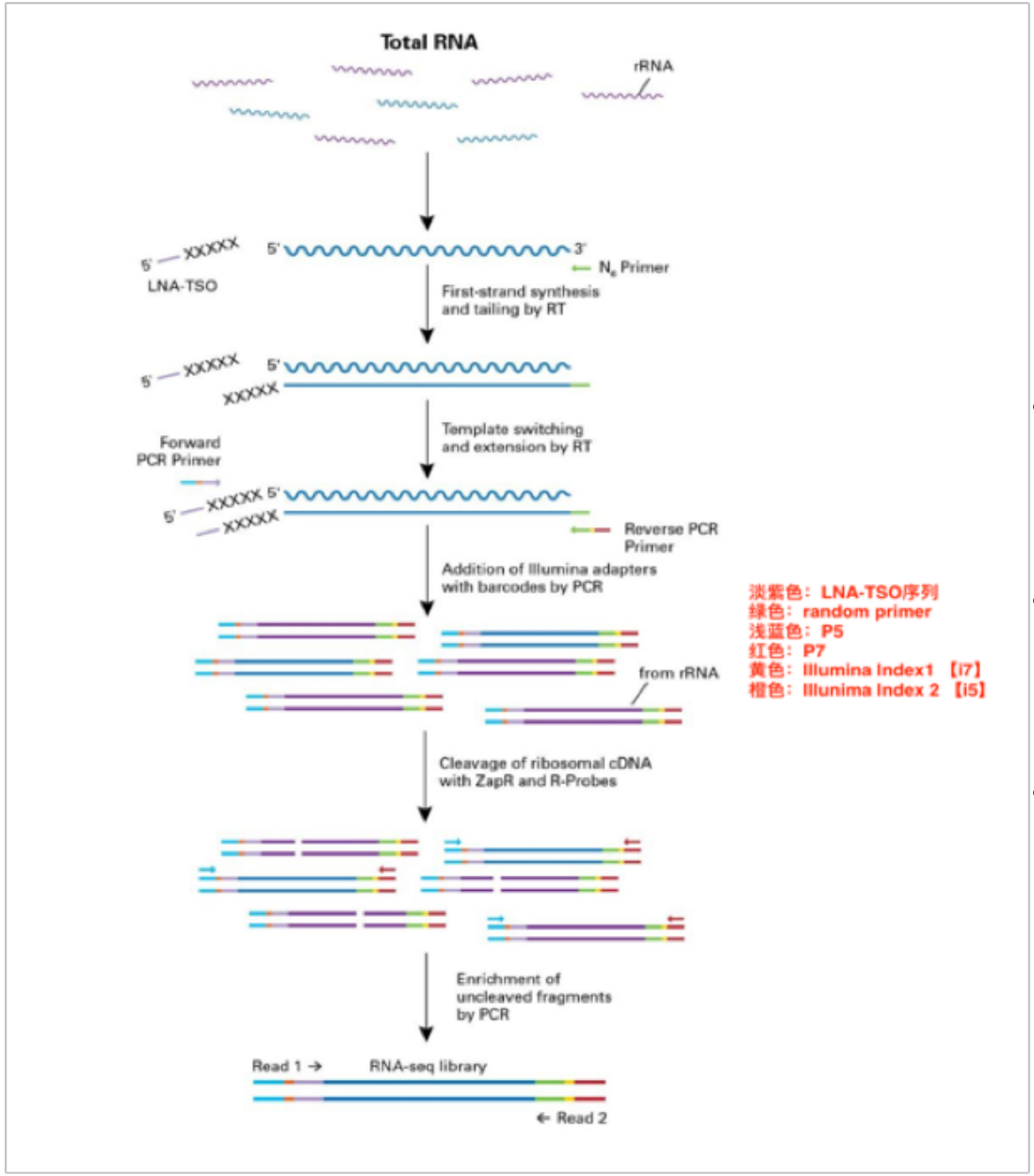
Research for lncRNA Library Generation kits TaKaRa _ SMARTer pico

Hu Xi

2018-12-28

Mechanism of remove ribosomal cDNA(SMARTer pico kit)



- Removal of ribosomal cDNA after cDNA synthesis using probes specific to mammalian rRNA.
- These R-Probes hybridize to ribosomal RNA and mitochondrial rRNA sequences
- The ribosomal cDNA is then cleaved by ZapR in the presence of the mammalian-specific R-Probes

Contents

1. Summary of current methods for removing rRNA (principle)
2. Summary and comparison with SMARTer pico kit

With Hongke Wang

3. Compare the principle of the two methods in the cDNA phase（PDD&DSN / DASH&Cas9）
4. Probe design method for PDD&DSN
5. Probe design & generation method for DASH&Cas9
6. **Compare the differences between the two methods**
7. **Designed a feasible test method**

1.Summary of current methods for removing rRNA (principle)

Depletion of abundant sequences

Summary from papers

- **Biotinylated capture probe —— rRNA**
- **RNase H & DNase I digestion — rRNA**

