Group Assignments

- Non-self-assembled
  - show up next week
- Self-Assembled (optional):
  - email josh@duke.edu by noon on Sunday (7/9)
    - proposed group members
    - expertise of each member (e.g. bio, stats, compute)

RNA-Seq Sample Preparation Theory

From Sample to Data

Overview

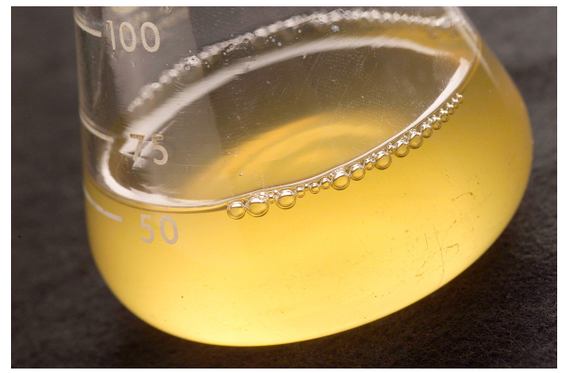
- From Library to Data
- Illumina Sequencing

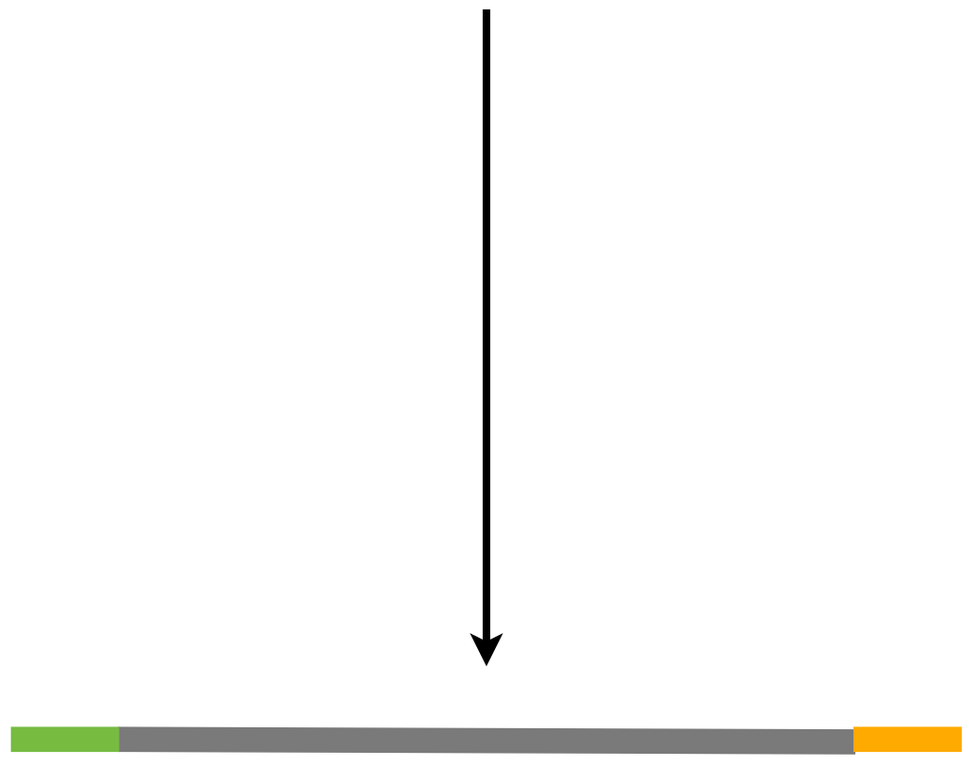
Not Appearing Today

- Data Analysis
- Software
- Details of our experiment

Major Experiment Components

1. Growth
2. Sample Collection
3. RNA Extraction
4. Ribosomal RNA Depletion
5. Library Preparation

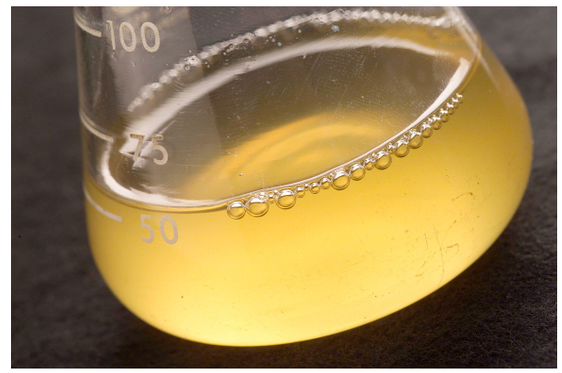


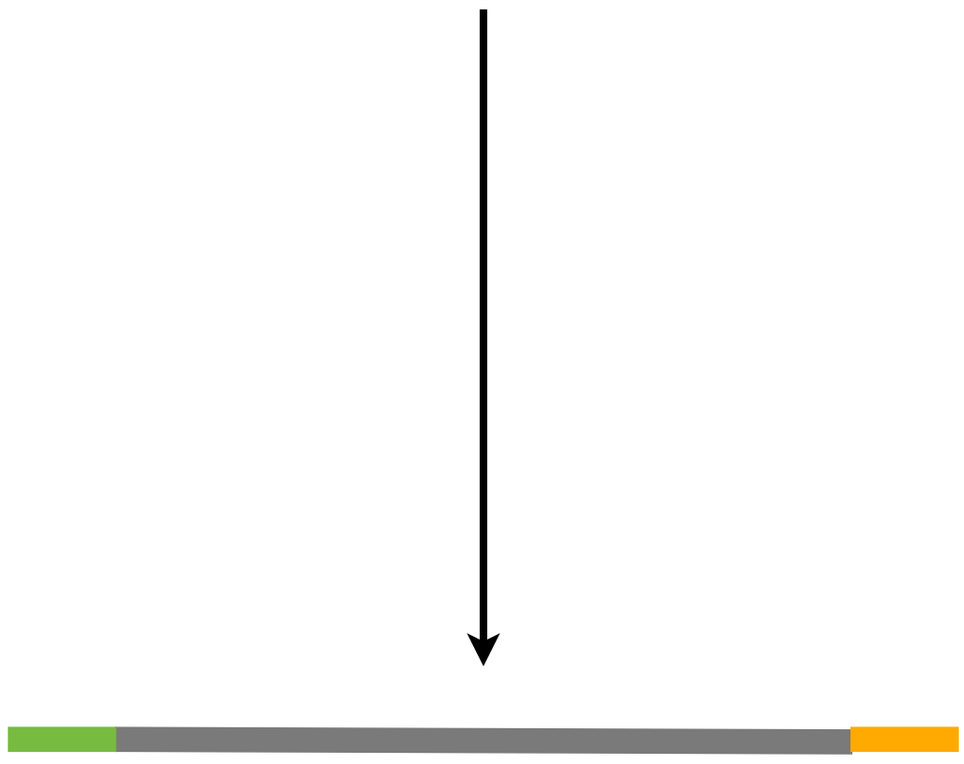


http://cse.ksu.edu/REU/S11/bvosburg/FLASK.jpg

Major Experiment Components

1. Growth
2. Sample Collection
3. RNA Extraction
4. Ribosomal RNA Depletion
5. Library Preparation





http://cse.ksu.edu/REU/S11/bvosburg/FLASK.jpg

Growth and Sample Collection

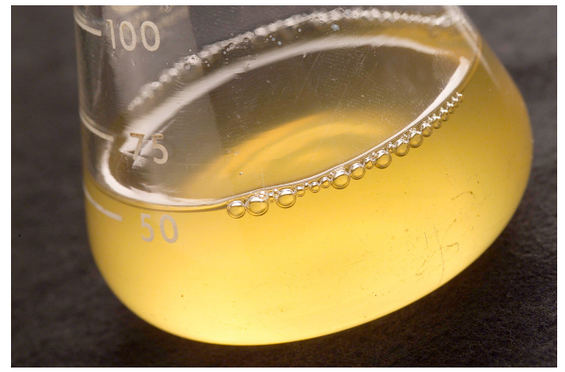
- Avoid Confounding Factors!
  - System Specific
  - Experiment Specific
  - Avoid RNA response to sample collection!

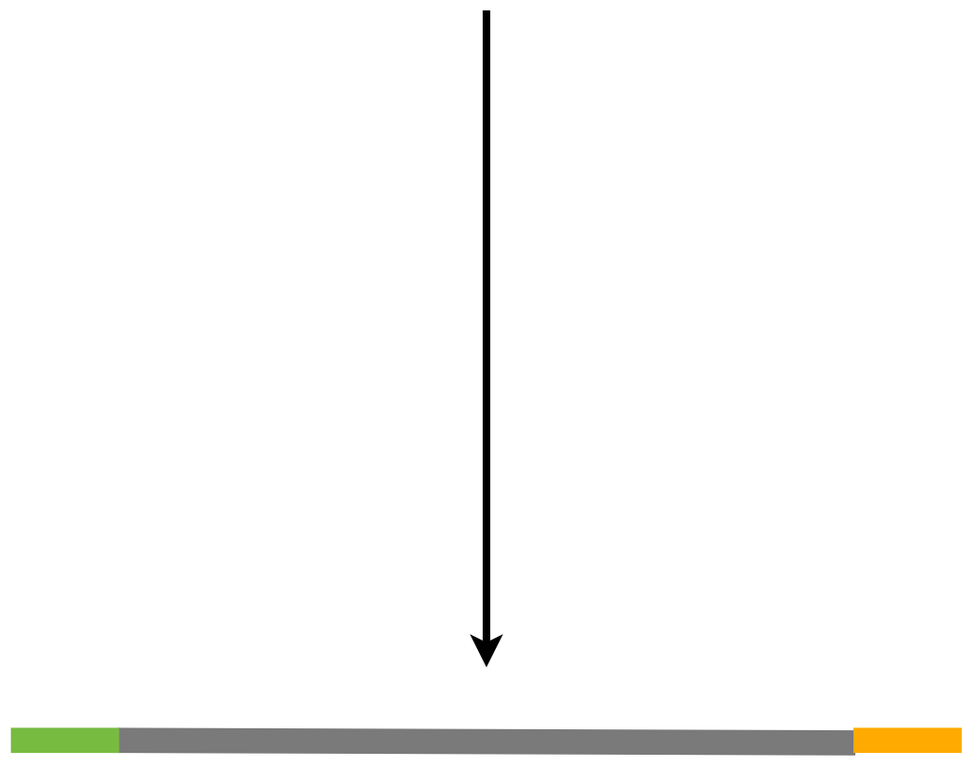
Sample Collection Options

- Flash freeze
- RNA stabilizers
    - RNA protect
    - RNAlater
- Phenol (hot acid phenol, trizol, etc)

Major Experiment Components

1. Growth
2. Sample Collection
3. RNA Extraction
4. Ribosomal RNA Depletion
5. Library Preparation





http://cse.ksu.edu/REU/S11/bvosburg/FLASK.jpg

RNA Extraction: Why?

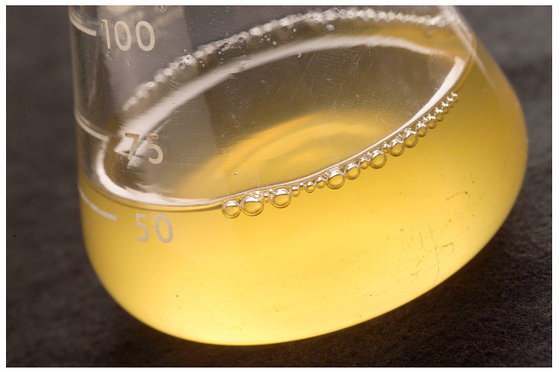
- Have cells, need RNA

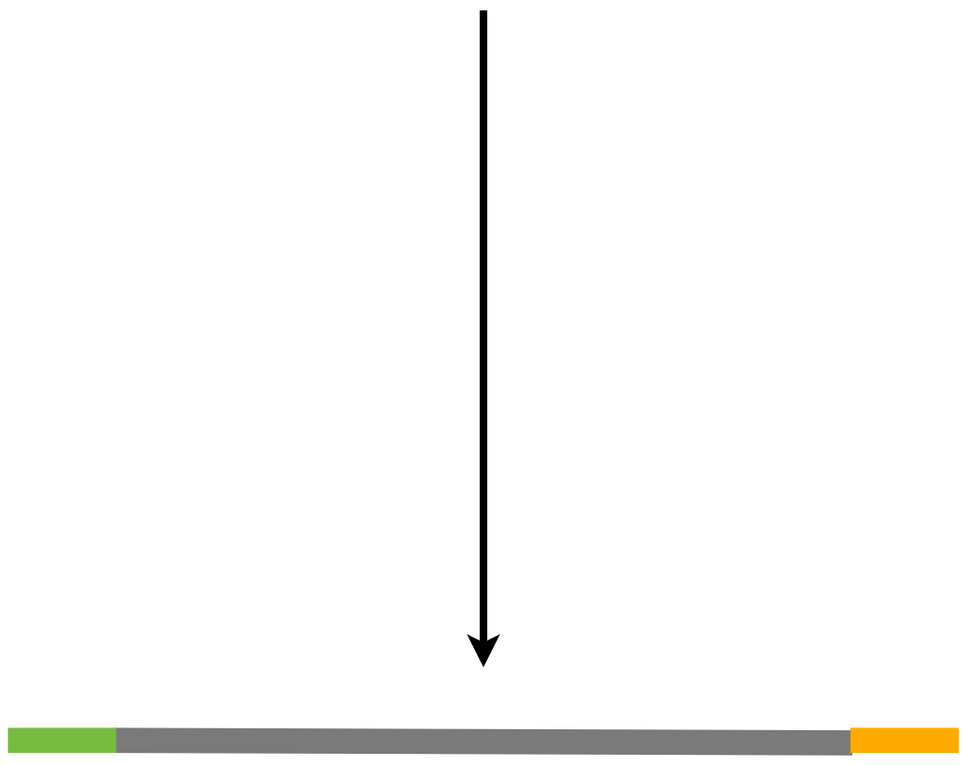
RNA Extraction Options

- Kits
  - Qiagen RNeasy Mini Kit
  - Etc
- Phenol (hot acid phenol, trizol, etc)

Major Experiment Components

1. Growth
2. Sample Collection
3. RNA Extraction
4. **Ribosomal RNA Depletion**
5. Library Preparation

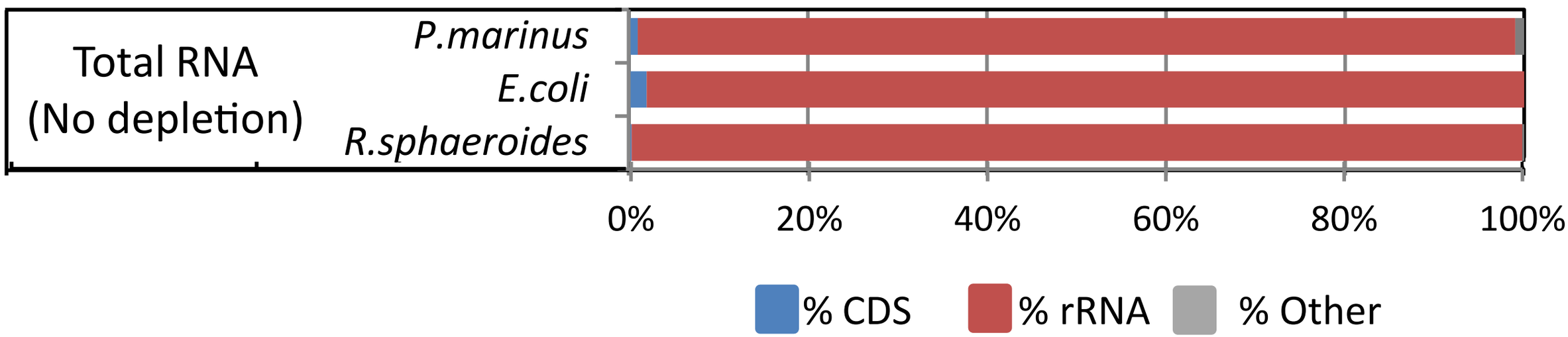




http://cse.ksu.edu/REU/S11/bvosburg/FLASK.jpg

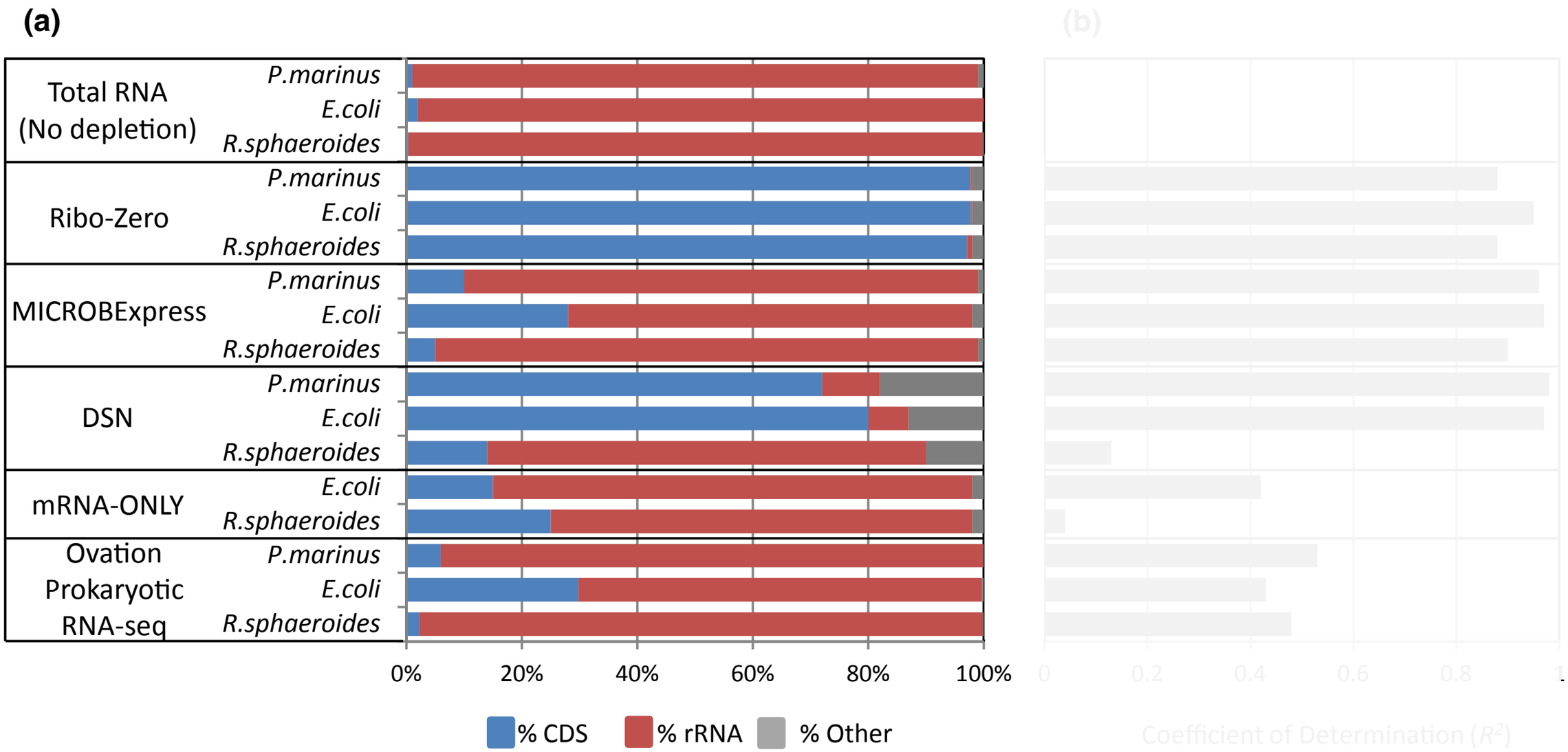
Ribosomal RNA Depletion

rRNA Depletion: Why?



