Complete SMART-Seq cDNA synthesis and library preparation from single cells

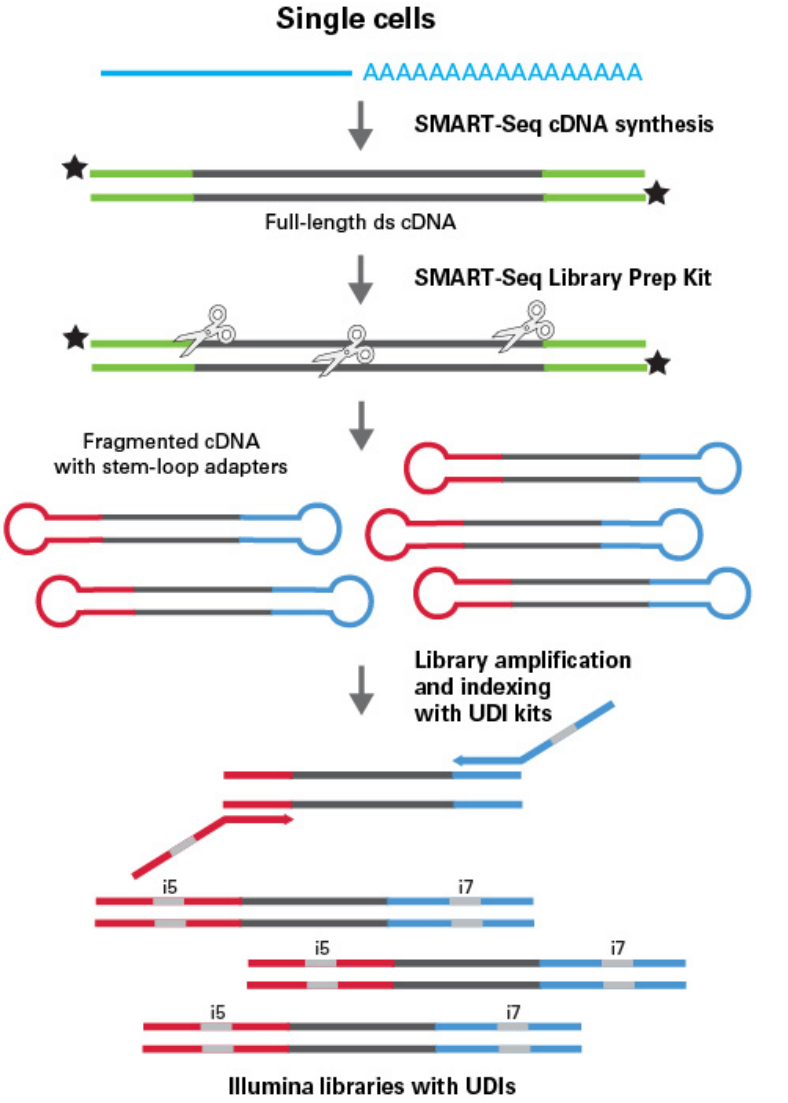
Introduction

The library input quantities (up to 310 µl of 2.5 nM pooled libraries) required to successfully load a high-throughput sequencer, such as the NovaSeq™ system, are much higher than the quantities (5 µl of 4 nM pooled libraries) required for a low-throughput sequencer like the NextSeq® system. For this reason, increased library yield is becoming a very desirable feature for library prep kits. The SMART-Seq Single Cell PLUS Kit (SSsc PLUS kit) is designed to synthesize high-quality, full-length cDNA directly from from single cells, especially those with low RNA content, and generate a high yield of Illumina-ready sequencing libraries. The kit includes the SMART-Seq Single Cell Kit for generating cDNA, the SMART-Seq Library Prep Kit (not sold separately) for fragmenting the cDNA and amplifying the library, and a unique dual index kit for allowing multiplexing of up to 96 samples (Figure 1, Panel A).

For cDNA generation, the SSsc PLUS kit incorporates our patented SMART (**S**witching **M**echanism **a**t 5′ end of **R**NA **T**emplate) technology. This technology relies on the template-switching activity of reverse transcriptase to enrich for full-length cDNAs and to add defined PCR adapters directly to both ends of the first-strand cDNA (Chenchik et al. 1998). This ensures that the final cDNA libraries contain the entire length of the mRNA, including the 5' end. This strategy maintains a true representation of the original mRNA transcripts, which is a critical factor for gene expression analysis. SMART technology offers unparalleled sensitivity and unbiased amplification of cDNA transcripts, and allows direct cDNA synthesis from intact cells.

