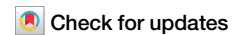


<https://doi.org/10.1038/s42003-026-09675-y>

A practical guide to targeted single-cell RNA sequencing technologies

Giulia Moro , Erich Brunner  & Konrad Basler

Current single-cell RNA sequencing (scRNA-seq) methods suffer from biases that restrict the detection of cellular transcripts to 10–40% of total RNAs. This hinders the identification of transcripts of interest. Additionally, information retrieved from most high-throughput scRNA-seq methods is confined to untranslated regions toward transcript ends, resulting in loss of detail in internal regions. In this review, we outline biases in scRNA-seq protocol steps that limit transcript and region detection. We then discuss the advantages and disadvantages of targeted sequencing solutions, grouped into five categories according to the protocol step they target. Finally, we present a decision tree that guides researchers in selecting the most suitable targeted method for their experiment.

Single-cell RNA sequencing (scRNA-seq) has emerged as a powerful tool to record and dissect gene expression of individual cells in tissues. Compared to bulk RNA-seq, where RNAs from all cells in a tissue or population are pooled and processed together, scRNA-seq technologies face a key challenge: retaining information at the individual cell level. This is achieved by tagging all RNAs (or cDNAs) from a single cell with a unique sequence (called an index or barcode), with which the origin of each transcript can be identified after sequencing. To do this, single cells must first be isolated into individual reaction chambers, typically droplets or microwells. In most protocols, the RNAs are then reverse transcribed into cDNA and PCR-amplified to generate sufficient material for high-throughput sequencing (HTS).

Over the past decade, the number of cells that can be sequenced, and the simplicity of scRNA-seq protocols have greatly increased^{1–6}. Ideally, every RNA in a single cell would be detected after sequencing. However, available scRNA-seq technologies capture only 10–40% of cellular transcripts^{1,5,7–11}, resulting in substantial information loss. This major limitation often impacts the detection of transcripts of interest (TOIs), such as cell markers, which are important for assigning cellular identity. Beyond detection of transcripts, the sequences of the RNAs themselves contain additional, highly valuable information for specific research or diagnostic applications. For example, RNA sequence data can reveal the presence of point mutations, splice junctions, fusion breakpoints, or uncover rearrangements at CRISPR target sites. We refer to these features as regions of interest (ROIs).

High-throughput, 3'/5'-based scRNA-seq methods are widely used for their ability to profile large numbers of cells and detect numerous transcripts efficiently. Popular platforms include 10x Chromium¹², DROP-seq¹³, inDrop¹⁴, and BD Rhapsody¹⁵, which use barcoded beads (microbeads covered with DNA oligonucleotides containing a set of sequences) to capture and tag mRNA molecules. However, a key limitation of these methods is their focus on transcript ends (either 3' ends or, with newer developments,

5' ends¹⁶), resulting in the lack of data from internal regions where ROIs are typically located.

Full-length protocols address the limitations of ROI detection by acquiring sequence profiles across the entire transcript length. These technologies include short-read based protocols (such as Smart-seq technologies^{17–22} and VASA-seq²³) which fragment transcripts or cDNAs before sequencing, and long-read sequencing technologies from PacBio or Oxford Nanopore Technologies (ONT) that directly sequence long nucleotide stretches^{24–27}. However, compared to 3'/5'-based methods, full-length technologies process significantly fewer cells (hundreds versus thousands), revealing an important tradeoff⁵: while they provide comprehensive coverage of internal transcript regions containing ROIs, they do so at the expense of cellular throughput.

The relevance of enhanced TOI and ROI detection in scRNA-seq experiments is emphasized by the numerous solutions developed to address this issue. Collectively known as targeted sequencing approaches, these methods are designed to detect specific ROIs within transcripts or to enrich for TOIs important to determine cellular traits.

Here, we provide a broad overview on the targeted scRNA-seq field. We first examine the biases affecting TOI and ROI detection in standard 3'/5'-based scRNA-seq experiments. We then give a comprehensive overview of targeted methods, highlighting the biases they address, their advantages and limitations. While the focus is on methods using barcoded beads, we also mention plate-based protocols. Depending on the targeting protocol used by the method, we group them into five classes: targeted capture, targeted priming, targeted amplification, dual-targeted polymerase chain reaction (PCR) and probe hybridization. Next, we explore long-read sequencing protocols which leverage targeting for improved transcript detection and characterization. Finally, we highlight targeted technologies in the field of spatial transcriptomics and provide examples of biological applications of targeted technologies. To summarize the review, we provide

a practical decision tree to guide researchers in selecting the most suitable method for their needs and discuss the future directions of targeted scRNA-seq approaches.

Biases affecting TOI and ROI detection

The detection of TOIs and ROIs in 3'/5'-based methods is influenced by various biases introduced throughout the scRNA-seq workflow (Fig. 1).

Many of these biases were described for bulk RNA-seq experiments²⁸ but also apply to scRNA-seq protocols as similar steps are used for the library generation. The specific biases occurring during a typical scRNA-seq workflow are detailed in Fig. 1b and described below, listed based on the protocol step they influence.

We further categorize these biases into three groups in Table 1: those inherent to scRNA-seq protocols, those related to transcript sequence

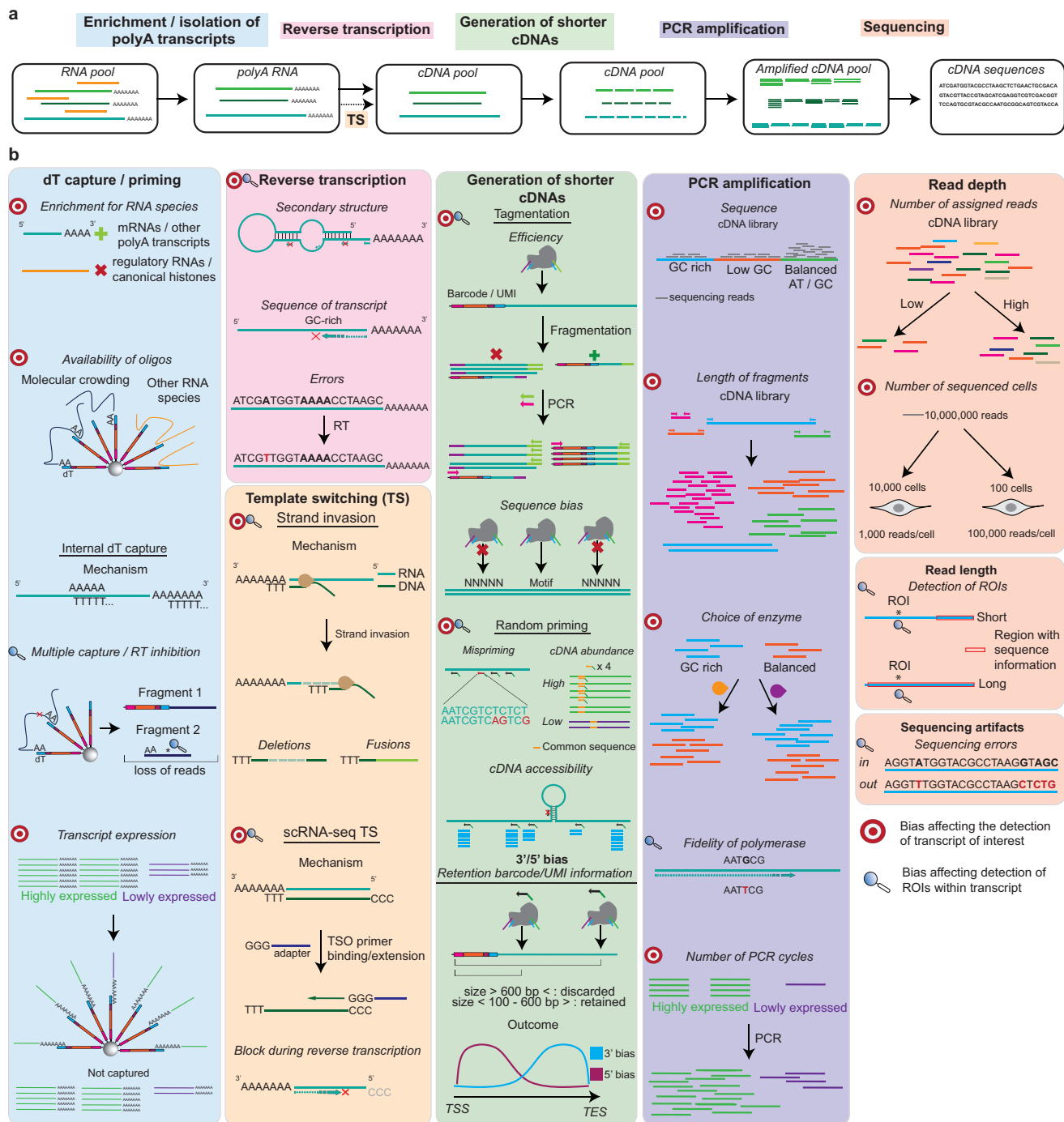


Fig. 1 | Main steps in 3'/5'-based scRNA-seq experiments and biases affecting each step. a Most scRNA-seq experiments start with enrichment for polyadenylated transcripts through dT-based capture or priming. RNAs are then converted into cDNAs through reverse transcription, which may include a template switching (TS) reaction. Shorter cDNAs suited for short-read sequencing are then generated, followed by PCR amplification and sequencing. Note: the order of the steps outlined at the top of the figure may vary between scRNA-seq protocols. b Bullseyes indicate

biases which affect detection of TOIs, while magnifying glasses indicate issues in detecting ROIs. Biases affecting ROI detection are only considered as such if the detection of TOIs is not affected. scRNA-seq TS refers to the reaction in which the terminal transferase activity adds untemplated Cs at the end of the template, followed by recognition by a TSO primer with a complementary stretch of Gs and subsequent extension to generate a second strand.

Table 1 | Summary of biases influencing the results of 3'/5'-based scRNA-seq experiments

Bias	Inherent to:	Consequence of bias	Strength of bias for TOI	Strength of bias for ROI
dT capture/priming	protocol	loss of information on non-polyA transcripts/loss of information on transcripts	strong	NA
Internal dT capture	transcript	loss of reads to non-informative transcripts/incomplete reverse transcription (generation of fragments)	low	low
Reverse transcription	transcript	incomplete across full transcript/loss of information on lowly expressed transcripts/sequence errors	strong	strong
Template switching	protocol	incomplete across full transcript/fusions and deletions of sequences/sequence artifacts	moderate	moderate
Generation of shorter cDNA fragments	protocol	loss of information on coding region/loss of information on cDNAs/long cDNAs (size excluded and lost)/sequence errors/loss of information on lowly expressed transcripts	moderate	high
PCR amplification	transcript	loss of information on transcripts/changes in relative abundance of cDNAs/sequencing errors	high	moderate
Sequencing depth	experimental design	loss of information on transcripts (especially lowly expressed)	high	high
Read length	experimental design	loss of information on regions of interest/assignment of read to transcript	moderate	high
Sequencing artifacts	transcript	false conclusions on original sequence of cDNAs	low	moderate

As the detection of a ROI directly depends on the detection of TOIs, we only consider biases affecting the detection of ROIs if the transcript to which it belongs to is detected. For this reason, the detection of the ROI is not available for the dT capture/priming (NA). The strength of the bias depends on the specific TOI and ROI itself, as well as the protocol conditions used and the experimental design.

characteristics, and those stemming from experimental design choices. Each bias creates limitations affecting TOI and ROI detection to varying degrees, depending on the specific properties of the TOIs and ROIs themselves.

Preparation of single cells

While often overlooked in scRNA-seq workflows, sample preparation and the isolation of single cells is a critical prerequisite for all experiments²⁹. However, these protocols are known to introduce various biases³⁰. Although dissociation protocols are tailored to the tissue type, the desired outcome is the same: preserving tissue composition, maintaining intact and viable cells, and minimizing alterations to the transcriptome during processing. Low-quality cells can result in fragmented RNA and an elevated proportion of mitochondrial transcripts, which can dominate sequencing reads and impact TOI and ROI detection. Several strategies have been developed to preserve native gene expression during dissociation, including the use of psychrophilic proteases³¹ which help preserve in vivo transcriptional states, or the isolation of single nuclei from tissue (snRNA-seq). Additionally, many scRNA-seq experiments begin with frozen or fixed tissue due to logistical constraints, which can also affect the cellular transcriptome. All of these steps introduce biases which may confound the output of the analysis³⁰ and impact TOI and ROI detection.

dT-based capture/priming and internal capture

The aim of most scRNA-seq experiments is the detection of polyadenylated transcripts, primarily protein-coding mRNAs, which constitute only 3–7% of all cellular RNAs by mass³². After cell lysis, mRNAs are captured through their polyA-tails, which bind to oligo(dT) stretches on barcoded beads or are targeted using oligo(dT) priming strategies. This polyA-dependent approach is employed by the majority of scRNA-seq methods, with notable exceptions being scFAST-seq³³, MATQ-seq²², SUPeR-Seq³⁴ and SMARTer (SMART-seq Total RNA-seq Single Cell)³⁵ which use alternative strategies.

dT-based capture/priming introduces several biases hampering TOI and ROI detection. One major limitation is the exclusion of non-polyadenylated RNAs, such as regulatory RNAs and canonical histone RNAs. Additionally, although there is a large excess of oligonucleotides on barcoded beads compared to mRNAs expressed in eukaryotic cells (10⁶–10⁷ DNA oligonucleotides vs 3–10 × 10⁵ mRNAs³²), binding may be impaired due to steric hindrance and molecular crowding (where high numbers of macromolecules occupy most of the free space; Fig. 1).