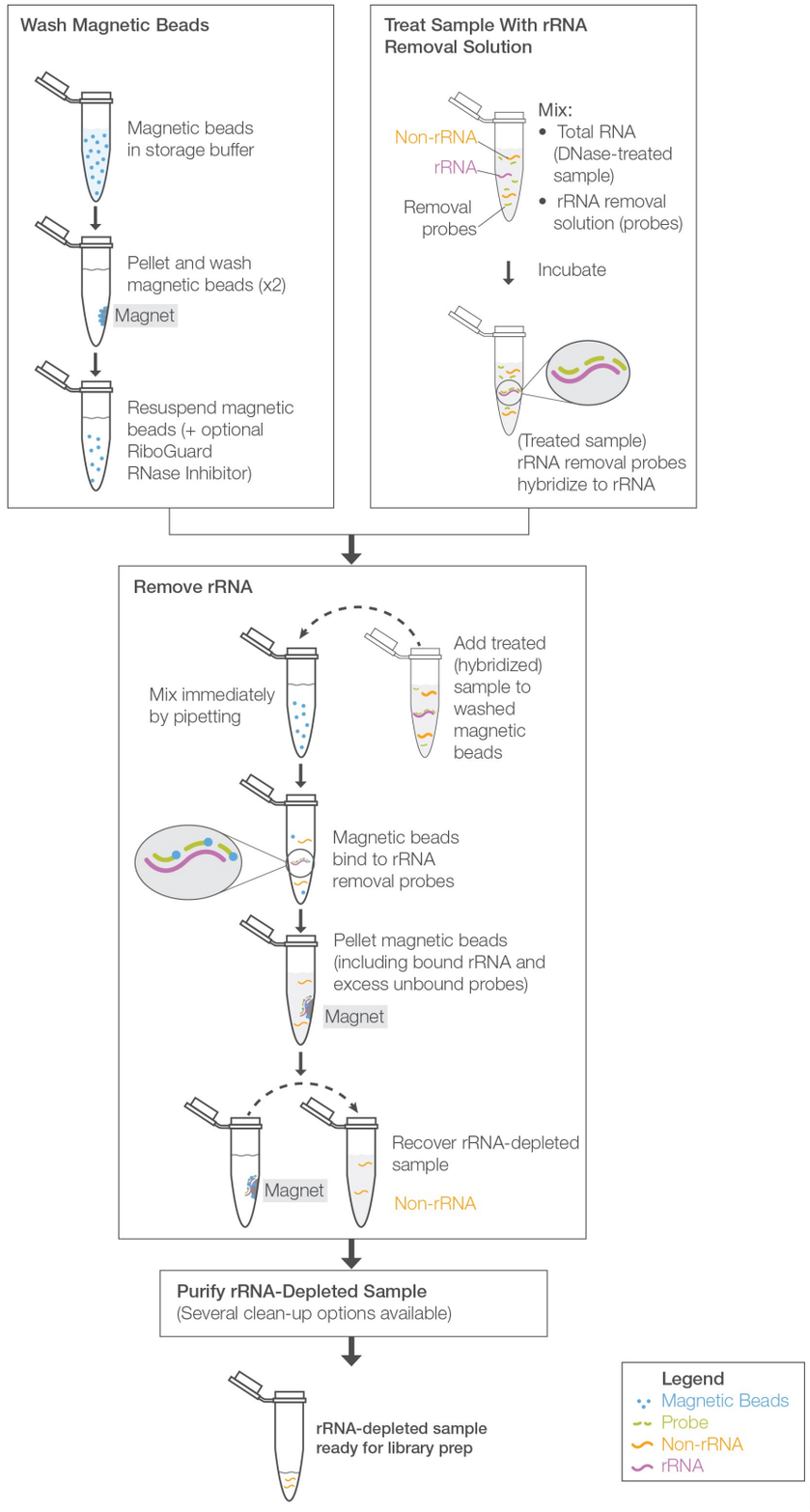
Genome Biol. 2012; 13(3): r23. 10.1186/gb-2012-13-3-r23

rRNA Depletion: Options



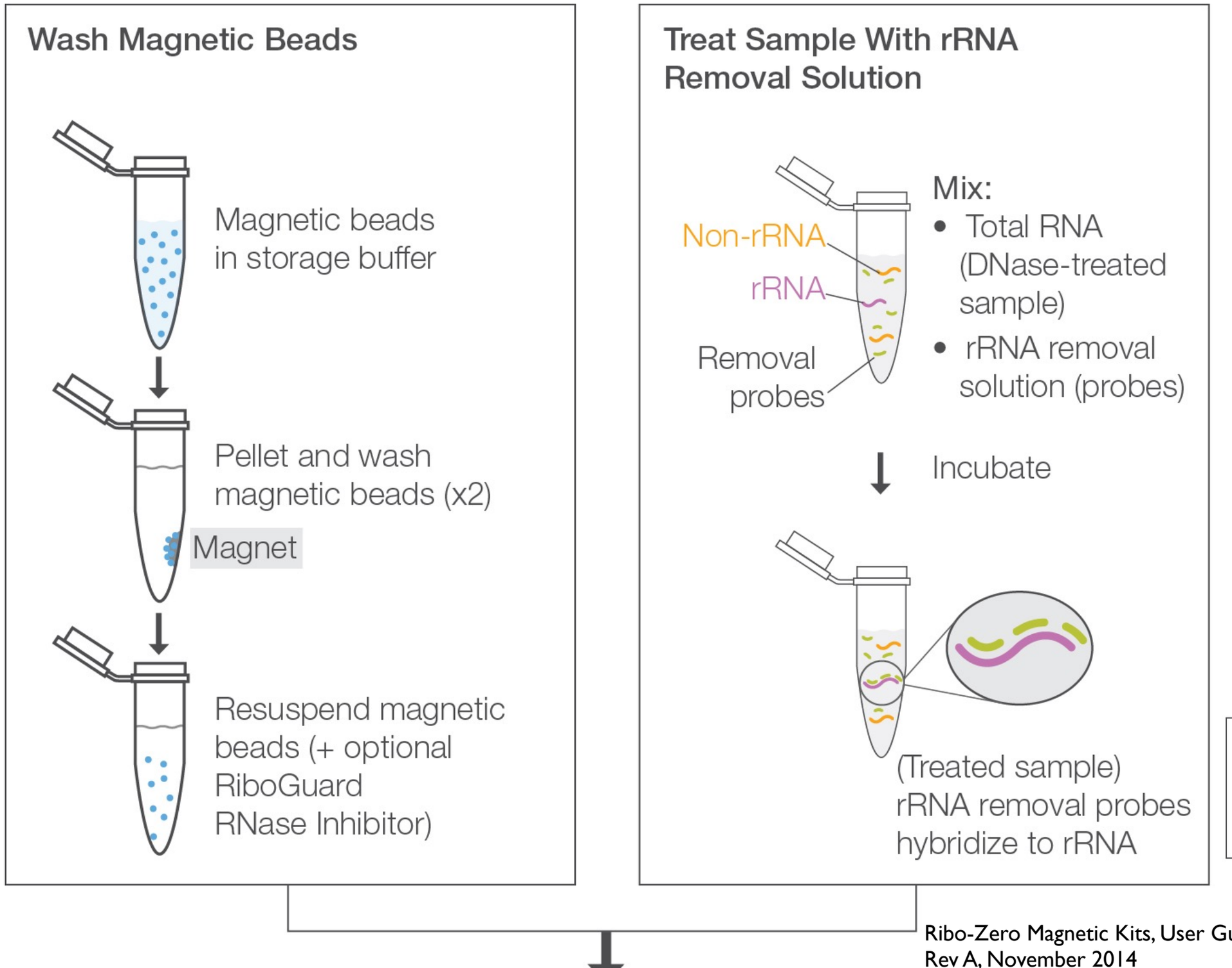
Genome Biol. 2012; 13(3): r23. 10.1186/gb-2012-13-3-r23

rRNA Depletion: How?



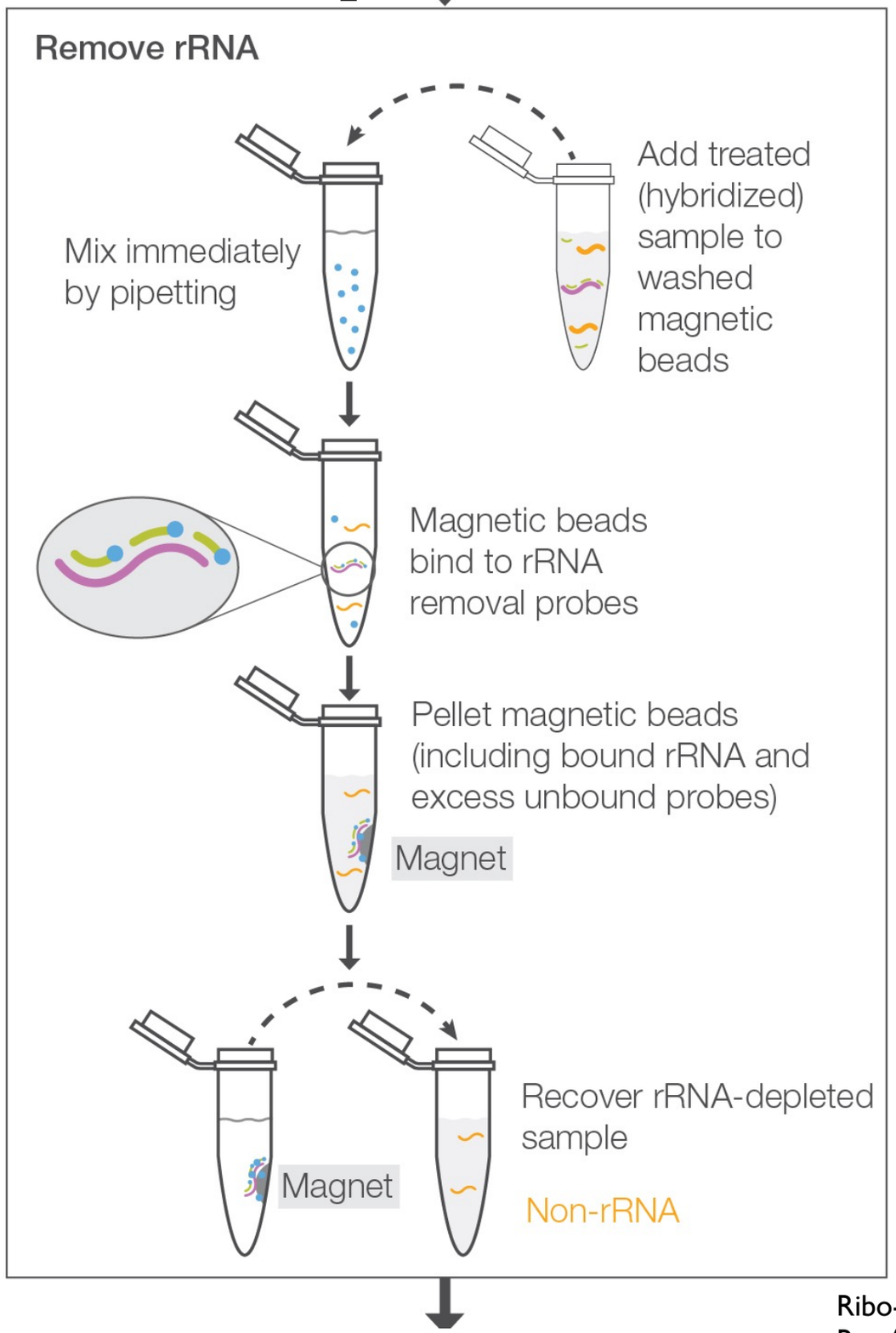
Ribo-Zero Magnetic Kits, User Guide, Part # 15065382, Rev A, November 2014

rRNA Depletion: How?



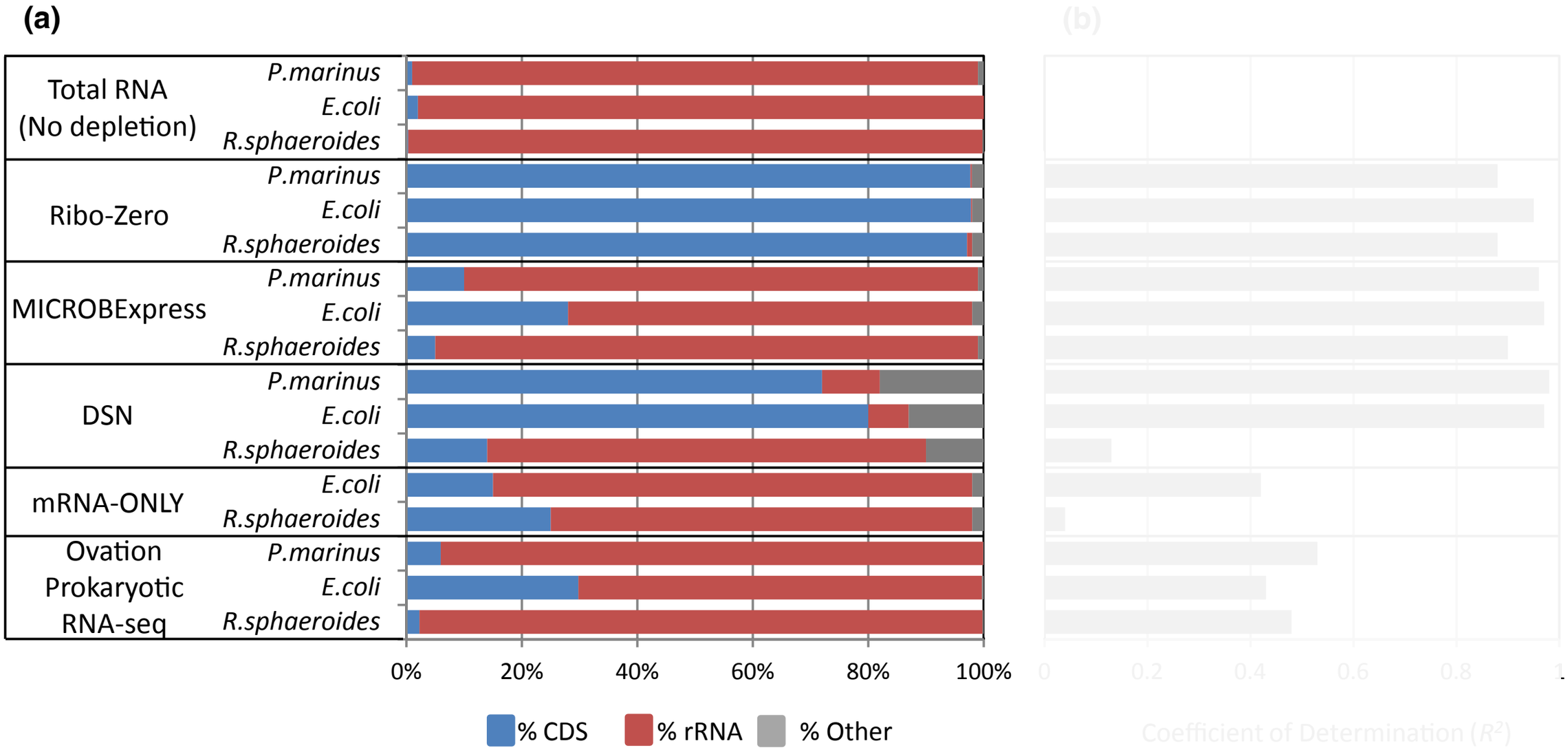
Ribo-Zero Magnetic Kits, User Guide, Part # 15065382, Rev A, November 2014

rRNA Depletion: How?



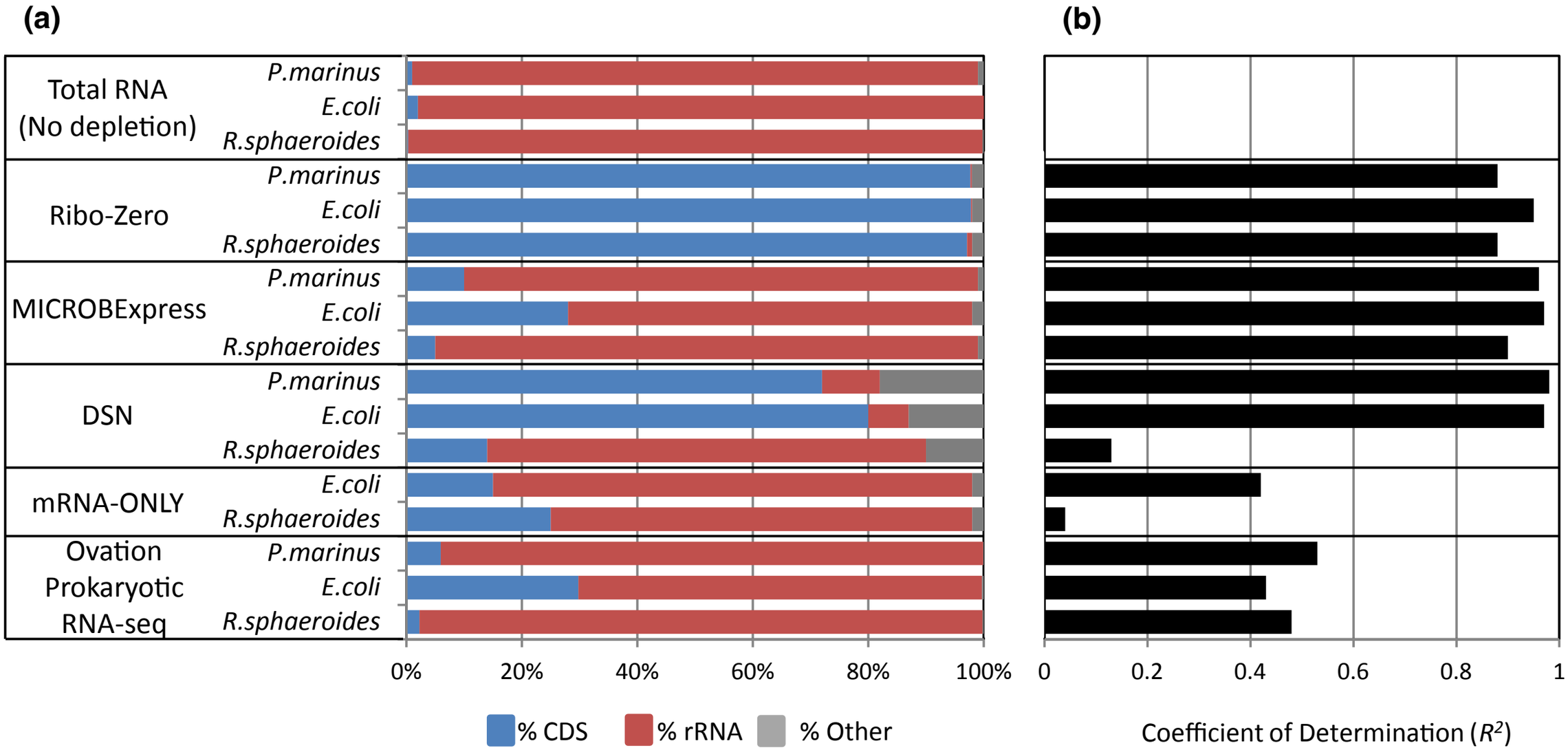
Ribo-Zero Magnetic Kits, User Guide, Part # 15065382, Rev A, November 2014

rRNA Depletion: Alternatives



Genome Biol. 2012; 13(3): r23. 10.1186/gb-2012-13-3-r23

rRNA Depletion: Alternatives



Genome Biol. 2012; 13(3): r23. 10.1186/gb-2012-13-3-r23

rRNA Depletion: Alternatives

- mRNA enrichment
  - poly(A) mRNA enrichment
  - Selective polyadenylation of mRNAs
  - Antibody capture of RNAs that interact with a specific protein
  - Non-random priming

- rRNA depletion
  - Ribosomal RNA capture
  - Duplex-specific nuclease (DSN) normalization
  - Degradation of processed RNA

rRNA Depletion: Alternatives

- mRNA enrichment
  - poly(A) mRNA enrichment
  - Selective polyadenylation of mRNAs
  - Antibody capture of RNAs that interact with a specific protein
  - Non-random priming

- rRNA depletion
  - Ribosomal RNA capture
  - Duplex-specific nuclease (DSN) normalization
  - Degradation of processed RNA

rRNA Depletion: Alternatives

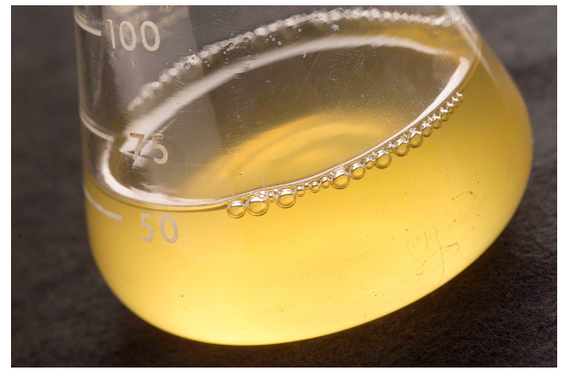
- mRNA enrichment

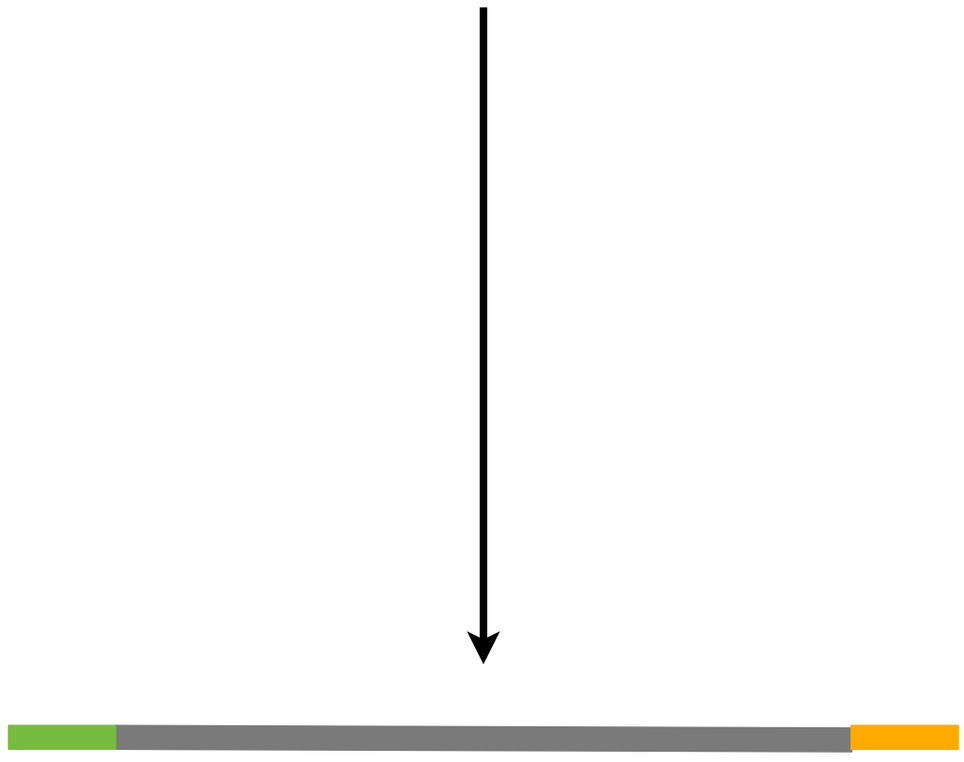
*** - poly(A) mRNA enrichment
  - Selective polyadenylation of mRNAs
  - Antibody capture of RNAs that interact with a specific protein
  - Non-random priming

