알기쉬운 NGS

성문우

The genomes of ill newborns can be sequenced in less than 24 hours to give clinicians a rapid diagnosis.

GENOMICS

Fast sequencing saves newborns

BY SARA REARDON

By two months of age, the boy was near death. He had spent his entire short life in the neonatal intensive care unit (NICU) at Children's Mercy Hospital in Kansas City, Missouri, while physicians tried to work out the cause of his abnormalities. When his liver failed in April 2013, the medical staff warned his parents that the outlook was grim.

Then geneticist Stephen Kingsmore and his team at Children's Mercy took on the case. Within three days, they had sequenced the genomes of the baby and his parents, and identified a rare mutation that was common to the child and both of his parents. The mutation turned out to be linked to a disease in which an overactive immune system damages the liver and spleen. Armed with a diagnosis, the baby's physicians put him on drugs to lower his immune response. The boy is now at home

NIH Studies Explore Promise of Sequencing Babies' Genomes

By **Jocelyn Kaiser** | Sep. 4, 2013 , 2:45 PM

In a few years, all new parents may go home from the hospital with not just a bundle of joy, but with something else—the complete sequence of their baby's DNA. A new research program funded at $25 million over 5 years by the National Institutes of Health (NIH) will explore the promise—and ethical challenges—of sequencing every newborn's genome.

To explore how newborn genomes might be used in medical care, as well as the ethical, legal, and social issues this raises, NICHD and the National Human Genome Research Institute (NHGRI) are funding **four projects**. Two separate teams at the University of California, San Francisco (UCSF), and the University of North Carolina, Chapel Hill, will sequence the exomes of babies, some with known diseases, to see if genetic data can firm up the results of standard newborn screening. A third group at Children's Mercy Hospital in Kansas City, Missouri, will study whether genome sequencing can speed the diagnosis of hundreds of genetic diseases in newborns sick enough to require neonatal intensive care.

차세대염기서열분석

- 차세대염기서열분석 = Next-generation sequencing (NGS)

- NGS의 다른 용어

  >> **대용량**염기서열분석(massively parallel sequencing)

MiSeq (Illumina)

국내 식약처 승인을 통과한 **진단용 차세대염기서열분석장비**

Ion PGM (Thermo Fisher)

크기: 61×51×53 cm
무게: 30 kg

국내 식약처 승인을 통과한 **진단용 차세대염기서열분석장비**

과연 얼마나 많은 검사를 할 수 있을까?

장비	회사	모드	데이터 생산량	소요시간
Ion PGM	Thermo	318 Chip v2, 400 bp	2 Gb	0.3d
Ion S5 XL	Thermo	Ion 540 Chip, 200 bp	15 Gb	0.3d
MiSeq	Illumina	Reagent Kit v3, 2×300bp	15 Gb	2.6d
HiSeq X	Illumina	2×150bp	1.8 Tb	3d

Google

average gene size in human

Web Images Videos News More Search tools

About 114,000,000 results (0.60 seconds)

The genes range in size from 994 to **41.8 kb** for mouse (average length = **7902 bp**, s.d. = 6391) and from 1148 to **37.7 kb** for human (average length = **8446 bp**,s.d. = 7124). The number of exons range from 1 to 41 per/gene (average = 8.4, s.d. = 6.9).

Comparative Analysis of Noncoding Regions of 77 ...
genome.cshlp.org/content/9/9/815.long

Feedback

가정

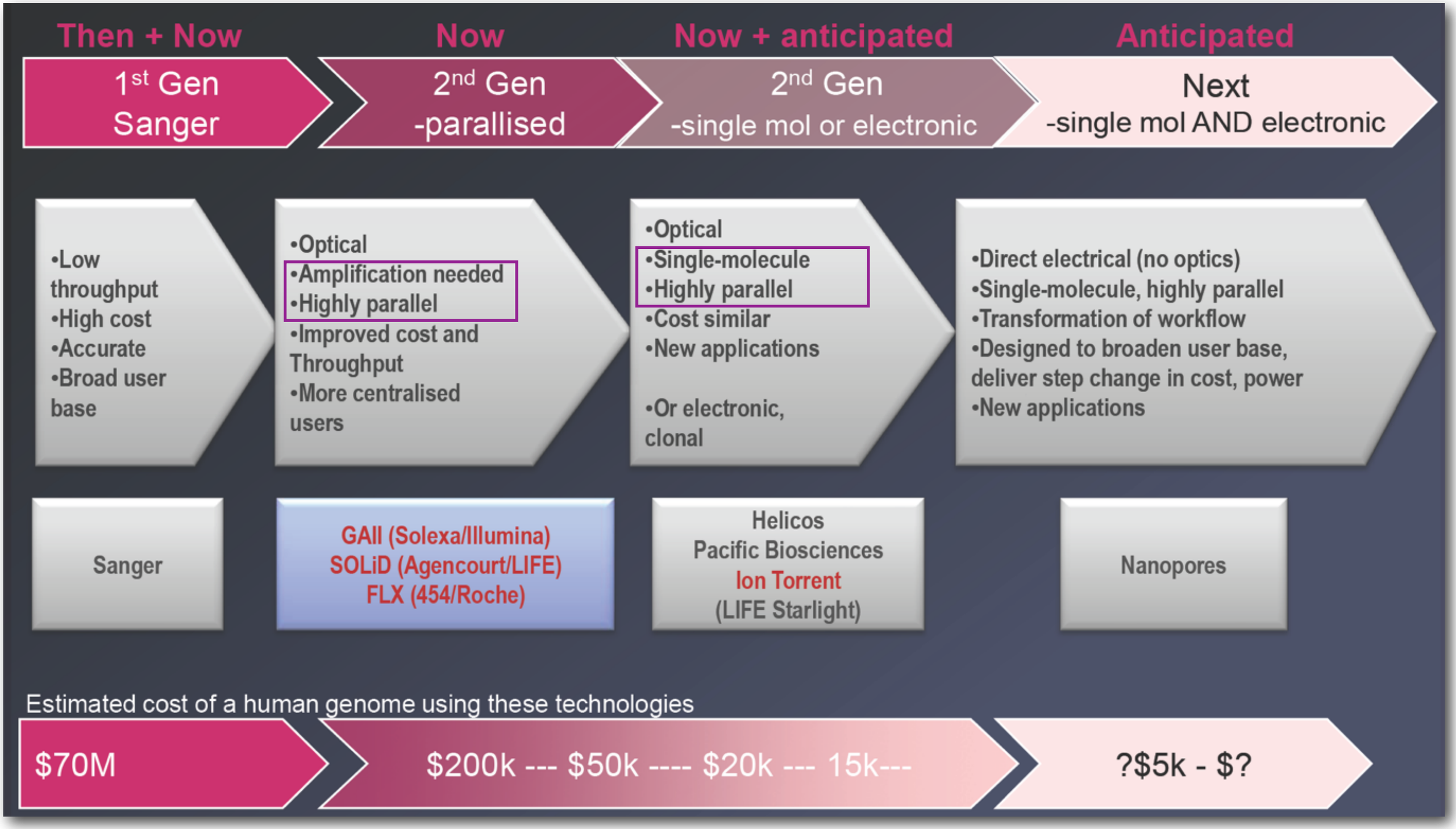
- 유전자 코딩 부위의 평균 크기는? 1.3 Kb (엑손수 9개, 145 bp/엑손)
- 인간 유전체내 엑솜의 총 길이는? 40 Mb (유전자수 30,000 개)
- 인간 유전체의 총 길이는? 3 Gb
- 정확도를 유지하려면 코딩 부위는 최소 100번, 유전체는 최소 30번 읽어야 함

