CELEMICS

Kit User Manual

NGS Library Preparation

With UMI Adapter for Cell-free DNA (cfDNA)

Illumina Platform Single / Dual Index

Including UMI Adapter Processing Pipeline

Version 2.0

Research Use Only

CELEMICS

Table of Contents

Materials Provided

Table 1. Library Preparation #1

Storage condition: -20°C

Box No.	Product Name	1 rxn (µL)	Cat No. LI1016 (16 rxn)	Cat No. LI1024 (24 rxn)	Cat No. LI1048 (48 rxn)	Cat No. LI1096 (96 rxn)
LI10	5X ER/A-Tailing Enzyme Mix	10	LI100416	LI100424	LI100448	LI100496
	10X ERA Buffer	5	LI100516	LI100524	LI100548	LI100596
	WGS Ligase	10	LI100616	LI100624	LI100648	LI100696
	5X Rapid Ligase Buffer	20	LI100716	LI100724	LI100748	LI100796

Table 2-1. Library Preparation #2 (Single Index)

Storage condition: -20°C

Box No.	Product Name	1 rxn (µL)	Cat No. LI2016 (16 rxn)	Cat No. LI2024 (24 rxn)
LI20	PCR Index Primer 1-24	2.5	LI200116 -LI201616	LI200124 -LI202424
	UMI Adapter	5	LI205816	LI205824
	USER	3	LI202616	LI202624
	Pre Capture PCR Forward Primer	2.5	LI202716	LI202724

Table 2-2. Library Preparation #2 (Dual Index)

Storage condition: -20°C

Box No.	Product Name	1 rxn (µL)	Cat No. LI2016 (16 rxn)	Cat No. LI2024 (24 rxn)	Cat No. LI2048 (48 rxn)	Cat No. LI2096 (96 rxn)
LI20	Dual Index 501-513	2.5	LI202816 - LI204016	LI202824 - LI204024	LI202848 - LI204048	LI202896 - LI204096
	Dual Index 701-716	2.5	LI204116 - LI205616	LI204124 - LI205624	LI204148 - LI205648	LI204196 - LI205696
	UMI Adapter	5	LI205816	LI205824	LI205848	LI205896
	USER	3	LI202616	LI202624	LI202648	LI202696

Additional Purchasable Materials

Table 3. Additional Purchasable Materials List

Product Name	Cat No.	Storage Condition (°C)	Comment
Celemics DNA purification bead (CeleMag Clean-up Bead)	CMCB50 (50 mL) CMCN500 (500 mL)	4°C	Included in All-in-one kit
Celemics Streptavidin bead (CeleMag Streptavidin Bead)	CMSB02 (2 mL) CMSB10 (10 mL)	4°C	
CLM Polymerase Amplification Kit PCR Ready Mix only	CMPFX.XX (X.XX: amount, mL)	-20°C	

Materials Required (Not Provided)

Table 4. Not Provided Materials List

Product Name	Suggested Supplier	
Reagents & Consumables		
Nuclease-free water (not DEPC-treated)	General lab supplier	-
Ethanol (200 proof)	General lab supplier	-
Conical centrifuge tubes (15 mL or 50 mL)	General lab supplier	-
P10,P20,P200 and P1000 pipettes	General lab supplier	-
Multichannel pipette	General lab supplier	-
Sterile, nuclease–free aerosol barrier pipette tips	General lab supplier	-
0.2 mL PCR Tube Strips, Low Profile and Cap Strips	General lab supplier	-
Powder-free gloves	General lab supplier	-
DNA LoBind Tubes, 1.5 mL PCR clean, 250 pieces	Eppendorf	p/n: 022431021 or equivalent
Microtubes 1.5ml Clear, Sterile	AXYGEN	p/n: MCT-150-C-S or equivalent
Eppendorf™ twin.tec™ 96 Well LoBind PCR Plates, Skirted	Eppendorf	p/n: 0030129512 or equivalent
Microseal 'B' adhesive seals	Bio-Rad	p/n: MSB-1001 or equivalent
Equipment		
TapeStation Instrument	Agilent Technologies, Inc.	p/n: G2991BA or equivalent
Magnetic separator	Invitrogen	p/n: 12321D or equivalent
Thermal Cycler	General lab supplier	-
Micro-centrifuge	General lab supplier	-
Vortex mixer	General lab supplier	-
Timer	General lab supplier	-

Before you begin

It is recommended to read the manual carefully to understand the experimental procedure. In particular, read and follow the text **highlighted in bold**.

Notice

The results of this experiment can be affected by the quality of the DNA. DNA extracted from FFPE tissue is highly variable in quality due to chemical damage and fragmentation, and often is available only in small amounts. There is a possibility that these steps will not lead to success with DNA extracted from FFPE samples. Therefore, it is recommended to use high-quality DNA samples.

