TaKaRa Clontech

that's GOOD science!®

New Technologies for Sensitive, Low-Input RNA-Seq

Clontech Laboratories, Inc.

Outline

Introduction

Single-Cell-Capable mRNA-Seq Using SMART Technology

- SMARTer® Ultra™ Low RNA Kit for the Fluidigm C_1 System
- SMART-Seq™ v4 Ultra Low Input RNA Kit for Sequencing

Total RNA-Seq Applications

- SMARTer Stranded RNA-Seq Kit
- SMARTer Stranded Total RNA Sample Prep Kit - HI Mammalian
- SMARTer Stranded Total RNA-Seq Kit - Pico Input Mammalian

Expanding Applications for SMART Technology

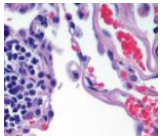
- DNA SMART™ ChIP-Seq Kit

Next-Gen Sequencing: RNA-Seq

cDNA synthesis for a complete representation of the transcriptome

Low-input total RNA samples

Single cells



Degraded, LCM, or FFPE samples

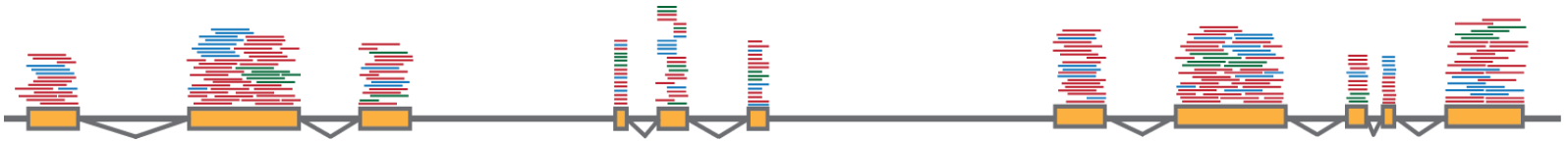
Prokaryotic samples

SMART Technology

SMARTer

- Industry standard for low-input and single-cell RNA-seq
- Full-length gene body coverage
- Simplified workflow
- Highest sensitivity and reproducibility
- High-quality sequencing libraries

Transcriptome Analysis with NGS



- RNA-seq produces millions of sequences from complex RNA samples. With this powerful approach, we can:
  - Measure gene expression/evaluate differential gene expression between different conditions, cell types, etc.
  - Discover and annotate complete transcripts
  - Discover and characterize alternative splicing (isoforms), polyadenylation & SNPs
- In recent years, more people have been interested in investigating such applications at the single-cell level

The Technology and Biology of Single-Cell RNA Sequencing
Kolodziejczyk A., *et al.* (2015) *Molecular Cell* **58**(4):610–620

SMARTer cDNA Synthesis for NGS

dT Primed

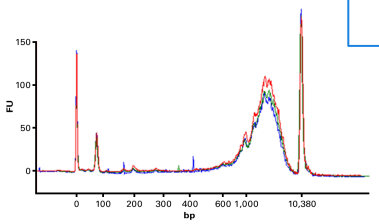
- mRNA
- Polyadenylated RNA
- Single cell capacity

N6 Primed

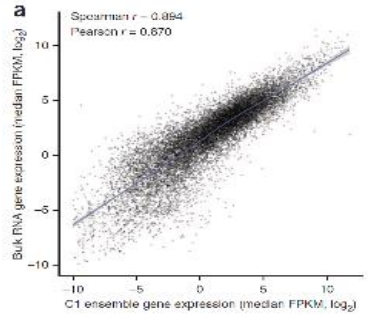
- Coding and non-coding RNA
- Non-polyadenylated RNA
- Degraded samples

SMART-Seq™ v4 Ultra™ Low Input RNA Kit for Sequencing

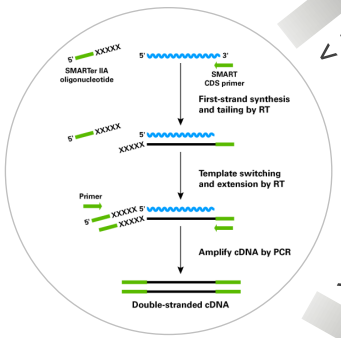
- Use with Ion Torrent or Illumina platforms



SMARTer Ultra Low RNA Kit for the Fluidigm C1™ System



- 96 single cells in parallel



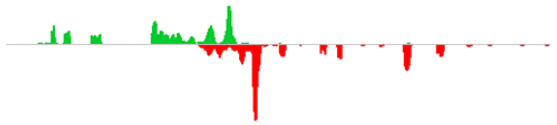
cDNA Synthesis

< 10 ng

> 10 ng

SMARTer Stranded Total RNA-Seq Kit - Pico Input Mammalian

- Use with Illumina platforms
- Use with highly degraded samples (FFPE)



SMARTer Stranded RNA-Seq Kit

- Use with Illumina platforms

RiboGone™ - Mammalian

SMARTer Stranded Total RNA Sample Prep- HI Mammalian

- Use with typical input RNA samples

TaKaRa Clontech

Importance of Studying Single Cells

Letter

Gene expression profiling in single cells from the pancreatic islets of Langerhans reveals lognormal distribution of mRNA levels

Martin Bengtsson,[1,2,4] Anders Ståhlberg,[2] Patrik Rorsman,[1,3] and Mikael Kubista[2]

Genome Research (2005) 15:1388–1392
Spring Harbor Laboratory Press

Cold

- Gene expression can vary significantly between cells (transcription occurs in bursts).
- Average expression of a population may not necessarily correlate with gene expression at the level of an individual cell.

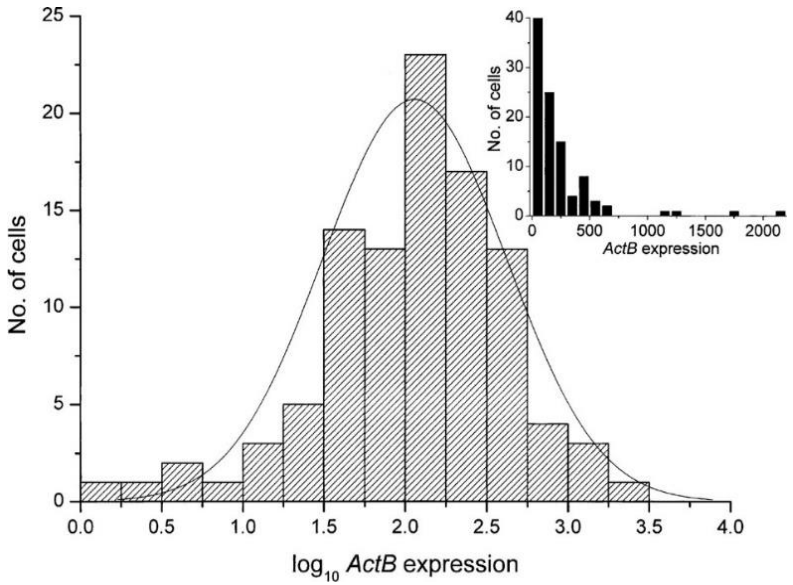
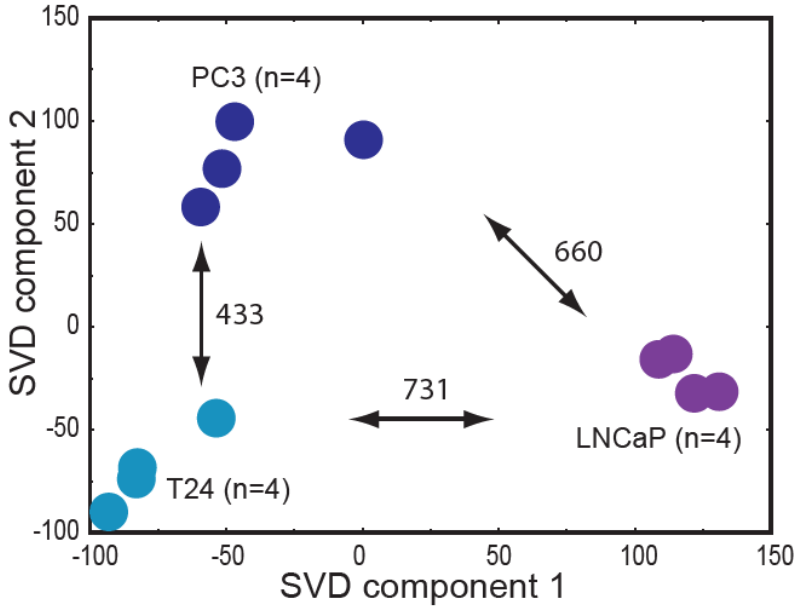


Fig 1. Histograms showing the expression levels of 96 cells expressing *ActB* in logarithmic and linear scale (inset).

TaKaRa Clontech

Transcriptome Analysis of Individual Cancer Cells

- Individual cells can be categorized according to their cell line of origin based on their transcriptome
- 12 individual cancer cells were isolated from three different cancer cell lines
  - Four cells each from prostate (PC3 and LNCaP) and bladder (T24) cell lines
- Global gene expression profiles were used to analyze each single-cell transcriptome



Full-length mRNA-seq from single-cell levels of RNA and individual circulating tumor cells.
Ramsköld, D., *et al.* (2012) *Nature Biotechnology* **30**(8):777–782.

