

RS-2021-I  
P821026

보안 과제( ), 일반 과제( O ) / 공개( O ), 비공개( )발간등록번호( O )  
기술사업화지원사업 2023년도 최종보고서

발간등록번호

11-1543000-004691-01

인공지능  
기반  
유전체  
분석  
프로그램  
및  
디지털  
육종  
플랫폼  
개발

# 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발

2024

2024.07.09.

한국생명공학연구원 /  
(주)유클리드소프트

농림축산식품부  
농림식품기술기획평가원

농림축산식품부  
(전문기관)농림식품기술기획평가원

## 제 출 문

농림축산식품부 장관 귀하

본 보고서를 “인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발”(개발  
기간 : 2021.04.01 ~ 2023.12.31)과제의 최종보고서로 제출합니다.

2024. 07. 09.

주관연구기관명 : 한국생명공학연구원 (대표자) 김 장 성 (인)

공동연구기관명 : (주)유클리드소프트 (대표자) 박 주 한 (인)

주관연구책임자 : 김 용 민 (인)

공동연구책임자 : 박 주 한 (인)

국가연구개발사업의 관리 등에 관한 규정 제18조에 따라 보고서 열람에 동의합니다.

## &lt; 요약 문 &gt;

※ 요약문은 5쪽 이내로 작성합니다.

사업명		기술사업화지원사업			총괄연구개발 식별번호 (해당 시 작성)		
내역사업명 (해당 시 작성)		공공기술 사업화 촉진			연구개발과제번호		821026-03
기술 분류	국가과학기술 표준분류	1순위 LB0202	50 %	2순위 LA0706	50 %	3순위	%
	농림식품 과학기술분류	1순위 AA0201	50 %	2순위 CA0302	50 %	3순위	%
총괄연구개발명 (해당 시 작성)		-					
연구개발과제명		인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발					
전체 연구개발기간		2021. 04. 01 - 2023. 12. 31 (2년 9개월)					
총 연구개발비		총 949,840 천원 (정부지원연구개발비: 825,000 천원, 기관부담연구개발비: 124,840 천원)					
연구개발단계		기초[ ] 응용[ ] 개발[ <input checked="" type="checkbox"/> ] 기타(위 3가지에 해당되지 않는 경우)[ ]		기술성숙도 (해당 시 작성)		착수시점 기준(3단계) 종료시점 목표(9단계)	
연구개발과제 유형 (해당 시 작성)							
연구개발과제 특성 (해당 시 작성)							
연구개발 목표 및 내용		최종 목표		인공지능 기술 기반의 유전체 분석 프로그램 개발과 인공지능 및 웹 기반의 디지털 육종 플랫폼 개발 및 서비스			
		전체 내용		<ul style="list-style-type: none"> <li>- 인공지능 분석 기반 유전체 데이터 분석 및 가공 프로그램 개발을 통한 유전체 통합 분석 프로그램 개발</li> <li>- 인공지능 기반 계통/개체 분류 및 유전자 네트워크 분석 기술개발</li> <li>- 인공지능 및 웹 기반 디지털 육종 플랫폼 개발 및 서비스</li> </ul>			
		1단계 (해당 시 작성)	목표	<ul style="list-style-type: none"> <li>■ 개체/계통 특이적 변이 탐색 프로그램</li> <li>■ 전사체 분석 자동화 프로그램 개발</li> <li>■ 유전체 변이 기반 개체/계통분류 프로그램 개발</li> <li>■ 표현형 데이터 생산</li> <li>■ 유전체 데이터베이스 구축</li> <li>■ 표현형 기반 분류 프로그램 개발</li> <li>■ 유전자 네트워크 분석 프로그램 개발</li> </ul>			
			내용	<ul style="list-style-type: none"> <li>■ 개체/계통 특이적 변이 탐색 프로그램의 범용화 및 자동화를 통한 프로그램 고도화(정확도 81.5% 이상)</li> <li>■ 전사체 기반 유전자 네트워크 분석 프로그램 개발을 위한 데이터 생산 및 분석 자동화 프로그램 개발</li> <li>■ 인공지능과 유전체 변이 데이터를 활용한 개체/계통분류 프로그램 개발</li> <li>■ 표현형 기반 분류 모델 프로그램 개발을 위한 표현형 데이터 생산 및 제공</li> <li>■ 웹 기반 디지털 육종 플랫폼을 서비스를 위한 유전체 DB 구축</li> <li>■ 인공지능을 활용한 표현형 기반 계통 분류 프로그램 개발</li> <li>■ 전사체 데이터를 활용한 유전자 네트워크 분석 프로그램 개발</li> </ul>			
2단계 (해당 시 작성)	목표	<ul style="list-style-type: none"> <li>■ 유전체 통합 분석 프로그램 개발 완료</li> <li>■ 유전체-표현형 연관 분석 프로그램 개발</li> <li>■ 디지털 육종 플랫폼 구축 및 서비스</li> </ul>					

		내용	<ul style="list-style-type: none"> <li>■ 개발된 유전체/전사체 분석 프로그램의 통합 분석 프로그램 개발 완료 및 정확성 검증</li> <li>■ 유전체 변이와 표현형의 연관 분석 프로그램 개발 완료</li> <li>■ 인공지능 및 웹 기반 디지털 육종 플랫폼 구축 및 서비스 개시를 통한 사업화 완료</li> </ul>
--	--	----	--

연구개발성과	<ul style="list-style-type: none"> <li>■ 개체/계통 특이적 변이 탐색 프로그램 자동화 완료</li> <li>■ 유전체 변이 기반 개체/계통분류 프로그램 개발 및 특허 출원 완료</li> <li>■ 표현형 데이터 생산 및 분류 프로그램 개발</li> </ul>
--------	--

연구개발성과 활용계획 및 기대 효과	<ul style="list-style-type: none"> <li>- 인공지능 기반 유전체/전사체 분석의 핵심 원천기술 확보</li> <li>- 핵심 원천기술 응용을 통한 유전체 분석 프로그램 개발을 통한 주요 농업자원의 종자순도 검증 마커 개발 및 계통분류 활용을 통한 디지털 육종 체계 구축</li> <li>- 고가의 유전체 분석 프로그램의 수입 대체 효과 및 유전체 분석 프로그램의 대중화를 통한 유전체 데이터 활용성 제고 및 분석 시장 확대를 통한 신규 시장 개척</li> </ul>
---------------------	--

연구개발성과의 비공개여부 및 사유	
--------------------	--

연구개발성과의 등록·기탁 건수	논문	특허	보고서 원문	연구 시설·장비	기술 요약 정보	소프트웨어	표준	생명자원		화합물	신품종	
								생명 정보	생물 자원		정보	실물
	4	2						30				
연구시설·장비 종합정보시스템 등록 현황	구입 기관	연구시설·장비명	규격 (모델명)	수량	구입 연월일	구입가격 (천원)	구입처 (전화)	비고 (설치장소)	ZEUS 등록번호			
국문핵심어 (5개 이내)	인공지능		유전체		전사체		계통/개체 분류		유전자 네트워크 분석			
영문핵심어 (5개 이내)	Artificial Intelligence		Genome		Transcriptome		Identification of variety/individual		Gene network analysis			

## < 목 차 >

1. 연구개발과제의 개요
2. 연구개발과제의 수행 과정 및 수행내용
3. 연구개발과제의 수행 결과 및 목표 달성 정도
4. 목표 미달 시 원인분석(해당 시 작성)
5. 연구개발성과 및 관련 분야에 대한 기여 정도
6. 연구개발성과의 관리 및 활용 계획

별첨 자료 (참고 문헌 등)

※ 각 항목에서 요구하는 정보를 포함하여 연구개발과제의 특성에 따라 항목을 추가하거나 항목의 순서와 구성을 변경하는 등 서식을 수정하여 사용하거나 별도의 첨부자료 활용이 가능합니다.

■ 농림축산식품연구개발사업 관리기준 [별지 제17호 서식] <제23조제3호 관련>

최종보고서							보안등급								
							일반[ <input checked="" type="checkbox"/> ], 보안[ <input type="checkbox"/> ]								
중앙행정기관명		농림축산식품부			사업명		기술사업화지원사업								
전문기관명 (해당 시 작성)		농림식품기술기획평가원			내역사업명 (해당 시 작성)		공공기술사업화촉진								
공고번호		제 농축2021-41 호			총괄연구개발 식별번호 (해당 시 작성)		-								
					연구개발과제번호				821026-03						
기술분류	국가과학기술 표준분류	1순위 LB0202	50 %	2순위 LA0706	50 %	3순위		%							
	농림식품과학기술분류	1순위 AA0201	50 %	2순위 CA0302	50 %	3순위		%							
총괄연구개발명 (해당 시 작성)		국문													
		영문													
연구개발과제명		국문		인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발											
		영문		Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform											
주관연구개발기관		기관명		한국생명공학연구원		사업자등록번호		314-82-06063							
		주소		(우)34141 대전광역시 유성구 과학로 125		법인등록번호		160171-0001632							
연구책임자		성명		김용민		직위		책임연구원							
		연락처		직장전화 전자우편		휴대전화		국가연구자번호							
연구개발기간		전체		2021. 04. 01 - 2023. 12. 31 (2년 9개월)											
		1단계[2년]		2021. 04. 01 - 2022. 12. 31 (1년 9개월)											
		2단계[1년]		2023. 01. 01 - 2023. 12. 31 (1년)											
연구개발비 (단위: 천원)		정부지원		기관부담		그 외 기관 등의 지원금		합계		연구개발 외 지원금					
		연구개발비		연구개발비		지방자치단체 기타( )									
		현금		현금		현금		현금		현금					
총계		825,000		8,734		116,106		833,734		116,106		949,840			
1단계		1년차		225,000		3,734		33,606		228,734		33,606		262,340	
		2년차		300,000		0		37,500		300,000		37,500		337,500	
2단계		1년차		300,000		5,000		45,000		305,000		45,000		350,000	
공동연구개발기관 등 (해당 시 작성)		기관명		책임자		직위		휴대전화		전자우편		비고			
공동연구개발기관		(주)유클리드소프트		박주한		대표이사						역할 수요		기관유형 중소기업	
위탁연구개발기관															
연구개발기관 외 기관															
연구개발담당자 실무담당자		성명		신아영		직위		선임연구원							
		연락처		직장전화 전자우편		휴대전화		국가연구자번호							

이 최종보고서에 기재된 내용이 사실임을 확인하며, 만약 사실이 아닌 경우 관련 법령 및 규정에 따라 제재처분 등의 불이익도 감수하겠습니다.

2024 년 2 월 29 일

연구책임자: 김 용 민 (인)

주관연구개발기관의 장: 김 장 성 (직인)

공동연구개발기관의 장: 채 은 경 (직인)

농림축산식품부장관·농림식품기술기획평가원장 귀하

■ 농림축산식품연구개발사업 관리기준 [별지 제17호 서식] <제23조제3호 관련>

최종보고서							보안등급		
							일반[ <input checked="" type="checkbox"/> ], 보안[ <input type="checkbox"/> ]		
중앙행정기관명	농림축산식품부			사업명	사업명		기술사업화지원사업		
전문기관명 (해당 시 작성)	농림식품기술기획평가원			내역사업명 (해당 시 작성)	공공기술사업화촉진				
공고번호	제 농축2021-41호			총괄연구개발 식별번호 (해당 시 작성)	-				
				연구개발과제번호	821026-03				
기술분류	국가과학기술 표준분류	1순위 LB0202	50 %	2순위 LA0706	50 %	3순위	%		
	농림식품과학기술분류	1순위 AA0201	50 %	2순위 CA0302	50 %	3순위	%		
총괄연구개발명 (해당 시 작성)	국문								
	영문								
연구개발과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발							
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform							
주관연구개발기관	기관명	한국생명공학연구원		사업자등록번호	314-82-06063				
	주소	(우)34141 대전광역시 유성구 과학로 125		법인등록번호	160171-0001632				
연구책임자	성명	김용민		직위	책임연구원				
	연락처	직장전화		휴대전화					
		전자우편		국가연구자번호					
연구개발기간	전체	2021. 04. 01 - 2023. 12. 31(2년 9개월)							
	1단계[2년]	2021. 04. 01 - 2022. 12. 31(1년 9개월)							
	2단계[1년]	2023. 01. 01 - 2023. 12. 31(1년)							
연구개발비 (단위: 천원)	정부지원	기관부담	그 외 기관 등의 지원금				합계		연구개발 외 지원금
	연구개발비	연구개발비	지방자치단체	기타( )		현금	현물	합계	
	현금	현금	현물	현금	현물	현금	현물	합계	지원금
총계	825,000	8,734	116,106			833,734	116,106	949,840	
1단계	1년차	225,000	3,734	33,606			228,734	33,606	262,340
	2년차	300,000	0	37,500			300,000	37,500	337,500
2단계	1년차	300,000	5,000	45,000			305,000	45,000	350,000
공동연구개발기관 등 (해당 시 작성)	기관명	책임자	직위	휴대전화	전자우편	비고			
		역할	기관유형						
공동연구개발기관	㈜유클라우드소프트	박주한	대표이사						수요 중소기업
위탁연구개발기관									
연구개발기관 외 기관									
연구개발담당자 실무담당자	성명	신아영		직위	선임연구원				
	연락처	직장전화		휴대전화					
		전자우편		국가연구자번호					

이 최종보고서에 기재된 내용이 사실임을 확인하며, 만약 사실이 아닌 경우 관련 법령 및 규정에 따라 제재처분 등의 불이익도 감수하겠습니다.

2024년 2월 29일

연구책임자:

김용민

주관연구개발기관의 장:

김장성

공동연구개발기관의 장:

채은경

위탁연구개발기관의 장:



농림축산식품부장관·농림식품기술기획평가원장 귀하

### 1. 연구개발과제의 개요

#### 1) 연구개발과제의 최종 목표

- 인공지능 분석을 위한 유전체 데이터 분석 및 가공 프로그램 개발을 통한 유전체 통합 분석 프로그램 개발
- 인공지능 분석을 활용한 계통/개체 분류(정확도 81.5% 이상) 및 계통 특이적 마커 발굴 프로그램 개발
- 전사체 데이터를 활용한 유전자 네트워크 분석 프로그램 개발
- 인공지능 및 웹 기반 디지털 육종 플랫폼 구축 기술의 개발 및 서비스화

#### 2) 연구개발과제의 내용

	연구목표	연구 범위 및 내용
1차년도 (2021)	<ul style="list-style-type: none"> <li>■ 개체/계통 특이적 변이 탐색 프로그램</li> <li>■ 전사체 분석 자동화 프로그램 개발</li> <li>■ 유전체 변이기반 개체/계통분류 프로그램 개발</li> </ul>	<ul style="list-style-type: none"> <li>■ 개체/계통 특이적 변이 탐색 프로그램의 범용화 및 자동화를 통한 프로그램 고도화(정확도 81.5% 이상)</li> <li>■ 전사체 기반 유전자 네트워크 분석 프로그램 개발을 위한 데이터 생산 및 분석 자동화 프로그램 개발</li> <li>■ 인공지능과 유전체 변이 데이터를 활용한 개체/계통분류 프로그램 개발</li> </ul>
2차년도 (2022)	<ul style="list-style-type: none"> <li>■ 표현형 데이터 생산</li> <li>■ 유전체 데이터베이스 구축</li> <li>■ 표현형 기반 분류 프로그램 개발</li> <li>■ 유전자 네트워크 분석 프로그램 개발</li> </ul>	<ul style="list-style-type: none"> <li>■ 표현형 기반 분류 모델 프로그램 개발을 위한 표현형 데이터 생산 및 제공</li> <li>■ 웹 기반 디지털 육종 플랫폼을 서비스를 위한 유전체 DB 구축</li> <li>■ 인공지능을 활용한 표현형 기반 계통 분류 프로그램 개발</li> <li>■ 전사체 데이터를 활용한 유전자 네트워크 분석 프로그램 개발</li> </ul>
3차년도 (2023)	<ul style="list-style-type: none"> <li>■ 유전체 통합 분석 프로그램 개발 완료</li> <li>■ 유전체-표현형 연관 분석 프로그램 개발</li> <li>■ 디지털 육종 플랫폼 구축 및 서비스</li> </ul>	<ul style="list-style-type: none"> <li>■ 개발된 유전체/전사체 분석 프로그램의 통합 분석 프로그램 개발 완료 및 정확성 검증</li> <li>■ 유전체 변이와 표현형의 연관 분석 프로그램 개발 완료</li> <li>■ 인공지능 및 웹 기반 디지털 육종 플랫폼 구축 및 서비스 개시를 통한 사업화 완료</li> </ul>

### 2. 연구개발과제의 수행 과정 및 수행 내용

- 개체/계통 특이적 변이 탐색 프로그램의 범용화 및 자동화를 통한 프로그램 고도화
  - 공동연구개발기관의 인공지능 기반 유전체 분석 프로그램 개발을 위한 개체/계통 특이적 변이 데이터 산출 및 서비스 제공을 위한 자동화 분석 파이프라인을 개발함.
  - 디지털 육종 플랫폼 탑재를 위해 공동 연구개발 기관으로 해당 파이프라인 이관 및 탑재를 완료함.
  - NCS 데이터를 활용한 자동화된 계통별 변이 산출 프로그램의 기술이전을 실시함(2022년).
- 인공지능 기반 개체/계통 분류 프로그램 개발
  - 기 개발된 인공지능 기반 개체/계통 분류 프로그램의 경우, 개체 특이적인 유전체 영역을 특정할 수 없는 기술적인 한계가 존재하여 문자 데이터인 변이정보를 이미지로 변환하여 계통특이적 영역을 식별하는 새로운 방법을 제시함(그림 1).
  - Variant Calling Format(VCF) 데이터의 genotype 변수를 분석에 사용했으며, CHROM, POS, Genotype data(REF, ALT)를 필터링하여 사용함.
  - 필터링한 데이터를 사용하여 문자의 sequence가 아닌 하나의 이미지로 표현하여 분석하고자 함. Genotype data의 A, T, G, C를 조합하여 16차원의 색상으로 표현하고 CHR과 POS로 유전체 이미지 내 픽셀의 위치를 지정함.

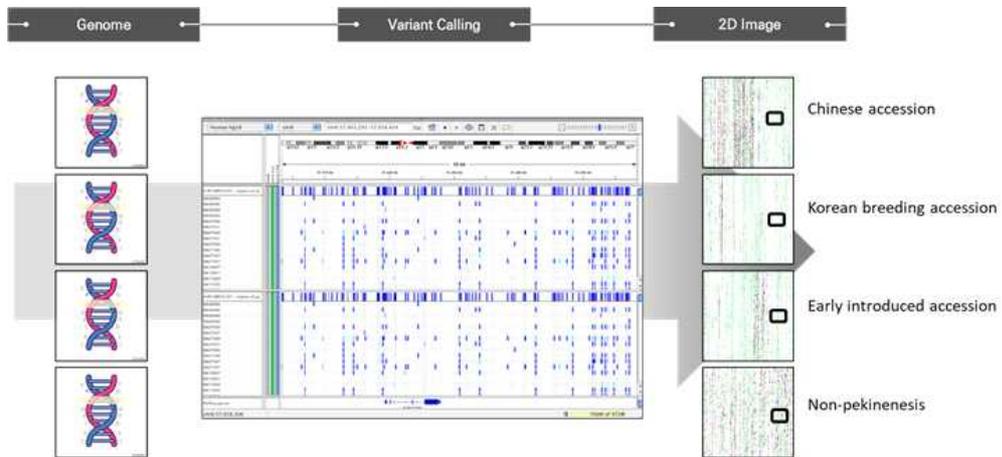


그림 1. 개발한 유전체 데이터 2D 이미지화 프로세스

- 유전체 변이지도의 디지털 이미지 변환 알고리즘의 개량을 통해, 유전체 변이를 특정하기 위한 연구를 수행 중임(그림 2).
- 한 개체의 유전체 전체 정보를 하나의 이미지 데이터로 생성하는 것에서 더 나아가 특이적 마커 추출을 위해 탐색의 범위를 좁혀가는 hierarchical 방법으로 분석을 시도함.
- 전체 sequence를 사용하여 변환한 이미지를 chromosome 및 genetic region 단위로 crop하여 분석을 수행함. 생성된 이미지는 224\*224 픽셀로 resize하여 분류 모델의 input으로 사용함.
- CNN 모델 학습시 데이터 셋을 train과 test set으로 활용하기 위해 전체 데이터를 8대2로 분류하고 train set에서 다시 train과 validation set을 8대2로 분류하여 사용함. Train set은 학습용도, validation set은 과적합을 막기 위한 용도, test set는 모델의 정확도를 테스트하기 위한 용도로 사용함.
- 하이퍼파라미터(학습률, 옵티마이저, 배치사이즈 등)를 최적화하여 정확도 성능을 고도화함. Hierarchical 분석을 위해 whole 이미지, chromosome 이미지, genetic region 순으로 분류 모델을 생성하고 test set(전체 데이터의 20%)으로 모델 테스트를 진행함.
- Test set의 결과를 정렬하여 가장 유사한 chromosome 및 region을 선별하여 계통 특이적 마커 추출 후보로 선정함.

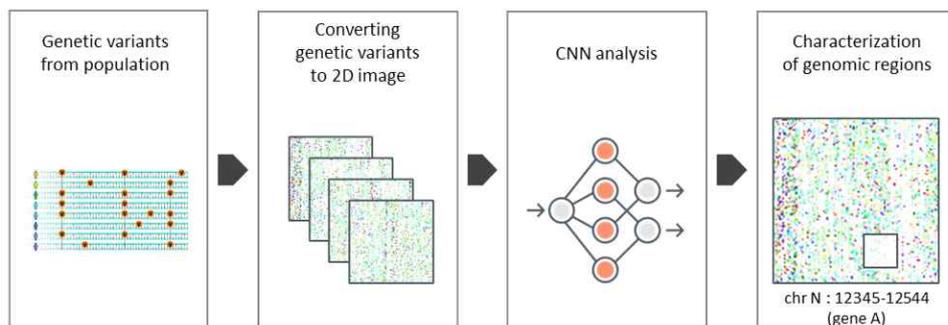


그림 2. 인공지능을 활용한 특이적 마커 추출 프로세스

- 전체 156개 배추 개체의 20%에 해당하는 32개 test set을 생성된 분류 모델에 입력한 결과, 과제 시작 이전의 선행 연구결과에서 나타난 정확도(81.5%)를 초과하는 새로운 이미지 분류 모델(분류 정확도 87.5%)을 생성하였음(그림 3).
- 구축한 이미지 분류 모델을 활용하여 예측한 결과를 실제 accession 정보와 대조한 혼동행렬을 구성하여 모델의 성능을 확인하였음(그림 4).

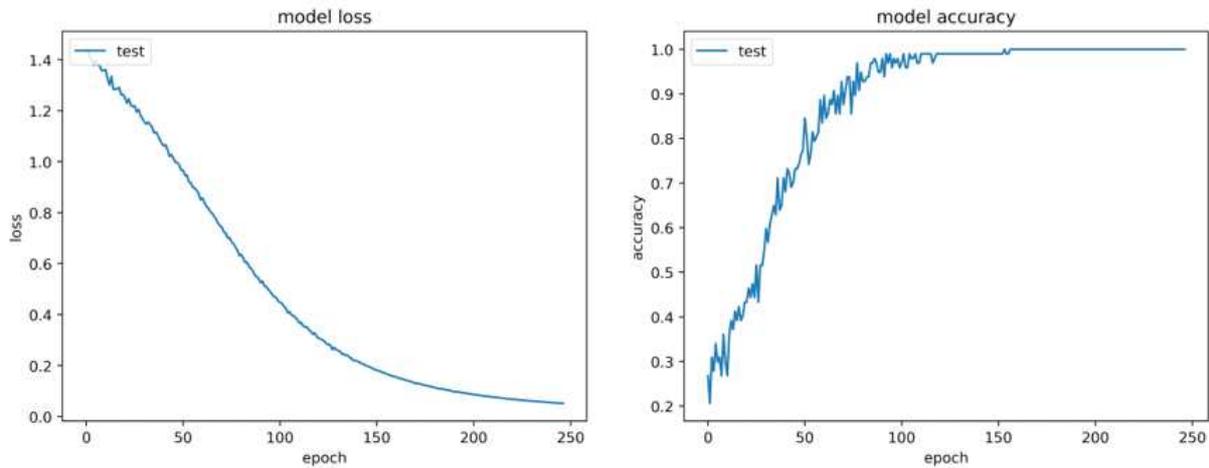


그림 3. 인공지능 기반 유전체 이미지 분석 모델의 test set 분류 결과

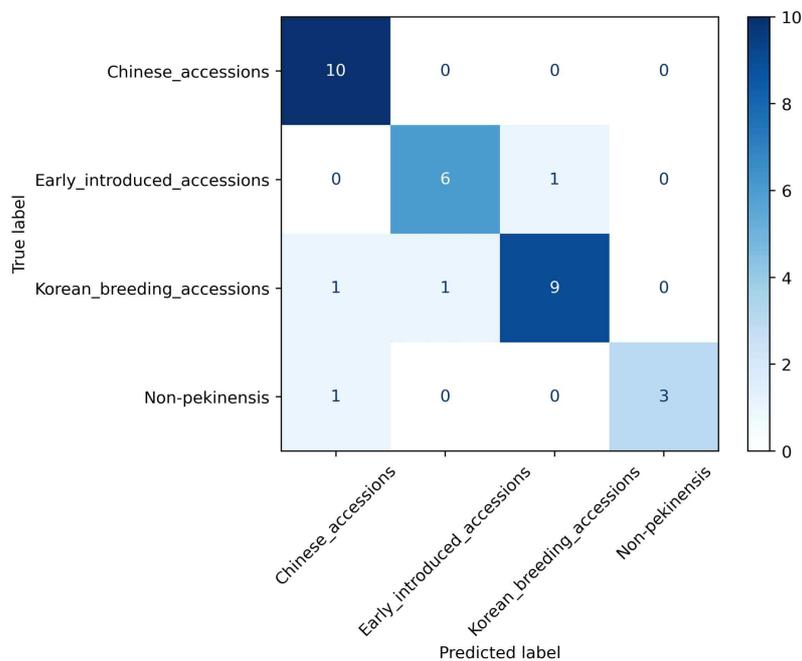


그림 4. 인공지능 기반 유전체 이미지 분석 모델의 테스트셋 분류 결과 - 혼동 행렬 (confusion matrix)

○ 그룹 특이적 변이 탐색 파이프라인의 구성

- 작물의 유전연구 및 육종을 위해 구축한 핵심집단의 경우, 다양한 야생종, 지역별 재래종, 및 육종활용 계통으로 구성된 것을 확인함.
- 식물의 집단유전학적 분석 및 문헌 조사를 통해 배추 핵심집단을 크게 4개의 그룹으로 구분이 가능한 것을 확인함(그림 5.A).
- 배추 156개 계통으로 구성된 집단에서 식별된 3,834,809개 SNP를 활용하여, 4개의 그룹에서 나타나는 6개 조합에서 조합별로 그룹을 구분할 수 있는 마커를 개발하였음(그림 5.B).