Directions for the use of this product

1. This product should be used for research use only.
2. Handle specimens cautiously to avoid potential infections.
    1) Handle reagents cautiously and avoid contact with any area of the user's skin, eyes, and mucous membrane. In the event of reagent contact, rinse with water immediately to remove. If the reagent spills from the cartridge, dilute with water and follow the duly accredited treatment.
    2) All blood serum materials from human sources should be handled carefully and treated as potential infectious sources.
3. Genomic DNA/RNA (or cfDNA) extracted from bio-specimens such as blood, saliva, FFPE, and fresh tissue should be used as sample specimens.
4. Reagents should be used in accordance with the suggested Usage and Volume within the expiry date.
5. Reagent storage should be carried out at their suggested temperature and condition. Temperature conditions are listed below:

    Room temperature (RT, 15-25°C), Freezer(-20°C), Deep freezer (-80°C)
6. Consult an authorized management company for proper management of used and expired reagents in accordance with the relevant laws.
7. Be cautious with specific allergenic reagents.

Please maintain a nuclease-free environment during the experiment. Wear a lab coat, disposable gloves and protective goggles. If you have any problems, questions, or related concerns, please email us at support@celemics.com

Protocol Overview

NGS Library Preparation workflow	Time
DNA Fragmentation	0.5 hour
End-repair & A-tailing / Adapter ligation & USER treatment / Purification using CeleMag Clean-up Beads	2 hours
STOPPING POINT	
Amplification of the Adapter-Ligated Library / Purification	1.5 hour
STOPPING POINT	

Sample Preparation

Step 1. DNA Fragmentation

[Note]

Cell-free DNA (cfDNA) sample does not need fragmentation process, as it already exists in fragmented form.

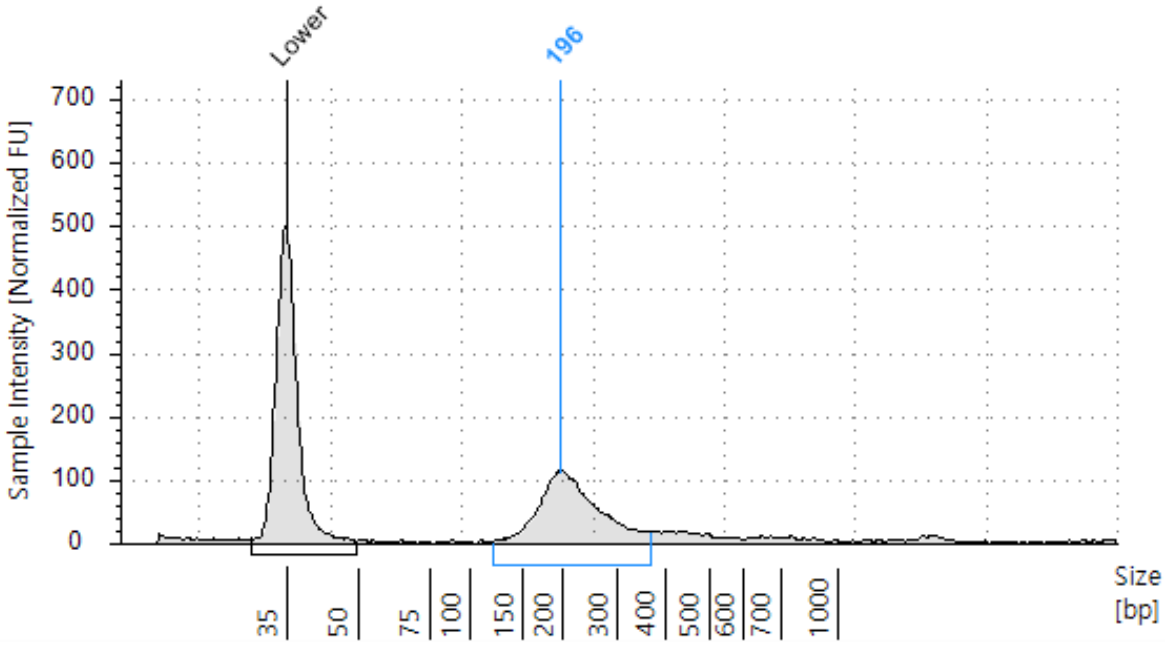

Figure 1. The electrophoresis diagram shows main peak size (196 bp).

1. Adjust the volume of the sample up to **35 µL** as below Table 5, if necessary.

Table 5. Adjusting the volume per sample

Reagent	Volume per sample (µL)
Nuclease-free water	35 - x
cfDNA	x
Total volume	35

Important
Proceed with the next step right away.

Step 2. Repair the Ends and A-tailing

[Preparation]

- From Library Preparation Box #1
- 5x ER/A-Tailing Enzyme Mix, 10x ERA buffer
- 0.2 mL PCR tubes and caps
- Ice bucket with ice or cooling block
- (Optional) Nuclease-free water

[Procedures]

1. Prepare **ERA buffer mix** in a PCR tube as described Table 6.
   * Important – Combining all reagents on ice and mix well by pipetting.