Internal capture artifacts^{36,37} further restrict the availability of oligonucleotides on barcoded beads. The same transcript may be captured at different positions, overrepresenting the analyte in the final dataset. Capture of non-polyadenylated transcripts further reduces the pool of oligonucleotides available for binding other transcripts. rRNAs are particularly problematic in this regard, as they constitute 80–90% of eukaryotic samples by mass³² and can make up 5–75% of reads in scRNA-seq experiments^{38,39}. Internal capture also limits ROI detection by generating multiple fragments which are then lost in downstream protocol steps, resulting in a depletion of information at the 5' end of RNAs^{36,37}. A class of TOIs that is particularly impacted is transcription factors, which are notoriously difficult to capture in scRNA-seq experiments^{40–42} due to their low expression.

Reverse transcription

The oligo(dT) sequences on barcoded beads serve two functions: they isolate polyadenylated RNAs and act as primers for reverse transcription. This critical step converts captured RNAs into cDNAs for subsequent PCR amplification but itself introduces several biases^{43,44}. Importantly, a one-size-fits-all type of reaction is performed to accommodate the broad range of transcripts being processed. However, reverse transcription efficiency is influenced by transcript-specific features such as RNA secondary structures^{45,46} and unfavorable sequence composition, including high GC content^{47,48}. These properties may cause premature enzyme dissociation from the transcribed RNA, generating incomplete cDNAs (i.e., not including the 5' end of a transcript), potentially impacting ROI information

depending on the position of the dissociation. In addition, reverse transcriptases lack proofreading activity^{49,50}, which can result in mutations or sequence errors during cDNA synthesis. If the change in sequence consistently occurs at the same position in the template, assignment of nucleotide polymorphisms may be incorrect⁵⁰, especially problematic if these sequences belong to ROIs. Finally, the efficiency of the reverse transcription step has been found to be dependent on the choice of reverse transcriptase enzymes^{48–54}.

Biases in reverse transcription of transcript pools are currently very difficult to overcome, but this step can be omitted by using direct RNA sequencing (DRS)^{55,56}, in which RNAs are directly sequenced using fluorescent oligonucleotides, or by sequencing with ONT⁵⁷. Additionally, the primer site can be influenced by ligating adapters to RNAs⁵⁸. While these methods circumvent the need for cDNA synthesis, they obtain information in bulk and currently cannot achieve high throughput.

Template switching

Most reverse transcriptases used in scRNA-seq protocols exhibit two distinct activities commonly referred to as template switching (TS): (1) strand invasion, in which the nascent cDNA dissociates from its original RNA template and anneals to a homologous region on a different transcript, and (2) scRNA-seq template switching (scRNA-seq TS), a term introduced in this work to describe a protocol-specific mechanism driven by the enzyme's terminal transferase activity⁵⁹.

Strand invasion can lead to the formation of chimeric cDNAs (e.g., as false fusion transcripts or circular RNAs) or deletions, complicating the accurate detection of alternative splicing events^{60–62}. In contrast, scRNA-seq TS involves the addition of untemplated cytosines to the 3' end of the newly synthesized cDNA when the reverse transcriptase reaches the 5' end of the RNA³⁹. A template switching oligonucleotide (TSO) then anneals to these cytosines and serves as a new template for cDNA extension, enabling capture of full-length transcripts or enrichment for 5' ends^{63–65}. However, this reaction is inherently inefficient⁶⁶ as it depends on the enzyme successfully transcribing the entire RNA molecule to its 5' end. While the template switching efficiency is difficult to quantify, Hughes et al.⁶⁶ implemented a randomly primed second-strand synthesis method (called Seq-Well S³) to recover cDNA molecules where scRNA-seq TS had failed. The authors were able to increase gene detection by tenfold compared to samples without recovery through random priming, highlighting the importance of this step.

Both strand invasion and scRNA-seq TS are influenced by the dissociation dynamics of reverse transcriptase, which in turn depend on transcript-specific properties such as sequence composition, secondary structure, and length.

Generation of shorter cDNAs

Generation of cDNAs between 200 and 600 base pairs (bp) in length is required for scRNA-seq libraries to be compatible with short-read sequencing^{67–69}. Two main approaches are used to obtain cDNAs of the appropriate length: tagmentation and random priming. During tagmentation, a transposase simultaneously fragments DNA and appends adapters needed for downstream amplification, whereas random priming involves primers with random sequences binding at different positions along cDNAs. In addition, cDNA size is controlled during library cleanup using size-selection beads, which remove unfavorably sized fragments (including primer dimers) that could bind to flow cells without producing usable data.

Tagmentation introduces several biases⁷⁰. Importantly, only fragments with the correct adapter attached at the correct end of the cDNA are compatible with amplification and sequencing (Fig. 1b). This occurs for 50% of fragments, while the remainder will undergo linear amplification or will not be amplified⁷⁰. Furthermore, transposase insertion sites are not entirely random but exhibit sequence preferences, which can lead to uneven coverage^{71–75}.

Random priming also introduces several biases, many of which are in common with reverse transcription reactions using random primers. Primers may partially anneal to non-complementary regions, typically with

perfect base pairing at the 3' end but mismatches at the 5' end. Upon extension, these primers will be incorporated into the second cDNA strand, resulting in cDNAs with altered or artifactually extended 5' ends⁷⁶. These ends should be removed during data analysis, as they can lead to false-positive mutation calling and biased mapping. Primers may also be sequestered by highly expressed or longer transcripts⁴⁸, contributing to the underrepresentation of lowly expressed transcripts in the final library. Finally, structural bias arises when stable RNA secondary structures prevent primer binding, resulting in localized coverage loss.

To be informative and amplifiable in scRNA-seq protocols, cDNAs generated by tagmentation or random priming must have the appropriate size and contain both cell barcode and unique molecular identifier (UMI) sequences. UMIs are unique sequences assigned to each captured or primed transcript and are used to distinguish individual molecules after PCR amplification. Both barcode and UMI sequences are located on barcoded beads adjacent to the captured region. As a result, sequence information is typically limited to regions within ~600 bp of the capture site which, in most cases, is the polyA tail. As the average length of 3' UTRs in vertebrates is over 800 bp and 100–200 bp for 5' UTRs^{77,78}, the information obtained with barcoded beads is mostly restricted to these regions, severely impacting retrieval of sequences from coding sequences. This bias is observed for all 3'/5'-based methods and is the most substantial limitation affecting ROI detection.

PCR biases

PCR amplification is needed to produce sufficient input for HTS. As a general rule, short cDNAs with a balanced GC content composition will be preferentially amplified⁷⁹, with regions having high or low GC content exhibiting a biased coverage after sequencing^{80–82}. This effect is dependent on the choice of polymerase⁸³. Amplification biases related to cDNA length have also been described⁸⁰, with efficiencies decreasing for longer cDNAs⁸⁴. Similarly to reverse transcription, PCR conditions are not tailored to amplify specific cDNAs⁸⁵, exacerbating the loss of information for unfavorable cDNAs. Transcript representation is also affected by the number of PCR cycles⁸⁵, with highly abundant cDNAs being preferentially amplified to the detriment of lowly abundant TOIs. Finally, detection of ROIs in scRNA-seq experiments is also influenced by the fidelity of the polymerase used⁸⁶. Taken together, PCR biases can strongly distort transcript abundance estimates and affect the detection of both TOIs and ROIs, depending on their specific sequences.

High-throughput sequencing (HTS)

One of the most important parameters to consider when designing RNA-seq experiments is sequencing depth⁸⁷ which strongly impacts ROI detection^{88,89}. As highly abundant cDNAs will be preferentially sequenced⁸⁸, the possibility of detecting lowly expressed transcripts underrepresented in the library increases with higher sequencing depth. Information obtained after HTS sequencing is also influenced by the choice of read length, particularly relevant for ROI detection. Depending on the distance of the ROI to the 3' end of cDNAs, the sequence may be detected by using longer reads. HTS sequencing also introduces sequencing errors, known to be influenced by GC-rich motifs which often precede these inaccuracies^{82,90,91}. The pattern of sequencing errors varies between short-read and long-read sequencing platforms such as PacBio and ONT. Although long-read technologies previously had higher error rates than short-read sequencing, recent advances have greatly improved their accuracy, with current error rates around 99% for ONT, 99.9% for PacBio, and 99.9% for Illumina⁹² (the most popular short-read sequencing platform to date⁹³). Careful optimization of sequencing depth and read length is thus essential to maximize recovery of low-abundance transcripts and reliably identify ROIs, particularly in complex single-cell datasets.

All of the biases described above impact the output from scRNA-seq experiments, influencing ROI and ROI detection. Several methods have been developed to mitigate these biases in a general, untargeted manner (Box 1). These methods also act at various levels, including sample

Box 1 | Beyond targeted methods: additional solutions to overcome biases in scRNA-seq experiments

Several solutions have been developed to address the biases described above. These methods do not focus on a given TOI or ROI, but aim to improve the overall information obtained from scRNA-seq experiments. They encompass a wide variety of methods acting at various levels during the scRNA-seq workflow.

Increasing information on cells of interest during sample preparation:

The main research focus of some scRNA-seq experiments is analyzing rare cell populations or specific cell types. Traditionally, such cells are enriched using FACS sorting, magnetic beads, or specialized isolation protocols. However, defining these populations can be challenging when transcriptomic signatures must be translated into protein markers, especially for transcription factors that often lack suitable antibodies. This is especially the case for FACS sorting, as this technology depends on the available antibodies or reporter strains. To overcome these challenges, recent methods have focused on isolating cells of interest based on the transcripts they express. For example, PERF-seq¹⁴⁸ and EnrichSci¹⁴⁹ use probe-specific detection of transcripts in cell, which are then FACS-sorted for downstream scRNA-seq analysis. By enriching for specific transcript-defined populations, these approaches enhance the recovery and resolution of information from cells of interest.

Depleting highly abundant RNAs: Several methods have been developed to deplete highly expressed, non-informative transcripts, mainly rRNAs. This is achieved through specific nucleases¹⁵⁰, by CRISPR/Cas9 degradation with guide RNAs against the transcripts to deplete^{38,151,152,153}, by digestion with RNaseH²³ or by using primers to block the extension or reverse transcription of selected transcripts during PCR amplification^{33,154}. All of these methods lead to similar results: depletion of highly expressed, non-informative transcripts increases the detection of lowly expressed RNAs in scRNA-seq experiments. For example, through scCLEAN¹⁵³ the authors were able to more than double the percentage of non-targeted UMIs (from 39% to 92%). Depleting highly abundant RNAs frees additional oligonucleotides on barcoded beads for mRNA capture (reducing dT capture/priming bias). It also

benefits lowly expressed transcripts by increasing the available read depth and preventing their underrepresentation during PCR amplification.

Using a transcript polyadenylation step to profile non-polyadenylated transcripts: Non-polyadenylated transcripts are missed in scRNA-seq experiments based on dT capture/priming. These transcripts are highly relevant for many biological questions, including detection of viral, bacterial and non-coding RNAs (such as microRNAs, snoRNAs and some lncRNAs). To overcome this challenge and obtain this information in combination with polyadenylated RNAs, several methods have included transcript polyadenylation. These include VASA-seq²³, where transcripts are first fragmented after which each fragment gets a polyA tail, Smart-seq-total¹⁵⁵, where all transcripts are polyadenylated (followed by depletion of rRNA transcripts) and SUPeR-seq³⁴, SMARTer³⁵ and MATQ-seq²², with reverse transcription based on random primers carrying an adaptor.

Overcoming PCR biases: PCR bias in scRNA-seq experiments is mitigated by using UMIs to control for over-amplification of cDNAs. Several additional methods have been developed for bulk RNA-seq libraries. Kozarewa et al.¹⁵⁶ designed a protocol relying on ligation of Illumina sequencing adapters directly to first-strand cDNAs. Compared to standard bulk RNA-seq protocols using PCR amplification, lower GC coverage bias and duplicate sequences were observed. An alternative method, smsDGE¹⁵⁷ relies on hybridization of first-strand cDNAs to flow cells, followed by direct sequencing using fluorescent oligonucleotides. Mamanova et al.¹⁵⁸ developed FRT-seq, which relies on reverse transcription of RNAs directly on flow cells, allowing sequencing of first cDNA strands. While these methods have not been extended to scRNA-seq experiments, linear amplification through in vitro transcription using a T7 DNA-based RNA polymerase has emerged as an alternative for minimizing the number of PCR cycles. Methods using this step include MARS-seq¹⁵⁹, CEL-seq⁸ and LAST-seq¹⁶⁰.

generation, capture on barcoded beads and PCR. However, they cannot specifically enrich for information on TOI and ROIs, which may still missed.

Targeted scRNA-seq strategies using short-read sequencing

The loss of information on TOIs and ROIs introduced by the biases described above is addressed by targeted methods acting at different steps in the scRNA-seq protocol. We have subdivided these methods into five different classes based on their targeted enrichment strategy (Fig. 2a): (1) targeted capture, (2) targeted priming, (3) targeted amplification, (4) dual targeted PCR and (5) probe hybridization. Each category of targeted methods addresses a certain subset of biases (Fig. 2a) and is associated with both advantages and inherent limitations, outlined in each section of the text. A timeline of the development of targeted scRNA-seq methods is outlined in Fig. 2b, while a schematic representation of their protocols is shown in Fig. 3. An overview of their key characteristics, including number of targets, number of cells and detection of the standard transcriptome is provided in Table 2.