>> 한 유전자를 읽는데 필요한 평균 데이터 양 = 1.3 Kb ×100 = 130 Kb
>> 엑솜(40 Mb)에 필요한 평균 데이터 양 = 40 Mb ×100 = 4 Gb
>> 유전체에 필요한 평균 데이터 양 = 3 Gb ×30 = 90 Gb

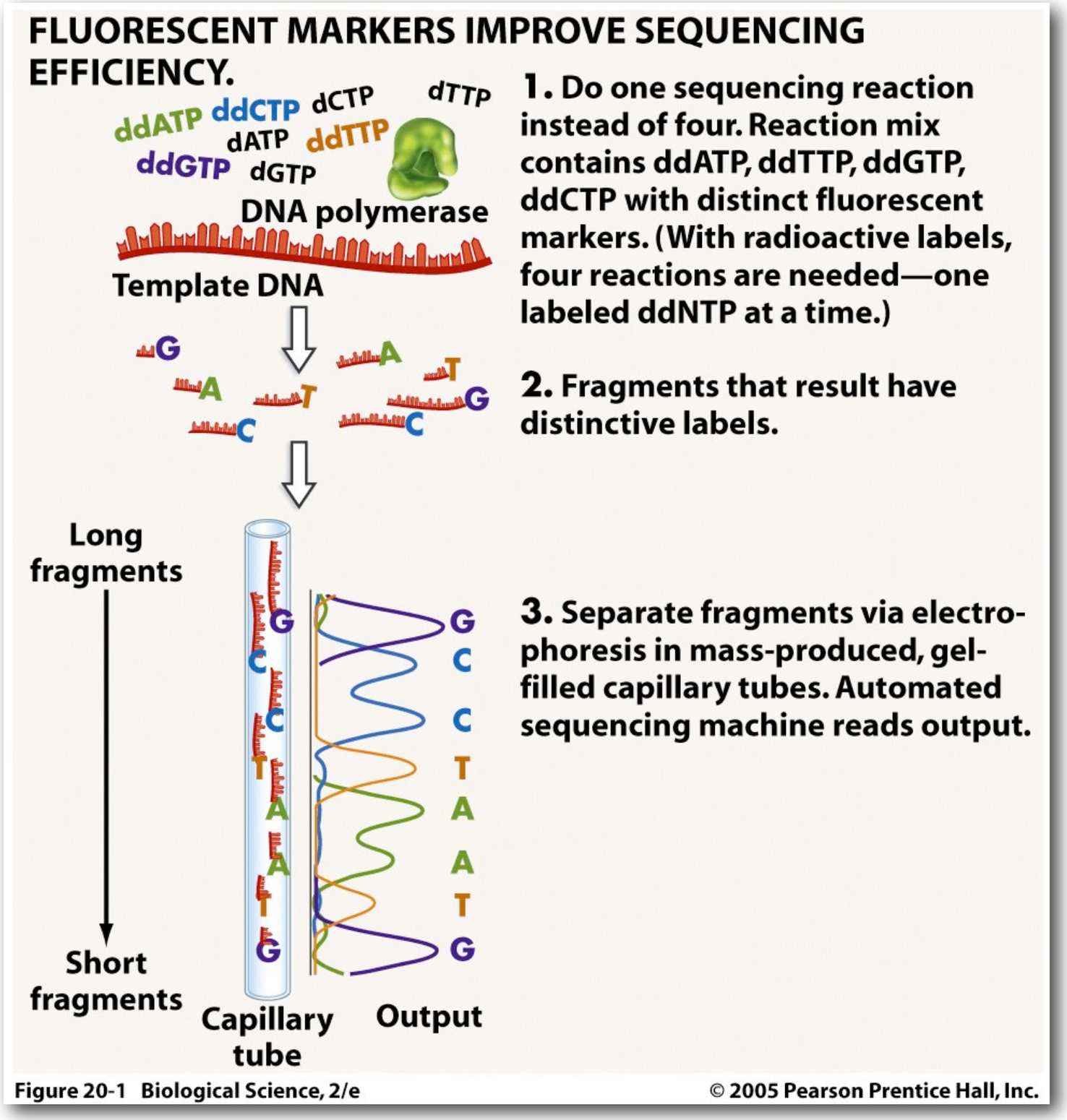
장비	회사	데이터 생산량	유전자 수	엑솜 수	유전체 수	소요시간
Ion PGM*	Thermo	2 Gb	15,000	1/2	1/45	0.3d
Ion S5	Thermo	15 Gb	115,000	3.8	1/6	0.3d
MiSeq*	Illumina	15 Gb	115,000	3.8	1/6	2.6d
HiSeq X**	Illumina	1.8 Tb	14×10^6	450	20	3d

* 진단용 장비
** 연구용 장비

염기서열분석법의 세대


Then + Now
Now
Now + anticipated
Anticipated
1st Gen Sanger
2nd Gen -parallised
2nd Gen -single mol or electronic
Next -single mol AND electronic
•Low throughput •High cost •Accurate •Broad user base
•Optical •Amplification needed •Highly parallel •Improved cost and Throughput •More centralised users
•Optical •Single-molecule •Highly parallel •Cost similar •New applications
•Or electronic, clonal
•Direct electrical (no optics) •Single-molecule, highly parallel •Transformation of workflow •Designed to broaden user base, deliver step change in cost, power •New applications
Sanger
GAII (Solexa/Illumina) SOLiD (Agencourt/LIFE) FLX (454/Roche)
Helicos Pacific Biosciences Ion Torrent (LIFE Starlight)
Nanopores
Estimated cost of a human genome using these technologies
$70M
$200k --- $50k ---- $20k --- 15k---
?$5k - $?

