Published in final edited form as:

Trends Cancer. 2016 June ; 2(6): 313–324. doi:10.1016/j.trecan.2016.05.001.

CRISPR/Cas9: From Genome Engineering to Cancer Drug Discovery

Ji Luo

Laboratory of Cancer Biology and Genomics, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, 20814, USA

Abstract

Advances in translational research are often driven by new technologies. The advent of microarrays, next-generation sequencing, proteomics and RNA interference (RNAi) have led to breakthroughs in our understanding of the mechanisms of cancer and the discovery of new cancer drug targets. The discovery of the bacterial clustered regularly interspaced palindromic repeat (CRISPR) system and its subsequent adaptation as a tool for mammalian genome engineering has opened up new avenues for functional genomics studies. This review will focus on the utility of CRISPR in the context of cancer drug target discovery.

The CRISPR/Cas9 Endonuclease System

In the post-genomic era of cancer drug discovery, a major challenge is to convert our increasingly detailed knowledge of the genetic and epigenetic alterations in the cancer genome into understanding about the functional vulnerabilities in cancer cells [1]. Functionalizing the cancer genome is critical for identifying appropriate oncogene and non-oncogene targets that could afford therapeutic benefit in cancer patients [2-4]. The oncogenic activity of driver mutations can be modeled through gene targeting using Cre and FLP recombinases as well as gene editing using zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) [5]. The discovery of RNA interference (RNAi) and its implementation as a loss-of-function genomic screening tool has dramatically accelerated our ability to interrogate the cancer genome for functional dependencies [6]. Recently, a new and powerful genome editing technology based on the bacterial clustered regularly interspaced palindromic repeat (CRISPR) endonuclease system has been discovered. CRISPR is a versatile platform that enables us to knockout, activate and introduce precise mutations in genes. Its deployment in cancer research both as a genome editing tool and as a genome screening tool will dramatically accelerate the pace of cancer target discovery and target validation.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

CRISPR/Cas9 in bacteria and its adaptation for mammalian genome editing

CRISPR is a RNA-guided nuclease system that bacteria use to protect against phage infection (Box 1) [7]. The CRISPR complex consists of two modules: a CRISPR associated (Cas) endonuclease module that introduces double-stranded DNA breaks, and a CRISPR RNA (crRNA) module that specifies the target DNA sequence. Functional analysis of the *S. pyrogenes* type II CRISPR system revealed that three components are sufficient to constitute CRISPR activity: the Cas9 endonuclease, a target specific crRNA, and a structural trans-activating CRISPR RNA (tracrRNA) [8]. This system can be further simplified by fusing the crRNA and tracrRNA to form a single guide RNA (**sgRNA**). In mammalian cells, the co-expression of Cas9 and sgRNA is sufficient to induce sequence-specific DNA cuts [9-11]. Because the *S. pyrogenes* CRISPR/Cas9 system is well characterized and simple to deploy, it is currently the most widely used CRISPR system for mammalian genome engineering.

Biochemical and structural studies have revealed the mechanism by which the Cas9/sgRNA complex binds to and cuts target DNA [12-16]. This complex recognizes a 20 nt DNA sequence that is complementary to the crRNA and upstream of a “NGG” protospacer adjacent motif (PAM). Binding between Cas9 and PAM is essential for the initiation of target recognition. Cas9 unwinds the DNA duplex upstream of the PAM to allow strand invasion and target interrogation by the crRNA 20-mer. Cas9 contains a RuvC and a HNH endonuclease domain, and duplex formation between crRNA and target DNA stimulates its nuclease activity to generate a blunt-end, double-stranded DNA break 3 nt upstream of the PAM (Figure 1A). In mammalian cells, this break can be repaired by two endogenous DNA repair pathways: **non-homologous end-joining** (NHEJ) and **homology-directed repair** (HDR). NHEJ repair is an imprecise mechanism and it often introduces small insertion and deletion (indel) mutations. This renders the target site no longer recognizable by the Cas9/sgRNA complex. HDR repair is error-free and it requires a homologous DNA template for repair. Both NHEJ and HDR have been exploited for CRISPR/Cas9-mediated genome engineering [7,17].

The CRISPR/Cas9 tool kit for genome engineering

The simplicity and modular nature of CRISPR/Cas9 makes it an ideal tool for genome engineering [7,17-19]. The most salient feature of CRISPR/Cas9 is its modular design. Because the targeting module (sgRNA) and the endonuclease module (Cas9) are encoded separately, each can be modified, evolved and optimized without affecting the function of the other. The short PAM requirement means that virtually all genomic loci are targetable with CRISPR/Cas9.

One major application for CRISPR/Cas9 in mammalian cells is the regulation of gene expression [9,10]. CRISPR/Cas9 is highly effective for gene knockout in mammalian cells when Cas9 is targeted to the exon regions of a gene. In this case, cutting by Cas9 and subsequent repair by NHEJ results in indel mutations that often cause frame-shift in the target gene. This leads to either the production of a non-functional, truncated protein or the degradation of mutant mRNA through **non-sense mediated mRNA decay**. Whereas **RNA interference** (RNAi) rarely achieves complete silencing, CRISPR/Cas9 can generate true nulls.