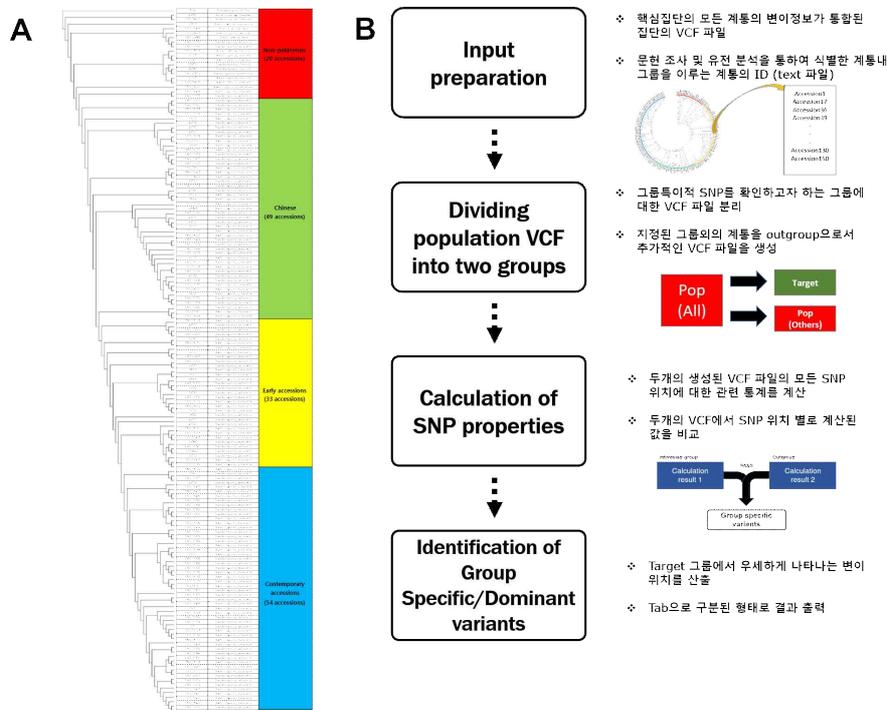


그림 5. 배추 156개 계통의 분류 정보를 활용한 그룹특이적 SNP의 개발 개요, A) 배추 156개 계통의 분류 정보, B) 배추 집단 변이 데이터를 활용한 그룹특이적 변이의 식별 파이프라인

- bcftools 소프트웨어를 활용, 집단 변이정보를 갖는 vcf 파일에서 지정된 2개 그룹에 해당하는 계통들의 변이정보만을 추출하여 새로운 vcf 파일을 구성함(그림 6).
- perl script를 활용, 새로운 vcf 파일 내 2개 그룹에서 minimum allele frequency, heterozygosity, missing rate 등을 계산 및 dominant allele 패턴을 구분함(표 1).
- Target group에서 alternative allele이 우세하고( $10\% \geq \text{MAF}$ ,  $10\% \geq \text{heterozygosity}$ ,  $10\% \geq \text{missing rate}$ ) control group에서 reference allele이 우세하게( $20\% \geq \text{MAF}$ ,  $10\% \geq \text{heterozygosity}$ ,  $10\% \geq \text{missing rate}$ ) 나타나는 SNP를 그룹특이적 SNP로서 식별

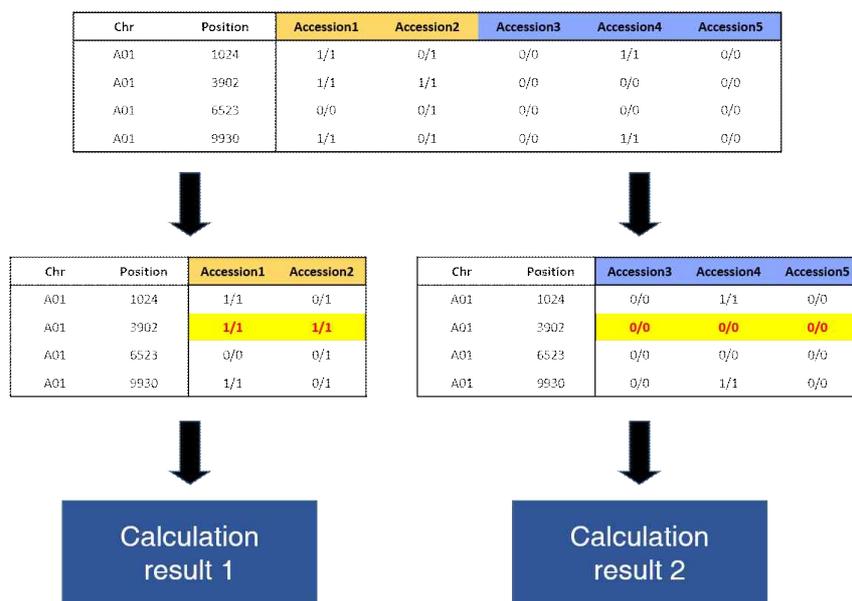


그림 6. 그룹 특이적 SNP 식별을 위한 VCF 파일 분리 및 dominant allele 식별 과정

**표 1.** 그룹 특이적 SNP 식별을 위해 반영되는 그룹별 변이 통계 정보

#	Category	Description	Value
1	SNP	크로모좀/스캐폴드상의 위치	Ex) A01_1024
2	Group name	파이프라인에 의해 입력된 그룹의 이름	입력된 텍스트
3	Majorallele	집단 내에서 우세한 allele의 개수	Number
4	Minor allele	집단 내에서 소수인 allele의 개수	Number
5	MAF	Minimum Allele Frequency (MAF)	Number
6	Missing allele	집단 내에서 genotype이 나타나지 않은 allele의 개수	Number
7	No.of homozygous line	집단 내에서 해당 SNP 위치가 동형접합인 계통의 개수	Number
8	No. of heterozygous line	집단 내에서 해당 SNP 위치가 이형접합인 계통의 개수	Number
9	No. of missing line	집단 내에서 해당 SNP 위치의 genotype이 나타나지 않은 계통의 개수	Number
10	No. of total line	VCF파일내의 총 계통의 개수	Number
11	Homozygous rate	해당 SNP 위치에서 homozygous allele을 갖는 계통의 비율	L %
12	Heterozygous rate	해당 SNP 위치에서 heterozygous allele을 갖는 계통의 비율	M %
13	Missing rate	해당 SNP 위치에서 missing인 계통의 비율	N %
14	Total proportion	전체 비율	100 (L+ M + N) %
15	Dominant alternative	Major allele이 alternative allele과 같음을 판정한 결과	O/X

- 배추 156개 계통은 *Brassica rapa*의 야생종 그룹(Non-pekinensis accession), 중국 계통(Chinese germplasm), 육종 초기 도입 계통(Early introduced accession), 및 현재 육종활용 계통(Korean breeding accession)로 구분됨.
- 배추 집단으로부터 총 6개의 비교 조합을 구성하였으며 조합별로 그룹특이적 SNP를 식별하였음(표 2).
- Group 6에 해당하는 Early introduced accessions와 Korean breeding accessions의 경우, 구축한 파이프라인으로부터 그룹 특이적 SNP가 나타나지 않아 target group과 control group의 MAF로 정렬한 SNP 정보에서 상위 40개 SNP를 식별함.
- 유전적 근연관계가 먼 Non-pekinensis와 다른 그룹 사이에서는 2,374 ~ 9,901개의 그룹 특이적 변이가 산출되었으며, 배추 육종 과정이 진전됨에 따라 식별된 그룹 특이적 변이의 개수가 증가하는 것으로 나타남.
- 각 그룹별 상위 40개 SNP를 대상으로 변이 데이터 상의 그룹 특이적 패턴을 확인함(그림 7).

**표 2.** 배추 집단내 분류의 조합별로 식별된 그룹특이적 SNP 개수

Group	Target group	Control group	Number of group specific SNPs
Group1	Non-pekinensis	Chinese	2,374
Group2	Non-pekinensis	Early introduced	4,383
Group3	Non-pekinensis	Korean breeding	9,901
Group4	Chinese	Early introduced	41
Group5	Chinese	Korean breeding	1,756
Group6	Early introduced	Korean breeding	40*

\*그룹 특이적 SNP 관련 계산 통계 상위 40개 SNP

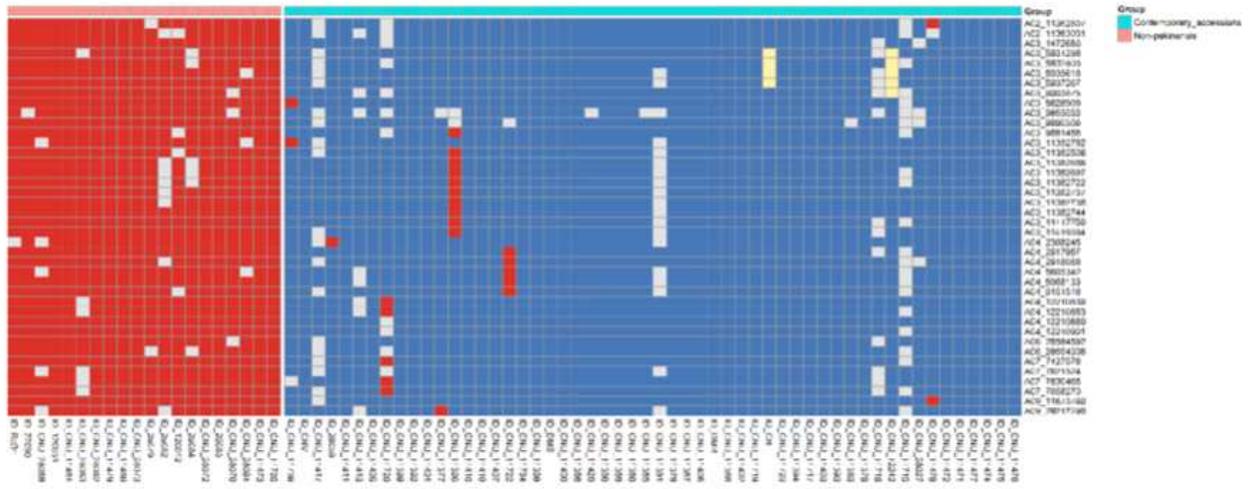


그림 7. Non-pekinesis와 Korean breeding group 간의 상위 40개 그룹 특이적 SNP(Red: Alternative allele, Blue: Reference allele, Yellow: Heterozygous allele, Grey: missing allele)

- Primer3를 활용, target SNP를 left primer의 end position에 고정된 primer 서열을 구성하였으며, 6개 그룹에서 총 63개의 프라이머 서열을 구성함(표 3).
- 본 성과를 활용한 특허 출원 절차가 진행 중

표 3. 배추 집단내 그룹 구별을 위한 KASP 마커 개발 결과

Group	SNP position	Forward(FAM)	Forward(HEX)	Reverse
Group1	SNP_A01_28209220	GTTCTGCTACGCATCAGTGTCTCT	GTTCTGCTACGCATCAGTGTCTCC	AACCGGGCGAGGAGTGTGTC
	SNP_A03_1440680	GACTAAAGCTTCTACGGGTATAATGC	GACTAAAGCTTCTACGGGTATAATGT	AGTTACTTACTTTCCCTCTGCT
	SNP_A03_822149	TCTTCTTTCTCGATTAAGTACGCC	TCTTCTTTCTCGATTAAGTACGCCA	CTGTATGCATTAGATCCAGCATTAT
	SNP_A03_5083670	ACTTGTGGTGACTTAGTTCAAAGAC	ACTTGTGGTGACTTAGTTCAAAGAT	AGGCTCGTGTCTGATGTTATCA
	SNP_A03_12209934	ACAAGTGTGTTATCATGTTATCAGTAG	ACAAGTGTGTTATCATGTTATCAGTAA	AGATGGCTTCACTTCTCTCGT
	SNP_A03_12214598	GATTATCCGAGITGAAATCGATGCTT	GATTATCCGAGITGAAATCGATGCTC	TGAGAAITGTTATGAGCGTTAAGTCA
	SNP_A04_12210838	CAITGGCCTACATGGTCTAATTCATA	CAITGGCCTACATGGTCTAATTCATT	ACCACACATGTAGTAAAAGGTGAC
	SNP_A06_1346319	AGCAAAAACGGAGGAAAGCTTCTTA	AGCAAAAACGGAGGAAAGCTTCTTG	GCGTAAGAGGATAGTGTGAAAGCG
	SNP_A07_7621524	TTCAAAGGGCTTCTCTTTCCGTCTC	TTCAAAGGGCTTCTCTTTCCGTCTA	GAGACTAGGTCGGAATAAGCT
	SNP_A07_7630465	GAGTAACAACTAACATCACATCGAAG	GAGTAACAACTAACATCACATCGAAT	TCATTCTAACTGGTTTTTCGATAACT
	SNP_A08_2627388	CGCATTTCAGATGGTGGTAAAGTGAA	CGCATTTCAGATGGTGGTAAAGTGAG	AGTCAAATTTAIGTAAAACCTCGGTCA
	SNP_A09_18791040	AACCCTATTGAAAAGGAAAGACTCC	AACCCTATTGAAAAGGAAAGACTCT	CATCAACAACAAGATAACTCAGGTCA
Group2	SNP_A06_21097459	TGTAATAACCAAGTTTGACCACAATT	TGTAATAACCAAGTTTGACCACAATA	TTTATCCGGACAACCTCCATGT
	SNP_A06_21097595	AACCTGAAACGATTTACTTGGAAGTT	AACCTGAAACGATTTACTTGGAAGTC	GACTTTTCTAGACGACTTACAG
	SNP_A06_21102003	CACTGATACAGCCGTAATAGTTGC	CACTGATACAGCCGTAATAGTTGT	GTAATGTGGTCATATCGAATCATAT
	SNP_A06_21102006	TGATACAGCCGTAATAGTTGCGTIG	TGATACAGCCGTAATAGTTGCGTIC	CTAGTAATGTGGTCATATCGAATCACA
	SNP_A06_21102014	CCGTAAATAGTTGCGTGAACCAAAT	CCGTAAATAGTTGCGTGAACCAAAC	TGGAGACTCTAGTAAATGTGGTCA
	SNP_A06_21102017	GTAATAGTTGCGTGAACCAAATGAG	GTAATAGTTGCGTGAACCAAATGAA	TGATGGAGACTCTAGTAAATGTGGT
Group3	SNP_A03_5935616	AACTTCATTGAAAGTTGTAAGCAAG	AACTTCATTGAAAGTTGTAAGCAAC	AACATTGCAGGTTCTATGACA
	SNP_A03_5937267	ATAACAAGACAATCTGGATAAGAAACA	ATAACAAGACAATCTGGATAAGAAACG	TGGAATCACTGCAATCATCTCG
	SNP_A03_11382697	CATGATAAGTAAAAGACAAGAAAGTC	CATGATAAGTAAAAGACAAGAAAGTT	TGGGTTTATCATGGGGTAGTIT
	SNP_A03_11352762	ACAATATCTCTAAACCAGTATGAACAG	ACAATATCTCTAAACCAGTATGAACAA	TGATGATGATCCAGTCCITATGTT

	SNP_A03_11382508	TGCAACAACGAGAAAAATCCTTCAGC	TGCAACAACGAGAAAAATCCTTCAGA	TAGAGTTGCCTCGGTATACTACTACT
	SNP_A04_12210838	CATTGGCTACATGGTCTATTTCATA	CATTGGCTACATGGTCTATTTCATT	ACCACACATGTAGTAAAAGGTGAC
	SNP_A07_7621524	TTCAAAGGGCTTCTCTTCCGTCTC	TTCAAAGGGCTTCTCTTCCGTCTA	GAGACTAGGGTCGGAAATAAGCT
	SNP_A07_7630465	GAGTAACAACTAACATCACATCGAAG	GAGTAACAACTAACATCACATCGAAT	TCATTCTAACTGGTTTTTGATAACT
Group4	SNP_A01_3301675	ATTACCGATGTGAATGGTCTCTCAC	ATTACCGATGTGAATGGTCTCTCAA	CAAACGTGATTAATACATACGTCCA
	SNP_A02_30651212	ACATCCTCAGCTTCAATCTCTTCAT	ACATCCTCAGCTTCAATCTCTTCAA	TTGTCTAACGTTCCACCCGGC
	SNP_A02_30651232	TCTTCATACAACGCATAGTTAATCCAC	TCTTCATACAACGCATAGTTAATCCAT	AGAAGTTTACGAGAGGGCTGT
	SNP_A02_30651277	GTAACATTCTCATTAGCCGGTGGGA	GTAACATTCTCATTAGCCGGTGGG	AGAGGAAGAGGTGGCTTCTTC
	SNP_A02_30651284	TTCTCATTAGCCGGTGGAACGTTAG	TTCTCATTAGCCGGTGGAACGTTAC	TTAGGTTAGAGGAAGAGGTGGC
	SNP_A03_18195518	ACACGATTACTCTAGAGATTCCCA	ACACGATTACTCTAGAGATTCCCG	CTCTCTGCACGCAAGCTGC
	SNP_A03_18195519	CACGATTACTCTAGAGATTCCCAA	CACGATTACTCTAGAGATTCCCAC	GCTCTCTGCACGCAAGCT
	SNP_A07_19139938	AAACCATGCATCAAACATAAACTAAAT	AAACCATGCATCAAACATAAACTAAAG	CAACGCAAGAGATTCCCCT
	SNP_A07_19140181	AAGTGTGTCGGTGAAGTGGGAATAA	AAGTGTGTCGGTGAAGTGGGAATAG	CGGAGTCGGATAAAGGCCTA
	SNP_A07_19140182	AGTGTGTCGGTGAAGTGGGAATAAC	AGTGTGTCGGTGAAGTGGGAATAAT	TCCGAGTCGGATAAAGGCCT
	SNP_A07_19173020	CAAGACCTCGAATCGTAATAATCTTTC	CAAGACCTCGAATCGTAATAATCTTTA	TCCTTGAACACCTTGCAAGC
Group5	SNP_A07_19156259	GCTTCTCCAAATGACTCTTATAAGTA	GCTTCTCCAAATGACTCTTATAAGTC	CACTGATTTAGAGAGGTTTTGTTCCA
	SNP_A08_14224373	TCATACAATGAGATTGTCTCTAGTGC	TCATACAATGAGATTGTCTCTAGTGA	CCAAGCCAATTCACATCATCATCA
	SNP_A08_14224831	ATGCATCTTCCAAATCGGTGAGTTA	ATGCATCTTCCAAATCGGTGAGTTG	AACATTAGACTCAACAACACATGT
	SNP_A08_14225137	GTTTGCTATGTGCCCTTTGCTTACAG	GTTTGCTATGTGCCCTTTGCTTACAA	AGAAAGAGGTGGAGACGATTCTG
	SNP_A08_14223077	ACTGACACCTGCAAAAATGTTGACT	ACTGACACCTGCAAAAATGTTGACG	AGTAAAGTTTCCATCTTTCCAACCTAC
	SNP_A08_14223929	AGATCCTGTTCAAGTAAGTCTGTGG	AGATCCTGTTCAAGTAAGTCTGTGA	TAGGAAGTAAAAGTCCCACATCCG
	SNP_A08_14224023	ACTTCTAACTTATTAGCTCCTTTTAC	ACTTCTAACTTATTAGCTCCTTTTAT	CCATAAGATAAAGCCACTTGTCTTT
	SNP_A08_14224436	AGGTTTTTCGTTGATGATGATGATGT	AGGTTTTTCGTTGATGATGATGATGG	TCTTCCAAGTCTCCAACCTCTCT
	SNP_A08_14224834	CATCTTCCAAATCGGTGAGTTATGA	CATCTTCCAAATCGGTGAGTTATGT	TTCAACATTAGACTCAACAACACA
	SNP_A08_14225035	TATGTACTCATGCCCTTTTATAGACT	TATGTACTCATGCCCTTTTATAGACA	TCAAACCTGGGACTAACACAA
	SNP_A08_14225072	CATTTACTACAGTTAAGTTCCCCAGG	CATTTACTACAGTTAAGTTCCCCAGA	TTTACATATCTGTAAGCAAAGGCACA
	SNP_A08_14225075	TACTACAGTTAAGTTCCCCAGGTTA	TACTACAGTTAAGTTCCCCAGGTTT	AGTCTTTACATATCTGTAAGCAAAGGC
	SNP_A08_14225118	GTTAGTCCCAGTTTGTATGTTTGC	GTTAGTCCCAGTTTGTATGTTTGG	TCTGACTCGCTCTCACATCA
	SNP_A08_14229644	GCTTAAGATTTTCGAAAAGGTAGGATC	GCTTAAGATTTTCGAAAAGGTAGGATG	TGTTCTGAGTTGAAAGCACATGC
Group6	SNP_A09_2859289	GCCATTCTATTAACAACTATCGTATCG	GCCATTCTATTAACAACTATCGTATCA	TCCTTGTCTCGTGTGTTGGTGT
	SNP_A09_2859730	CCAGGGCTTAATCTCGACTTTACAA	CCAGGGCTTAATCTCGACTTTACAG	CACTGTGACGTGGCCAA
	SNP_A09_2859833	AAACAAGTCTTCTCGAATACGGA	AAACAAGTCTTCTCGAATACGGC	GAGTCCAAGTGGGATTTCG
	SNP_A09_2866527	AAGAGTGAACATAAGAGCCGTGAAG	AAGAGTGAACATAAGAGCCGTGAAA	GAGATCTCTGTAATCACTGGAGG
	SNP_A09_2866578	GGCTTACACTATCTACATACACATGAC	GGCTTACACTATCTACATACACATGAT	ACCAATCTTGACTTCTCTTGGT
	SNP_A09_5862593	GATGGAAGGGTAAACGTTGCAATA	GATGGAAGGGTAAACGTTGCAATG	GACCAAACTCAAACAAGGCTT
	SNP_A09_5862617	AGACTCTGCTGATGTTTTGAATAGC	AGACTCTGCTGATGTTTTGAATAGT	CGACATCGAAACTCGTTTTTGT
	SNP_A09_5870163	TAAATGAACGGGAAGAAGTGAATA	TAAATGAACGGGAAGAAGTGAAT	TGCTCGATTGAAAAGGGACCT
	SNP_A09_2852967	CTCTGTCTTAAAGTTAATATGACCGTT	CTCTGTCTTAAAGTTAATATGACCGTG	TAATCATACGTAACCAATCTCTTAGT
	SNP_A09_2860541	TGATTTTTGAATTTTCTTGGCACACG	TGATTTTTGAATTTTCTTGGCACACT	TTTTGTGACGGTTAATCTGAATAACT
	SNP_A09_2863077	GATCGACCGACATTAACGTGAAATC	GATCGACCGACATTAACGTGAAATT	CAACAAAAGGAAATGAAATCACAGA
SNP_A09_37236608	AACACCGATTATCTTGATTACGTAA	AACACCGATTATCTTGATTACGTAC	GTGCGATATATTATAACCACACT	

○ 표현형 데이터 생산 및 표현형 분류 프로그램의 개발

- 배추의 발달 단계별 표현형 데이터 및 외부 공개 데이터베이스의 식물의 병징과 관련된 이미지를 생산 및 확보하였음.
- 배추의 표현형 기반 계통의 그룹 분류 프로그램 개발에 관련된 논문을 출판하였음(Frontiers in Plant Science. 2021; 12: 738685., 그림 8).

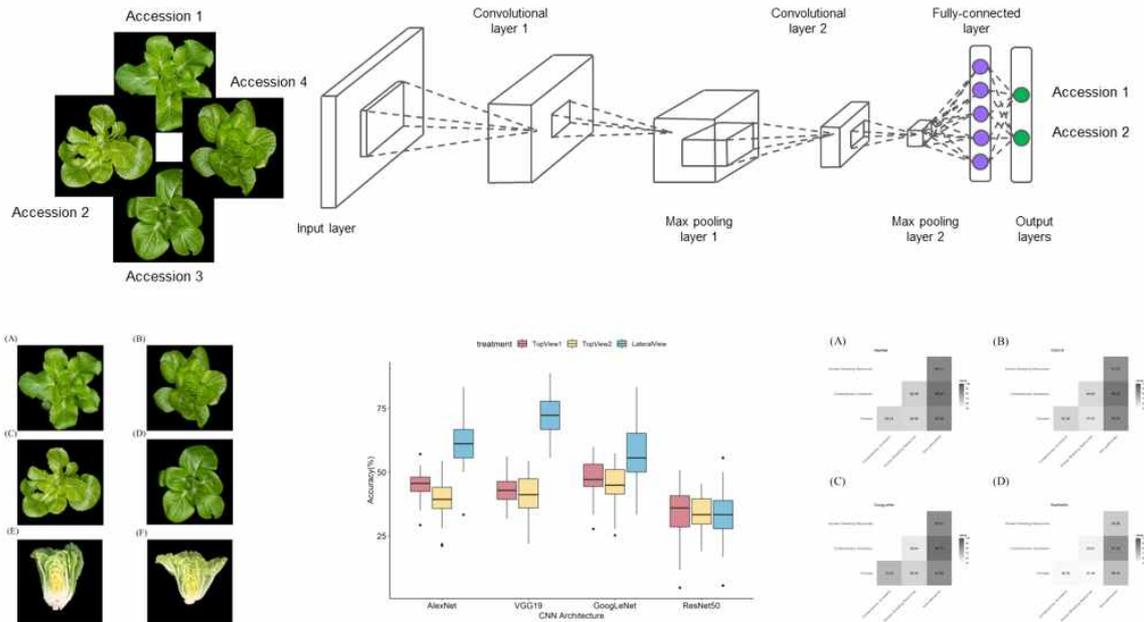


그림 8. 출판된 논문의 주요 결과

- 타 작물의 표현형 데이터 분류 모델 구축을 위해 산림청, 경상대학교, 연암대학교와의 협업을 통해 작물의 이미지 데이터의 추가 생산을 진행 중임.
- 인공지능을 활용한 표현형 데이터 분류 모델 생성을 위해 normal과 disease(특정 질병)의 데이터를 획득하여 noise 제거 등의 전처리 후 annotation을 수행하였음.
- 공개 데이터베이스인 Plant village에서 식물의 병징과 관련된 이미지를 확보하였음(그림 9, 표 4).