Category 1: Targeted capture

Targeted capture is based on hybridization of cellular RNAs to barcoded beads modified with oligonucleotides complementary to internal regions of targets of interest (TOIs). These capture sequences physically bind TOIs from the diverse RNA pool, enriching them prior to reverse transcription and enhancing their detection while maintaining coverage of the standard

cellular transcriptome (the transcriptome obtained in the non-targeted version of the method). Targeted capture methods tackle five biases (Table 1, Fig. 2a): dT capture/priming, reverse transcription, 3'/5' bias, read depth and read length. The dT capture/priming step is addressed with the capture sequence on the beads by influencing the position from which reverse transcription initiates. By initiating reverse transcription at internal regions, rather than at the transcript ends, information which would be limited to the 3' or 5' end of transcripts can be shifted, increasing ROI detection and decreasing the read length needed to reach the ROI(s).

Three examples of methods using targeted capture are DART-seq⁹⁴, RoCKseq⁹⁵ and direct capture Perturb-seq⁹⁶. DART-seq and RoCK-seq are both based on solid barcoded beads. In DART-seq, a subset of the poly(dT) oligonucleotides on Drop-seq beads is modified with capture sequences, allowing simultaneous enrichment of TOIs and detection of the standard transcriptome. In contrast, RoCK-seq modifies only the template-switch oligonucleotides (TSOs) on BD Rhapsody beads, leaving all poly(dT) oligonucleotides available for standard transcriptome capture. The two methods use different strategies for bead modification: DART-seq uses a ligase to attach a partially double-stranded oligonucleotide, whereas RoCK-seq employs a polymerase-based approach using a single-stranded splint oligonucleotide as a template. These differences influence the properties of the two methods. DART-seq achieves a variable modification rate ranging between 20 and 40%, while RoCKseq achieves modification rates near completion, which can be titrated, multiplexed or used in combination with different assays. As DART-seq uses the same dT oligonucleotides for both

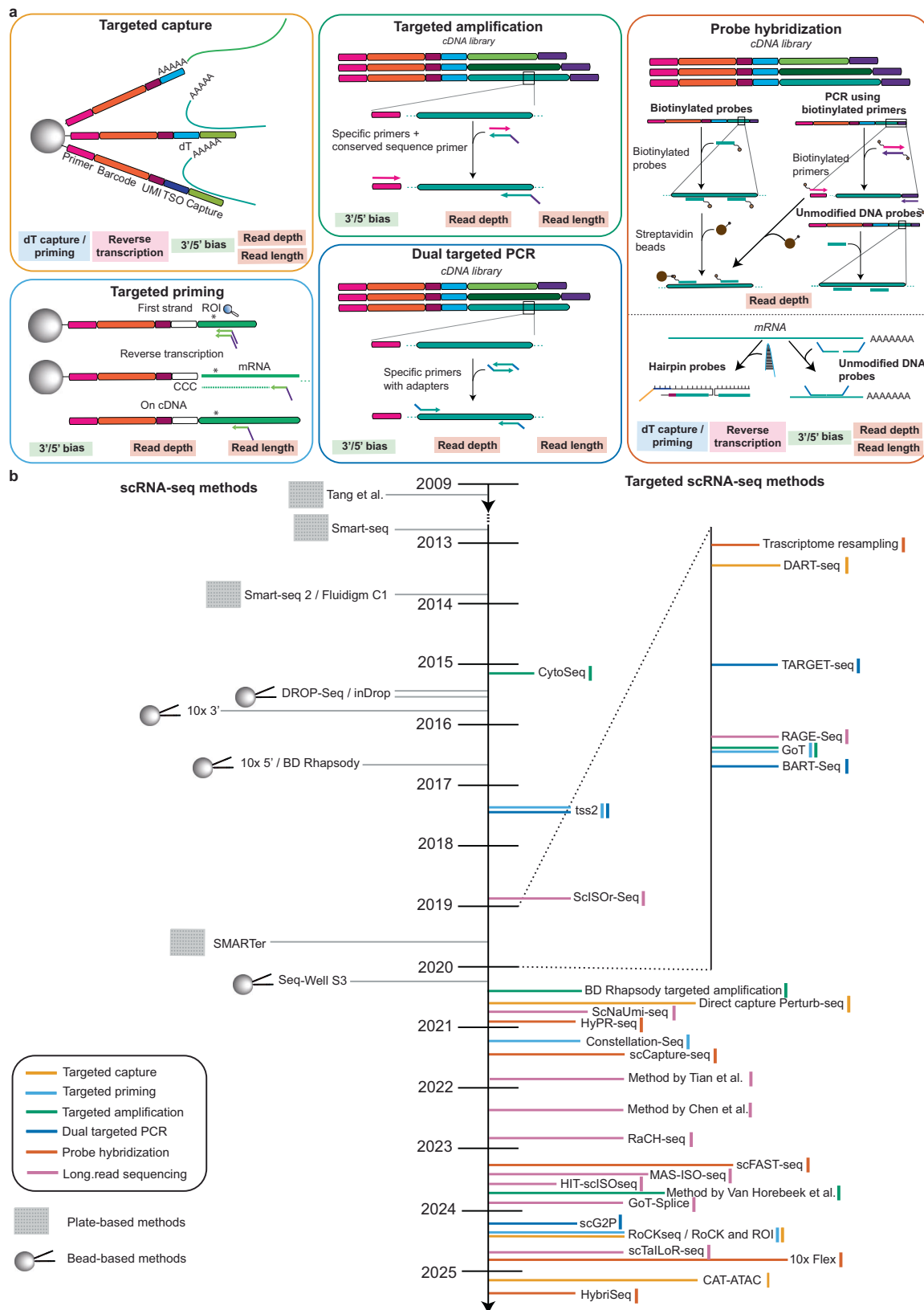


Fig. 2 | Categorization of five targeting strategy in scRNA-seq experiments. a Five categories of targeted scRNA-seq methods using short-read sequencing. The biases which they address are listed below. All of the methods except for targeted capture can be directly applied in the context of current scRNA-seq technologies using either barcoded beads or dT primers for mRNA capture. **b** Timeline representing the targeted scRNA-seq methods and the release of the platforms which they are based on. For all methods the release date of the preprint or of publishing of the paper was used. Methods shown in the figure include CytoSeq¹⁵, tss2⁹⁹, SclSO-Seq¹¹⁹,

Transcriptome resampling¹⁰⁶, DART-seq⁹⁴, TARGET-seq¹⁰⁴, RAGE-Seq¹¹⁸, GoT⁹⁸, BART-Seq¹⁰³, BD Rhapsody targeted amplification¹⁰⁰, Direct capture Perturb-seq⁹⁶, ScNaUmi-seq¹²⁰, HyPR-seq¹¹⁰, Constellation-Seq⁹⁷, scCapture-seq¹⁰⁷, Method by Chen et al.¹²⁶, RaCH-seq¹²², scFAST-seq³³, MAS-ISO-seq¹²⁴, HIT-sclSOseq¹²³, Method by Van Horebeek et al.¹⁰², GoT-Splice¹⁰¹, scG2P¹⁰⁵, RoCKseq / RoCK and ROI⁹⁵, scTalLoR-seq¹²⁵, 10x Flex¹⁰⁹, CAT-ATAC¹⁴⁵, HybriSeq¹⁰⁸, Method by Tian et al.¹²¹.

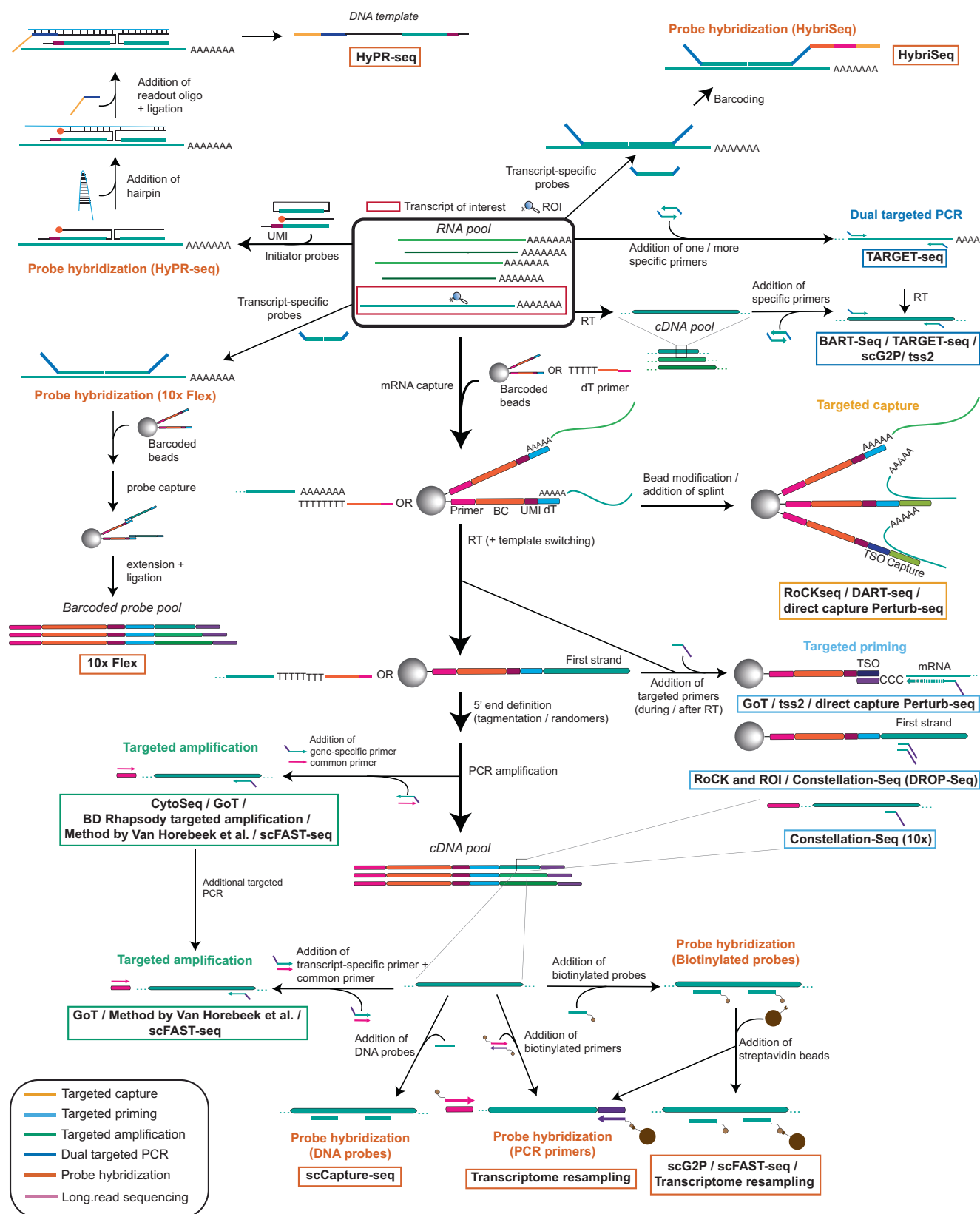


Fig. 3 | Summary of targeted scRNA-seq methods. RT: reverse transcription. Methods belonging to the five categories of the scRNA-seq protocol and are subdivided into dT priming and capture, while lighter arrows indicate additional steps which are needed for specific targeting towards TOI and ROI detection. Methods shown in the

figure include CytoSeq¹⁵, tss2⁹⁹, Transcriptome resampling¹⁰⁶, DART-seq⁹⁴, TARGET-seq¹⁰⁴, GoT⁹⁸, BART-Seq¹⁰³, BD Rhapsody targeted amplification¹⁰⁰, Direct capture Perturb-seq⁹⁶, HyPR-seq¹¹⁰, Constellation-Seq⁹⁷, scCapture-seq¹⁰⁷, scFAST-seq³³, Method by Van Horebeek et al.¹⁰², scG2P¹⁰⁵, RoCKseq / RoCK and ROI⁹⁵, 10x Flex¹⁰⁹, HybriSeq¹⁰⁸.

