SMART-Seq® Single Cell Kit: easy to use and outperforms homebrew method

Overview

Key features of the SMART-Seq Single Cell Kit:

- Robust full-length chemistry optimized for superior performance on single cells with very low RNA content
- Greater gene identification and capture of rare transcripts resulting from higher cDNA yield with fewer PCR cycles
- Streamlined workflow that is easily automated and miniaturized, saving time and resources
- Greater confidence for intra-laboratory comparisons due to strict quality standards for kit manufacturing
- Obtain the complete solution from a single provider with the SMART-Seq Single Cell PLUS Kit, includes the cDNA synthesis kit, a library preparation kit, and a single-use unique dual index (UDI) plate to generate Illumina-compatible sequencing libraries

Introduction

The SMART-Seq Single Cell Kit (SSsc) is powered by robust cDNA synthesis chemistry that provides unparalleled sensitivity and reproducibility for single-cell and nuclei sequencing applications. This kit generates high-quality, full-length cDNA directly from single cells known to have low RNA content (e.g., peripheral blood mononuclear cells, T cells, B cells, etc.). In addition, Takara Bio provides a complete library preparation solution with its SMART-Seq Single Cell PLUS Kit, which includes the SSsc cDNA synthesis kit, an enzymatic fragmentation library preparation kit, and a single-use UDI plate.

Recently, Rickard Sandberg's group (Karolinska Institute, Sweden) updated its homebrew Smart-seq2 (SS2) protocol (Hagemann-Jensen et al. 2020). This new homebrew protocol, known as "Smart-seq3" (SS3), is promoted to have a higher sensitivity than SS2. Additionally, it features unique molecular identifier (UMI) tags to help collapse PCR-derived errors, while still enabling the reconstruction of alternatively spliced and allele-specific mRNA isoforms *in silico*.

Here, we perform a head-to-head comparison of our well-validated SSsc kit versus the SS3 homebrew protocol to understand which method provides the best data from precious single-cell samples.

In order to compare performance and sensitivity in the most equitable manner, we isolated and generated cDNA from T cells—known to harbor low RNA content—with both protocols. The sequences from all samples were then analyzed using an identical UMI-agnostic bioinformatics pipeline since the SSsc kit does not contain UMI tags.

Results

SSsc yields more cDNA from single cells than SS3

The production of large quantities of full-length cDNA from low-input levels of native RNA is essential for detecting rare transcripts during single-cell RNA sequencing. It also permits better resolution of cell-to-cell expression variation and a more accurate assessment of differentially expressed genes. For T cells, SSsc generated nearly 3X as much cDNA, using fewer PCR cycles, indicating a greater capture of the starting material than SS3 (Figure 1).

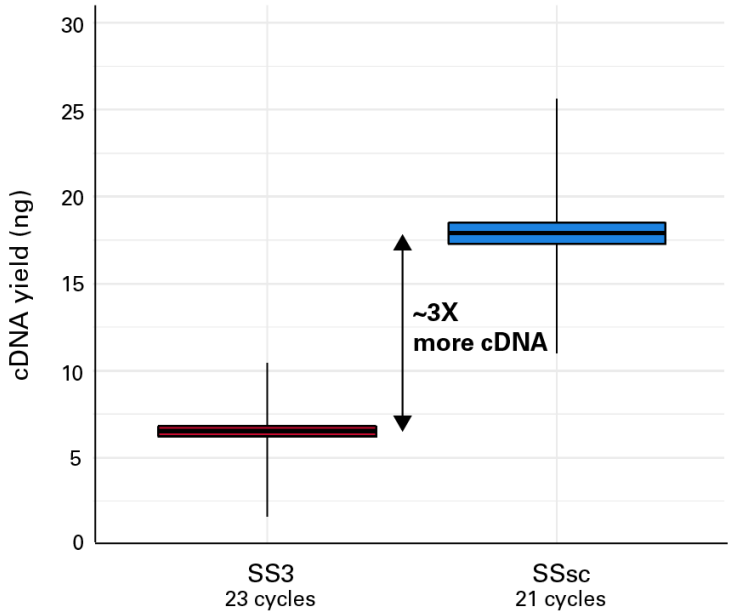

Figure 1. SSsc consistently yields a higher quantity of cDNA than SS3, with fewer PCR cycles. Single T cells isolated from PBMCs were processed with SSsc (40 cells) or the SS3 protocol (40 cells) using fewer PCR cycles. As described in the methods, RNA-seq libraries were generated, and the sequences were analyzed. The central lines in the boxplots indicate the mean cDNA yield, and the lower and upper bounds of each box show the mean-1 standard error of the mean (SEM), and the mean+1SEM, respectively. The whiskers indicate the highest and lowest cDNA value per sample.