Transcriptome Analysis of Individual Neurons

Single-cell RNA-seq discovers a wealth of new RNA markers that define discrete groups of neurons

Posted by: RNA-Seq Blog in Commentary 6 days ago 667 Views

from Bio-IT World By Aaron Krol

An adult mouse's brain, an object not much bigger than the last joint of your pinky finger, contains around 75 million neurons. At the Allen Institute for Brain Science in Seattle, the Mouse Cell Types program, led by Hongkui Zeng, is trying to figure out just how many varieties of neurons make up this vast complex, and what makes each one unique.

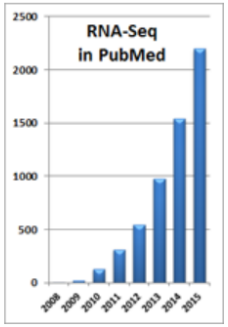
Zeng's research focuses on the primary visual cortex, a tiny sliver of the brain where signals from the eyes are processed and interpreted. Because vision is a relatively well-defined process, it's thought to be a good model for connecting the behavior of individual neurons to larger brain functions.

"You really can't understand a system until you understand its parts," says Bosiljka Tasic, a founding member of the Mouse Cell Types program.

To a shocking extent, those parts are still a mystery. Many supposed cell types are based on little more than what you can see through a microscope: a neuron's shape, or the pattern of rootlike dendrites extending from its body. These morphological traits, though important, are hard to see in full, and even harder to track methodically across thousands or millions of cells.

This month, Zeng's team published a study in *Nature Neuroscience* that takes advantage of new technological developments to get a fine-grained look at the molecular toolkits of single neurons. Using newly refined methods to isolate single cells, Zeng's lab collected over 1,600 brain cells from the visual cortexes of adult mice, intact and in good shape for sequencing. With advances in highly parallel, unbiased RNA sequencing, the group was able to measure each cell's entire "transcriptome"—the array of RNA molecules that indicate which genes are actively producing proteins—at a depth that reveals even the scarcest RNA traces.

"We think this is probably the most comprehensive survey of a cortical area," says Tasic, who co-led the study with her colleague Vilas Menon. "Many studies that are coming out now do very shallow sequencing



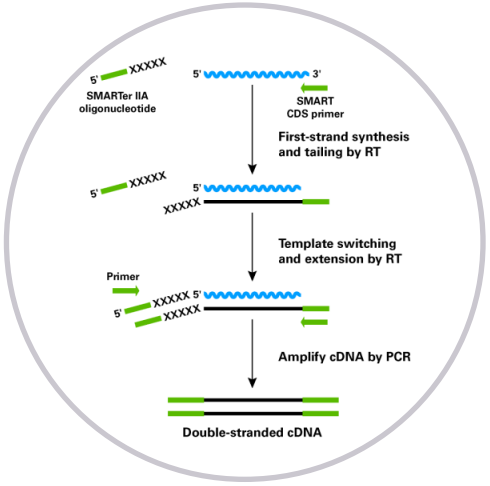
http://casestudies.brain-map.org/celltax#section_introa

Our Solutions for mRNA-Seq

dT Priming

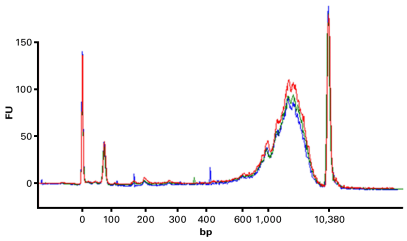
- mRNA
- Single-cell capacity
- Polyadenylated RNA
- With no strand information

Ultra-low-input total RNA and 1–1,000 cells



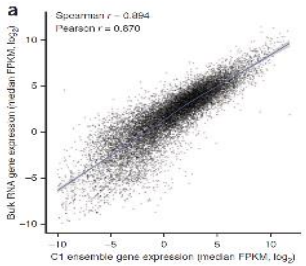
For the Fluidigm C1 cell-capture system

SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing
Input = 10 pg–10 ng; 1–1,000 cells



- Ultra-low-input total RNA, poly(A$^+$) RNA
- Single-cell capacity
- Compatible with Ion Torrent and Illumina® platforms

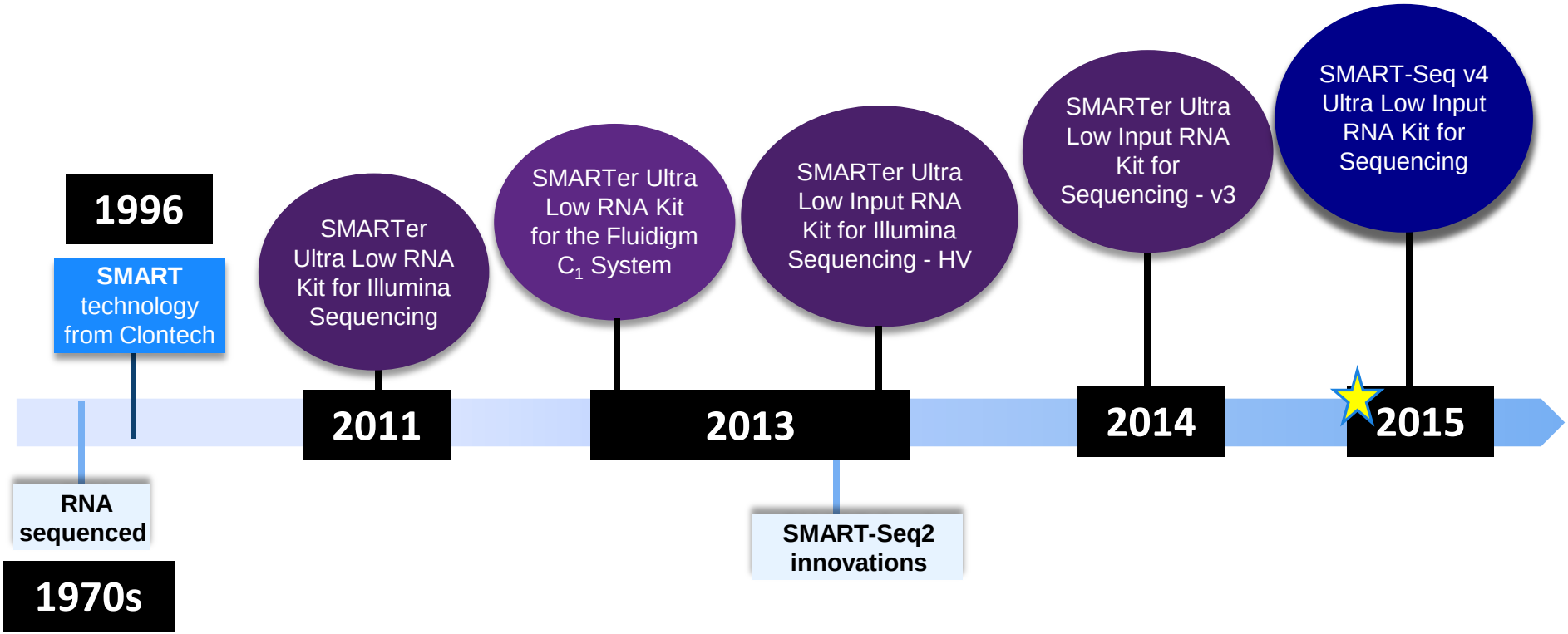
SMARTer Ultra Low RNA Kit for the Fluidigm C_1 System
Input = 1–1,000 cells