그림 9. Plant village의 식물 병 이미지 예시, 건강한 잎(상)과 병든 잎(하)의 비교

표 4. 식물 병 이미지 확보 결과(출처: Plant-village)

Species	Disease	Count
Apple	Healthy	1,645
	Apple scab	630
	Black rot	621
	Cedar apple rust	275
Cherry	Healthy	854
	Powdery mildew	1,052
Corn	Healthy	1,162
	Cercosporaleaf spot	513
	Common rust	1,192
	Northern leaf blight	985
Grape	Healthy	423
	Black rot	1,180
	Esca(Black Measles)	1,384
	Leaf blight	1,076
Peach	Healthy	360
	Bacterial spot	2,297
Pepper	Healthy	1,456
	Bacterial spot	929
Potato	Healthy	148
	Early blight	929
	Late blight	734
Strawberry	Healthy	456
	Leaf scorch	1,109
Tomato	Healthy	1,217
	Bacterial spot	1,670
	Early blight	302
	Late blight	467
	Leaf mold	334
	Septorialeaf spot	497
	Spider mites	1,354
	Target spot	740
	Tomato mosaic virus	281
	Tomato yellow leaf curl virus	1,704

- 인공지능 분석 방법을 이용하여 농작물의 잎에서 발생하는 질병을 판별하는 모델을 생성 및 합성곱신경망분석을 활용하여 1) 식물의 종 분류, 2) 식물 병 발생 여부 및 3) 식물 병 분류 모델 구축을 수행하였음
- PlantVillage에서 확보한 공개 DB의 이미지는 서로 다른 작물 9종에서의 건강한 잎과 병든 잎으로, 이 중 가지과 작물인 토마토, 감자, 고추의 이미지 데이터를 활용하여 단계별 질병 탐색 모델을 구축하였음(그림 10).

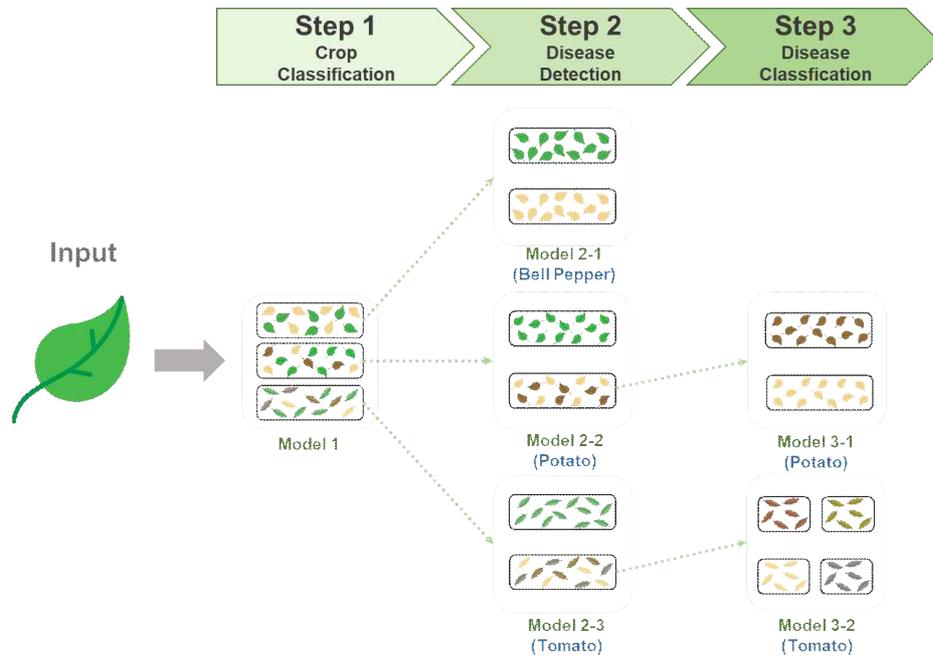


그림 10. 인공지능을 활용한 단계별 작물 질병 탐색 모델

- 다양한 architecture의 pre-trained CNN model을 사용하여 단계별로 분류 모델을 생성하였고, fine-tuning을 통해 각 단계별로 최적의 분류 모델을 생성하였음(표 5).
- 실제 사용을 가정하여 병반만을 취한 이미지, field 생산 이미지뿐만 아니라 분류 모델에 사용하지 않은 작물의 이미지를 이용하여 질병 분류 모델을 검증하였음(그림 11).

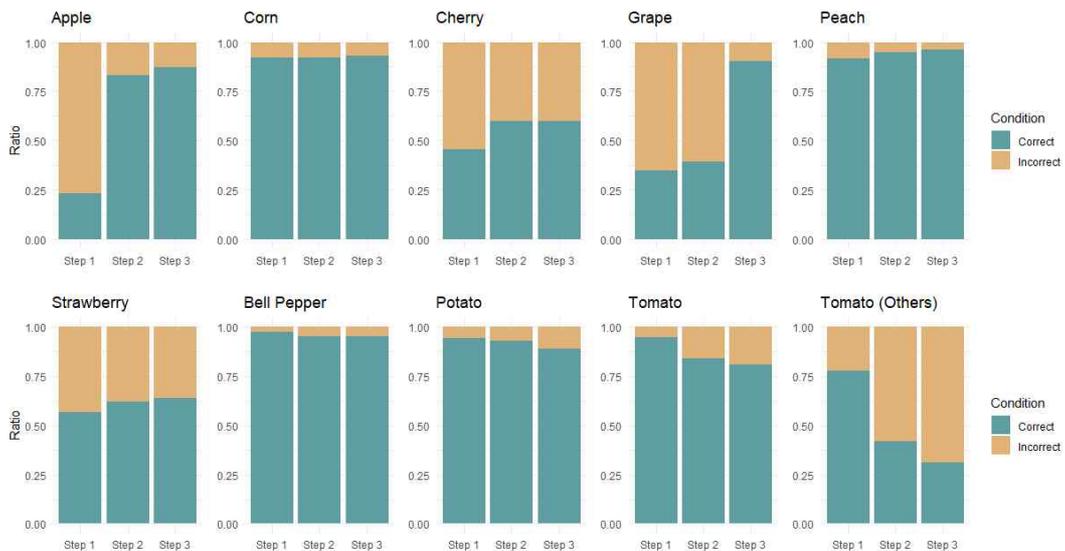


그림 11. 모델 구축에 사용하지 않은 데이터를 이용한 분류 분석 테스트 결과

- 생산한 단일 데이터 및 외부에서 생성된 다른 데이터셋을 활용하여 분류 모델을 검증함.
- 한국지능정보사회진흥원(National Information Society Agency, NIA)의 AI 통합플랫폼 AI-hub(<https://www.aihub.or.kr/>)에서 제공하는 인공지능학습용 데이터를 활용하여 테스트를 실시함.
- 배경이 제거되지 않은 노지 생산의 식물 작물 질병 진단 이미지 데이터 셋을 활용하여 검증을 실시한 결과, 74.55%의 정확도로 분류됨(그림 12).

표 5. 복수의 pre-trained CNN model을 사용한 검증 결과

Step		Crop	Pre-trained Model	Performance			
				Accuracy	Precision	Recall	F1-score
I	Crop Classification	All	ResNet50	91.84%	92.86%	91.84%	91.66%
			AlexNet	96.87%	96.87%	96.87%	96.86%
			GoogLeNet	99.08%	99.08%	99.08%	99.08%
			VGG19	98.71%	98.71%	98.71%	98.71%
			EfficientNet	99.33%	99.33%	99.33%	99.32%
II	Disease Detection	Bell Pepper	ResNet50	98.32%	95.90%	100.00%	97.91%
			AlexNet	99.16%	99.46%	98.40%	98.93%
			GoogLeNet	100.00%	100.00%	100.00%	100.00%
			VGG19	99.58%	99.47%	99.47%	99.47%
			EfficientNet	99.58%	100.00%	98.93%	99.46%
		Potato	ResNet50	99.45%	100.00%	99.40%	99.70%
			AlexNet	98.90%	99.70%	99.10%	99.40%
			GoogLeNet	99.45%	99.70%	99.70%	99.70%
			VGG19	100.00%	100.00%	100.00%	100.00%
			EfficientNet	99.45%	100.00%	99.40%	99.70%
		Tomato	ResNet50	99.75%	99.75%	99.75%	99.75%
			AlexNet	99.45%	99.50%	99.50%	99.49%
			GoogLeNet	99.62%	99.62%	99.62%	99.62%
			VGG19	99.62%	99.62%	99.62%	99.62%
			EfficientNet	98.23%	98.33%	98.23%	98.24%
III	Disease Classification	Potato	ResNet50	98.80%	99.31%	97.96%	98.63%
			AlexNet	99.40%	99.32%	99.32%	99.32%
			GoogLeNet	99.40%	100.00%	98.64%	99.32%
			VGG19	99.40%	100.00%	98.64%	99.32%
			EfficientNet	99.40%	100.00%	98.64%	99.32%
		Tomato	ResNet50	87.80%	90.31%	87.80%	88.39%
			AlexNet	95.45%	95.32%	95.45%	95.35%
			GoogLeNet	95.81%	96.29%	95.81%	95.82%
			VGG19	95.08%	94.92%	95.08%	94.96%
			EfficientNet	97.09%	97.19%	97.09%	97.12%

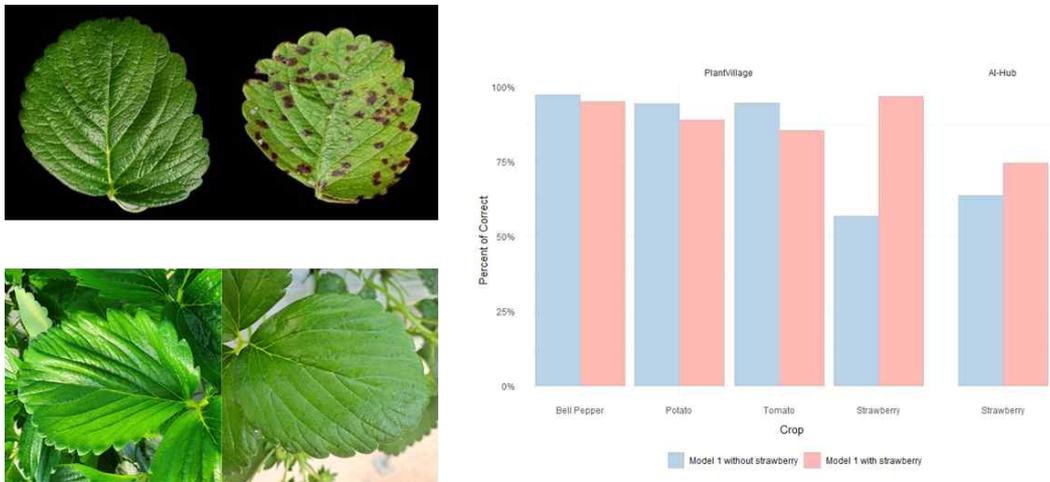


그림 12. 일반 field에서 촬영된 AI-hub의 딸기 이미지 데이터(왼쪽 아래)를 이용하여 질병 분류 모델의 검증 결과(오른쪽)

- 병반이 드러난 식물 잎의 일부분을 이용하여 구축한 질병 분류 모델을 검증함(그림 13). 감자와 토마토 질병 데이터에서 병반만을 취하여 분류 모델에 적용하여, 해상도 및 이미지 크기와 정확성의 상관관계를 검증함.
- 비교적 병반의 형태와 색이 선명한 질병 데이터(Early blight)의 경우는 온전하지 않은 이미지 또는 낮은 해상도의 이미지임에도 불구하고 높은 정확도의 분류가 가능하였으나 (정확도: 95.56%), 병반의 경계가 모호하고, 색이 불분명한 경우(Late blight)에는 분류 성능의 편차가 큰 것을 확인함(정확도: 70.62%)(그림 13).
- 따라서 다양한 작물의 질병 분류 모델의 구축 이전에 양질의 데이터를 확보가 선행되어야 함.

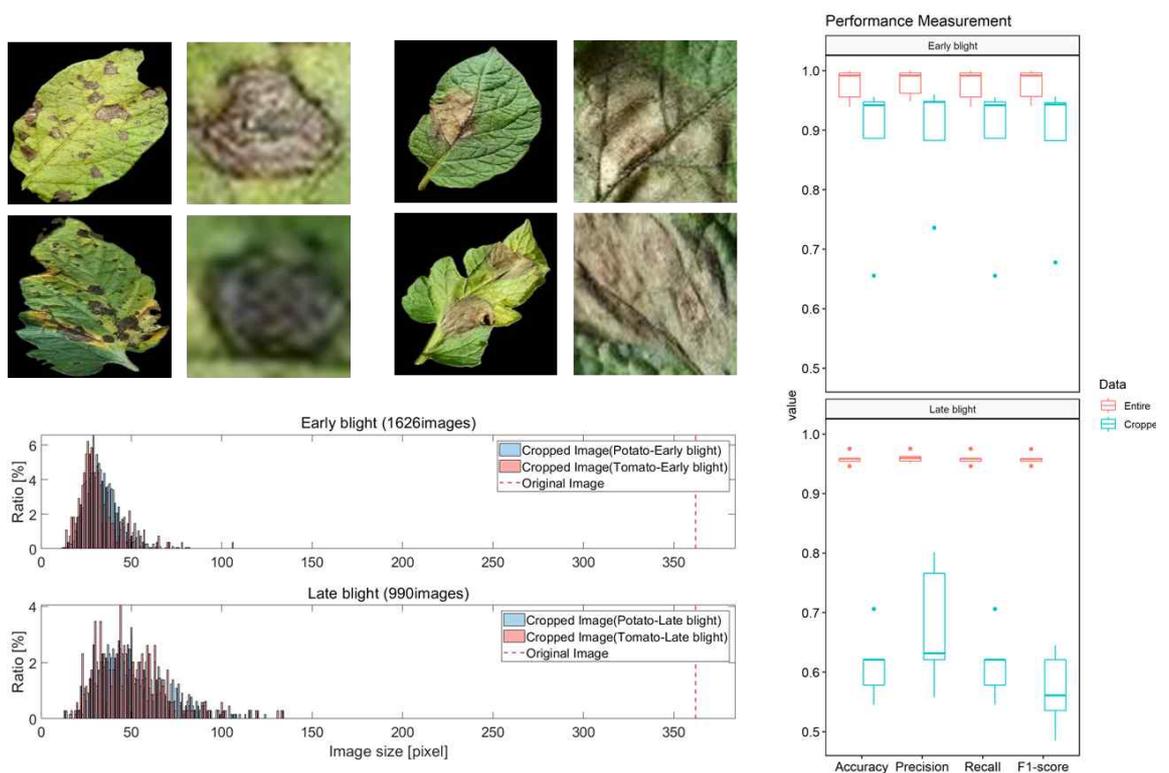


그림 13. 두 가지 질병의 병반 이미지만을 이용한 분류 모델의 성능 비교

○ 고추 표현체 데이터 확보 및 가공

- 경상대학교와의 공동연구를 통해 고추의 서로 다른 두 계통인 수비초와 MC4의 F1 집단에서 육성된 F2 집단의 과실에 대한 사진 이미지 데이터를 확보함.
- F2 집단 309개 개체에서 수확된 과실에 대하여 총 389장의 고화질 이미지 데이터가 균일한 조건으로 생산되었으며 이는 과실의 선단 부분의 형태에 따라 표현형이 총 3개의 형태로 분류가 되었음(그림 14, 표 6).



그림 14. 수비초와 MC4 계통의 F2 집단에서 나타나는 과실의 형태에 따른 분류의 예시

표 6. 수비초 X MC4의 F2 집단의 과실 표현형 분류 결과 (경상대학교 제공)

분류	둥근 과실	둥근 과실#	중간 과실	중간 과실#	각진 과실	각진 과실#	계
개수	60	12	155	40	90	32	389

# 해당 형태에 근접한 형질

- 확보된 고화질 이미지 데이터(Step 1)는 배경의 제거와 동일 축척의 설정(Step 2), 과실의 시작 위치 일치 및 이미지의 정사각형화(Step 3)로 작업 순서를 설정하고 표준화 작업을 수행 중임(그림 15).
- 본 이미지 데이터 및 편집 결과는 추후 생산 예정인 수비초 X MC4의 F2 집단의 변이 데이터와 함께, 유전체-표현체 연관분석 프로그램 개발에 활용될 예정임.

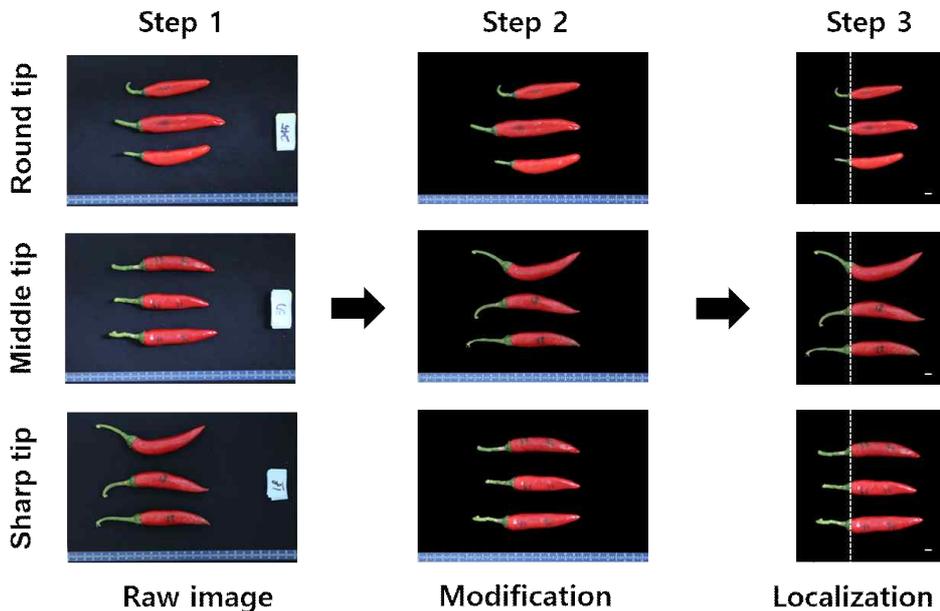


그림 15. 확보된 고추 과실 이미지의 계통별 편집 과정 개요(Step 1: 계통별 이미지 데이터의 확보, Step 2: 이미지 데이터의 배경 제거, Step3: 과실 시작 부분 및 축척 일치)

○ 오이 표현체 데이터 확보 및 인공지능 기반 질병 중증도 판별 모델 구축

- 연암대학교와의 공동연구를 통해 오이에서 발생한 흰가루병(Powdery mildew)의 질병 이미지 데이터 392장을 확보함.
- 작물 질병의 조기 진단을 위하여 질병의 중증도 구분 작업이 필요함. 오이 흰가루병 데이터를 전문가의 평가를 거쳐 초기, 중기, 말기의 3가지 단계로 구분함.
- 한국지능정보사회진흥원의 AI학습용 데이터 공개 데이터베이스인 AI-hub에서 오이의 정상 생육 이미지 데이터를 확보하여 AI 기반 작물 질병의 조기 진단 모델 구축 작업을 진행중임.
- 정상 샘플과 질병 초기 샘플을 AI 모델로 분류 분석을 실시한 결과, 89.5%의 정확도로 분류되었음(그림 16). 모델의 개선을 통해 작물 질병의 조기진단 모델을 구축할 예정임.

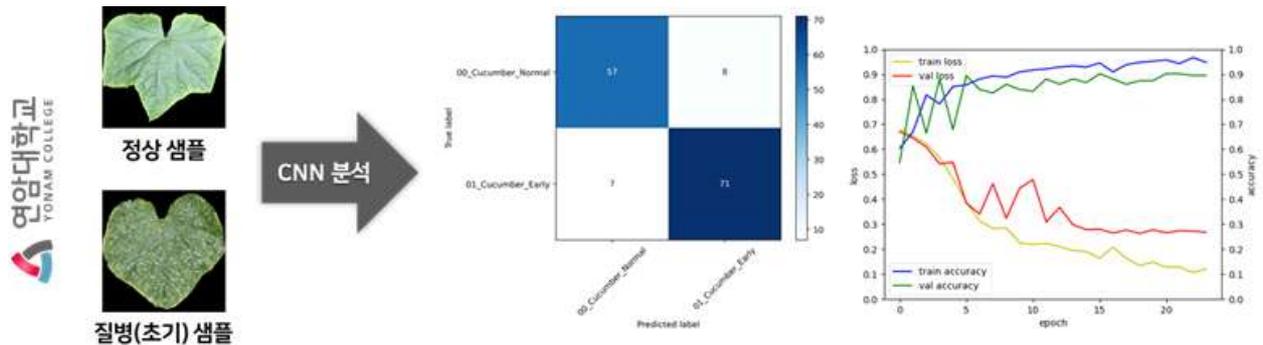


그림 17. 작물의 질병 중증도 구분을 위한 인공지능 분석

○ 유전체 데이터를 이용한 인공지능 기반 작물 표현형 예측 분석

- 배추의 158개 핵심집단의 유전체 데이터와 표현형 측정 데이터를 이용하여 인공지능 알고리즘 기반 표현형 예측 분석을 실시함.
- 핵심집단의 plant weight 표현형에 대한 수치형 데이터의 기준값을 설정하고, 세 가지 범주로 분류하였음. 측정이 누락된 샘플을 제외하고 총 137개 샘플을 47, 46, 44개의 High, Medium, Low 범주로 구분하여 분석함.
- pre-trained CNN model(VGG19)을 활용하여, training set과 test set을 각각 80%, 20%로 나누어 80%의 training set으로 표현형 예측 모델을 학습한 후, 20%의 test set으로 모델을 검증하였음(그림 17).
- 총 28개의 test set을 이용하여 모델을 검증한 결과, 64.29%의 정확도로 예측함.

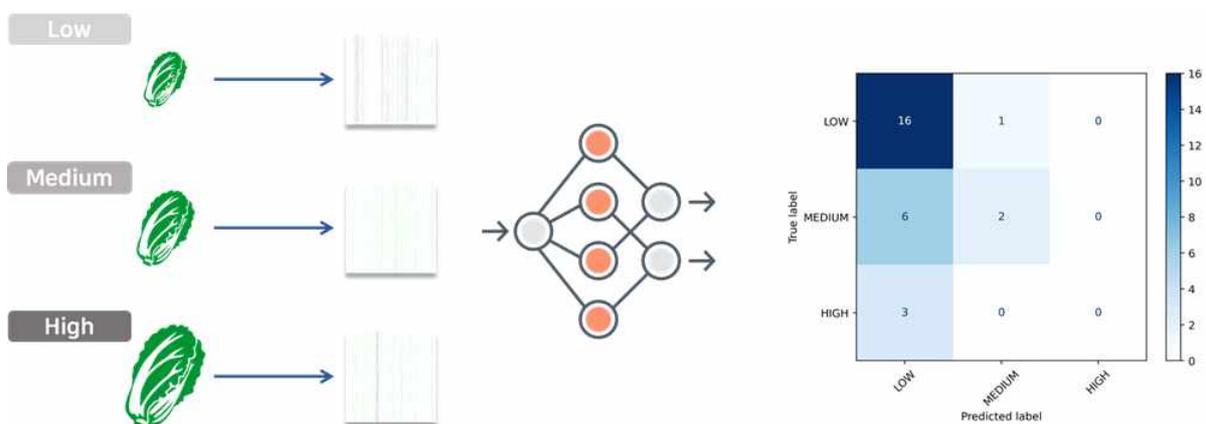


그림 18. 유전체 데이터를 이용한 표현형 예측 분석

○ 전사체 네트워크 분석 프로그램 개발

- 기능적으로 중요한 많은 수의 유전자들은 보통 발현 패턴이 일정하게 유지되고 있으며, 이러한 유전자들이 기존의 DEG 기반 분석에서는 제외되어 정확성이 떨어지는 단점이 있음.
- 예측된 유전자 사이의 상호작용 연구가 잘되어 있는 인간유전체와 달리, 상호작용의 규명이 진행 중인 식물에서는 DEG 기반의 네트워크 분석 결과가 낮은 정확성을 보일 것으로 우려됨. 이를 활용하여 개발된 형질전환 식물체의 경우, 예상치 못한 부작용(side effect)을 보일 수 있음.
- 이러한 문제점 해결을 위해, DEG 정보와 유전자 기능 기반(예: 생합성 경로)의 클러스터링을 활용한 유전자 상호작용 분석을 융합하여 전사체 네트워크 분석 프로그램을 개발하고자 하였음.
- 이러한 전사체 네트워크 분석 프로그램 개발을 위해, 일반 배추와 기능성 자색 배추와의 비교 분석을 통한 자색 형질 관련 유전자의 분리·동정을 위한 전사체 세트를 설계함(그림 18).

