Review Article
CRISPR screen in cancer: status quo and future perspectives

Chenglong He¹*, Siqi Han¹*, Yue Chang², Meijuan Wu², Yulu Zhao³, Cheng Chen¹,², Xiaoyuan Chu¹,²,³

¹Department of Medical Oncology, Jinling Hospital, The First School of Clinical Medicine, Southern Medical University, Nanjing 210002, China; ²Department of Medical Oncology, Jinling Hospital, School of Medicine, Nanjing University, Nanjing 210002, China; ³Department of Medical Oncology, Jinling Hospital, Nanjing Medical University, Nanjing 210002, China. *Equal contributors.

Received October 20, 2020; Accepted January 19, 2021; Epub April 15, 2021; Published April 30, 2021

Abstract: Clustered regularly interspaced short palindromic repeats (CRISPR) system offers a powerful platform for genome manipulation, including protein-coding genes, noncoding RNAs and regulatory elements. The development of CRISPR screen enables high-throughput interrogation of gene functions in diverse tumor biologies, such as tumor growth, metastasis, synthetic lethal interactions, therapeutic resistance and immunotherapy response, which are mostly performed in vitro or in transplant models. Recently, direct in vivo CRISPR screens have been developed to identify drivers of tumorigenesis in native microenvironment. Key parameters of CRISPR screen are constantly being optimized to achieve higher targeting efficiency and lower off-target effect. Here, we review the recent advances of CRISPR screen in cancer studies both in vitro and in vivo, with a particular focus on identifying cancer immuno-therapy targets, and propose optimizing strategies and future perspectives for CRISPR screen.

Keywords: CRISPR-Cas9, CRISPR screen, high-throughput, cancer biology, target discovery

Introduction

Cancer is characterized by multiple genetic and epigenetic alterations in oncogenes and tumor suppressor genes. Although cancer genomics studies have contributed to the discovery of cancer-associated genes, many of them are not ‘driver genes’, but only ‘passenger genes’. Functional genetic screens are powerful tools for identifying the causal relationship between genotype and phenotype [1, 2]. Compared to conventional screens conducted with RNAi or cDNA library, CRISPR-Cas9 (clustered regularly interspaced short palindromic repeats-CRISPR associated nuclease 9) based screen demonstrates stronger genetic editing capability [3-5], less off-target effect and is more versatile as it can be used in multiple formats and can target coding and non-coding regions throughout the genome.

CRISPR screens are mainly categorized into knockout screen, knockdown screen and activation screen, which share a common workflow. Firstly, designed single-guide RNAs (sg-RNAs) are cloned into a lentivirus library, and transduced into Cas9-expressing or dCas9-expressing cells at a low multiplicity of infection to ensure that only one copy of sgRNA is integrated per cell. Secondly, CRISPR library-transduced cells are subjected to biology assay-based screening. If the target gene alters cell fitness in the context of a selection pressure, cells harboring the sgRNA will drop out or enrich among the population. Lastly, CRISPR screens leverage the unique sgRNA sequences and next-generation sequencing (NGS) to identify shifts in sgRNA frequency after a phenotypic selection.

As CRISPR screen continues to improve, its applications in cancer research are becoming increasingly extensive. In this review, we introduce the approaches of CRISPR-based (epi) genome editing, summarize the recent applications of CRISPR screen in cancer studies both in vitro and in vivo, and propose potential optimization strategies and future directions for CRISPR screen.
CRISPR screen in cancer research

CRISPR-based (epi) genome manipulation: coding and noncoding regions

**CRISPRing protein-coding genes**

For knockout of coding genes, generation of a targeted double-strand DNA break (DSB) is the key event [4-6]. The repair of DSB by nonhomologous end joining (NHEJ) will result in frameshift indel mutations, potentially leading to knockout of the target gene [1]. When designing sgRNAs, it is generally preferred to target the exon region within 200 bp downstream of the start codon (Figure 1A). For genes that have more than two isoforms, a public region as close as possible to the 5’ end is usually chosen to ensure that each isoform will successfully undergo frameshift indel mutation. To date, several genome-wide or sub-pooled sgRNA libraries for CRISPR knockout screen have been constructed (GeCKOv1 library, GeCKOv2 library, Brunello library etc.), and has been widely used for investigating functions of protein-coding genes in diverse tumor biology [7-10].

Nevertheless, in addition to frameshift mutation, this approach often generates in-frame variants which produce nearly full-length proteins with alterations at the sgRNA-targeted site. In-frame mutations in the functional region of a protein may result in loss-of-function or gain-of-function that alters cell fitness in the context of a selection pressure. In this regard, researchers used the CRISPR screen to identify functional regions within LSD1 that modulate sensitivity to LSD1 inhibitors, and revealed a nonenzymatic role of LSD1 [11].

CRISPRi/CRISPRa screens have been applied to repress or activate the transcription of protein-coding genes by localizing dCas9-KRAB repressor fusion protein or synergistic activation mediator (SAM) to the transcriptional start site (TSS), respectively [12, 13]. In order to achieve high efficiency and low off-target effect, sgRNAs usually target a window of DNA from -50 to 300 bp relative to the TSS for CRISPRi (Figure 1B), and -400 to -50 bp upstream of the TSS for CRISPRa (Figure 1C). To date, genome-scale sgRNA libraries for CRISPRi or CRISPRa screen have been established (hCRISPRi-v2 library, hCRISPRa-v2 library, SAM library etc.) and carried out to identify essential coding genes for cancer cell growth and drug resistance [10, 13].

**CRISPRing non-coding RNAs**

Up to 70% of human genome is transcribed into non-coding RNAs [14, 15], including microRNAs (miRNAs) and long non-coding RNAs (lncRNAs), which are involved in diverse physiological and pathological progresses. The emergence of CRISPR-Cas9 system provides researchers with new ideas to evaluate the biological functions of non-coding RNAs [16-18].

