

Dissecting cellular ecosystem with single-cell CRISPR screens

Zhirui Liu^{a,b,c}, Zilin Lan^{a,b,c}, Xiaoli Kang^b, Yao Yao^{a,b,c}, Shuquan Rao^{a,b,c,*}

^aState Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Tianjin 300020, China; ^bHaihe Laboratory of Cell Ecosystem, Tianjin 300051, China; ^cTianjin Institutes of Health Science, Tianjin 301600, China

Abstract

Clustered regularly interspaced short palindromic repeats (CRISPR) screens represent a transformative force in biological discovery, enabling the unbiased interrogation of gene function in a wide range of applications. Traditional screening approaches predominantly hinge on cell fitness or established markers, which inherently constrain their abilities for unbiased biological discovery. By contrast, single-cell CRISPR screening technologies, which combine pooled CRISPR screens with an array of sophisticated single-cell omics platforms, permit comprehensive profiling of the transcriptome and epigenome following individual genetic manipulations within complex cellular ecosystems. Over the past decade, a panoply of single-cell CRISPR platforms has emerged, each tailored to address specific experimental challenges. Iterative refinements in protocols have bolstered precision, scalability, and reproducibility, thereby enormously advancing functional genomics and translational research. However, technical obstacles such as perturbation efficiency, scalability, and data integration persist, necessitating cross-disciplinary collaboration and innovation. As single-cell CRISPR platforms evolve to incorporate spatial resolution, multi-omics integration, and AI-guided design, they are poised to bridge the gap between genetic perturbation and system-level interpretation. Here, we summarize recent advances in single-cell CRISPR technologies, outline their applications, and provide a comparative framework to guide platform selection (Perturb-seq, CROP-seq, ECCITE-seq, Direct-seq, and Mosaic-seq).

Key Words: Functional genomics; Single-cell CRISPR screen

The clustered regularly interspaced short palindromic repeat (CRISPR) renaissance was ignited by the revelation that RNA-guided prokaryotic CRISPR-associated (Cas) effectors could induce double-strand breaks (DSBs) within eukaryotic genomes.^{1,2} An enormous diversity of discovered effectors, categorized into multi-subunit effector systems (class 1, comprising types I, III, IV, and VII) and single-protein effector systems (class 2, comprising types II, V, and VI), has propelled the development of innovative

tools.³ Among these, the Type II CRISPR effector Cas9 has been extensively explored, which has been reviewed elsewhere.⁴ By engineering single guide RNAs (sgRNAs), CRISPR-based approaches can precisely direct Cas9 to specific genomic loci, enabling site-specific DSBs. These DSBs subsequently rely on intricate cellular DNA repair pathways, such as the competing non-homologous end-joining (NHEJ), microhomology-mediated end-joining (MMEJ), and homology-directed repair (HDR) mechanisms, to achieve genome editing.⁵ Beyond leveraging the cell intrinsic repair mechanisms for genome editing, inactivation of HNH and/or RuvC nuclease domains yields catalytically deficient variants (dCas9 and Cas9 nickase) that can in turn recruit other effector domains.⁶ These effector domains include transcription activation or transcription repression modulators (CRISPRa or CRISPRi),^{7,8} catalytic domains of DNA deaminase (base editor),^{9,10} reverse transcriptase (prime editor),¹¹ and integrases (PASTE).¹² These evolved genome engineering toolkits are capable of achieving more precise and predictable genome editing outcomes.

Beyond the canonical SpCas9, other orthologues, including *Streptococcus thermophilus* Cas9 (StCas9), *Campylobacter jejuni* Cas9 (CjCas9), and *Staphylococcus aureus* Cas9 (SaCas9), have expanded the repertoire of tools adaptable for eukaryotic genome manipulation.^{13–16} Furthermore, additional Cas effectors, such as Cas12¹⁷ and Cas13¹⁸ from Type V and VI CRISPR systems, respectively, as well as Cas6e¹⁹ and Csm–Cmr complexes²⁰ from Type I and III CRISPR systems, have been characterized, further enriching the genome engineering toolkit.

The CRISPR technology empowers researchers to simultaneously interrogate myriad perturbations and pinpoint genes intricately involved in specific biological processes through a forward paradigm of genetics. This methodology leverages pooled or arrayed sgRNA libraries with each sgRNA targeting an individual gene, thereby enabling comprehensive genome-wide or

*Address correspondence: Shuquan Rao, State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Tianjin 300020, China. E-mail address: raoshuquan@ihcams.ac.cn (S. Rao).

Conflict of interest: The authors declare that they have no conflict of interest.

Z.L., Z.L., and X.K. contributed equally to this work.

This work was supported by the National Key Research and Development Program of China (2021YFA1102300), the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (2021-RC310-015), the CAMS Innovation Fund for Medical Sciences (CIFMS, 2024-I2M-ZH-015 and 2023-I2M-2-007), and the National Natural Science Foundation of China (82370118 and 82470240).

Blood Science (2025) 7, 1–7:e00266.

Received April 14, 2025; Accepted November 5, 2025.

<http://dx.doi.org/10.1097/BS9.0000000000000266>

Copyright © 2025 The Authors. Published by Wolters Kluwer Health Inc., on behalf of the Chinese Medical Association (CMA) and Institute of Hematology, Chinese Academy of Medical Sciences & Peking Union Medical College (IHCAMS). This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

focused loss-of-function, gain-of-function, or base-editing perturbations across cellular populations.^{21–25} Upon the delivery of sgRNA library in conjunction with Cas9 or derived variants, cells are subjected to phenotypic selection, such as survival challenge, proliferation, and fluorescence- or image-based sorting.^{26–28} Deep sequencing of sgRNA abundance pre- and post-selection identifies genes whose perturbation confers either a selective advantage or disadvantage, thus establishing a profound connection between genotypes and phenotypes. CRISPR screens hold the potential to unravel the intricate mechanisms of gene function and regulation in an unprecedentedly parallel fashion, and to enable exploration of how genes work together to modulate complex phenotypes.

