생명정보개론 제3강

Week03 - 차세대 염기서열 해독법

2018.01.26

0. Nic Volker 이야기

The First Child Saved By DNA Sequencing

Matthew Herper, FORBES STAFF ✔
I cover science and medicine, and believe this is biology's century. **FULL BIO** ∨

Since he was a toddler, six-year-old Nicholas Volker's intestine had been dangerously inflamed, necessitating a hundred surgeries including the removal of his colon. No one knew the cause, but it seemed certain that the boy was dying. In a desperate attempt to figure out what was wrong, doctors at the Medical College of Wisconsin did something desperate and unproven: they sequenced his DNA.

- 면역 결핍이 있음

- 확인이 불가한 심각한 질병임

- 기존치료법에 장기적 위험성이 있음

- 적절한 면역 재건 치료가 논의되었지만 매우 위험

치료의 성공은 근본적인 mechanism의 파악에 달려있음
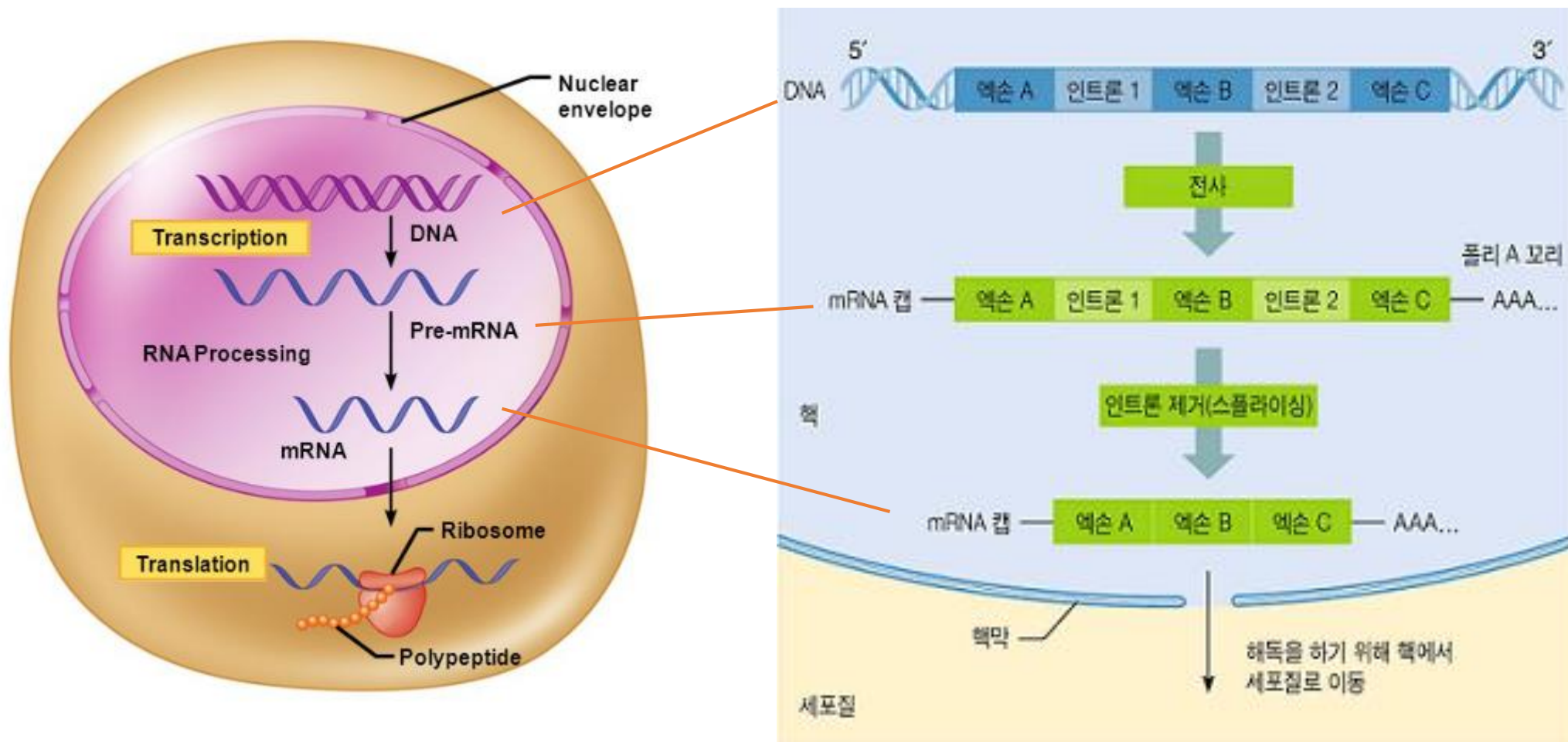
0. Nic Volker 이야기
Nic Volker논문 Introduction

