Trends in Genetics

CellPress

Feature Review

Next-generation forward genetic screens: uniting high-throughput perturbations with single-cell analysis

John A. Morris,[1,2] Jennifer S. Sun,[1,2] and Neville E. Sanjana[1,2,*]

Programmable genome-engineering technologies, such as CRISPR (clustered regularly interspaced short palindromic repeats) nucleases and massively parallel CRISPR screens that capitalize on this programmability, have transformed biomedical science. These screens connect genes and noncoding genome elements to disease-relevant phenotypes, but until recently have been limited to individual phenotypes such as growth or fluorescent reporters of gene expression. By pairing massively parallel screens with high-dimensional profiling of single-cell types/states, we can now measure how individual genetic perturbations or combinations of perturbations impact the cellular transcriptome, proteome, and epigenome. We review technologies that pair CRISPR screens with single-cell multiomics and the unique opportunities afforded by extending pooled screens using deep multimodal phenotyping.

Highlights

Forward genetic screens coupling CRISPR (clustered regularly interspaced short palindromic repeats) perturbations with single-cell sequencing have rapidly advanced in the past few years. Recent improvements include increased scale of perturbations – from screening hundreds to now millions of cells – and multimodal phenotypic readouts beyond transcriptomes, such as open chromatin and cell-surface proteins.

Analysis of CRISPR perturbation screens with single-cell sequencing requires connecting cells with the delivered perturbations. For this purpose, initial screens used barcodes delivered in tandem with CRISPR guide RNAs (gRNAs) but were plagued by high rates of barcode swapping. It is now possible to directly capture the functional CRISPR gRNAs, thus improving the fidelity and utility of these screens.

Multimodal single-cell CRISPR screens represent a major development in the toolbox of the molecular geneticist, and enable high-throughput studies of cell function in healthy and disease states.

CRISPR-based tools for pooled screens

Over the past decade diverse CRISPR/Cas (CRISPR-associated) systems have been harnessed for targeted genome modification [1–4]. A major feature of these systems is their programmability enabling high-throughput functional genomic screens such as genome-wide knockout screens [5,6].

The best-known CRISPR nuclease, *Streptococcus pyogenes* Cas9, uses a 20 nt guide RNA (gRNA) to induce genetic alterations at specific genomic locations that are complementary to the gRNA sequence [7–9]. Cas9 creates double-strand breaks (DSBs) [7], but engineered Cas9 mutants can create strand-specific nicks (nCas9) [10], modulate gene expression with a catalytically inactive enzyme (dCas9) tethered to other functional domains [11–14], and bind regions with different protospacer adjacent motif (PAM) target requirements [15–17]. Another CRISPR nuclease used in pooled screens is Cas12a (also known as Cpf1) which recognizes a different (T-rich) PAM and allows easier multiplexing of multiple gRNAs [18]. Some CRISPR systems, such as with Cas13 [19] and Cas7–11 [20], target RNA for cleavage instead of DNA. Modifications to Cas9 include fusing domains which enable gene expression control through inhibition (CRISPRi) [12,21], activation (CRISPRa) [22,23], or other types of chromatin remodeling [24]. Cas9 and other DNA-targeting CRISPRs have also been engineered to create single-nucleotide changes via cytosine [25] and adenine base editors [26], or via prime editors [27] (Figure 1).

For functional genomic screens, CRISPR systems and associated gRNA libraries can be delivered in an arrayed fashion, for studies of the effects of typically a few perturbations individually, or a pooled fashion, for genome-scale studies with thousands of perturbations and/or combinations

[1]New York Genome Center, New York, NY 10013, USA
[2]Department of Biology, New York University, New York, NY 10003, USA

*Correspondence:
neville@sanjanalab.org (N.E. Sanjana).

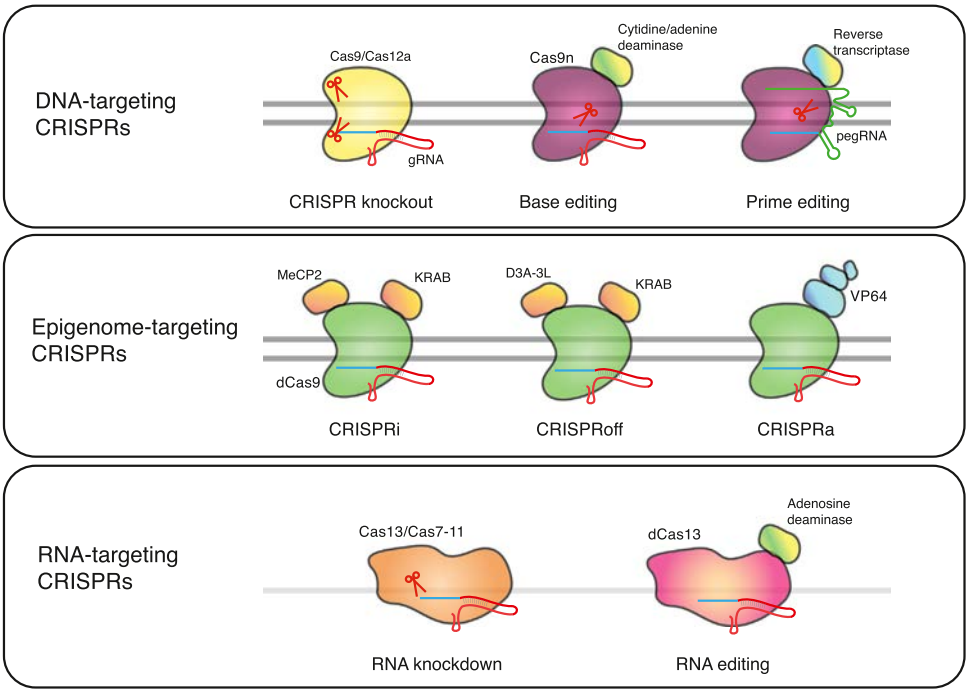

Figure 1. CRISPR systems for targeted DNA, epigenome, and RNA editing. DNA-targeting CRISPR systems can induce double-stranded breaks (Cas9 or Cas12a) to inhibit gene function or induce strand-specific nicks to perform targeted nucleotide mutagenesis (Cas9 nickases, Cas9n) through base editing or prime editing and specialized prime editing gRNAs (pegRNAs). Epigenome-targeting CRISPR systems use a nuclease-dead Cas protein (dCas9) to recruit transcriptional activators or repressors to specific loci. RNA-targeting CRISPR systems use single-strand cutting (Cas13 or Cas7–11) or nuclease-dead Cas proteins (dCas13) to knockdown or perform targeted RNA nucleotide mutagenesis, respectively. Abbreviations: CRISPR, clustered regularly interspaced short palindromic repeats; D3A-3L, DNA methyltransferases DNMT3A/DNMT3L; gRNA, guide RNA; KRAB, Krüppel-associated box repression domain; MeCP2, methyl CpG binding protein 2; VP64, herpes simplex virus protein 16 (VP16) transactivation domain tetrameric repeat.

of perturbations [28]. High-content CRISPR screening has been discussed at length by Bock *et al*. [29], and in this review we discuss recent advances in single-cell readouts for CRISPR-based pooled screens that match the large scale of genetic perturbations with similarly high-dimensional multiomic profiling.

Technologies for CRISPR guide capture in single cells

The first pooled CRISPR screens relied on cell survival to reveal genes required by organisms under particular environmental conditions through measuring the depletion or enrichment of gRNAs targeting specific genes [5,30]. Pooled CRISPR screens can also use cell sorting to map how gRNA perturbations impact specific phenotypes. For example, genome-wide CRISPR knockout screens have identified genes that are important for T cell activation, polarization, or differentiation using fluorescence-activated cell sorting (FACS) [31,32]. After CRISPR-mediated knockout, T cells are sorted based on markers of T cell activation or differentiation to identify key genes that regulate these aspects of T cell biology. For example, Shifrut *et al*. [31] found that loss of *FAM105A* increased the resistance of cytotoxic T cells to adenosine receptor-mediated immunosuppression, a key mechanism of immune evasion in cancer. Using naive $CD4^+$ T cells, Henriksson *et al*. [32] identified multiple genes, including *Pparg* and *Bhlhe40*, that have broad effects on helper T cell activation and differentiation. Cell sorting is not limited to FACS-based approaches either, as genome-wide CRISPR

knockout screens have also been used to identify genes that are important for phagocytosis by measuring gRNA abundances after magnetic selection for cells that had phagocytosed magnetic substrates [33]. One particularly powerful extension of sorting-based screens, called Flow-FISH (fluorescence *in situ* hybridization), uses fluorescent labeling of RNA combined with *in situ* hybridization for phenotypic selection based on gene expression [34]. Flow-FISH combined with tiling CRISPRi libraries can identify regulatory elements, such as enhancer–gene pairs, but its throughput is somewhat limited as each locus requires its own screen [35]. Importantly, these methods rely on bulk sequencing of gRNA perturbations after selection via growth, cell sorting, or other means. These methods do not provide explicitly the transcriptomes of individual cells or the perturbations they received, and typically require follow-on experiments with individual perturbations to quantify changes in gene expression. Single-cell RNA sequencing (scRNA-seq) and related technologies provide an integrated approach to directly connect genetic perturbations to key molecular phenotypes such as gene expression changes.

