

# 보도자료



미래를 개척하는 지식 공동체



서울대학교  
SEOUL NATIONAL UNIVERSITY

보도일시	즉시 보도 가능
	2024. 6. 14.(금)
문의	연구책임자 농생명공학부 김희발 교수(02-880-4822) / 교신저자

## ■ 제목/부제

제목	서울대 농생명공학부 김희발 교수팀, 유전체 주석 프로그램 정확도 최대 20% 향상 알고리즘 발표
Journal Link	<a href="http://doi.org/10.1101/gr.278566.123">http://doi.org/10.1101/gr.278566.123</a>

## ■ 본문

### □ 문단 1

최근 생물학 연구에서 시퀀싱 기술의 발전으로 막대한 양의 유전체 데이터가 생성되고 있다. 지구바이옴게놈 프로젝트 (Earth BioGenome Project)와 같은 대규모 프로젝트는 생명의 다양성을 지도로 그려내기 위해 수많은 종의 유전체를 해독하고 있으며, 첫 3년 동안 9,000종, 10년 동안 180만 종의 진핵생물 유전체 데이터를 생산하고자 한다. 이러한 막대한 양의 유전체 데이터 생산 조립 후, 가장 중요한 과정은 유전자의 밝혀내는 유전체 주석달기(Genomic Annotation)이다. 이는 유전자의 위치와 기능을 식별하는 과정으로, 생물의 유전 정도 이해를 위해 필수적이다.

### □ 문단 2

서울대 김희발 교수 연구팀은 최근 특정 경로에 있는 유전자를 신속하고 정확하게 식별할 수 있는 유전체 주석달기 프로그램 페이지서치 (PaGeSearch)를 개발했다. 이 결과는 유전체 연구의 최상위급 저널인 지놈 리서치 (Genome Research) 5월호에 게재되었다.

### □ 문단 3

페이지서치는 초기 시퀀스 유사성 검색을 통해 관련 유전체 영역을 식별한 후, 타깃 유전자 예측과 신경망 기반 결과 필터링을 결합하여 작동한다. 기존에 이미 높은 수준의 정확도를 가진 유전체 주석달기 프로그램이 있는 상태에서 이를 개선하는 것은 큰 도전이었지만 개발팀은 기계 학습 기법을 통해 종에 따라 최대 20%까지 정확도를 향상시켰다.

### □ 연구결과

현재는 동물과 식물 종에 초점을 맞추고 있지만, 향후 이 프로그램은 곰팡이를 비롯해 원생생물, 조류, 플랑크톤 등 더 다양한 생물에 적용될 수 있다. 이는 진핵생물 전반에 걸친 정확한

유전자 모델링을 가능하게 하며 앞으로 생산될 수많은 유전체 데이터를 해석하는 데에 유용하게 쓰일 수 있을 것으로 전망된다.