1. SINGLE-CELL CRISPR SCREENING

A key limitation of bulk CRISPR screening is that this approach can only provide information on sgRNA enrichment or depletion, without characterizing the cell components that are either enriched or depleted.²⁶ A promising alternative is to integrate CRISPR screening with single-cell RNA-sequencing (scRNA-seq), hereafter referred to as single-cell CRISPR screening. This approach dissects the transcriptome of individual cells, revealing functional nuances previously inaccessible (Fig. 1). Notable approaches include Perturb-seq,^{29,30} CRISP-seq,³¹ CROP-seq,³² and Mosaic-seq.³³ For a detailed summary of representative single-cell CRISPR screening platforms and their applications, please refer to Supplementary Table 1, <https://links.lww.com/BS/A138>. While these single-cell CRISPR screening approaches offer unprecedented insights, they face technical complexity and throughput limitations. Split-pool ligation-based transcriptome sequencing (SPLiT-seq)³⁴ addresses these using combinatorial barcoding to analyze >100,000 cells/run without microfluidics overcoming low throughput (<10⁴ cells) and costs (>\$1/cell). Compatible with fixed samples, it integrates multi-omics, minimizes batch effects, and enhances studies of gene perturbations^{35,36} (Fig. 2).

The readout for single-cell CRISPR screens extend far beyond RNA-seq, encompassing epigenetic profiling and protein detection to elucidate chromatin architecture or the expression patterns of cell surface proteins, respectively^{37,38} (Supplementary Table 1, <https://links.lww.com/BS/A138>). Epigenetic methodologies, such as Perturb-ATAC,³⁹ CRISPR-sciATAC,³⁷ or SPEAR-ATAC,³⁸ unveil the impact of individual gene or regulatory element perturbations on the chromatin landscape. Simultaneous quantification of transcriptomes and surface protein expression within individual cells has been previously achieved through innovative techniques like CITE-seq⁴⁰

or REAP-seq.⁴¹ The approaches harness oligonucleotide-barcoded antibodies, which specifically bind to cell surface proteins, and integrate these signals with scRNA-seq data within a cohesive and unified workflow. Advanced frameworks such as Expanded CRISPR-compatible CITE-seq (ECCITE-seq)^{42,43} or direct-capture Perturb-seq⁴⁴ allow CRISPR screening while simultaneously capturing multiple modalities, including transcriptome, clonotype, sgRNAs, surface protein, or cell hashing. These groundbreaking technologies have unlocked the potential for pooled CRISPR screens at an unprecedented single-cell resolution, greatly facilitating our understanding of biological processes and disease progression.

2. DETECTION OF sgRNA SEQUENCES

The cornerstone of the process lies in correlating sgRNA information with single-cell omics profiling. Given that sgRNAs are transcribed by RNA polymerase III and lack polyadenylation (polyA) tails, traditional polyA-based capture methods prove inadequate for their detection.⁴⁵ To overcome this, 2 general strategies have been developed: polyA-adapted methods, which embed the sgRNA or an associated barcode into a polyadenylated transcript compatible with standard scRNA-seq capture; and polyA-independent methods, which directly capture sgRNAs using modified scaffolds or custom primers without relying on polyA tails. An engineered vector that allows detection of either the sgRNA or a tethered barcode sequence is required (Fig. 3, Supplementary Table 1, <https://links.lww.com/BS/A138>).

Strategically, sgRNAs coupled unique barcodes can be positioned under a constitutive promoter and immediately upstream of the poly-A signal, thus enabling their efficient detection through scRNA-seq (Fig. 3). However, a significant challenge arises with this approach: the potential uncoupling of sgRNAs and their corresponding barcodes during viral packaging due to template switching.⁴⁶ Beyond DNA-based barcoding, combinations of antibody-detectable epitopes can be employed to generate protein-based barcodes (ProCodes).⁴⁷ Nevertheless, protein-based barcoding remains susceptible to barcode swapping observed in previously described methodologies.

The challenge of uncoupling can be elegantly resolved by directly interrogating the sgRNA rather than relying on a barcode (Fig. 3). CROP-seq employs an ingeniously re-engineered vector, embedding the sgRNA within the 3' long terminal repeat (LTR). This design ensures that the sgRNA in the 3' LTR is co-transcribed with other viral genes by RNA polymerase II, thus enabling detection via conventional scRNA-seq protocols.³² Direct-seq introduces an innovative 8A8G sequence into the