Single gene testing대신

<u>Exome</u> sequencing을 통해

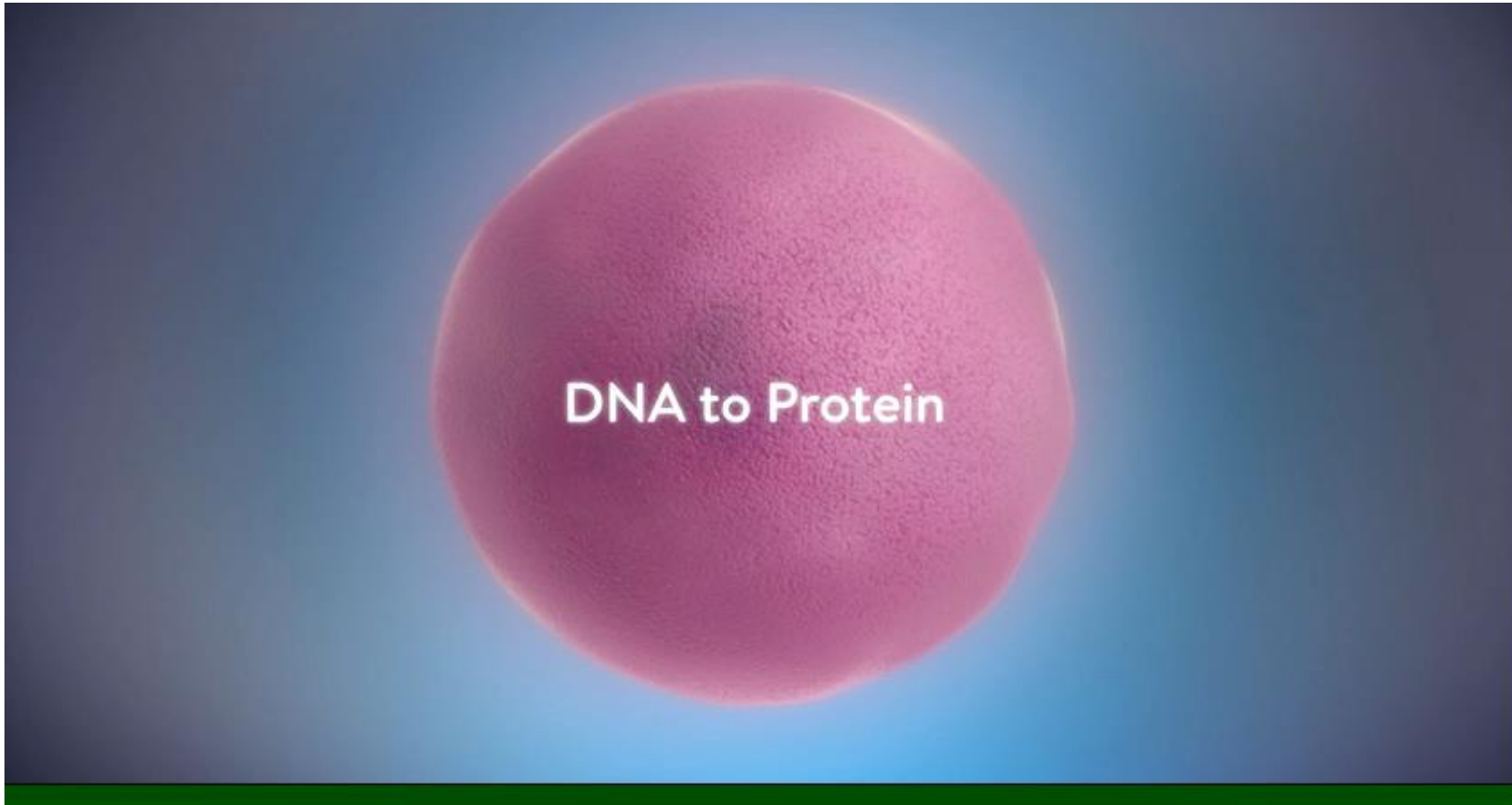원인이 되는 돌연변이 동정이 필요

Exon의 총 집합체 (-ome)
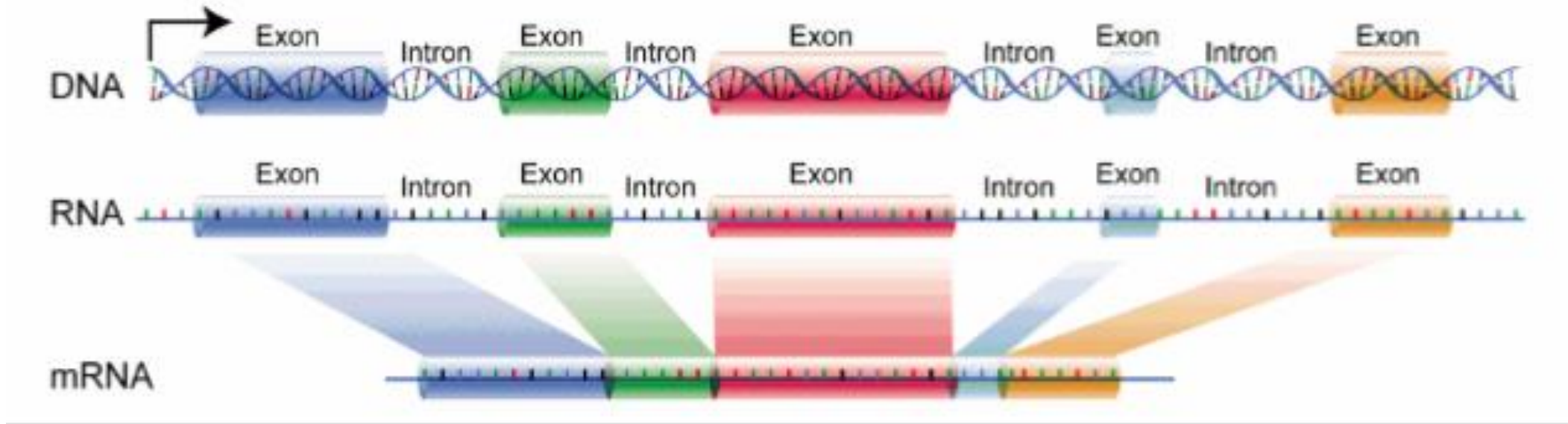Proteome, Genome

1. 엑솜시퀀싱의 개념 들어가기 전 기초설명

1. 엑솜시퀀싱의 개념