- 96 single cells in parallel
- Compatible with Illumina platforms

TaKaRa Clontech

Single-Cell and Ultra-Low Input RNA-Seq Solutions—Timeline



SMARTer Ultra Low RNA Kit for the Fluidigm C_1 System

* Slide from Fluidigm

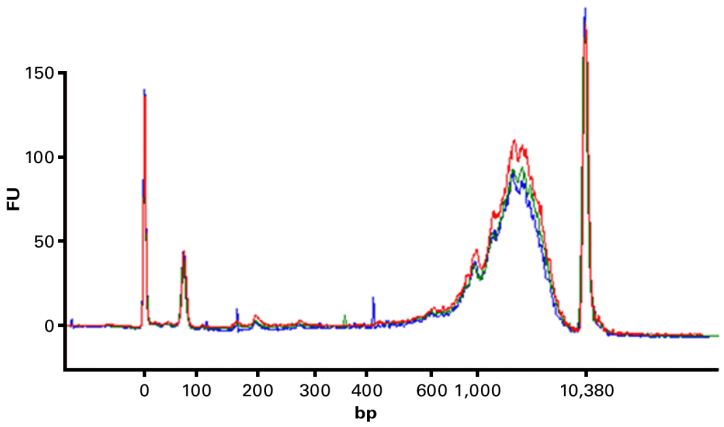
SMART-Seq v4—Advancements

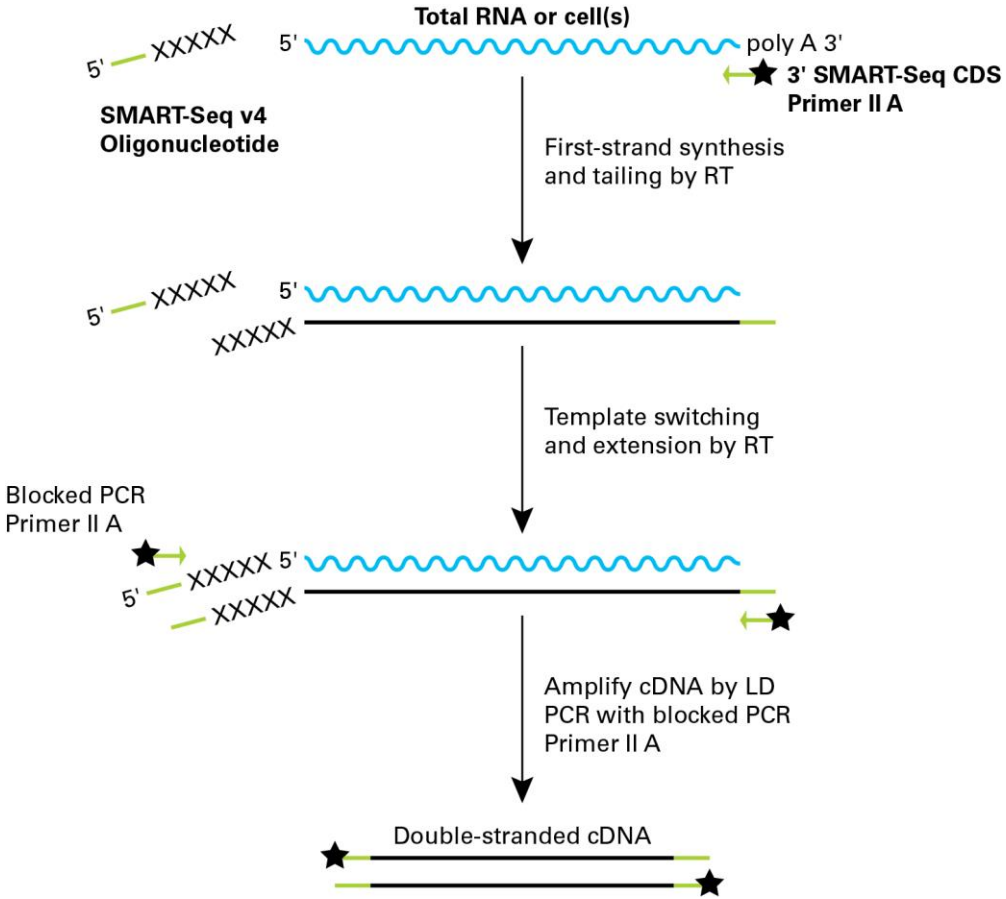
- Input = 10 pg–10 ng; 1–1,000 cells for mRNA-seq
- Optimized template-switching oligo: Incorporates LNA and with our proprietary knowledge of template switching
- Improved sensitivity and reproducibility
  - More genes identified
  - Higher yield
- Simplified protocol

The SMART-Seq v4 kit outperforms all previous generations of SMARTer Ultra Low kits by increasing sensitivity and reproducibility.

SMART-Seq v4—Technology

- The SMARTScribe™ Reverse Transcriptase (RT) makes cDNA.
- When the SMARTScribe RT reaches the 5' end of the RNA, its terminal transferase activity adds a few nucleotides.
- The SMART-Seq Oligonucleotide base-pairs with the non-templated nucleotide stretch, creating an extended template to allow the SMARTScribe RT to continue replicating.
- The SMART-Seq primer and oligo serve as universal priming sites for cDNA amplification by PCR.





TaKaRa Clontech

Comparing ULv3, SMART-Seq v4, and SMART-Seq2

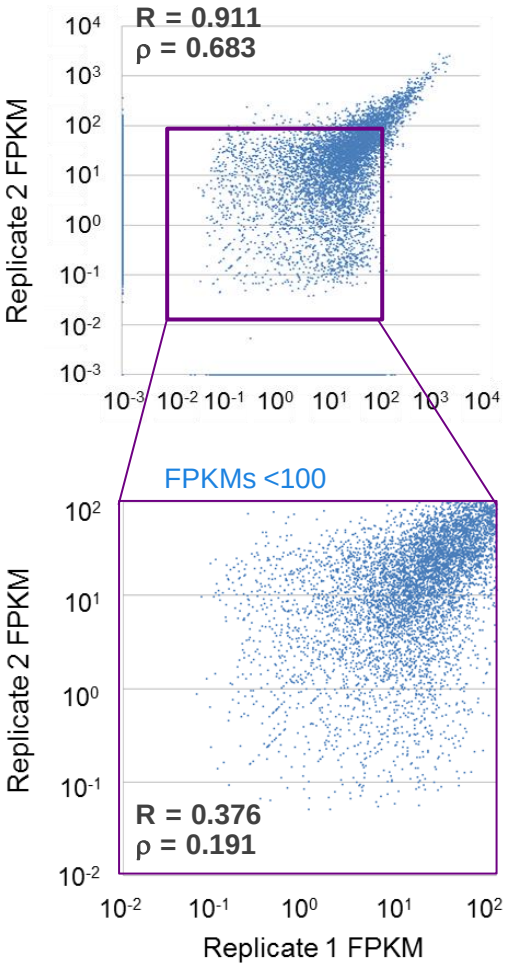
Improvements to the template-switching oligo

- Reduced background
- Improved sensitivity and reproducibility

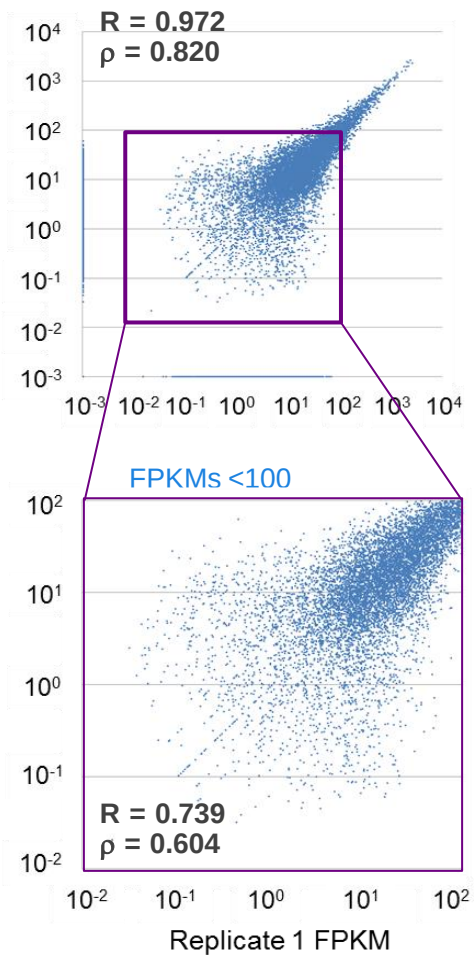
Sequencing Data Comparing cDNA Synthesis Methods						
Input	10 pg Mouse Brain RNA					
Protocol	ULv3		SMART-Seq v4		SMART-Seq2	
Number of PCR cycles	18		17		18	
Number of reads (Millions)	4.0 (paired-end)					
Replicate	1	2	1	2	1	2
No. of transcripts identified	11,647	10,885	14,731	14,813	12,080	12,039
Percentage of reads (%):						
Mapped to genome	96	97	96	95	72	93
Mapped to exons	73	73	76	76	66	67
Mapped to introns	21	21	19	20	28	27
Mapped to intergenic regions	6.0	6.2	4.7	4.7	5.8	5.8