- rRNA depletion
  - Ribosomal RNA capture
  - Duplex-specific nuclease (DSN) normalization
  - Degradation of processed RNA

Major Experiment Components

1. Growth
2. Sample Collection
3. RNA Extraction
4. Ribosomal RNA Depletion
5. Library Preparation





http://cse.ksu.edu/REU/S11/bvosburg/FLASK.jpg

Illumina Sequencing

From Library to Data

Library Preparation

DNA Fragment

200-1000 bp

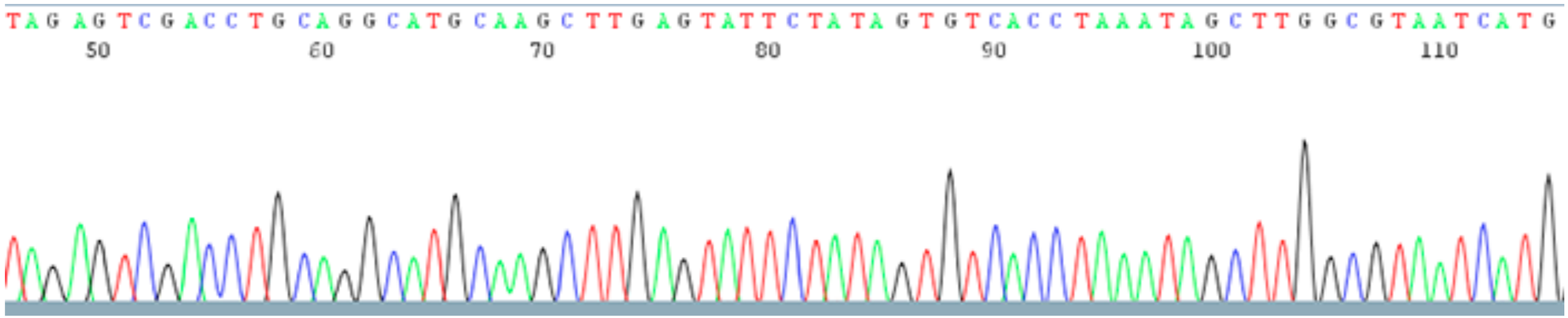
+

Adapters

=

Sequencing "Library"

Dye-terminator Sanger Sequencing



https://dnacore.mgh.harvard.edu/new-cgi-bin/site/pages/sequencing_pages/seq_troubleshooting.jsp

Sequencing

Sequencing

Sequencing

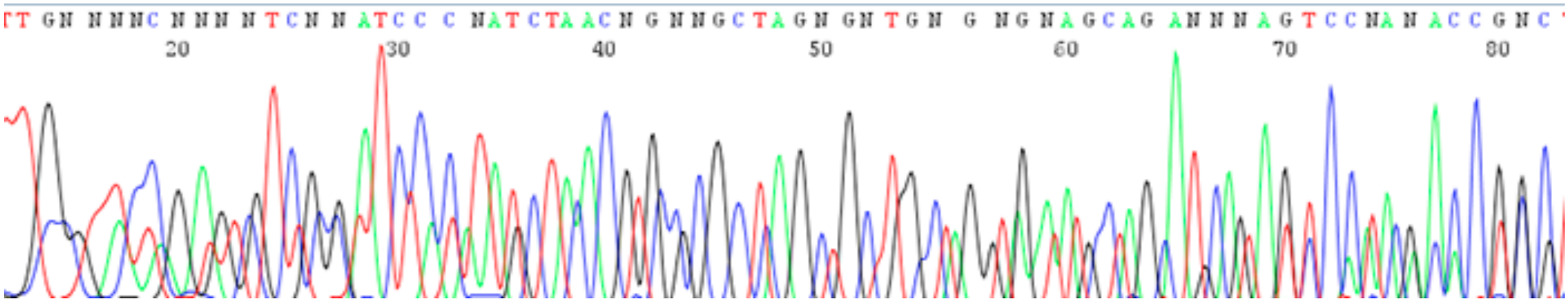
Sequencing

Sequencing

Sequencing

Sequencing

Dye-terminator Sanger Sequencing



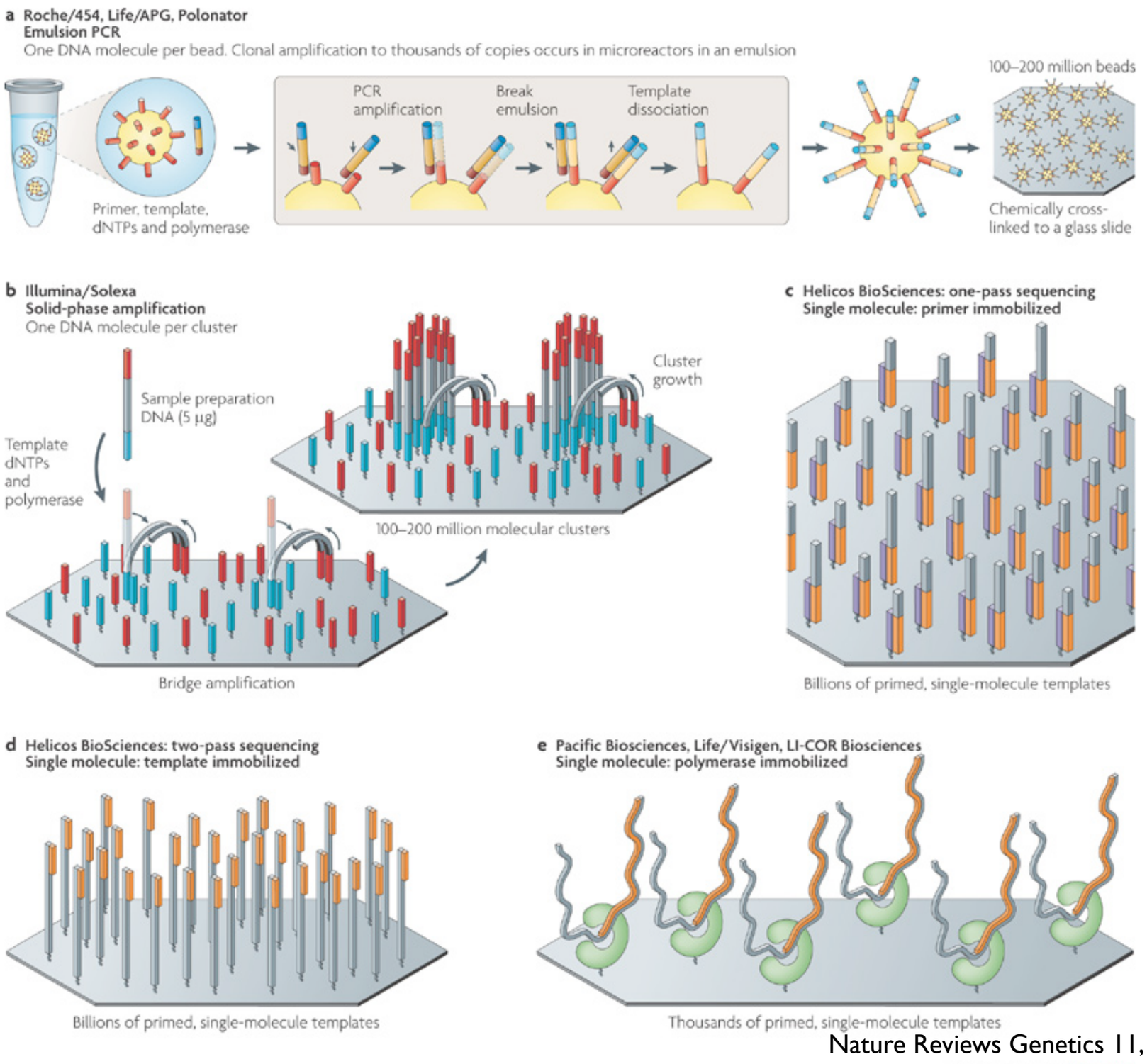
https://dnacore.mgh.harvard.edu/new-cgi-bin/site/pages/sequencing_pages/seq_troubleshooting.jsp

How?

How?

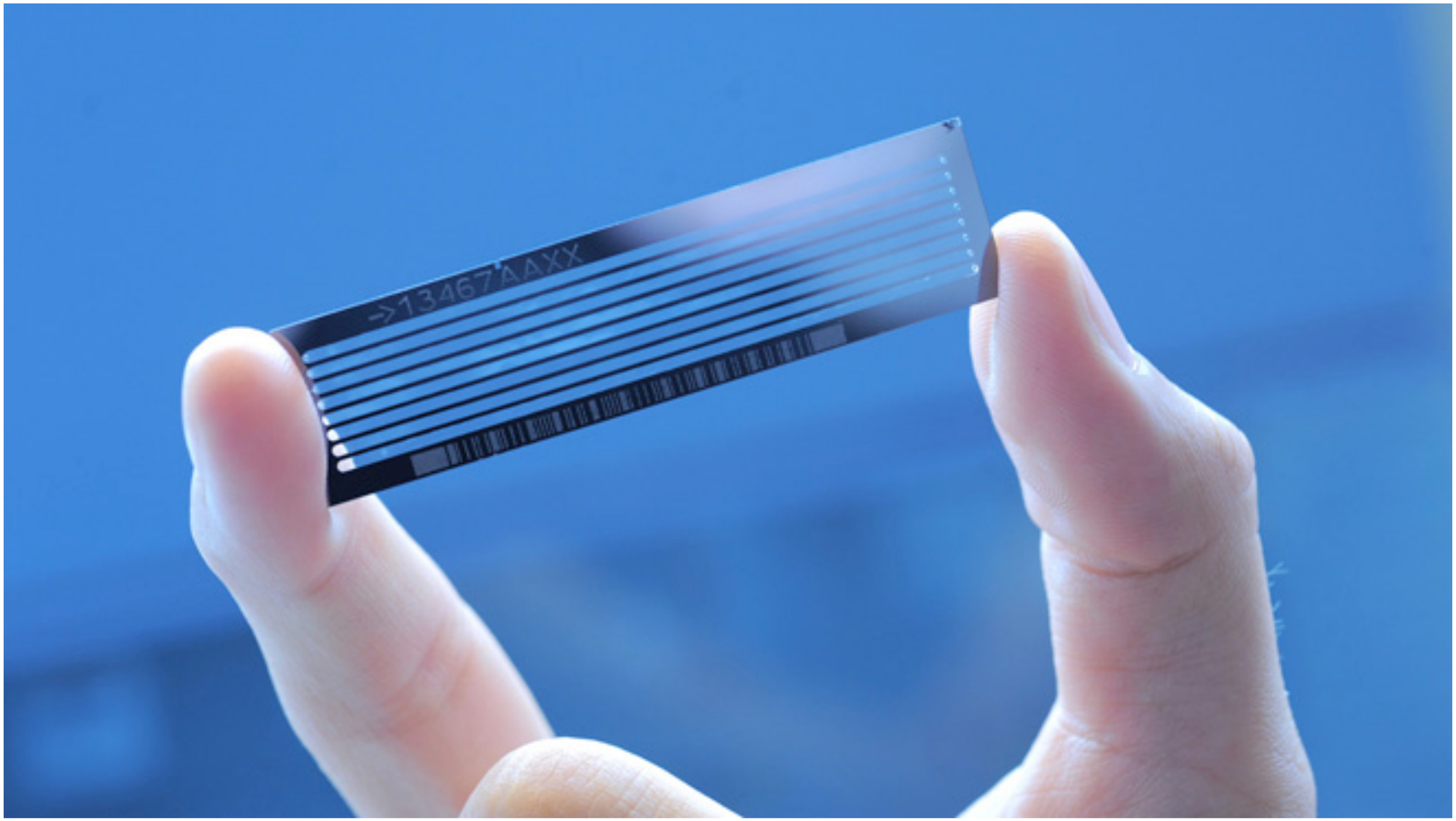
- Separate
- Detect

Template immobilization



Nature Reviews Genetics 11, 31-46, doi:10.1038/nrg2626

A Flow Cell



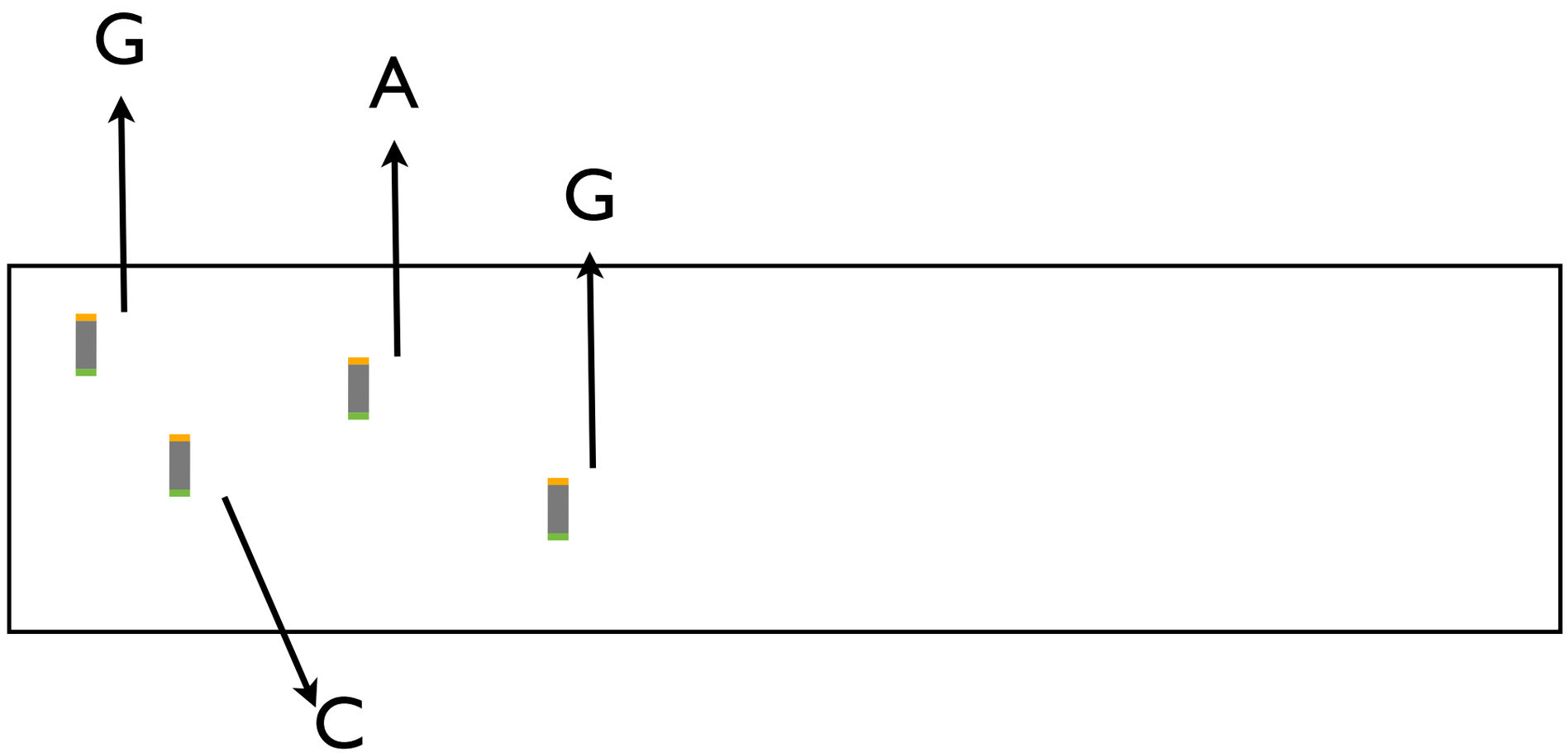
http://global.fncstatic.com/static/managed/img/Scitech/neanderthal%20genome%205.jpg

Pass Around Flow Cells!!!