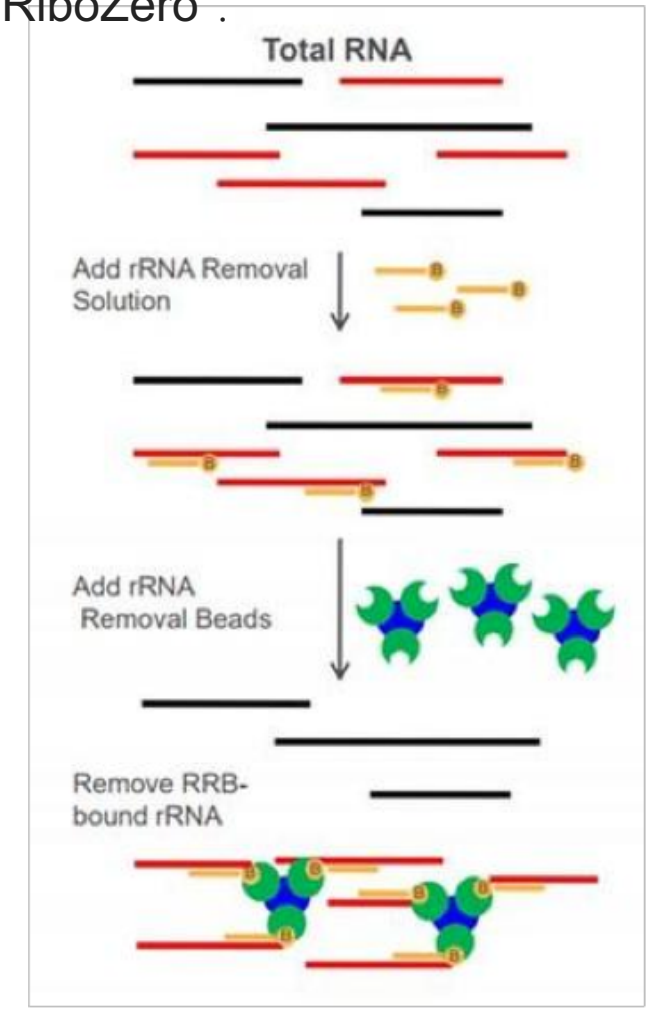
  → Commercial kits

  Illumina' s RiboZero/Qiagen GeneRead rRNA depletion /Lexogen RiboCop /NEBNext rRNA depletion /Roche /Kapa RiboErase /Takara/Clontech' s RiboGone

- **Probe-Directed Degradation (PDD)&Duplex-specific nuclease(DSN)**
- **Depletion of Abundant Sequences by Hybridization (DASH)& CRISPR/Cas9**