CRISPR screens with scRNA-seq readouts (i.e., Perturb-seq [36], CRISP-seq [37], Mosaic-seq [38], CROP-seq [39], and enhanced CRISPR-compatible cellular indexing of transcriptomes and epitopes (ECCITE-seq) [40]) facilitate exploration of gene function and systematic delineation of gene regulatory networks (Table 1). These single-cell CRISPR screening approaches require lentiviral delivery of pooled gRNAs to single cells, where functional gRNAs are expressed from RNA polymerase III (Pol III) promoters and thus lack poly-A tails [9]. Methods such as Perturb-seq, CRISP-seq, and Mosaic-seq utilize a separate barcode sequence with a poly-A tail to indirectly capture the gRNA, whereas ECCITE-seq [40] and later direct Perturb-seq [41] use direct gRNA capture (Figure 2). We briefly describe the differences between indirect and direct gRNA capture.

Indirect gRNA capture methods were designed to be compatible with massively parallel droplet-based scRNA-seq assays that capture mRNA by their poly-A tails using poly-T sequences [42,43]. For droplet-based scRNA-seq assays, cells are loaded into a partitioner device alongside gel beads containing barcoded reverse transcription (RT) oligo-dT primers, the beads and cells are encapsulated in droplets along with RT enzymes, cellular transcriptomes are amplified within each droplet, and sequencing libraries are prepared. Droplet-based methods attach droplet-barcodes and unique molecular identifiers (UMIs) to each mRNA captured so that they can be assigned to specific droplets (i.e., cells) and counted. Indirect gRNA capture methods use this poly-A capture mechanism to their advantage by using lentiviral plasmids with polyadenylated barcodes that would be captured within droplets and sequenced, allowing the identification of gRNAs per cell (Figure 2). However, a major caveat of the plasmid design for indirect gRNA capture is that the polyadenylated barcode (transcribed via RNA Pol II) and gRNA (transcribed via RNA Pol III) were 2.5 kb apart, resulting in high barcode-swapping frequencies (1 event per kb) due to lentiviral recombination, causing a ~4.8-fold decrease in gRNA representation [44–46]. This problem is especially insidious because only the barcode is captured, making it impossible to know whether this barcode comes from the correct gRNA or not. A rigorous analysis of this problem showed that ~50% of all gRNAs with Perturb-seq have barcode swapping between gRNAs [45].

An improved CRISPR droplet sequencing protocol developed by Datlinger *et al*. [39] (CROP-seq) addressed the barcode-swapping challenge by clever use of the molecular processes underlying lentiviral integration. Datlinger *et al*. engineered a lentiviral vector that used an *EF-1*α promoter to transcribe the antibiotic resistance gene and the gRNA. The U6 promoter and its gRNA are placed downstream of the *EF-1a* promoter in the 3′ long terminal repeat (LTR) region of the vector. Upon integration, the 3′ LTR is copied, thus creating an independent U6-driven gRNA

Table 1. Single-cell CRISPR screening methods and modalities[a]

Method	Guide RNA capture method	CRISPR screening applications	Single-cell methodology
ECCITE-seq [40]	Does not require a specialized gRNA plasmid; requires a direct capture spike-in oligonucleotide	CRISPR Cas9-based screens (e.g., gene knockout, activation, inhibition, base editing)	Droplet-based single-cell experiments relying on 5′ capture of transcripts
CROP-seq [39]	Requires a specialized CROP-seq plasmid to capture polyadenylated gRNA barcodes	CRISPR Cas9-based screens (e.g., gene knockout, activation, inhibition, base editing)	Combinatorial indexing and droplet-based single-cell experiments relying on 3′ poly-A tail capture
Direct Perturb-seq, Perturb-CITE-seq[b] [41,62]	Requires specialized gRNA plasmids with encoded capture sequences; requires a direct capture spike-in oligonucleotide	CRISPR Cas9-based screens (e.g., gene knockout, activation, inhibition, base editing)	Droplet-based single-cell experiments relying on 3′ or 5′ capture of transcripts
TAP-seq [52]	Can be coupled with gRNA capture method of choice	Can be coupled to a CRISPR screening method of choice; requires nested primers designed to enrich single-cell sequencing libraries for transcripts of choice	Droplet-based single-cell experiments relying on 3′ or 5′ capture of transcripts
CaRPool-seq [103]	Requires specialized gRNA plasmids with encoded capture sequences in a cleavable gRNA array	CRISPR/Cas13-based screens (e.g., RNA knockout, inhibition, base editing)	Droplet-based single-cell experiments relying on 3′ capture of transcripts
OverCITE-seq[c] [104]	Requires a direct capture spike-in oligonucleotide to capture open reading frames	Alternative screening approach for CRISPR activation	Droplet-based single-cell experiments relying on 5′ capture of transcripts
CRISPR-sciATAC [47]	Does not require a specialized gRNA plasmid; requires tagging integrated gRNAs with RT and PCR	CRISPR/Cas9-based screens (e.g., gene knockout, activation, inhibition, base editing)	Combinatorial indexing-based single-cell experiments relying on DNA tagmentation
Perturb-ATAC [70]	Does not require a specialized gRNA plasmid; requires a direct capture spike-in oligonucleotide	Cas9-based screens (e.g., gene knockout, activation, inhibition, base editing)	Physically isolated single cells relying on DNA tagmentation
Spear-ATAC [71]	Requires a specialized gRNA plasmid with Nextera read adapters flanking the gRNA and a direct capture spike-in oligonucleotide	CRISPR/Cas9-based screens (e.g., gene knockout, activation, inhibition, base editing)	Droplet-based single-cell experiments relying on DNA tagmentation

[a]Each method includes a description of the modalities captured, the CRISPR gRNA capture method, the types of CRISPR screening they enable, and the single-cell partitioning and chemistry.
[b]Direct Perturb-seq captures the transcriptome only, Perturb-CITE-seq captures both transcriptome and cell-surface markers.
[c]Uses lentivirally transduced open reading frames as an alternative to CRISPR activation screening.

in the 5′ region of the genomically integrated provirus. The resulting RNA Pol II transcript (driven by *EF-1a*) encodes both the antibiotic resistance gene and a polyadenylated gRNA detectable as a barcode (Figure 2 and Table 1). Although CROP-seq is a highly versatile approach given its compatibility with single-cell technologies that relies on poly-A tail capture and avoidance of barcode recombination issues seen with Perturb-seq, the CROP-seq approach results in a significantly lower titer virus due to the modified 3′ LTR which may require optimization for some applications [47].

Indirect gRNA capture restricted the scale of previous studies and can be incompatible with the delivery of multiple gRNAs; therefore, direct gRNA capture alongside single-cell transcriptomes offers a more versatile single-cell CRISPR screen [40,41] (Figure 2 and Table 1). Mimitou *et al*. [40] performed the first direct gRNA capture through the development of ECCITE-seq that uses gRNA RT primers to directly capture gRNAs by leveraging a template-switching oligonucleotide (TSO) with a 5′ scRNA-seq droplet-based approach. Replogle *et al*. [41] later developed Direct Perturb-seq in which appending a capture sequence to the stem-loop of gRNA oligonucleotides does not affect their function in CRISPRi screens, and separate gel beads containing RT oligo-dT primers complementary to the capture sequence can capture gRNAs. Therefore, for direct

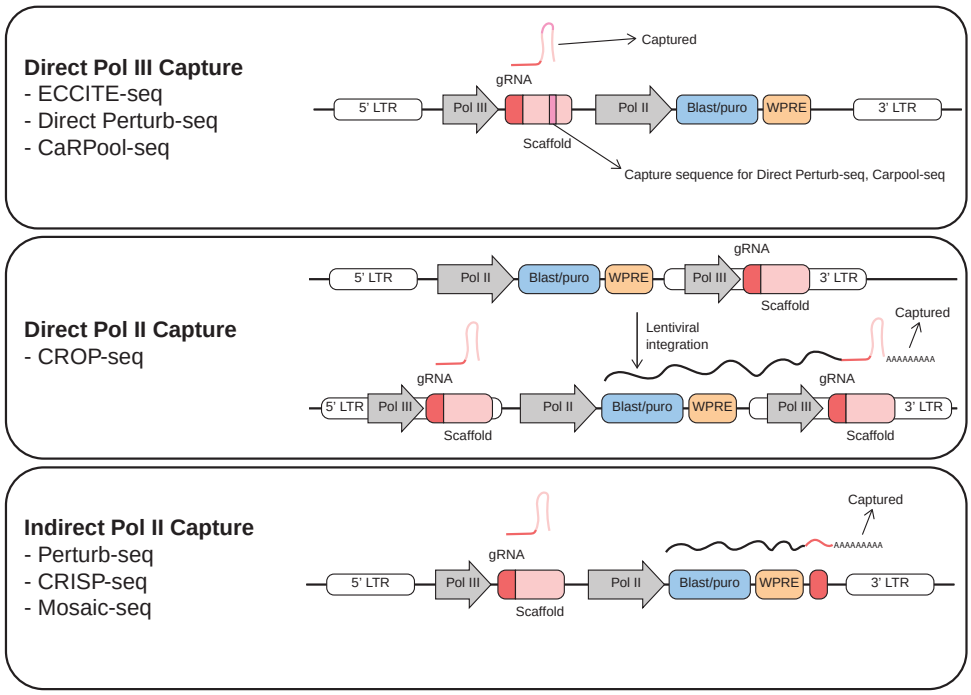

Figure 2. Lentiviral vectors for single-cell pooled CRISPR screens. Direct Pol III capture vectors can be used to directly recover the functional guide RNA (gRNA) in a pooled CRISPR screen. Direct Pol II capture allows direct recovery of the gRNA sequence with a poly-A tail matching the functional gRNA. Indirect Pol II capture allows inference of the gRNA sequence by capturing a barcoded capture sequence with a poly-A tail. Abbreviations: Blast, blasticidin *S deaminase (blasticidin selection marker); CRISPR, clustered regularly interspaced short palindromic repeats; ECCITE-seq, enhanced CRISPR-compatible cellular indexing of transcriptomes and epitopes sequencing; LTR, long terminal repeat; Pol II/Pol III, RNA polymerase II/III promoters; puro, puromycin N-acetyltransferase (puromycin selection marker); WPRE, woodchuck hepatitis virus post-transcriptional regulatory element.*