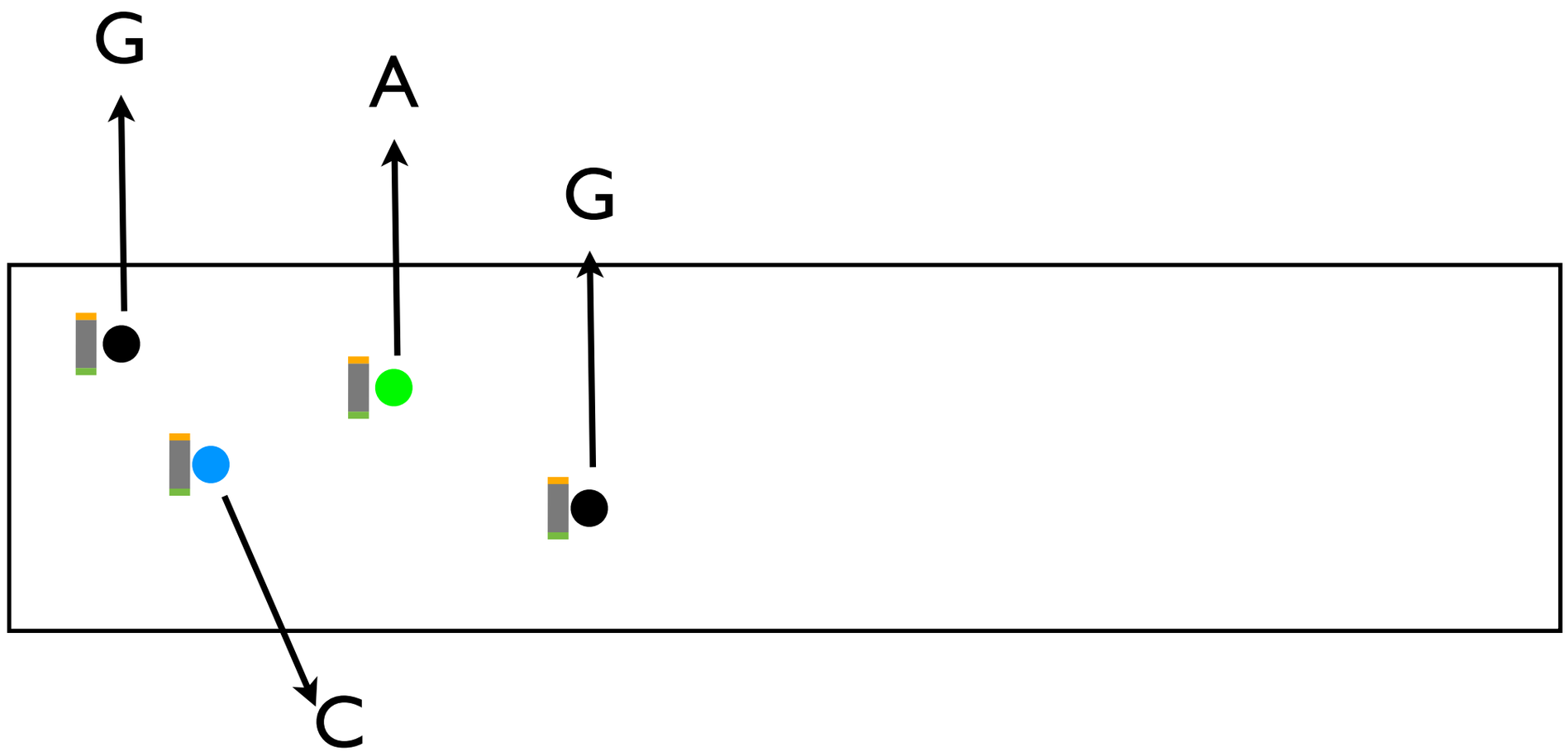
A Flow Cell

Bind Library

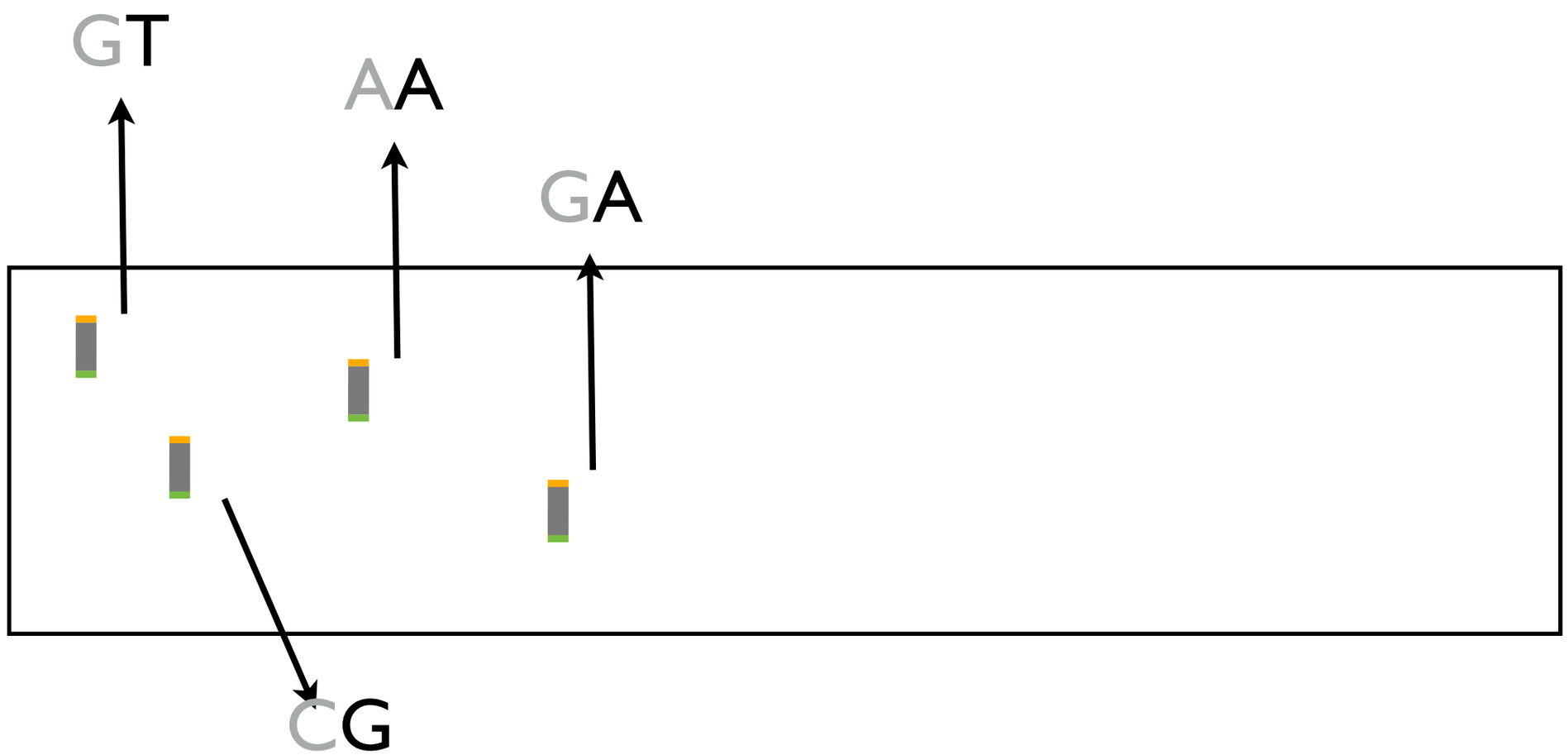
1st Cycle



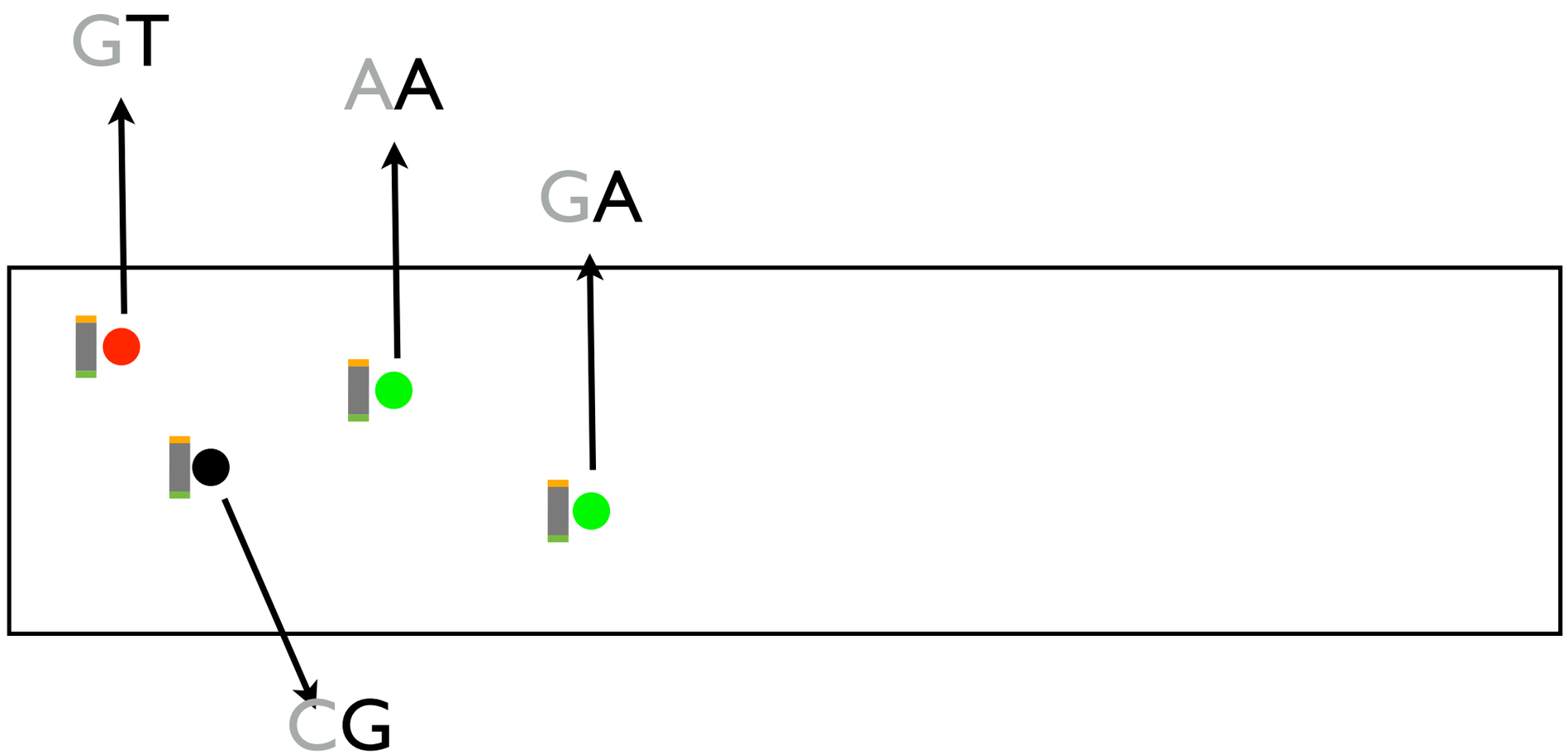
1st Cycle



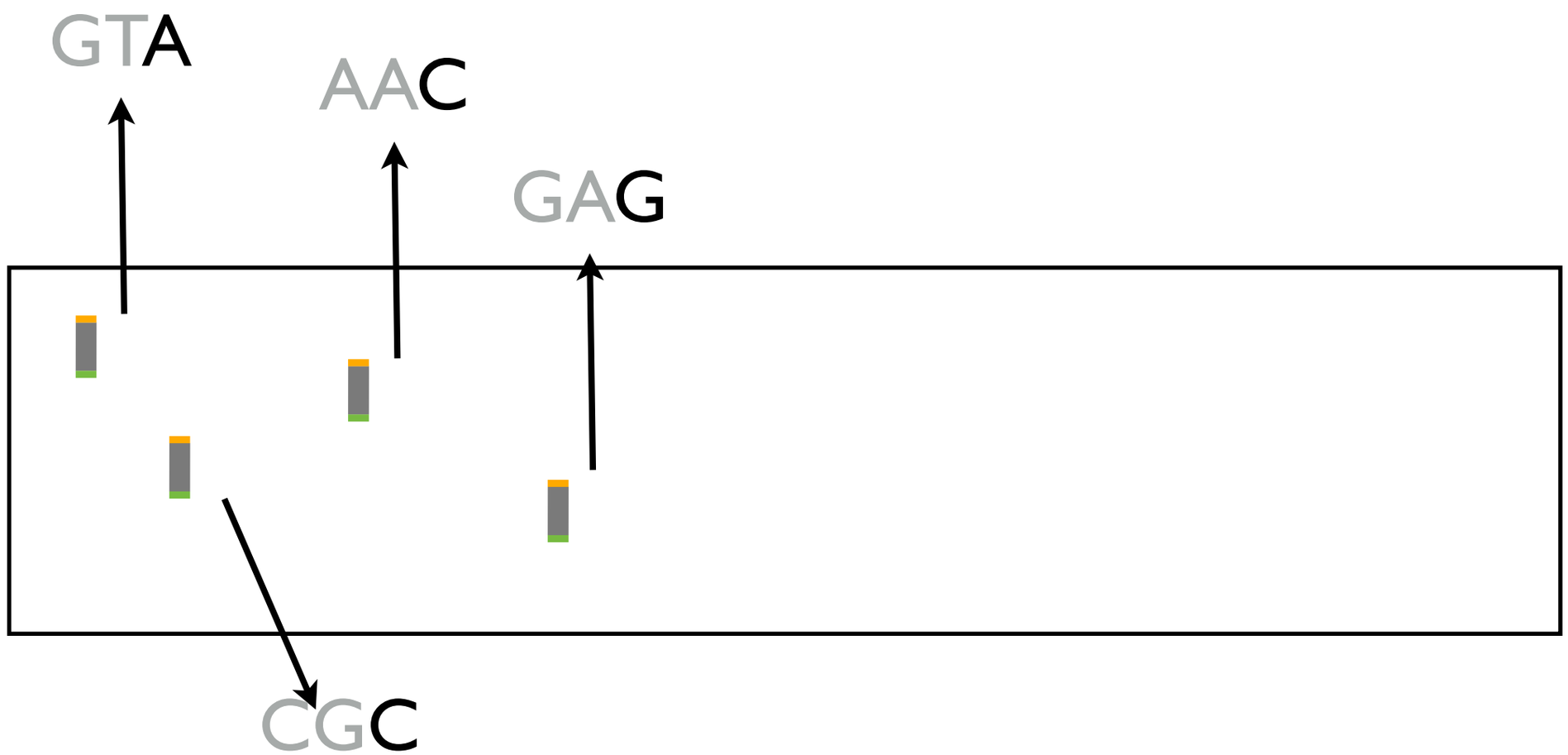
2nd Cycle



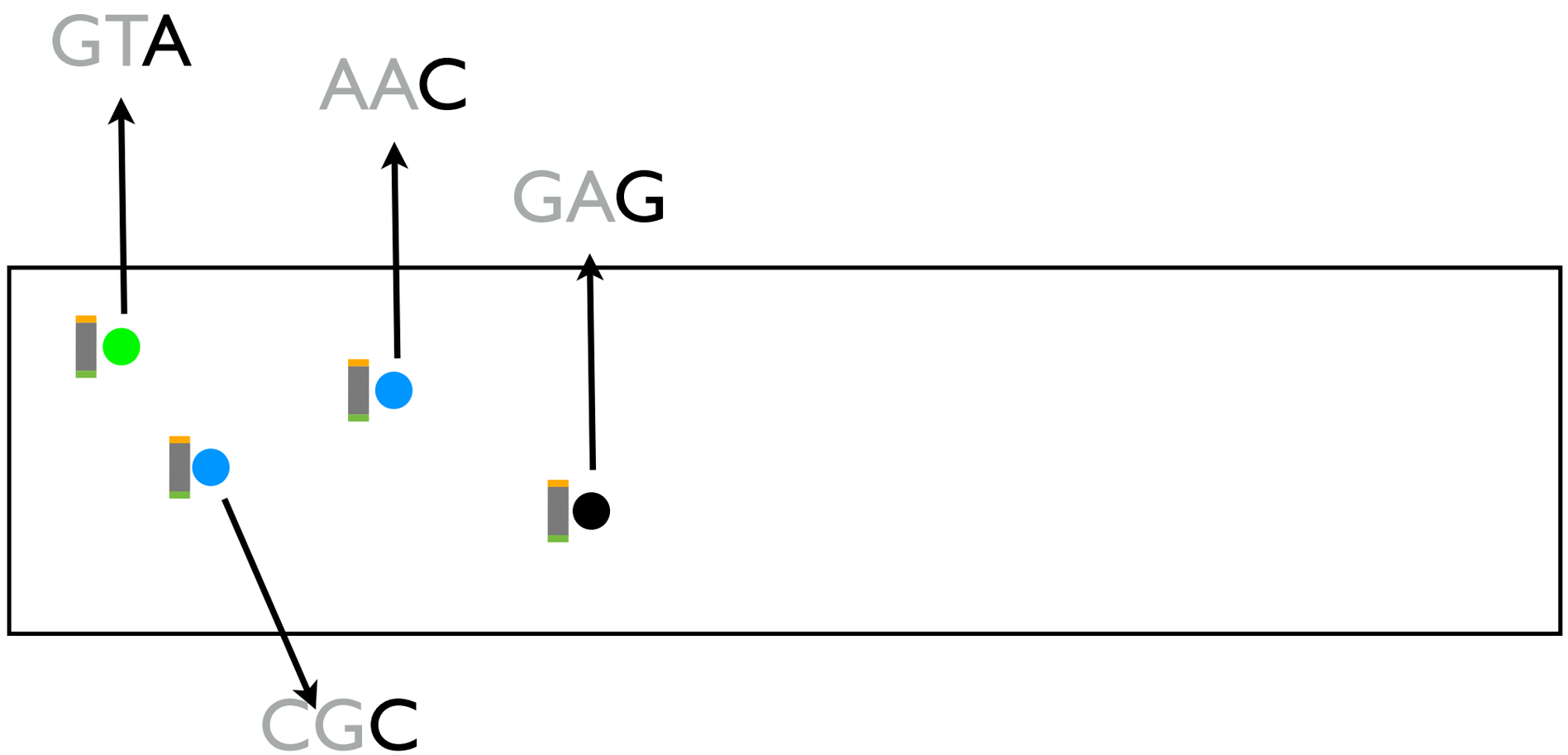
2nd Cycle



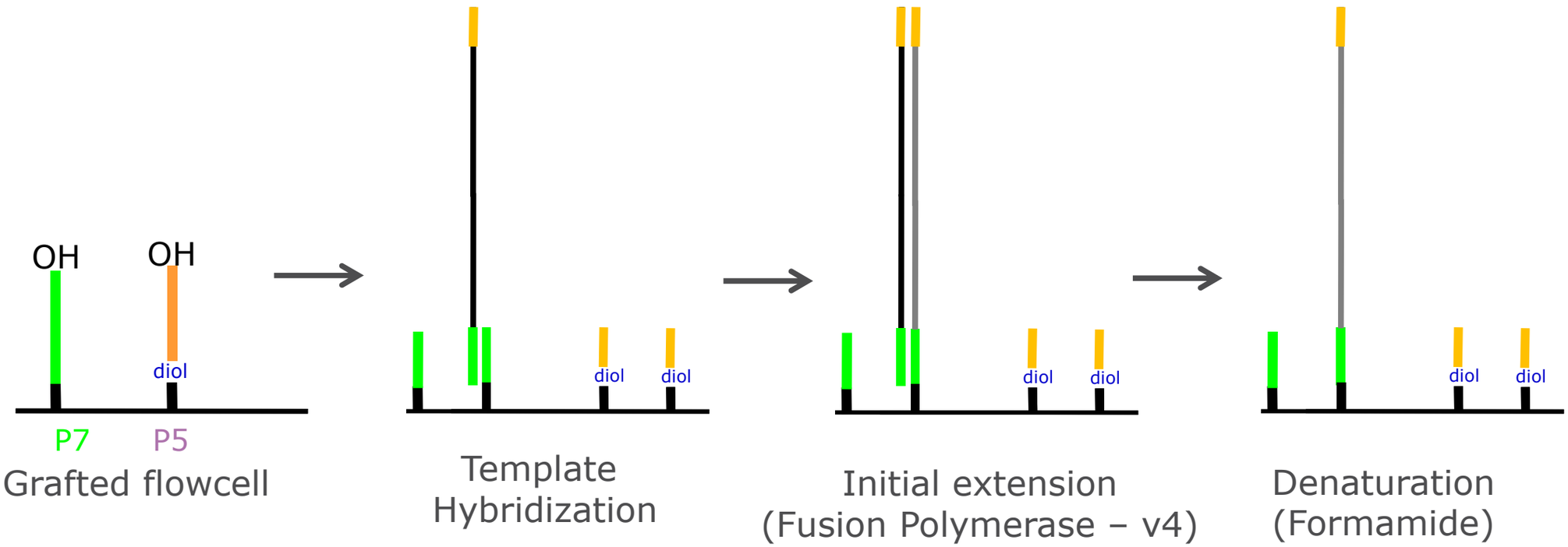
3rd Cycle



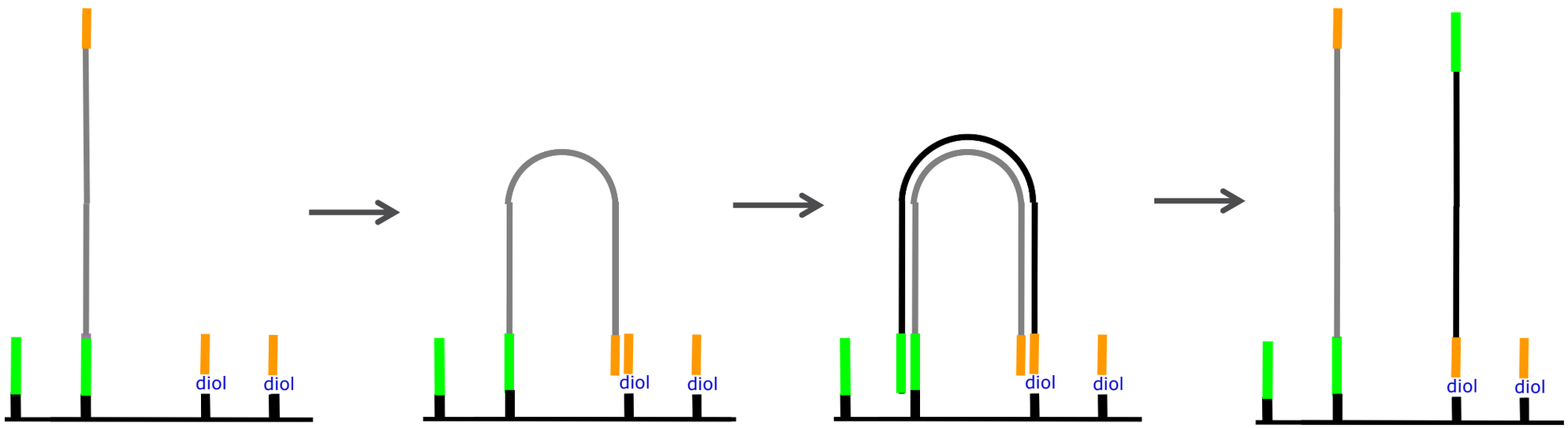
3rd Cycle



Cluster generation – hybridization and amplification



Cluster generation – hybridization and amplification



illumina

Sequencing Library

Library Preparation: Why?

We have RNA. We need a DNA library.

Library Preparation: How?

NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (#E7420)

Library Preparation: How?

