for screens using some of the more recent CRISPR technology.

JULIAN: A good example of this is in the prime-editing space.

Prime editing has significant potential for precise perturbation in large-scale screening, but doing so requires libraries of really long sgRNA sequences. Prime editing sgRNA (pegRNA) is necessarily longer than a sgRNA as it includes both a sgRNA and an additional 30+ nucleotides that code for both a primer binding site and a repair template[8]. To facilitate this, pegRNA libraries must be precisely built using long

oligonucleotide synthesis. Few companies can do this well, Twist Bioscience is one of them.

Custom libraries also enable you to play around with the plasmid backbone, the tracrRNA, the placement of barcodes, and other features that may help you improve your screening, whereas premade libraries typically come with these features set in stone.

We are also seeing researchers combine pre-made libraries with custom solutions. Broad scale screens can lead to hundreds-to-thousands of perturbations that need to be followed up. One of the most efficient ways to do so uses a focused secondary screen utilizing a custom guide library designed to targeting all the genes identified in the primary genome-wide screen.

So, researchers can get a lot of use out of both custom sgRNA libraries and pre-made ones. It depends on your timelines, resources, and needs. But whatever library you choose, you want to make sure it's uniform and created with a low error rate.

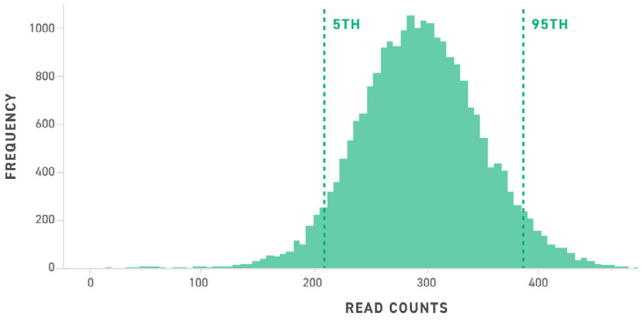
Where can researchers get uniform, low error CRISPR sgRNA libraries?

SIYUAN: Twist Bioscience. Twist's silicon-based oligonucleotide synthesis platform enables the rapid generation of low-error, highly uniform oligo-nucleotides which can then be turned into sgRNA libraries. These qualities of uniformity and low-error are extremely appealing for any CRISPR screening experiment.

JULIAN: I agree, uniformity is key. Uniformity is a description of how frequently each guide occurs relative to the rest of the library. A highly uniform library will have each guide represented equally, whereas a low uniformity library may contain an overrepresentation of a few select guides.

When screening, uniformity is important because a non-uniform library can lead to a perceived loss or diminished effect from underrepresented guides. Put another way, non-uniform libraries reduce screening sensitivity and may result in biased results.

NGS quality control data from a typical oligo pool containing 23,000 90mer oligos show the uniformity of the pool at 300x read coverage. The corresponding table indicates the uniformity metrics for this pool. Twist oligos are synthesized bias-free with high uniformity and complete oligo representation.



UNIFORMITY METRICS	
95th / 5th Percentile	1.9
99.5th / 0.5th Percentile	3.5
Oligos Represented <1/10th mean	0.03% (6)
Oligos Represented >10x mean	0% (0)

To compensate for non-uniformity, screens require a larger number of cells to ensure that the effect of underrepresented guides can be detected. Therefore uniform libraries enable successful screens with fewer cells—an important feature when working with primary cell lines that may be limited in number.

That's why it's extremely important to ensure you select a vendor that offers oligo pools with a high degree of uniformity when designing sgRNA libraries. We've written a bit more about this in our white paper on the importance of uniformity.

SIYUAN: Error rates are similarly important. CRISPR systems are designed to precisely target a sequence of DNA based on the spacer portion of the sgRNA. Mismatches in the latter 10 nucleotides of the sgRNA are tolerable to a point; however, mismatches increase the likelihood of off-target effects. Errors during sgRNA generation can thus affect sgRNA specificity. It's also worth noting that for applications that use templates for repair, such as prime editing, it's critical that the template be coded without errors.

With Twist Bioscience's oligo pools, Researchers can design custom libraries that can then be cloned into your vector of choice. Twist's CRISPR experts can also offer some guidance on the library design, vector and NGS strategy if needed.

Are there any final considerations for sgRNA design?

JULIAN: I think it's also important to call out an often overlooked aspect to sgRNA design: The need to protect your experiment from primer binding site contamination.

NGS sequencing sgRNAs is a fast and simple method for assessing which sgRNAs are present at the end of a screen. Critical to this process is having unique primer binding sites that allow you to limit amplification to just the sgRNAs in your screen, rather than any other sgRNA expressing vectors that may have snuck in. Primer binding site contamination can occur far too easily in labs that work with CRISPR often. It's much easier to amplify from a plasmid than from genomic DNA, resulting in most of your NGS reads coming from the contaminating vector.

This type of contamination can occur for many reasons. Typically labs use primer binding sites that are located within the U6 promoter that's used to promote sgRNA transcription. If, for example, you want to perform a knockout in your cell line before carrying out your screen, it's easy to reuse the same CRISPR plasmid constructs, including the same U6 promoter, and accidentally insert your primer binding site with it.

To protect your experiment, it's important to consider adding a primer binding site that is unique to your screening library. One way to do this is to insert a unique primer binding site in the screening library plasmid.

Any last words of encouragement for researchers getting into screening?

SIYUAN: Don't be afraid to ask for help and be creative. A lot of people want to just follow what's been done before. But there are creative ways to perform screens that can help improve your experiment. And, there's always experts available to help in places like Twist Bioscience. Feel free to reach out to us at twistbioscience.com/contact ■

References

1. Michlits, Georg, *et al.* "Multilayered VBC Score Predicts SgRNAs That Efficiently Generate Loss-of-Function Alleles." *Nature Methods*, vol. 17, no. 7, 8 June 2020, pp. 708–716, 10.1038/s41592-020-0850-8. Accessed 18 Aug. 2021.

2. Doench, John G. "Am I Ready for CRISPR? A User's Guide to Genetic Screens." *Nature Reviews Genetics*, vol. 19, no. 2, 4 Dec. 2017, pp. 67–80, 10.1038/nrg.2017.97. Accessed 5 Dec. 2019.

3. Nuñez, James K., *et al.* "Genome-Wide Programmable Transcriptional Memory by CRISPR-Based Epigenome Editing." *Cell*, vol. 0, no. 0, 9 Apr. 2021, www.cell.com/cell/fulltext/S0092-8674(21)00353-6, 10.1016/j.cell.2021.03.025. Accessed 10 Apr. 2021.

4. Datlinger, Paul, *et al.* "Pooled CRISPR Screening with Single-Cell Transcriptome Readout." *Nature Methods*, vol. 14, no. 3, 18 Jan. 2017, pp. 297–301, 10.1038/nmeth.4177.

5. Hill, Andrew J., *et al*. On the Design of CRISPR-Based Single Cell Molecular Screens. 29 Jan. 2018, 10.1101/254334. Accessed 3 Sept. 2021.

6. Xie, Shiqi, *et al.* "Frequent SgRNA-Barcode Recombination in Single-Cell Perturbation Assays." *PLOS ONE*, vol. 13, no. 6, 6 June 2018, p. e0198635, 10.1371/journal.pone.0198635. Accessed 3 Sept. 2021.

7. Mohr, Stephanie E., *et al.* "CRISPR Guide RNA Design for Research Applications." *The FEBS Journal*, vol. 283, no. 17, 22 June 2016, pp. 3232–3238, www.ncbi.nlm.nih.gov/pmc/articles/PMC5014588/, 10.1111/febs.13777. Accessed 6 Sept. 2019.

8. Anzalone, Andrew V *et al.* "Search-and-replace genome editing without double-strand breaks or donor DNA." *Nature* vol. 576,7785 (2019): 149-157. doi:10.1038/s41586-019-1711-4

Five Powerful CRISPR Applications for Next Generation Functional Genomic Screening

In its most basic form, knockout screening is a "break it and see what happens" approach to genetics that helps uncover genotype-phenotype relationships. Early geneticists like Herman Muller exemplified this in 1926 by blasting fruit flies with x-ray radiation, inducing measurable phenotypic effects from random mutations[1]. These early screens ultimately helped reveal the mutagenic properties of x-rays and strongly supported the idea that chromosomes carried discrete genetic units that could be modified in heritable ways.

Muller's Nobel prize-winning work foreshadowed the powerful role that functional genomic screening would come to play in modern genetics. Through the methodical perturbation of gene expression on a large scale, researchers can uncover tumor drivers and suppressors, discover actionable drug targets, and explore pathway interactions among many other applications.

The modern toolkit for genomic screening is sophisticatedand powerful, in large part due to the rapid evolution of CRISPR technology. Far from simply knocking out genes at random, CRISPR enables targeted gene knockout, gene overexpression and repression, defined mutations, and hyper- and hypomethylation with minimal off-target effects.

