Comparative analysis of single-cell RNA-sequencing methods

*Christoph Ziegenhain[1], Swati Parekh[1], Beate Vieth[1], Björn Reinius[2], Martha Smets[3], Heinrich Leonhardt[3], Ines Hellmann[1] and Wolfgang Enard[1]**

[1] Anthropology & Human Genomics, Department of Biology II, Ludwig-Maximilians University, Großhaderner Str. 2, 82152 Martinsried, Germany.

[2] Ludwig Institute for Cancer Research, Box 240, and the Department of Cell and Molecular Biology, Karolinska Institutet, 171 77 Stockholm, Sweden.

[3] Human Biology & Bioimaging, Department of Biology II, Ludwig-Maximilians University, Großhaderner Str. 2, 82152 Martinsried, Germany.

*** Corresponding author**
Wolfgang Enard
Anthropology and Human Genomics
Department of Biology II
Ludwig-Maximilians University
Großhaderner Str. 2
82152 Martinsried, Germany
Phone: +49 (0)89 / 2180 - 74 339
Fax: +49 (0)89 / 2180 - 74 331
E-Mail: enard@bio.lmu.de

Keywords: single-cell RNA-seq, transcriptomics, differential expression analysis

Abstract

Single-cell RNA sequencing (scRNA-seq) offers exciting possibilities to address biological and medical questions, but a systematic comparison of recently developed protocols is still lacking. Here, we generated data from 447 mouse embryonic stem cells using Drop-seq, SCRB-seq, Smart-seq (on Fluidigm C1) and Smart-seq2 and analyzed existing data from 35 mouse embryonic stem cells prepared with CEL-seq. We find that Smart-seq2 is the most sensitive method as it detects the most genes per cell and across cells with the most even coverage, well suited for annotating transcriptomes. However, we also find that unique molecular identifiers (UMIs), available for CEL-seq, Drop-seq and SCRB-seq, reduce the measurement noise considerably, which is most relevant for quantifying transcriptomes. Importantly, we show by power simulations that SCRB-seq and Drop-seq are the most cost-efficient methods for detecting differentially expressed genes. Our analyses offer a solid basis for an informed choice among five prominent scRNA-seq protocols and for future evaluations of protocol improvements.

Introduction

Genome-wide quantification of mRNA transcripts can be highly informative for the characterization of cellular states and to understand regulatory circuits and processes[1,2]. Ideally, such data are collected with high spatial resolution, and scRNA-seq now allows for transcriptome-wide analyses of individual cells, revealing new and exciting biological and medical insights[3–5]. scRNA-seq requires the isolation of single cells and the conversion of their RNA into cDNA libraries that can be quantified using high-throughput sequencing[4,6]. How well single-cell transcriptomes can be characterized depends on many factors, including the sensitivity of the method, i.e. which and how many mRNAs can be detected, its accuracy, i.e. how well the quantification corresponds to the actual concentration of mRNAs and its precision, i.e. with how much technical noise mRNAs are quantified. Of high practical relevance is also the efficiency of the method, i.e. the monetary cost to characterize single cells e.g. at a certain level of precision. In order to make a well-informed choice among available scRNA-seq methods, it is important to estimate these parameters comparably. Each method is likely to have its own strengths and weaknesses. For example, it has previously been shown that scRNA-seq conducted in the small volumes available in the automated microfluidic platform from Fluidigm (Smart-seq protocol on the C1-platform) performs better than Smart-seq or other commercially available kits in microliter volumes[7]. Furthermore, the Smart-seq protocol has been optimized for sensitivity, even full-length coverage, accuracy and cost[8] and this improved “Smart-seq2” protocol[9] has also become widely used[10–14].

Other protocols have sacrificed full-length coverage for 3’ or 5’ sequencing of mRNAs in order to sequence part of the primer used for cDNA generation. This enables early barcoding of libraries, i.e. the incorporation of well-specific or cell-specific barcodes, allowing to multiplex cDNA amplification and library generation and thereby increasing the throughput of scRNA-seq library generation by one to three orders of magnitude[15–19]. Additionally, this approach allows the incorporation of Unique Molecular Identifiers (UMIs), random nucleotide sequences that tag individual mRNA molecules and hence allow for the distinction between original molecules and amplification duplicates that derive from the cDNA or library amplification[18,20,21]. Utilization of UMI information leads to improved quantification of mRNA molecules[22,23] and has been implemented in several scRNA-seq protocols, such as STRT[22], CEL-seq[23], Drop-seq[17], inDrop[19], MARS-seq[16] or SCRB-seq[15].

However, a thorough and systematic comparison of scRNA-seq methods, evaluating sensitivity, accuracy, precision and efficiency is still lacking. To address this issue, we analyzed 482 scRNA-seq libraries from mouse embryonic stem cells (mESCs) generated using five different methods with two technical replicates for each method (Fig. 1).

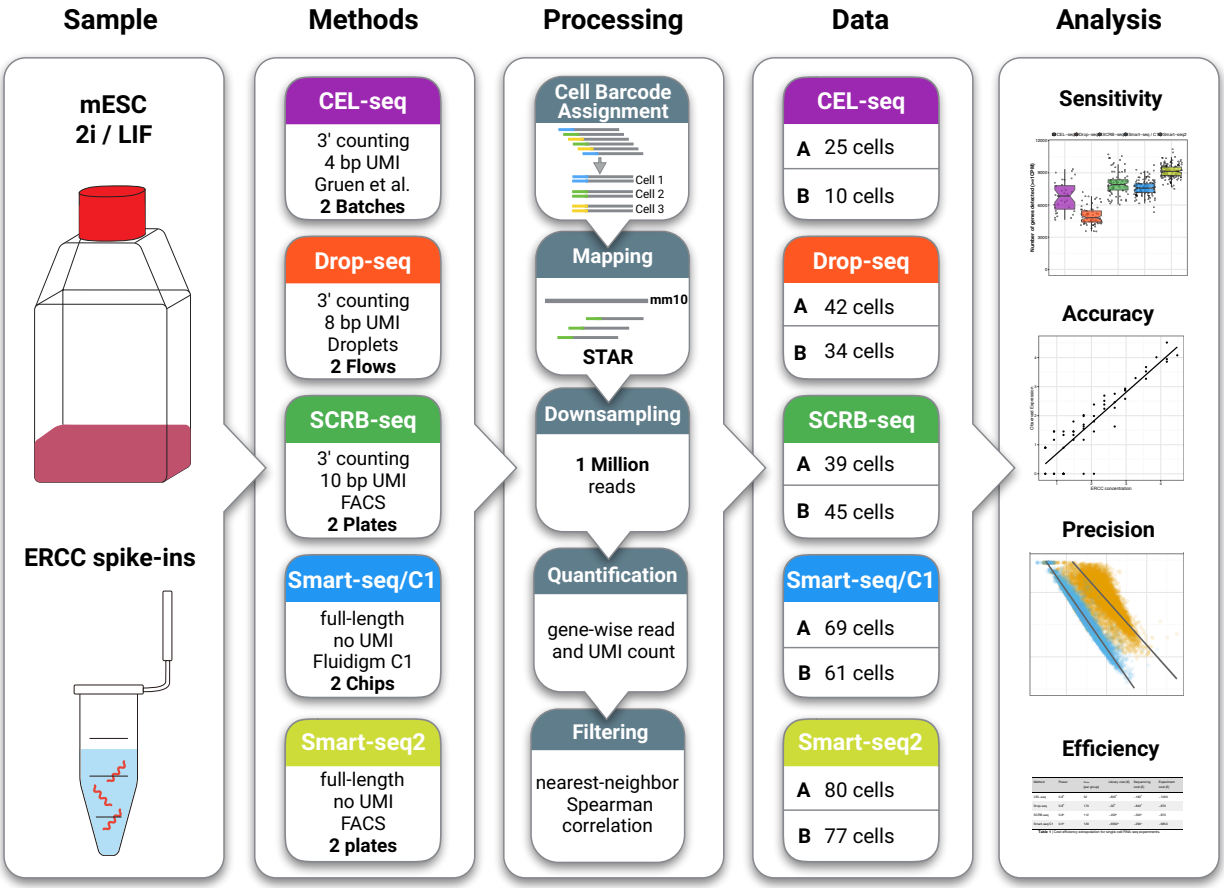

Figure 1 | Schematic overview of the experimental and computational workflow. Mouse embryonic stem cells (mESCs) cultured in 2i/LIF and ERCC spike-in RNA were used to generate single-cell RNA-seq data with five different library preparation methods (CEL-seq, Drop-seq, SCRB-seq, Smart-seq/C1 and Smart-seq2). The methods differ in the usage of unique molecular identifier sequences (UMI), which allow the discrimination between reads derived from original mRNA molecules and duplicates during cDNA amplification. Data processing was identical across methods and analyzed cell numbers per method and replicate are given, which were used to compare sensitivity, accuracy, precision and cost-efficiency. The five scRNA-seq methods are denoted by color throughout the figures of this study: purple - CEL-seq, orange - Drop-seq, green SCRB-seq, blue - Smart-seq, yellow - Smart-seq2.