For library preparation from the generated full-length cDNA, the SSsc PLUS kit also incorporates our patented library preparation technology. Unlike other library preparation kits, which are based on ligation of Y-adapters, our approach uses stem-loop adapters to construct high-quality libraries. The easy two-step workflow takes place in a single tube and is completed in about two hours. Sample mix-up, other handling errors, and sample loss are minimized since no intermediate purification steps or sample transfers are necessary (Figure 1, Panel B). The SSsc PLUS kit also includes unique dual index (UDI) primers for amplification of indexed NGS-ready libraries, allowing the multiplexing of up to 96 samples. Once purified and quantified, the resulting libraries are ready for use on Illumina NGS instruments using standard Illumina sequencing reagents and protocols.

A



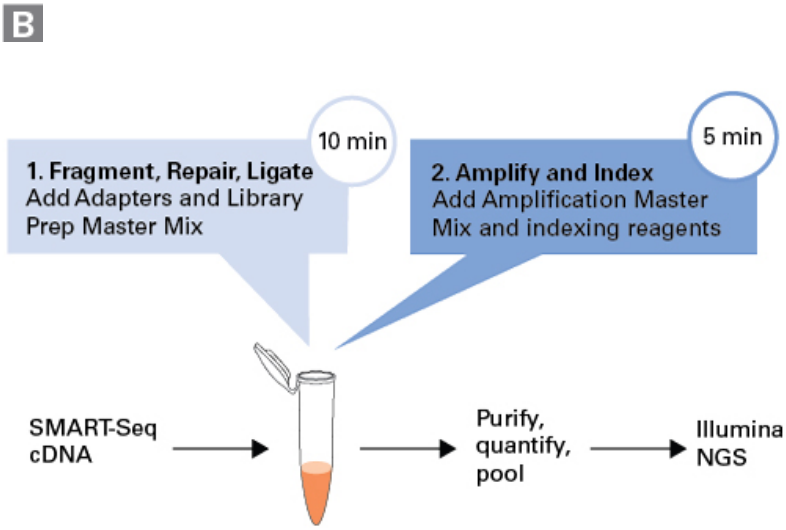


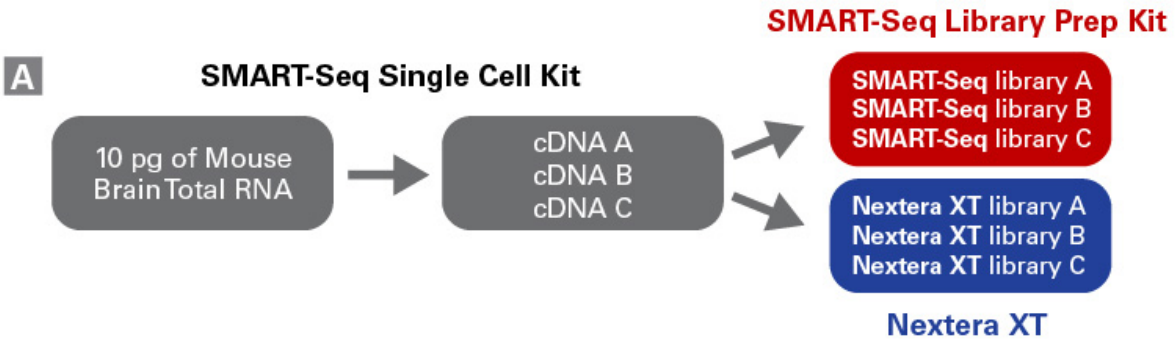
Figure 1. SMART-Seq Single Cell PLUS Kit protocol overview. Panel A. Double-stranded cDNA generated with SMART-Seq single-cell chemistry is enzymatically fragmented and ligated to stem-loop adapters in a single step. Libraries are then amplified and indexed, creating Illumina-compatible libraries with unique dual indexes. **Panel B.** The library preparation method features a new streamlined two-step protocol that generates sequencing-ready libraries with

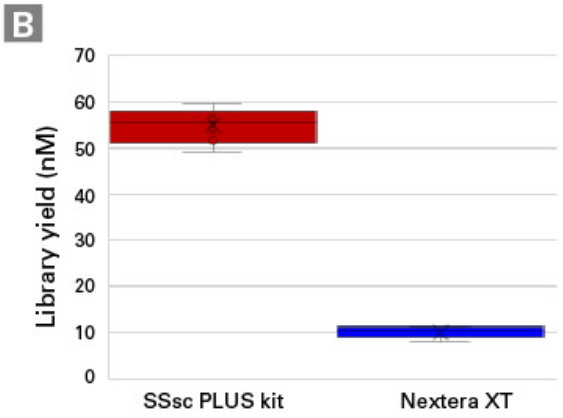
minimal hands-on time.