1. RNA Fragmentation
2. cDNA Synthesis
3. Adapter Ligation
4. PCR Enrichment

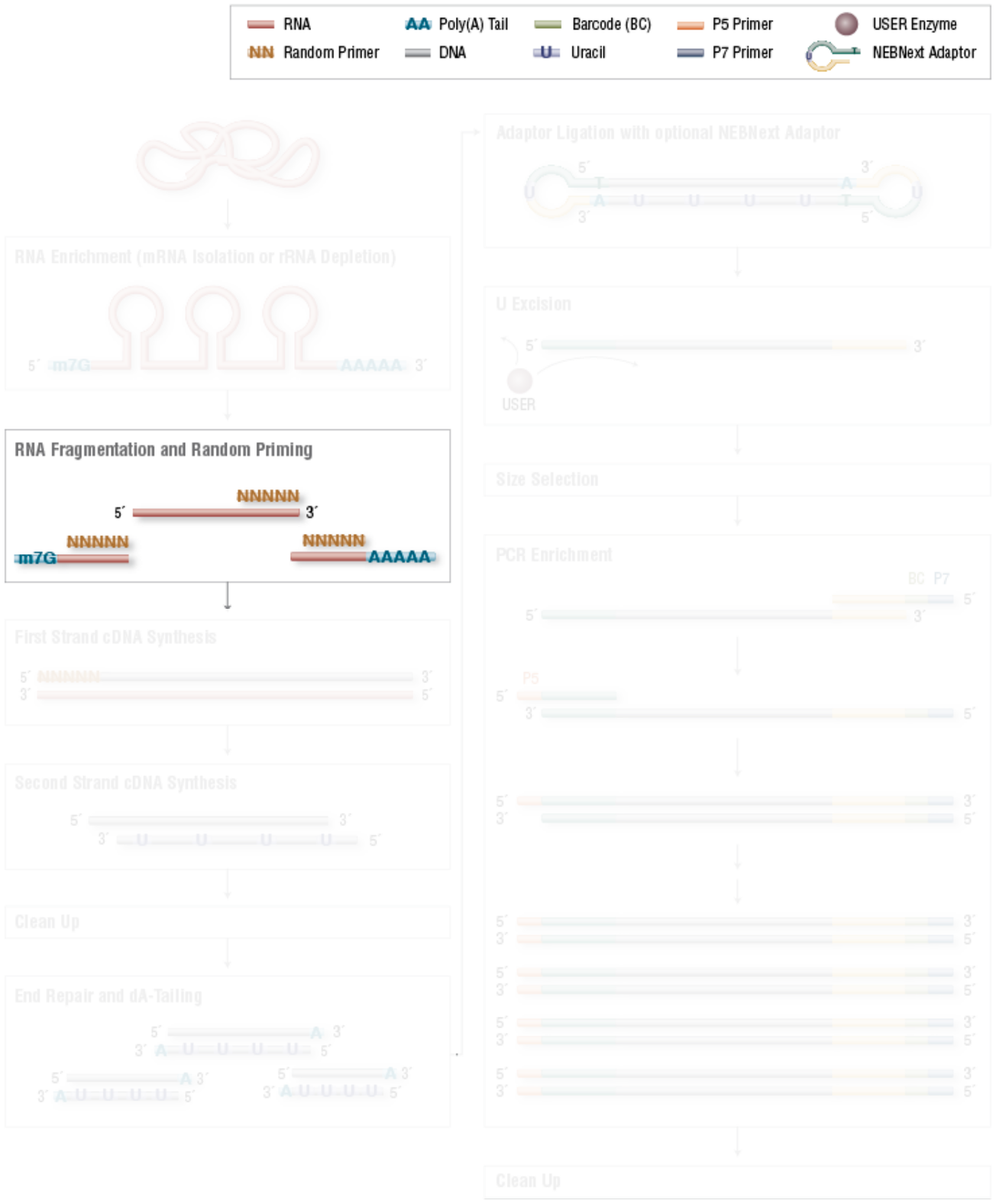
Library Prep Workflow

NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

Fragmentation

NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

Fragmentation



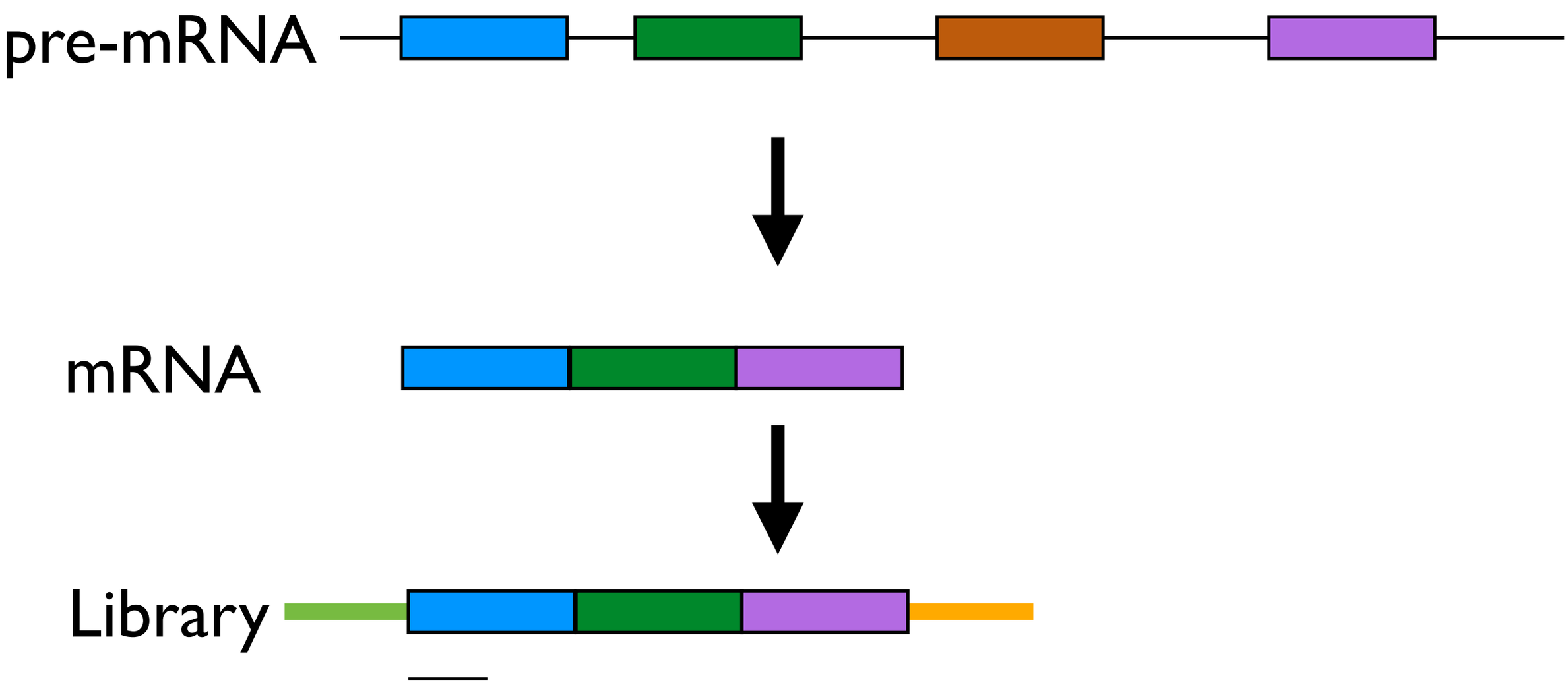
NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

Library Prep Fragmentation: Why?

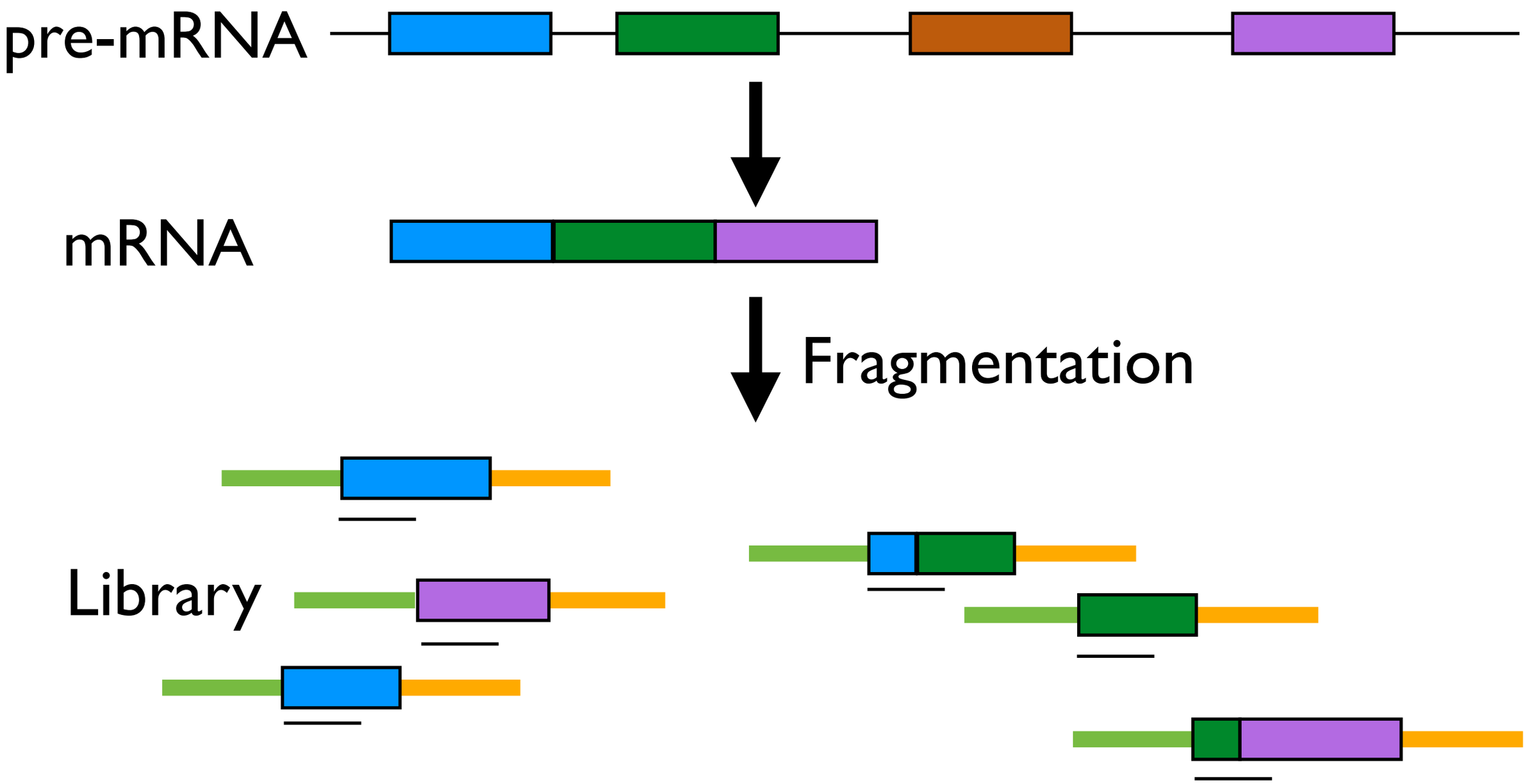
- Efficient cluster generation and sequencing
- Distribution of reads across mRNA

NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

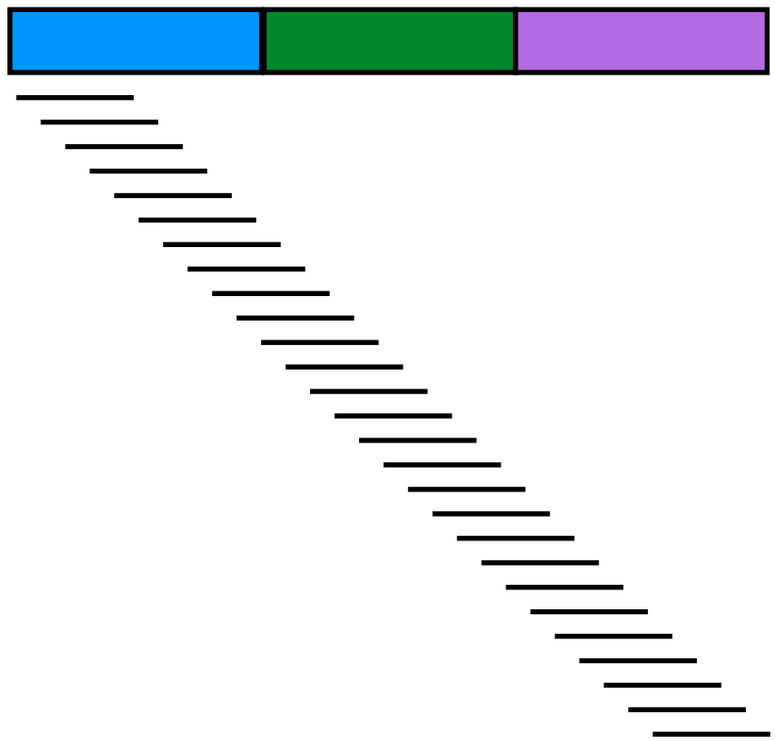
Library Prep Fragmentation: Why?



Library Prep Fragmentation: Why?

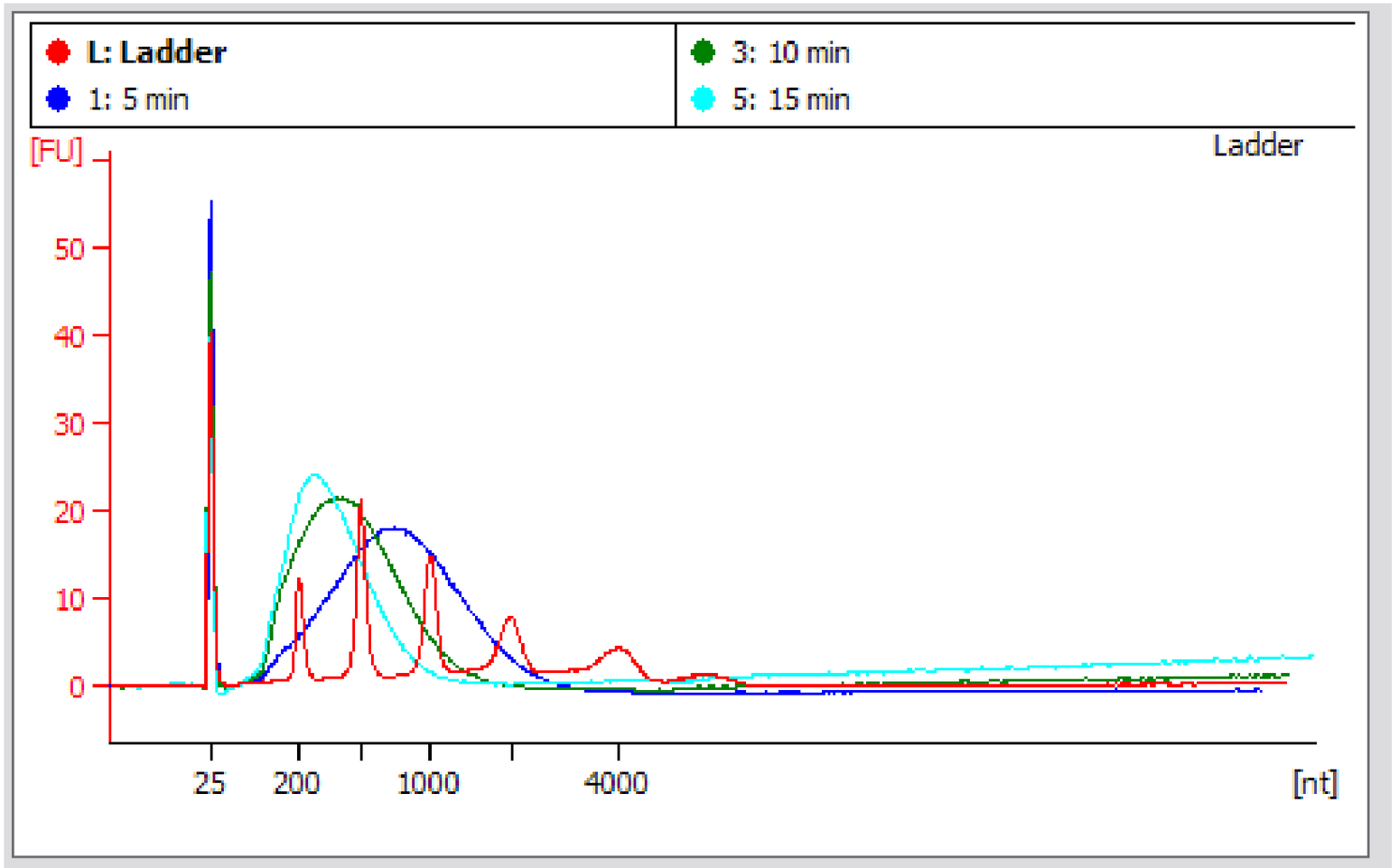


Library Prep Fragmentation: Why?



Library Prep Fragmentation: How?

- Heat with divalent metal cation (Chemical)



NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

Library Prep Fragmentation: Alternatives?

- Degraded RNA
- Small RNAs
- DNA Fragmentation uses Physical or Enzymatic methods
- Needs to be Random!!!

NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

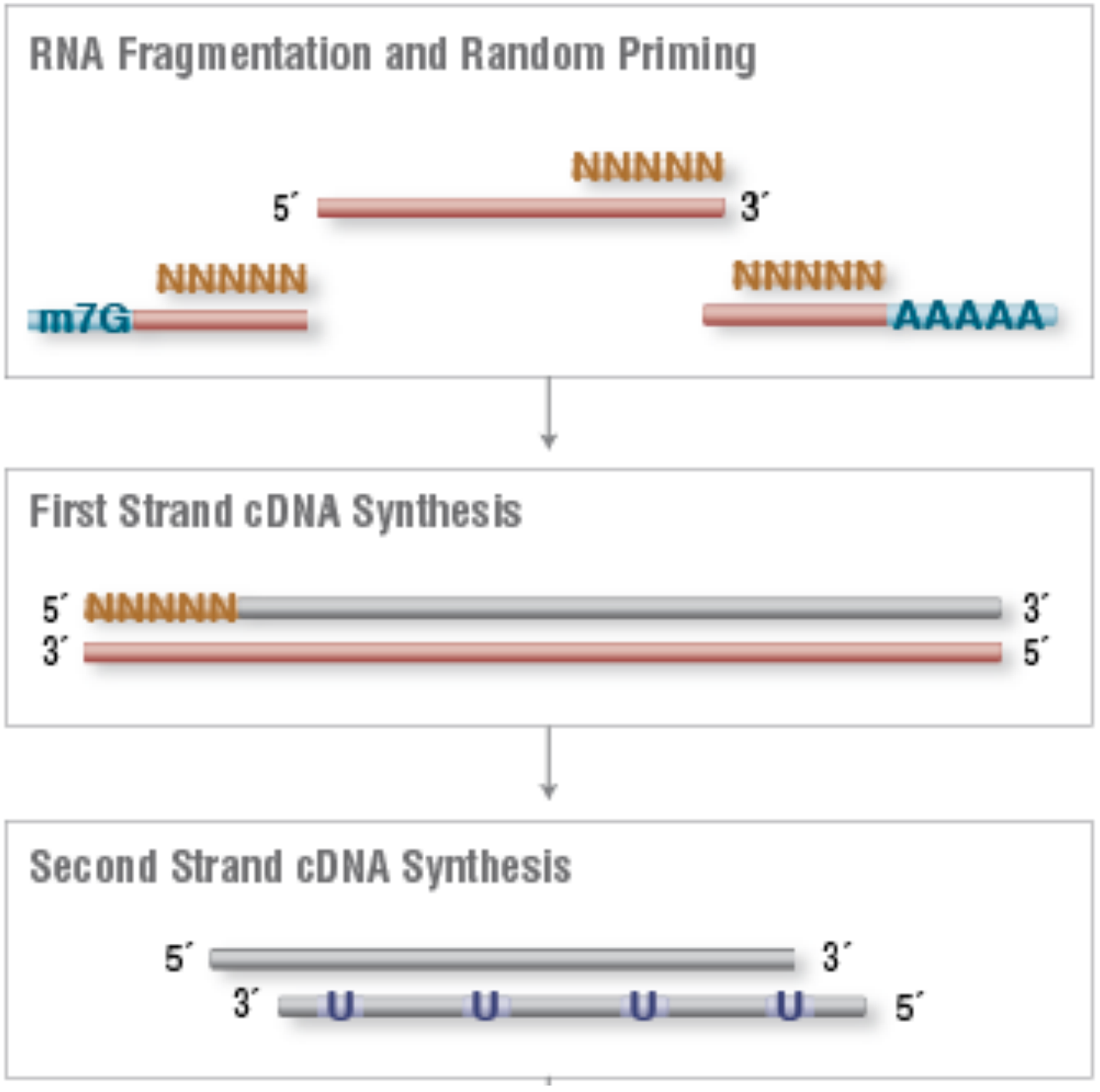
cDNA Synthesis

NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

cDNA Synthesis

NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

cDNA Synthesis



NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

cDNA Synthesis: Why?

- Have RNA, need DNA

cDNA Synthesis: How?