Pri-miRNAs are processed in two steps by RNase III protein Drosha and Dicer successively to become mature miRNAs. CRISPR-Cas9 knockout system can inhibit miRNAs expression by targeting their biogenesis processing sites [19, 20] (Figure 1D). It is also feasible to deplete a miRNA by targeting its 5’ region, including the Drosha processing site and seed region [21]. Recently, a miRNA-focused CRISPR-Cas9 library, which targeted stem-loops of 1594 (85%) annotated human miRNAs, has been constructed to screen for miRNAs that affect cell fitness of HeLa or NCI-N87 cells [22].

LncRNAs is a diverse class of non-coding RNAs with length over 200nt [23]. Similar to targeting coding genes, CRISPRi/CRISPRa screens have been utilized to identify functional lncRNAs (Figure 1B, 1C). Liu et al. developed a CRISPRi Non-Coding Library (CRiNCL), targeting 16,401 lncRNA loci with 10 sgRNAs targeting each lncRNA TSS. They identified 499 lncRNA loci required for robust cellular growth [24]. Genome-scale activation screens were also performed by using sgRNA libraries targeting the TSS of lncRNAs [25, 26]. However, in terms of CRISPR knockout system, unlike targeting protein-coding genes, sgRNA induced-small indels in noncoding regions do not generally disrupt the lncRNA function. Nonetheless, large-fragment deletion of lncRNA using two gRNAs can be effective at eliminating lncRNA genes [27] (Figure 1E), and has been extended to perform high-throughput screening. Wei’s group constructed a lentiviral paired-guide RNA (pgRNA) library targeting about 700 human lncRNA genes, identifying 51 lncRNAs that can positively or negatively regulate human cancer cell growth [28]. Recently, they constructed a novel CRISPR library with sgRNAs that specifically target the splice sites of lncRNAs, achieving exon skipping or intron retention (Figure 1F). This method may provide an effective tool for systematic discovery of lncRNA functions [29].
CRISPR screen in cancer research

**Figure 1.** CRISPR-based (epi)genome editing for coding and noncoding regions. A. CRISPR knockout of coding gene by sgRNAs targeting the exon region within 200 bp downstream of start codon, and directing Cas9 to generate DSB on the target DNA sequence. B. CRISPRi-based transcriptional repression of coding gene or IncRNA using dCas9 fused to transcriptional repressor such as KRAB. The dCas9 is directed by sgRNAs targeting a window of DNA from -50 to 300 bp relative to TSS. C. CRISPRa-based transcriptional activation using dCas9 tethered to transcriptional activator such as VP64, and sgRNA-recruited MS2-p65-HSF1. The dCas9 is directed by sgRNAs targeting -400 to -50 bp upstream of TSS. D. CRISPR knockout of miRNA by sgRNA targeting the biogenesis processing sites. E. CRISPR-mediated deletion of lncRNA by paired gRNAs targeting promoter or gene body of lncRNA. F. CRISPR-mediated perturbation of lncRNA function by sgRNAs targeting splice sites, which lead to exon skipping or intron retention. G. CRISPR-based interrogation of regulatory element by sgRNAs targeting the regulatory region, usually distant 5’ flanking region.
CRISPR screen in cancer research

Despite these impressive results, it should be noted that CRISPR approaches are not suitable to target all IncRNAs, because many IncRNAs are derived from bidirectional promoters or overlap with promoters or bodies of genes. Researchers found only about 1/3 of IncRNA loci are safely amenable to CRISPR applications, while the rest of IncRNA loci are at risk of deregulating neighbor genes [30].

**CRISPRing regulatory elements**

Regulatory elements in the genome, such as enhancers, regulate transcription of distantly located genes and are critical for human disease. The extension of CRISPR screens into the regulatory regions provides novel insights for interrogation of genome (Figure 1G). To explore whether mutations in the regulatory regions around three genes involved in vemurafenib resistance could similarly affect drug resistance, Zhang et al. designed three sgRNA libraries targeting 100-kb 5' and 100-kb 3' of each gene. Sequencing analysis revealed that sgRNAs was most enriched at the 5' noncoding region flanking CUL3, which was validated to regulate CUL3 expression and drug resistance [31]. Reuven Agami’s group focused on the functional genetic screens of enhancer elements. They constructed sgRNA libraries targeting senescence-induced enhancers that were putatively bound by p53 or AP-1, and discovered key enhancers for the establishment of oncogene-induced senescence. Their work demonstrates the power of CRISPR screens in deciphering the function of regulatory elements [32, 33]. To design a suitable sgRNA library for mapping of noncoding regulatory regions, a technology called Molecular Chipper was developed. This technique combines random fragmentation with a class III restriction enzyme to generate a densely packed sgRNA library from DNA. Using this method, researchers identified regions critical for miR-142 production, including the pre-miR-142 region and two new cis-regulatory regions [34].

### Applications of CRISPR screen in cancer

**Discovery of growth essential genes for drug targets**

CRISPR screen enables high-resolution detection of growth essential genes, which represent viable therapeutic targets [9, 24, 35-43]. *In vitro* or indirect *in vivo* (transplant model) screens with negative selection are commonly used (Table 1). For *in vitro* experiments, cancer cells transduced with a library are directly cultured *in vitro* for a period of time, followed with deep sequencing to identify depleted sgRNAs (Figure 2A). Shi et al. constructed sgRNA libraries targeting exons that encode functional protein domains, achieving a higher proportion of