DNA to Protein 비디오

1. 엑솜시퀀싱의 개념

· Intron이 훨씬 길다 (1000bp : 140bp)

· Coding gene 약 20,000개 ·

· Average Gene Exon 개수 : 약 10개

· Exome 개수 : 약 200,000개
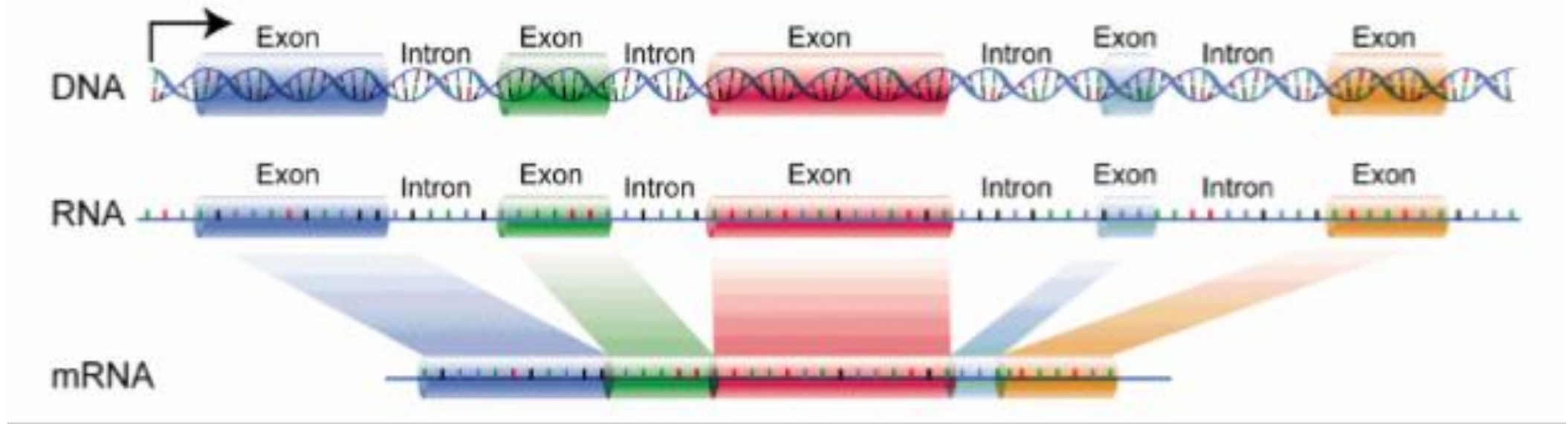
Whole Genome Sequencing

