**Introduction**

Understanding cell heterogeneity at the isoform level is critical for both basic and disease research. Short reads can only capture gene-level information, while other long-read technologies lack the accuracy for accurate unique molecular identifiers (UMI) and cell barcode (CBC) identification. PacBio® HiFi reads sequence full-length RNA isoforms along with single-cell barcode and UMI information, revealing extraordinary insight into single-cell biology.

The Kinnex™ single-cell RNA kit takes as input single-cell cDNA and outputs a sequencing-ready library that results in a 16-fold throughput increase compared to regular single-cell Iso-Seq® libraries. Combined with isoform-aware single-cell analysis SMRT® Link software, PacBio offers cost-effective single-cell isoform sequencing that does not require orthogonal sequencing methods. The SMRT Link software supports bioinformatics analysis to produce an isoform-level single-cell data matrix compatible with tertiary analysis software.
**Single-cell RNA sequencing**

Single-cell RNA sequencing (scRNA-seq) emerged to characterize gene expression differences between individual cells derived from a complex tissue, allowing a higher-resolution view of the transcriptome.

Most single-cell experiments are done with short reads, which only capture the ends of molecules due to fragmentation. Sequencing fragments limits expression information to the gene level, missing important isoform diversity that could be important for disease or biological function.

PacBio HiFi reads sequence full-length RNA isoforms along with single-cell barcode and UMI information (Figure 1), revealing isoform diversity at the single-cell level.

### HiFi sequencing advantages for single-cell RNA sequencing

- Full-length isoform information
- Accurate cell barcode and UMI detection
- Variant detection

![Figure 1. Single-cell isoform sequencing with PacBio long reads. Short reads only capture gene ends, missing isoform diversity. PacBio HiFi reads cover the entire isoform along with the single-cell barcode and UMI information with high accuracy.](image)

### Kinnex for single-cell isoform sequencing

The Kinnex single-cell RNA kit utilizes the MAS-Seq method for throughput increase. **MAS-Seq** (Al’Khafaji et al., 2021) is a concatenation method for increasing throughput by joining cDNA molecules into longer concatenated fragments. HiFi reads generated from sequencing the concatenated molecules can then be bioinformatically broken up to retrieve the original cDNA sequences. The result is higher throughput and reduced sequencing needs for cost-effective single-cell isoform sequencing. Traditionally, orthogonal short-read scRNA-Seq is used to supplement the lower throughput of regular single-cell Iso-Seq method — with Kinnex, no orthogonal sequencing data are required.

The **PacBio Single-cell Iso-Seq workflow** processes the full-length cDNA sequences to classify them against a reference annotation (e.g., GENCODE) to identify novel genes and isoforms. The output consists of gene- and isoform-level count matrices that are compatible with tertiary analysis software.

The Kinnex kit is compatible with cDNA generated using the 10x Chromium Next GEM Single Cell 3’ kit (v3.1) and Single Cell 5’ kit (v2) and is intended for use on a 3,000 to 10,000 cell library with 15-75 ng of cDNA as input.
Kinnex single-cell library workflow

The Kinnex library workflow begins with single-cell cDNA and produces a Kinnex library that is ready for sequencing.

![Diagram of the Kinnex single-cell library workflow]

Single-cell cDNA molecules are first removed of TSO (template-switching oligos) artifacts, then concatenated to form ordered arrays. Kinnex adapter ligation ensures full arrays are enriched, which are then sequenced on PacBio long-read sequencers (Figure 3). Additionally, the barcoded Kinnex adapters allow multiple Kinnex libraries to sequence on one SMRT® Cell.

With proper full array formation and adequate sequencing, one SMRT Cell on the Sequel® II/Ile and Revio™ systems are expected to achieve 40 million and 80 million cDNA sequences, respectively (Table 1).

![Table 1. Target Kinnex single-cell library performance]

Kinnex single-cell bioinformatics workflow

The SMRT Link Read Segmentation and Single-cell Iso-Seq workflow processes the HiFi reads generated from the Kinnex library to produce gene- and isoform-level count matrices that are compatible with tertiary single-cell analysis tools (Figure 4).

