the FEBS Journal

STATE-OF-THE-ART REVIEW

High-throughput screens in mammalian cells using the CRISPR-Cas9 system

Jingyu Peng, Yuexin Zhou, Shiyou Zhu and Wensheng Wei

Biodynamic Optical Imaging Centre (BIOPIC), Peking-Tsinghua Center for Life Sciences, State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences, Peking University, Beijing, China

Keywords
CRISPR-Cas9 system; high-throughput; knockout; screening; sgRNA

Correspondence
W. Wei, Biodynamic Optical Imaging Centre (BIOPIC), Peking-Tsinghua Center for Life Sciences, State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences, Peking University, Beijing 100871, China
Fax: +86 10 62757131
Tel: +86 10 62757227
E-mail: wswei@pku.edu.cn

(Received 1 December 2014, revised 14 February 2015, accepted 26 February 2015)

doi:10.1111/febs.13251

As a powerful genome-editing tool, the clustered regularly interspaced short palindromic repeats (CRISPR)-clustered regularly interspaced short palindromic repeats-associated protein 9 (Cas9) system has been quickly developed into a large-scale function-based screening strategy in mammalian cells. This new type of genetic library is constructed through the lentiviral delivery of single-guide RNA collections that direct Cas9 or inactive dead Cas9 fused with effectors to interrogate gene function or regulate gene transcription in targeted cells. Compared with RNA interference screening, the CRISPR-Cas9 system demonstrates much higher levels of effectiveness and reliability with respect to both loss-of-function and gain-of-function screening. Unlike the RNA interference strategy, a CRISPR-Cas9 library can target both protein-coding sequences and regulatory elements, including promoters, enhancers and elements transcribing microRNAs and long noncoding RNAs. This powerful genetic tool will undoubtedly accelerate the mechanistic discovery of various biological processes. In this mini review, we summarize the general procedure of CRISPR-Cas9 library mediated functional screening, system optimization strategies and applications of this new genetic toolkit.

Introduction

Gene knockout is the ultimate strategy for investigating functions of transcripts and pathways critical for biological processes and disease mechanisms [1]. Systematic knockouts in yeast have been successfully employed to study how genotype directs phenotype [2,3]. It has been a daunting task, however, to create large- or genome-scale loss-of-function mutations in mammalian cells as a result of the diploid or even polyploid nature of eukaryotic genomes. One way to introduce genetic mutations in mammalian cells is to use the standard homologous recombination (HR) technique to target both alleles in mouse embryonic stem cells [4]. However, this procedure is time-consuming and has low efficiency. By contrast, RNA interference (RNAi) used to be the most favourable choice for genome-wide screening to characterize gene function because small interfering RNAs or short hairpin RNAs (shRNAs) can potentially give rise to detectable

Abbreviations
6-TG, 6-thioguianine; Cas9, clustered regularly interspaced short palindromic repeats-associated protein 9; CRISPRa, CRISPR activation; CRISPR, clustered regularly interspaced short palindromic repeats; CRISPRi, CRISPR interference; crRNA, clustered regularly interspaced short palindromic repeats RNA; dCas9, dead clustered regularly interspaced short palindromic repeats-associated protein 9; DSB, double-strand break; HR, homologous recombination; indel, insertion or deletion; MOI, multiplicity of infection; NHEJ, nonhomologous end joining; RNAi, RNA interference; sgRNA, single-guide RNA; shRNA, short hairpin RNA; tracrRNA, trans-activating clustered regularly interspaced short palindromic repeats RNA.

phenotypic changes by repressing gene expression through sequence-specific degradation of mRNA regardless of the copy number of the target gene [5]. Nevertheless, the RNAi approach has inherent limitations because of its off-target feature and the fact that partial suppression of gene expression is often not sufficient to create noticeable changes in phenotype, resulting in high false-negative and false-positive rates. An alternative technique has emerged that takes advantage of the human near-haploid cell line KBM7 to implement genetic screening using the gene-trap system [6]. However, the application of the haploid system has been greatly hindered by the instability of its karyotype and the technical difficulty of identifying targeted loci. Taken together, there remains an urgent need for a more reliable and highly efficient screening strategy for large-scale target identification in mammalian cells.