Table 2 | Main features of targeted scRNA-seq methods

Method	Platform they are based on	Throughput	Type of targeting	Show targeting of TOIs	Show targeting of ROIs	# targets described	Standard transcriptome	Show targeting of TOIs	Show targeting of ROIs
CytoSeq	Custom	medium	Targeted amplification	yes	no	111	no	yes	no
tss2	Smart-seq 2/ Fluidigm C1	low	Targeted priming / dual targeted PCR	yes	yes	1	yes	yes	yes
BART-Seq	Custom	low	Dual targeted PCR	yes	yes	15	no	yes	yes
DART-seq	DROP-Seq	high	Targeted capture	yes	yes	2	yes	yes	yes
GoT	5'/(D)J 10x Chromium and 3' 10x Chromium	high	Targeted priming at reverse transcription level / targeted amplification	yes	yes	5	yes	yes	yes
TARGET-seq	Custom	medium	Dual targeted PCR	no	yes	12	yes	no	yes
Transcriptome resampling	3' 10x Chromium	low	Probe hybridization (biotin)	yes	no	NA	yes	yes	no
BD Rhapsody targeted amplification	BD Rhapsody	high	Targeted amplification	yes	no	492	no	yes	no
Direct capture Perturb-seq	5'/(D)J 10x Chromium and 3' 10x Chromium	high	Targeted capture / targeted priming	yes	yes	1	yes	yes	yes
HyPR-seq	Custom	medium	Probe hybridization (hairpin probes)	yes	yes	179	no	yes	yes
Constellation-Seq	DROP-Seq/3' 10x Chromium	high	Targeted priming on amplified cDNAs or first strands	yes	no	127	no	yes	no
scCapture-seq	Smart-seq, Fluidigm C1 and Smarter-seq	low	Probe hybridization (DNA probes)	yes	no	972	yes	yes	no
scFAST-seq	Custom	medium	Targeted amplification / probe hybridization (biotin)	yes	yes	2	yes	yes	yes
Method by Van Horebeek et al.	3' 10x Chromium	high	Targeted amplification	yes	yes	1	yes	yes	yes
RoCK and ROI	BD Rhapsody	high	Targeted capture / targeted priming	yes	yes	8	yes	yes	yes
scG2P	Custom	medium	Dual targeted PCR / probe hybridization (biotin)	yes	yes	56	no	yes	yes
10x Flex	3' 10x Chromium	high	Probe hybridization	yes	no	19,000	(no)*	yes	no
HybriSeq	Custom	medium	Probe hybridization	yes	no	95	no	yes	no

The throughput relates to the number of cells. Low throughput is in the range of a few hundreds of cells. Medium throughput is up to 10,000 cells, and high throughput is upwards of 10,000 cells. Methods shown in the table include CytoSeq¹⁵, tss2²⁹, Transcriptome resampling¹⁰⁶, DART-seq¹⁰⁴, GoT⁴⁶, BART-Seq¹⁰³, BD Rhapsody targeted amplification¹⁰⁰, Direct capture Perturb-seq⁸⁶, HyPR-seq¹⁰, Constellation-Seq⁸⁷, scCapture-seq¹⁰⁷, scFAST-seq¹⁰², Method by Van Horebeek et al.¹⁰², scG2P¹⁰⁵, RoCKseq / RoCK and ROI⁸⁵, 10x Flex¹⁰⁸, HybriSeq¹⁰⁸, 10 Flex protocol: although over 18'000 genes are targeted through probes, the transcriptome information depends on the choice and position of probes.

targeted and whole-transcriptome capture, its multiplexing capacity is limited. However, in both cases targeted capture was shown to increase the amount of information compared to samples using unmodified beads (430-fold increase for DART-seq and 379.7 increase for RoCKseq across different regions of interest). Direct capture Perturb-seq is based on the $10 \times 3'$ Chromium platform. Similar to DART-seq, it relies on addition of a partially double-stranded oligonucleotide. Since $10 \times$ beads are not solid, modification cannot be done in advance. Reagents are added directly to the reverse transcription mix, and the modification rate cannot be assessed prior to the scRNA-seq run. Direct capture Perturb-seq modifies both cs1 or cs2 oligonucleotides on $10 \times$ beads, leaving the dT oligonucleotides free for polyA capture. While the method was designed for guide RNAs in CRISPR/Cas9 perturbation assays, it can also be expanded to other targets.

A key advantage of targeted capture methods is that they address the earliest step of library generation, ensuring TOI incorporation into the sequencing library. Regions unfavorable for reverse transcription (e.g., GC-rich sequences) can be bypassed, and both polyadenylated and non-polyadenylated transcripts can be targeted. Limitations include the potential for off-target capture at sequence motifs resembling the designed capture sequence. This issue is exacerbated by the hybridization conditions of RNAs on barcoded beads, which occurs at low temperatures (for the $10 \times$ and BD Rhapsody platform room temperature with ice-cold reagents), below the melting temperatures of the capture sequences. Additionally, hybridization at TOIs may be inefficient as the targeted sequences may be buried inside RNA secondary structures.

Despite these caveats, targeted capture remains a valuable approach, offering efficient enrichment for TOIs while preserving comprehensive transcriptome coverage. As all of these methods are built on widely used high-throughput commercial platforms, they can be easily integrated into existing workflows with minimal adjustments. The optimal choice of targeted capture method ultimately depends on the available platform and experimental resources.

Category 2: Targeted priming

Targeted priming methods use primers which bind to either transcripts or cDNAs to enrich for the information on ROIs and/or TOIs, followed by reverse transcription or second strand synthesis, respectively. By defining the 5' ends of cDNAs and shifting the information to regions within TOIs, they mitigate the 3'/5' bias and require a decreased read length to reach ROIs.

Targeted priming methods are distinguishable based on the step in which the targeting occurs: first strand cDNAs for RoCK and ROI⁹⁵ and Constellation-Seq⁹⁷, reverse transcription for GoT⁹⁸, direct capture Perturb-seq⁹⁶ and tss2⁹⁹ and amplified cDNAs for Constellation-Seq. The Constellation-Seq authors applied their method to both $3' 10 \times$ Chromium and DROP-Seq technologies (Fig. 3). For DROP-Seq libraries, specific primers are added after first-strand cDNA synthesis, and the authors report an increase of UMI counts of 2.7 times compared to the untargeted condition. Targeted priming on $3' 10 \times$ Chromium libraries is directly applied to PCR-amplified cDNAs by using a transcript-specific primer. Compared to standard $10 \times$, the authors calculated a 22-fold greater sensitivity through targeting. The authors of direct capture Perturb-seq described TOI enrichment on both $3'$ -based $10 \times$ Chromium libraries (through targeted capture) and $5'/V(D)J 10 \times$ Chromium (through targeted priming). tss2⁹⁹ is the only out of the targeted priming methods which is not based on $3'/5'$ methods but on Smart-seq technology and thus associated with lower cellular throughput. However, this method is highly sensitive, with the authors increasing ROI detection from 25% to 100%.

Advantages of targeted priming include ease of implementation and flexibility, as the specific primers can be spiked into a pre-existing reaction. However, these methods strongly depend on the sequence of the transcript itself, which may not be favorable for binding of primers. Priming to other transcripts may also occur, leading to the detection of off-targets⁹⁷. Additionally, targeted priming at the level of reverse transcription cannot be applied when using barcoded beads and $3'$ -based library generation

protocols, as the reverse transcription reaction initiates directly on the beads to preserve critical barcode and UMI information. Applying targeted priming to already amplified cDNAs may introduce additional biases, as PCR amplification can already skew transcript representation. Consequently, targeted priming is most applicable at the level of reverse transcription (in the context of $5'$ or $V(D)J 10 \times$ Chromium workflows) or during first-strand cDNA synthesis.

Category 3: Targeted amplification

Targeted amplification methods use PCR to enrich for TOIs by combining a universal primer sequence with a transcript-specific primer, thereby retaining information on the cellular barcode and UMI of each TOI (Fig. 2a). The universal primer sequence is added to all cDNAs after reverse transcription and is used for PCR amplification; it is located adjacent to the barcode. The transcript-specific primer has the same adapter appended during fragmentation or random priming, allowing subsequent amplification of the cDNAs during standard library generation. Targeted amplification methods tackle three biases: 3'/5' bias, read length and sequencing depth. These methods can recover information on multiple TOIs, but the standard transcriptome may not always be simultaneously profiled. This is the case for CytoSeq¹⁵, which uses a set of primers to amplify transcript panels after first-strand synthesis on barcoded beads (Fig. 3). The same principle is used by the BD Rhapsody targeted amplification system¹⁰⁰ which scaled up the number of targets to several hundred. Other scRNA-seq methods using targeted amplification, which also profile the standard transcriptome, include GoT⁹⁸ and its derivative GoT-Splice¹⁰¹. As described above, in GoT a specific primer is added during reverse transcription while generating $5'/V(D)J 10 \times$ Chromium libraries. Libraries are then split into two: one part of the sample undergoes standard $10 \times$ Chromium indexing with PCR and fragmentation, while the second part (around 10%) undergoes a round of targeted amplification. Through targeted amplification and priming, the GoT-seq authors were able to show that for their mutation of interest in the *CALR* transcript, genotyping data were available for 88.7% of cells, compared to only 1.4% of cells in the untargeted condition. The same protocol (without targeted priming) is followed by GoT-Splice. Targeted amplification was also used by Van Horebeek et al.¹⁰², who were able to increase multiple targets from below 1% to over 80%. Similar to GoT, this method is based on using part of the $3' 10 \times$ Chromium output after first-strand synthesis to perform two rounds of targeted amplification, thus enriching for TOIs.

Targeted amplification is cost-effective, flexible, and compatible with existing cDNA libraries. Additionally, compared to targeted capture, annealing of primers occurs close to their melting temperature, with lower chances of off-target binding. However, it introduces PCR-related biases, and targeting efficiency depends on cDNA sequence and reaction conditions. Multiplexed detection can be uneven, as uniform PCR conditions may favor some targets over others. The BD Rhapsody targeted amplification authors in fact showed that 25% of the amplified transcripts actually had a lower detection level compared to the standard protocol, even after multiple PCR rounds¹⁰⁰. Importantly, if the initial RNA capture or reverse transcription was inefficient for a given transcript, further PCR cannot recover the information. Despite these limitations, targeted amplification remains a suitable approach for amplifying selected targets from pre-generated cDNA libraries, such as when retrieving additional information after sequencing. Nonetheless, this strategy is not recommended for large panels (e.g., >10 targets), as PCR conditions cannot be optimized for each primer set.

Category 4: Dual targeted PCR

Dual targeted PCR methods use transcript-specific forward and reverse primers for TOI amplification. In contrast to the previously described targeting categories, these methods do not retain information on the cell barcode and UMI but can instead introduce cellular indices attached to transcript-specific primers. This strategy is used by BART-Seq¹⁰³, a plate-based technology described for both bulk and scRNA-seq. The TARGET-seq¹⁰⁴ plate-based, multiomic method uses a combination of targeting

genomic DNA, as well as cDNA, for the detection of mutated genes and their expression in cancer cells. TARGET-seq is based on two rounds of targeted PCR, one directly on mRNAs while they are being reverse-transcribed and one on previously generated cDNAs. Similar to TARGET-seq, scG2P¹⁰⁵ is a plate-based method combining sequencing of genomic DNA with targeted PCR, which occurs after reverse transcription using primers harboring cell-specific barcodes. Finally, tss2, which as outlined above uses targeted priming at the reverse transcription level, further enhances the information on ROIs using targeted PCR by appending specific adapters used for subsequent PCR amplification.

As both targeted amplification and dual-targeted PCR methods are based on PCR amplification, they share common advantages and disadvantages. However, methods using dual-targeted PCR do not retain cellular barcode information and require the addition of cellular indices to the primers used for amplification. This implies that UMI-based correction for PCR amplification cannot be performed. Additionally, all dual-targeted PCR methods rely on custom set ups and, except for scG2P, are limited in the number of profiled cells and thus not compatible with high-throughput technologies. Given these factors, dual-targeted PCR methods currently appear less scalable and less broadly applicable than other targeting strategies.

Category 5: Probe hybridization

Targeted scRNA-seq approaches using probe hybridization to enhance TOI detection are the most diverse category of strategies. There are three main types of probe hybridization methods based on the type of enrichment: biotin-based enrichment, unmodified DNA probes or hairpin probe hybridization.

In approaches using biotinylated probes, the probes first hybridize to their target cDNAs, which are then selectively recovered using streptavidin-coated beads that bind to the biotin moiety (Fig. 3). This strategy is used by Transcriptome resampling¹⁰⁶. The goal of the study was to obtain information on a specific subset of cells in a dataset generated with 3' 10x Chromium chemistry. Biotinylated probes recognizing the respective cell barcodes were used to re-sequence them at a higher depth. The authors also used biotin probes as well as PCR primers conjugated with biotin to isolate TOIs from libraries. Through their probe hybridization strategy, the authors were able to increase the number of UMIs by up to 2.85-fold compared to the non-targeted cells. Additionally, the percentage of positive cells increased from 59.7 to 100% for a given TOI. Another scRNA-seq method using biotinylated probes for the detection of TOIs is scFAST-seq³³. It relies on either two rounds of targeted amplification, or enrichment of transcripts via biotinylated probes followed by PCR amplification.

The second type of probe hybridization method uses linear DNA probes without modification, as exemplified by scCapture-seq¹⁰⁷. Following hybridization to the target cDNAs, the resulting double-stranded DNA complexes are isolated and subsequently PCR-amplified to increase their abundance. Through this technology, the authors showed an increase in reads mapped to transcription factors (their TOIs) from 2.2 to 78.3% (36-fold), with target gene expression increasing by 150-fold on average. In contrast to scCapture-seq, HybriSeq¹⁰⁸ is a plate-based method using DNA probes which recognize two adjacent regions on RNAs. The bound probes are then ligated, followed by split-pool barcoding to tag each cell with a specific barcode, PCR amplification and sequencing. A similar strategy is used by the 10x Flex protocol¹⁰⁹, in which probes also carrying a specific index sequence and adapters, are hybridized to mRNAs of interest. Compared to the HybriSeq protocol, barcodes are added through 10x beads, as the probes contain an adapter which is complementary to a sequence on the barcoded beads and can thus be subsequently captured. As of fall 2025, the 10x Flex probe library includes probes against more than 18,000 genes for mouse and 19,000 for human. While this is still a targeted method, the resolution is extremely high and close to the information obtained in the standard transcriptome. Additionally, by having multiple probes per gene this technology is less dependent on efficiency of probe design (which might for example be impacted by mutations in transcripts). Both the 10x Flex

protocol and HybriSeq can be applied to fixed and permeabilized tissues, where RNA quality is often lower compared to fresh tissue samples.

