Clontech Laboratories, Inc.
A Takara Bio Company

SMART-Seq v4 Ultra Low Input RNA Kit for the Fluidigm C1 System: improved chemistry for single cell transcriptome studies

Magnolia Bostick, Sangwon Lee, Yevgeniy Gindin, Andrew Farmer[1] [1] Corresponding Author: Andrew_Farmer@clontech.com

Abstract

In single cell transcriptome studies, it is critical to obtain high-quality cDNA libraries from individual cells that represent the original *in vivo* mRNAs as closely as possible. To achieve this, we improved the existing commercially available mRNA-Seq chemistry for single cell transcriptome studies (the SMARTer® Ultra® Low Input RNA Kit for the Fluidigm® C1™ System), which is incorporated in the "mRNA Seq" script in the Fluidigm C1 system.

The new chemistry employs locked nucleic acid (LNA) technology integrated in our proven SMART® technology, which greatly enhances the first strand cDNA synthesis reaction. The new chemistry also utilizes efficient SeqAmp™ polymerase, which excels even in GC- and AT-rich regions. All procedures (including lysis, RT, and all PCR steps) are designed to be performed in the Fluidigm C1 System using Open App IFC.

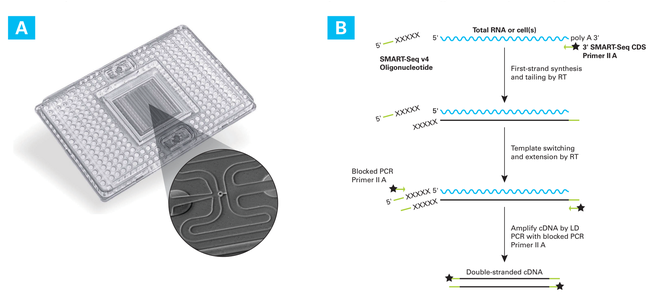
In this poster, we present data derived from K562 cells (human leukemia cell line). The results show a higher yield of cDNA, a larger number of genes detected, and an improvement in high-GC gene detection compared to the existing chemistry. The new chemistry also produces higher consistency among replicates. This is important because the reduced technical variations will increase the likelihood of discovering true biological variations. Altogether, these results indicate that the new chemistry can robustly produce high-quality and highly reproducible cDNA from single cells for meaningful transcriptome analysis.

Introduction and Methods

Gaining the ability to identify and quantify the mRNA from a single cell has been a substantial benefit to many scientific fields, especially those where homogeneous populations are elusive, such as cancer research, developmental biology, neurobiology, and immunology. Here we introduce the "SMART-Seq v4" C1 script for single-cell RNA-seq employing the Open App IFC and new chemistry (SMART-Seq® v4 Ultra Low Input RNA Kit for the Fluidigm C1 System), which has remarkable sensitivity and precision.

In this poster, we will refer to the new chemistry used for the SMART-Seq v4 script as "UL-v4" and the previous chemistry used for the mRNA Seq script as "UL-v1".

cDNA synthesis with new UL-v4 chemistry on the Fluidigm C1 Open App IFC platform



Construction of cDNA libraries from single cells using the SMART-Seq v4 Ultra Low Input RNA Kit for the Fluidigm C1 System (UL-v4 chemistry). Panel A. In the experiments described here, K562 cells (human leukemia cell line) were captured, lysed and processed for cDNA library construction on the Fluidigm C1 platform. The C1 Single-Cell Auto Prep IFC for Open App (for medium-sized cells) was used as described here for the SMART-Seq v4 script (UL-v4 chemistry), while the C1 Single-Cell Auto Prep IFC (for medium-sized cells) was used according to the Fluidigm Protocol (1) for experiments performed with the existing mRNA Seq script and UL-v1 chemistry. For all experiments, the cells were cultured to a log phase, mixed with suspension reagent at a 6:4 ratio for the best buoyancy, and 500–700 cells were loaded into an IFC. Every capture site was examined using microscopy to determine if a single cell was captured. **Panel B.** Improved SMART technology (UL-v4) for cDNA synthesis. SMART technology enables us to synthesize cDNA from total RNA using a poly dT primer and the SMART-Seq v4 Oligonucleotide, and eliminates ligation steps by using sequencing adapters in downstream steps (Ramsköld *et al.*, 2012). Our first generation SMART technology, UL-v1, is currently available on the market (the SMARTer Ultra Low RNA Kit for the Fluidigm C1 System) and used in the "mRNA-Seq" script built into the Fluidigm C1. The UL-v4 chemistry improves upon UL-v1 chemistry and the SMART-Seq2 method (Picelli *et al.*, 2013) and incorporates modified locked nucleic acid (LNA) technology and SeqAmp DNA Polymerase. The black star on the 3′ SMART-Seq CDS Primer II A and PCR Primer II A indicates a chemical block on the 5′ end of the oligonucleotide.