SSsc captures more genes per cell than SS3

Sensitivity, which is measured by the number of genes mapped in a cDNA library for a given RNA input, is the metric most frequently used to assess the efficiency of single-cell RNA sequencing chemistries. High sensitivity can only be achieved by the production of diverse cDNA sequencing libraries, which in turn requires high cDNA yields during the initial phase of reverse transcription. The chemistry behind SSsc excels in all these points and results in higher cDNA yields, more diverse sequencing libraries, and a higher number of genes captured per cell than SS3 (Figure 2).

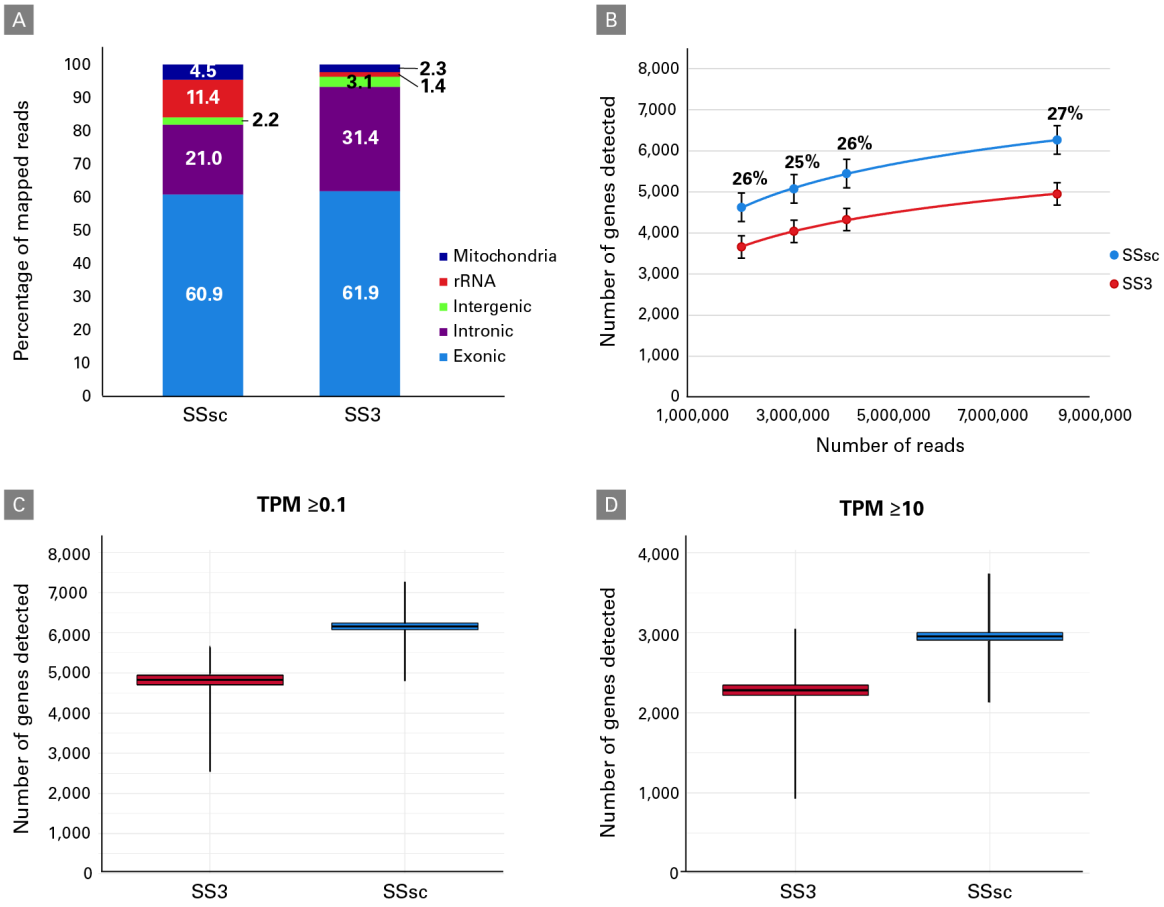

Figure 2. Detect a greater number of genes with similar read distributions at equivalent sequencing depths. **Panel A.** The read distribution was similar between the SSsc and SS3 chemistries. Both exhibited comparable proportions of reads mapped to exonic regions. SSsc does have a higher rRNA mapping percentage, but this is due to the increased sensitivity of the chemistry for low RNA content single cells. **Panel B.** About 25% more genes (TPM cut-off of ≥0.1) were detected in the cells processed with SSsc, regardless of the number of reads used for the analysis, saving sequencing costs while obtaining meaningful data. Error bars show the SEM of number of genes detected for a given number of reads. **Panel C.** Over 28% more low-expression genes (TPM ≥0.1) were identified in the cells processed with SSsc than SS3. Combined with the increased cDNA yield, there is a greater chance of these genes being detected from experiment to experiment. The central lines in the boxplots indicate the mean number of genes detected at TPM ≥0.1 , and the lower and upper bounds of each box show the mean-1SEM, and the mean+1SEM, respectively. The whiskers indicate the highest and lowest counts of genes detected at TPM ≥0.1 per sample. **Panel D.** Over 29% more high-expression genes (TPM ≥10) genes were detected in the cells processed with SSsc than SS3. The central lines in the boxplots indicate the mean number of genes detected at TPM ≥10, and the lower and upper bounds of each box show the mean-1SEM, and the mean+1SEM, respectively. The whiskers indicate the highest and lowest counts of genes detected at TPM ≥10 per sample.

Identify a higher number of unique genes at all expression levels