gRNA capture, depending on the transcriptome capture strategy, the gRNA capture sequence can be linked to the oligo-dT beads (3′) or spiked in alongside the oligo-dT beads (5′) (Figure 3A). As with CROP-seq, these direct capture strategies avoid the barcode-swapping issues with Perturb-seq.

These recent innovations in single-cell CRISPR screens have also been paired with multiplexed CRISPR technologies to simultaneously activate and repress multiple genes in the same single cells [41], thus enabling linked profiling of perturbations with the transcriptome, proteome, and clonotype, all at single-cell resolution [40]. ECCITE-seq is particularly notable as it uses existing CRISPR lentiviral gRNA vectors without any special modification for scRNA-seq studies. We next explore how these technologies have been used to provide novel insights into regulatory and disease biology.

Applications of single-cell CRISPR screens

Single-cell pooled CRISPR screens have yielded tremendous insights into gene function *in vivo* and *in vitro*. By capturing both gRNAs and the transcriptome in a single cell, researchers have identified causal links between perturbations and gene expression patterns at-scale. In an early Perturb-seq study, Dixit *et al*. [36] performed single-cell pooled CRISPR knockout (CRISPRko) screens to study the consequences of perturbing transcription factors (TFs) in mouse bone marrow-derived dendritic cells (BMDCs), and discovered altered gene expression networks in response to lipopolysaccharide (LPS) stimulation. By dissecting these networks, the authors identified TFs with known target genes for antiviral responses and were able to nominate new

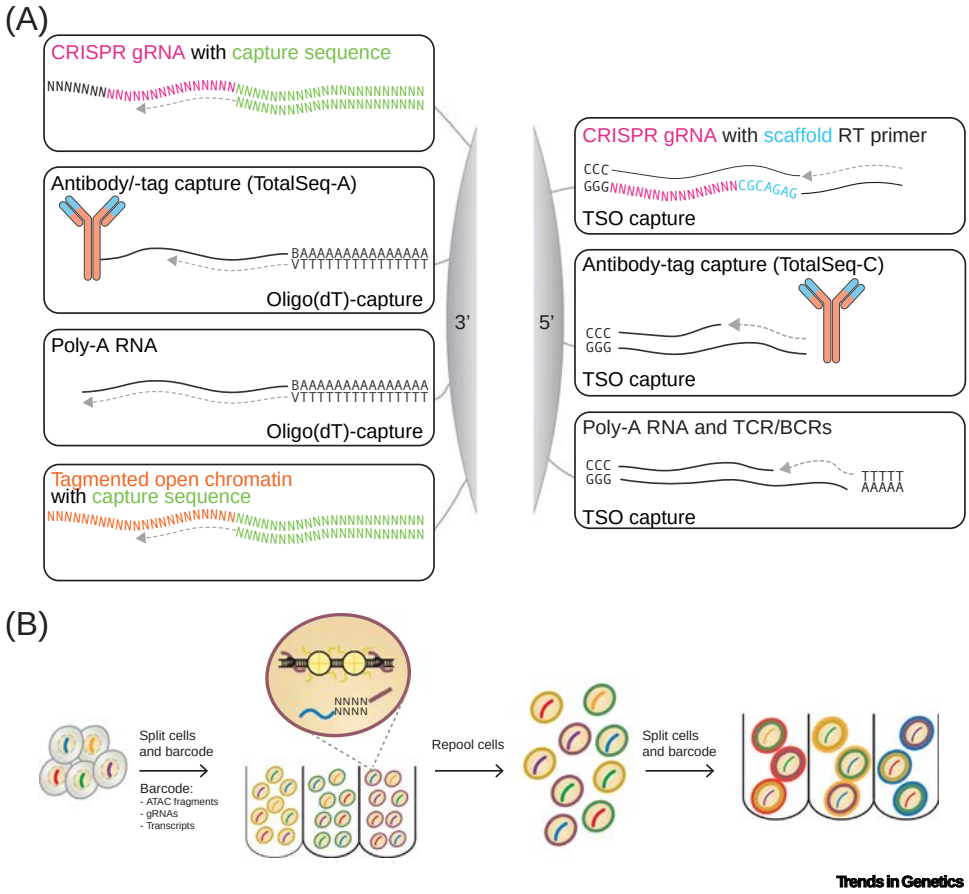

Figure 3. Modalities beyond the transcriptome for single-cell pooled screens. (A) Droplet-based single-cell capture methods currently allow CRISPR guide RNA (gRNA), cell-surface antibodies/tags/receptors, transcriptome, and open chromatin modalities. (B) Combinatorial indexing-based single-cell capture methods currently allow CRISPR gRNA, transcriptome, and open chromatin modalities. Abbreviations: ATAC, assay for transposase accessible chromatin; BCR, B cell receptor; CRISPR, clustered regularly interspaced short palindromic repeats; RT, reverse transcription; TCR, T cell receptor; TSO, template-switching oligonucleotide.

gene functions by examining causal pathways of gene activation [e.g., *Stat2* controls guanylate binding protein (GBP) genes through *Irf8*]. They also used Perturb-seq to study the temporal regulation of proliferation of K562, a human erythroid progenitor cell model, and correctly predicted TF function in which *GABPA* represses mitochondrial function and *YY1* reduces oxidative phosphorylation. Jaitin *et al*. [37] performed CRISP-seq single-cell pooled CRISPRko screens in mouse immune cells to study innate immunity genes. Jaitin *et al*. mapped regulatory networks of differentiation and immune responses and, for example, identified known roles of *Rela* in regulating the monocyte inflammatory response and a novel role for *Cebpa* for regulating differentiation into dendritic or monocytic cells. Shifrut *et al*. [31] also performed a single-cell pooled CRISPRko screen in primary human T cells to better understand the mechanisms underlying sustained T cell proliferation. Single-cell pooled CRISPRko screens continue to be applied to great effect today and, for example, have provided novel insights into host immunity in response to coronavirus disease 2019 (COVID-19) infection. Daniloski *et al*. [48] performed a genome-scale pooled CRISPRko screen in human A549 alveolar cells to identify host factors required for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection. ECCITE-seq of the top-ranked genes identified six genes (*ATP6AP1*, *ATP6V1A*, *CCDC22*, *NPC1*, *PIK3C3*, and *RAB7A*) that led to upregulation of pathways

affecting lipid and cholesterol homeostasis, and found that increases in cholesterol led to resistance to SARS-CoV-2 and that cholesterol modulation using existing FDA-approved therapies could block viral infection.

Using transcriptional modulation instead of nuclease-driven knockout, Adamson *et al*. [49] performed Perturb-seq using CRISPRi (knockdown) to identify genes required for endoplasmic reticulum (ER) homeostasis in the mammalian unfolded protein response (UPR) pathway [49]. Adamson *et al*. also developed a triple gRNA vector, and tested single, double, and triple combinations of gRNAs targeting UPR genes, thus enabling the delineation of *IRE1A*-, *PERK*-, and *ATF6*-controlled transcriptional programs. Schmidt *et al*. [50] performed Perturb-seq using CRISPRa to validate 14 genes from their genome-scale CRISPRi and CRISPRa screens in T cells. Schmidt *et al*. analyzed $CD4^+$ and $CD8^+$ T cells clustered by expression patterns and found that regulators of cytokine production tune T cell activation and program cells into different stimulation-responsive states. Using these insights, they were able to nominate genes that enhance the efficacy of adoptive T cell therapies [50].