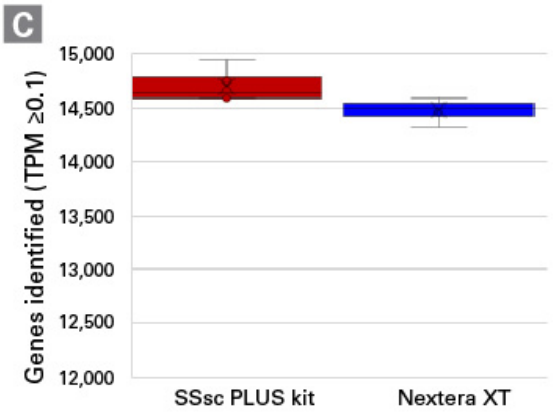
Results

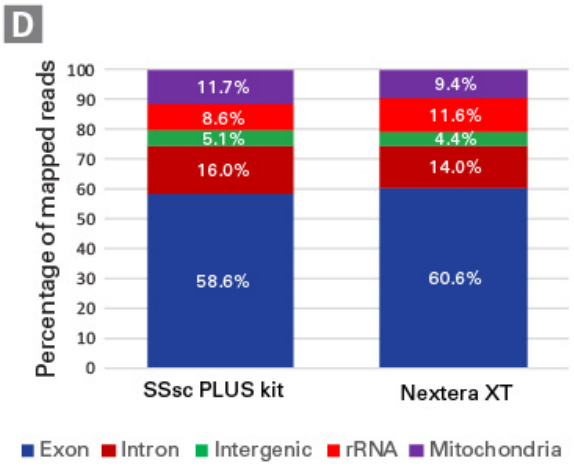
Greater sensitivity and reproducibility for cDNA synthesis

The SMART-Seq Single Cell PLUS Kit was designed to provide a streamlined workflow from single cells to Illumina-ready libraries. For this experiment, 10 pg of Mouse Brain Total RNA was input into the SMART-Seq Single Cell Kit to generate cDNA in triplicate. Libraries were then produced with the SMART-Seq Library Prep Kit portion of the SSsc PLUS kit or Nextera® XT, a gold-standard library preparation kit, and the performances were compared (Figure 2, Panel A). The library yields obtained with the SSsc PLUS kit were, on average, five times higher than yields obtained when using Nextera XT (Figure 2, Panel B). Next, the libraries were then sequenced on a NextSeq 500, and the number of genes identified with over 0.1 transcripts per million (TPM) was reported (Figure 2, Panel C). The SSsc PLUS kit enabled the detection of more genes than Nextera XT. The percentage of reads mapped to the exonic regions, intronic regions, intergenic regions, ribosomal RNA (rRNA), and mitochondria were identified (Figure 2, Panel D) and found to be comparable for the two methods. Finally, the Pearson and Spearman's correlations calculated from the TPM values were similarly high when comparing the triplicates prepared by each kit (Figure 2, Panel E), indicating high confidence in the reproducibility of the data.