The third class of probe hybridization methods employs hairpin probes, as in HyPR-seq¹¹⁰. This approach targets RNAs directly using two initiator probes that hybridize to adjacent sequences on the transcript of interest (TOI). After addition of further DNA oligonucleotides (plate-based), including a hairpin probe, cDNA constructs containing the sequence of the targeted RNAs, adapters, and UMIs are generated. The final cDNA constructs are then annealed to barcoded microbeads, thus retaining the information at the single-cell level. By comparing their method to existing 10x Chromium data, the authors calculated that using 10x they would need over 1,000,000 cells at 20,000 reads per cell to detect 25% of expression changes in 90% of genes expressed above 1 transcript per million. In contrast, their method achieved the same power with only about 25,000 cells at 5000 reads per cell.

Similar to targeted amplification and dual-targeted PCR, one of the main advantages of probe hybridization methods acting on cDNA is their flexibility and compatibility with different platforms. A key challenge across all probe hybridization methods lies in designing probes with sufficient target specificity, as sequence variations in RNA or cDNA can impair probe binding. For methods applied to cDNA, enrichment of TOIs is further constrained by earlier steps in library preparation and overall library complexity.

In contrast, probe hybridization methods targeting RNA act at the earliest stage of the scRNA-seq workflow. Historically, these approaches were labor-intensive, required custom setups, and thus have not been widely adopted. The 10x Flex protocol represents a major exception, enabling scalable, high-throughput targeting of thousands of transcripts on the 10x Chromium platform. However, it is not suited for ROI detection, as it sequences ligated probes rather than the RNAs themselves. Moreover, probes are currently available only for human and mouse samples. Because tens of thousands of genes are targeted simultaneously, this approach still entails some information loss for specific TOIs. Finally, non-templated ligation events may occur, contributing to background noise¹¹¹. Despite these limitations, the 10x Flex protocol introduced a breakthrough in the use of fixed and permeabilized tissues for scRNA-seq experiments.

While targeted methods tackle most biases described in the section above, none of them comprehensively addresses PCR biases (Fig. 2a). This would require changing the PCR conditions to specifically fit TOIs, which is not possible when multiple transcripts with different properties are amplified in the same reaction and under the same conditions. Such an approach however is a valuable solution and intrinsically very sensitive for single targets of interest.

Targeted sequencing combined with long-read technologies

The length of a single ROI which can be detected with the methods described above is limited to ~600 bp or restricted to a low number of specific, but short transcript regions. For example, studying novel isoforms in a long transcript using short-read targeted scRNA-seq methods requires targeted capture or priming using several primers or PCR reactions in order to recover information from multiple ROIs⁹⁴. To overcome these limitations, long-read sequencing has been recently combined with 3'/5'-based methods, obtaining full-length transcript profiles^{24,112–116}. Information obtained through long-read sequencing and 3'/5'-based methods is connected through detection of the barcode sequence.

Combining long-read sequencing technologies with 3'/5'-end-based short-read sequencing platforms presents two major challenges (Fig. 4). Several biases limit recovery of information from the full transcript length (internal capture/priming and dissociation of the reverse transcriptase enzyme). Additionally, similar to short-read platforms and as long-read sequencing experiments typically yield lower read depth than short-read sequencing¹¹⁷, TOIs need to be enriched for in the library prior to sequencing. Both challenges have been addressed using targeted enrichment

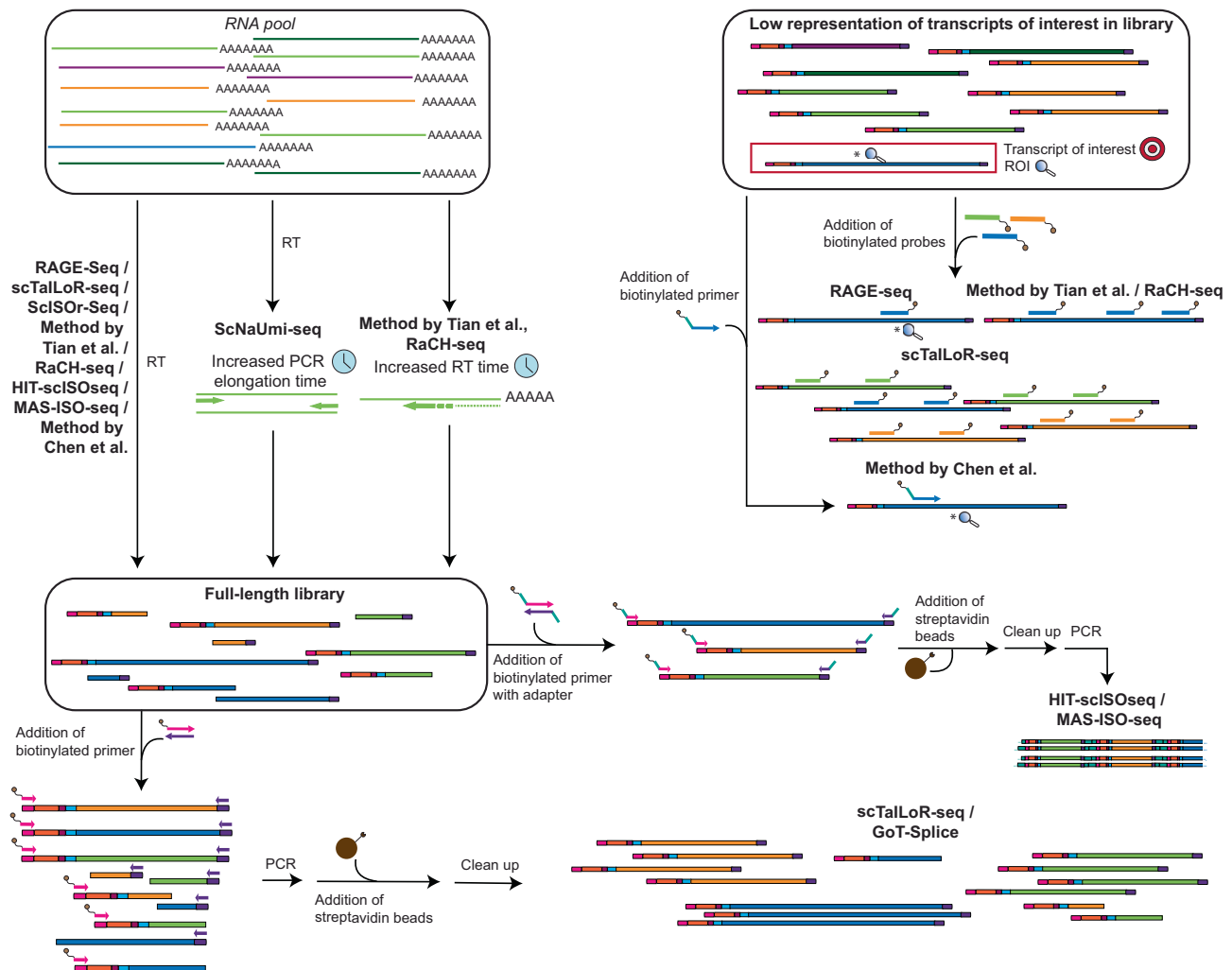


Fig. 4 | Summary of technologies coupling long-read sequencing with 3'/5' based library generation. The full-length library also shows fragmented transcripts which were generated by biases during early library generation steps, including reverse

transcription. Methods shown in the figure include ScISOR-Seq¹¹⁹, RAGE-Seq¹¹⁸, ScNaUmi-seq¹²⁰, Method by Chen et al.¹²⁶, RaCH-seq¹²², MAS-ISO-seq¹²⁴, HIT-sciISOseq¹²³, GoT-Splice¹⁰¹, scTalLoR-seq¹²⁵, Method by Tian et al.¹²¹.

strategies. An overview of the key features and implementation stages of the methods discussed below is provided in Fig. 4.

Application of targeting strategies to long-read protocols

Early long-read sequencing methods adapted to 3'/5'-based technologies lacked effective strategies for enriching full-length cDNAs with barcode sequences. Examples include RAGE-Seq¹¹⁸ and ScISOR-Seq¹¹⁹. In RAGE-Seq, only 18.7% of reads contained barcode information, while in ScISOR-Seq, 58% of reads carried barcodes, out of 61.6% that exhibited a detectable oligo(dT) tail marking the 3' end of the cDNA.

To improve the recovery of full-length, barcode-containing transcripts, several protocol modifications were introduced. ScNaUmi-seq¹²⁰ increased PCR elongation time from 1 to 3 min, while Tian et al.¹²¹ and RaCH-seq¹²² extended the reverse transcription step to 2 h. Other approaches, including HIT-sciISOseq¹²³, MAS-ISO-seq¹²⁴, and GoT-Splice¹⁰¹, used targeted priming with a biotinylated universal primer to selectively enrich cDNA fragments that contain both the cell barcode and UMI. Similarly, scTalLoR-seq¹²⁵ employed biotinylated PCR primers to enhance the recovery of barcode-containing sequences.

Beyond barcode enrichment, several long-read methods have leveraged biotinylated probes or primers to target TOIs (Fig. 4). RAGE-Seq uses biotinylated capture probes directed at specific transcript

regions, while ScISOR-Seq and RaCH-seq extend this strategy across the full cDNA length. Chen et al.¹²⁶ implement targeted priming using a biotinylated primer positioned adjacent to a fusion breakpoint of interest in 10× Chromium libraries. In scTalLoR-seq, biotinylated probes are used to enrich for over 1000 specific genes, boosting TOI recovery from 5% to 95%. Notably, these enrichment strategies do not require precise targeting of ROIs, as long-read sequencing enables recovery of internal transcript regions.

Despite these advancements, long-read protocols adapted to 3'/5'-based technologies require additional PCR amplification to increase input material, potentially introducing further bias. Additionally, these methods are all built upon libraries generated through oligo(dT)-based capture or priming, template switching, and reverse transcription. As detailed in Table 1, these three steps introduce sequence-dependent biases, including the generation of fragmented constructs, which may affect transcript coverage and interpretation.

Targeted methods in spatial biology

Targeted technologies are essential in the field of spatial biology, which has emerged as a standard technology to obtain information on RNAs expressed by cells in combination with their spatial position. The field has been rapidly evolving from multi-cell to single-cell resolution and to broad transcript

coverage¹²⁷. Methods can be largely classified into two categories: imaging-based and sequencing-based^{128,129} technologies.

Image-based technologies

Image-based spatial transcriptomic technologies ultimately rely on microscopy-based readouts to obtain spatial information on RNAs. Two types of probe-based, targeted approaches are prevalent in the field (reviewed in Tian et al.¹²⁷). Early methods, such as seqFISH¹³⁰ and MERFISH¹³¹ were based on in situ fluorescent hybridization using FISH probes complementary to RNAs of interest¹³². In the second strategy, RNA molecules or in-situ reverse transcribed cDNAs are hybridized with specific probes which are then amplified using rolling circle amplification and visualized with fluorescent probes¹²⁷. As both types of image-based technologies use targeted probe hybridization methods, they suffer from the same biases described above, in particular probe design.

Sequencing-based technologies

Many sequencing-based spatial transcriptomic platforms, such as the 10× Visium technology¹³³, rely on capture of RNAs on spatially indexed surfaces covered with DNA oligonucleotides with a barcode specific to a given location on the slide¹²⁷. Alternative solutions to introduce a spatially-resolved barcode are Slide-seq¹³⁴ and Stereo-seq¹³⁵. As these methods rely on dT-based capture, information is restricted to the UTRs of transcripts. After RNA capture, the next steps follow similar workflows as the 3'/5'-based methods, with reverse transcription, PCR amplification and sequencing. Targeted amplification and probe hybridization methods can be applied to these libraries to enrich for TOI information, while retaining the spatially-resolved information. This was for example implemented by McKellar et al.¹³⁶, who used biotinylated probes to enrich for viral transcripts. Similarly, Engblom et al.¹³⁷, used biotinylated probes to enrich for *BCR-TCR* transcripts, retaining spatial information. B and T cell repertoires were also enriched for by Sudmeier et al.¹³⁸, who used targeted amplification with a transcript-specific reverse primer. A second strategy used in sequencing-based spatial transcriptomic methods is probe hybridization to RNAs, followed by sequencing of the bound probes¹³⁹. This strategy is used by the 10× Xenium technology, which is also the basis for the 10× Flex protocol. In this case, the information is not limited to the UTRs of transcripts, but internal ROIs cannot be accessed as the probes themselves are sequenced and not the RNAs themselves.

These developments illustrate how targeted technologies in scRNA-seq and spatial transcriptomics have co-evolved, with targeting strategies transferable between both domains.

Biological applications of targeted scRNA-seq methods

Targeted scRNA-seq methods have been applied across a wide range of contexts, from detecting non-polyadenylated transcripts to studying cancer mutations, alternative splicing and capturing information on guide RNAs during pooling screens.

Non-polyadenylated transcripts

As viral and prokaryotic RNAs lack poly(A) tails, they are often invisible in conventional 3'/5'-based scRNA-seq. Targeted strategies resolve this gap. For example, DART-seq enriched reads from reovirus-infected cells by 430-fold compared to standard DROP-seq technologies by directing sequencing towards an internal viral region, enabling mutation-level resolution of viral heterogeneity. Focusing on enhancing detection of bacterial transcripts, INVADE-seq¹⁴⁰ combined targeted priming and PCR to amplify bacterial 16S rRNA alongside host cell transcriptomes in patient tumor samples. This issue is particularly relevant as bacterial cells are estimated to contain more than 100 times less RNA compared to eukaryotic cells, and transcriptional turnover is much faster compared to eukaryotes¹⁴¹. The dual profiling revealed that bacteria in the tumor microenvironment are primarily localized within myeloid cells and can modulate host immune responses, with potential implications for immunotherapy outcomes.