Results

Generation of scRNA-seq libraries

We generated scRNA-seq libraries from mouse embryonic stem cells (mESCs) in two independent replicates using Smart-seq[24], Smart-seq2[8], Drop-seq[17] and SCRB-seq[15]. Additionally, we used available scRNA-seq data[23] from mESCs that was generated using CEL-seq[18]. An overview of the employed methods and their library generation workflows is provided in Figure 2 and in Supplementary Table 1.

For each replicate of the Smart-seq protocol, we performed a run on the C1 platform from Fluidigm (Smart-seq/C1) using the 10-17 µm mRNA-seq Integrated Fluidic Circuit (IFCs) microfluidic chips that can automatically capture up to 96 cells[7]. We imaged the cells to identify doublets (see below) and added lysis buffer together with External RNA Control Consortium spike-ins (ERCCs) that consist of 92 poly-adenylated synthetic RNA transcripts spanning a range of concentrations[25]. We used the commercially available Smart-seq kit (Clontech) that uses oligo-dT priming, template switching and PCR amplification to generate full-length double-stranded cDNA. We harvested the amplified cDNAs and converted them into 96 different sequenceable libraries by tagmentation (Nextera, Illumina) and PCR amplification using indexed primers for multiplexing. Advantages of this system include that single cell isolation and cDNA generation is automated, that captured cells can be imaged, that reaction volumes are small and that full-length cDNA libraries are sequenced.

For each replicate of the Smart-seq2 protocol, we sorted mESCs by flow cytometry into 96-well PCR plates containing lysis buffer and ERCCs. We generated cDNA as described[8,9] and used an in-house produced Tn5 transposase[26] to generate 96 libraries by tagmentation. While Smart-seq/C1 and Smart-seq2 are very similar protocols that generate full-length libraries they differ in how cells are isolated, the reaction volume and in that Smart-seq2 has been systematically optimized[8,9]. The main disadvantage of both protocols is that the generation of full-length cDNA libraries precludes an early barcoding step and the incorporation of UMIs.

For each replicate of the SCRB-seq protocol[15], we also sorted mESCs by flow cytometry into 96-well PCR plates containing lysis buffer and ERCCs. Also similar to Smart-seq2, cDNA is generated by oligo-dT priming, template switching and PCR amplification of full-length cDNA. However, the oligo-dT primers contain well-specific (i.e. cell-specific) barcodes and UMIs. Hence, cDNA from one plate can be pooled and then be converted into RNA-seq libraries, whereas a modified transposon-based fragmentation approach is used that enriches for 3' ends. The protocol is optimized for small volumes and few handling steps, but it does not generate full-length RNA-seq profiles and its performance compared to other methods is unknown.

The fourth method evaluated was Drop-seq, a recently developed microdroplet-based approach[17]. Similarly to SCRB-seq, each cDNA molecule is labeled with a cell-specific

multiplexing barcode and an UMI to count original mRNA molecules. In the case of Drop-seq, over 10^8 of such barcoded oligo-dT primers are immobilized on beads with each bead carrying a unique cell barcode. A flow of beads suspended in lysis buffer and a flow of a single-cell suspension are brought together in a microfluidic chip that generates nanoliter-sized emulsion droplets. Cells are lysed within these droplets, their mRNA binds to the oligo-dT-carrying beads, and after breaking the droplets reverse transcription, template switching and library generation is performed for all cells in parallel in a single tube. The ratio of beads to cells (20:1) ensures that the vast majority of the beads have either no (>95% expected when Poisson distributed) or one single cell (4.8% expected) in their droplet and hence ensures that doublets are rare (<0.12% expected)[17]. We benchmarked our Drop-seq setup as recommended[17] and determined the doublet rate by mixing mouse and human T-cells (~2.5% of sequenced cell transcriptomes; Supplementary Fig. 1a), confirming that the Drop-seq protocol works well in our setup. The main advantage of the protocol is that many scRNA-seq libraries can be generated at low costs. One disadvantage is that the simultaneous inclusion of ERCC spike-ins is not practical for Drop-seq, as their addition would generate cDNA from ERCCs also in all beads that have no cell and hence would approximately double the sequencing costs. As a proxy for the missing ERCC data, we used a published dataset[17], where ERCC spike-ins were sequenced by the Drop-seq method without single-cell transcriptomes.

Finally, we re-analyzed data[23] generated using CEL-seq[18] for which two replicates of scRNA-seq libraries were available for the same cell type and culture conditions (mESCs in 2i/LIF). Similarly to Drop-seq and SCRB-seq, cDNA is tagged with multiplexing barcodes and UMIs. As opposed to the four PCR-based methods described above, CEL-seq relies on linear amplification by in-vitro transcription (IVT) for the initial pre-amplification of single-cell material.

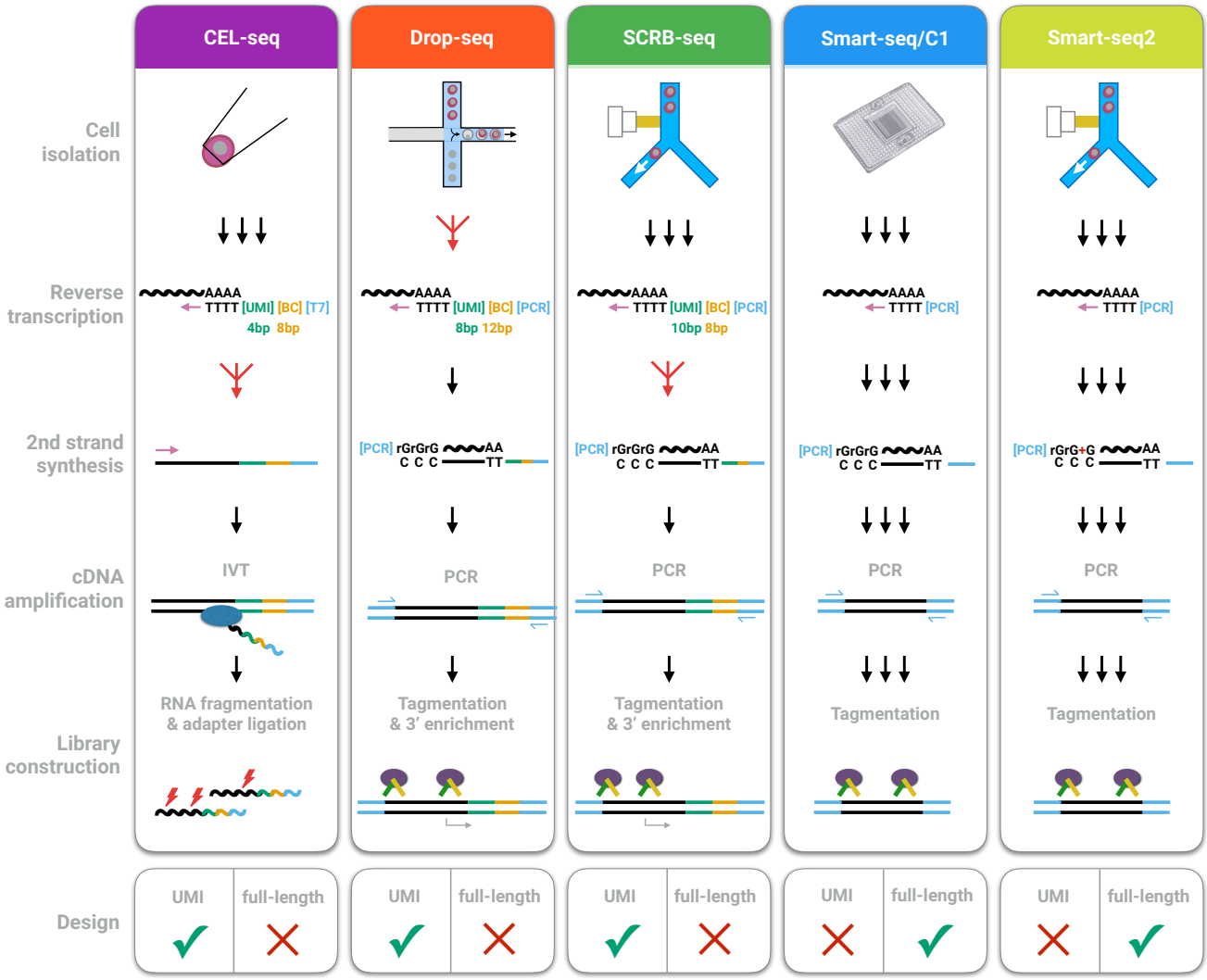

Figure 2 | Schematic overview of key library preparation steps in each method analyzed in this study.