Until recently single-cell pooled CRISPR screens had not been applied at a scale comparable with pooled screens with bulk readouts. Given the hundreds of single cells needed to interrogate the effects of perturbations, and the sequencing depth necessary to detect significant effects, genome-scale scRNA-seq perturbation screens are limited primarily by their high cost. Replogle *et al*. [51] performed the first genome-scale CRISPR screens using single-cell readouts. They targeted all expressed genes in K562 cells (~10 000), and all essential genes shared between K562 and the retinal pigment epithelial cell line RPE1 (~2000). In all, they sequenced ~2.5 million cells to achieve ~100 cells per perturbation, with 800–3000 reads per cell. This study identified functions for several poorly annotated genes (e.g., ribosome biogenesis, transcription, and mitochondrial respiration) by observing that their transcriptional profiles clustered tightly with genes with known functions. Given the enormous costs of library preparation and sequencing, it can be helpful to develop methods with a similar perturbation scale but a more focused readout: Targeted Perturb-seq (TAP-seq) represents one potential solution for overcoming sequencing depth requirements for single-cell studies (Table 1) [52]. In TAP-seq the sequencing library is first enriched for transcripts of interest (e.g., 100–1000 specific transcripts), which provides a balance between multidimensional phenotyping and cost, as it can reduce the reads-per-cell requirements by up to 50-fold.

Instead of directly targeting genes, single-cell pooled CRISPR screens can also target noncoding *cis*-regulatory elements (CREs) or enhancers. CREs typically regulate gene transcription by binding to enhancer elements and TFs. By interrogating CREs with single-cell screens, we can identify putative gene targets and understand the epigenetic changes that precede gene regulation in development and disease. Xie *et al*. [38] developed Mosaic-seq, one of the first indirect gRNA capture methods for droplet-based approaches, targeted gRNAs to enhancers, and quantified enhancer penetrance on target gene expression. Later studies by Xie *et al*. [53] and Gasperini *et al*. [54] used CRISPRi coupled CROP-seq to further identify hundreds of enhancer–gene pairs. These studies found that, although most enhancers target a single gene, some enhancers target multiple genes. In addition, multiple enhancers can target the same gene and enhancer dosage is a major determinant of gene expression. Xie *et al*. [53] explicitly investigated putative enhancers near TFs and found that enhancers for the same TF can modulate different submodules within a regulatory network. Therefore, by targeting enhancers, researchers can flexibly control entire regulatory networks. Recently, Morris *et al*. [55] used CRISPRi with scRNA-seq and direct gRNA capture to inhibit noncoding loci identified from genome-wide association studies of blood traits. This approach (systematic targeting and inhibition of noncoding GWAS loci with single-cell sequencing, STING-seq) starts by targeting common genetic variants in CREs

and, using ECCITE-seq, can identify CRE target genes in *cis* and entire regulatory networks in *trans*. This study underscores an important utility of single-cell pooled CRISPR screens because targeted perturbations can reveal dynamic regulatory networks for human complex traits and common diseases, and help to analyze noncoding loci with human genetic evidence.

The aforementioned studies were conducted in *in vitro* systems, which are often not representative of intact organisms. In 2020, Jin *et al*. [56] performed the first *in vivo* single-cell CRISPR screen. They knocked out 35 autism spectrum disorder (ASD) and neurodevelopmental delay (ND) genes in mouse embryos to study developmental processes of the early postnatal brain. ASD/ND genes are predicted to act in multiple brain cell types. The authors linked specific gene knockouts to perturbed regulatory networks, and revealed novel functions of ASD/ND genes in specific types of neurons. A recent similar study used a different viral vector (adeno-associated virus) to study a broader set of cell types and timepoints in 22q11.2 deletion syndrome [57]. Taken together, these studies demonstrate the power of single-cell pooled CRISPR screens to functionally dissect genes involved in disorders of mammalian brain development.

Beyond the transcriptome: CRISPR screens with single-cell proteomics and genome accessibility

Although single-cell transcriptomics provides insight into regulatory networks, cell states, and differentiation, multimodal single-cell CRISPR pooled screens can enhance our understanding of the underlying biology by measuring protein or epigenetic modifications. For example, cellular indexing of transcriptomes and epitopes (CITE-seq) [58] combines protein measurements with droplet-based scRNA-seq (Figure 3A). CITE-seq leverages single-cell barcoding to capture information encoded on the cell surface, and, by using antibodies tagged with an oligonucleotide sequence, it marks specific samples with a unique barcode. This allows multiple samples to be pooled in a single assay and sequencing run, reducing potential batch effects and technical confounders. Recently, CITE-seq and similar approaches have been expanded in several directions. Wroblewska *et al*. [59] developed Pro-Codes, which use lentiviral vectors encoding epitope barcodes that can be expressed on the cell surface for cell indexing. McGinnis *et al*. [60] developed MULTI-seq, which uses lipid-tagged oligonucleotides, to increase sample multiplexing and single-cell recovery. Mylka *et al*. [61] benchmarked CITE-seq and MULTI-seq, and found that both performed well for cell line, nuclei, and tissue staining; however, given the experimental design of the study, one method may outperform the other under specific circumstances. Mimitou *et al*. [40] developed ECCITE-seq, which combines direct gRNA capture, whole transcriptomes, and cell-surface protein quantification. More recently, a similar method called Perturb-CITE-seq was developed using Direct Perturb-seq gRNA capture sequences [62].

Given that CITE-seq was initially designed to target cell-surface proteins, further innovation has been necessary to target intranuclear or intracellular proteins. For example, inCITE-seq [63], NEAT-seq [64], and INs-seq [65] are able to capture intranuclear proteins in single cells, and Phospho-seq [66] can capture phosphorylated intranuclear and intracellular proteins in single cells. Although these methods have enabled measurement of protein levels at both the cell surface and within the cell, they do not yet include gRNA capture. Altogether, advances to the CITE-seq toolkit have enabled simultaneous measurement of the effects of CRISPR perturbations on the transcriptome and cell-surface proteins, and intranuclear and intracellular proteins represent the next major advances on the horizon.

Another key cellular modality is accessible (open) chromatin, which indicates functionally active coding and noncoding genomic regions [67]. One such method for identifying open chromatin

is assay for transposase-accessible chromatin (ATAC)-seq, which uses transposases to fragment and attach sequencing primers to accessible DNA [68]. ATAC-seq has been adapted for single-cell sequencing (scATAC-seq) [69], and combining CRISPR perturbations with CRISPR gRNA capture and scATAC-seq can reveal the impact of individual genes or noncoding regulatory elements on chromatin accessibility (Figure 3). Perturb-ATAC [70] relies on physical isolation of single cells, whereas CRISPR-sciATAC [47] and Spear-ATAC [71] capitalize on more scalable methods (combinatorial indexing and microfluidic droplets, respectively) to analyze larger numbers of single cells, albeit with less sequencing depth (Table 1). These studies profiled chromatin accessibility in response to TF perturbations in GM12878 cells (immortalized B lymphoblasts), primary human keratinocytes, and K562 cells (blood progenitors). Briefly, single-cell combinatorial indexing (sci) approaches rely on split-pooling cells for unique barcodes (Figure 3B), often using easily obtained 96-well plates [72], and have been applied to a range of modalities such as sciRNA-seq [73] and sciATAC-seq [74].

Liscovitch-Brauer *et al*. [47] developed CRISPR-sciATAC, which combines single-cell combinatorial indexing with CRISPR gRNA capture. In the first step of CRISPR-sciATAC, they performed barcoded tagmentation on open chromatin from nuclei and used RT to apply the same barcode to gRNAs, allowing the recovery of both modalities in multiwell plates. After pooling and re-splitting, a second set of barcodes was applied via PCR using primers specific for each molecular species (ATAC or gRNA). Using this method, they profiled the impact of knockout of every chromatin modifier in the human genome on chromatin accessibility, and mapped how each chromatin-modifying complex and individual subunit modulates the chromatin accessibility of specific TF binding sites across the genome. For example, loss of *EZH2* increased the accessibility of genes in the *HOX* cluster, highlighting the key role played by *EZH2*, a histone methyltransferase, in human development.

Several other combinational indexing approaches combining CRISPR gRNA capture and additional modalities have been developed recently. For example, Xu *et al*. [75] developed PerturbSci-Kinetics, a pooled CRISPRi screening approach that enables the capture of CRISPR gRNAs, whole transcriptomes, and nascent transcriptomes. RNA kinetics, such as RNA synthesis and degradation, can be studied by labeling nascent transcripts in cells with a metabolic label, 4-thiouridine (4sU) [76]. After inducing C>T conversions in 4sU-labeled transcripts via thiol (SH) alkylation [77], the difference in C>T conversions between nascent transcripts and unlabeled transcripts enables per-transcript kinetic estimates. PerturbSci-Kinetics was also used to identify a novel role for *AGO2*, a known post-transcriptional silencer functioning in RNA interference [78], as a regulator of RNA synthesis [75].

Single-cell pooled CRISPR screens that combine all of the aforementioned modalities – RNA, protein, chromatin accessibility, and CRISPR gRNA capture – do not currently exist, but recent advances suggest they are possible. Mimitou *et al*. [79] developed DOGMA-seq, a variant of CITE-seq, which simultaneously profiles RNA, cell-surface proteins, and chromatin accessibility in single cells. Given that other groups have interrogated these modalities individually alongside CRISPR gRNA capture, it should be possible to combine them in a single assay – an exciting area for future technology development.

Emerging technologies for CRISPR screens with multimodal readouts

Most single-cell CRISPR screens are performed with short-read sequencing, which captures either the 5′ or the 3′ end of each transcript; however, emerging studies leverage long-read sequencing to capture entire RNA molecules for a more informative output. Long-read sequencing enables the detection of transcript isoforms which can result from perturbations. However, none of

these long-read sequencing technologies have been combined with CRISPR gRNA capture. Instead, single-cell arrayed CRISPR screens implement long-read sequencing in cells where the CRISPR gRNA identity is already known [80]. Although powerful for examining the effects of specific targets of interest, arrayed screens are low-throughput given the requirement to perform a CRISPR screen for each individual target. Future studies should aim to combine long-read sequencing with single-cell pooled CRISPR screens and indirect or direct gRNA capture to enable scalability.