   Table 6. Composition for ERA buffer mix

Reagent	1 Reaction (µL)
10x ERA buffer	5
Eluted DNA sample	35
Total volume	40

2. **Add 10 µL of 5x ER/A Enzyme Mix** to each reaction tube and gently mix well.
   * Important – Combining it on ice and mix well by pipetting
   * Important – The final reaction volume is 50 µl

3. Spin down the sample tube and keep it on ice

4. Make the following program into thermal cycler (**lid temperature: 70˚C).**

   Table 7. Thermal cycler program

Step	Temperature	Time
1	4°C	1 min
2	20°C	30 min
3	65°C	30 min
4	4°C	Hold

5. Transfer the sample tube to thermocycler. Run the thermal cycler program.

   When the reaction is finished and block temperature reaches at 4˚C, transfer the End repaired and A-tailed sample **to ice**.

Important
Proceed with the next step right away.

Step 3. Adapter Ligation and USER Treatment

[Preparation]

- From Library Preparation Box #1 - 5x Rapid Ligase Buffer, WGS Ligase
- From Library Preparation Box #2 - UMI Adapter, USER
- 0.2 mL PCR tubes and caps
- Nuclease-free water

[Procedures]

1. **Add 5 µL of UMI Adapter** directly into the PCR tube where End repaired and A-tailed reaction was performed. Mix well with pipetting.

   Table 8. UMI Adapter sequence information

No.	Sequence	No.	Sequence	No.	Sequence	No.	Sequence
1	CGAATCGA	2	ATGGCGCA	3	ACCTGATC	4	CCAGTCTG
5	TTACTGTC	6	AATCCGAC	7	TAGCGGCA	8	CAACGTAG
9	TCCTTCCA	10	GATTATGC	11	GTGAACGG	12	AGTACACA
13	GCCTCAAG	14	GGTGAAGA	15	GTAGGTGA	16	GCTCTGCG
17	ATCCTGGC	18	AACCTCAG	19	TACGTCGA	20	ATCGCATA
21	CTATAGGC	22	ACTGGAAC	23	TAGCCAGG	24	CTTGTCGG
25	CAGTGCTG	26	AGGTTGCG	27	TCATGCTA	28	GGAGCACC
29	CAGCTGAA	30	TGTGATTC	-	-	-	-

2. Prepare a ligation reaction mix in a separate tube **on ice** and mix well with vortexing (add 5% of additional amount of reagents for multiple samples).
   * Important – Combining all reagents on ice

   Table 9. Composition for ligation reaction mix

Reagent	1 Reaction (µL)
5x Rapid Ligase Buffer	20
WGS Ligase	10
Nuclease-free water	15
Total volume	45

3. **Add 45 µL of the ligation reaction mix** to the sample from procedures 1.
   * Important – The final reaction volume is 100 µl

4. Incubate the ligation reaction mix at 20°C for 15 minutes (**lid temp.: 40°C**).

5. **Add 3 µL of USER** directly to the sample from step 4 and mix well with pipetting.

6. Incubate the sample at 37°C for 15 minutes (**lid temperature: 40°C).**

Important
Proceed with the next step right away.

Step 4. Purification of the Sample Using CeleMag Clean-up Beads

[Preparation]

- CeleMag Clean-up Beads
  Keep the beads at room temperature for at least 30 minutes before use.
  Store CeleMag Clean-up Beads at 4°C after use.
- 80% ethanol solution
- Nuclease-free water

[Procedure]

1. Thoroughly vortex CeleMag Clean-up Beads in order to achieve a homogeneous state.
2. **Add 100 µL CeleMag Clean-up Beads** to a new 1.5 mL LoBind tube and **add 103 µL adapter ligated sample** to the tube. Combine well on a vortex mixer and pulse-spin down the tube. Incubate at room temperature for 5 minutes (*avoid spinning down the CeleMag Clean-Up Beads too strongly to prevent precipitation*).

   Table 10. Volumes of CeleMag Clean-up Beads (1X)

Reagent	Volume Ratio (Beads : Sample)
CeleMag Clean-up Beads	**1 : 1** 100 µL : 103 µL

3. Place the tube in a magnetic separator for **3-5 minutes** until the solution is clear.
4. Keep the tube in the magnetic separator and discard the supernatant *(be careful not to touch the beads while you remove the supernatant)*.
5. Keeping the tube in the magnetic separator, **add 500 µL (180 µL for PCR tube or 96 well plate) fresh 80% ethanol** to each sample tube.
6. Incubate for 30 seconds at room temperature and then discard the 80% ethanol *(be careful not to touch the beads while you remove the supernatant)*.
7. Repeat 80% ethanol wash (**step 5-6**).
8. Spin down and put the tube in a magnetic separator and discard residual ethanol.
9. **Dry the sample tube** at room temperature for **2 minutes** or **until residual ethanol has completely evaporated** (*avoid drying beads until pellets appear cracked, as excessive drying decreases elution efficiency*).
10. **Add 22 µL of nuclease-free water** to the sample and mix well on a vortexer. Pulse-spin down and incubate the sample at room temperature for **2 minutes**.
11. Put the tube in the magnetic separator for **2 minutes** until solution is clear.
12. **Transfer the supernatant (20 µL)** to a new 1.5 mL LoBind Tube.

Stop Point

If do not continuing to the next step, samples can be stored at -20℃.

Step 5-1. Amplification of the Adapter-Ligated Library (Single Index)

[Note]

Determine the appropriate index primers for each sample (Table 14-1). Use different index primers in case of each sample to be sequenced in the same lane.

