CRISPR Screening 101

Strategies for Target Identification

SYNTHEGO

Introduction to Genetic Screens

Functional genetic screens are a powerful tool for understanding the genetic underpinnings of biological pathways at a systems level. Through the induction of hundreds to thousands of genome-wide modulations, these screens enable researchers to form hypotheses about genetic associations with normal or disease phenotypes. Causal relationships between the genes and phenotypes are then validated through further screening and experimentation.

There are two modalities of genetic modulation used in screens: gain-of-function and loss-of-function. The former involves overexpressing genes so that more mRNA/protein is produced, whereas the latter involves reducing mRNA levels or completely terminating gene function. For both types of manipulation, changes in resulting phenotypes indicate the involvement of a gene in the pathway or disease state of interest.

Over the past decade, new methodological breakthroughs have greatly advanced screening technologies. In particular, the emergence of CRISPR has enabled researchers to deactivate genes in a robust and highly specific manner. This eBook provides an introduction to CRISPR-mediated loss-of-function (LOF) screens, with a particular focus on identifying and validating gene targets for drug discovery. We will first discuss the role of screens in target identification for new drugs, review available gene repression technologies, and then describe two formats of CRISPR screens and their components.

Genetic Screens in Drug Discovery

Loss-of-function screens play an important role in **drug discovery** (Fig 1), the process through which putative therapeutic compounds are identified to treat diseases. In the target-based approach to drug discovery the pipeline begins with **target identification**, a process that identifies genes (or mRNA/proteins) that associate with a disease and can thus serve as "targets" for a potential drug. While many factors can contribute to identifying putative targets—including data mining biomedical research and clinical studies—LOF screens often play a crucial role. For instance, screens using healthy cells can be used to identify genes that, when disrupted, recapitulate a disease. Alternatively, genetic disruptions in diseased cells (e.g., cancer) that cause a normal phenotype can mimic the therapeutic effect of a drug. Ultimately, these screens can identify genes that affect disease sensitivity or viability, or uncover completely new members of a disease pathway.

LOF screens for target identification require thoughtful planning, including the choice of a cell model, a method of gene perturbation, and an assay. Assay development is a necessary prerequisite, as it is needed to measure changes in a desired phenotype resulting from genetic perturbation. This process may include defining an assay window,

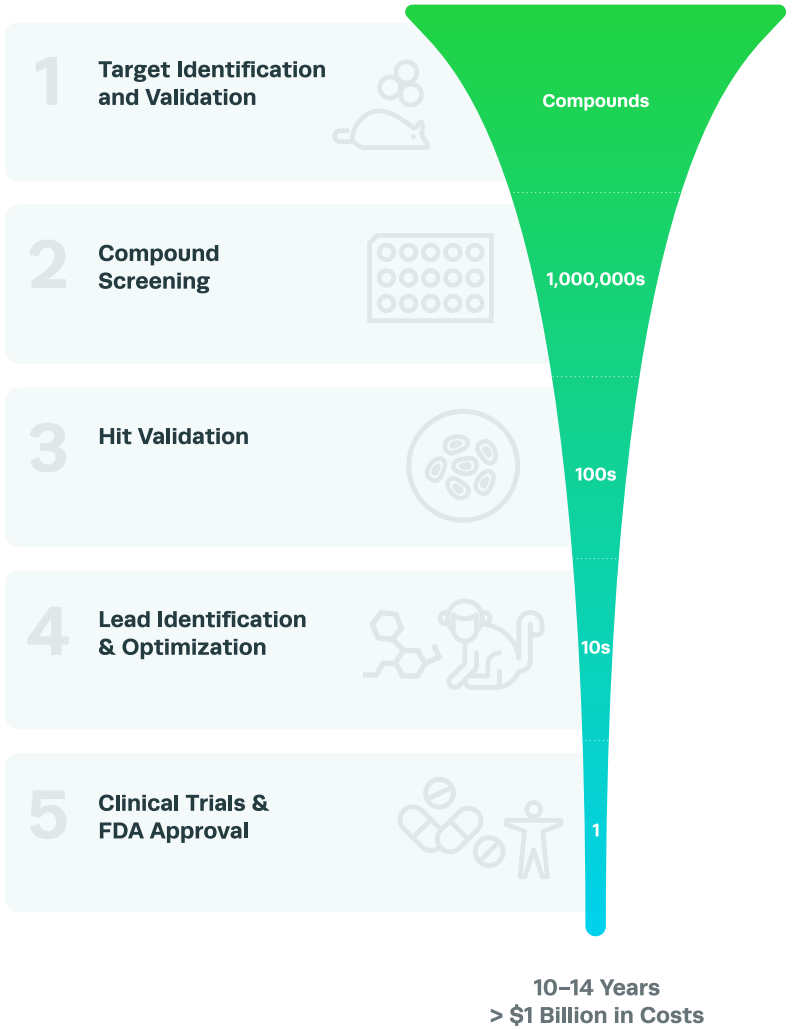


Figure 1. Target-based drug discovery pipeline.
1) Potential drug targets for a disease of interest are identified. Then, comprehensive validation by secondary screens confirm the genetic contribution to a disease phenotype. **2)** Selected targets are used to screen a large library of compounds to identify hits that interact with the target. **3)** Hits are validated using disease models and relevant cell types. **4)** Promising hits (leads) are optimized and tested for safety. **5)** Leads go to clinical trials and, finally, assessment by the FDA. A typical assessment lasts 10–14 years and costs more than $1 billion.

> **Drug discovery:** *the process of identifying potential new drugs. Target-based drug discovery begins by identifying a gene, mRNA, or protein associated with a disease prior to searching for a drug (in contrast to phenotypic drug discovery, which identifies a therapeutic compound first and then its target).*

> **Target identification:** *the first step of target-based drug discovery in which genes (or gene products) that can serve as possible therapeutic targets for drugs are identified.*

optimizing guide RNA delivery, and adapting the assay to a plate format, among other factors (Fig 2).

Once this step has been completed, a **primary screen** is conducted in which a large set of genes is perturbed with the goal of broadly identifying disease-associated targets. Candidate genes identified in the primary screen then undergo a rigorous process of **target validation** in order to confirm their involvement in the disease of interest. This process may include a variety of *in vitro* and *in vivo* approaches to increase confidence in target-disease relationships.

For instance, a subset of candidates identified in a primary screen are often assessed using a **secondary screen** to confirm that they are associated with a disease. Secondary screens may aim to reproduce results in biologically relevant cell types (iPS cells, primary cells, 3D cultures) or by using different gRNA sequences for each target. Alternatively, one can use an orthogonal method of gene repression. For instance, if RNAi (or CRISPRi) is used in a primary screen and CRISPR-mediated knockouts can be used in a secondary screen, or vice versa.