E

Pearson

	A	B	C
A	1.00	0.97	0.97
B	0.97	1.00	0.97
C	0.97	0.97	1.00

SSsc PLUS kit

	A	B	C
A	1.00	0.97	0.97
B	0.97	1.00	0.97
C	0.97	0.97	1.00

Nextera XT

Spearman

	A	B	C
A	1.00	0.81	0.82
B	0.81	1.00	0.82
C	0.82	0.82	1.00

SSsc PLUS kit

	A	B	C
A	1.00	0.81	0.81
B	0.81	1.00	0.81
C	0.81	0.81	1.00

Nextera XT

Figure 2. Library preparation with the complete SMART-Seq Single Cell PLUS Kit increases library yield and genes identified compared to Nextera XT. Panel A. The SMART-Seq Single Cell Kit was used to produce cDNA, in triplicate, from 10 pg of Mouse Brain Total RNA. Illumina-compatible libraries were then generated from 1 ng or 125 pg of SMART-Seq cDNA using either the SMART-Seq Library Prep Kit portion of the SSsc PLUS kit or Nextera XT, respectively. Libraries were sequenced on a NextSeq 500. The reads were normalized to 4M paired-end reads and analyzed as described in the methods. **Panel B.** Library yields obtained with the SSsc PLUS kit are higher than yields produced with Nextera XT. **Panel C.** The SSsc PLUS kit also identified more genes compared to Nextera XT. **Panel D.** As expected, the distribution of the mapped reads is similar between the two library preparation methods. **Panel E.** Pearson and Spearman's correlations were calculated from the TPM values obtained from triplicate cDNAs (A, B, and C) processed with either the SSsc PLUS or Nextera XT kit.

Switch with confidence

To evaluate any bias introduced by the library preparation step, the same SSsc-generated cDNA was processed in duplicate with the SMART-Seq Library Prep Kit portion of the SSsc PLUS kit or Nextera XT (Figure 3, Panel A). The TPM values were plotted and used to calculate correlations between the replicates. Looking at the SSsc PLUS libraries, the correlations obtained when using the same starting cDNA (P: 0.999 & S: 0.92 for cDNA A1 versus cDNA A2; Figure 3, Panel B, left) were found to be higher than those obtained when using different input cDNAs (P: 0.97 & S: 0.81 for cDNA A versus cDNA B; Figure 2, Panel E). We, therefore, concluded that the main source of variation in the TPM values

was coming from the cDNA generation.

We also assessed the different genes identified with SSsc PLUS versus Nextera XT. The Spearman's correlation for these DNA replicate processed with the two different kits (S: 0.91 for SSsc PLUS cDNA A1 vs. Nextera XT cDNA A1; Figure 3, Panel B, right) was significantly higher than correlations for different cDNA replicates prepared with the same kit (0.81/0.82 [SSsc PLUS] & 0.81 [Nextera XT]; Figure 2, Panel E). This high degree of correlation demonstrates the similarity in gene-identification performance between the two library prep technologies. Additionally, as shown in Figure 3, Panel C, the number of genes identified in the replicates of the SSsc PLUS-generated libraries (13,226 genes) was higher compared to those found with the Nextera XT replicates (12,975 genes). Most of the genes identified that did not overlap between replicates showed low TPM values as depicted by the dots overlying the X- and Y-axis of the plots (Figure 3, Panel B). The same 13,059 genes were detected by both kits at similar TPM values as represented by the correlation plot (Panel B) and the Venn diagram (Panel C). Therefore, existing Nextera XT users should see minimal effects on their data when using the new SMART-Seq Single Cell PLUS Kit.