![Diagram of the Kinnex analysis using Read Segmentation and Single-cell Iso-Seq workflow]

**Read segmentation**

HiFi reads are segmented into individual segmented reads (S-reads) that represent the original cDNA sequences.

**Primer/UMI/BC extraction**

Primers and polyA tails are removed, but also used to orient the read into 5' – 3' orientation. Single-cell barcode and UMI information are extracted.

**Barcode correction + UMI deduplication**

Cell barcodes are corrected given an expected barcode list. Real cells — cell barcodes that represent encapsulated single cells (as opposed to ambient RNA) are also identified at this step. Reads are then deduplicated based on cell barcodes and UMIs.

**Mapping and transcript classification**

Deduplicated reads are mapped to the reference genome and classified against a transcript annotation (e.g., GENCODE). Finally, a gene- and isoform-level single-cell matrix is output for tertiary analysis.
Kinnex example: PBMC dataset

PBMC cDNA generated using the 10x Chromium Next GEM Single Cell 3’ kit was made into Kinnex libraries and sequenced with one SMRT Cell each and analyzed using the Read Segmentation and Single-Cell Iso-Seq workflow in SMRT Link v11.1 (Table 2). The output results were then processed with tertiary tools to identify cell types (Figure 5).

Kinnex for single-cell isoform sequencing: summary

The Kinnex single-cell RNA kit offers an end-to-end solution for single-cell RNA isoform sequencing from sample preparation to bioinformatics analysis.

- Supports cDNA generated from 10x Chromium Next GEM Single Cell 3’ kit (v3.1) and Single Cell 5’ kit (v2)
- Supports multiplexing with barcoded Kinnex adapters
- 15-75 ng input cDNA
- Target 3,000 to 10,000 cell library
- 16-fold throughput increase compared to non-Kinnex methods
- No orthogonal sequencing data required

Table 2. Read, cell, and transcript statistics of Kinnex single-cell library runs on Sequel II/Ile and Revio systems. Each library was run on one SMRT Cell. Data can be downloaded here.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Sequel II PBMC 1</th>
<th>Sequel II PBMC 2</th>
<th>Revio PBMC 3</th>
<th>Revio PBMC 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>HiFi reads</td>
<td>2,622,891</td>
<td>2,789,502</td>
<td>7,261,030</td>
<td>7,439,608</td>
</tr>
<tr>
<td>Segmented reads (S-reads)</td>
<td>40,131,832</td>
<td>43,788,133</td>
<td>110,127,016</td>
<td>114,238,613</td>
</tr>
<tr>
<td>S-read length distribution</td>
<td>300–2,500 bp</td>
<td>300–2,500 bp</td>
<td>300–2,500 bp</td>
<td>300–2,500 bp</td>
</tr>
<tr>
<td>S-reads with valid barcodes</td>
<td>37,634,585 (94%)</td>
<td>41,369,598 (94%)</td>
<td>105,309,352 (96%)</td>
<td>109,213,167 (96%)</td>
</tr>
<tr>
<td>Deduplicated reads</td>
<td>23,883,685</td>
<td>30,903,830</td>
<td>60,421,293</td>
<td>61,831,404</td>
</tr>
<tr>
<td>Estimated number of cells</td>
<td>3,966</td>
<td>9,316</td>
<td>8,822</td>
<td>8,815</td>
</tr>
<tr>
<td>Reads in cells</td>
<td>91.6%</td>
<td>87.84%</td>
<td>94.99%</td>
<td>94.82%</td>
</tr>
<tr>
<td>Mean reads per cell</td>
<td>8,708</td>
<td>3,905</td>
<td>11,352</td>
<td>11,762</td>
</tr>
<tr>
<td>Median UMIs per cell</td>
<td>4,821</td>
<td>5,861</td>
<td>6,002</td>
<td>957</td>
</tr>
<tr>
<td>Median genes per cell</td>
<td>704</td>
<td>938</td>
<td>957</td>
<td></td>
</tr>
<tr>
<td>Median transcripts per cell</td>
<td>818</td>
<td>705</td>
<td>1,110</td>
<td>1,130</td>
</tr>
</tbody>
</table>

Figure 5. UMAP cell clustering of Revio-PBMC 2 using CellTypist after SMRT Link analysis.

References