*1*세대 염기서열분석*: Sanger* 법



염기별로 다른 형광 표지, terminator chemistry 사용 및 전기영동 크기에 따른 분획

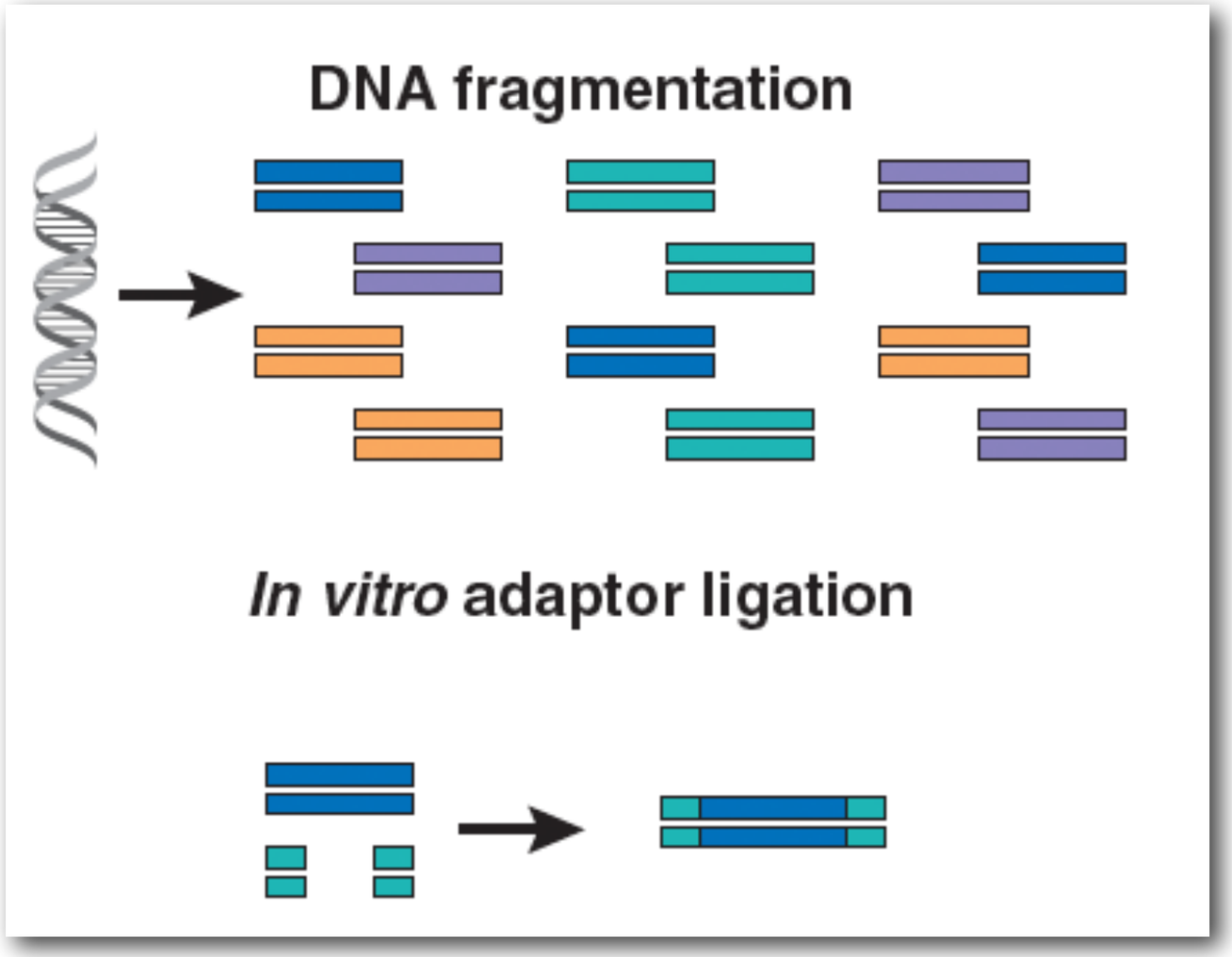
염기서열분석의 단계

1. 검체 준비(sample preparation)

2. 클론 증폭(clonal amplification)

3. 염기서열분석 반응(sequencing reaction)

1. 검체 준비

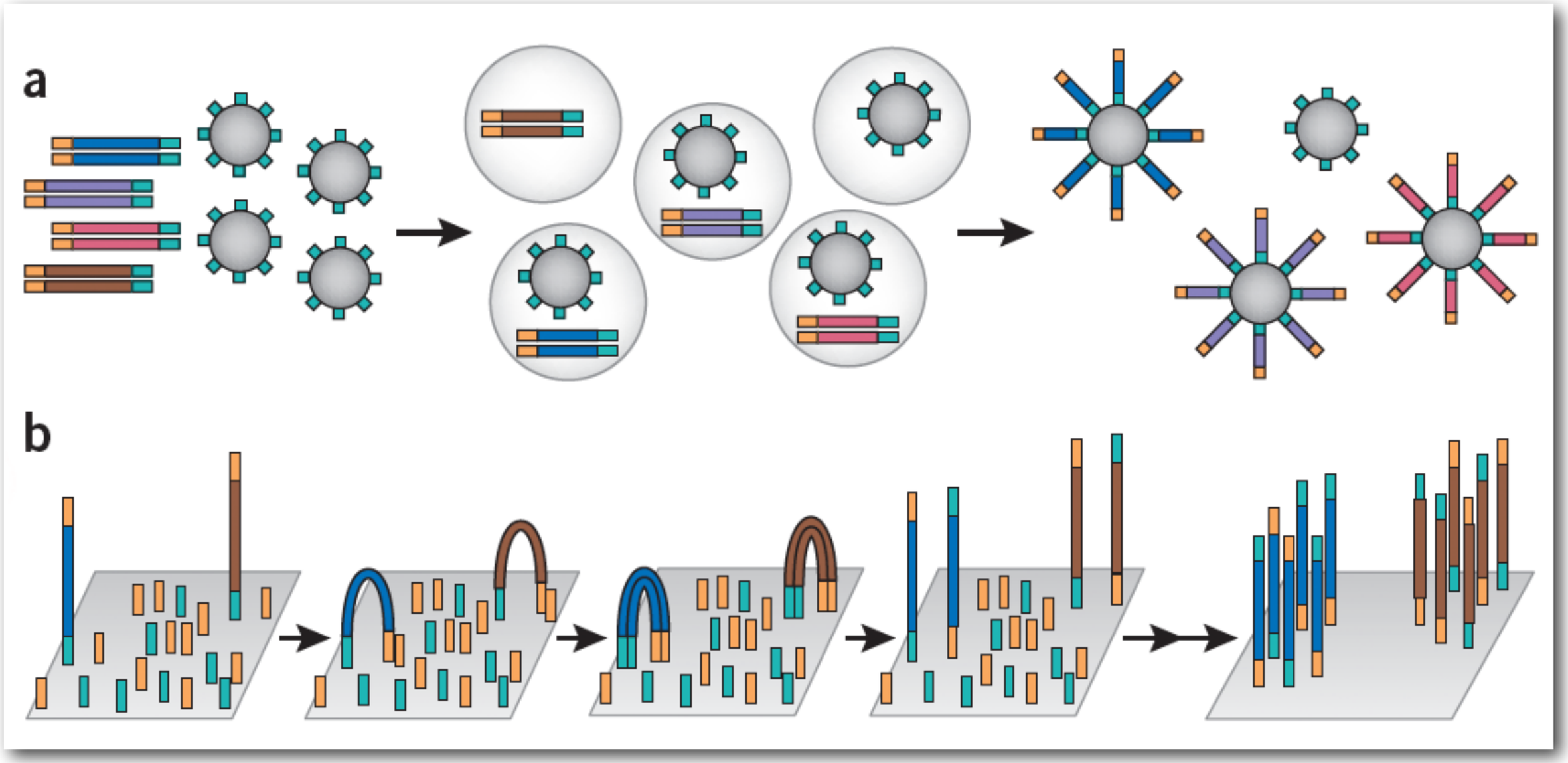


Adaptor: 증폭 및 염기서열분석에 사용되는 공통 primer sequence

1) Whole genome preparation: fragmentation ➔ adaptor ligation
2) Specific target enrichment: exome, gene set, chromosomal regions
    - Capture by probe ➔ adaptor ligation
    - Amplification by PCR ➔ (fragmentation) ➔ adaptor ligation

