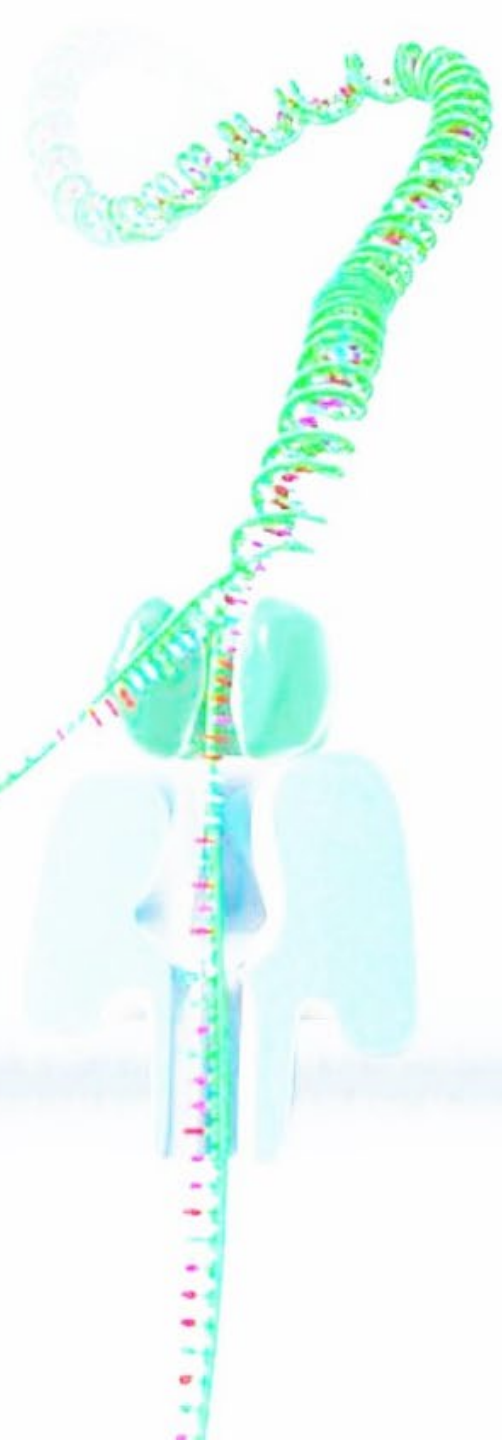


2025 한국식물분류학회 Technical Workshop

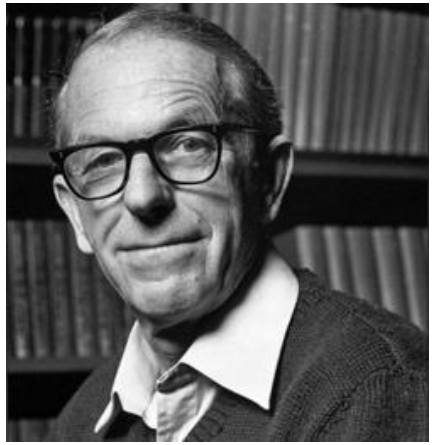
Long-read Nanopore sequencing의 활용한 식물 미토콘드리아 유전체 결정

성신여자대학교 바이오생명공학과 김상태

생물학의 발달 과정

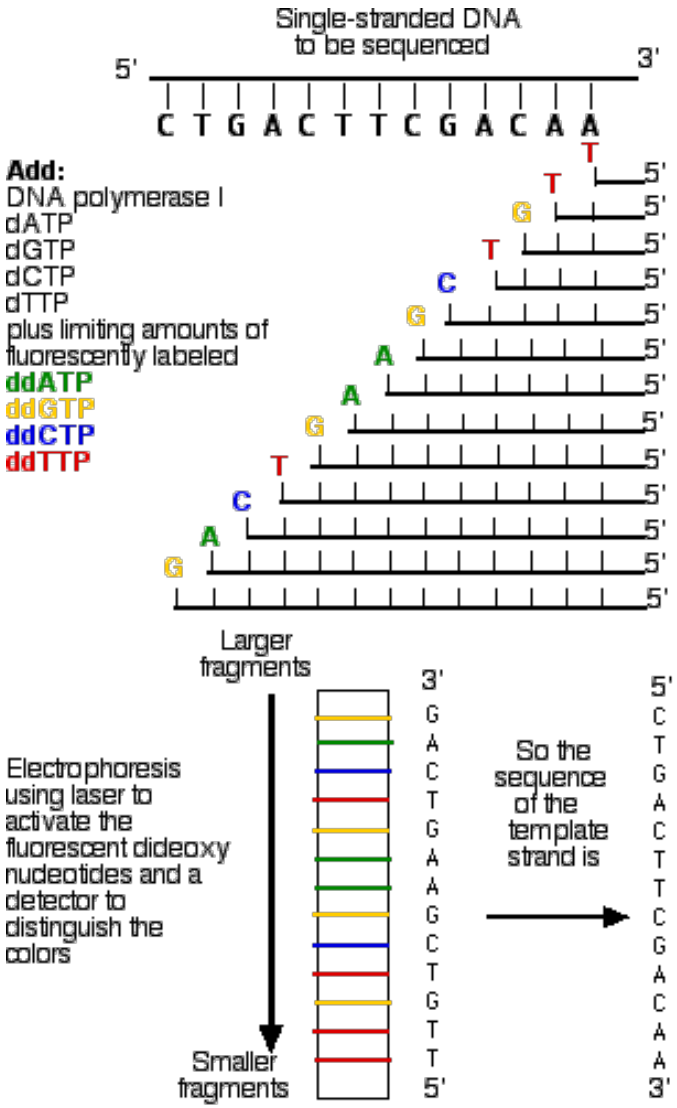
Ⅰ. Introduction: 염기서열 결정의 진화

The first generation sequencing: Sanger sequencing



Frederick Sanger (영국)

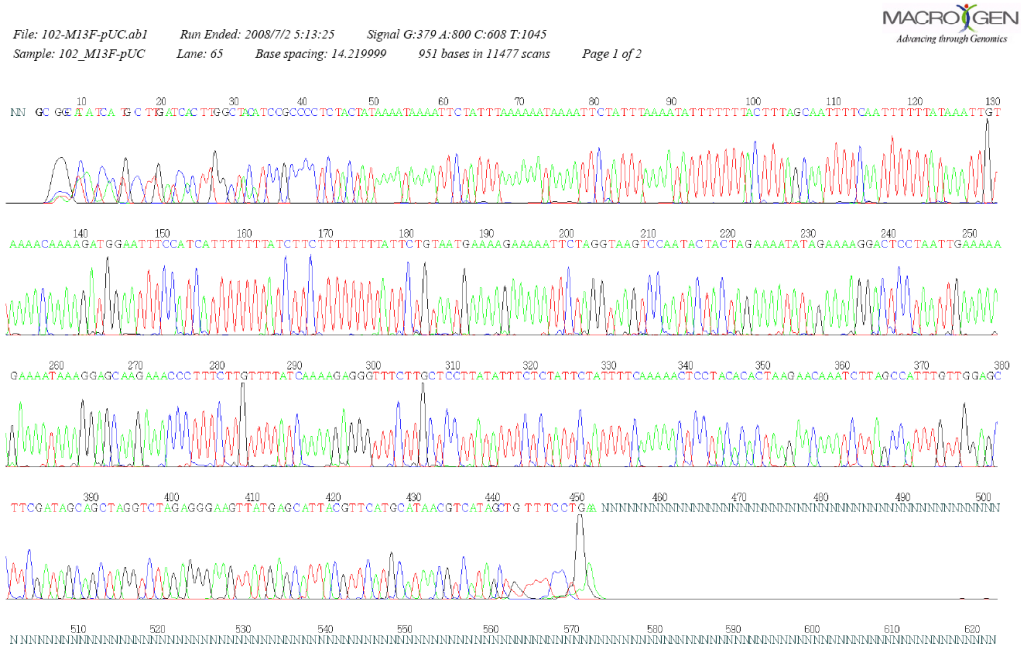
- 두 번의 노벨상 수상자
- "termination" method
- One-dye four lane system에서 four-dye one lane system으로 발전.



현재 가장 많이 쓰는 모델: ABI 3730 (Thermo Fisher)

- Applied Biosystems Co.
  → Thermo Fisher Co.

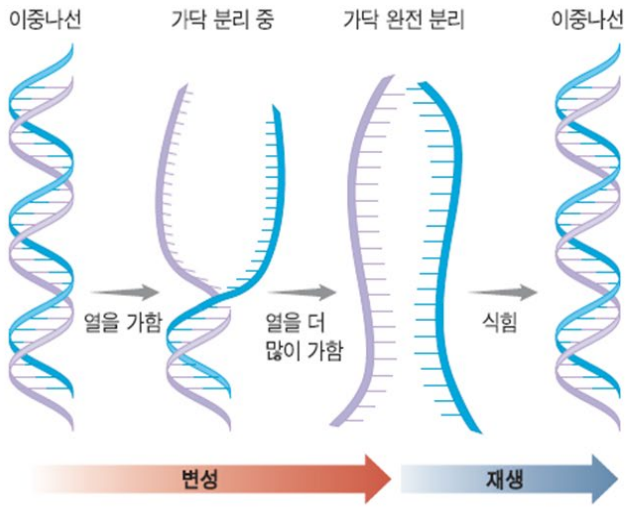
- Ca. 900 bp/capillary
- 384 capillary/run
- 900 X 384
  = ca. 350 kbp

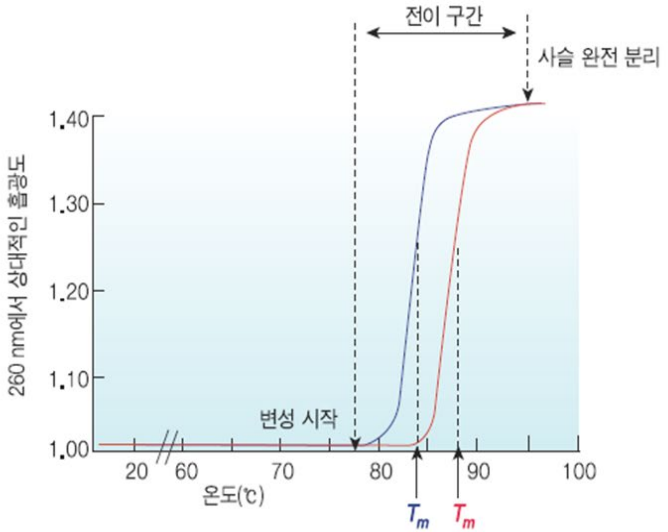


※ 참고
일반생물학

Denaturation and Renaturation

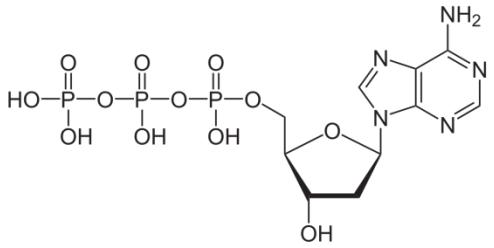
DNA 사슬의 뼈대를 이루는 인산과 당의 결합은 매우 단단한 공유결합이지만, 두 가닥의 사슬이 서로 붙는 염기쌍이 이루는 결합은 상대적으로 약한 수소결합으로 이루어져 있다. 그러므로 높은 온도에서 또는 높은 pH에서 DNA 두 사슬은 서로 분리되어 단일사슬이 되며, 이를 DNA의 **변성(denaturation)**이라고 한다. 높은 온도에서 변성이 일어난 DNA를 다시 온도를 낮추면 염기들의 배열 순서에 따라 서로 상보적인 DNA 사슬들이 결합하여 다시 이중나선의 구조를 이루게 되고, 이를 **재생(renaturation)**이라 한다. G와 C에 의한 DNA의 결합은 3중 수소결합이기 때문에 전체 DNA에서 G와 C의 비율이 높아지면 DNA가 단일사슬로 변하는 중간온도(Tm 값)가 더 높아진다.





※ 참고
일반생물학

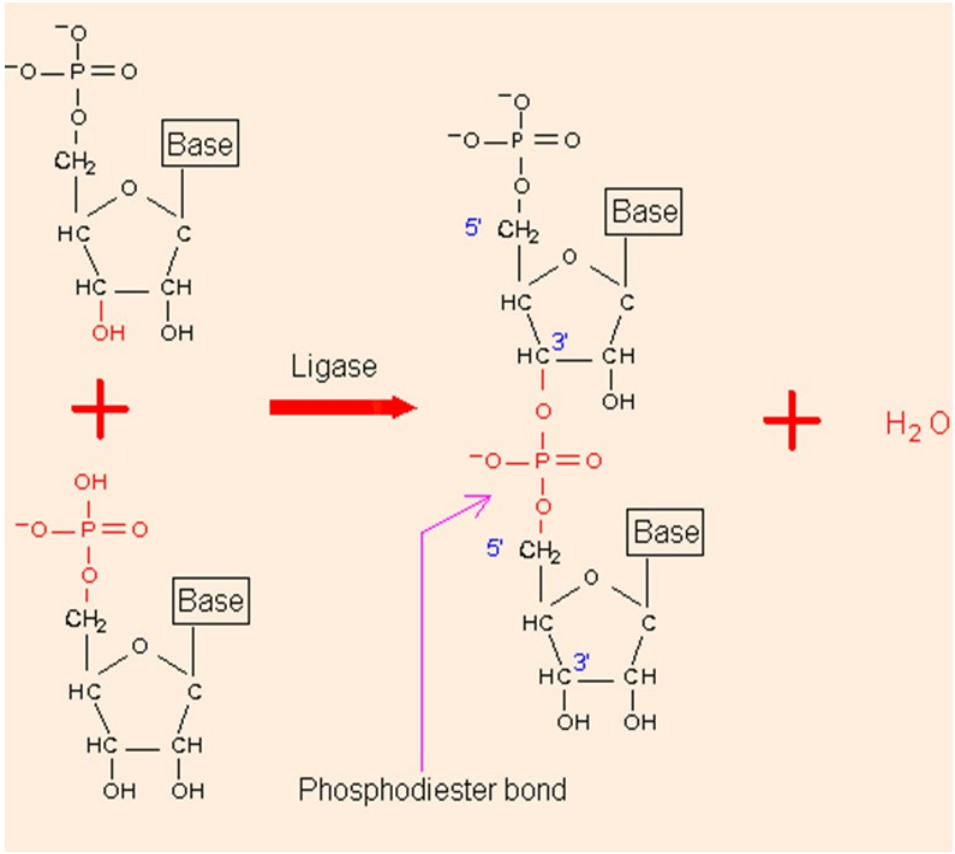
뉴클레오티드들의 탈수축합에 의한 중합과정



dATP (deoxyadenosine triphosphate) 의 구조

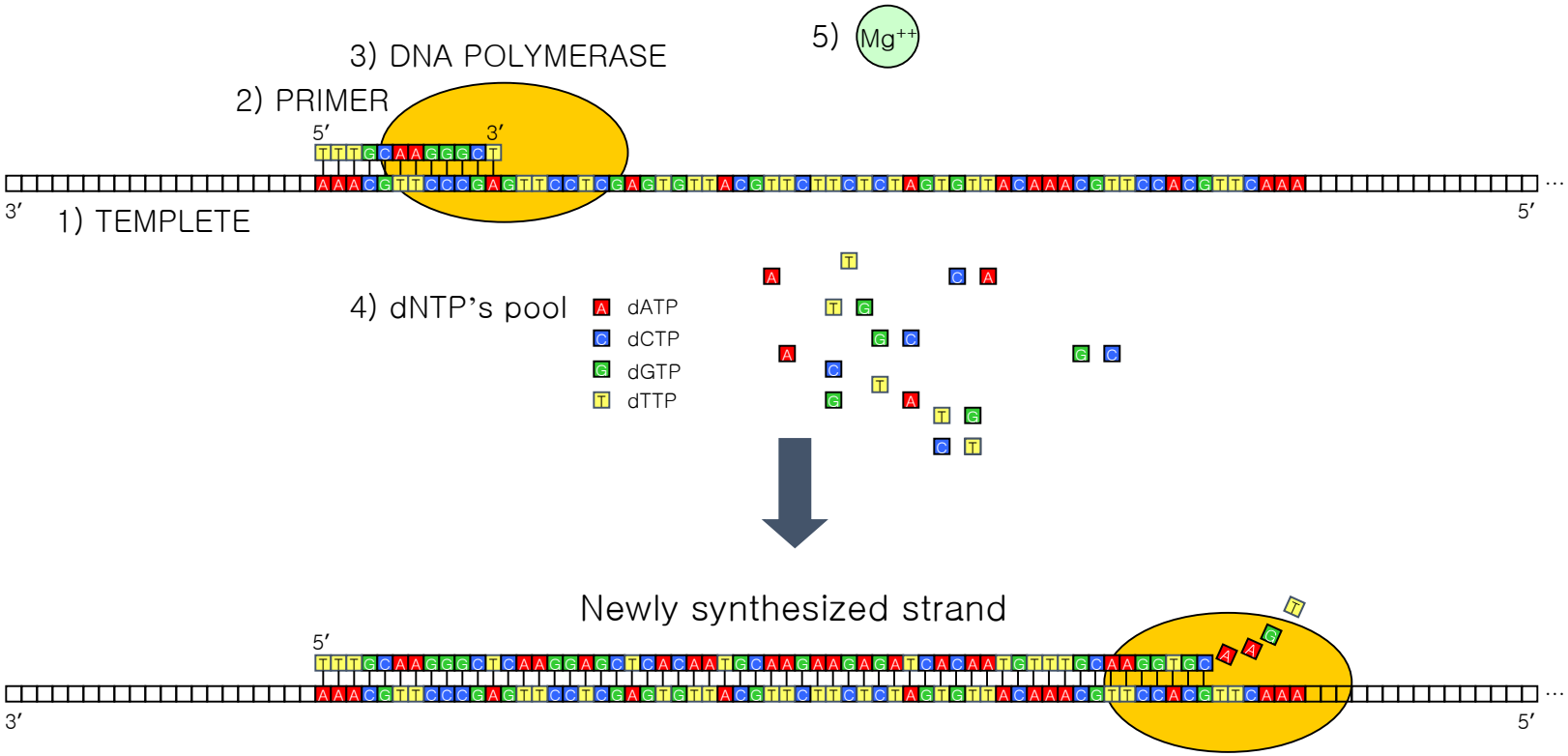
dATP, dCTP, dGTP, dTTP 를 합쳐서 dNTP 라고 함.

- DNA의 합성은 반드시 앞선 뉴클레오티드의 3번 탄소와 뒤쪽 뉴클레오티드의 5번 탄소간에 일어난다. 즉 DNA는 뉴클레오티드의 3번탄소 방향으로만 합성이 일어난다.



※ 참고
일반생물학

DNA 중합효소에 의한 DNA 합성과정

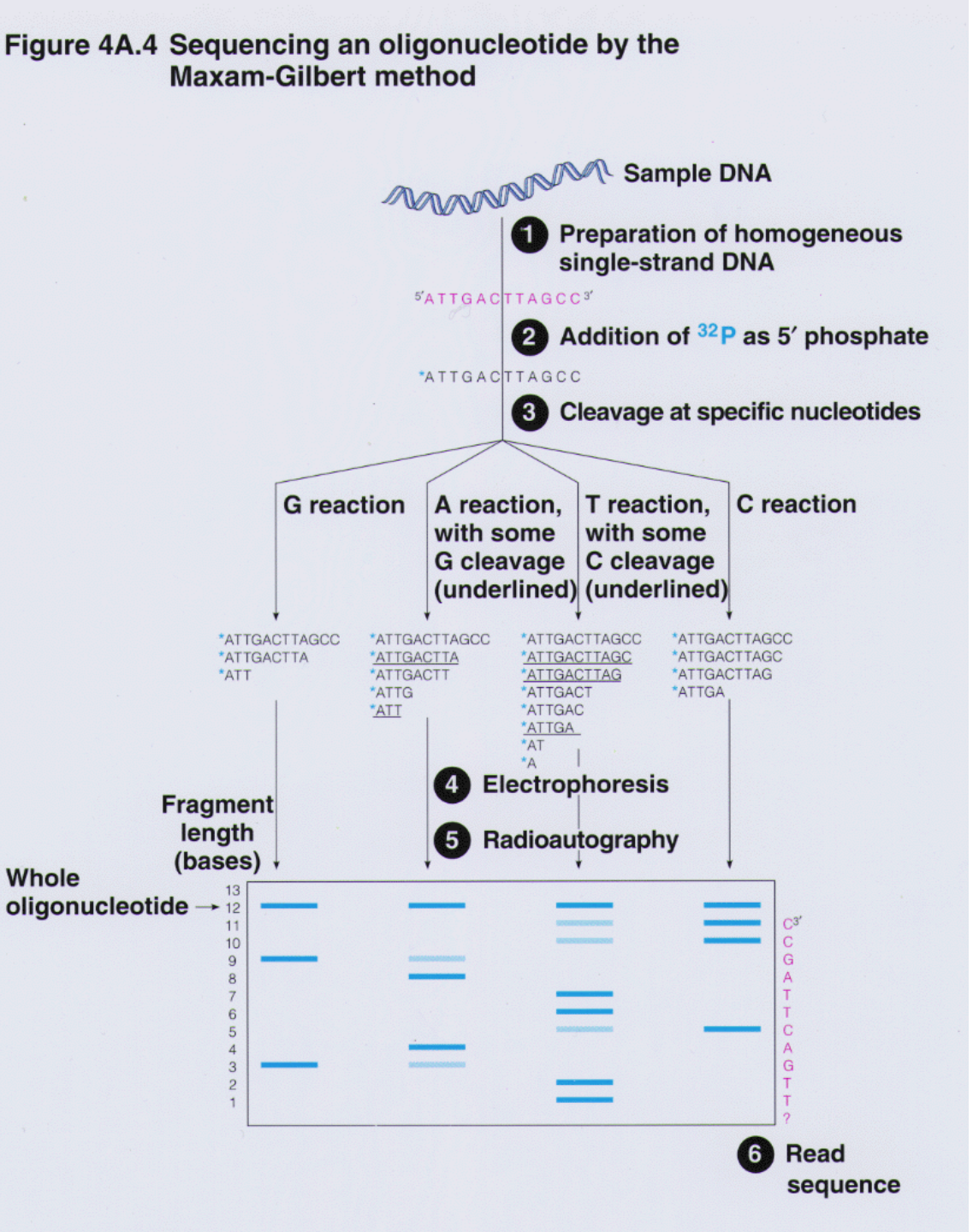


DNA가 합성되려면 다음과 같은 다섯 가지의 요소가 필요하다.

