

# illumina® Bio-Rad® SureCell™ WTA 3' Library Prep Kit for the ddSEQ™ System

Gain unprecedented insight into gene expression with a sensitive, scalable, and cost-effective high-throughput next-generation sequencing platform

## Highlights

- Sensitive and Unbiased Characterization of Transcriptional Signatures**  
 Highly sensitive and reproducible gene detection in single cells
- Comprehensive Single-Cell RNA Sequencing Workflow**  
 Fully supported workflow developed in collaboration by technology innovators
- Powerful Next-Generation Sequencing Integrated with Simple Data Analysis**  
 Proven Illumina sequencing combined with streamlined, user-friendly analysis software

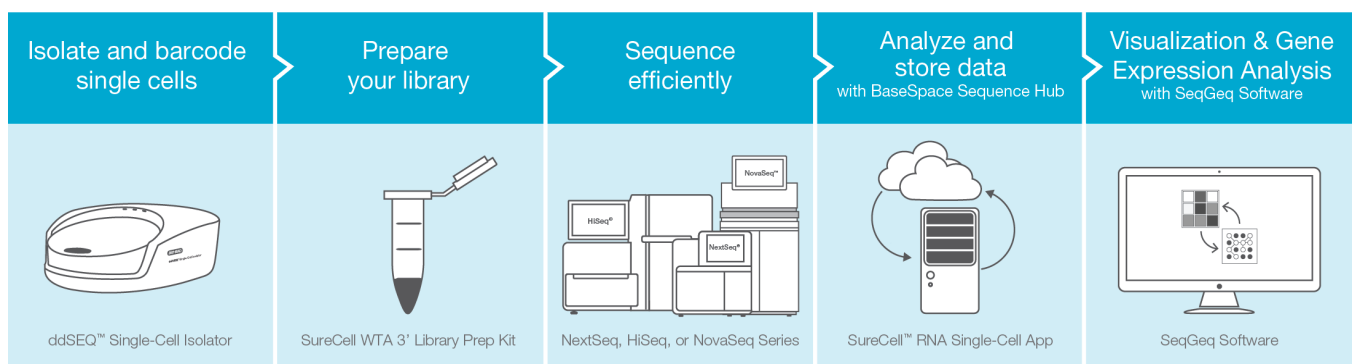
To deliver on the promise of single-cell biology, the Illumina Bio-Rad Single-Cell Sequencing Solution combines the highly innovative Bio-Rad Droplet Digital™ technology(3) with Illumina NGS library preparation, sequencing, and analysis technologies. This new platform provides a comprehensive, user-friendly workflow for single-cell RNA-Seq (Figure 1) that enables controlled experiments with multiple samples, treatment conditions, and time points. Built and supported in collaboration between technology leaders, the Illumina Bio-Rad Single-Cell Sequencing Solution enables transcriptome analysis of hundreds to tens of thousands of single cells in a single experiment. This scalable, robust single-cell NGS sample prep methodology enables more researchers to apply the sensitivity and precision of RNA-Seq to questions that can only be answered by interrogating individual cells.

## Introduction

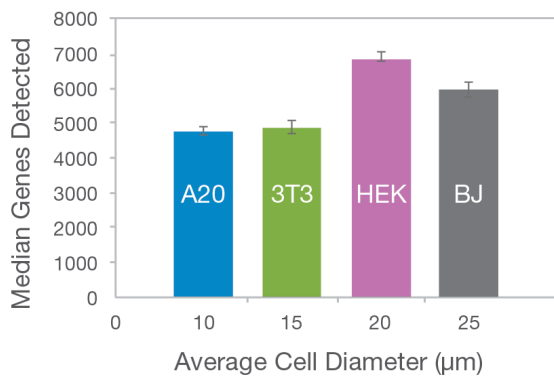
Complex biological systems are determined by the coordinated functions of individual cells. Conventional technologies providing bulk transcriptome data are unable to reveal the transcriptional heterogeneity that drives this complexity (1). Single-cell RNA sequencing (RNA-Seq) enables in-depth gene expression analysis, providing insight into cell function, disease progression, and therapeutic efficacy (2). However, generating thousands of single-cell next-generation sequencing (NGS) libraries in an affordable, high-throughput, and user-friendly manner remains challenging (2).