Conclusions

The SMART-Seq v4 Ultra Low Input RNA Kit for the Fluidigm C1 System (UL-v4) provides an improved single-cell, full-length mRNA-seq method. The new SMART-Seq v4 script with UL-v4 chemistry demonstrates great sensitivity and delivers high quality, robust and reproducible transcriptomic data.

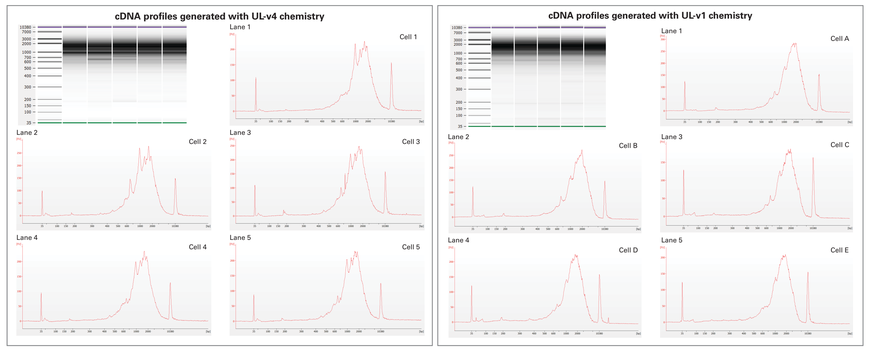
- **High sensitivity**—the new chemistry (UL-v4) produced more cDNA, detected significantly more genes, and achieved a significantly higher exon to intron mapped ratio than the UL-v1 chemistry (Figure 3).
- **Improved detection of GC-rich genes**—more genes with high GC content were identified in the cDNA libraries constructed with UL-v4 relative to the cDNA libraries constructed with UL-v1 (Figure 4).
- **Reduced technical variability**—transcriptomic data generated by UL-v4 were more consistent between replicates than the data generated by UL-v1 (Figure 5). The protocol reduced technical variability, which may increase the opportunity to detect biological variabilities.

1 Improved UL-v4 chemistry and SMART-Seq v4 script for full-length cDNA synthesis on the Fluidigm C1 system

	UL-v4 (SMART-Seq v4 script)	UL-v1 (mRNA Seq script)
Lysis Conditions	3 min 72°C 10 min 4°C 1 min 25°C	
Lysis Reagents	10X Lysis Buffer	Dilution Buffer
RT Conditions	90 min 42°C 10 min 70°C	
RT Reagents	SMART-Seq v4 Oligonucleotide	SMARTer II A Oligonucleotide
PCR Conditions	1 cycle: 1 min 98°C 5 cycles: 20 sec 98°C, 4 min 59°C, 6 min 68°C 9 cycles: 20 sec 95°C, 30 sec 65°C, 6 min 68°C 7 cycles: 30 sec 95°C, 30 sec 65°C, 7 min 68°C 1 cycle: 10 min 72°C	1 cycle: 1 min 95°C 5 cycles: 20 sec 95°C, 4 min 58°C, 6 min 68°C 9 cycles: 20 sec 95°C, 30 sec 64°C, 6 min 68°C 7 cycles: 30 sec 95°C, 30 sec 64°C, 7 min 68°C 1 cycle: 10 min 72°C
PCR Reagents	SeqAmp DNA Polymerase	Advantage® 2 Polymerase

Comparison of sample preparation conditions between the SMART-Seq v4 script (UL-v4 chemistry) and mRNA Seq script (UL-v1 chemistry) for cDNA library construction from single cells on the Fluidigm C1 platform. To employ the UL-v4 chemistry on C1 IFCs, a SMART-Seq v4 script was written using the C1 Script Builder (v.2.0.10). The existing Fluidigm mRNA Seq script that is built into the C1 system was used for the UL-v1 experiments. The lysis and RT thermal conditions are the same for the UL-v4 and UL-v1 protocols; however a variety of reagents, enzymes (for lysis, RT, and PCR) and the PCR conditions vary between the two protocols. Key differences are highlighted here: for example, the SMART-Seq v4 script uses SeqAmp polymerase and conducts PCR at a higher temperature.