Cancer mutations and fusion events

Another major application of targeted scRNA-seq technologies is cancer research, where mutations often reside in internal transcript regions missed by 3'/5'-based methods. Targeted methods are thus fundamental to not only focus on these mutations, but also to answer relevant questions such as which cell types harbor the mutation and the changes in the cellular transcriptome they introduce. GoT-seq, for instance, enabled mutation detection in *CALR* and *XBPI*, linking them to transcriptional changes in specific cell types and highlighting IRE1 signaling as a potential therapeutic target. The focus of the study conducted by Van Horeebek et al. and scFAST-seq was also the detection of disease-causing mutations and their assignment to specific cell types. Of particular interest in many cancer types are fusion breakpoints, as these are often disease-causing mutations. These include the breakpoint of the *BCR::ABL1* gene, which is found at more than 3 kb away from the 3' end of the transcript. Given its relevance in the context of leukemia, several targeted methods focused on directing their information to this ROI. These methods include tss2 and other methods such as scFAST-seq, RoCK and ROI and Nilsson et al.¹⁴², who adapted a version of GoT-seq for the detection of *BCR::ABL1*.

Alternative splicing

Targeted methods also make it possible to study alternative splicing with single-cell resolution. These events exclusively occur within internal transcript regions and thus cannot be analyzed with standard 3'/5'-based technologies. scTallOR-seq, for example, examined over 2200 targeted transcripts to characterize isoform diversity in ovarian cancer, revealing differential isoform usage between cell types. As shown by RoCK and ROI, alternative splicing can also be analyzed with short-read based methods. The authors of this method directed reads to exon-exon junctions of a selected transcript (*Pdgfra*), providing a framework for studying splicing variation in defined cell types across TOI.

Rare cell types and markers genes

Targeted enrichment can also improve detection of rare cell populations and lowly expressed genes, which often serve as cell-type-specific markers. Transcriptome Resampling recovered more information from megakaryocytes, which represent only ~0.1% of bone marrow samples, by resequencing reads from selected barcodes. Constellation-seq increased marker recovery in PBMCs, expanding dendritic cell representation from 51 to 127 cells. Likewise, scCapture-seq enriched transcription factor transcripts, which are notoriously under-detected in scRNA-seq, enabling finer resolution of neuronal subpopulations and increasing the number of differentially expressed genes distinguishing annotated clusters from 129 to 155.

Capture of guide RNAs

Capturing guide RNAs is essential in pooled CRISPR screens, as it enables linking cellular transcriptome profiles to the corresponding perturbations. Perturb-seq¹⁴³ is a widely used approach in which guide constructs are engineered with a poly(A) tail, allowing capture on barcoded beads. Guide recovery is then enhanced through targeted amplification, which associates each perturbation with transcriptomic changes. To increase sensitivity for genes of interest that may be affected by perturbations, TAP-seq¹⁴⁴ introduced targeted amplification of both guide RNAs and selected transcripts. Targeted capture has also been applied specifically to enrich guide RNAs, as in direct-capture Perturb-seq⁹⁶ and CAT-ATAC¹⁴⁵. This last method is a multiomic technology integrating information on guide RNAs, open chromatin and RNA. Another example of guide capture in multiomic assays is ECCITE-seq¹⁴⁶, which combines 5' scRNA-seq, protein detection, and targeted guide RNA capture using both reverse-transcription priming and targeted amplification.

Together, these examples highlight the versatility of targeted scRNA-seq methods. By enriching for TOIs and ROIs, they not only overcome blind spots of standard scRNA-seq but also allow integration with information on the standard transcriptome. This combination enables researchers to link specific mutations, isoforms, or microbial signals to broader transcriptional changes, expanding the scope of biological questions addressable at single-cell resolution.



Fig. 5 | Decision tree for the choice of targeted scRNA-seq method. Red box indicates start point for decision. Boxes and arrows in grey indicate methods which can be used for targeted sequencing but which we do not recommend due to the outlined reasons. The cutoff of 10 TOIs / ROIs was chosen somewhat arbitrarily, as the maximum number of transcripts which can be targeted depends also on the properties of the transcripts themselves. Standard scRNA-seq methods are also included, as they may suffice for the detection of TOIs and ROIs. Methods are colored based on the categories of scRNA-seq methods outlined above. Methods

shown in the figure include CytoSeq¹⁵, tss⁹⁹, Transcriptome resampling¹⁰⁶, DART-seq⁹⁴, TARGET-seq¹⁰⁴, GoT⁹⁸, BART-Seq¹⁰³, BD Rhapsody targeted amplification¹⁰⁰, Direct capture Perturb-seq⁹⁶, ScNaUmi-seq¹²⁰, HyPR-seq¹¹⁰, Constellation-Seq⁹⁷, scCapture-seq¹⁰⁷, scFAST-seq³³, Method by Van Horebeek et al.¹⁰², scG2P¹⁰⁵, RoCKseq / RoCK and ROI⁹⁵, 10x Flex¹⁰⁹, HybriSeq¹⁰⁸, ScISOR-Seq¹¹⁹, RAGE-Seq¹¹⁸, Method by Chen et al.¹²⁶, RaCH-seq¹²², MAS-ISO-seq¹²⁴, HIT-scISOseq¹²³, GoT-Splice¹⁰¹, scTaILoR-seq¹²⁵, Method by Tian et al.¹²¹.

While many targeted scRNA-seq methods using short-read sequencing have been described, we only included a selection based on the considerations described above. Methods requiring custom setups were excluded, as they tend to be more labor-intensive and costly. Instead, we focused on approaches that can be implemented using widely available commercial platforms.

The first node in the decision tree is whether to generate a novel library or apply targeted solutions to previously synthesized cDNAs. Wherever possible, we recommend generating a new library, as TOIs may not be captured in the existing ones. The second major node in the decision tree is the concomitant detection of the standard transcriptome, which is essential if the research question relies on the identification of cell types. Key considerations then include the number of TOIs/ROIs, the technologies available in-house, as well as the actual sequences of the targets. Because the choice of targeted method also depends on the technologies available in-house, nodes may have multiple recommended methods. An example is targeted capture methods, which are based on different commercially available technologies.

Depending on the number of targets, the 10x Flex protocol, RoCKseq and DART-seq remain the methods of choice for TOI detection as they do not introduce additional amplification and target the first step in the scRNA-seq protocol. Despite inherent PCR biases, targeted amplification remains the only solution to recover information on a high number of ROIs (>10), as multiplexing of targeted capture methods has not been extensively shown or validated.

Targeting methods described for a given platform may also be applied to a similar technology, and the methods themselves can be combined (e.g., TOI targeting for long-read sequencing technologies can be extended to most (if not all) protocols, same as technologies acting on final cDNA libraries, while targeted capture and priming can be used in the same experiment). Additionally, the methods described in Box 1, in particular those increasing information on cells of interest during sample preparation and depleting highly abundant RNAs, can also be applied in combination with targeted technologies to further enhance the detection of TOIs and ROIs.

Importantly, while we propose a set of methods based on the reasoning outlined in this review, a systematic, side-by-side comparison of all targeted scRNA-seq methods would be highly valuable for the field. Ideally, this study should be conducted on the same biological sample, preferably on the same day, to minimize batch effects. All methods should target the same TOI(s)/ROI(s), and protocols should be carried out by experts to reduce technical biases. However, such a study would be challenging, as different targeted methods are often optimized for distinct subsets of TOI(s)/ROI(s) and comparing them across divergent targets or scenarios would compromise the benchmarking objective. Our decision tree thus serves as a valuable tool to guide researchers toward both the most appropriate targeting strategy and the specific method best aligned with their research question.

Perspectives

Currently, TOI and ROI detection in scRNA-seq experiments is limited by the fact that only a certain portion of the transcriptome is detected with current technologies. While improvements in read depth per cell (facilitated by decreased cost of sequencing) aid with the detection of TOIs, the resulting data from many such experiments is still biased towards the 3'/5' end of transcripts. All scRNA-seq approaches still suffer from biases introduced during protocol steps such as reverse transcription and PCR. These

technologies have been used for over 50 years and have already undergone multiple rounds of optimization, and it is thus unlikely that they will change. A promising solution to potentially overcome these biases is DRS, which allows bypassing the need for reverse transcription and PCR amplification. This topic has been addressed by a previous call for direct sequencing of full-length RNAs with the aim to determine modifications on RNA bases¹⁴⁷, but to our knowledge current methods have not been applied yet to single cells and are still associated with low throughput, high cost and challenging set ups. While further innovation in this area is expected, progress is likely to be gradual, as these approaches are currently largely conceptual.

While multiple targeting strategies have been developed, they have not been systematically assessed or compared experimentally. This benchmarking study, outlined above, would be highly relevant for the field, further crystallizing the optimal methods for TOI and ROI enhancement. The observation that targeted capture of RNAs can lead to an increase in the detection of TOIs is highly promising for future development of targeted scRNA-seq methods. Capture/priming of transcripts is the first part of scRNA-seq protocols, and biases occurring at this step are exacerbated during the downstream protocol. While this step has received relatively little attention, future improvements in scRNA-seq performance may benefit from optimizing targeted capture and priming conditions. Given the relaxed binding conditions of RNAs on barcoded beads, these solutions should focus on enhancing the specificity of capture, thus also reducing internal priming artifacts. Potential solutions could be using LNA oligonucleotides, or addition of chemicals such as formamide to reduce the melting temperature. In parallel, probe-based approaches targeting the same step, such as 10x Flex, have shown great promise. Although these methods are already highly efficient, further refinement of probe design could enhance target specificity and performance.

Another promising direction is long-read sequencing, which has seen rapid growth in the past years through the development of methods combining 3'/5'-based technologies with long-read sequencing to profile the cellular transcriptome. With the decrease in cost of long-read sequencing platforms, increased read throughput and lower error rate, the need for 3'/5'-based methods may be fully circumvented. These improvements would allow full transcripts, including ROIs, to be sequenced and the standard cellular transcriptome, including TOIs, to be profiled at the same time in single cells. Although we believe major technological advances will derive from the field of long-read sequencing, currently targeted scRNA-seq methods remain indispensable for detection of TOIs and ROIs.

Taken together, systematic improvements in targeted methods at the RNA level including targeted capture and probe hybridization, adaptation of DRS, and integration with long-read platforms represent the most promising paths to comprehensively profile TOIs and ROIs at single-cell resolution.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Received: 23 October 2025; Accepted: 29 January 2026;

Published online: 14 February 2026

References

1. Pan, Y., Cao, W., Mu, Y. & Zhu, Q. Microfluidics facilitates the development of single-cell RNA sequencing. *Biosensors* **12**, 450 (2022).

