10x GENOMICS®

TECHNICAL NOTE

Base Composition of Sequencing Reads of Chromium™ Single Cell V(D)J Libraries

INTRODUCTION

The Chromium™ Single Cell V(D)J Protocol (CG000086) produces Single Cell V(D)J libraries, ready for Illumina® sequencing. Chromium™ Single Cell V(D)J libraries incorporate standard Illumina® paired-end constructs with P5 and P7 sequences at opposite ends. The 16 bp 10x™ Barcode and the 10bp UMI are located at the start of Read 1, while sample index sequence information is incorporated into the i7 index read. Read 1 and Read 2 are standard Illumina® sequencing primer sites used in paired-end sequencing (Figure 1). This Technical Note describes base percentage profiles and Phred quality scores (shown as %Q30 in Illumina® SAV software) that are characteristic of Chromium™ Single Cell V(D)J libraries for our recommended sequencing run configuration.

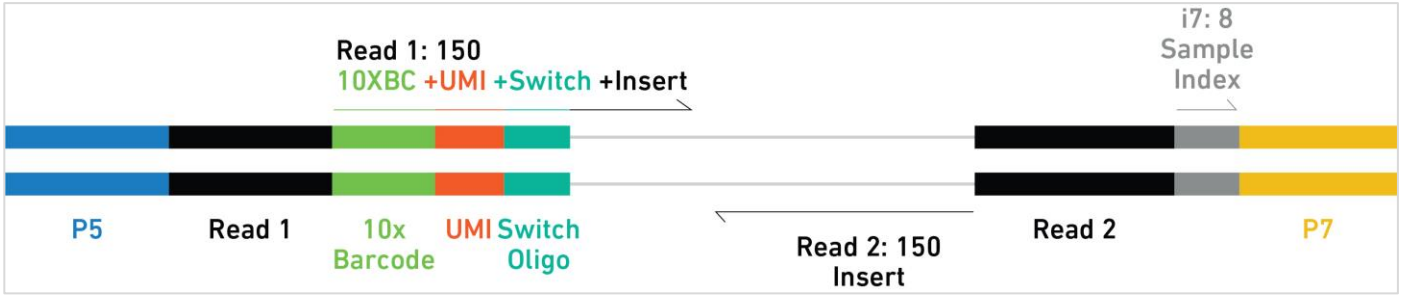

Fig. 1. Schematic overview of a fragment from a final Chromium™ Single Cell V(D)J library.

METHOD

We prepared 14 Chromium™ Single Cell V(D)J libraries with cells listed in Table 1 following the *Single Cell V(D)J Reagent Kits User Guide* (CG000086). A pool of 14 libraries was sequenced on Illumina® HiSeq® 4000 with approximately 5,000 read pairs per cell (one lane). Libraries were run using paired-end sequencing (150 bp Read 1 and 150 bp Read 2) with a single index (8 bp). Raw and processed data from a subset of replicate libraries are freely available from: https://support.10xgenomics.com/single-cell-vdj/datasets.

Cell Type	# Libraries
Pan T cells	4
Peripheral blood mononuclear cells (PBMCs)*	2
Jurkat (lymphoblast cell line)*	4
Anti EBV specific T cells*	4

Table 1. Cells used to generate 14 Single Cell V(D)J libraries. *Data from replicate libraries are available on the 10x Genomics® Support Website.

DISCUSSION

Chromium™ Single Cell V(D)J libraries are run using paired-end sequencing with a single index read per sample. We recommend the sequencing run parameters listed in Table 2.

Sequencing Read	Recommended Number of Cycles
Read 1	150 cycles
i7 index	8 cycles
i5 index	0 cycles
Read 2	150 cycles

Table 2. Recommended sequencing run parameters for Chromium™ Single Cell V(D)J libraries.

