

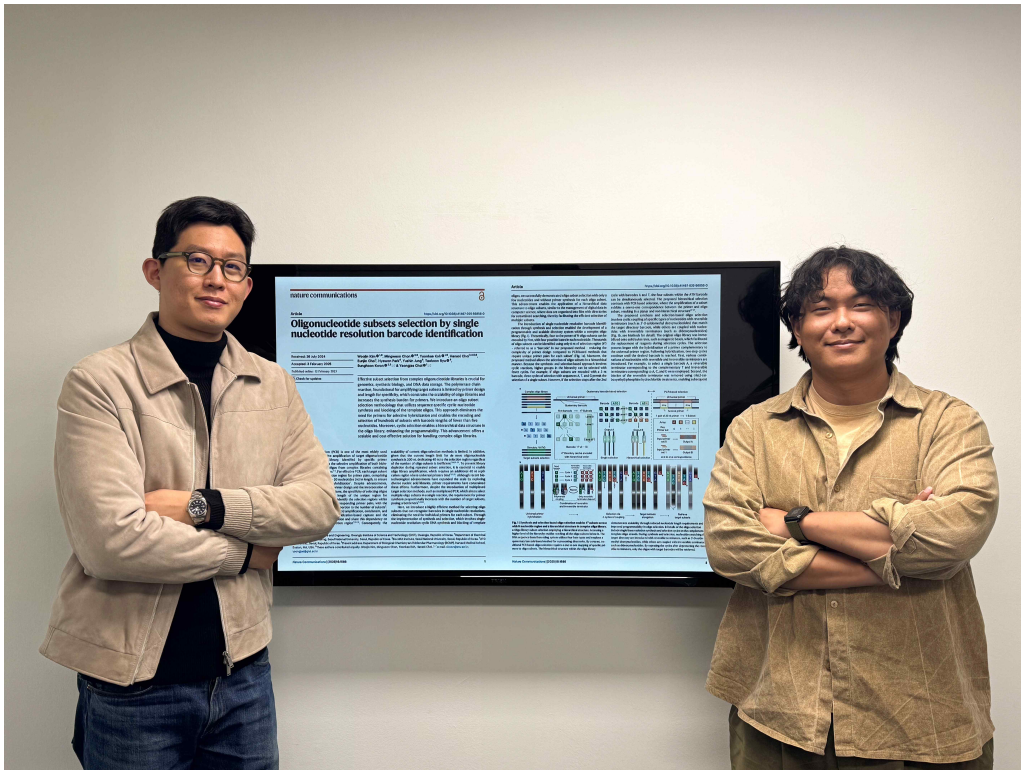
“꿈의 저장 기술 현실로 성큼”

GIST, 차세대 DNA 저장 기술 개발

데이터 센터 수준의 저장 효율 및 확장성 확보 가능

프라이머 없이도 4000억 개 이상 DNA 파일에 자유자재 접근

- GIST·서울대·(주)에이티지라이프텍 공동연구팀 '순환적 DNA 합성 및 선택' 방식 개발... 단 4개 염기로 원하는 DNA 파일 접근, 기존 방식 대비 7400만 배 높은 데이터 집적도
- “DNA 저장 시스템 상용화의 중요한 돌파구 기대” 국제학술지 《Nature Communications》 게재



▲ (왼쪽부터) 신소재공학과 최영재 교수, 김우진 석박통합과정 학생

전 세계적으로 DNA* 응용 기술 연구가 활발히 진행되고 있는 가운데, 반영구적이며 유지 비용이 적게 드는 **DNA 기반 저장 방식**이 차세대 메모리 기술로 주목받고 있다.

국내 연구진이 개발한 **새로운 DNA 파일 접근 기술**이 기존 실리콘 반도체 메모리의 한계를 뛰어넘어 데이터 집적도를 ‘데이터 센터’ 수준까지 획기적으로 높일 수 있을 것으로 기대된다.

