NIPPON Genetics Co.,Ltd.

Application

Preparation of a library from minor RNA (50pg) and single cell-equivalent RNA using KAPA HyperPlus Kit altered protocol.

Product Name: KAPA HyperPlus Kit (for illumina) (KK8510, KK8512, KK8514)

Manufacturer: KAPA BIOSYSTEMS

The below data is provided by courtesy of Dr. Masahito Hosokawa, the Research Organization for Nano & Life Innovation, Waseda University, Japan.

Outline

The cost for library preparation is significant in the RNA-seq experiment in which a large number of sample are used, and this may limit the scale of experiment. Many kits recommend to use more than 1ng input cDNA as a requirement. However, the yield is small for cDNA derived from the single-cell RNA and the consumption of valuable samples is concerned.
In order to resolve the issue of costs and input volume limit, we considered a method to prepare an altering library by combining SMART-seq2, which is an RNA-seq (amplifier cDNA) library preparation protocol for one cell, with the KAPA HyperPlus Kit, and compared this with products of other companies.

Conditions and procedure of experiment

We prepared cDNA through the SMART-seq2 (Nature Protocols 9, 171 – 181 (2014)) method using 50pg of total RNA extracted from cells. Using this, library was prepared in several conditions in which the input cDNA volume and the reaction liquid volume of Ligation and Amplification were changed.

- Species: human cancer cell line
- Initial sample volume: total RNA 50pg
- RNA extracting method: RNeasy mini kit (Qiagen) DNase I treated
- cDNA preparation: SMART-seq2 (Nature Protocols 9, 171-181 (2014))
- Library preparation:
  KAPA HyperPlus Kit (ligation-based)
  Kit N of Company I (tagmentation-based)
- Sequencer: Miseq

KAPA HyperPlus reaction conditions

Sample ID	Protocol	Input cDNA			Adapter (nM)	Group
		Concentration (ng/uL)	Liquid volume (uL)	cDNA volume (ng)		
S01	Standard (recommended method)	0.2	5	1	300	K1
S02						
S03						
S04	L15_A12	0.2	1	0.2	1500	K2
S05						
S06						
S07		1	1	1		K3
S08						
S09						
S10		0.2	1	0.2	15000	K4
S11						
S12						
S13		1	1	1		K5
S14						
S15						
S16	L15_A15	0.2	1	0.2	1500	K6
S17						
S18						
S19		1	1	1		K7
S20						
S21						

Kit N of Company I reaction conditions

Sample ID	Input cDNA (ng)	Reaction liquid volume	Group
N.S01	1 (recommended method)	1x	N1
N.S02			
N.S03			
N.S04	0.25	0.25x	N2
N.S05			
N.S06			
N.S07	0.5		N3
N.S08			
N.S09			
N.S10	1		N4
N.S11			
N.S12			

＊Points of consideration for reaction conditions＊

KAPA HyperPlus kit
Input DNA volume: 0.2ng 1.0ng
Reaction liquid volume:
L15_A12: Reacts with 1/5 Ligation volume and 1/2 Amplification volume
L15_A15: Reacts with 1/5 Ligation volume and 1/5 Amplification volume
Number of amplification cycles: 15 cycles in all conditions

Kit N of Company I
Input DNA volume: 0.25ng 0.5ng 1.0ng
Reaction volume: Reacts with 1/4 volume
Number of amplification cycles:
12 cycles in all conditions

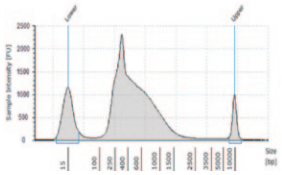
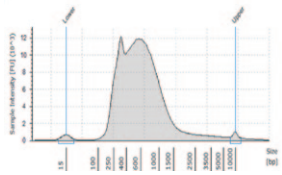
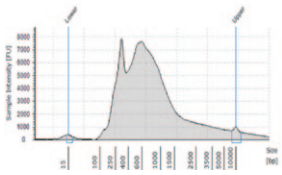
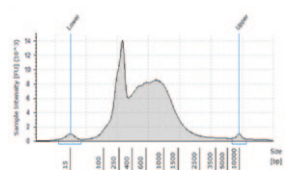
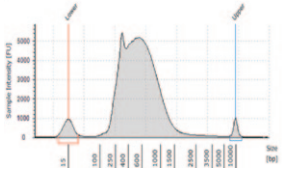
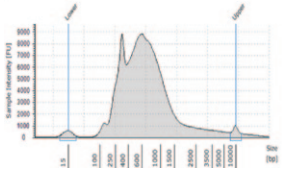
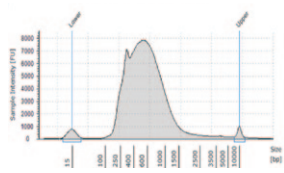
Database and analysis tool