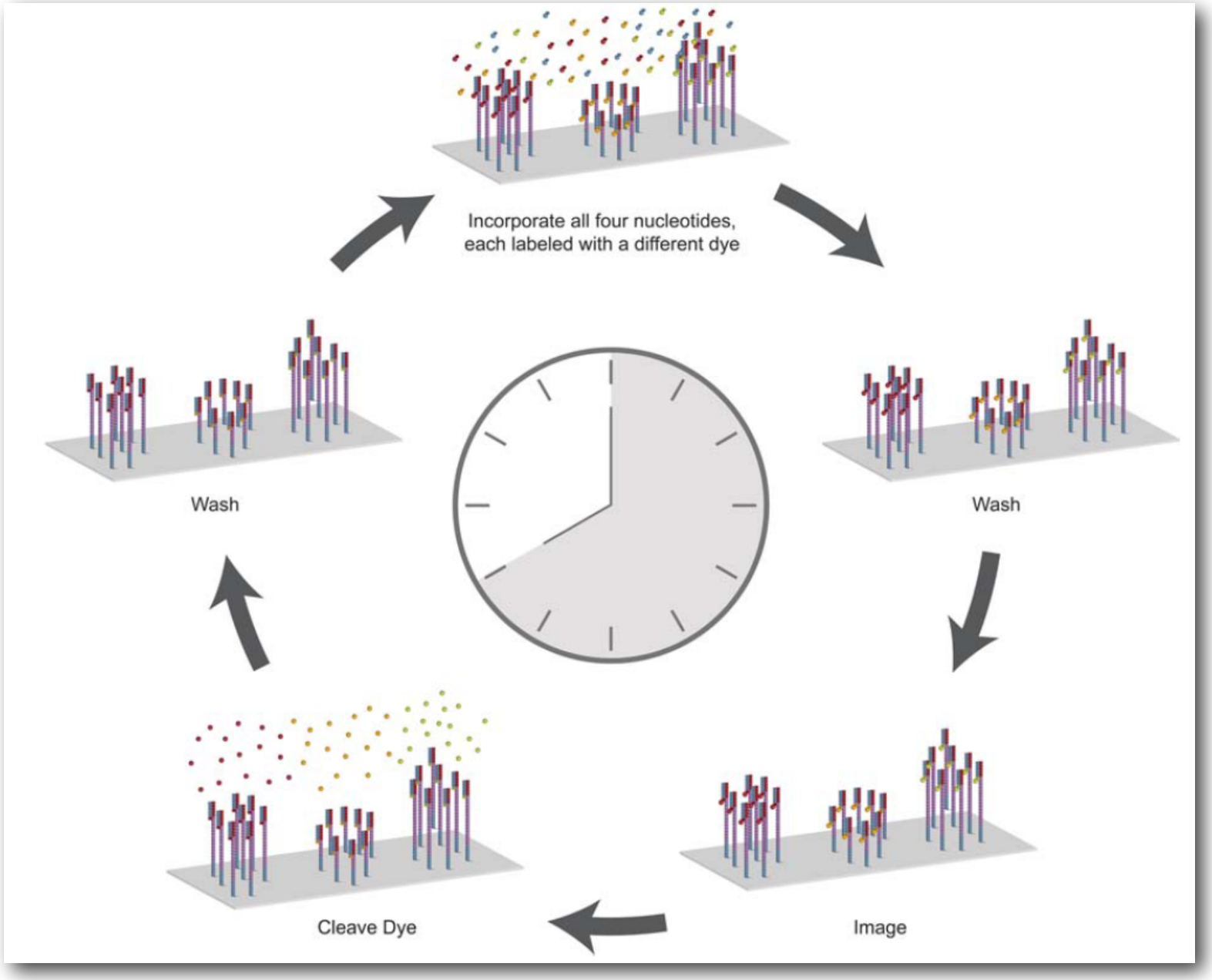
2. 클론 증폭

Emulsion PCR (GS, Roche)



Cluster PCR (GA, Illumina)

3. 염기서열분석 반응



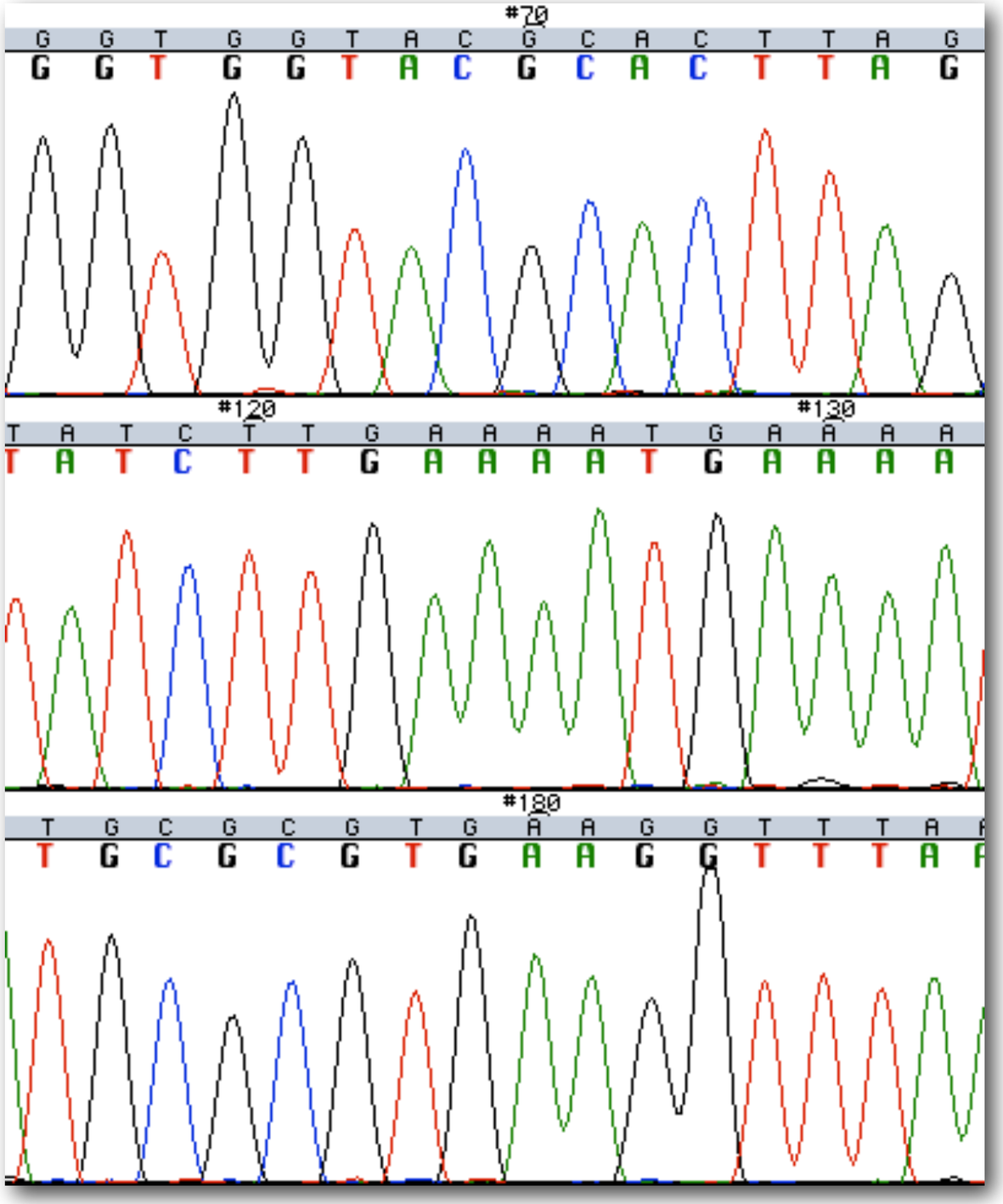
Wash-and-scan sequencing by synthesis

어떤 데이터가 생성되는가?