### Table 1. Application of CRISPR screen for growth essential genes

<table>
<thead>
<tr>
<th>Selection context</th>
<th>Cancer type</th>
<th>CRISPR type</th>
<th>Library</th>
<th>Top hits</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>In vitro</strong></td>
<td>CML, lymphoma (KBM7, K562, Raji, Jiyoye)</td>
<td>Knockout</td>
<td>custom library</td>
<td>C16of80, C3or f17, C9</td>
<td>[10]</td>
</tr>
<tr>
<td></td>
<td>HCC (Huh7, SMMC-7721)</td>
<td>Knockout</td>
<td>GeCKOv2</td>
<td>SGOL1</td>
<td>[35]</td>
</tr>
<tr>
<td></td>
<td>CML (K562)</td>
<td>Knockout</td>
<td>GeCKOv1</td>
<td>VHL</td>
<td>[36]</td>
</tr>
<tr>
<td></td>
<td>AML (RN2)</td>
<td>Knockout</td>
<td>Custom library</td>
<td>Dot1I, Ehmt1, Ehmt2, Ezh2, Brd4, Kdm1a</td>
<td>[37]</td>
</tr>
<tr>
<td></td>
<td>AML (MOLM-13, MV4-11, HL-60, OCIAML2, OCI-AML3)</td>
<td>Knockout</td>
<td>murine lentiviral gRNA library (version 2)</td>
<td>BRD4, DOT1L, MEN1</td>
<td>[38]</td>
</tr>
<tr>
<td></td>
<td>AML (L-MEPs)</td>
<td>Knockout</td>
<td>Custom library</td>
<td>Elf1 and Sp1</td>
<td>[39]</td>
</tr>
<tr>
<td></td>
<td>30 cancer types (324 human cancer cell lines)</td>
<td>Knockout</td>
<td>Human CRISPR Library v.1.0 and v.1.1</td>
<td>Werner syndrome ATP-dependent helicase</td>
<td>[40]</td>
</tr>
<tr>
<td></td>
<td>CML, glioblastoma, cervical cancer, breast cancer</td>
<td>CRISPRi</td>
<td>non-coding library</td>
<td>499 IncRNA loci</td>
<td>[24]</td>
</tr>
<tr>
<td><strong>Indirect in vivo</strong></td>
<td>NSCLC (Primary cell)</td>
<td>Knockout</td>
<td>mGeCKOa</td>
<td>Cdkn2b, Nf2, Pten, Pdgfra, Itgax</td>
<td>[41]</td>
</tr>
<tr>
<td></td>
<td>epithelial ovarian cancer (SKOV3)</td>
<td>Knockout</td>
<td>GeCKOv2 library A</td>
<td>KPNB1</td>
<td>[42]</td>
</tr>
<tr>
<td></td>
<td>AML (Primary cell)</td>
<td>Knockout</td>
<td>GeCKOv2</td>
<td>DCPS</td>
<td>[43]</td>
</tr>
</tbody>
</table>

CML, chronic myelocytic leukemia; HCC, hepatocellular carcinoma; AML, acute myelocytic leukemia; NSCLC, non-small cell lung cancer.
CRISPR screen in cancer research

Figure 2. Schematics of CRISPR screen in diverse processes of cancer. A. The mutagenized cancer cells are cultured in vitro or transplanted into mice. Through sequencing and analysis of the depleted sgRNAs in cancer cells, potential drug targets are identified. B. Cancer cells with low metastatic potential are transduced with sgRNA library, and transplanted or injected into mice for metastasis formation. The metastatic tumors are dissected and subjected to NGS for sgRNA abundance. C. Cancer cells with mutation (MUT) or non-mutation (WT) are infected with sgRNA library and cultured in vitro. Candidate genes that are synthetic lethal with the mutated gene are identified based on sgRNA depletion in MUT cells compared to WT cells. D. Cas9-expressing cancer cells are infected with a lentiviral sgRNA library, and subjected to drug treatment. Surviving cells with resistant phenotype are subjected to NGS to identify candidate genes for drug resistance. E. Cancer cells are transduced with a reporter to reflect the activities of certain signaling pathway or biological process. The reporter-expressing cells are transduced with sgRNA library, followed by FACS sorting of fluorescence high or low populations. Then selected cell pools are subjected to deep sequencing for sgRNA abundance.
null mutations and increasing the potency of negative selection [37]. In a recent study, Behan et al. used genome-wide CRISPR knock-out library to carry out the largest screen in 324 human cancer cell lines from 30 cancer types. Thousands of key cancer genes entered into a prioritized system that outputs approximately 600 genes with the greatest potential for drug development. This study provided a wealth of reliable information for the initial stages of anticancer drug development [40].

For indirect in vivo experiments, mutagenized cancer cells are transplanted into mice for tumor formation, and then the DNA of tumor tissues are extracted for amplification and deep sequencing (Figure 2A). Jenkins et al. performed loss-of-function screen in epithelial ovarian cancer cells using both pooled shRNA library and genome-wide sgRNA library. Transduced cell pools were injected intraperitoneally into mice for tumor formation. These screens identified 10 high-confidence candidate genes as drug targets for epithelial ovarian cancer, including a novel oncogene KPNB1 [42]. Another group performed a genome-wide CRISPR-Cas9 screen and identified 2,256 dropout genes for AML cells. Those genes were filtered for potentially actionable targets by database search, among which 470 genes were subjected to a second in vivo CRISPR screen. Overall, 130 genes necessary for AML cell survival both in vitro and in vivo were identified, including the mRNA decapping enzyme scavenger (DCPS), validated as a target for AML therapy [43].