vs

Exome Sequencing

-> Exon만 모으면 Compact한 데이터량 / 비용절감

-> 특히 엑솜 영역에서도 아미노산 변화가 단백질 기능에 영향을 미치는 non-synonymous한 변이에 초점을 맞출 수 있음.
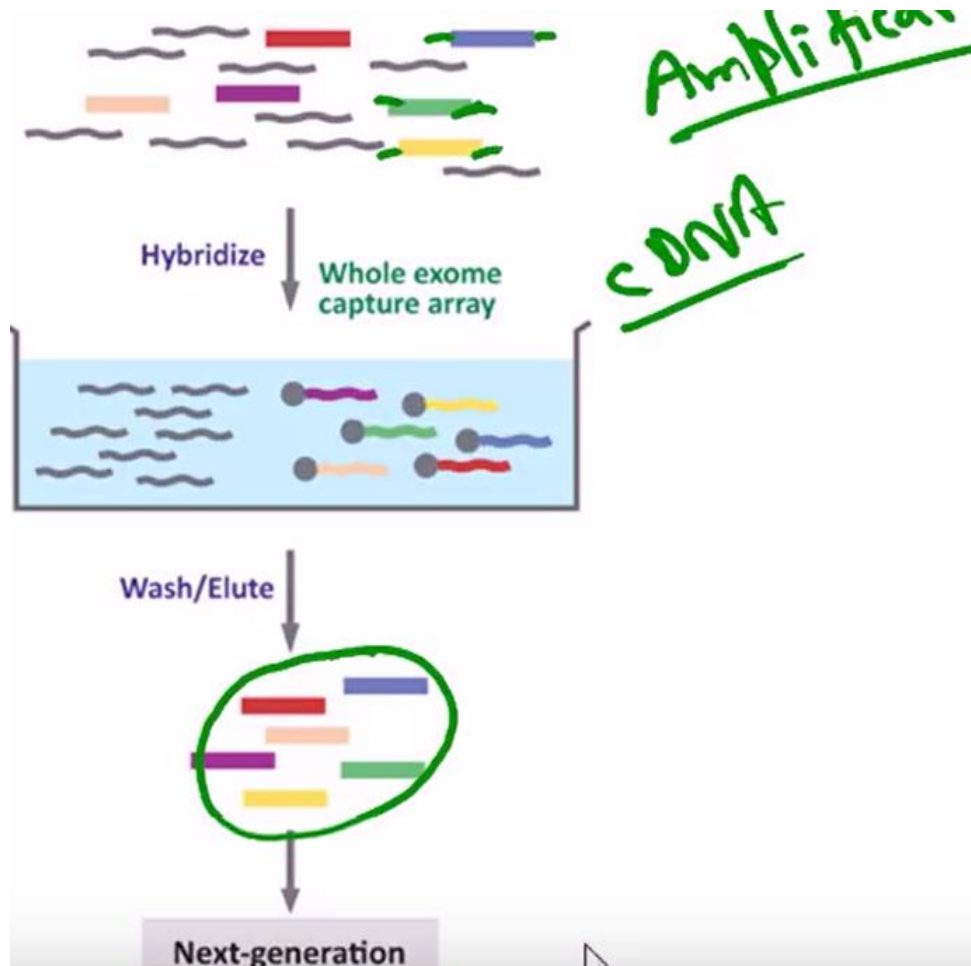
1. 엑솜시퀀싱의 개념
Exome Capture Process

1. 어디가 Exon인지 아직 몰라! DNA를 무작위로 듬성듬성 자른다.

1. 엑솜시퀀싱의 개념
Exome Capture Process

2) 잘라진 Nucleotide중 Exon만 스크리닝 하기 위해 c DNA Library를 이용한다.