Figure 2 illustrates the distribution of base composition along Read 1, i7 index read, and Read 2 that we typically observe after a successful sequencing run of a Single Cell V(D)J library that was prepared according to the *Single Cell V(D)J Reagent Kits User Guide* (CG000086). The profiles are characteristic for Chromium™ Single Cell V(D)J libraries that are sequenced with the recommended number of cycles and can be explained as follows:

- Cycle 1 to 26 (Read 1): Base percentages fluctuate due to sequences from the 16 bp 10x™ barcode and the 10 bp UMI that are attached to the Gel Bead primers. These non-normally distributed sequences will have a different distribution than the human transcriptome, which causes the apparent shift in percentage of each base.
- Cycle 27 to 39 (Read 1): Base percentages are of low diversity due to the Switch Oligo sequence.
- Cycle 40 to 150 (Read 1): Base percentages are expected to fluctuate throughout the V sequence region.
- Cycle 151 to 158 (i7 index): Base percentages are dissimilar due to sequences from the 8 bp sample index.
- Cycle 159 to 308 (Read 2): Base percentages are expected to fluctuate throughout the V(D)J sequence region.

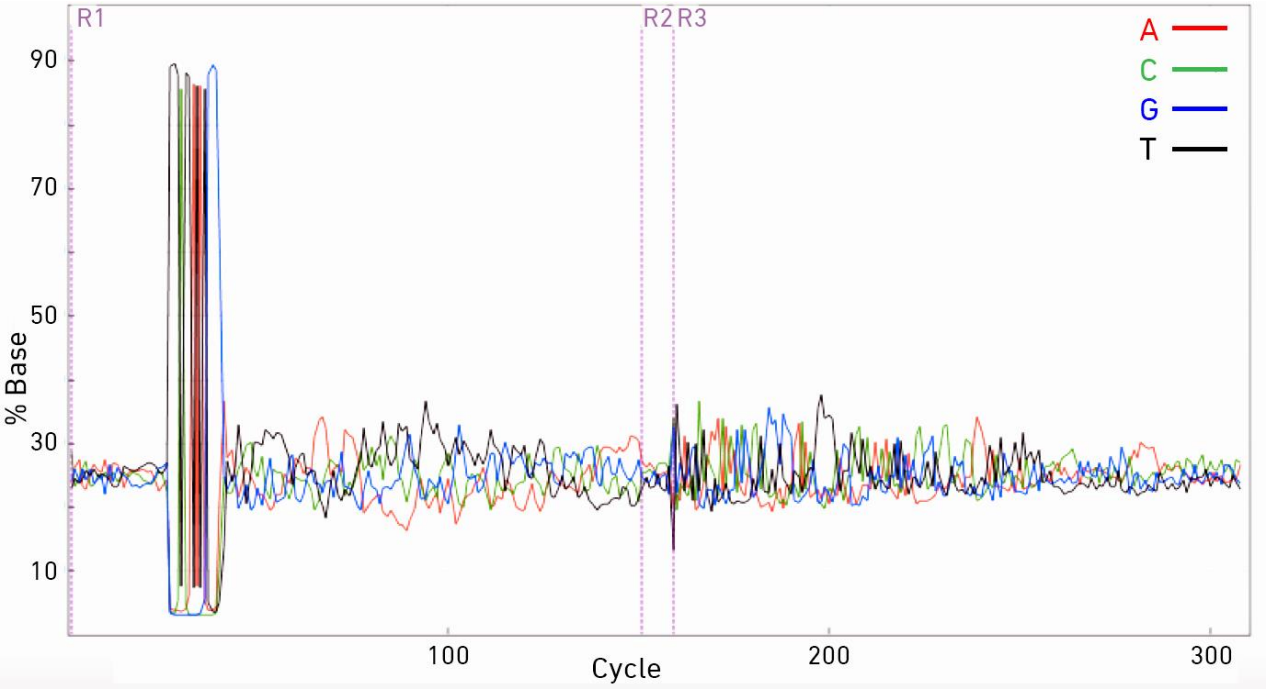

Fig. 2. Representative example of the 'Data by Cycle' plot in the Sequencing Analysis Viewer software (Illumina®). Shown is the percentage of clusters for which the selected base has been called (% base: y axis) along the sequencing length (x axis). The profile is based on the sequencing of a Chromium™ Single Cell V(D)J library by itself with no other library type sequenced alongside.

The Phred quality score assesses base calling accuracy and is typically used to determine how much of the data from a given sequencing run can be used. Sequencing data with lower quality scores can result in a

significant portion of reads being unusable. Figure 3 and Table 3 outline the Q30 quality metrics that we typically achieve with Chromium™ Single Cell V(D)J libraries run on the Illumina® HiSeq® 4000. Percentages of Q30 are relatively stable at the beginning of Read 1. The percentages drop near the end of Read 1, but recover for the i7 Read and remain relatively stable through the beginning of Read 2.