Comparing ULv3, SMART-Seq v4, and SMART-Seq2

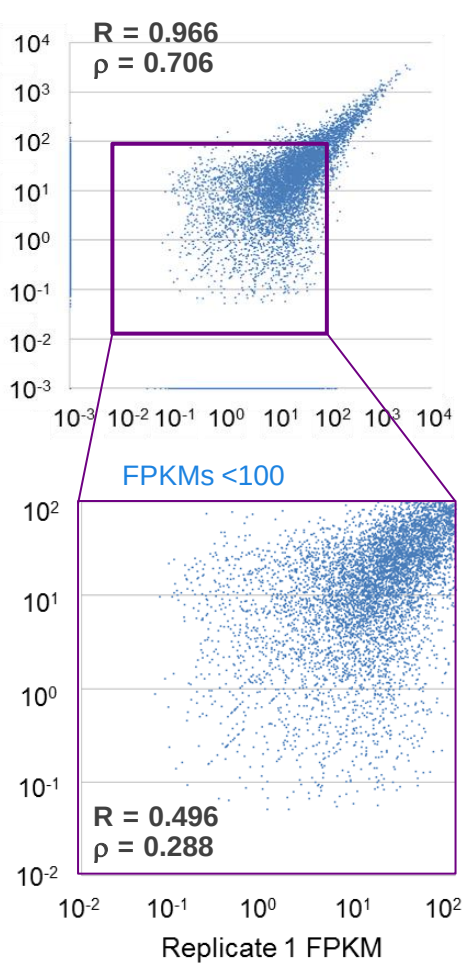
ULv3



SMART-Seq v4



SMART-Seq2



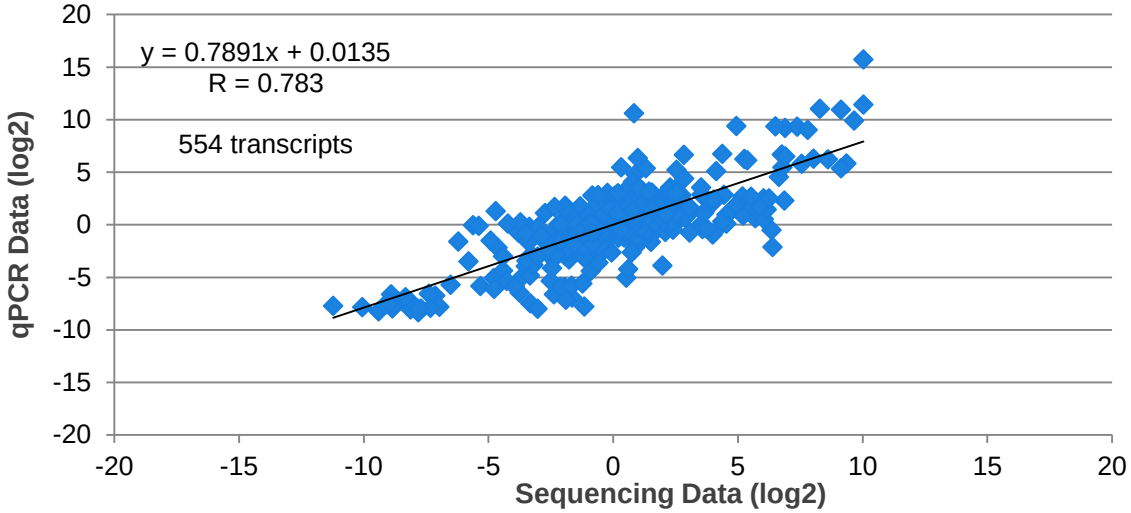
TaKaRa Clontech

Mapping Statistics from 10 pg–10 ng of MAQC Controls

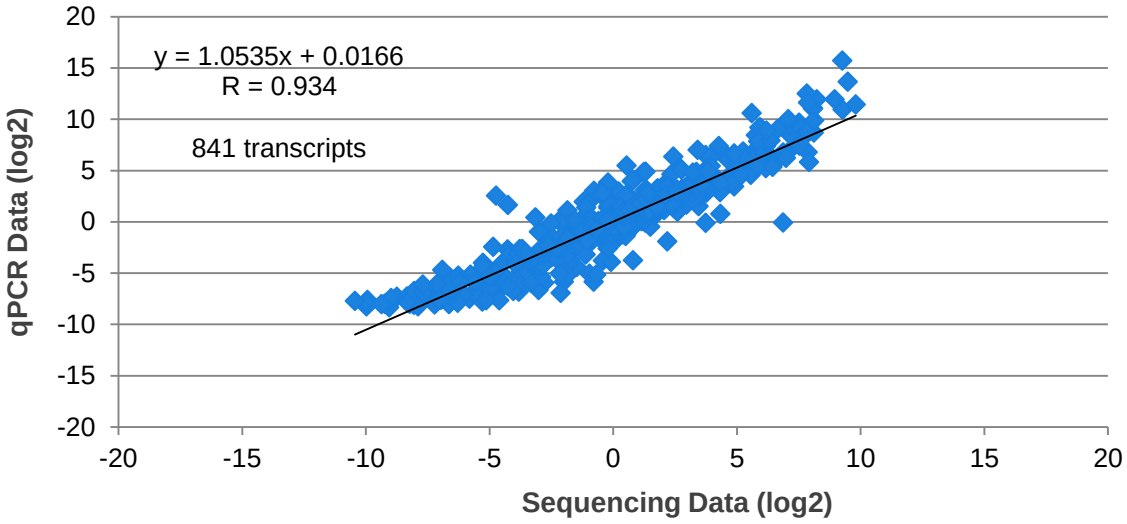
Sequencing Metrics from MAQC and ERCC Controls												
RNA source	HBRR			HURR			HBRR			HURR		
Input amount	10 pg						10 ng					
Number of paired-end reads (Millions)	2.3						3.4					
Number of PCR cycles	18						8					
Transcripts with FPKM >0.1	15,482	15,338	15,421	17,612	17,756	17,588	25,111	25,218	25,128	25,075	25,118	25,146
Transcripts with FPKM >1	12,598	12,561	12,707	14,491	14,516	14,445	18,346	18,163	18,386	18,061	17,792	17,907
Percentage of reads (%):												
Mapped to rRNA	1.2	1.2	1.1	0.7	0.7	0.6	5.6	5.4	5.6	4.0	4.0	4.1
Mapped to mitochondria	9.0	9.1	8.8	3.5	3.4	3.3	8.7	8.7	8.8	3.2	3.5	3.5
Mapped to genome	92	90	92	94	94	94	94	94	94	96	96	96
Mapped to exons	77	79	78	80	80	81	73	73	73	77	77	77
Mapped to introns	18	17	17	14	14	14	21	21	21	18	17	18
Mapped to intergenic regions	5.1	4.9	5.1	5.4	5.4	5.2	5.8	5.8	5.8	5.8	5.6	5.7
Duplicates	21	22	22	20	20	20	18	18	18	21	20	19
Mapped to ERCC	3.7	3.8	3.8	1.2	1.2	1.1	2.9	2.8	2.9	0.95	0.96	0.97

Differential Expression Compared to MAQC qPCR Data

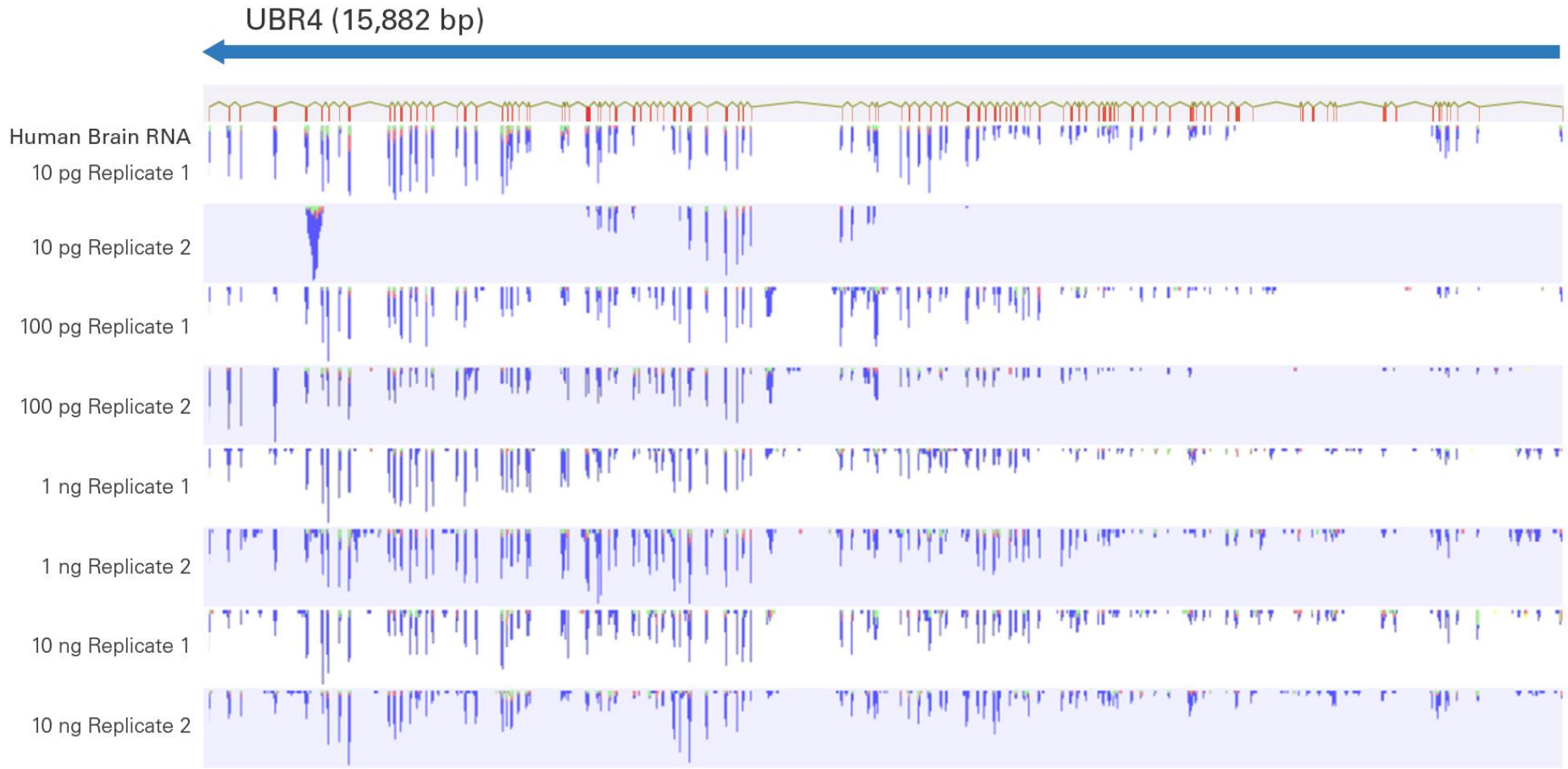
10 pg



10 ng



Identification of Long Transcripts



SMARTer cDNA Synthesis for NGS

dT Primed

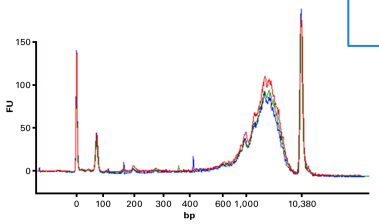
- mRNA
- Polyadenylated RNA
- Single cell capacity