3) Exon만 Capture 후 붙지 않는 Arrary 들은 씻어 목표하는 Array만 남게 한다.
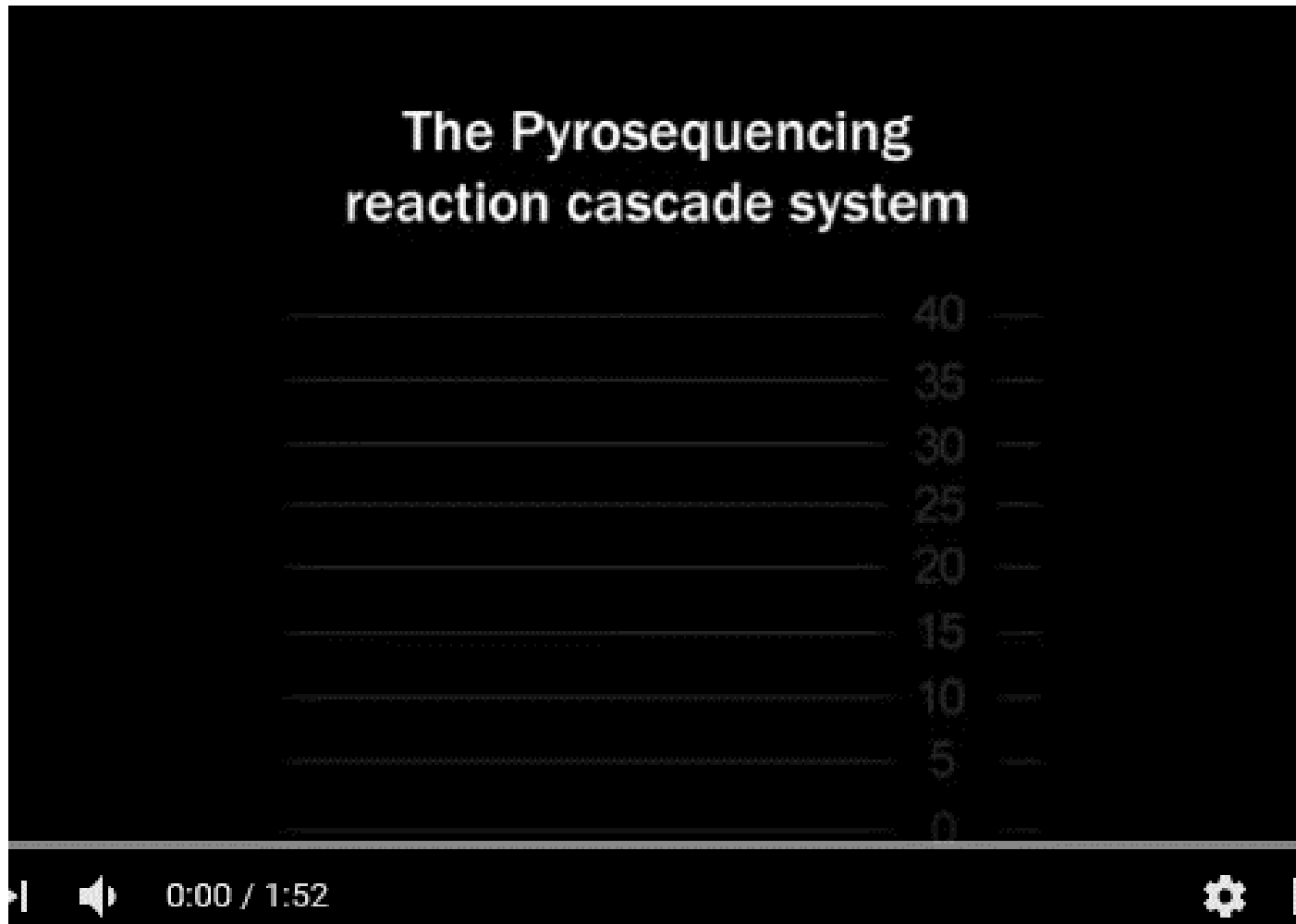
4) NGS로 exome의 서열 밝힘