[Preparation]

- From Library Preparation Box #2 → Pre Capture Forward Primer, PCR index primer
- CLM Polymerase
- 0.2 mL PCR tubes and caps

[Procedures]

1. For **1** library
    a. Prepare 0.2 mL PCR tube and PCR mix as described in Table 11-1. Mix **PCR mix (30 µL)** and **Adapter-ligated DNA library (20 µL)**.

    *** Important** – Combining all reagents **on ice**.

    b. Run the thermal cycler with the PCR program as shown in Table 12-1 (see the number of PCR cycles in Table 13-1).

2. For **multiple** libraries
    a. Prepare new 0.2 mL PCR tubes and PCR mix as described in Table 11-1 (add 5% of additional amount of reagents for multiple libraries).

    b. **Add 25 µL of the CLM Polymerase** to each tube.

    c. **Add 2.5 µL of PCR Index Primer and 2.5 µL of Pre Capture PCR Forward Primer** to each tube.

    d. **Add 20 µL of Adapter-ligated DNA library** to each PCR tube, mix well and spin down.

    *** Important** – Combining all reagents **on ice**.

    e. Run the thermal cycler with the PCR program as shown in Table 12-1 (see the number of PCR cycles in Table 13-1).

Table 11-1 (Single Index). Composition for PCR mix

Reagent	Volume for 1 Library (µL)
PCR Index Primer (select one)	2.5
Pre Capture PCR Forward Primer	2.5
CLM Polymerase	25
Total volume	30

Table 12-1 (Single Index). PCR program

Step	Temperature	Time
Step 1	98°C	45 seconds
Step 2	98°C	15 seconds
Step 3	65°C	30 seconds
Step 4	72°C	1 minute
Step 5	-	Repeat step 2 to step 4 (Total PCR cycles, See in Table 13-1)
Step 6	72°C	10 minutes
Step 7	4°C	Hold

Table 13-1 (Single Index). The number of PCR cycles

Input DNA amount (ng), from page 7 - Step 1	Total PCR cycles
500 ng ≤ input DNA amount	6
100 ng ≤ input DNA amount < 500 ng	8
50 ng ≤ input DNA amount < 100 ng	10
input DNA amount < 50 ng	12

Table 14-1 (Single Index). Index sequences

Index	Sequence	Index	Sequence
PCR Index Primer 1	ATCACG	PCR Index Primer 13	AGTCAA
PCR Index Primer 2	CGATGT	PCR Index Primer 14	AGTTCC
PCR Index Primer 3	TTAGGC	PCR Index Primer 15	ATGTCA
PCR Index Primer 4	TGACCA	PCR Index Primer 16	CCGTCC
PCR Index Primer 5	ACAGTG	PCR Index Primer 17	GTCCGC
PCR Index Primer 6	GCCAAT	PCR Index Primer 18	GTGAAA
PCR Index Primer 7	CAGATC	PCR Index Primer 19	GTGGCC
PCR Index Primer 8	ACTTGA	PCR Index Primer 20	GTTTCG
PCR Index Primer 9	GATCAG	PCR Index Primer 21	CGTACG
PCR Index Primer 10	TAGCTT	PCR Index Primer 22	GAGTGG
PCR Index Primer 11	GGCTAC	PCR Index Primer 23	ACTGAT
PCR Index Primer 12	CTTGTA	PCR Index Primer 24	ATTCCT

Step 5-2. Amplification of the Adapter-Ligated Library (Dual Index)

[Note]

Determine the appropriate index primers for each sample (Table 14-2). Use different index primers in case of each sample to be sequenced in the same lane.

[Preparation]

- From Library Preparation Box #2 → Dual Index
- CLM Polymerase
- 0.2 mL PCR tubes and caps

[Procedures]

1. For **1** library
    a. Prepare 0.2 mL PCR tube and PCR mix as described in Table 11-2. Mix **PCR mix (30 µL)** and **Adapter-ligated DNA library (20 µL)**.

    *** Important** – Combining all reagents **on ice**.

    b. Run the thermal cycler with the PCR program as shown in Table 12-2 (see the number of PCR cycles in Table 13-2).

2. For **multiple** libraries
    a. Prepare new 0.2 mL PCR tubes and PCR mix as described in Table 11-2 (add 5% of additional amount of reagents for multiple libraries).

    b. **Add 25 µL of the CLM Polymerase** to each tube.

    c. **Add each 2.5 µL of dual index PCR Primers** to individual tube.

    * Important – Check appropriate combination of dual index in Table 14-2.

    d. **Add 20 µL of Adapter-ligated DNA library** to each PCR tube, mix well and spin down.

    * Important – Combining all reagents on ice.

    e. Run the thermal cycler with the PCR program as shown in Table 12-2 (see the number of PCR cycles in Table 13-2).

