10x GENOMICS®

TECHNICAL NOTE

Chromium™ Genome Libraries – Sequencing Metrics for Illumina® NovaSeq®

INTRODUCTION

The Chromium™ Genome v2 Protocol (CG00043) produces Genome v2 libraries, ready for Illumina® sequencing. Genome libraries incorporate standard Illumina paired-end constructs with P5 and P7 sequences at opposite ends. The 16bp 10x™ Barcode is encoded 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). The libraries have been validated on the following sequencing instruments: HiSeq® 2500 Rapid Run (RR), HiSeq® 3000/4000, and HiSeq® X Ten/Five. With the introduction of the Illumina NovaSeq® Series, we have validated the performance of Genome v2 libraries on the NovaSeq sequencing platform. This Technical Note describes key sequencing metrics of the Illumina NovaSeq platform and is intended to provide general guidance on the expected range of sequencing metrics. Individual results may still vary, depending on the particular sample and loading characteristics.

Fig. 1. Schematic of a fragment from a final Chromium™ Genome v2 library.

METHOD

We prepared five Chromium Genome v2 libraries from the sample NA12878. Libraries were prepared following the *Chromium™ Genome Reagent Kits v2 User Guide* - CG00043. Libraries were pooled (Library ID 1) at equimolar ratios and run on the Illumina NovaSeq using paired-end sequencing with a single index read per sample.

RESULTS

Libraries were sequenced using a NovaSeq 5000/6000 S2 Reagent Kit (300 cycles) and the sequencing run parameters listed in Table 1.

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

Table 1. Recommended NovaSeq® sequencing run parameter for Chromium Genome v2 libraries.

We report the following sequencing metrics to assess sequencing run performance (Table 2):

- "Percentage of Clusters Passing Filers (%PF)"
- Yield per Lane for Read 1 and Read 2 in Gb
- Phred quality scores (shown as %Q30) for Read 1 (R1), i7 index and Read 2 (R2)

Library ID	Loading Conc. (pM)	Instrument	%PF	Yield per Lane (Gb)		%>=Q30		
				R1	R2	R1	i7	R2
1	300	NovaSeq	77.51	665.29	665.32	93.88	96.07	90.82

Table 2. Reported sequencing metrics for Genome v2 libraries on the Illumina NovaSeq® instrument with recommended loading concentration.

Figure 2 illustrates the distribution of base composition along R1, the i7 index read, and R2 that we typically observe after a successful sequencing run of a Chromium™ Genome v2 library that was prepared according to the *Chromium™ Genome Reagent Kits v2 User Guide*. The profiles are characteristic for Chromium Genome v2 libraries that are sequenced with the recommended number of cycles.

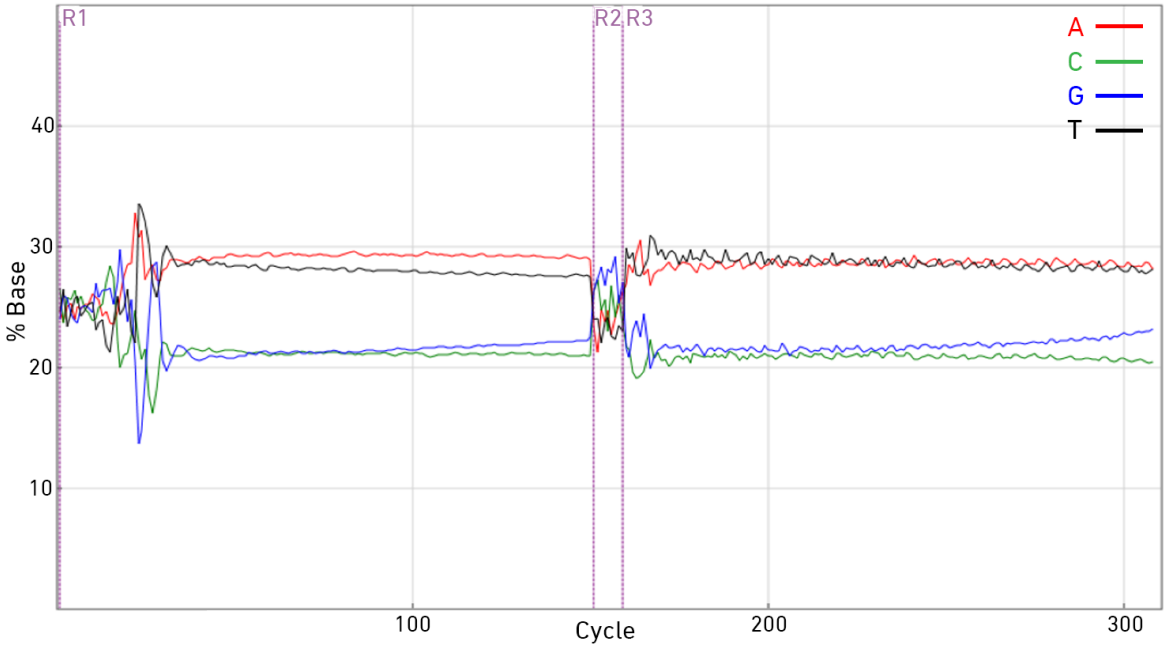

Fig. 2. Representative example of the 'Data by Cycle' plot in the Sequencing Analysis Viewer software (Illumina). Shown is the percentage of clusters for which the selected base has been called (% base: y axis) along the sequencing length (x axis). Profile is based on sequencing 10x library by itself 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 outlines the Q30 quality metrics that we typically achieve with Genome v2 libraries run on the Illumina® NovaSeq®. Percentages of Q30 are relatively stable across cycles for R1, R2 and the i7 index read. Percentages drop slightly at ~120 cycles and ~82 cycles for R1 and R2, respectively, but remain at > 80%.