An alternative approach for CRISPR-mediated gene silencing, known as CRISPR interference (CRISPRi), takes advantage of the high binding affinity of Cas9 for its target sequence. Mutating both of the nuclease domains of Cas9 results in a nuclease-dead protein (dCas9) that binds to target DNA without cleaving it [20,21]. When targeted to the transcription start site of a gene, dCas9 can suppress its expression by interfering with the transcription machinery (Figure 1B). Gene silencing with dCas9 can be further enhanced by fusing dCas9 with a transcriptional repressor module such as the **Krüppel associated box (KRAB) domain** of Kox1 protein (Figure 1B). This Cas9-KRAB fusion protein can recruit chromatin modifying factors to the target site to achieve sustained gene silencing [21]. Similar to RNAi, CRISPRi affords gene suppression but not necessarily complete silencing. However, unlike wild type (WT) Cas9, dCas9-based CRISPRi does not permanently alter the target gene's sequence and its effect is therefore reversible.

The property of dCas9 as a sequence-specific DNA binding module has also been exploited for gene activation. This approach, known as CRISPR activator (CRISPRa), utilizes either a fusion between dCas9 and a transcriptional activator module, a stem-loop modified sgRNA capable of recruiting transcriptional factors, or a combination of both to activate gene expression (Figure 1C) [21-24]. When targeted to the promoter region of a gene, CRISPRa promotes the assembly of transcription factors to drive robust gene over-expression. Impressively, CRISPRi and CRISPRa can be deployed in the same cell to achieve combinatorial gene regulation [25].

Another key application of CRISPR/Cas9 is precision genome engineering. By supplying the cell with a DNA template for HDR following Cas9 cleavage, virtually all forms of sequence alterations can be precisely engineered. These include point mutation, small indel, large deletion, and chromosome rearrangement [17]. In the context of cancer biology, this approach is particularly useful for the construction of cell line and animal models with complex genetic alterations that mimic the mutational landscape of human tumors.

CRISPR Screens for Cancer Drug Target Discovery

High throughput genetic screening is a powerful tool for target discovery: it can be used to systematically identify genes that support cancer cell viability and regulate drug sensitivity. Genetic screens can also be used to discover new members of a pathway or new interactions between pathways. Such information can provide valuable mechanistic insights for the development of new therapeutic approaches.

Limitations with RNAi and cDNA screens

Prior to the discovery of CRISPR, high throughput loss-of-function (LOF) and gain-of-function (GOF) screens in mammalian cells were made possible by RNAi-based gene knockdown libraries and cDNA-based gene overexpression libraries, respectively. Methodologies for well-by-well siRNA library screen and **pooled shRNA library** screen have been well established, although two technical issues remain. The first is that RNAi often gives incomplete gene knockdown. As a result, this could give rise to a high **false negative rate** due to insufficient gene silencing. The second issue is that RNAi has a significant off-target effect, and this could lead to a high **false positive rate**. These technical

issues can be mitigated by pre-validating shRNAs/siRNAs and by the use of ultra-deep shRNA libraries [26-28]. However, these approaches are associated with a significant increase in the cost and/or size of the library. Alternatively, LOF screens have been carried out with lentiviral insertion-mediated gene silencing [29], but this method is limited to rare cell lines with a haploid genome.

Both cDNA and ORF overexpression libraries have been successfully deployed for high-throughput GOF screens [30]. However, these libraries are time consuming to construct as cDNAs and ORFs vary greatly in size and each must be cloned individually. In addition, these libraries are not truly genome wide because some cDNAs are simply too large for cloning and expression with existing library vectors.

Because sgRNA sequences are short and uniform in size, they are amenable for high throughput, array-based oligonucleotide synthesis. Pooled CRISPR libraries targeting the entire genome can be readily generated [31,32]. The availability of CRISPR knockout, CRISPRi and CRISPRa libraries have enabled both LOF and GOF screen at the genome scale. In the context of cancer biology, the principle of screening pooled CRISPR libraries is similar to that of pooled shRNA libraries (Box 2) [18]. For example, in the most common assay using cell viability as readout, individual sgRNAs that positively or negatively regulate cell proliferation are identified by comparing sgRNA library composition before and after a defined number of cell doublings using deep-sequencing. This approach can be extended to identify genes that modulate drug sensitivity (Figure 2).

Loss-of-function CRISPR library screens

Several recent studies have demonstrated the power of LOF CRISPR library screens for both positive and negative selection assays [31-37]. In these studies, both CRISPR knockout and CRISPRi approaches have been employed to screen very large sgRNA library pools. These early screens have already yielded impressive results. Negative selection screens in cancer cells have been successfully deployed to identify both essential genes that are required for cell viability in general and cancer-lethal genes that are selectively required for cancer cell viability in a context-dependent manner. Among cell essential genes are those known to play a role in critical cellular processes including DNA replication, RNA processing and proteolysis. These genes encode components of core protein complexes such as the ribosome, spliceosome, kinetochore and proteasome [31-33,36,37]. Interestingly, hundreds of genes with little functional information have been found to be cell essential [37]. These studies demonstrated CRISPR as a powerful discovery platform for the rapid annotation of gene function and revealed significant gaps in our knowledge about fundamental cellular processes. Establishing a comprehensive list of cell-essential genes is valuable for target discovery in two ways. First, this reference gene set is useful for evaluating the quality of future LOF lethality screens, as the majority of these genes should be reproducibly lethal across most cell lines. Second, target discovery effort that avoids cell-essential genes might lead to better therapeutic window and lower on-target toxicity in normal tissues.