With the advance of genome-editing technologies, various tools have been developed to effectively introduce mutations at specific loci in eukaryotes [7]. Although zinc-finger nucleases and transcription activator-like effector nucleases have been successfully employed to modify loci with precision, they are not suitable for generating large-scale knockout events because of technical challenges in library construction and conflicts with lentiviral delivery [8–10]. By contrast, the clustered regularly interspaced short palindromic repeats (CRISPR)-clustered regularly interspaced short palindromic repeats-associated protein 9 (Cas9) system, derived from the adaptive immune system in *Streptococcus pyogenes* bacteria, demonstrates unique advantages for genome engineering [11–14]. The natural CRISPR-Cas9 system works through the type II Cas9 nuclease and two individual RNA components, CRISPR RNA (crRNA) and trans-activating crRNA (tracrRNA), which have been fused experimentally as a single-guide RNA (sgRNA) for ease of use in genome engineering [11,13,14]. The specific sgRNA directs the Cas9 protease to the matching DNA sequence to create double strand breaks (DSBs) that recruit native DNA repair mechanisms, HR [15] or error-prone nonhomologous end joining (NHEJ) [16], resulting in indels (insertions and deletions) at the targeted site (Fig. 1). Through Watson–Crick base pairing [11], the sgRNA directs Cas9 nuclease to the targeted genomic locus followed by the protospacer adjacent motif to create DSBs, thus allowing for a much greater simplicity of vector construction. The high efficiency of indels induced by the CRISPR-Cas9 system ensures bi-allelic or multi-allelic mutations in eukaryotes, paving the way for the creation of large-scale mutations.

In this review, we discuss the principles and prospects of CRISPR-based screening, especially the first four studies [17–20] that have demonstrated the feasibility of high-throughput CRSIPR-Cas9 screening in mammalian cells, providing the hallmark for a new era of functional genomics. Potential optimization of CRISPR-based screening, as well as new technical advances, such as CRISPR interference (CRISPRi)- and CRISPR activation (CRISPRa)-mediated gene identification [21,22], will also be discussed.

Case study of screens using the CRISPR system

Five groups have published four studies using the CRISPR-Cas9 system to conduct large-scale knockout screens in eukaryotic cells. Wang *et al.* [17] built a library of 73 151 sgRNAs targeting 7113 genes in two cell lines, the near-haploid line KBM7 and the diploid line HL-60. This library was used to identify essential genes related to cell growth and genes involved in resistance to the nucleotide analogue 6-thioguianine (6-

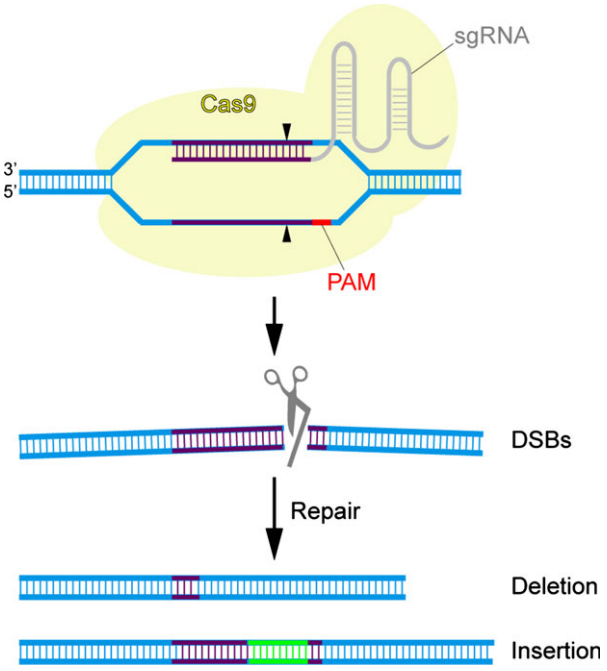

Fig. 1. Genome editing mediated by the type II CRISPR-Cas9 system. sgRNA directs Cas9 endonuclease to the target locus through its guide RNA (purple). Cas9 causes DSBs (black arrows) via its interaction with the protospacer-adjacent motif (PAM; red) located on the strand complementary to the sgRNA binding sequence. DSBs induce repair systems, especially the error-prone NHEJ pathway, resulting in indels of variable length and, consequently, the complete disruption of gene expression.