Other genetic manipulations include knock-in and overexpression experiments that aim to reverse a knockout phenotype. Importantly,

> **Primary screen:** *a screen that interrogates a large number of genes (hundreds to ten thousands) to understand which targets associate with a disease of interest.*

> **Target validation:** *the multifaceted process of confirming the association of gene targets with a disease, and that the modulation of the target will result in a therapeutic effect.*

> **Secondary screen:** *a screen aimed at validating a subset of putative targets identified in a primary screen.*

Ready to Design and Optimize Your Experiment?

Don't miss our Optimizing an Arrayed CRISPR Screen: A Guide to Success Using sgRNA Libraries eBook for more detailed instructions and things to consider on experimental design and optimization.

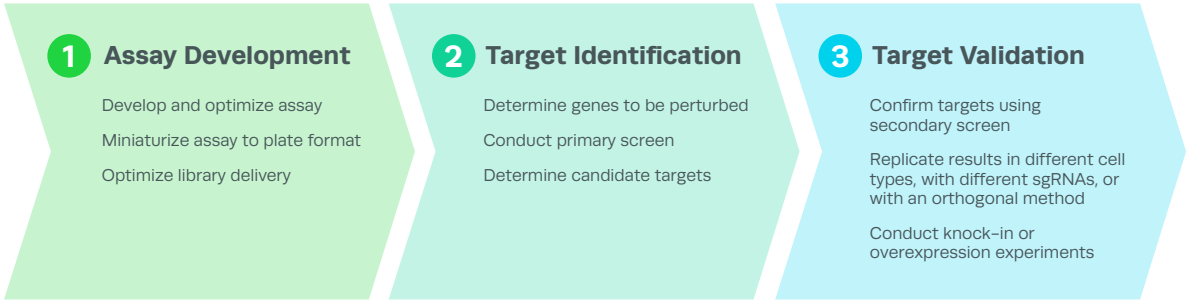


Figure 2. Steps for target identification and validation. 1) Assay Development: prior to a screen, an assay must be developed and optimized. It is also important to optimize the delivery of sgRNA (in your chosen delivery format) to your cells. **2) Target Identification:** a primary screen is conducted using a large gene set and candidate targets are identified. **3) Target Validation:** a subset of targets are validated using a variety of methods, including secondary LOF screens and overexpression experiments. A validated target is then selected for compound screening (Fig 2, step 2).

validated targets must be 1) safe, 2) efficacious, and 3) druggable- meaning that they must be accessible to modulation by pharmacological compounds.

Once a validated target is chosen, the search for a drug begins. High-throughput screens are used to rapidly assess millions of compounds and identify those that interact with the target and elicit a therapeutic effect. Hit compounds are validated in disease models and cell-based assays, and promising hits (leads) are optimized for efficiency and tested for safety. A few qualified leads (~10 of fewer) progress to clinical trials and those that achieve FDA approval become new drugs (Fig 1).

Although the drug discovery process is critical to ensuring the development of safe and efficacious medications, it is notoriously long and expensive, with a high attrition rate. Poor validation of targets is a major reason why drug discovery programs fail. Thus, ensuring that targets truly associate with a disease early in the drug discovery process is crucial for minimizing financial risk and increasing the chance of success.

For the remainder of this eBook, we will explore the various aspects of LOF screens that researchers must consider for successful target identification. First, let's look at some methods of genetic perturbation that are used in screens.

Other Uses of Loss-of-function Screens

In addition to discovering novel drug targets, LOF screens can be used to improve existing therapies. For instance, one may want to identify genes that confer resistance or increase sensitivity to an existing drug. Both types of information can be used to make potent drug combinations that are more effective treatments for a disease of interest. Combinatorial screens, in which multiple genes are knocked out, can uncover genetic interactions that can be leveraged for therapy development. Using these screens helps researchers gain insights on the complexities of different diseases and enable the development of more personalized medicines.

CRISPR vs. RNAi

Screens for target identification require the ability to specifically deactivate a large number of genes in a systematic fashion. Over the past several decades, researchers have relied heavily on **RNA interference (RNAi)** to repress gene function. The strategy involves introducing double-stranded RNA to trigger an innate cellular defense mechanism intended for viral pathogens. After processing, one strand of the RNA is incorporated into an RNA-induced silencing complex (RISC), which then targets and cleaves complementary mRNA species. Depletion of mRNA causes temporary gene silencing (**knockdown**).

Libraries of shRNA or siRNA are often used for large-scale LOF screens and have led to many important discoveries. Because RNAi only reduces (does not eliminate) RNA transcript abundance, scientists can study the effects of perturbing essential genes without killing cells. Also, because RNAi can repress genes in a dosage-dependent manner, it can be used to more realistically mimic the effect of a drug than through complete gene ablation.

Despite these advantages, RNAi-based screens have several drawbacks (see Table 1). For instance, knockdown efficiency can be variable and may produce inadequate signal for some assays, resulting in false negatives. Silencing is also transient, limiting the timeframe that assays can be conducted. Another potential issue is that sh/siRNA have a short region of sequence complementarity, thus often resulting in off-target silencing if sequences match regions of the 3′ UTR of transcripts.[1,2,3] High off-targets can lead to false negatives or false positives. Both of these issues can confound results and thus require extensive validation.

Table 1. Benefits and drawbacks of RNAi and CRISPR.

	RNAi	CRISPR
Benefits	· Pre-designed reagents readily available · Useful for studying the effect of essential genes on phenotypes · Studies where temporary loss-of-function is desired (e.g., to mimic the effect of a drug)	· Precise gene targeting with fewer off-target effects · Permanent gene disruption results in robust signal · Lower risk of immune response (some formats) · Flexible time frame for assay
Drawbacks	· Temporary gene disruption may require a narrow assay window · Incomplete silencing (knockdown) may not produce a strong signal · Associated with more off-target effects · Silencing of multiple transcripts possible (introducing noise) · Introduced RNA may stimulate immune response · Laborious analysis and verification of true hits	· Cannot be used to study essential genes

RNA interferece (RNAi): *a reversible and transient biological process in whereby double-stranded RNA reduces gene expression by cleaving mRNA.*

Knockdown: *a reduction of gene function, with some functional protein still produced.*

CRISPR (clustered regularly interspaced palindromic repeats) technology has emerged as a powerful alternative to RNAi. Instead of repressing gene function at the post-transcriptional level, CRISPR introduces mutations directly to genomic DNA.