Among the cancer lethal genes discovered by CRISPR screen are those involved in signaling, differentiation, survival and regulatory processes [35-37]. These genes represent functional dependencies in individual cancer cell lines that are exploitable as potential drug

targets. Known oncogene additions have been reliably uncovered in these screens. For example, *BCR* and *ABL* are lethal hits in the chronic myelogenous leukemia cell line KBM7 that harbors *BCR-ABL* translocation [31]; and *KRAS* and *PIK3CA* are lethal hits in the colorectal cancer cell lines DLD-1 and HCT116, both harbor mutations in these genes [36]. In addition, CRISPR screen has revealed unexpected genetic dependencies in cancer cells. For example, the Burkitt's lymphoma cell line Raji is selectively dependent on the DEAD-box helicase *DDX3Y* located on the Y-chromosome due to the loss of its paralog *DDX3X* on the X-chromosome [37]. The *KRAS* mutant colorectal cancer cell lines DLD-1 and HCT116 are sensitive to the loss of genes encoding subunits of the mitochondrial ribosome, indicating their heightened dependency on mitochondria function [36]. Thus, future expansion of cancer lethal CRISPR screen to a large number of cancer cell lines [38], when coupled with their genomic and drug sensitivity information [39,40], could provide a comprehensive view of the target landscape in cancer cells.

Positive selection screen using LOF CRISPR libraries has been fruitful in identifying genes involved in cancer drug resistance [31,32]. For example, a screen against the antimetabolite 6-thioguanine in KBM7 cells recovered genes in the DNA mismatch repair pathway that constitute known mechanisms of resistance. Similarly, a resistance screen against the topoisomerase IIa inhibitor etoposide in the promyelocytic leukemia cell line HL60 recovered the drug target *TOP2A* itself as well as the G1 cell cycle kinase *CDK6* [31]. CRISPR screens using the BRAF inhibitor vemurafenib in the melanoma cell line A375 have revealed new mechanisms of resistance to this drug. It was found that deletion of the tumor suppressor *NF2*, the cullin E3 ligase *CUL3* and several members of the STAGA histone acetyltransferase complex could all confer vemurafenib resistance [32,37]. These unexpected findings, if validated in patients, could guide the design of future therapies in BRAF mutant melanomas.

Gain-of-function CRISPR library screens

GOF screen using CRISPRa library is a convenient alternative approach to cDNA based overexpression library [24,33]. CRISPRa has been used to identify genes that both positively and negatively regulate cancer cell proliferation. It was shown that proliferation of the CML cell line K562 could be inhibited by CRISPRa-mediated activation of tumor suppressor genes and genes involved in lineage differentiation [33]. This finding suggests that in cancer cells, anti-proliferative genes could remain functional despite being down-regulated, and drugs that restore their expression might be useful at halting tumor growth. CRISPRa screen has also been applied to identify genes whose over-expression confers resistance to vemurafenib in A375 cells. This study found that BRAF inhibition can be bypassed through the over-expression of genes that either re-activate the MAPK pathway through BRAF-independent mechanisms or activate parallel survival pathways independent of the MAPK pathway [24]. Thus co-targeting of these resistance mechanisms might help preventing the rapid acquisition of BRAF inhibitor resistance. By employing both LOF (CRISPR) and GOF (CRISPRa) libraries in the same assay such as the BRAF inhibitor assay, CRISPR screen can discovery both positive and negative regulators of a pathway and thus provide better hit saturation and deeper biological insight into the question at hand.

Advantages and limitations of CRISPR libraries

Early experience with CRISPR screen has revealed several advantages over shRNA screens. A comparison of CRISPR and shRNA results from similar screens indicates that CRISPR library could be more robust than shRNA library in several ways [32,33,37]. First, CRISPR library tends to have less intra-screen variations such that hit sgRNAs are more consistently recovered across biological replicates. This improves the confidence of hits calling at the hairpin level. Second, for top-scoring genes in the screen, CRISPR library tends to recover more independent hairpins for each gene than shRNA library of comparable size. This increases the confidence of hits calling at the gene level. Third, CRISPR library appears to have better penetrance than shRNA library and therefore could potentially recover more hits from a screen. This last attribute is likely due to the fact that CRISPR results in gene knockout whereas shRNA often leads to incomplete gene knockdown. Thus, genes whose biological function can be fulfilled at very low levels of expression will have a lower probability of scoring in a shRNA screen than in a CRISPR screen. For this reason, CRISPR library could potentially have an overall lower false negative rate than shRNA library.

One important parameter that is not fully characterized is the off-target rate of CRISPR libraries. Both RNAi and CRISPR/Cas9 utilize short RNA guides of comparable lengths to specify their targets: 19-22 mers for RNAi and 20-mer for CRISPR/Cas9. In both cases, a shorter seed sequence within the guide dominates sequence specificity and mismatches outside the seed sequence can be tolerated to various degrees [41,42]. Thus, CRISPR/Cas9 can cause off-target cutting in the genome and the silencing of unintended genes. For shRNA libraries, both off-target effect and false negative rate can be reduced by increasing the number of shRNAs per gene through the creation of ultra-deep libraries that contain ~25 shRNAs per gene [43]. Analogous approach can be used to mitigate the off-target effect and further reduce the false negative rate of CRISPR libraries. However, increasing the number of sgRNAs will inflate the size of the library and thus increase the cost and effort associated with the screen.