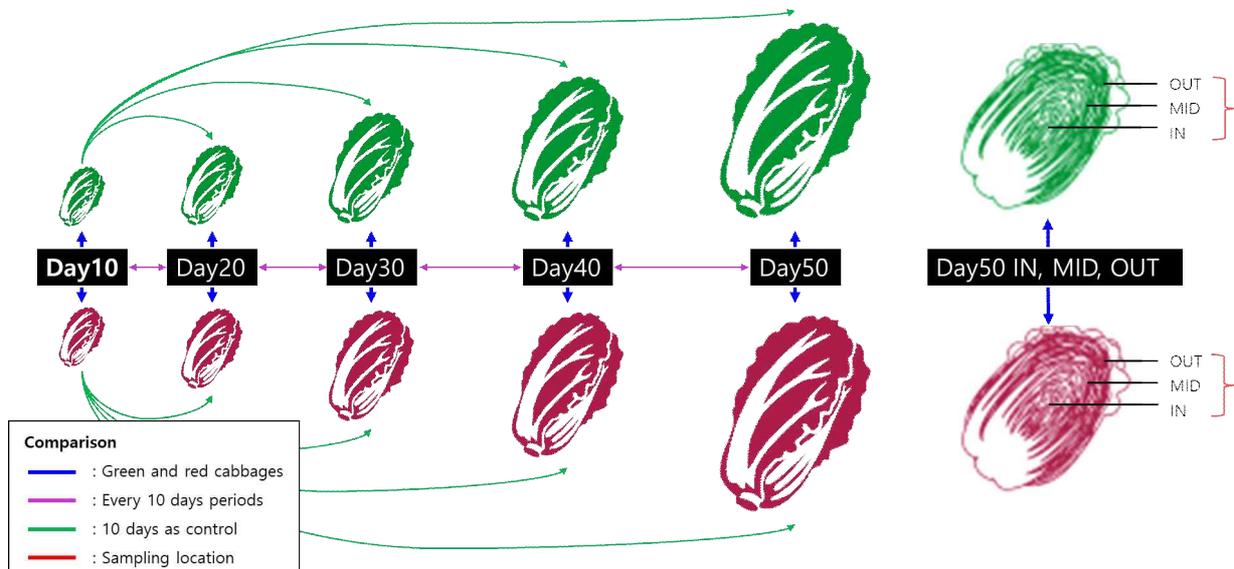


그림 18. 자색 배추 분석을 위한 전사체 데이터 세트 설계

- 샘플당 3개의 biological replicate를 두어 총 30개의 전사체 데이터를 생산하였으며, 샘플당 6Gb 내외의 데이터를 생산하였음.
- 전사체 데이터를 활용하여 Differential Expressed Gene(DEG)와 유전자의 발현 패턴 분석을 통해 일반 배추와 자색 배추의 생육 일자에 따른 DEG를 식별함(그림 19, 표 7).

표 7. 일반 및 자색 배추의 발달 단계에 따른 DEG의 산출결과

발달 단계	DEG 수
Day 10	3,116
Day 20	4,246
Day 30	2,851
Day 40	2,046
Day 50	1,938

Intersection		
Pattern	Green	Red
DDDU	0	2
DDUD	19	8
DDUU	1	1
DUDD	1	1
DUDU	1	10
DUUD	4	31
DUUU	0	18
UDDU	3	1
UDDU	3	4
UDUD	37	6
UDUU	2	1
UUDD	0	9
UUUU	5	0
UUUU	2	0
SUM	78	92

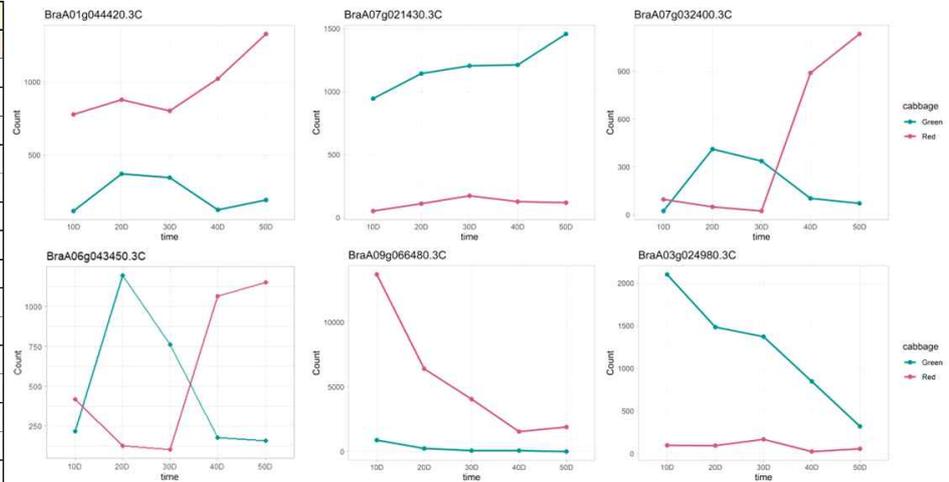


그림 19. 발달 단계에 따른 유전자 발현 패턴의 확인

- 산출된 DEG 유전자의 GO term 분석과 KEGG Pathway 분석을 수행하여, DEG를 clustering 하고, clustering 된 DEG를 기반으로 생합성 경로를 탐색함(그림 20).

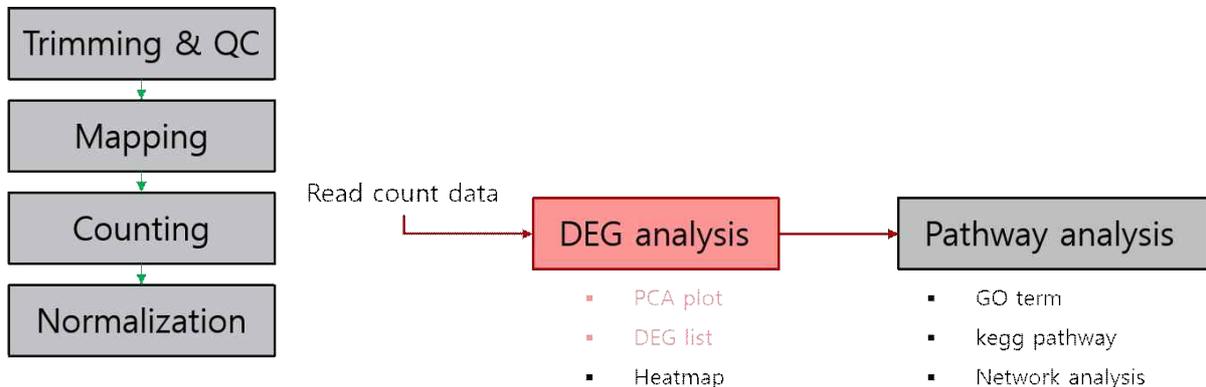


그림 20. 전사체 분석에서 산출된 DEG 분석의 work flow

- 다양한 조건에서 관련된 hub 유전자의 발굴을 위해 DEG의 GO term 및 유전자의 네트워크 분석을 수행하는 프로그램을 개발하고 있음(그림 21).

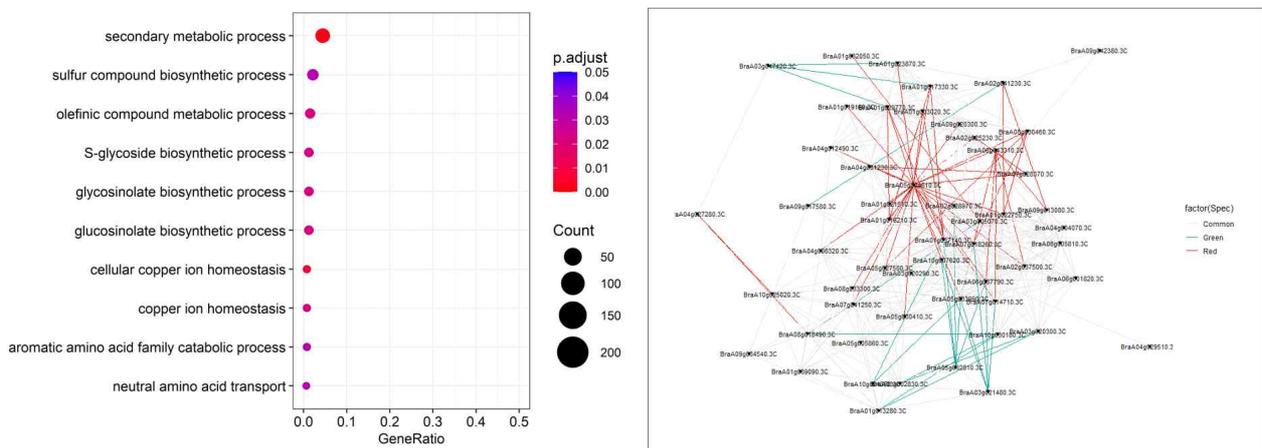


그림 21. 현재 개발 중인 유전자 네트워크 분석 프로그램 결과

- 실험 결과로 생성된 sequencing raw data를 전처리한 후 비교 그룹 간의 차등발현 유전자 (Differentially Expressed Gene, DEG) 분석을 실시하였으며(그림 22) 조건에 따른 DEG 결과(그림 23, 표 8) 및 GO term set을 구성함(그림 24). 기능 중심의 GO term set과 관심 유전자인 DEG set을 이용하여 그 연관성과 중요도를 탐색함. 연관성을 네트워크로 표현함.

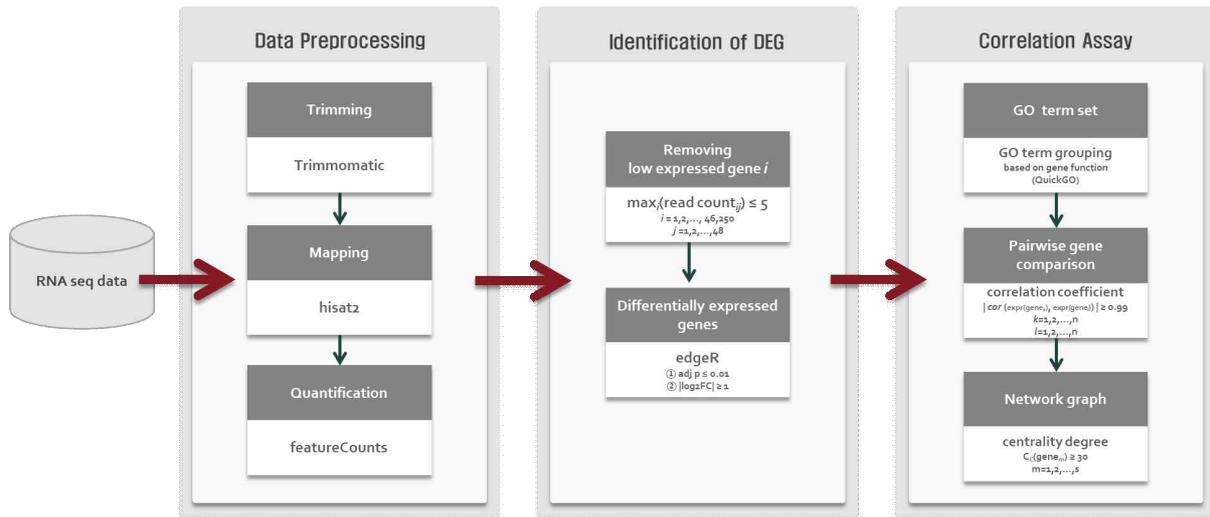


그림 22. 전사체 유전자 발현량 데이터를 이용한 네트워크 분석의 workflow

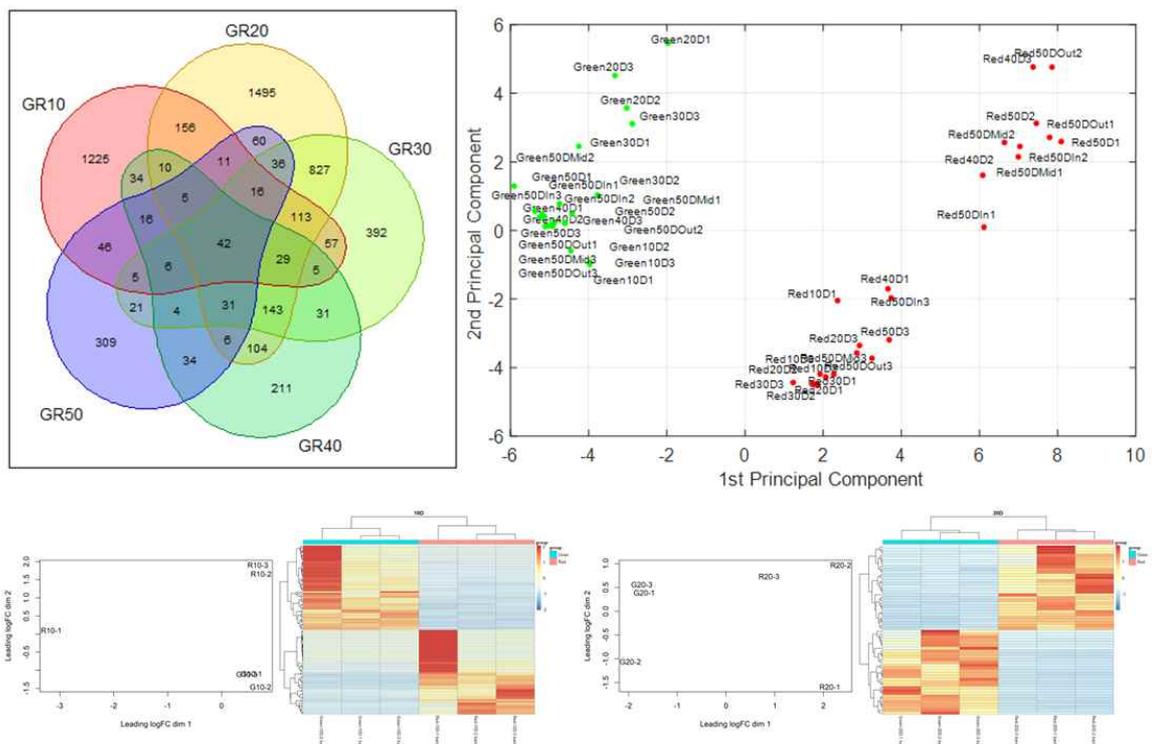


그림 23. 배추의 발달 단계별 유전자 발현량 데이터를 이용한 DEG 분석 결과

표 8. 일반 배추와 자색 배추의 발달 단계별 차등발현 유전자의 수

		Up	Down	Total	
Green vs. Red	Day 10	850	926	1,776	
	Day 20	1,483	1,601	3,084	
	Day 30	820	938	1,758	
	Day 40	358	353	711	
	Day 50	402	246	648	
	Day 50 In	716	428	1,144	
	Day 50 Mid	262	202	464	
	Day 50 Out	728	516	1,244	
Green	Developmental Stage	Day 10 vs. 20	3,707	3,498	7,205
		Day 10 vs. 30	3,764	3,639	7,403
		Day 10 vs. 40	724	343	1,067
		Day 10 vs. 50	2,070	1,910	3,980
		Day 20 vs. 30	105	45	150
		Day 30 vs. 40	747	393	1,140
		Day 40 vs. 50	0	4	4
Position	In vs. Mid	645	780	1,425	
	In vs. Out	2,037	2,464	4,501	
	Mid vs. Out	52	183	235	
Red	Developmental Stage	Day 10 vs. 20	679	1,189	1,868
		Day 10 vs. 30	525	811	1,336
		Day 10 vs. 40	865	1,522	2,387
		Day 10 vs. 50	532	957	1,489
		Day 20 vs. 30	39	8	47
		Day 30 vs. 40	1,150	727	1,877
		Day 40 vs. 50	48	64	112
Position	In vs. Mid	63	36	99	
	In vs. Out	157	366	523	
	Mid vs. Out	6	40	46	

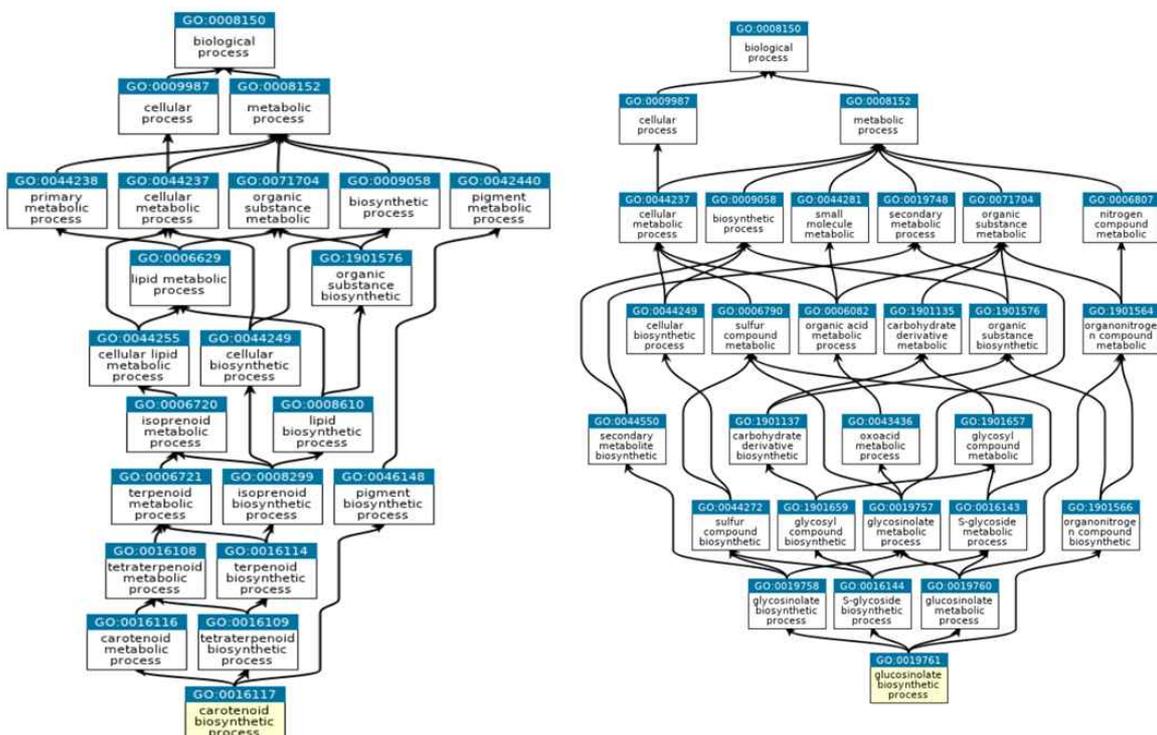


그림 24. 배추의 차등발현 유전자를 활용한 pathway 기반 GO term 분석 결과

- GO term set을 이용하여 연관분석을 시행하여 클러스터내의 유전자간의 연관성을 확인하였으며, 유전자를 점(node), 연관 관계를 선(edge)로 표현하여 네트워크 이미지를 구현함(그림 25).
- 시계열로 수집된 데이터 또는 서로 다른 샘플을 활용하여 구축된 네트워크의 경우, 관련 처리구에서의 유전자간의 상호작용을 서로 다른 색으로 표시하여 연구자가 시각적으로 쉽게 확인할 수 있게 구현하였음(그림 25).
- 관심 유전자군 개별로 몇 개의 유전자와 상호 연관되어 있는지 확인하여, hub gene target을 선정할 수 있도록 정량적으로 표시하였음.

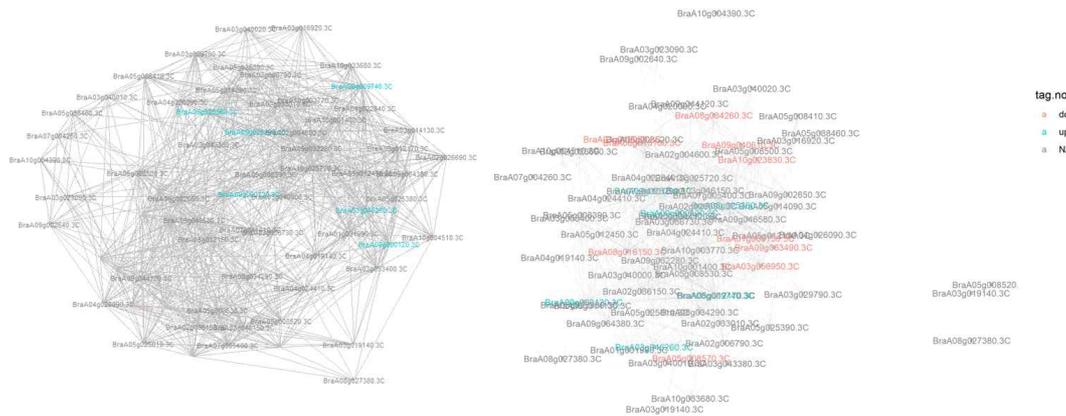


그림 25. GO term 유전자 set을 이용한 네트워크 이미지

○ 담배 전사체 데이터를 활용한 유전자 네트워크 분석

- 배추에 이어, 담배(*Nicotiana benthamiana*)의 전사체 데이터를 활용하여 유전자 네트워크 분석 프로그램을 고도화 하였음(그림 26).

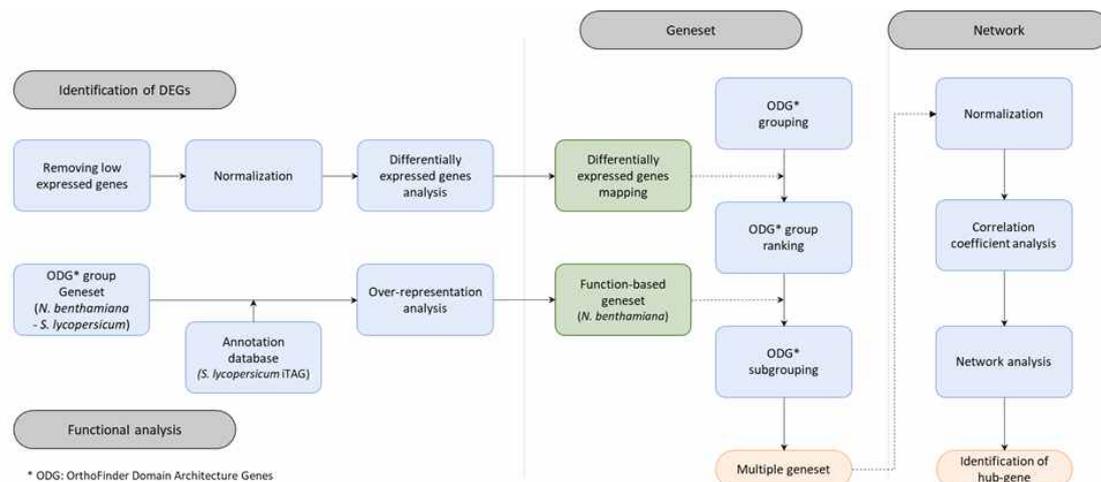


그림 27. 담배의 전사체 데이터를 활용한 네트워크 분석 개요도

- 전사체 데이터의 전처리시 발현량이 특정 기준 (read count 5 미만)인 유전자는 분석 대상에서 제외하였음.
- 분석 파이프라인은 담배 전사체에서 설정된 비교군을 활용하여, 발현 비교를 실시하여 차등유전자를 탐색하고 후속 분석에 활용될 수 있게 함.
- 관심 유전자의 cluster를 구성하여, 담배의 차등발현유전자와 그 연관 유전자의 상관관계를 확인함. 관심 유전자 cluster는 연관 품종(토마토, 고추)와 함께 Domain architecture를 비교하여 유사성이 있는 유전자세트(OrthoFinder Domain Architecture Genes (ODG))로 구성함.

- Gene annotation database와 domain architecture를 기반으로 구성된 ODG 유전자세트에서 관심 유전자인 차등발현유전자를 확인하여, 차등발현유전자가 많이 mapping된 ODG 유전자 세트의 기능 분석(functional analysis)을 실시하여, ODG 유전자세트를 한층 더 세분화 함(그림 27).

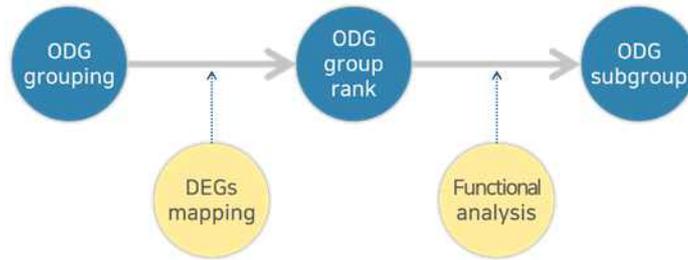


그림 28. 네트워크 분석을 위한 유전자세트 구성 과정

- Domain architecture와 gene function을 기반으로 구성된 유전자세트에서 관심 유전자인 차등발현유전자를 확인하여, 해당 유전자세트 내에 속한 차등발현유전자와 상관관계가 있는 유전자를 모두 검토하여 네트워크 분석을 실시함(그림 28).

Rank	Orthogroup	n_Tomato	n_Tobacco	n_DEGs
1	OG0000002DAGNO	0	168	9
2	OG0000007DAG1612	37	61	8
3	OG0000092DAG1288	11	11	6
4	OG0000002DAG457	0	154	5
5	OG0000022DAG1288	26	23	5
6	OG0000124DAG1288	9	7	5
7	OG0000041DAG4653	11	14	4
8	OG0000056DAG1288	12	13	4
9	OG0000213DAG3165	5	6	4
10	OG0000827DAG1668	1	4	4

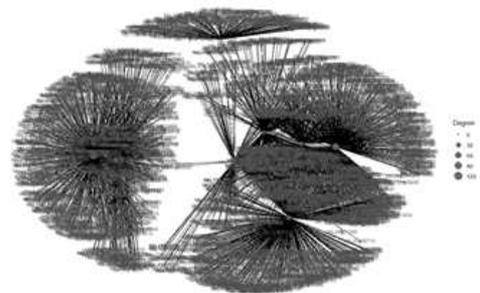


그림 29. Domain architecture의 유사성을 바탕으로 생성된 유전자세트(좌)와 해당 유전자세트를 활용한 네트워크 분석 결과(우)

- 네트워크 분석 결과 유전자들의 상호관계를 시각적으로 확인하고, 여러 유전자와 상호 연관된 유전자를 hub gene target으로 선정하여 그 목록을 제시하였음 (그림 29).

