

# NGS데이터분석

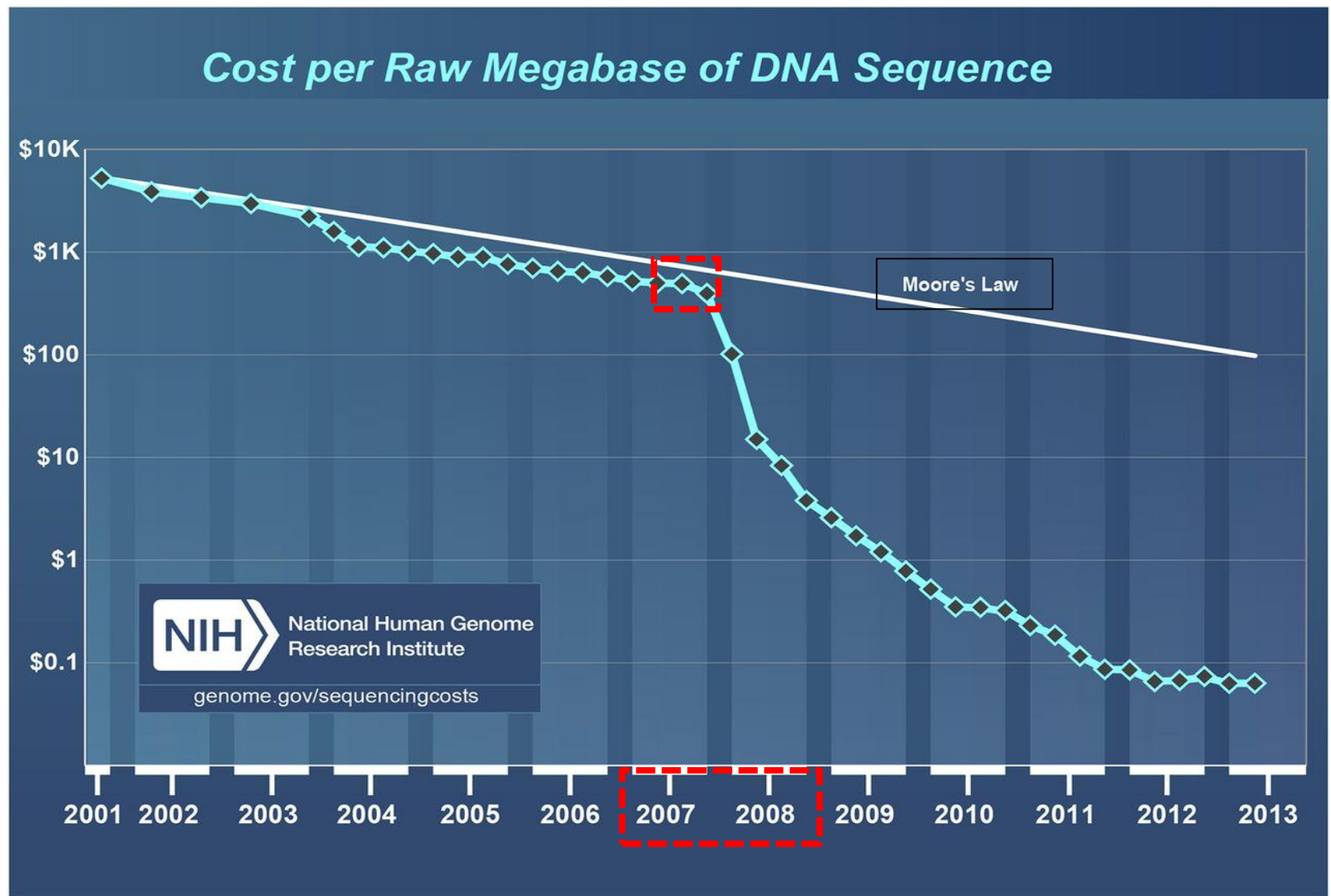
## NGS(Next Generation Sequencing)