A particular form of false positive associated with CRISPR/Cas9 screen in cancer cells can arise from lethality caused by Cas9-induced DNA damage. It has been observed that sgRNAs targeted to a region of chromosome amplification in a cancer cell line can reproducibly kill these cells. This is caused by DNA damage-induced cell death as a result of the many on-target DNA breaks introduced by Cas9 [37]. This problem is likely to be specific to aneuploid cancer cell lines, particularly those with regions of high-copy gene amplification. Thus, the interpretation of CRISPR/Cas9 screen data in cancer cells should be taken cautiously in light of copy number variation data. Alternatively, CRISPRi libraries can be used to mitigate this problem, since CRISPRi-mediated gene silencing does not result in DNA cuts.

Arrayed siRNA libraries have been particularly useful for high-content screen in multi-well plates. Analogous CRISPR/Cas9 libraries are under development and they are based on either arrayed Cas9/sgRNA plasmids or purified Cas9/sgRNA nuclease complex [44]. The assembly of these arrayed CRISPR/Cas9 libraries on a genome scale is likely to be time consuming and costly. In addition, in many cancer cell lines, the transfection efficiency of such libraries is likely to be substantially lower than that with siRNA libraries. Thus, new

methods to construct arrayed CRISPR/Cas9 libraries that are low-cost, easily scalable, and with high transfection efficiency need to be developed in order for arrayed CRISPR/Cas9 libraries to achieve the same level of success as arrayed siRNA libraries did in high-content screening.

GOF CRISPRa libraries are easier to construct and deploy than cDNA libraries, they also enable access to large genes that are absent from cDNA libraries. However, the degree of gene over-expression achieved by CRISPRa could be more variable than cDNA over-expression. A cDNA library uses the same synthetic promoter to drive uniform transgene expression, whereas CRISPRa-mediated gene activation is dependent on endogenous promoters and is therefore more sensitive to chromatin structure and basal promoter activity for a given promoter in a given cell line. For this reason, the efficiency of CRISPRa libraries could be more variable across different cell lines. Thus, CRISPR libraries provide an orthogonal platform to RNAi and cDNA libraries, rather than replacing them, to enable a more comprehensive characterization of the functional landscape of the cancer genome.

Target Validation and Cancer Modeling Using CRISPR/Cas9

Target validation and cell line models

In addition to its utility in high throughput screening, CRISPR/Cas9 is enabling the rapid construction of cell line and animal cancer models for mechanistic studies. CRISPR/Cas9-mediated genome editing can be deployed to verify oncogene addiction, evaluate drug target, and model drug resistance. For example, CRISPR/Cas9 has been used to tag endogenous genes with a chemically regulatable **degtron** for the purpose of target validation. This approach places the target gene under the control of a chemical switch to assess its dependency status in cancer cells (Figure 3A) [45]. Furthermore, the WT and mutant alleles of the gene can be tagged separately to distinguish their functional differences. This approach was used to assess the RNA splicing factor *SF3B1* as a potential drug target. In two cancer cell lines carrying *SF3B1* mutations, loss of the mutant alleles did not affect cell proliferation whereas loss of the WT alleles was lethal. This unexpected finding indicates that cancer cells may not be functionally addicted to mutant SF3B1 protein, and potential SF3B1 inhibitors need to spare the WT protein to avoid significant on-target toxicity in normal cells [45].

Oncogenic chromosome translocation events can be modeled with CRISPR/Cas9 at a high degree of precision and efficiency (Figure 3B, left). For examples, three forms of chromosomal rearrangements that are found in lung adenocarcinoma, *EML4-ALK* paracentric inversion, *KIF5B-RET* pericentric inversion, and *CD74-ROS1* translocation, have all been successfully modeled in human cell lines using pairs of sgRNA that co-target the affected genes [46]. An analogous approach was used to model *EWSR1-FLI1* translocation in Ewing's sarcoma and *RUNX1-ETO* translocation in acute myeloid leukemia [47].

CRISPR/Cas9 is particularly useful for the rapid engineering of compound genetic alterations in cell lines (Figure 3B, right). In human intestinal **organoid culture** models, CRISPR/Cas9-mediated genome editing was used to introduce various combinations of

mutations in five genes associated with colorectal cancer: *APC*, *TP53*, *KRAS*, *SMAD4* and *PIK3CA* [48,49]. Comparison of these isogenic organoid cultures with one or more mutations revealed that each mutation obviates the cell's need for a specific growth factor in the culture media. Furthermore, these mutations can cooperate to drive tumor aggressiveness [48,49]. Importantly, sequencing analysis found that CRISPR/Cas9 manipulation introduced only a few off-target mutations and did not affect chromosome stability [48,49]. In a separate study, a pool of sgRNAs targeting the mouse orthologs of genes frequently deleted in human myeloid malignancies were transduced into mouse hematopoietic stem cells. Subsequent *in vivo* selection yielded the outgrowth of leukemic clones harboring 4 or 5 gene deletion events [50]. Thus, CRISPR/Cas9 is a powerful tool for the creation of **isogenic cancer cell lines** that harbor defined, combinatorial genetic lesions. These cell lines will provide a better model for understanding the functional interaction among oncogenes and tumor suppressors.