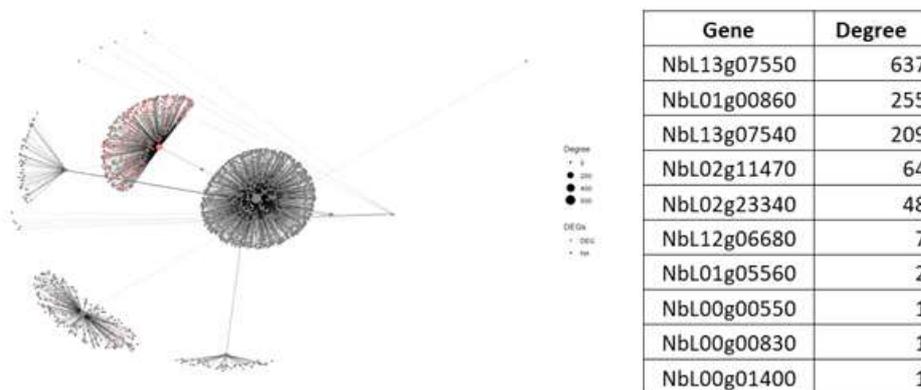


그림 30. domain architecture 기반 유전자 그룹의 네트워크 분석 결과로 나타낸 네트워크 이미지와 hub gene 목록

○ 디지털 육종 플랫폼 구축

- 현재, 다양한 연구 및 관련 사업을 통해 국내에서 중요하게 취급되고 있는 작물을 대상으로 유전체 및 표현체 정보가 지속적으로 생산 및 공개가 되고 있음.
- 그러나 중요한 시장적 가치를 가진 작물의 표준유전체 및 변이정보와 작물의 핵심집단에서 얻어진 표현체 데이터를 효율적으로 연관 분석할 수 있는 플랫폼은 전무한 상태임.
- 본 과제에서는 다양한 생물 데이터를 이용하여 인공지능 분석이 가능한 웹 플랫폼인 Demeter를 구축하고 있으며, 본 과제의 1단계에서 확보된 데이터 및 개발 완료된 프로그램의 탑재를 진행하고 있음(그림 30, <http://demeter.euso.kr/>)

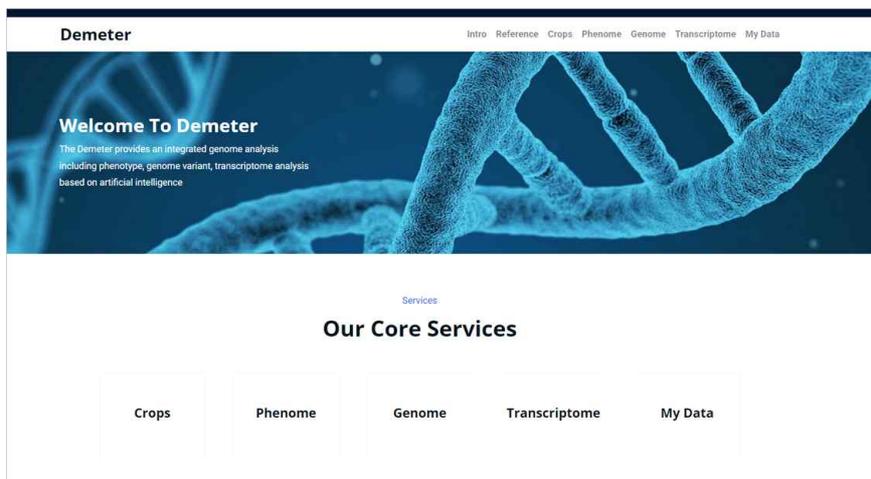


그림 30. 디지털 육종 플랫폼 Demeter의 메인 웹페이지 및 주요 기능

- 배추, 고추, 토마토, 및 무궁화를 대상 작물로 하여 Demeter의 구축이 진행 중이며 각 작물에 대한 일반적 정보, 표현체 데이터, 변이(유전체) 데이터, 전사체 데이터를 확보하여 Demeter에 탑재 중임(그림 31).
- 표현체 데이터의 경우, 표현형의 측정치 데이터와 실제 촬영된 이미지 데이터를 분류하고 딥러닝 모델 구축에 활용하기 위한 데이터의 전처리 및 편집을 추가적으로 수행함.
- 배추의 경우, 156개 핵심 계통에서 생산된 표현체(이미지 및 측정된 표현형 수치) 데이터 및 인공지능 기반 계통분류를 위한 2D 디지털 이미지로 변환된 변이지도 정보 데이터가 데이터베이스화 되어 현재 열람이 가능함.
- 본 연구에서 수행된 전사체 데이터 세트로부터 생산된 DEG 및 관련 파이프라인으로 전사체(Transcriptome) 부분의 웹페이지 구성 및 관련 콘텐츠의 구성을 진행하고 있음.

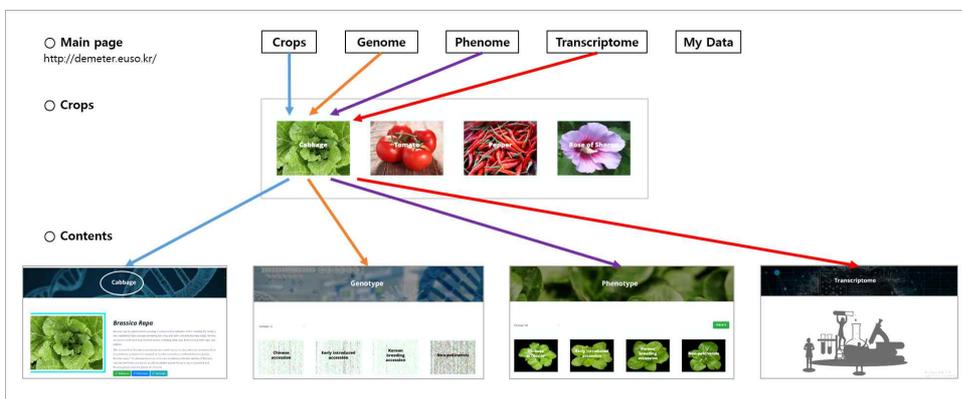


그림 31. Demeter에서 제공하는 주요 contents의 구축 현황

- Crops는 본 데이터베이스에서 다루는 작물의 일반적 사항 및 표준유전체의 세부 정보의 열람이 가능함.
- Genome은 작물의 계통별 변이정보 및 이의 인공지능 분석을 위한 이미지 변환 정보를 제공하고, 보유 변이정보 이미지를 활용한 사용자 보유 작물의 계통 분류를 지원할 예정임.
- Phenome은 연구과정에서 구축된 CNN 기반 분류 모델을 활용하여 사용자가 입력한 배추 계통의 사진 데이터를 통해 배추의 기원 및 그룹을 분류할 수 있는 기능을 구성하고 있음(그림 32).

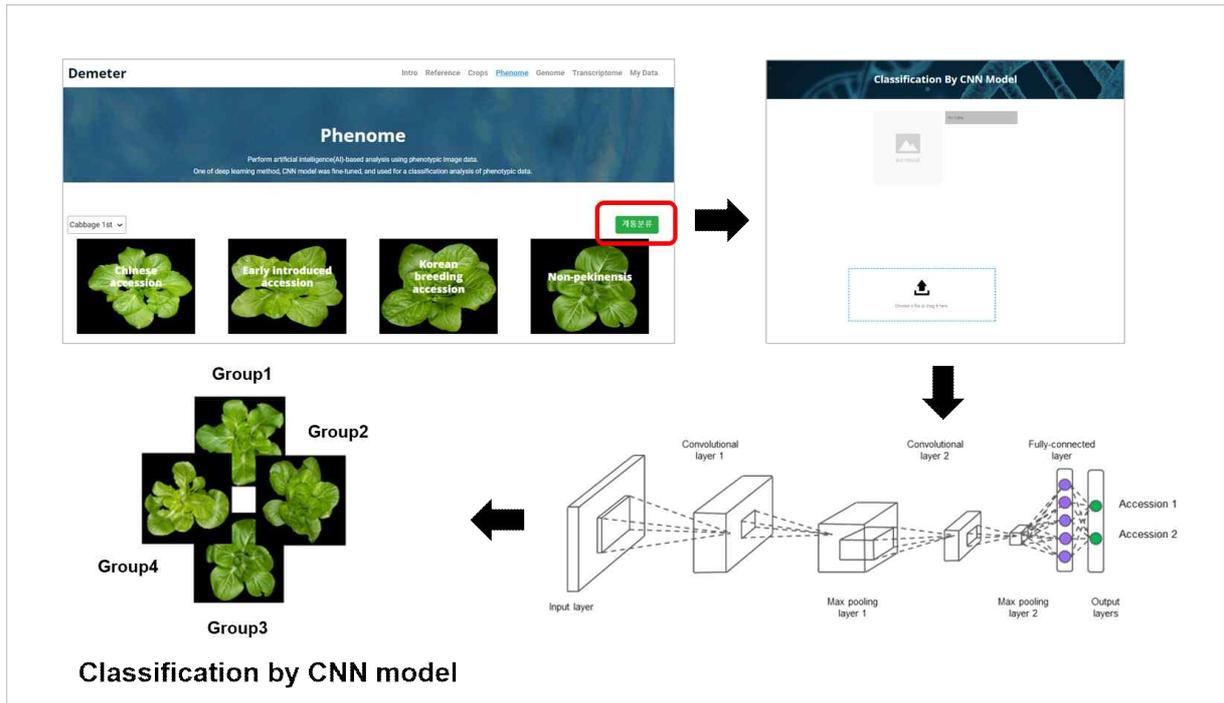


그림 33. 계통 이미지를 활용한 작물 계통의 분류 기능 개요도

- My Data 페이지에서는 Demeter에서 보유하고 있는 다양한 데이터(표현체, 유전체, 전사체)를 활용하여, 연구자의 데이터 분석을 지원할 수 있는 1) 유전체 변이 지도 이미지 기반 계통 분류 프로그램, 2) 표현체 기반 계통분류 프로그램, 3) 전사체 기반 유전자 네트워크 분석 프로그램 및 4) 계통 특이적 변이 발굴 프로그램 등을 지원할 예정임.
- 일례로, 연구자는 Demeter 보유 작물의 변이정보와 계통 특이적 변이 발굴 프로그램을 활용하여, 사용자 보유 작물의 유전체 재분석 데이터(fastq 파일)를 활용하여 보유 작물의 계통 특이적 변이를 탐색할 수 있음. 해당 프로그램은 기술이전을 통해, Demeter에 탑재되어 있음.
- Demeter의 변이 탐색 프로그램 'Run Pipeline'은 1회 수행시 최소 30분에서 1시간 이상 소요되는 작업으로 고성능 서버 도입 및 파이프라인 구축을 통해 최대 4개의 생성 프로세스를 동시에 진행할 수 있음(그림 33).

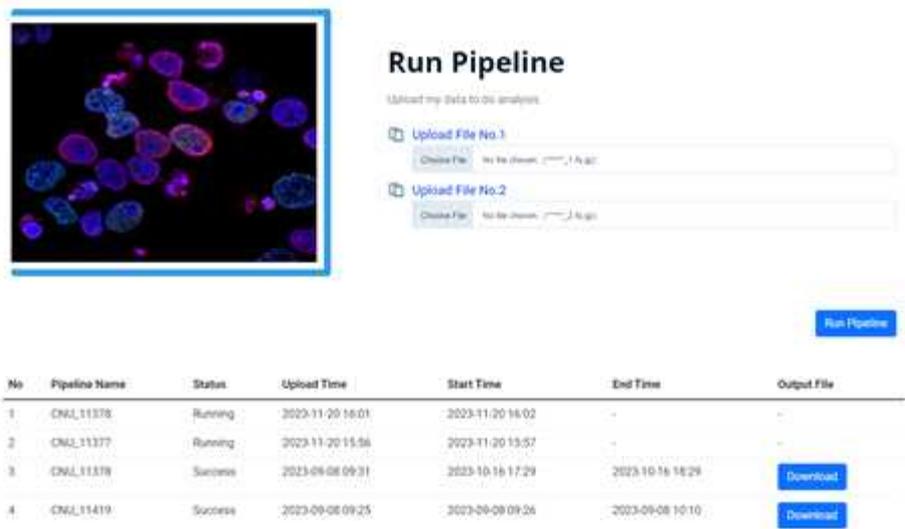


그림 34. 디지털 육종 플랫폼 Demeter의 계통 특이적 변이 탐색 과정

- 유전자 네트워크 분석 프로그램의 웹 플랫폼 구현
  - 서로 다른 두 품종 작물의 전사체 데이터 비교 분석 과정을 파이프라인 형태의 프로그램으로 제작함
  - 1차년도에 생산된 일반 배추와 기능성 자색 배추의 전사체 데이터의 비교 분석 결과는 GO term 유전자 set 단위로 웹플랫폼에서 검색하도록 구현함(그림 34)

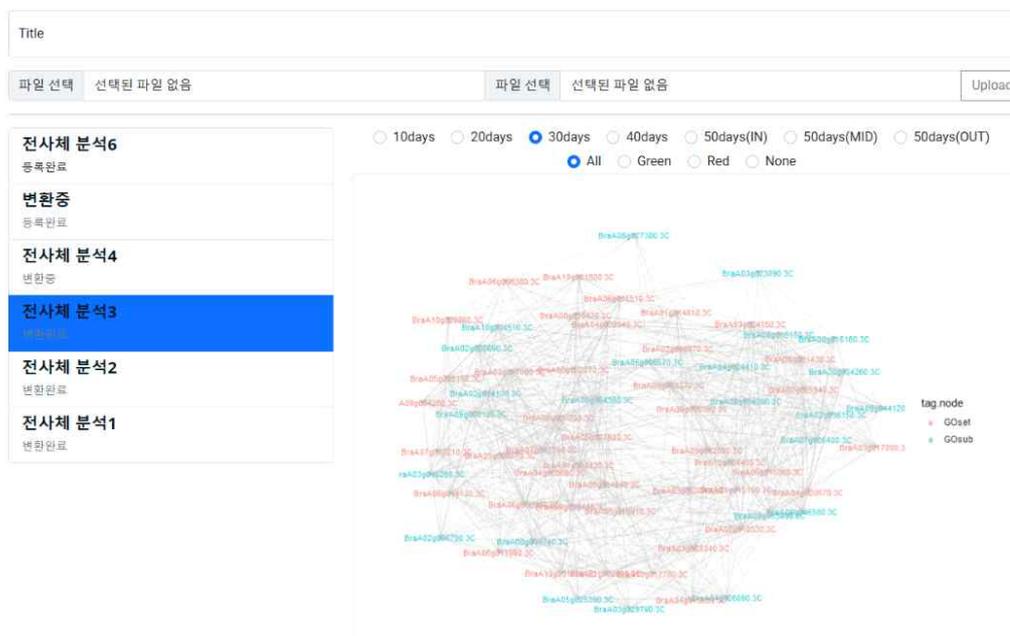


그림 35. 전사체 네트워크 분석 결과 검색 화면

- 유전체 기반 knowledge 데이터베이스 구축
  - 배추, 고추, 토마토의 공개된 표준유전체의 정보에 기반한 knowledge 데이터베이스를 구축하였음(그림 35).
  - 수집 가공된 표준유전체의 정보는 Gene Archive 페이지에서 확인을 할 수 있으며 이는 표 및 도식을 통해 표준유전체의 버전, 출처, 관련된 종의 taxonomy 정보를 제공함.
  - 추후 같은 종내의 다양한 아종의 표준유전체가 구축되어있는 배추와 고추의 경우, 이를 아종별로 검색이 가능하도록 조치할 계획임.



표 9. Demeter Knowledge 데이터베이스에서 배추, 고추, 토마토를 대상으로 수집한 표준유전체의 정보

Species	Version	Database	Detail	Institute/University	Release date
<i>Brassica rapa</i>	Brapa_sequence_v1.5	BRAD	<i>Brassica rapa</i> cv. <i>Chiuftu</i>	institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences	2011-03-07
	BrapaV.2.5_chr	BRAD	<i>Brassica rapa</i> cv. <i>Chiuftu</i>	institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences	2017-04-03
	Brara_chiiftu_V.3.1	BRAD	<i>Brassica rapa</i> cv. <i>Chiuftu</i>	institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences	2021-03-01
	Brara_chiiftu_V.3.5	BRAD	<i>Brassica rapa</i> cv. <i>Chiuftu</i>	institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences	2021-11-28
	GCA_008629595.1_ASM862959v1	NCBI	<i>Brassica rapa</i> ssp. <i>pekinensis</i>	Kyung Hee University	2019-09-20
<i>Capsicum annuum</i>	Annum.v.1.6	PEPPER GENOME	<i>Capsicum annuum</i> cv. CM334	Seoul National University	2012-12-30
	Baccatum.v.1.2	PEPPER GENOME	<i>C. annuum</i> spp. <i>Baccatum</i>	Seoul National University	2017-10-27
	Chinese.v.1.2	PEPPER GENOME	<i>Capsicum chinense</i>	Seoul National University	2017-08-14
	ASM51225v2	NCBI	<i>Capsicum annuum</i> cv. CM334	Seoul National University	2017-10-30
	GCA_011745865.1_SNU_SF	NCBI	<i>Capsicum annuum</i> cv. SF	Seoul National University	2020-03-25
	GCA_021292125.1_CA_59_1.0	NCBI	<i>Capsicum annuum</i> cv. 59	South China Agriculture University	2021-12-28
	GCF_000710875.1_Pepper_zunla_1_REF_v1.0_genomic	NCBI	<i>Capsicum annuum</i> cv. Zunla-1	Beijing Genome Institute	2015-03-11
SNU_ECW.v.1.0	NCBI	<i>Capsicum annuum</i> cv. ECW	Seoul National University	2020-03-25	
<i>Solanum lycopersicum</i>	1	Sol genmics	<i>Solanum lycopersicum</i> cv. Heinz 1706	SOL Genomics Network(SGN)	2009-11-27
	1.03	Sol genmics	<i>Solanum lycopersicum</i> cv. Heinz 1706	SOL Genomics Network(SGN)	2010-01-22
	1.5	Sol genmics	<i>Solanum lycopersicum</i> cv. Heinz 1706	SOL Genomics Network(SGN)	2010-06-16
	2.1	Sol genmics	<i>Solanum lycopersicum</i> cv. Heinz 1706	SOL Genomics Network(SGN)	2010-06-25
	2.3	Sol genmics	<i>Solanum lycopersicum</i> cv. Heinz 1706	SOL Genomics Network(SGN)	2010-08-09
	2.3.1	Sol genmics	<i>Solanum lycopersicum</i> cv. Heinz 1706	SOL Genomics Network(SGN)	2011-11-15
	2.4	Sol genmics	<i>Solanum lycopersicum</i> cv. Heinz 1706	SOL Genomics Network(SGN)	2011-02-02
	2.5	Sol genmics	<i>Solanum lycopersicum</i> cv. Heinz 1706	SOL Genomics Network(SGN)	2014-02-03

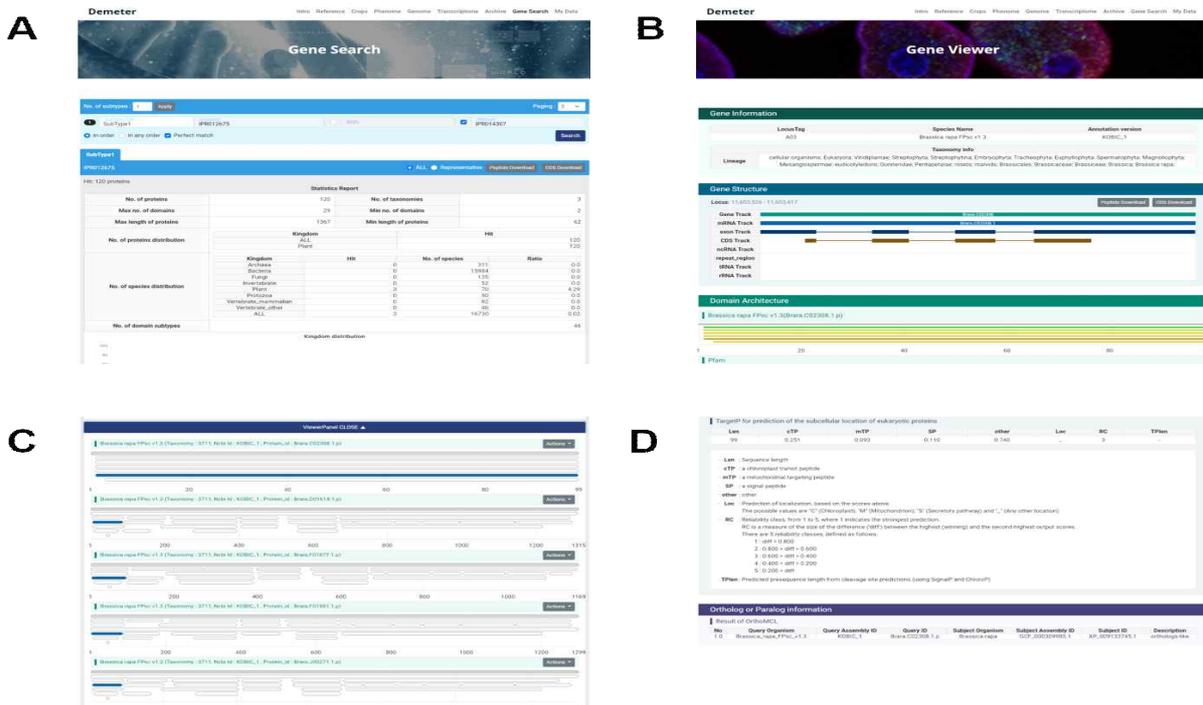


그림 36. Demeter knowledge 데이터베이스의 Gene search 페이지의 구성;

A) Gene search 메인 페이지, B) 검색된 유전자의 서열 상의 엑손/인트론 구조 확인 기능, C) 검색된 유전자의 도메인 구조 확인 기능, D) 검색된 유전자의 다른 종에서의 orthologue 및 paralogue 확인 기능.

### 3. 연구개발과제의 수행 결과 및 목표 달성 정도

#### 1) 연구수행 결과

##### (1) 정성적 연구개발성과

- 개체/계통 특이적 변이 탐색 프로그램의 범용화 및 자동화 프로그램 개발 완료
- 배추 전사체 데이터 세트 설계 및 전사체 데이터 생산 완료
- 전사체의 DEG 분석 자동화 프로그램 개발 중임
- DEG 정보를 활용한 네트워크 분석 프로그램 개발 중임
- 배추와 고추의 표현형 데이터 확보 및 표현형 분류 프로그램 개발 완료
- 인공지능과 유전체 데이터를 활용한 개체/계통 분류 프로그램의 정확도 87.5%이상 달성 (단계 목표로 설정된 정확도는 81.5%로 6% 이상 초과 달성)
- 인공지능 분석기능이 탑재된 디지털 육종 플랫폼 구축함
- 기술이전 1건 진행 완료

##### (2) 정량적 연구개발성과(해당 시 작성하며, 연구개발과제의 특성에 따라 수정이 가능합니다)

- SCI급 논문 출판 4건 완료(Impact factor 4.26 (1건), 4.60 (1건), 6.627 (2건))
- 특허 출원 2건 완료
- 배추 전사체 데이터 30건 생산 및 NABIC의 농생명 오믹스정보 등록시스템에 등록
- 학술대회 발표 8건 수행
- 기술이전 1건 달성(정액기술료 3,000만원)
- 신규 인력 11명 채용 달성

< 정량적 연구개발성과표 >

(단위 : 건, 천원)

성과지표명		연도	1단계	2단계	계	가중치 (%)
			(2021~2022)	(2023)		
전담기관 등록·기탁 지표 <sup>1)</sup>	논문	목표(단계별)	3	2	5	
		실적(누적)	3	1	4	
	논문평균IF	목표(단계별)	4	4	4 (평균)	10
		실적(누적)	5.9	4.6	5.56 (평균)	
	특허(출원)	목표(단계별)	2	1	3	
		실적(누적)	2	0	2	
	생명정보	목표(단계별)				
		실적(누적)	30		30	
연구개발과제 특성 반영 지표 <sup>2)</sup>	기술실시	목표(단계별)	0	1	1	
		실적(누적)	1	-	1	
	기술료	목표(단계별)	30,000	-	30,000	10
		실적(누적)	30,000	-	30,000	
	채용	목표(단계별)	3	1	4	10
		실적(누적)	8	3	11	
계						

\* 1) 전담기관 등록·기탁 지표: 논문[SCI Expanded(SCIE), 비SCIE, 평균Impact Factor(IF)], 특허, 보고서원문, 연구시설·장비, 기술요약정보, 저작권(소프트웨어, 서적 등), 생명자원(생명정보, 생물자원), 표준화(국내, 국제), 화합물, 신물질 등을 말하며, 논문, 학술발표, 특허의 경우 목표 대비 실적은 기재하지 않아도 됩니다.

\* 2) 연구개발과제 특성 반영 지표: 기술실시(이전), 기술료, 사업화(투자실적, 제품화, 매출액, 수출액, 고용창출, 고용효과, 투자유치), 비용 절감, 기술(제품)인증, 시제품 제작 및 인증, 신기술지정, 무역수지개선, 경제적 파급효과, 산업지원(기술지도), 교육지도, 인력양성(전문 연구인력, 산업연구인력, 졸업자수, 취업, 연수프로그램 등), 법령 반영, 정책활용, 설계 기준 반영, 타 연구개발사업에의 활용, 기술무역, 홍보(전시), 국제화 협력, 포상 및 수상, 기타 연구개발 활용 중 선택하여 기재합니다

(연구개발과제 특성별로 고유한 성과지표를 추가할 수 있습니다).