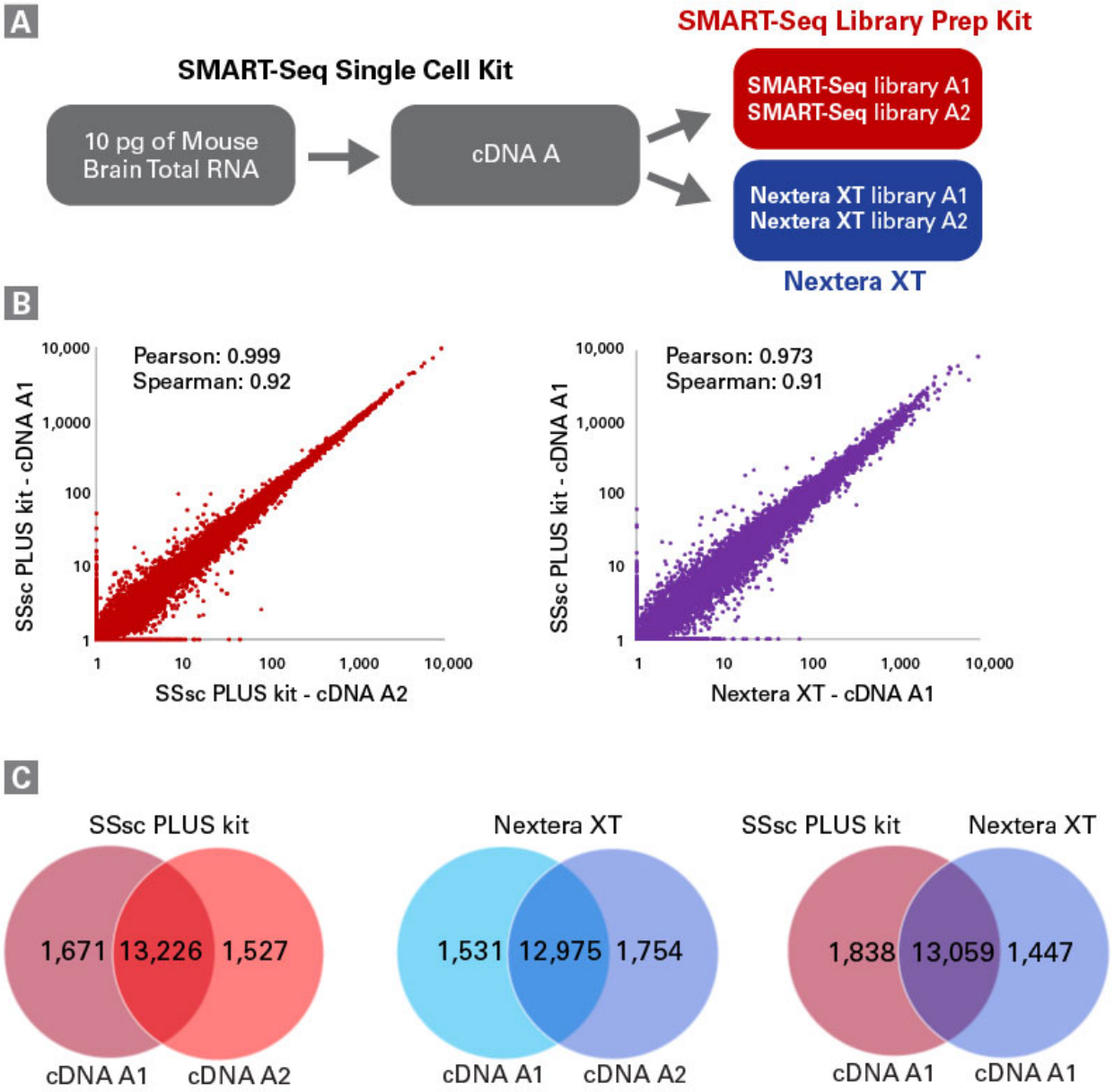


Figure 3. Comparable data between SMART-Seq library prep and Nextera XT kits enable existing users to easily make the switch. Panel A. Libraries were prepared in duplicate from the same SSsc cDNA using either the SMART-Seq Library Prep Kit portion of the SSsc PLUS kit or Nextera XT. **Panel B.** Correlation plots between libraries made with SMART-Seq library prep and Nextera XT kits. Pearson and Spearman's correlations were calculated from the TPM values obtained. Scatter plots showed the TPM values from all genes with a log10+1 scale. **Panel C.** Venn diagram showing the number of genes identified from processed cDNA, in duplicate, with the SSsc PLUS kit, Nextera XT kit, or between both chemistries.

SSsc PLUS outperforms the NEBNext Single Cell Kit

Like SSsc PLUS, the NEBNext Single Cell/Low Input RNA Library Prep Kit for Illumina (NEBNext kit) is an all-in-one method for generating sequencing-ready Illumina-compatible RNA libraries from single cells. Both methods reduce the need to purchase multiple kits and provide a single protocol from cDNA synthesis to sequence-ready solutions. Both the SSsc PLUS and NEBNext kits generate cDNA based on RT with a poly-A targeting primer, template switching for cDNA generation, and PCR amplification for high cDNA yield. The cDNA is then directly used in the accompanying library preparation steps to generate Illumina compatible libraries. Each method's sequencing library preparation includes processes for fragmentation, adapter addition, and amplification.

Although both kits are very similar with regards to library preparation, the SSsc PLUS technique provides a notably simpler workflow. While the sequencing-ready, Illumina-compatible NEBNext kit requires two separate steps for fragmentation and adapter ligation, thereby doubling the handling for each sample, we have optimized these two critical reactions into a single step. Additionally, the SSsc protocol reduces the required number of bead cleanups by half relative to the NEBNext protocol (two versus four).With fewer rounds of cleanup, the SSsc PLUS chemistry by comparison reduces handling time, risk of contamination, and potential sample mix-up. The SSsc PLUS protocol also provides detailed guidance for a pooling strategy that enables excellent normalization intended to ensure confidence when comparing across samples and controls.