The CRISPR system consists of a programmable guide RNA (gRNA) and a Cas9 nuclease (*S. pyogenes*) that together form a ribonucleoprotein (RNP), as shown in Figure 3. The RNP binds to a region of genomic DNA that 1) has a compatible protospacer-adjacent motif (PAM) to Cas9, and 2) has complementary target sequence to the gRNA. If both requirements are met, Cas9 makes a double-strand break (DSB) in the DNA. This action triggers innate cellular repair that can be harnessed for gene editing. Non-homologous end joining (NHEJ), the most common type of DNA repair in mammalian cells, is often exploited to terminate gene function. This mechanism fixes DSBs by ligating the DNA ends together, but occasionally inserts or deletes nucleotides (indels) in the process. Indels that cause a frameshift mutation will disrupt the gene so that it no longer makes a functional protein (**knockout**).

CRISPR-mediated screens have several advantages over RNAi technology (Table 1), including higher consistency and fewer off-target effects.[4,5,6] Also, because CRISPR-Cas9 editing permanently abolishes protein expression, it can produce a stronger phenotypic signal and allow for a longer time frame for analysis. Thus, CRISPR-Cas9 technology is increasingly becoming adopted for functional genomic screens.[7]

CRISPR: *a set of DNA sequences involved in the prokaryotic adaptive immune system that form the basis of the CRISPR genome engineering tool.*

Knockout: *a mutation in a genetic sequence that causes it to be inoperative (i.e., no functional protein is made).*

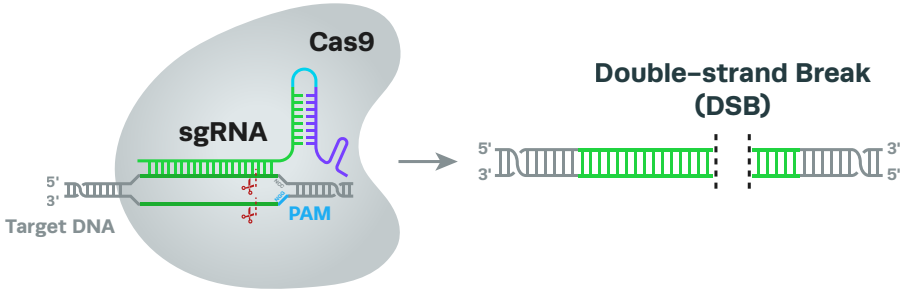


Figure 3. The CRISPR-Cas9 system.
The CRISPR-Cas9 system comprises a guide RNA (sgRNA shown here) and Cas9 nuclease, which together form a ribonucleoprotein (RNP) complex. The presence of a specific protospacer adjacent motif (PAM) in the genomic DNA is required for the gRNA to bind to the target sequence. The Cas9 nuclease then makes a double-strand break (DSB) in the DNA (denoted by the scissors). Non-homologous end joining triggered by the break may result in gene knockout via a frameshift mutation.

What is Single Guide RNA (sgRNA)?

Guide RNAs are composed of a crRNA containing the gene–targeting sequence and a tracrRNA that interacts with Cas9. These components can be annealed together (cr:tracr) or seamlessly linked as single guide RNA (sgRNA). The latter option is popular for screening because it does not require an extra annealing step (saving time for libraries comprising 100s–1000s of gene targets!) and also achieves higher editing efficiencies.

Functional Assays

As mentioned earlier, genetic screens rely on an assay to qualitatively or quantitatively evaluate the effects of knockouts. Each assay should be customized according to the biological question that is addressed. In this section, we will go over two broad categories of assays: binary and multi-parametric (Table 2).

Binary Assays

Binary assays identify cells based on the presence or absence of a desired phenotype. These simple "yes/no" assays can be conducted by either using a selective pressure to kill off a subset of cells (viability assays), or by physically sorting cells based on a biomarker (FACS-based assays). Let's explore these two options below.

Viability Assays

Viability assays separate cells based on whether they live/ proliferate or die/ senesce. These assays apply a positive or negative selection, such as a drug, toxin, or time, to select for a fitness-associated phenotype.

Positive selection enriches cells with knockouts that provide a survival or growth advantage under a certain selective pressure, such as a drug, infection, or adverse condition. The selective pressure kills most of the cells, and the remaining cells are deep-sequenced to identify the gene perturbations associated with survival. These types of screens are often used for identifying mutations that confer drug resistance.

Negative selection selects *against* cell survival. These screens, often referred to as "dropout" screens, identify genes that (when perturbed) die after the pressure is applied. This process is similar to positive selection in that a selective pressure is applied and a subset of cells survive. However, the sequences of surviving cells must be compared to the initial population (if the selective pressure is time) or to a control population in order to identify genes that, when ablated, caused cell death.

Negative selection is often used to identify essential genes [8] and genetic vulnerabilities of diseases. For instance, a negative screen can be used to identify genes that (when ablated) sensitize diseases to an existing drug.[9] Modulating these additional targets through a cocktail of several pharmacological agents can make the original drug more effective. An additional approach that is commonly used in cancer research is to study synthetic lethality: cell death caused by the simultaneous disruption of two (or more) genes. In these screens, gRNAs that are depleted compared to a control reveal potential gene targets for cancer therapies.

FACS–based Assays

Another type of binary assay involves using fluorescence-activated cell sorting (FACS) to physically sort the edited cells into two subpopulations,

Binary assay: *an assay that measures the presence or absence of a single paramter. Cells are catagorized as having a phenotype of interest or not.*

Positive selection: *a strategy that identifies genes that confer a growth or survival advantage in the presence of a selective agent.*

Negative selection: *a strategy that identifies genes that confer a growth or survival disdvantage in the presence of a selective agent.*

those that exhibit desired phenotype and those that do not. The edited cells are first stained with a fluorescent antibody specific for a biomarker of interest and then FACS is used to sort the cells based on the presence or absence of the marker. These assays can be used to investigate the effects of gene knockouts on cell surface markers or reporter proteins.

Multi–parametric Assays

More sophisticated than their binary counterparts, **multi-parametric assays** measure multiple parameters simultaneously. Thanks to advancements in imaging, microscopy, and other technologies, there are a diverse array of assays that can measure everything from morphological features to the location of proteins in the cell. Some assays can even be used to measure markers of cellular processes over time. Below, we discuss some commonly used assay approaches (also listed in Table 2).

High–Content Imaging

One of the major advantages of arrayed screens is that they can be used with high-content imaging, an analysis platform that combines automated microscopy, fluorescent detection, and advanced software algorithms. Screens that leverage this type of analysis, often called high-content screens, can quantitatively assess complex cellular phenotypes. For instance, these screens can measure changes in cell physiology and morphology, even at the subcellular level, by visualizing endogenous tags or fluorescent antibodies. This compelling technology enables researchers to capture a wealth of phenotypic information and, as a consequence, gain more meaningful insights about the effects their manipulations have on cell models.