2. Shen, X., Zhao, Y., Wang, Z. & Shi, Q. Recent advances in high-throughput single-cell transcriptomics and spatial transcriptomics. *Lab Chip* **22**, 4774–4791 (2022).
3. Baysoy, A., Bai, Z., Satija, R. & Fan, R. The technological landscape and applications of single-cell multi-omics. *Nat. Rev. Mol. Cell Biol.* **24**, 695–713 (2023).
4. Vandereyken, K., Sifrim, A., Thienpont, B. & Voet, T. Methods and applications for single-cell and spatial multi-omics. *Nat. Rev. Genet.* **24**, 494–515 (2023).
5. Conte, M. I., Fuentes-Trillo, A. & Domínguez Conde, C. Opportunities and tradeoffs in single-cell transcriptomic technologies. *Trends Genet.* **40**, 83–93 (2024). **This review provides an overview of recent advances in the single-cell sequencing field.**
6. De Jonghe, J. et al. scTrends: A living review of commercial single-cell and spatial 'omic technologies. *Cell Genomics* **4**, 100723 (2024).
7. Islam, S. et al. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res.* **21**, 1160–1167 (2011).
8. Hashimshony, T., Wagner, F., Sher, N. & Yanai, I. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep.* **2**, 666–673 (2012).
9. Islam, S. et al. Highly multiplexed and strand-specific single-cell RNA 5' end sequencing. *Nat. Protoc.* **7**, 813–828 (2012).
10. Shalek, A. K. et al. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* **510**, 363–369 (2014).
11. Svensson, V. et al. Power analysis of single-cell RNA-sequencing experiments. *Nat. Methods* **14**, 381–387 (2017).
12. Zheng, G. X. Y. et al. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017).
13. Macosko, E. Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
14. Klein, A. M. et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187–1201 (2015).
15. Fan, H. C., Fu, G. K. & Fodor, S. P. A. Combinatorial labeling of single cells for gene expression cytometry. *Science* **347**, 1258367 (2015).
16. Adiconis, X. et al. Comprehensive comparative analysis of 5'-end RNA-sequencing methods. *Nat. Methods* **15**, 505–511 (2018).
17. Ramsköld, D. et al. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat. Biotechnol.* **30**, 777–782 (2012).
18. Picelli, S. et al. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10**, 1096–1098 (2013).
19. Hagemann-Jensen, M. et al. Single-cell RNA counting at allele and isoform resolution using Smart-seq3. *Nat. Biotechnol.* **38**, 708–714 (2020).
20. Hagemann-Jensen, M., Ziegenhain, C. & Sandberg, R. Scalable single-cell RNA sequencing from full transcripts with Smart-seq3xpress. *Nat. Biotechnol.* **40**, 1452–1457 (2022).
21. Hahaut, V. et al. Fast and highly sensitive full-length single-cell RNA sequencing using FLASH-seq. *Nat. Biotechnol.* **40**, 1447–1451 (2022).
22. Sheng, K., Cao, W., Niu, Y., Deng, Q. & Zong, C. Effective detection of variation in single-cell transcriptomes using MATQ-seq. *Nat. Methods* **14**, 267–270 (2017).
23. Salmen, F. et al. High-throughput total RNA sequencing in single cells using VASA-seq. *Nat. Biotechnol.* **40**, 1780–1793 (2022).
24. Byrne, A., Cole, C., Volden, R. & Vollmers, C. Realizing the potential of full-length transcriptome sequencing. *Philos. Trans. R. Soc. B* **374**, 20190097 (2019).
25. Rhoads, A. & Au, K. F. PacBio sequencing and its applications. *Genomics Proteom. Bioinform.* **13**, 278–289 (2015).
26. Amarasinghe, S. L. et al. Opportunities and challenges in long-read sequencing data analysis. *Genome Biol.* **21**, 30 (2020).
27. Wang, Y., Zhao, Y., Bollas, A., Wang, Y. & Au, K. F. Nanopore sequencing technology, bioinformatics and applications. *Nat. Biotechnol.* **39**, 1348–1365 (2021).
28. Shi, H. et al. Bias in RNA-seq library preparation: current challenges and solutions. *BioMed. Res. Int.* **2021**, 6647597 (2021). **This review provides an overview of biases in bulk RNA-seq.**
29. Machado, L., Relaix, F. & Mourikis, P. Stress relief: emerging methods to mitigate dissociation-induced artefacts. *Trends Cell Biol.* **31**, 888–897 (2021).
30. Denisenko, E. et al. Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus RNA-seq workflows. *Genome Biol.* **21**, 130 (2020).
31. Adam, M., Potter, A. S. & Potter, S. S. Psychrophilic proteases dramatically reduce single cell RNA-seq artifacts: a molecular atlas of kidney development. *Development* dev.151142 <https://doi.org/10.1242/dev.151142> (2017).
32. Palazzo, A. F. & Lee, E. S. Non-coding RNA: what is functional and what is junk? *Front. Genet.* **6**, 2 (2015).
33. Sang, G. et al. High throughput detection of variation in single-cell whole transcriptome through streamlined scFAST-seq. Preprint at *bioRxiv* <https://doi.org/10.1101/2023.03.19.533382> (2023).
34. Fan, X. et al. Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos. *Genome Biol.* **16**, 148 (2015).
35. Verboom, K. et al. SMARTer single cell total RNA sequencing. *Nucleic Acids Res.* **47**, e93–e93 (2019).
36. Nam, D. K. et al. Oligo(dT) primer generates a high frequency of truncated cDNAs through internal poly(A) priming during reverse transcription. *Proc. Natl. Acad. Sci. USA* **99**, 6152–6156 (2002).
37. Svoboda, M., Frost, H. R. & Bosco, G. Internal oligo(dT) priming introduces systematic bias in bulk and single-cell RNA sequencing count data. *NAR Genomics Bioinform.* **4**, lqac035 (2022). **This article analyzes the effects of oligo(dT) priming in bulk and single-cell RNA sequencing.**
38. Wang, K.-T. & Adler, C. E. CRISPR/Cas9-based depletion of 16S ribosomal RNA improves library complexity of single-cell RNA-sequencing in planarians. *BMC Genomics* **24**, 625 (2023).
39. What fraction of reads map to ribosomal proteins? *10X Genomics* <https://kb.10xgenomics.com/hc/en-us/articles/218169723-What-fraction-of-reads-map-to-ribosomal-proteins>.
40. Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. A. & Luscombe, N. M. A census of human transcription factors: function, expression and evolution. *Nat. Rev. Genet.* **10**, 252–263 (2009).
41. Lambert, S. A. et al. The human transcription factors. *Cell* **172**, 650–665 (2018).
42. Wheat, J. C. et al. Single-molecule imaging of transcription dynamics in somatic stem cells. *Nature* **583**, 431–436 (2020).
43. Boone, M., De Koker, A. & Callewaert, N. Capturing the 'ome': the expanding molecular toolbox for RNA and DNA library construction. *Nucleic Acids Res.* **46**, 2701–2721 (2018).
44. Verwilt, J., Mestdagh, P. & Vandesompele, J. Artifacts and biases of the reverse transcription reaction in RNA sequencing. *RNA* **29**, 889–897 (2023). **This review provides an overview of biases introduced at the reverse transcription step.**
45. Brooks, E. M., Sheflin, L. G. & Spaulding, S. W. Secondary structure in the 3' UTR of EGF and the choice of reverse transcriptases affect the detection of message diversity by RT-PCR. *Biotechniques* **19**, 814–815 (1995).
46. Moqtaderi, Z., Geisberg, J. & Struhl, K. Secondary structures involving the poly(A) tail and other 3' sequences are major determinants of mRNA isoform stability in yeast. *MIC* **1**, 137–139 (2014). **Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA.**

47. Bustin, S. Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *J. Mol. Endocrinol.* **29**, 23–39 (2002).
48. Minshall, N. & Git, A. Enzyme- and gene-specific biases in reverse transcription of RNA raise concerns for evaluating gene expression. *Sci. Rep.* **10**, 8151 (2020).
49. Gout, J.-F., Thomas, W. K., Smith, Z., Okamoto, K. & Lynch, M. Large-scale detection of in vivo transcription errors. *Proc. Natl. Acad. Sci. USA* **110**, 18584–18589 (2013).
50. Yasukawa, K. et al. Next-generation sequencing-based analysis of reverse transcriptase fidelity. *Biochem. Biophys. Res. Commun.* **492**, 147–153 (2017).
51. Schwaber, J., Andersen, S. & Nielsen, L. Shedding light: The importance of reverse transcription efficiency standards in data interpretation. *Biomol. Detect. Quant.* **17**, 100077 (2019).
52. Jia, E. et al. Optimization of library preparation based on SMART for ultralow RNA-seq in mice brain tissues. *BMC Genomics* **22**, 809 (2021).
53. Ståhlberg, A., Kubista, M. & Pfaffl, M. Comparison of reverse transcriptases in gene expression analysis. *Clin. Chem.* **50**, 1678–1680 (2004).
54. Štovićek, A., Cohen-Chalamish, S. & Gillor, O. The effect of reverse transcription enzymes and conditions on high throughput amplicon sequencing of the 16S rRNA. *PeerJ* **7**, e7608 (2019).
55. Ozsolak, F. et al. Direct RNA sequencing. *Nature* **461**, 814–818 (2009).
56. Ozsolak, F. & Milos, P. M. Single-molecule direct RNA sequencing without cDNA synthesis. *WIREs RNA* **2**, 565–570 (2011).
57. Garalde, D. R. et al. Highly parallel direct RNA sequencing on an array of nanopores. *Nat. Methods* **15**, 201–206 (2018).
58. Hoque, M. et al. Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing. *Nat. Methods* **10**, 133–139 (2013).
59. Zajac, P., Islam, S., Hochgerner, H., Lönnerberg, P. & Linnarsson, S. Base preferences in non-templated nucleotide incorporation by MMLV-derived reverse transcriptases. *PLoS ONE* **8**, e85270 (2013).
60. Cocquet, J., Chong, A., Zhang, G. & Veitia, R. A. Reverse transcriptase template switching and false alternative transcripts. *Genomics* **88**, 127–131 (2006).
61. Mader, R. M. et al. Reverse transcriptase template switching during reverse transcriptase–polymerase chain reaction: artificial generation of deletions in ribonucleotide reductase mRNA. *J. Lab. Clin. Med.* **137**, 422–428 (2001).
62. Schulz, L. et al. Direct long-read RNA sequencing identifies a subset of questionable exons likely arising from reverse transcription artifacts. *Genome Biol.* **22**, 190 (2021).
63. Carninci, P. High efficiency selection of full-length cDNA by improved biotinylated cap trapper. *DNA Res.* **4**, 61–66 (1997).
64. Matz, M. Amplification of cDNA ends based on template-switching effect and step- out PCR. *Nucleic Acids Res.* **27**, 1558–1560 (1999).
65. Zhu, Y. Y., Machleder, E. M., Chenchik, A., Li, R. & Siebert, P. D. Reverse transcriptase template switching: a SMART™ approach for full-length cDNA library construction. *BioTechniques* **30**, 892–897 (2001).
66. Hughes, T. K. et al. Second-strand synthesis-based massively parallel scRNA-Seq reveals cellular states and molecular features of human inflammatory skin pathologies. *Immunity* **53**, 878–894.e7 (2020).
67. Quail, M. A., Swerdlow, H. & Turner, D. J. Improved protocols for the illumina genome analyzer sequencing system. *CP Human Genet.* **62**, 18.2.1–18.2.27 (2009).
68. Tan, G., Opitz, L., Schlapbach, R. & Rehrauer, H. Long fragments achieve lower base quality in Illumina paired-end sequencing. *Sci. Rep.* **9**, 2856 (2019).
69. Modi, A., Vai, S., Caramelli, D. & Lari, M. The Illumina Sequencing Protocol and the NovaSeq 6000 System. in *Bacterial Pangenomics* (eds Mengoni, A., Bacci, G. & Fondi, M.) Vol. 2242 15–42 (Springer US, New York, NY, 2021).
70. Adey, A. C. Tagmentation-based single-cell genomics. *Genome Res.* **31**, 1693–1705 (2021).
71. Lodge, J. K., Weston-Hafer, K. & Berg, D. E. Transposon Tn5 target specificity: preference for insertion at G/C pairs. *Genetics* **120**, 645–650 (1988).
72. Green, B., Bouchier, C., Fairhead, C., Craig, N. L. & Cormack, B. P. Insertion site preference of Mu, Tn5, and Tn7 transposons. *Mob. DNA* **3**, 3 (2012).
73. Adey, A. et al. Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome Biol.* **11**, R119 (2010).
74. Parkinson, N. J. et al. Preparation of high-quality next-generation sequencing libraries from picogram quantities of target DNA. *Genome Res.* **22**, 125–133 (2012).
75. Kia, A. et al. Improved genome sequencing using an engineered transposase. *BMC Biotechnol.* **17**, 6 (2017).
76. Van Gorp, T. P., McIntyre, L. M. & Verhoeven, K. J. F. Consistent errors in first strand cDNA due to random hexamer mispriming. *PLoS ONE* **8**, e85583 (2013).
77. Mignone, F., Gissi, C., Liuni, S. & Pesole, G. Untranslated regions of mRNAs. *Genome Biol.* **3**, reviews0004 (2002).
78. Hong, D. & Jeong, S. 3'UTR diversity: expanding repertoire of RNA alterations in human mRNAs. *Mol. Cells* **46**, 48–56 (2023).
79. Quail, M. A., Corton, C., Uphill, J., Keane, J. & Gu, Y. Identifying the best PCR enzyme for library amplification in NGS. *Microbial. Genomics* **10**, 001228 (2024).
80. Benjamini, Y. & Speed, T. P. Summarizing and correcting the GC content bias in high-throughput sequencing. *Nucleic Acids Res.* **40**, e72–e72 (2012).
81. Aird, D. et al. Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol.* **12**, R18 (2011).
82. Dohm, J. C., Lottaz, C., Borodina, T. & Himmelbauer, H. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Res.* **36**, e105 (2008).
83. Dabney, J. & Meyer, M. Length and GC-biases during sequencing library amplification: a comparison of various polymerase-buffer systems with ancient and modern DNA sequencing libraries. *BioTechniques* **52**, 87–94 (2012).
84. Arezi, B., Xing, W., Sorge, J. A. & Hogrefe, H. H. Amplification efficiency of thermostable DNA polymerases. *Anal. Biochem.* **321**, 226–235 (2003).
85. Kanagawa, T. Bias and artifacts in multitemplate polymerase chain reactions (PCR). *J. Biosci. Bioeng.* **96**, 317–323 (2003).
86. McInerney, P., Adams, P. & Hadi, M. Z. Error rate comparison during polymerase chain reaction by DNA polymerase. *Mol. Biol. Int.* **2014**, 1–8 (2014).
87. Sims, D., Sudbery, I., Illott, N. E., Heger, A. & Ponting, C. P. Sequencing depth and coverage: key considerations in genomic analyses. *Nat. Rev. Genet.* **15**, 121–132 (2014).
88. Tarazona, S., Garcia-Alcalde, F., Dopazo, J., Ferrer, A. & Conesa, A. Differential expression in RNA-seq: A matter of depth. *Genome Res.* **21**, 2213–2223 (2011).
89. Dueck, H. R. et al. Assessing characteristics of RNA amplification methods for single cell RNA sequencing. *BMC Genomics* **17**, 966 (2016).
90. Meacham, F. et al. Identification and correction of systematic error in high-throughput sequence data. *BMC Bioinformatics* **12**, 451 (2011).
91. Nakamura, K. et al. Sequence-specific error profile of Illumina sequencers. *Nucleic Acids Res.* **39**, e90–e90 (2011).