For the genome, the 3D structure and folding of DNA is crucial to understand gene regulatory compartments and which promoters might be regulated by specific distal enhancers [81–83]. Chromosome conformation capture (3C) and related sequencing-based methods, such as Hi-C, capture DNA fragments which physically interact via crosslinking, restriction enzyme digestion, and then ligation [84,85]. The earliest single-cell studies adapting 3C/Hi-C technology were relatively low-throughput by current standards, as they required manual isolation of cells [86,87]. The development of combinatorial indexing-based sci-Hi-C approaches enabled the simultaneous profiling of thousands of single cells [88,89]. Recently, Liu *et al*. [90] developed a multimodal combinatorial indexing-based approach, termed Hi-C and RNA-seq employed simultaneously (HiRES), to simultaneously capture Hi-C and RNA-seq in single nuclei from thousands of cells. Single-cell Hi-C approaches continue to advance and incorporate additional modalities, but they do not yet incorporate CRISPR gRNA capture, thus restricting the scale of perturbation studies and their consequences for 3D chromosome interactions. For example, Guo *et al*. [91] performed arrayed CRISPR perturbations of CTCF binding sites with 3D conformation capture, and found that inverting the orientation of these binding sites can reconfigure chromatin looping and alter enhancer–gene interactions. Adding 3C/Hi-C as a modality to single-cell pooled CRISPR screens promises to functionally profile TF binding domains, 3D compartmentalization junctions, and pathogenic variants.

Although not strictly a single-cell method, Repair-seq quantifies DSB outcomes using an innovative approach with UMIs and bulk sequencing [92]. By introducing a gRNA together with a CRISPR target site and flanking restriction enzyme sites, the sites of gRNA and target region integration can be extracted from genomic DNA and tagged with a UMI before PCR. Deep sequencing then allows the assignment of gRNAs to specific DSB repair outcomes in a single cell labeled by a single UMI, and enables the study of genetic modulators of DNA repair [92].

Imaging-based approaches can be used to select cells with distinct cellular and subcellular features, such as nuclear size, thereby enabling CRISPR screens for genes that regulate cell morphology, cellular dynamics, or cell-to-cell interactions [93–95]. For example, Feldman *et al*. [93] developed an optical pooled CRISPR screening approach where cells are transduced with lentivirus that contain gRNAs that include 12 nt barcodes. Upon selecting cells by physical features under a microscope, single cells can then be matched to their perturbation by *in situ* sequencing of the 12 nt barcode. Briefly, dyes are annealed, imaged, and cleaved to identify nucleotides individually (*in situ* sequencing-by-synthesis). Feldman *et al*. [93] examined p65 translocation dynamics upon knocking out 952 genes involved in NF-κB signaling, and uncovered a new role for Mediator complex subunits. Another imaging-based pooled screen is BARC-FISH (barcode amplification by rolling circle and FISH), which uses FISH to image pooled, barcoded CRISPR gRNAs in single cells coupled with DNA FISH chromatin tracing to examine 3D chromatin changes [96]. BARC-FISH was used to identify several new modifiers of 3D genome organization, including *CHD7*, a gene where *de novo* mutations lead to specific birth defects and CHARGE (coloboma, heart defects, atresia choanae, restriction of growth and development, genital abnormalities, ear abnormalities) syndrome.

Recent single-cell CRISPR screens with multimodal readouts have expanded beyond targeting DNA by capitalizing on the tremendous metagenomic diversity of CRISPR systems. For example, the class II type VI CRISPR-Cas13 family can be used to directly target RNA for degradation [19,97,98] or base editing [99] (Figure 1). Wessels *et al.* [100,101] and Méndez-Mancilla *et al*. [102] determined Cas13 RNA-targeting rules, and optimized Cas13 gRNA sequence and chemical modifications, respectively. Compared with DNA-targeting CRISPRs, the Cas13d nuclease has several unique advantages: it does not require a specific PAM sequence (e.g., NGG for Cas9) for target recognition and acts in a strand-specific manner [97], thus allowing Cas13d to target any sequence in the transcriptome. To combine Cas13d with droplet-based single-cell methods, Wessels *et al*. developed CaRPool-seq for Cas13d-based single-cell pooled RNA-targeting CRISPR screens (Table 1) [103]. Since Cas13d can cleave gRNA arrays into independent perturbations, this enables combinatorial perturbations with multimodal single-cell profiling. Arrays were cloned with up to three distinct gRNAs and a single barcode, and not only resulted in separate targets being encoded in the same array but also enabled consistent and efficient gRNA capture. The authors found that CaRPool-seq outperformed the Cas9-based direct-capture Perturb-seq with respect to combinatorial perturbation detection. CaRPool-seq uses a single barcode for up to three gRNAs, and in cells with a barcode detected there was a 99% concordance rate of the gRNAs assigned and the barcode label. Direct-capture Perturb-seq with dual-gRNA delivery requires the detection of individual gRNAs per pair, and in cells with at least one gRNA detected there was a 67% detection rate for the expected paired gRNA. Given these differences in capture efficiency, Cas13-based screens may be a preferred approach for combinatorial gene perturbations over Cas9-based screens.

Other notable single-cell pooled screens include other non-CRISPR-based genetic perturbations. For example, Legut *et al*. [104] developed OverCITE-seq by building upon the existing CITE-seq toolkit to enable the direct capture of lentivirally delivered gene overexpression via open reading frames (ORFs). Instead of using CRISPRa to constitutively express genes within a cell, OverCITE-seq lentivirally delivers ORFs under the control of a constitutive promoter (cytomegalovirus, CMV) to express genes directly (Table 1). In addition to driving higher levels of gene expression than CRISPRa, a major benefit of OverCITE-seq is the reduced size of the viral payload that makes it easier to transduce cells that have been traditionally difficult to study, such as primary human T cells. Demonstrating the utility of OverCITE-seq, Legut *et al*. identified *LTBR*, a gene not canonically expressed in $CD4^+$ and $CD8^+$ T cells, as a driver of T cell effector functions that makes them resistant to exhaustion in chronic stimulation settings. Using OverCITE-seq, Legut *et al*. also captured single-cell T cell receptor (TCR) clonotypes, and this enabled them to show that modifier genes such as *LTBR* are truly drivers of the phenotype rather than reflecting a clonal effect due to the lentiviral integration site (e.g., by disrupting the gene at the integration site [105]). Similarly, orthogonal perturbation methods can be combined with CRISPR-based approaches. For example, Li *et al*. [106] developed a combinatorial indexing-based approach to capture lentivirally delivered short-hairpin RNAs (shRNAs) for gene knockdown flanked by a bacteriophage T7 promoter and an RT primer-binding site to study modifiers of CRISPR prime editing efficiency. This dual-perturbation strategy enabled the identification of *trans*-acting regulators of prime editing efficiency; for example, the authors found that *HLTF* inhibition improves prime editing efficiency [106].

Analytical methods and tools for single-cell pooled screens

The analysis of pooled CRISPR screens is a developing field, and new methodologies are tackling the challenges inherent to distilling key phenotypes from high-dimensional, noisy measurements of single cells. The analysis workflow for single-cell pooled CRISPR screens is conceptually

straightforward: (i) CRISPR gRNA libraries are sequenced – these can be either the direct sequence or their barcodes, which can be referred to as feature barcodes; (ii) the sequencing data are aligned to reference gRNA sequences or feature barcodes; (iii) quality control is then performed, such as only keeping those cells with a single gRNA detected and with a minimum number of gRNA reads; followed by (iv) differential expression (or other modality of interest) analysis (Figure 4).