## Scalable, Flexible Single-Cell Isolation

The Bio-Rad ddSEQ Single-Cell Isolator (SCI) encapsulates and partitions single cells into subnanoliter droplets on a disposable cartridge. Each cartridge can accommodate multiple samples, and multiple cartridges can be processed in parallel, allowing for isolation of hundreds to tens of thousands of cells per day. Cell lysis and cell barcoding occur inside individual droplets, enabling tracking of individual cells throughout the workflow. This droplet-based method is agnostic to mammalian cell size, providing unbiased profiling of diverse cell populations (Figure 2). The Bio-Rad ddSEQ SCI can process samples in < 5 minutes. By eliminating lengthy experimental workflows that can alter transcriptional signatures, acute transcriptional responses can be detected and tracked in time course experiments.



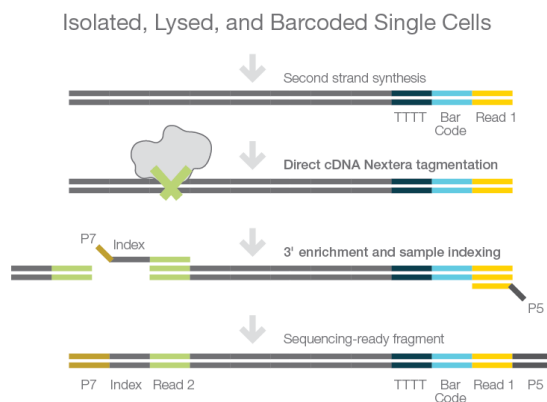
**Fig. 1. Single-Cell RNA-Seq Workflow.** The workflow integrates proven cell isolation using the Bio-Rad ddSEQ SCI, followed by library preparation using the SureCell WTA 3' Library Prep Kit with Nextera technology, Illumina sequencing, and data analysis with BaseSpace Sequence Hub and SeqGeq Analysis Software from FlowJo, LLC.



**Fig. 2. Unbiased, Cell-Size Agnostic RNA-Seq.** Mouse cell lines (A20, NIH 3T3) and human cell lines (HEK, BJ) were processed using the SureCell WTA 3' Library Prep Kit. Consistently high numbers of detected genes demonstrates that recovery of transcripts is not limited by cell size.

## Advanced NGS Chemistry for Library Prep

A single cell suspension is loaded into a ddSEQ M Cartridge and cells are encapsulated and barcoded by the ddSEQ Single-Cell Isolator. Lysis and barcoding takes place in each droplet. Droplets are disrupted and cDNA is pooled for second strand synthesis in bulk. Libraries are generated with direct cDNA tagmentation using Nextera<sup>®</sup> technology, without the need for shearing or preamplification (Figure 3). Tagmentation is followed by 3' enrichment and sample indexing to prepare up to 24 indexed, sequencing-ready libraries with minimal hands-on time.



**Fig. 3. Overview of The SureCell WTA 3' Library Prep Kit for the ddSEQ System.** After cells are isolated, lysed, and barcoded, cDNA pooled for second strand synthesis in bulk. Libraries are generated with direct cDNA tagmentation followed by 3' enrichment and sample indexing.

## Reliable Sequencing with Illumina NGS Systems

Prepared single-cell libraries are loaded directly onto an Illumina MiSeq<sup>®</sup>, NextSeq<sup>®</sup>, or HiSeq<sup>®</sup> Series System for sequencing. These sequencing systems harness industry-leading Illumina sequencing by

synthesis (SBS) chemistry, used in >90% of all NGS around the world, resulting in over 26,000 peer-reviewed publications.\* This common technology provides consistency and reliability in sequencing data across all Illumina NGS platforms. The MiSeq, NextSeq, and HiSeq Series of sequencing systems offer flexibility for a broad range of applications, and scalable throughput to support various study sizes.

Easy-to-use instrument control software guides researchers through sample loading and run setup with intuitive user-interfaces, and enables real-time monitoring of fully automated sequencing runs either on-instrument or remotely.