* **DNA(Deoxyribo Nucleic Acid, 디옥시리보핵산)**: 대부분의 생명체(일부 바이러스 제외)의 유전 정보를 담고 있는 화학 물질의 일종이다.

광주과학기술원(GIST, 총장 임기철)은 신소재공학과 최영재 교수 연구팀이 서울대학교 권성훈 교수, (주)에이티지라이프텍 연구팀과 공동으로 ‘순환적 DNA 합성 및 선택’ 방식*을 개발했다고 밝혔다.

이 기술을 활용하면 **DNA 데이터 내 특정 파일을 보다 정밀하게 찾아내고 자유자재로 조작할 수 있다.**

* **순환적 DNA 합성 및 선택(Cyclic DNA Synthesis and Selection) 방식**: 특정 DNA 서열을 선택하기 위해 DNA 합성과 선택 과정을 반복적으로 수행하는 방법이다. 먼저, DNA 합성 단계에서 단일 나선 주형 DNA 위에서 목표 서열을 포함하는 DNA만 계속 합성되도록 조절하고, 그렇지 않은 서열은 멈추게 한다. 이후, 선택 과정에서 원하는 서열만 남기고 나머지는 제거하며, 이 과정을 여러 번 반복하여 최종적으로 특정 DNA 서열만 남긴다. 이를 통해 프라이머 없이도 원하는 DNA 서열을 정밀하게 선택할 수 있다.

PCR(polymerase chain reaction)*, 혼성화 캡처(Hybridization Capture) 등 기존의 DNA 파일 접근 기술은 특정 DNA를 증폭하거나 물리적으로 분리하기 위해 서로 다른 프라이머*를 설계해야 한다.

그러나 프라이머는 최소 20개의 염기를 포함해야 하며, 이 때문에 특정 DNA를 인식하려면 긴 서열을 추가적으로 할당해야 하는 구조적 한계가 있다. 또한, 구분해야 하는 DNA 파일의 종류가 늘어날수록 이를 구별하기 위한 프라이머 설계 및 합성 비용이 기하급수적으로 증가하는 문제가 있었다.

따라서 프라이머를 사용하지 않더라도 효율적으로 DNA 파일을 식별하고 저장할 수 있으며, 다양한 데이터 규모와 복잡한 구조에도 대응할 수 있는 새로운 생화학적 방식이 필요하다.