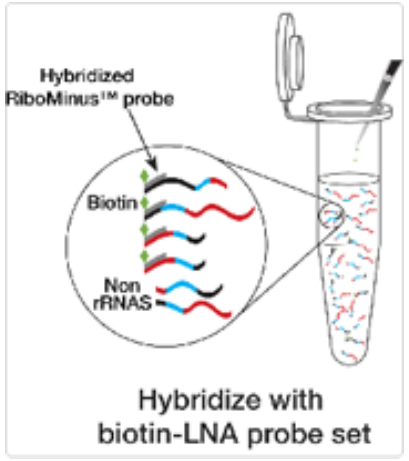
Biotinylated capture probe

Illumina RiboZero：

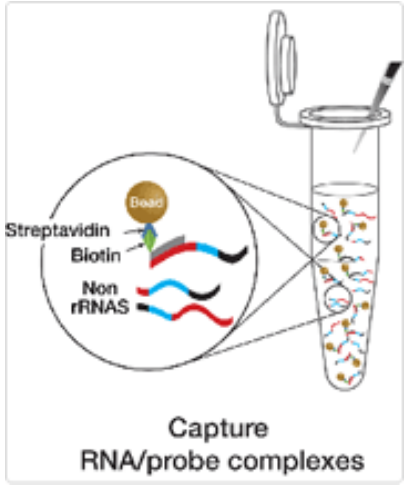


Thermo RiboMinus：

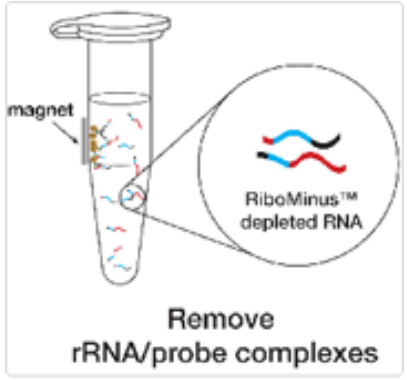
Step 1



Step 2

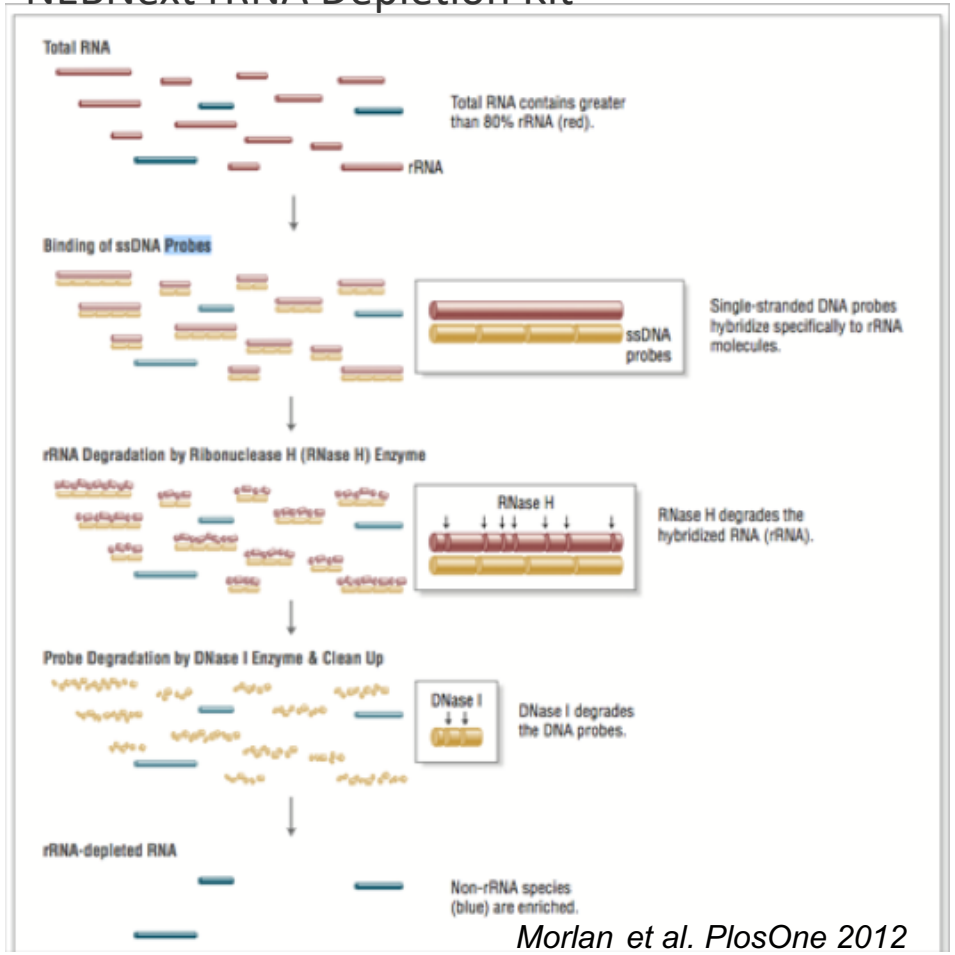


Step 3



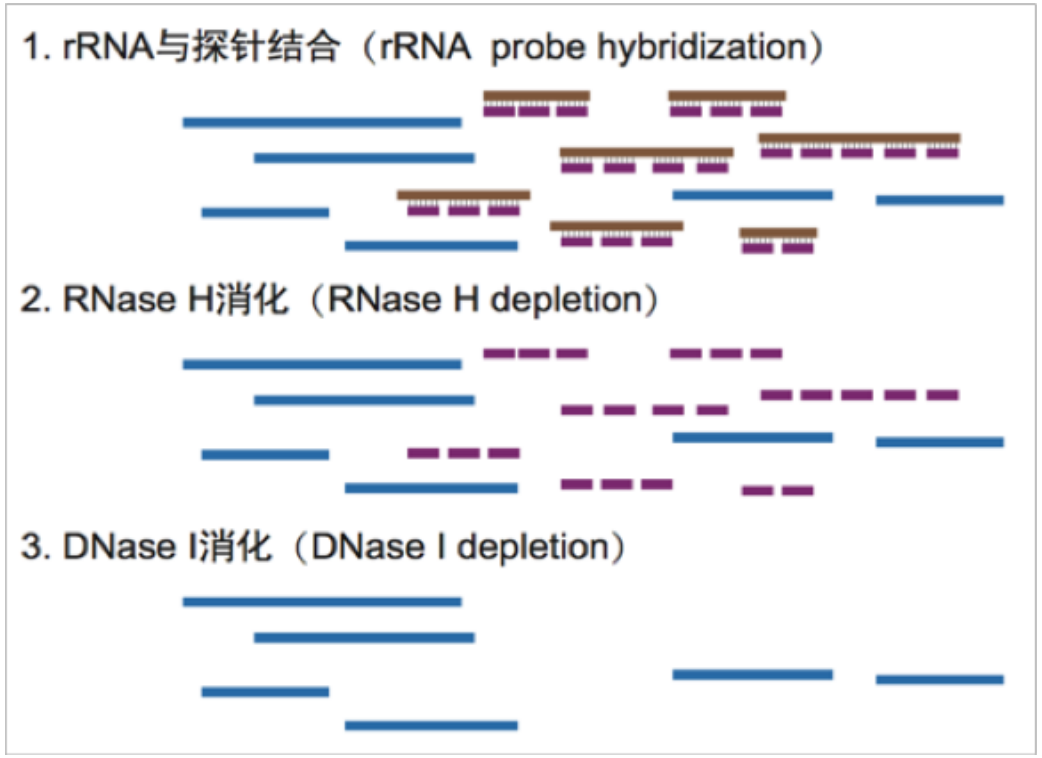
RNase H & DNase I digestion

NEBNext rRNA Depletion Kit