Construction of complex mouse cancer models with CRISPR/Cas9

CRISPR/Cas9-mediated genome editing is dramatically reducing the time and cost associated with the creation of complex mouse cancer models. Traditionally, **Cre-lox mediated gene recombination** is the method of choice to generate tissue-specific, single gene-driven cancer models. This requires the creation of germline animals with a floxed gene allele and crossing them with animals expressing a tissue-specific *Cre* transgene. To generate a cancer model that involves multiple genes, extensive breeding is necessary to bring all the desired alleles together and the cost and time associated with breeding increase significantly with each additional allele. CRISPR/Cas9 has been successfully used to introduce, in a single step, mutations in up to five genes in ES cells (Figure 3C, left) [51], this dramatically simplifies the construction of multi-gene cancer models in mice. Furthermore, CRISPR/Cas9-mediated gene editing works in the zygote [51], thus novel cancer models can be generated in animal species for which ES cells are either difficult to work with or are unavailable.

CRISPR/Cas9 has been combined with Cre-lox models to rapidly assess the effect of compound mutations in a defined genetic background. Using this approach, mice carrying floxed *TP53* and mutant *Kras* alleles were treated intratracheally with a lentiviral vector co-expressing Cre, Cas9 and a sgRNA targeting either *Pten*, *Nkx2.1* or *Apc* [52]. This led to the generation of lung tumors harboring triple mutations in *Kras*, *TP53* and one additional tumor suppressor without the need for additional animal crossing.

CRISPR/Cas9-mediated genome editing can be applied directly to somatic tissues of adult animals to induce tumor formation (Figure 3C, middle). For example, hydrodynamic injection of plasmids encoding Cas9 and sgRNA targeting *Pten* and *TP53* led to the deletion of both genes in hepatocytes *in vivo*. Injection of plasmids encoding Cas9 and *Ctnnb1* sgRNA together with a mutant *Ctnnb1* HDR template successfully introduced point mutation in the *Ctnnb1* gene in the liver [53]. Using lung cancer as a model, it has been shown that intratracheal delivery of adenoviral or lentiviral vectors co-expressing Cas9 and sgRNAs targeting *Eml4* and *Alk* can successfully trigger *Eml4-Alk* inversion and tumor

formation in the lung [54,55]. Thus, CRISPR/Cas9 can effectively model different types of oncogenic lesions in adult tissues without the need for germline mice.

The creation of transgenic mice with tissue-specific and inducible Cas9 expression has further facilitated *in vivo* cancer modeling (Figure 3C, right) [56,57]. In these mice, co-delivery of sgRNAs targeting *Tp53*, *Kras* and *Lkb1*, together with a mutant *Kras* HDR template, gave rise to lung tumors with *Tp53* and *Lkb1* double deletion and *Kras* activating mutation in a single step [56]. By engineering sgRNAs into the germline, CRISPR/Cas9 can be used to delete genes in tissues such as the intestinal epithelium where direct viral sgRNA delivery is difficult [57]. Together, these *in vivo* CRISPR/Cas9 approaches will rapidly expand the repertoire of mouse cancer models and facilitate mechanistic studies, target validation, and drug evaluation.

Concluding Remarks

CRISPR has emerged as a powerful tool for genetic screens and for genome editing. The technology is fast evolving and there are still many challenging problems that CRISPR could help solve (see Outstanding Questions). Currently, much effort is being devoted to improving the potency and specificity of CRISPR. Large-scale analyses of sgRNA screening data have improved algorithm for picking sgRNA sequences with high on-target efficiency and low off-target effect [31,42,58-61]. The off-target effect of Cas9 is further reduced through the development of **Cas9 nickase** that requires a pair of sgRNAs to create double-stranded breaks [22,62]. More recently, the discovery of high-fidelity Cas9 mutants with reduced non-specific binding to the DNA backbone provides new means to reduce Cas9 off-target cutting (Figure 1D) [63,64]. These advances should greatly increase the efficiency of CRISPR libraries and improve the precision of CRISPR/Cas9-mediated genome editing. The exploration of alternative CRISPR enzymes such as Cpf1 will diversify the CRISPR tool kit to access sequences in the genome not targetable by Cas9 (Figure 1E) [65,66]. Lastly, CRISPR/Cas9 itself might serve as a useful therapeutic modality and it has already been tested as gene therapy platform to correct genetic defects in mice [67-69]. Whether it can be extended to inactivate currently “undruggable” oncogenes such as Ras and Myc in tumors, or to edit a patient's immune cells for improved anti-tumor response, are only some of the exciting applications that remain to be explored. Thus, the wide adoption of CRISPR technology in cancer research will accelerate the pace of drug discovery and aid the development of more effective cancer therapies.

References

1. Vogelstein B, et al. Cancer genome landscapes. *Science*. 2013; 339:1546–1558. [PubMed: 23539594]
2. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell*. 2011; 144:646–674. [PubMed: 21376230]
3. Pagliarini R, et al. Oncogene addiction: pathways of therapeutic response, resistance, and road maps toward a cure. *EMBO Rep*. 2015; 16:280–296. [PubMed: 25680965]
4. Luo J, et al. Principles of cancer therapy: oncogene and non-oncogene addiction. *Cell*. 2009; 136:823–837. [PubMed: 19269363]
5. Kim H, Kim JS. A guide to genome engineering with programmable nucleases. *Nat Rev Genet*. 2014; 15:321–334. [PubMed: 24690881]