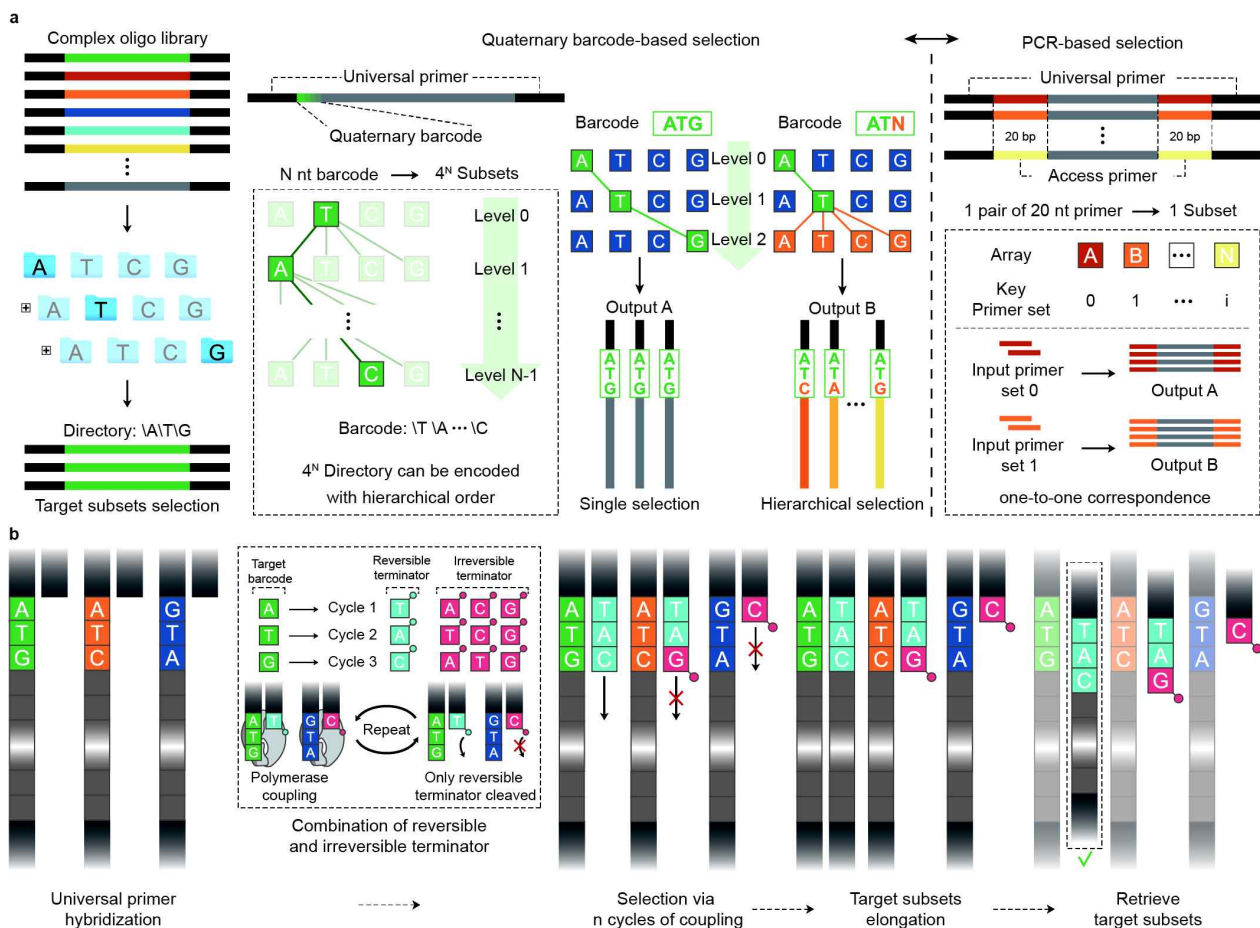
* **PCR(중합효소 연쇄 반응, Polymerase Chain Reaction)**: 특정 DNA 서열을 증폭하는 기술로, 짧은 시간 내에 원하는 DNA를 대량으로 복제할 수 있도록 한다. 이 과정은 변성(denaturation), 결합(annealing), 연장(extension)의 세 단계를 반복하며 진행된다. 먼저, 열을 가해 DNA의 이중나선을 분리(변성)하여 단일 가닥 형태로 만든다. 이후, 특정 서열과 상보적으로 결합할 수 있는 프라이머가 DNA와 결합(결합)하며, DNA 중합효소가 프라이머를 시작점으로 새로운 DNA 가닥을 합성(연장)한다. 이 세 단계를 반복하면서 각 사이클마다 DNA의 양이 기하급수적으로 증가하며, 최종적으로 원하는 DNA 서열이 대량으로 증폭된다.

* **프라이머**: 일반적으로 20개의 염기로 이루어진 짧은 DNA 조각으로 DNA 복제나 증폭 과정에서 특정 부분을 찾아주는 역할을 한다. 프라이머가 특정 서열을 인식하여 결합하면 DNA 복제 과정이 그 지점에서부터 시작된다.

연구팀이 개발한 '순환적 DNA 합성 및 선택' 기술은 단일 염기 수준의 바코드를 활용하여 프라이머 없이도 DNA 파일을 계층 구조(계층적 선택 방식)*로 탐색할 수 있도록 설계되었다.

이는 컴퓨터의 폴더 탐색 방식과 유사한 개념으로, 기존 PCR 방식 대비 비용이 10배 절감되고 접근 효율은 3배 이상 향상되었다. 또한, 구분할 수 있는 DNA 파일의 수가 최소 7400만 배 이상 증가했으며, 특정 DNA 파일을 제거하고 새로운 DNA 파일을 삽입하는 등 파일 교체도 가능해졌다.