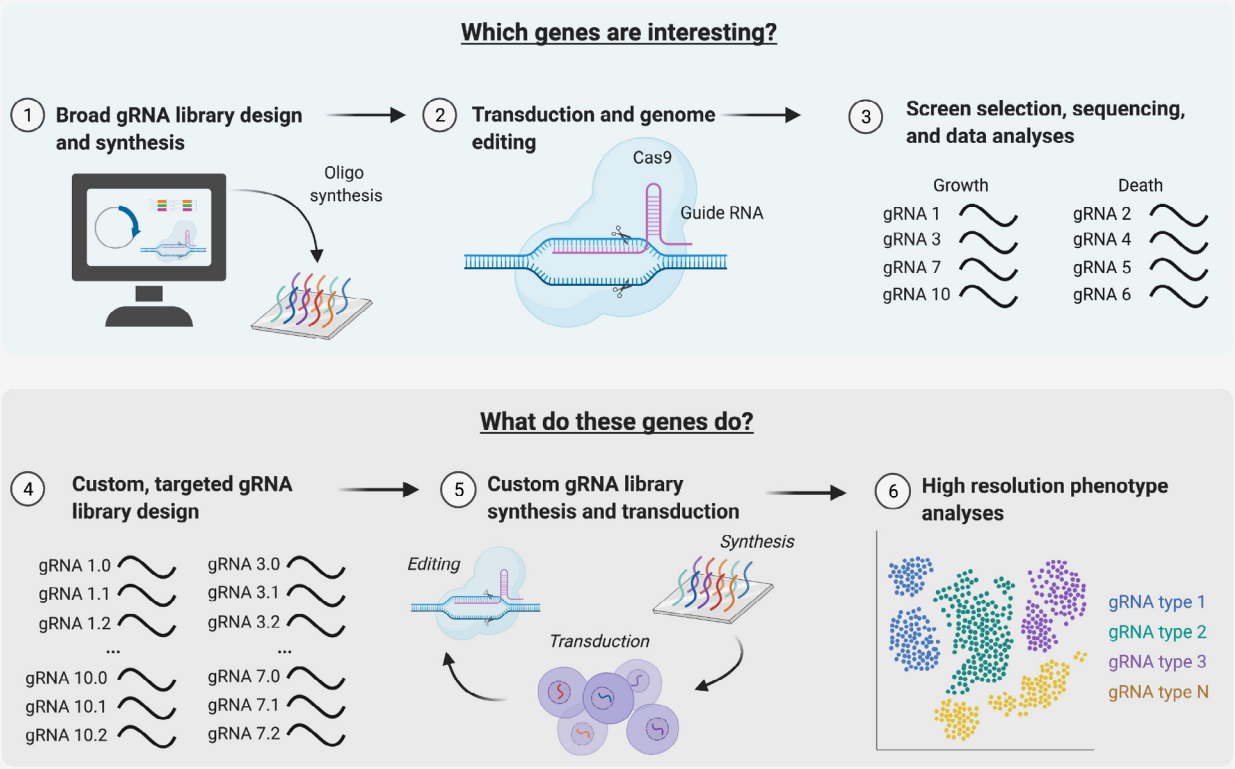
In this chapter, we'll explore five promising developments in CRISPR technology that are helping to usher in a new era of functional genomics.

Application 1—Focused CRISPR screening: Applying the discovery power of broad screening to specific questions

Functional genomic screening is often performed at the genome-wide scale, such that every gene is independently and redundantly targeted for manipulation. These genome-wide screens offer a powerful approach to unbiased hit discovery, but they also come with important limitations, including:

- The production of many hundreds of statistical hits that require resource-intensive validation.
- A need to balance breadth with depth, often resulting in fewer sgRNAs per gene and a potentially higher rate of false positives or negatives due to poor statistical power at the gene level or guide loss during sgRNA library preparation and delivery.
- Targeting genes on this scale requires significant cell counts. This isn't an issue for in vitro screens on immortal cell lines, but is problematic for more complex model systems, involving finite amounts of cellular resources.

Example CRISPR Screen workflow



Adapted from "CRISPR Screening Protocol", by BioRender.com (2021). Retrieved from https://app.biorender.com/biorender-templates.

An example workflow for a focused CRISPR screen. Focused CRISPR screens allow you to pursue more detailed phenotypic analyses, such as single-cell RNA sequencing, that may not be feasible for broader genome wide screening. Because of this, genome wide screens are often used to identify guide RNAs that contribute to a general phenotype like growth or death. A focused screen can then be designed based on the results of the genome wide screen in order to provide finer resolution detail about guide activity and gene function.

Unlike genome-scale libraries, focused CRISPR screening libraries concentrate resources on a select set of genomic targets to ensure a high depth of coverage, an increased likelihood of true hit detection, and an economical way to validate hits from genome-wide screens. With fewer targets, small sgRNA libraries also require fewer cells for the screening process which means screening can be done in more complex systems, such as primary cell lines, in vivo models, and in single-cell CRISPR screens (described in more detail below).

Focused screens are helping researchers overcome the aforementioned challenges and more confidently validate potential hits. For example,

Parnas *et al*. interrogated the innate immune response circuitry in dendritic cells, initially finding 2000 potential regulators in a genome-wide screen. The team narrowed their list down to 176 potential regulators for validation by using a focused CRISPR screen that provided better target coverage, specificity, and sensitivity, as well as a lower false discovery rate[2]. Similarly, Lindner *et al*. used CRISPR screening to supplement a larger CRISPR screen that had identified 835 potential hits. Through focused screening, this list was reduced to 75 hits with a low calculated false discovery rate[3].

Together with massively parallel oligonucleotide synthesis technology, focused CRISPR screens enable researchers to apply the discovery power of genome-wide screens to a select pathway or cluster of targets in an efficient and sensitive manner.

Application 2—Manipulating expression with CRISPRi, CRISPRa, CRISPR on, or CRISPR off

A recent wave of developments in CRISPR has expanded the technology's potential beyond gene editing by forming versions of CRISPR that alter epigenetic signatures and reversibly throttle gene expression, all without affecting the DNA sequence. The technologies center around the use of a catalytically dead Cas9 which can deliver fused proteins in a targeted fashion throughout the genome. These tools are referred to as: CRISPRi, CRISPRa, CRISPRoff, and CRISPRon.

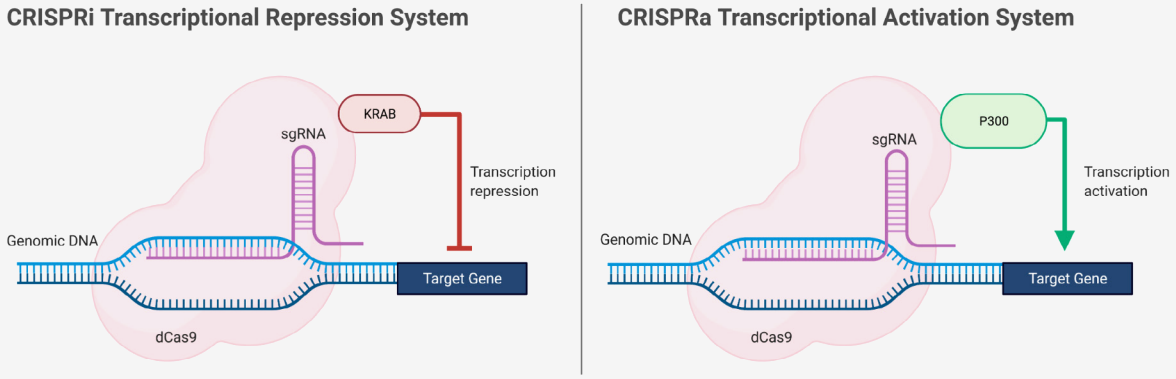
If the name CRISPRi is reminiscent of RNAi, it's for good reason: Both methods allow researchers to repress gene expression without modifying the DNA sequence.

CRISPRi targets the DNA proximal to the transcription start sites where it physically and reversibly blocks transcriptional machinery. To achieve this, CRISPRi fuses a transcriptional repressor (such as KRAB) to the nuclease-inactivated Cas protein, often referred to as dCas9[4-7]. Reversibility is gained by including an inducible element in the CRISPRi vector, such that repression only occurs in the presence of the inducing agent. Doing so expands the potential applications of genomic screens by enabling the temporal effects of gene perturbation to be studied[8,9].

Targeting non-protein coding regions is an important development to the CRISPR toolkit. Traditional CRISPR techniques that use nuclease active Cas proteins are largely ineffective for probing targets such a long non-coding RNA (lncRNA) or transcriptional regulatory elements, as these regions may have a higher functional tolerance for mutation[10].

Nonetheless, transcriptional regulatory elements are known to play a role in the development of many cancer types, however, the mechanisms remain unclear[11]. CRISPRi offers researchers an opportunity to query these regions and begin to elucidate their function, as was demonstrated by Lopes *et al*[11]. Using more than 34,000 CRISPRi sgRNAs, Lopes *et al*.

CRISPRi and CRISPRa



Adapted from "dCas9-KRAB Transcriptional Repression System", by BioRender.com (2021). Retrieved from https://app.biorender.com/biorender-templates.

An illustration of CRISPRi and CRISPRa mediated transcriptional regulation. Fusing nuclease dead (dCas9) with transcriptional regulatory proteins, such as KRAB and P300, can be leveraged for large CRISPR screens evaluating the effect of gene knockdown. While P300 is shown here, there have been several transcriptional activators explored for use in CRISPRa, for example VPR is also commonly used[4].

interrogated a series of 15,000 regulatory elements targeted by estrogen receptor. This analysis led to the identification of several regulatory regions that may play a significant role in breast cancer development[11].

Whereas CRISPRi prevents gene expression with the aid of a transcriptional repressor, CRISPRa amplifies gene expression with the aid of a transcriptional activator, such as P300. Having the ability to amplify native gene expression is also novel in the genomic perturbation field and allows researchers to study the effects of abnormal gene activation, as can be seen with tumor drivers.