TG) (concerning DNA mismatch repair function) and to the chemotherapy drug etoposide. They were able to obtain known important genes from both negative (for candidates that are significantly decreased or even lost) and positive screens (for candidates that are enriched). Shalem *et al.* [18] reported a one-vector library consisting of 64 751 unique sgRNAs that target 18 080 human genes, and this was used to screen for essential genes involved in cell growth in human embryonic stem cell line HUES72 and resistance to the chemotherapeutic agent vemurafenib in melanoma cell line A375. They also demonstrated that CRISPR knockout screening is highly efficient and has a lower false discovery rate in comparison with the shRNA strategy.

During the same period, Koike-Yusa *et al.* [19] reported their genome-wide screen in mouse embryonic stem cell line JM8 including 87 897 sgRNAs targeting 19 150 mouse genes. They obtained all known positive genes involved in cell resistance to 6-TG or *Clostridium septicum* α-toxin, demonstrating the effectiveness of the screen. Our group has also reported a study using a rather small library, 869 sgRNAs targeting 291 genes, for identifying genes important for the cellular toxicity of chimaeric anthrax toxin and diphtheria toxin [20]. In our screen, we identified known genes encoding receptors of these two toxins that ranked at the very top of the list, and we also obtained and validated novel hits, indicating that the screening of a small knowledge-based library could yield high quality data with outstanding efficiency. Table 1 summarizes the key information obtained from these four studies.

Although the library size and the cell type vary in different studies, these four studies unanimously chose the lentiviral system to deliver the sgRNA into cells at a low multiplicity of infection (MOI). A lower MOI indicates a lower chance that two or more types of sgRNA will enter the same cell, which is critical for ensuring low levels of misattribution of sgRNAs. Importantly, the ease of the screening is inversely associated with the library size. It might be technically and economically challenging to ensure a fair chance for every sgRNA in the library to create a gene knockout in genome-scale screening. It is therefore beneficial to choose a smaller sized library, if possible, to improve the quality of the screen, especially for knowledge-based studies [20].

There are two ways to express Cas9 nuclease in cells for library screening: the one-vector system in which a Cas9-expressing cassette and the sgRNA are located in the same lentiviral backbone, as proposed by Zhang *et al.* [18], and the two-component system in which Cas9-expressing cells are pre-generated before the introduction of the sgRNA library [17,19,20,23]. In particular, Wang *et al.* [17] also used an inducible Cas9-expressing vector that could eliminate potential effects of the constitutive expression of Cas9.

After the construction of a gene knockout library in target cells and the corresponding phenotypic screening, all studies [17,19,20,23] employed the deep sequencing technique to decode the enriched sgRNAs, similar to the pooled screening used for shRNA libraries [24]. Deep sequencing, also known as next generation sequencing, provides a cost-effective method to massively acquire parallel sequencing data within a short period [24,25]. Different methods have been chosen to analyze sequence data and obtain a ranked list of candidates, including those that were used by the studies discussed here [26–28] (Table 1). A new analytic tool has also been developed specifically for CRISPR-mediated screening [29].

In general, the CRISPR screening protocol can be divided into two sections: library construction in the chosen cell line and functional screening based on a specific biological assay followed by high-throughput sequencing analysis (Fig. 2). The quality of the library and the effectiveness of the phenotypic assay are both key to the success of a library screen.

Ways to improve the quality of a CRISPR library

- Design of sgRNAs
- Optimization of the sgRNA scaffold
- Selection of a Cas9-expressing clone from the chosen cell line

The first step toward sgRNA library construction is the design of sgRNAs for oligo synthesis. The efficacy with which the sgRNAs cause DSBs and consequently indels is apparently critical for the library's quality. According to large-scale evaluation studies, there are certain rules worthy of notice. For example, extreme GC content might weaken the effectiveness of an sgRNA, and sgRNAs containing homopolymers, especially 'UUUUs', likely have lower efficiency [21]. Any difference in the efficiency of an sgRNA depending on whether it targets the sense or template strand remains to be clarified. Wang *et al.* [17] concluded that targeting the sense strand was less effective than the template strand based on an evaluation of 2741 sgRNAs targeting 43 positive ribosomal genes essential for the cell growth, whereas Gilbert *et al.* [21] suggested that targeting either strand of DNA is equally effective based on data on 49 000 sgRNAs targeting 49 positive genes essential for cell resistance to ricin. Although

Table 1. Comparison of the CRISPR knockout screens.