Category	Version
Reference genome	Ensembl Human Genome,Release 90
Annotation file	Ensembl Human GTF, Release 90, GTF
QC, filtering	flexbar 2.4, fastq-mcf, FastQC 0.11.2
Genome mapping	HISAT2 2.0.5
Gene expression amount	RSEM 1.3.0
Mapping region identification	bedtools 2.26.0
Calculation of cover rate	RSeQC (bam2wig.py)

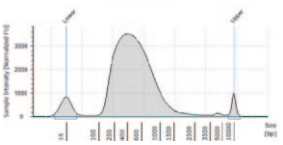
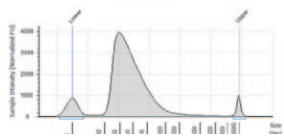
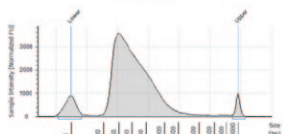
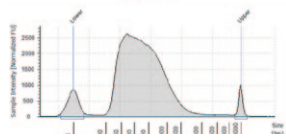
Results

1. Checking library size using TapeStation

① Library prepared using KAPA HyperPlus

Adapter Stock / Input DNA volume	Library Amplification 50μL	Library Amplification 25μL		Library Amplification 10μL
	300nM	1500nM	15000nM	1500nM
	X 5μL	X 1μL	X 1μL	X 1μL
1ng	K1 (control) 	K3 	K5 	K7 
0.2ng		K2 	K4 	K6 

② Library prepared using Kit N of Company I

Input DNA volume	1ng	0.25ng	0.5ng	1.0ng
	N1 (control) 	N2 	N3 	N4 

Although there is a peak around 250bp in KAPA HyperPlus, it didn't cause problems in sequence.

2. Sequence analysis data

＊Refer to "Each Analysis Data" after page 4.

Summary

Category	Results of KAPA HyperPlus	Kit N of Company I results	Evaluation
Accuracy of sequence	Peaks in Phred Score 38, the proportion of which is less than Kit N of Company I.	Peaks in Phred Score 38	KAPA ≦ Kit N of Company I
Mismatch	About 12%	About 8%	KAPA < Kit N of Company I
Insert	About 0.3%	About 0.3%	KAPA = Kit N of Company I
Defect	About 0.5%	About 0.5%	KAPA = Kit N of Company I
GC content	Peaks at 60%	Peaks at 50 – 60% (profiles change by control and conditional change sample)	KAPA > Kit N of Company I
Sequence lead length	Several peaks due to inclusion of cDNA amplification adapter sequence*1	Only 1 peak as cDNA amplification adapters can hardly be detected*1	KAPA > Kit N of Company I
Genome map rate	40 – 60% (possible to improve by theparameter of mapping tool)	70 – 80%	KAPA ≦ Kit N of Company I
Insert size	150 -200 bp	50 – 100bp	KAPA > Kit N of Company I
Cover rate of copied materials	Consistent on the whole	Inclines toward 3'	KAPA > Kit N of Company I
Number of genes identified	About 10,000	About 10,000	KAPA = Kit N of Company I
Repeatability of data	Repeatability can be low depending on conditions	Repeatability is high if conditions are the same	KAPA < Kit N of Company I
Consistency with conditional changes	Mostly consistent even if conditions change	Varies if the expression level is high.	KAPA > Kit N of Company I
Variance	CV limit: 0.25 (1000TPM) CV limit: 0.5 (100TPM) CV limit: 1.0 (10TPM)	CV limit: 0.25 (1000TPM) CV limit: 0.5 (100TPM) CV limit: 1.5 (10TPM)	KAPA = Kit N of Company I (excluding low-expression genes) KAPA < Kit N of Company I (low-expression genes)
Dependency on examination concentration	Almost the same as control even if concentration is changed.	If the concentration is changed, the repeatability of high-expression genes slightly drops. Expression level changes significantly depending on GC content.	KAPA > Kit N of Company I (repeatability) KAPA > Kit N of Company I (Changes of expression level depending on GC content)

*1 (Supplement)
For sequence lead length, the lead length after removing adapter sequence when cDNA and PCR used for SMART-seq2 are made is analyzed.
Therefore, short lead in particular was a sequence originally including an amplicon terminal after PCR.
It is assumed that, as KAPA HyperPlus is a ligation-basis, these sequences appear as peaks.
On the other hand, Kit N of Company I is a tagmentation-basis and it is assumed that the amplicon terminals do not appear as a lead in principle.

■ General evaluation

It is judged that the following point is especially evaluated and KAPA HyperPlus is superior.
It is judged that it can operate with K6 (1/5 Ligation volume, 1/5 Amplification, volume, input cDNA 0.2ng) as a condition.

- Deviations from GC contents are small.
- Insert size is long.