Processing of scRNA-seq libraries

For Smart-seq2, Smart-seq/C1, SCRB-seq and Drop-seq we generated libraries from 192, 192, 192 and ~200 cells in the two independent replicates and sequenced a total of 852, 437, 443 and 866 million reads, respectively. The data from CEL-seq consisted of 102 million reads from a total of 74 cells (Fig. 1, Supplementary Fig. 1b). All data were processed identically, with cDNA reads clipped to 45 bp, mapped using STAR[27] and UMIs being quantified using the Drop-seq pipeline[17]. To adjust for differences in sequencing depths, we used only cells with at least one million reads, resulting in 40, 79, 93, 162, 187 cells for CEL-seq, Drop-seq, SCRB-seq, Smart-seq/C1 and Smart-seq2, respectively. To exclude doublets (libraries generated from two or more cells) in the Smart-seq/C1 data, we analyzed microscope images of the microfluidic chips and identified 16 reaction chambers with multiple cells that were excluded from further analysis. For the three UMI methods, we calculated the number of UMIs per library and found that - at least in our case of a rather homogenous cell population - doublets can be readily identified as libraries that have more than twice the mean total UMI count (Supplementary Fig. 1c), which lead to the removal of 0, 3 and 9 cells for CEL-seq, Drop-seq and SCRB-seq, respectively.

Finally, to remove low-quality libraries, we used a method that exploits the fact that transcript detection and abundance in low-quality libraries correlate poorly with high-quality libraries as well as with other low-quality libraries[28]. We therefore determined the maximum Spearman correlation coefficient for each cell in all-to-all comparisons, which readily allowed the identification of low-quality libraries by visual inspection of the distributions of correlation coefficients (Supplementary Fig. 1c). This filtering led to the removal of 5, 16, 30 cells for CEL-seq, Smart-seq/C1, Smart-seq2, respectively, while no cells were removed for Drop-seq and SCRB-seq. The higher number for the two Smart-seq methods is consistent with the notion that in the early barcoding methods (CEL-seq, Drop-seq, SCRB-seq), low-quality cells are probably outcompeted by high-quality cells so that they do not pass our one million reads filter. As Smart-seq/C1 and Smart-seq2 libraries are generated in separate reactions, filtering by correlation coefficient is more important for these methods.

In summary, we processed and filtered our data so that we could use a total of 482 high-quality, equally sequenced scRNA-seq libraries for a fair comparison of the sensitivity, accuracy, precision and efficiency of the methods.

Single-cell libraries are sequenced to reasonable saturation at one million reads

For all five methods >50% of the reads mapped to the mouse genome (Fig. 3a), comparable to previous results[7,16]. Overall, between 48% (Smart-seq2) and 32% (CEL-seq) of all reads were exonic and thus used to quantify gene expression levels. However, the UMI data showed that only 12 %, 5 % and 15 % of the exonic reads were derived from independent mRNA molecules for CEL-seq, Drop-seq and SCRB-seq, respectively (Fig. 3a).

This indicates that - at the level of mRNA molecules - most of the libraries complexity has already been sequenced at one million reads. To quantify the relationship between the number of detected genes or mRNA molecules and the number of reads in more detail, we downsampled reads to varying depths and estimated to what extend libraries were sequenced to saturation (Supplementary Fig. 2). The number of unique mRNA molecules plateaued at 28,632 UMIs per library for CEL-seq, increased only marginally at 17,207 UMIs per library for Drop-seq and still increased considerably at 49,980 UMIs per library for SCRB-seq (Supplementary Fig. 2c). Notably, CEL-seq showed a steeper slope at low sequencing depths than both Drop-seq and SCRB-seq, potentially due to a less biased amplification by in vitro transcription. Hence, among the UMI methods we found that SCRB-seq libraries had the highest complexity of mRNA molecules that was not yet sequenced to saturation at one million reads. To investigate saturation for non-UMI-based methods, we applied a similar approach at the gene level by counting the number of genes detected by at least one read. By downsampling, we estimated that ~90% (Drop-seq, SCRB-seq) to 100% (CEL-seq, Smart-seq/C1, Smart-seq2) of all genes present in the library were detected at 1 million reads (Fig. 3b, Supplementary Fig. 2a). In particular, the deep sequencing of Smart-seq2 libraries showed clearly that the number of detected genes did not change when increasing the sequencing depth from one million to five million reads per cell (Supplementary Fig. 2b).

All in all, these analyses show that single-cell RNA-seq libraries are sequenced to a reasonable level of saturation at one million reads, a cut-off that has also been previously suggested for different scRNA-seq datasets[7,29]. While it is important to keep in mind that it can be more efficient to analyze scRNA-seq data at lower coverage (see analyses below), comparing scRNA-seq methods at one million reads per cell represents a reasonable choice for the method comparison of this study.

Smart-seq2 has the highest sensitivity

Taking the number of detected genes per cell as a measure to compare the sensitivity of the five methods, we found that Drop-seq had the lowest sensitivity with a median of 4811 genes detected per cell, while with CEL-seq, SCRB-seq and Smart-seq/C1 6839, 7906 and 7572 genes per cell were detected, respectively (Fig. 3c). Smart-seq2 detected the highest number of genes per cell, with a median of 9138. To compare the total number of genes detected across several cells, we pooled 35 cells per method and detected ~16,000 genes for CEL-seq and Drop-seq, ~17,000 for SCRB-seq, ~18,000 for Smart-seq/C1 and ~19,000 for Smart-seq2 (Fig. 3d). While the vast majority of genes (~12,000) were detected by all methods, ~500 genes were specific to each of the 3' counting methods, but ~1000 genes were specific to each of the two full-length methods (Supplementary Fig. 3a,b). That the full length methods detect more genes in total is also apparent when

plotting the genes detected in all available cells as the 3' counting methods level off well below 20,000 genes while the two full-length methods well above 20,000 genes (Fig. 3d).

How even reads cover mRNAs can be regarded as another measure of sensitivity. As expected, the 3' counting methods showed a strong bias of reads mapped to the 3' end (Supplementary Fig. 4a). However, it is worth mentioning that a considerable fraction of reads also covered more 5' regions, probably due to internal oligo-dT priming[30]. Smart-seq2 showed a more even coverage than Smart-seq, confirming previous findings[8]. A general difference between the 3'-counting and the full-length methods can also be seen in the quantification of expression levels as they are separated by the first principal component explaining 75% of the total variance (Supplementary Fig. 4b).

As an absolute measure of sensitivity, we compared the probability of detecting the 92 spiked-in ERCCs, for which the number of molecules available for library construction is known (Supplementary Fig. 5). We determined the detection probability of each ERCC mRNA as the proportion of cells with non-zero read or UMI counts[31]. For the CEL-seq data, Gruen et al. noted that their ERCCs were likely degraded[23] and we also found that ERCCs from the CEL-seq data are detected with a ten-fold lower efficiency than for the other methods (data not shown). Therefore, we did not consider the CEL-seq libraries for any ERCC-based analyses. For Drop-seq, we used the ERCC-only data set[17] and for the other three methods, 2-5% of the one million reads per cell mapped to ERCCs, which were sequenced to complete saturation at that level (Supplementary Fig. 5b). For Smart-seq2, an ERCC RNA molecule was detected on average in half of the libraries when ~7 molecules were present in the sample, while Smart-seq/C1 required ~11 molecules for detection in half of the libraries. Drop-seq and SCRB-seq has estimates of ~16-17 molecules per cell (Supplementary Fig. 5c-e).