NEBNext is promoted as a robust and highly sensitive method to generate high-quality sequencing data from single cells, resulting in unmatched detection of low-abundance transcripts. To test the quality of the sequencing data generated from single cells fully processed with the SSsc PLUS and NEBNext protocols, we performed unbiased single-cell analysis using each kit's protocol according to the manufacturer's instructions in parallel (Figure 4).

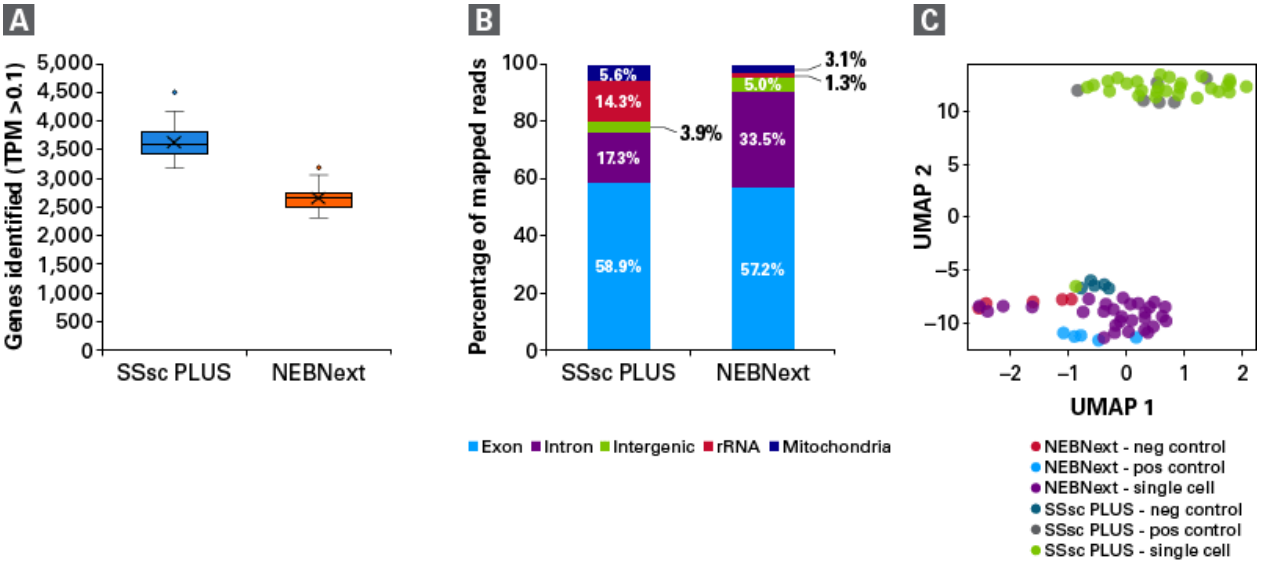


Figure 4. The SMART-Seq Single Cell PLUS Kit demonstrates greater sensitivity and discrimination between control and single-cell samples than NEBNext, while displaying comparable read mapping. Panel A. The SSsc PLUS method resulted in the identification of 40% more genes (SSsc median = 3,610 versus NEBNext median = 2,659), allowing for a greater depth of understanding of the sample. The box plot denotes the interquartile range (IQR), i.e., the 25th and 75th quartiles, and the whiskers are 1.5 x IQR from the median value and represent the extremes of the data. **Panel B.** The exon mapping is comparable between both chemistries, while the intron mapping is higher for NEBNext. As consistently seen in SSsc data sets, SSsc PLUS does have a slightly higher percentage of rRNA mapping; this is due to the increased sensitivity of the SSsc cDNA chemistry to optimize capture of information from single cells. It should be noted that despite the slightly higher rRNA, the number of genes identified is still significantly higher for SSsc PLUS. **Panel C.** By UMAP analysis, there is a clear separation between the negative controls and the positive controls/single-cell samples for SSsc PLUS. As expected, the positive controls (grey) and the single-cell samples (green) cluster together for the SSsc PLUS while the negative controls (dusty blue) cluster together near the bottom of the chart. Interestingly, the SSsc PLUS negative controls and all of the NEBNext samples cluster together: negative (red), positive (light blue), and the single cells (purple). The clustering of the SSsc single cell and positive controls and significantly distinct discrimination of the negative controls provides confidence in the data generated with SSsc PLUS. These analyses indicate that experimental noise, even when working with such small amounts of starting material, will not affect confidence in the biological import of the results.