N6 Primed

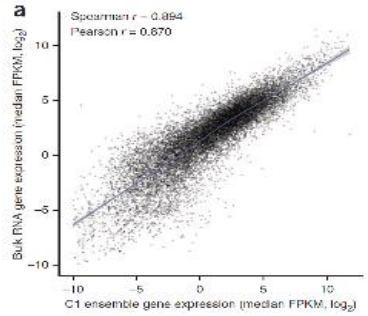
- Coding and non-coding RNA
- Non-polyadenylated RNA
- Degraded samples

SMART-Seq™ v4 Ultra™ Low Input RNA Kit for Sequencing

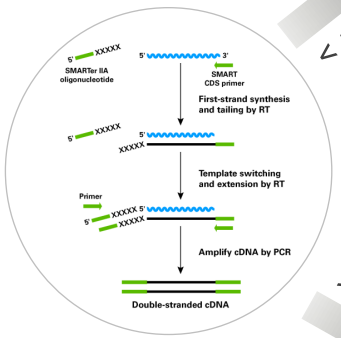
- Use with Ion Torrent or Illumina platforms



SMARTer Ultra Low RNA Kit for the Fluidigm C1™ System



- 96 single cells in parallel



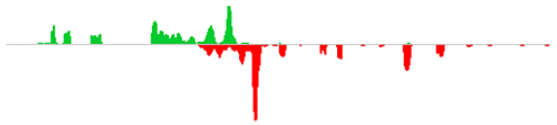
cDNA Synthesis

< 10 ng

> 10 ng

SMARTer Stranded Total RNA-Seq Kit - Pico Input Mammalian

- Use with Illumina platforms
- Use with highly degraded samples (FFPE)



SMARTer Stranded RNA-Seq Kit

- Use with Illumina platforms

RiboGone™ - Mammalian

SMARTer Stranded Total RNA Sample Prep- HI Mammalian

- Use with typical input RNA samples

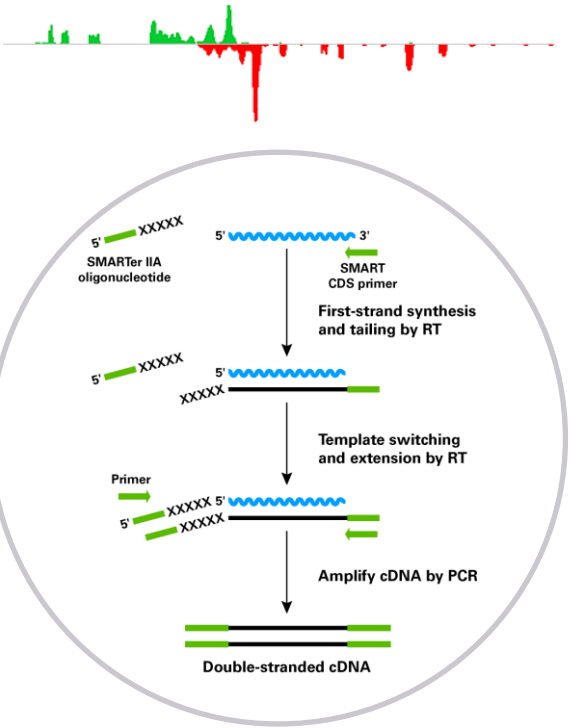
TaKaRa Clontech

Our Solutions for Total RNA-Seq

N6 Priming

- Coding and non-coding RNA
- Non-polyadenylated RNA
- Degraded, FFPE, and LCM samples

With Strand Information

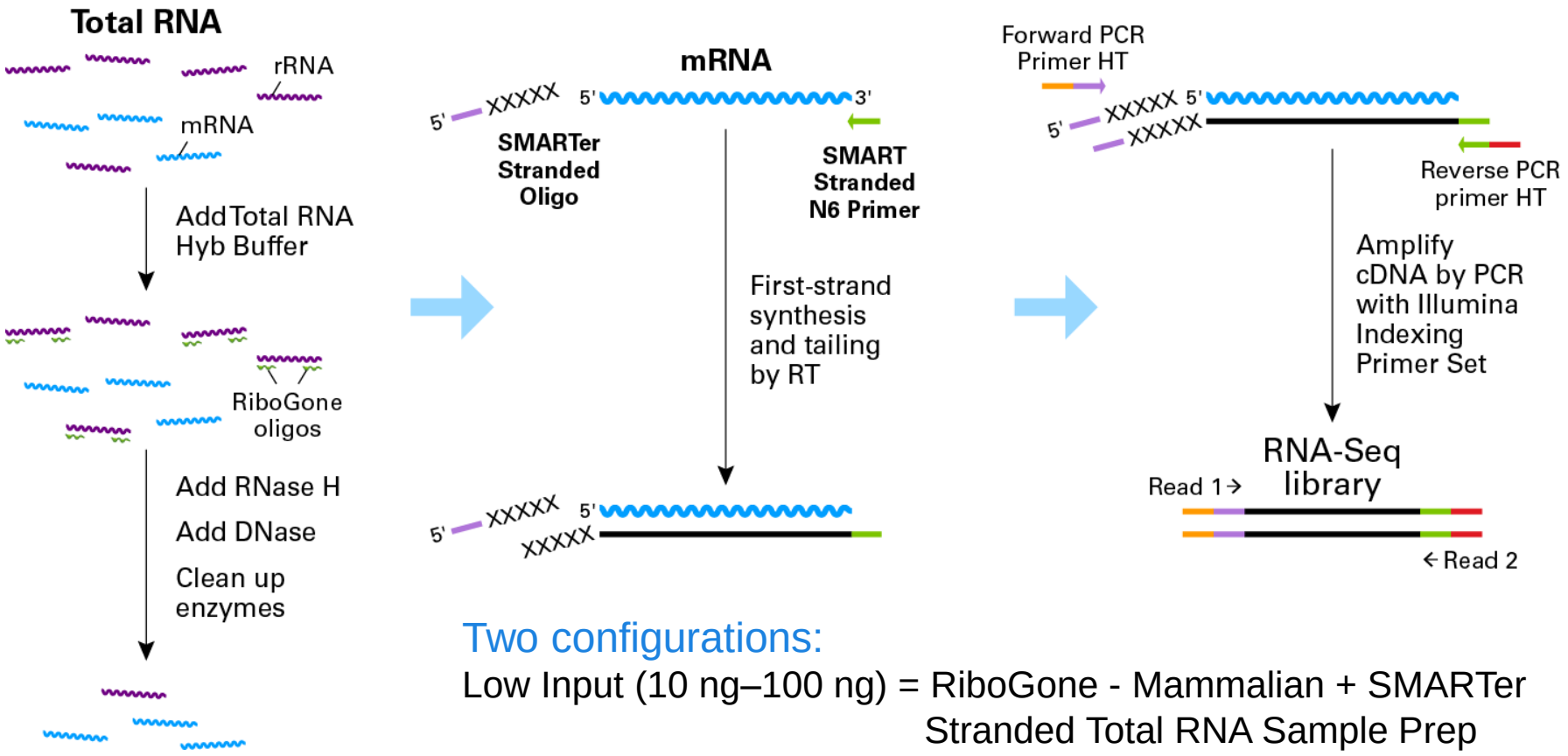


SMARTer Stranded Total RNA Sample Prep Kit - HI Mammalian
(RiboGone™ - Mammalian kit built in)
Input = 100 ng–1 μg of total RNA

SMARTer Stranded Total RNA Sample Prep Kit - Low Input Mammalian
(with RiboGone - Mammalian kit)
Input = 10 ng–100 ng of total RNA

SMARTer Stranded Total RNA-Seq Kit - Pico Input Mammalian
Input = 250 pg–10 ng of total RNA

SMARTer Stranded RNA-Seq with rRNA Depletion



Two configurations:

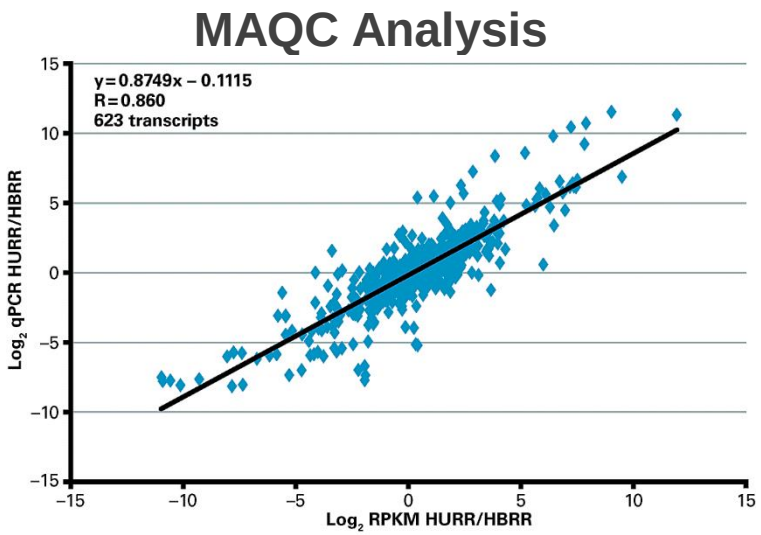
Low Input (10 ng–100 ng) = RiboGone - Mammalian + SMARTer Stranded Total RNA Sample Prep

High Input (100 ng–1 mg) = SMARTer Stranded Total RNA Sample Prep Kit - HI

SMARTer Stranded RNA-Seq for Low-Input Total RNA Samples

Analyses of sequencing data

Sequence Alignment Metrics (Input: 10 ng)		
RNA source	Human Universal	Human Brain
Number of reads (Millions)	6.8	7.7
Number of genes identified	14,563	13,839
Percentage of reads (%):		
Mapped to rRNA	0.9%	0.7%
Mapped to mtRNA	4.7%	2.9%
Mapped uniquely to genome	76%	75%
Mapped to exons	70%	66%
Mapped to introns	47%	49%
Mapped to intergenic regions	53%	51%

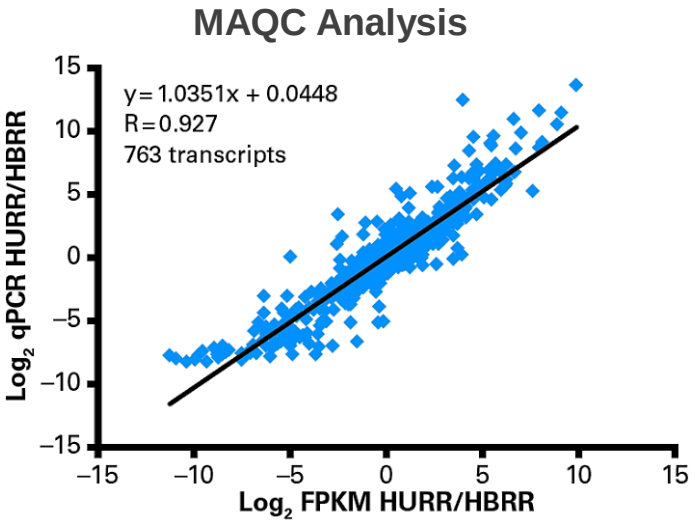


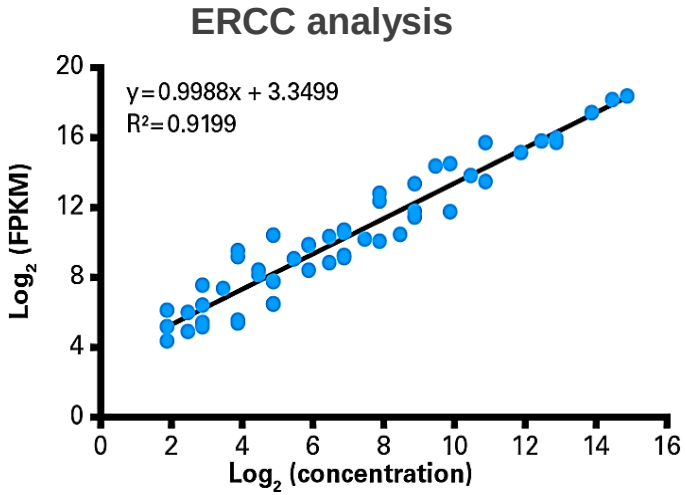
TaKaRa Clontech

SMARTer Stranded Total RNA Sample Prep Kit - HI Mammalian

Analyses of sequencing data

Sequence Alignment Metrics (Input: 400 ng)		
RNA source	Human Universal	Human Brain
Number of reads (Millions)	8.5 (paired-end)	
Number of genes identified	17,570	17,600
Percentage of reads (%):		
Mapped to rRNA	0.3%	5.3%
Mapped to genome	94%	88%
Mapped uniquely to genome	91%	84%
Mapped to exons	43%	50%
Mapped to introns	43%	33%
Mapped to intergenic regions	14%	12%
ERCC transcripts with correct strand	99.3%	98.8%





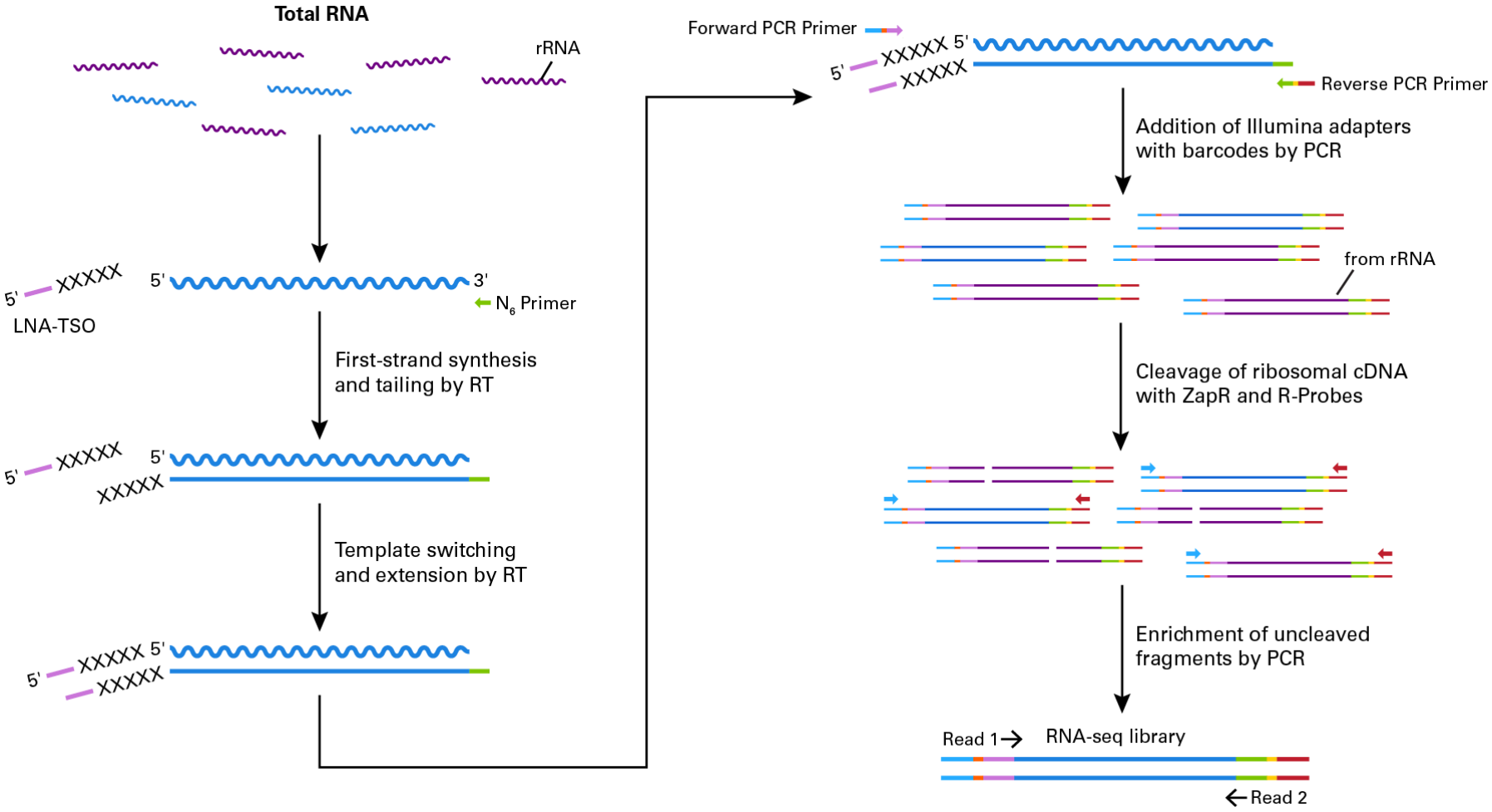
TaKaRa Clontech

SMARTer Stranded Total RNA-Seq Kit - Pico Input Mammalian

SMARTer Stranded Total RNA-Seq Kit - Pico Input Mammalian

- RNA-seq library prep kit directly resulting in Illumina-compatible libraries
- **Input = 250 pg–10 ng of mammalian total RNA**
- Incorporates LNA technology, leading to greater sensitivity
- Streamlined workflow, including the depletion of rRNA in the form of ribosomal cDNA using a novel, proprietary technology
- Maintains strand information
- Uses random priming to generate information from coding and non-coding RNA
- Compatible with a range of RNA qualities (e.g., FFPE & LCM samples)