Customers' comments

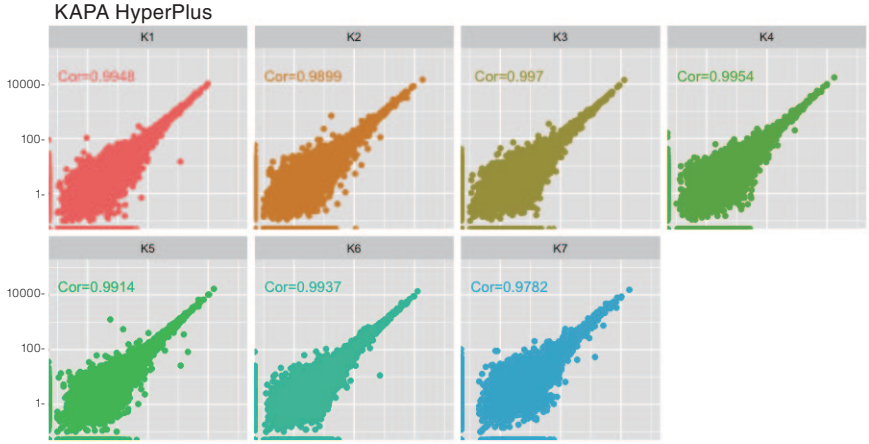
KAPA HyperPlus comprises simple workflows based on fragmentation using an enzymatic treatment. Therefore, it could easily be introduced in the experiment without installing a special DNA fragmentation device, an automated fluid dispensing machine or a nucleic acid extraction device in the laboratory. This product produces stable data even when reaction conditions are adjusted and enables a (relatively) reasonable sequence library preparation by saving reaction fluid volumes. This product can be recommended even for beginners who are not used to library preparation when installing.

Analysis data

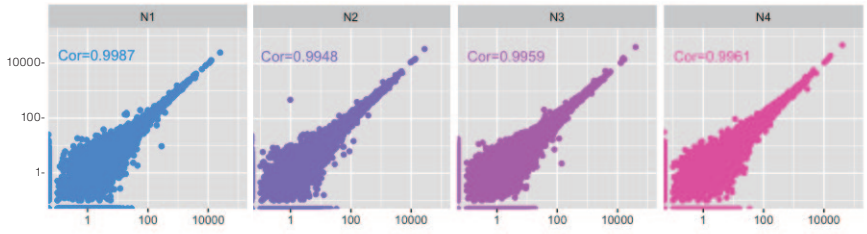
1. Repeatability of gene expression amount

- Comparing variations between samples in the same condition.
- The correlation factor was more than 0.99 in all conditions excluding K7, and variations could hardly be seen.

KAPA HyperPlus



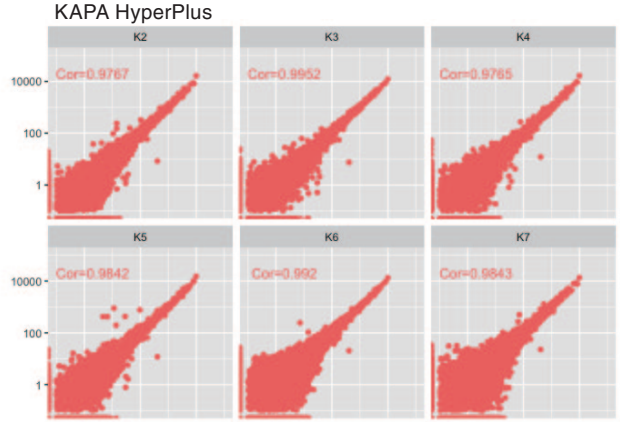
Kit N of Company I



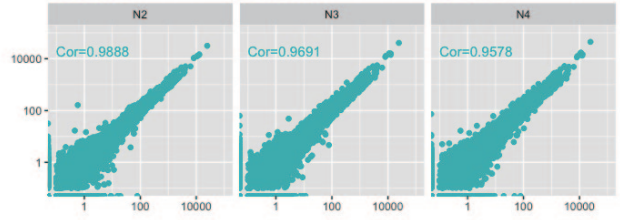
2. Comparison with control

- Comparing repeatability with K1 and N1 as control.
- There was a high consistency with control for KAPA HyperPlus.
- The repeatability of genes in high expression areas was low for Kit N of Company I.