6. Cheung HW, et al. Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineage-specific dependencies in ovarian cancer. *Proc Natl Acad Sci USA*. 2011; 108:12372–12377. [PubMed: 21746896]
7. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science*. 2014; 346:1258096–1258096. [PubMed: 25430774]
8. Jinek M, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012; 337:816–821. [PubMed: 22745249]
9. Cong L, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013; 339:819–823. [PubMed: 23287718]
10. Mali P, et al. RNA-guided human genome engineering via Cas9. *Science*. 2013; 339:823–826. [PubMed: 23287722]
11. Cho SW, et al. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol*. 2013; 31:230–232. [PubMed: 23360966]
12. Sternberg SH, et al. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*. 2014; 507:62–67. [PubMed: 24476820]
13. Nishimasu H, et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*. 2014; 156:935–949. [PubMed: 24529477]
14. Anders C, et al. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature*. 2014; 513:569–573. [PubMed: 25079318]
15. Jinek M, et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science*. 2014; 343:1247997. [PubMed: 24505130]
16. Szczelkun MD, et al. Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proc Natl Acad Sci USA*. 2014; 111:9798–9803. [PubMed: 24912165]
17. Hsu PD, et al. Development and applications of CRISPR-Cas9 for genome engineering. *Cell*. 2014; 157:1262–1278. [PubMed: 24906146]
18. Shalem O, et al. High-throughput functional genomics using CRISPR-Cas9. *Nat Rev Genet*. 2015; 16:299–311. [PubMed: 25854182]
19. Dominguez AA, et al. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat Rev Mol Cell Biol*. 2016; 17:5–15. [PubMed: 26670017]
20. Qi LS, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*. 2013; 152:1173–1183. [PubMed: 23452860]
21. Gilbert LA, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*. 2013; 154:442–451. [PubMed: 23849981]
22. Mali P, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol*. 2013; doi: 10.1038/nbt.2675
23. Tanenbaum ME, et al. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell*. 2014; 159:635–646. [PubMed: 25307933]
24. Konermann S, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*. 2015; 517:583–588. [PubMed: 25494202]
25. Zalatan JG, et al. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell*. 2015; 160:339–350. [PubMed: 25533786]
26. Fellmann C, et al. Functional Identification of Optimized RNAi Triggers Using a Massively Parallel Sensor Assay. *Molecular Cell*. 2011; 41:733–746. [PubMed: 21353615]
27. Knott SRV, et al. A Computational Algorithm to Predict shRNA Potency. *Molecular Cell*. 2014; 56:796–807. [PubMed: 25435137]
28. Bassik MC, et al. Rapid creation and quantitative monitoring of high coverage shRNA libraries. *Nat Meth*. 2009; 6:443–445.
29. Carrette JE, et al. Global gene disruption in human cells to assign genes to phenotypes by deep sequencing. *Nat Biotechnol*. 2011; 29:542–546. [PubMed: 21623355]
30. ORFeome Collaboration. The ORFeome Collaboration: a genome-scale human ORF-clone resource. *Nat Meth*. 2016; 13:191–192.

31. Wang T, et al. Genetic screens in human cells using the CRISPR-Cas9 system. *Science*. 2014; 343:80–84. [PubMed: 24336569]
32. Shalem O, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*. 2014; 343:84–87. [PubMed: 24336571]
33. Gilbert LA, et al. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*. 2014; 159:647–661. [PubMed: 25307932]
34. Koike-Yusa H, et al. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat Biotechnol*. 2014; 32:267–273. [PubMed: 24535568]
35. Shi J, et al. Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat Biotechnol*. 2015; doi: 10.1038/nbt.3235
36. Hart T, et al. High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell*. 2015; 163:1515–1526. [PubMed: 26627737]
37. Wang T, et al. Identification and characterization of essential genes in the human genome. *Science*. 2015; 350:1096–1101. [PubMed: 26472758]
38. Luo B, et al. Highly parallel identification of essential genes in cancer cells. *Proc Natl Acad Sci USA*. 2008; doi: 10.1073/pnas.0810485105
39. Barretina J, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*. 2012; 483:603–607. [PubMed: 22460905]
40. Garnett MJ, et al. Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature*. 2012; 483:570–575. [PubMed: 22460902]
41. Mohr SE, et al. RNAi screening comes of age: improved techniques and complementary approaches. *Nat Rev Mol Cell Biol*. 2014; 15:591–600. [PubMed: 25145850]
42. Hsu PD, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*. 2013; 31:827–832. [PubMed: 23873081]
43. Kampmann M, et al. Next-generation libraries for robust RNA interference-based genome-wide screens. *Proc Natl Acad Sci USA*. 2015; doi: 10.1073/pnas.1508821112
44. McClelland ML, et al. CCAT1 is an enhancer-templated RNA that predicts BET sensitivity in colorectal cancer. *J Clin Invest*. 2016; 126:639–652. [PubMed: 26752646]
45. Zhou Q, et al. A chemical genetics approach for the functional assessment of novel cancer genes. *Cancer Research*. 2015; 75:1949–1958. [PubMed: 25788694]
46. Choi PS, Meyerson M. Targeted genomic rearrangements using CRISPR/Cas technology. *Nat Commun*. 2014; 5:3728. [PubMed: 24759083]
47. Torres R, et al. Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR-Cas9 system. *Nat Commun*. 2014; 5:3964. [PubMed: 24888982]
48. Matano M, et al. Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. *Nat Med*. 2015; 21:256–262. [PubMed: 25706875]
49. Drost J, et al. Sequential cancer mutations in cultured human intestinal stem cells. *Nature*. 2015; doi: 10.1038/nature14415
50. Heckl D, et al. Generation of mouse models of myeloid malignancy with combinatorial genetic lesions using CRISPR-Cas9 genome editing. *Nat Biotechnol*. 2014; 32:941–946. [PubMed: 24952903]
51. Wang H, et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*. 2013; 153:910–918. [PubMed: 23643243]
52. Sanchez-Rivera FJ, et al. Rapid modelling of cooperating genetic events in cancer through somatic genome editing. *Nature*. 2014; 516:428–431. [PubMed: 25337879]
53. Xue W, et al. CRISPR-mediated direct mutation of cancer genes in the mouse liver. *Nature*. 2014; doi: 10.1038/nature13589
54. Maddalo D, et al. In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. *Nature*. 2014; 516:423–427. [PubMed: 25337876]
55. Blasco RB, et al. Simple and rapid in vivo generation of chromosomal rearrangements using CRISPR/Cas9 technology. *Cell Rep*. 2014; 9:1219–1227. [PubMed: 25456124]
56. Platt RJ, et al. CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling. *Cell*. 2014; 159:440–455. [PubMed: 25263330]