**Identification of metastasis regulators**

CRISPR screening for metastasis-related genes are reported relatively less [41, 44, 45]. Tumor cells with low metastatic potential are usually used for transduction, for cells tend to become highly metastatic after being mutagenized with sgRNA library. Metastatic models could be established through subcutaneous or orthotopic transplantation, as well as tail vein or intrasplenic injection, according to study purpose [46, 47] (Figure 2B). In this process, a large amount of mutagenized cells (usually over $6 \times 10^7$) will be dividedly transplanted into a number of mice. However, the random sampling during transplantation influences sgRNA dynamics. Besides, each step toward metastasis turns as a bottleneck, together leading to the loss of a large number of sgRNAs in the metastatic tumors. As reported in a study of Sidi Chen, they mutagenized a non-metastatic mouse lung cancer cell line with a mouse genome-wide library and transplanted it subcutaneously into immunocompromised mice. Infected cells were found to form primary tumors more quickly than controls and metastasize into lungs. Notably, deep sequencing revealed a high sgRNA dropout rate during tumor evolution, with less than half of the sgRNAs retained in the early tumors, and even less in the metastases [41]. Thus, maybe only those genes that are quite critical for metastasis could be screened out, due to a high dropout rate in this process. In a recent report, Chen performed double perturbations using a massively parallel CRISPR-Cpf1/Cas12a crRNA array profiling (MCAP) to identify genetic interactions that drive metastasis. They found that certain gene pairs may be synergistic in promoting metastasis, such as Nf2 plus Trim72. This study highlights the power of MCAP for high-throughput interrogation of genetic interactions [45].

**Finding synthetic lethal interactions**

Synthetic lethality involves the interaction of two genes, in which one mutated gene alone does not affect cell viability, but the combinatorial mutation or loss leads to cell death. CRISPR screen provides strong support for identifying synthetic lethal genes, giving new insights into targeted therapies [48-58] (Table 2). To identify synthetic lethality associated with a certain mutant gene, cells with mutation and non-mutation (wild-type) are infected with sgRNA library and subjected to in vitro culture or transplantable tumor formation. Candidate genes that were synthetic lethal with the mutated gene are identified based on sgRNA depletion in the mutant cells compared to wild-type cells (Figure 2C). Edwin H. Yau et al. performed a genome-wide CRISPR screen in KRAS$^{\text{MUT}}$ and KRAS$^{\text{WT}}$ HCT116 colorectal cancer cell lines, and found that hit genes were associated with the MAPK signaling pathway and metabolic pathways, including SUCLA2, NADK and KHK. Moreover, they performed a secondary focused screen with higher depth for validation, revealing potential synthetic lethal partners of KRAS mutation with greater power [48]. If iso-
genic mutant and wild-type cells are not ready-made, doxycycline-inducible vectors could be used to generate cells with certain gene loss. For example, cells lacking RB1 or ΔNp63α were established upon doxycycline treatment, and were subjected to CRISPR screen for respective synthetic lethal hits [50, 51]. To more efficiently screen out synthetic lethal genes, researchers developed a CRISPR-based double knockout (CDKO) system that comprised 490,000 double-sgRNAs against 21,321 pairs of drug targets. They identified synthetic lethal drug-target pairs from CRISPR-deleted gene pairs, and validated that corresponding drug combinations exhibit synergistic killing effect [53].

Small-molecule inhibitors can phenocopy the effect of specific mutations. Thus, finding synthetic lethal drug-mutation interactions will help discover specific mutated cancer cells that are selectively sensitive to the drug or combinatorial drugs that improve therapeutic index. To identify synthetic lethality associated with a certain drug, library transduced cells are treated with the drug or vehicle control, and passaged for doubling. Then deep sequencing is conducted to identify selective dropouts in drug treatment group compared with vehicle control group. Resistance to asparaginase is a common problem for leukemia in the clinic. Using a genome-wide CRISPR-Cas9 screen, Laura Hinze et al. found a synthetic lethal interaction between Wnt pathway activation and asparaginase in resistant leukemia cells. Inhibition of GSK3α profoundly sensitized drug-resistant leukemia to asparaginase [54]. Similarly, researchers identified synthetic lethal interactions of SHOC2 deletion with MEK inhibition in Ras-mutant cancer [55] and RNASEH2 deficiency with ATR inhibition [56].

Identification of genes involved in drug resistance

Pooled lentiviral CRISPR-Cas9 screen provides a powerful platform for identifying genes involved in drug resistance. In general, Cas9-expressing cancer cells are transduced with a lentiviral sgRNA library, and subjected to drug treatment. After a period of culture, surviving cells with resistant phenotype were subjected to NGS to identify candidate genes for drug resistance (Figure 2D). The genetic CRISPR screen was first developed in mammalian cells to identify genes whose loss-of-function mutations conferred resistance to 6-thioguanine and etoposide in leukemic cells or to vemurafenib in melanoma cells [7, 8]. Thereafter, CRISPR screen approach is widely used to investigate critical and novel mechanisms underlying drug resistance in cancers [26, 59-65] (Table 3). Bester et al. developed an integrated genome-wide platform based on a dual protein-coding and IncRNA CRISPRa screening, with a focus on IncRNA. They developed a CRISPRa of IncRNA (CaLR) library, targeting 14,701 IncRNAs, among which GAS6-AS2 IncRNA was identified and validated to mediate cytarabine resistance by activating GAS6/TAM pathway. This study represents a powerful approach to