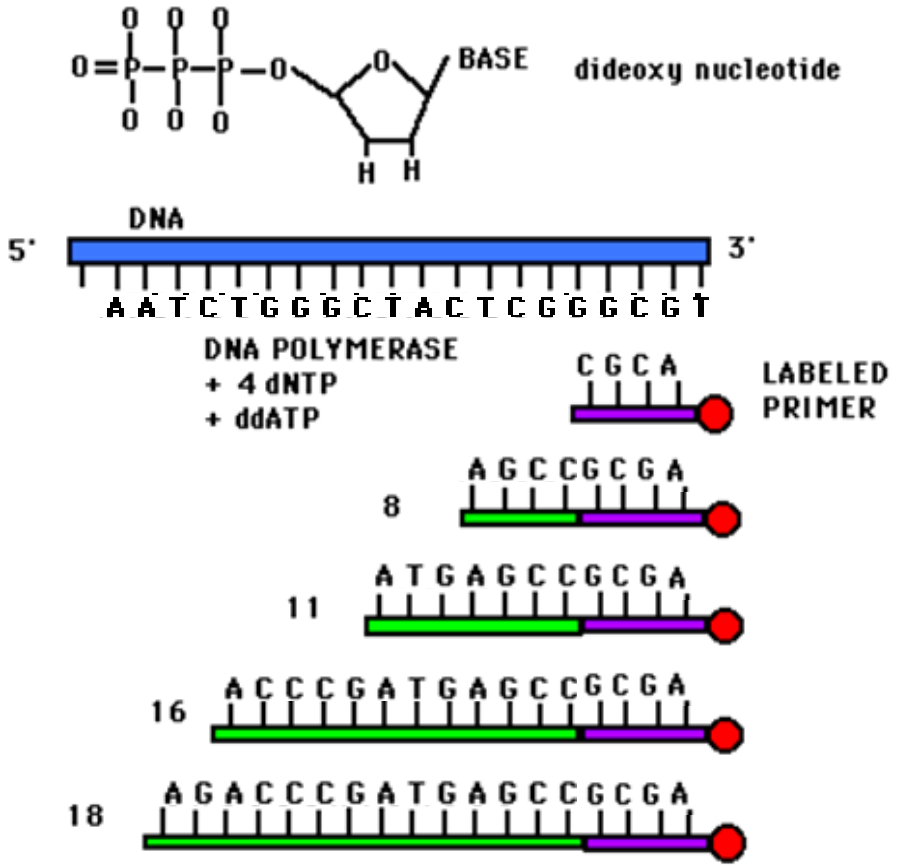
이중 나선이 풀어진 1) **DNA의 단일 사슬**은 DNA 합성을 위한 주형이 된다. 18개 정도의 단일 사슬 염기로 이루어진 2) **프라이머**가 주형 DNA의 특정부위에 상보적으로 부착하여 이중사슬을 형성하면 3) **DNA 중합효소**가 이를 인식하여 붙게 된다. 이후 반응이 일어나기 위해서는 일정농도 이상의 4) **dATP, dCTP, dGTP, dTTP**의 pool이 존재하여야 하고, 조효소로서 5) **Mg^{2+}**가 필요하다.

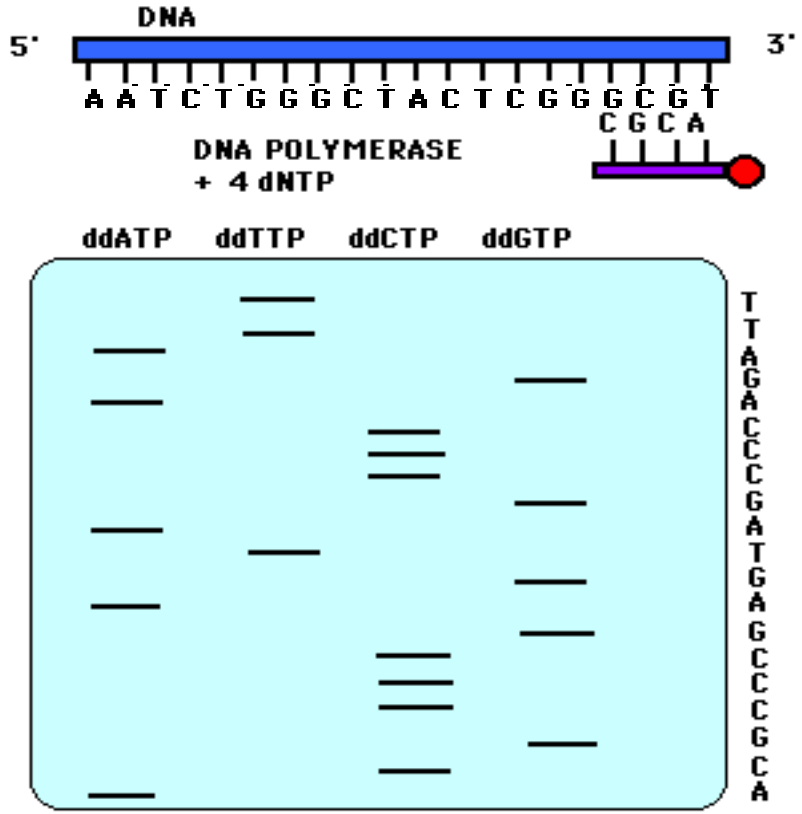
Maxam-Gilbert method

최초의 염기서열 결정 방법으로 여러가지 조건의 용액으로 염기서열을 "partial cleavage" 하는 것을 기본 원리로 한다. 0.1N NaOH를 5초간 처리하면 전체 염기서열에서 특정 염기 (예를 들어 A와 G)를 한 개 또는 두 개 정도 자르는 역할을 함. 다른 조건들은 G 만을, T와 C를, C만을 자르게 디자인 되어 있다.



Sanger method



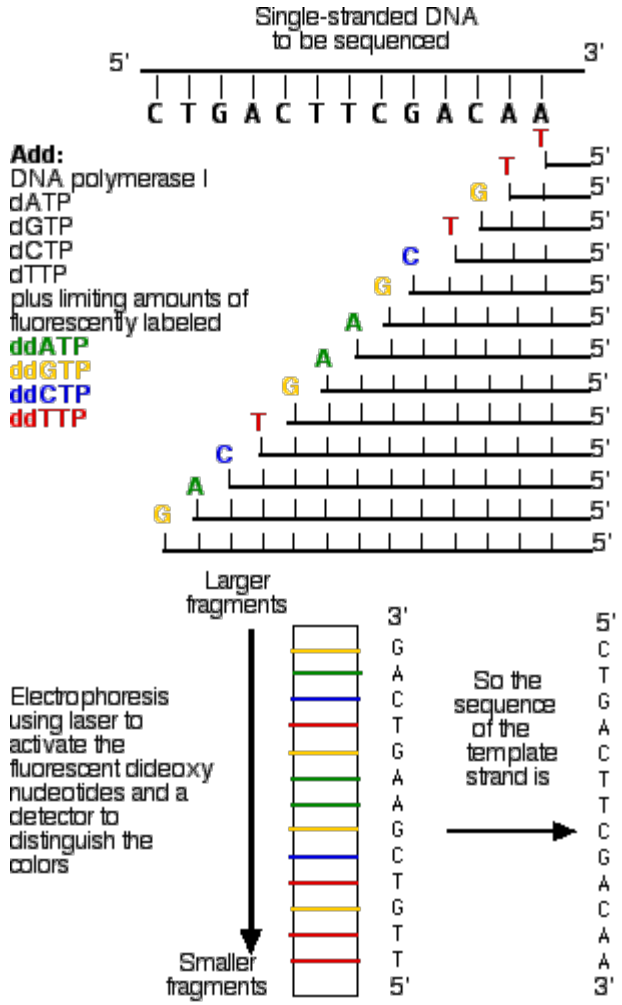


Sanger method의 특징은 ddNTP를 이용하여 polymerization을 "termination" 시키는 것이다!

One-dye (or isotope) four-lane system

List of developed florescence dye

FLUOROPHORE	ALTERNATE DYES	DYE-5'-T_{10} EX	EM	RECOMMENDED QUENCHER	BHQ Dye (QUENCHING RANGE)
Biosearch Blue™		352	447	BHQ-1	BHQ-0 430-520 nm
FAM		495	520	BHQ-1	
TET		521	536	BHQ-1	
CAL Fluor® Gold 540	VIC/TET/JOE	522	544	BHQ-1	
JOE		529	555	BHQ-1	
VIC		538	554		
HEX		535	556	BHQ-1	BHQ-1 480-580 nm
CAL Fluor Orange 560	VIC/HEX/JOE	538	559	BHQ-1	
Quasar® 570	CY3	548	566	BHQ-2	
Cy™ 3		549	566		
NED		546	575		
TAMRA		557	583	BHQ-2	
CAL Fluor Red 590	TAMRA	569	591	BHQ-2	BHQ-2* 559-670 nm
Cy 3.5		581	596		
ROX		586	610	BHQ-2	
CAL Fluor Red 610	TEXAS RED/ROX/ ALEXA FLUOR® 594	590	610	BHQ-2	
Texas Red®		597	616		
CAL Fluor Red 635	LC RED® 640	618	637	BHQ-2	
Pulsar® 650		460	650	BHQ-2	
Cy 5		646	669		
Quasar 670	CY5	647	670	BHQ-2*, BHQ-3	
Cy 5.5		675	694		
Quasar 705	CY5.5	690	705	BHQ-2*, BHQ-3	BHQ-3 620-730 nm

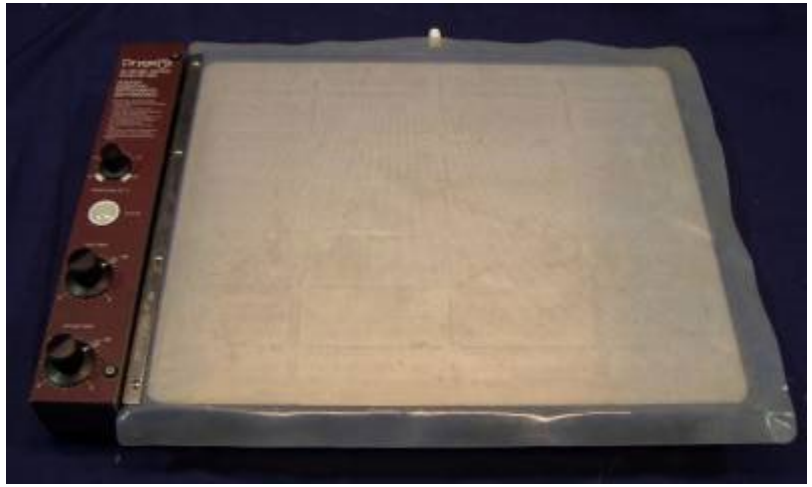


Four-dye one-lane system

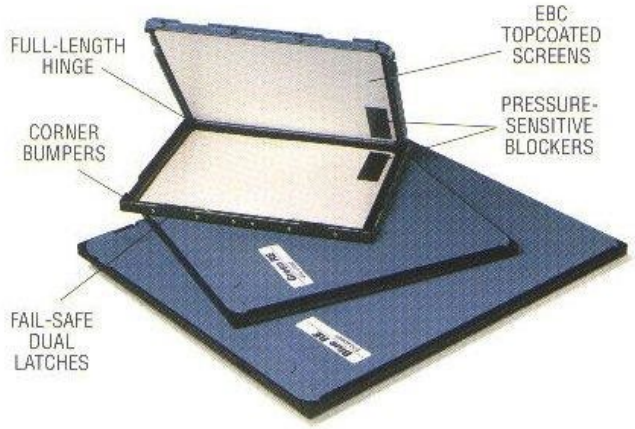
~1990년대 말

Radioisotope + gel type manual Sanger sequencing

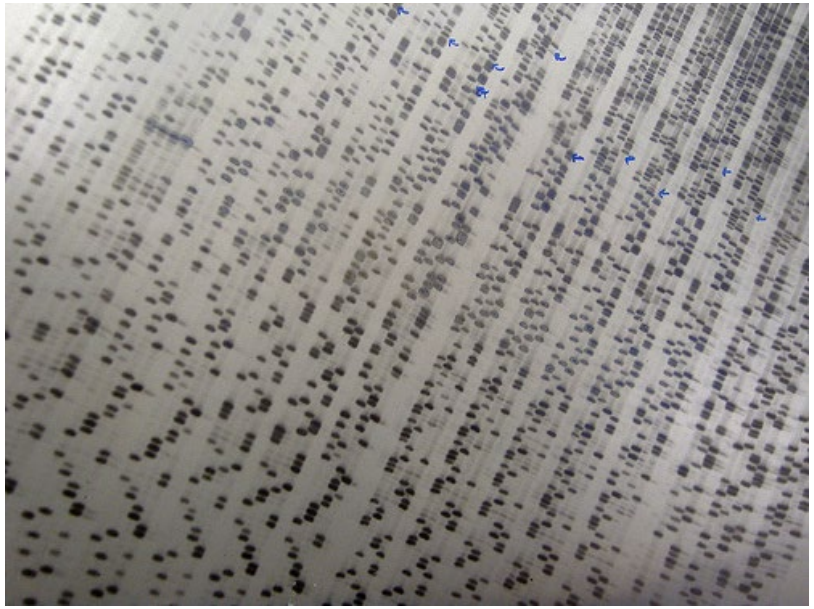
Vertical electrophoresis kit



Gel dryer



Intensifying screen

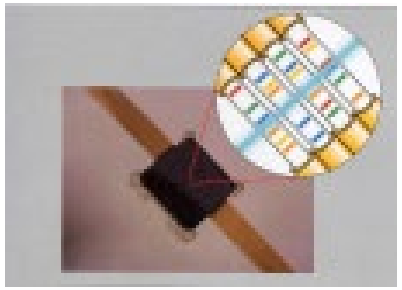


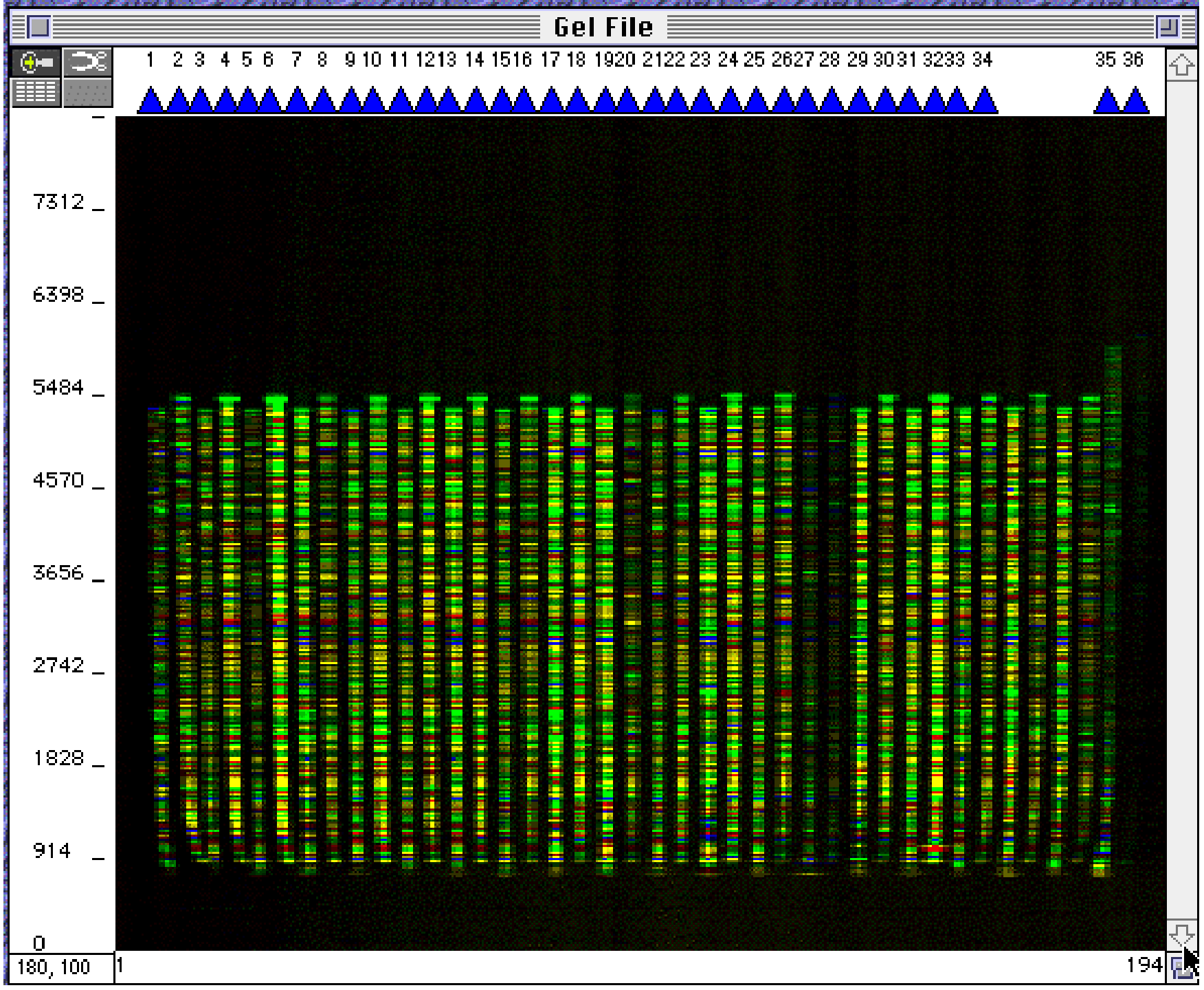
1990년대 중반~약 10년

Gel-type Automatic sequencer

- One lane four dye 의 florescent dye에 의한 gel running의 detection 방법을 자동화 한 것

ABI 377




Gel File
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34
35 36
7312
6398
5484
4570
3656
2742
1828
914
0
180, 100
1
194

2000년대 중반~현재

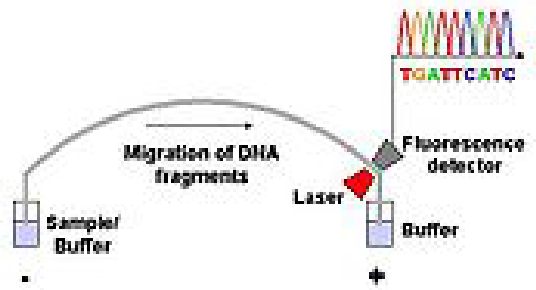
Capillary-type Automatic sequencer

- Gel 이 아닌 capillary를 이용한 전기영동 방법으로 정확도를 증가시키고, running 시간을 획기적으로 줄였다. 전체 기기의 염기서열 결정 용량은 capillary 수에 따라 다르다. 주로 96 well plate를 이용하여 96개 단위로 running이 이루어지며, 현재는 대부분 각각의 well을 4등분하여 96X4=384 well plate를 이용한다.

ABI 3730:

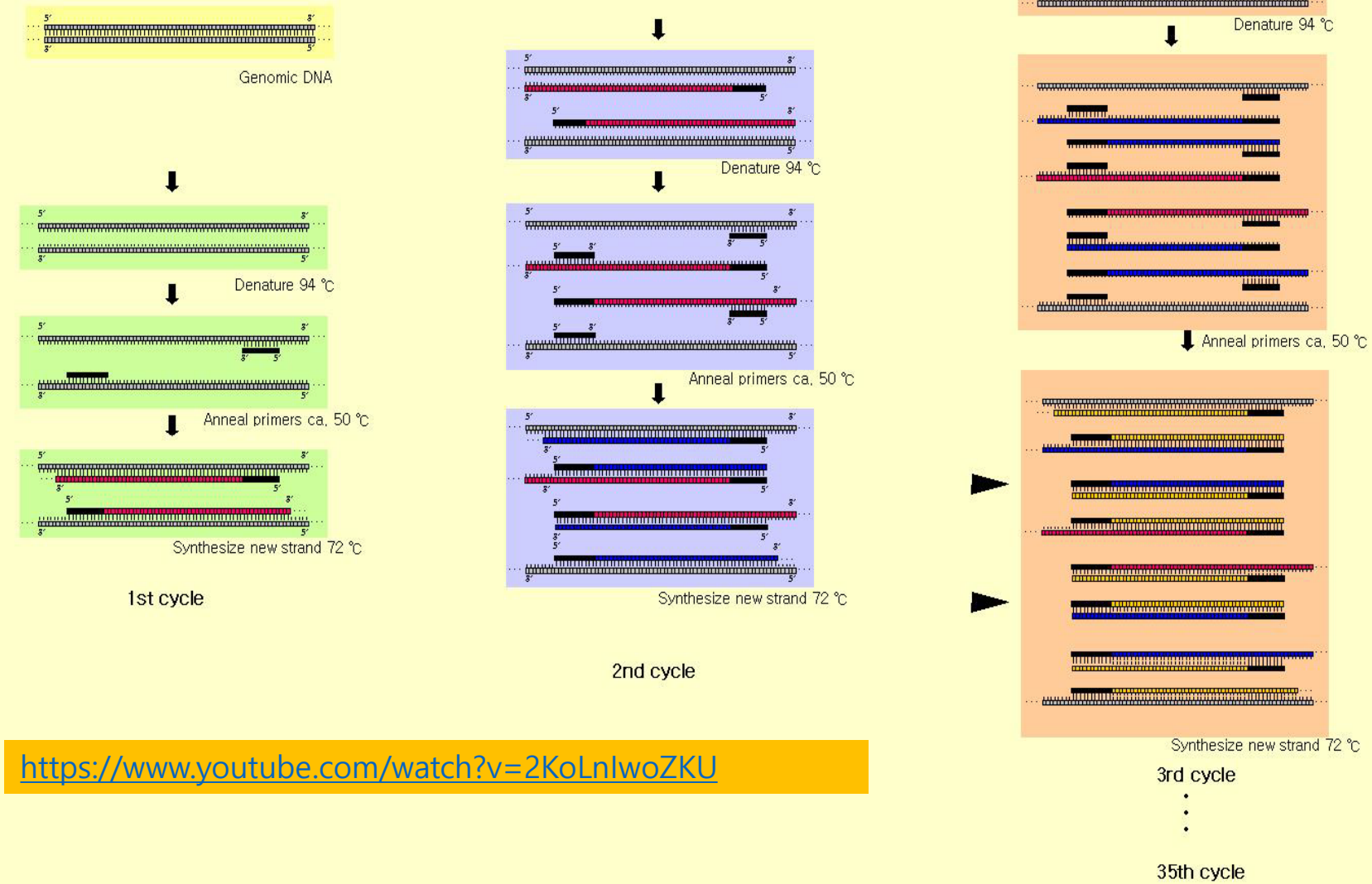
ABI 3100

- ABI (Applied Biosystem) 3730기기는 현재도 Sanger sequencing 서비스에 사용되고 있다. 한 개의 반응에서 얻어질 수 있는 sequence는 현재 약 900bp 로서 384 plate를 이용한 한번의 반응에 약 350kbp를 얻게 된다.



※ 참고
일반생물학

Schematic diagram of PCR



https://www.youtube.com/watch?v=2KoLnIwoZKU

Next Generation Sequencing (차세대 염기서열 결정)
1) The second generation sequencing:

1) emulsion based clonal amplification (emPCR)의 기술,

2) DNA 분자가 합성될 때 형광을 발하는 염기서열 결정기술 (pyrosequencing)

3) 광섬유들을 평행하게 붙여 만든 pico-titer plate

등의 신기술을 이용하여 염기서열 결정 용량은 획기적으로 증가.

대표적 기업 / 기술:

Roche (454) / 454 (최초의 NGS; 2005)

Solexa / Illumina → MGI (Illumina 유사 기술에 의한 중국 기업/제품)

ABI / SOLiD

Hilicos / Ion Torrent

→ 대부분 없어지고, 현재 **Illumina (and MGI)**가 가장 널리 쓰임.