	Wang *et al.* [17]		Shalem *et al.* [18]		Koike-Yusa *et al.* [19]		Zhou *et al.* [20]	
Species	Human		Human		Mouse		Human	
Cell line	KBM7 HL60		A375 HUES72		Mouse JM8 ESCs		HeLa	
sgRNAs	73,151		64,751		87,897		869	
Target genes	7114		18,080		19,150		291	
sgRNAs/gene	10		Average 3–4		Average 4–5		2–3	
Coverage	NA		15×		25×		1000×	
MOI	Low		0.3–0.4		0.2–0.3		0.025–0.05	
Cas9 expression	Inducible/constitutive, single clone		Constitutive, together with sgRNA		Constitutive, single clone		Constitutive, single clone	
Marker of library	Blasticidin		Puromycin		BFP		GFP	
Screening aims and assays	Cell proliferation	Resistance to 6-TG or etoposide	Cell proliferation	Resistance to PLX	Resistance to α-toxin	Resistance to 6-TG	Resistance to chimeric anthrax toxin	Resistance to diphtheria toxin
Results	Many	4 and 2 known	Many	2 known, 4 novel	16 known, 4 novel	5 known	1 known, 4 novel	1 known, 1 novel
Analysis tool	Gene set enrichment analysis [26]		RIGER algorithm [27]		N/A		DESeq2 [28]	

ESCs, embryonic stem cells; NA, not available; BFP, blue fluorescent protein; GFP, green fluorescent protein; PLX, BRAF protein kinase inhibitor vemurafenib; 6-TG, 6-thioguianine.

some mathematical suggestions for sgRNA design have been proposed by bioinformaticians [30], a mature model is expected especially after copious experimental data are available.

A modification of the sgRNA scaffold may potentially improve its gene targeting efficiency. The tracrRNA tail of sgRNA is important for Cas9-sgRNA mediated DSBs, and an increase of tracrRNA length was reported to boost the efficiency of sgRNA in generating DSBs [31]. Extension of the sgRNA stem loop was also shown to enhance the assembly of sgRNA with Cas9 protein, resulting in an improved NHEJ-mediated mutation rate at the targeted sequence [32]. Furthermore, the existence of a potential Pol III terminator (four consecutive Us) in the stem loop of an sgRNA might cause the pre-mature termination of sgRNAs transcribed from the U6 promoter [33]. Taken together, a modified sgRNA structure with an A-U flip in the Pol III terminator sequence and a hairpin structure extension (Fig. 3A) was reported to promote sgRNA stability and dead Cas9 (dCas9) association in a study of CRISPR-Cas system-based dynamic imaging [34]. Aiming to determine whether these modifications would also improve the efficiency of sgRNA recruitment of Cas9 and cleavage of matching sequences, we randomly chose seven targeting sites on five genes and compared the indels caused by these sgRNAs without and with (designated as sgRNA2.0) the above scaffold modifications. The indels were increased for five sgRNAs containing these modifications, suggesting that sgRNA2.0 might improve the efficiency of sgRNA at creating NHEJ-mediated gene mutations, although the other two low-efficiency sgRNAs did not show a difference (Fig. 3B). It is highly possible that there is still room for further optimization of the scaffold of sgRNA to increase its stability and/or binding affinity for Cas9.

The expression of the Cas9 nuclease presumably affects the mutation efficiency mediated by specific sgRNA sequences on the chromosome. Pre-generating a Cas9-expressing cell line prior to the sgRNA library construction appears to be advantageous for library screening because it provides an unbiased background for the screening as indicated above, albeit with a longer preparation time compared to the one-vector system. We would like to emphasize that the clonal variation of Cas9-expressing cells in creating DSBs is often not trivial [20]. It is therefore critical to pre-select the best single clone for library construction and the subsequent screening. In addition to HeLa cells, we also observed similar phenomena in Huh7.5 [35], HEK293T and HT1080 cell lines (data not shown) in which a selected clone showed a higher level of mutagenesis activity than the others. Although it is inconclusive whether the clonal variation is exclusively a result of differences in Cas9 expression, it is advantageous in principle to start the library from a selected cell clone.