- 각 절편의 염기서열정보가 텍스트 파일(fastq)로 정리됨
  - 예) 한 줄의 길이가 150 bp 이고, 수백만개-수억개 줄이 있는 텍스트 파일(파일 용량이 수백 Mb~수십 Gb)
  - 분석을 위해 고성능 컴퓨터 및 전문 프로그램 필요

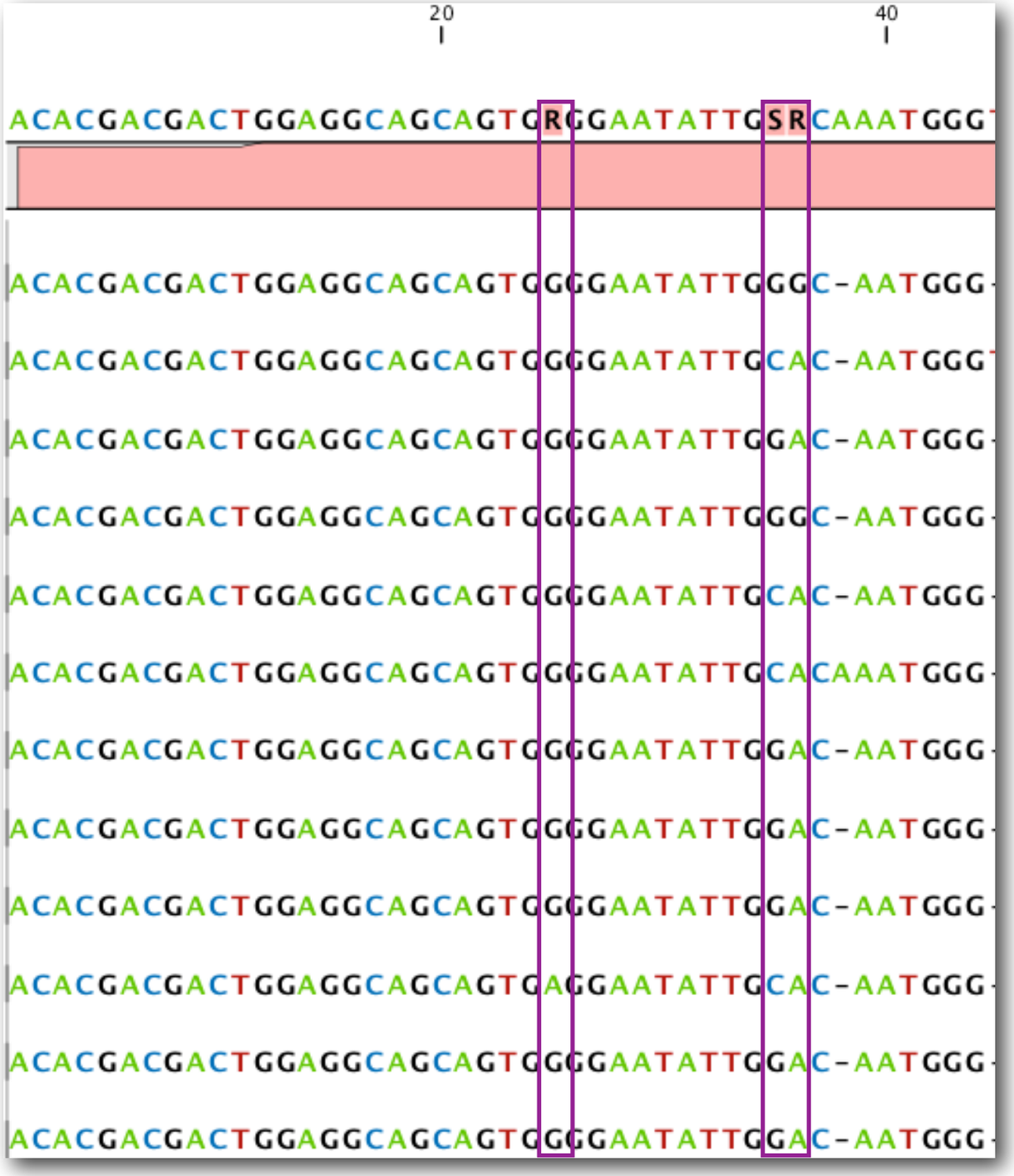
@MG00HS19:301:H9DJTADXX:1:1101:1611:2242 1:N:0:CGTACTAGTATCCTCT
GCACTGTGAAAATGGACTAATACACTTGGAAATGGCTAAGTTCCCTTTTTAAATTGGTTTTCAATTTTCCCTGCTCTTAATACTTTAAAATTCTCCCAAGAAGTTAAGGCTGGGTGAACATTAAACCTTCCAGAAGAACATGCTAAAGGGT
+
CCCFFFFFHHHHHJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJJHJJJJJJIJJJJJIJIJIGHIHHFFHHHFFFFEFEEEEEEDDDDDDDDEEDDDDDDDCDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDCDDDDDB8

*1*세대 염기서열분석



Sum of all reads
(all cells & both parental strands)

*2*세대 염기서열분석



Read by read
(paternal or maternal strand)