- First Strand: Reverse Transcriptase
- Second Strand:
  - RNase H: generate RNA primers
  - DNA polymerase I: DNA synthesis
  - DNA ligase: ligate fragments

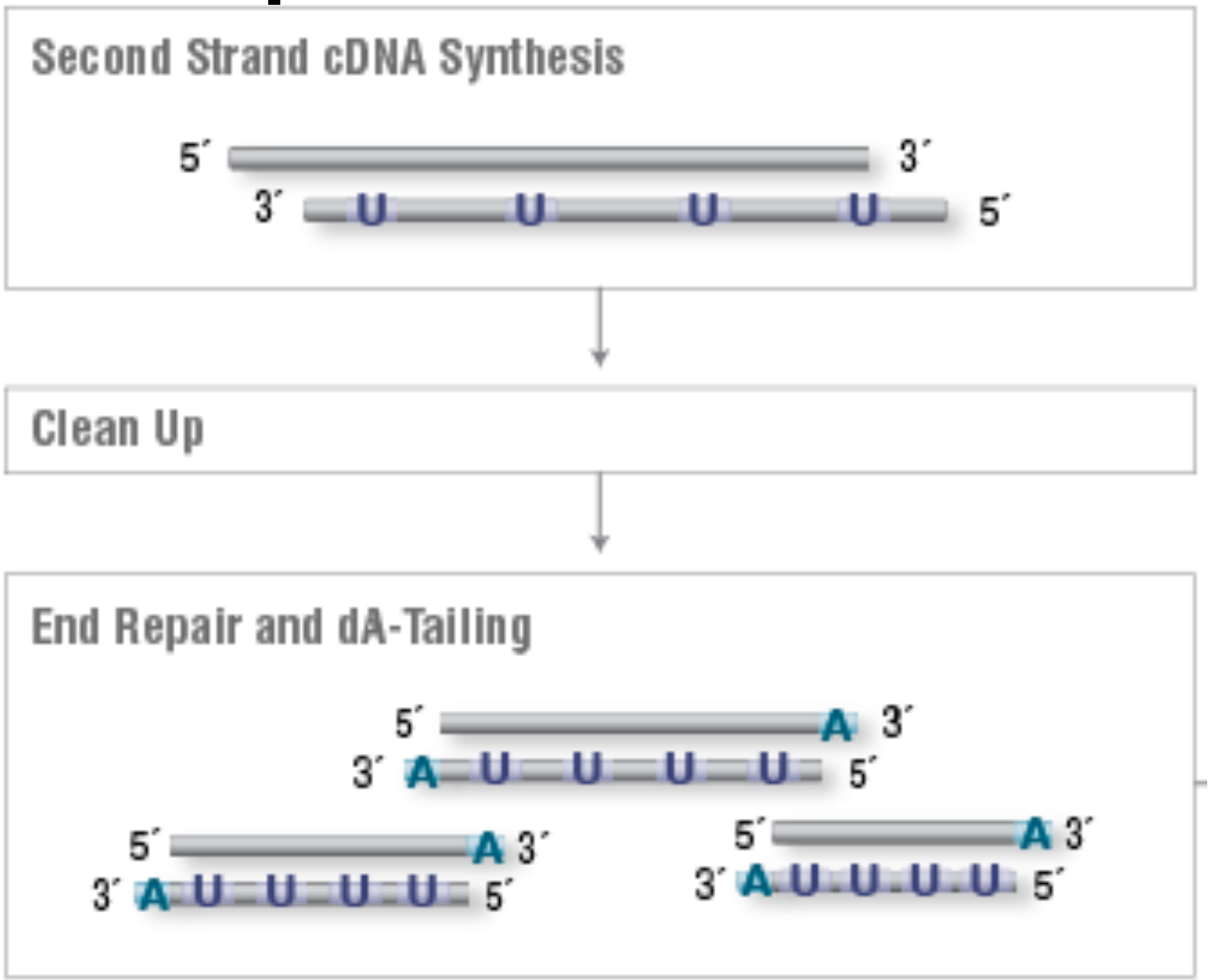
End-Repair and dA-Tailing

NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

End-Repair and dA-Tailing

NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

End-Repair and dA-Tailing



NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

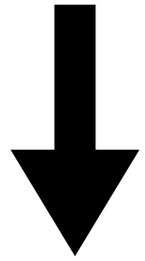
End-Repair and dA-Tailing

- Prepare fragments for adapter ligation
- Generate blunt ends
- Then generate 3’ A overhang

End Repair

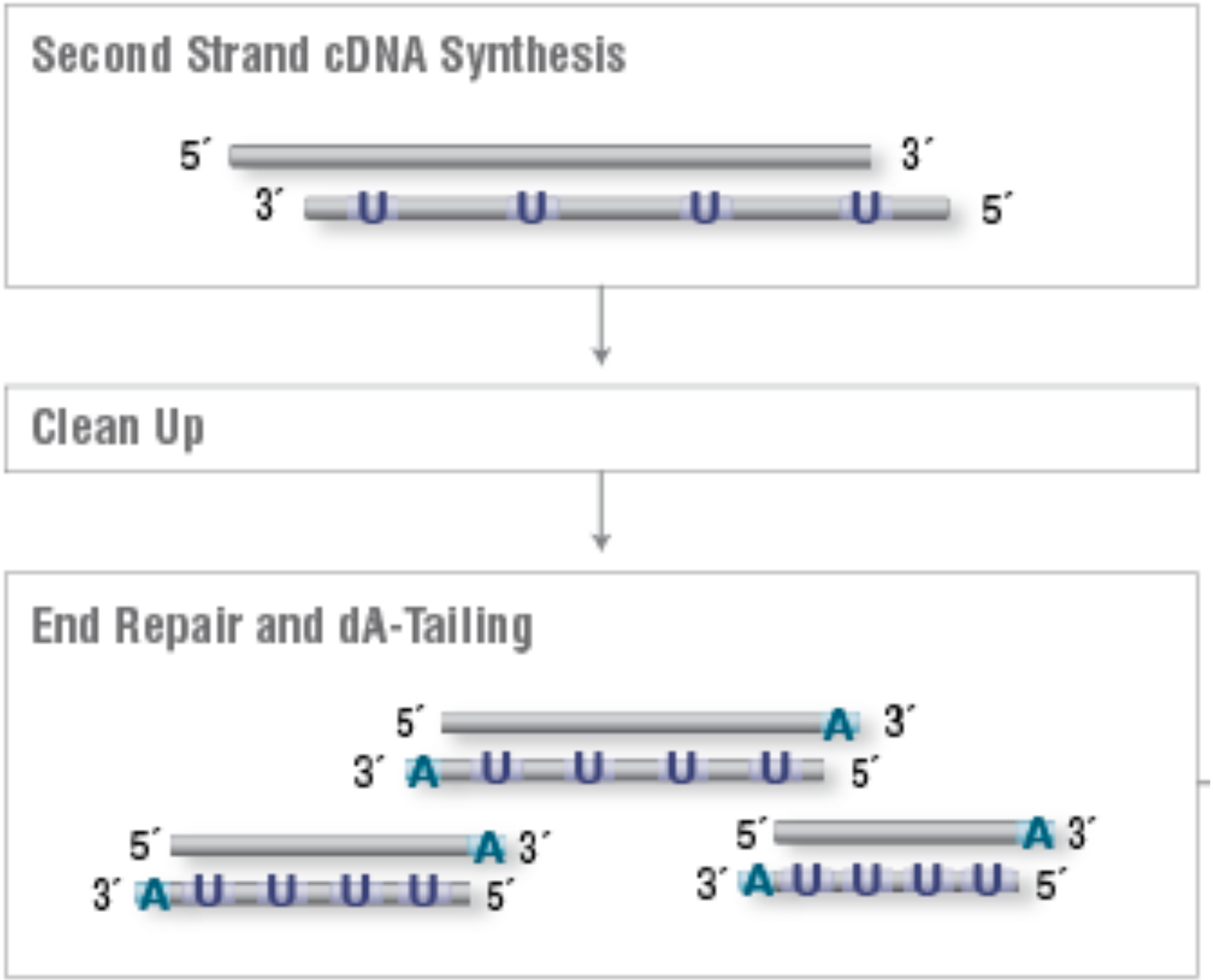
End Repair: What

```
5′-CTGATCTGACT     -3'
3′-GACTAGACTGACTAC-5′
```



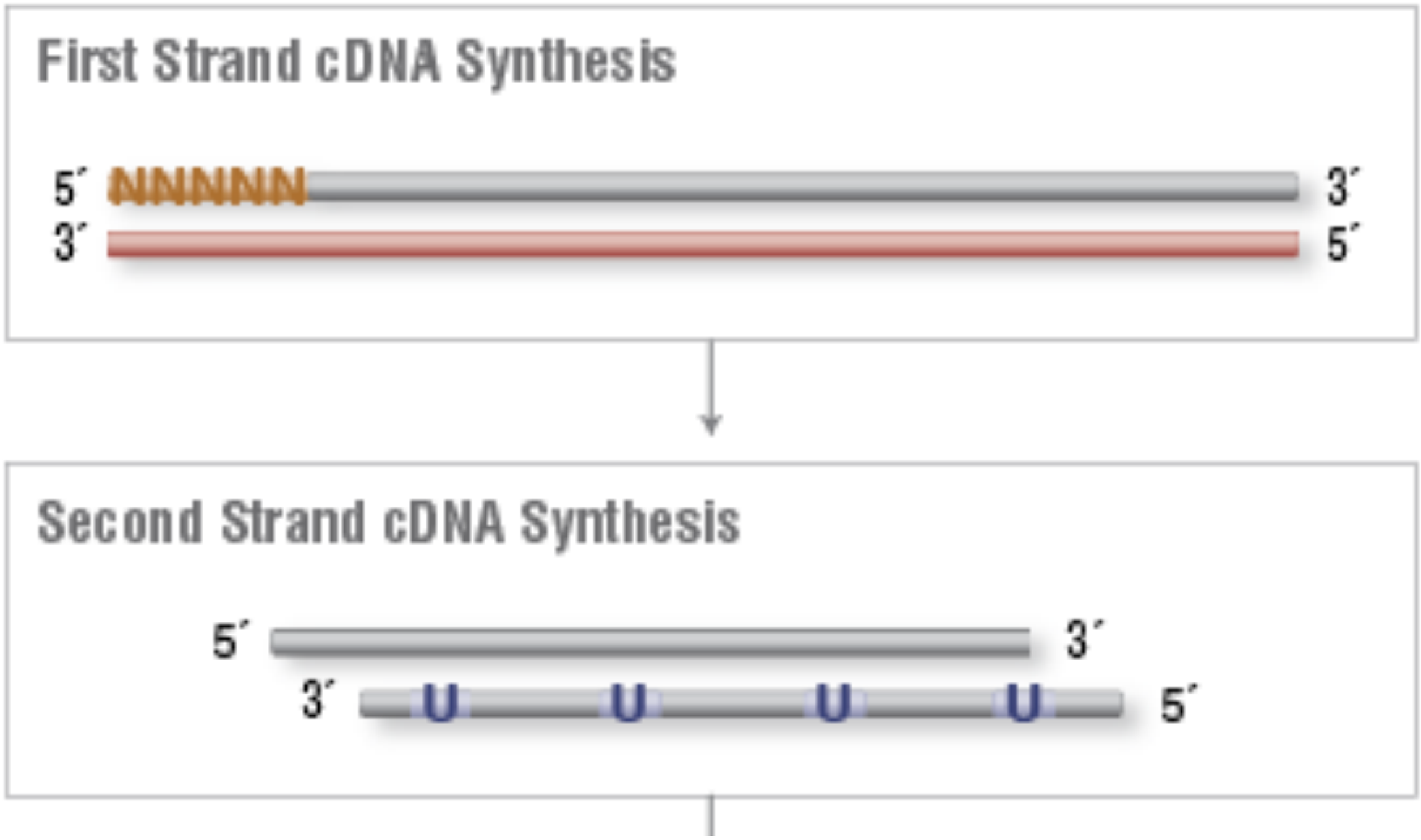
```
5′-CTGATCTGACTGATG-3′
3′-GACTAGACTGACTAC-5′
```

Why are ends NOT blunt?



NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

Why are ends NOT blunt?

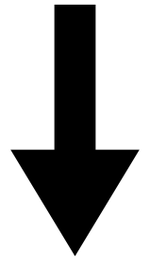


NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

dA-Tailing

dA-Tailing: What

```
5′-CTGATCTGACTGATG-3'
3′-GACTAGACTGACTAC-5′
```



```
5′- CTGATCTGACTGATGA-3'
3′-AGACTAGACTGACTAC -5′
```

dA-Tailing: Why

- Allow sticky-end ligation to a “universal fragment”

dA-Tailing: Why

```
5′- GATGATTGCTGAAGA-3'
3′-ACTACTAACGACTTC -5′
```

```
5′- AGTACTGTTCTTTATA-3'
3′-ATCATGACAAGAAATA -5′
```

+

```
5′- CCATG-3'
3′-TGGTAC-5′
```

=

```
5′- GATGATTGCTGAAGACCATG-3'
3′-ACTACTAACGACTTCTGGTAC-5′
```

```
5′- AGTACTGTTCTTTATACCATG-3'
3′-ATCATGACAAGAAATATGGTAC-5′
```

dA-Tailing: Why?

NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

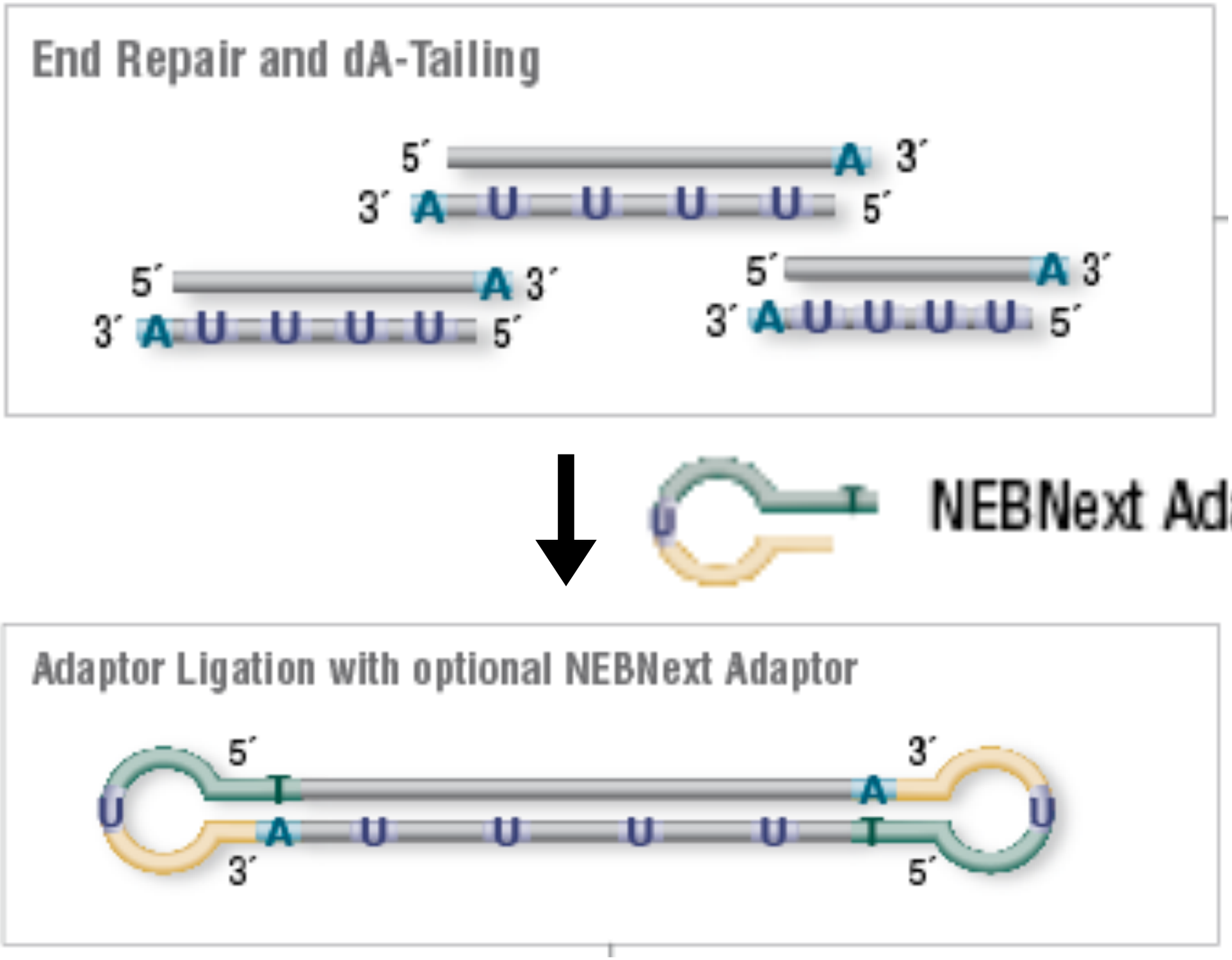
Adapter Ligation

NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

Adapter Ligation

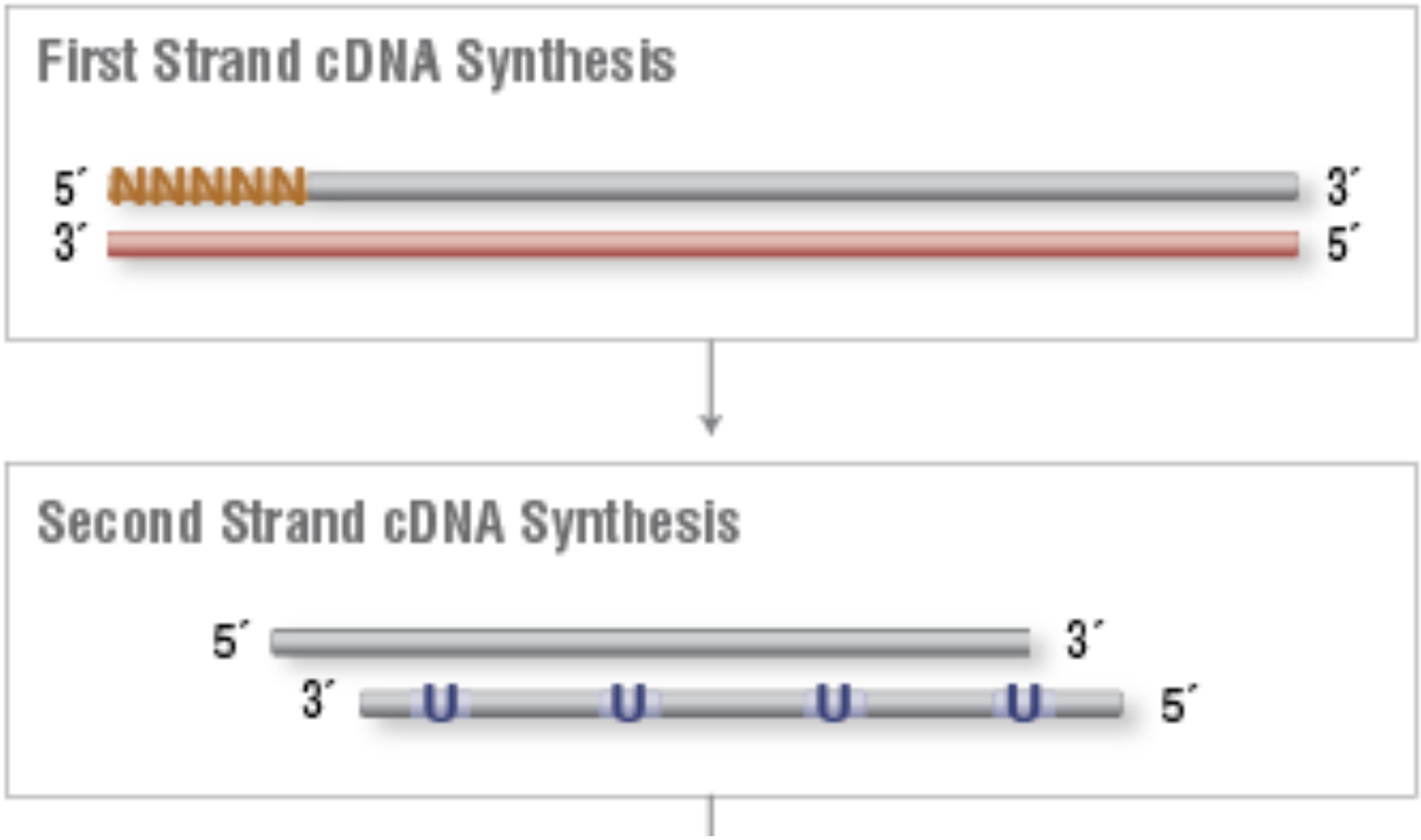
NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

Adapter Ligation



NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

U Excision



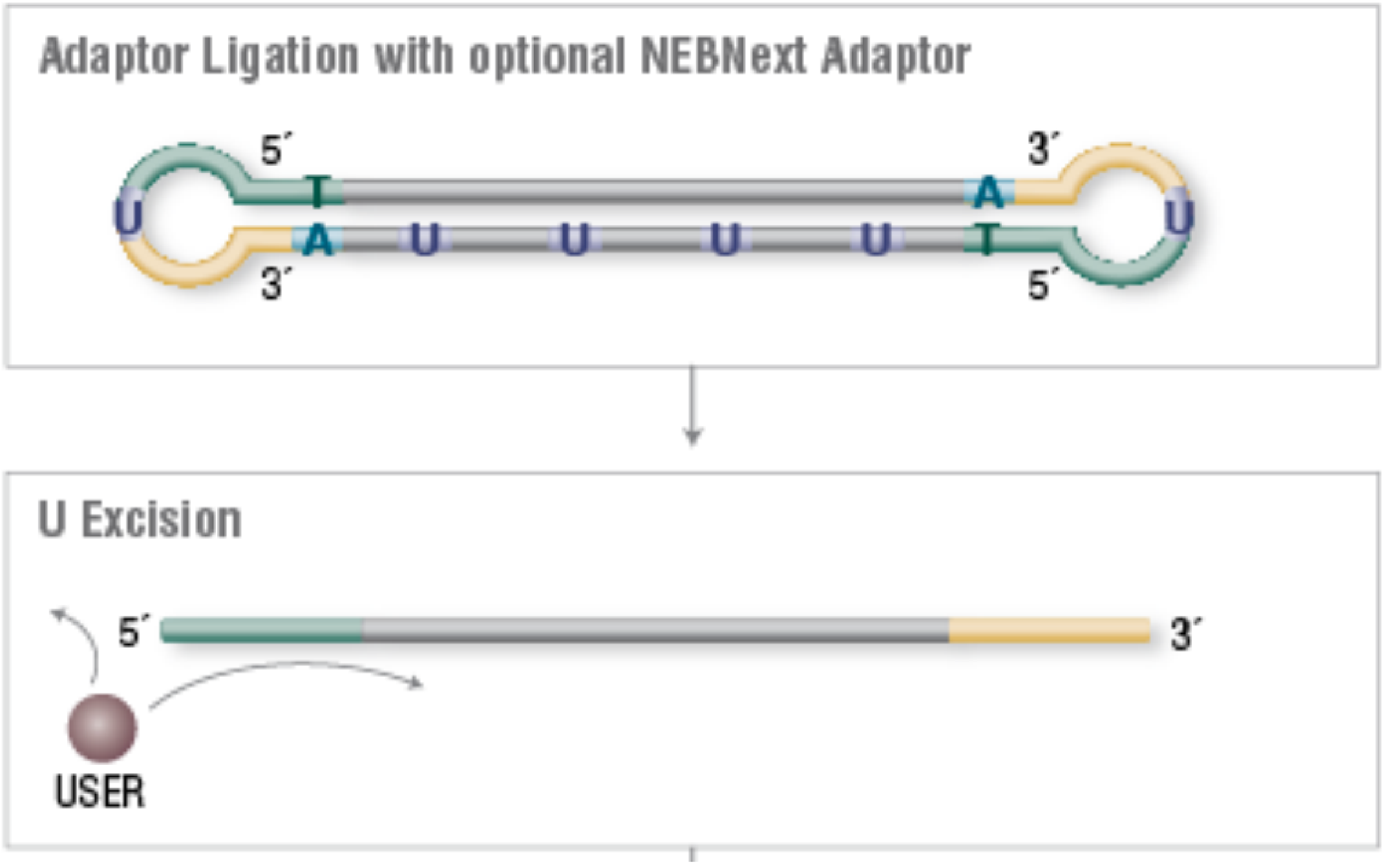
NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

U Excision

Why?

NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

U Excision



NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

Size Selection

NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

Size Selection

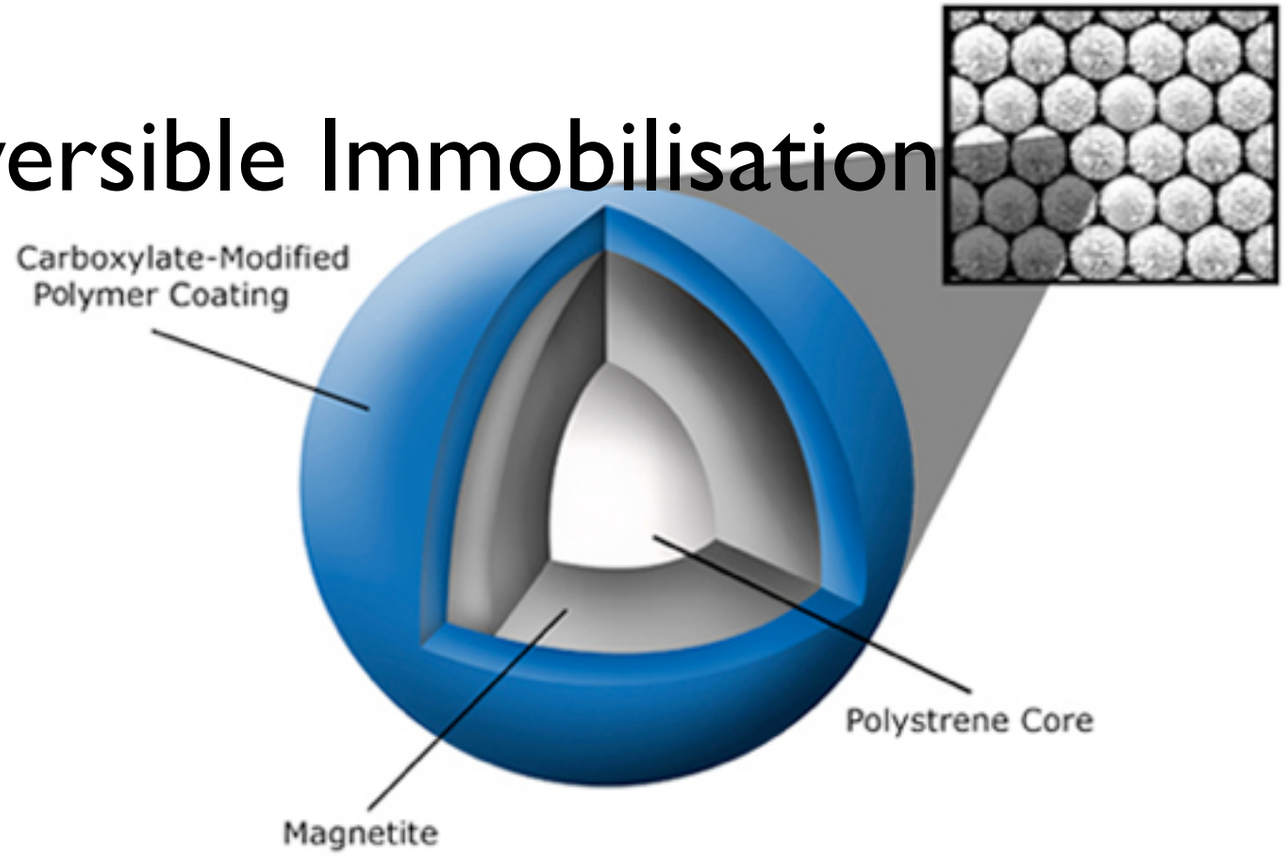
NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

Clean Up and Size Selection: Why?

- Remove regents from previous step
- Eliminate unwanted fragments
  - Unligated adapter
  - adapter dimers
  - fragments without adapter
- Efficient cluster generation and sequencing

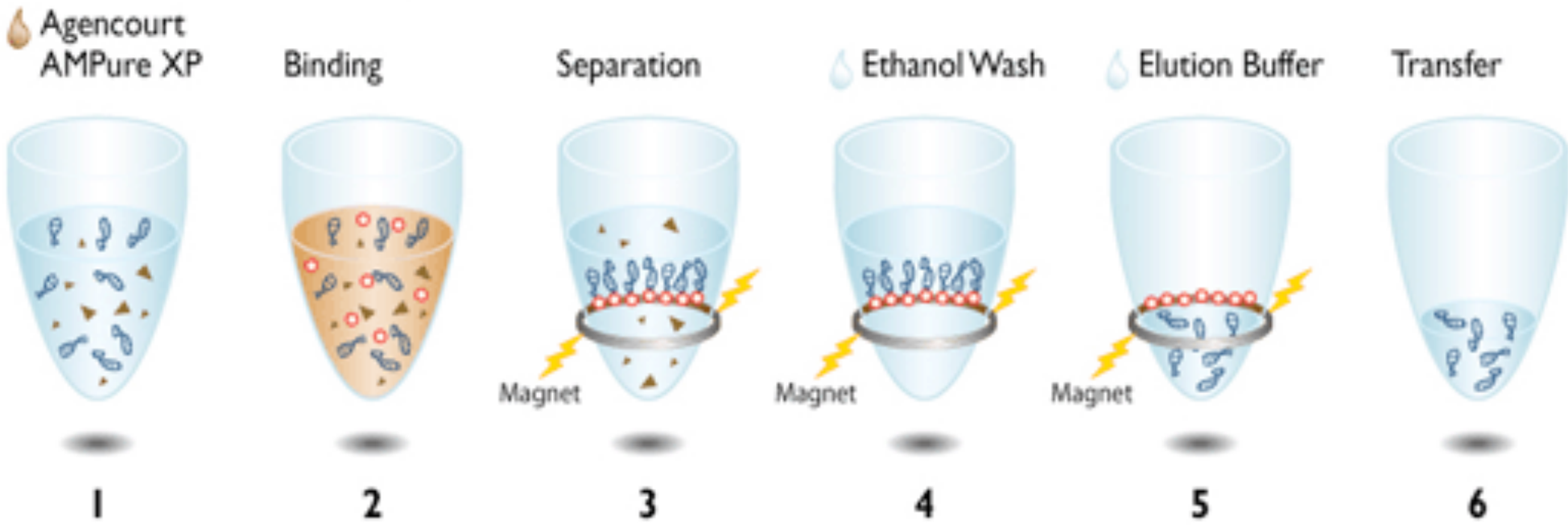
Clean Up and Size Selection: How?

- Solid Phase Reversible Immobilisation (SPRI) beads



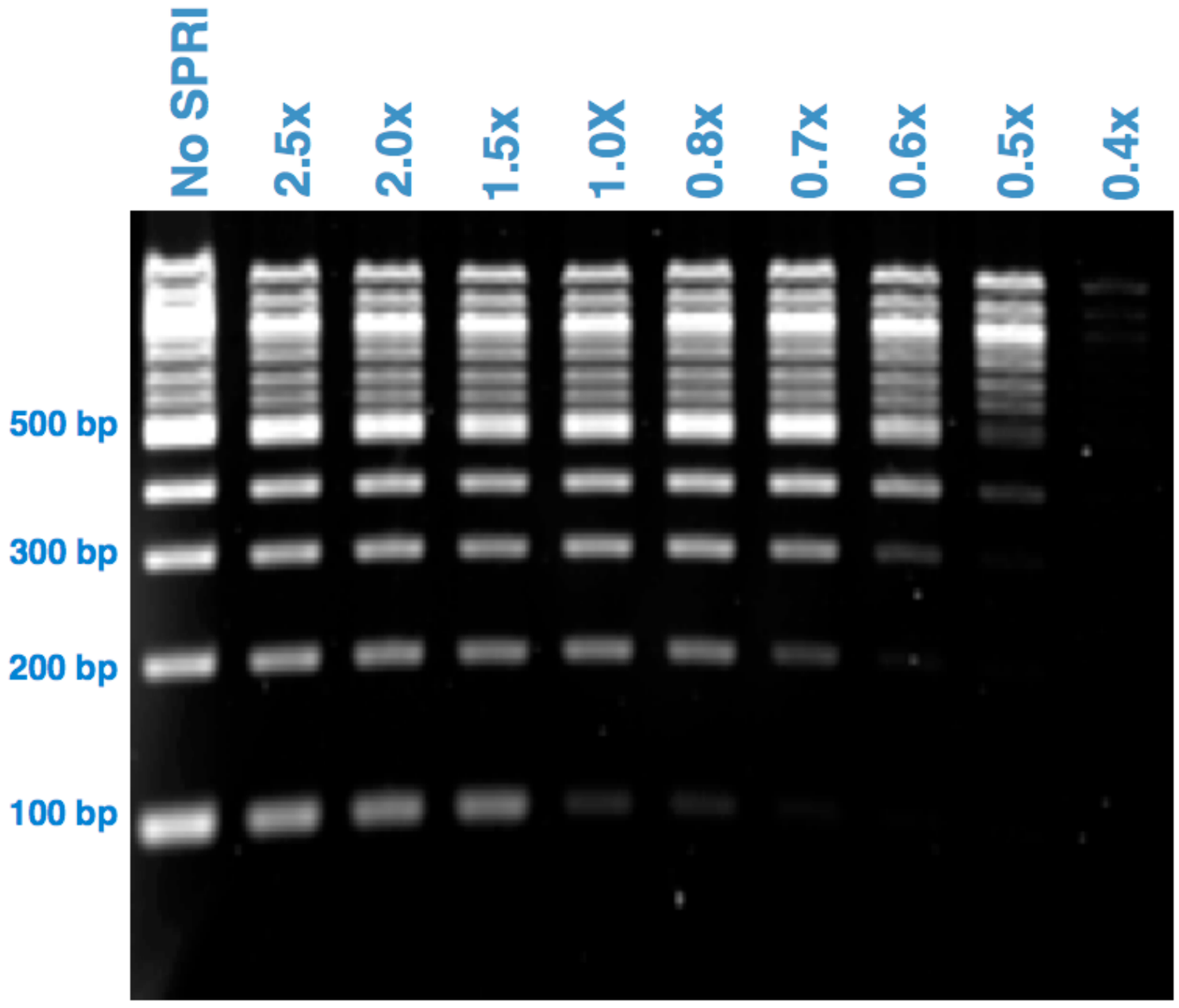
http://core-genomics.blogspot.com/2012/04/how-do-spri-beads-work.html

Clean Up and Size Selection: How?



https://www.beckmancoulter.com/ucm/idc/groups/public/documents/webasset/glb_bci_003489.gif

Size Selection: How?



http://core-genomics.blogspot.com/2012/04/how-do-spri-beads-work.html

Size Selection: How?

Option 2: Double-Sided SPRI

- By implementing a combination of good shearing with SPRI and "reverse" SPRI, one can select a fairly tight size range *with no gel*:

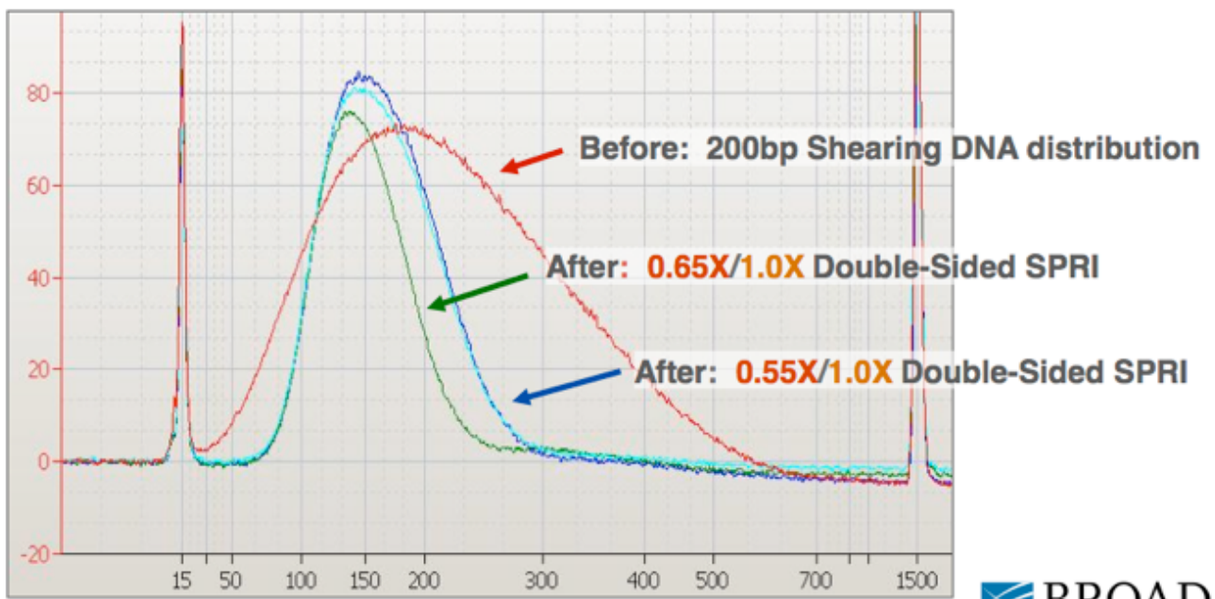
0.65X SPRI
Fragments >250 bp on beads
Fragments <250 bp in supernatant

KEEP SUPERNATANT
0-250 bp
DISCARD BEADS
"Reverse SPRI"

1.0X SPRI
Fragments >100 bp on beads
Fragments <100 bp in supernatant

KEEP BEADS
100-250 bp
DISCARD SUPERNATANT

Results:



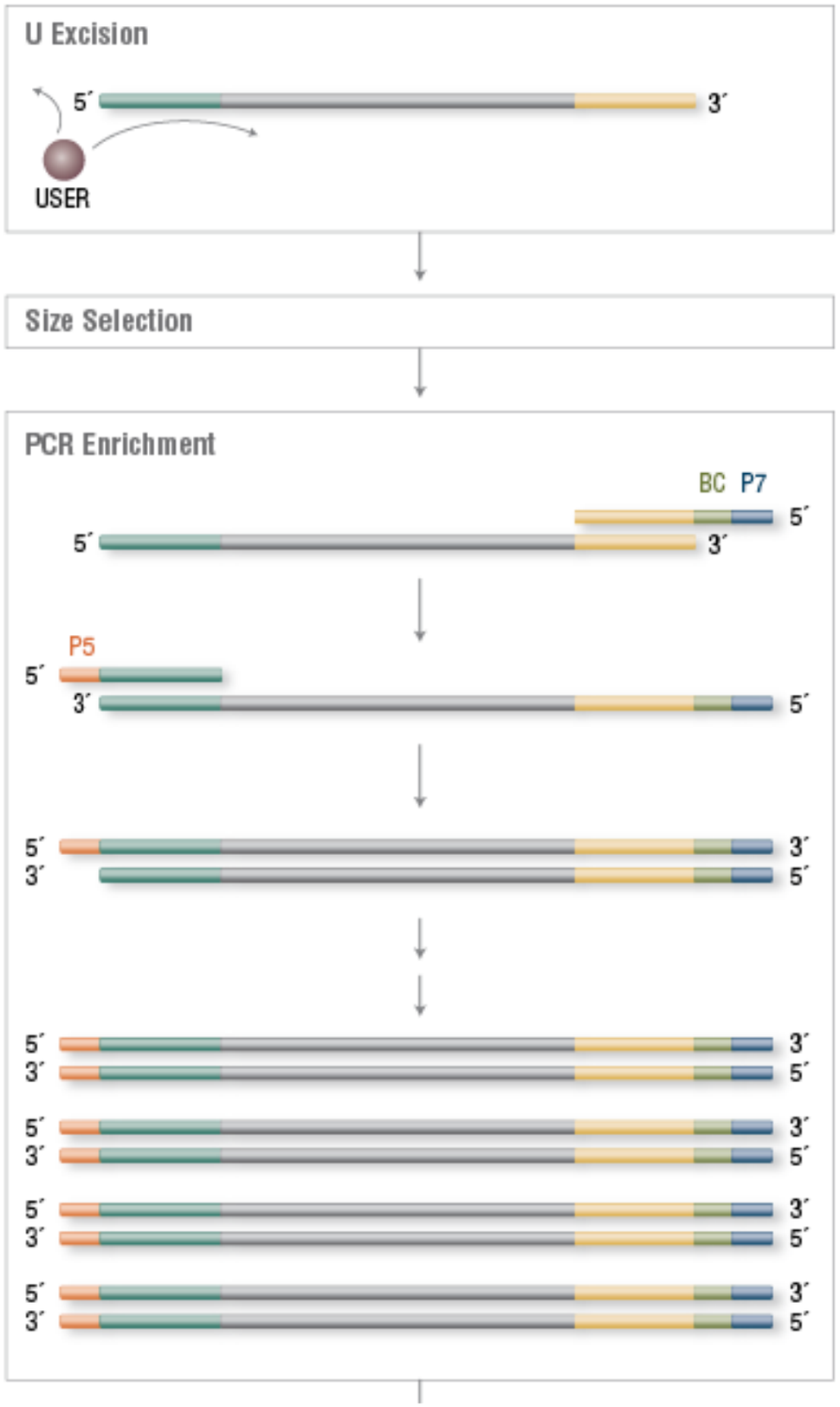
BROAD INSTITUTE illumina

http://core-genomics.blogspot.com/2012/04/how-do-spri-beads-work.html

Clean Up and Size Selection: Alternatives

- Spin Columns
- Gel Purification
- DIY SPRI

PCR Enrichment



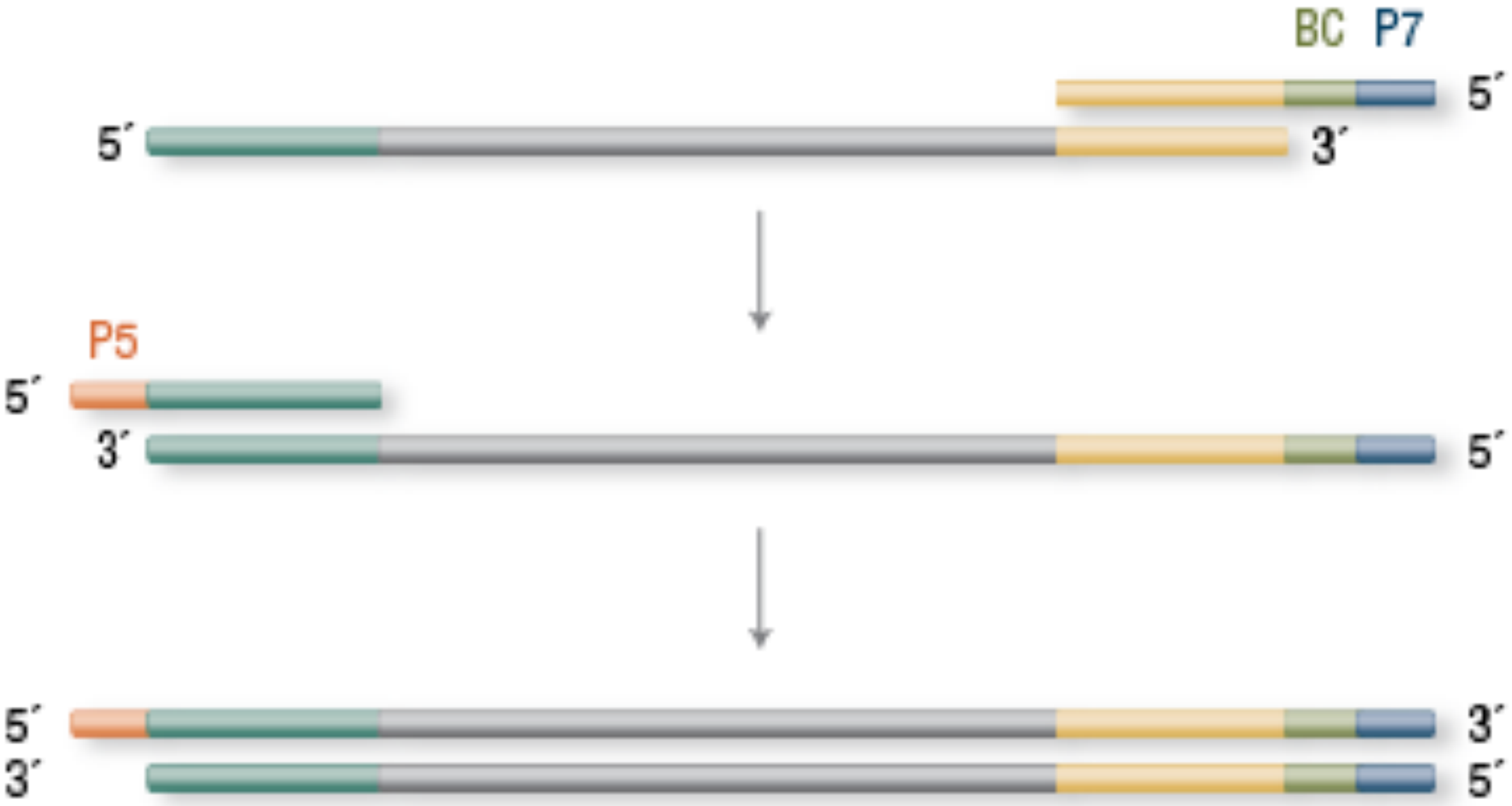
NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