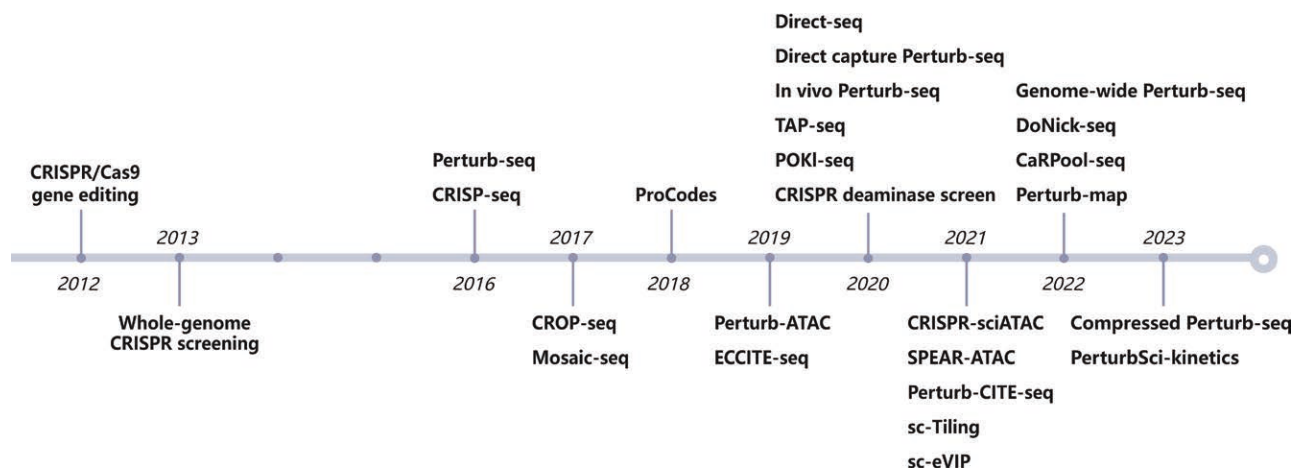


Figure 1. Summary of evolved single-cell CRISPR screening technologies. CRISPR = clustered regularly interspaced short palindromic repeats.

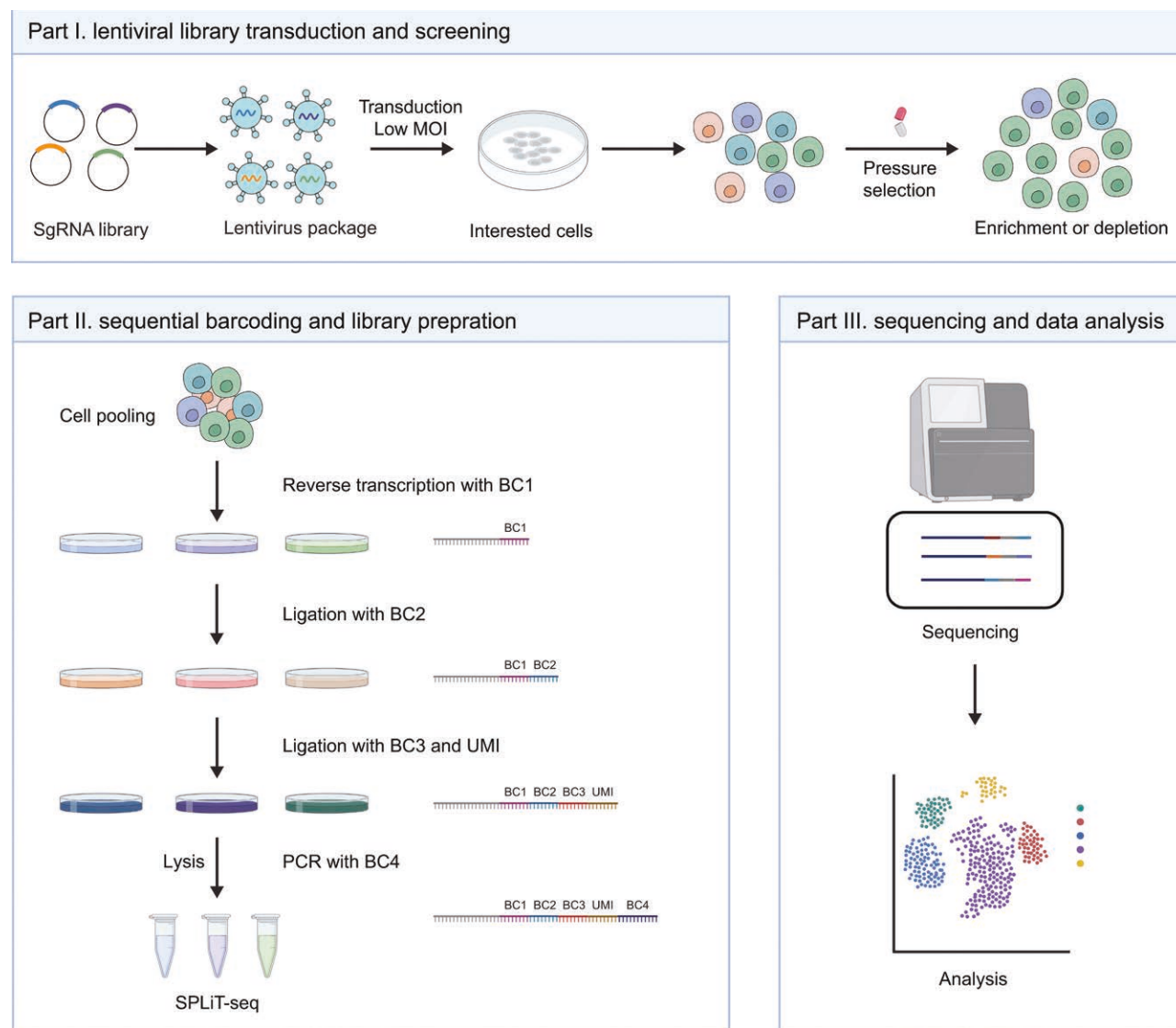


Figure 2. Workflow of single-cell CRISPR screening. The single-cell CRISPR workflow begins with sgRNA library packaging, followed by viral delivery to target cells at low MOI. After selection, cells undergo sequential barcoding steps (reverse transcription with first barcode, ligation with second and third barcodes/UMI), cell lysis, and PCR with a fourth barcode before SPLiT-seq and analysis. BC = barcode, CRISPR = clustered regularly interspaced short palindromic repeats, MOI = multiplicity of infection, sgRNA = single guide RNA, UMI = unique molecular identifier.

sgRNA scaffold, allowing for direct capture using polyT primers under standard scRNA-seq conditions.⁴⁸ Direct-capture Perturb-seq utilizes reverse transcription primers tailored to target specific scaffold sequences, compatible with scRNA-seq workflows.⁴⁴ However, its capture efficiency may be compromised by complex structures such as stem-loops within the scaffold.