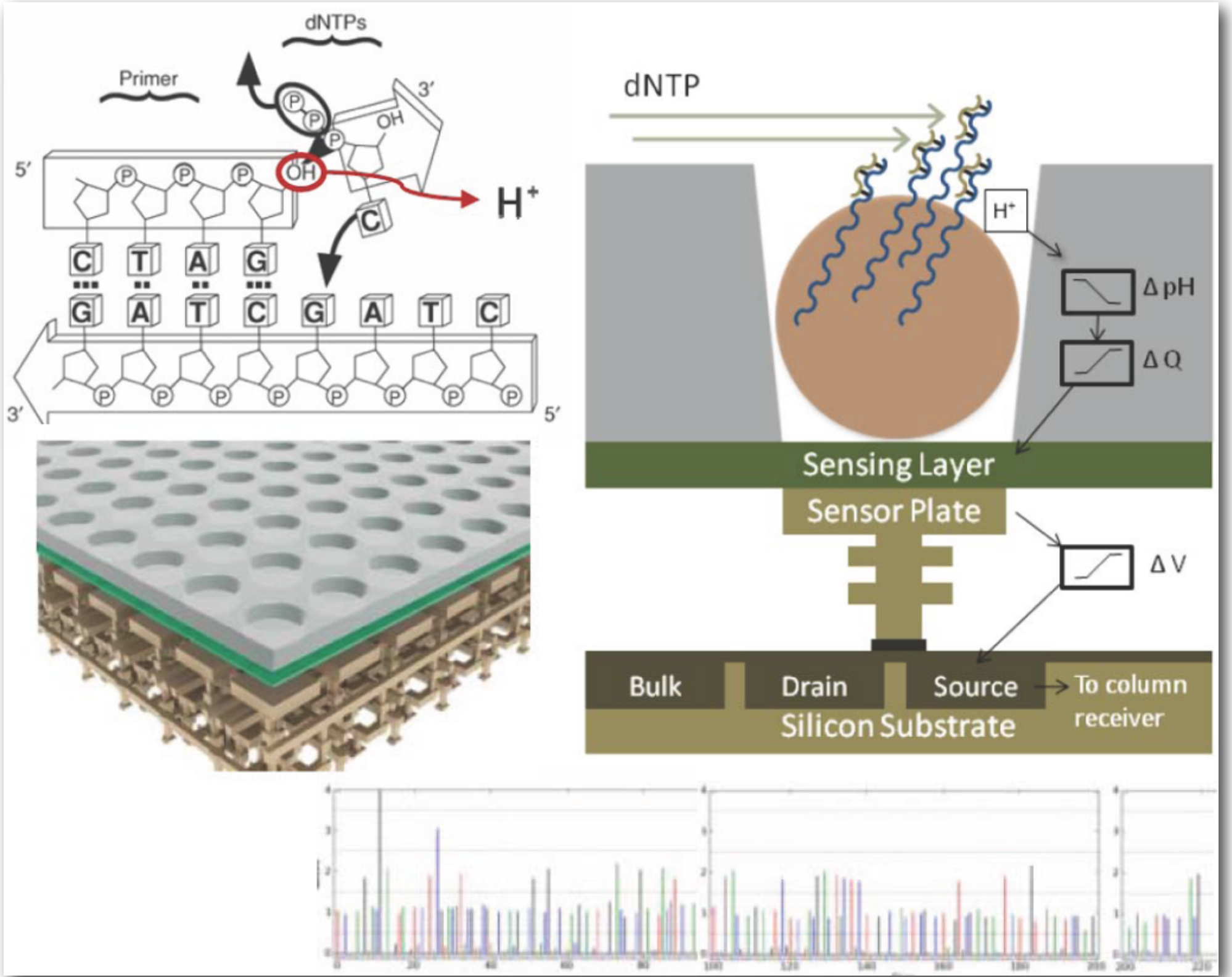
NGS 기반 분자진단검사의 단계

NGS Sequencing	• Library preparation • Clonal amplification • Sequencing reaction	NGS 장비
Analytic Interpretation	• Read alignment • Variant identification • Variant annotation	분석소프트웨어
Clinical Interpretation	• Variant validation by Sanger sequencing • Sanger sequencing for gapped region • Clinical significance of variant	판독

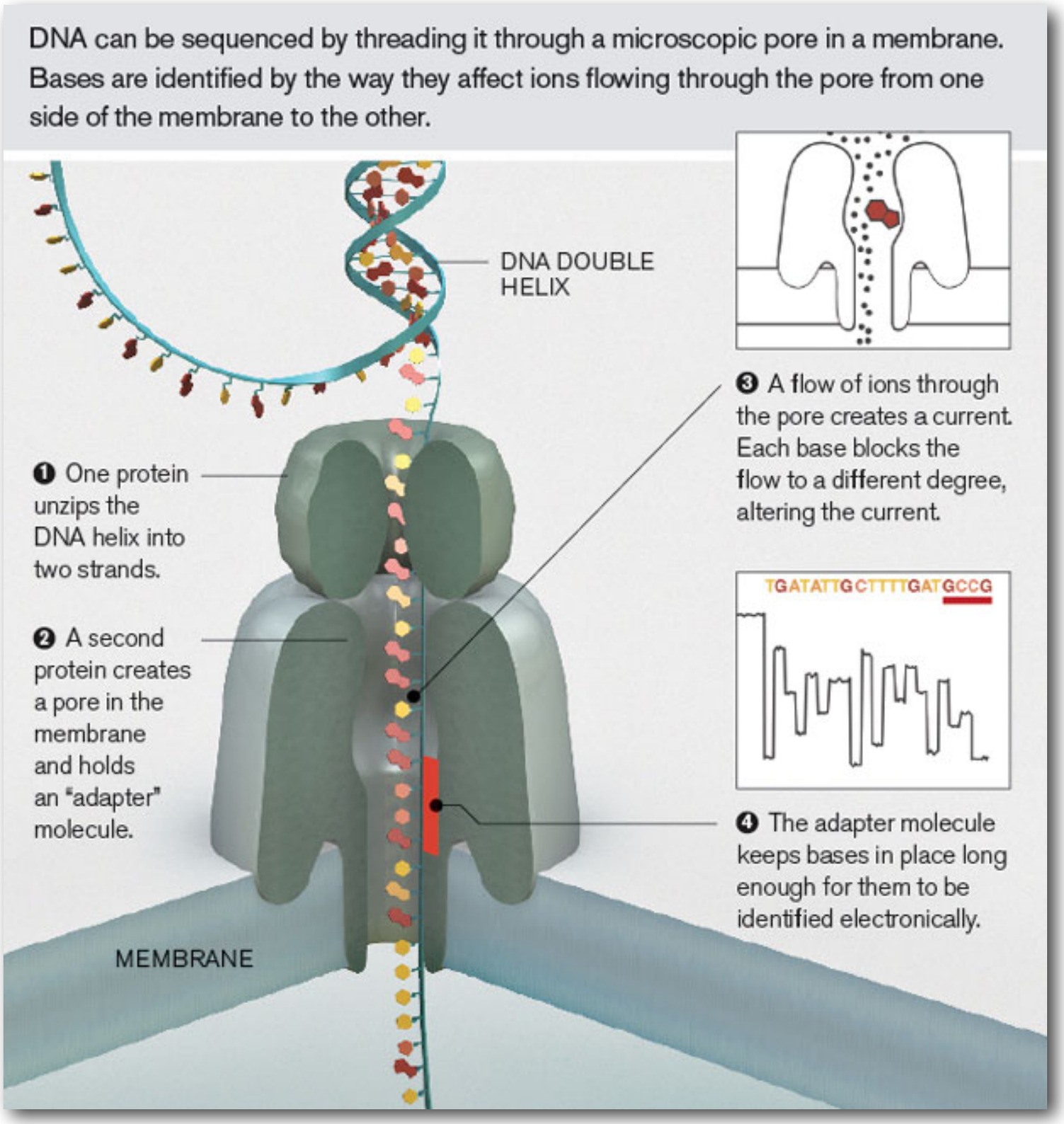
다양한 NGS 장비

- Polony sequencer: Illumina/Solexa
- Dinucleotide sequencer: AB SOLiD
- Pyrosequencer: 454 GS FLX
- Single molecule sequencing
  - Helicos
  - Pacific Bioscience
  - Oxford Nanopore
  - Complete Genomics

*Ion Torrent*의 원리



*Nanopore*의 원리



*Nanopore*의 종류

USB-sized, up to 512 channels

10-20 Gb/48 hrs

144,000 nanopores

NGS 기반의 분자진단검사

- 대상 유전자
  - 질환특이적 유전자패널검사, 엑솜검사, 유전체검사
- 대상 정보
  - 정성 정보: 염기변이 및/또는 구조변이
  - 정량 정보: 예) 태아 염색체수적이상 검출용 비침습적 산전진단검사(NIPT)
- 임상 상황
  - 유전질환(단일유전질환, 복합유전질환) 진단, 선별
  - 혈액암/고형암 진단, 예후, 치료 방침 결정
  - 감염질환의 진단

NGS 패널검사 수가

- 2017년 3월 개시
- 질환유형: 혈액암/고형암/유전질환
- 레벨 I/II
  - 레벨I: 유전자 수 2-30개 또는 target size ≤150kb
  - 레벨II: 유전자수 >30개 또는 target size >150kb
- 수가산정 방법
  - 조건부 선별급여: 본인부담률 50%(입원, 외래 불문)
  - 승인된 요양기관에서 실시한 경우에 인정(위탁은 승인된 실시기관만 가능)
  - 식약처 허가(신고)받은 시약/장비를 사용하지 않는 경우는 수가의 90% 인정

유전질환 패널검사 인증기관

- 총 17개 기관
- 서울/경기: 강남세브란스병원, 고대구로병원, 고대안산병원, 분당서울대병원, 삼성서울병원, 서울대병원, 서울성모병원, 서울아산병원, 순천향대서울병원, 세브란스병원, 아주대병원, 인천성모병원
- 그 외 지역: 부산대병원
- 전문수탁기관: 녹십자의료재단, 랩지노믹스, 서울의과학연구소, 이원의료재단

유전자패널검사의 종류(100 여종: 12개 기관을 대상으로 시행한 설문조사)