1. 엑솜시퀀싱의 개념
Next Generation Sequencong Roche 454 GS FLX

1. 엑솜시퀀싱의 개념
Pyrosequencing 원리

2. Nic은 어떤 절차를 밟았을까?
Nic Volker논문 Introduction

Single gene testing대신

Exome sequencing을 통해

원인이 되는 돌연변이 동정이 필요

Exon의 총 집합체 (-ome)
Proteome, Genome

2. Nic은 어떤 절차를 밟았을까?
Nic Volker논문 Material/Method

· **Exome Capture/서열 해독**

- 닉의 genomic DNA 5µg서 엑손캡쳐

- 180,000 exons, 700 miRNA exons을 타깃으로

- 탐지가 가능한 양으로 증폭시킨 captured exon

  조각을 Roche 454 GS FLX로 서열해독

2. Nic은 어떤 절차를 밟았을까?

Nic Volker논문 Material/Method

- Roche 454 gsMapper software tool

· **Exome analysis**

- **Mapping** : 해독된 DNA read를 **human genome reference sequence** 에 정렬
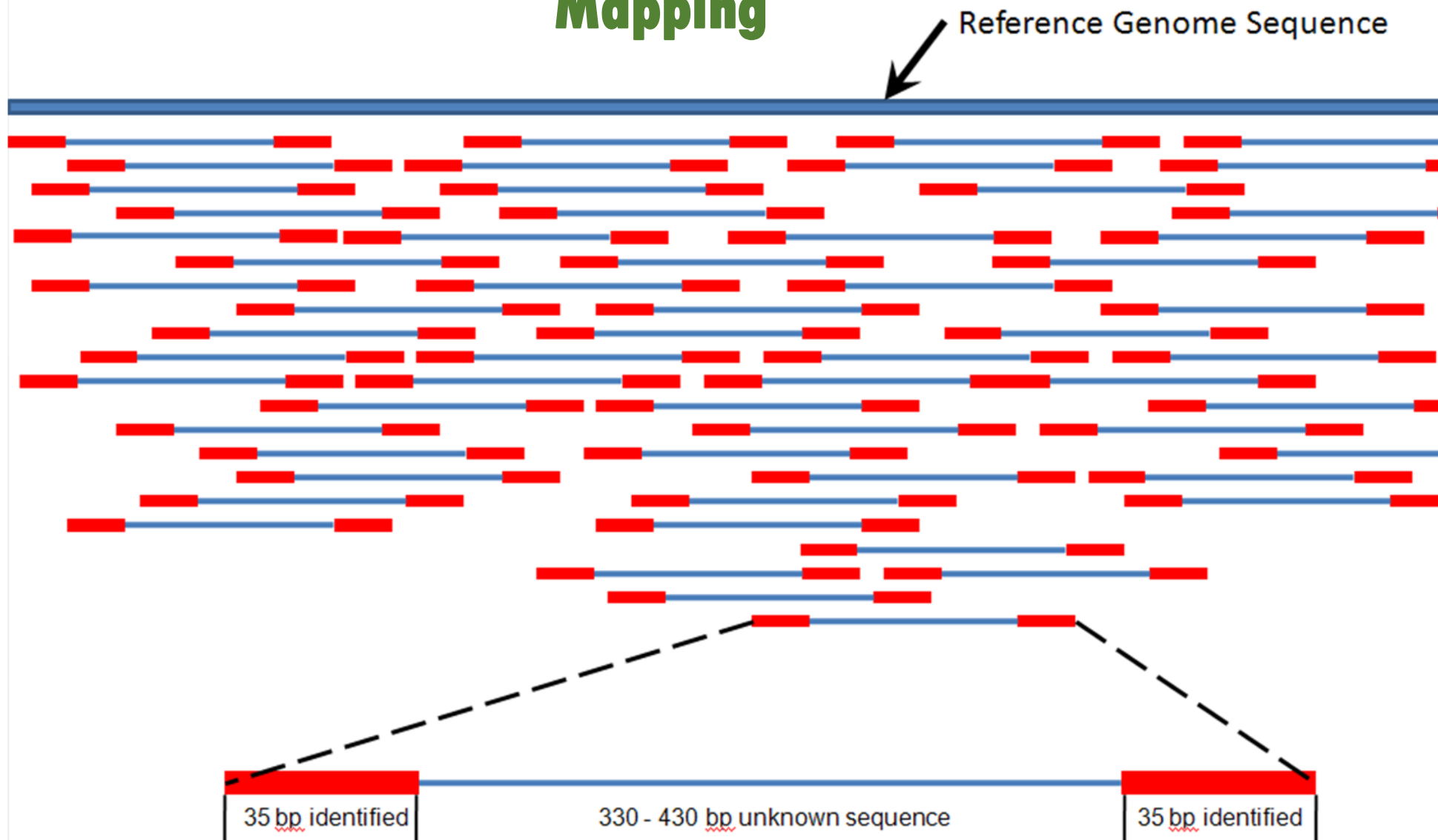
- **Variant calling** : Reference 서열과 환자의 서열 사이에 차이가 발생한 서열 분석
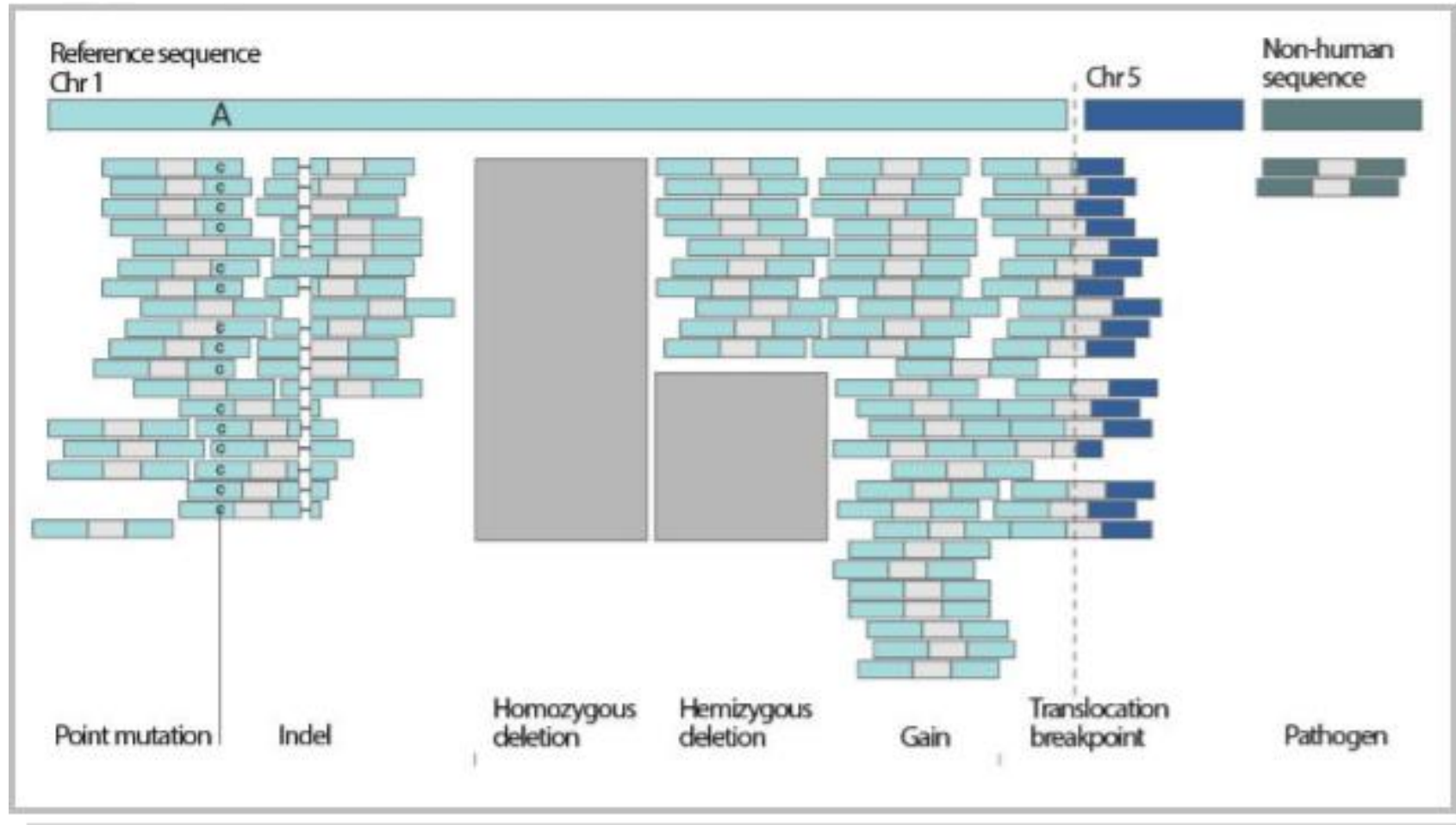
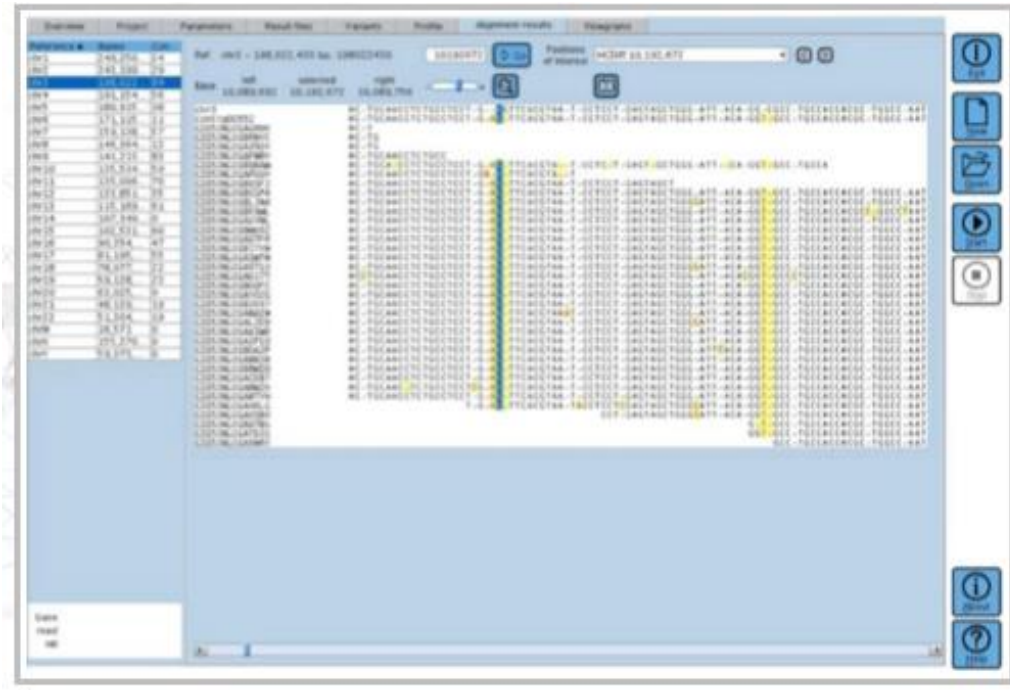(insertions, deletions and substitutions)