Methylation is a repressive epigenetic mark that plays an important role in gene regulation and cell-fate determination. In conditions like cancer, changes to a cell's methylome are believed to be early in the transformation process[12]. However, the lack of good tools for targeted manipulation of epigenetic patterns on a large-scale have limited researchers' abilities to screen the effects of gene hyper- and hypomethylation.

CRISPRon/off makes it possible to include epigenetic modulation in functional genomics screens.

CRISPRoff makes use of the dCas9 protein that has been fused with DNA methyltransferase proteins (such as DNMT3A), allowing it to perform highly-targeted methylation of genes and regulatory elements. Methylation, particularly in the CpG islands that precede transcription start sites, can inhibit polymerase binding and result in repressed gene expression.

CRISPRon and CRISPRoff

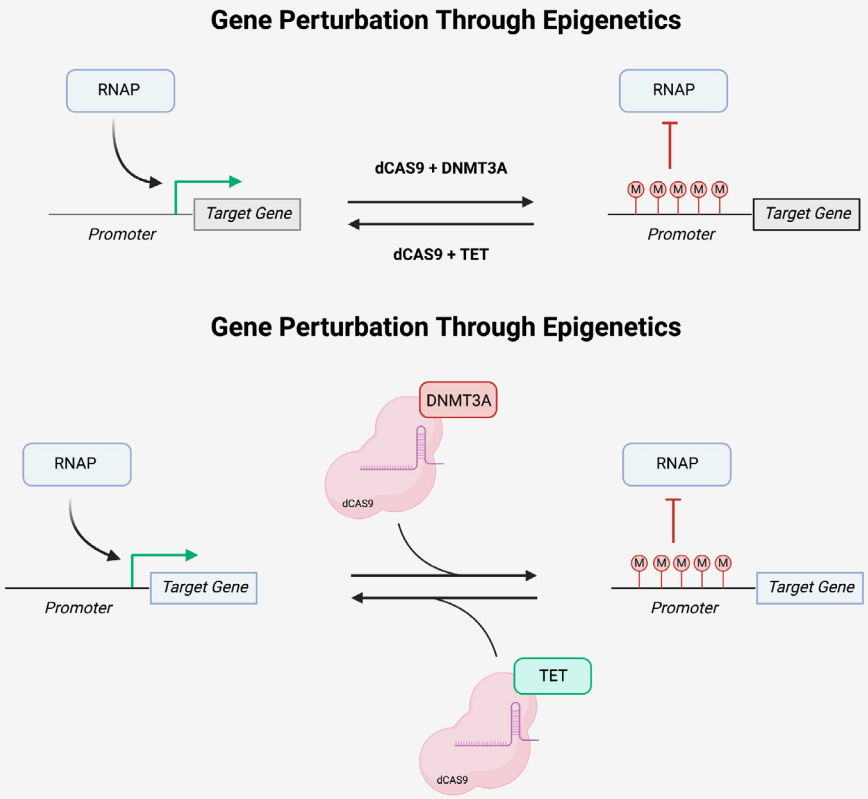


Figure created with BioRender.com.

A CRISPR-mediated perturbation of DNA methylation. Tethering Ten Eleven Translocase (TET) proteins or DNA (Cytosine-5) Methyltransferase 3A (DNMT3A) with nuclease dead Cas9 (dCas9) enables targeted DNA demethylation and methylation, respectively. Hypermethylated DNA typically is associated with gene repression, preventing recruitment of RNA polymerase (RNAP). Hypomethylation is associated with increased gene expression. In perturbing DNA methylation, CRISPRon/off enables researchers to explore the effect of reduced gene expression and allows targeting of regulatory elements, such as promoters, enhancers, and repressors.

By using sgRNAs targeting loci near the transcription start site of target genes, CRISPRoff can effectively and reversibly reduce gene expression. The inverse of CRISPRoff is CRISPRon, a technique that similarly uses a dCas9 protein that has been fused with a TET protein to decrease methylation of DNA[6,13].

With the ability to durably alter methylation patterns, researchers are better equipped to use high-throughput functional genomics to identify epigenetic drivers of myriad phenotypes.

Application 3—Dual-Guide Screens

As the name suggests, dual-guide CRISPR systems target two different regions of the genome in the same cell, guided by two sgRNAs.

Dual guide CRISPR is particularly interesting for use in synthetic lethality studies, wherein the effects of gene knockout may be masked by compensation of other proteins. By knocking out multiple genes at once, researchers can induce mutations that are only found in cancer cells, and then use this as a screening background to identify cancer-specific therapeutic targets[14].

Uses of dual-guide CRISPR include:

- Synthetic lethality studies, wherein the knockout of multiple proteins can help identify synergistic drug targets, for example synthetic lethal partners unique to cancer.
- Increasing knockout efficiency by targeting the same gene multiple times. Similarly, this can be used with CRISPRi/a to increase inhibition or activation efficiency.
- Introduction of point mutations through homology-directed repair by targeting Cas nickases to flanking regions around the target site.
- Excision of regulatory regions.
- To improve accuracy of Cas systems by using the Cas nickases to introduce single-strand breaks at complementary sites, such that both guides have to bind to the target loci for a double-stranded cut—and subsequent repair—to occur.

Illustrative of this is an elegant study from Thompson *et al*. wherein the authors used dual-guide screening to dissect synthetic lethal interactions between gene paralogues. From a screen of 1,191 gene pairs, 27 gene pairs were identified as robust synthetic lethal targets across multiple cell lines[15]. Synthetic lethality is highly context-dependent, and the variable spectrum of somatic mutations observed in cancer cells has made it difficult to systematically identify synthetic lethal targets. Thompson *et al.*'s work demonstrates how combinatorial CRISPR screening can be an effective tool for cataloging epistatic gene pairs and surfacing potential therapeutic targets.

Application 4—Moving beyond knockouts with base editing

CRISPR has historically been used to introduce knockout mutations wherein a frameshift renders a gene non-functional. Yet, many conditions like sickle cell anemia and progeria can be traced back to individual point mutations[16,17]. Functional genomics thus requires researchers to go beyond knockout mutations to investigate a wide spectrum of potential perturbations.

> "We did not detect any off-target or unintended bystander editing effects, which means the prime editing is highly precise. This high precision would minimize adverse effects originating from unintended edits, which have often been observed when CRISPR nucleases or base editors were used in animal models of human genetic diseases."
>
> — Professor Henry Kim
> Department of Pharmacology
> Yonsei University College of Medicine

Cytosine Base Editing with APOBEC

Base editing is a technique that combines the targeting potential of nuclease dead CRISPR with the editing potential of deaminase proteins, such as APOBEC (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like) proteins. APOBECs are a family of cytosine deaminases that catalyze the deamination of cytosine residues, resulting in the formation of uracil. During DNA replication, uracil is paired with an adenine, resulting in a G to A transition on non-target strands, and a C to T transition on target strands[17].

To improve CRISPR's ability to introduce single nucleotide polymorphisms (SNPs), APOBEC can be fused to a nuclease dead Cas9 protein which is then targeted to a specific genomic location by a sgRNA. Once CRISPR is bound to a complementary sequence, the APOBEC protein will deaminate cytosines within a short distance of the PAM sequence.

In addition to enabling single-base edits, APOBEC-CRISPR editing has two significant benefits over classic CRISPR perturbation: Base editing does not introduce potentially detrimental double-strand DNA breaks. And, without a reliance on non-homologous end joining, researchers can reliably predict the outcome of base editing.

Use of APOBEC-CRISPR fusions in large genomic screens has significant potential for studying drug resistance, cancer development, and other diseases that may result, not from gene knockout, but from point mutations. In February of 2021, Hanna *et al.*, demonstrated the utility of cytosine base editing in a study involving multiple large-scale base editing CRISPR screens[19]. Through the methodical mutation of more than 50,000 targets, the team identified point mutations that either enhanced or reduced the sensitivity of chemotherapeutic agents in primary cell lines. Base editors allowed the team to replicate clinically relevant mutations identified in ClinVar, as well as identify

mutations that may drive therapeutic resistance.

Multiple base editors targeting each nucleotide base have been developed and the application of these in gene editing and functional genomics is expanding[18].

Prime editing

Prime editing is a newer, state of the art method for introducing specific mutations, like insertions and deletions, into a DNA sequence. The technology leverages a Cas9 nickase fused to a reverse transcriptase, both of which are guided to the target site by a prime editing guide RNA (pegRNA). PegRNA is a guide RNA whose spacer sequence has been extended to include a template for repair of the DNA single-strand break induced by the nickase[20].

By incorporating a repair template, prime editing enables high fidelity editing relative to other forms of base editing. Deaminase-based editing, for example, is prone to "bystander edits" wherein bases proximal to the target site may be inappropriately edited. Prime editing avoids this issue by coding the desired change directly into the repair template, significantly reducing the likelihood of unintended mutations. Professor Henry Kim, a prime editing expert from Yonsei University recently published a paper in *Nature Biomedical Engineering* using prime editing to correct liver and eye diseases in mice[21]. "We did not detect any off-target or unintended bystander editing effects, which means the prime editing is highly precise" states Kim. "This high precision would minimize adverse effects originating from unintended edits, which have often been observed when CRISPR nucleases or base editors were used in animal models of human genetic diseases."

In Prof. Kim's paper, prime editing was used to prevent disease phenotypes in hereditary tyrosinemia and Leber congenital amaurosis[21]. In vitro prime editing screens were first used to narrow in on optimized pegRNA designs that were most likely to succeed in vivo. Subsequent in vivo testing showed a low editing frequency between 1% and 11% which was nonetheless effective at preventing disease phenotypes in both mouse models. This study demonstrates a powerful workflow leveraging in vitro prime editing screens to guide decisions and drive successful in vivo prime editing, potentially foreshadowing the technologies usefulness in future therapeutics applications.