Alzheimer's disease	Epilepsy	Marrow failure syndrome
Amyotrophic lateral sclerosis	Eye disorder	Metabolic disorder
Androgenetic alopecia	**Familial thoracic aortic aneurysms**	Metabolic myopathy & channelopathy
Aortopathy	**Fanconi anemia**	Microdeletion syndrome
Arrhythmia	Focal epilepsy	Mitochondria disorder
Ataxia	Glomerulopathy	Moyamoya disease
Autism	Glycogen storage disease	Muscular disorder
Brugada syndrome	Heart disorder	Muscular dystrophy
Cardiomyopathy	Hemolytic anemia	Myofibrillar myopathy
Cerebral cavernous malformations	Hemophagocytic lymphohistiocytosis	Myopathy
Charcot-Marie-Tooth disease	Hereditary anemia	Neonatal cholestasis
Chorea	**Hereditary breast/ovary cancer synd**	Neurodevelopmental disorder
Coagulopathy	**Hereditary cancer syndrome**	Neuropathy
Congenital adrenal hyperplasia	Hereditary endocrine disorder	Osteogenesis imperfecta
Congenital cholestasis	**Hereditary hearing loss**	**Parkinson disease**
Congenital diarrhea	**Hereditary pheochromocytoma synd**	Pediatric retinal vascular disease
Congenital glaucoma	**Hereditary spastic paraplegia**	Polycystic kidney
Congenital muscular dystrophy	Hirschsprung's disease	Proportionate short stature
Congenital myopathy	Hypertrophic cardiomyopathy	Pulmonary hypertension
Connective tissue disorder	Hypogonadism	**Rasopathy**
Corneal dysplasia	Hypothyroidism	Retinal dystrophy
Craniosynostosis	Immune deficiency	**Retinitis pigmentosa**
Cytopenia	Inborn error of metabolism	Sexual development
Dermatology	Kallman syndrome	Short QT syndrome
Dilated cardiomyopathy	Leukoencephaly	Short stature
Dominant optic atrophy (DOA)	Leukopenia	**Skeletal dysplasia**
Dyskinesia	Limb girdle muscular dystrophy	Skin disorder
Dyskinesia dystonia paralysis	Long QT syndrome	Spondyloepimetaphyseal dysplasia
Dystonia	Lysosomal storage disease	Thrombocytopenia
Early onset epilepsy	Macular dystrophy	Thrombotic microangiopathy
Ehleres-Danlos syndrome	Malformation of cortical development	Very–Early-Onset Inflammatory Bowel ds

서울대학교병원, NGS 패널검사

Acid-base-electro	28	Glycogen storage disease	10
Amyotrophic lateral sclerosis	27	Hemic treatment	76
Androgenetic alopecia	30	Hereditary hearing loss	143
Arrythmia	32	Hereditary paraganglioma-pheochromocytoma	10
Ataxia	27	Hereditary spastic paraplegia	18
Ca-Pi-Mg	20	Hirschsprung's disease	13
Cerebral cavernous malformations	3	Hypertrophic cardiomyopathy	36
Charcot Marie Tooth (CMT) disease	31	Hypogonadism	10
Chorea	25	Limb girdle muscular dystrophy	43
Congenital diarrhea	21	Macular dystrophy	32
Congenital muscular dystrophy	30	Malformation of cortical development	28
Congenital myopathy	30	Metabolic myopathy & channelopathy	29
Craniosynostosis	25	Moyamoya disease	6
Dilated cardiomyopathy	49	Myofibrillar myopathy	30
Dominant optic atrophy (DOA)	3	Neonatal cholestasis	34
Dyskinesia-dystonia-paralysis	28	Noonan syndrome/Rasopathy	19
Dystonia	29	Osteogenesis imperfecta	24
Early onset epilepsy	25	Parkinson	22
EDS, Classic type	2	Pediatric retinal vascular disease	6
Familial thoracic aortic aneurysms/dissections	25	Retinitis pigmentosa	41
Fanconi anemia	16	Spondyloepi/metaphyseal dysplasia	31
Focal epilepsy	22	Thrombotic microangiopathy	15
Glomerulopathy	23	Very–Early-Onset inflammatory bowel disease	43

근육병 유전자패널

- 유전자패널
  - 지대형 근이형성증 유전자패널, 43종 유전자
  - 선천성 근이형성증 유전자패널, 30종 유전자
  - 선천성 근육병증 유전자패널, 36종 유전자
  - 대사성 근육병증 유전자패널, 29종 유전자
  - 근섬유 근육병 유전자패널, 36종 유전자

NGS 기반 분자진단검사의 보고서 예제

관리번호	01-001-236
검사종목	Limb-Girdle Muscular Dystrophy

Sequence Analysis

Gene	NT change	AA change	Zygosity	Clinical Significance
DYSF	c.1284+2T>C		Heterozygote	Pathogenic
DYSF	c.2086delC	p.Leu696*	Heterozygote	Pathogenic
TTN	c.61490T>C	p.Leu20497Ser	Heterozygote	Uncertain Significance
TTN	c.54293T>C	p.Ile18098Thr	Heterozygote	Uncertain Significance

결과해석 및 권장사항

1. 본 환자의 *DYSF* 유전자에서 기보고된 pathogenic variant인 c.1284+2T>C, heterozygote 및 기보고가 없는 pathogenic variant인 c.2086delC, p.Leu696* 가 검출되었습니다.
2. c.1284+2T>C, heterozygote는 Miyoshi myopathy 환자에게서 수차례 기보고되었습니다 (PMID: 12767493, 27363342, 25591676)

검사정보

1. Method: Sequence analysis
2. Gene: 31 genes (ANO5, CAPN3, CAV3, DAG1, DES, DNAJB6, DYSF, FHL1, FKRP, FKTN, GAA, GMPPB, HNRPDL, ISPD, ITGA7, LIMS2, LMNA, MYOT, PLEC, POMGNT1, POMT1, POMT2, SGCA, SGCB, SGCD, SGCG, TCAP, TNPO3, TRAPPC11, TRIM32, TTN)
3. Analysis region: Entire coding exons and their flanking regions of the genes described above
4. Reference sequence: Please refer to the table below
5. Nucleotide numbering: c.1 is the "A" of the initiation codon
6. Amino acid numbering: p.1 is the initiation codon, ATG
7. Mutation database: http://www.hgmd.cf.ac.uk/

혈액암치료 유전자패널

- 목적
  - 진단검사의학과/혈액종양내과/삼성SDS 공동 개발
  - 재발 혈액암 환자에서 맞춤치료를 위한 유전자 변이별 치료약제에 대한 정보 제공
- 대상유전자
  - 76종 유전자(SNV/indel, CNV, translocation, ITD)
- 현황
  - 2017년 6월 개시
  - 총 50 건 의뢰

HEMATOLOGIC CANCER PANEL ANALYSIS REPORT

SNUH 서울대학교병원

Medical Record #	----------	**Specimen #**	G1703572_S3	**Clinical Information**	AML
Report #	201708211323-800	**Specimen Type**	NA	**Report Date**	2017-08-21-13:23

▣ Test Summary

Hematologic Cancer Panel is an end-to-end test that comprises of an assay based on Next Generation Sequencing (NGS) and bioinformatics analysis that detects and provides clinical evidences on genomic alterations from a patient's genome.