3. 엑솜 시퀀싱 데이터 처리

Mapping

Reference Genome Sequence

35 bp identified | 330 - 430 bp unknown sequence | 35 bp identified

3. 엑솜 시퀀싱 데이터 처리
Varient Calling
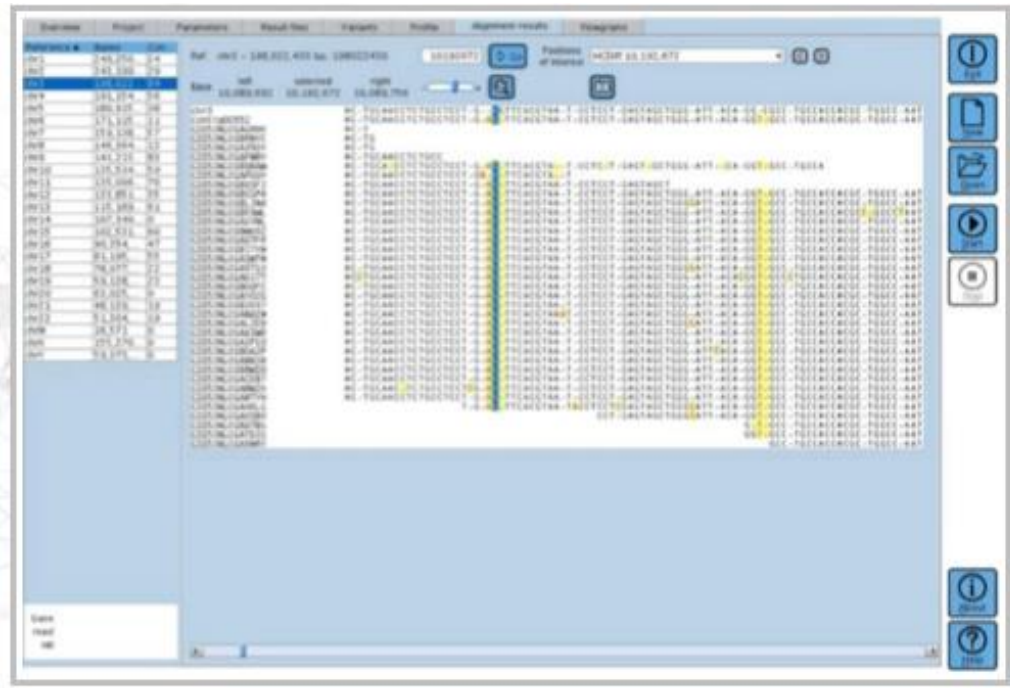
3. 엑솜 시퀀싱 데이터 처리 – gsMapper

· **Roche 454 얻은 reads를 Human reference sequence에 매핑해주는 소프트웨어**

3. 엑솜 시퀀싱 데이터 처리 – gsMapper

시퀀싱 에러

두 개 나와야지 에러? 세 개 나와야지 에러?

많은 데이터를 생산하기 때문에

통계적 접근이 필요함

소프트웨어에서는 오류의 Cut off정해서 판단함.

감사합니다 ^^

-조은해-