The use of prime editing in large genomic screens thus has significant potential for studying and potentially rectifying hereditary disease, drug resistance, cancer development, and other phenotypes that may result, not from gene knockout, but from point mutations.

Retrons

A retron is a DNA cassette that codes for a reverse transcriptase and a highly structured RNA-DNA hybrid molecule. The RNA portion is critical for the

creation of single-stranded DNA from the retron element. This DNA may be nearly identical in sequence to the target DNA in the host's genome but for a single polymorphism. Incorporation of the retron DNA into the host's genome during replication (which can be aided by the inclusion of a single-stranded annealing protein) thus introduces the point mutation with precision[22].

Recent work from the Church lab at Harvard University showed that retron libraries can be generated for high-throughput screening in bacteria, demonstrating accurate target mutation with an almost 90% efficiency[23]. Unique DNA templates can be added to the retron cassette forming a barcode that can be differentiated during sequencing. The simplicity of producing retrons, their ease of barcoding, and their lack of dependence on nuclease proteins all make them amenable to high-throughput screening.

Thus far, demonstrations of retrons as a genome editing tool in mammalian cells utilize a combination of retron sequences and CRISPR[24]. In this approach, Cas9 is used to introduce double strand breaks at the target site, prompting homologous recombination repair with retron DNA as the template.

While still a nascent technology, retrons have drawn considerable excitement. When paired with CRISPR for large-scale screening, retrons may prove to be an important tool for studying the functional impact of point mutations and more.

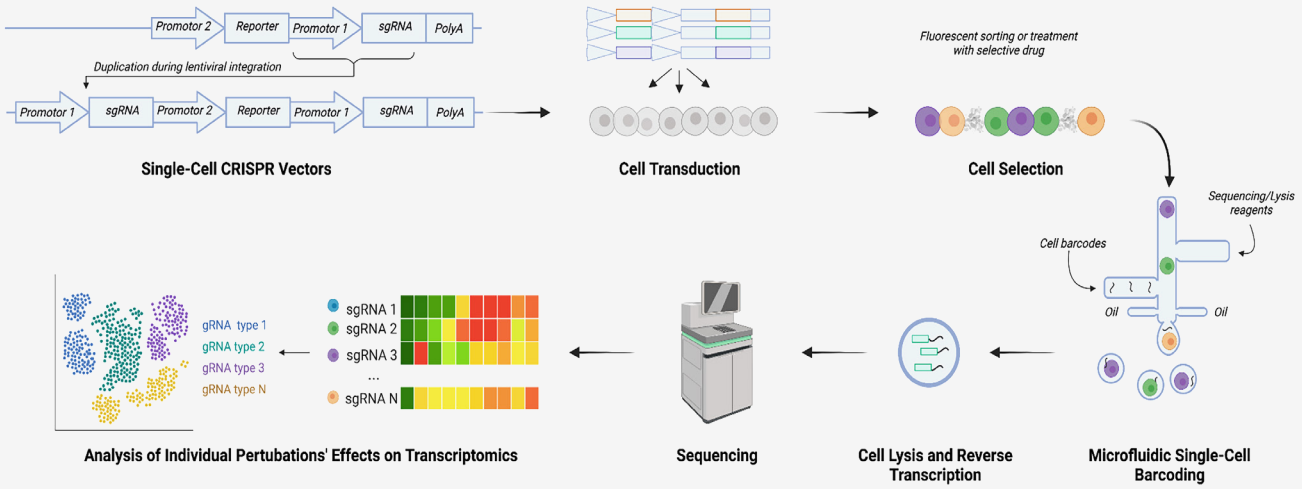
Application 5—Finer resolution functional genomics through Single-cell CRISPR screening

Single-cell CRISPR screening enables researchers to analyze more complex phenotypes resulting from genetic manipulation. CRISPR screening often relies on generalized phenotypes—such as growth, death, or the expression of a reporter gene. As a result, researchers are left with a pool of phenotypically similar cells that may each carry different mutations. Most perturbations do not affect cell viability or proliferation. By going beyond these and extending phenotypic analysis to include transcriptomic data, researchers can learn how perturbations affect cellular states and identities.

Through single-cell CRISPR screening, each perturbation can be linked to a transcriptomic profile, providing researchers with valuable insight into the individual effects of each sgRNA. This is achieved by leveraging techniques developed for single-cell RNA sequencing, wherein microfluidics and molecular barcoding are used to identify each individual cell's transcriptome and subsequent use of custom target enrichment panels allows for sensitive detection of targets[25-27].

Methods like CROP-seq, Perturb-seq, and direct capture Perturb-seq link transcriptomic profiles to their respective perturbations by assigning a unique molecular index sequence to each sgRNA (albeit through varying mechanisms) which can later be

Single-Cell CRISPR Screening General Workflow



This figure was created with BioRender.com and is adapted from workflows outlined for single-cell transcriptomics[24] and direct capture perturb-seq[27].

A Typical Single-Cell CRISPR Screening Workflow.
Workflows for single-cell CRISPR screening may differ, but generally follow the above diagram. A key step is designing each sgRNA to have an associated barcode that can later be used to identify which cell received which perturbation. The sgRNA vector design shown here is inspired by CROP-seq. Direct capture Perturbseq circumvents barcoding (see chapter 7 for more). Single-cell transcriptomics enable more complex phenotype analyses to be done, providing a more holistic picture of how each perturbation may influence cell behavior.

targeted during sequencing[25-27].

Dixit *et al*. demonstrated the power of Perturb-seq in lipopolysaccharide (LPS) stimulated mouse bone marrow-derived stem cells (BDMSCs)[27]. In response to LPS, BDMSCs undergo a differentiation process that involves a milieu of transcription factors. Perturb-seq was used to query the role of each transcription factor and provided insight into this complex cellular process.

A drawback of Perturb-seq is the potential for barcodes to become dissociated from their respective sgRNAs, leading to incorrect associations between phenotypes and perturbations. One way to overcome this limitation is to design the CRISPR vector in a way that decreases the likelihood of dissociation, or else relies on the sgRNA to serve as a molecular barcode.

Building on the work of Dixit *et al*., Replogle *et al*. used direct capture Perturb-seq to perform multiplexed high-throughput screening using CRISPRi and CRISPRa to study synthetic lethality between cholesterol synthesis pathway inhibition and DNA repair proteins[26]. Not only did this work provide valuable insight into a synthetic lethal interaction, it also demonstrated utility of direct capture Perturb-seq in combinatorial, dual-guide applications. This method's

use of target capture sequencing of guides enabled efficient single-cell screening while mitigating the likelihood of dissociation between sgRNAs and their barcodes.

CROP-seq similarly can be used for single-cell screening[27]. Here, CRISPR vectors are designed with the sgRNA template placed just upstream of a polyA tail. During lentiviral integration, a duplicate of the sgRNA and its promoter are inserted upstream of the reporter gene. This duplicate serves as a functional sgRNA, while the original template serves as a barcode.

These works demonstrate that single-cell CRISPR screening is a powerful method to begin functionally characterizing target genes, mapping gene networks, and highlighting cell-to-cell heterogeneity in high-throughput CRISPR screens. ■

References

1. "Hermann Joseph Muller's Study of X-Rays as a Mutagen, (1926-1927) | the Embryo Project Encyclopedia." Asu.edu, 2017, embryo.asu.edu/pages/hermann-joseph-mullers-study-x-rays-mutagen-1926-1927.
2. Parnas, Oren *et al*. "A Genome-wide CRISPR Screen in Primary Immune Cells to Dissect Regulatory Networks." *Cell* vol. 162,3 (2015): 675-86. doi:10.1016/j.cell.2015.06.059
3. Lindner, Benjamin, *et al*. "A Genome-Wide CRISPR/Cas9 Screen to Identify Phagocytosis Modulators in Monocytic THP-1 Cells." *Scientific Reports*, vol. 11, no. 1, 21 June 2021, 10.1038/s41598-021-92332-7. Accessed 18 Aug. 2021.
4. Chavez, Alejandro, *et al*. "Comparison of Cas9 Activators in Multiple Species." *Nature Methods*, vol. 13, no. 7, 23 May 2016, pp. 563–567, 10.1038/nmeth.3871. Accessed 10 Sept. 2019.
5. Meaker, Grace A *et al*. "Advances in engineering CRISPR-Cas9 as a molecular Swiss Army knife." *Synthetic biology* (Oxford, England) vol. 5,1 ysaa021. 24 Oct. 2020, doi:10.1093/synbio/ysaa021
6. Doench, John G. "Am I Ready for CRISPR? A User's Guide to Genetic Screens." *Nature Reviews Genetics*, vol. 19, no. 2, 4 Dec. 2017, pp. 67–80, 10.1038/nrg.2017.97.
7. le Sage, Carlos, *et al*. "CRISPR: A Screener's Guide." *SLAS DISCOVERY: Advancing the Science of Drug Discovery*, vol. 25, no. 3, 29 Oct. 2019, pp. 233–240, 10.1177/2472555219883621. Accessed 18 Aug. 2021.
8. Mandegar, Mohammad A *et al*. "CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs." *Cell stem cell* vol. 18,4 (2016): 541-53. doi:10.1016/j.stem.2016.01.022
9. Li, Rui *et al*. "Generation and validation of versatile inducible CRISPRi embryonic stem cell and mouse model." *PLoS biology* vol. 18,11 e3000749. 30 Nov. 2020, doi:10.1371/journal.pbio.3000749
10. Kampmann, Martin. "CRISPRi and CRISPRa Screens in Mammalian Cells for Precision Biology and Medicine." *ACS chemical biology* vol. 13,2 (2018): 406-416. doi:10.1021/acschembio.7b00657
11. Lopes, Rui, *et al*. "Systematic Dissection of Transcriptional Regulatory Networks by Genome-Scale and Single-Cell CRISPR Screens." *Science Advances*, vol. 7, no. 27, July 2021, p. eabf5733, 10.1126/sciadv.abf5733. Accessed 18 Aug. 2021.
12. De Almeida, Melanie, *et al*. "AKIRIN2 Controls the Nuclear Import of Proteasomes in Vertebrates." *Nature*, vol. 599, no. 7885, 28 Oct. 2021, pp. 491–496, 10.1038/s41586-021-04035-8. Accessed 13 Dec. 2021
13. Nuñez, James K., *et al*. "Genome-Wide Programmable Transcriptional Memory by CRISPR-Based Epigenome Editing." *Cell*, vol. 0, no. 0, 9 Apr. 2021, www.cell.com/cell/fulltext/S0092-8674(21)00353-6?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS0092867421003536%3Fshowall%3Dtrue, 10.1016/j.cell.2021.03.025. Accessed 10 Apr. 2021.