Stranded Total RNA-Seq - Pico Input with ribosomal cDNA depletion



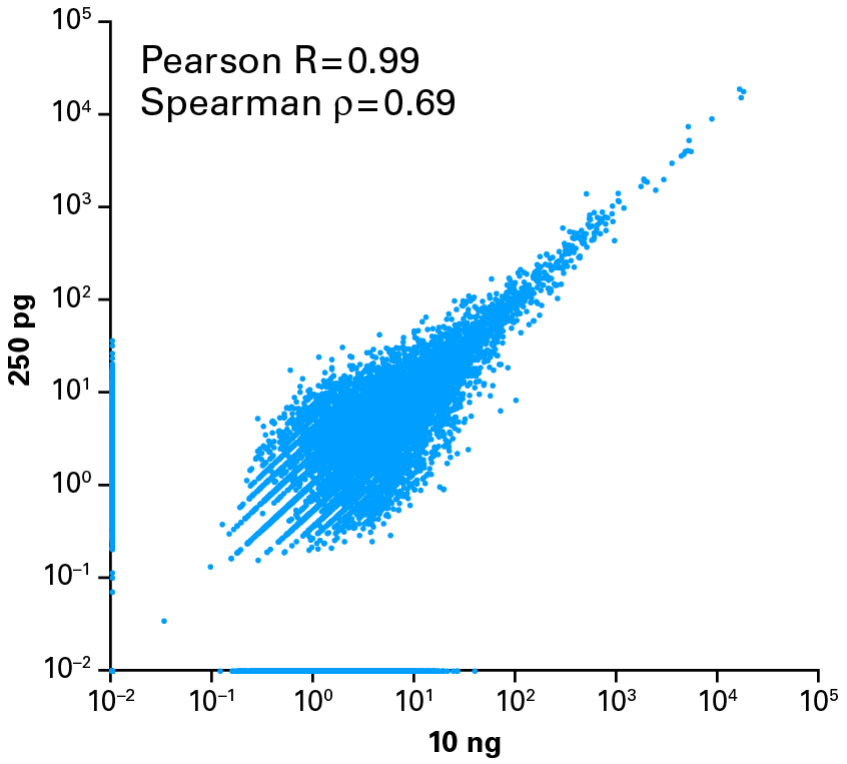
Consistent Sequencing Metrics Across a 100-Fold Input Range

Sequencing Alignment Metrics								
RNA source	Mouse Brain Total RNA							
Input amount (ng)	**10**		**1**		**0.25**		**0.1**	
Library yield (ng/µl)	10.5	14.8	9.93	8.3	6.91	7.48	5.76	7.26
Number of reads	2.6 million (paired-end)							
Number of transcripts FPKM >1	**12,714**	**12,709**	**12,744**	**12,725**	**12,540**	**12,615**	**12,286**	**12,528**
Pearson/Spearman correlations	**0.99/0.93**		**0.99/0.93**		**0.98/0.92**		**0.97/0.90**	
Correct strand per biol. annotation (%)	97.7	97.7	97.7	97.7	97.7	97.7	97.7	97.6
Proportion of total reads (%)								
Exonic	22.6	22.8	23.4	23.5	23.3	23.1	23.1	22.8
Intronic	35.6	35.7	35.3	36.2	35.9	35.5	36.1	35.1
Intergenic	8.3	8.2	8.2	8.2	8.0	8.0	7.8	7.8
rRNA	**11.2**	**10.5**	**10.8**	**9.9**	**9.7**	**9.7**	**8.8**	**9.5**
Mitochondrial	8.8	8.7	8.3	8.5	8.3	8.4	7.5	7.9
Overall mapping (%)	86.4	85.9	86.1	86.2	85.1	84.8	83.3	83.1
Duplicate rate (%)	12.8	12.5	17.3	17.8	31.3	28.8	44.2	40.2

Stranded Total RNA-Seq - Pico Input from FFPE Samples

Sequencing Alignment Metrics				
RNA source	Human Liver Total RNA - FFPE			
Input amount (ng)	**10**	**1**	**0.25**	**0.25**
Ribosomal cDNA removal	**Yes**			**No**
Number of reads	1 million (paired-end)			
Number of transcripts FPKM >1	**11,752**	**11,360**	**10,368**	**4,501**
Number of transcripts FPKM >0.1	**15,358**	**14,680**	**12,793**	**4,507**
Correct strand per biol. annotation (%)	98.3	98.1	98.3	97.1
Proportion of total reads (%)				
Exonic	23.9	24.3	21.3	1.7
Intronic	18.4	19.2	17.1	1.7
Intergenic	2.8	2.7	2.5	1.2
rRNA	36.7	34.3	34.4	90.1
Mitochondrial	4.0	3.5	3.1	1.4
Overall mapping (%)	85.8	84.0	78.5	96.1
Duplicate rate (%)	22.6	39.0	52.5	44.0