> **Multi-parametric assay:** *an assay that measures multiple parameters simultaneously or over time interval.*

Protein or Metabolite Abundance

Arrayed screens that interrogate metabolic networks often have fixed endpoint readouts (a single measurement after an incubation period), where the abundance of a given protein or metabolite is assessed. The output is typically measured as a gradient of colorimetric or fluorescent signal, such as in ELISAs, β-gal assays or luciferase assays. Mass spectrometry can also be used to evaluate various molecules and compounds within cells.

Time–course Monitoring

If the aim of a study is to measure biomarkers/ morphology over a time interval, arrayed screens enable the collection of data over the course of a screen. These screens are used for studies aimed at identifying genes involved in cellular processes such as cell differentiation, migration, and phagocytosis.

Table 2. Examples of binary and multi-parametric assays.

Binary Assays	Multi–Parametric Assays
Viability screen • Negative selection • Positive selection	**High–content imaging** • Cell morphology • Cellular uptake • Protein localization
FACS–based screen • Expression of a surface marker • Expression of a reporter protein	**Protein/ metabolite abundance** • ELISA • Luciferase • B-gal
	Time–course monitoring of cellular processes • Differentiation • Migration • Phagocytosis

Pooled vs. Arrayed Workflows

Now that we have explored different types of functional assays, let's look at how they are applied to CRISPR screening workflows. There are two types of CRISPR screening formats: pooled and arrayed. While both can be used for target identification and validation, they differ in methodology, equipment, and assay compatibility. **Pooled screens** involve delivering a mixed population of sgRNA-containing viral constructs into a single tube of cells, while **arrayed screens** target each gene separately with sgRNA across a multiwell plate. Below, we provide an overview of each.

Pooled Screens

Pooled screens involve introducing a "pool" of sgRNAs into a single population of cells. This is accomplished by packaging sgRNA-containing plasmids into lentiviral particles (one per vector), and then transducing host cells. The stable expression of guide (along with Cas9) facilitates the knockouts of targeted genes. Because knockouts of all the targets occur in a single tube of cells, it is challenging to link the phenotype of each individual cell with the underlying genetic perturbation. Pooled screens are thus only compatible with binary assays that physically separate edited cells exhibiting a phenotype of interest from those that do not. After the cells are sorted, the integrated sgRNAs are deep sequenced and data must be deconvoluted. Enriched or depleted sgRNAs in the population confer information about the involvement of corresponding genes with the phenotype. See Figure 4 for the steps of a pooled screen that utilizes a positive/negative viability assay.

> **Pooled CRISPR screen:** *a screen that involves delivering a pool of sgRNAs to cells in a single tube via lentiviral transduction.*

> **Arrayed CRISPR screen:** *a screen that invovles targeting each gene separately in a multiwell plate format (1 gene per well).*

Arrayed Screens

Arrayed screening is a newer technology that is more versatile in both methodology and analysis than its pooled counterpart. Arrayed screens involve targeting one gene per well in a multiwell plate format. Library delivery may be accomplished through transient transfection or lentiviral transduction. Because gene targets are separated across wells, phenotypes do not need to be selected for and sequencing/ data deconvolution are not required to associate phenotypes with genotypes. Arrayed screens are compatible with both binary and multiparametric assays. See Figure 5 for a step-by-step description of an arrayed screen workflow.

Workflow for Pooled Screens

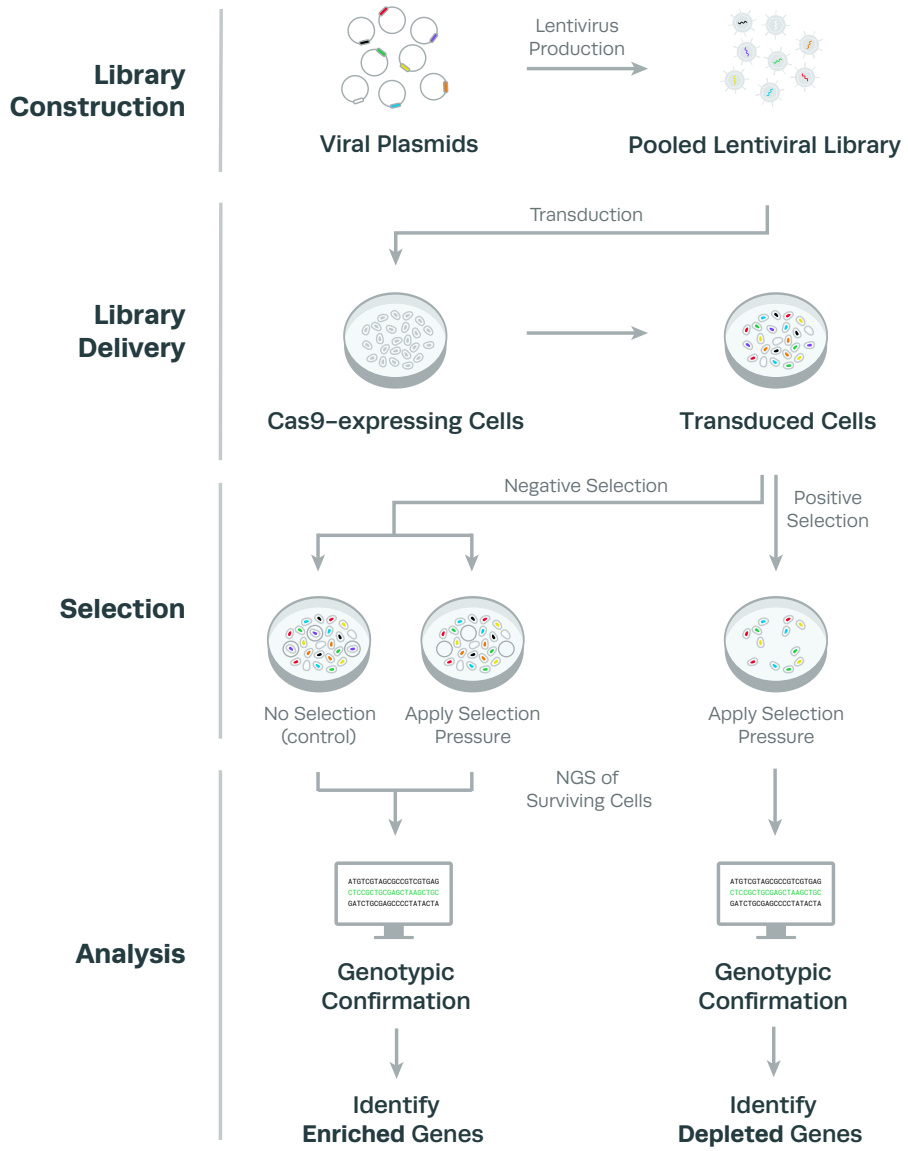


Figure 4. Steps of a pooled CRISPR screen.