There are multiple decisions to be made at each step; for example, the gRNA vector and the single-cell method will determine whether gRNAs are captured directly and the specific capture strategy. After sequencing, multiple tools can be used for alignments, and each has its own strengths and weaknesses. The goal is to use a list of gRNA sequences or barcodes as a reference genome for mapping sequenced reads and generating UMI count matrices (UMIs per gRNA per cell). 10X Genomics provides a proprietary software suite to perform read alignments and generate UMI count matrices for the transcriptome as well as additional modalities such as Cell Ranger [43]. Cell Ranger uses the STAR aligner [107] for antibody or gRNA capture analysis, and allows up to one mismatch in sequence alignments. CITE-seq-Count, originally designed for processing CITE-seq antibody capture data, can be easily adapted to gRNA sequencing reads and allows the user to specify the maximum number of mismatches that are allowed. Alternatively, kallisto-bustools [108] and alevin-fry [109], which can perform computationally rapid and low-memory scRNA-seq read alignment via pseudoalignment [110], can be repurposed for gRNA read alignment and generating UMI count matrices. Once UMI count matrices have been generated, quality control checks should be performed to determine UMI thresholds per gRNA per cell. Cell Ranger will perform a secondary analysis upon generating UMI count matrices, where for each gRNA it

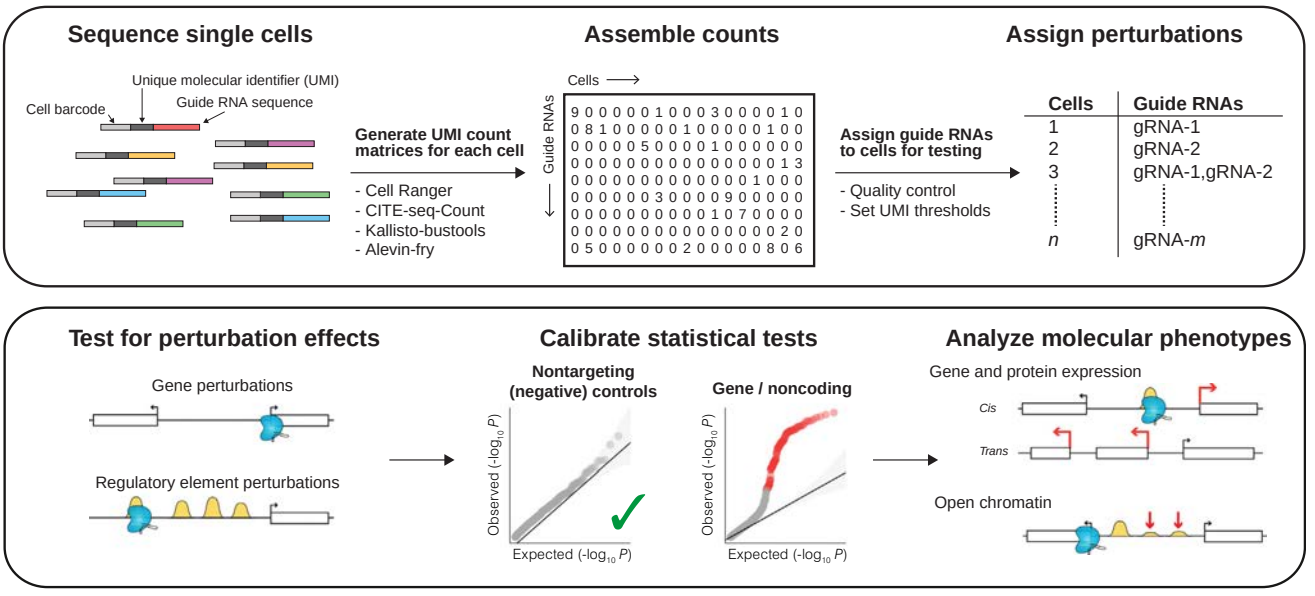

Figure 4. Analysis of single-cell pooled perturbation screens. Single-cell perturbations [e.g., guide RNA (gRNA) or barcoded overexpression libraries] can be represented by a $n \times m$ matrix, where n is the number of single cells and m is the number of perturbations, and the matrix is populated with the number of unique molecular identifiers (UMIs) per perturbation per cell. After defining UMI thresholds and assigning perturbations to cells for follow-up analyses, changes in different modalities of interest (e.g., open chromatin or gene expression changes) are measured. First, nontargeting (negative) controls should be inspected for potential inflation of observed test statistics. This can be performed by comparing the observed P values with the P values that would be expected at random given the number of tests performed, and by examining the results for systematic deviations from the null expectation. Upon verifying that there is no systematic inflation, targeting gRNAs are examined for their effects on measured molecular phenotypes (e.g., gene expression changes in *cis* and *trans*, or open chromatin). Abbreviation: CITE-seq, cellular indexing of transcriptomes and epitopes sequencing.

calculates its UMI threshold to call cells bearing that gRNA using a Gaussian mixture model. Barry *et al*. [111] developed a model for determining gRNA UMI thresholds – GLM-based errors in variables (GLM-EIV) – and observed through simulation studies and applications to real data that, in the absence of modeling a UMI threshold, a minimum of three UMIs can be used to reliably call a cell bearing a gRNA.

Last, after generating quality-controlled UMI count matrices, and assigning gRNAs to cells, pairwise differential tests can be performed. For low multiplicity of infection (MOI) experiments, where most cells express a single gRNA, differential expression testing generally involves dividing cells into a nonperturbed set (e.g., cells bearing nontargeting control gRNAs) and a perturbed set (e.g., cells targeting a specific gene), and performing differential expression tests [36,112,113] on a selected gene for differences between the sets. Given that CRISPR knockout may have incomplete penetrance, methods are also available for identifying populations of cells bearing gRNAs that remained unperturbed, such as Mixscape [114] or MELD [115]. For high MOI experiments, where most cells express multiple gRNAs, these approaches can suffer from test statistic inflation [116]. For these cases, SCEPTRE [116,117] or Normalisr [118] approaches outperform other methods that are not explicitly designed for high MOI experiments [53,54,112]. SCEPTRE couples a conditional resampling test with a negative binomial approach for differential expression testing, whereas Normalisr uses a linear regression approach. Regardless of the differential expression testing method, a good practice is to first examine negative controls and verify a null distribution of test statistics (e.g., examining *P* values on a $-\log_{10}$ scale with a quantile–quantile plot). This demonstrates that the results do not suffer from model miscalibration or systematic test statistic inflation or deflation (Figure 4). It is important to verify that nontargeting gRNAs are not systematically inflated or deflated compared with a null distribution; otherwise, significant observations may be due to a high false positive rate. After these steps, perturbations can then be confidently tested for their effects on gene expression in *cis* and in *trans*, or for other single-cell modalities such as open chromatin (Figure 4).

Concluding remarks and future perspectives

Over the past decade since we and others developed the first CRISPR pooled screens [5,30], there has been widespread adoption of these methods throughout academia and biopharma for genetic discovery at a genome-wide scale. We expect a similar increase in the adoption of single-cell CRISPR screens over the next decade and the development of new technologies that will drive higher throughput, lower cost, and simplified experimental and analytical workflows. For example, current droplet-based approaches, while they allow more modalities than combinatorial indexing approaches, are limited by the number of single cells that can be generated in a single assay. Datlinger *et al*. [119] developed single-cell combinatorial fluidic indexing RNA-seq (scifi-RNA-seq) by performing combinatorial indexing barcoding before droplet-based barcoding, and generated 100-fold more single cells than standard droplet-based scRNA-seq approaches. The field is rapidly evolving to a place where we will no longer be limited by the number of single cells we can generate. Future hurdles will be the modalities we can capture within the same cells and the depth to which we can sequence them (see Outstanding questions).

The next few years will likely see multiple genome-scale and combinatorial single-cell CRISPR screens that will yield new insights into gene function and a deeper understanding of how combinations of genes function and interact at scale. We expect that these genome-scale datasets will include millions of single cells, and will include not only transcriptome and CRISPR gRNA capture modalities but also cell-surface and intracellular proteins, chromatin accessibility, DNA and RNA modifications, histone post-translational modifications, TF activity and binding, and

Outstanding questions

How many modalities can one capture from a single cell? As the number of modalities continues to increase (e.g., DNA, RNA, protein), innovations in single-cell library preparation and sequencing will be needed for simultaneous readout. These innovations will deepen our understanding of causal effects of genetic perturbations on diverse cellular phenotypes, and will link different molecular species to a more complete understanding of the inner workings of cells.

How many cells should be sequenced for a particular single-cell forward genetic screen? Current single-cell sequencing studies are performed with thousands to millions of single cells, but rigorous power calculations are rarely done. Determining the minimum number of cells necessary for a given pooled CRISPR screen to achieve sufficient representation of different perturbations is important for designing well-powered studies.

What sequencing depth is necessary to overcome sparsity in single-cell data? Current technologies are limited in that they do not capture the entirety of the intended modality per cell. Studies have sequenced single cells to depths of thousands to tens of thousands of reads per cell to overcome this sparsity, but additional innovations in single-cell capture and reduced costs for sequencing will be necessary to improve the recovery of information from cells for forward genetic screens.

What computational and analytical approaches are best suited for multiomic single-cell perturbation experiments? Single-cell datasets continue to grow in terms of the number of cells studied and the number of modalities generated. To identify biological insights efficiently and accurately, these increasingly complex study designs will require sophisticated pipelines and mathematical frameworks. Examples of challenges in this space include appropriate normalization methods that preserve biological variability, multiple modality integration and comparisons, and statistical methods that strike the right balance between discovery and penalties for multiple-hypothesis testing.

chromosome conformation. For example, recent advances in applying nanobody-tethered transposases can allow simultaneous profiling of multiple chromatin states within a single cell [120].

Similarly, on the perturbation side, new CRISPR platforms will broaden the set of genome elements that can be reliably perturbed. For example, RNA-targeting Cas13 can be combined with single-cell pooled screens to profile targeted knockdowns of specific transcript isoforms or RNAs that do not code for proteins (e.g., long noncoding RNAs, enhancer RNAs, miRNAs, etc.). All these applications depend, crucially, on sequencing: as the number of single cells increases in an experiment, the amount of sequencing required scales linearly (e.g., if 10 000 reads are required per cell, each additional cell requires an additional 10 000 reads) until saturation is reached. The past few years have seen many exciting developments to reduce sequencing cost through the latest advances in short-read sequencing chemistry (e.g., Ultima Genomics USD $100 per genome [121]) and the increased throughput and accuracy of long-read sequencing (e.g., Oxford Nanopore low-cost devices [122]) (see Outstanding questions). Taken together, these innovations in multimodal capture, new perturbation capabilities, sequencing technologies, and analytical frameworks will enable new types of cutting-edge pooled single-cell screens, thus bringing us closer to understanding the functions of all genes and noncoding elements in the genome [123].