Next Generation Sequencing 기기들

GS-Titanium; Roche 454

SOLiD; ABI

Solexa; Illumina

Helicos; Helicos Bioscience

nature Vol 437|15 September 2005|doi:10.1038/nature03959

ARTICLES

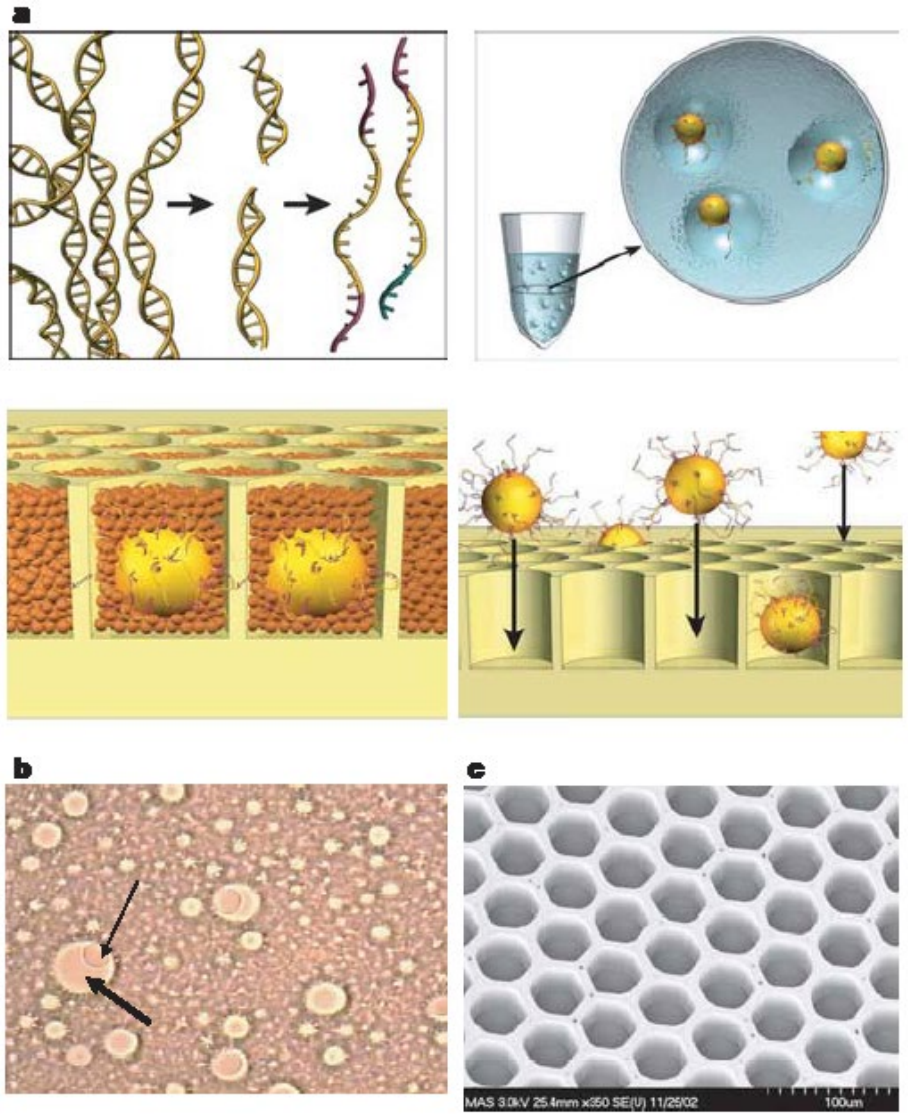
Nature 지에 출판된 최초의 NGS 시스템(454)

Genome sequencing in microfabricated high-density picolitre reactors

Marcel Margulies[1]*, Michael Egholm[1]*, William E. Altman[1], Said Attiya[1], Joel S. Bader[1], Lisa A. Bemben[1], Jan Berka[1], Michael S. Braverman[1], Yi-Ju Chen[1], Zhoutao Chen[1], Scott B. Dewell[1], Lei Du[1], Joseph M. Fierro[1], Xavier V. Gomes[1], Brian C. Godwin[1], Wen He[1], Scott Helgesen[1], Chun He Ho[1], Gerard P. Irzyk[1], Szilveszter C. Jando[1], Maria L. I. Alenquer[1], Thomas P. Jarvie[1], Kshama B. Jirage[1], Jong-Bum Kim[1], James R. Knight[1], Janna R. Lanza[1], John H. Leamon[1], Steven M. Lefkowitz[1], Ming Lei[1], Jing Li[1], Kenton L. Lohman[1], Hong Lu[1], Vinod B. Makhijani[1], Keith E. McDade[1], Michael P. McKenna[1], Eugene W. Myers[2], Elizabeth Nickerson[1], John R. Nobile[1], Ramona Plant[1], Bernard P. Puc[1], Michael T. Ronan[1], George T. Roth[1], Gary J. Sarkis[1], Jan Fredrik Simons[1], John W. Simpson[1], Maithreyan Srinivasan[1], Karrie R. Tartaro[1], Alexander Tomasz[3], Kari A. Vogt[1], Greg A. Volkmer[1], Shally H. Wang[1], Yong Wang[1], Michael P. Weiner[4], Pengguang Yu[1], Richard F. Begley[1] & Jonathan M. Rothberg[1]

The proliferation of large-scale DNA-sequencing projects in recent years has driven a search for alternative methods to reduce time and cost. Here we describe a scalable, highly parallel sequencing system with raw throughput significantly greater than that of state-of-the-art capillary electrophoresis instruments. The apparatus uses a novel fibre-optic slide of individual wells and is able to sequence 25 million bases, at 99% or better accuracy, in one four-hour run. To achieve an approximately 100-fold increase in throughput over current Sanger sequencing technology, we have developed an emulsion method for DNA amplification and an instrument for sequencing by synthesis using a pyrosequencing protocol optimized for solid support and picolitre-scale volumes. Here we show the utility, throughput, accuracy and robustness of this system by shotgun sequencing and *de novo* assembly of the *Mycoplasma genitalium* genome with 96% coverage at 99.96% accuracy in one run of the machine.

454 Technology



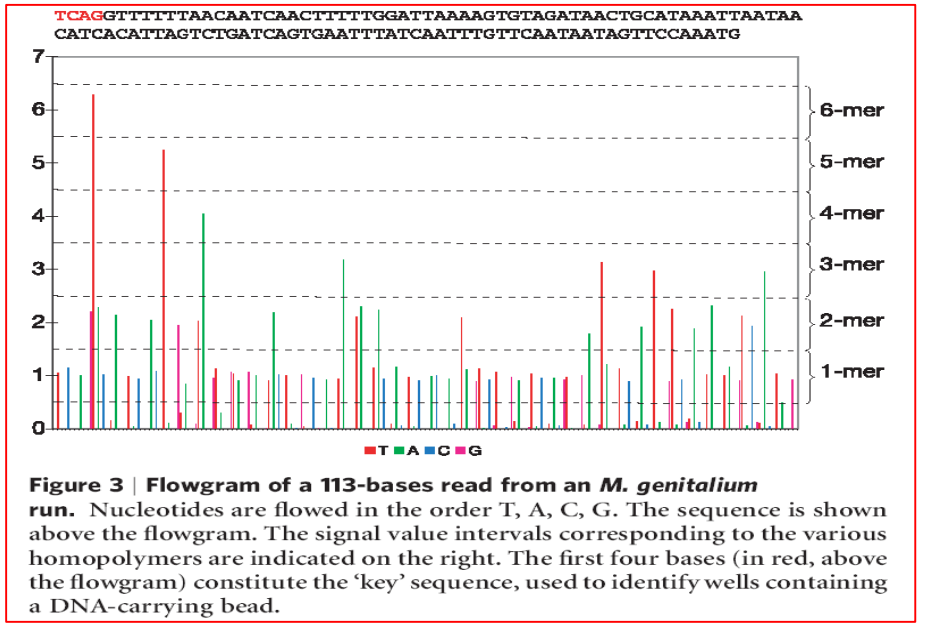


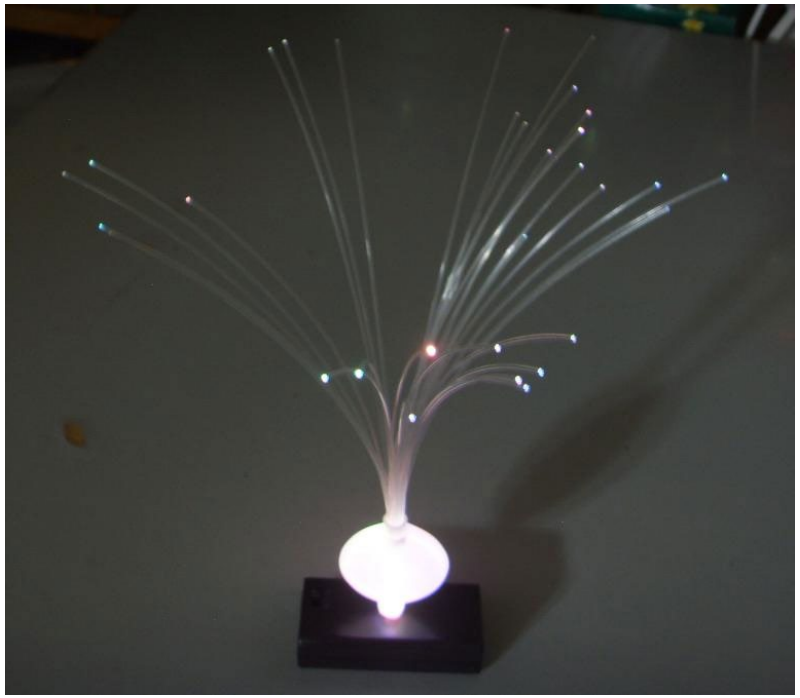
Figure 3 | Flowgram of a 113-bases read from an *M. genitalium* run. Nucleotides are flowed in the order T, A, C, G. The sequence is shown above the flowgram. The signal value intervals corresponding to the various homopolymers are indicated on the right. The first four bases (in red, above the flowgram) constitute the 'key' sequence, used to identify wells containing a DNA-carrying bead.

- 454 technique

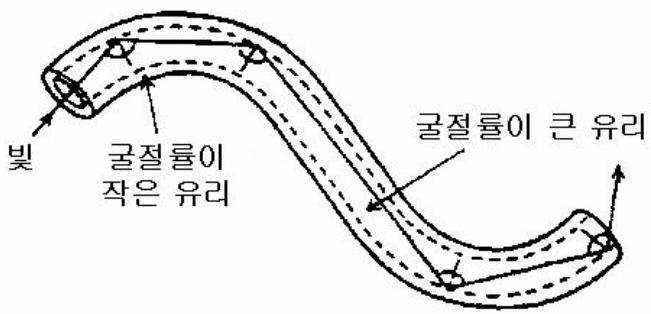
단점: 동일한염기서열이 길게 반복될 때 (polyN) 반복 수를 정확히 판단하기 어려워 에러를 발생시킬 수 있다.

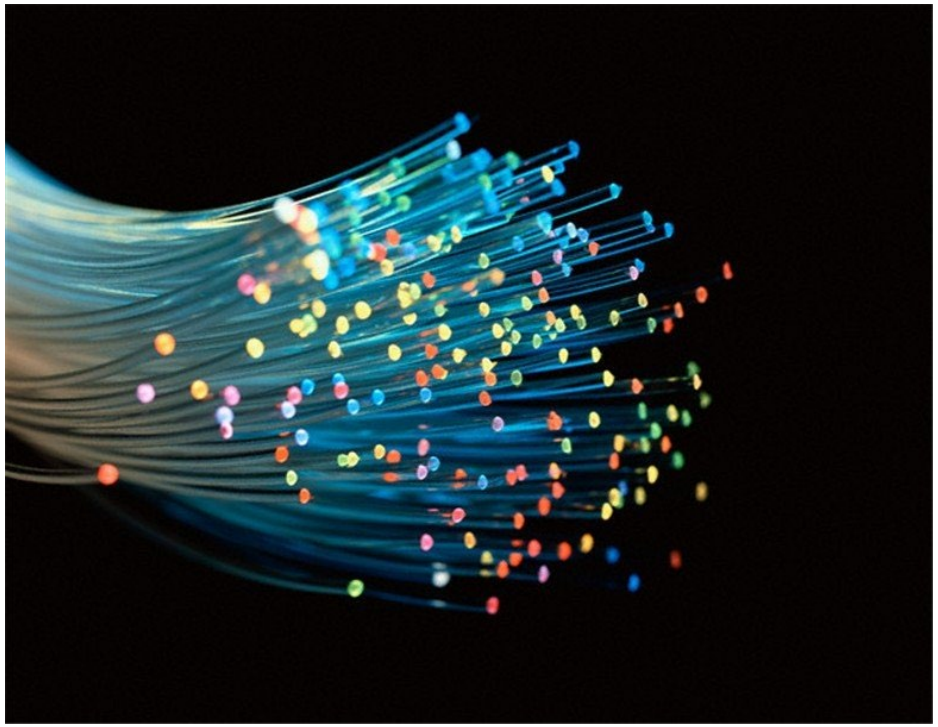
장점: 경쟁 기술인 Solexa/Illumina 기술에 비해 한번에 읽어 낼 수 있는 sequence의 길이가 길다(현재 약 450bp 정도 임)

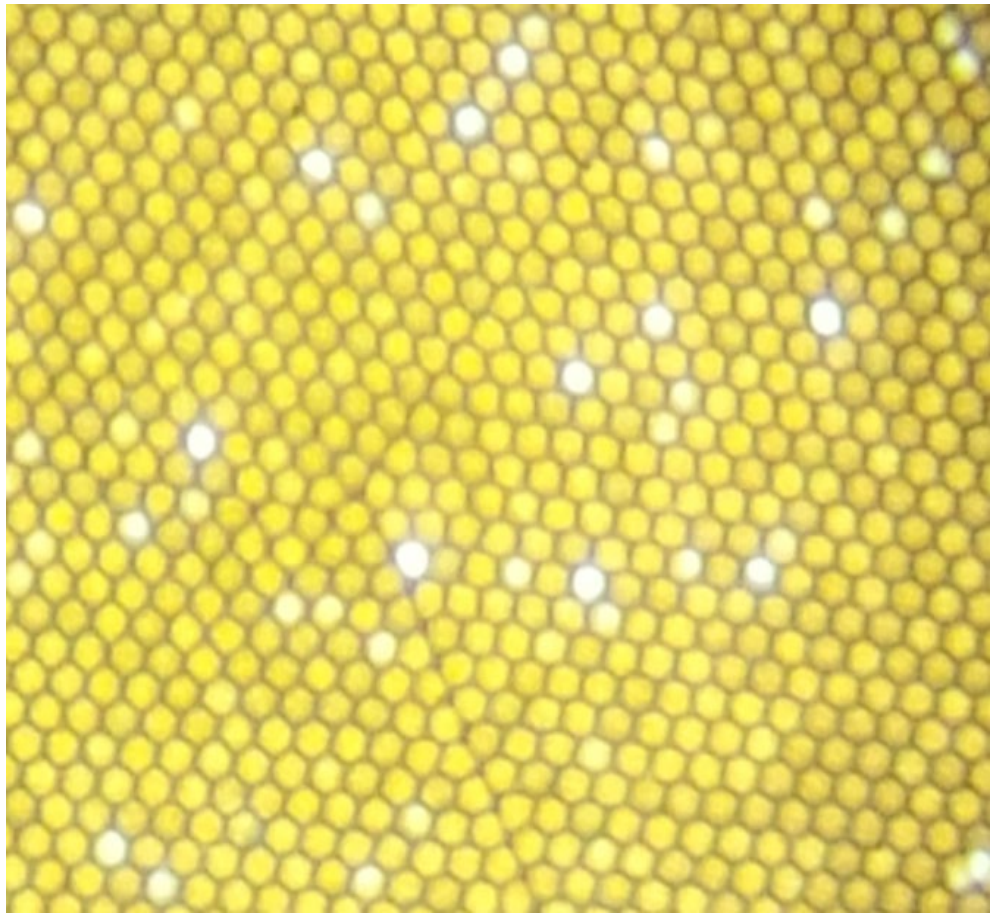
454 기술은 2016년 현재 이미 더 이상 기술 개발이 없이 단종된 기술이 되고 말았다.



Micro-titer plate의 원료:
Optic fibers (광섬유). 매우 작은 well을 만들어내어 한 개의 bead가 들어갈 수 있어 bead의 위치를 고정시킨다.



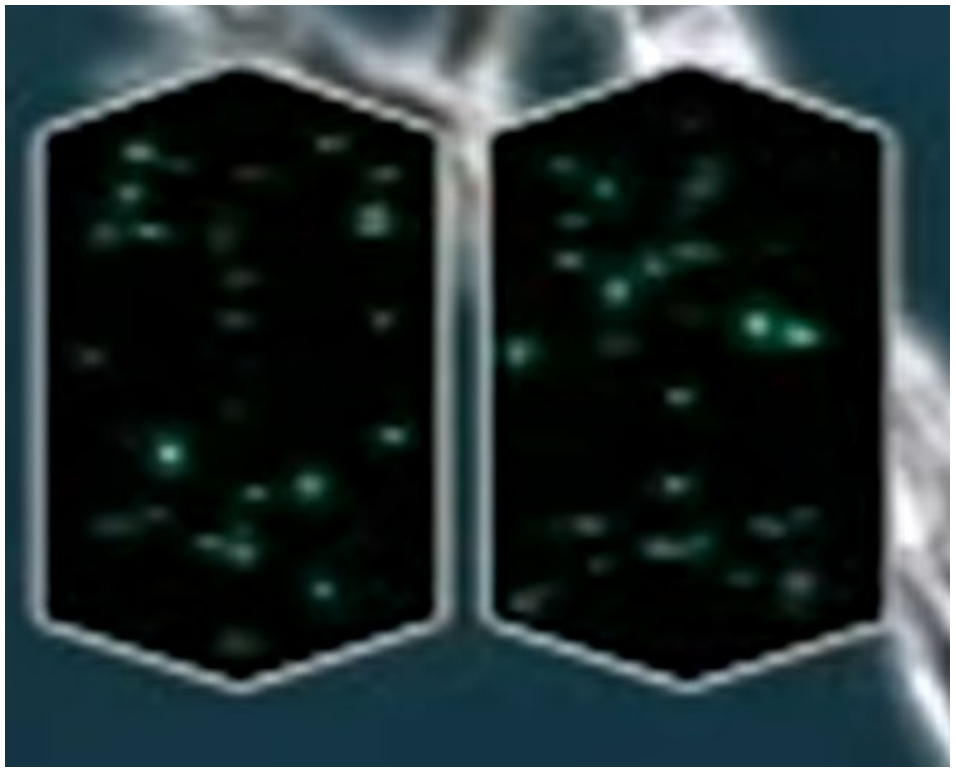




Pico-titer plate 의 현미경 사진(80배).

육각형의 격자에 PCR 반응이 일어난 구슬들이 각각 들어간다.

- CCD camera가 잡은 454 system의 발광 사진

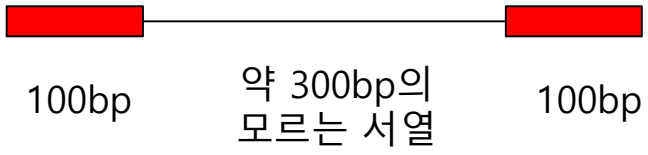


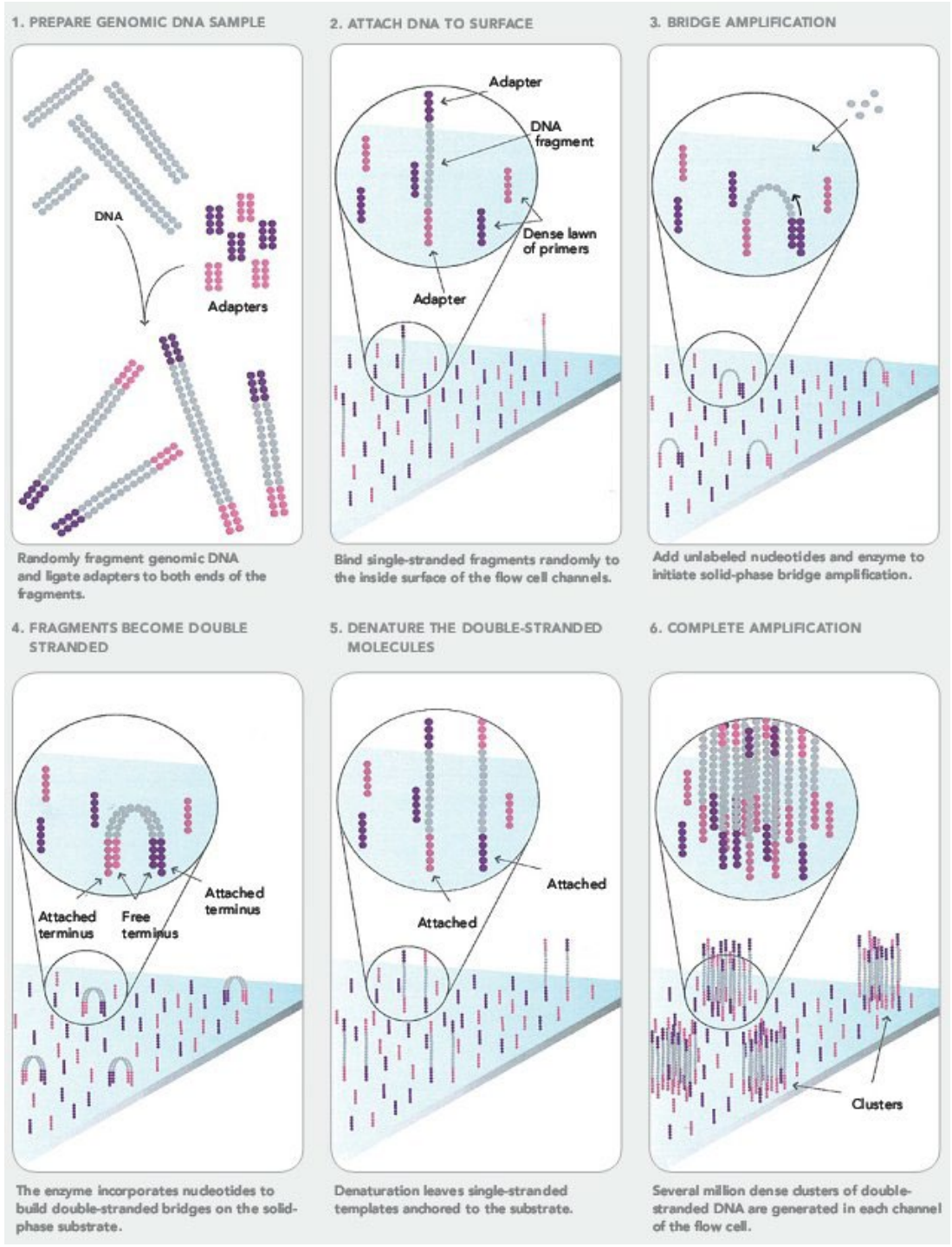
Solexa / Illumina Technology

Illumina를 이용한 결과의 특징:

- 한 가닥의 DNA로 부터 염기서열을 결정할 때 순방향과 역방향으로 각각 약 100~150 bp 정도 읽게 됨.
- 그러므로 DNA를 일정크기로 잘라 만든 조각이 500bp일 경우 좌우로 100bp씩을 얻게 되고, 중간의 300bp는 모르는 서열로 연결되게 됨.