1. **Library Construction:** Plasmids encoding sgRNAs are PCR-amplified and validated via NGS to ensure that equal representation is maintained. The plasmids are then packaged into lentiviral particles (one guide per vector) containing a selectable marker (e.g., antibiotic resistance gene). Typically, more than one sgRNA is designed per gene to increase confidence in genotype to phenotype correlations. Libraries can also be purchased as pre-packaged viral particles.
2. **Library Delivery:** The viral particles are then introduced to a single group of cells at low multiplicity of infection (MOI) to ensure that, on average, only one viral particle will enter each cell and insert the sgRNA sequence into the genome. Cas9 is expressed by using a Cas9-expressing cell line or through co-transduction. The population is then enriched for transduced cells (antibiotic selection) and subsequently expanded. Cells that proliferate pass on the sgRNA sequence to successive generations.
3. **Selection:** A positive or negative selective pressure is applied to select for the desired viability phenotype. Positive selection identifies knockouts that provide a growth advantage in the presence of the selective agent, whereas negative selection identifies knockouts that confer a survival disadvantage in the presence of the selective agent.
4. **Analysis:** The frequency of each sgRNA in the cell population is measured via next-generation sequencing (NGS). The enrichment or depletion of particular sgRNAs following selection iimplicates gene perturbations that desensitize or sensitize cells to the selective agent. Note that the identity and relative abundance of sgRNAs in the manipulated cell population is the readout and not the edits made to the targeted genes.

Workflow for Arrayed Screens

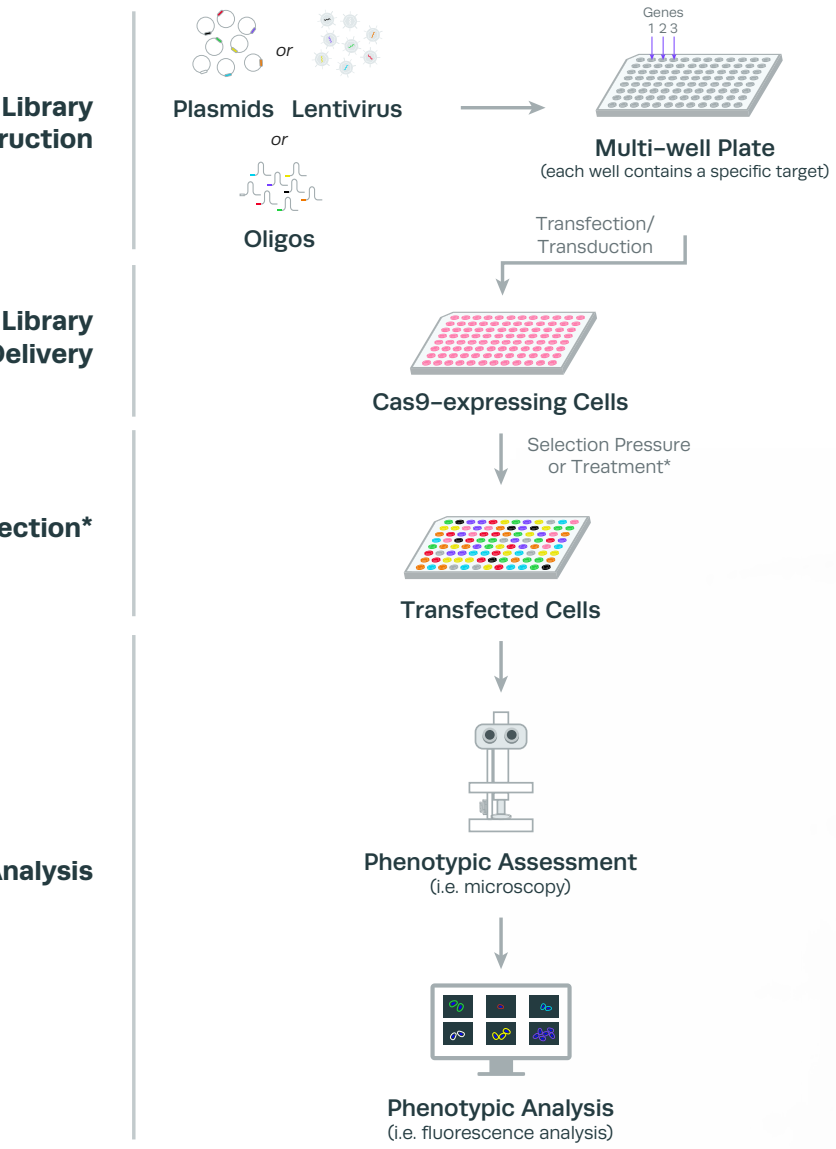


Figure 5: Steps of an arrayed CRISPR screen.

1. **Library Construction:** sgRNA libraries can be produced in a variety of formats: plasmid, lentivirus, or RNA oligonucleotides (synthetic sgRNA).
2. **Library Delivery:** The sgRNA is introduced to cells (one target per well) through one of a variety of methods. Cas9 is either co-transfected as plasmid or protein (RNP), or a Cas9-expressing cell line is used.
3. **Selection (*optional):** A selective pressure, may be applied to arrayed screens in order to test what genes, when ablated, affect one or more cellular phenotypes when under certain contexts.
4. **Analysis:** Edited cells are assessed using an appropriate functional assay. Because each target is separated across wells, direct genotype-phenotype associations can be made.

When to Use a Pooled or Arrayed Screen?

Choosing a screening format based on a number of considerations, including the assay, cell model, and labor, among others. Below, we outline some of the advantages and disadvantages of each (also summarized in Table 3).

Assay

As described above, pooled screens are restricted to binary assays, with simple readouts, such as a cell survival/death or sortable biomarkers. Arrayed screens are compatible with binary and multiparametric assays, including those that assess complex phenotypes. Given these differences in versatility, it is important to consider what phenotype(s) would be most informative for answering your research question. For instance, if identifying genes that sensitize/desensitize a disease cell type to a given drug is the goal, then a pooled screen may be adequate. However, if the objective is to identify gene disruptions that cause changes in multiple morphological features, then an arrayed screen may be more appropriate.

Cell Model

The type of screen one chooses also depends on the cell model used. Appropriate cell models often depend on the stage of the screening workflow and disease of interest. For instance, whereas an immortalized cell line may be sufficient for primary screens, a more clinically relevant cell type may be preferred for secondary screens. Because pooled screens require the integration of sgRNA into the genome and passage to daughter cells, they are most appropriate for use with actively-dividing cells. Pooled screens are thus not well-suited for primary cells and neurons, which have a limited capacity to proliferate. Alternatively, arrayed screens do not require an extended period of expansion and can be used with a wide variety of cell types.