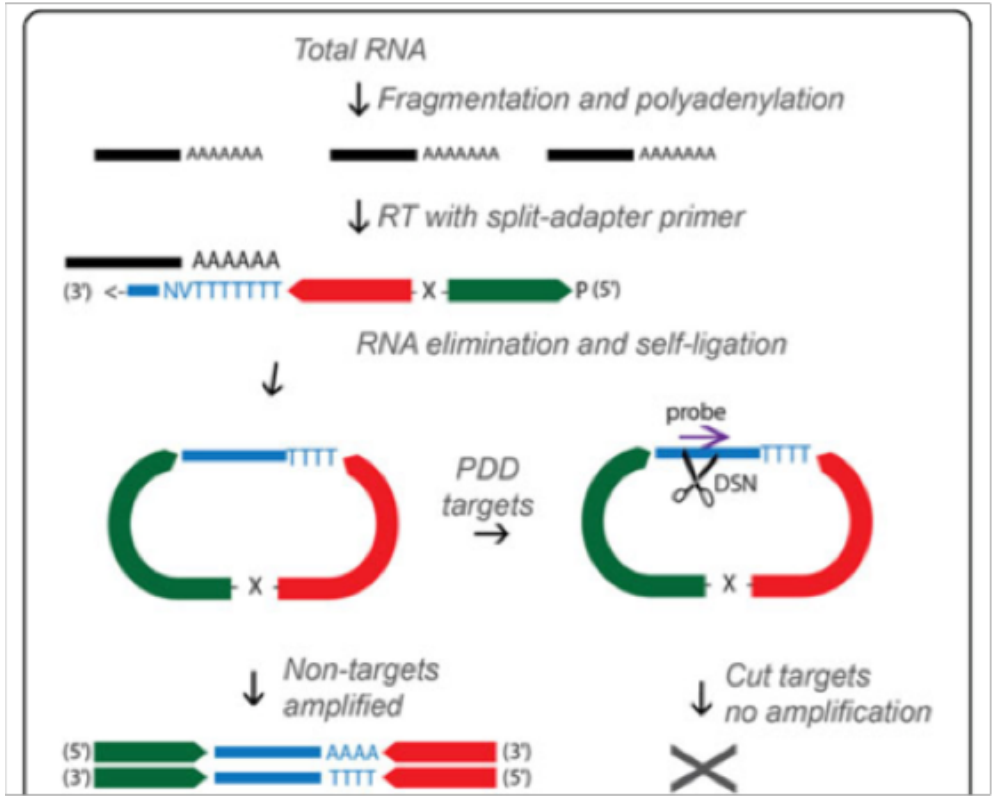
Morlan et al. PlosOne 2012

Vazyme Ribo-off rRNA Depletion Kit



- RNase H is an endoribonuclease that **specifically hydrolyzes the phosphodiester bonds of RNA which is hybridized to DNA**. This enzyme does not digest single or double-stranded DNA.

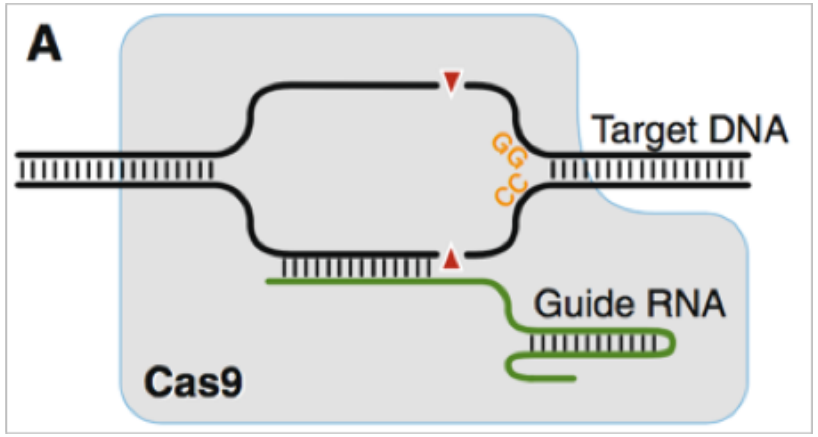
Probe-Directed Degradation (PDD)&Duplex-specific nuclease