Illumina read의 예



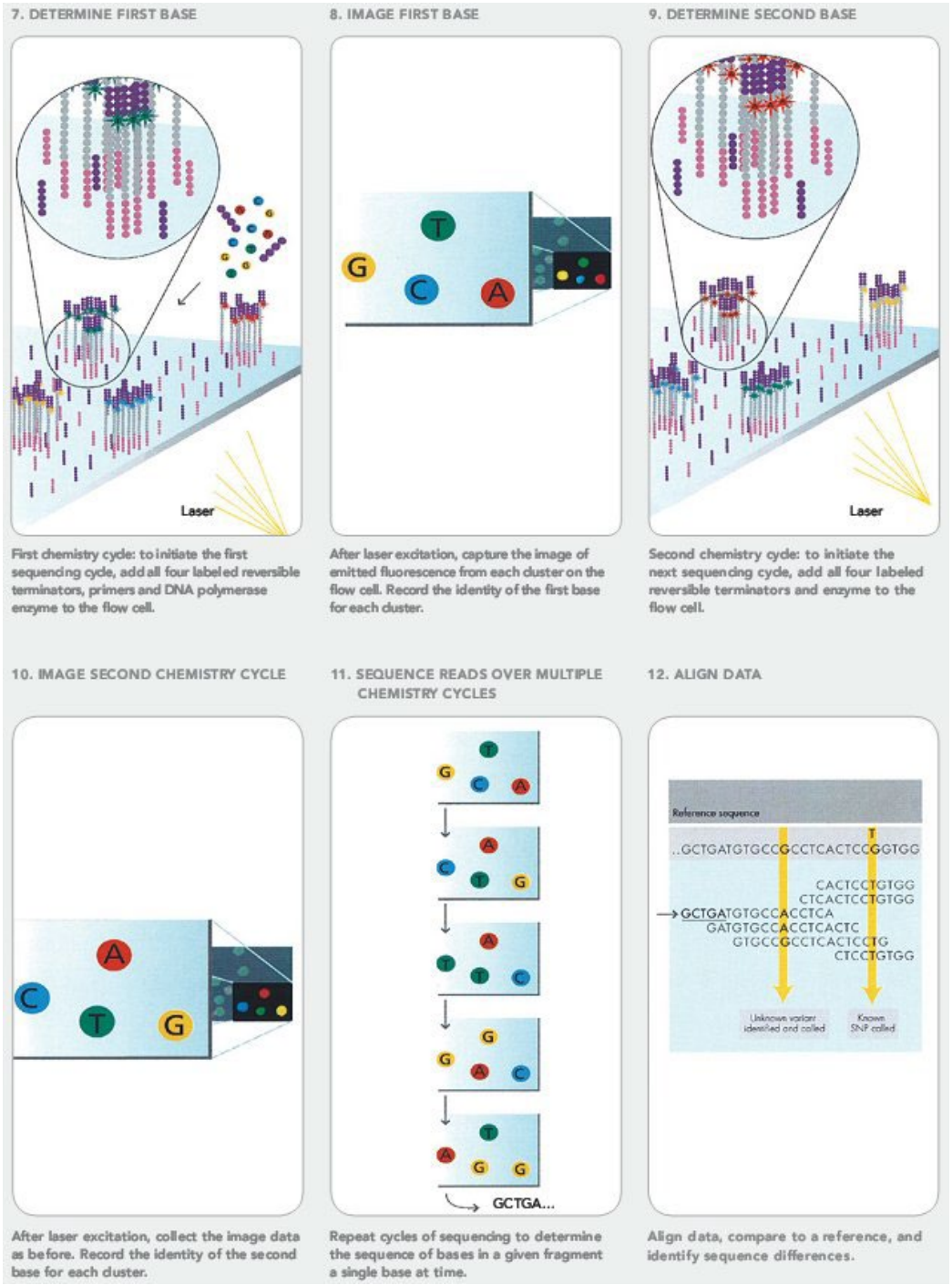


Solexa / Illumina Technology

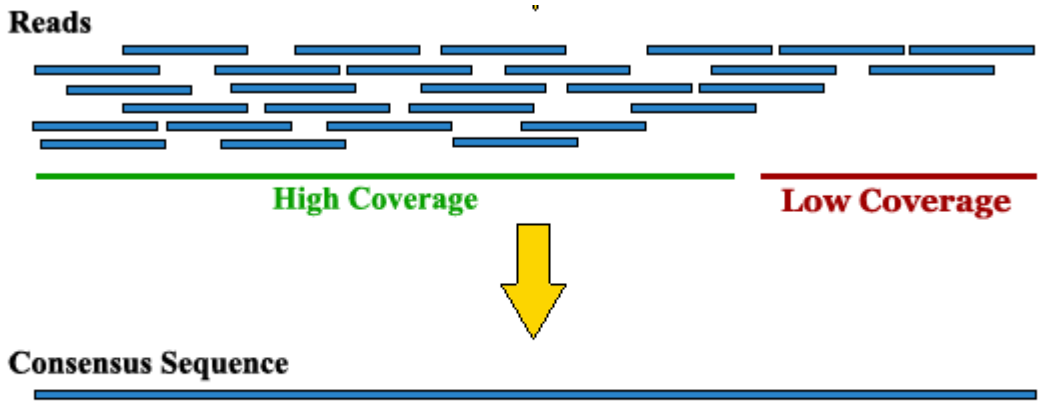
https://www.youtube.com/watch?v=fCd6B5HRaZ8

※ Illumina 기술에 의한 염기서열 결정은 "매우" 많은 염기서열을 한번에 얻을 수 있음. 그러므로 종종 많은 시료를 섞어서 염기서열을 결정하기도 한다. 이때 섞은 시료들을 구분하기 위하여 시료 각각을 구분하는 index sequence 를 각각의 시료에 붙임. 전체 염기서열 결정 후 index sequence로 시료들은 구분한 후 각각 정렬하여 결과를 얻게 됨.

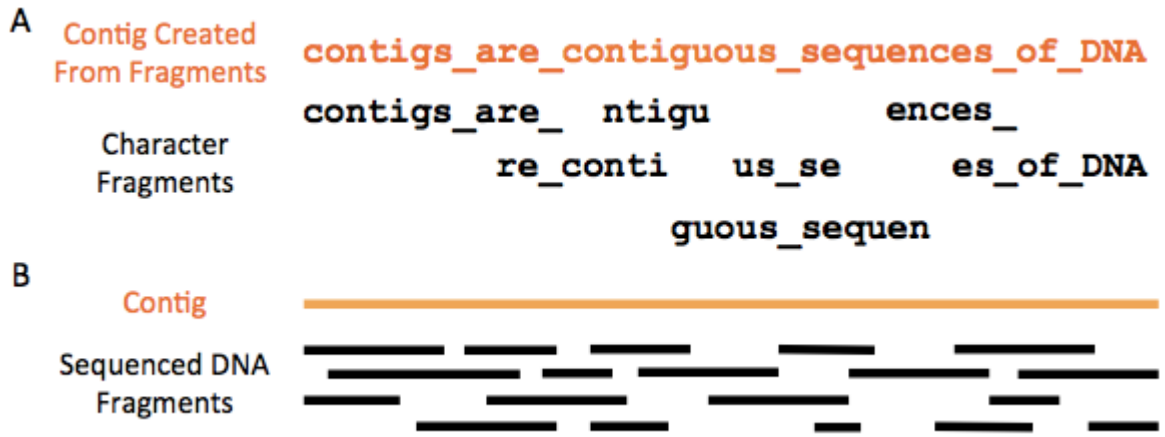
※ **index sequence:** 여러 시료를 섞어 실험할 때 각각의 시료를 구분할 수 있는 짧은 염기서열. 실험의 첫 단계인 adaptor를 붙이는 과정에 index sequence를 삽입한다.



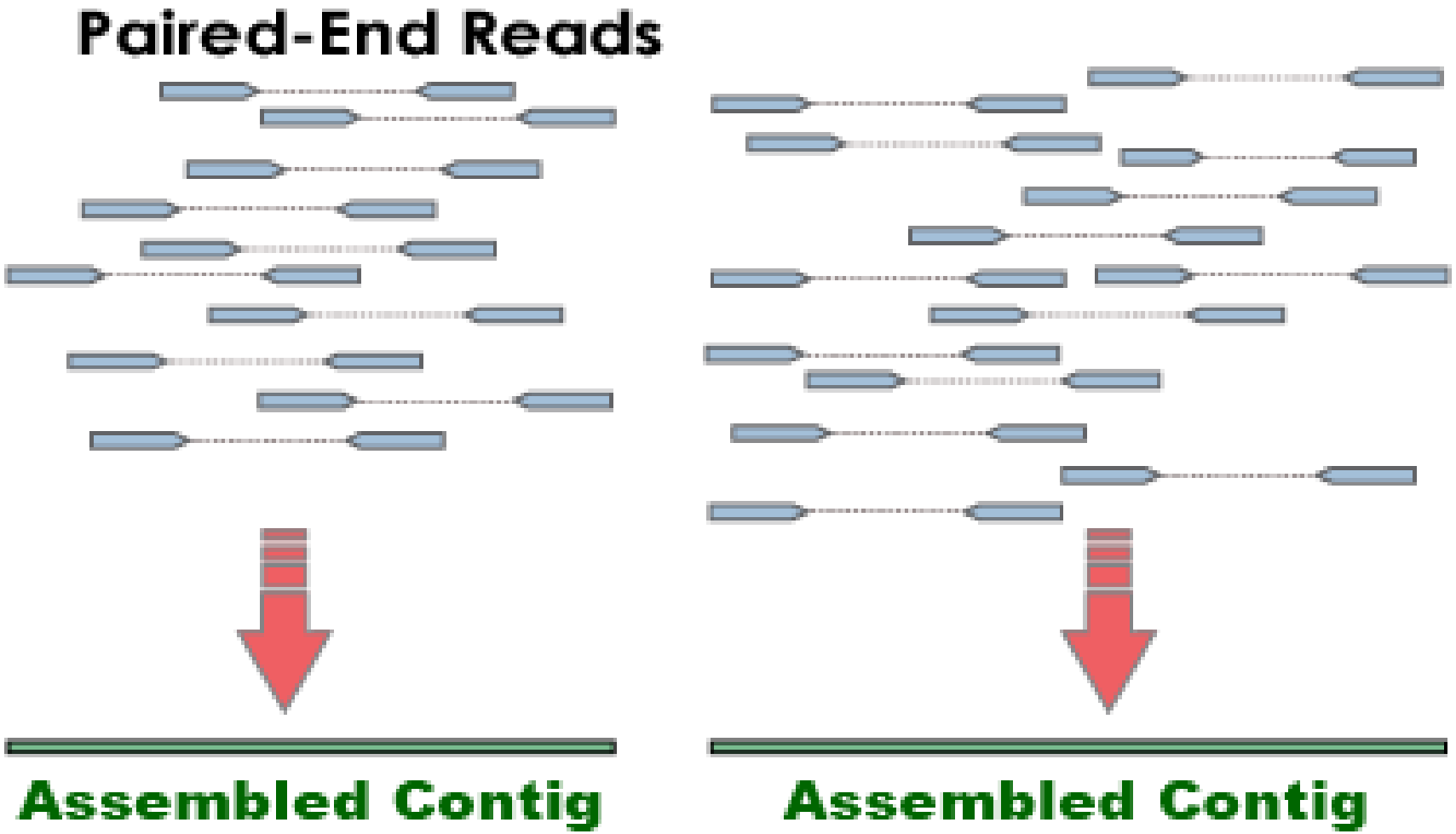
많은 "짧은" read 들로서 하나의 consensus sequence를 만들어내는 과정



Contig: 결정된 짧은 염기서열들을 조합하여 만들어낸 긴 염기서열
Coverage: contig의 각 부분에 original read 들이 얼마나 많이 중복적으로 기여했는지를 나타냄



Illumina data 의 assemble 과정



- Illumina sequence는 최근까지 유전체 결정에 있어서 주된 data로 사용되어 옴.
- Macrogen Co. service:

  HiSeq X ten, HiSeq4000, NextSeq

- 진정한 인간유전체 결정 $1,000 시대

- 최근에는 유사한 원리이지만 특허를 우회한 MGI seq 또한 널리 보급됨.

HiSeq X System Performance Parameters	
Key Application	Large Whole-Genome Sequencing (human, plant, animal)
Output per Run	Dual flow cell: 1.6-1.8 Tb
Single Reads Passing Filter	Dual flow cell: 5.3-6 billion
Maximum Read Length	2 x 150 bp
Run Time	< 3 days
Quality	≥75% of bases above Q30 at 2 x 150 bp

하지만… 무엇이 문제인가?

→ 100~150 bp의 너무 짧은(short-read) 서열을 제공하고 있기 때문에, 이것을 이어 붙여 유전체를 조립하면 구조적인 에러를 발생할 확률이 높아짐.

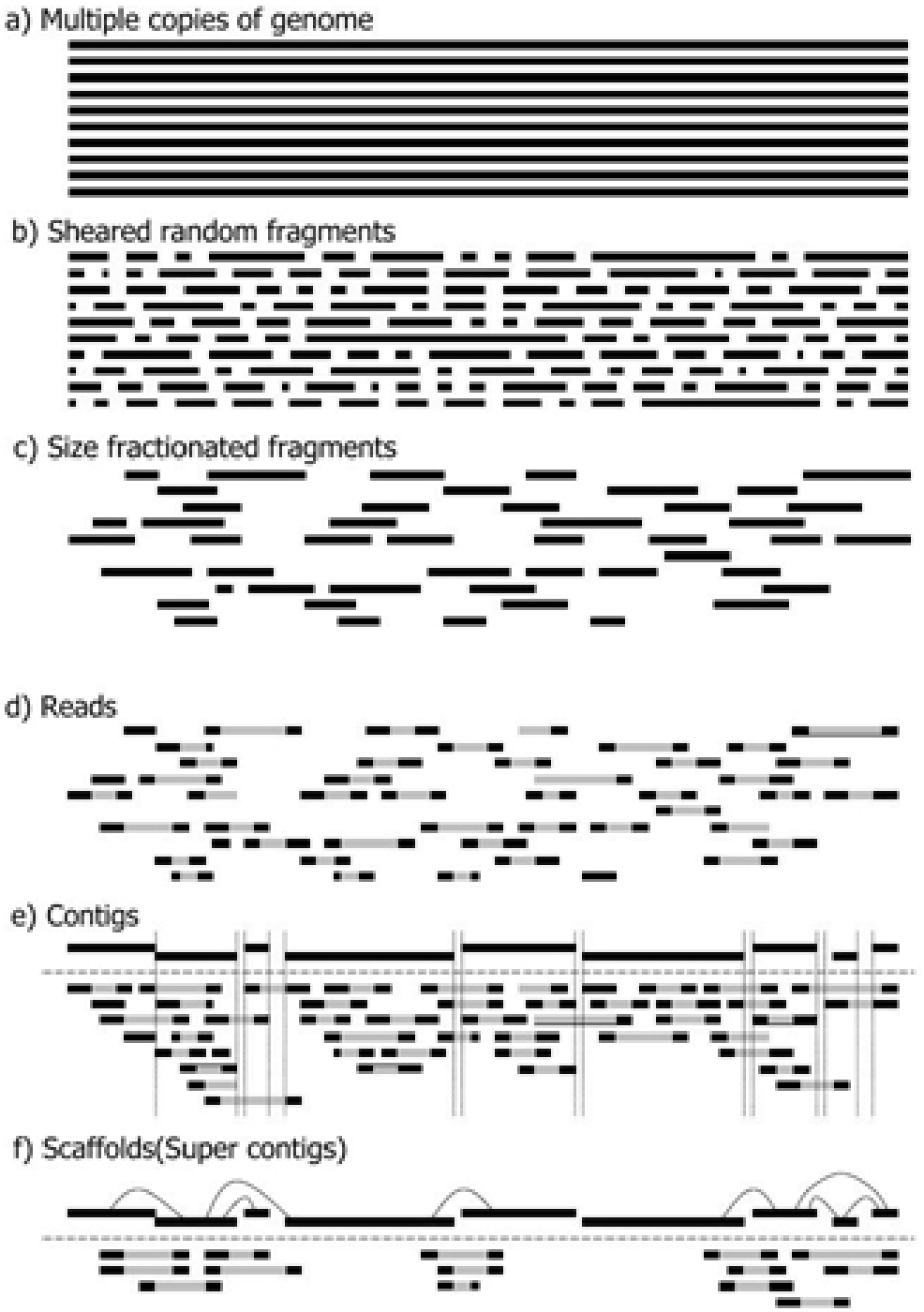
1) Mate-pair sequencing에 의한 해결

2) **Long-read data**를 생산하는 것이 궁극적 해결 방법임.

Mate-pair sequencin에 의한

Genome assembl의 과정 (Illumina)

1) 추출된 DNA를 적당한 크기로 자른다.
2) 잘린 DNA 절편들 중 약 500 bp 의 크기를 갖는 절편들만 정제한다.
3) Illumina **paired-end sequencing**에서는 잘린 절편의 양쪽 끝의 각 각 약 100bp 정도의 데이터를 얻을 수 있다(reads).
4) Assembly program을 이용하여 각각의 read들을 정렬하여 **contig** 들을 생성해 낸다.
5) 이와는 별도로 추출된 DNA를 보다 긴 길이의 절편으로 잘라 (1K, 2K, 5K...) 이들 절편의 양쪽 끝 100bp 의 염기서열들을 결정한다 (**mate-pair library**).
6) 일반 paired-end sequencing을 통해 생성된 contig들과 mate-pair 결과를 합쳐 assembly한다. 이를 통해 결과적으로 전체 유전체는 부분적으로 염기서열이 완성된 contig들이 mate-pair 에 의해 연결되어 많은 gap을 포함하지만 서로의 위치관계가 명확해진 긴 염기서열을 얻게 된다. 이를 **scaffold** 또는 **super contig**라 한다



Next Generation Sequencing (차세대 염기서열 결정)
2) The Third generation sequencing:

- 100-150 bp 길이의 short-read를 생성해 내는 2세대 염기서열결정장치와는 달리 수십 kbp 에 달하는 long-read를 생성해 내는 기술
- 대표적 기업 / 기술:

1) PacBIO / PacBioHiFi

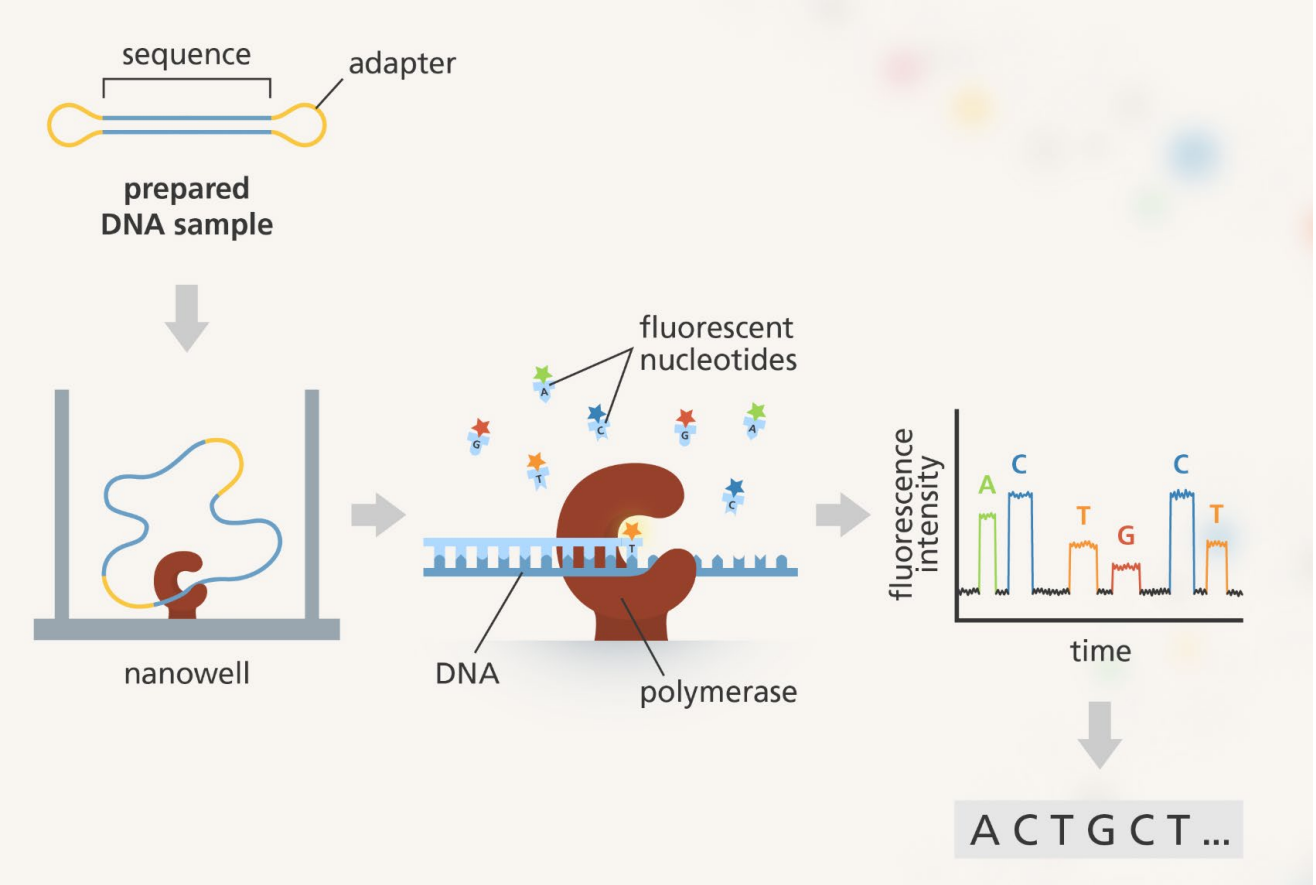
2) **Oxford Nanopore Technology (ONT)** / Nanopore

- **PacBioHiFi** 주로 10~15 kbp (up to 25 kbp)의 제한된 길이의 대용량 데이터를 생산하여 고가의 장비로 일반적으로 sequencing service 업체에 의해 운영됨.
- **Nanopore**는 연구실 단위의 실험이 가능하고, 추출된 DNA의 길이만큼 훨씬 긴 서열의 생성이 가능함.