250 pg vs. 10 ng Human Liver Total FFPE RNA

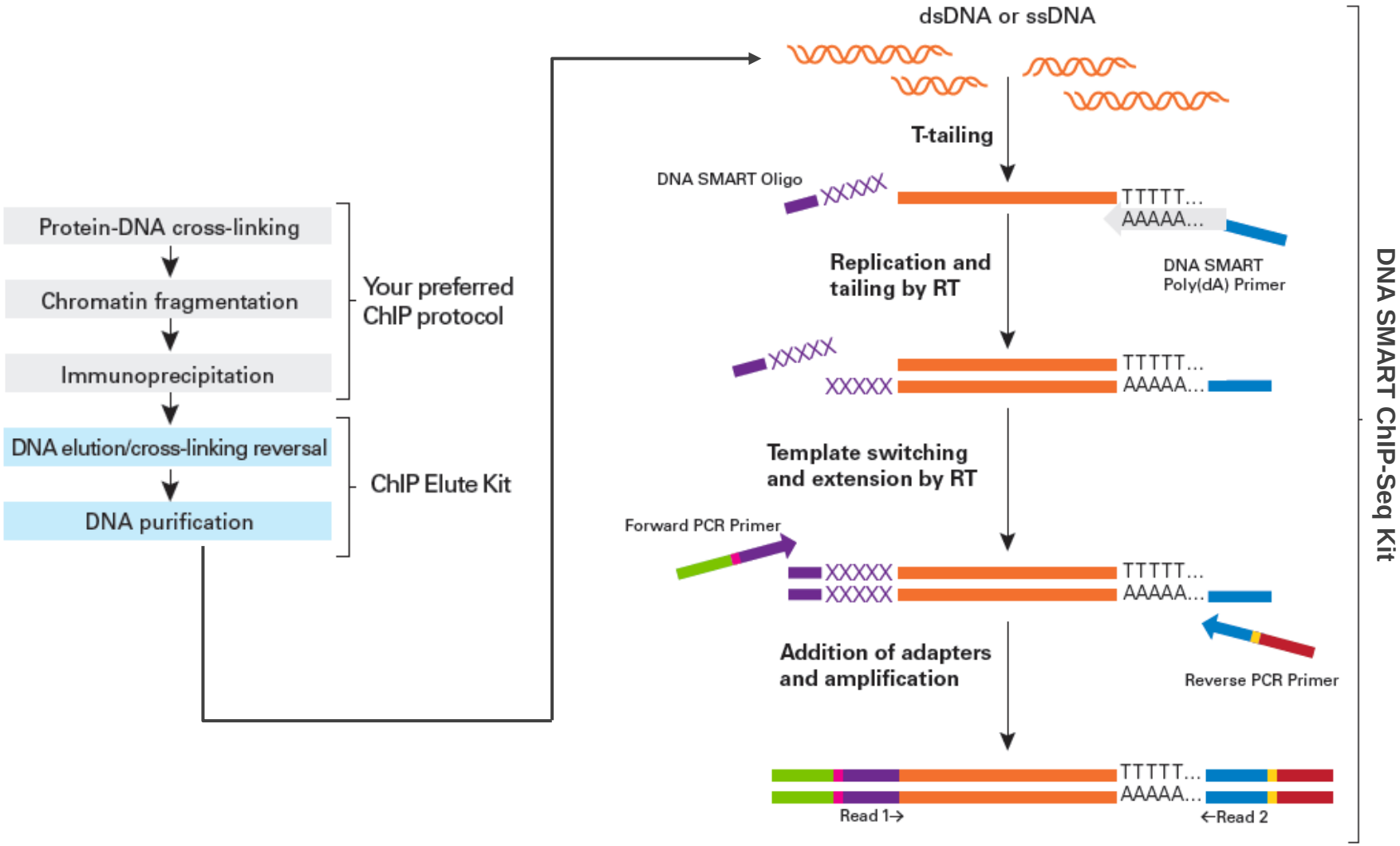


SMART Technology for DNA Sequencing—Expanding Applications

DNA SMART™ ChIP-Seq Kit

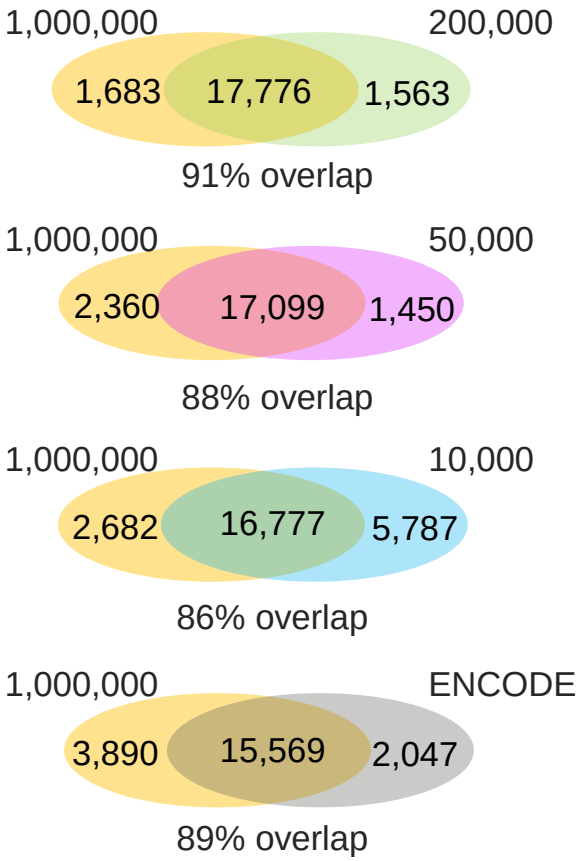
- For low-input ChIP-seq for Illumina platforms
- Single-tube workflow; under 4 hours
- Compatible with dsDNA or ssDNA (100 pg–10 ng)
- Ligation-free addition of Illumina adapters
- Generates high-complexity libraries from picogram amounts of input DNA

Combined ChIP Elute and ChIP-Seq Kits—Workflow



High-Quality Data from Low-Cell Number ChIP-Seq Experiments

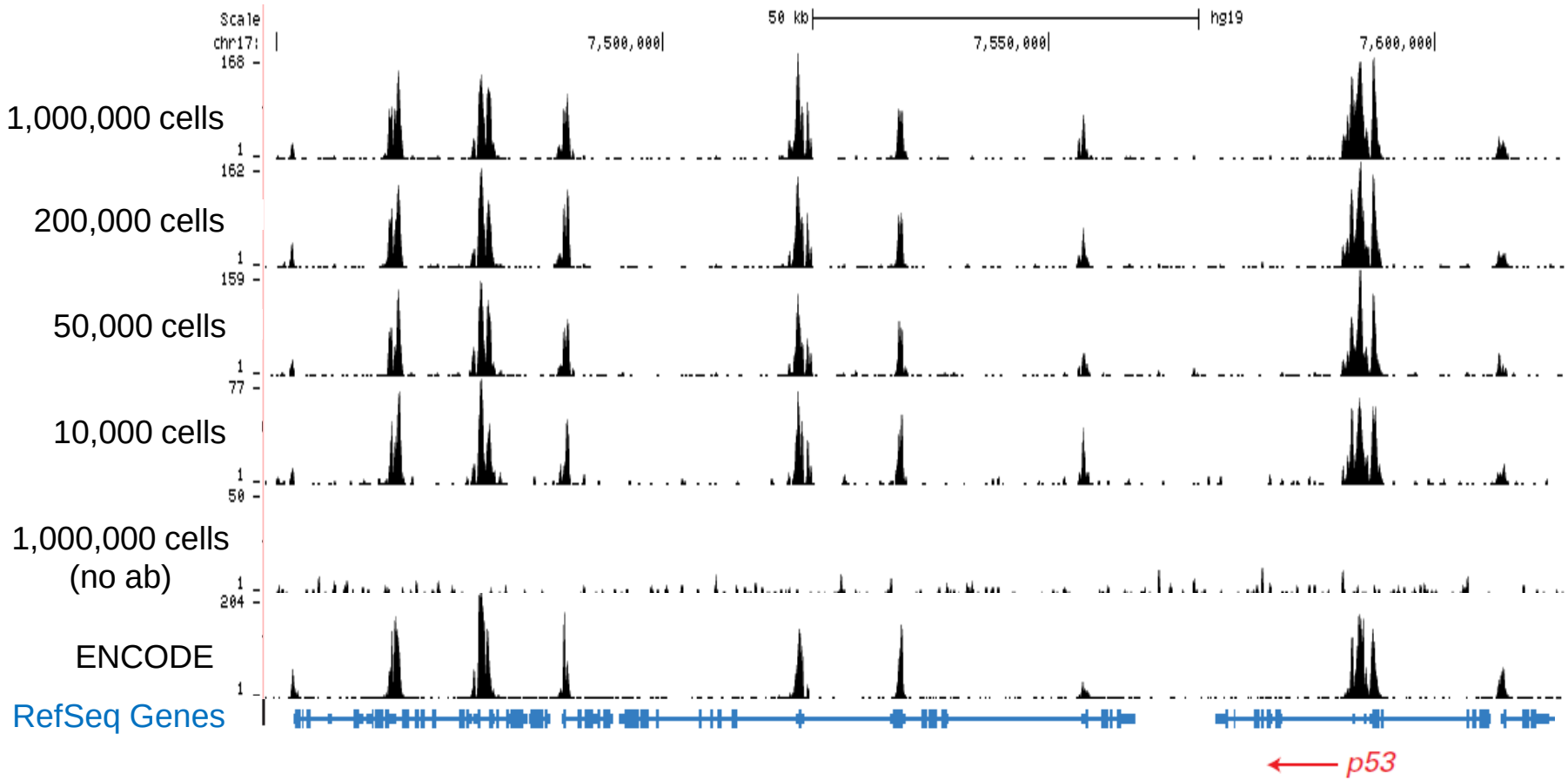
Sequencing Metrics from Specified Numbers of Cells				
ChIP antibody	H3K4me3			
Input (293T cells)	1,000,000	200,000	50,000	10,000
PCR cycles	15	18	18	18
Library yield (nM)	86.7	101	44.6	20.5
Peaks identified	19,459	19,339	18,549	22,564
Percentage of reads (%):				
Reads mapped	92.7	88.6	84.3	75.8
Uniquely mapped reads	79.0	74.8	70.4	59.6
Useful reads (uniquely mapped, non-duplicates)	66.8	63.5	49.9	34.0
Non-redundant rate	0.85	0.85	0.71	0.57



(ENCODE data from U. Washington—293T cells)
Analysis performed with 15–18 million reads per sample

Robust Libraries from Low-Cell-Number ChIP-Seq Experiments

Reproducibility is maintained for low cell numbers



Summary

Single-Cell-Capable mRNA-Seq Using SMART Technology

- SMARTer Ultra Low RNA Kit for the Fluidigm C_1 System
- SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing

Total RNA-Seq Applications

- SMARTer Stranded RNA-Seq Kit
- SMARTer Stranded Total RNA Sample Prep Kit - HI Mammalian
- SMARTer Stranded Total RNA-Seq Kit - Pico Input Mammalian
- SMARTer Universal Low Input RNA Kit for Sequencing

Expanding Applications for SMART Technology

- DNA SMART ChIP-Seq Kit

NGS Learning Resources

NGS Resource Portal

TECH NOTES

mRNA-seq with the highest sensitivity ▸

High-input total RNA-seq for Illumina® Platforms ▸

ChIP-seq library preparation ▸

See all tech notes ▸

WEBINARS

Stranded RNA-seq for low input samples ▸

New and improved single-cell RNA-seq ▸

Low-input RNA-seq with Ion Torrent platforms ▸

See all available webinars ▸

SCIENTIFIC POSTERS

SMARTer Ultra Low cDNA synthesis ▸

Comparison of low-input RNA-seq kits by ABRF ▸

SMARTer Universal application for FFPE samples ▸

See all scientific posters ▸

NEW!

The latest generation of ultra-low input mRNA-seq

Learn more ❯

FAQs

Answers to your questions about:

RNA-seq ▸

ChIP-seq ▸

TIPS TO GET SMARTer

Learn useful tricks for cDNA synthesis, like how to make a magnetic separator or identify the ideal pellet ▸

CITATION LISTS

Publications that have used SMARTer kits for transcriptome analysis ▸

SELECTION GUIDE

Which RNA-seq kit is right for you? View the Selection Guide ▸

NGS Resource Portal: www.clontech.com/ngs

TaKaRa Clontech