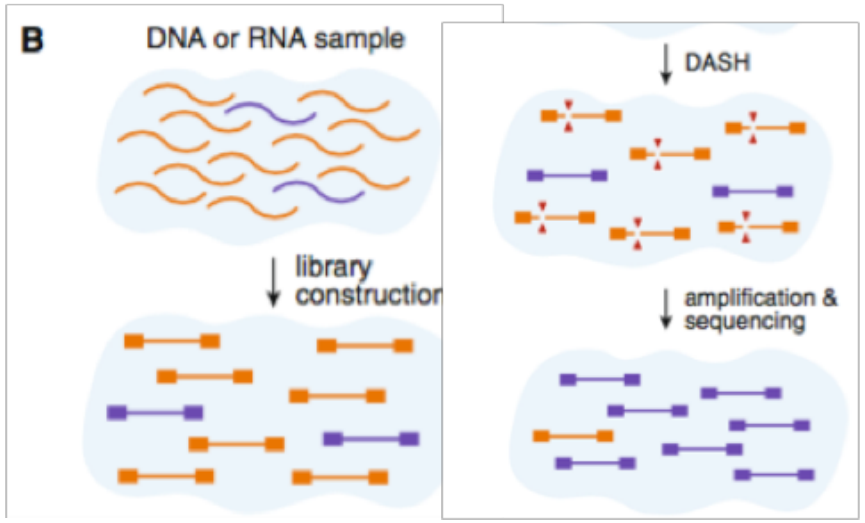


Archer et al. BMC Genomics 2014

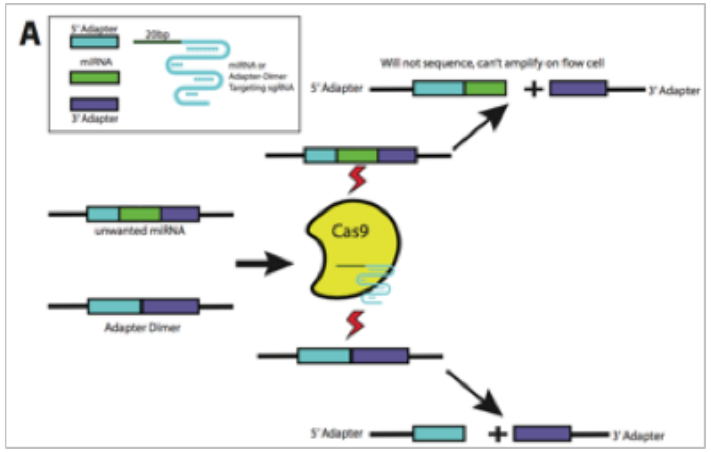
DSN(Duplex-specific nuclease)shows a strong preference for cleaving double-stranded **(ds) DNA** and **DNA in DNA-RNA hybrid duplexes**, and is practically inactive towards single-stranded (ss) DNA or single or double-stranded RNA.

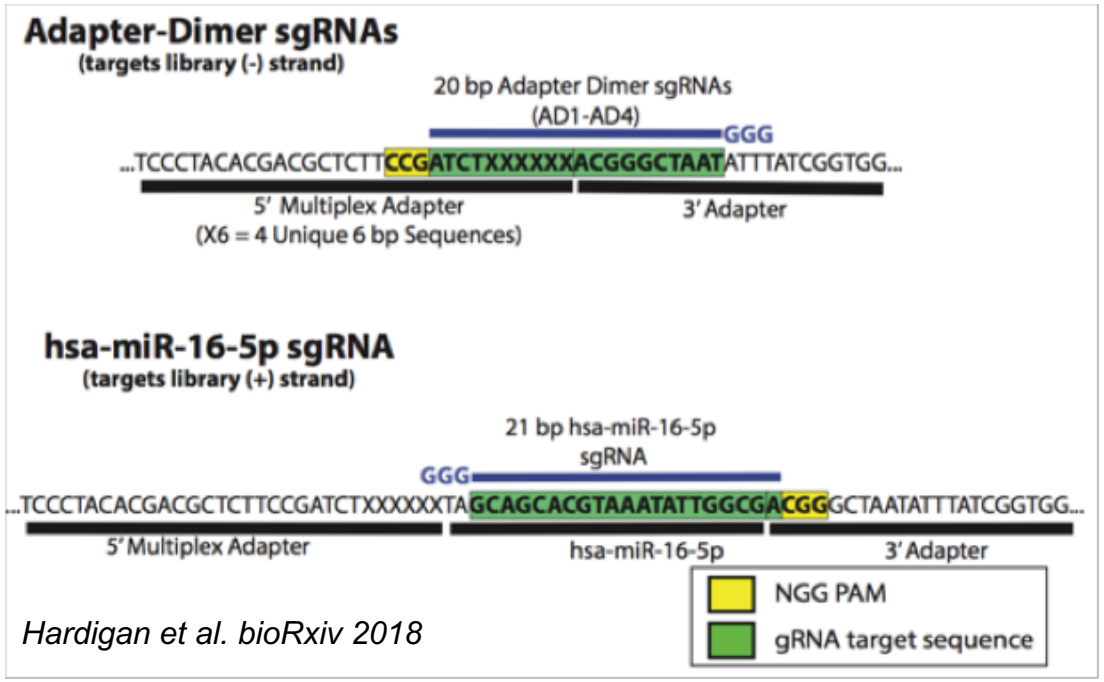
Depletion of abundant sequences _ CRISPR/Cas9