Table 3. Benefits and drawbacks of different screening formats.

	Pooled Screen	Arrayed Screen
Benefits	· Cost-effective · High transduction efficiency (lentivirus) in most cell types · Suitable for viability assays · Easy to interrogate entire genome · Does not require special equipment	· Clear genotype–phenotype correlation, no data deconvolution · Can be used for high-content screening/ sophisticated assays · Multi-parametric readouts possible · Suitable for primary cells and neurons
Drawbacks	· Requires a selection step · Requires NGS · Data deconvolution needed to identify genotype–phenotype correlations · Restricted to binary assays · Not suitable for non-dividing cells (e.g., primary cells) · Library preparation and viral packaging is time consuming · Biosafety requirements due to working with virus	· Higher upfront cost (though can save money in the long run) · Requires special equipment and some automation · Requires optimization to identify optimal transfection conditions for editing

Time and Labor

Time and labor is another important factor. Pooled screens require a considerable amount of upfront work if viral packaging is done oneself. Also, because cells with different gene knockouts are mixed together in pooled screens, data deconvolution is necessary to untangle the genotype-phenotype relationships. Alternatively, arrayed screens have options that can save time and labor. For instance, some library formats (e.g., synthetic sgRNA) require little preparation and analysis can be relatively straightforward. However, optimization is required for these screens in order to ensure high editing efficiencies.

Equipment

Pooled and arrayed screens require different equipment. Whereas pooled screens can be conducted with a standard laboratory setup, arrayed screens may require specialized equipment and automation capabilities. In addition, a high-content system may be necessary to analyze image-based phenotypes.

Financial Cost

Pooled screens are relatively cost-effective to run, and thus offer a feasible way to interrogate the genome. Arrayed screens have a higher upfront cost, but can provide an abundance of high-quality information. Thus, these screens can ultimately save money in the long run if researchers do not need to conduct as many follow-up studies.

Thinking Big Picture: Leveraging a Combination of Screen Types

It is important to remember that pooled and arrayed screens can both be useful in screening workflows. For instance, if one aims to identify new drug targets, a pooled format may be appropriate as a primary screen to identify a broad set of target genes in an easy-to-transfect cell model (e.g., immortalized cell line). An arrayed format may then be used in a secondary screen to validate the hits using a more realistic model (e.g., primary cells). When designing a target identification experiment, consider all the tools at your disposal.

CRISPR Reagents and Delivery Formats for Arrayed Screens

All CRISPR screens involve the introduction of sgRNAs and Cas9 into cells. Whereas pooled screens always deliver sgRNA libraries using viral transduction, arrayed screens can deliver CRISPR components in a variety of ways. Let's take a look at the different options (Table 4).

Guide RNA Format

DNA (plasmid)

One way to deliver sgRNA libraries is through expression plasmids. A main benefit of this format is the large packaging capacity, such that multiple sgRNAs targeting a single gene can be cloned into one vector. Additionally, selection markers can be included in the backbone so that transfected cells can be easily isolated. However, the plasmid format has several drawbacks that should be taken into consideration. For instance, plasmids can remain active in cells for several weeks and thus can increase off-target effects.[10,11] Unintended integration events into the genome can also disrupt surrounding genes. Lastly, plasmids have low transfection efficiency in some cell types, including primary cells.

Lentivirus

Relative to plasmids, virus is a more popular method of library delivery for CRISPR screens. As in pooled screens, lentivirus can be used to stably transduce cells for arrayed screens. An advantage of lentivirus is its efficiency at delivering sgRNA libraries to a variety of cell types. Like plasmids, selectable markers can be used to isolate transduced cells. However, conducting large-scale arrayed screens using lentivirus is challenging, as vectors for hundreds of targets must be packaged separately. In addition, these screens are prone to cross contamination and often have high well-to-well titer variability. Moreover, continuous sgRNA expression can increase off-target effects, and random genomic integration can disrupt essential genes. Lastly, because additional time is required for selection of transduced cells and for transcription of sgRNA, the time to assay may be long. This presents a challenge for working in multiwell plates that are not conducive to long culture periods.

Adeno-associated viruses (AAV) are another option for transgene delivery. These vectors usually remain in an extrachromosomal state, but have limited cargo capacity (~ 4.5 kb). Working with either virus requires extra safety requirements, including a BSL2 facility with appropriate training and infrastructure.

Synthetic sgRNA

Single guide RNA can be produced at scale through synthetic manufacturing. A main benefit of synthetic production (as opposed to *in vitro* transcription) is that guides can be chemically modified to increase stability and prevent the triggering of innate immune responses within host cells.[12] These modified sgRNAs have higher editing efficiencies than unmodified guides in a variety of cell types, including primary cells. Another benefit of the RNA format is that assays can be performed relatively quickly (sometimes only days post-transfection) compared to viral and plasmid formats. A drawback of synthetic guides is that transfection optimization is required to ensure high editing efficiencies.

Table 4. Benefits and drawbacks of guide RNA formats.

	Plasmid	Lentivirus	Synthetic gRNA
Benefits	· Cost-effective · High packaging capacity	· Efficient delivery in a variety of cell types	· High editing efficiency · Chemical modifications · Reduced off-target risk
Drawbacks	· Risk of genomic integration · Increased off-targets risks · Low efficiency in primary cells	· Extra safety requirements · Increased off-target risks · Laborious preparation & limited scalability in arrayed format	· Transfection requires optimization

Cas9 Format

Cas9-expressing cell lines are often preferred for arrayed screens, as it simplifies the workflow (i.e., only the sgRNA library needs to be delivered). However, because primary cells have limited capabilities for expansion, immortalized cells and iPS cells can only be used to develop stable Cas9-expressing cell lines.

Alternatively, Cas9 can be co-transfected in a plasmid format. However, plasmid transfection risks cytotoxicity in sensitive cells and requires more time between transfection and assay (to allow transcription and translation of Cas9).

Lastly, Cas9 can be introduced as pre-complexed ribonucleoproteins (RNPs). One major advantage of this format is that editing is limited to a

short time frame (reducing the risk of off-target effects). Also, RNPs yield high editing efficiencies across a variety of cell types, including primary cells, iPS cells, and immortalized cells.[10,11,12]

Transfection Method

Arrayed screens are compatible with viral transduction, as well as several non-viral transfection methods. Lipid-based transfection (lipofection) uses cationic lipid reagents to deliver CRISPR components (e,g, plasmid, RNPs) into cells. The process first involves constructing lipid-soluble structures, called liposomes, around the components. The components are then transported into the cell through endocytosis, a process in which the plasma membrane surrounds the components on the exterior side, and buds off inside the cell. The CRISPR components then escape the endosomal pathway and diffuse through the cytoplasm. Lipid-based transfection is cost-effective, does not require special equipment, and is well-suited for high-throughput workflows. While this technique is commonly used for screening immortalized cells, it is not recommended for primary or stem cells.