PCR Enrichment

1. Extend adapter to full length
   A. add barcodes
   B. add priming sites
2. Amplify library
   A. Make more of the good fragments
   B. Leave the garbage in the dust

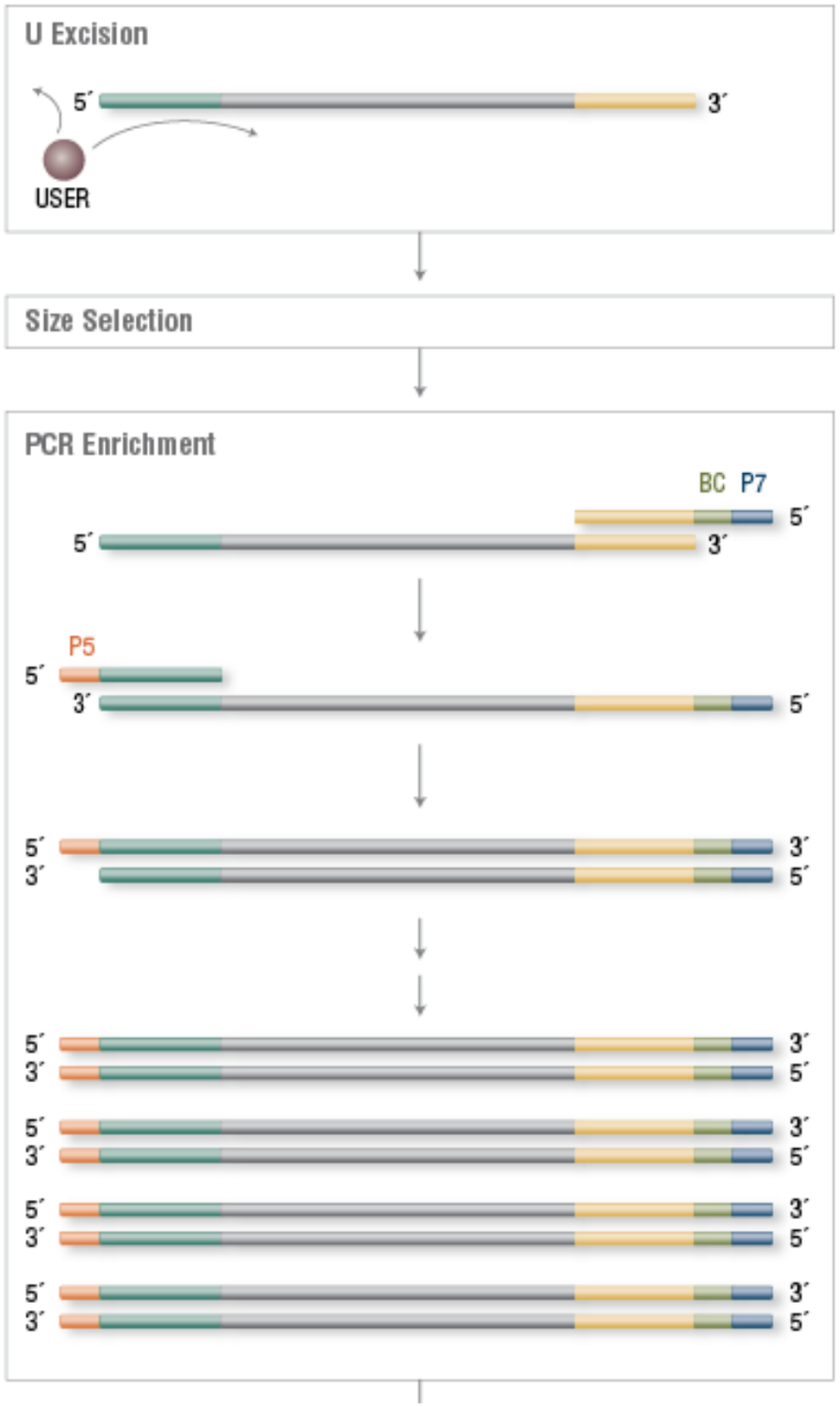
NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

PCR Enrichment



NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

PCR Enrichment



NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

Barcodes

Why?

Barcodes

Sample_Name	I7_Index_ID	index
1_A	P49-E1	AAGACCGT
2_A	P50-E2	TTGCGAGA
3_A	P51-E3	GCAATTCC
4_A	P52-E4	GAATCCGT
5_A	P53-E5	CCGCTTAA
6_A	P54-E6	TACCTGCA
7_B	P55-E7	GTCGATTG
8_B	P56-E8	TATGGCAC
9_B	P57-E9	CTCGAACA
10_B	P58-E10	CAACTCCA
11_B	P59-E11	GTCATCGT
12_B	P60-E12	GGACATCA
13_C	P61-F1	CAGGTTCA
14_C	P62-F2	GAACGAAG
15_C	P63-F3	CTCAGAAG
16_C	P64-F4	CATGAGCA
17_C	P65-F5	GACGAACT
18_C	P66-F6	AGACGCTA
19_D	P67-F7	ATAACGCC
20_D	P68-F8	GAATCACC
21_D	P69-F9	GGCAAGTT
22_D	P70-F10	GATCTTGC
23_D	P71-F11	CAATGCGA
24_D	P72-F12	GGTGTACA
25_E	P73-G1	TAGGAGCT
26_E	P74-G2	CGAATTGC
27_E	P75-G3	GTCCTAAG
28_E	P76-G4	CTTAGGAC
29_E	P77-G5	TCCACGTT
30_E	P78-G6	CAACACAG
31_F	P79-G7	GCCTTAAC
32_F	P80-G8	GTAAGGTG
33_F	P81-G9	AGCTACCA
34_F	P82-G10	CTTCACTG
35_F	P83-G11	GGTTGAAC
36_F	P84-G12	GATAGCCA
37_G	P85-H1	TACTCCAG
38_G	P86-H2	GGAAGAGA
39_G	P87-H3	GCGTTAGA
40_G	P88-H4	ATCTGACC
41_G	P89-H5	AACCAGAG
42_G	P90-H6	GTACCACA
43_H	P91-H7	GGTATAGG
44_H	P92-H8	CGAGAGAA
45_H	P93-H9	CAGCATAC
46_H	P94-H10	CTCGACTT
47_H	P95-H11	CTTCGGTT
48_H	P96-H12	CCACAACA

Barcodes

Sample_Name	I7_Index_ID	index
1_A	P49-E1	AAGACCGT
2_A	P50-E2	TTGCGAGA
3_A	P51-E3	GCAATTCC
4_A	P52-E4	GAATCCGT
5_A	P53-E5	CCGCTTAA
6_A	P54-E6	TACCTGCA
7_B	P55-E7	GTCGATTG
8_B	P56-E8	TATGGCAC
9_B	P57-E9	CTCGAACA
10_B	P58-E10	CAACTCCA

Nasty Stuff

- Sodium Azide
- Actinomycin D

Library Preparation: Alternatives

1. Illumina Kits
2. Other Kits
3. DIY

Additional Sequencing Details

WhY Adapter?

DNA Fragment

200-1000 bp

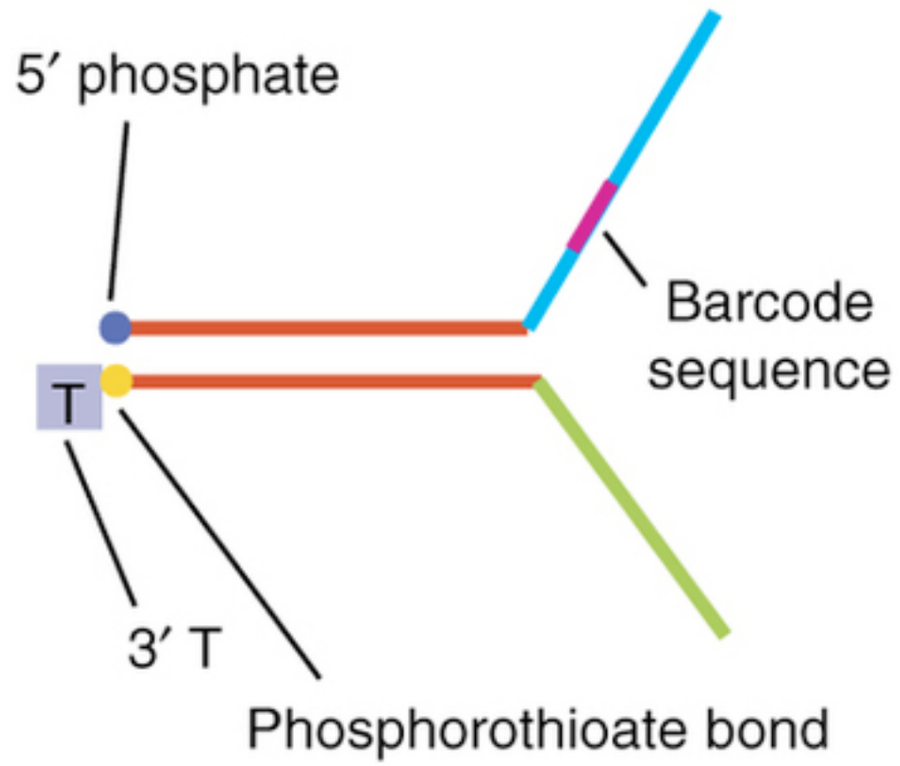
+

Adapters

=

Sequencing "Library"

Y Adapter?

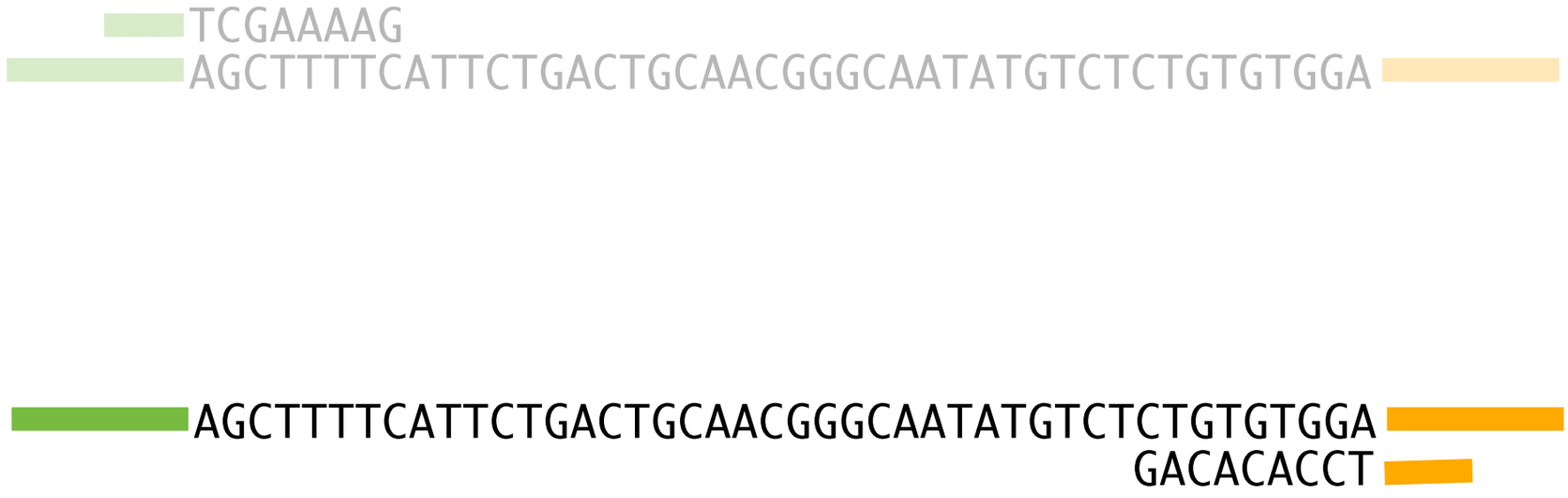


Nature Neuroscience 17, 1463–1475 (2014) doi:10.1038/nn.3814

Paired-End


TCGAAAAG
AGCTTTTCATTCTGACTGCAACGGGCAATATGTCTCTGTGTGGA

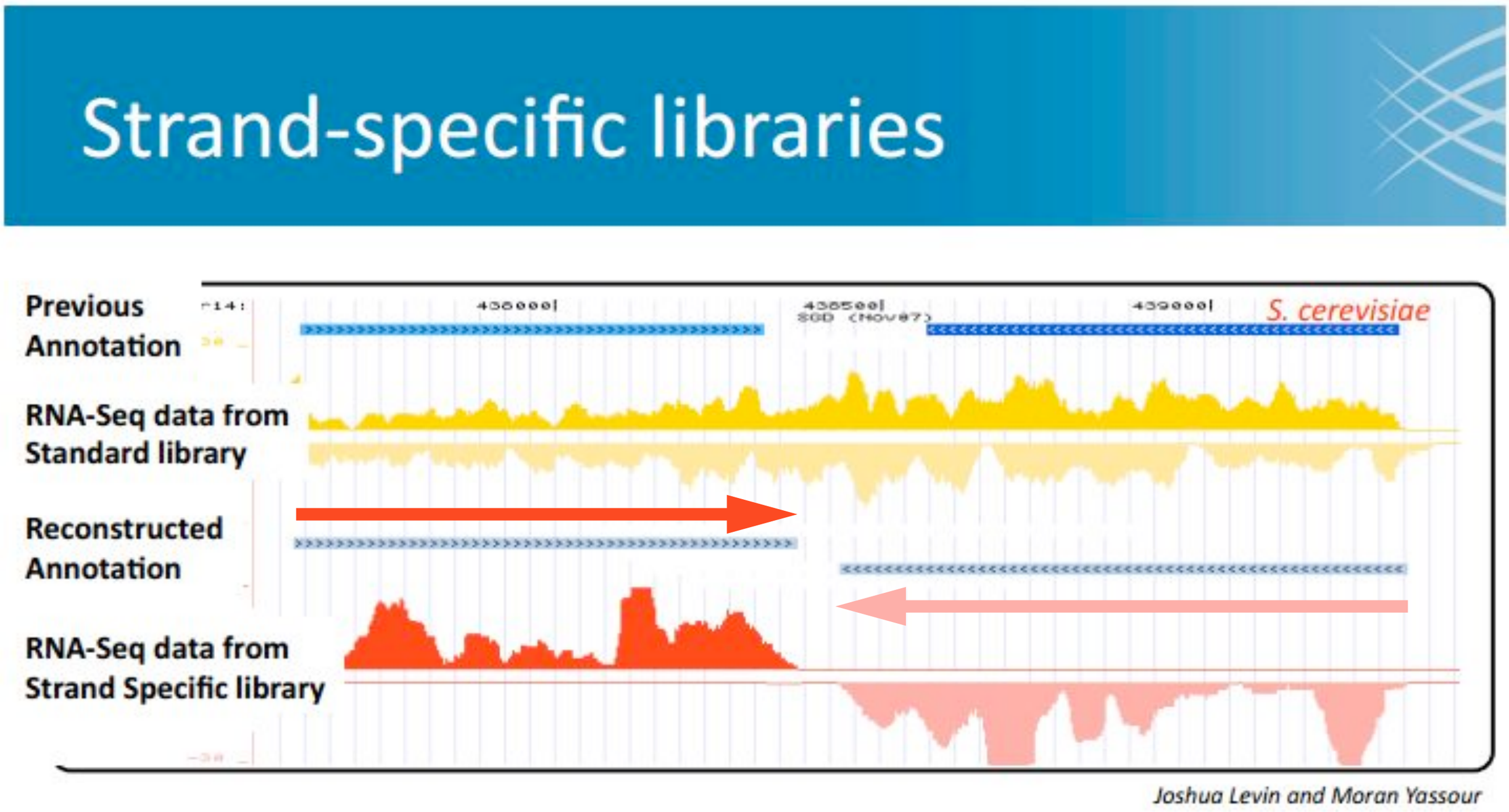
Paired-End


TCGAAAAG
AGCTTTTCATTCTGACTGCAACGGGCAATATGTCTCTGTGTGGA
AGCTTTTCATTCTGACTGCAACGGGCAATATGTCTCTGTGTGGA
GACACACCT

Strand-Specific Library

- Why Bother?

Strand-Specific Library



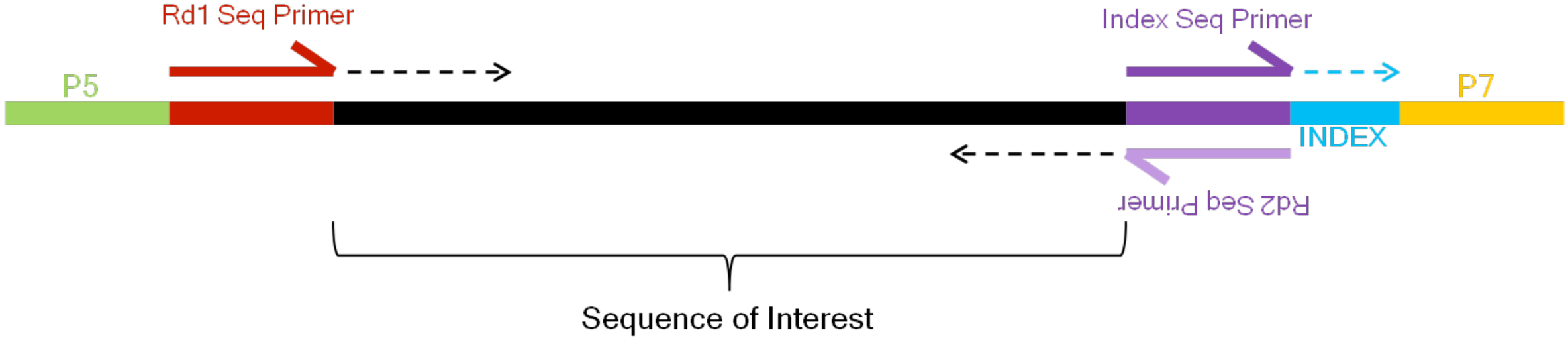
http://evomicsorg.wpengine.netdna-cdn.com/wp-content/uploads/2014/01/SSlib.jpg

Strand Specific Prep

NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 4.0

Multiplexing (Barcodes)

STRUCTURE DETAILS



http://nextgen.mgh.harvard.edu/IlluminaChemistry.html

HiSeq vs. MiSeq

	MiSeq	HiSeq
Maximum Output	15 Gb	1500 Gb
Maximum Reads per Run	25 million	5 billion
Maximum Read Length	2 × 300 bp	2 × 150 bp
Run Time	4–55 hours	7 hours – 6 days
Cost	$939	$1053
Cost/Mbp	$7.51	$0.0042

Illumina Video

https://www.youtube.com/watch?v=HMyCqWhwB8E

Acknowledgements

- NEB
- Illumina

Evaluation!