* **계층적 선택 방식**: 컴퓨터에서 파일을 개별적으로 다루는 것이 아닌 폴더를 이용하는 것과 같이 데이터 구조적으로 매우 효율적인 방식이다.



▲ **순환적 DNA 합성 및 선택 기술의 개요.** DNA 라이브러리는 계층적 구조로 구성되어 있어 상위 단계에 접근하면 그 하위에 포함된 모든 DNA 집합에 쉽게 도달할 수 있다. DNA 염기서열 기반의 바코딩 시스템은 4개의 염기를 이용한 쿼터너리(4진법) 방식으로 디렉토리를 표현하며 기존의 PCR 방식처럼 특정 프라이머와 일대일 대응할 필요가 없다. 또한, 상보적 템플릿 DNA 가닥 위에서 단일 염기 수준의 합성과 선택 사이클이 진행된다. 타겟 디렉토리에 부합하는 염기는 가역적 종결제(예: 3'-O-아지도메틸 디옥시뉴클레오타이드)를 사용하여 추가되고, 타겟이 아닌 염기는 비가역적 종결제(예: 디데옥시뉴클레오타이드)를 사용하여 결합된다. 비가역적 종결제를 제거한 후 사이클을 반복함으로써 최종적으로 원하는 바코드를 가진 DNA만을 선택적으로 회수한다.

기존 PCR 방식에서는 DNA 파일마다 별도의 프라이머(최소 20개 염기) 한 쌍을 설계하고 합성해야 했지만, 이 기술은 **4개의 염기만으로 특정 DNA 파일에 접근할 수 있다.**

이때, 염기의 개수를 추가하는 방식으로 이론적으로 **4^N 개의 DNA 파일을 구분할 수 있어 확장성이 뛰어나다.** 예를 들어, 4개 염기의 바코드는 256종의 DNA 파일을, 8개 염기의 바코드는 65,536종의 DNA 파일을 구분할 수 있다. 또한 **일반적인 프라이머의 길이인 20개의 염기를 이용하면 이론적으로 4,150억 개의 하위 집합을 인코딩할 수 있다.**

최영재 교수는 “이번 연구를 통해 **프라이머 없이도 특정 DNA 파일에 접근할 수 있는 새로운 방법론**을 제시하여 **계층적 바코드 시스템을 적용함으로써 기존 PCR 방식의 한계를 극복했다**”고 설명했다. 이어 “향후 바코드 설계 최적화 및 자동화된 시스템과 결합하면 **차세대 DNA 파일 접근 기술로서 DNA 기반 저장 시스템 상용화에 중요한 돌파구가 될 것으로 기대한다**”고 밝혔다.

이번 연구는 GIST 신소재공학과 최영재 교수가 지도하고 김우진 석박통합과정생과 고윤희 석사과정생이 함께 수행했으며, 과학기술정보통신부 STEAM연구사업(미래유망융합기술파이오니어)의 지원을 받았다. 연구 결과는 국제학술지 《네이처 커뮤니케이션즈(Nature Communications)》에 2025년 2월 12일 온라인 게재되었다.

논문의 주요 정보

1. 논문명, 저자정보

- 저널명 : Nature Communications (IF: 14.7, 2023년 기준)
- 논문명 : Oligonucleotide subsets selection by single nucleotide resolution barcode identification
- 저자 정보 : 김우진(제1저자, GIST 신소재공학과), 고윤희(공동 제1저자, GIST 신소재공학과), 전민권(공동 제1저자, 서울대학교), 최한솔 박사(공동 제1저자, 서울대학교) 최은진 박사(GIST 신소재공학과), 류태훈 박사(에이티지 라이프텍), 정유신 박사(에이티지 라이프텍), 박혜원(에이티지 라이프텍), 권성훈 교수(교신저자, 서울대학교), 최영재 교수(교신저자, GIST 신소재공학과)