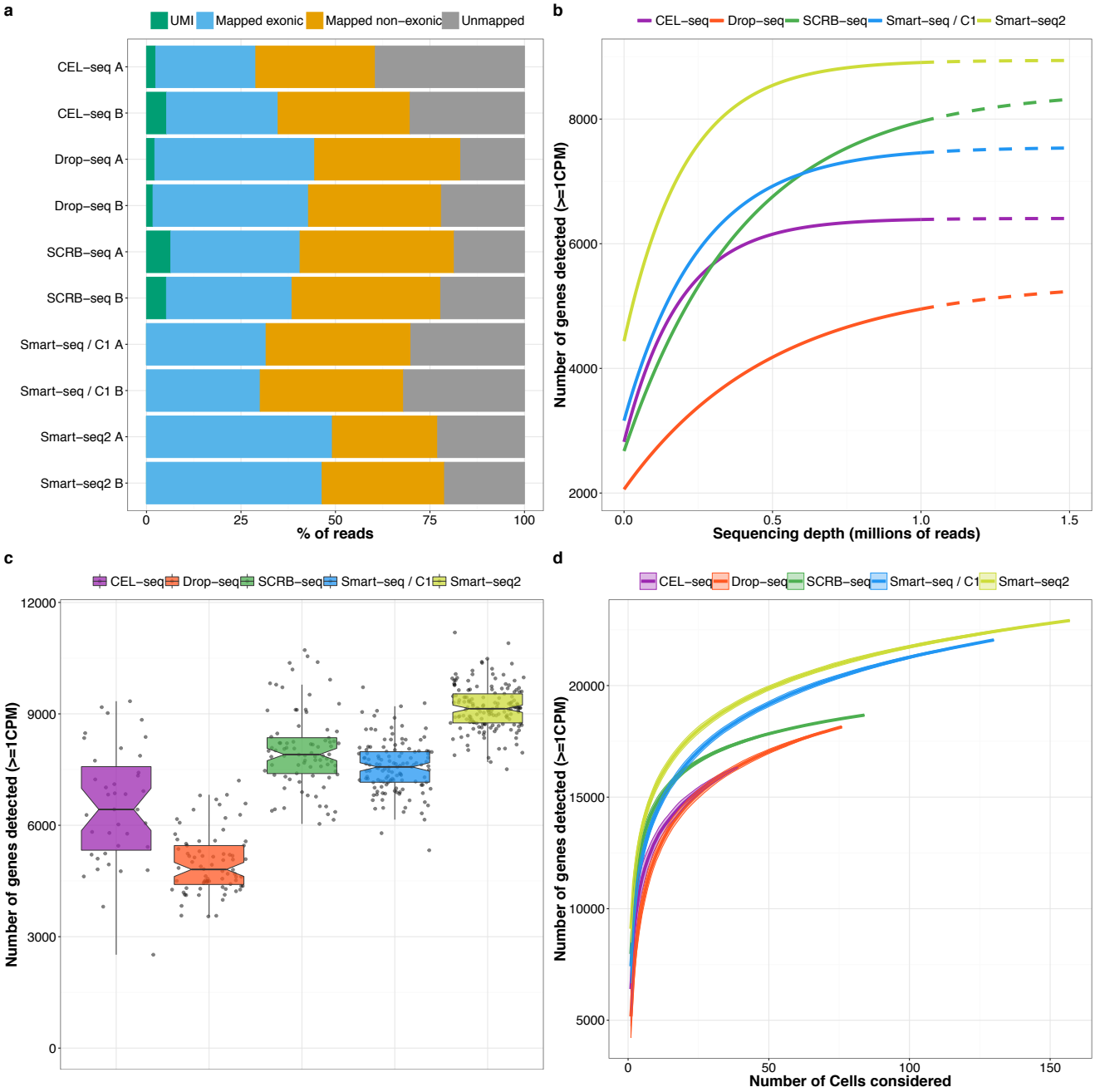

Figure 3 | Sensitivity of scRNA-seq methods. (**a**) Percentage of the 1 million downsampled reads that can not be mapped to the mouse genome (grey), are mapped to regions outside exons (orange), inside exons (blue) and carry a unique UMI (green). For UMI methods, blue denotes the duplicated exonic reads. (**b**) Median number of genes detected (CPM>=1) per cell when downsampling total read counts to indicated depths. Dashed line above 1 million reads represents the extrapolated asymptotic fit. (**c**) Number of genes detected (CPM>=1) per cell. Each dot represents a cell and each boxplot represents the median, first and third quartile per replicate and method. (**d**) Cumulative number of genes detected as more cells are added. The order of cells considered was drawn randomly 100 times to display mean ± standard deviation (shaded area).

Notably, the sensitivity estimated from the number of detected genes does not fully agree with the comparison based on ERCCs. While Smart-seq2 is the most sensitive method in both cases, Drop-seq performs better and SCRB-seq performs worse when using ERCCs. The reasons for this discrepancy are unclear, but several have been noted before[32–34] including that ERCCs do not model endogenous mRNAs perfectly since they are shorter, have shorter poly-A tails, lack a 5' cap and can show batch-wise variation in concentrations as observed for the CEL-seq data. In the case of Drop-seq, it should be kept in mind that ERCCs were sequenced separately as discussed above and in this way leading to a higher efficiency. Therefore, while it is still useful to estimate the absolute range in which molecules are detected, for our purpose of comparing the sensitivity of methods using the same cells, we regard the number of detected genes per cell as the more reliable estimate of sensitivity in our setting, as it sums over many, non-artificial genes.

In summary, we find that Smart-seq2 is the most sensitive method as it detects the highest number genes per cell, the most genes in total across cells and has the most even coverage of transcripts. Smart-seq/C1 is slightly less sensitive per cell, but detects the same number of genes across cells, if one considers its lower fraction of mapped exonic reads (Fig. 3a). Among the 3' counting methods, SCRB-seq is most sensitive, closely followed by CEL-seq, whereas Drop-seq detects considerably fewer genes.

Accuracy is similar across scRNA-seq methods

In order to quantify the accuracy of transcript level quantifications, we compared observed expression values with annotated molecule concentration of the 92 ERCC transcripts (Supplementary Fig. 5a). For each cell, we calculated the correlation coefficient (R^2) for a linear model fit (Fig 4). While the median accuracy did differ among methods (Kruskal-Wallis test, p<2.2e-16), all methods had high accuracies ranging between 0.86 and 0.91. As discussed above, CEL-seq was excluded from the ERCC analyses due to the potential degradation of the ERCCs in this data set[23]. The original publication for CEL-seq from 10 pg of total RNA input and ERCC spike-in reported a mean correlation coefficient of R^2=0.87[18], similar to the correlations reported for the other four methods. Hence, we find that the accuracy is similarly high across all five methods and also because absolute expression levels are rarely of interest, the small differences in accuracy will rarely be a decisive factor when choosing among the five methods.

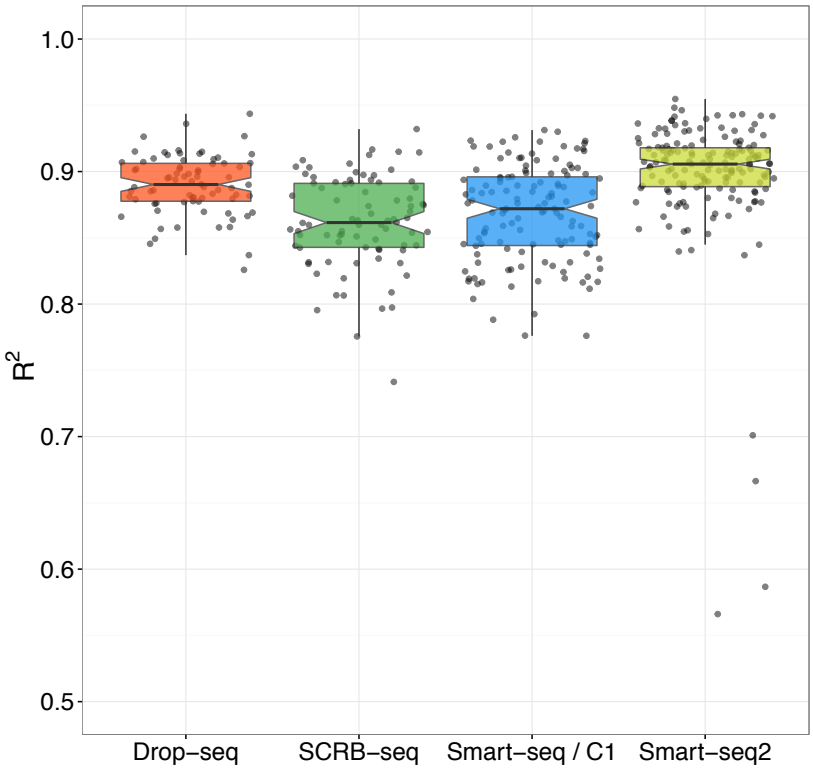

Figure 4 | Accuracy of scRNA-seq methods. ERCC expression values were correlated to their annotated molarity. Shown are the distributions of correlation coefficients (adjusted R^2 of linear regression model) across methods. Each dot represents a cell/bead and each boxplot represents the median, first and third quartile.

Precision is highest for CEL-seq and SCRB-seq and strongly increased by UMIs

While a high accuracy is necessary to quantify absolute expression values, one of the most common experimental aims is to compare relative expression levels in order to identify differentially expressed genes or biological variation between cells. Hence, the precision of a method - i.e. its reproducibility or its amount of technical variation - is the major factor that determines the power to detect differentially expressed genes. As we used the same cells under the same culture conditions for all five methods, we assume that the amount of biological variation is the same across all five methods. Therefore, differences in the total variation among methods can be interpreted as differences in their technical variation. In order to compare variation for the same set of genes across all methods, we analyzed all 12942 genes that were detected in 25% of the cells by at least one method (Supplementary Fig. 6). In this way we also include genes that are hardly or not at all detected by less sensitive methods and allow for a fair comparison of technical variation.

To visualize the amount of variation as proposed before[35], we plotted for each gene its squared coefficient of variation (CV^2) against its mean expression count. As expected[35], we found a strong negative correlation, i.e. that lowly expressed genes vary more (Fig. 5). Importantly, we find that all three UMI methods clearly outperform the two non-UMI methods (Fig. 5b) and that this depends on the usage of UMIs, as the UMI methods show much higher variation when counting reads instead of UMIs (Fig. 5a). This reduction in variation due to UMIs has been described before for CEL-seq[23] and seems even stronger for SCRB-seq and Drop-seq, fitting with the notion that in vitro amplification is more precise than PCR amplification.