PacBio Technology

- 대형 genome project에서 현재 가장 많이 활용됨.
- 비싸고 대용량! 장점이자 단점.
- 두 가지 모드:
  - **CLR (Continuous Long Read):**
    최대 길이로 읽을 수 있지만, 정확도 85% 수준
  - **CCS (Circular Consensus Sequencing):**
    15-20kbp reads, 정확도 >99%

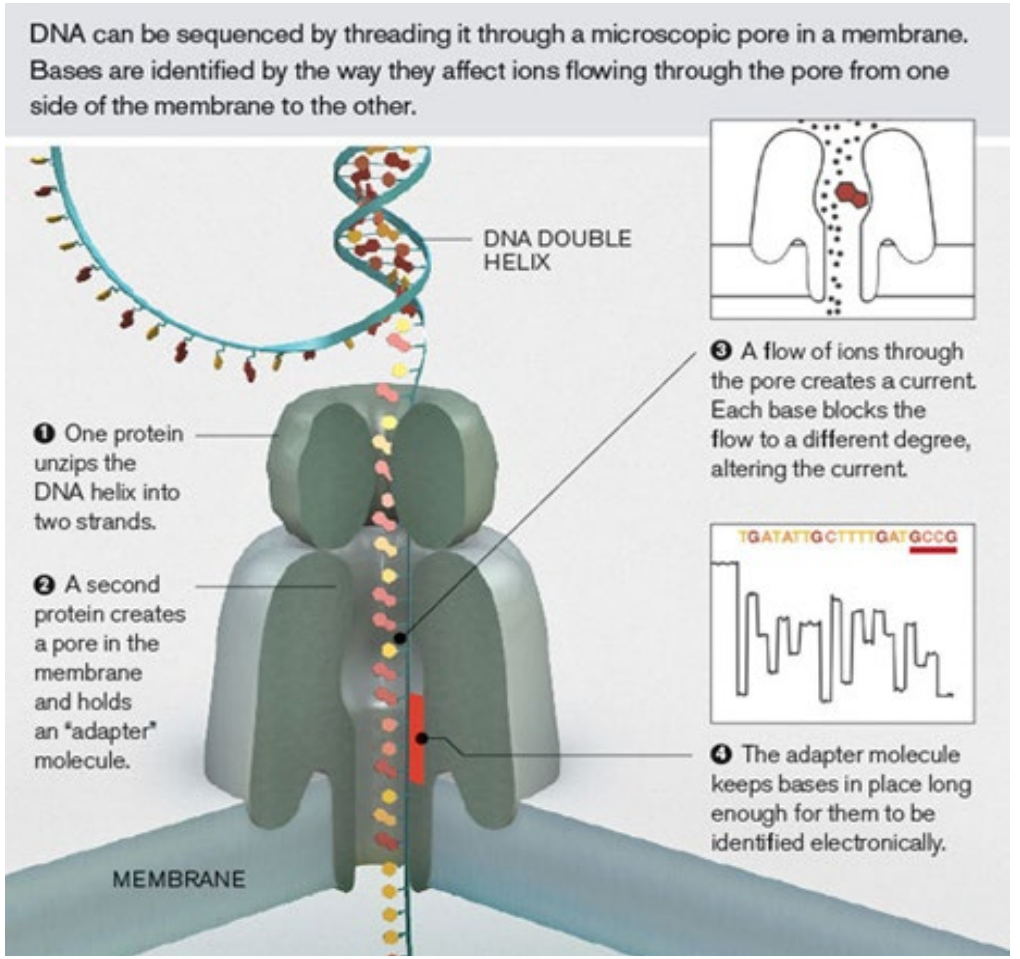
https://www.youtube.com/watch?v=_ID8JyAbwEo

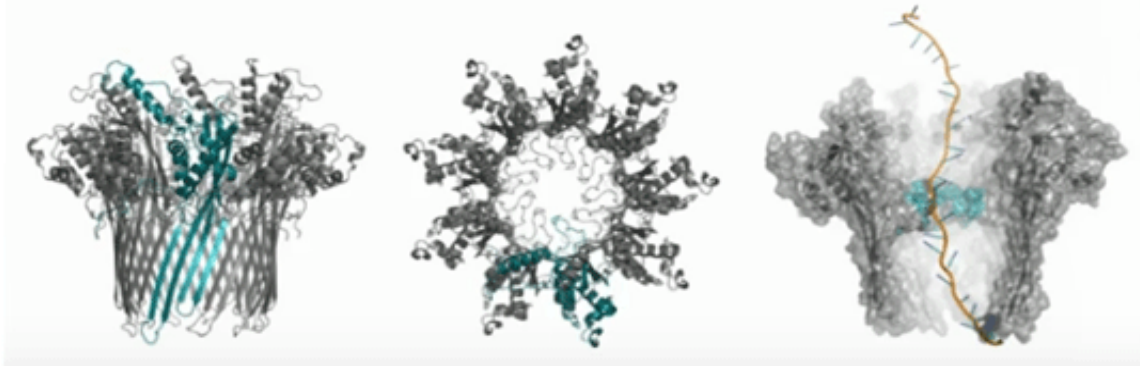


ONT / Nanopore technology

- **Oxford Nanopore Technologies** (ONT): 2005년에 University of 로부터 설립.
- "Nanopore (나노포어)"라는 막단백질을 이용한 획기적인 NGS.
- 막단백질의 pore 직경이 수 나노미터를 이루기 때문에 이런 이름이 붙여짐.
- 세포에서는 막에 위치하는 나노포어를 통해 이온들이 막 내외로 이동하는데, 이온이 지나가면서 막에 전류를 발생시키게 됨. 그런데 나노포어로 DNA나 RNA를 통과시키면, 이들이 이온의 흐름을 방해하여 여기서 발생하는 전류에도 변화가 생기게 되어 이런 막 단백질의 전류의 변화를 측정하는 것임.
- 즉, A, C, G, T 각 염기서열마다 이온의 흐름을 방해시키는 정도가 다르고 이에 의해 전류가 변화하는 정도도 달라지게 되므로, 이를 분석하여 서열을 파악할 수 있음.

https://www.youtube.com/watch?v=E9-Rm5AoZGw

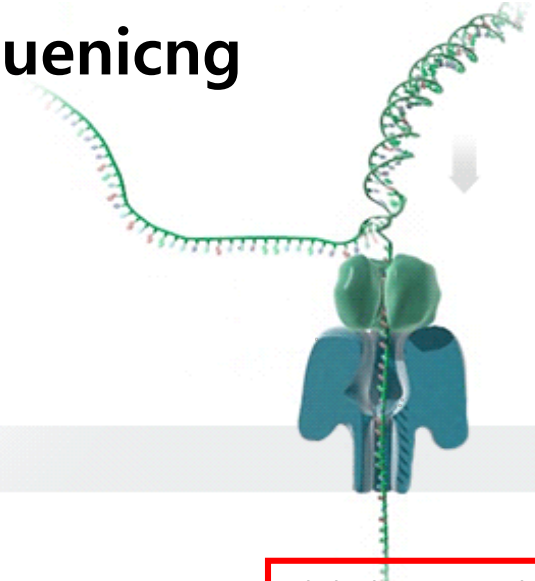


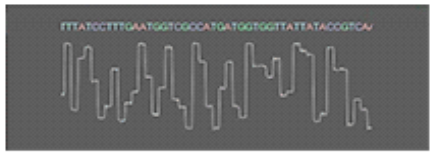


ONT 시퀀싱에 사용되는 나노포어인 CsgG의 구조

Nanopore Sequenicng

기기 모델들





가장 작은 model인 MinION은 노트북에 연결하여 현장사용도 가능하다.

MinION
Mk1C

Available to pre-order

Complete sequencing, analysis, and viewing device

Up to 30 Gb data / flow cell
512 channels*

MinION
Mk1B

Commercially available

Portable, USB powered biological analysis

Up to 30 Gb data / flow cell
512 channels*

GridION
Mk1

Commercially available

Five flow cell capacity and integrated computing

Up to 150 Gb (5 x 30 Gb) data / device
5 x 512 channels*

PromethION
P24 P48

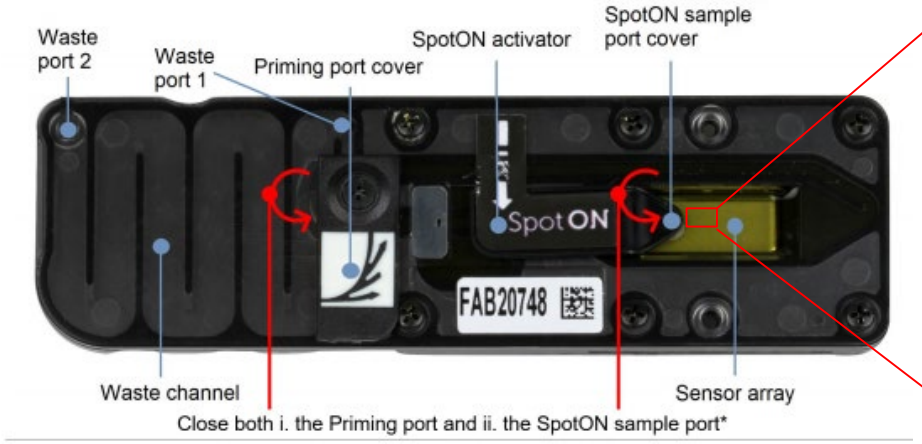
Commercially available

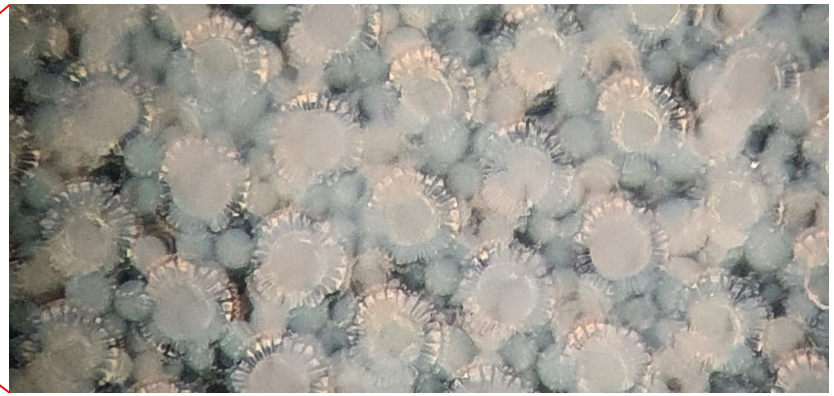
High-throughput, versatile benchtop system (P24 or P48)

P24	P48
>3.5 Tb data / device	**>7 Tb data / device**
24 x 3,000 channels*	48 x 3,000 channels*

Flongle

Adapter for MinION/GridION, supports smaller single-use flow cells.
Up to 1.8 Gb currently; towards 3 Gb.





Molecular weight
T: 126.11 g/mol
A: 135.13 g/mol
G: 151.13 g/mol
C: 190.24 g/mol

Summary: Comparison of NGS Platforms

	Platform	Amplification Method	Sequencing Method	Detection Method	Average read length	Errors
O	Illumina (HiSeq, MiSeq etc)	bridge PCR	sequencing by synthesis	Light	100-200 bp	~ 0.1 %
X	Life Tech Ion Torrent / Proton	emulsion PCR	Ion semiconductor sequencing	pH	200-400 bp	~ 1 %
X	Roche 454	emulsion PCR	Pyrosequencing, cleavage of released pyrophosphate	light	700 bp	~1 % High error rate in homopolymer
X	Life Tech SOLiD	emulsion PCR	sequencing by ligation of hybridizing labeled oligos	light	100 bp	~ 0.1 %
O	Pacific Biosciences PacBio	No amplification, single-molecule sequencing	polymerase incorporating colored NTPs	light	4.5 – 8 kb	< 1 %
O	Oxford Nanopore MinION	No amplification, single molecule nanopore sequencing	DNA molecule traverses pore	current	> 5.4 kb	< 0.2 %

Further reading, great lecture: Sequencing technology - Past, Present and Future, http://www.molgen.mpg.de/899148/OWS2013_NGS.pdf

X 현재 사장된 기술들

II. 실험 스케쥴

최종 목표: 대상 식물의 미토콘드리아 유전체 조립

<u>강의 및 실습 스케쥴</u>

8월 4일(월)

9:00~12:00 성신여대 김상태 (A235 실험실)
- DNA 추출 이론 강의

13:00~16:00 성신여대 김상태 (A235 실험실)
- High-Molecular-Weight DNA 추출 1일차

8월 5일(화)

9:00~12:00 성신여대 김상태 (A235 실험실)
- High-Molecular-Weight DNA 추출 2일차
- DNA 전기영동, Qubit에 의한 정량, 정성 분석 의뢰
- Nanopore 이론 강의

13:00-16:00 성신여대 최상철 (A235 실험실)
- Organelle DNA assembly 이론 및 실습

8월18일(월)

9:00~16:00 ONT 안종화, ㈜필코리아 한종민 (A235 실험실)
- Library작성 실습
- Nanopore sequencing running

8월 25일(월)

9:00-12:00 ONT 안종화, ㈜필코리아 한종민 (A235 실험실)
- 결과 생성 및 sequencing 결과 분석

13:00-16:00 성신여대 최상철 (B-109 PC실)
- Long-read에 의한Mitochondrial genome assembly
- Short-read에 의한 polishing

(주) 필코리아:

Oxford Nanopore Technology의

Korean distributor

- 조교: 식물분자계통학실 서정우(석박통합과정), 서지예(석사과정)
- DNA 추출 실험: 7개 조로 한 조당 1개 시료로 추출 진행.
- Nanopore sequencing:

  추출한 DNA의 QC 이후 가장 품질이 좋은 DNA 네 개의 시료를 선발하여 running.

1조 정세인 정회수 (충남대)

2조 허준현 곽도현 (카톨릭대)

3조 강현준 최윤경 (대전대학교)

4조 김윤수 서화정 (대전대학교)

5조 김이선 (성균관대학교)

6조 김우호 임지은 (한림대학교)

7조 문현지 (성신여대)

N50

- 일반적으로 유전체 수준에서 assembly가 잘 되었는지 안되었는지는 **N50** 값으로 제시함.
- Nanopore 실험으로 생성된 데이터 또한 **N50 값으로 얻은 길이에 대한 평가**가 가능.
- 전체 contig 들을 크기순으로 배열하여 큰 것으로부터 크기를 차례로 더하여 더한 값이 전체 유전체 크기의 50%를 넘는 순간의 contig의 크기를 N50라 함.
- N50 는 평균값 또는 중간값과는 다른 의미의 수치임!!!
- 3 3 4 6 7 8 8 9 9 9 10 11 13 25 의 길이의 contig들이 있을 때

Mean = 125/14= 8.93 (sum=125) Median = (3+25)/2= 14

125/2= 62.5 sum of contig lengths reach to 62.5 (from the largest to the smallest)

25+13+11+10=59

25+13+11+10+9=68

Therefore, N50 = 9

N50 계산의 실제 예

- NGS data에 의한 de novo assembly 의 결과 아래와 같이 16개의 contig들을 얻었다고 가정 할 때 아래 표는 이들을 길이 순으로 정렬한 것이다.
- 이때 전체 contig들의 합은 105,721이고, 이 길이의 50%에 해당하는 수치는 52,860.5이다. 길이가 긴 것부터 차례로 누적값을 만들어 이 누적값이 52,860.5를 넘을 때의 contig의 길이가 바로 N50이다.

Contig 번호	각각의 contig 길이	누적값		
1	33407	33407		
7	15243	48650		
6	14172	62822	52860.5	N50= 14172
8	9250	72072		
2	8275	80347		
10	5714	86061		
3	5406	91467		
13	5227	96694		
4	3683	100377		
11	2849	103226		
9	1251	104477		
14	663	105140		
5	442	105582		
12	136	105718		
16	2	105720		
15	1	105721		
합계	105721			
합계 /2	52860.5			

Experiment group

Position	Flow cell ID	Sample ID	Health	Run time	Run state	Reads	Bases	Baseca
MN35716	FAO37153	no_sample		44.8 H / 72 H	Active	2.52 M	14.9 Gb basecalled 16.32 Gb estimated	100%

되도록 긴 서열들을 얻는 것이 좋음! (N50가 높아야 함)

Scroll right

Read length histogram

SAMSUNG

그런데…

재료가 좋아야 좋은 결과가 나옴.

→ 일반 DNA 추출과정으로는 조각난 DNA가 추출됨.

→ 그러므로 되도록이면 “긴” DNA를 추출하는 것이 성공적인 Nanopore sequencing의 1차적 관건임!

※ Long DNA = **High-Molecular Weight (HMW) DNA**

III. HMW DNA 추출 이론

DNA 추출 기초지식: DNA의 추출을 위한 버퍼와 DNA의 성질

- 조직을 파쇄 하면 즉시 DNase가 방출되어 DNA를 파괴하게 된다. 그러나 다음과 같은 buffer의 작용으로 DNA 파괴를 막는다.

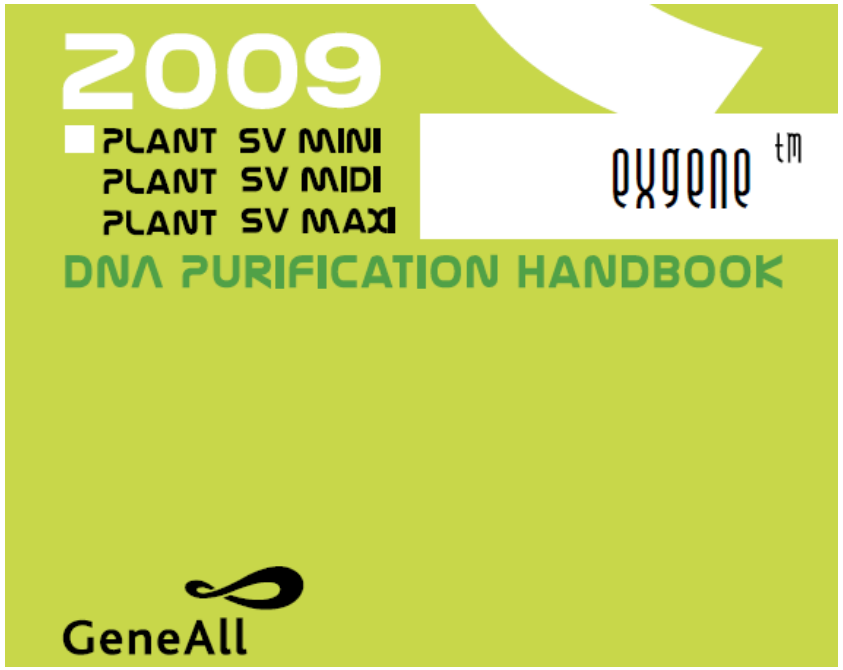
1) Extraction buffer 속의 **CTAB은 계면활성제**의 일종으로 계면활성제는 모든 물질들을 각각 물리적으로 분리시켜 서로 반응 하지 못하도록 한다. 또한 세포막의 주성분인 인지질을 녹이는 역할을 한다.

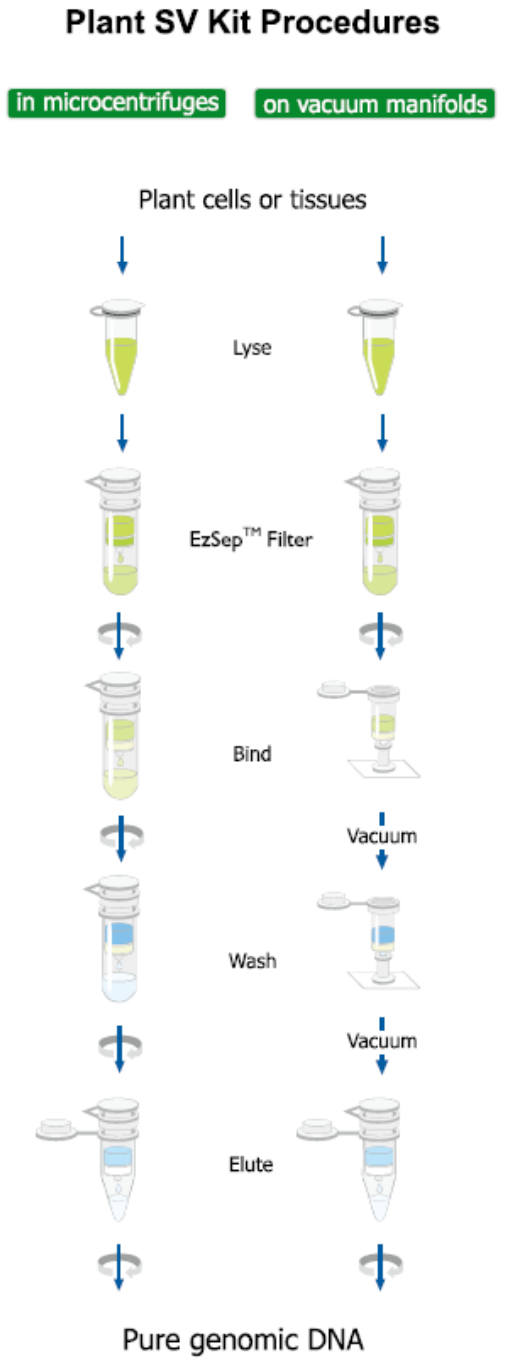
2) Extraction buffer 속의 EDTA는 **킬레이트(chelate) 작용**을 하며, 효소작용에 필수적인 Mg2+ 같은 금속이온을 둘러싸서 효소를 불활성화 시킨다.

※ 킬레이트: 한 개의 리간드가 금속 이온과 두 자리 이상에서 배위결합을 하여 생긴 착이온을 뜻한다.

3) 이후 chloroform등의 **단백질 비활성화 물질을 처리**함으로서 모든 효소작용을 정지시킨다.

- DNA는 염(salt)의 존재 하에서 70%정도의 EtOH에서 엉기는 (pellet을 형성) 성질을 갖고 있다. 이 때 원심분리를 하면 엉긴 DNA는 가라앉고 다른 이물질은 용액속에 남아있게 된다. 이러한 성질을 이용하여 순수한 DNA를 추출할 수 있다.
- DNA는 물과 TE (Tris-EDTA) buffer에서 용해 되는데, 특히 TE에서는 매우 잘 용해된다.
- 전통적인 식물 DNA 추출방법: CTAB method http://amborella.net/2024-Nanopore-Lecture/04-CTAB-DNAextraction-Doyle-Sangtae.pdf
- 최근에는 소량의 추출 시 주로 commercial kit 이용 http://amborella.net/2024-Nanopore-Lecture/05-Kit-DNAextraction-GeneAll.pdf





상업용 킷트에 의한 일반적인 DNA 추출법:

Filter-binding method에 의한 방법

그러나 매우 짧은 단편화된 DNA가 추출됨.

→ PCR 등 일반 실험을 위해서는 문제 없지만,

Nanopore로는 적합하지 않음.

최근 성신여자대학교 식물분자계통학실에서 개발하여 출판된 식물 HMW DNA 추출 방법
(2023년 6월 APPS 출판)

본 프로그램의 첫 part는 이 논문의 protocol에 준하여 수행 됨.

Applied in Plant Sciences의 DNA추출을 위한 특별호에 개제됨.

https://bsapubs.onlinelibrary.wiley.com/toc/21680450/2023/11/3

Received: 26 September 2022 | Accepted: 4 May 2023

DOI: 10.1002/aps3.11528

PROTOCOL NOTE

Applications in Plant Sciences

High-molecular-weight DNA extraction for long-read sequencing of plant genomes: An optimization of standard methods

Myoungbo Kang[1] | Andre Chanderbali[2] | Seungyeon Lee[1] | Douglas E. Soltis[2,3] | Pamela S. Soltis[2] | Sangtae Kim[1]

[1]Department of Biotechnology, Sungshin Women's University, Seoul 01133, Republic of Korea

[2]Florida Museum of Natural History, University of Florida, Gainesville, Florida 32611, USA

[3]Department of Biology, University of Florida, Gainesville, Florida 32611, USA

Correspondence

Sangtae Kim, Department of Biotechnology, Sungshin Women's University, Seoul 01133, Republic of Korea.
Email: amborella@sungshin.ac.kr

This article is part of the special issue "Emerging Methods in Botanical DNA/RNA Extraction."

Abstract

Premise: Developing an effective and easy-to-use high-molecular-weight (HMW) DNA extraction method is essential for genomic research, especially in the era of third-generation sequencing. To efficiently use technologies capable of generating long-read sequences, it is important to maximize both the length and purity of the extracted DNA; however, this is frequently difficult to achieve with plant samples.

Methods and Results: We present a HMW DNA extraction method that combines (1) a nuclei extraction method followed by (2) a traditional cetyltrimethylammonium bromide (CTAB) DNA extraction method for plants with optimized extraction conditions that influence HMW DNA recovery. Our protocol produced DNA fragments (percentage of fragments >20 kbp) that were, on average, ca. five times longer than those obtained using a commercial kit, and contaminants were removed more effectively.