- 생물개체의 유전정보를 갖고 있는 DNA의 서열을 밝히는 과정을 의미

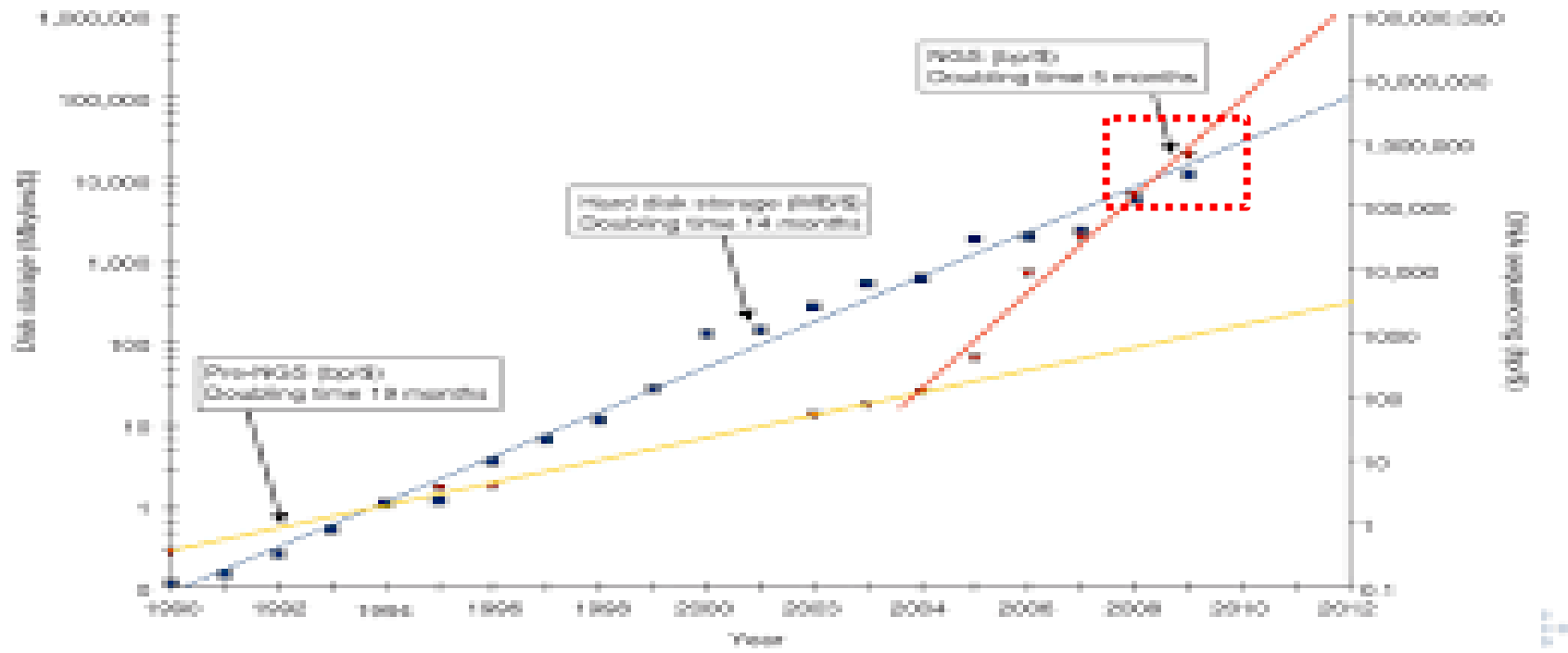
## NGS economy

- 1977년 – Sanger sequencing 개발
- 인간게놈프로젝트 시작: 1990년 – 완료: 2003년
- 1990년대 말 automatic sequencing 기기 개발 (1-2년만에 대부분의 시퀀싱 데이터를 생산)
- 미국보건원 산하 인간게놈연구소 프로젝트 공모: \$1000 - 한 명의 인간게놈 시퀀싱

