TECHNICAL NOTE

Sequencing Metrics & Base Composition of Visium Spatial Gene Expression Libraries

Introduction

The Visium Spatial Gene Expression Reagent Kits workflow produces Visium Spatial Gene Expression libraries for measuring total mRNA in intact tissue sections. This Technical Note presents a comparison of sequencing metrics for pooled Visium Spatial Gene Expression libraries across Illumina platforms. The expected base percentage profiles and Phred quality scores based on a control library are described to provide general guidance on the expected range of sequencing metrics on Illumina platforms. Individual results may vary depending on the specific sequencing instrument and/or particular sample and loading characteristics.

Visium Spatial Gene Expression Libraries

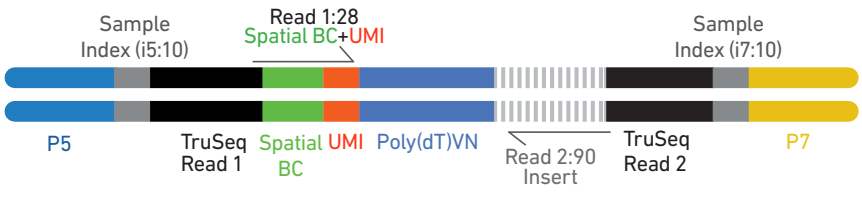
The dual index library that can be generated using the Visium Spatial Gene Expression reagents and protocols is shown in the schematic below.

Visium Spatial Gene Expression libraries comprise standard Illumina paired-end constructs that are flanked with P5/P7, necessary for binding to the Illumina flow cell. TruSeq Read 1 is used for priming and sequencing the 16 bp Spatial Barcode and 12 bp UMI, and TruSeq Read 2 is used for priming and sequencing the cDNA insert. The two 10 bp sample indexes are sequenced in the i5 and i7 read respectively.

Methods Overview

Eight Visium Spatial Gene Expression libraries were generated from eight immunostained human breast tissue samples following the Methanol Fixation, Immunofluorescence Staining & Imaging Demonstrated Protocol (Document CG000312) and the Visium Spatial Gene Expression Reagents Kits User Guide (Document CG000239) and pooled for sequencing to generate the % Base and % ≥Q30 plots shown in Table 1. Sequencer compatibility, shown in Table 2, was determined using four Visium Spatial Gene Expression libraries from two mouse brain and two human heart samples fixed and stained using the Methanol Fixation, H&E Staining & Imaging Demonstrated Protocol (Document CG000160). All libraries were quantified with the KAPA DNA Quantification Kit and sequenced with 1% PhiX.

Visium Spatial Gene Expression library



Results Overview

Table 1 shows base composition data derived from the indicated libraries. A representative 'Data by Cycle' plot from the Illumina SAV software displaying the percentage of base calls and Q30 quality scores is shown. Table 2 shows representative sequencing metrics derived from the indicated libraries. Due to expected lower output from iSeq and MiSeq flow cells, data from these platforms did not meet the minimum sequencing depth requirement for Visium libraries and should be used for quality control purposes only. Individual results may vary depending on the specific sequencing instrument and/or particular sample and loading characteristics.

Base percentage fluctuation throughout sequencing, as shown on the Illumina SAV 'Data by Cycle' plot in Table 1, are due to the following:

- Read 1, 1-28 cycles - fluctuation is due to sequences from the 16 bp Spatial Barcode and 12 bp UMI.
- i7, 29-38 cycles - fluctuation is due to the 10 bp sample index.
- i5, 39-48 cycles - fluctuation is due to the 10 bp sample index.
- Read 2, 49-139 cycles - fluctuation may be due to sequences from the TSO (30 bp) from the 5' end of the cDNA from a small subset of library fragments. Base percentages reflect the expected base composition of the transcript read. An increase in "A" is expected towards the end as the proportion of sequences containing the poly-A tail increases

Sequencing configuration & run parameters:

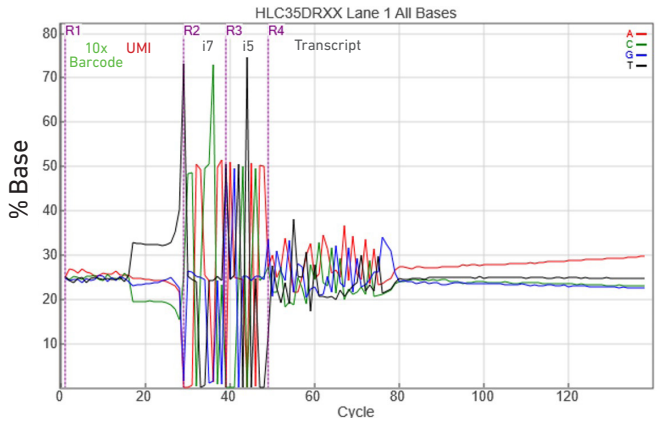
Minimum sequencing depth: 50,000 read pairs per tissue covered spot

Paired-end, dual indexing

- Read 1: 28 cycles
- i7 Index: 10 cycles
- i5 Index: 10 cycles
- Read 2: 90 cycles

Table 1: Representative Plots

Plots shown are from a pool of eight Visium Spatial Gene Expression libraries sequenced on a NovaSeq SP flowcell.