14. McCarty, Nicholas S., *et al.* "Multiplexed CRISPR Technologies for Gene Editing and Transcriptional Regulation." *Nature communications*, vol. 11, no. 1, 9 Mar. 2020, 10.1038/s41467-020-15053-x. Accessed 29 Nov. 2020.

15. Thompson, Nicola A *et al.* "Combinatorial CRISPR screen identifies fitness effects of gene paralogues." *Nature communications* vol. 12,1 1302. 26 Feb. 2021, doi:10.1038/s41467-021-21478-9

16. Koblan, Luke W *et al.* "In vivo base editing rescues Hutchinson-Gilford progeria syndrome in mice." *Nature* vol. 589,7843 (2021): 608-614. doi:10.1038/s41586-020-03086-7

17. Newby, Gregory A., *et al.* "Base Editing of Haematopoietic Stem Cells Rescues Sickle Cell Disease in Mice." *Nature*, vol. 595, no. 7866, 2 June 2021, pp. 295–302, 10.1038/s41586-021-03609-w.

18. Anzalone, Andrew V., *et al.* "Genome Editing with CRISPR–Cas Nucleases, Base Editors, Transposases and Prime Editors." *Nature Biotechnology*, vol. 38, no. 7, 22 June 2020, pp. 824–844, 10.1038/s41587-020-0561-9.

19. Hanna, Ruth E., *et al.* "Massively Parallel Assessment of Human Variants with Base Editor Screens." *Cell*, vol. 184, no. 4, Feb. 2021, pp. 1064-1080.e20, 10.1016/j.cell.2021.01.012. Accessed 13 Dec. 2021.

20. Anzalone, Andrew V *et al.* "Search-and-replace genome editing without double-strand breaks or donor DNA." *Nature* vol. 576,7785 (2019): 149-157. doi:10.1038/s41586-019-1711-4

21. Jang, Hyewon, *et al.* "Application of Prime Editing to the Correction of Mutations and Phenotypes in Adult Mice with Liver and Eye Diseases." *Nature Biomedical Engineering*, 26 Aug. 2021, 10.1038/s41551-021-00788-9. Accessed 7 Sept. 2021.

22. Simon, Anna J *et al.* "Retrons and their applications in genome engineering." *Nucleic acids research* vol. 47,21 (2019): 11007-11019. doi:10.1093/nar/gkz865

23. Schubert, Max G., *et al.* "High-Throughput Functional Variant Screens via in Vivo Production of Single-Stranded DNA." *Proceedings of the National Academy of Sciences*, vol. 118, no. 18, 27 Apr. 2021, p. e2018181118, 10.1073/pnas.2018181118. Accessed 18 Aug. 2021.

24. Zhao, Bin, *et al.* Bacterial Retrons Enable Precise Gene Editing in Human Cells. 29 Mar. 2021, 10.1101/2021.03.29.437260. Accessed 18 Aug. 2021.

25. Zilionis, Rapolas, *et al.* "Single-Cell Barcoding and Sequencing Using Droplet Microfluidics." *Nature Protocols*, vol. 12, no. 1, 8 Dec. 2016, pp. 44–73, 10.1038/nprot.2016.154. Accessed 20 Feb. 2021.

26. Replogle, Joseph M *et al.* "Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing." *Nature biotechnology* vol. 38,8 (2020): 954-961. doi:10.1038/s41587-020-0470-y

27. Dixit, Atray *et al.* "Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens." *Cell* vol. 167,7 (2016): 1853-1866.e17. doi:10.1016/j.cell.2016.11.038

29. Datlinger, Paul, *et al.* "Pooled CRISPR Screening with Single-Cell Transcriptome Readout." *Nature Methods*, vol. 14, no. 3, 18 Jan. 2017, pp. 297–301, 10.1038/nmeth.4177. Accessed 5 Dec. 2019.

CRISPRi/CRISPRa Screens Reveal Neuron-Specific Pathways That May Lead to Dementia

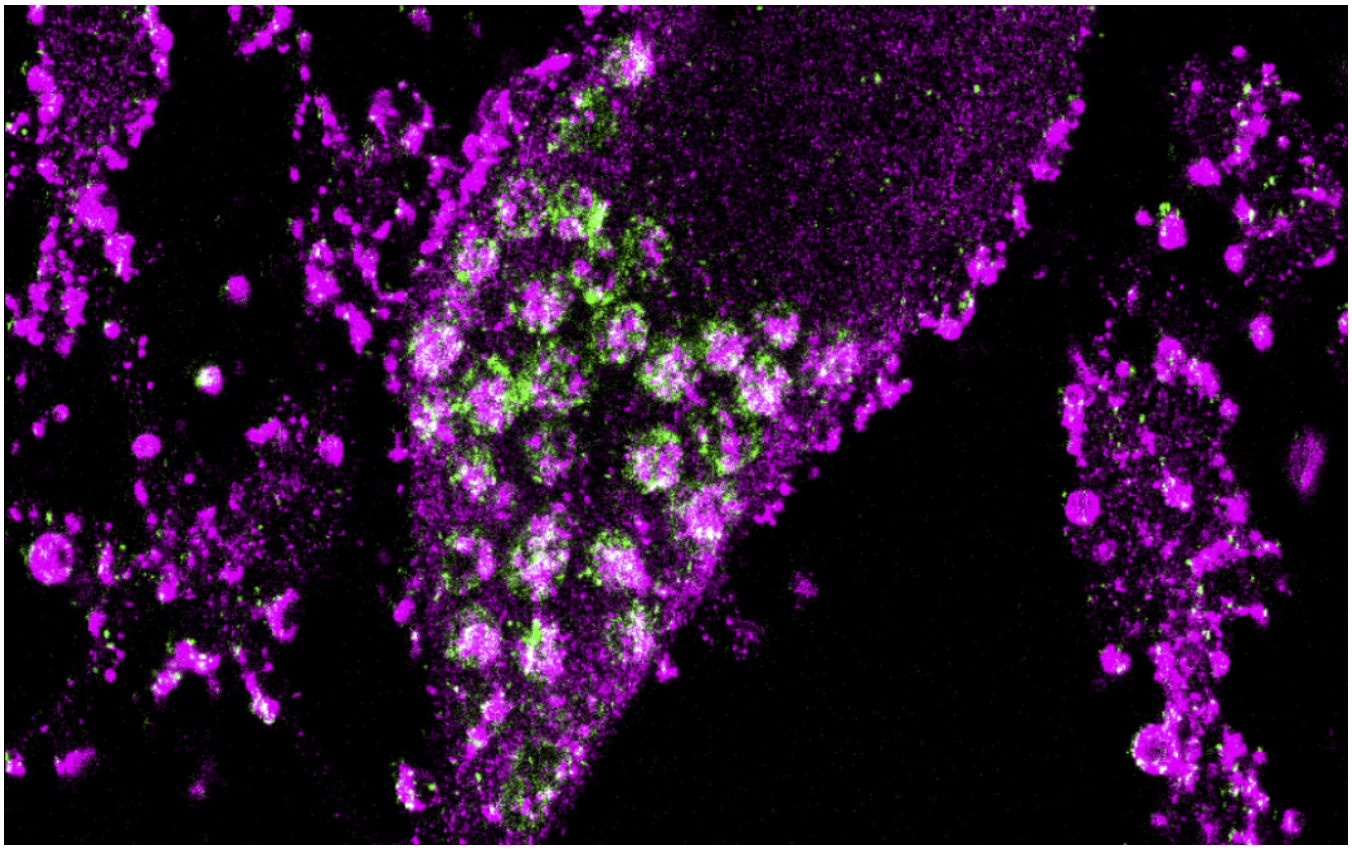
By integrating CRISPR-based functional genomics and stem cell technology, researchers based at the University of California, San Francisco (UCSF), have uncovered pathways that control the neuronal response to chronic oxidative stress, which is implicated in neurodegenerative diseases. The researchers, led by Martin Kampmann, PhD, associate professor at UCSF, determined how individual genes in human-stem-cell-generated neurons could, upon inactivation or activation, affect the ability of the neurons to cope with toxic, oxygen-containing molecules.