Acknowledgments

We thank the entire Sanjana laboratory for support and advice. J.A.M. is supported by the National Institutes of Health (NIH)/National Human Genome Research Institute (NHGRI) (K99HG012792). N.E.S. is supported by New York University and New York Genome Center startup funds, the NIH/NHGRI (DP2HG010099, R01HG012790), the NIH/National Cancer Institute (R01CA218668, R01CA279135), the NIH/National Institute of General Medical Sciences (R01GM138635), the NIH/National Institute of Allergy and Infectious Diseases (R01AI176601), the NIH/National Institute of Neurological Disorders and Stroke (R01NS124920), the MacMillan Center for the Study of the Noncoding Cancer Genome, and the Simons Foundation for Autism Research Initiative (Genomics of ASD 896724).

Declaration of interests

N.E.S. is an adviser to Qiagen and is a cofounder and adviser of OverT Bio. The other authors declare no conflicts of interest.

References

1. Hsu, P.D. *et al.* (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157, 1262–1278
2. Mohanraju, P. *et al.* (2016) Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* 353, aad5147
3. Makarova, K.S. *et al.* (2020) Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* 18, 67–83
4. Doudna, J.A. and Charpentier, E. (2014) The new frontier of genome engineering with CRISPR-Cas9. *Science* 346, 1258096
5. Shalem, O. *et al.* (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343, 84–87
6. Sanjana, N.E. *et al.* (2014) Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* 11, 783–784
7. Jinek, M. *et al.* (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821
8. Mali, P. *et al.* (2013) RNA-guided human genome engineering via Cas9. *Science* 339, 823–826
9. Cong, L. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823
10. Gasiunas, G. *et al.* (2012) Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 109, E2579–E2586
11. Qi, L.S. *et al.* (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152, 1173–1183
12. Gilbert, L.A. *et al.* (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154, 442–451
13. Bikard, D. *et al.* (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.* 41, 7429–7437
14. Perez-Pinera, P. *et al.* (2013) RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* 10, 973–976
15. Nishimasu, H. *et al.* (2018) Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* 361, 1259–1262
16. Hu, J.H. *et al.* (2018) Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556, 57–63
17. Walton, R.T. *et al.* (2020) Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* 368, 290–296
18. Swarts, D.C. *et al.* (2017) Structural basis for guide RNA processing and seed-dependent DNA targeting by CRISPR-Cas12a. *Mol. Cell* 66, 221–233
19. Abudayyeh, O.O. *et al.* (2017) RNA targeting with CRISPR–Cas13. *Nature* 550, 280–284
20. Özcan, A. *et al.* (2021) Programmable RNA targeting with the single-protein CRISPR effector Cas7-11. *Nature* 597, 720–725
21. Alerasool, N. *et al.* (2020) An efficient KRAB domain for CRISPRi applications in human cells. *Nat. Methods* 17, 1093–1096
22. Konermann, S. *et al.* (2015) Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* 517, 583–588
23. Chavez, A. *et al.* (2015) Highly-efficient Cas9-mediated transcriptional programming. *Nat. Methods* 12, 326–328
24. Liu, P. *et al.* (2018) CRISPR-based chromatin remodeling of the endogenous Oct4 or Sox2 locus enables reprogramming to pluripotency. *Cell Stem Cell* 22, 252–261
25. Komor, A.C. *et al.* (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424

26. Gaudelli, N.M. *et al.* (2017) Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* 551, 464–471
27. Anzalone, A.V. *et al.* (2019) Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576, 149–157
28. Canver, M.C. *et al.* (2018) Integrated design, execution, and analysis of arrayed and pooled CRISPR genome-editing experiments. *Nat. Protoc.* 13, 946–986
29. Bock, C. *et al.* (2022) High-content CRISPR screening. *Nat. Rev. Methods Primer* 2, 9
30. Wang, T. *et al.* (2014) Genetic screens in human cells using the CRISPR-Cas9 system. *Science* 343, 80–84
31. Shifrut, E. *et al.* (2018) Genome-wide CRISPR screens in primary human T cells reveal key regulators of immune function. *Cell* 175, 1958–1971
32. Henriksson, J. *et al.* (2019) Genome-wide CRISPR screens in T helper cells reveal pervasive crosstalk between activation and differentiation. *Cell* 176, 882–896
33. Haney, M.S. *et al.* (2018) Identification of phagocytosis regulators using magnetic genome-wide CRISPR screens. *Nat. Genet.* 50, 1716–1727
34. Porichis, F. *et al.* (2014) High throughput detection of miRNAs and gene-specific mRNA at the single-cell level by flow cytometry. *Nat. Commun.* 5, 5641
35. Fulco, C.P. *et al.* (2019) Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. *Nat. Genet.* 51, 1664–1669
36. Dixit, A. *et al.* (2016) Perturb-Seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* 167, 1853–1866
37. Jaitin, D.A. *et al.* (2016) Dissecting immune circuits by linking CRISPR-pooled screens with single-cell RNA-Seq. *Cell* 167, 1883–1896
38. Xie, S. *et al.* (2017) Multiplexed engineering and analysis of combinatorial enhancer activity in single cells. *Mol. Cell* 66, 285–299
39. Datlinger, P. *et al.* (2017) Pooled CRISPR screening with single-cell transcriptome readout. *Nat. Methods* 14, 297–301
40. Mimitou, E.P. *et al.* (2019) Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat. Methods* 16, 409–412
41. Replogle, J.M. *et al.* (2020) Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nat. Biotechnol.* 38, 954–961
42. Macosko, E.Z. *et al.* (2015) Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 161, 1202–1214
43. Zheng, G.X.Y. *et al.* (2017) Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* 8, 14049
44. Feldman, D. *et al.* (2018) Lentiviral co-packaging mitigates the effects of intermolecular recombination and multiple integrations in pooled genetic screens. *BioRxiv* Published online February 8, 2018. https://doi.org/10.1101/262121
45. Hill, A.J. *et al.* (2018) On the design of CRISPR-based single-cell molecular screens. *Nat. Methods* 15, 271–274
46. Xie, S. *et al.* (2018) Frequent sgRNA-barcode recombination in single-cell perturbation assays. *PLoS ONE* 13, e0198635
47. Liscovitch-Brauer, N. *et al.* (2021) Profiling the genetic determinants of chromatin accessibility with scalable single-cell CRISPR screens. *Nat. Biotechnol.* 39, 1270–1277
48. Daniloski, Z. *et al.* (2021) Identification of required host factors for SARS-CoV-2 infection in human cells. *Cell* 184, 92–105
49. Adamson, B. *et al.* (2016) A multiplexed single-cell CRISPR screening platform enables systematic dissection of the unfolded protein response. *Cell* 167, 1867–1882
50. Schmidt, R. *et al.* CRISPR activation and interference screens decode stimulation responses in primary human T cells. *Science* 375, eabj4008.
51. Replogle, J.M. *et al.* (2022) Mapping information-rich genotype–phenotype landscapes with genome-scale Perturb-seq. *Cell* 185, 2559–2575
52. Schraivogel, D. *et al.* (2020) Targeted Perturb-seq enables genome-scale genetic screens in single cells. *Nat. Methods* 17, 629–635
53. Xie, S. *et al.* (2019) Global analysis of enhancer targets reveals convergent enhancer-driven regulatory modules. *Cell Rep.* 29, 2570–2578
54. Gasperini, M. *et al.* (2019) A genome-wide framework for mapping gene regulation via cellular genetic screens. *Cell* 176, 377–390
55. Morris, J.A. *et al.* (2023) Discovery of target genes and pathways at GWAS loci by pooled single-cell CRISPR screens. *Science* 380, eadh7699
56. Jin, X. *et al.* (2020) In vivo Perturb-Seq reveals neuronal and glial abnormalities associated with autism risk genes. *Science* 370, eaaz6063
57. Santinha, A.J. *et al.* (2023) Transcriptional linkage analysis with in vivo AAV-Perturb-seq. *Nature* 622, 367–375
58. Stoeckius, M. *et al.* (2017) Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* 14, 865–868
59. Wroblewska, A. *et al.* (2018) Protein barcodes enable high-dimensional single-cell CRISPR screens. *Cell* 175, 1141–1155
60. McGinnis, C.S. *et al.* (2019) MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nat. Methods* 16, 619–626
61. Mylka, V. *et al.* (2022) Comparative analysis of antibody- and lipid-based multiplexing methods for single-cell RNA-seq. *Genome Biol.* 23, 55
62. Frangieh, C.J. *et al.* (2021) Multimodal pooled Perturb-CITE-seq screens in patient models define mechanisms of cancer immune evasion. *Nat. Genet.* 53, 332–341
63. Chung, H. *et al.* (2021) Joint single-cell measurements of nuclear proteins and RNA in vivo. *Nat. Methods* 18, 1204–1212
64. Chen, A.F. *et al.* (2022) NEAT-seq: simultaneous profiling of intra-nuclear proteins, chromatin accessibility and gene expression in single cells. *Nat. Methods* 19, 547–553
65. Katzenelenbogen, Y. *et al.* (2020) Coupled scRNA-Seq and intracellular protein activity reveal an immunosuppressive role of TREM2 in Cancer. *Cell* 182, 872–885
66. Blair, J.D. *et al.* (2023) Phospho-seq: Integrated, multi-modal profiling of intracellular protein dynamics in single cells. *BioRxiv* Published online March 28, 2023. https://doi.org/10.1101/2023.03.27.534442
67. Thurman, R.E. *et al.* (2012) The accessible chromatin landscape of the human genome. *Nature* 489, 75–82
68. Buenrostro, J. *et al.* (2015) ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* 109, 21.29.1–21.29.9
69. Buenrostro, J.D. *et al.* (2015) Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* 523, 486–490
70. Rubin, A.J. *et al.* (2019) Coupled single-cell CRISPR screening and epigenomic profiling reveals causal gene regulatory networks. *Cell* 176, 361–376
71. Pierce, S.E. *et al.* (2021) High-throughput single-cell chromatin accessibility CRISPR screens enable unbiased identification of regulatory networks in cancer. *Nat. Commun.* 12, 2969
72. Amini, S. *et al.* (2014) Haplotype-resolved whole genome sequencing by contiguity preserving transposition and combinatorial indexing. *Nat. Genet.* 46, 1343–1349
73. Cao, J. *et al.* (2017) Comprehensive single cell transcriptional profiling of a multicellular organism. *Science* 357, 661–667
74. Cusanovich, D.A. *et al.* (2015) Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* 348, 910–914
75. Xu, Z. *et al.* (2023) PerturbSci-Kinetics: dissecting key regulators of transcriptome kinetics through scalable single-cell RNA profiling of pooled CRISPR screens. *BioRxiv* Published online May 11, 2023. https://doi.org/10.1101/2023.01.29.526143
76. Garibaldi, A. *et al.* (2017) Isolation of newly transcribed RNA using the metabolic label 4-thiouridine. *Methods Mol. Biol.* 1648, 169–176
77. Herzog, V.A. *et al.* (2017) Thiol-linked alkylation of RNA to assess expression dynamics. *Nat. Methods* 14, 1198–1204
78. Iwakawa, H.-O. and Tomari, Y. (2022) Life of RISC: formation, action, and degradation of RNA-induced silencing complex. *Mol. Cell* 82, 30–43
79. Mimitou, E.P. *et al.* (2021) Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells. *Nat. Biotechnol.* 39, 1246–1258