Gu et al. Genome Biology 2016





Hardigan et al. bioRxiv 2018

2.Summary and comparison with SMARTer pico kit

Method	Stage	Input	Performance		Note
Biotinylated capture probe	RNA	>100ng	~ 87%	Biotinylated probe ：rRNA / Streptavidin Beads	Ribo-Zero™ Magnetic Gold Kit(Human/Mouse/Rat)，24 Rxns
RNase H & DNase I digestion	RNA	5ng to 1ug	~ 99%	Probes ：rRNA / RNase H	NEBNext rRNA Depletion Kit (Human/Mouse/Rat) 6 Rxns
PDD & DSN（2014）	ssDNA	500ng（start）	~ 94%	Probes：cDNA / DSN；	**Incubation temperature：75℃—48℃（Gradient cooling），1h** **Termination condition：stop solution**
DASH&Cas9 *	dsDNA	<10ng	~ 99%	sgRNA：cDNA / “NGG” PAM motif： Cas9	**Incubation temperature：37℃，2h；** **Termination condition：heat or stop solution**
TaKaRa/Clontech SMARTer pico**	dsDNA	250 pg to 10 ng	~ 95%	R-Probes：cDNA / ZapR	**Incubation temperature：37℃，1h；** **Termination condition：heat**

*samples from plasma ** E-mail consultation

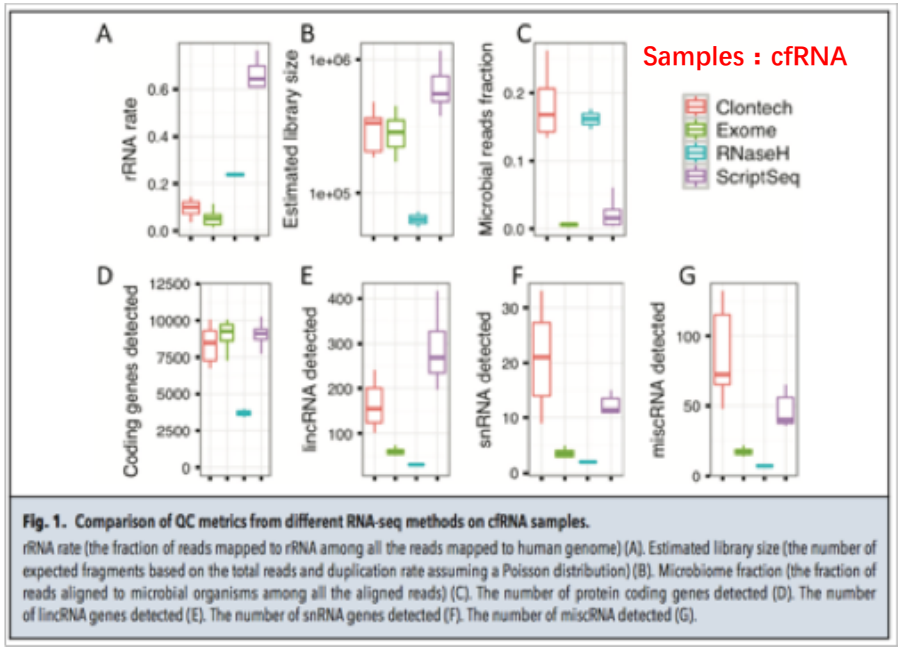
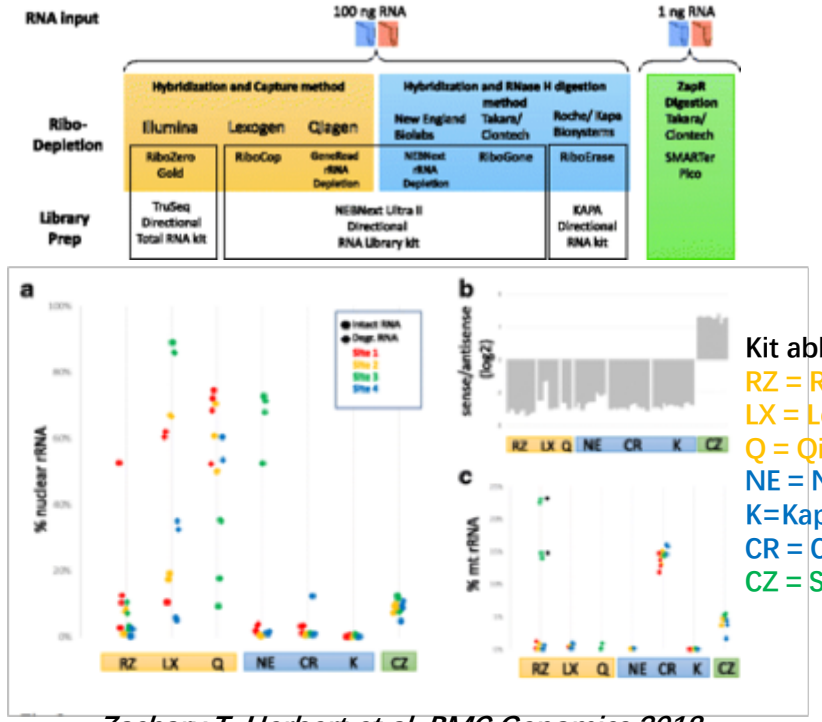


Fig. 1. Comparison of QC metrics from different RNA-seq methods on cfRNA samples.
rRNA rate (the fraction of reads mapped to rRNA among all the reads mapped to human genome) (A). Estimated library size (the number of expected fragments based on the total reads and duplication rate assuming a Poisson distribution) (B). Microbiome fraction (the fraction of reads aligned to microbial organisms among all the aligned reads) (C). The number of protein coding genes detected (D). The number of lincRNA genes detected (E). The number of snRNA genes detected (F). The number of miscRNA detected (G).