<table>
<thead>
<tr>
<th>Cancer type</th>
<th>Mutation/Drug</th>
<th>CRISPR type</th>
<th>Library</th>
<th>Synthetic lethal hits</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRC (HCT116)</td>
<td>KRAS</td>
<td>Knockout</td>
<td>GeCKOv2</td>
<td>NADK, HKH</td>
<td>[48]</td>
</tr>
<tr>
<td>PDAC (HPAF-II)</td>
<td>RNF43</td>
<td>Knockout</td>
<td>TKO library</td>
<td>Wnt-FZD5 signaling</td>
<td>[49]</td>
</tr>
<tr>
<td>SCC (H226)</td>
<td>ΔNp63α</td>
<td>Knockout</td>
<td>GeCKOv2</td>
<td>RHOA</td>
<td>[50]</td>
</tr>
<tr>
<td>SCLC (NCI-H82)</td>
<td>RB1 loss</td>
<td>Knockout</td>
<td>custom library</td>
<td>Aurora B kinase</td>
<td>[51]</td>
</tr>
<tr>
<td>HCC (PLC/PRF/S)</td>
<td>ATRX loss</td>
<td>Knockout</td>
<td>GeCKOv2</td>
<td>WEE1</td>
<td>[52]</td>
</tr>
<tr>
<td>CML (K562)</td>
<td>-</td>
<td>Double</td>
<td>sgRNA library</td>
<td>BCL2L1 and MCL1</td>
<td>[53]</td>
</tr>
<tr>
<td>T-ALL (CCRF-CEM)</td>
<td>Asparaginase</td>
<td>Knockout</td>
<td>GeCKO</td>
<td>ASNS, NKD2, LG6</td>
<td>[54]</td>
</tr>
<tr>
<td>pancreatic cancer, NSCLC</td>
<td>MEK inhibition</td>
<td>Knockout</td>
<td>Avana-4 barcoded</td>
<td>SHOC2</td>
<td>[55]</td>
</tr>
<tr>
<td>(CFPAC-1, A549, NCIH23)</td>
<td></td>
<td></td>
<td>sgRNA library</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRC, breast cancer (HCT116, MCF10A)</td>
<td>ATR inhibition</td>
<td>Knockout</td>
<td>TKOv3 library</td>
<td>RNASEH2</td>
<td>[56]</td>
</tr>
<tr>
<td>(293A)</td>
<td>pan-Aurora kinase inhibitor</td>
<td>Knockout</td>
<td>TKOv3 library</td>
<td>GSG2</td>
<td>[57]</td>
</tr>
<tr>
<td>TNBC (SUM159, SUM149)</td>
<td>BET bromodomain inhibitor</td>
<td>Knockout</td>
<td>H1 and H2 libraries</td>
<td>CDK4 and BRD2</td>
<td>[58]</td>
</tr>
</tbody>
</table>

CRC, colorectal cancer; PDAC, pancreatic ductal adenocarcinoma; SCC, squamous cell carcinoma; HCC, hepatocellular carcinoma; CML, chronic myelocytic leukemia; NSCLC, non-small cell lung cancer; T-ALL, T-cell acute lymphoblastic leukemia.
identify integrated coding and non-coding pathways of therapeutic relevance [26].

**Reporter-based CRISPR screen**

To enable CRISPR screen for modulators of certain signaling pathway or biological process, corresponding fluorescence-based reporters are used. In general, reporter-expressing cells are transduced with lentiviral sgRNA library, and then fluorescence high or low populations are sorted by fluorescence-activated cell sorting (FACS) and subjected to deep sequencing respectively (Figure 2E). To identify regulators of YAP activity, researchers used a doxycycline inducible Strep-YAP5SA allele and a turboRFP (red fluorescent protein) reporter under the control of a CTGF promoter. They found that the transcriptional repressor protein TRPS1 acted as a repressor of YAP-dependent transactivation [66]. Super-enhancers (SEs) play pivotal roles in oncogenesis. To identify regulators of Epstein-Bar virus bound SEs (ESEs) in lymphoblastoid cell, genome-wide CRISPR screens were performed in lymphoblastoid cells expressing GFP reporters driven by MYC ESE which located 525 kb upstream of MYC TSS. Cells that lost GFP signals were sorted by FACS and analyzed for deleted genes, among which TAF family was identified as essential regulators of ESE activity [67]. To elucidate the mechanisms underlying cancer senescence, Wang et al. performed a CRISPR screen for senescence-inducing genes based on a miR146a-EGFP reporter, as miR146a expression has been demonstrated to upregulate during senescence. The chromatin remodeler SMARCB1 was identified as top candidate, whose suppression induced senescence through activating MAPK pathway in melanoma [68]. Similar approach has been used to screen regulators of autophagy as well [69-71].

**CRISPR screen for cancer immunotherapy targets**

Immunotherapy has increasingly become an effective means of treatment for late-stage cancer. For the past few years, some strategies of CRISPR screen, such as one cell type (1CT)-CRISPR screen, two cell type (2CT)-CRISPR screen, and transplantable in vivo CRISPR screen, has been conducted to find targets for enhancing the efficacy of immunotherapy [72-82] (Table 4).

**1CT-CRISPR screen**

One strategy is to identify regulators of molecules that mediate immune evasion in tumor cells (Figure 3A). Loss of MHC class I (MHC-I) in cancer cells can elicit immune evasion. Lotte et al. transduced a genome-wide sgRNA library into neuroblastoma cells bearing a NF-κB reporter, and then isolated NF-κB negMHC neg and NF-κB posMHC pos populations for sequencing. N4BP1 and TNIP1 were identified as inhibitors of NF-κB-mediated MHC-I expression in neuroblastoma, presenting potential targets to enhance the effect of therapeutic T cells [72]. In a recent study, cancer cells with no MHC-I expression were infected with a pooled lentiviral library and MHC-I high cells were enriched by FACS for analysis. The epigenetic repressive complex PRC2 was identified to silence MHC-I expression, providing a rationale for combining PRC2 inhibitors with immunotherapy to treat MHC-I-deficient malignancies [73].
sion of programmed death-1 ligand 1 (PD-L1) on cancer cells surface helps cancer cells evade immune destruction. Elucidating the molecular regulation of PD-L1 may propose targets to enhance tumor immunoreactivity. Marian et al. applied a genome-wide sgRNA library in cells with endogenous PD-L1 expression. Subsequently, PD-L1 low cells were enriched by FACS for sequencing and CMTM6 was identified as a key regulator of PD-L1. Mechanistic study revealed that CMTM6 co-localized with PD-L1 and prevented PD-L1 from lysosome mediated degradation [74].