57. Dow LE, et al. Inducible in vivo genome editing with CRISPR-Cas9. *Nat Biotechnol.* 2015; 33:390–394. [PubMed: 25690852]
58. Wu X, et al. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol.* 2014; 32:670–676. [PubMed: 24752079]
59. Doench JG, et al. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol.* 2014; 32:1262–1267. [PubMed: 25184501]
60. Doench JG, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol.* 2016; doi: 10.1038/nbt.3437
61. Chari R, et al. Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach. *Nat Meth.* 2015; 12:823–826.
62. Ran FA, et al. Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell.* 2013; doi: 10.1016/j.cell.2013.08.021
63. Slaymaker IM, et al. Rationally engineered Cas9 nucleases with improved specificity. *Science.* 2015; doi: 10.1126/science.aad5227
64. Kleinstiver BP, et al. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature.* 2016; doi: 10.1038/nature16526
65. Esvelt KM, et al. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat Meth.* 2013; 10:1116–1121.
66. Zetsche B, et al. Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell.* 2015; doi: 10.1016/j.cell.2015.09.038
67. Long C, et al. Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science.* 2016; 351:400–403. [PubMed: 26721683]
68. Nelson CE, et al. In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science.* 2016; 351:403–407. [PubMed: 26721684]
69. Tabebordbar M, et al. In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science.* 2016; 351:407–411. [PubMed: 26721686]

Glossary

Cas9 nickase

a mutant of Cas9 with a single functional endonuclease domain that can introduce single-stranded DNA nicks

Cre-lox and Flp-FRT mediated gene recombination

a site-specific recombination method that uses the bacteriophage Cre recombinase and the yeast Flp recombinase to trigger gene recombination at sites flanked by *loxP* and *FRT* sequences, respectively

Degron

a amino acid sequence in a protein that specify the protein for proteolytic degradation

False positive rate and false negative rate

the probability of falsely rejecting and accepting the null hypothesis, respectively

High-fidelity Cas9

mutant Cas9 enzymes with reduced cleavage efficiency at off-target sites

Homology-directed repair (HDR)

a DNA repair mechanism that uses a DNA template that is homologous to the site of DNA double-stranded break to repair the break

Isogenic cell lines

cell lines with identical genetic background other than the specified mutations

Krüppel associated box (KRAB) domain

a transcription repressor domain found in many zinc finger transcription factors

Non-homologous end-joining (NHEJ)

a DNA repair mechanism that directly ligates the ends of DNA double-stranded breaks

Non-sense mediated mRNA decay

a cellular process that degrades mutant mRNAs that contain premature stop codons, often as a result of frame-shift mutations

Off-target effect

phenotype arising from perturbation of unintended genes rather than the specific target gene of interest

Organoid culture

an in vitro cell culture that mimics or retains certain structural and functional properties of its organ of origin

Pooled and arrayed libraries

in a pooled library members of the library are kept in a homogenous mixture, whereas in an arrayed library each member is kept separately (typically in multi-well plates)

RNA interference (RNAi)

a cellular mechanism that uses short RNA guides (siRNAs or shRNAs) to suppress the expression of mRNAs through either degradation or translation inhibition of target mRNA

sgRNA

a synthetic, single guide RNA sequence that is a fusion of crRNA and tracrRNA

Target validation

the process to verify that a predicted molecular target would have the desired therapeutic benefit when modulated by small molecules or other means

Transcription activator-like effector nuclease (TALEN)

A synthetic restriction endonuclease that consists of several transcription activator-like effector DNA binding domains for sequence recognition and an endonuclease domain for DNA cutting

Zinc finger nuclease (ZFN)

A synthetic restriction endonuclease that consists of several zinc finger DNA binding domains for sequence recognition and an endonuclease domain for DNA cutting

Box 1**CRISPR in bacteria**

In bacteria, the CRISPR RNA and endonuclease modules are encoded in the CRISPR locus where the crRNAs are transcribed from a CRISPR array containing alternating element of direct repeats and spacer sequences. The spacer sequences are acquired from phage DNA and they serve as templates for the CRISPR system to recognize phage DNA in the host genome. Processing of the precursor RNA transcript from the CRISPR array yields short, mature crRNAs that complex with the Cas endonuclease module to form the holoenzyme. This endonuclease complex interrogates the host genome for DNA sequences that match the crRNA and proceeds to cleave these DNA sites. Thus, CRISPR serves as a form of adaptive immunity in bacteria. Bacterial CRISPR systems are divided into three classes based on the structural organization and sequence homology of their constituent protein and RNA subunits [7,17]. Type I and III CRISPR systems use multiple protein subunits for the endonuclease module, whereas the Type II CRISPR system utilizes a single endonuclease subunit Cas9.