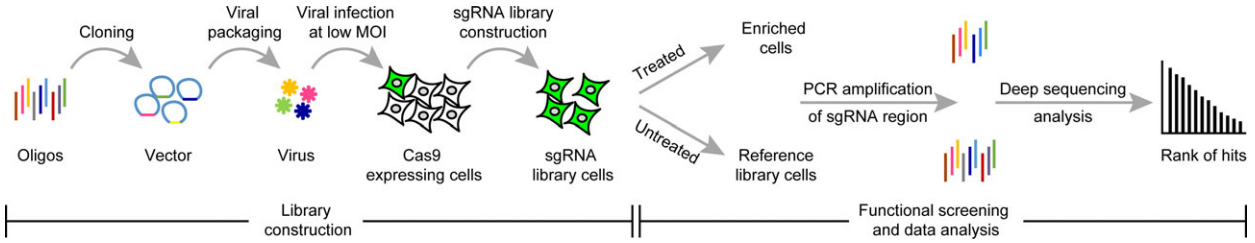

Fig. 2. The common procedure of CRISPR-based library construction and a pooled approach for function-based screening in mammalian cells. After design and synthesis, the sgRNA-encoding sequences are PCR-amplified and cloned into a lentiviral backbone under the control of a Pol III promoter such as U6. The lentiviruses are produced and titrated through the standard packaging process. After viral infection (at a low MOI) of a pre-generated Cas9-expressing cell line, an sgRNA cell library is established through either antibiotics or fluorescence selection. Prolonged culturing for 7–14 days is recommended for the sgRNA cell library to maximize the chance of gene knockout. The library cells undergo customized treatment based on a phenotype of interest, and the selected cell pools are harvested for genomic DNA extraction. The original library cells (untreated) are commonly used as a reference. sgRNA-coding fragments are PCR-amplified from genomic DNA, and they are subsequently decoded through high-throughput sequencing analysis. Deep sequencing data provide a ranking of the screened sgRNAs, revealing the biological relevance and significance of their corresponding genes in the phenotype of interest.

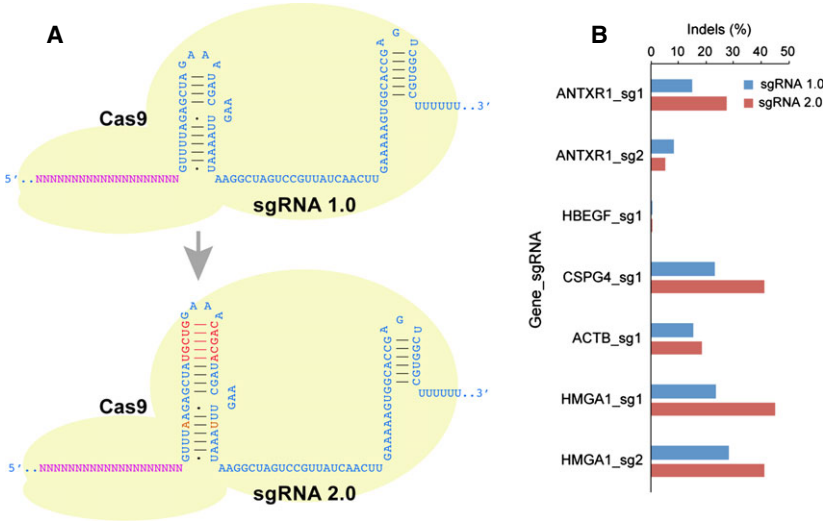

Fig. 3. Effect of a modified scaffold (sgRNA2.0) on the efficiency of sgRNA at creating sequence-specific indels. (A) Modified sgRNA2.0 structure with an A-U flip and hairpin extension relative to the conventional structure (sgRNA1.0). (B) Comparison of indels caused by seven sgRNAs at the targeted loci between sgRNA1.0 and sgRNA2.0 scaffolds.