# NGS(Next Generation Sequencing)

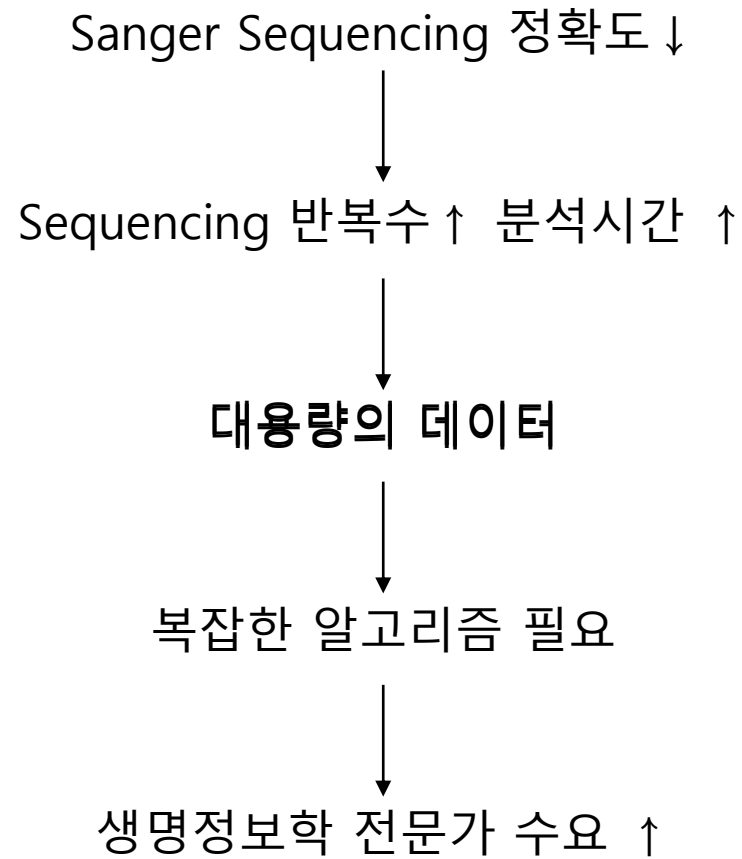


# NextGen Sequencing a Game-Changer

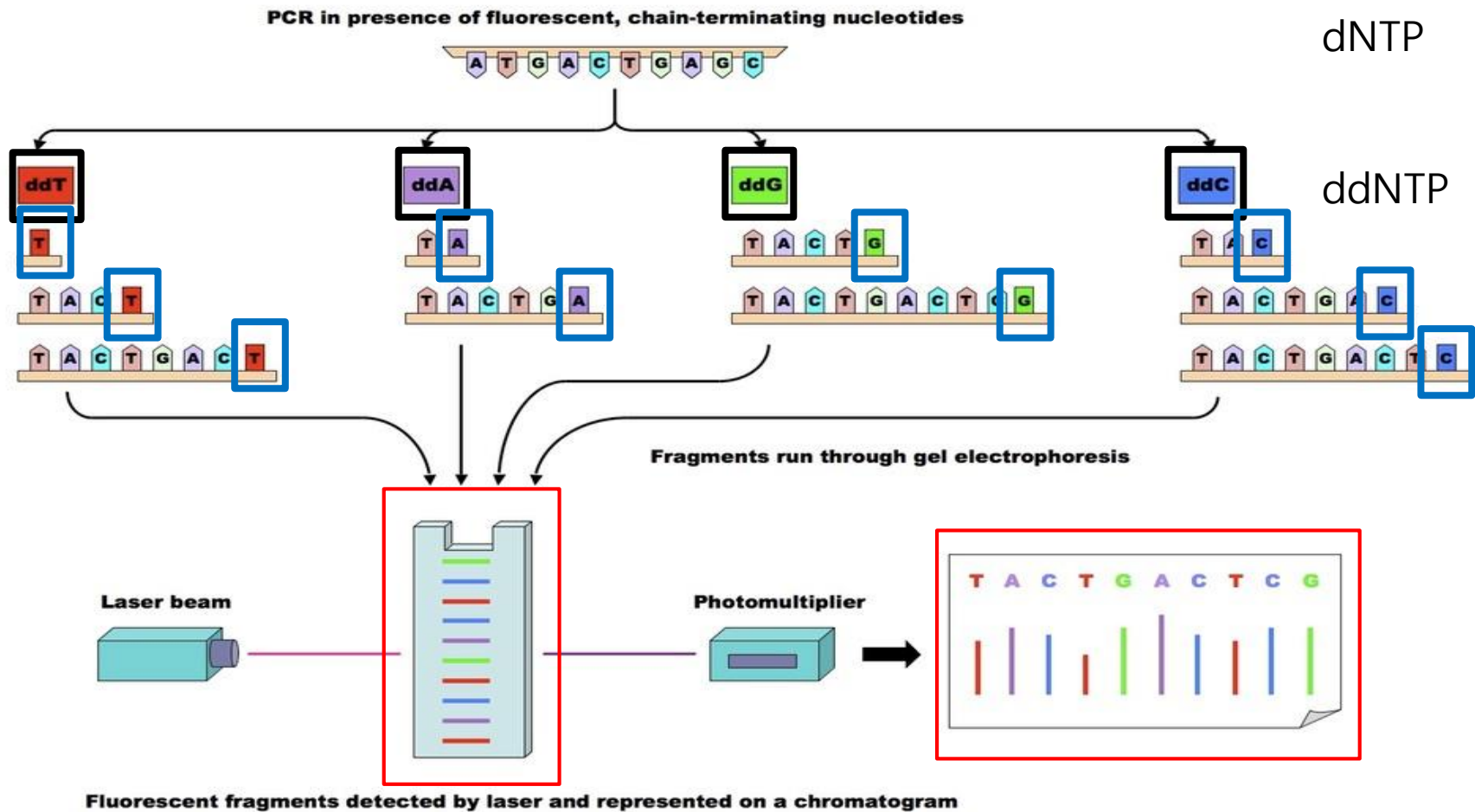


- NGS 개발 이전 – **19**개월 마다 2배씩 증가
- Hard disk - **14**개월 마다 2배씩 증가
- NGS – **5**개월 마다 2배씩 증가