▣ Result of Gene Mutation Analysis and Therapeutic Implications

Gene	Alteration (Evidence)	Therapies (Patient tumor type)	Therapies (Other tumor type)	Potential Clinical Trial	Potential Prognostic Outcome
DNMT3A	* p.Arg882His (48.90% (87/178))	Decitabine* (B)	none	none	Poor (B)
NPM1	E11 p.286Leu_287Trpfs (47.30% (69/146))	(-) All-trans Retinoic Acid (B) Daunorubicin* (B) Valproic acid* (B)	none	none	Better (B)
NRAS	* p.Gly12Cys (45.00% (224/498))	none	(-) Anti-EGFR Antibody (B) COL	Binimetinib (B) MEL Docetaxel+Trametinib (B) MEL RO4987655 (B) MEL Sorafenib+Tivantinib (B) MEL Radioiodine+Selumetinib (C) THY Ribociclib (C) MEL	none

(-) Patient may show resistance to the drug
* mark in the therapy column indicates KFDA approved drug
* mark in the alteration column indicates non-specific target by the drug
+ indicates combination therapy.

Clinical Whole-Exome Sequencing for the Diagnosis of Mendelian Disorders

BACKGROUND

Whole-exome sequencing is a diagnostic approach for the identification of molecular defects in patients with suspected genetic disorders.

METHODS

We developed technical, bioinformatic, interpretive, and validation pipelines for whole-exome sequencing in a certified clinical laboratory to identify sequence variants underlying disease phenotypes in patients.

RESULTS

We present data on the first 250 probands for whom referring physicians ordered whole-exome sequencing. Patients presented with a range of phenotypes suggesting potential genetic causes. Approximately 80% were children with neurologic phenotypes. Insurance coverage was similar to that for established genetic tests. We identified 86 mutated alleles that were highly likely to be causative in 62 of the 250 patients, achieving a 25% molecular diagnostic rate (95% confidence interval, 20 to 31). Among the 62 patients, 33 had autosomal dominant disease, 16 had autosomal recessive disease, and 9 had X-linked disease. A total of 4 probands received two nonoverlapping molecular diagnoses, which potentially challenged the clinical diagnosis that had been made on the basis of history and physical examination. A total of 83% of the autosomal dominant mutant alleles and 40% of the X-linked mutant alleles occurred de novo. Recurrent clinical phenotypes occurred in patients with mutations that were highly likely to be causative in the same genes and in different genes responsible for genetically heterogeneous disorders.

CONCLUSIONS

Whole-exome sequencing identified the underlying genetic defect in 25% of consecutive patients referred for evaluation of a possible genetic condition. (Funded by the National Human Genome Research Institute.)

From the Departments of Molecular and Human Genetics (Y.Y., A.W., P.A.W., A.B., J.B., F.X., Z.N., M.H., R.P., M.R.B., M.S.L., A.K., J.S., S.E.P., J.R.L., A.L.B., C.M.E.) and Pediatrics (S.E.P., J.R.L.) and the Human Genome Sequencing Center (D.M.M., J.G.R., M.N.B., P.P., M.W., Y.D., J.R.L., R.A.G.), Baylor College of Medicine, Houston. Address reprint requests to Dr. Eng at the Department of Molecular and Human Genetics, NAB 2015, Baylor College of Medicine, Houston, TX 77030, or at ceng@bcm.edu.

N Engl J Med 2013;369:1502-11.
DOI: 10.1056/NEJMoa1306555

국제진료센터: 엑솜검사

- 대상 및 목적
  - 유전질환이 의심되는 환자 및 부모
  - 원인 유전자/변이 진단, 진료 방침 및 유전 상담
- 수가
  - singleton WES 300, trio WES 600
- 현황
  - 2017년 7월 개시
  - 총 5건 의뢰 및 검사 진행

비침습적 산전진단검사

- 태아조직(양수, 융모막) 대신 산모혈액 이용
  - 산모혈액에 태아유리핵산(cell-free fetal DNA)이 약 10% 존재
- 태아 염색체의 수적 이상
  - Trisomy 13 (파타우증후군), 18 (에드워드증후군), 21 (다운증후군) 및 성염색체 수적 이상
    - 산모혈청검사 및 초음파 대비 높은 검출율: 81-96% vs. 98-99% (Trisomy 13)
- 멘델유전질환 등으로 적용 범위 확대 예상

진단목적의 검사와 연구목적의 분석

- 양성 결과뿐 아니라 음성 결과의 임상적 의미 해석
- 분석민감도/특이도 및 임상민감도/특이도, 제한점
  - 해당 검사에서 어떤 유전자 및 부위가 누락되었는가?
  - 어떤 유형의 변이를 놓칠 수 있는가?
  - 위양성 및 위음성율은?
  - 재현 가능한가?

진단검사시 고려사항

- 검사 평가
  - 임상 적용 전 평가 필요: 민감도, 특이도, 재현성
  - 다양한 변이 유형 및 검체 유형에 대해 평가 필요
- 최적화 및 검사 표준 프로토콜 확립
  - 검사 조건 및 환경에 대한 최적화
  - 검체의 질, 질관리, 분석파이프라인, 데이터 저장 및 소급성, 표준물질 및 신빙도평가에 대한 표준 프로토콜

진단검사시 고려사항

- 보고 표준 지침
  - 검사소요시간: 임상적으로 적절해야 함
  - 결과 해석: 근거 기반의 평가
  - 대상 유전자 외에서 발견된 이상 소견의 보고
  - 보고서 포함 내용: 변이 목록, 해석, 변이 분류의 근거, 검사 유전자, 유전자가 실제 읽힌 횟수(coverage), 해당 검사법의 진단율, 데이터 처리 과정

Test development

Sample prep + target enrichment * → Sequence → Analysis

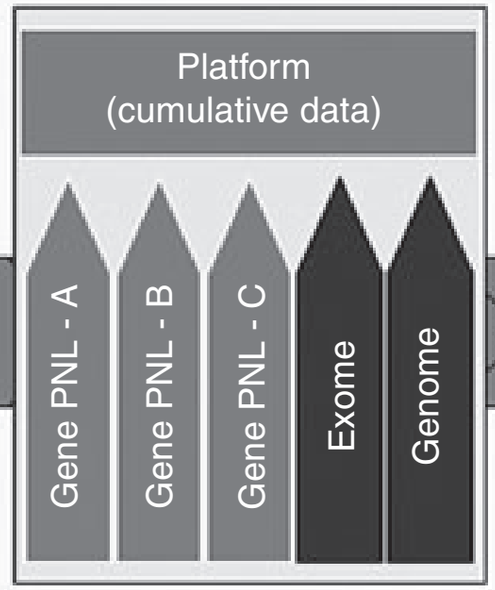
Optimize components and entire test

- General assay conditions
- Coverage
- Sample pooling
- Analysis setting/thresholds

Establish protocol for entire workflow

* Target enrichment applies to targeted assays only (gene panels and whole exome)

Test validation



Validate entire test using set conditions

- Sensitivity
- Specificity
- Robustness
- Reproducibility

Include samples representing

- Different variant types (in/dels, substitutions, CNVs)
- Disease-specific aspects (common pathogenic variants)

Quality management

QC PT

Patient testing

Quality control (QC)

- Every run

Proficiency testing (PT)

- Periodically

ACMG clinical laboratory standards for NGS. Genet Med 2013;15:733-47

NGS 기반 분자진단검사의 미래는?

- 유전자패널
  - 특정 질환용 유전자패널: 유전비후성심근병증 패널, 긴QT증후군 패널
  - 광의의 유전자패널: 심장질환 유전자패널, 근골격질환 유전자패널
- 엑솜검사
- 유전체검사
- DNA 염기서열 외 유전 요인: 후성유전체, 전사체 등