2CT-CRISPR screen

The second screen strategy is conducted by coculturing mutagenized tumor cells with cytotoxic T lymphocytes (CTLs), usually CD8+ T cells, followed by sequencing of sgRNA representation (Figure 3B). Using the ‘two cell type’ (2CT) CRISPR assay, Shashank et al. profiled genes whose loss in tumor cells impaired the function of CTLs, and validated that loss-of-function mutations in APLNR rendered tumor cells resistant to T cell-mediated cytotoxicity [75]. On the contrary, Pan et al. focused on sensitive detection of depleted sgRNAs, and identified the inactivation of PBRM1, a chromatin regulator, sensitized tumor cells to T cell-mediated killing [76]. A recent study applied the coculture assay for high-throughput screen with small molecule library and genome-wide CRISPR library to identify both compounds and target genes that affect T-cell killing of tumor cells. EGFR inhibitor was the top compound, paralleled with the identification of sgRNAs targeting EGFR to sensitize tumor cells to T cell-mediated killing. This study provides a tool to rationally identify promising drug combinations to enhance immunotherapy [77].

Transplantable in vivo CRISPR screen

The third strategy is in vivo screen by transplantation of mutagenized tumor cells or T cells into mice, with immunotherapy treatment or not (Figure 3C and 3D). Robert et al. performed a pooled CRISPR knockout screen in transplantable tumors in mice treated with immunotherapy. They found that deletion of PTPN2 enhanced IFN-γ-mediated antigen presentation and growth inhibition, promoting immunotherapy efficacy [79]. Another group conducted an epigenetic-focused in vivo (transplantable) CRISPR screen using KrasG12D/P53-/- mouse lung cancer cell line. They revealed that loss of Asf1a in tumor cells induced immunogenic macrophage differentiation, promoting T cell activation in combination with anti-PD-1 treatment [80]. To facilitate the identification of T cell targets, Sidi Chen developed an adenoassociated viruses (AAVs)-Sleeping Beauty (SB) hybrid vector that enabled both efficient gene editing in primary murine T cells and genomic integration of sgRNA for screen readout. They transduced Cas9+CD8+ T cells with an AAV-SB library focusing on membrane protein, and adoptive transferred the mutant T cells into glioblastoma-engrafted mice via tail vein injection. The hits identified in this study may serve as targets for antibodies or T-cell engineering [81].

---

Table 4. Application of CRISPR knockout screen in identifying immunotherapy targets

<table>
<thead>
<tr>
<th>Cell type</th>
<th>Library</th>
<th>Selection context</th>
<th>Immunotherapy</th>
<th>Top hits</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>neuroblastoma (GIMEN)</td>
<td>custom library</td>
<td>MHC-I expression</td>
<td>-</td>
<td>N4BP1 and TNIPI</td>
<td>[72]</td>
</tr>
<tr>
<td>CML (K562)</td>
<td>Bassik Human CRISPR KO library</td>
<td>MHC-I expression</td>
<td>-</td>
<td>EED, SUZ12</td>
<td>[73]</td>
</tr>
<tr>
<td>pancreatic cancer (BxPC-3)</td>
<td>custom library</td>
<td>PD-L1 expression</td>
<td>-</td>
<td>CMTM6</td>
<td>[74]</td>
</tr>
<tr>
<td>melanoma (Me1624)</td>
<td>GeCK0v2</td>
<td>co-culture with CTL</td>
<td>-</td>
<td>APLNR</td>
<td>[75]</td>
</tr>
<tr>
<td>melanoma (B16F10)</td>
<td>CRISPR Brie lentiviral pooled library</td>
<td>co-culture with CTL</td>
<td>-</td>
<td>Pbrm1, Arid2, Bcd7</td>
<td>[76]</td>
</tr>
<tr>
<td>ovarian cancer (ID8)</td>
<td>custom library</td>
<td>co-culture with CTL</td>
<td>-</td>
<td>EGFR</td>
<td>[77]</td>
</tr>
<tr>
<td>melanoma (IFNFR1-KO D10)</td>
<td>GeCK0</td>
<td>co-culture with CTL</td>
<td>-</td>
<td>TNF pathway</td>
<td>[78]</td>
</tr>
<tr>
<td>melanoma (B165)</td>
<td>custom library</td>
<td>in vivo xenograft</td>
<td>anti-PD-1</td>
<td>Pipl2</td>
<td>[79]</td>
</tr>
<tr>
<td>lung cancer (KP lung cancer cell)</td>
<td>epigenetic-focused sgRNA library</td>
<td>in vivo xenograft</td>
<td>anti-PD-1</td>
<td>Asf1a</td>
<td>[80]</td>
</tr>
<tr>
<td>mouse primary CD8+ T cell</td>
<td>membrane protein-focused sgRNA library</td>
<td>adoptive transfer to GBM-engrafted mice</td>
<td>-</td>
<td>Pdia3, Mgtat5, Emp1, Lag3</td>
<td>[81]</td>
</tr>
<tr>
<td>mouse primary CD8+ T cell</td>
<td>MKO library</td>
<td>adoptive transfer to TNBC-engrafted mice</td>
<td>-</td>
<td>Dlx37</td>
<td>[82]</td>
</tr>
</tbody>
</table>

* CML, chronic myelocytic leukemia.
Direct in vivo CRISPR screen for cancer research

To date, most of the CRISPR screens are based on in vitro cultures or cellular transplant models, which are not in native tissue microenvironment involving complex interactions of multiple cell types. To more faithfully recapitulate the development of human cancer, researchers began to explore direct in vivo CRISPR screens, i.e. directly mutagenizing target tissues in vivo. However, CRISPR libraries are mostly constructed in lentiviral vectors, limiting their in vivo applications due to difficulties in delivery. Early attempt for large-scale in vivo CRISPR screen used piggyBac (PB) transposon as an alternative to deliver gRNA library. High-pressure tail vein injection of the PB-CRISPR-M2 library and pCAG-PBase induced mutagenesis in liver. Sequencing of tumors formed in
the liver revealed known and unknown tumor suppressor genes [83].