There is always a desire to identify as many genes as possible to fully understand the biology of precious or rare samples. In a head-to-head comparison, it is evident that SSsc enables you to capture a greater number of unique genes than SS3 (Figure 3). Higher-expressing genes are well-known and easy to identify. SSsc detects at least 38% more high-expressing genes (TPM ≥10) than SS3. On the opposite end of the spectrum, the low-expressing genes (TPM ≥0.1)—like chemokines, cytokines, or peptidoglycan recognition proteins—may be unique and interesting genes but are also the ones that are difficult to capture. As seen below, at the single-cell level, SSsc easily identifies a greater portion of unique genes compares to SS3. In addition, the detection of a greater number of genes overall indicates a more efficient capture of the original starting material from precious samples. This means you can be confident knowing that you can learn more about your sample using the SSsc kit.

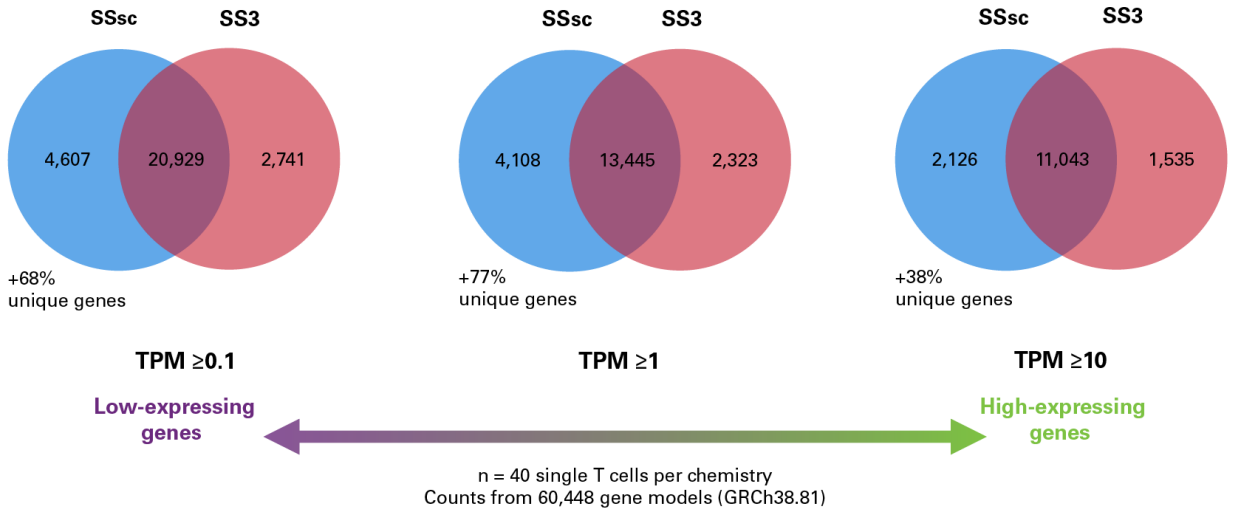

Figure 3. Uncover a greater number of unique genes, particularly low-expressing genes. An examination of unique versus shared genes at three TPM thresholds reveals a greater number of genes unique to SSsc at each threshold. (Genes counted if observed in ≥1 of the 40 samples per protocol). Upon a closer inspection of the additional genes detected by SSsc, these genes are unique to T cells and thus contribute to providing a better functional understanding of the sample. By contrast, the SS3 additional genes are related to broad categories of biological function.

Taking a closer look into the unique genes detected by SSsc, the list of low-expression genes (TPM ≥0.1) uniquely detected by SSsc in T cells reveals a 2.5X positive enrichment in gene ontology classes related to immunity, as evidenced by PANTHER overrepresentation tests (GO; www.geneontology.org) focusing on biological processes. Specifically, SSsc-processed T cells contained transcripts from 40 of 142 genes related to antimicrobial humoral response (GO:0019730; $p = 6.08e^{-7}$; median TPM = 0.61), 23 of 85 genes related to "killing cells of other organisms" (GO:0031640; $p = 1.92e^{-4}$; median TPM = 0.95), and 27 of 103 genes related to "antimicrobial humoral immune response mediated by antimicrobial peptide" (GO:0061844; $p = 1.10e^{-4}$; median TPM = 0.81). These transcripts were undetected in the T cells processed using SS3.