Conclusions: This effective HMW DNA extraction protocol can be used as a standard protocol for a diverse array of taxa, which will enhance plant genomic research.

KEYWORDS

CTAB, DNA extraction, Femto Pulse system, high-molecular-weight DNA, nuclei extraction

Kang et al. (2023) 내용 요약:

The successful application of third-generation sequencing technologies for sequencing nuclear genomes requires high-molecular-weight (HMW) DNA in sufficient quantity and quality for library preparation and sequencing (Healey et al., 2014). These DNA requirements are often challenging for non-model plant species and represent an important bottleneck for plant genome research; therefore, the development of an efficient HMW DNA extraction method is essential for the plant genomics community. Although several approaches have recently been provided for HMW DNA extraction from plants, they were only applied to a few taxa, required additional purification steps, or the essential factors influencing the process were not adequately discussed (Healey et al., 2014; Mayjonade et al., 2016; Li et al., 2020; Cai et al., 2021; Jones et al., 2021; Mavrodiev et al., 2021; Zerpa-Catanho et al., 2021). Therefore, there is a need for an easy-to-use protocol that can produce HMW DNA from a wide range of plant taxa at a low cost.

In this study, we propose a HMW plant DNA extraction method that combines two classic protocols: (1) a nuclei extraction method (Green et al., 1987) and (2) a cetyltrimethyl-ammonium bromide (CTAB) plant DNA extraction method (Doyle and Doyle, 1987), with modifications. The nuclear extraction step reduces the ratio of organelle genomes in the extracted DNA (Hanania et al., 2004). The CTAB method has been modified in our protocol to solve the problems associated with phenolics and polysaccharides: polyvinylpyrrolidone (PVP) was added to isolate genomic DNA, as suggested by Healey et al. (2014). To more efficiently meet the needs of genome sequencing, our combined protocol includes (1) improvements to optimize time and reagent requirements and (2) suggestions of favorable conditions for factors influencing the results

(number of pipetting steps, grinding time in liquid nitrogen, and centrifugation force in *g*). A combination of these two classic protocols has already been proposed for high-quality DNA extraction from *Vitis vinifera* L. (Hanania et al., 2004), but not with regard to HMW DNA and applicability in other taxa. Similarly, a method combining the nuclear isolation process and sodium dodecyl sulfate (SDS)-based DNA extraction protocol has recently been proposed for HMW DNA extraction (Zerpa-Catanho et al., 2021); however, its effectiveness has only been confirmed in a few plant taxa (six genera in three families), and it requires an extra purification step (QIAGEN Genomic Tip 20/G columns; QIAGEN, Hilden, Germany). By contrast, we have assessed the broad applicability of our protocol in species representing 18 orders of flowering plants from all major angiosperm lineages (Angiosperm Phylogeny Group, 2016), as well as a gymnosperm, *Pinus* L.

To confirm the effectiveness of our HMW DNA extraction method, we compared the results with those obtained using a commercial plant DNA extraction kit. The DNA length distributions and purity were evaluated as validation criteria for comparing the two methods. We also discuss factors influencing the results, such as the number of pipetting steps, grinding time in liquid nitrogen, and centrifugation force in *g*.

METHODS

HMW DNA extraction method

We sampled leaves of species from each of 18 major angiosperm orders and one gymnosperm to test the taxon-specific efficiency of our protocol. For details of all samples used in this study, see Appendix 1. Reagents, recipes, and a stepwise protocol are provided in Appendix 2. Our HMW DNA extraction protocol consists of three major steps: (1) grinding and nuclei isolation, (2) nuclear DNA extraction using CTAB buffer, and (3) RNase A and proteinase K treatment. We started with 2 g of tissue (preferably fresh, young leaves) and used a vacuum-aided cell strainer (40 µm and 100 µm; pluriSelect Life Science, Leipzig, Germany) to collect the nuclei suspension. We also conducted additional DNA extractions using the same samples from our HMW DNA extraction protocol. For this, we employed the Exgene Plant SV kit (GeneAll Biotechnology, Seoul, Republic of Korea), a commercial plant DNA extraction kit based on the DNA-binding filter method. Following the instructions in the manufacturer's manual, we used 0.1 g of leaf tissue, which is the recommended amount for fresh leaves.

Grinding and nuclei isolation

The protocol starts with 2 g of fresh, young leaves. We ground the leaves into a powder in liquid nitrogen (−80°C) and placed the powder in 20 mL of nuclei isolation buffer (IB). After 30 s of vortexing, we added Triton X-100 (20 µL) and β-mercaptoethanol (1.5 mL). This step should be conducted inside a fume hood as β-mercaptoethanol is toxic. The samples were placed on ice for 10 min, and then the mixture was filtered through a 100-µm cell strainer (pluriStrainer 100 µm; pluriSelect Life Science) seated in a 50-mL conical tube to collect the nuclear suspension. During filtration, gently scraping plant material accumulated on the filter with the side of a 1000-µL pipette tip may facilitate a smoother filtration. The filtering step was repeated with a 40-µm cell strainer (pluriStrainer 40 µm; pluriSelect Life Science), and Triton X-100 (200 µL) was added to the obtained nuclear suspension. This process lyses the cell and organellar membranes but not the nuclear membrane (Peterson et al., 1997). As a non-ionic detergent, Triton X-100 facilitates the release of nuclei from cells and prevents nuclei from clumping (Loureiro et al., 2007). To pellet the nuclei, the samples were centrifuged, and the supernatant was discarded. Centrifugation for 10 min at 3000 × *g* (4°C) is recommended to prevent fragmenting long DNA molecules (see Results).

Nuclear DNA extraction using CTAB buffer

The nuclei pellet was resuspended in 5 mL of Carlson Lysis Buffer (Carlson et al., 1991). Adding β-mercaptoethanol (12.5 µL) denatures globular proteins to make them insoluble in water (Jadhav et al., 2015). An incomplete resuspension can reduce yield; thus, we incubated the samples at 65°C for a minimum of 15 min for efficient resuspension. If the pellet still does not suspend, crushing the pellet with a pipette tip might be helpful. For easy handling, we transferred the suspended nuclei pellet to a 15-mL tube instead of proceeding with the 50-mL tube. We added 5 mL (equal volume) of chloroform:isoamyl alcohol (24:1 [v/v]) to remove impurities. During this step, chloroform ($CHCl_3$; a non-polar 3-hydrophobic solvent) dissolves non-polar proteins and lipids to promote the partitioning of lipids and cellular debris into the organic phase. Isoamyl alcohol ($C_5H_{12}O$) prevents the emulsification of the solution (Jadhav et al., 2015). After centrifugation (3000 × *g* for 10 min at 4°C), the aqueous upper phase containing DNA was collected and transferred into a new tube, while the organic phase containing lipids, proteins, and other impurities was discarded. The separation of a pure aqueous phase is critical for the purity of the end product, and we recommend collecting just four-fifths of the upper liquid volume to avoid including any cellular debris. Adding the proper ratio of sodium acetate (NaOAc) and isopropanol to the acquired supernatant is essential for precipitating the DNA: for every 10 mL of supernatant, a 1/10 volume of 3 M NaOAc (1 mL) and the same volume (including NaOAc) of room-temperature isopropanol (11 mL) are needed. It is essential to use room-temperature isopropanol for this step; otherwise, both polysaccharides and DNA will precipitate (Shepherd and McLay, 2011). The precipitated DNA was separated from other solvents through centrifugation (3000 × *g* for 10 min at 4°C), and the resulting DNA pellet was washed with 70% cold ethanol, recentrifuged (3000 × *g* for 10 min at 4°C), and thoroughly dried. We recommend rapidly drying samples using room-temperature air blown by a hair dryer.

RNase A and proteinase K treatment

The DNA pellet was dissolved in 2 mL Tris-EDTA (TE) buffer. To remove RNA and protein efficiently, which account for most of the impurities in extracted DNA, we treated the samples with RNase A (10 mg/mL) and proteinase K (>600 units/mL), respectively. For each treatment, the proper incubation time and enzyme activation temperature are important: 5 min at 37°C for RNase A and 15 min at 50°C for proteinase K. The enzymes used in each step are removed by a treatment with 2 mL of chloroform: isoamyl alcohol (24:1 [v/v]). After treatment with RNase A and proteinase K, the same precipitation procedure as for the CTAB extraction is followed. The resulting pellet is dissolved using an appropriate amount of deionized water (50–500 µL) according to the size of the pellet (recommended final concentration is ca. 200 ng/µL). If it is difficult to dissolve the pellet, we recommend incubating the tube at 50°C. If the pellet remains after incubation at 50°C, it is recommended to take only the dissolved aqueous layer after brief centrifugation.

Quality evaluation of extracted DNA

The quantity and purity (A_{260}/A_{280} and A_{260}/A_{230} ratios) of the extracted DNA were measured using a Qubit 4 Fluorometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific), respectively. The length distribution of the extracted DNA was evaluated using a Femto Pulse system (Agilent Technologies, Santa Clara, California, USA).

Optimization of conditions for HMW DNA recovery

We tested three factors influencing the results: (1) the number of pipetting steps, (2) the grinding time in liquid nitrogen, and (3) the centrifugation force in *g*. Three independent experiments were performed on different taxa in each case to evaluate each factor. First, we tested the impact of high *g* forces during centrifugation on DNA damage by comparing the setting in our protocol (3000 × *g*; control group) and a higher setting (5000 × *g*; experimental group). Second, the amount of grinding was compared. The control group was subjected to one minute of grinding (ensuring the sample was fully chilled before grinding began). The experimental group was subjected to an additional two minutes of grinding after adding extra liquid nitrogen. Third, we assessed whether high-speed and multiple pipetting steps could potentially damage DNA. We conducted pumping at the maximum-achievable speed 200 times in a tube using a P200 tip (experimental group) and compared the resulting DNA size distribution with the original DNA (control group).

RESULTS

DNA quantity, size, and purity measurements

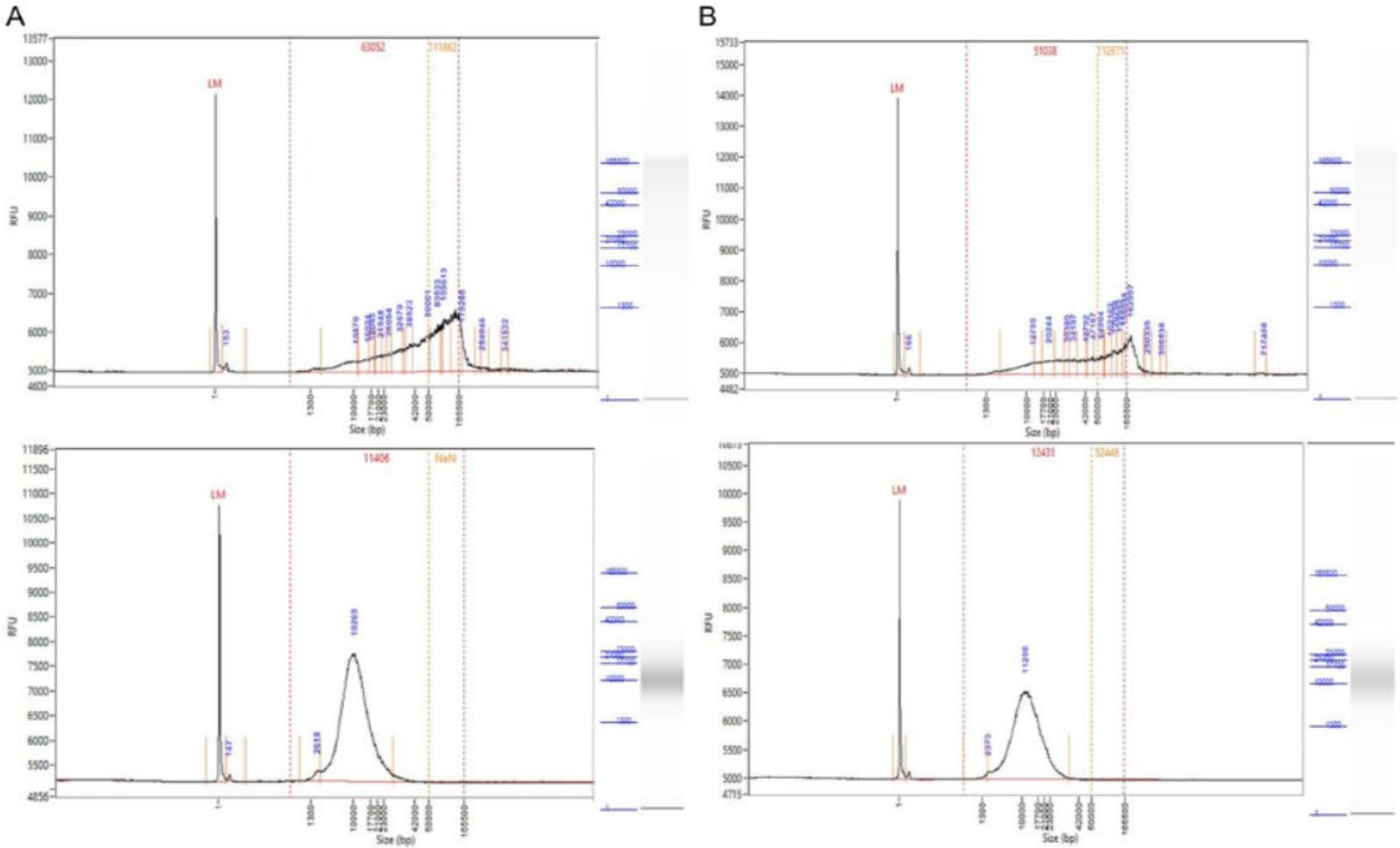
Usually, the quantity of the end DNA product per extraction is enough to generate 4–5 libraries (8–15 µg) for long DNA sequencing with MinION or GridION (Oxford Nanopore Technologies, Oxford, United Kingdom), based on the library construction protocol (Ligation Sequencing Kit). The measurements obtained through the Femto Pulse system (peak height and percentages of fragments >20 kbp in the fragment-length distribution graph) confirm that our protocol successfully produced DNA fragments an average of five times longer than those generated using the commercial kit (Table 1, Figure 1), although the results of our standard HMW DNA extraction protocol showed different patterns depending on the taxon (Figure 2A, B). With our protocol, the taxon with the highest portion of >20-kbp fragments was *Chloranthus fortunei* Solms (Chloranthales; 83.6%), and the longest peak of DNA fragment distribution was obtained from *Alisma plantago-aquatica* subsp. *orientale* (Sam.) Sam. (Alismatales; 183.0 kbp) (Table 1). In the most efficient instance, our protocol yielded 35 times more DNA fragments over 20 kbp (77.1%) in *Lysimachia davurica* Ledeb. (Ericales) than the commercial kit, for which only 2.2% of fragments were greater than 20 kbp.

The quality of DNA extracted using the HMW method was superior to that obtained using the kit method in most samples. In the context of next-generation sequencing, high-quality DNA is characterized as predominantly HMW with an A_{260}/A_{280} ratio over 1.8 and without contaminating substances, such as polysaccharides or phenolics (Kasem et al., 2008; Desjardins and Conklin, 2010). With both methods, the A_{260}/A_{280} absorbance ratio, which measures protein contamination, showed similar results with low contamination (both averaged 1.83); however, our standard protocol more effectively removed carbohydrates and organic solvents (average A_{260}/A_{230} ratio = 1.88) than the commercial kit (average A_{260}/A_{230} ratio = 1.49) (Table 2; Figure 2C, D). Generally, A_{260}/A_{230} values between 1.8–2.2 indicate DNA is free of carbohydrates and organic solvents (Kasem et al., 2008; Desjardins and Conklin, 2010).

To address the statistical difference between the results from our protocol and a commercial kit, we performed paired *t*-tests on all pairs of DNA length and quality, with $P < 0.05$ considered significant. In the DNA length

TABLE 1 A comparison between our HMW DNA extraction method and a commercial kit. Fragment lengths were estimated using the Femto Pulse system.

	HMW method		Commercial kit		
Taxon	**Peak (kbp)**	**% of >20 kbp (a)**	**Peak (kbp)**	**% of >20 kbp (b)**	**Ratio (a)/(b) × 100 (%)**
Platycladus orientalis	21.57	58.8%	17.70	32.4%	181.5%
Nymphaea tetragona var. *minima*	22.10	59.7%	14.04	14.4%	414.6%
Chloranthus fortunei	38.21	83.6%	26.87	54.2%	154.2%
Asarum sieboldii	22.74	69.1%	28.81	66.5%	104.0%
Alisma plantago-aquatica subsp. *orientale*	183.00	56.8%	11.21	10.8%	525.9%
Hemerocallis fulva	24.80	66.2%	31.48	70.2%	94.3%
Carex breviculmis	107.36	83.2%	22.19	52.4%	159.8%
Epimedium koreanum	169.21	67.2%	10.60	10.2%	658.8%
Euonymus alatus	27.04	67.0%	22.96	58.1%	115.3%
Viola collina	142.52	75.1%	10.69	14.9%	504.0%
Spiraea prunifolia var. *simpliciflora*	165.50	70.1%	10.27	7.2%	973.6%
Pelargonium inquinans	22.45	65.2%	21.20	55.7%	117.1%
Aesculus turbinata	24.00	66.1%	15.38	26.1%	253.3%
Lysimachia davurica	154.32	77.1%	9.08	2.2%	3504.5%
Isodon inflexus	23.88	71.9%	21.05	49.1%	145.5%
Ipomoea nil	23.65	68.4%	25.78	42.1%	162.5%
Adenophora erecta	132.21	67.2%	13.36	19.1%	351.8%
Cicuta virosa	17.70	45.5%	21.05	54.4%	83.6%
Sambucus williamsii	157.43	64.8%	20.67	44.0%	147.2%
Average	77.88 ± 64.48	67.53% ± 0.09%	18.65 ± 6.67	36.00% ± 0.21%	455.34% ± 7.55%


A
B
RFU
Size (bp)
LM
63052
11406
NaN
51038
12431
10269
11266

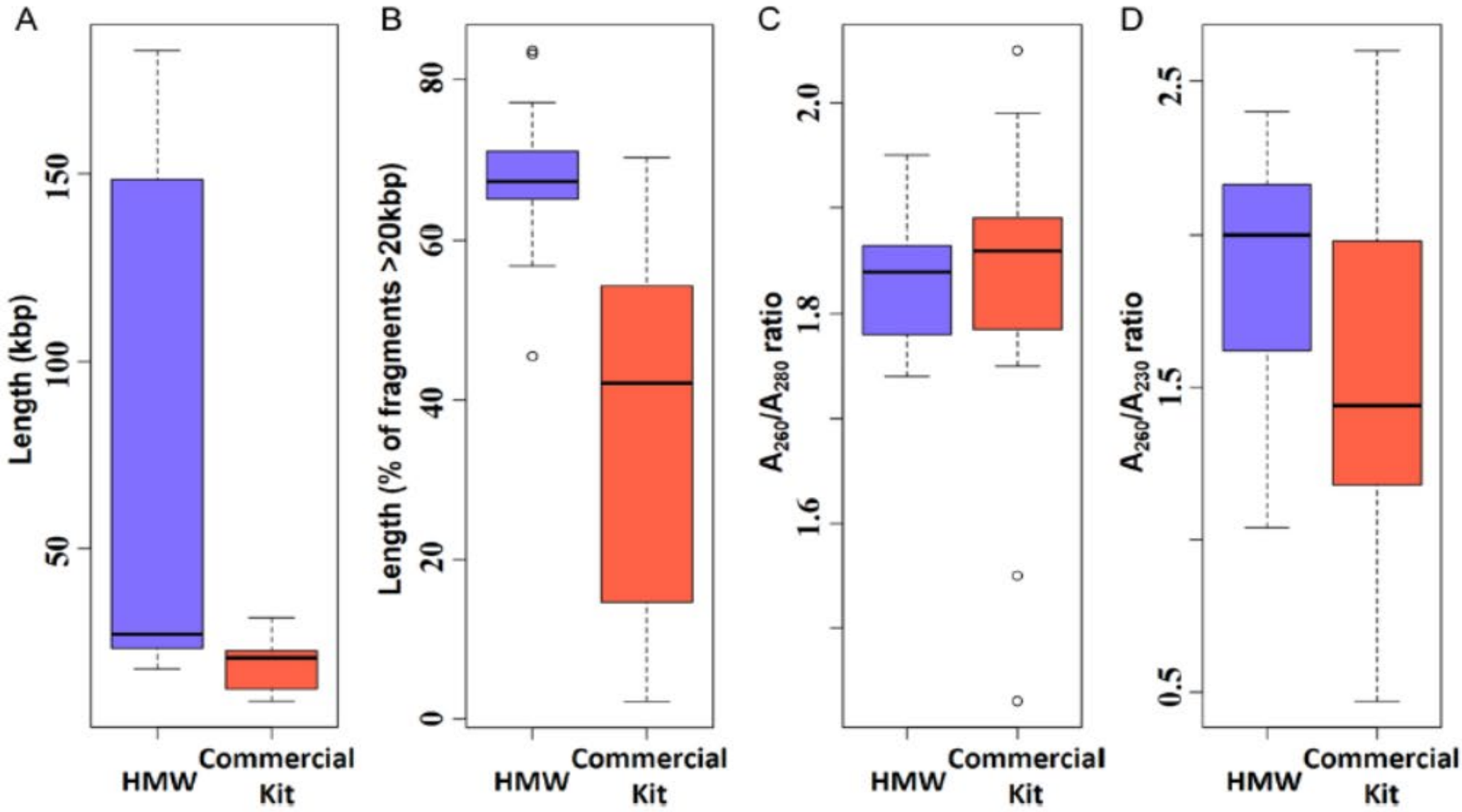


FIGURE 2 Comparison of the size and quality of DNA extracted using the two methods. (A, B) Size comparisons of (A) the highest peak and (B) the percentage of fragments >20 kbp. (C, D) Quality comparisons using (C) A_{260}/A_{280} ratio and (D) A_{260}/A_{230} ratio. The bold horizontal line in the middle of the box plot is the median value, and the lower and upper boundaries indicate the 25th and 75th percentiles, respectively.

Appendix 2: An optimized protocol for high-molecular-weight (HMW) DNA extraction in plant genomic studies.

Note: This protocol starts with 2 g of fresh, young leaves. Usually, the end product of one extraction process is sufficient to generate 4–5 libraries for sequencing with MinION or GridION (Oxford Nanopore Technologies, Oxford, United Kingdom).

I. Preparation of solutions

1. Preparation of nuclei isolation buffer (IB) (for 10 reactions)

- For 200 mL of nuclei IB, dissolve the following in ca. 100 mL of water:
  - 3 mL Tris-HCl (1 M stock, pH 9.5; final concentration: 15 mM)
  - 4 mL EDTA (0.5 M stock; final concentration: 10 mM)
  - 1.94 g KCl (final concentration: 130 mM)
  - 0.8 mL NaCl (5 M stock; final concentration: 20 mM)
- Gradually add 16 g of polyvinylpyrrolidone (PVP)-10 while rapidly stirring the solution with a magnetic stir bar.
- Use water to increase the volume to 200 mL.
- Add 0.05 g of spermine and 0.07 g of spermidine. Store IB at 4°C.
- Prepare 20 µL of Triton X-100 and 1.5 mL of β-mercaptoethanol, to be added after mixing the IB with the ground tissue (final concentrations of 0.1% and 7.5%, respectively; this constitutes IBTB).

Note: Store at 4°C until use, or for a maximum of two weeks.