80. Kim, H.S. *et al.* (2021) Single-cell characterization of CRISPR-modified transcript isoforms with nanopore sequencing. *Genome Biol.* 22, 331
81. Rao, S.S.P. *et al.* (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680
82. Oh, S. *et al.* (2021) Enhancer release and retargeting activates disease-susceptibility genes. *Nature* 595, 735–740
83. Popay, T.M. and Dixon, J.R. (2022) Coming full circle: on the origin and evolution of the looping model for enhancer–promoter communication. *J. Biol. Chem.* 298, 102117
84. Dekker, J. *et al.* (2002) Capturing chromosome conformation. *Science* 295, 1306–1311
85. Lieberman-Aiden, E. *et al.* (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289–293
86. Nagano, T. *et al.* (2013) Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 502, 59–64
87. Flyamer, I.M. *et al.* (2017) Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature* 544, 110–114
88. Ramani, V. *et al.* (2017) Massively multiplex single-cell Hi-C. *Nat. Methods* 14, 263–266
89. Ramani, V. *et al.* (2020) Sci-Hi-C: a single-cell Hi-C method for mapping 3D genome organization in large number of single cells. *Methods* 170, 61–68
90. Liu, Z. *et al.* (2023) Linking genome structures to functions by simultaneous single-cell Hi-C and RNA-seq. *Science* 380, 1070–1076
91. Guo, Y. *et al.* (2015) CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* 162, 900–910
92. Hussmann, J.A. *et al.* (2021) Mapping the genetic landscape of DNA double-strand break repair. *Cell* 184, 5653–5669
93. Feldman, D. *et al.* (2019) Optical pooled screens in human cells. *Cell* 179, 787–799
94. Kanfer, G. *et al.* (2021) Image-based pooled whole-genome CRISPRi screening for subcellular phenotypes. *J. Cell Biol.* 220, e202006180
95. Yan, X. *et al.* (2021) High-content imaging-based pooled CRISPR screens in mammalian cells. *J. Cell Biol.* 220, e202008158
96. Hu, M. *et al.* (2023) High-content image-based CRISPR screening reveals regulators of 3D genome architectures. *BioRxiv* Published online February 1, 2023. https://doi.org/10.1101/2023.01.31.525983
97. Konermann, S. *et al.* (2018) Transcriptome engineering with RNA-targeting type VI-D CRISPR effectors. *Cell* 173, 665–676
98. Abudayyeh, O.O. *et al.* (2016) C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* 353, aaf5573
99. Cox, D.B.T. *et al.* (2017) RNA editing with CRISPR-Cas13. *Science* 358, 1019–1027
100. Wessels, H.-H. *et al.* (2020) Massively parallel Cas13 screens reveal principles for guide RNA design. *Nat. Biotechnol.* 38, 722–727
101. Wessels, H.-H. *et al.* (2023) Prediction of on-target and off-target activity of CRISPR–Cas13d guide RNAs using deep learning. *Nat. Biotechnol.* Published online July 3, 2023. https://doi.org/10.1038/s41587-023-01830-8
102. Méndez-Mancilla, A. *et al.* (2022) Chemically modified guide RNAs enhance CRISPR-Cas13 knockdown in human cells. *Cell Chem. Biol.* 29, 321–327
103. Wessels, H.-H. *et al.* (2023) Efficient combinatorial targeting of RNA transcripts in single cells with Cas13 RNA Perturb-seq. *Nat. Methods* 20, 86–94
104. Legut, M. *et al.* (2022) A genome-scale screen for synthetic drivers of T cell proliferation. *Nature* 603, 728–735
105. Fraietta, J.A. *et al.* (2018) Disruption of TET2 promotes the therapeutic efficacy of CD19-targeted T cells. *Nature* 558, 307–312
106. Li, X. *et al.* (2023) Chromatin context-dependent regulation and epigenetic manipulation of prime editing. *BioRxiv* Published online April 12, 2023. https://doi.org/10.1101/2023.04.12.536587
107. Dobin, A. *et al.* (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21
108. Melsted, P. *et al.* (2021) Modular, efficient and constant-memory single-cell RNA-seq preprocessing. *Nat. Biotechnol.* 39, 813–818
109. He, D. *et al.* (2022) Alevin-fry unlocks rapid, accurate and memory-frugal quantification of single-cell RNA-seq data. *Nat. Methods* 19, 316–322
110. Bray, N.L. *et al.* (2016) Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* 34, 525–527
111. Barry, T. *et al.* (2022) Exponential family measurement error models for single-cell CRISPR screens. *ArXiv* Published online January 6, 2022. http://arxiv.org/abs/2201.01879
112. Yang, L. *et al.* (2020) scMAGeCK links genotypes with multiple phenotypes in single-cell CRISPR screens. *Genome Biol.* 21, 19
113. Finak, G. *et al.* (2015) MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* 16, 278
114. Papalexi, E. *et al.* (2021) Characterizing the molecular regulation of inhibitory immune checkpoints with multimodal single-cell screens. *Nat. Genet.* 53, 322–331
115. Burkhardt, D.B. *et al.* (2021) Quantifying the effect of experimental perturbations at single-cell resolution. *Nat. Biotechnol.* 39, 619–629
116. Barry, T. *et al.* (2021) SCEPTRE improves calibration and sensitivity in single-cell CRISPR screen analysis. *Genome Biol.* 22, 344
117. Morris, J.A. *et al.* (2021) Discovery of target genes and pathways of blood trait loci using pooled CRISPR screens and single cell RNA sequencing. *BioRxiv* Published online April 8, 2021. https://doi.org/10.1101/2021.04.07.438882
118. Wang, L. (2021) Single-cell normalization and association testing unifying CRISPR screen and gene co-expression analyses with Normalisr. *Nat. Commun.* 12, 6395
119. Datlinger, P. *et al.* (2021) Ultra-high-throughput single-cell RNA sequencing and perturbation screening with combinatorial fluidic indexing. *Nat. Methods* 18, 635–642
120. Stuart, T. *et al.* (2022) Nanobody-tethered transposition enables multifactorial chromatin profiling at single-cell resolution. *Nat. Biotechnol.* 41, 806–812
121. Almogy, G. *et al.* (2022) Cost-efficient whole genome-sequencing using novel mostly natural sequencing-by-synthesis chemistry and open fluidics platform. *BioRxiv* Published online August 10, 2022. https://doi.org/10.1101/2022.05.29.493900
122. Wang, Y. *et al.* (2021) Nanopore sequencing technology, bioinformatics and applications. *Nat. Biotechnol.* 39, 1348–1365
123. Green, E.D. *et al.* (2020) Strategic vision for improving human health at the forefront of genomics. *Nature* 586, 683–692