When taking a closer look at the additional genes detected by SS3, these genes were either related to broad categories of biological function (such as "cellular metabolic processes,"; GO:0044237; $p = 7.25e^{-12}$; median TPM = 0.31), or they belonged to extremely large gene families such as olfactory receptors ("detection of chemical stimulus involved in sensory perception of smell,"; GO:0050911; $p = 6.58e^{-5}$; median TPM = 0.68). Thus, there was no appreciable enrichment of immune-related genes detected by SS3 to provide a deeper understanding of the T cells.

Conclusion

Extracting meaningful biological information from the small amount of mRNA present in individual cells requires an RNA-seq preparation method with exceptional sensitivity and reproducibility. To address the need for improvement with extremely challenging samples, such as cells with very low RNA content or nuclei, we modified our core technology to create a new chemistry with higher sensitivity: the SMART-Seq Single Cell Kit. This kit continues to outperform all current commercial and noncommercial full-length methods, including the recently released Smart-seq3 protocol, which our kit outshines in terms of convenience, higher cDNA yield (which is extremely useful when dealing with difficult cells such as clinical research samples), and higher sensitivity (including detection of low-expression genes which are difficult to capture). In addition to the higher sensitivity, the additional genes detected by SSsc help to provide a deeper biological understanding of precious cells.

Methods

Cells were labeled with CD3-FITC antibody and FVD 660 (for distinguishing live from dead cells) prior to sorting using a BD FACS Jazz cell sorter to isolate single T-lymphocyte cells into 96-well plates. All cells were from the same single healthy donor sample of PBMCs. After sorting, cells were flash-frozen on dry ice and then stored at –80°C until ready to use. Libraries created with the SMART-Seq Single Cell Kit were processed at full volume per the user manual. For comparison, cells used to create libraries with the Smart-seq3 protocol were sorted in the same manner as above and processed at 2X the miniaturized volumes (1/4X) detailed in the authors' publicly available protocol (Hagemann-Jensen et al. 2020).

The total yield of each cDNA sample was measured by Quant-iT PicoGreen dsDNA detection assay (Invitrogen, Life Technologies). Concentrations of cDNA were calculated by plotting the standard curve of the known DNA standard and subtracting of background fluorescence using protocol-specific negative controls. A total of 21 PCR cycles were used to amplify SSsc cDNA, whereas 23 PCR cycles were used to amplify SS3 cDNA to produce enough material for the construction of sequencing libraries. Fragment length distributions of cDNA (post-PCR and SPRI bead purification at a 0.6:1 bead-to-sample ratio) for both protocols were inspected using the Agilent Bioanalyzer High Sensitivity DNA assay.

Sequencing libraries were generated using 125 pg of cDNA and the Nextera® XT DNA Library Preparation Kit (Illumina®) with a quarter of the recommended volume, as described in the SMART-Seq Single Cell Kit User Manual. Although Takara Bio now offers its own PLUS enzymatic fragmentation library prep kits (optimized to work with our SMART® cDNA synthesis kits), the Nextera XT chemistry was used here to enable a fair comparison with which to benchmark the SSsc and SS3 protocols. Libraries were sequenced on a NextSeq® 500 instrument using 2 x 150 bp paired-end reads, and analysis was performed using CLC Genomics Workbench (Qiagen; using genome build GRCh38.81 and Ensembl annotations).

All percentages shown—including the number of reads that map to introns, exons, or intergenic regions—are percentages of mapped reads in each library.

References

Hagemann-Jensen, M. et al. Single-cell RNA counting at allele and isoform resolution using Smart-seq3. *Nat. Biotechnol.* 1–7 (2020).

Products

Cat. #	Product	Size
634470	SMART-Seq Single Cell Kit	12 Rxns
634471	SMART-Seq Single Cell Kit	48 Rxns
634472	SMART-Seq Single Cell Kit	96 Rxns
R400750	SMART-Seq Single Cell PLUS Kit*	48 Rxns
R400751	SMART-Seq Single Cell PLUS Kit*	96 Rxns

*The PLUS kits provide an end-to-end solution, containing the cDNA synthesis kit, an enzymatic fragmentation library preparation kit, and single-use UDIs.

Takara Bio USA, Inc.
United States/Canada: +1.800.662.2566 • Asia Pacific: +1.650.919.7300 • Europe: +33.(0)1.3904.6880 • Japan: +81.(0)77.565.6999