Emerging sgRNA capture technologies are advancing and offer promising alternatives. For example, antibody-based protein barcoding methods, such as Perturb-map with ProCode, link sgRNAs to fluorescent epitopes and enable spatial analysis.⁴⁹ Collectively, these strategies not only enhance sensitivity and specificity but also sharpen spatial resolution, unlocking pathways for transformative insights into intricate physiological and pathological conditions.

3. APPLICATION OF SINGLE-CELL CRISPR SCREENS

Single-cell CRISPR screens, such as Perturb-seq³⁰ and CROP-seq,³² provide pooled information on genetic perturbations and associated multivariate phenotypic changes at the individual

cell and, therefore, are promising tools for genotype-phenotype mapping. Herein, we highlight recent advancements in the application of single-cell CRISPR screens within biological and medical research (Supplementary Table 1, <https://links.lww.com/BS/A138>).

Adamson et al harnessed the power of perturb-seq for a comprehensive dissection of the mammalian unfolded protein response (UPR).²⁹ They decoupled the 3 UPR branches governed by IRE1 α , ATF6, and PERK, respectively, revealed bifurcated UPR branch activation and uncovered responses among cells exposed to the same perturbation. Dixit et al showcased the versatility of perturb-seq by analyzing 200,000 cells, with a focus on elucidating the intricate regulatory mechanisms of transcription factors (TFs) governing dendritic cell responses to lipopolysaccharide (LPS).³⁰ Pacalin et al further developed CRISPRai Perturb-seq that enables simultaneous activating (CRISPRa) and repressive (CRISPRi) perturbations at 2 distinct genomic loci within the same cell.⁵⁰ They subsequently employed this approach to unravel the intricate genetic interplay between *SPI1* and *GATA1* during erythroid differentiation, and the complex regulatory dynamics of multiple

Indirect gRNA Capture

① Perturb-seq

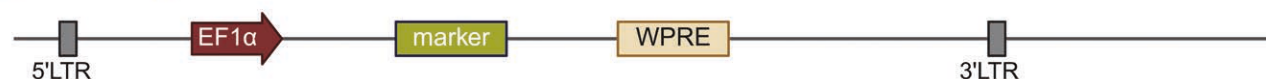


② Mosaic-seq



Direct gRNA Capture

① CROP-seq



② Direct capture Perturb-seq



Figure 3. SgRNA capture approaches. Schematic representation of indirect and direct gRNA capture approaches. Indirect capture methods (Perturb-seq, Mosaic-seq) rely on barcodes (shown in red) to link sgRNAs with transcriptomes. Direct capture methods (CROP-seq, Direct capture Perturb-seq) enable direct sequencing of sgRNAs (shown in blue). Text labels for indirect methods are shown in red, while direct methods are shown in blue to highlight the 2 distinct strategies. BGHpA = bovine growth hormone polyadenylic acid, CS = capture sequence, EF1α = elongation factor 1 alpha, LTR = long terminal repeat, sgRNA = single guide RNA, WPRE = woodchuck hepatitis virus post-transcriptional regulatory element.

enhancers in modulating the expression of the *IL2* gene in T cells.⁵⁰ To uncover the genetic determinants underlying cancer cell susceptibility to natural killer (NK) cell activity, Dufva et al conducted genome-wide CRISPR screens and targeted CROP-seq analyses in cancer cells, highlighting the multifaceted mechanisms shaping NK cell susceptibility.⁵¹ Rapiteanu et al conducted a CROP-seq analysis targeting 45 non-coding regulatory elements and 35 transcription start sites enriched within T-cell-specific open chromatin regions.⁵² By profiling approximately 250,000 single-cell T-cell transcriptomes, this study unveiled the genes and molecular programs modulated by these regulatory elements in primary CD4⁺ T cells. Recently, single-cell CRISPR screening was applied to human brain organoids to uncover developmental defects associated with autism spectrum disorder, highlighting the power of this approach in modeling complex neurodevelopmental diseases.⁵³

Additional breakthroughs in single-cell CRISPR technologies have emerged and been deployed with remarkable ingenuity. Khavari et al introduced Perturb-ATAC, an innovative method that merges multiplexed CRISPR interference or knock-out with transposase-accessible chromatin sequencing (ATAC-seq).³⁹ The team then applied Perturb-ATAC to investigate TFs, chromatin-modifying factors, and non-coding RNAs (ncRNAs), revealing a hierarchical network of TFs orchestrating B cell state.³⁹ Similarly, Liscovitch-Brauer et al³⁷ devised a comparable technique called CRISPR-sciATAC. In human myelogenous leukemia cells, CRISPR-sciATAC was employed to target 105 chromatin-related genes, constructing an extensive chromatin accessibility atlas for approximately 30,000 single cells. Furthermore, Morris et al pioneered base-editing STING-seq

(beeSTING-seq), combining base-editing CRISPR screens and scRNA-seq, and employed it to quantify the effect size and direction of blood trait GWAS variants on gene expression in human erythroid progenitors.⁵⁴

Collectively, single-cell CRISPR technologies hold a transformative potential for deciphering genetic regulation and functional heterogeneity at unprecedented resolution.⁵⁵ Moving forward, advancements will likely focus on amplifying the scalability of single-cell CRISPR technologies, thereby enabling genome-wide screens within single-cell frameworks. Emerging machine learning algorithms tailored for single-cell CRISPR data will serve as indispensable tools in deciphering high-dimensional datasets and forecasting synergistic gene interactions.⁴² Furthermore, the application of spatial CRISPR screens promises could bridge molecular perturbations with tissue architecture, shedding light on how microenvironmental cues sculpt cellular behavior.