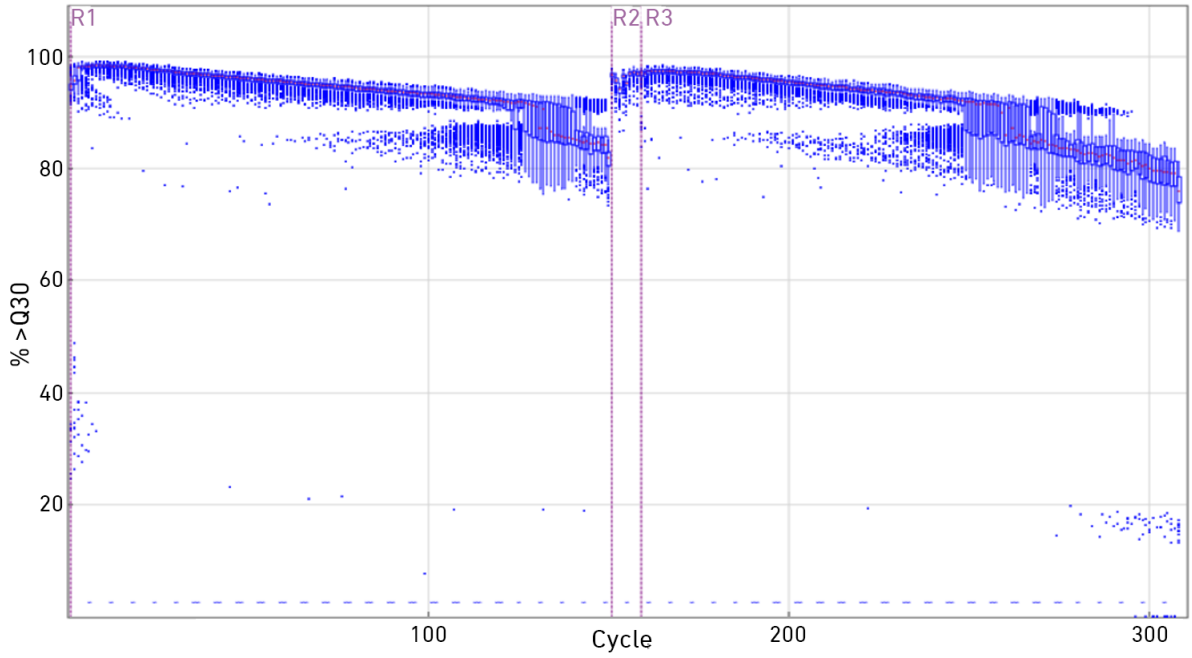

Fig. 3. Representative example of the 'Data by Cycle' plot in the Sequencing Analysis Viewer software (Illumina). Shown is the Q30 percentage along the sequencing length. Profile is based on sequencing 10x library by itself with no other library type sequenced alongside.

DISCUSSION

As expected, libraries sequenced on the NovaSeq produced high quality data. Overall, sequencing quality on the NovaSeq is highest compared to data obtained from the HiSeq® 2500 and HiSeq 4000 platforms, particularly for Read 2 sequencing quality. Please refer to the Technical Note *Chromium™ Genome v2 Libraries – Sequencing Metrics for Illumina® Sequencers* – CG000126 for more details. Q30 quality scores for R1 start at >90% and as expected decreased to ~85% towards the end of R1. Q30 quality scores for R2, similar to R1, start at >90% but decreased more rapidly to ~80% towards the end of R2. The sample index read (i7) consistently stayed at >90%. We do recommend to load Chromium Genome v2 libraries at a concentration of 300 pM with 1% PhiX spike-in.

CONCLUSION

We have discussed sequencing parameters for Chromium™ Genome v2 libraries sequenced on the Illumina® NovaSeq. Illumina's NovaSeq sequencing platform is compatible with Chromium Genome v2 libraries and may be used as an alternative for sequencing projects that require increased sample sequencing throughput. The representative example profiles and sequencing performance metrics of Chromium Genome v2 Libraries demonstrated here serve as a reference for what constitutes a successful sequencing run using this library type.

REFERENCES

- *Chromium™ Genome Reagent Kits v2 User Guide* (CG00043)
- *Chromium™ Genome v2 Libraries – Sequencing Metrics for Illumina® Sequencers* (CG000126)

Notices

Document Number

CG00122 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