Table 11-2 (Dual Index). Composition for PCR mix

Reagent	Volume for 1 Library (µL)
Dual index primer 501-513 (select one)	2.5
Dual index primer 701-716 (select one)	2.5
CLM Polymerase	25
Total volume	30

Table 12-2 (Dual Index). PCR program

Step	Temperature	Time
Step 1	98°C	45 seconds
Step 2	98°C	15 seconds
Step 3	65°C	30 seconds
Step 4	72°C	1 minute
Step 5	-	Repeat step 2 to step 4 (Total PCR cycles, See in Table 13-2)
Step 6	72°C	10 minutes
Step 7	4°C	Hold

Table 13-2 (Dual Index). The number of PCR cycles

Input DNA amount (ng), from page 7 - Step 1	Total PCR Cycles
500 ng ≤ input DNA amount	6
100 ng ≤ input DNA amount < 500 ng	8
50 ng ≤ input DNA amount < 100 ng	10
input DNA amount < 50 ng	12

Table 14-2 (Dual Index). Index sequences

Index	Sequence	Index	Sequence
Dual index 501 primer	TATAGCCT	Dual index 701 primer	ATTACTCG
Dual index 502 primer	ATAGAGGC	Dual index 702 primer	TCCGGAGA
Dual index 503 primer	CCTATCCT	Dual index 703 primer	CGCTCATT
Dual index 504 primer	GGCTCTGA	Dual index 704 primer	GAGATTCC
Dual index 505 primer	AGGCGAAG	Dual index 705 primer	ATTCAGAA
Dual index 506 primer	TAATCTTA	Dual index 706 primer	GAATTCGT
Dual index 507 primer	CAGGACGT	Dual index 707 primer	CTGAAGCT
Dual index 508 primer	GTACTGAC	Dual index 708 primer	TAATGCGC
Dual index 509 primer	GACCTGTA	Dual index 709 primer	CGGCTATG
Dual index 510 primer	CGGTGGTA	Dual index 710 primer	TCCGCGAA
Dual index 511 primer	GTTGGACT	Dual index 711 primer	TCTCGCGC
Dual index 512 primer	CTAAGATC	Dual index 712 primer	AGCGATAG
Dual index 513 primer	AATGGTTC	Dual index 713 primer	GCATAGTG
-	-	Dual index 714 primer	CATTCCAG
-	-	Dual index 715 primer	GGCAGGAT
-	-	Dual index 716 primer	CAACAACA

Important

Proceed with the next step right away.

Step 6. Purification of the Sample Using CeleMag Clean-up Beads

[Preparation]

- CeleMag Clean-up Beads
  Keep the beads at room temperature for at least 30 minutes before use.
  Store CeleMag Clean-up Beads at 4°C after use.
- 80% ethanol solution
- Nuclease-free water

[Procedure]

1. Thoroughly vortex CeleMag Clean-up Beads in order to achieve a homogeneous state.
2. **Add 50 µL CeleMag Clean-up Beads** to a new 1.5 mL LoBind tube and **add 50 µL amplified adapter-ligated sample** to the tube. Combine well on a vortex mixer and pulse-spin down the tube. Incubate at room temperature for 5 minutes (*avoid spinning down the CeleMag Clean-Up Beads too strongly to prevent precipitation*).

   Table 15. Volume of CeleMag Clean-up Beads (1X)

Reagent	Volume Ratio (Beads : Sample)
CeleMag Clean-up Beads	**1 : 1** 50 µL : 50 µL

3. Place the tube in a magnetic separator for **3-5 minutes** until the solution is clear.
4. Keep the tube in the magnetic separator and discard the supernatant *(be careful not to touch the beads while you remove the supernatant)*.
5. Keeping the tube in the magnetic separator, **add 500 µL (180 µL for PCR tube or 96 well plate) fresh 80% ethanol** to each sample tube.
6. Incubate for 30 seconds at room temperature and then discard the 80% ethanol *(be careful not to touch the beads while you remove the supernatant)*.
7. Repeat 80% ethanol wash (**step 5-6**).
8. Spin down and put the tube in the magnetic separator and discard residual ethanol.
9. **Dry the sample tube** at room temperature for **2 minutes** or **until residual ethanol has completely evaporated** (*avoid drying beads until pellets appear cracked, as excessive drying decreases elution efficiency*).
10. **Add 32 µL of nuclease-free water** to the sample and mix well on a vortexer. Pulse-spin down and incubate the sample at room temperature for **2 minutes**.
11. Put the tube in the magnetic separator for **2 minutes** until the solution is clear.
12. **Transfer the supernatant (30 µL)** to a new 1.5 mL LoBind Tube.

Stop Point
If not continuing to the next step, the sample can be stored at -20°C.

Celemics UMI Analysis System

1. Design of UMI Adapter

[Feature and uses of UMI Adapter]

- UMI Adapters are used to identify PCR duplicates through sequence combinations unique to each DNA molecule.
- UMI Adapter sequence is inserted at the end of the NGS adapter sequence and read from the front of the read (it does not affect Illumina sample indexing system).