< 연구개발성과 성능지표(예시) >

평가 항목 (주요성능 <sup>1)</sup> )	단위	전체 항목에서 차지하는 비중 <sup>2</sup> (%)	세계 최고		연구개발 전 국내 성능수준	연구개발 목표치		목표설정 근거
			보유국/보유기관	성능수준	성능수준	1단계(21~22)	2단계(23)	
1 시스템 응답 속도	sec		미국/Oracle	500ms	-	1		
2 트래픽 처리 속도	TPS		미국/Google	∞	-	10		
3 시스템 가동률	%		-	-	-	99		
4 정확도	%					81.5		

\* 1」 정밀도, 인장강도, 내충격성, 작동전압, 응답시간 등 기술적 성능판단기준이 되는 것을 의미합니다.

\* 2」 비중은 각 구성성능 사양의 최종목표에 대한 상대적 중요도를 말하며 합계는 100%이어야 합니다.

**(3) 세부 정량적 연구개발성과(해당되는 항목만 선택하여 작성하되, 증빙자료를 별도 첨부해야 합니다)**

**[과학적 성과]**

**□ 논문(국내외 전문 학술지) 게재**

번호	논문명	학술지명	주저자명	호	국명	발행기관	SCIE 여부 (SCIE/비SCIE)	게재일	등록번호 (ISSN)	기여율
1	Genome-Wide Comparative Analysis of Flowering-Time Genes; Insights on the Gene Family Expansion and Evolutionary Perspective	Frontiers in Plant Science	홍성민 (제1저자), 신아영 (교신), 김용민 (교신)	-	스위스	FRONTIERS MEDIA SA	SCIE	2021.07.05	1664-462X	30
2	Deep Learning Algorithms Correctly Classify <i>Brassica rapa</i> Varieties Using Digital Images	Frontiers in Plant Science	정민아 (제1저자), 김용민 (교신)	-	스위스	FRONTIERS MEDIA SA	SCIE	2021.09.29	1664-462X	40
3	Identification of accession-specific variants and development of KASP markers for assessing the genetic makeup of <i>Brassica rapa</i> seeds	BMC Genomics	홍성민 (제1저자), 신아영 (교신), 김용민 (교신)	23	영국	BioMed Central	SCIE	2022.04.25	1471-2164	
4	Construction of deep learning-based disease detection model in plants	Scientific Reports	정민아 (제1저자), 신아영 (제1저자), 김용민 (교신)	13	영국	Nature Publishing Group UK	SCIE	2023.05.05	2045-2322	40

**□ 국내 및 국제 학술회의 발표**

번호	회의 명칭	발표자	발표 일시	장소	국명
1	한국식물생명공학회	정민아, 김용민	2021.08.11	제주 라마다 호텔	대한민국
2	한국분자세포생물학회	홍성민, 김용민	2021.11.04	제주 ICC	대한민국
3	한국유전체학회	홍성민, 김용민	2022.02.09	홍천 비발디파크	대한민국
4	한국육종학회	신아영, 김용민	2022.06.30	제주 라마다 호텔	대한민국
5	한국식물생명공학회	고상진, 김용민	2022.08.11	제주 라마다 호텔	대한민국
6	한국분자세포생물학회	홍성민, 김용민	2022.09.29	제주 ICC	대한민국
7	한국유전체학회	홍성민, 김용민	2023.02.01	홍천 비발디파크	대한민국
8	한국유전체학회	고상진, 김용민	2023.02.01	홍천 비발디파크	대한민국



□ 신기술 지정

번호	명칭	출원일	고시일	보호 기간	지정 번호

□ 기술 및 제품 인증

번호	인증 분야	인증 기관	인증 내용		인증 획득일	국가명
			인증명	인증 번호		

□ 표준화

○ 국내표준

번호	인증구분 <sup>1)</sup>	인증여부 <sup>2)</sup>	표준명	표준인증기구명	제안주체	표준종류 <sup>3)</sup>	제안/인증일자

- \* 1) 한국산업규격(KS) 표준, 단체규격 등에서 해당하는 사항을 기재합니다.
- \* 2) 제안 또는 인증 중 해당하는 사항을 기재합니다.
- \* 3) 신규 또는 개정 중 해당하는 사항을 기재합니다.

○ 국제표준

번호	표준화단계구분 <sup>1)</sup>	표준명	표준기구명 <sup>2)</sup>	표준분과명	의장단 활동여부	표준특허 추진여부	표준개발 방식 <sup>3)</sup>	제안자	표준화 번호	제안일자

- \* 1) 국제표준 단계 중 신규 작업항목 제안(NP), 국제표준초안(WD), 위원회안(CD), 국제표준안(DIS), 최종국제표준안(FDIS), 국제표준(IS) 중 해당하는 사항을 기재합니다.
- \* 2) 국제표준화기구(ISO), 국제전기기술위원회(IEC), 공동기술위원회1(JTC1) 중 해당하는 사항을 기재합니다.
- \* 3) 국제표준(IS), 기술시방서(TS), 기술보고서(TR), 공개활용규격(PAS), 기타 중 해당하는 사항을 기재합니다.

[경제적 성과]

□ 시제품 제작

번호	시제품명	출시/제작일	제작 업체명	설치 장소	이용 분야	사업화 소요 기간	인증기관 (해당 시)	인증일 (해당 시)

□ 기술 실시(이전)

번호	기술 이전 유형	기술 실시 계약명	기술 실시 대상 기관	기술 실시 발생일	기술료 (해당 연도 발생액)	누적 징수 현황
1	통상실시	전장유전체 데이터를 활용한 프라이머 제작 파이프라인	(주)유클리드 소프트	2022.07.18	3,000 만원	3,000 만원

- \* 내부 자금, 신용 대출, 담보 대출, 투자 유치, 기타 등

□ 사업화 투자실적

번호	추가 연구개발 투자	설비 투자	기타 투자	합계	투자 자금 성격*

□ 사업화 현황

번호	사업화 방식 <sup>1)</sup>	사업화 형태 <sup>2)</sup>	지역 <sup>3)</sup>	사업화명	내용	업체명	매출액		매출 발생 연도	기술 수명
							국내 (천원)	국외 (달러)		

- \* 1) 기술이전 또는 자기실시

- \* 2) 신제품 개발, 기존 제품 개선, 신공정 개발, 기존 공정 개선 등
- \* 3) 국내 또는 국외

□ 매출 실적(누적)

사업화명	발생 연도	매출액		합계	산정 방법
		국내(천원)	국외(달러)		
합계					

□ 사업화 계획 및 무역 수지 개선 효과

성과					
사업화 계획	사업화 소요기간(년)				
	소요예산(천원)				
	예상 매출규모(천원)	현재까지	3년 후	5년 후	
	시장 점유율	단위(%)	현재까지	3년 후	5년 후
		국내 국외			
향후 관련기술, 제품을 응용한 타 모델, 제품 개발계획					
무역 수지 개선 효과(천원)	수입대체(내수)	현재	3년 후	5년 후	
	수출				

□ 고용 창출

순번	사업화명	사업화 업체	고용창출 인원(명)			합계
			2021년	2022년	2023년	
1	기술사업화과제	(주) 유클리드소프트	2	6	3	11
합계						

□ 고용 효과

구분		고용 효과(명)	
고용 효과	개발 전	연구인력	
		생산인력	
	개발 후	연구인력	
		생산인력	

□ 비용 절감(누적)

순번	사업화명	발생연도	산정 방법	비용 절감액(천원)
합계				

□ 경제적 파급 효과

(단위: 천원/년)

구분	사업화명	수입 대체	수출 증대	매출 증대	생산성 향상	고용 창출 (인력 양성 수)	기타
해당 연도							
기대 목표							

산업 지원(기술지도)

순번	내용	기간	참석 대상	장소	인원

기술 무역

(단위: 천원)

번호	계약 연월	계약 기술명	계약 업체명	계약업체 국가	기 징수액	총 계약액	해당 연도 징수액	향후 예정액	수출/ 수입

[사회적 성과]

법령 반영

번호	구분 (법률/시행령)	활용 구분 (제정/개정)	명 칭	해당 조항	시행일	관리 부처	제정/개정 내용

정책활용 내용

번호	구분 (제안/채택)	정책명	관련 기관 (담당 부서)	활용 연도	채택 내용

설계 기준/설명서(시방서)/지침/안내서에 반영

번호	구분 (설계 기준/설명서/지침/안내서)	활용 구분 (신규/개선)	설계 기준/설명서/ 지침/안내서 명칭	반영일	반영 내용

전문 연구 인력 양성

번호	분류	기준 연도	현황															
			학위별				성별		지역별									
			박사	석사	학사	기타	남	여	수도권	충청권	영남권	호남권	기타					

산업 기술 인력 양성

번호	프로그램명	프로그램 내용	교육 기관	교육 개최 횟수	총 교육 시간	총 교육 인원

다른 국가연구개발사업에의 활용

번호	중앙행정기관명	사업명	연구개발과제명	연구책임자	연구개발비

국제화 협력성과

번호	구분 (유치/파견)	기간	국가	학위	전공	내용

□ 홍보 실적

번호	홍보 유형	매체명	제목	홍보일

□ 포상 및 수상 실적

번호	종류	포상명	포상 내용	포상 대상	포상일	포상 기관

[인프라 성과]

□ 연구시설·장비

구축기관	연구시설/ 연구장비명	규격 (모델명)	개발여부 (○/×)	연구시설·장비 종합정보시스템* 등록여부	연구시설·장비 종합정보시스템* 등록번호	구축일자 (YY.MM.DD)	구축비용 (천원)	비고 (설치 장소)

\* 「과학기술기초법 시행령」 제42조제4항제2호에 따른 연구시설·장비 종합정보시스템을 의미합니다.

[그 밖의 성과](해당 시 작성합니다)

---



---

(4) 계획하지 않은 성과 및 관련 분야 기여사항(해당 시 작성합니다)

- 농림식품부에서 진행하는 농식품 R&D 우수사례 선정
  - 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발
  - 생물학 및 농업적 연구에 인공지능을 융합한 첫 번째 사례로 인정받음

---



---

## 2) 목표 달성 수준

추진 목표	달성 내용	달성도(%)
○ 개체/계통 특이적 변이 탐색 프로그램	○ 개체/계통 특이적 변이 탐색 프로그램의 범용화 및 자동화를 통한 프로그램 고도화	○ 100
○ 전사체 데이터 생산	○ 전사체를 활용한 유전자 네트워크 분석을 위한 데이터 세트 설계 및 데이터 생산	○ 100
○ 전사체 분석 자동화 프로그램	○ 전사체 데이터 분석 및 분석 자동화 프로그램 개발	○ 80
○ 개체/계통 분류 프로그램	○ 인공지능 기술과 유전체 변이 데이터를 활용한 개체/계통 분류 프로그램 개발(정확도 87.5% 이상)	○ 100
○ 표현형 데이터 제공	○ 표현형 기반 분류 모델 프로그램 개발을 위한 표현형 데이터 생산, 공개 데이터 수집 및 제공	○ 100
○ 유전체 기반 knowledge 데이터 베이스 구축	○ 웹 기반 디지털 육종 플랫폼을 서비스를 위한 유전체 기반 knowledge DB 구축	○ 90
○ 표현형 기반 분류 모델 분석 프로그램 개발	○ 인공지능을 활용한 표현형 기반 계통 분류 프로그램 개발	○ 100
○ 유전자 네트워크 분석 프로그램	○ 전사체 데이터를 활용한 유전자 네트워크 분석 프로그램 개발	○ 70

## 4. 목표 미달 시 원인분석(해당 시 작성합니다)

### 1) 목표 미달 원인(사유) 자체분석 내용

---

- 1단계 2년차 등록특허의 경우, 1년차 때 출원을 완료하였으나 아직 심사 대기 중으로 나타나 2년차에 등록을 관련 절차를 시작하지 못하였음.
  - 전사체 분석 자동화 프로그램의 경우, R 스크립트 기반으로 관련 데이터를 통한 작업 파이프라인을 구축이 완료된 상태이며, 현재 Demeter 탑재를 위한 작업 스크립트의 python 변환이 수행되고 있음.
  - 유전자 네트워크 분석의 경우, 네트워크 분석 결과의 검증 방법 설계 및 검증 과정을 수행 중으로, 계획에 비해 다소 일정이 지체되었음.
- 

### 2) 자체 보완활동

---

- 2단계 1년차 특허출원의 경우, 현재 원내 자체 출원 심사를 통과하여 출원 절차 진행 중임.
  - 유전자 네트워크 분석의 경우, 계산의 효율화를 위해 GO term category를 활용한 방식으로 진행 중이며, 현재 식물 gene ID에 대한 작업을 완료 후 파이프라인 검증 중임.
- 

### 3) 연구개발 과정의 성실성

---

- 논문 출판 4건 완료 (계획 목표 5건, 추가 2건 진행 중, 출판 논문 4건의 impact factor 평균 5.56으로 기준 impact factor 4를 초과)
  - 특허출원 2건 완료 (계획 목표 3건)
  - 학술대회 발표 8건 수행으로 계획 목표 초과 달성(계획 목표 6건)
  - 신규인력 11명 채용으로 계획 목표를 초과 달성(계획 목표 4건)
  - 단계 성과 계획에 없는 생물정보 등록 및 기탁 30건을 달성(NABIC, 농생명 오믹스정보 등록 시스템)
  - Demeter 업데이트를 위한 신규 생물정보 데이터 및 표현체 데이터의 지속적인 수집 진행 중
-

## 5. 연구개발성과의 관련 분야에 대한 기여 정도

---

- 생물학 및 농업적 연구에 인공지능을 융합한 첫 번째 사례
    - 본 연구 수행을 통해, 작물의 핵심 집단 내 유전적 다양성을 갖는 156개 계통에 대한 표현체 데이터를 생산한 후 정제 및 인공지능을 직접적으로 활용하여 분류모델을 개발한 첫 사례라는 점에서 의의가 크다.
    - 이에 더하여 아직 국내에서 생소한 생물학 및 농업적 연구에 인공지능을 융합한 사례로서 현재 산업 및 학계에서 대두되는 디지털 육종에 방향성을 제시할 수 있는 연구라고 할 수 있다.
  - 사회 경제적으로도 기여도
    - 첫째, 현장에서 대량으로 생산된 이미지를 활용하여 본 연구에서 구축한 모델의 정확성을 개선하게 되면 국내 농업 관련 산업의 다양한 분야에 적용이 가능할 것으로 예상된다.
    - 둘째, 이는 농작물의 이미지를 통한 비파괴 검사, 원산지 파악 및 이물 분류와 같은 분야에 적용하여 산물 분류에 필요한 비용의 감축을 이끌어 생산비 절감에 기여할 수 있을 것으로 기대된다.
    - 셋째, 정확한 예측 및 분류를 수행하는 인공지능 모델의 구축을 통하여 과거 숙련자들에 의해 소수로 운영되던 업종의 진입 장벽을 낮출 수 있을 것으로 전망된다.
- 

## 6. 연구개발성과의 관리 및 활용 계획

---

- 디지털 육종 플랫폼(Demeter)의 지속적인 업그레이드를 통한 사용자층 확대 및 내실화
    - 주요 작물의 다양한 유전체/표현형 데이터의 지속적인 수집 및 정제·가공 및 디지털 육종 플랫폼에 탑재하여 다양한 작물의 표현형-유전체 연관 분석을 지원할 계획임.
    - 사용자 데이터를 분석할 수 있는 다양한 생물정보학 분석 프로그램 개발 및 탑재를 통한 사용자 편의성을 높이고자 함.
    - 이러한 일련의 업그레이드를 통해 다양한 작물 육종에 활용될 수 있는 플랫폼 구축 및 사용자층 확대를 통한 이윤을 창출하고자 함.
  - 경제적 효용이 뛰어난 작물 및 표현형 분석 다양화에 기반한 서비스 대상 확장을 통한 사업화 추진
    - 최근 식물 유래 이차 대사산물을 활용한 기능성 식품 및 의약품 개발 연구가 활발히 진행되고 있으며, 이러한 연구 및 시장 상황을 반영하여 사업화 방향성을 정립하고자 함.
    - 유용 이차 대사산물 함량이 높고 유전체 정보 및 표현형 정보 확보가 상대적으로 용이한 병풀(centelloside), 인삼(saponin), 포도(resveratrol) 등의 작물에서의 이차대사산물 연구에 표현형 예측 프로그램을 적용할 수 있는 연구를 진행하여 프로그램의 고도화 및 사업화를 추진하고자 함.
  - 디지털 육종 플랫폼(Demeter) 홍보 방안 수립
    - 사업화 대상인 디지털 육종 플랫폼(Demeter) 서비스의 성과창출을 위한 홍보 방안을 마련할 예정임.
    - 관련 학회 참석, 발표 및 홍보 부스 설치 등을 실시하여 잠재적 고객인 연구자들 대상 홍보 활동 예정임.
    - Private 데이터 관리 방안 등을 적극 홍보하여 기관 또는 개인의 자발적 참여를 유도함.
    - 네트워킹을 통해 공공/민간사업에 적극적으로 참여하여 공동 참여 기관 대상으로 플랫폼 서비스를 소개하여 사용자 효능감을 고취시킴.
-





# Deep Learning Algorithms Correctly Classify *Brassica rapa* Varieties Using Digital Images

Minah Jung<sup>1,2†</sup>, Jong Seob Song<sup>2†</sup>, Seongmin Hong<sup>3,4</sup>, SunWoo Kim<sup>5</sup>, Sangjin Go<sup>3,6</sup>, Yong Pyo Lim<sup>4</sup>, Juhan Park<sup>2</sup>, Sung Goo Park<sup>1,7</sup> and Yong-Min Kim<sup>3\*</sup>

<sup>1</sup> Department of Functional Genomics, KRIBB School of Biological Science, Korea University of Science and Technology, Daejeon, South Korea, <sup>2</sup> Euclidsoft Co., Ltd, Daejeon, South Korea, <sup>3</sup> Genome Editing Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon, South Korea, <sup>4</sup> Molecular Genetics and Genomics Laboratory, Department of Horticulture, College of Agriculture and Life Science, Chungnam National University, Daejeon, South Korea, <sup>5</sup> Department of Bio-AI Convergence, Chungnam National University, Daejeon, South Korea, <sup>6</sup> Department of Horticulture, Gyeongsang National University, Jinju, South Korea, <sup>7</sup> Disease Target Structure Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon, South Korea

## OPEN ACCESS

### Edited by:

Brian Cullis,  
University of Wollongong, Australia

### Reviewed by:

Marcin Wozniak,  
Silesian University of  
Technology, Poland  
Julie N. Malcoff,  
University of California, Davis,  
United States

### \*Correspondence:

Yong-Min Kim

†These authors share first authorship

### Specialty section:

This article was submitted to  
Plant Bioinformatics,  
a section of the journal  
Frontiers in Plant Science

**Received:** 09 July 2021

**Accepted:** 31 August 2021

**Published:** 29 September 2021

### Citation:

Jung M, Song JS, Hong S, Kim S,  
Go S, Lim YP, Park J, Park SG and  
Kim Y-M (2021) Deep Learning  
Algorithms Correctly Classify *Brassica*  
*rapa* Varieties Using Digital Images.  
*Front. Plant Sci.* 12:738685.  
doi: 10.3389/fpls.2021.738685

Efficient and accurate methods of analysis are needed for the huge amount of biological data that have accumulated in various research fields, including genomics, phenomics, and genetics. Artificial intelligence (AI)-based analysis is one promising method to manipulate biological data. To this end, various algorithms have been developed and applied in fields such as disease diagnosis, species classification, and object prediction. In the field of phenomics, classification of accessions and variants is important for basic science and industrial applications. To construct AI-based classification models, three types of phenotypic image data were generated from 156 *Brassica rapa* core collections, and classification analyses were carried out using four different convolutional neural network architectures. The results of lateral view data showed higher accuracy compared with top view data. Furthermore, the relatively low accuracy of ResNet50 architecture suggested that definition and estimation of similarity index of phenotypic data were required before the selection of deep learning architectures.

**Keywords:** artificial intelligence, deep learning, classification model, phenotypic analysis, *Brassica rapa* (Brassicaceae)

## INTRODUCTION

One of the major features of modern science is convergent analyses using heterogeneous technologies from multiple and independent fields to analyze huge amounts of data. To manipulate these data, artificial intelligence (AI) technology has come into the spotlight. Deep learning is a type of AI that uses computer algorithms based on artificial neural networks (ANNs), which mimic the principles and structure of human neural networks to emulate human cognitive processes (Chauhan et al., 2018). In an ANN, artificial neurons (nodes) combine synapses to form a network and strengthen synapses through learning, thus acquiring problem-solving capabilities. An ANN consists of three major components: an input layer that receives data, an output layer that presents the results of analysis, and hidden layers that exist between the input and output layers. To construct an analytic model that uses deep learning, the numbers of nodes and hidden layers must be specified.

better performance compared with different development stages (Supplementary Tables 9–11). In addition, more images are needed to construct classification models for various timepoint datasets. The prediction accuracies for the lateral view dataset were higher than those for the top view datasets. This is because the lateral view images showed diverse colors and more features than the top view images, such as the thickness of the leaves and the entire shape and stem of the plant.

The classification accuracies for the validation sets (step 1) and the test sets (step 2) depended on the type of architecture used (Figure 4). In the validation step, VGG19 showed the highest accuracy for the lateral view dataset, whereas GoogLeNet had the highest accuracy for the top view datasets. GoogLeNet showed the highest accuracy for all top view datasets in the test step. ResNet50 showed the lowest accuracy for all the datasets and was about 20% less accurate than GoogLeNet using the top view dataset in the test step. In previous studies, ResNet50 showed a relatively low performance to classify homogenous or highly similar images (Rudakov et al., 2018; Hassan et al., 2021). Therefore, the low accuracy of ResNet50 in our experiments might have been caused by a high degree of similarity among the images.

To investigate that hypothesis, we tested the architectures using pairs of accession types. The pairwise tests indicated that the classification of non-pekinensis together with any of the other three accession types was highly accurate, whereas that of the early introduced accessions was relatively inaccurate regardless of the other type of accession used. The classification accuracy for the Korean breeding accessions increased gradually depending on the other accession type in the pairwise test, with the early introduced accessions producing the lowest accuracy, the Chinese accessions producing higher accuracy, and the non-pekinensis accessions yielding the highest accuracy. According to the phylogenetic tree, the Korean breeding accessions are genetically close to the early introduced accessions and genetically distant from the non-pekinensis accessions (Figure 1). This suggests that phenotypic variances between the non-pekinensis accessions and the other accession types due to genetic dissimilarity led to high performance in the classification tests. The Korean breeding accessions were often classified accurately regardless of the deep learning architecture used, whereas the classifications of the Chinese accessions and early introduced accessions were inconsistent. These results further suggest that low accuracy in the classifications was caused by genetic differences between the accession types (Figure 1). Almost all of the early introduced accessions were imported from China in the early 1900s and have since been used as breeding sources. These accessions showed heterogenous phenotypes. On the other hand, the Korean breeding accessions have acquired homogenous phenotypes due to long-term breeding activities. In addition, the image depth for classification might affect the accuracy. The numbers of individual accessions used in the current analysis were not sufficient to train and evaluate deep learning models with very high accuracy for phenotypic classification. More images of various developmental stages would improve the accuracy of the models. Furthermore, tissue-specific and trait-specific images can be used in the

future to identify trait-associated genes by correlation with genotypic data.

For applications of deep learning classification models in industrial fields, top view images are more suitable than lateral view images, although we achieved higher accuracy using the lateral view. Classification models like the ones presented in this study can be applied for nondestructive inspection of accessions and cultivars; however, improved models need to be developed, which can identify different types of plants with high accuracy using top view images.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

## AUTHOR CONTRIBUTIONS

Y-MK conceived the project, designed the experiments, and organized the manuscript. MJ, JS, and SP performed the deep learning analysis. SH, SG, SK, and YL generated the phenotypic data. MJ and Y-MK wrote the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the Korea Forest Service of the Korean government through its R&D Program for Forestry Technology (Project No. 2014071H10-2122-AA04), and by the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, and Forestry through the Golden Seed Project (213006-05-5-SBG30) and the **Technology Commercialization Support Program (821026-03)**, which is funded by a grant from the Ministry of Agriculture, Food, and Rural Affairs, the Ministry of Oceans and Fisheries, the Rural Development Administration, and the Korea Forest Service to Y-MK.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online in the Dryad repository via the DOIs indicated below.

- Supplementary Figure 1** | Available via doi: 10.6084/m9.figshare.16620523.
- Supplementary Figure 2** | Available via doi: 10.6084/m9.figshare.16620538.
- Supplementary Figure 3** | Available via doi: 10.6084/m9.figshare.16620568.
- Supplementary Figure 4** | Available via doi: 10.6084/m9.figshare.16620571.
- Supplementary Figure 5** | Available via doi: 10.6084/m9.figshare.16620580.
- Supplementary Figure 6** | Available via doi: 10.6084/m9.figshare.16620592.
- Supplementary Figure 7** | Available via doi: 10.6084/m9.figshare.16620598.
- Supplementary Tables 1–13** | Available via doi: 10.6084/m9.figshare.16620604.