To facilitate informed decision-making in platform selection, we present a curated comparative benchmarking table (Supplementary Table 2, <https://links.lww.com/BS/A138>), which encapsulates the essential experimental parameters across 5 single-cell CRISPR screening platforms. The data compiled herein are primarily drawn from peer-reviewed publications and reflect empirically validated experimental applications, rather than speculative theoretical capacities.

4. COMPUTATIONAL TOOLS FOR ANALYZING SINGLE-CELL CRISPR SCREENING DATA

Single-cell CRISPR screening has revolutionized the field of functional genomics by seamlessly integrating precise genetic

perturbations with high-resolution, genome-wide transcriptomic profiling. However, the resulting datasets pose formidable analytical challenges due to their high dimensionality, sparsity, and inherent biological and technical noise. In response, a sophisticated and multifaceted array of computational tools has emerged, spanning the entire analytical pipeline, from initial data preprocessing and quality control to advanced functional interpretation and predictive modeling. We offer a detailed overview of the currently available tools in Supplementary Table 3, <https://links.lww.com/BS/A138>.

A resilient analytical workflow commences with the accurate assignment of sgRNAs to individual cells, a process that can be systematically optimized using tools such as *crispat*.⁵⁶ It is essential to account for both biological variability and technical artifacts. *Mixscape* uses Gaussian mixture models to distinguish perturbed cells from non-responders,⁴² while *SCEPTRE* applies conditional resampling to control confounders and provide calibrated inference.⁵⁷ Following normalization, a variety of analytical strategies are employed to quantify the impact of genetic perturbations. *scMAGeCK*, for example, adapts negative binomial and rank-based models originally developed for bulk screens, while *MIMOSCA* (Multi-Input-Multi-Output-Single-Cell-Analysis) leverages regularized linear modeling to elucidate the intricate relationships between perturbations and transcriptomic responses.³⁰ At the systems biology level, *MUSIC* (Model-based understanding of scCRISPR screening) applies topic modeling to uncover latent biological programs that are influenced by genetic interventions.⁵⁸ The analytical frontier is now evolving from descriptive analysis toward predictive and causal modeling. Tools such as *Pando* reconstruct gene regulatory networks by applying principles of causal inference.⁵⁹ Meanwhile, deep learning frameworks like *GEARS* (Graph-Enhanced Gene Activation and Repression Simulator) and *PerturbNet* integrate biological context to forecast the phenotypic consequences of previously unobserved perturbations.^{60,61}

5. CHALLENGES AND FUTURE DIRECTIONS

Despite the remarkable strides made in single-cell CRISPR technologies, several technical and interpretational challenges persist. A fundamental limitation lies in the low sgRNA multiplicity per cell inherent to most current platforms, which constrains the capacity to interrogate combinatorial gene perturbations in a high-throughput manner.^{32,48} Additionally, off-target effects, occurring at both DNA and RNA levels, pose significant interpretational hurdles, particularly in the absence of stringent controls or orthogonal validation strategies.^{62,63} Another critical constraint pertains to lineage tracing. Although CRISPR-based recording methodologies show promise, existing approaches often lack sufficient temporal resolution or compatibility with concurrent transcriptomic profiling.

From a computational perspective, the inference of perturbation effects from high-dimensional single-cell data remains fundamentally challenged by technical artifacts such as transcriptional noise, dropout events, and batch effects.^{64–67} Addressing these confounding factors is therefore indispensable for deriving biologically meaningful insights. To this end, a suite of advanced computational strategies, including Bayesian models such as *scVI* and integration algorithms like *Seurat*'s anchor-based correction,^{64,68} have been deployed to harmonize heterogeneous datasets and generate robust, standardized representations. This foundational normalization process lays the groundwork for specialized downstream tools, such as *SCEPTRE* and *scMAGeCK*, to accurately dissect the biological ramifications of genetic perturbations.^{57,69}

Standardization initiatives in single-cell CRISPR screening are of paramount importance to ensure reproducibility and cross-platform comparability. Recent benchmarking efforts,

such as *scPerturb*, have systematically assessed key performance metrics including sgRNA detection sensitivity, perturbation effect estimation, and inter-platform variability.⁷⁰ Collectively, these endeavors are forging a path toward more rigorous, reproducible, and universally applicable single-cell perturbation assays.

Protocols for single-cell CRISPR screening have been thoroughly and meticulously detailed in prior literature.^{71,72} Here, we do not reiterate step-by-step procedures, but instead offer a focused comparative overview of 5 widely adopted methodologies, including *Perturb-seq*,^{29,30} *CROP-seq*,³² *ECCITE-seq*,⁴³ *Direct-seq*,⁴⁸ and *Mosaic-seq*.³³

6. sgRNA LIBRARY CONSTRUCTION

6.1. Perturb-seq

The construction of the *Perturb-seq* guide-barcode (GBC) library (pBA571, Addgene #85968) was based on the pBA439 vector backbone (Addgene #85967).^{29,30} An 18-nucleotide random barcode was seamlessly integrated via Gibson assembly between the Blue Fluorescent Protein (BFP) and the Bovine Growth Hormone (BGH) polyadenylation signal. Individual sgRNA protospacer sequences were precisely cloned into the barcoded vector at the *Bst*XI and *Blp*I restriction sites.

6.2. CROP-seq

To construct the sgRNA expression library, the *CROPseq*-Guide-Puro vector (Addgene #86708) was linearized through *Bsm*BI digestion.³² sgRNAs were synthesized as 74-nt single-stranded oligonucleotides featuring 18 and 35 bp of homology to the hU6 promoter and sgRNA scaffold, respectively, and were seamlessly assembled into the vector via Gibson assembly.