Electroporation is another popular transfection method for high-throughput screening. This technique involves suspending cells in a conductive solution and briefly applying high-voltage electrical pulses. The application of electricity induces the formation of temporary pores in the plasma membrane and the electrical potential across the membrane causes charged molecules (e.g., sgRNA/RNPs) to enter the cytoplasm through the pores. Transfection is often highly efficient and is appropriate for a variety of cell types (including primary and stem cells). However, it does require access to specialized equipment, such as an electroporator or Nucleofector™.

Still Have Questions About Transfection?

Read more about guide transfection methods for arrayed screens in our Optimizing an Arrayed CRISPR Screen: A Guide to Success Using sgRNA Libraries eBook.

Library Design

Guide RNA Design

Guide design can have a profound effect on the outcome of a screen. For instance, sgRNA libraries that do not robustly knock out targets may not produce enough detectable signal, and hence lead to false negatives. Additionally, sgRNAs that have off-target editing can introduce noise into the system and ultimately complicate the analysis of results.

Several factors can be taken into account to increase the robustness and specificity of sgRNA libraries. For instance, it is recommended that guides target early exons of protein-coding genes (to ensure protein function is completely terminated) and that guide sequences be evaluated *in silico* to minimize off-target editing. Also, multiple guides may be designed per target to increase the likelihood of a knockout.

Guides may also be designed to strategically bias a type of genomic disruption. Most libraries are designed to depend on the generation of traditional frameshift-induced indels. However, because some indels are a multiple of three nucleotides (and do not shift the reading frame), the associated proteins may retain their function. Thus, knockout efficiencies may vary widely.

Synthego has addressed this issue by designing libraries that comprise **multi-guide sgRNA**, in which each gene is strategically targeted by up to three guides. The guides are designed to concurrently cut a target locus in an early exon, causing one or more **fragment deletions** (Fig 6). This strategy results in higher knockout efficiencies (and lower variance) than when the same guides are introduced individually (Fig 7).

Now that we have outlined some factors of library design, let's look at various libraries that are available for purchase.

Types of Libraries

Depending on the goal, researchers may want to target a large collection of genes that span the genome, or just or a small subset of candidates. To meet these needs, several vendors offer libraries in standard and custom formats.

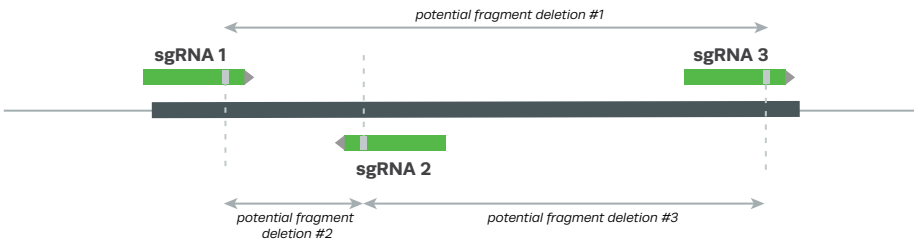


Figure 6. Synthego's multi-guide design.
Synthego's multi-guide sgRNA design includes up to 3 modified sgRNAs (green bars) that target each gene. When co-transfected, the sgRNAs create concurrent double-strand breaks (vertical dotted lines) at the targeted genomic locus and consequently induce one or more fragment deletions that effectively terminate gene function. The above diagram depicts 3 possible deletions.

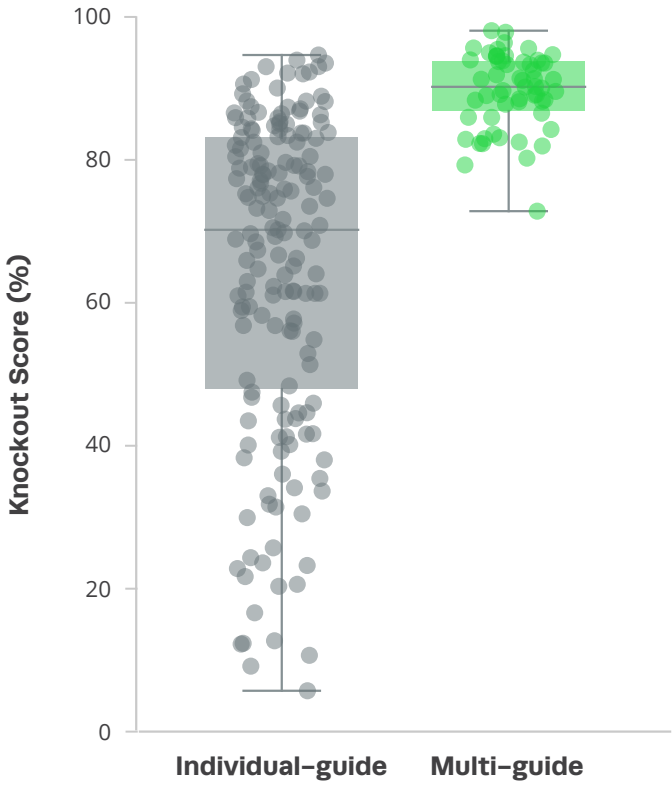


Figure 7. Robust knockouts with multi-guide sgRNA.
Across 32 genetic targets, 96 sgRNAs were designed (three guides targeting each gene) and transfected as multi-guide (3 guides/ gene) and individual-guide (1 guide/ gene) formats. Multi-guide sgRNA had a 29.2% better median knockout efficiency and lower variance (89.9% ± 1.40 95% CI; green) relative to guides transfected individually (69.6% ± 3.48; gray). Knockout Score: the percentage of sequences that result in a putative knockout (i.e., fragment deletions & frameshift-inducing indels).

Multi-guide sgRNA: *up to three sgRNAs designed to concurrently cut a target gene, causing one or more fragment deletions.*

Fragment deletion: *a 21+ bp deletion induced by multi-guide sgRNA.*

Standard Libraries

Standard libraries are pre-made “off the shelf” collections of sgRNA. These libraries are often used for primary screens that aim to assess a large number of genes for their involvement in a disease or pathway. The most comprehensive standard libraries are updated frequently to account for new genes, gene reclassifications, and SNV re-definitions. Synthego currently offers 30 standard libraries, including the whole human (19,753 genes), druggable, deubiquitinating enzymes (DUBS), GPCRs, kinases, and immuno-oncology targets, among others. All of these libraries are curated using multiple gene ontology and drug databases.

Custom Libraries

Custom libraries are subsets of gene targets that are chosen by the researcher. These libraries are often used in secondary screens to validate gene target hits from primary screens. Small custom libraries are also used in assay development to optimize conditions and timing, as well as to miniaturize the assay to a plate format.