# NGS 문제점

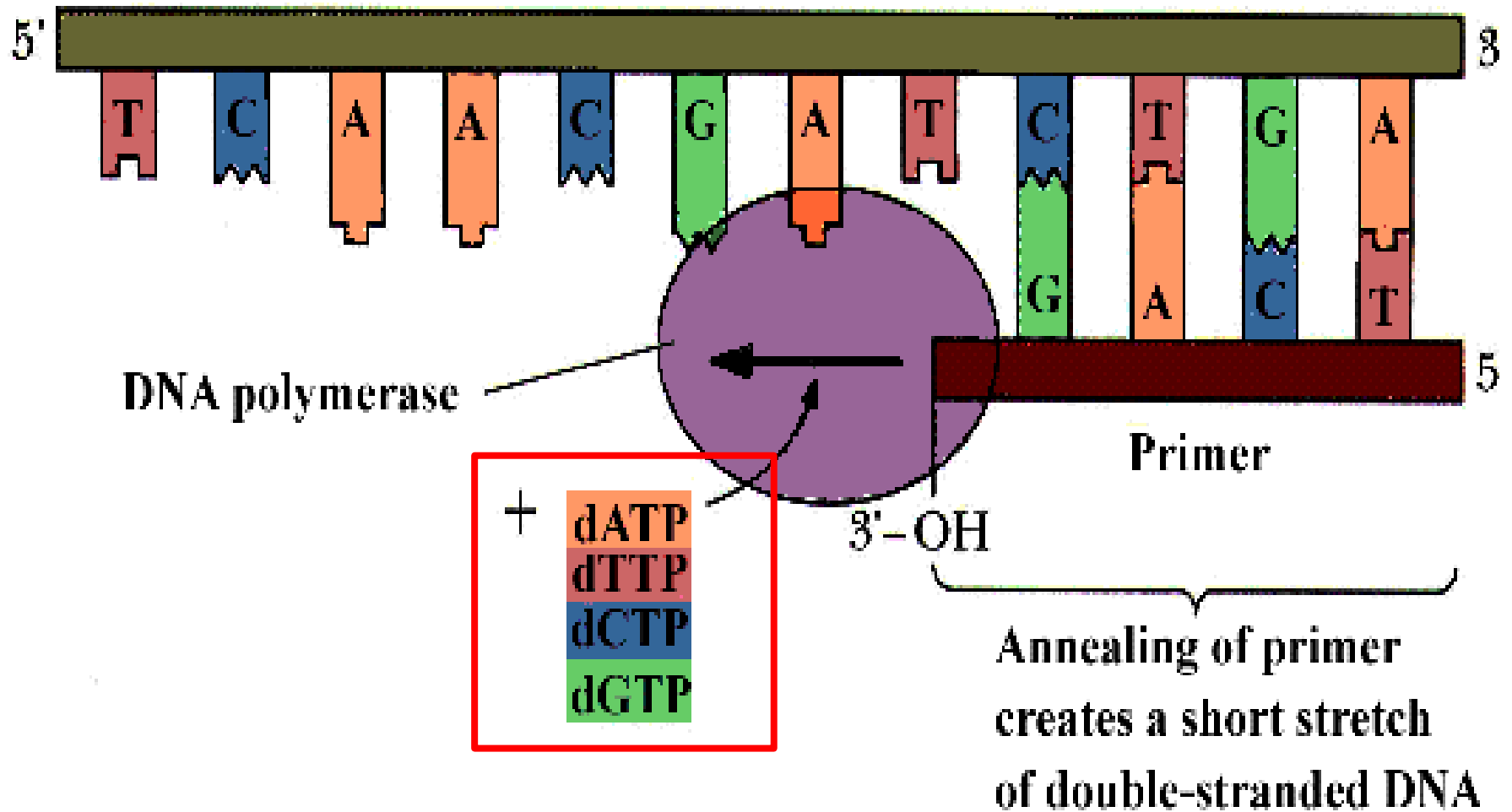


- Sanger Sequencing

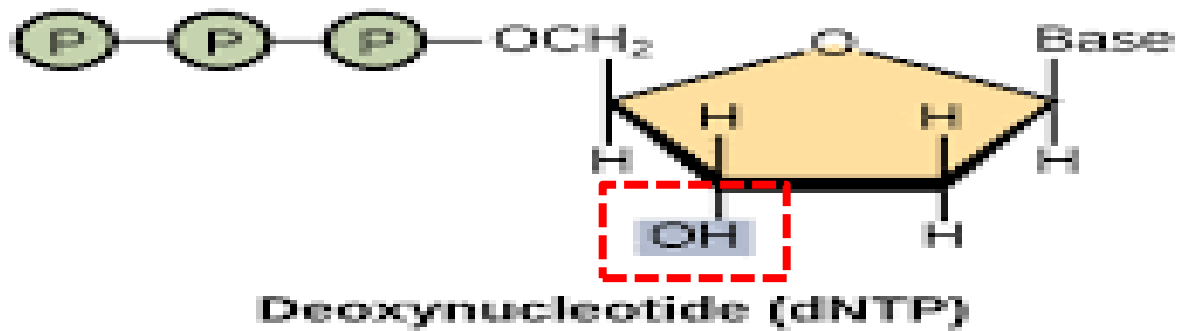
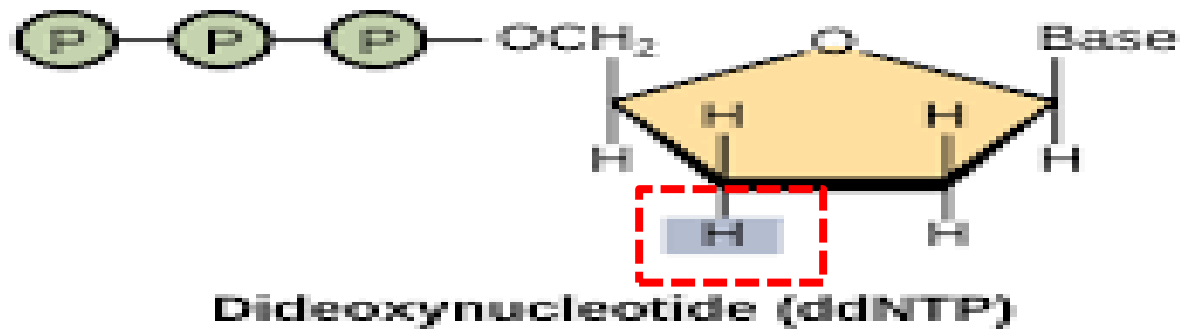


- DNA Polymerase

## Single-Stranded DNA



- dNTP / ddNTP



- 

- Trimming: 1% 이상 연속적인 분포이면 제거
- Phred score =  $\log(x) \times -10$