2. Preparation of Carlson Lysis Buffer (Carlson et al., 1991)

- Carlson Lysis Buffer = 2× cetyltrimethylammonium bromide (CTAB) buffer + 1% polyethylene glycol (PEG) 6000
- For 100 mL of Carlson Lysis Buffer:
  - 10 mL Tris-HCl (1 M stock, pH 9.5; final concentration: 100 mM)
  - 4 mL EDTA (0.5 M stock; final concentration: 20 mM)
  - 8.2 g NaCl (final concentration: 1.4 M)
  - 2 g CTAB (final concentration: 2%)
  - 1 g PEG (final concentration: 1%)

Note: Store at room temperature until use, or for up to two weeks.

3. Tris-EDTA buffer (TE) (1×)

- TE buffer = 10 mM Tris-HCl (pH 8.0) + 1 mM EDTA

Note: Store at 4°C until use.

II. Grinding and nuclei isolation (modified from Hanania et al., 2004)

1. Chill mortar and pestle at −80°C before beginning the extraction procedure. Grind 2 g of fresh, young leaves in liquid nitrogen for 1 min.

Note: Ensuring the sample is fully chilled before grinding.

2. Add 2 g of ground leaf powder to 20 mL of IB in a 50-mL conical tube and mix by inverting.

Note: Over-grinding negatively affects the extraction of HMW DNA. Grinding for 1 min is fine; additional grinding with extra liquid nitrogen is not needed.

Note: Increase the sample amount for succulent plants, and increase the volume of IB when the mixture becomes viscous.

3. Immediately add 20 µL of Triton X-100 and 1.5 mL of β-mercaptoethanol and mix by inverting.
4. Keep on ice for 10 min.

Note: This step should be conducted inside a fume hood because the IBTB contains β-mercaptoethanol, which is toxic.

5. Filter the mixture through a vacuum-aid cell strainer (pore size: 100 µm) seated on a 50-mL conical tube to collect the nuclei suspension (Figure A1).

Note: To aid filtration, gently scrape away plant tissue from the filter with the top of a 1000-µL (blue) pipette tip. The filtrate should be light green.

6. Repeat the filtering step with a 40-µm pore cell strainer.
7. Add 200 µL Triton X-100 to the nuclei suspension.

Note: This step lyses cell and organellar membranes, but not the nuclear membrane.

8. To pellet the nuclei, centrifuge for 10 min at 3000 × *g* at 4°C.
9. Discard the supernatant.

III. Nuclear DNA extraction using CTAB buffer (modified from Doyle and Doyle, 1987)

1. Add 5 mL of Carlson Lysis Buffer and 12.5 µL β-mercaptoethanol to the tube and resuspend the nuclei pellet with brief tapping.

Note: Incomplete resuspension could reduce the yield as many nuclei will not have been lysed by CTAB. Briefly pipetting the pellet with an end-cut 1000-µL pipette tip and gentle vortexing may aid resuspension.

2. Incubate at 65°C for 15 min (maximum 2 h).

Note: If the pellet is not completely resuspended after incubation, a brief centrifugation (3000 × *g* for 5 min) followed by only the use of the supernatant will help speed up processing.

3. Transfer the suspended nuclei pellet to a 15-mL polypropylene tube and add an equal volume (5 mL) of chloroform:isoamyl alcohol (24:1 [v/v]) solution.
4. Invert several times to mix.
5. Centrifuge (3000 × g) for 10 min at 4°C.

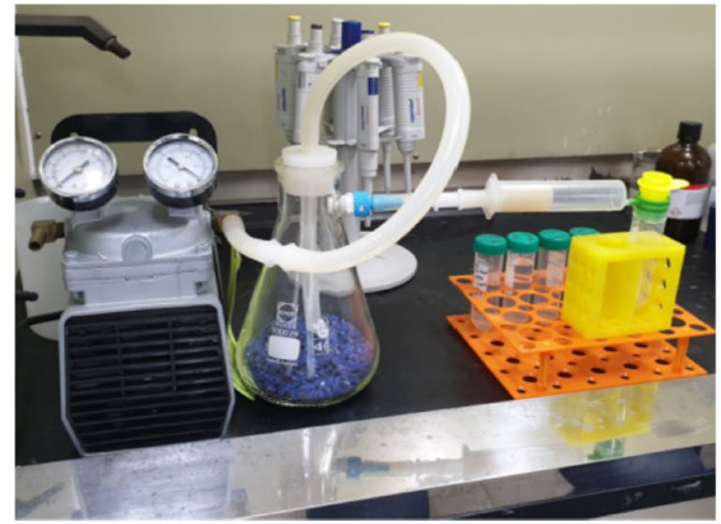


FIGURE A1 The setup of the vacuum-aid filtration including a liquid-overflow trap filled with silica gel. Using this setup can shorten extraction times.

6. Transfer the aqueous upper phase to a new tube using a P1000 pipette.

Note: Take only 80% of the supernatant to avoid the inclusion of cellular debris. Take care removing the supernatant as this step is highly correlated with the quality of extracted DNA.

Note: If the supernatant is viscous, slow pipetting will help avoid sucking up the plant tissue.

7. Repeat steps 3–6 (optional but highly recommended).
8. Add a 1/10 volume of 3 M sodium acetate (NaOAc), mix gently, add the same volume of isopropanol (room temperature), and gently invert several times.

Note: For 4.5 mL of supernatant, add 0.45 mL of 3 M NaOAc and 4.95 mL of isopropanol.

9. Precipitate at −20°C for more than 1 h.

Note: If precipitates are visible, moving to the next step is possible for faster processing. For highly viscous extracts, cold treatment makes the extract more viscous and more difficult to work with.

10. Centrifuge (3000 × *g*) for 10 min at 4°C.
11. Discard supernatant.
12. Wash pellets with 70% cold ethanol (ca. 20 mL per tube).
13. Centrifuge (3000 × *g*) for 10 min at 4°C.
14. Discard supernatant.

Note: Keep the tube inverted for 1 min, and wipe out the tube wall with a Kimwipe.

15. Dry the pellet completely.

Note: This step is very important for the quality of the DNA. The smell of alcohol is a good indicator of incomplete drying.

IV. RNase A and proteinase K treatment

1. Dissolve the pellet with 2 mL of TE buffer.

Note: If the pellet is difficult to dissolve, incubate in a 50°C water bath for up to 10 min.

Note: Gently crushing the pellet with a pipette tip might be helpful for faster resuspension, but never vortex the sample. If the pellet is not completely resuspended after incubation, a brief centrifugation (3000 × *g* for 5 min) followed by only the use of the supernatant will help speed up processing.

2. Add 20 µL (10 µL/mL) of RNase A (10 mg/mL conc.).
3. Incubate at 37°C for 5 min.
4. Add 20 µL (10 µL/mL) of proteinase K (>600 units/mL conc.).
5. Incubate at 50°C for 15 min.
6. Add an equal volume (2 mL) of chloroform:isoamyl alcohol (24:1 [v/v]).
7. Invert several times to mix.
8. Centrifuge (3000 × *g*) for 10 min at 4°C.
9. Transfer the aqueous upper phase to a new 15-mL tube.

Note: Taking only 90% of the supernatant is best to avoid the inclusion of cellular debris. This is highly correlated with the quality of the extracted DNA.

10. Repeat steps 6–9 (optional).
11. Add a 1/10 volume of 3 M NaOAc, mix gently, add an equal volume of room-temperature isopropanol, and gently invert several times.

Note: For 3.5 mL of supernatant, add 0.35 mL of 3 M NaOAc and add 3.85 mL of isopropanol.

12. Precipitate at −20°C for more than 1 h.

Note: If aggregates are visible, moving to the next step is possible for faster processing.

13. Centrifuge (3000 × *g*) for 10 min at 4°C.
14. Discard supernatant.
15. Wash pellets with 70% cold ethanol (ca. 5 mL per tube).
16. Centrifuge (3000 × *g*) for 10 min at 4°C.
17. Discard supernatant.

Note: Keep the tube inverted for 1 min, and wipe out the tube wall with a Kimwipe.

18. Dry the pellet completely.

Note: This step is very important for the quality of DNA. The smell of alcohol is a good indicator of incomplete drying.

19. Add 50–500 µL of deionized water to each tube to dissolve the pellet.

Note: If it is difficult to dissolve, incubate in a 50°C water bath for up to 10 min.

Note: Crushing the pellet with a pipette tip might be helpful for faster resuspension, but never vortex the sample. If the pellet is not completely resuspended after incubation, a brief centrifugation (3000 × *g* for 5 min) followed by only the use of the supernatant will help speed up processing.

V. DNA size and quality measurements

1. Check the quality (A_{260}/A_{280} and A_{260}/A_{230} ratios) and quantity of extracted DNA using a NanoDrop 2000 (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and a Qubit 4 Fluorometer (Thermo Fisher Scientific).
2. Check the length distribution of the DNA fragments using the Femto Pulse system (Agilent Technologies, Santa Clara, California, USA).

VI. Special reagents and consumables

1. Reagents

- PVP-10: MilliporeSigma (Burlington, Massachusetts, USA) CAS 9003-39-8
- Spermine: MilliporeSigma S2876
- Spermidine: MilliporeSigma S2501
- Triton X-100: MilliporeSigma T8787
- PEG 6000: MilliporeSigma 81260
- RNase A: MilliporeSigma R6513
- Proteinase K: MilliporeSigma P2308

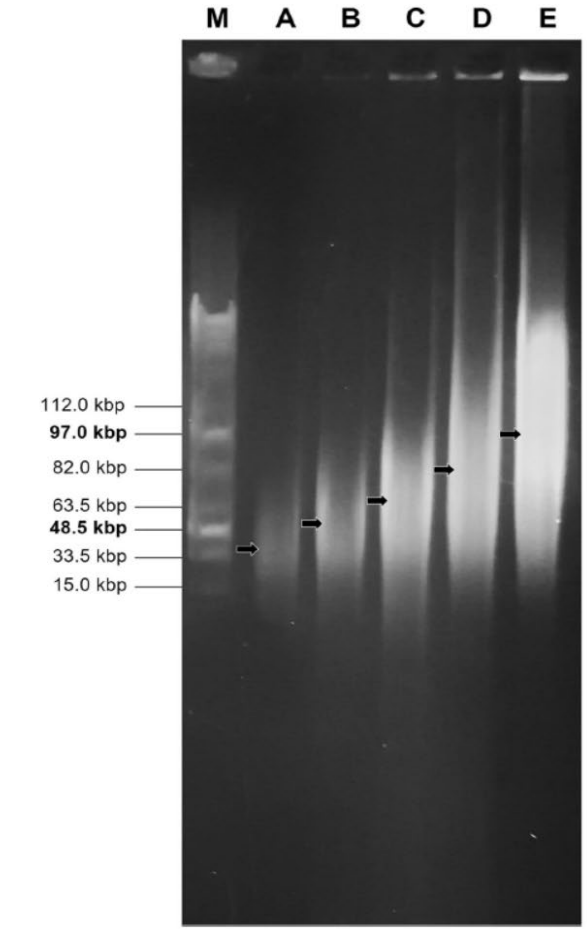
2. Consumables

- Vacuum-aid cell strainer (100 µm): pluriSelect Life Science (pluriSelect Life Science, Leipzig, Germany) 43-50100-51 yellow 100 µm
- Vacuum-aid cell strainer (40 µm): pluriSelect Life Science 43-50040-51 blue 40 µm
- Connector ring: pluriSelect Life Science 41-50000-03

Appendix 3: Evaluation of factors influencing the DNA extraction process, using three taxa each as examples. Effects of (A) pipetting repeats (experimental group: additional 200 pipetting pumps with P200 tip), (B) degree of grinding (experimental group: additional 2 min of grinding with a second pour of liquid nitrogen), and (C) centrifugation force (control group: 3000 × *g*; experimental group: 5000 × *g*).

	DNA fragment length			
	Peak (kbp)		% of fragments >50 kbp	
Example taxa for each factor	Control group (a)	Experimental group (b)	Control group (a)	Experimental group (b)
(A) Pipetting				
Chloranthus fortunei	25.99	22.00	12.5%	7.7%
Alisma plantago-aquatica subsp. *orientale*	165.50	91.6	26.1%	19.1%
Scutellaria insignis	37.89	36.18	12.3%	13.3%
Average ± standard deviation	76.46 ± 63.15	49.93 ± 30.03	17.0% ± 0.06%	13.4% ± 0.05%
Average (a) – average (b)	26.53 ± 41.04		3.6% ± 0.04%	
(B) Grinding				
Chloranthus fortunei	32.37	29.30	21.5%	0.0%
Carex breviculmis	107.36	88.95	34.7%	35.8%
Viola collina	142.52	147.49	46.8%	42.9%
Average ± standard deviation	94.08 ± 45.94	88.58 ± 48.25	34.3% ± 0.10%	26.2% ± 0.18%
Average (a) – average (b)	5.50 ± 11.88		8.10% ± 0.12%	
(C) Centrifugation				
Chloranthus fortunei	151.12	92.37	42.4%	31.1%
Scutellaria salviifolia	132.27	46.21	41.4%	27.7%
Sambucus williamsii	157.42	127.62	35.3%	31.4%
Average ± standard deviation	146.94 ± 10.69	88.73 ± 33.33	39.7% ± 0.03%	30.0% ± 0.02%
Average (a) – average (b)	58.20 ± 22.97		9.63% ± 0.04%	

Appendix 4: Effect of DNA amount on the pulse-field gel electrophoresis image (0.7% agarose). Different quantities of the same extracted DNA (*Magnolia grandiflora*) were loaded in each lane. The length of the brightest position (peak; arrows) and the quantity of DNA show a positive relationship. M = MidRange I PFG marker (New England Biolabs, Ipswich, Massachusetts, USA); A = 125 ng; B = 250 ng; C = 500 ng; D = 1 µg; E = 2 µg.



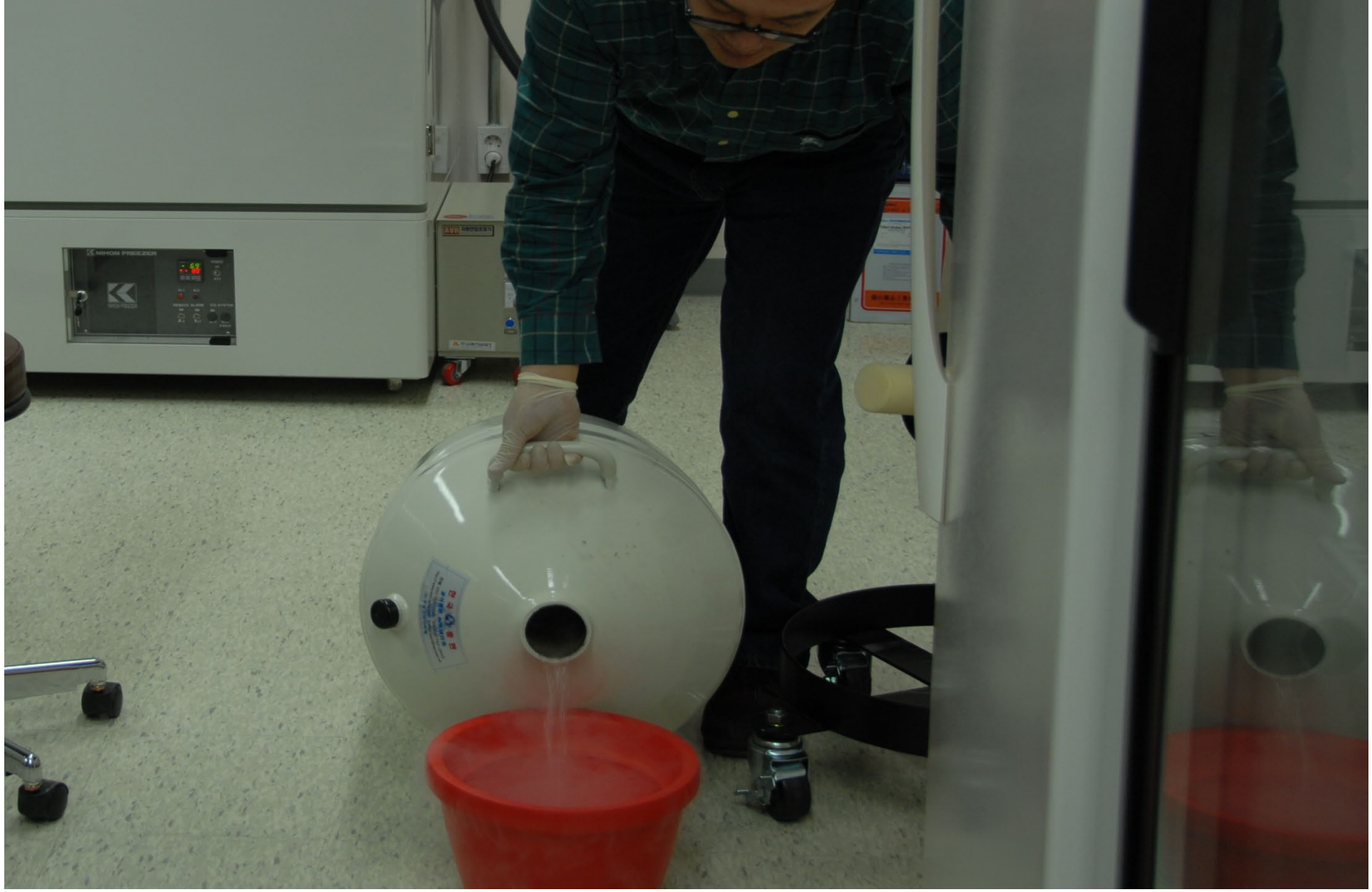
※ 참고

일반생물학 실험

CATB에 의한

DNA extraction

과정

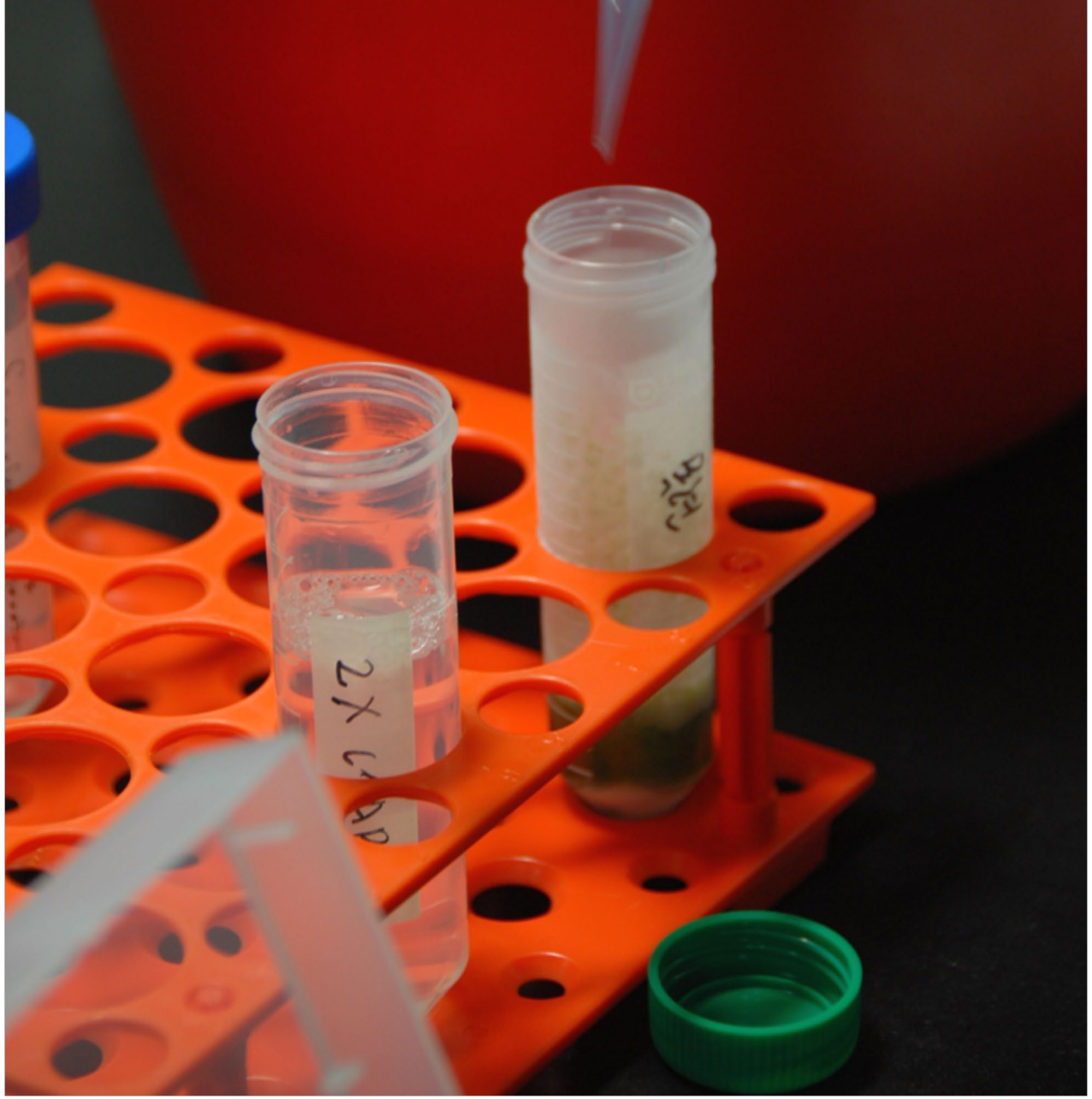


액체 질소를 전용 버켓에 붓는다.

시료를 파쇄하기 전 막자사발에 액체 질소를 부어 충분히 온도를 낮춰 준다.

시료를 액체질소에 넣고 막자사발로 파쇄 한다.
면장갑 두개를 겹쳐 끼고, 그 위에 폴리그로브를 끼고 파쇄 한다.

약수저도 액체질소에 보관하여 충분히 온도를 낮추어 시료를 옮기는데 사용한다.
2g의 시료 사용. 스페츌라도 액체질소에 보관!