# Genome-Wide Comparative Analysis of Flowering-Time Genes; Insights on the Gene Family Expansion and Evolutionary Perspective

Seongmin Hong<sup>1,2</sup>, Yong Pyo Lim<sup>2</sup>, Suk-Yoon Kwon<sup>3</sup>, Ah-Young Shin<sup>3\*†</sup> and Yong-Min Kim<sup>1\*†</sup>

<sup>1</sup> Genome Editing Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIIBB), Daejeon, South Korea, <sup>2</sup> Molecular Genetics and Genomics Laboratory, Department of Horticulture, College of Agriculture and Life Sciences, Chungnam National University, Daejeon, South Korea, <sup>3</sup> Plant Systems Engineering Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon, South Korea

## OPEN ACCESS

### Edited by:

Jen-Tsung Chen,  
National University of Kaohsiung,  
Taiwan

### Reviewed by:

David Sankoff,  
University of Ottawa, Canada  
Rafaqat Ali Gill,  
Oil Crops Research Institute, Chinese  
Academy of Agricultural Sciences,  
China

### \*Correspondence:

Ah-Young Shin

Yong-Min Kim

† These authors have contributed  
equally to this work and share last  
authorship

### Specialty section:

This article was submitted to  
Plant Systematics and Evolution,  
a section of the journal  
Frontiers in Plant Science

**Received:** 30 April 2021

**Accepted:** 14 June 2021

**Published:** 05 July 2021

### Citation:

Hong S, Lim YP, Kwon S-Y,  
Shin A-Y and Kim Y-M (2021)  
Genome-Wide Comparative Analysis  
of Flowering-Time Genes; Insights on  
the Gene Family Expansion  
and Evolutionary Perspective.  
*Front. Plant Sci.* 12:702243.  
doi: 10.3389/fpls.2021.702243

In polyploids, whole genome duplication (WGD) played a significant role in genome expansion, evolution and diversification. Many gene families are expanded following polyploidization, with the duplicated genes functionally diversified by neofunctionalization or subfunctionalization. These mechanisms may support adaptation and have likely contributed plant survival during evolution. Flowering time is an important trait in plants, which affects critical features, such as crop yields. The flowering-time gene family is one of the largest expanded gene families in plants, with its members playing various roles in plant development. Here, we performed genome-wide identification and comparative analysis of flowering-time genes in three plant families i.e., Malvaceae, Brassicaceae, and Solanaceae, which indicate these genes were expanded following the event/s of polyploidization. Duplicated genes have been retained during evolution, although genome reorganization occurred in their flanking regions. Further investigation of sequence conservation and similarity network analyses provide evidence for functional diversification of duplicated genes during evolution. These functionally diversified genes play important roles in plant development and provide advantages to plants for adaptation and survival in response to environmental changes encountered during evolution. Collectively, we show that flowering-time genes were expanded following polyploidization and retained as large gene family by providing advantages from functional diversification during evolution.

**Keywords:** flowering-time gene, whole genome duplication, copy number variation, polyploidization, functional diversification

## INTRODUCTION

Flowering time is an important trait in plants that has wide-ranging effects on features such as growth, development, and crop yields (Huang et al., 2011), and evolution of new crop types (Gill et al., 2021). Thus, numerous studies using approaches that include genetic mapping, T-DNA insertional mutagenesis, and functional study have been conducted on flowering-time genes in

importance for survival in plants. In addition, the multiple copies of duplicated genes were functionally diversified and have contributed to the formation of distinctive phenotypes for individual species.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

## AUTHOR CONTRIBUTIONS

A-YS and Y-MK wrote the manuscript, conceived the project, designed the experiments, and organized the manuscript. SH, YL, S-YK, and A-YS generated the data and performed bioinformatic analyses. All authors contributed to the article and approved the submitted version.

## REFERENCES

- Akhtar, J., Goyal, A., Kaur, N., Atri, C., Mittal, M., Singh, M. P., et al. (2021). Genome wide association analyses to understand genetic basis of flowering and plant height under three levels of nitrogen application in Brassica juncea (L.) Czern & Coss. *Sci. Rep.* 11:4278.
- Argout, X., Salse, J., Aury, J. M., Guiltinan, M. J., Droc, G., Gouzy, J., et al. (2011). The genome of theobroma cacao. *Nat. Genet.* 43, 101–108.
- Bouche, F., Lobet, G., Tocquin, P., and Perilleux, C. (2016). FLOR-ID: an interactive database of flowering-time gene networks in *Arabidopsis thaliana*. *Nucleic Acids Res.* 44, D1167–D1171.
- Buckler, E. S., Holland, J. B., Bradbury, P. J., Acharya, C. B., Brown, P. J., Browne, C., et al. (2009). The genetic architecture of maize flowering time. *Science* 325, 714–718.
- Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A., Tang, H., Wang, X., et al. (2014). Plant genetics. early allopolyploid evolution in the post-Neolithic brassica napus oilseed genome. *Science* 345, 950–953.
- Chen, Z. J., Sreedasyam, A., Ando, A., Song, Q., De Santiago, L. M., Hulse-Kemp, A. M., et al. (2020). Genomic diversifications of five gossypium allopolyploid species and their impact on cotton improvement. *Nat. Genet.* 52, 525–533. doi: 10.1038/s41588-020-0614-5
- Cheng, C. Y., Krishnakumar, V., Chan, A. P., Thibaud-Nissen, F., Schobel, S., and Town, C. D. (2017). Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *Plant J.* 89, 789–804. doi: 10.1111/tpj.13415
- Dodsworth, S., Chase, M. W., and Leitch, A. R. (2015). Is post-polyploidization diploidization the key to the evolutionary success of angiosperms? *Bot. J. Linn. Soc.* 180, 1–5. doi: 10.1111/boj.12357
- Emms, D. M., and Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 16:157. doi: 10.1186/s13059-015-0721-2
- Flayel, L. E., and Wendel, J. F. (2009). Gene duplication and evolutionary novelty in plants. *N. Phytol.* 183, 557–564. doi: 10.1111/j.1469-8137.2009.02923.x
- Freeling, M., Scanlon, M. J., and Fowler, J. E. (2015). Fractionation and subfunctionalization following genome duplications: mechanisms that drive gene content and their consequences. *Curr. Opin. Genet. Dev.* 35, 110–118. doi: 10.1016/j.gde.2015.11.002
- Gao, B. Z., Liu, B., Li, S., Liang, J., Cheng, F., Wang, X., et al. (2017). Genome-wide association studies for flowering time in brassica rapa. *Sci. Agric. Sin.* 50, 3375–3385.

## FUNDING

This work was supported by the Korea Forest Service (KFS) of the Korean Government, through its R&D Program for Forestry Technology (Project No. 2014071H10-2122-AA04), and the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, and Forestry (IPET), through the Golden Seed Project (213006-05-5-SBG30) and Technology Commercialization Support Program (821026-03), which is funded by a grant from the Ministry of Agriculture, Food and Rural Affairs (MAFRA), Ministry of Oceans and Fisheries (MOF), Rural Development Administration (RDA), and Korea Forest Service (KFS) to Y-MK.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found on Figshare. Supplementary Tables 1 and 5, and Supplementary Figures 1–11 (doi: 10.6084/m9.figshare.14822070). Supplementary Tables 2–4 and 6–9 (doi: 10.6084/m9.figshare.14822061).

- Gill, R. A., Scossa, F., King, G. J., Golicz, A., Tong, C., Snowdon, R. J., et al. (2021). On the role of transposable elements in the regulation of gene expression and subgenomic interactions in crop genomes. *Crit. Rev. Plant Sci.* 40, 157–189. doi: 10.1080/07352689.2021.1920731
- Grover, C. E., Gallagher, J. P., and Wendel, J. F. (2015). Candidate gene identification of flowering time genes in cotton. *Plant Genome* 8:lantgenome2014.12.0098.
- Halary, S., Mcinerney, J. O., Lopez, P., and Baptiste, E. (2013). EGN: a wizard for construction of gene and genome similarity networks. *BMC Evol. Biol.* 13:146. doi: 10.1186/1471-2148-13-146
- Hao, Z., Lv, D., Ge, Y., Shi, J., Weijers, D., Yu, G., et al. (2020). Rldeogram: drawing SVG graphics to visualize and map genome-wide data on the idiograms. *PeerJ Comput. Sci.* 6:e251. doi: 10.7717/peerj-cs.251
- Hu, Y., Chen, J., Fang, L., Zhang, Z., Ma, W., Niu, Y., et al. (2019). Gossypium barbadense and *Gossypium hirsutum* genomes provide insights into the origin and evolution of allotetraploid cotton. *Nat. Genet.* 51, 739–748. doi: 10.1038/s41588-019-0371-5
- Huang, X., Zhao, Y., Wei, X., Li, C., Wang, A., Zhao, Q., et al. (2011). Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat. Genet.* 44, 32–39. doi: 10.1038/ng.1018
- Islam, M. S., Saito, J. A., Emdad, E. M., Ahmed, B., Islam, M. M., Halim, A., et al. (2017). Comparative genomics of two jute species and insight into fibre biogenesis. *Nat. Plants* 3:16223.
- Kersey, P. J. (2019). Plant genome sequences: past, present, future. *Curr. Opin. Plant Biol.* 48, 1–8. doi: 10.1016/j.pbi.2018.11.001
- Kim, S., Park, J., Yeom, S. I., Kim, Y. M., Seo, E., Kim, K. T., et al. (2017). New reference genome sequences of hot pepper reveal the massive evolution of plant disease-resistance genes by retroduplication. *Genome Biol.* 18:210.
- Kim, S., Park, M., Yeom, S. I., Kim, Y. M., Lee, J. M., Lee, H. A., et al. (2014). Genome sequence of the hot pepper provides insights into the evolution of pungency in Capsicum species. *Nat. Genet.* 46, 270–278.
- Kim, Y. M., Kim, S., Koo, N., Shin, A. Y., Yeom, S. I., Seo, E., et al. (2017). Genome analysis of *Hibiscus syriacus* provides insights of polyploidization and indeterminate flowering in woody plants. *DNA Res.* 24, 71–80.
- Koo, N., Shin, A. Y., Oh, S., Kim, H., Hong, S., Park, S. J., et al. (2020). Comprehensive analysis of Translationally Controlled Tumor Protein (TCTP) provides insights for lineage-specific evolution and functional divergence. *PLoS One* 15:e0232029. doi: 10.1371/journal.pone.0232029

## RESEARCH

## Open Access



# Identification of accession-specific variants and development of KASP markers for assessing the genetic makeup of *Brassica rapa* seeds

Seongmin Hong<sup>1,2</sup>, Su Ryun Choi<sup>2,3</sup>, Jihyeong Kim<sup>4</sup>, Young-Min Jeong<sup>4</sup>, Ju-Sang Kim<sup>5</sup>, Chun-Hee Ahn<sup>5</sup>, Suk-Yoon Kwon<sup>6</sup>, Yong Pyo Lim<sup>2†</sup>, Ah-Young Shin<sup>6†</sup> and Yong-Min Kim<sup>1\*†</sup>

## Abstract

**Background:** Most crop seeds are F1 hybrids. Seed providers and plant breeders must be confident that the seed supplied to growers is of known, and uniform, genetic makeup. This requires maintenance of pure genotypes of the parental lines and testing to ensure the genetic purity of the F1 seed. Traditionally, seed purity has been assessed with a grow-out test (GOT) in the field, a time consuming and costly venture. Early in the last decade, seed testing with molecular markers was introduced as a replacement for GOT, and Kompetitive allele specific PCR (KASP) markers were recognized as promising tools for genetic testing of seeds. However, the markers available at that time could be inaccurate and applicable to only a small number of accessions or varieties due to the limited genetic information and reference genomes available.

**Results:** We identified 4,925,742 SNPs in 50 accessions of the *Brassica rapa* core collection. From these, we identified 2,925 SNPs as accession-specific, considering properties of flanking region harboring accession-specific SNPs and genic region conservation among accessions by the Next Generation Sequencing (NGS) analysis. In total, 100 accession-specific markers were developed as accession-specific KASP markers. Based on the results of our validation experiments, the accession-specific markers successfully distinguished individuals from the mixed population including 50 target accessions from *B. rapa* core collection and the outgroup. Additionally, the marker set we developed here discriminated F1 hybrids and their parental lines with distinct clusters.

**Conclusions:** This study provides efficient methods for developing KASP markers to distinguish individuals from the mixture comprised of breeding lines and germplasm from the resequencing data of Chinese cabbage (*Brassica rapa* spp. *pekinensis*).

**Keywords:** Seed purity assessment, KASP marker, Accession-specific marker, *Brassica rapa*, *B. rapa* breeding

## Background

Most growers of vegetable crops rely on F1 hybrid seeds, and suppliers of these seeds must maintain genetically pure stocks. Not only do the suppliers need to keep seeds of known genetic makeup for sales but also for their ongoing breeding programs. Until the late 1990s, seed providers relied on what is known as the grow-out test (GOT), in which the seeds were planted in the field and the traits of the test plants were assessed by investigation [1]. However, this method is time consuming, requires

\*Correspondence:

<sup>†</sup>Yong Pyo Lim, Ah-Young Shin and Yong-Min Kim contributed equally to this work and share last authorship.

<sup>1</sup> Genome Editing Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, Republic of Korea

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

**Acknowledgements**

Not applicable.

**Authors' contributions**

Y.-M.K. conceived the project, designed the experiments, and organized the manuscript. S.R.C. and Y.P.L. generated and maintained the core collection. J.-S.K. and C.-H.A. provided breeding lines for KASP marker assay. S.H., A.-Y.S., S.Y.K., Y.P.L., and Y.-M.K. performed the bioinformatics analysis and development of the KASP markers. J.K. and Y.-M.J. performed validation of KASP markers. A.-Y.S. and Y.-M.K. wrote the manuscript. The author(s) read and approved the final manuscript.

**Funding**

This work was supported by the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, and Forestry through the Golden Seed Project (213006-05-5-SBG30) and the **Technology Commercialization Support Program (821026-03-2-SB010)** funded by a grant from the Ministry of Agriculture, Food, and Rural Affairs, the Ministry of Oceans and Fisheries, the Rural Development Administration, and the Korea Forest Service supported to Y.-M.K. The latter is funded by the Korea Forest Service of the Korean government through its Research and Development Program for Forestry Technology (Project No. 20140711H10-2122-AA04) supported to Y.-M.K.

**Availability of data and materials**

The datasets have been deposited at NCBI under BioProject number PRJNA787013. The whole genome resequencing data for 50 *B. rapa* accessions are available through the NCBI Sequence Read Archive (<https://www.ncbi.nlm.nih.gov/sra/SRP354760>) with Bam file format.

**Declarations****Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

S.H., S.R.C., J.K., Y.-M.J., J.-S.K., C.-H.A., S.Y.K., Y.P.L., A.-Y.S., and Y.-M.K. declare that they have no competing interest.

**Author details**

<sup>1</sup>Genome Editing Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, Republic of Korea. <sup>2</sup>Molecular Genetics and Genomics Laboratory, Department of Horticulture, College of Agriculture and Life Science, Chungnam National University, Daejeon, Republic of Korea. <sup>3</sup>XENOTYPE Inc., 118 Jungang-ro, Jung-gu, Daejeon 34912, Republic of Korea. <sup>4</sup>Seed Industry Promotion Center, Foundation of Agri. Tech. Commercialization & Transfer (FACT), Gimje, Republic of Korea. <sup>5</sup>Dayi International Seed Co., Ltd., 16-35 Ssiat-gilJeollabuk-do, Baeksan-myeon, Gimje 54324, Korea. <sup>6</sup>Plant Systems Engineering Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon, Republic of Korea.

Received: 27 November 2021 Accepted: 19 April 2022

Published online: 25 April 2022

**References**

- Ballester J, de Vicente MC. Determination of F1 hybrid seed purity in pepper using PCR-based markers. *Euphytica*. 1998;103(2):223–6.
- Saxena RK, Saxena K, Varshney RK. Application of SSR markers for molecular characterization of hybrid parents and purity assessment of ICPH 2438 hybrid of pigeonpea [*Cajanus cajan* (L.) Millspaugh]. *Mol Breed*. 2010;26(2):371–80.
- Sundaram RM, Naveenkumar B, Biradar SK, Balachandran SM, Mishra B, IlyasAhmed M, Viraktamath BC, Ramesha MS, Sarma NP. Identification of informative SSR markers capable of distinguishing hybrid rice parental lines and their utilization in seed purity assessment. *Euphytica*. 2008;163(2):215–24.
- Komori T, Nitta N. A simple method to control the seed purity of japonica hybrid rice varieties using PCR-based markers. *Plant Breed*. 2004;123(6):549–53.
- Tomato Genome C. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*. 2012;485(7400):635–41.
- Kim S, Park M, Yeom SJ, Kim YM, Lee JM, Lee HA, Seo E, Choi J, Cheong K, Kim KT, et al. Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. *Nat Genet*. 2014;46(3):270–8.
- Li Q, Li H, Huang W, Xu Y, Zhou Q, Wang S, Ruan J, Huang S, Zhang Z. A chromosome-scale genome assembly of cucumber (*Cucumis sativus* L.). *Gigascience*. 2019;8(6):giz072.
- Garcia-Mas J, Benjak A, Sansaverino W, Bourgeois M, Mir G, Gonzalez VM, Henaff E, Camara F, Cozzuto L, Lowy E, et al. The genome of melon (*Cucumis melo* L.). *Proc Natl Acad Sci U S A*. 2012;109(29):11872–7.
- Shin AY, Koo N, Kim S, Sim YM, Choi D, Kim YM, Kwon SY. Draft genome sequences of two oriental melons, *Cucumis melo* L. var. makuwa. *Sci Data*. 2019;6(1):220.
- International Wheat Genome Sequencing C, investigators IRp, Appels R, Eversole K, Feuillet C, Keller B, Rogers J, Stein N, investigators Iw-gap, Pozniak CJ, et al. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*. 2018;361(6403):eaar7191.
- Zhang L, Cai X, Wu J, Liu M, Grob S, Cheng F, Liang J, Cai C, Liu Z, Liu B, et al. Improved *Brassica rapa* reference genome by single-molecule sequencing and chromosome conformation capture technologies. *Hortic Res*. 2018;5:50.
- Semagn K, Babu R, Hearne S, Olsen M. Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. *Mol Breeding*. 2014;33(1):1–14.
- Rasheed A, Wen W, Gao F, Zhai S, Jin H, Liu J, Guo Q, Zhang Y, Dreisigacker S, Xia X, et al. Development and validation of KASP assays for genes underpinning key economic traits in bread wheat. *Theor Appl Genet*. 2016;129(10):1843–60.
- Jagtap AB, Vikal Y, Johal GS. Genome-Wide Development and Validation of Cost-Effective KASP Marker Assays for Genetic Dissection of Heat Stress Tolerance in Maize. *Int J Mol Sci*. 2020;21(19):7386.
- Chen Q, Song J, Du WP, Xu LY, Jiang Y, Zhang J, Xiang XL, Yu GR. Identification, Mapping, and Molecular Marker Development for Rgsr8.1: A New Quantitative Trait Locus Conferring Resistance to Gibberella Stalk Rot in Maize (*Zea mays* L.). *Front Plant Sci*. 2017;8:1355.
- Du X, Huang G, He S, Yang Z, Sun G, Ma X, Li N, Zhang X, Sun J, Liu M, et al. Resequencing of 243 diploid cotton accessions based on an updated A genome identifies the genetic basis of key agronomic traits. *Nat Genet*. 2018;50(6):796–802.
- Lu K, Wei L, Li X, Wang Y, Wu J, Liu M, Zhang C, Chen Z, Xiao Z, Jian H, et al. Whole-genome resequencing reveals *Brassica napus* origin and genetic loci involved in its improvement. *Nat Commun*. 2019;10(1):1154.
- Yuan J, Wang X, Zhao Y, Khan NU, Zhao Z, Zhang Y, Wen X, Tang F, Wang F, Li Z. Genetic basis and identification of candidate genes for salt tolerance in rice by GWAS. *Sci Rep*. 2020;10(1):9958.
- Su T, Li P, Yang J, Sui G, Yu Y, Zhang D, Zhao X, Wang W, Wen C, Yu S, et al. Development of cost-effective single nucleotide polymorphism marker assays for genetic diversity analysis in *Brassica rapa*. *Mol Breed*. 2018;38(4):1–3.
- Pang W, Li X, Choi SR, Dhandapani V, Im S, Park MY, Jang CS, Yang M-S, Ham IK, Lee EM, et al. Development of a leafy *Brassica rapa* fixed line collection for genetic diversity and population structure analysis. *Mol Breeding*. 2015;35(1):54.
- Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–20.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754–60.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. Genome Project Data Processing S: The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–9.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297–303.



# OPEN Construction of deep learning-based disease detection model in plants

Minah Jung<sup>1,2,9</sup>, Jong Seob Song<sup>2,9</sup>, Ah-Young Shin<sup>3,4,9</sup>, Beomjo Choi<sup>3,5,9</sup>, Sangjin Go<sup>3</sup>, Suk-Yoon Kwon<sup>3,6</sup>, Juhan Park<sup>2</sup>, Sung Goo Park<sup>1,7,8</sup> & Yong-Min Kim<sup>3,4,8</sup>✉

Accurately detecting disease occurrences of crops in early stage is essential for quality and yield of crops through the decision of an appropriate treatments. However, detection of disease needs specialized knowledge and long-term experiences in plant pathology. Thus, an automated system for disease detecting in crops will play an important role in agriculture by constructing early detection system of disease. To develop this system, construction of a stepwise disease detection model using images of diseased-healthy plant pairs and a CNN algorithm consisting of five pre-trained models. The disease detection model consists of three step classification models, crop classification, disease detection, and disease classification. The 'unknown' is added into categories to generalize the model for wide application. In the validation test, the disease detection model classified crops and disease types with high accuracy (97.09%). The low accuracy of non-model crops was improved by adding these crops to the training dataset implicating expendability of the model. Our model has the potential to apply to smart farming of Solanaceae crops and will be widely used by adding more various crops as training dataset.

Crop disease management is important in agriculture to increase yield and quality by reducing the economic and aesthetic damage caused by plant diseases. Although research into the causes and effective treatments for crop diseases is actively underway, monitoring plant health and early detection of pathogens are critical to reduce disease spread and facilitate effective management<sup>1</sup>. Detecting and protecting crops from pathogens is labor-intensive and time-consuming, making it virtually impossible for humans to analyze each plant<sup>2</sup>. Therefore, research on combining and applying new technologies to efficiently detect diseases has been conducted, and recently, research on detecting plant diseases in leaves using artificial intelligence (AI) is in progress<sup>3</sup>. Continued development of improved classification models, such as disease detection, or plant health monitoring, may enable AI-supported decision-making systems for smart agriculture<sup>4</sup>. Various studies have been carried out to apply deep learning algorithms more precisely to disease detection, such as applying newly developed architectures<sup>5,6</sup>, automatically detecting and classifying lesions in plant images<sup>7</sup>, or conducting research on preprocessing methods for incomplete images<sup>8</sup> for practical use.