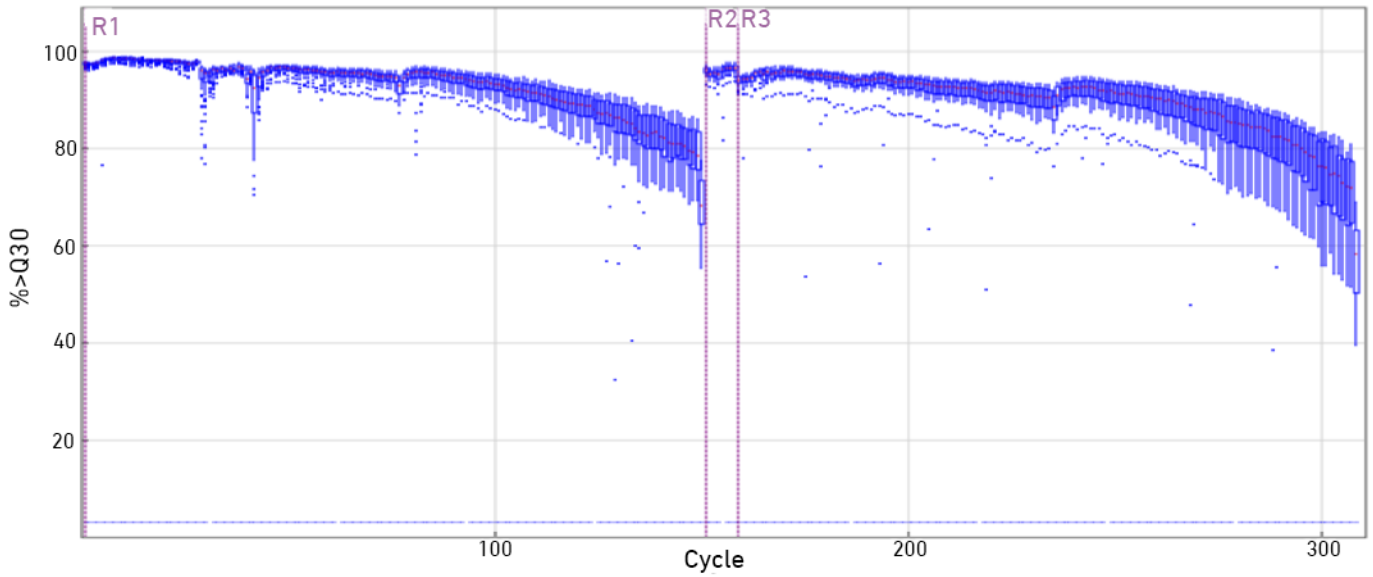

Fig. 3. Representative example of the 'Data by Cycle' plot in the Sequencing Analysis Viewer software (Illumina®). Shown is the Q30 percentage along the sequencing length. Profile is based on sequencing Chromium™ Single Cell V(D)J library by itself with no other library type sequenced alongside.

Run Format	Read 1 > Q30	i7 > Q30	Read 2 > Q30
150 x 150	93.1 %	96.0%	89.6 %

Table 3. Q30 percentages for Read 1, i7 and Read 2 for the recommended sequencing run format.

CONCLUSION

We have discussed sequencing parameters for Chromium™ Single Cell V(D)J libraries and the resulting apparent base composition of sequencing reads generated from our recommended sequencing run configuration. Base percentages and quality scores across the library fragments are correlated with the sequences flanking the library insert. Base percentages will trend away from "normal" as a function of increased sample clonality (the more clonal the sample, the smaller the effective amplicon diversity). The representative example profiles for Chromium™ Single Cell V(D)J libraries demonstrated here will serve as a reference for what constitutes a successful sequencing run using this library type.

REFERENCES

- Chromium™ Single Cell V(D)J Reagent Kits User Guide (CG000086)

Notices

Document Number

CG000103 Rev A *Technical Note*

Legal Notices

Customer Information and Feedback

For technical information or advice, please contact our Customer Technical Support Division online at any time.

Email: support@10xgenomics.com

10x Genomics 7068 Koll Center Parkway

Suite 401

Pleasanton, CA 94566 USA