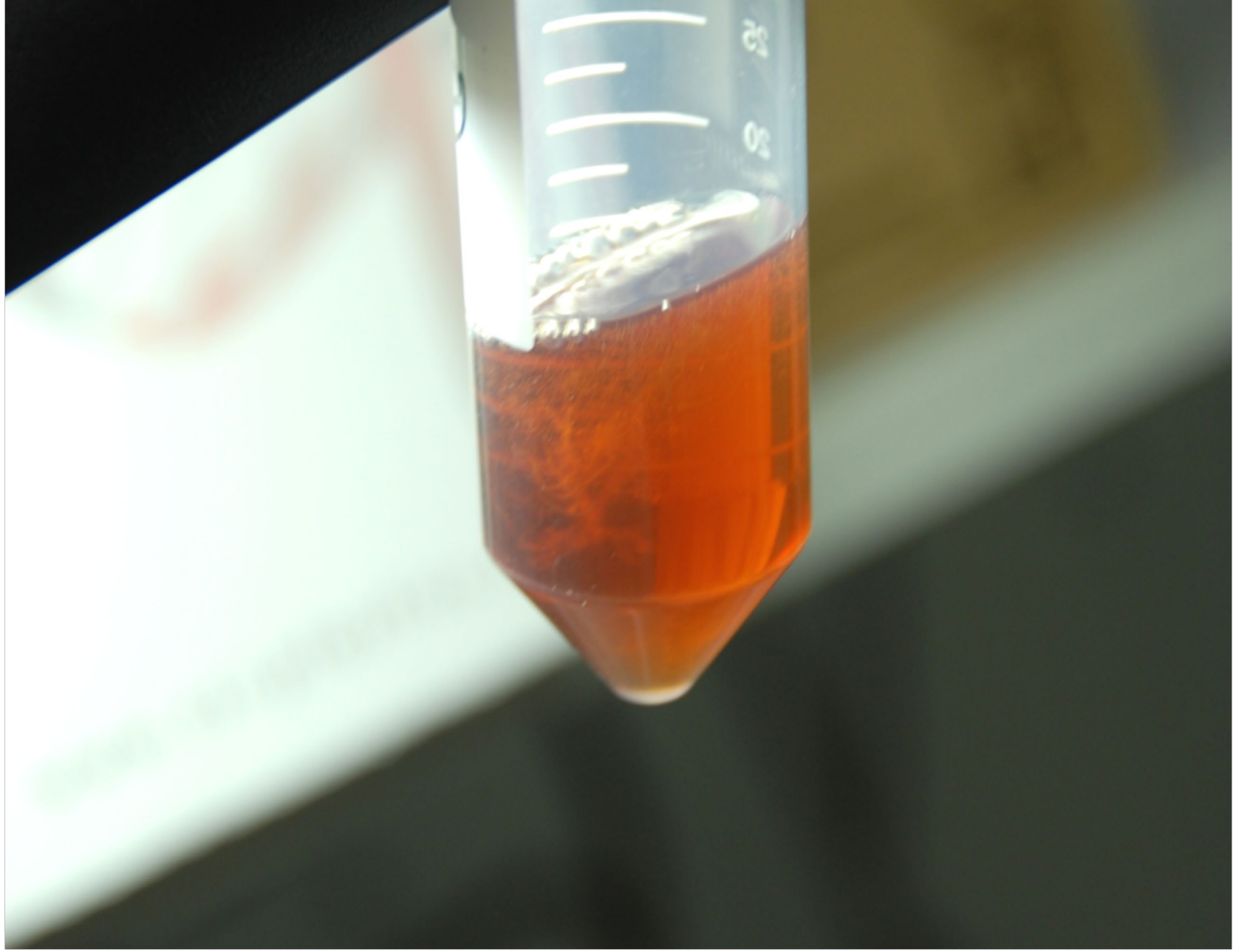


5ml의 2X CTAB buffe를 넣고(1ml pipette 사용), 65도에서 10분간 중탕한다.

Chloroform:Isoamyl alchole=24:1 solutio을 넣고 뚜껑을 잘 닫고 vortaxing.

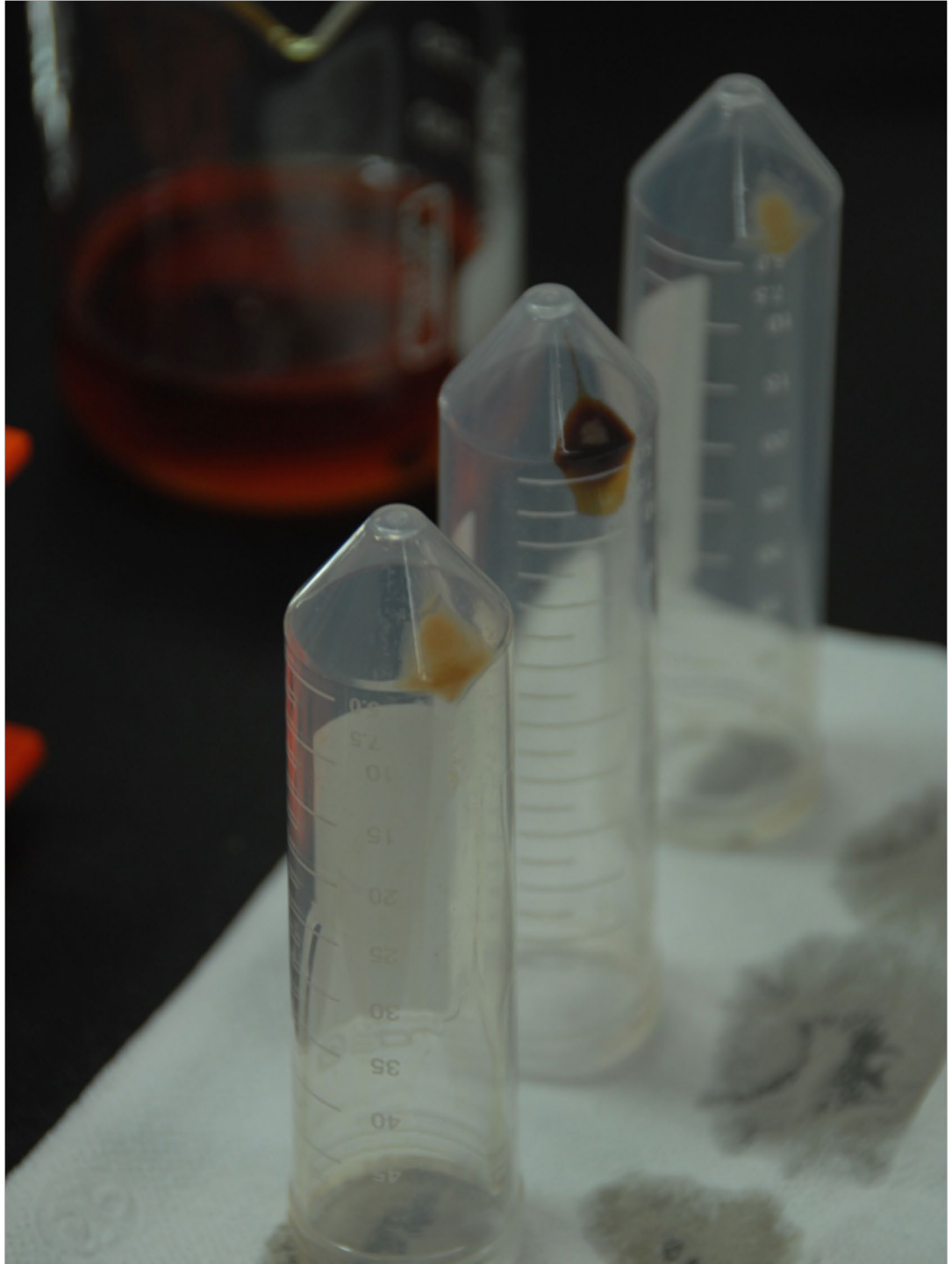
Centrifuge 후 상등액 만을 새로운 tube로 옮긴다.

2/3 volum의 isopropanol을 넣는다.



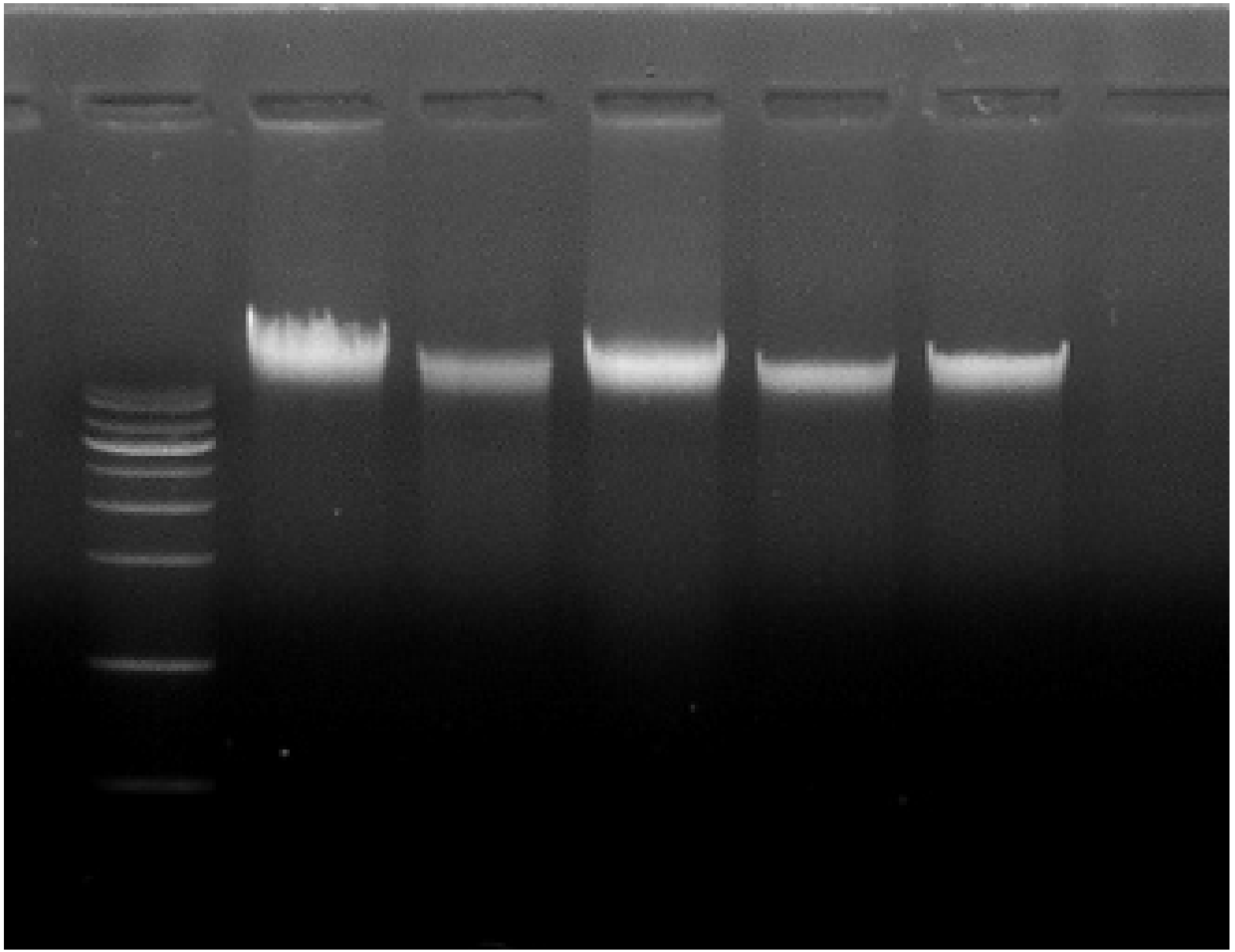
뚜껑을 닫고 tube를 뒤집어 섞어 주면 엉긴 DNA를 확인할 수 있다.

Desktop centrifuge로 원심분리하여 pellet을 가라앉힌다.



Tube를 뒤집어 상등액을 버리면 가라앉아 있는 DNA pellet을 확인할 수 있다.
Pellet이 아래로 떨어져 없어질 수 있으니 계속 지켜보면서 tube 안쪽을 kimwipe로 닦아준다.

DNA를 TE로 다시 녹인 후
1/10 volume의 ammonium acetate를 넣고,
2 volume의 99.9% EtOH 첨가하여 침전시킴.



Ⅳ. HMW DNA 추출 실험

Ⅴ. HMW DNA 추출 결과의 QC

Agarose gel electrophoresis에 의한 정량

- 100 ng/ul의 농도를 갖고있는 DNA를 serial dilution 하여 다양한 농도의 DNA를 만든 후 이를 한꺼번에 전기영동하여 gel 상에서의 DNA 농도 차이를 확인한다.

- **Serial dilution**

	Lambda DNA	DW	총량
① 1/3 시료 (33.3 ng/ul):	원래 농도의 시료 3ul	+ 6ul	= 9 ul
② 1/9 시료 (11.1 ng/ul):	①번 시료 3ul	+ 6ul	= 9 ul
③ 1/27 시료: (3.7 ng/ul):	②번 시료 3ul	+ 6ul	= 9 ul
④ 1/81 시료: (1.2 ng/ul):	③번 시료 3ul	+ 6ul	= 9 ul

- **Loading 하는 방법**

1) Parafilm을 잘라 놓고 loading dye (BPB; 푸른색)를 loading할 시료 수 만큼 1ul씩 분주하여 방울을 만들어 놓는다.

2) 1번 lane에 준비된 ladder를 3ul loading

3) 준비한 각 시료(① ~ ④번 시료) 5ul를 취하여 1ul의 loading dye 와 pipette으로 잘 섞고 (기포가 생기면 안됨; 전체 6ul) 2~5 번 lane에 loading함.

Band test: 농도를 알고있는 standard DNA의 serial dilution과 농도를 알고자 하는 DNA의 비교를 통한 DNA 정량

정확한 DNA 정량법

- 일반적으로 추출된 DNA를 정확히 정량 하기 위해서는 spectrophotometer를 이용한다. DNA는 염기는 공명구조를 가지는 화학적 특성상 UV를 흡수하는데 특히 260nm에서 나타나는 **OD (optical density)값이 DNA 양과 비례한다.** 이를 이용하여 DNA의 양을 측정한다.

- 단백질은 aromatic amino acid(방향족 아마노산) 때문에 UV를 흡수하며 **단백질** 존재를 확인하는 흡수파장은 **280nm** 이다. 따라서 **OD 260/280** 값은 추출된 DNA가 얼마나 순수한지를 측정하는 지표로 사용된다. 일반적으로 260/280의 값이 1.8 이상이면 순도가 좋다고 취급한다. **1.8** 이하의 수치일 때는 단백질오염이 의심된다.

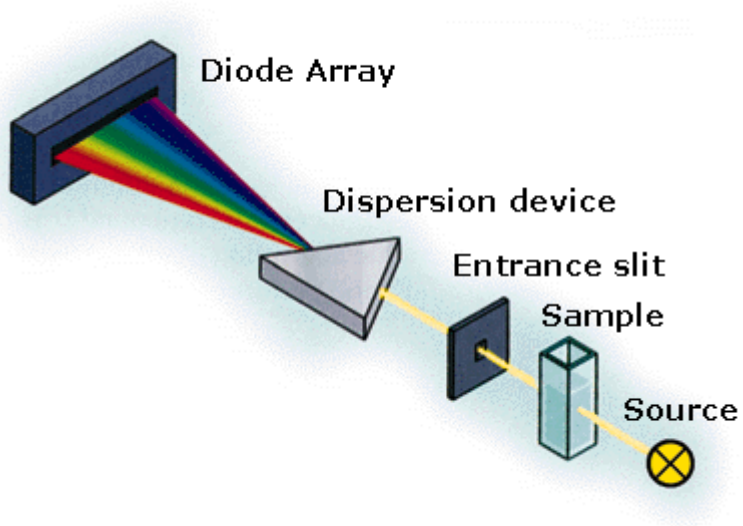
- **OD 230**은 **carbohydrate**나 추출과정에서 남은 **phenol**의 오염 측정을 위한 indicator 이다. **OD 260/230**의 값이 **2.0**이상인 경우 순도가 좋다고 취급한다.

- Spectrophotometer를 이용하기 위해서는 일정부피 (cuvette의 크기에 따라 약 1ml~100ul 정도)이상이 필요하며, OD 값 또한 측정을 위한 유효한 범위가 존재하기 때문에 일반적으로 추출된 DNA는 1/10~1/100정도로 희석하여 측정한다.

- Spectrophotometer는 소량의 DNA의 정량에는 한계가 있다.

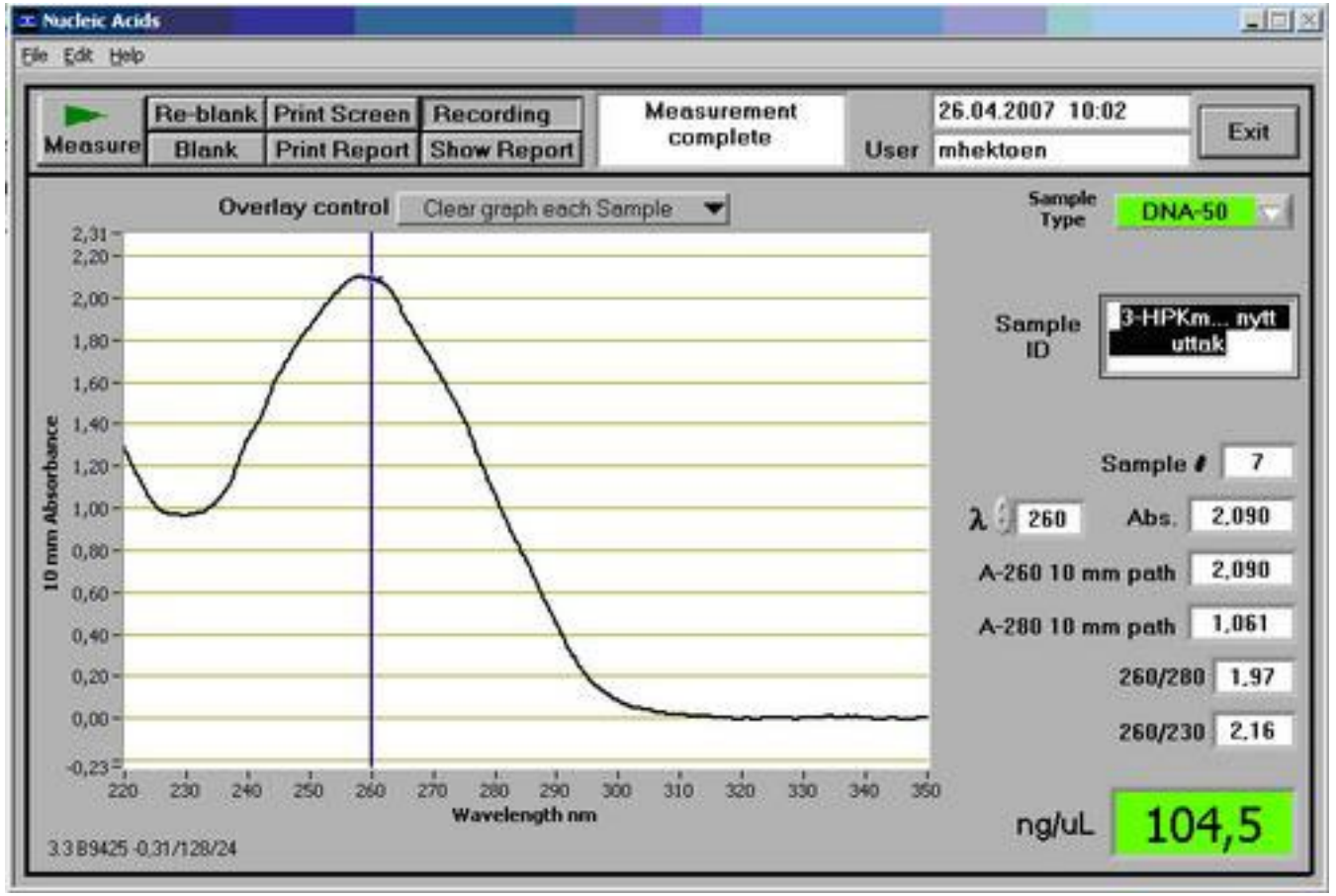
- 최근 개발된 특수한 형태의 spectrophotometer인 **Nanodrop**과 같은 측정장치는 1ul정도로도 정량이 가능하므로 소량의 DNA의 정밀한 정량이 가능해졌다.

Spectrophotometer:
전통적인 방법의 DNA 정량 도구.
측정을 위해 너무 많은 부피의 DNA가 필요하다.



Nanodrop:

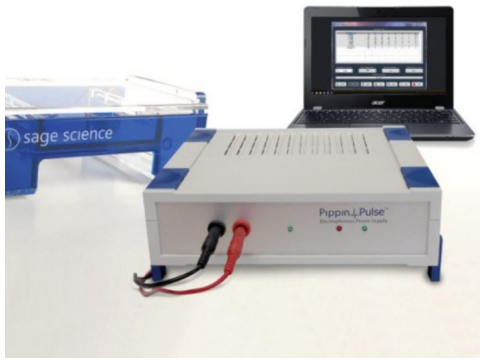
Spectrophotometer의 일종이지만, 1ul 정도의 소량의 DNA도 정밀 측정이 가능하도록 설계되어 DNA를 이용한 분자생물학 실험에 널리 쓰인다.

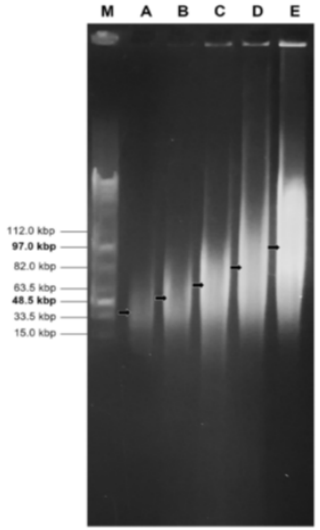


Qubit:
DNA와 결합하는 형광물질을 이용하여 DNA 농도를 측정하는 방법으로 spectrophotometer에 의한 방법 보다 오차가 적다. 이에 따라 차세대염기서열 분석장치(NGS) 실험 등을 위한 DNA 정량에 반드시 필요한 장치이다.

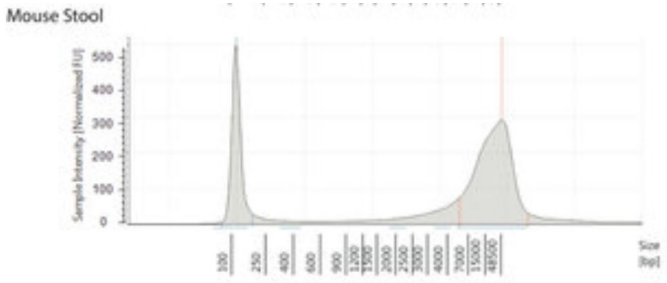
다양한 DNA 길이 측정 도구들: HMW DNA 측정을 위해서는 반드시 Femto Pulse 사용!

Pulse-Field gel electrophoresis

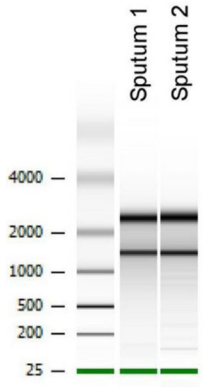


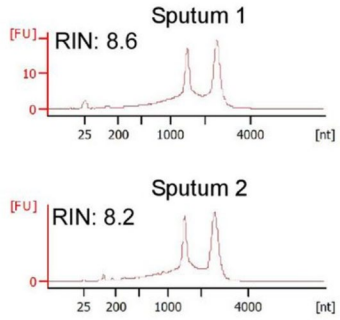


TapeStation

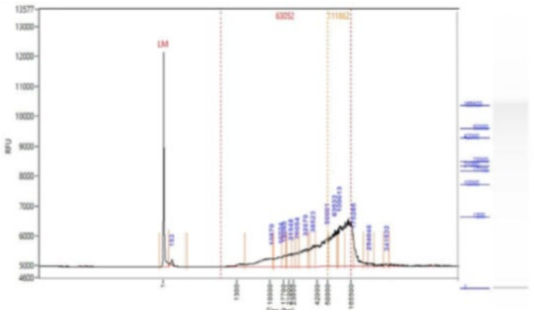


Bioanalyzer





Femto Pulse System



Ⅳ. Nanopore long-read에 의한 mitochondrial genome assembly 이론

8월 5일 오후. 성신여자대학교 최상철

IV. Nanopore Sequencing running

8월 18일 오전, 오후. Oxford Nanopore Technologies 안종화

IIV. Nanopore Sequencing 결과 분석

8월 25일 오전. Oxford Nanopore Technologies 안종화

IIIV. Nanopore long-read에 의한 mitochondrial genome assembly 실습

8월 25일 오후. 성신여자대학교 최상철