With advances in genetically engineered mouse models, Cas9 transgenic mice have been generated [84, 85], simplifying the delivery of CRISPR library and direct in vivo mutagenesis (Figure 4). Sidi Chen and colleagues performed direct in vivo CRISPR screen by intracranially or intravenously injecting AAV sgRNA library into conditional-Cas9 mice, efficiently inducing glioblastomas or liver tumors respectively. Subsequent capture sequencing of sgRNA target regions revealed functional landscape of tumor suppressors for gliomagenesis or liver tumorigenesis [86, 87]. Additionally, the application of CRISPR dCas9-activator mice will enable the functional identification of oncogenes in vivo [88].

Compared to in vitro or transplant-based approaches, direct in vivo CRISPR screen has the following advantages: (a) the native microenvironment of tumor is retained (b) the tumor entity is derived from the endogenous target tissue (c) the immune system remains. However, the direct in vivo CRISPR screens also have limitations. Random sampling errors from low viral transduction rate will lead to many false positive or negative results. Thus, the library size must be controlled to ensure adequate coverage in vivo. In addition, conventional cell transfection and lentiviral delivery methods often have difficulty achieving efficient gene targeting in vivo. In this regard, AAV delivery of sgRNA can achieve higher in vivo transduction efficiency and minimal immune rejection, but since AAV usually do not integrate into the genome, capture sequencing are used to reveal mutational profiles across tumors.

**Optimization strategies for CRISPR screen**

**Library size**

For high-throughput screen, the size of library is critical for the screen outcome. Genome-wide screen tend to provide more candidates and novel targets, but the large size of the library will lead to low coverage for each sgRNA in a fixed number of cells. A smaller focused library can achieve higher coverage for each sgRNA and improve data quality. Thus, in some studies, researchers used a genome-wide screen first, and then a secondary focused screen to achieve higher accuracy. If the initial study purpose is on specific pathway or particular biological process, such as cancer-related genes, kinases, membrane proteins or RNA binding proteins, a knowledge-based focused sgRNA library will be a better choice [43, 89]. In terms of in vivo screen, a focused library with more sgRNA targeting each gene is more suitable, for random sampling may result in low coverage. At present, ready-made genome-wide sgRNA libraries and sub-pool libraries with different functions for human and mouse are commercially available, such as in Addgene (https://

---

**Figure 4.** Schematics of direct in vivo CRISPR screen. The lentivirus and adenoassociated virus (AAV) approaches are used for CRISPR mutagenesis at target native organ site. Intravenous, intracranial and intratracheal injections of viral sgRNA library into conditional-Cas9 mice can drive tumorigenesis from the liver, brain and lung, respectively. Then the mice are subjected to MRI, histology, and deep sequencing (if lentivirus) or capture sequencing (if AAV) for sgRNA readout.
www.addgene.org/search/catalog/pooled-libraries), greatly saving labor and cost for researchers worldwide. These libraries are constantly being optimized and upgraded [90].

**Library design**

To improve the performance of sgRNA libraries, sgRNA sequence may be optimized for greater activity and less off-target effect. Main points of optimizing the design of sgRNA are summarized as follows: (1) reducing the complementary sequence of the 5’ end of sgRNA to 17-18nt, which can greatly reduce the off-target effect without affecting the targeting activity [91]. Because sgRNA can tolerate a mismatch at the 5’ end, but not at 3’ end [92]; (2) extension of the duplex length by about 5 bp, and mutating the fourth thymine in Pol III terminator (four consecutive thymines) to cytosine or guanine, which will significantly increase knockout efficiency [93]; (3) using RNA hairpins to construct modular scaffold RNA (scRNA) that can simultaneously activate and repress multiple genes in one cells [94].

**Cas9 variants selection**

Many efforts have also devoted to the development of Cas9 variants, such as eSpCas9 [95], SpCas9-HF1 [96], HypaCas9 [97]. In addition, researchers have constantly explored new CRISPR systems, such as CRISPR-Cas12a (also known as CRISPR-Cpf1) [98], CRISPR-Cas12b [99], CRISPR-CasX and CRISPR-CasY [100]. Cpf1 has unique multiplexing capabilities of genetic targeting, due to its independence from a tracrRNA [45, 47]. In this regard, this technology would enable high-throughput screening of mutation combinations. It has already been used to screen genetic interactions in tumor growth and metastasis in vivo [101] and remains to be extended to other aspects of cancer, such as identification of novel synergistic or synthetically lethal interactions.

**Delivery strategy**

The typical delivery strategy of CRISPR-Cas9 system into cells is plasmid vectors expressing Cas9 and sgRNA, either in an “all-in-one” vector that contains both Cas9 and sgRNA or a dual system that separates the two. Pre-selecting the best single clone expressing Cas9 with a high level of mutagenesis activity prior to transfection of the sgRNA library appears to be more advantageous for CRISPR screen, as it provides an unbiased background [102]. In contrast, the all-in-one approach may be more suited to screens in primary cells, which is not feasible to establish a stable expression of Cas9 clone in advance.

In addition to plasmid vectors, new delivery systems have been developed, including delivering Cas9 mRNA or directly Cas9 protein along with sgRNAs into cells. These new systems present transient Cas9 expression, thus achieving low off-target effect and saving time for vector construction and viral package. Michiko Kodama et al. performed a genetic screen by sgRNA lentiviral infection with Cas9 protein electroporation (SLICE), efficiently identifying functional gene targets in primary cells [103].