Wenying Pan, et al. Clinical chemistry. 2017

Roche NimbleGen SeqCap EZ Library SR

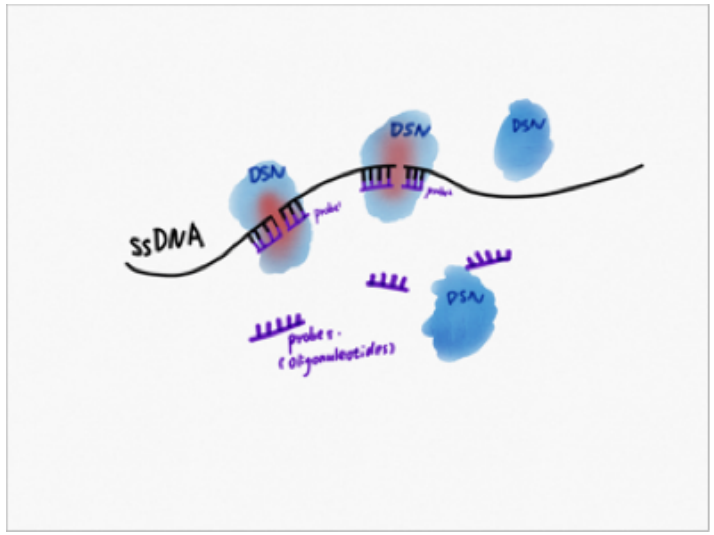
Both the Exome and Clontech methods were effective and sensitivity.

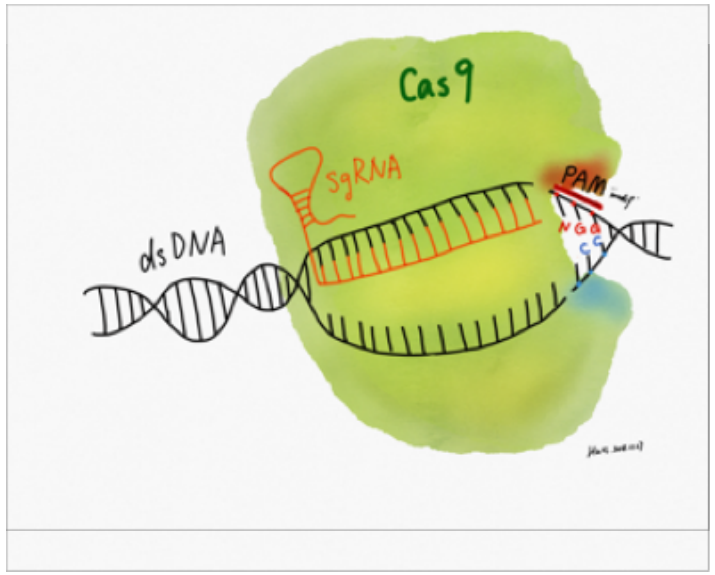


Kit abbreviations:
RZ = RiboZero Gold,
LX = Lexogen RiboCop,
Q = Qiagen GeneRead rRNA Depletion,
NE = NEBNext rRNA Depletion,
K=Kapa RiboErase,
CR = Clontech Ribogone,
CZ = SMARTer Pico total RNA

Zachary T. Herbert et al. BMC Genomics 2018

3.Compare the principle of the two methods in the cDNA phase（PDD&DSN / DASH&Cas9）





INPUT

cDNA Library form
Reverse Transcription

cDNA Library form
PCR Amplification

PDD&DSN

RNase A remove RNA from cDNA library

Hybrid incubation and DSN treatment

Purification of cDNA

DASH&Cas9

Hybrid incubation with sgRNA and Cas9

RNase A remove sgRNA

Purification of cDNA

OUTPUT

PCR Amplification

Gel separation selection wanted size sequence

4.PDD&DSN

Probes design

NCBI BLAST/ ApE (A Plasmid Editor, v2.0.47)/Biopython (www.biopython.org)

Define number of PDD probes

Calculate optimal probe density

The maximum inter-probe distance=the minimum library insert size - 20 nt

The period (distance from the start of one probe to the start of the next) =this distance +the minimum probe length

Allow coverage of nearly 10 kb of target sequence using 50 probes

Identify regions of cross-homology in the target transcript

(www.ncbi.nlm.nih.gov/blast)

Enter
the target sequence

Database:refseq_rna

Choose the species

"Optimize for" → "somewhat similar sequences (blastn)"

Algorithm parameters:
"Max target sequences" to 10000;
"Expect threshold" to 100000;
"word size" to 7;
"match/mismatch scores" to 1/−4;
"Gap costs" to "existence: 5 extension: 2"