So while it is evident from this analysis that CEL-seq and SCRB-seq are the most precise methods (Fig. 5b), it is not equally clear how this affects the detection of differentially expressed genes and how it can inform the experimental design. Therefore, we conducted power simulations that used - for each method across the set of 12942 genes - the observed mean-variance relationship and the dropout probability to simulate read counts when 5% of genes are differently expressed between two groups of single cells. To this end we first estimated the mean and dispersion parameter (i.e. the shape parameter of the gamma mixing distribution) for each gene per method. Next, we fitted a spline to the resulting pairs of mean and dispersion estimates (Supplementary Fig. 7a) in order to predict the dispersion of a gene given its mean. Moreover, we included the sensitivity of each scRNA-seq method in the power simulations by modeling a gene-wise dropout parameter from the observed detection rates also dependent on the mean expression (Supplementary Fig. 7b).

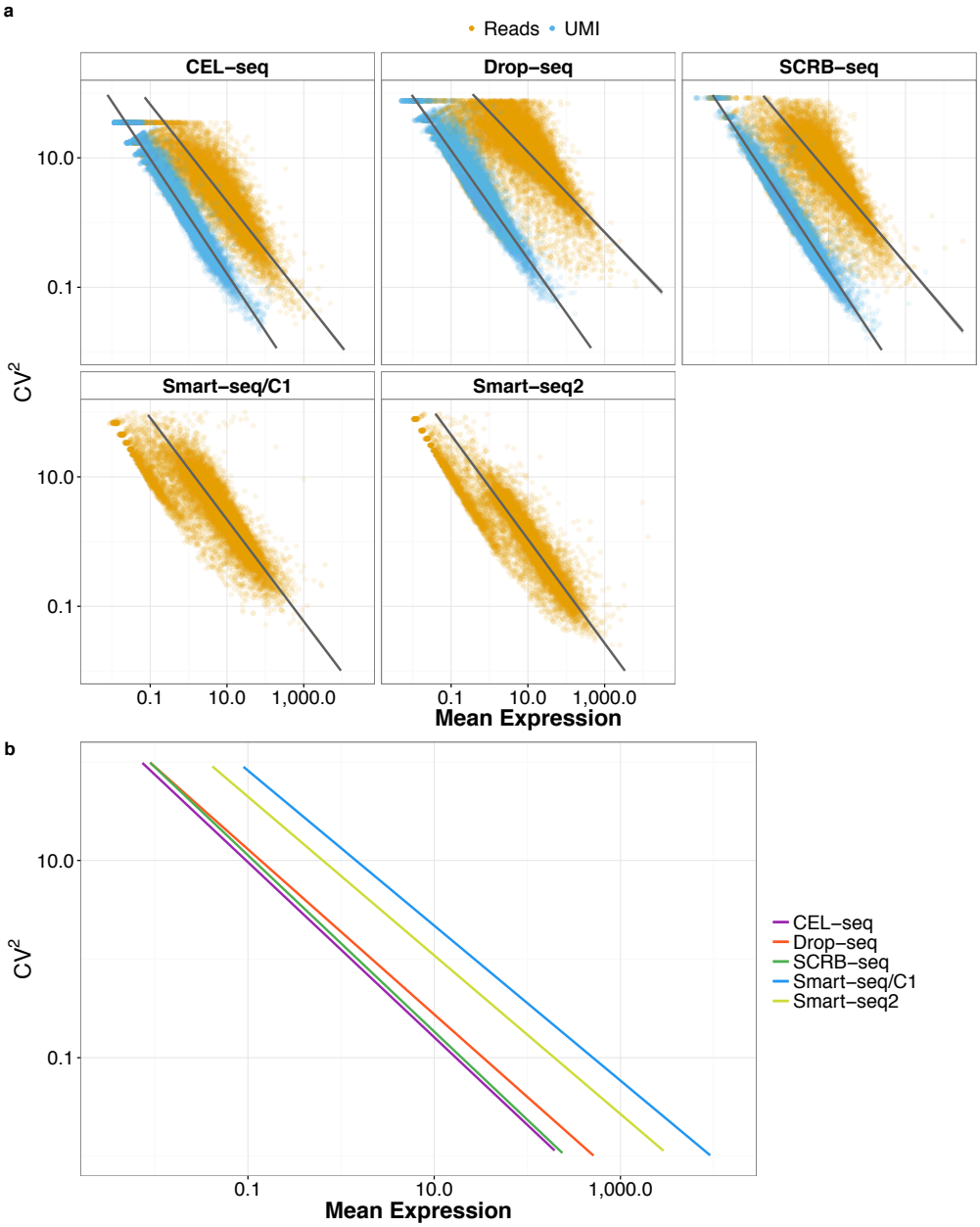

Figure 5 | Precision of scRNA-seq methods. (**a**) Gene-wise mean and squared coefficient of variation (CV) of scRNA-seq data were calculated. Scatter plots are shown for UMI (blue) and read-count (orange) based quantification. (**b**) Linear model (lm) fits of CV^2 for UMI (CEL-seq, Drop-seq, SCRB-seq) and read-count (Smart-seq/C1, Smart-seq2) based quantification.

Subsequently, we simulate read counts for two groups that mimic realistic gene expression changes by adding log-fold changes to 5% of the genes. These log-fold changes were drawn from observed differences between microglial subpopulations from a previously published dataset[36] to mimic a biologically realistic scenario The simulated datasets were then tested for differential expression using limma[37] from which the average true positive rate (TPR) and the average false discovery rate (FDR) could be calculated.

First, we analyzed how the number of cells affects TPR and FDR by running 100 simulations each for a range of 16 to 512 cells per group. As expected, SCRB-seq and CEL-seq performed best, reaching a TPR of 80% with 72 and 77 cells, respectively (Fig. 6a). While Drop-seq reached 80% power with a slightly lower estimate of 111 cells, Smart-seq2 and Smart-seq/C1 reached 80% power only for larger cell numbers of 139 and 190 per group, respectively. FDRs were similar in all methods and just slightly above 5% (Suppl. Fig. 8).

Accordingly, when simulating the power to detect expression differences with a fixed sample size of 64 cells per group, SCRB-seq and CEL-seq performed best (Fig. 6b), followed by Drop-seq, while Smart-seq/C1 had the lowest power. As expected from the mean-variance plots above, when analyzing power without using UMIs, SCRB-seq and CEL-seq performed similar to Smart-seq/C1, while Drop-seq performed very poorly (Fig. 6b). Hence, UMIs strongly increase the power to detect differentially expressed genes by scRNA-seq. For the read-count based Smart-seq methods, Smart-seq2 consistently performed better than Smart-seq/C1.

Next, we asked how TPR and FDR depend on the sequencing depth. We repeated our simulation studies as described above, but estimated the mean-dispersion and mean-dropout relationships from data downsampled to 500,000 or 250,000 reads per cell. Overall, the decrease in power was very moderate (Fig. 6c, Supplementary Fig. 8). Interestingly, not all methods respond to downsampling at similar rates. While CEL-seq was nearly unaffected in power by downsampling to either 500,000 or 250,000 reads, power decreased a bit more for SCRB-seq, when downsampling to 250,000 reads. For Drop-seq, there is a clear reduction in power when downsampling from 1,000,000 to 500,0000 reads, but no further loss is seen for 250,000 reads. For the Smart-seq methods, we do not see a strong effect of downsampling, although a slight variance in the power estimate is observed.

In summary, when using power simulations to compare precision among methods, CEL-seq and SCRB-seq performed best, followed by Drop-seq. Smart-seq2 and especially Smart-seq/C1 performed considerably worse, due to their lack of UMIs.