Focused library versus genome-wide library

For high-throughput screening, the size of the library is always a critical factor affecting the screening outcome. It is generally more appealing to conduct genome-scale screening because there would be more chances of identifying genes of interest. However, a genome-wide sgRNA library usually contains more than 100 000 sgRNAs depending on the average number of sgRNAs designed for each gene. The screening of a large-size library is both labour- and cost-intensive. Given a fixed number of cells for screening, each individual sgRNA receives less coverage as more sgRNAs are included in the library. Thus, a knowledge-based focused library with a limited number of sgRNAs would be an alternative and, occasionally, a better choice. If a study is specifically investigating a particular group of targets, such as kinases or membrane proteins, it would be more efficient to screen a small-sized library than to conduct a genome-wide screen. Moreover, it is much easier to achieve high coverage for each sgRNA with smaller sized library, resulting in improved data quality.

Although the off-target effects caused by the lentiviral delivery of sgRNA and Cas9 in the CRISPR

library are not severe, false positive results do occur as a result of the misattribution of two or more different sgRNAs in one cell [17–20]. In addition, the efficiency of sgRNA targeting the same gene varies. To maximize the chance of knockout events for any given gene, more sgRNAs are required for library construction. We have therefore purposely designed an average of 10 sgRNAs targeting each gene so that we could combine both functional screening and candidate validation in one step. In light of this, we categorized all human genes into 10 sub-libraries based on their annotated functions or localization (Table 2). Some genes without annotation were arbitrarily distributed into different groups, and there is also a certain percentage of overlapping of sgRNAs among groups because of the multi-functionality of their targeted genes. We hope that this type of categorization will offer added flexibility for usage of the library.

Application of CRISPR-Cas9 screening

CRISPR-Cas9 library screening enables the identification of critical components in a variety of biological processes in mammals. Two criteria must be fulfilled to ensure the success of the screen: the quality of the lentiviral delivered sgRNA library and an appropriate assay or treatment that separates or enriches mutant cells with the expected phenotypic change. Cell death or cell growth is an ideal type of selection that usually gives rise to lower false-positive rates [17–20]. If the phenotype of interest does not lead to drastic changes in cell viability, a specific reporter system is often required to establish the screen. For example, a fluorescence signal is a popular choice, especially if the desired mutation specifically switches on a fluorescence signal that could be sorted by flow cytometry.

In addition to CRISPR-mediated gene knockout screening, a CRISPRi library has also been developed for functional genomics based on the suppression of transcripts [21]. Taking advantage of a catalytically inactive Cas9 (dCas9) fused with the transcription suppression domain of KRAB [36], the CRISPRi method appears superior to the RNAi strategy, and it is particularly useful when studying genes essential for cell growth. The off-target frequency has been further reduced in CRISPRi strategy. In addition, CRISPRa libraries have also been developed to conduct gain-of-function screens [21,37].

Despite the broad application of large-scale CRISPR screens to study gene function on a massive scale, this technique still has inherent limitations because it relies on single sgRNA-mediated gene knockout, suppression or activation. In the study of regulatory DNA elements whose sequences are very long, such as long noncoding RNAs and super enhancers, a single-site mutagenesis or frameshift often does not lead to the inactivation of these regulatory elements. It is therefore highly desirable to develop such libraries in which sgRNAs work in pairs to generate potential deletions of a targeted region. Encouragingly, the delivery of two sgRNAs has been reported to program Cas9 to create a precise deletion [13,38,39].

Summary

As a powerful genome-editing tool, the CRISPR-Cas9 system has revolutionized the way most biological laboratories conduct their research. Now with the

Table 2. The design of ten human sgRNA sub-libraries. Because of the multifunctional nature of proteins, different sub-libraries overlap in sgRNAs in various degrees. In total, 241 883 sgRNAs (not counting negative control sgRNAs) were synthesized to construct the 10 sub-libraries, which contain 178 267 unique sgRNAs targeting 18 852 genes in total.