6.3. ECCITE-seq

The sgRNA expression library was constructed using the *LentiCRISPR v2* plasmid (Addgene #52961), following previously established cloning strategies.⁷³ Notably, *ECCITE-seq* does not require any modification of this standard vector.⁴³ Instead, sgRNA capture is achieved through a thoughtfully designed reverse transcription primer that hybridizes to the 3' end of the sgRNA scaffold. Amplification of the sgRNA library was carried out using the same strategy as described for *CROP-seq*.

6.4. Direct-seq

A 34-nucleotide sequence, composed of a 4-nt linker and an "8A8G" capture motif, was strategically inserted between the 3' end of the sgRNA scaffold and the polyT termination signal, and subsequently cloned into the *lentiGuide-puro* backbone (Addgene #52963) via Golden Gate assembly.⁴⁸ Plasmid amplification and library quality control were carried out following the same procedures as described for *CROP-seq*.

6.5. Mosaic-seq

The sgRNA expression library was constructed using the *lenti-sgRNA(MS2)*-puro plasmid (Addgene #73795), which was linearized by *Bsr*GI and *Eco*RI digestion.³³ A 12-bp barcode sequence, flanked by homologous arms containing the restriction sites, was precisely inserted between the 3' LTR and the woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) of the linearized vector via Gibson assembly. The sgRNA insertion strategy followed a previously established

protocol.⁷³ In the study by Xie and Hon,⁷⁴ this technique was further refined by integrating the CROP-seq design for sgRNA expression, enabling direct detection of sgRNAs.

7. sgRNA ASSIGNMENT

7.1. Perturb-seq

The original protocol introduced a normalization strategy for GBC counts within each cell, effectively filtering out low-abundance sgRNA contaminants.²⁹ More advanced frameworks have since refined this process by incorporating precise quantitative thresholds.⁷⁵ A widely adopted strategy now excludes GBC signals that meet any of the following criteria: (i) low sequencing coverage (READ_counts/unique molecular identifier [UMI]_counts <60), (ii) supported by only a single UMI, or (iii) representing <10% of the total GBC-UMI counts within the nucleus. Cells harboring multiple high-confidence GBCs are classified as co-transductions of multiple sgRNAs.

7.2. CROP-seq

sgRNA identities are assigned to individual cells by referencing a curated whitelist of designed sgRNAs.⁷⁶ Initially, reads are matched exactly to the whitelist sequences. In cases where no exact match is found, local alignments are performed using the Smith–Waterman algorithm. If a candidate sgRNA sequence falls within an edit distance no greater than half the minimum pairwise edit distance among all sgRNAs (rounded down), the assignment is retained. To mitigate chimeric artifacts, UMI–sgRNA pairs contributing <20% of the UMI reads within a given cell are excluded. A sgRNA is confidently assigned to a cell if it is supported by more than 10 reads and constitutes over 7.5% of the total sgRNA reads for that cell.

7.3. ECCITE-seq

The original ECCITE-seq protocol did not elaborate on the sgRNA assignment methodology.⁴³ However, due to the similarity in sgRNA capture strategies, the CROP-seq protocol serves as a reliable reference for implementation.

7.4. Direct-seq

In the study by Song et al.,⁴⁸ the sgRNA assignment procedure was not explicitly detailed. As with ECCITE-seq, the CROP-seq framework can be effectively adapted for use in this context.

7.5. Mosaic-seq

SgRNA assignment in Mosaic-seq is executed through a probabilistic model designed to infer the most probable number of true sgRNA barcodes present in each cell.³³ For any given cell, all detected barcodes are ranked by read count in descending order. The likelihood of each possible barcode count (j) is then evaluated under 2 competing models: a null model assuming uniform sampling probability across all barcodes, and an alternative model positing exactly j true barcodes, with the remainder classified as background noise. Under the alternative model, probabilities are calculated using multinomial distributions for the top j barcodes, incorporating an empirical decay factor that diminishes successive barcode probabilities. Additionally, contamination rate corrections are integrated into the model. The final number of assigned barcodes is determined as the value of j that yields the highest likelihood, provided it is at least twice as probable as the null model.

ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China (2021YFA1102300), the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (2021-RC310-015), the CAMS Innovation Fund for Medical Sciences (CIFMS, 2024-I2M-ZH-015 and 2023-I2M-2-007), and the National Natural Science Foundation of China (82370118 and 82470240).

AUTHOR CONTRIBUTIONS

Z. Liu, Z. Lan, and X.K. collaboratively gathered literature and crafted the initial manuscript. S.R. and Y.Y. envisioned the project and refined the manuscript with rigorous revisions. All authors approved the final submitted version.