BLAST

Results export as an "XML"

find short stretches of high homology to other transcripts in the target RNA

Highlight cross-homology in ApE

Python script: blast2ape.py https://github.com/stu2/xml2ape

In ApE:

Paste the target RNA sequence → **save it**

Download the BLAST hit table text file

Add the xml2ape.Py script

Execute the script

Produce an ApE file

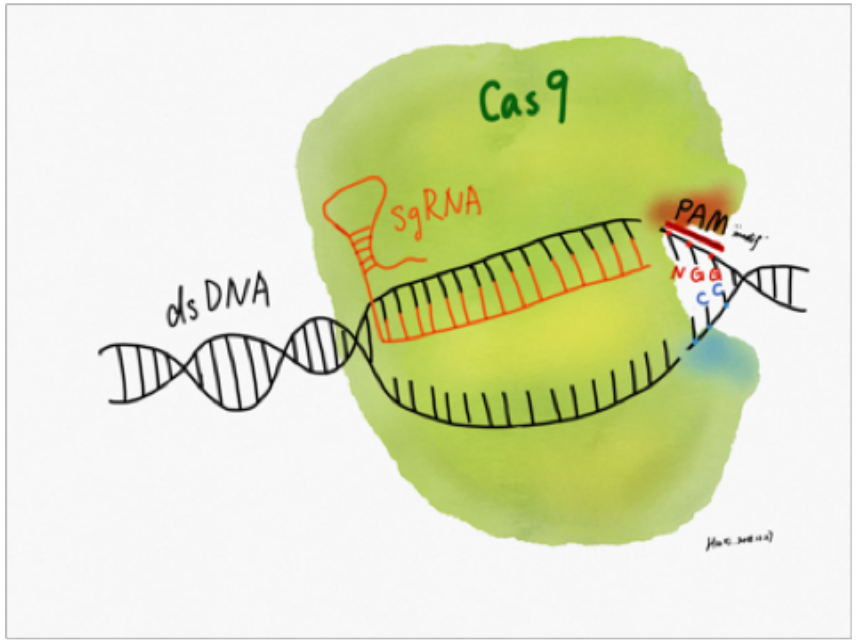
Select probes in ApE

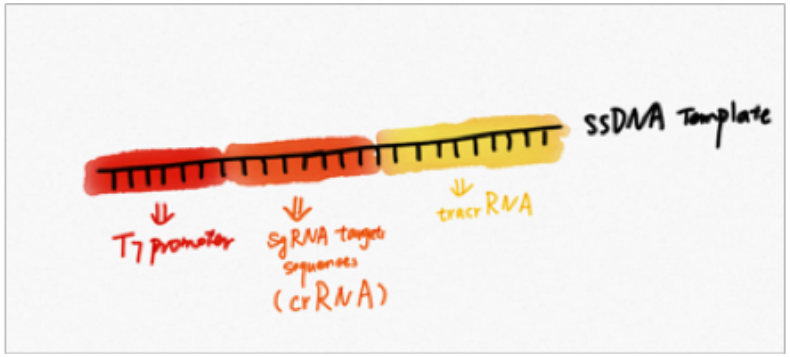
5.DASH&Cas9

sgRNAs design & generation

https://zlab.bio/guide-design-resources

SnapGene
display design rules





Find and identify targeting sequences

↓

Design sgRNA Transcription template

↓

Synthetic DNA template

↓

PCR template amplification

↓

In vitro transcription (IVT)

↓

Digesting DNA

↓

Purified RNA

6.Compare the differences between the two methods

CONTENT	PDD&DSN	DASH &Cas9
Probes types	DNA，~15nt	RNA,~20nt（targets sequences）
Probes preparation method	Company synthesis	Company synthesis & Self preparation
Probe storage condition	-20C°，12 months	-80C°，6 months
Target sequence sequence condition	>35nt, Suitable for lncRNA-seq	no request，Suitable for lncRNA-seq & smRNA-seq
Ability to remove Adapter dimer**	YES	YES
Key reagents	DSN(Evrogen,EA003-10U):¥1820/10 reactions HEPES-KOH (Sigma, H0527-25G):¥644.67	MEGAscript™ T7 Transcription Kit（Invitrogen，AM1333):¥3122/25 reactions TURBO™ DNase（Invitrogen，AM2238):¥1805/500 reactions MEGAclear™ Transcription Clean-Up Kit（Invitrogen，AM1908):¥1720/50 reactions S. pyogenes Cas9(PNABio,CPO2-250ug):¥2540.27/50 reactions NEBuffer™ 3.1(B7203S):¥235.2/2500 reactions

**Adapter dimer vs Primer dimer

~¥10/probe　~¥182/reaction　~¥220（~¥85_TaKaRa）/probe*　~¥52/reaction

*idea：Design universal primers to amplify DNA templates

7.Designed a feasible test method

Samples： Standard RNA，such as Spike In

Methods： Choose two RNA sequences

rS1=Control sequences

rS2=Target sequences → Probes design

Reverse Transcription / RT-PCR

DSN / Cas9

PCR Amplification/qPCR

Gel electrophoresis/Quantitative analysis