Sub-library	Category of sgRNA-targeting gene products (based on annotated function, localization, etc.)	Gene	sgRNA	Control sgRNA
Sub-01	Calcium-binding protein, chaperone, cytoskeletal protein, storage protein, structural protein, cell cycle, etc.	2382	22 509	150
Sub-02	Cell adhesion molecule, cell junction protein, extracellular protein, membrane traffic protein, surfactant, etc.	1924	18 213	150
Sub-03	Defense immunity protein, signaling molecule, etc.	2451	23 056	150
Sub-04	Enzyme modulator, peptidase, etc.	1795	17 439	150
Sub-05	Hydrolase, protease, unknown function	2971	25 889	150
Sub-06	Isomerase, kinase, ligase, lyase, phosphatase, oxidoreductase, etc.	2467	24 092	150
Sub-07	Nucleic acid binding proteins	2487	24 269	150
Sub-08	Receptor, transmembrane receptor regulatory adaptor protein, other membrane proteins	2989	28 811	150
Sub-09	Transcription factor, nucleic proteins	3289	32 056	150
Sub-10	Transfer carrier protein, transferase, transporter	2609	25 549	150

establishment of high-throughput screening strategies, we foresee explosive progress in the study of functional genomics related to broad biological processes and disease mechanisms. Despite its unprecedented power, the customized construction of a CRISPR-Cas9 library is relatively easy. Similar to its predecessor RNAi technique, CRISPR libraries can be applied to both positive and negative screens. Current screening strategies all share a common workflow, including CRISPR library construction through lentiviral delivery of sgRNAs into Cas9-expressing cells, biology assay-based screening, and the decoding of sgRNA enrichment through high-throughput sequencing analysis. There are at least three ways to further improve the quality of a CRISPR screen: better design of the sgRNAs, optimization of the sgRNA scaffold, and selection of the best pre-generated Cas9-expressing cell clone for the library construction. In addition, it is important to choose wisely between a genome-scale library versus a knowledge-based focused library. The application of CRISPR-Cas9 screening in mammalian cell lines will undoubtedly accelerate research into understanding important diseases and biological mechanisms. We anticipate an explosive research advance driven by CRISPR-Cas9 screening in mammalian cells and exponential growth of information on the functional characterization of genetic elements in the years to come.

Acknowledgements

We thank all members of the Wei laboratory for their helpful discussions. This work was supported by funds from the National Science Foundation of China (NSFC31170126, NSFC31430025, NSFC81471909) and the Peking-Tsinghua Centre for Life Sciences.

Author contributions

P.J., Z.Y. and Z.S. summarized the literature. P.J. and W.W. prepared the tables. Z.Y and Z.S performed the experiments. P.J., Z.Y., Z.S. and W.W. prepared the figures and wrote the paper.

References

1 Liberali P, Snijder B & Pelkmans L (2015) Single-cell and multivariate approaches in genetic perturbation screens. *Nat Rev Genet* **16**, 18–32.
2 Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B *et al.* (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**, 387–391.
3 Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, Ding H, Koh JL, Toufighi K, Mostafavi S *et al.* (2010) The genetic landscape of a cell. *Science* **327**, 425–431.
4 Capecchi MR (2005) Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nat Rev Genet* **6**, 507–512.
5 Berns K, Hijmans EM, Mullenders J, Brummelkamp TR, Velds A, Heimerikx M, Kerkhoven RM, Madiredjo M, Nijkamp W, Weigelt B *et al.* (2004) A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* **428**, 431–437.
6 Carette JE, Guimaraes CP, Varadarajan M, Park AS, Wuethrich I, Godarova A, Kotecki M, Cochran BH, Spooner E, Ploegh HL *et al.* (2009) Haploid genetic screens in human cells identify host factors used by pathogens. *Science* **326**, 1231–1235.
7 Gaj T, Gersbach CA & Barbas CF 3rd (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* **31**, 397–405.
8 Moscou MJ & Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. *Science* **326**, 1501.
9 Urnov FD, Rebar EJ, Holmes MC, Zhang HS & Gregory PD (2010) Genome editing with engineered zinc finger nucleases. *Nat Rev Genet* **11**, 636–646.
10 Holkers M, Maggio I, Liu J, Janssen JM, Miselli F, Mussolino C, Recchia A, Cathomen T & Goncalves MA (2013) Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic Acids Res* **41**, e63.
11 Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA & Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821.
12 Cho SW, Kim S, Kim JM & Kim JS (2013) Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* **31**, 230–232.
13 Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823.
14 Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE & Church GM (2013) RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826.
15 Capecchi MR (1989) Altering the genome by homologous recombination. *Science* **244**, 1288–1292.
16 Bibikova M, Golic M, Golic KG & Carroll D (2002) Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. *Genetics* **161**, 1169–1175.