2 Similar Bioanalyzer cDNA profiles were produced by the SMART-Seq v4 script (UL-v4) and the mRNA Seq script (UL-v1)



Bioanalyzer electropherograms of the cDNA harvested from the C1 IFCs. One µl of the harvested cDNA (from 3–3.5 µl cDNA diluted in 10 µl C1 dilution reagent) was analyzed using an Agilent Bioanalyzer with Agilent's High Sensitivity DNA chip and 2100 Expert software. Most cDNAs ranged between 400 bp and 7,000 bp in length with peaks at 1,650–1,900 bp. The cDNA electropherograms from the UL-v4 and UL-v1 libraries were similar.

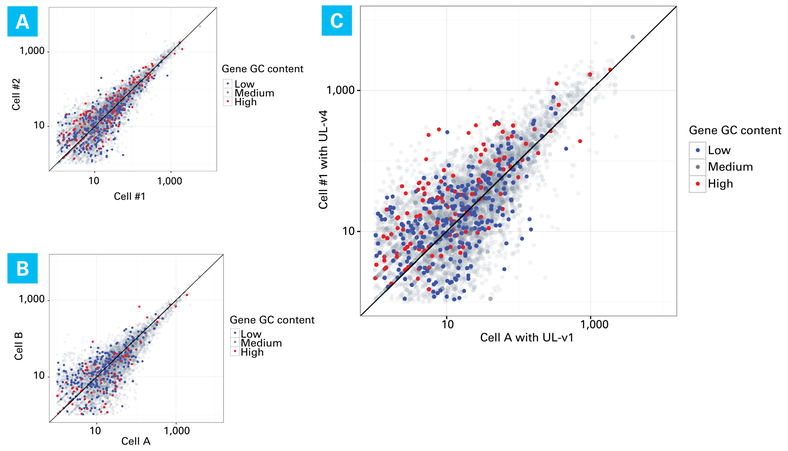
3 High sensitivity of the UL-v4 chemistry and SMART-Seq v4 script

Sequencing Alignment Metrics			
Chemistry used	UL-v4		UL-v1
Replicate	T1	T2	
Number of single cells captured	65	75	81
Number of single cells that generated >6.5 ng cDNA	65	74	80
Average cDNA yield (ng per cell)	38 (±5.4)	31 (±5.6)	23 (±9.4)
Number of single cells sequenced	11		
Number of sequence reads per library	2M*		
Proportion of reads (%)			
Mapped to human genome	98 (±0.3)	98 (±0.2)	97 (±0.5)
Mapped to rRNA	1.6 (±0.4)	0.9 (±0.2)	0.5 (±0.1)
Mapped to mitochondria	11 (±1.5)	7 (±1.6)	6 (±0.9)
Mapped to NCBI-RefSeq	86 (±1.6)	90 (±1.7)	91 (±1.0)
Exonic	89 (±2.5)	90 (±2.2)	71 (±8.2)
Intronic	8 (±2.1)	7 (±1.8)	22 (±7.1)
Intergenic	3 (±0.5)	3 (±0.5)	7 (±2.5)
Number of genes identified with >0.1 RPKM	8,419 (±297)	7,669 (±538)	6,644 (±606)

The UL-v4 method generated more cDNA and detected more genes than the UL-v1 method. Most single cells (96–100%) successfully produced good amounts (>6.5 ng) of cDNA. The UL-v4 method significantly increased cDNA yield, with results 38–68% higher than the UL-v1 method. Eleven cDNA libraries per experiment were sequenced using the Illumina® Nextera® XT DNA Library Preparation Kit and a MiSeq® sequencer. The FASTQ files were analyzed using CLC Genomics Workbench (v.8.5.1, Qiagen Aarhus). The paired sequence reads were trimmed using quality score (< 0.05) and primer and adapter sequences. In mapping, the paired end sequences that mapped to rRNA were removed first. Next, the sequences that mapped to mitochondria were excluded from the data set. The remaining sequences were tested for mapping against the reference genome (GRCH37). The UL-v4 sequencing data showed a higher percentage of reads mapped to exons and a much lower percentage of reads mapped to introns compared to UL-v1. UL-v4 also detected a higher number of genes compared to UL-v1 (a 15–27% increase).