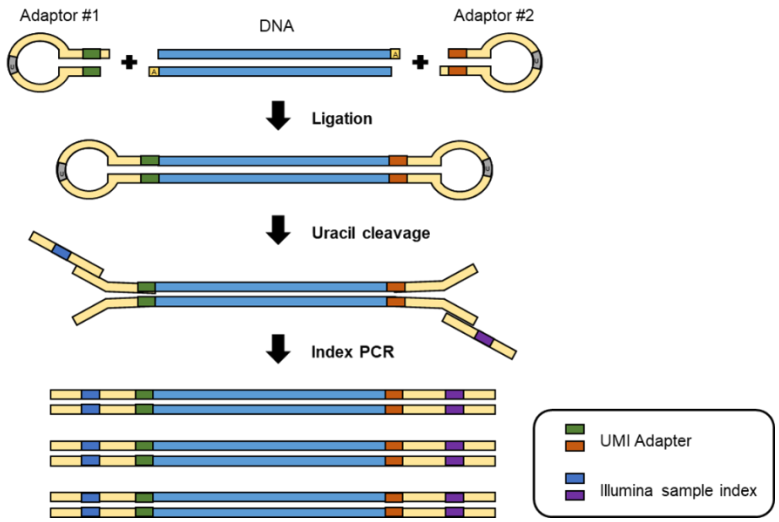

Figure 2. The process of how UMI adapter is inserted in NGS preparation

Figure 3. NGS data sample showings UMI Adapter attached to the front of the read

- No additional reagents or experiment steps needed to introduce UMI Analysis System as below Figure 4.

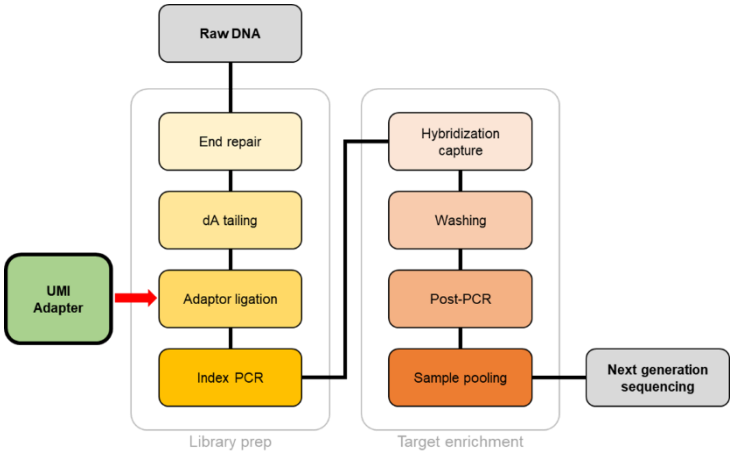

Figure 4. NGS workflow using UMI

2. UMI Analysis Algorithm Overview

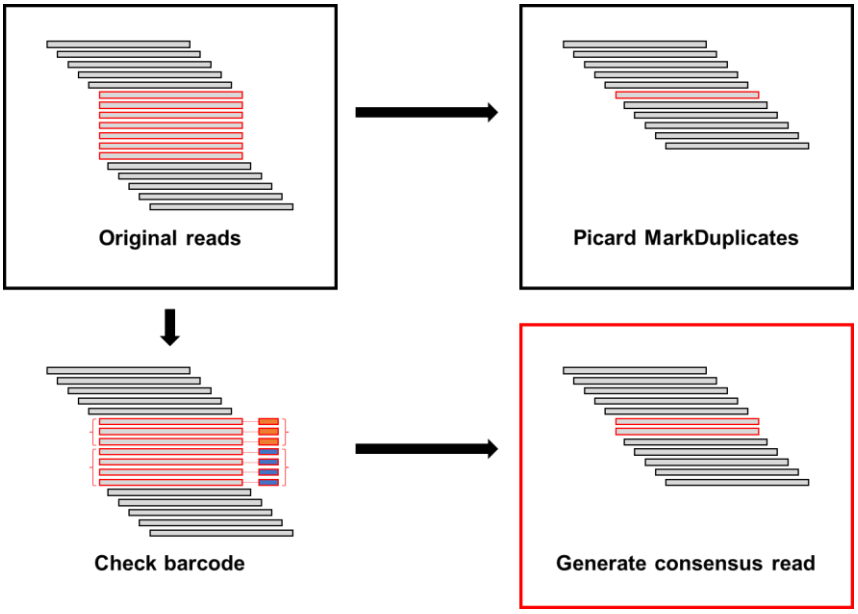

Figure 6. Comparison of conventional analysis methods and UMI analysis methods in the process of removing duplicates