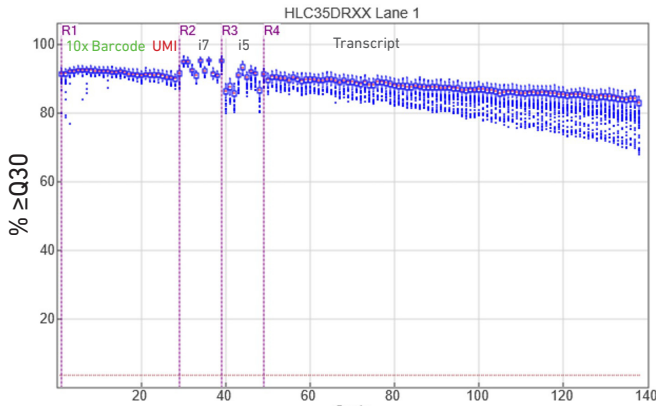
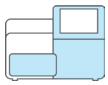
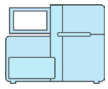
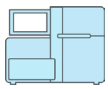


Table 2: Sequencing Data								
		% ≥Q30				Yield per Lane (Gb)		Reads Mapped to Genome (%)
		R1	i7	i5	R2	R1	R2	
iSeq								
	Loading Conc. (pM): 60 % PF*: 72.8 Phix (%): 1	96.26	89.87	77.63	95.4	0.15	0.50	94.08
MiSeq								
	Loading Conc. (pM): 18 Cluster Density: 1528 K/mm^2 Phix (%): 1	96.17	95.95	96.04	79.2	0.81	2.67	88.43
HiSeq 2500 RR								
	Loading Conc. (pM): 12 Cluster Density: 1295 K/mm^2 Phix (%) 1	95.75	89.55	94.36	77.7	11.85	39.16	88.83
HiSeq 4000								
	Loading Conc. (pM): 240 Cluster Density: % PF*: 78.9 Phix (%): 1	98.02	89.73	88.19	75.30	82.41	272.46	90.29
NextSeq 500								
	Loading Conc. (pM): 1.8 Cluster Density: 265 K/mm^2 Phix (%): 1	93.68	92.34	88.03	76.70	15.46	51.02	92.01
NovaSeq (SP flow cell)								
	Loading Conc. (pM): 300 % PF*: 77.4 Phix (%) 1	94.82	95.41	94.27	91.90	26.68	92.87	94.51

*Percent Pass Filter (% PF) is reported for iSeq, HiSeq 4000, and NovaSeq instead of cluster density due to the patterned flow cell

Conclusions

In summary, % Base by cycle and % ≥Q30 Quality Score distribution showed highly consistent profiles for all sequencing platforms tested. The data serve as guidelines for assessing the quality of Visium Spatial Gene Expression library sequencing. Additional factors that may contribute to overall success of a sequencing run and impact downstream application performance metrics include:

- Starting with a high quality tissue block. The Visium Spatial Gene Expression assay was optimized using samples with a RIN score of ≥7.
- Determining optimal permeabilization time for the tissue type, section thickness, and region of interest using the Tissue Optimization kit prior to starting the Gene Expression experiment.
- Final libraries with fragment length of 300 to 600 bp and a significant number of inserts between 400-500 bp for optimal cluster formation on Illumina flowcells.
- Reliable and accurate library quantification using the KAPA DNA Quantification Kit and the average insert size determined by Agilent Bioanalyzer QC.
- Sequencing platform loading concentration.

References

- Visium Spatial Gene Expression Reagent Kits User Guide (CG000239)
- Methanol Fixation, Immunofluorescence Staining, and Imaging for Visium Spatial Protocols (CG000312)
- Methanol Fixation, H&E Staining, & Imaging for Visium Spatial Protocols (CG000160)

Contact:
support@10xgenomics.com
10x Genomics
6230 Stoneridge Mall Road
Pleasanton, CA 94588 USA