## Simple Yet Powerful Data Analysis

Single-cell sequencing data can be instantly transferred, stored, and analyzed securely in BaseSpace Sequence Hub, the Illumina cloud-based genomics computing environment. BaseSpace Sequence Hub provides a large collection of BaseSpace Apps. Commercial and open-source tools support a range of common data analysis needs such as alignment, variant calling, and more. These Apps feature intuitive push-button user interfaces designed to be used without the need for bioinformatics expertise.

The SureCell RNA Single-Cell App was designed to support data analysis for the Illumina Bio-Rad Single-Cell Sequencing Solution. The SureCell RNA Single-Cell App enables streamlined data analysis for up to 96 samples across multiple sequencing runs and includes:

- Convenient sequencing quality control (QC) metrics
- Easy assignment of unique transcripts to single cells
- Exportable and downloadable gene expression matrices and reports
- Various analysis options for identification of subpopulations and differentially expressed genes

SeqGeq Software is a desktop application for data analysis, exploration, and visualization that enables researchers to extract insight from the individuality of single-cell gene expression. Built for and by biologists with more than 19 years of experience with single-cell biology, SeqGeq Software features powerful data reduction and population identification tools.

SeqGeq Software directly integrates with BaseSpace Sequence Hub and enables visualization and analysis of expression data with statistic color-mapping of individual cells, summary heatmaps, and drag and drop report editors.

## High-Quality Data

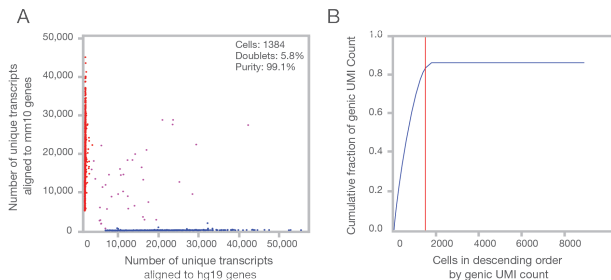
To demonstrate the high-quality single-cell RNA-Seq data achieved with the Illumina Bio-Rad Single-Cell Sequencing Solution, *in vitro* experiments were performed with cells mixed from different species. Human embryonic kidney 293 (HEK 293) cells and NIH 3T3 mouse embryonic fibroblasts were mixed at a 1:1 ratio (unless otherwise noted) and loaded across four sample chambers of a single Bio-Rad

ddSEQ M Cartridge. The ddSEQ SCI encapsulated and barcoded 1,384 single cells. Barcoded transcripts were processed for single-cell sequencing using the Illumina Bio-Rad SureCell WTA 3' Library Prep Kit for the ddSEQ System, and sequenced on the Illumina NextSeq 550 System. Sequencing results were analyzed using the Illumina SureCell RNA Single-Cell App.

\*Data calculations on file. Illumina, Inc. 2015.

### Confident Gene Detection in a Heterogeneous Population of Cells

Data analysis and plotting of human (hg19) and mouse (mm10) unique transcripts detected in the mixed species samples shows a low percentage of cell barcodes (5.8%) containing transcripts mapping to both species, which represent cell doublets (Figure 4A). This demonstrates efficient single-cell isolation at high purity (99%). By graphing the cumulative fraction of gene transcripts assigned to cell barcodes, the inflection point (red line) is used to determine the number of barcoded cells detected (Figure 4B). This indicates that a high fraction of transcripts are assigned to single cells.

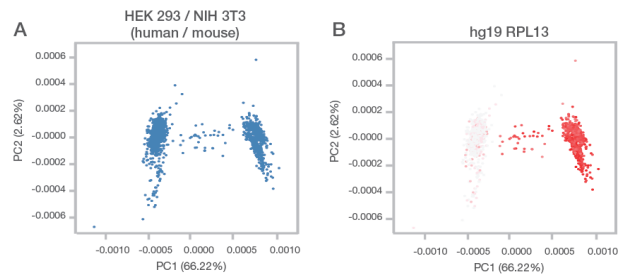


**Fig. 4. Confident Single-Cell Transcript Identification.** **A.** BaseSpace generated plots of the number of unique molecular identifiers (UMI), ie transcripts, assigned to the mouse/mm10 (red) and human/hg19 (blue) genomes for each cell barcode. Unique transcripts mapping to both mouse and human (purple) represent cell doublets. **B.** Cumulated fraction of unique transcripts assigned to cell barcodes (linear scale). The inflection point (red line) determines the number of barcoded cells detected in the sequencing run.