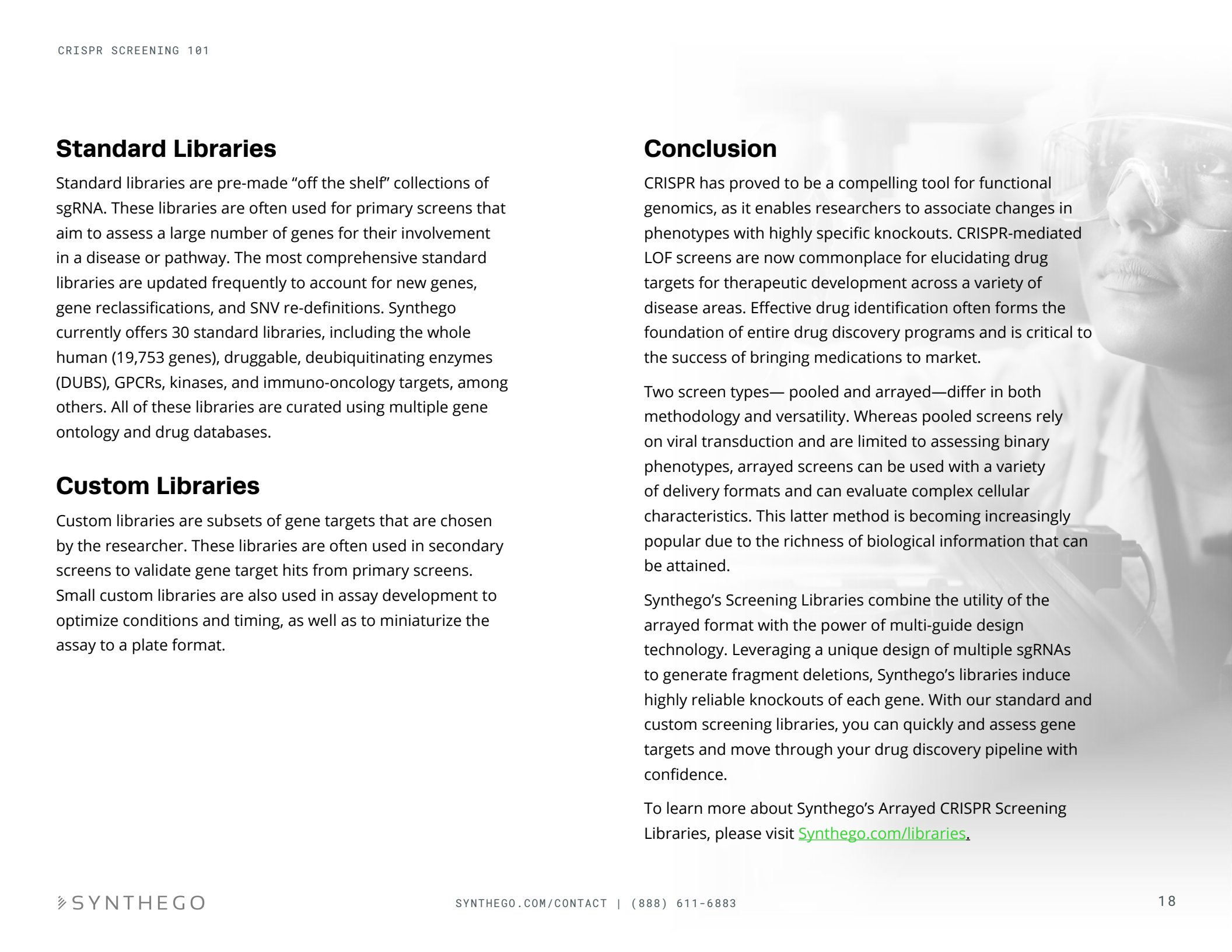
Conclusion

CRISPR has proved to be a compelling tool for functional genomics, as it enables researchers to associate changes in phenotypes with highly specific knockouts. CRISPR-mediated LOF screens are now commonplace for elucidating drug targets for therapeutic development across a variety of disease areas. Effective drug identification often forms the foundation of entire drug discovery programs and is critical to the success of bringing medications to market.

Two screen types— pooled and arrayed—differ in both methodology and versatility. Whereas pooled screens rely on viral transduction and are limited to assessing binary phenotypes, arrayed screens can be used with a variety of delivery formats and can evaluate complex cellular characteristics. This latter method is becoming increasingly popular due to the richness of biological information that can be attained.

Synthego’s Screening Libraries combine the utility of the arrayed format with the power of multi-guide design technology. Leveraging a unique design of multiple sgRNAs to generate fragment deletions, Synthego’s libraries induce highly reliable knockouts of each gene. With our standard and custom screening libraries, you can quickly and assess gene targets and move through your drug discovery pipeline with confidence.

To learn more about Synthego’s Arrayed CRISPR Screening Libraries, please visit Synthego.com/libraries.



More Screening Solutions

Synthego offers libraries of CRISPR-edited cell pools in arrayed format, called Engineered Cell Libraries. These cell pools are readily available for downstream functional assays, thereby bypassing transfection optimization, and operational logistics around handling and storage. To learn more about Synthego's Engineered Cell Libraries, please visit our website at Synthego.com/products/engineered-cells/engineered-cell-libraries.

Learn how scientists are using Engineered Cell Libraries to rapidly validate targets for drug discovery. Read our Accelerating Target Validation for COVID-19 Therapeutics using Engineered Cell Libraries Case Study.

Conclusion

CRISPR has proved to be a compelling tool for functional genomics, as it enables researchers to associate changes in phenotypes with highly specific knockouts. CRISPR-mediated LOF screens are now commonplace for elucidating drug targets for therapeutic development across a variety of disease areas. Effective drug identification often forms the foundation of entire drug discovery programs and is critical to the success of bringing medications to market.

Two screen types— pooled and arrayed—differ in both methodology and versatility. Whereas pooled screens rely on viral transduction and are limited to assessing binary phenotypes, arrayed screens can be used with a variety of delivery formats and can evaluate complex cellular characteristics. This latter method is becoming increasingly popular due to the richness of biological information that can be attained.

Synthego's Screening Libraries combine the utility of the arrayed format with the power of multi-guide

Additional Information

For an up-to-date list of all Synthego Application Notes and other resources, please visit: Synthego.com/resources

For technical assistance, contact our Scientific Support Team:
Phone: 1-844-4-CRISPR
Email: support@synthego.com

About Synthego

Synthego is a genome engineering company that enables the acceleration of life science research and development in the pursuit of improved human health. With its foundations in engineering, the company leverages machine learning, automation, and gene editing to build platforms to advance both basic research and therapeutic development programs.