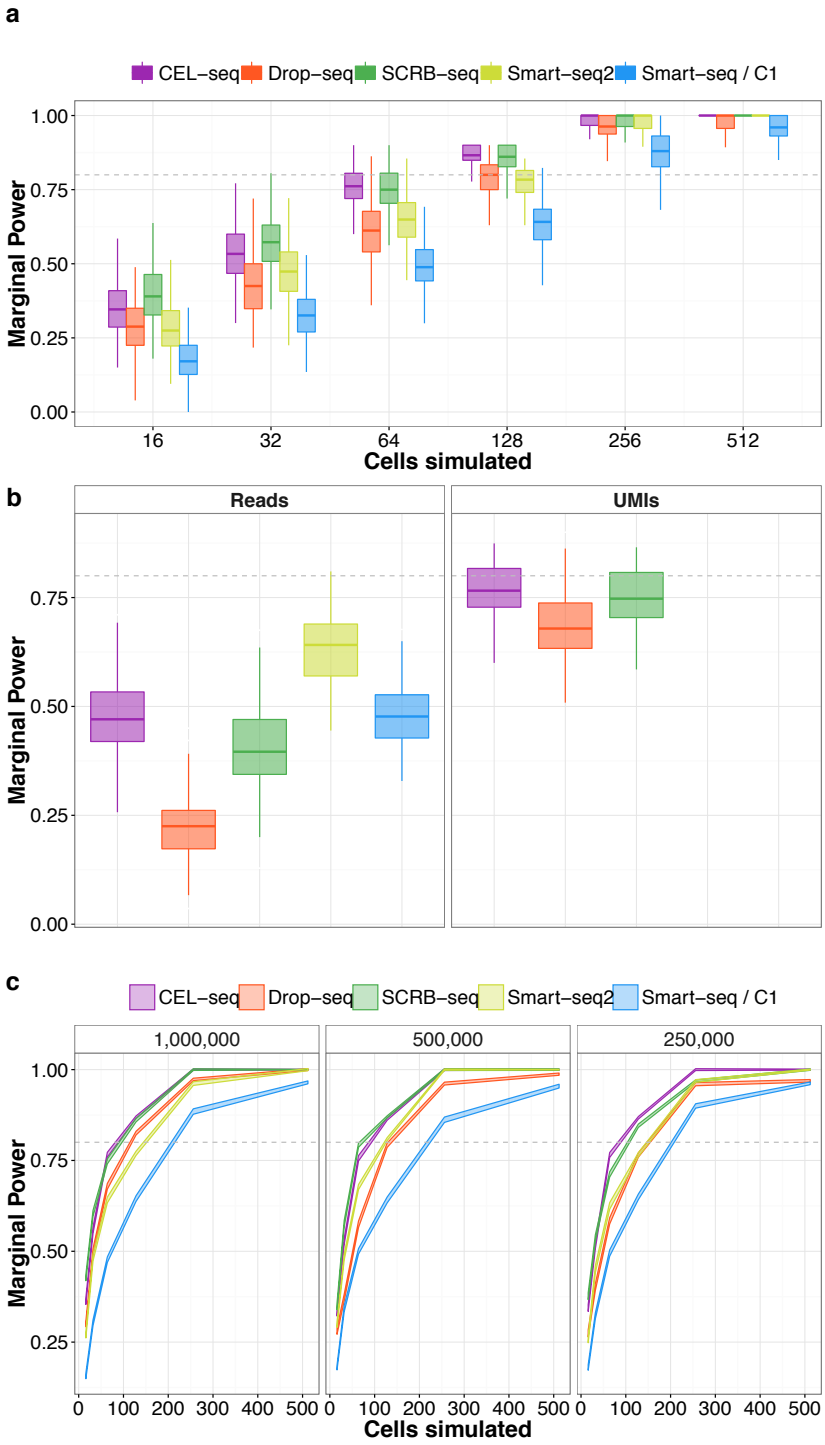

Figure 6 | Power analysis. (**a**)-(**b**) Power simulations using empirical mean/dispersion and mean/dropout relationships from single-cell data at 1 million reads. Simulated effect sizes of at least 0.25 log2 fold-change are considered. Boxplots represent the median, first and third quartile of 100 simulations. (**a**) Power analysis of variable sample sizes n=16, n=32, n=64, n=128, n=256 and n=512. Shown are power calculations of 100 simulations for each method. (**b**) For a fixed sample size of n=64, we computed the power of read-count and UMI quantification over 100 simulations. (**c**) Power simulations using mean/dispersion and mean/dropout estimates from empirical single-cell data at 1 million, 0.5 million and 0.25 million reads. Line areas indicate the median power with standard error over 100 simulations.

Efficiency is highest for SCRB-seq and Drop-seq when considering costs and power

In practice, the costs of a method also matter when judging the performance of different scRNA-seq methodologies. We estimated the cost-efficiency by calculating the costs for generating scRNA-seq data at a given amount of power. Given the number of single cells that are needed per group to reach 80% power as simulated above for three sequencing depths (Fig. 6a), we calculated the costs to generate and sequence these libraries. For example, at one million reads, SCRB-seq requires 72 cells per group and generating 144 SCRB-seq libraries costs ~290€ plus sequencing costs of ~720€. We assume that generating paired-end reads for CEL-seq, SCRB-seq and Drop-seq is done with a 50 cycles single end kit and thus costs the same as sequencing single-end Smart-seq libraries.

Method	Marginal Power	FDR (%)	n_{cells} (per group)	Library cost per cell (€)	Experiment cost (€) (incl. sequencing[b])
CEL-seq	0.8[a]	~5.3	77[c] | 80[d] | 80[e]	~8	~ 2000[c] | 1680[d] | 1480[e]
Drop-seq	0.8[a]	~7.1	111[c] | 135[d] | 132[e]	~0.1	~ 1130[c] | 700[d] | 360[e]
SCRB-seq	0.8[a]	~5.8	72[c] | 70[d] | 94[e]	~2	~ 1010[c] | 630[d] | 610[e]
Smart-seq/C1	0.8[a]	~5.8	190[c] | 184[d] | 195[e]	~25	~ 11380[c] | 10130[d] | 10260[e]
Smart-seq2 (commercial)	0.8[a]	~5.1	139[c] | 123[d] | 134[e]	~30	~ 15280[c] | 12890[d] | 13720[e]
Smart-seq2 (in-house Tn5)	0.8[a]	~5.1	139[c] | 123[d] | 134[e]	~3	~ 2220[c] | 1350[d] | 1140[e]

Table 1 | Cost efficiency extrapolation for single-cell RNA-seq experiments.

[a] Based on simulations (Fig. 6a) for detection of $\log_2$ (fold-change) >= 0.25

[b] assuming 5 € per million reads

[c] sequencing depth 1 million reads

[d] sequencing depth 0.5 million reads

[e] sequencing depth 0.25 million reads

When we do analogous calculations for the four other methods (Table 1, Supplementary Fig. 9), we find that at a sequencing depth of 1 million reads SCRB-seq is most cost-effective, followed by Drop-seq and CEL-seq, while Smart-seq/C1 is almost ten-fold less efficient due its high library costs that arise from the microfluidic chips and the costs for generating independent libraries. Smart-seq2 is similarly expensive when relying on commercial library preparation reagents. However, these costs can be significantly reduced by in-house production of Tn5 transposase[26] as was also done in our experiments.

Interestingly, we find that a relatively low sequencing depth seems sufficient for differential expression analysis, confirming previous findings[15,38]. Hence, methods with low per-cell reagent costs are in advantage, despite their lower sensitivity and precision. Indeed, at lower sequencing depths of a quarter million reads, we find that Drop-seq is clearly the most efficient method.

The estimates should be understood as a lower boundary for the real costs since many factors are not considered such as costs to set-up the methods, costs to isolate single cells or costs due to practical constraints in generating a fixed number of scRNA-seq libraries. In particular, it is important that independent biological replicates are needed when investigating particular factors such as genotypes or developmental timepoints and some methods such as Smart-seq/C1 and Drop-seq are less flexible in distributing scRNA-seq libraries across replicates. This said, we do think that our simulations do allow a fair and realistic comparison of the five methods and reveal that SCRB-seq and Drop-seq are the most efficient methods analyzed here.

Discussion

Single-cell RNA-sequencing (scRNA-seq) is a powerful technology to tackle a multitude of biomedical questions. To facilitate choosing among the many approaches that were recently developed, we systematically compared five scRNA-seq methods and assessed their sensitivity, accuracy, precision and cost-efficiency. We chose a leading commercial platform (Smart-seq/C1), one of the most popular full-length methods (Smart-seq2), a method that uses in-vitro transcription for amplification from manually isolated cells (CEL-seq), a PCR-based method with a very high throughput (Drop-seq) and a UMI-based method that allows single-cell isolation by FACS (SCRB-seq). Protocols are available for all these methods and can therefore be set up by any molecular biology lab.

We find that SCRB-seq, Smart-seq/C1 and CEL-seq detect a similar number of genes per cell, while Drop-seq detects nearly 50% less than the most sensitive method Smart-seq2 (Fig. 3b,c). Despite this lower per cell sensitivity, Drop-seq does not generally detect fewer genes since the total number of detected genes converges around 18,000, similar as for SCRB-seq and CEL-seq (Fig. 3d). A potential explanation could be that a fraction of mRNA molecules gets detached from the beads when droplets are broken up for reverse transcription. It will be interesting to see whether this step could be optimized in the future.

While the three 3' counting methods detect largely the same set of genes, Smart-seq/C1 and Smart-seq2 detect around 3000 additional genes (Fig. 3d, Supplementary Fig. 3b), suggesting that some 3' ends of cDNAs might be difficult to convert to sequenceable molecules. When using ERCCs to compare absolute sensitivities, we again find Smart-seq2 to be the most sensitive method. However, we also find that sensitivity estimates from ERCCs do not perfectly correlate with estimates from endogenous genes, suggesting that they might not always be an ideal benchmark for comparing different methods. In summary,

we find that Smart-seq2 is the most sensitive method based on its gene detection rate per cell and in total. In addition, Smart-seq2 shows the most even read coverage across transcripts (Supplementary Fig. 4a), making it the most appropriate method for detecting alternative splice forms. Hence, it would in general be the most suitable method when an annotation of single cell transcriptomes is the focus.