Box 2**Pooled shRNA and CRISPR library screens**

Large number of shRNA or sgRNA hairpins can be screened in a pooled format without the need to array them individually in multi-well plates. In this approach, tens of thousands of hairpin are synthesized on oligonucleotide arrays and cloned, as a pool, into either retroviral or lentiviral vectors. The library is packaged into a pool of viruses and is used to transduce cells at low multiplicity of infection (MOI), typically less than 1. This results in the majority of cells receiving only one hairpin per cell. During the transduction process, the representation of the library is kept relatively high such that for a given hairpin there are typically hundreds of independent integration events. The cell population carrying the library is then subjected to selection criteria, which could be proliferation, killing by a drug, or cell sorting for a desired reporter status. The composition of the library in the cell population before and after the selection process is assessed by PCR recovery of the hairpins from genomic DNA followed by deep-sequencing.

Outstanding Questions Box

- How to further reduce the false negative and false positive rates of CRISPR library screens with improved Cas9 and sgRNA on-target efficiency?
- Could other CRISPR endonuclease systems, such as CRISPR/Cpf1, which has distinct sgRNA structure, PAM and DNA cutting property than Cas9, be effectively deployed for CRISPR library screens?
- What are the novel genetic screens that genome-wide pooled and arrayed CRISPR libraries enable that were previously difficult or impossible to carry out?
- What are the novel cancer models that CRISPR technology could help create to better mimic human cancer?
- Could CRISPR/Cas9 itself be a useful therapeutic modality for cancer?

Trend Box

- CRISPR/Cas9 is a versatile genomic engineering tool for gene inactivation, activation, mutation and chromosome rearrangements.
- Genome-wide lentiviral CRISPR libraries have been created for both gain- and loss-of-function screens. They are powerful tools for discovering genes that could serve as cancer drug targets.
- CRISPR/Cas9-mediated genome editing reduces the time and effort associated with creating cell line and animal models of cancer with complex genetic alterations. It therefore enables the generation of better cancer models for target validation and drug evaluation.

Figure 1. CRISPR/Cas9 tools for genome engineering

A. Principle of CRISPR/Cas9 mediated genome editing. The Cas9/sgrRNA complex binds to a 20 nt target sequence that precedes an “NGG” protospacer adjacent motif. Complementarity between the sgrRNA and DNA sequence induces double-stranded cutting 3 nt upstream of the PAM. DNA repair through the non-homologous end joining (NHEJ) pathway often results in insertion or deletion (indel) mutations that silence the target gene when they occur in exons. DNA repair through the homology directed repair (HDR) pathway enables the introduction of point mutations and precision indels when a DNA template is supplied. **B.** CRISPR inactivation (CRISPRi)-mediated gene silencing with nuclease-dead Cas9 (dCas9). When targeted to the transcription start site (TSS) of a gene, dCas9 can interfere with gene expression by blocking RNA polymerase (left). Gene silencing can be further enhanced by fusing dCas9 with a KRAB domain to promote epigenetic silencing (right). **C.** CRISPR activator (CRISPRa)-mediated gene over-expression. Direct fusion of dCas9 with a transcriptional activator domain such as VP64 (left), or the use of epitope tags (middle) and modified sgrRNA stem loops (right) to recruit transcriptional factors (VP64 and MS2-p65-HSF1) can drive high levels of target gene expression. **D.** The specificity of CRISPR/Cas9 can be improved by using paired Cas9 nickase (Cas9n) to introduce DNA breaks (left) or by using high-fidelity Cas9 mutants

(Cas9-HF) that have reduced affinity for DNA backbone (right). **E.** Alternative CRISPR enzymes such as Cpf1 recognize different PAMs and thus can target sequences not accessible by Cas9.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 2. Cell-based high-throughput screening using pooled CRISPR/Cas9 libraries
Lentiviral CRISPR, CRISPRi or CRISPRa libraries are transduced into cells at a low multiplicity of infection such that each cell receives a single sgRNA (represented by different colors). Proliferation assays can be used to detect sgRNAs that either affect generally cell viability (left), or the viability of specific cancer cell lines in a context-dependent manner (middle), or the sensitivity of cancer cells to drug treatment (right).

Figure 3. Creation of cell line and mouse cancer models with CRISPR/Cas9

A. CRISPR/Cas9-mediated knock-in of a chemically regulatable degon-tag (red bar) in a gene enables target validation by placing its expression under the control of a small molecule. **B.** CRISPR/Cas9 can be used to engineer different types of mutation including small deletions and point mutations (left), as well as large alterations such as chromosomal translocation (right), to create panels of isogenic cell lines that harbor different combinations of oncogenic lesions. **C.** Germline mouse models carrying multiple genetic mutations can be created with CRISPR/Cas9 edited ES cells (left). Virus or plasmid vectors expressing Cas9 and sgRNAs can be delivered by either airway instillation or tail vein injection to create somatic mutations in the lung epithelium and in hepatocytes, respectively (middle). In Cas9 transgenic animals, lung-specific mutations can be created by viral sgRNA delivery through the airway. In transgenic mice co-expressing an inducible Cas9 together with sgRNAs, tissue-specific gene knockout can be achieved in the intestinal epithelium. In each of these models, multiple genes can be simultaneously manipulated to create models with compound genetic mutations.