To their surprise, the researchers found that neurons became more vulnerable to oxidative stress if a gene encoding a lysosomal protein was disabled. This discovery was reported May 24 in the journal *Nature Neuroscience*, in an article titled, "Genome-wide CRISPRi/a screens in human neurons link lysosomal failure to ferroptosis."

"Unexpectedly, knockdown of the lysosomal protein prosaposin strongly sensitizes neurons, but not other cell types, to oxidative stress by triggering the formation of lipofuscin, a hallmark of aging, which traps iron, generating reactive oxygen species and triggering ferroptosis," the article's authors wrote. "We also determine transcriptomic changes in neurons after perturbation of genes linked to neurodegenerative diseases."

> ""By simply inactivating a single gene, in only days we could generate a hallmark of aging that would normally take decades to develop in the human body."
>
> — Martin Kampmann, PhD,
> Associate Professor
> University of California, San Francisco

The article described how the UCSF researchers turned individual genes off and on by using genetic screens that incorporated CRISPR inactivation (CRISPRi) and CRISPR activation (CRISPRa) machinery. The article also asserted that the UCSF researchers are the first to present the results of genome-wide CRISPRi and CRISPRa screens



UCSF

in human neurons.

The article added that the CRISPRi/CRISPRa approach could be applied with many different human cell types, not just neurons. To help realize this possibility, the UCSF researchers have established a data commons named CRISPRbrain.

"This is the key next step in uncovering the mechanisms behind disease genes," said Kampmann. "There are lots of human genetics studies linking specific genes to specific diseases. The work we're doing can provide insight into how changes in these genes lead to disease and allow us to target them with treatments."

To identify genes that might be involved in neurodegenerative diseases such as Alzheimer's and related forms of dementia, Kampmann and colleagues evaluated stem-cell-generated human neurons after individual genes had been turned on and off. The researchers were looking specifically for downstream changes in gene expression that would produce oxidative stress in the cell. Such stress is thought to contribute to neurodegeneration.

Among the most interesting of the team's findings was that switching off the gene for a protein called prosaposin, which normally assists with the cell's recycling of waste products, greatly increased the levels of oxidative stress. In neurons, prosaposin is associated with a part of the cell called the lysosome, where biological molecules and toxins are sorted through and dealt with in a variety of ways.

"At first glance, prosaposin should have nothing to do with oxidative molecules," Kampmann noted. "It caught our attention because this gene had recently been linked to Parkinson's disease. What was really exciting was that now, with the results from this CRISPR screen, we had a cell-based model to help us understand what's behind that linkage."

The team then embarked on what Kampmann called a "detective story" to find out how the lack of prosaposin is linked to neurodegeneration. The researchers found that suppression of the gene led to buildup of a substance called age pigment, which has been seen in aging cells whose lysosomes no longer degrade material as efficiently. The researchers discovered that age pigment trapped iron, generating reactive oxygen molecules that triggered ferroptosis, an iron-dependent process that leads to cell death.

"By simply inactivating a single gene," Kampmann emphasized, "in only days we could generate a hallmark of aging that would normally take decades to develop in the human body."

The cascade of changes Kampmann and colleagues observed are specific to the function of neurons and are related to just one set of conditions. He said the results make a case for using CRISPRi/CRISPRa to perform similar screens looking for changes that prompt other kinds of disease-related environments in neurons and other types of differentiated cells.

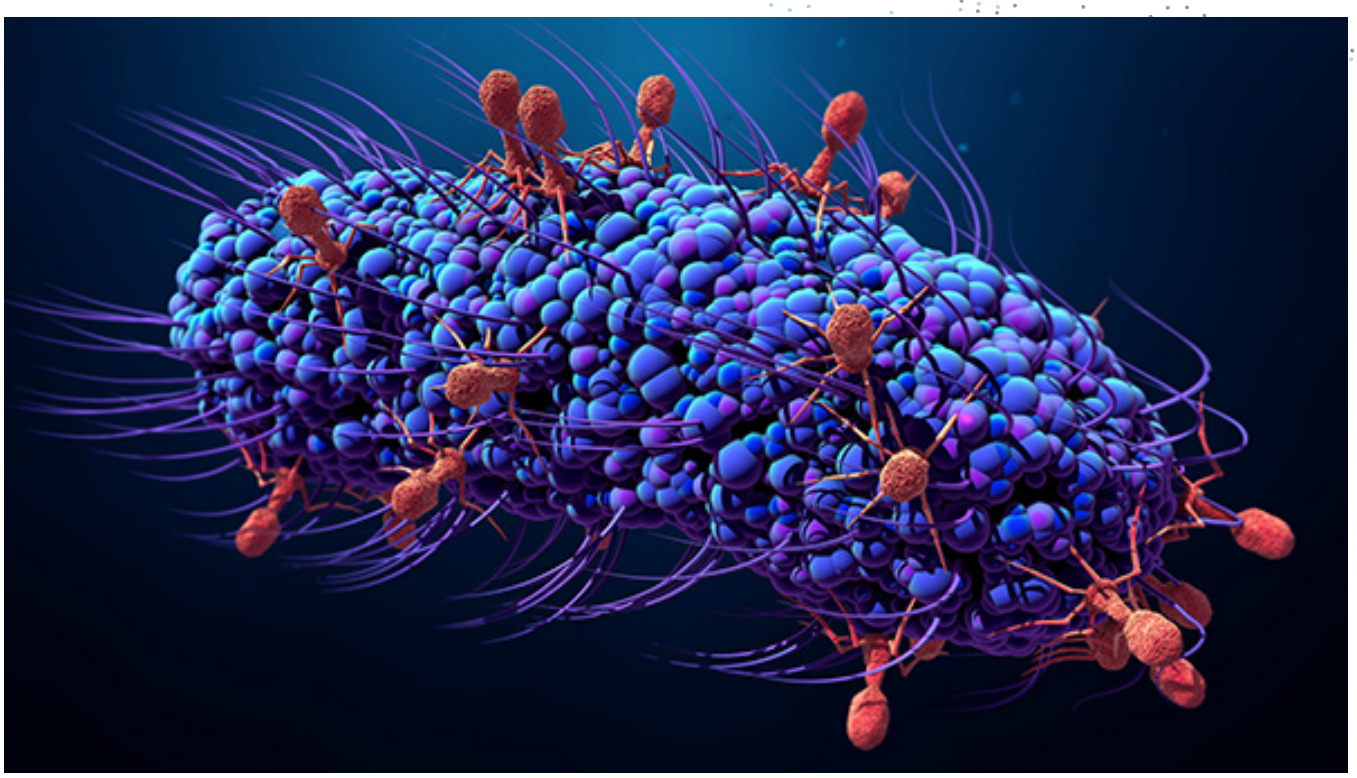
To that end, the team created CRISPRbrain, an open-access database that is designed to let scientists share and study large-scale data sets like the ones generated in the current study. Applying advanced computational technology such as machine learning can then detect patterns in this sea of data.

"By becoming the data commons for screens of many different cell types from many different labs and in different disease contexts, we can achieve a critical mass of information," Kampmann said. "There's enormous power in aggregating and cross-analyzing all of this."

The UCSF team's next step is to perform similar screens on neurons made from stem cells derived from patients with mutations known to contribute to neurodegeneration, as well as look at other cells such as astrocytes and microglia that play roles in brain disease.

Kampmann's hope is that the technology and database are widely adopted: "Now that we can do this in a systematic way, we can really interpret the underlying processes of how genes contribute to disease and find pathways to treat those conditions." ■

Retrons Display Genome Editing Strengths Even CRISPR Might Envy



Design Cells / Getty Images

Retron-based genome editing has been inviting comparisons with CRISPR-based genome editing, especially since researchers have learned that retrons, like CRISPR systems, function as a sort of immune system in bacteria. However, the comparisons may be a little premature. Retron-based genome editing has yet to work in mammalian cells. Nonetheless, retrons have already been shown to offer unique advantages in genome editing applications.

The latest advance in retron-based genome editing comes from researchers at the Wyss Institute for Biologically Inspired Engineering at Harvard University and Harvard Medical School (HMS). They report that they have created a new tool for retron-based genome editing. The tool, which is called Retron Library Recombineering (RLR), may be used to perform high-throughput functional screens that surpass the scale and specificity of the screens enabled by CRISPR-Cas technology.

RLR generates up to millions of mutations simultaneously, and it inserts "barcodes" into mutant cells so that the entire pool can be screened at once, enabling massive amounts of data to be easily generated and analyzed. A description of how RLR performed in bacterial cells appeared April 29 in

the *Proceedings of the National Academy of Sciences (PNAS)*, in a paper titled, "High-throughput functional variant screens via in vivo production of single-stranded DNA."

"We use the targeted reverse-transcription activity of retrons to produce single-stranded DNA (ssDNA) in vivo, incorporating edits at >90% efficiency and enabling multiplexed applications," the article's authors wrote. "RLR simultaneously introduces many genomic variants, producing pooled and barcoded variant libraries addressable by targeted deep sequencing.

"We use RLR for pooled phenotyping of synthesized antibiotic resistance alleles, demonstrating quantitative measurement of relative growth rates. We also perform RLR using the sheared genomic DNA of an evolved bacterium, experimentally querying millions of sequences for causal variants, demonstrating that RLR is uniquely suited to utilize large pools of natural variation."