We find that accuracy is similarly high across methods. However, because absolute quantification of mRNA molecules is rarely of interest, accuracy is not an important criterion for choosing among the five methods. In contrast, relative quantification of gene expression levels is of interest for most scRNA-seq studies and hence the precision of the method is an important benchmark. When approximating technical variation and reproducibility as gene-wise coefficient of variation, we find that SCRB-seq and CEL-seq have the highest precision. Both the variance analysis and our power simulations show that UMIs increase the power for differential gene expression considerably, especially for the PCR-based methods (Fig. 5a, Fig. 6b). This is due to the large amount of amplification needed for scRNA-seq libraries, as the effect of UMIs on power for bulk RNA-seq libraries is neglectable[39]. Although CEL-seq and SCRB-seq have the highest precision for a fixed set of cells, what matters in practice is the trade-off between costs and power. Using simulations considering the mean-variance relationship and the dropout probabilities for a realistic level of differential gene expression, we find that at one million reads, SCRB-seq is the most efficient method. When reducing the number of sequenced reads per cell to 250,000, it is Drop-seq (Table 1, Supplementary Fig. 9). Interestingly, Smart-seq2 using in-house produced transposase is - despite its lower precision - still reasonably efficient, probably due to its higher sensitivity. In contrast, Smart-seq/C1 and Smart-seq2 using commercial transposase are by far the least efficient methods due to their high costs and their lower precision.

As mentioned above, the cost estimates are a lower boundary and many additional aspects will be important when choosing a method. Despite its lower efficiency to quantify expression levels, Smart-seq2 might be preferable if one can produce transposase and is interested also in splice variants. SCRB-seq or Smart-seq2 might be also preferable when rare subpopulations are isolated by FACS as Drop-seq in its current setup requires a large amount of cells. Another advantage of these two methods is that they can also be used for generating bulk RNA-seq libraries. So while such factors will be differently weighted by each individual lab and for each research question, our analyses provide a solid basis for such considerations when choosing among the five analysed methods.

Our analysis also provide a basis for evaluating further improvements of scRNA-seq methods. For example, the efficiency of the Fluidigm C1 platform can be increased considerably, when implementing UMI-based protocols on the C1 platform[22,40], or using early barcoding to increase the number of cells per chip, as available in the HT mRNA-seq IFC. The CEL-seq variant MARS-seq combines the high precision of UMIs and IVT-based

409 amplification with higher throughput[16] and SCRB-seq is even more cost-efficient on a 384-
410 well format[15]. Other recent developments, such as CEL-seq2[40] promise to further increase
411 the performance of scRNA-seq protocols.

412 In summary, we find that Drop-seq is the most efficient method when quantifying
413 transcriptomes of a large numbers of cells with low sequencing depth. Smart-seq2 is
414 probably preferred when annotation of transcriptomes and alternative splicing is of
415 relevance. Smart-seq2 using in-house produced transposase and SCRB-seq might be
416 preferable for analyzing fewer numbers of cells, especially when cells need to be
417 preselected by FACS and bulk RNA-seq needs to be generated with the same
418 methodology.

Methods

Published data

CEL-seq data for J1 mESC cultured in 2i/LIF condition[23] were obtained under accession GSE54695. Drop-seq ERCC[17] data were obtained under accession GSE66694. Raw fastq files were extracted using the SRA toolkit (2.3.5). We trimmed cDNA reads to the same length and processed raw reads in the same way as data sequenced for this study.

Cell culture of mESC

J1 mouse embryonic stem cells were maintained on gelatin-coated dishes in Dulbecco's modified Eagle's medium supplemented with 16% fetal bovine serum (FBS, Sigma-Aldrich), 0.1 mM β-mercaptoethanol (Invitrogen), 2 mM L-glutamine, 1x MEM non-essential amino acids, 100 U/ml penicillin, 100 µg/ml streptomycin (PAA Laboratories GmbH), 1000 U/ml recombinant mouse LIF (Millipore) and 2i (1µM PD032591 and 3µM CHIR99021 (Axon Medchem, Netherlands). J1 embryonic stem cells were obtained from E. Li and T. Chen and mycoplasma free determined by a PCR-based test. Cell line authentication was not recently performed.

Single cell RNA-seq library preparations

Drop-seq

Drop-seq experiments were performed as published[17] and successful establishment of the method in our lab was confirmed by a species-mixing experiment (Supplementary Fig. 1a). For this work, J1 mES cells (100/µl) and barcode-beads (120/µl, Chemgenes) were co-flown in Drop-seq PDMS devices (Nanoshift). Emulsions were broken by addition of perfluoroctanol (Sigma-Aldrich) and mRNA on beads was reverse transcribed (Maxima RT, Thermo Fisher). Unused primers were degraded by addition of Exonuclease I (New England Biolabs). Washed beads were counted and aliquoted for pre-amplification (2000 beads / reaction). Nextera XT libraries were constructed from 1 ng of pre-amplified cDNA with a custom P5 primer (IDT).

SCRB-seq

RNA was stabilized by resuspending cells in RNAprotect Cell Reagent (Qiagen) and RNAse inhibitors (Promega). Prior to FACS sorting, cells were diluted in PBS (Invitrogen). Single cells were sorted into 5 µl lysis buffer consisting of a 1/500 dilution of Phusion HF buffer (New England Biolabs) and ERCC spike-ins (Ambion), spun down and frozen at -80 °C. Plates were thawed and libraries prepared as described previously[15]. Briefly, RNA was desiccated after protein digestion by Proteinase K (Ambion). RNA was reverse transcribed using barcoded oligo-dT primers (IDT) and products pooled and concentrated.

Unincorporated barcode primers were digested using Exonuclease I (New England Biolabs). Pre-amplification of cDNA pools were done with the KAPA HiFi HotStart polymerase (KAPA Biosystems). Nextera XT libraries were constructed from 1 ng of pre-amplified cDNA with a custom P5 primer (IDT).

Smart-seq/C1

Smart-seq/C1 libraries were prepared on the Fluidigm C1 system according to the manufacturer's protocol. Cells were loaded on a 10-17 µm RNA-seq microfluidic IFC at a concentration of 200,000/ml. Capture site occupancy was surveyed using the Operetta (Perkin Elmer) automated imaging platform.

Smart-seq2

mESCs were sorted into 96-well PCR plates containing 2 µl lysis buffer (1.9 µl 0.2% TritonX-100; 0.1 µl RNAseq inhibitor (Lucigen)) and spike-in RNAs (Ambion), spun down and frozen at -80 °C. To generate Smart-seq2 libraries, priming buffer mix containing dNTPs and oligo-dT primers was added to the cell lysate and denatured at 72 °C. cDNA synthesis and pre-amplification of cDNA was performed as described previously[8,9]. Sequencing libraries were constructed from 2.5 ng of pre-amplified cDNA using an in-house generated Tn5 transposase[26]. Briefly, 5 µl cDNA was incubated with 15 µl tagmentation mix (1 µl of Tn5; 2 µl 10x TAPS $MgCl_2$ Tagmentation buffer; 5 µl 40% PEG8000; 7 µl water) for 8 min at 55 °C. Tn5 was inactivated and released from the DNA by the addition of 5 µl 0.2% SDS and 5 min incubation at room temperature. Sequencing library amplification was performed using 5 µl Nextera XT Index primers (Illumina) that had been first diluted 1:5 in water and 15 µl PCR mix (1 µl KAPA HiFi DNA polymerase (KAPA Biosystems); 10µl 5x KAPA HiFi buffer; 1.5 µl 10mM dNTPs; 2.5µl water) in 10 PCR cycles. Barcoded libraries were purified and pooled at equimolar ratios.

DNA sequencing

For SCRB-seq and Drop-seq, final library pools were size-selected on 2% E-Gel Agarose EX Gels (Invitrogen) by excising a range of 300-800 bp and extracting DNA using the MinElute Kit (Qiagen) according to the manufacturer's protocol.

Smart-seq/C1, Drop-seq and SCRB-seq library pools were sequenced on a Illumina HiSeq1500 using High Output mode. Smart-seq2 pools were sequenced on Illumina HiSeq2500 (Replicate A) and HiSeq2000 (Replicate B) platforms. Smart-seq/C1 and Smart-seq2 libraries were sequenced 45 cycles single-end, whereas Drop-seq and SCRB-seq libraries were sequenced paired-end with 20 cycles to decode cell barcodes and UMI from read 1 and 45 cycles into the cDNA fragment. Similar sequencing qualities were confirmed by FastQC v0.10.1 (Supplementary Fig. 1b).

Basic data processing and sequence alignment

Smart-seq/C1/Smart-seq2 libraries (i5 and i7) and Drop-seq/SCRB-seq pools (i7) were demultiplexed from the Nextera barcodes using deML[41]. All reads were trimmed to the same length of 45 bp by cutadapt[42] and mapped to the mouse genome (mm10) including mitochondrial genome sequences and unassigned scaffolds concatenated with the ERCC spike-in reference. Alignments were calculated using STAR 2.4.0[27] using all default parameters.

For libraries containing UMIs, cell- and gene-wise count/UMI tables were generated using the published Drop-seq pipeline (v1.0)[17]. We discarded the last 2 bases of the Drop-seq cell and molecular barcodes to account for bead synthesis errors.