17 Wang T, Wei JJ, Sabatini DM & Lander ES (2014) Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84.
18 Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, Heckl D, Ebert BL, Root DE, Doench JG *et al.* (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87.
19 Koike-Yusa H, Li Y, Tan EP, Velasco-Herrera Mdel C & Yusa K (2014) Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat Biotechnol* **32**, 267–273.
20 Zhou Y, Zhu S, Cai C, Yuan P, Li C, Huang Y & Wei W (2014) High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. *Nature* **509**, 487–491.
21 Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C, Panning B, Ploegh HL, Bassik MC *et al.* (2014) Genome-scale CRISPR-mediated control of gene repression and activation. *Cell* **159**, 647–661.
22 Zalatan JG, Lee ME, Almeida R, Gilbert LA, Whitehead EH, La Russa M, Tsai JC, Weissman JS, Dueber JE, Qi LS *et al.* (2015) Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* **160**, 339–350.
23 Sanjana NE, Shalem O & Zhang F (2014) Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* **11**, 783–784.
24 Sims D, Mendes-Pereira AM, Frankum J, Burgess D, Cerone MA, Lombardelli C, Mitsopoulos C, Hakas J, Murugaesu N, Isacke CM *et al.* (2011) High-throughput RNA interference screening using pooled shRNA libraries and next generation sequencing. *Genome Biol* **12**, R104.
25 Aird D, Ross MG, Chen WS, Danielsson M, Fennell T, Russ C, Jaffe DB, Nusbaum C & Gnirke A (2011) Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol* **12**, R18.
26 Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* **102**, 15545–15550.
27 Luo B, Cheung HW, Subramanian A, Sharifnia T, Okamoto M, Yang X, Hinkle G, Boehm JS, Beroukhim R, Weir BA *et al.* (2008) Highly parallel identification of essential genes in cancer cells. *Proc Natl Acad Sci USA* **105**, 20380–20385.
28 Anders S & Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* **11**, R106.
29 Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, Irizarry RA, Liu JS, Brown M & Liu X (2014) MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol* **15**, 554.
30 Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, Sullender M, Ebert BL, Xavier RJ & Root DE (2014) Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol* **32**, 1262–1267.
31 Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O *et al.* (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* **31**, 827–832.
32 Jinek M, East A, Cheng A, Lin S, Ma E & Doudna J (2013) RNA-programmed genome editing in human cells. *eLife* **2**, e00471.
33 Nielsen S, Yuzenkova Y & Zenkin N (2013) Mechanism of eukaryotic RNA polymerase III transcription termination. *Science* **340**, 1577–1580.
34 Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, Park J, Blackburn EH, Weissman JS, Qi LS *et al.* (2013) Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**, 1479–1491.
35 Ren Q, Li C, Yuan P, Cai C, Zhang L, Luo G & Wei W (2015) A dual-reporter system for real-time monitoring and high-throughput CRISPR/Cas9 library screening of the hepatitis C virus. *Sci Rep* **5**, 8865.
36 Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA *et al.* (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451.
37 Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu PD, Habib N, Gootenberg JS, Nishimasu H *et al.* (2014) Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583–588.
38 Canver MC, Bauer DE, Dass A, Yien YY, Chung J, Masuda T, Maeda T, Paw BH & Orkin SH (2014) Characterization of genomic deletion efficiency mediated by CRISPR/Cas9 in mammalian cells. *J Biol Chem* **289**, 21312–21324.
39 Zheng Q, Cai X, Tan MH, Schaffert S, Arnold CP, Gong X, Chen CZ & Huang S (2014) Precise gene deletion and replacement using the CRISPR/Cas9 system in human cells. *Biotechniques* **57**, 115–124.

Podcast

This article is accompanied by a podcast, listen now. Or listen in iTunes.