**Statistical analysis**

Data analysis is also a challenging aspect of CRISPR screens. There is no standard analysis pipeline exists yet, researchers have developed many bioinformatics tools for CRISPR screen to choose from, including MAGeCK [104], MAGeCK-VISPR [105], ScreenBEAM [106], BAGEL [107], CasTLE [108], ENCoRE [109], PBNPA [110] and JACKS [111]. Researchers can use one or multiple algorithms for data analysis to identify valuable hits according to their analysis needs. These algorithms are continuously being optimized. For example, Wang et al. combined the MAGeCK and MAGeCK-VISPR algorithms to develop a new algorithm called MAGeCKFlute and added downstream analysis functionalities, distinguished from other current tools [112]. In addition, the automatic improvement of algorithm performance through machine learning methods shows great potential [107, 113]. Recently, a tool called Mean Alterations Using Discrete Expression (MAUDE) has been created for data analysis of sorting-based CRISPR screens. MAUDE quantifies guide-level effects by modeling the distribution of cells across sorting expression bins [114]. It is believed that the continuous development of artificial intelligence will bring more convenience to data analysis.

**Conclusions and perspectives**

The versatility and specificity of CRISPR screen renders it a promising player in genetic researches. It has been applied to various cancer genomic studies, such as cancer cell growth,
metastasis, synthetic lethal interactions, therapeutic resistance and immunotherapy response. Future improvements in library selection, sgRNA/Cas9 design, delivery strategy and statistical analysis will help to improve the efficiency of CRISPR screen.

A recent study has revealed that CRISPR screens in 3D more accurately recapitulated the in vivo tumors than in 2D [115]. CRISPR screens have also been extended to 3D intestinal or colon organoids, which better recapitulate human colorectal cancer. Michels et al. performed a tumor suppressor gene (TSG)-focused CRISPR screen in pre-malignant human APC-/-; KRASG12D organoids in vitro and in vivo. They identified TGFBR2 as the most prevalent TSG, followed by known and uncharacterized mediators of colorectal cancer growth [116]. However, applying genome-scale CRISPR screen to organoids is hampered by technical limitations and requirement of extensive cell numbers. In this regard, Ringel et al. improved the accuracy and robustness for CRISPR screens by capturing sgRNA integrations in single organoids, substantially reducing required cell numbers for genome-scale screening [117]. In the future, CRISPR screen may be broadly applied to various organoid models and open new avenues for genetically dissecting mechanisms of human disease.

Recently, integrated approach combining CRISPR screens with single-cell RNA sequencing (scRNA-seq) has been developed for profiling the perturbation and transcriptome in the same cell [103, 118]. Researchers have even developed “Direct-seq”, a framework to combine scRNA-seq with CRISPR screen by introducing a capture sequence into the gRNA scaffold, leading to a streamlined workflow for readouts of CRISPR perturbation and transcriptome [119]. Roth et al. developed a pooled knockin screen method that integrated non-viral DNA templates into the T cell receptor (TCR)-locus in human T cells. They further developed pooled knockin sequencing (PoKl-seq), combining scRNA-seq and pooled knockin screening to measure cell abundance in vitro and in vivo. The pooled knockin screen presents a new CRISPR platform to enable scalable gain-of-function screening, and to test knockin candidate therapeutic constructs in primary human cells, which would accelerate the development of cellular therapies [120].

With further improvements and modifications, novel CRISPR screen formats have been developed. Thomas et al. presented a hybrid Cas9-Cas12a platform, named Cas Hybrid for Multiplexed Editing and screening Applications (CHyMErA), to achieve combinatorial genetic manipulation. This system outperforms genetic screens using Cas9 or Cas12a alone. Application of CHyMErA achieves systematic mapping of genetic interactions and uncovers phenotypes normally masked by functional redundancy [121]. Cas13 has been reported to cleave single-stranded RNA targets with complementary spacers in mammalian cells. Chen et al. identified that RfxCas13d enabled effective and specific knockdown of circRNAs by using gRNAs targeting sequences spanning backsplicing junction (BSJ) sites in RNA circles. They further proved RfxCas13d-BSJ-gRNA screening as a useful tool to discover functional circRNAs in a large-scale level [122].

In conclusion, CRISPR screen provides a practical and high-throughput way for functional genomic studies in cancer. As technologies continue to develop, we envisage that CRISPR screen will accelerate researches on the functional characterization of genetic elements and the identification of novel therapeutic targets.

Acknowledgements

This work was supported by grants from National Natural Science Foundation of China (81872042 and 81972333), Foundation of Jiangsu Province (BK20181238) and Jiangsu Postdoctoral Research Foundation (1701156C).

Disclosure of conflict of interest

None.

Abbreviations

CRISPR-Cas9, clustered regularly interspaced short palindromic repeats-CRISPR associated nuclease 9; CRISPRi, CRISPR inhibition; CRISPRa, CRISPR activation; sgRNA, single-guide RNA; DSB, double-strand break; NHEJ, nonhomologous end joining; NGS, next-generation sequencing; SAM, synergistic activation mediator; TSS, the transcriptional start site; lncRNA, long non-coding RNA; pgRNA, paired-guide RNA; FACS, fluorescence-activated cell sorting; PD-L1, programmed death-1 ligand 1; CTLs, cytotoxic T lymphocytes; AAVs, adenoassociat-
CRISPR screen in cancer research

ed viruses; SB, Sleeping Beauty; scRNA, scaffold RNA; SLICE, sgRNA lentiviral infection with Cas9 protein electroporation.

Address correspondence to: Drs. Xiaoyuan Chu and Cheng Chen, Department of Medical Oncology, Jinling Hospital, The First School of Clinical Medicine, Southern Medical University, 305 East Zhongshan Road, Nanjing 210002, China. Tel: +86-25-8086-4750; Fax: +86-25-80864750; E-mail: chuxiaoyuan000@163.com (XYC); chencheng1289@126.com (CC).

References


CRISPR screen in cancer research


CRISPR screen in cancer research


CRISPR screen in cancer research


CRISPR screen in cancer research


CRISPR screen in cancer research