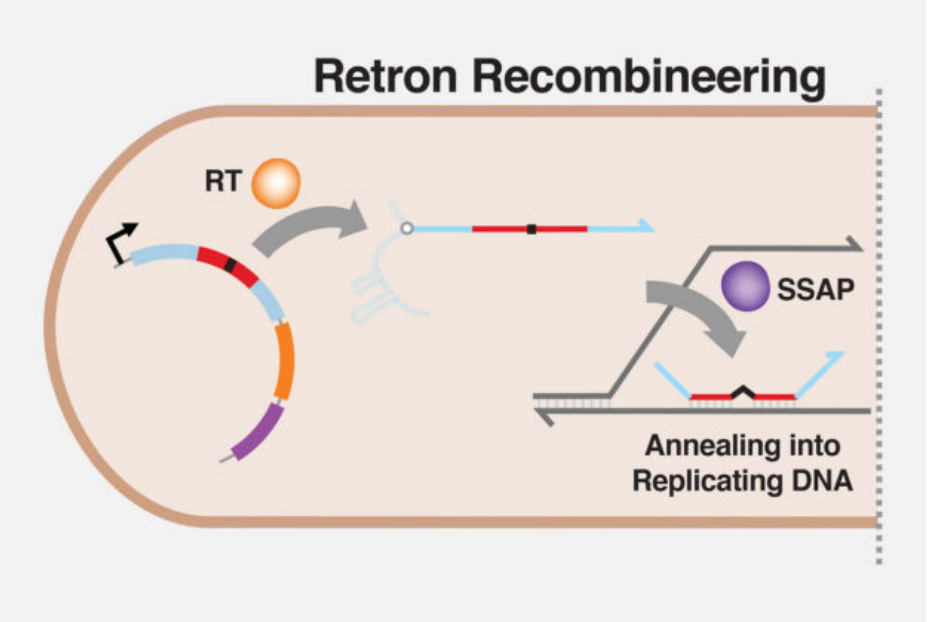
Retrons—complexes of DNA, RNA, and protein—are poorly understood bacterial retroelements that undergo targeted reverse-transcription, producing single-stranded multicopy satellite DNA (msDNA). Several research teams have shown that msDNA can function as a recombineering donor, creating specific edits in the genome. This capability suggests that retrons could function as part of a "donor only" genome editing method, unlike CRISPR, which is typically a "donor + guide" method.

Another attraction of retrons is that their sequences themselves can serve as "barcodes" that identify which individuals within a pool of bacteria have received each retron sequence, enabling dramatically faster, pooled screens of precisely created mutant strains.

Retron recombineering has been used in various studies, including synthetic biology studies, but editing rates were low—too low for retron-based genome editing to be practical for studying mutants of physiological interest.

To see if retron recombineering could be more efficient, the Wyss scientists first created circular plasmids of bacterial DNA that contained antibiotic resistance genes placed within retron sequences, as well as a single-stranded annealing protein (SSAP) gene to enable integration of the retron sequence into the bacterial genome. They inserted these retron plasmids into *Escherichia coli* bacteria to see if the genes were successfully integrated into their genomes after 20 generations of cell replication. Initially, less than 0.1% of *E. coli* bearing the retron recombineering system incorporated the desired mutation.

To improve this disappointing initial performance, the team made several genetic tweaks to the bacteria. First, they inactivated the cells' natural mismatch repair machinery, which corrects DNA replication errors and could therefore be "fixing" the desired mutations before they were able to



Max Schubert / Wyss Institute at Harvard University

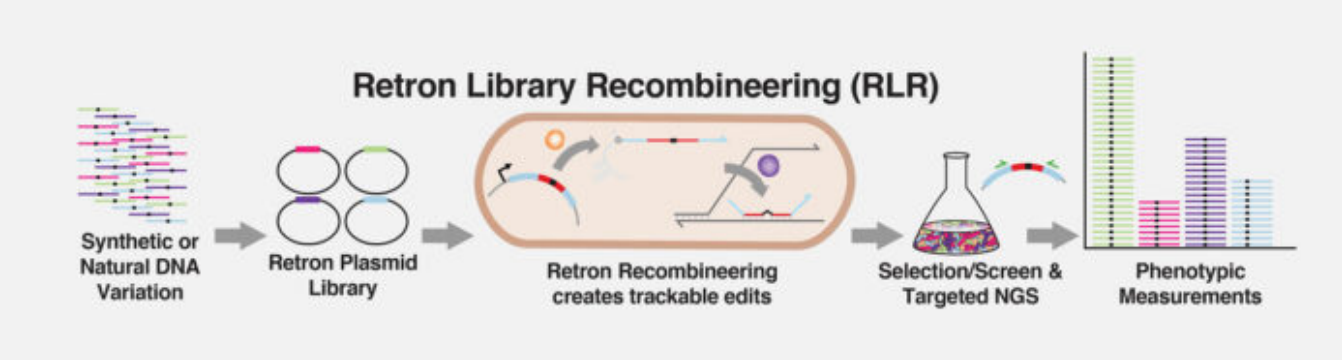
Retron sequences (red) containing a mutation of interest (black notch) are introduced into a bacterial cell along with the enzyme reverse transcriptase (RT). The retron then produces ssDNA that is inserted into replicating DNA with the help of another enzyme called single-stranded annealing protein (SSAP).

be passed on to the next generation. They also inactivated two bacterial genes that code for exonucleases—enzymes that destroy free-floating ssDNA. These changes dramatically increased the proportion of bacteria that incorporated the retron sequence, to more than 90% of the population.

Now that they were confident that their retron ssDNA was incorporated into their bacteria's genomes, the team tested whether they could use the retrons as a genetic sequencing "shortcut," enabling many experiments to be performed in a mixture. Because each plasmid had its own unique retron sequence that can function as a "name tag," they reasoned that they should be able to sequence the much shorter retron rather than the whole bacterial genome to determine which mutation the cells had received.

First, the team tested whether RLR could detect known antibiotic resistance mutations in *E. coli*. They found that it could—retron sequences containing these mutations were present in much greater proportions in their sequencing data compared with other mutations. The team also determined that RLR was sensitive and precise enough to measure small differences in resistance that result from very similar mutations. Crucially, gathering these data by sequencing barcodes from the entire pool of bacteria rather than isolating and sequencing individual mutants, dramatically speeds up the process.

Then, the researchers took RLR one step further to see if it could be used on randomly fragmented DNA, and find out how many retrons they could use at once. They chopped up the genome of a strain of *E. coli* highly resistant to another antibiotic, and used those fragments to build a library of tens of millions of genetic sequences contained within retron



Max Schubert / Wyss Institute at Harvard University

Retrons enable the rapid production and screening of millions of trackable DNA variations and their effects on bacteria simultaneously.

sequences in plasmids.

"The simplicity of RLR really shone in this experiment, because it allowed us to build a much bigger library than what we can currently use with CRISPR, in which we have to synthesize both a guide and a donor DNA sequence to induce each mutation," said Max Schubert, PhD, the co-first author of the PNAS article and a postdoc in the lab of Wyss core faculty member George Church, PhD, senior author of the paper and leads the Wyss Institute's synthetic biology focus area and is also a professor of genetics at HMS.

This library was then introduced into the RLR-optimized *E. coli* strain for analysis. Once again, the researchers found that retrons conferring antibiotic resistance could be easily identified by the fact that they were enriched relative to others when the pool of bacteria was sequenced.

"Being able to analyze pooled, barcoded mutant libraries with RLR enables millions of experiments to be performed simultaneously, allowing us to observe the effects of mutations across the genome, as well as how those mutations might interact with each other," said Church. "This work helps establish a road map toward using RLR in other genetic systems, which opens up many exciting possibilities for future genetic research." ■

Tools for Conducting CRISPR Screens at the Single-cell Resolution

Abstract

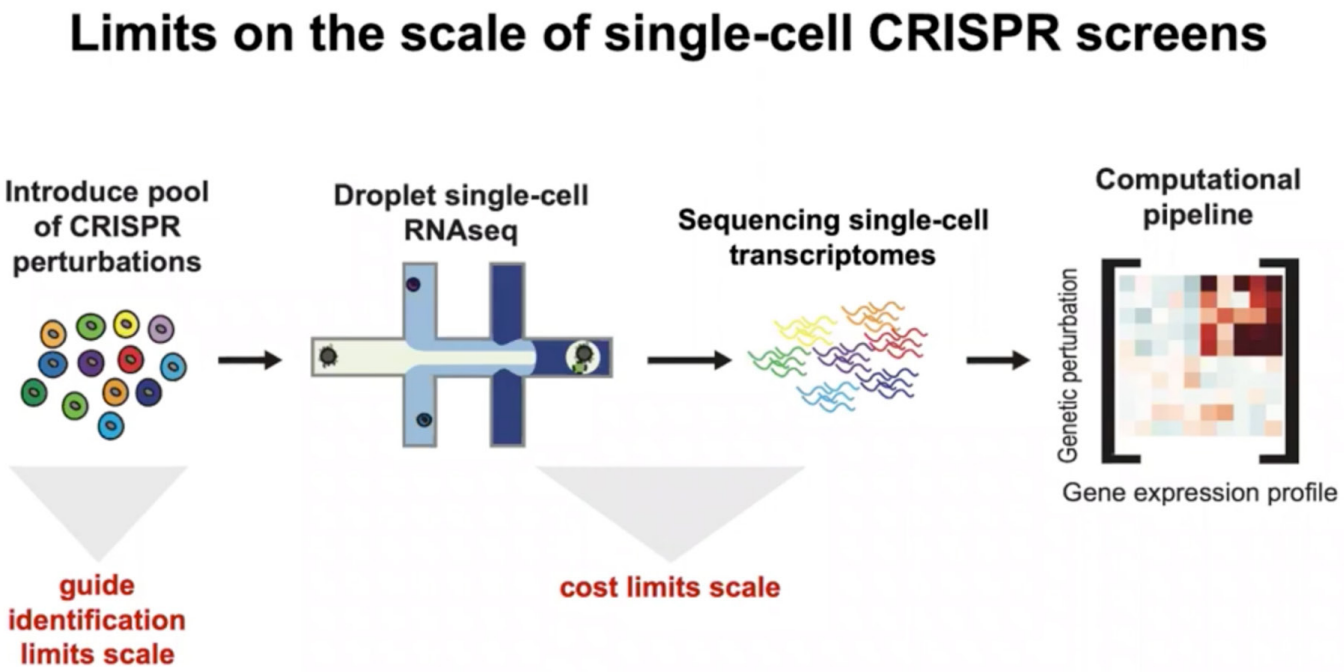
Single-cell transcriptomics enables researchers to see populations of cells in finer detail, breaking phenotypically similar cells into discrete categories based on transcriptomic profiles. Technical limitations—such as a reliance on polyadenylated transcripts for sequencing—have hindered the use of single-cell transcriptomics in large-scale CRISPR screening. In recent years, however, new strategies for sequencing and sgRNA design have made it possible to use single-cell technologies to track the transcriptional effects of CRISPR perturbations on a large scale. Here, we detail how Joseph Replogle, an M.D.-Ph.D. student in the Weissman lab at the University of California at San Francisco, and Britt Adamson, Assistant Professor at Princeton University, developed a method for genome-scale single-cell CRISPR screening—a method that offers considerable advancements for functional genomic screening.