92. Espinosa, E., Bautista, R., Larrosa, R. & Plata, O. Advancements in long-read genome sequencing technologies and algorithms. *Genomics* **116**, 110842 (2024).
93. Polonis, K. et al. Innovations in short-read sequencing technologies and their applications to clinical genomics. *Clin. Chem.* **71**, 97–108 (2025).
94. Saikia, M. et al. Simultaneous multiplexed amplicon sequencing and transcriptome profiling in single cells. *Nat. Methods* **16**, 59–62 (2019).
95. Moro, G. et al. RoCK and ROI: single-cell transcriptomics with multiplexed enrichment of selected transcripts and region-specific sequencing. *Nat. Commun.* **16**, 10991 (2025). **This article introduces a targeting method combining both targeted capture and targeted priming to enhance the detection of TOIs and ROIs in scRNA-seq experiments.**
96. Replogle, J. M. et al. Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nat. Biotechnol.* **38**, 954–961 (2020).
97. Vallejo, A. F. et al. Resolving cellular systems by ultra-sensitive and economical single-cell transcriptome filtering. *iScience* **24**, 102147 (2021).
98. Nam, A. S. et al. Somatic mutations and cell identity linked by Genotyping of Transcriptomes. *Nature* **571**, 355–360 (2019). **This study introduces a targeting method combining targeted priming and dual targeted PCR to enhance TOI and ROI detection in scRNA-seq experiments.**
99. Giustacchini, A. et al. Single-cell transcriptomics uncovers distinct molecular signatures of stem cells in chronic myeloid leukemia. *Nat. Med.* **23**, 692–702 (2017).
100. Mair, F. et al. A targeted multi-omic analysis approach measures protein expression and low-abundance transcripts on the single-cell level. *Cell Rep.* **31**, 107499 (2020).
101. Cortés-López, M. et al. Single-cell multi-omics defines the cell-type-specific impact of splicing aberrations in human hematopoietic clonal outgrowths. *Cell Stem Cell* **30**, 1262–1281.e8 (2023).
102. Van Horebeek, L. et al. A targeted sequencing extension for transcript genotyping in single-cell transcriptomics. *Life Sci. Alliance* **6**, e202301971 (2023). **This paper uses targeted amplification to obtain transcript level genotyping information, in combination with information on the cellular transcriptome.**
103. Uzbass, F. et al. BART-Seq: cost-effective massively parallelized targeted sequencing for genomics, transcriptomics, and single-cell analysis. *Genome Biol.* **20**, 155 (2019).
104. Rodríguez-Meira, A. et al. Unravelling intratumoral heterogeneity through high-sensitivity single-cell mutational analysis and parallel RNA sequencing. *Mol. Cell* **73**, 1292–1305.e8 (2019).
105. Yuan, D. J. et al. Genotype-to-phenotype mapping of somatic clonal mosaicism via single-cell co-capture of DNA mutations and mRNA transcripts. <https://doi.org/10.1101/2024.05.22.595241> (2024).
106. Riemondy, K. A. et al. Recovery and analysis of transcriptome subsets from pooled single-cell RNA-seq libraries. *Nucleic Acids Res.* **47**, e20–e20 (2019). **This paper uses several targeting methods to enhance information on TOIs and specific cells of interest in previously generated scRNA-seq libraries.**
107. Pokhilko, A. et al. Targeted single-cell RNA sequencing of transcription factors enhances the identification of cell types and trajectories. *Genome Res.* **31**, 1069–1081 (2021).
108. Foyt, D. et al. HybriSeq: probe-based device-free single-cell RNA profiling. *Commun. Biol.* **8**, 1250 (2025).
109. Janesick, A. et al. High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis. *Nat. Commun.* **14**, 8353 (2023).
110. Marshall, J. L. et al. HyPR-seq: Single-cell quantification of chosen RNAs via hybridization and sequencing of DNA probes. *Proc. Natl. Acad. Sci. USA* **117**, 33404–33413 (2020).
111. Scacchetti, A. et al. A ligation-independent sequencing method reveals tRNA-derived RNAs with blocked 3' termini. *Mol. Cell* **84**, 3843–3859.e8 (2024).
112. Oikonomopoulos, S. et al. Methodologies for transcript profiling using long-read technologies. *Front. Genet.* **11**, 606 (2020).
113. Gupta, P., O'Neill, H., Wolvetang, E. J., Chatterjee, A. & Gupta, I. Advances in single-cell long-read sequencing technologies. *NAR Genomics Bioinform.* **6**, lqae047 (2024).
114. Kumari, P. et al. Advances in long-read single-cell transcriptomics. *Hum. Genet.* <https://doi.org/10.1007/s00439-024-02678-x> (2024).
115. Yuan, C. U., Quah, F. X. & Hemberg, M. Single-cell and spatial transcriptomics: bridging current technologies with long-read sequencing. *Mol. Asp. Med.* **96**, 101255 (2024).
116. Monzó, C., Liu, T. & Conesa, A. Transcriptomics in the era of long-read sequencing. *Nat. Rev. Genet.* <https://doi.org/10.1038/s41576-025-00828-z> (2025). **This review highlights recent advances in the long-read sequencing field, as well as relevant applications and future perspectives.**
117. Zajac, N. et al. Comparison of single-cell long-read and short-read transcriptome sequencing via cDNA molecule matching: quality evaluation of the MAS-ISO-seq approach. *NAR Genomics Bioinform.* **7**, lqaf089 (2025).
118. Singh, M. et al. High-throughput targeted long-read single cell sequencing reveals the clonal and transcriptional landscape of lymphocytes. *Nat. Commun.* **10**, 3120 (2019). **This paper introduces a method to combine long-read sequencing with cell-type information obtained through 3'/5'-based methods.**
119. Gupta, I. et al. Single-cell isoform RNA sequencing characterizes isoforms in thousands of cerebellar cells. *Nat. Biotechnol.* **36**, 1197–1202 (2018).
120. Lebrigand, K., Magnone, V., Barbry, P. & Waldmann, R. High throughput error corrected Nanopore single cell transcriptome sequencing. *Nat. Commun.* **11**, 4025 (2020).
121. Tian, L. et al. Comprehensive characterization of single-cell full-length isoforms in human and mouse with long-read sequencing. *Genome Biol.* **22**, 310 (2021).
122. Thijssen, R. et al. Single-cell multiomics reveal the scale of multilayered adaptations enabling CLL relapse during venetoclax therapy. *Blood* **140**, 2127–2141 (2022).
123. Shi, Z.-X. et al. High-throughput and high-accuracy single-cell RNA isoform analysis using PacBio circular consensus sequencing. *Nat. Commun.* **14**, 2631 (2023).
124. Al'Khafaji, A. M. et al. High-throughput RNA isoform sequencing using programmed cDNA concatenation. *Nat. Biotechnol.* **42**, 582–586 (2024).
125. Byrne, A. et al. Single-cell long-read targeted sequencing reveals transcriptional variation in ovarian cancer. *Nat. Commun.* **15**, 6916 (2024). **This paper introduces probe hybridization to enhance the detection of TOIs and cDNAs containing a barcode sequence in long-read sequencing libraries combined with 3'/5'-based methods.**
126. Chen, C. et al. Single-cell multiomics reveals increased plasticity, resistant populations, and stem-cell-like blasts in KMT2A-rearranged leukemia. *Blood* **139**, 2198–2211 (2022).
127. Tian, L., Chen, F. & Macosko, E. Z. The expanding vistas of spatial transcriptomics. *Nat. Biotechnol.* **41**, 773–782 (2023).
128. Jain, S. & Eadon, M. T. Spatial transcriptomics in health and disease. *Nat. Rev. Nephrol.* **20**, 659–671 (2024).
129. Cheng, M. et al. Spatially resolved transcriptomics: a comprehensive review of their technological advances, applications, and challenges. *J. Genet. Genomics* **50**, 625–640 (2023).

130. Lubeck, E., Coskun, A. F., Zhiyentayev, T., Ahmad, M. & Cai, L. Single-cell in situ RNA profiling by sequential hybridization. *Nat. Methods* **11**, 360–361 (2014).
131. Moffitt, J. R. et al. High-performance multiplexed fluorescence in situ hybridization in culture and tissue with matrix imprinting and clearing. *Proc. Natl. Acad. Sci. USA* **113**, 14456–14461 (2016).
132. Valihrach, L., Zucha, D., Abaffy, P. & Kubista, M. A practical guide to spatial transcriptomics. *Mol. Asp. Med.* **97**, 101276 (2024).
133. Ståhl, P. L. et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78–82 (2016).
134. Rodriques, S. G. et al. Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463–1467 (2019).
135. Chen, A. et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* **185**, 1777–1792.e21 (2022).
136. McKellar, D. W. et al. Spatial mapping of the total transcriptome by in situ polyadenylation. *Nat. Biotechnol.* **41**, 513–520 (2023).
137. Engblom, C. et al. Spatial transcriptomics of B cell and T cell receptors reveals lymphocyte clonal dynamics. *Science* **382**, ead8486 (2023).
138. Sudmeier, L. J. et al. Distinct phenotypic states and spatial distribution of CD8⁺ T cell clonotypes in human brain metastases. *Cell Rep. Med.* **3**, 100620 (2022).
139. Chen, T.-Y., You, L., Hardillo, J. A. U. & Chien, M.-P. Spatial transcriptomic technologies. *Cells* **12**, 2042 (2023).
140. Galeano Niño, J. L. et al. INVADEseq to identify cell-adherent or invasive bacteria and the associated host transcriptome at single-cell-level resolution. *Nat. Protoc.* **18**, 3355–3389 (2023).
141. Homberger, C., Barquist, L. & Vogel, J. Ushering in a new era of single-cell transcriptomics in bacteria. *microLife* **3**, uqac020 (2022).
142. Nilsson, M. S. et al. Multiomic single-cell analysis identifies von Willebrand factor and TIM3-expressing BCR-ABL1⁺ CML stem cells. <https://doi.org/10.1101/2023.09.14.557507> (2023).
143. Dixit, A. et al. Perturb-Seq: dissecting molecular circuits with scalable single-Cell RNA profiling of pooled genetic screens. *Cell* **167**, 1853–1866.e17 (2016).
144. Schraivogel, D. et al. Targeted Perturb-seq enables genome-scale genetic screens in single cells. *Nat. Methods* **17**, 629–635 (2020).
145. Shevade, K. et al. Simultaneous capture of single cell RNA-seq, ATAC-seq, and CRISPR perturbation enables multiomic screens to identify gene regulatory relationships. *Cell Rep. Methods* **5**, 101222 (2025).
146. Mimitou, E. P. et al. Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat. Methods* **16**, 409–412 (2019).
147. Alfonzo, J. D. et al. A call for direct sequencing of full-length RNAs to identify all modifications. *Nat. Genet.* **53**, 1113–1116 (2021).
148. Abay, T. et al. Transcript-specific enrichment enables profiling of rare cell states via single-cell RNA sequencing. *Nat. Genet.* **57**, 451–460 (2025).
149. Liao, A. et al. EnrichSci: transcript-guided targeted cell enrichment for scalable single-cell RNA sequencing. Preprint at *bioRxiv* <https://www.biorxiv.org/content/10.1101/2025.05.02.651937v1> (2025).
150. Fang, N. & Akinci-Tolun, R. Depletion of ribosomal RNA sequences from single-cell RNA-sequencing library. *CP Mol. Biol.* **115**, 7.27.1–7.27.20 (2016).
151. Loi, D. S. C., Yu, L. & Wu, A. R. Effective ribosomal RNA depletion for single-cell total RNA-seq by scDASH. *PeerJ* **9**, e10717 (2021).
152. Nishimura, M., Takeyama, H. & Hosokawa, M. Enhancing the sensitivity of bacterial single-cell RNA sequencing using RamDA-seq and Cas9-based rRNA depletion. *J. Biosci. Bioeng.* **136**, 152–158 (2023).
153. Pandey, A. C. et al. A CRISPR/Cas9-based enhancement of high-throughput single-cell transcriptomics. *Nat. Commun.* **16**, 4664 (2025).
154. Everaert, C. et al. Blocking abundant RNA transcripts by high-affinity oligonucleotides during transcriptome library preparation. *Biol. Proced. Online* **25**, 7 (2023).
155. Isakova, A., Neff, N. & Quake, S. R. Single-cell quantification of a broad RNA spectrum reveals unique noncoding patterns associated with cell types and states. *Proc. Natl. Acad. Sci. USA* **118**, e2113568118 (2021).
156. Kozarewa, I. et al. Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes. *Nat. Methods* **6**, 291–295 (2009).
157. Lipson, D. et al. Quantification of the yeast transcriptome by single-molecule sequencing. *Nat. Biotechnol.* **27**, 652–658 (2009).
158. Mamanova, L. et al. FRT-seq: amplification-free, strand-specific transcriptome sequencing. *Nat. Methods* **7**, 130–132 (2010).
159. Jaitin, D. A. et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* **343**, 776–779 (2014).
160. Lyu, J. & Chen, C. LAST-seq: single-cell RNA sequencing by direct amplification of single-stranded RNA without prior reverse transcription and second-strand synthesis. *Genome Biol.* **24**, 184 (2023).

Acknowledgements

We thank members of the Basler laboratory, including G. Hausmann, T. Dalessi, F. Fazilat, J. Little, and T. Valenta, for their valuable input on the manuscript. Given the broad scope of this review and the wide range of papers offering diverse targeting strategies, we may not have cited all relevant work and sincerely apologize to any authors whose contributions we may have overlooked.

Author contributions

G.M. researched the data and wrote the first draft of the manuscript. E.B. and K.B. contributed to the conception of the study and provided supervision. G.M., E.B., and K.B. contributed to discussions of the content and reviewed and edited the manuscript prior to submission.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-026-09675-y>.

Correspondence and requests for materials should be addressed to Giulia Moro.

Peer review information *Communications Biology* thanks Jo Vandesompele, Brecht Soulliaert and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: George Inglis. [A peer review file is available].

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026