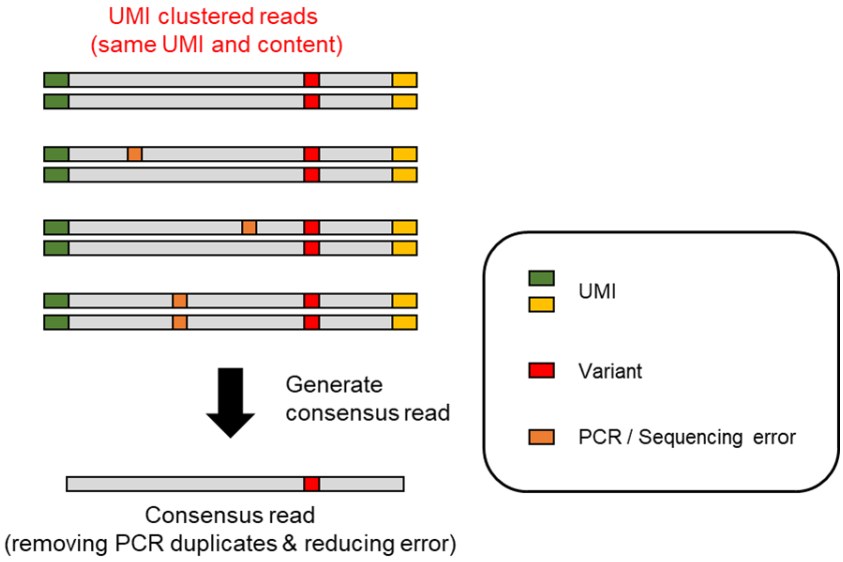

Figure7. Error correction using UMI

- Reads with same alignment position and UMI combination are clustered.
- Consensus read are generated from clustered read groups by considering the number of base calls and the quality score.
- Higher depth can be obtained than conventional analysis methods (Picard MarkDuplicates).
- PCR or sequencing errors can be eliminated during generating consensus reads because they occur randomly within each read

Celemics UMI Processing Pipeline

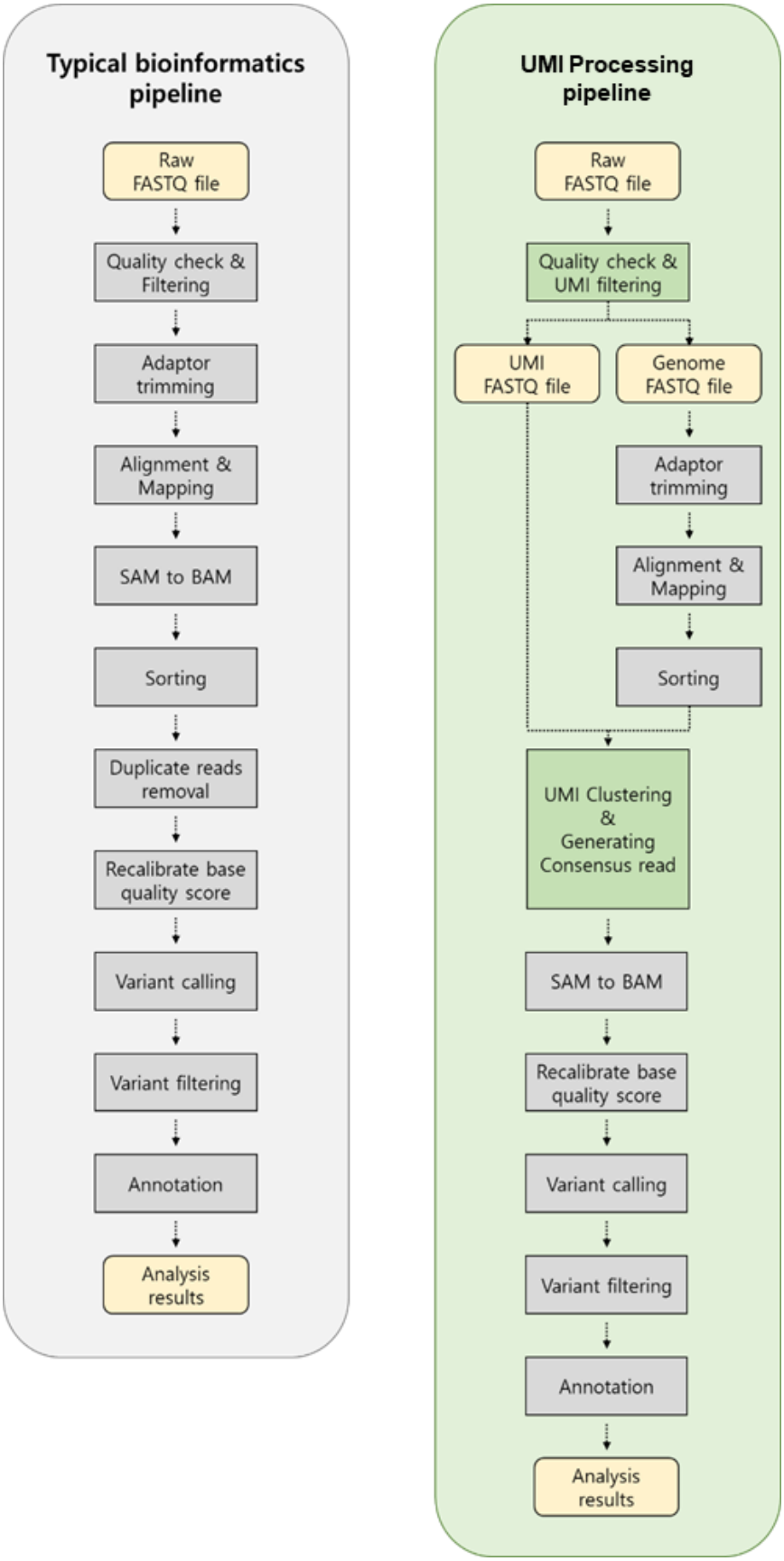

Figure 4. Whole workflow of UMI processing pipeline

- UMI analysis program is also provided (for Linux, CLI program)
- UMI analysis program can be easily inserted into typical analysis pipelines

[Celemics UMI Analysis program]

UMI program in Celemics consists of two programs.