For the short time that it's been around, CRISPR screening has maintained the same basic premise: methodically alter or ablate gene expression and see what happens to the cell. Technological advances have made it possible for researchers to alter the expression of a gene in increasingly complex ways, advancing from simple knockouts to epigenetic perturbations. Yet, the "see what happens" part of CRISPR has largely stayed the same as researchers are often limited to screening perturbed cells for simple, one-dimensional phenotypes, such as growth, death, or the expression of a reporter gene. The lack of nuance in these phenotypes means that the molecular cascade linking a sgRNA to the phenotype is hard to decipher and may differ dramatically between cells.

To advance CRISPR screening technology, researchers are developing ways to integrate more complex and informative phenotypes into the screening process. One promising approach is known as direct capture Perturb-seq, the latest advancement in single-cell CRISPR screening methodology.

What is single-cell CRISPR screening?

Single-cell CRISPR screening is a type of CRISPR screen that's designed to capture RNA sequencing data from individual cells within the test population. By analyzing cells' transcriptomic profiles, researchers can design screens to go beyond the simplistic phenotypes centered on cell viability and instead begin to analyze relationships between genes, perturbations, and phenotypes.



The process of single-cell CRISPR screening, highlighting two hurdles to scaling this technology: guide identification, and cost.

Take for instance the characterization of regulatory elements in the genome. Promoters, enhancers, repressors, and other regulatory elements are peppered throughout the genome, yet many of them have an unknown function, while many more have yet to be discovered. Single-cell CRISPR screening can be used in combination with tools such as CRISPRi/a/on/off to identify relationships between regulatory elements and the genes they influence[1].

Single-cell CRISPR screening can thus be a critical tool in advancing our understanding of human genetics, health, and disease. However, to realize this potential, methodologies will need to be developed that allow single-cell CRISPR screening to be carried out with high fidelity at genome-scale.

A team of researchers in Jonathon Weissman's lab at UCSF/MIT and Britt Adamson's lab at Princeton have collaborated on a new approach to single-cell CRISPR screening, one that is amenable to multiplexed large-scale screening, while offering increased editing efficiency using multiple guides per cell. Their approach is known as direct capture Perturb-seq.

What is Direct Capture Perturb-seq

First reported in *Nature Biotechnology* and later elaborated in a **webinar** on the Twist Bioscience platform, direct capture Perturb-seq is a method for

performing single-cell RNA sequencing on CRISPR modified cells at a large scale[2].

Direct capture Perturb-seq is the latest iteration of Perturb-seq—a method first reported in 2016 by many of the same authors. In its original design, Perturb-seq offered researchers an efficient way to capture sgRNAs for RNA-sequencing.

A major hurdle in single-cell CRISPR screening is linking sgRNAs to their respective transcriptomic profiles. In a recent webinar detailing the development of direct capture Perturb-seq, Joseph Replogle explained: "The challenge is that single-cell RNA-seq generally relies on capturing 3'-polyadenylated transcripts, but guide RNAs are expressed by RNA Pol III and therefore are not polyadenylated, and are not captured on conventional, single-cell RNA-seq platforms. The solution that made single-cell CRISPR screens possible was to express barcoded transcripts alongside the guide RNAs."

Perturb-seq inserts guide-specific barcodes in the 3' untranslated region of a selection gene that is often included in sgRNA vectors. Unlike the sgRNA, this selection gene would produce a polyadenylated transcript that contained the unique molecular barcode, enabling researchers to identify which sgRNA had been used.

While effective, Replogle explains that this approach had one major drawback: "There's this long stretch of homology between the sgRNA and the guide barcode. And, because lentiviruses are pseudodiploid, they can recombine their genomes in those homologous regions. If lentiviruses are prepared with many different sgRNAs in a pooled format, then you can have recombination leading to uncoupling of the guide from the guide barcodes that prevents you from identifying the genotype of every cell. So that meant that you needed to prepare your lentivirus in an array format which is not scalable beyond a few hundred genes at a time."

Direct capture Perturb-seq is a modified version of Perturb-seq that removes the need for a 3' barcode by adding a capture sequence directly into the sgRNA constant region, enabling direct capture of the sgRNA and opening the door to larger-scale projects.

When put to the test, Replogle and the team found that direct capture Perturb-seq enabled assignment of guides to 90% of their cell population, and was able to delineate gene-relationships between 30 well-studied genes with equivalent performance to the current gold-standard (Perturb-seq).

With capture sequences coded into the sgRNA, vectors can now be designed to hold multiple sgRNAs. One application of dual-sgRNA screens is to study the mechanism underlying genetic interactions in cancer, as the UCSF team demonstrated in a paper in Science[3]. According to Replogle, dual-guide CRISPR screens have significant potential to increase the power of Single Cell CRISPR screens "because they give you a greater chance of having

an active guide in each cell," and because dual-targeting of a gene "could increase the efficacy of CRISPRi/a by multiplexing the guide activity, creating synergy between guides." Highlighting this, Replogle and his colleagues showed that single-guide CRISPRi gene targeting with direct capture Perturb-seq resulted in 80% gene knockdown, whereas dual-guide targeting led to 89% knockdown.

Together, this data suggests that direct capture Perturb-seq performs as well as the original Perturb-seq, but is free of its predecessor's scalability limitations. ■

Considerations when performing direct capture Perturb-seq

There are several considerations to keep in mind when designing a direct capture Perturb-seq experiment. For a full list of considerations, read through the team's paper[2] and listen to Replogle's recent **webinar**. In the meantime, here are some highlights:

- Consider using biotinylated target enrichment libraries to increase RNA-seq sensitivity. Replogle and the team showed that transcripts that had previously represented only 6% of the sequenced targets could be boosted to 90% with a Twist Bioscience target enrichment library. Enriching for the targets you're interested in can greatly decrease the cost of single-cell CRISPR screening.
- Pay close attention to where you place your capture sequences. Replogle reported that the capture sequence that performed best in the stem-loop did not perform as well when placed at the 3′ end, indicating that sequence performance is location-dependent.
- Consider the applications you're interested in to determine whether to use 3′ or 5′ capture. Both are possible with direct capture Perturb-seq, and both can be advantageous under specific circumstances (outlined in the webinar).

References

1. Gasperini, Molly, et al. "A Genome-Wide Framework for Mapping Gene Regulation via Cellular Genetic Screens." *Cell*, vol. 176, no. 1-2, Jan. 2019, pp. 377-390.e19, www.cell.com/cell/fulltext/S0092-8674(18)31554-X, 10.1016/j.cell.2018.11.029.
2. Replogle, Joseph M., et al. "Combinatorial Single-Cell CRISPR Screens by Direct Guide RNA Capture and Targeted Sequencing." *Nature Biotechnology*, vol. 38, no. 8, 30 Mar. 2020, pp. 954–961, 10.1038/s41587-020-0470-y.
3. Norman, Thomas M., et al. "Exploring Genetic Interaction Manifolds Constructed from Rich Single-Cell Phenotypes." *Science*, vol. 365, no. 6455, 23 Aug. 2019, pp. 786–793, 10.1126/science.aax4438.

Precision. Performance. Flexibility.

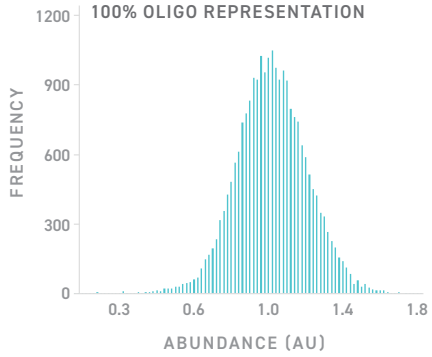
Twist Oligo Pools enable groundbreaking science

Performance without limits

Do more: Enable better experimental design with oligo lengths up to 300mers

Don't compromise: Unmatched uniformity drives confidence in results

Sacrifice nothing: Precise synthesis ensures maximal sequence representation



Eliminate the noise, expand your idea of what's possible, and design without limits!

LEARN MORE AT TWISTBIOSCIENCE.COM

TWIST BIOSCIENCE