KAPA HyperPlus

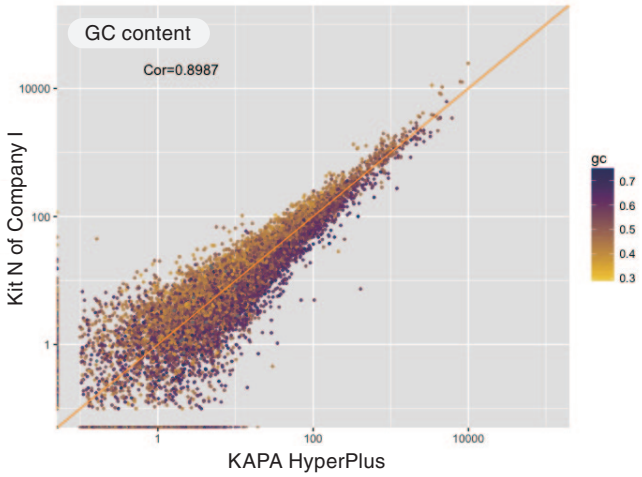


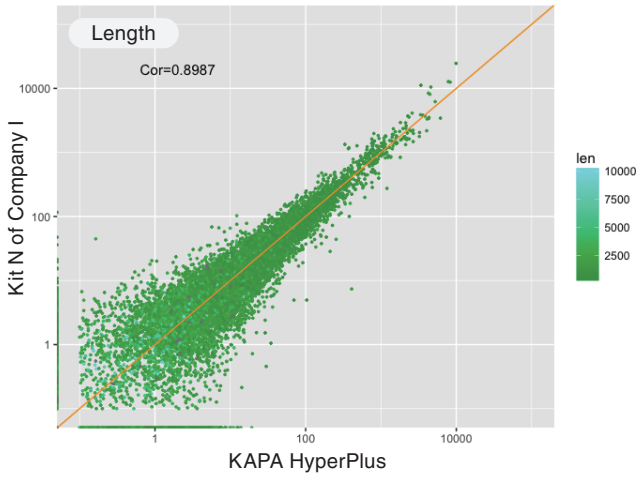
Kit N of Company I



3. Comparison between KAPA HyperPlus and Kit N of Company I (GC content and impact of length)

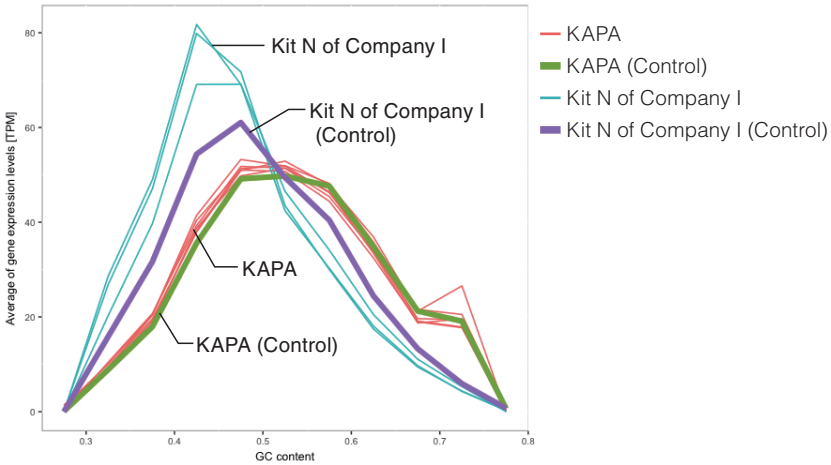
- Comparison between K1 (KAPA HyperPlus) and N1 (Kit N of Company I).
- There was a deviation in the expression level depending on GC content.
- No impact of length could be seen.





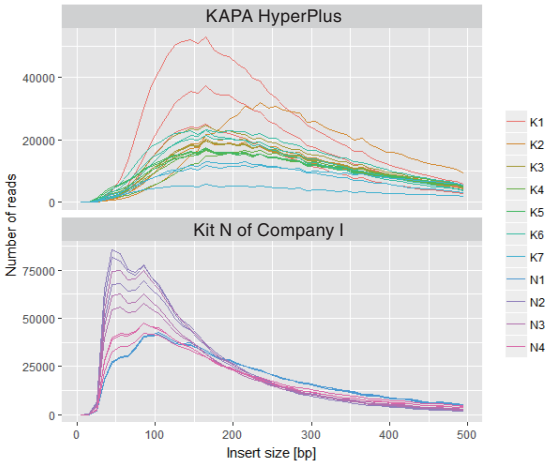
4. Average gene expression level by GC content for each sample

- There was a deviation in the expression level of genes with a low GC content in Kit N of Company I.
- KAPA HyperPlus stably showed the same trend as control in various conditions.



5. Distribution of insert size in each sample

- Insert sizes were calculated from the results of mapping.
- The insert size of KAPA HyperPlus was longer.
  Kit N of Company I: most frequent value of insert size: less than 100bp
  KAPA HyperPlus: most frequent value of insert size: 150 – 200bp



Nippon Genetics Co.,Ltd.
http://www.n-genetics.com
TEL 03(3813)0961 FAX 03(3813)0962 info@genetics-n.co.jp