For Smart-seq/C1 and Smart-seq2, features were assigned and counted using the Rsubread package (v1.20.2)[43].

Power Analysis

We developed a custom R package for statistical power evaluation of differential gene expression. For each method, we estimated the mean expression, dispersion and dropout probability per gene from the same number of cells per method. In the read count simulations, we followed the framework proposed in Polyester[44], i.e. we retained the observed mean-variance dependency by applying a smooth spline fit. Furthermore, we included a log-logistic function for the binomial mean-dropout relationship. In each iteration, we simulated count measurements for 12942 genes (shared gene set) for sample sizes of 2^4, 2^5, 2^6, 2^7, 2^8 and 2^9 cells per group.

The read count for a gene *i* in a cell *j* is modeled as a product of a binomial and negative binomial distribution:

$$X_{ij} \sim B(p = 1 - p_0) * NB(\mu, \theta)$$

The mean expression magnitude μ was randomly drawn from the empirical distribution. 5 percent of the genes were defined as differentially expressed with an effect size drawn from the observed fold changes between microglial subpopulations in Zeisel et al[36]. The dispersion θ and dropout probability p_0 were predicted by above mentioned fits.

For each method, 100 RNA-seq experiments were simulated and tested for differential expression using limma[37] in combination with voom[45] (v3.26.7).

ERCC capture efficiency

To estimate the single molecule capture efficiency, we assume that the success or failure of detecting an ERCC is a binomial process, as described before[31]. Detections are independent from each other and are thus regarded as independent Bernoulli trials. We recorded the number of cells with nonzero and zero read or UMI counts for each ERCC per

method and applied a maximum likelihood estimation to fit the probability of successful detection. The fit line was shaded with the 95% Wilson score confidence interval.

Cost efficiency calculation

We based our cost efficiency extrapolation on the power simulations starting from empirical data at different sequencing depths (250,000 reads, 500,000 reads, 1,000,000 reads; Fig. 6c). We determined the number of cells required per method and depth for adequate power (80%) by an asymptotic fit to the median powers. For the calculation of sequencing cost, we assumed 5€ per million raw reads, independent of method. Although UMI-based methods need paired-end sequencing, we assumed a 50 cycle sequencing kit is sufficient for all methods.

Data accession

The raw and analyzed data files can be obtained in GEO under accession number GSE75790.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We thank Rickard Sandberg for facilitating the Smart-seq2 sequencing. We thank Christopher Mulholland for assistance with FACS sorting, Dominik Alterauge for help establishing the Drop-seq method and Stefan Krebs and Helmut Blum from the LAFUGA platform for sequencing. We are grateful to Magali Soumillon and Tarjei Mikkelsen for providing the SCRB-seq protocol. This work was supported by the Deutsche Forschungsgemeinschaft (DFG) through LMUexcellent and the SFB1243 (Subproject A14/ A15) as well as a travel grant to CZ by the Boehringer Ingelheim Fonds.

Author's contributions

CZ and WE conceived the experiments. CZ prepared scRNA-seq libraries and analyzed the data. SP helped in data processing and power simulations. BV implemented the power simulation framework and estimated ERCC capture efficiencies. BR prepared Smart-seq2 scRNA-seq libraries. MS performed cell culture of mESC. WE and HL supervised the experimental work and IH provided guidance in data analysis. CZ, IH and WE wrote the manuscript. All authors read and approved the final manuscript.

References

1. Kim, H. D., Shay, T., O'Shea, E. K. & Regev, A. Transcriptional regulatory circuits: predicting numbers from alphabets. *Science* **325,** 429–432 (2009).
2. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489,** 57–74 (2012).
3. Sandberg, R. Entering the era of single-cell transcriptomics in biology and medicine. *Nat. Methods* **11,** 22–24 (2014).
4. Kolodziejczyk, A. A., Kim, J. K., Svensson, V., Marioni, J. C. & Teichmann, S. A. The technology and biology of single-cell RNA sequencing. *Mol. Cell* **58,** 610–620 (2015).
5. Eberwine, J., Sul, J.-Y., Bartfai, T. & Kim, J. The promise of single-cell sequencing. *Nat. Methods* **11,** 25–27 (2014).
6. Saliba, A.-E., Westermann, A. J., Gorski, S. A. & Vogel, J. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* **42,** 8845–8860 (2014).
7. Wu, A. R. *et al.* Quantitative assessment of single-cell RNA-sequencing methods. *Nat. Methods* **11,** 41–46 (2014).
8. Picelli, S. *et al.* Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10,** 1096–1098 (2013).
9. Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9,** 171–181 (2014).
10. Macaulay, I. C. *et al.* Single-Cell RNA-Sequencing Reveals a Continuous Spectrum of Differentiation in Hematopoietic Cells. *Cell Rep.* **14,** 966–977 (2016).
11. Macaulay, I. C. *et al.* G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat. Methods* **12,** 519–522 (2015).
12. Swiech, L. *et al.* In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat. Biotechnol.* **33,** 102–106 (2015).
13. Thomsen, E. R. *et al.* Fixed single-cell transcriptomic characterization of human radial glial diversity. *Nat. Methods* **13,** 87–93 (2016).

14. Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352,** 189–196 (2016).
15. Soumillon *et al.* Characterization of directed differentiation by high-throughput single-cell RNA-Seq. *bioRxiv* (2014). doi:10.1101/003236
16. 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).
17. Macosko, E. Z. *et al.* Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161,** 1202–1214 (2015).
18. Hashimshony, T., Wagner, F., Sher, N. & Yanai, I. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep.* **2,** 666–673 (2012).
19. Klein, A. M. *et al.* Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161,** 1187–1201 (2015).
20. Fu, G. K., Hu, J., Wang, P.-H. & Fodor, S. P. A. Counting individual DNA molecules by the stochastic attachment of diverse labels. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 9026–9031 (2011).
21. Kivioja, T. *et al.* Counting absolute numbers of molecules using unique molecular identifiers. *Nat. Methods* **9,** 72–74 (2012).
22. Islam, S. *et al.* Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat. Methods* **11,** 163–166 (2014).
23. Grün, D., Kester, L. & van Oudenaarden, A. Validation of noise models for single-cell transcriptomics. *Nat. Methods* **11,** 637–640 (2014).
24. 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).
25. Jiang, L. *et al.* Synthetic spike-in standards for RNA-seq experiments. *Genome Res.* **21,** 1543–1551 (2011).
26. Picelli, S. *et al.* Tn5 transposase and tagmentation procedures for massively-scaled sequencing projects. *Genome Res.* (2014). doi:10.1101/gr.177881.114

27. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29,** 15–21 (2013).
28. Petropoulos, S. *et al.* Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell* **0,** (2016).
29. Shalek, A. K. *et al.* Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* **510,** 363–369 (2014).
30. Nam, D. K. *et al.* Oligo(dT) primer generates a high frequency of truncated cDNAs through internal poly(A) priming during reverse transcription. *Proceedings of the National Academy of Sciences* **99,** 6152–6156 (2002).
31. Marinov, G. K. *et al.* From single-cell to cell-pool transcriptomes: Stochasticity in gene expression and RNA splicing. *Genome Res.* **24,** 496–510 (2014).
32. Grün, D. & van Oudenaarden, A. Design and Analysis of Single-Cell Sequencing Experiments. *Cell* **163,** 799–810 (2015).
33. Stegle, O., Teichmann, S. A. & Marioni, J. C. Computational and analytical challenges in single-cell transcriptomics. *Nat. Rev. Genet.* **16,** 133–145 (2015).
34. Risso, D., Ngai, J., Speed, T. P. & Dudoit, S. Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat. Biotechnol.* **32,** 896–902 (2014).
35. Brennecke, P. *et al.* Accounting for technical noise in single-cell RNA-seq experiments. *Nat. Methods* **10,** 1093–1095 (2013).
36. Zeisel, A. *et al.* Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* (2015). doi:10.1126/science.aaa1934
37. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43,** e47 (2015).
38. Pollen, A. A. *et al.* Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nat. Biotechnol.* (2014). doi:10.1038/nbt.2967
39. Parekh, S., Ziegenhain, C., Vieth, B., Enard, W. & Hellmann, I. The impact of amplification on differential expression analyses by RNA-seq. *Sci. Rep.* **6,** 25533

(2016).

40. Hashimshony, T. *et al.* CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol.* **17,** 77 (2016).
41. Renaud, G., Stenzel, U., Maricic, T., Wiebe, V. & Kelso, J. deML: robust demultiplexing of Illumina sequences using a likelihood-based approach. *Bioinformatics* **31,** 770–772 (2015).
42. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17,** 10–12 (2011).
43. Liao, Y., Smyth, G. K. & Shi, W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* **41,** e108 (2013).
44. Frazee, A. C., Jaffe, A. E., Langmead, B. & Leek, J. T. Polyester: simulating RNA-seq datasets with differential transcript expression. *Bioinformatics* **31,** 2778–2784 (2015).
45. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15,** R29 (2014).