- Celemics_UMI_filter : Separates UMI from FASTQ file after demultiplexing.
- Celemics_UMI_processor : produce UMI clustering and consensus read

The following is an analysis pipeline example for utilizing **Celemics UMI Analysis program**.

Programs except Celemics programs in pipelines can be replaced by other program with the same function.

* The input name used in the code example is in the following format.

Both gzip compressed file and FASTQ file are available.

- Sample1_R1.fastq.gz
- Sample1_R2.fastq.gz

  or

- Sample1_R1.fastq
- Sample2_R2.fastq

1. Quality check & UMI filtering (using Celemics_UMI_Filter)

Celemics_UMI_filter program is used as the first course for UMI analysis.

In this process, low quality reads or unrecognized UMI reads are discarded.

```
$ Celemics_UMI_filter ./Sample1
```

The following files are generated as a result of running a program

Sample1_filtered_R1.fastq

Sample1_filtered_R2.fastq

Sample1_UMI_R1.fastq

Sample1_UMI_R2.fastq

2. Adapter trimming

The process is to remove the adaptor sequence if it is read at the end of the read due to the short insert DNA.

The following example uses the 'Adapterremoval' program and can be replaced by another program with the same functionality.

The input files of the program to be used in this process are as follows.

Sample1_filtered_R1.fastq

Sample1_filtered_R2.fastq

```
$ AdapterRemoval --file1 ./Sample1_filtered_R1.fastq --file2 ./Sample1_filtered_R2.fastq ₩
--basename ./Sample1 ₩
--trimns --trimqualities --minquality 20 --minlength 30 ₩
--adapter1 AGATCGGAAGAGCACACGTCTGAACTCCAGTCACNNNNNNNNATCTCGTATGCCGTCTTCTGCTTG ₩
--adapter2 GATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTNNNNNNNNGTGTAGATCTCGGTGGTCGCCGTATCATT ₩
--output1 ./Sample1_trimmed_R1.fastq --output2 ./Sample1_trimmed_R2.fastq
```

The following files are generated as a result of running a program.

Sample1_trimmed_R1.fastq

Sample1_trimmed_R2.fastq

3. Alignment & Mapping

The process is to arrange the sequencing read to each position on the reference genome.

The following example uses the 'bwa mem' program and can be replaced by another program with the same functionality.

The input files of the program to be used in this process are as follows.

Sample1_trimmed_R1.fastq

Sample1_trimmed_R2.fastq

```
$ bwa mem -M ₩

-R '@RG₩tID:RunID₩tLB:Library1₩tPL:SeqPlatform₩tPU:RunName₩tSM:Sample1' ₩

./hg19.fa ₩

./Sample1_trimmed_R1.fastq ./Sample1_trimmed_R2.fastq ₩

>./Sample1.sam
```

The following files are generated as a result of running a program.

Sample1.sam

4. Sorting SAM files

The process is to **arrange SAM files**, as a result of Alignment & Mapping, **in coordinate order**.

The following example uses the 'samtools' program and can be replaced by another program with the same functionality.

The input files of the program to be used in this process are as follows.

Sample1.sam

```
$ samtools sort -O sam -o ./Sample1.sorted.sam -T ./Sample1.temp ./Sample1.sam
```

The following files are generated as a result of running a program.

Sample1.sorted.sam

5. UMI clustering & Generating consensus read (using Celemics_UMI_processor)

This is the process of using a Celemics program to perform UMI clustering and to create a consensus read within each cluster.

The input files of the program to be used in this process are as follows.

Sample1_UMI_R1.fastq

Sample1_UMI_R2.fastq

Sample1.sorted.sam

```
$ Celemics_UMI_processor ₩
-1 ./Sample1_UMI_R1.fastq ₩
-2 ./Sample1_UMI_R2.fastq ₩
-S ./Sample1.sorted.sam -o ./Sample1_css
```

The following files are generated as a result of running a program.

Sample1_css.processed.sam

6. SAM to BAM conversion

The process of creating a Consensus read and converting SAM files into BAM files.

The following example uses the 'samtools' program and can be replaced by another program with the same functionality.

The input files of the program to be used in this process are as follows.

Sample1_css.processed.sam

```
$ samtools view -bt ./hg19.fa -o ./Sample1.bam ./Sample1_css.consensus.sam
```

The following files are generated as a result of running a program.

Sample1.bam

7. Sorting BAM

The process is to arrange BAM files in coordinate order after creating consensus read.

The following example uses the 'samtools' program and can be replaced by another program with the same functionality.

The input files of the program to be used in this process are as follows.

Sample1.bam

```
$ samtools sort -O bam -o ./Sample1.sorted.bam -T ./Sample1.temp ./Sample1.bam
```

The following files are generated as a result of running a program.

Sample1.sorted.bam

The BAM file generated after performing the above pipeline is the same as the standard BAM format, which can then be applied to a typical bioinformatics pipeline.