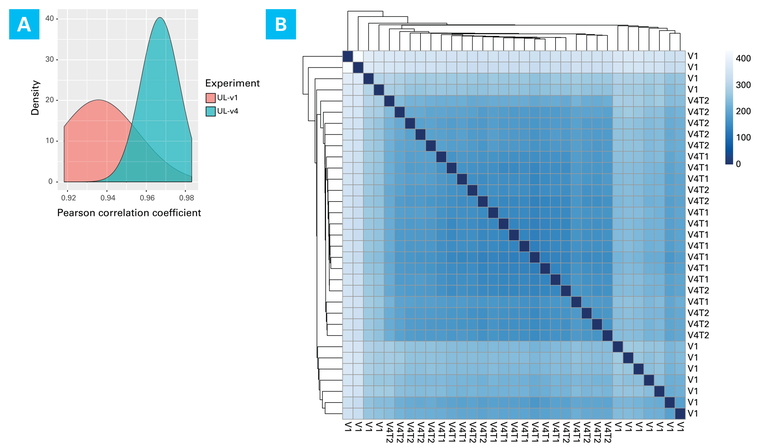
***Sequence reads were resampled to 2 million reads per library.** Two out of eleven UL-v1 libraries had lower numbers of reads and were analyzed as is (1.12 million and 0.54 million reads, respectively).

4 High sensitivity of the UL-v4 chemistry and SMART-Seq v4 script for GC-rich genes



Comparing expression levels by gene GC content for two cDNA synthesis protocols. Libraries made from single K562 cells using either the UL-v4 protocol or the UL-v1 protocol were compared. Genes were binned by GC content (Low: 0–36% GC, Medium: 37–62% GC, High: 63–100% GC) and correlation plots were used to evaluate the two protocols. The RPKM values were very reproducible for pairs of cells using the UL-v4 chemistry (**Panel A**) or the UL-v1 chemistry (**Panel B**), with even distribution of genes across all three GC content categories. **Panel C.** When the two protocols were compared, genes with high GC content (shown in red) showed higher expression with the UL-v4 protocol while genes with medium or low GC content (shown in gray and blue, respectively) showed an even distribution.

5 Improved precision of the UL-v4 chemistry and SMART-Seq v4 script



Reduced variability among cDNA libraries synthesized using the UL-v4 protocol. Sequencing data were generated with the UL-v4 method (11 libraries x 2 replicate experiments) and the UL-v1 method (11 libraries) as described in Figure 3. The FASTQ files were aligned against the human reference genome (Gencode release GRCh38) using STAR v. 2.5 (Dobin *et al.*, 2013) with default options. Gene counts generated by STAR were subjected to regularized log transformation (DESeq2; v.1.10.1; Love *et al.*, 2014). These transformed gene expression values were then used in the gene expression analyses. The analysis results were visualized using R. **Panel A.** The Pearson correlation coefficients were higher and tighter in the cDNA libraries generated with UL-v4. **Panel B.** The hierarchical clustering heatmap demonstrates that the Euclidean distances between the UL-v4 cDNA libraries were smaller than the distances between the UL-v1 cDNA libraries. This indicates that the cDNA libraries generated with UL-v4 were more similar to each other than the cDNA libraries generated with UL-v1.

References

1. Using C1 to Generate Single-Cell cDNA Libraries for mRNA Sequencing. Protocol PN 100-7168, Fluidigm Corporation.
2. Ramsköld, D., Luo, S., Wang, Y.-C., Li, R., Deng, Q., Faridani, O. R., Daniels, G. A., Khrebtukova, I., Loring, J. F, Laurent, L. C., Schroth, G. P. & Sandberg, R. (2012) Full-length mRNA-seq from single-cell levels of RNA and individual circulating tumor cells. *Nat. Biotechnol.* **30**(8):777–782.
3. Picelli, S., Björklund, Å. K., Faridani, O. R., Sagasser, S., Winberg, G., Sandberg, R. (2013) Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Meth.* **10**(11):1096–1098.
4. Dobin, A., Davis, C. A. , Schlesinger, F, Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. & Gingeras, T. R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**(1):15–21.
5. Love, M. I., Huber W. & Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**(12):550.

To download this poster visit:
www.clontech.com/SMART-Seq-for-Fluidigm-Poster