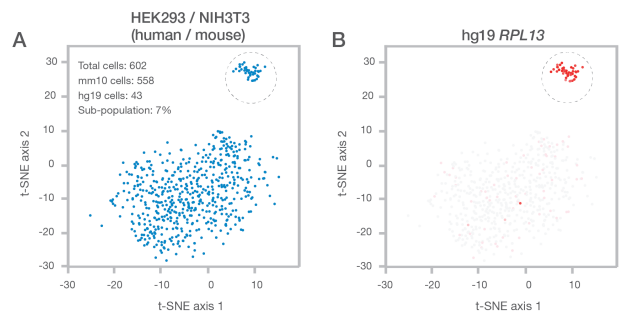
### Clear Identification of Cell Subpopulations in Mixed Samples

Principal component analysis (PCA) clustering of sequenced single cells from the HEK 293/NIH 3T3 mixture shows two distinct cell populations (Figure 5A). Cells color coded by gene expression of human ribosomal protein L13 (hg19 RPL13) clearly discriminate human from mouse cells (Figures 5B).

An additional *in vitro* species mixing experiment was performed in which mouse NIH 3T3 cells were spiked with a small number of human HEK 293 cells. With t-distributed stochastic neighbor embedding (t-SNE) analysis, human cells from this mixture were identified as a distinct cluster representing 7% of the total cell population (Figure 6A). Cells color coded by gene expression of human *RPL13* confirm the identity of the subpopulation as human cells (Figures 6B).



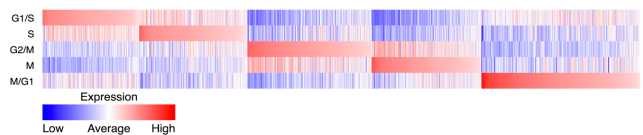
**Fig. 5. Confident Cell Population Identification.** **A.** PCA clustering of 1384 single cells sorted from a 1:1 ratio mixture of HEK 293 and NIH 3T3 cells, sequenced and analyzed with the SureCell RNA Single-Cell App. **B.** Cell populations cluster based on expression of the human (hg19) *RPL13* gene.



**Fig. 6. Confident Cell Subpopulation Identification.** **A.** Analysis in the SureCell RNA Single-Cell App using the t-SNE algorithm of a mixture of NIH 3T3 and HEK 293 cells identifies a distinct subpopulation of cells. **B.** Cells color coded by gene expression of hg19 *RPL13* confirms the identity of the subpopulation as human.

### Analysis of Single Cells Enables Identification of Cell Cycle Phases

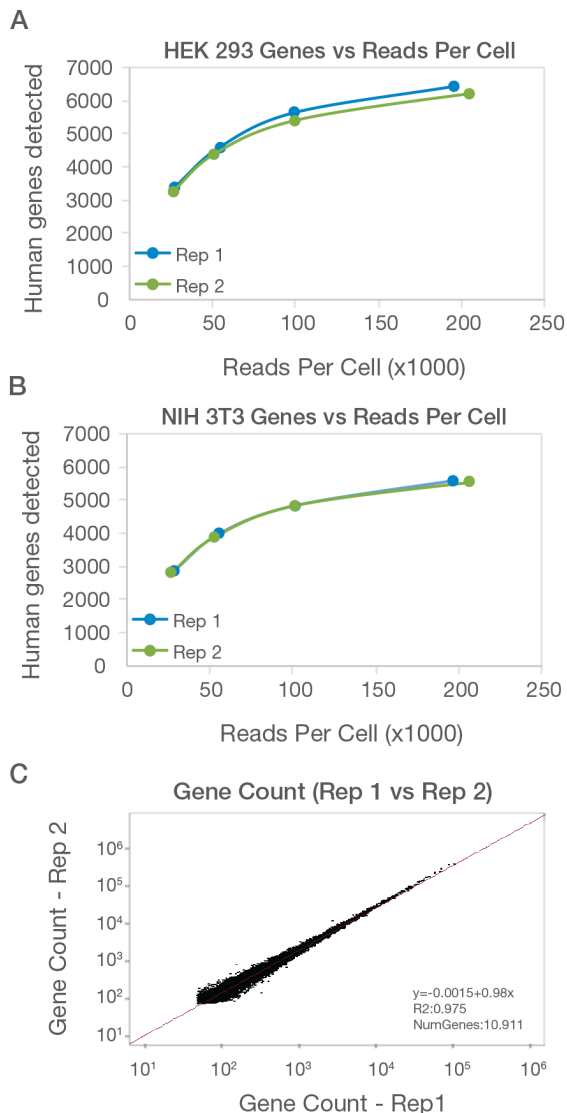
Gene expression analysis of markers associated with each major phase of the cell cycle enables accurate cell cycle analysis. Proliferating HEK 293 cells were sorted, prepared, and sequenced as before. Each phase of the cell cycle can be visualized as a heat map generated by the SureCell RNA Single-Cell App (Figure 7).