REFERENCES

- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 2012;337(6096):816–821.
- Cong L, Ran FA, Cox D, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* 2013;339(6121):819–823.
- Koonin EV, Gootenberg JS, Abudayyeh OO. Discovery of diverse CRISPR-Cas systems and expansion of the genome engineering toolbox. *Biochemistry* 2023;62(24):3465–3487.
- Hille F, Richter H, Wong SP, Bratovic M, Ressel S, Charpentier E. The biology of CRISPR-Cas: backward and forward. *Cell* 2018;172(6):1239–1259.
- Yeh CD, Richardson CD, Corn JE. Advances in genome editing through control of DNA repair pathways. *Nat Cell Biol* 2019;21(12):1468–1478.
- Ran FA, Hsu PD, Lin CY, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 2013;154(6):1380–1389.
- Qi LS, Larson MH, Gilbert LA, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 2013;152(5):1173–1183.
- Gilbert LA, Horlbeck MA, Adamson B, et al. Genome-Scale CRISPR-mediated control of gene repression and activation. *Cell* 2014;159(3):647–661.
- Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 2016;533(7603):420–424.
- Gaudelli NM, Komor AC, Rees HA, et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* 2017;551(7681):464–471.
- Anzalone AV, Randolph PB, Davis JR, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 2019;576(7785):149–157.
- Yarnall MTN, Ioannidi EI, Schmitt-Ulms C, et al. Drag-and-drop genome insertion of large sequences without double-strand DNA cleavage using CRISPR-directed integrases. *Nat Biotechnol* 2022;41(4):500–512.
- Kim E, Koo T, Park SW, et al. In vivo genome editing with a small Cas9 orthologue derived from *Campylobacter jejuni*. *Nat Commun* 2017;8(1):14500.
- Koo B-K, Hu Z, Wang S, et al. A compact Cas9 ortholog from *Staphylococcus aureus* (SauriCas9) expands the DNA targeting scope. *PLoS Biol* 2020;18(3):e3000686.
- Esvelt KM, Mali P, Braff JL, Moosburner M, Yaung SJ, Church GM. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat Methods* 2013;10(11):1116–1121.
- Chatterjee P, Jakimo N, Jacobson JM. Minimal PAM specificity of a highly similar SpCas9 ortholog. *Sci Adv* 2018;4(10):eaau0766.
- Zetsche B, Gootenberg JS, Abudayyeh OO, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 2015;163(3):759–771.
- East-Seletsky A, O'Connell MR, Knight SC, et al. Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. *Nature* 2016;538(7624):270–273.
- Li M, Cai Z, Song S, et al. EcCas6e-based antisense crRNA for gene repression and RNA editing in microorganisms. *Nucleic Acids Res* 2024;52(14):8628–8642.
- Hale CR, Zhao P, Olson S, et al. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* 2009;139(5):945–956.