Conclusions

Having a sensitive and reproducible end-to-end approach is crucial to efficiently and easily analyze samples and extract meaningful biological information from the sequencing data. The SMART-Seq Single Cell PLUS Kit combines the proven SMART-Seq cDNA generation approach and

user-friendly library preparation chemistry to generate Illumina-compatible libraries ready for sequencing. By using the SMART-Seq Single Cell PLUS Kit, researchers can produce about 5X higher library yields than traditional methods. An increase in the number of genes identified is also observed, which is extremely useful when dealing with cells containing low RNA levels or difficult cells, including clinical research samples. In addition, the SMART-Seq Single Cell PLUS Kit continues to have the highest sensitivity compared to similar commercial kits, such as the NEBNext kit.

Methods

cDNA and Illumina libraries were generated from 10 pg of Mouse Brain Total RNA using the SMART-Seq Single Cell PLUS Kit. For comparison, sequencing libraries were generated using 125 pg of cDNA and the Nextera XT DNA Library Preparation Kit (Illumina) with a quarter of the recommended volumes, as described in SMART-Seq Single Cell Kit User Manual. The cDNA and sequencing libraries were quantified using Bioanalyzer HS reagents and Qubit, respectively.

Libraries were sequenced on a NextSeq 500 instrument using 2 x 75 bp paired-end reads, and analysis was performed using CLC Genomics Workbench. Reads were downsampled to 4M reads and mapped to the mouse [GRCm38] genome with 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. TPMs were reported using CLC.

T cells were co-labeled with anti-CD3-FITC and 7-AAD, to distinguish live from dead cells. Single cells were sorted using a BD FACSJazz Cell Sorter into a 96-well plate pre-prepared with either the SSsc PLUS or the NEBNext kit lysis buffer, respectively. After sorting, cells were flash-frozen on dry ice, then stored at –80°C until ready to use. The cDNA libraries were processed at full volume, per the SSsc PLUS or NEBNext user manuals. The NEBNext kit was tested twice, with a standard 20 PCR cycles and the 25 cycles NEB recommends to produce sufficient cDNA for library preparation. The cDNA generated with the 25 cycles was used as input for sequencing library production. The positive controls were five cells, and the negative controls were no cells, sorted into the sample plates used for the single-cell tests for both the SSsc PLUS and the NEBNext single cell kits. For the SSsc PLUS cDNA libraries, 1 ng of cDNA was processed into sequencing libraries per the protocol recommendations. The NEBNext sequencing libraries were processed with 20 ng of NEBNext cDNA (at the upper level of the recommended range, as lower inputs failed to produce sufficient library for sequencing). After making each sequencing library for the SSsc PLUS and NEBNext samples, the libraries were pooled together at equal input amounts and purified with AMPure beads. The combined library was then diluted to 4 nM. Of the total library, 1.7 pmol of the total library was used for a NextSeq® 500 run using a NextSeq® 550 System Mid-Output Kit (150 Cycles).

References

Chenchik, A., Zhu, Y., Diatchenko, L., Li., R., Hill, J. & Siebert, P. Generation and use of high-quality cDNA from small amounts of total RNA by SMART PCR. In *RT-PCR Methods for Gene Cloning and Analysis*. Eds. Siebert, P. & Larrick, J. (BioTechniques Books, MA), pp. 305–319 (1998).

Related Products

Cat. #	Product	Size	Price	License	Quantity	Details
R400750	SMART-Seq® Single Cell PLUS Kit	48 Rxns	USD $3696.00			

The SMART-Seq Single Cell PLUS Kit is a complete kit designed to first generate high-quality cDNA from single cells and then high-quality Illumina sequencing-ready libraries. Indexes are added using a unique dual index kit (Cat. # R400744 or R400745). This kit supports up to 48 reactions.

Documents | Components

R400751	SMART-Seq® Single Cell PLUS Kit	96 Rxns	USD $6816.00			

Add to Cart

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