**Fig. 7. Cell Cycle Analysis of Single Cells by RNA-Seq.** Analysis of cell cycle state using the SureCell RNA Single-Cell App is based on unique transcript counts of genes associated with each phase of the cell cycle, normalized by the total count for each cell. Expression is centered by the median and scaled by the median absolute deviation for each cell cycle.

## Highly Sensitive and Reproducible Results

Replicate samples of HEK 293 cells were processed and sequenced on a NextSeq 550 System as described. Sequencing reads were subsampled to various reads per cells ranging from 25,000–200,000 reads. The median number of reads detected per cell at each sequencing depth shows highly sensitive and reproducible gene detection across replicates of different cell lines (Figure 8). Total gene counts for each human gene were summed across all HEK 293 cells. A linear regression fit of the summed gene counts between two replicates processed on a single ddSEQ M Cartridge further demonstrates highly reproducible results (Figure 8C).



**Fig. 8. Sensitivity and Reproducibility of Gene Detection Across Varied Cell Lines.** **A.** The median number of genes detected per cell are plotted at sequencing depths from 25,000–200,000 reads per cell for two replicates of human HEK 293

cells. **B.** The median number of genes detected per cell for two replicates of mouse NIH 3T3 cells. **C.** A linear regression of gene expression is plotted for genes with  $\geq 50$  counts, summed across all HEK 293 cells in two replicate samples.

## Summary

To advance understanding of the transcriptional heterogeneity that drives complex biological systems, researchers need a scalable, high-throughput, and user-friendly method for generating thousands of single-cell NGS libraries. The Illumina Bio-Rad Single-Cell Sequencing Solution is a comprehensive workflow developed in collaboration by industry experts in droplet-based cell isolation and NGS technologies. It reveals new types of single-cell information by facilitating analysis of multiple samples in parallel, under multiple treatment conditions, and at multiple time-points. The simple yet powerful data analysis options in the SureCell RNA Single-Cell App can resolve heterogeneous cell populations and identify subpopulations of interest using gene expression profiles and data visualization tools. Analysis of cell cycle markers allows for cell cycle analysis of individual cells in complex tissues. The Illumina Bio-Rad Single-Cell Sequencing Solution enables highly sensitive and reproducible interrogation of single-cell transcriptomes from hundreds to tens of thousands of single cells in a single experiment.

## Ordering Information

| Product  | Catalog No. |
|--|-------------|
| SureCell WTA 3' Library Prep Kit (2 cartridge kit) | 20014279    |
| SureCell WTA 3' Library Prep Kit (6 cartridge kit) | 20014280    |

## Learn More

To learn more about the Illumina Bio-Rad Single-Cell Sequencing Solution, visit:

[www.illumina.com/surecell](http://www.illumina.com/surecell)

[www.bio-rad.com/ddSEQ](http://www.bio-rad.com/ddSEQ)

[www.flowjo.com/seqgeq](http://www.flowjo.com/seqgeq)

## References

1. Kolodziejczyk AA, Kim JK, Svensson V, Marioni JC, Teichmann SA. The technology and biology of single-cell RNA sequencing. *Mol Cell.* 2015;58(4):610–620.
2. Shapiro E, Biezuner T, Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet.* 2013;14(9):618–630.
3. Hindson BJ, Ness, KD, Masquelier DA, et al. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal Chem.* 2011;83(22):8604–8610.

For Research Use Only. Not for use in diagnostic procedures.