- [21] Shalem O, Sanjana NE, Hartenian E, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 2014;343(6166):84–87.
- [22] Doench JG, Fusi N, Sullender M, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* 2016;34(2):184–191.
- [23] Hanna RE, Hegde M, Fagre CR, et al. Massively parallel assessment of human variants with base editor screens. *Cell* 2021;184(4):1064–1080.e20.
- [24] Rao S, Yao Y, Soares de Brito J, et al. Dissecting ELANE neutropenia pathogenicity by human HSC gene editing. *Cell Stem Cell* 2021;28(5):833–845.e5.
- [25] Yao Y, Zhou Z, Wang X, et al. SpRY-mediated screens facilitate functional dissection of non-coding sequences at single-base resolution. *Cell Genom* 2024;4(7):100583.
- [26] Przybyla L, Gilbert LA. A new era in functional genomics screens. *Nat Rev Genet* 2022;23(2):89–103.
- [27] Yan X, Stuurman N, Ribeiro SA, et al. High-content imaging-based pooled CRISPR screens in mammalian cells. *J Cell Biol* 2021;220(2):e202008158.
- [28] Bock C, Datlinger P, Chardon F, et al. High-content CRISPR screening. *Nat Rev Methods Primers* 2022;2(1):9.
- [29] Adamson B, Norman TM, Jost M, et al. A multiplexed single-cell CRISPR screening platform enables systematic dissection of the unfolded protein response. *Cell* 2016;167(7):1867–1882.e21.
- [30] Dixit A, Parnas O, Li B, et al. Perturb-Seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* 2016;167(7):1853–1866.e17.
- [31] Jaitin DA, Weiner A, Yofe I, et al. Dissecting immune circuits by linking CRISPR-pooled screens with single-cell RNA-seq. *Cell* 2016;167(7):1883–1896.e15.
- [32] Datlinger P, Rendeiro AF, Schmidl C, et al. Pooled CRISPR screening with single-cell transcriptome readout. *Nat Methods* 2017;14(3):297–301.
- [33] Xie S, Duan J, Li B, Zhou P, Hon GC. Multiplexed engineering and analysis of combinatorial enhancer activity in single cells. *Mol Cell* 2017;66(2):285–299.e5.
- [34] Rosenberg AB, Roco CM, Muscat RA, et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* 2018;360(6385):176–182.
- [35] Oshima K, Zhao J, Pérez-Durán P, et al. Mutational and functional genetics mapping of chemotherapy resistance mechanisms in relapsed acute lymphoblastic leukemia. *Nat Cancer* 2020;1(11):1113–1127.
- [36] Autry RJ, Paugh SW, Carter R, et al. Integrative genomic analyses reveal mechanisms of glucocorticoid resistance in acute lymphoblastic leukemia. *Nat Cancer* 2020;1(3):329–344.
- [37] Liscovitch-Brauer N, Montalbano A, Deng J, et al. Profiling the genetic determinants of chromatin accessibility with scalable single-cell CRISPR screens. *Nat Biotechnol* 2021;39(10):1270–1277.
- [38] Pierce SE, Granja JM, Greenleaf WJ. High-throughput single-cell chromatin accessibility CRISPR screens enable unbiased identification of regulatory networks in cancer. *Nat Commun* 2021;12(1):2969.
- [39] Rubin AJ, Parker KR, Satpathy AT, et al. Coupled single-cell CRISPR screening and epigenomic profiling reveals causal gene regulatory networks. *Cell* 2019;176(1–2):361–376.e17.
- [40] Stoeckius M, Hafemeister C, Stephenson W, et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods* 2017;14(9):865–868.
- [41] Peterson VM, Zhang KX, Kumar N, et al. Multiplexed quantification of proteins and transcripts in single cells. *Nat Biotechnol* 2017;35(10):936–939.
- [42] Papalexli E, Mimitou EP, Butler AW, et al. Characterizing the molecular regulation of inhibitory immune checkpoints with multimodal single-cell screens. *Nat Genet* 2021;53(3):322–331.
- [43] Mimitou EP, Cheng A, Montalbano A, et al. Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat Methods* 2019;16(5):409–412.
- [44] Replogle JM, Norman TM, Xu A, et al. Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nat Biotechnol* 2020;38(8):954–961.
- [45] Wessels H-H, Méndez-Mancilla A, Hao Y, et al. Efficient combinatorial targeting of RNA transcripts in single cells with Cas13 RNA Perturb-seq. *Nat Methods* 2022;20(1):86–94.
- [46] Sack LM, Davoli T, Xu Q, Li MZ, Elledge SJ. Sources of error in mammalian genetic screens. *G3 (Bethesda, Md)* 2016;6(9):2781–2790.
- [47] Wroblewska A, Dhainaut M, Ben-Zvi B, et al. Protein barcodes enable high-dimensional single-cell CRISPR screens. *Cell* 2018;175(4):1141–1155.e16.
- [48] Song Q, Ni K, Liu M, et al. Direct-seq: programmed gRNA scaffold for streamlined scRNA-seq in CRISPR screen. *Genome Biol* 2020;21(1):136.
- [49] Dhainaut M, Rose SA, Akturk G, et al. Spatial CRISPR genomics identifies regulators of the tumor microenvironment. *Cell* 2022;185(7):1223–1239.e20.
- [50] Pacalin NM, Steinhart Z, Shi Q, et al. Bidirectional epigenetic editing reveals hierarchies in gene regulation. *Nat Biotechnol* 2024;43(3):355–368.
- [51] Dufva O, Gandolfi S, Huuhtanen J, et al. Single-cell functional genomics reveals determinants of sensitivity and resistance to natural killer cells in blood cancers. *Immunity* 2023;56(12):2816–2835.e13.
- [52] Alda-Catalinas C, Ibarra-Soria X, Flouri C, et al. Mapping the functional impact of non-coding regulatory elements in primary T cells through single-cell CRISPR screens. *Genome Biol* 2024;25(1):42.
- [53] Li C, Fleck JS, Martins-Costa C, et al. Single-cell brain organoid screening identifies developmental defects in autism. *Nature* 2023;621(7978):373–380.
- [54] Morris JA, Caragine C, Daniloski Z, et al. Discovery of target genes and pathways at GWAS loci by pooled single-cell CRISPR screens. *Science* 2023;380(6646):eadh7699.
- [55] Meyers S, Demeyer S, Cools J. CRISPR screening in hematology research: from bulk to single-cell level. *J Hematol Oncol* 2023;16(1):107.
- [56] Braunger JM, Velten B. Guide assignment in single-cell CRISPR screens using crisprat. *Bioinformatics* 2024;40(9):btac535.
- [57] Barry T, Wang X, Morris JA, Roeder K, Katsevich E. SCEPTRE improves calibration and sensitivity in single-cell CRISPR screen analysis. *Genome Biol* 2021;22(1):344.
- [58] Duan B, Zhou C, Zhu C, et al. Model-based understanding of single-cell CRISPR screening. *Nat Commun* 2019;10(1):2233.
- [59] Fleck JS, Jansen SMJ, Wolny D, et al. Inferring and perturbing cell fate regulomes in human brain organoids. *Nature* 2023;621(7978):365–372.
- [60] Roohani Y, Huang K, Leskovec J. Predicting transcriptional outcomes of novel multigene perturbations with GEARS. *Nat Biotechnol* 2024;42(6):927–935.
- [61] Yu H, Qian W, Song Y, Welch JD. PerturbNet predicts single-cell responses to unseen chemical and genetic perturbations. *Mol Syst Biol* 2025;21(8):960–982.
- [62] Grünwald J, Zhou R, Garcia SP, et al. Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature* 2019;569(7756):433–437.
- [63] Chen Q, Chuai G, Zhang H, et al. Genome-wide CRISPR off-target prediction and optimization using RNA-DNA interaction fingerprints. *Nat Commun* 2023;14(1):7521.
- [64] Kharchenko PV, Silberstein L, Scadden DT. Bayesian approach to single-cell differential expression analysis. *Nat Methods* 2014;11(7):740–742.
- [65] Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. *Cell* 2021;184(13):3573–3587.e29.
- [66] Lopez R, Regier J, Cole MB, Jordan MI, Yosef N. Deep generative modeling for single-cell transcriptomics. *Nat Methods* 2018;15(12):1053–1058.
- [67] Zheng GX, Terry JM, Belgrader P, et al. Massively parallel digital transcriptional profiling of single cells. *Nat Commun* 2017;8:14049.
- [68] Stuart T, Butler A, Hoffman P, et al. Comprehensive integration of single-cell data. *Cell* 2019;177(7):1888–1902.e21.
- [69] Yang L, Zhu Y, Yu H, et al. scMAGeCK links genotypes with multiple phenotypes in single-cell CRISPR screens. *Genome Biol* 2020;21(1):19.
- [70] Peidl S, Green TD, Shen C, et al. scPerturb: harmonized single-cell perturbation data. *Nat Methods* 2024;21(3):531–540.
- [71] Zheng X, Thompson PC, White CM, Jin X. Massively parallel in vivo Perturb-seq screening. *Nat Protoc* 2025;20(7):1733–1767.
- [72] Alda-Catalinas C, Eckersley-Maslin MA, Reik W. Pooled CRISPR-activation screening coupled with single-cell RNA-seq in mouse embryonic stem cells. *STAR Protoc* 2021;2(2):100426.
- [73] Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* 2014;11(8):783–784.
- [74] Xie S, Hon GC. Experimental and computational approaches for single-cell enhancer perturbation assay. *Methods Mol Biol* 2019;1935:203–221.
- [75] Santinha AJ, Klingler E, Kuhn M, et al. Transcriptional linkage analysis with in vivo AAV-Perturb-seq. *Nature* 2023;622(7982):367–375.
- [76] Hill AJ, McFaline-Figueroa JL, Starita LM, et al. On the design of CRISPR-based single-cell molecular screens. *Nat Methods* 2018;15(4):271–274.