Artificial Neural Network (ANN) is an AI technology with an analytic system inspired by the nerve system of the human brain that mimics the way the brain processes information<sup>9</sup>. ANN contains a three component processing unit consisting of input, hidden, and output layers<sup>10</sup>. Nodes of individual layers are connected to nodes of adjacent layers. Convolutional Neural Network (CNN) is a specialized method to recognize or assign images and consists of fully connected layers, numerous convolution layers, and pooling layers<sup>11</sup>. Three types of layers are arranged and connected differently depending on the model architectures and model performance is affected by this architecture<sup>12</sup>. AlexNet<sup>13</sup>, VGG19<sup>14</sup>, GoogLeNet<sup>15</sup>, ResNet<sup>16</sup>, and EfficientNet<sup>17</sup> are pre-trained

<sup>1</sup>Department of Functional Genomics, KRIBB School of Biological Science, Korea University of Science and Technology (UST), Daejeon, Republic of Korea. <sup>2</sup>Euclidsoft Co., Ltd, Daejeon, Republic of Korea. <sup>3</sup>Plant Systems Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, Republic of Korea. <sup>4</sup>Department of Bioinformatics, KRIBB School of Bioscience, Korea University of Science and Technology (UST), Daejeon, Republic of Korea. <sup>5</sup>Department of Environmental Horticulture, University of Seoul, Seoul, Republic of Korea. <sup>6</sup>Biosystems and Bioengineering Program, KRIBB School of Bioscience, Korea University of Science and Technology (UST), Daejeon, Korea. <sup>7</sup>Disease Target Structure Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, Republic of Korea. <sup>8</sup>Digital Bioinnovation Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, Republic of Korea. <sup>9</sup>These authors contributed equally: Minah Jung, Jong Seob Song, Ah-Young Shin and Beomjo Choi. ✉email: [yongmin.kim@krribb.ac.kr](mailto:yongmin.kim@krribb.ac.kr)

7. Albattah, W., Nawaz, M., Javed, A., Masood, M. & Albahli, S. A novel deep learning method for detection and classification of plant diseases. *Complex Intell. Syst.* **8**, 507–524 (2022).
8. Harakannavar, S. S., Rudagi, J. M., Puranikmath, V. I., Siddiqua, A. & Pramodhini, R. Plant leaf disease detection using computer vision and machine learning algorithms. *Glob. Trans. Proc.* (2022).
9. Maind, S. B. & Wankar, P. Research paper on basic of artificial neural network. *Int. J. Recent Innov. Trends Comput. Commun.* **2**, 96–100 (2014).
10. Keskin, M. E. & Terzi, Ö. Artificial neural network models of daily pan evaporation. *J. Hydrol. Eng.* **11**, 65–70 (2006).
11. Yamashita, R., Nishio, M., Do, R. K. G. & Togashi, K. Convolutional neural networks: An overview and application in radiology. *Insights Imaging* **9**, 611–629 (2018).
12. Zia, T. & Zahid, U. Long short-term memory recurrent neural network architectures for Urdu acoustic modeling. *Int. J. Speech Technol.* **22**, 21–30 (2019).
13. Krizhevsky, A., Sutskever, I. & Hinton, G. E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **25**, 1097–1105 (2012).
14. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556) (2014).
15. Szegedy, C. *et al.* In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 1–9.
16. He, K., Zhang, X., Ren, S. & Sun, J. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 770–778.
17. Tan, M. & Le, Q. In *International Conference on Machine Learning*. 6105–6114 (PMLR).
18. Russakovsky, O. *et al.* Imagenet large scale visual recognition challenge. *Int. J. Comput. Vis.* **115**, 211–252 (2015).
19. Dhaka, V. S. *et al.* A survey of deep convolutional neural networks applied for prediction of plant leaf diseases. *Sensors* **21**, 4749 (2021).
20. Jung, M. *et al.* Deep learning algorithms correctly classify *Brassica rapa* varieties using digital images. *Front. Plant Sci.* **12** (2021).
21. Shrestha, A. & Mahmood, A. Review of deep learning algorithms and architectures. *IEEE Access* **7**, 53040–53065 (2019).
22. Lu, J., Tan, L. & Jiang, H. Review on convolutional neural network (CNN) applied to plant leaf disease classification. *Agriculture* **11**, 707 (2021).
23. Udayananda, G. & Kumara, P. In *2022 3rd International Conference for Emerging Technology (INCET)*. 1–7 (IEEE).
24. Ubbens, J. R. & Stavness, I. Deep plant phenomics: A deep learning platform for complex plant phenotyping tasks. *Front. Plant Sci.* **8**, 1190 (2017).
25. Lee, S. H., Chan, C. S., Wilkin, P. & Remagnino, P. In *2015 IEEE International Conference on Image Processing (ICIP)*. 452–456 (IEEE).
26. Minarno, A. E. *et al.* In *2022 10th International Conference on Information and Communication Technology (ICOICT)*. 99–104 (IEEE).
27. Wu, S. G. *et al.* In *2007 IEEE International Symposium on Signal Processing and Information Technology*. 11–16 (IEEE).
28. Moyazoma, R., Hossain, M. A. A., Anuz, M. H. & Sattar, A. In *2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)*. 526–529 (IEEE).
29. Bhagat, M., Kumar, D. & Kumar, S. Bell pepper leaf disease classification with LBP and VGG-16 based fused features and RF classifier. *Int. J. Inf. Technol.* **15**, 1–11 (2022).
30. Sholihati, R. A., Sulistijono, I. A., Risnumawan, A. & Kusumawati, E. In *2020 International Electronics Symposium (IES)*. 392–397 (IEEE).
31. Kaushik, M., Prakash, P., Ajay, R. & Veni, S. In *2020 5th International Conference on Communication and Electronics Systems (ICCES)*. 1125–1132 (IEEE).
32. Jasim, M. A. & Al-Tuwaijari, J. M. In *2020 International Conference on Computer Science and Software Engineering (CSASE)*. 259–265 (IEEE).
33. Kurmi, Y. *et al.* Deep CNN model for crops' diseases detection using leaf images. *Multidimension. Syst. Signal Process.* **33**, 981–1000 (2022).
34. Singh, L., Pandey, M. & Lakra, S. In *2022 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES)*. 487–492 (IEEE).
35. Al-Tuwaijari, J. M., Jasim, M. A. & Raheem, M. A.-B. In *2020 2nd Al-Noor International Conference for Science and Technology (NICST)*. 7–12 (IEEE).
36. Liu, J., Wang, M., Bao, L. & Li, X. In *Journal of Physics: Conference Series*. 012148 (IOP Publishing).
37. Srinidhi, V., Sahay, A. & Deebea, K. In *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*. 1119–1127 (IEEE).
38. Dheeraj, G., Anumala, P. K., Sagar, L. R., Krishna, B. V. & Bala, I. In *2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS)*. 1429–1434 (IEEE).
39. Lawal, M. O. Tomato detection based on modified YOLOv3 framework. *Sci. Rep.* **11**, 1–11 (2021).
40. Waheed, A. *et al.* An optimized dense convolutional neural network model for disease recognition and classification in corn leaf. *Comput. Electron. Agric.* **175**, 105456 (2020).
41. Ucat, R. C. & Cruz, J. C. D. In *2019 International Symposium on Multimedia and Communication Technology (ISMAT)*. 1–6 (IEEE).
42. Hughes, D. & Salathé, M. An open access repository of images on plant health to enable the development of mobile disease diagnostics. arXiv preprint [arXiv:1511.08060](https://arxiv.org/abs/1511.08060) (2015).
43. Moussafir, M. *et al.* Design of efficient techniques for tomato leaf disease detection using genetic algorithm-based and deep neural networks. *Plant Soil* **479**, 251–266 (2022).
44. Ahmad, I., Hamid, M., Yousaf, S., Shah, S. T. & Ahmad, M. O. Optimizing pretrained convolutional neural networks for tomato leaf disease detection. *Complexity* **2020** (2020).
45. Hong, M. *et al.* Multi-class classification of lung diseases using CNN models. *Appl. Sci.* **11**, 9289 (2021).

### Author contributions

Y.-M.K. conceived the project, designed the analysis, and organized the manuscript. M.J., J.S.S., A.-Y.S., and Y.-M.K. performed the deep-learning analysis. B.C., A.-Y.S., S.G., J.P., S.-Y.K., and S.G.P. generated the phenotypic data. M.J., A.-Y.S., S.G.P., and Y.-M.K. wrote the manuscript.

### Funding

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education [NRF-2021R111A2044678], the **Technology Commercialization Support Program [821026-03]** (funded by a grant from the Ministry of Agriculture, Food, and Rural Affairs; the Ministry of Oceans and Fisheries; the Rural Development Administration; and the Korea Forest Service), and the Korea Forest Service of the Korean government through its R&D Program for Forestry Technology [Project No. 2014071H10-2122-AA04] to Y.-M.K.

P3-32

## Construction of classification models for *Brassica rapa* core collections using deep learning algorithms

Minah Jung<sup>1,2†</sup>, Jong Seob Song<sup>2†</sup>, Seongmin Hong<sup>3,4</sup>, SunWoo Kim<sup>5</sup>, Sangjin Ko<sup>3</sup>, Yong Pyo Lim<sup>4</sup>, Juhan Park<sup>2</sup>, Sung Goo Park<sup>1,6</sup> and Yong-Min Kim<sup>3\*</sup>

<sup>1</sup>Department of Functional Genomics, KRIBB School of Biological Science, Korea University of Science and Technology (UST), Daejeon, Republic of Korea

<sup>2</sup>Euclidsoft Co., Ltd, Daejeon 35214, Republic of Korea

<sup>3</sup>Genome Editing Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, Republic of Korea

<sup>4</sup>Molecular Genetics and Genomics Laboratory, Department of Horticulture, College of Agriculture and Life Science, Chungnam National University, Daejeon, Korea

<sup>5</sup>Department of Bio-AI Convergence, Chungnam National University, Daejeon, Korea

<sup>6</sup>Disease Target Structure Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, Republic of Korea

Huge amount of biological data was accumulated in various research fields including genomics, phenomics, and genetics. As amount of biological data has been largely generated and accumulated, efficient and accurate analysis methods to be required. Artificial intelligence-based analysis is one of promising methods to manipulate biological data. To archive this end, various algorithms have been developed and applied for various research fields such as disease diagnosis, species classification, and object prediction. In phenomics area, classification of accessions or variants are one of important for basic science and industrial applications. To construct artificial intelligence-based classification models, three types of phenotypic image data were generated from 156 *Brassica rapa* core collections and classification analyses were carried out using four different Convolutional Neural Network architectures. Results of lateral view data showed higher accuracy compared to top view data. Furthermore, relatively low accuracy of ResNet50 architecture suggested that definition and estimation of similarity index of phenotypic data were required before selection of deep learning architectures.

**Keywords:** Artificial Intelligence, Deep learning, Classification model, Phenotypic analysis, *Brassica rapa*

\*Corresponding author: e-mail: [REDACTED]

†These authors contributed equally to this work



**2021 INTERNATIONAL CONFERENCE**  
OF THE KOREAN SOCIETY FOR  
MOLECULAR AND CELLULAR BIOLOGY

NOV 3-5, 2021  
JEJU ICC, JEJU, KOREA

**K S M C B**

## Certificate of Poster Presentation

**NAME :** Seongmin Hong<sup>1,2</sup>, Yong Pyo Lim<sup>2</sup>, Suk-Yoon Kwon<sup>3</sup>, Ah-Young Shin<sup>3,\*</sup>, and Yong-Min Kim<sup>1,\*</sup>

**AFFILIATION :** <sup>1</sup>Genome Editing Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, South Korea, <sup>2</sup>Molecular Genetics and Genomics Laboratory, Department of Horticulture, College of Agriculture and Life Sciences, Chungnam National University, Daejeon, South Korea, <sup>3</sup>Plant Systems Engineering Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon, South Korea

**TITLE :** Genome-Wide Comparative Analysis of Flowering-Time Genes; Insights on the Gene Family Expansion and Evolutionary Perspective

The Korean Society for Molecular and Cellular Biology hereby certificate that above individual had presented in the International Conference of Korean Society for Molecular and Cellular Biology which was held on Nov. 3 - Nov. 5, 2021, ICC JEJU, Jeju, Korea.

November 24, 2021

  
Seong Rho Hyun, Ph.D.  
2021 President  
Korean Society for Molecular and Cellular Biology

International Conference  
**KSMCB 2021**  
KOREAN SOCIETY FOR MOLECULAR AND CELLULAR BIOLOGY

COPYRIGHT © KOREAN SOCIETY FOR MOLECULAR AND CELLULAR BIOLOGY. ALL RIGHTS RESERVED.  
The Korea Science Technology Center (Rm. 1105), 22, 7Gil, Teheran-ro, Gangnam-gu, Seoul 06130, Korea  
Tel. 82-2-568-4490 | Fax: 82-2-558-0131 | home@ksmcb.or.kr | Business No. : 220-82-02552

## ▶▶▶ F-05

**De novo assembly and annotation of Hibiscus syriacus Baekdansim genome using PacBio sequencing**Seongmin Hong<sup>1</sup>, Young Mi Sim<sup>2</sup>, Ah-Young Shin<sup>3</sup>, Suk-Yoon Kwon<sup>3</sup>, \*Yong-Min Kim<sup>1</sup><sup>1</sup>Genome Editing Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), <sup>2</sup>Korean Bioinformation center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), <sup>3</sup>Plant Systems Engineering Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB)

Whole genome duplication or triplication have been occurred in plant genomes during their evolution and these whole genome duplication and triplication also have played an important role in emergence of new species in plants. Hibiscus syriacus genome belonged to Malvaceae is a polyploidy genome caused by repetitive whole genome duplication. This features of H. syriacus genome becomes a barrier to construct high quality reference genome by the bioinformatics formation of chimeric sequences and fragmented contigs. Long reads or ultra-long reads from third-generation sequencing platforms become a solution to construct high quality reference genomes in plant genomes containing large amount of duplication regions. We report on high quality draft genome of H. syriacus Baekdansim known as cold resistance plants using Pac-Bio sequencing platform. The draft genome comprised of 1.82 Gb of assembly covers ~ 92.8% of the genome with a contig N50 of 2.52 Mb and 88,467 genes were predicted. These draft genome sequence of H. syriacus Baekdansim will be used for further analysis to investigate cold resistance trait.

## ▶▶▶ F-06

**Comprehensive Analysis of Genes and MicroRNAs Affecting Pre-harvest Sprouting Between two Rice (*Oryza sativa* L.) Accessions**Sang-Yoon Shin<sup>1</sup>, Minsu Park<sup>2</sup>, Woochang Choi<sup>2</sup>, Hongman Moon<sup>2</sup>, Dowhan Lee<sup>2</sup>, Yun-Shil Gho<sup>4</sup>, Ki-Hong Jung<sup>4</sup>, Jong-Seong Jeon<sup>4</sup>, \*Chanseok Shin<sup>3</sup><sup>1</sup>Interdisciplinary Program in Agricultural genomics, Seoul National University, Seoul, Korea, <sup>2</sup>Department of Agricultural Biotechnology, Seoul National University, Seoul, Korea, <sup>3</sup>Department of Agricultural Biotechnology, Seoul National University, Seoul, Korea, <sup>4</sup>Research Center for Plant Plasticity, Seoul National University, Seoul, Korea, <sup>4</sup>Graduate School of Biotechnology and Crop Biotech Institute, Kyung Hee University, Yongin, Korea

Pre-harvest sprouting (PHS) is one of the primary problems caused by the failure in maintaining seed dormancy in rice (*Oryza sativa* L.). PHS causes yield loss and reduces grain quality under unpredictable humid conditions at the ripening stage, leading to damages in productivity and economic values of the rice. Drastic changes in world-wide climate in recent years would deepen the problems caused by prolonged rainy seasons or unpredicted rainfalls during seed maturation stage. To understanding the molecular mechanism underlying seed dormancy in rice, we analyzed the transcriptome and small RNAome of embryo and endosperm of seeds from two contrasting japonica rice accessions, PHS susceptible (with low seed dormancy) and PHS resistant (with high seed dormancy), at three different maturation stages. Differential expression analysis between embryos / endosperms of two accessions found that 9,068 genes and 35 microRNAs (miRNAs) were differentially expressed in the embryo, whereas 360 genes were differentially expressed in the endosperm. Furthermore, we identified and verified the candidate genes associated with seed dormancy and heat stress-related responses in rice through quantitative real-time PCR, leading to selection of eight hormone-related genes, four heat-shock protein-related genes, and two miRNAs potentially involved in PHS. These would provide a foundation for understanding the dynamics of transcriptome and small RNAome of hormone- and heat stress-related genes affecting regulation of PHS during seed maturation. This work was carried out with the support of 'Cooperative Research Program for Agriculture Science and Technology Development (Project No. PJ01577601)' Rural Development Administration, Republic of Korea.

0021

## Construction of Deep Learning-Based Disease Detection Model in Plants

Minah Jung<sup>1,2</sup>, Jong Seob Song<sup>2</sup>, Ah-Young Shin<sup>3</sup>, Beomjo Choi<sup>4</sup>, Sangjin Go<sup>3</sup>, Suk-Yoon Kwon<sup>3</sup>, Juhan Park<sup>2</sup>, Sung Goo Park<sup>1,5</sup>, Yong-Min Kim<sup>3</sup>

<sup>1</sup>Department of Functional Genomics, KRIBB School of Biological Science, Korea University of Science and Technology (UST), Daejeon, Republic of Korea

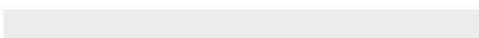
<sup>2</sup>Euclidsoft Co., Ltd, Daejeon, Republic of Korea

<sup>3</sup>Plant Systems Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, Republic of Korea

<sup>4</sup>Department of Environmental Horticulture, University of Seoul, Seoul, Republic of Korea

<sup>5</sup>Disease Target Structure Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, Republic of Korea

Construction of deep learning-based disease detection model using images of diseased-healthy plant pairs and a CNN algorithm consisting of five pre-trained models. The disease detection model consists of three step classification models, crop classification, disease detection, and disease classification. In the validation test, the disease detection model classified crops and disease types with high accuracy (97.09%). The low accuracy of non-model crops was improved by adding these crops to the training dataset for model construction. Also, cropped images allow higher accuracy with more obvious disease symptoms or lesion patterns. The accuracy of the classification model depended on the depth, quality, and variety of image data for training during model construction. Thus, to apply this classification model, more high quality images of various crops are required for model construction, as well as an approach for developing a platform to apply this model to smart farming.

\*Corresponding Author: 

P3-33

### Construction of high-contiguity reference genome sequences of two *Hibiscus syriacus* accessions (Gangneung and Baekdansim)

Sangjin Go, Seongmin Hong, Ah-Young Shin and Yong-Min Kim\*

*Plant Systems Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB),  
Daejeon, Republic of Korea*

Rose of sharon (*Hibiscus syriacus*) one of the most widespread garden shrubs in the world, showed a complexed structure of genome generated by multiple number of polyploidization during its diversification and evolution. We constructed two reference genome assemblies for *H. syriacus* accessions (Gangneung and Baekdansim) with long read sequencing and Hi-C scaffolding, resulting in 1.87 Gbp and 1.82 Gbp size of assembly. Total 88,741 and 87,438 gene models were predicted from Gangneung and Baekdansim genome assemblies respectively. To identify genetic varieties putatively associated to horticultural traits of the *H. syriacus*, we re-sequenced 10 accessions showed different phenotypes related to cold-resistance and floral architecture. As a result, we identified total 4,339,048 SNPs based on the reference genome of Gangneung. The phylogenetic trees drawn by four-fold degenerate sites (4D SNPs) clustered accessions corresponding their known origins and crossing information. Total 1,051 number of SNPs were located on 268 number of cold-related genes(CRGs) searched by sequence similarity against 114 *A. thaliana* CRGs, resulting in detection of 19 cold-resistant accession/group specific SNPs. The reference genome will serve as a platform for developing cold-tolerance varieties and improved floral traits in the *H. syriacus* breeding program.

\*Corresponding author: e-mail:



## Certificate of Poster Presentation

**NAME :** Seongmin Hong, Suk-Yoon Kwon, Ah-Young Shin, and Yong-Min Kim\*

**AFFILIATION :** *Plant Systems Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon, Republic of Korea*

**TITLE :** Identification of accession-specific variants and development of KASP markers for assessing the genetic makeup of *Brassica rapa* seeds

The Korean Society for Molecular and Cellular Biology hereby certificate that above individual had presented in the International Conference of Korean Society for Molecular and Cellular Biology which was held on Sep. 28 - Sep. 30, 2022, ICC JEJU, Jeju, Korea.

October 4, 2022

Goo Taeg Oh, Ph.D.  
2022 President  
Korean Society for Molecular and Cellular Biology



# The 19<sup>th</sup> KOGO Winter Symposium

The Genomic Language of Life

2023.02.01(수)-02.03(금)

📍 강원도 홍천 비발디파크

주 최: **KOGO**  
KOREA GENOME ORGANIZATION

공동주관:  국립보건연구원  
National Institute of Health

**kobic** 국가생명연구자원정보센터  
Korea Bioinformation Center

## ▶▶▶ F-02

**Pseudo-chromosomal genome assembly of warm-season turfgrass (*Zoysia sinica*) surviving in wetlands**

Hyeonseon Park, \*Donghwan Shim

*College of Bioscience & Biotechnology, Chungnam National University*

The genus *Zoysia* is warm-season turfgrass and has been commonly used in golf courses, gardens, and parks. *Zoysia* species tolerate various abiotic stresses, including salt, and drought. Therefore, *Zoysia* species are economically and biologically important resources. Especially, *Z. sinica* is a salt marsh species, which grows on the seashore, estuary marsh, and even in the intertidal zone. Waterlogging stress induces inhibition of root respiration and accumulation of phytotoxic compounds in roots. In some species in the Poaceae family, constitutive or inducible aerenchyma and barriers to radial oxygen loss (ROL) contribute to waterlogging tolerance. However, little is known about waterlogging responses in *Zoysia* species, possibly due to a lack of genome information. In this study, we present a pseudo-chromosomal genome of allotetraploidy grass, *Z. sinica* using the long-read Oxford Nanopore sequencing and Illumina sequencing platform. The genome had 20 pseudo-molecules with a total length of 313.7 Mb and a high contiguous with the scaffolds N50 value of 18.9Mb. The *Z. sinica* genome contained 30,476 protein-coding genes and 41.64 % of repeat sequences. Comparing orthologous genes involved in waterlogging responses among five Poaceae species, *Arabidopsis thaliana*, and *Z. sinica*, revealed that orthologs of *Z. sinica* had copy number variants (CNVs) and protein domain architectures related to promoting ROS accumulation and formations of aerenchyma in roots. Among *Z. sinica*-specific genes, 13 orthologs of class III peroxidases (CIII Prxs) associated with cell wall remodeling and responses under flooding conditions were identified. Our genome will provide genetic resources for the molecular breeding of new cultivars, particularly those with abiotic tolerances under flood risks. Although waterlogging mechanisms of zoysiagrasses are largely unknown, our genomic information could provide genomic evidence of *Z. sinica* survival in waterlogged environments.

Keywords: Genomics, Genome assembly, Zoysiagrass, Allotetraploidy plant

## ▶▶▶ F-03

**The comparative analyses of the *Brassica rapa* revealed diversification pattern of Cold-related gene families and selective pressure during the breeding history**Seongmin Hong<sup>1</sup>, Sangjin Go<sup>2</sup>, Ah-Young Shin<sup>1</sup>, Hyunjin Koo<sup>1</sup>, Minah Jung<sup>3</sup>, Ye Na Kim<sup>4</sup>, \*Yong-Min Kim<sup>1</sup>

<sup>1</sup>Plant Systems Engineering Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), <sup>2</sup>Department of Bioinformatics, KRIBB School of Bioscience, Korea Research Institute of Bioscience and Biotechnology (KRIBB), <sup>3</sup>College of Agriculture and Life Sciences, Chungnam National University, Institute of Agricultural Science, Chungnam National University, <sup>4</sup>Department of Horticulture, College of Agriculture and Life Sciences, Gyeongsang National University

The continuous cold stress triggers physiological changes inducing irreversible transition or cell death. As it acts as a critical event in whole plant life cycle, plants have evolved elaborate pathways to survive under cold temperature by cold acclimation. By gaining cold tolerance, plants minimize physiological damage induced by low temperature and prepare ordinary transition from vegetative stage to reproduction. Brassicaceae, one of important plant family includes important domesticated crops and landraces adapted to extremely harsh climate changes during the evolutionary history. Here, we identified total 69,113 CRG analogues by two-way gene search process based on sequence similarity and domain architecture from the known *A.thaliana* 114 CRGs. The comparative analyses showed that the majority of CRGs in the *B. rapa* were generated by whole genome duplication and the changes of selective pressure on the CRGs in the Chinese cabbage (*B. rapa* spp. *pekinensis*) core collection during the commercial breeding process by the human.

Keywords: Comparative analysis, Cold-related genes, Whole genome duplication, *Brassica rapa*, Core collection

The 19<sup>th</sup> KOGO Winter Symposium

## ▶▶ F-04

**Construction of High-contiguity reference genome sequences of two *Hibiscus syriacus* accessions (Gangneung and Baekdansim)**Sangjin Go<sup>1</sup>, Ah-Young Shin<sup>2</sup>, Seongmin Hong<sup>2</sup>, Hyunjin Koo<sup>2</sup>, Minah Jung<sup>3</sup>, Ye Na Kim<sup>4</sup>, \*Yong-Min Kim<sup>2</sup>

<sup>1</sup>Department of Bioinformatics, KRIBB School of Bioscience, Korea University of Science and Technology (UST), <sup>2</sup>Plant Systems Engineering Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), <sup>3</sup>Department of Bio Environmental Chemistry, College of Agriculture and Life Sciences, Chungnam National University, <sup>4</sup>Department of Horticulture, College of Agriculture and Life Sciences, Gyeongsang National University

*Hibiscus syriacus* called Rose of Sharon is one of the most widespread garden shrubs in the world. *H. syriacus* contains a complexed genome structure generated by multiple rounds of polyploidization during evolution. Two high quality reference genomes of *H. syriacus* varieties (Gangneung and Baekdansim) were constructed using long read sequencing and Hi-C scaffolding. In Gangneung and Baekdansim genome assemblies, a total size of 1.87 Gbp, 1.82 Gbp and 88,741, 87,438 gene models were predicted, respectively. In order to identify genetic varieties, associated to horticultural traits of the *H. syriacus*, we re-sequenced 32 accessions showing different phenotypes related to floral architecture and cold-resistance. As a result, we discovered 10,424 SNPs based on the reference genome of Gangneung. The phylogenetic trees constructed by four-fold degenerate sites (4D SNPs) clustered accessions corresponding crossing information and their known origins. Total 1,051 number of SNPs were located in 268 number of cold-related genes(CRGs) based on sequence similarity against 114 *A. thaliana* CRGs, resulting in 19 cold-resistant accession/group specific SNPs. The reference genome will serve as a platform for developing cold-tolerance varieties and floral traits in the *H. syriacus* breeding program.

Keywords: Long read sequencing, Genome assembly, *Hibiscus syriacus*, Population study

## ▶▶ F-05

**SoyPedi: pedigree database for breeding design based on Korean ideal elite soybean cultivars**Min-Gyun Jeong<sup>1</sup>, Chaeyoung Lee<sup>2</sup>, Jin Hyun Kim<sup>3</sup>, Yoram Choi<sup>1</sup>, Hye Jeong Kim<sup>4</sup>, Young-Soo Chung<sup>4</sup>, \*Hong-Kyu Choi<sup>4</sup>

<sup>1</sup>Department of Applied Bioscience, Dong-A University, <sup>2</sup>Department of Medical Bioscience, Dongnam Institute of Radiological & Medical Sciences, <sup>3</sup>Department of Agricultural Biotechnology, National Institute of Agricultural Sciences, <sup>4</sup>Department of Molecular Genetics, Dong-A University

For breeding design, parental information of any domesticated species is an important source because it's an indicator of the genetic relationship among bred lines. As the accumulation of genome and phenome data on cultivars rapidly increases, well-structured pedigree and system to utilize it have played a key role. For this reason, we intended to construct a pedigree-based web interface to link the phenotypic information and genealogical relationship of soybean cultivars. A well-structured pedigree of 178 Korean soybean cultivars and its phenotype information were used as key data for database construction. Three submodules consisting of pedigree-, tabular- and single cultivar-based search pages constitute the web interface. Single Page Application (SPA) makes users to quickly switch between pedigree and tabular list mode still on a same filter status. Furthermore, comparison of the parental information and a SNP-based phylogeny from cultivars in its pedigree presented that there is a certain degree of correlation between them. Our web page is developed as an intuitive auxiliary system for breeding design in the future and plans to expand itself to a global database by integrating international soybean data. This web interface is available at <https://tgil.donga.ac.kr/SoyPedi/index.html>.

Keywords: Soybean, Pedigree, Bioinformatics, Webpage

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김봉민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처				
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
A TIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7855			
	KORD 번호	NN-7855-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	빨간 배추 생육 10일차 반복 1			
	용량	4.03 GB 7.68 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 03일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31IID020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	신입연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처				
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7863			
	KORD 번호	NN-7863-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	빨간 배추 생육 10일차 반복 2			
	용량	3.88 GB 7.42 Gbp			
<p>"생명연구자원의 확보 관리 및 활용에 관한 법률" 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 03일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p>					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등 록 자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
ATIS 연계정보	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
등 록 내용	생명정보 내분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
	등록필증번호	NN-7865			
	KORD 번호	NN-7865-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	빨간 배추 생육 10일차 반복 3			
	용량	4.07 GB 7.79 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 03일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국분	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 내분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7869			
	KORD 번호	NN-7869-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용 용량	빨간 배추 생육 20일차 반복 1 3.66 GB 6.99 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 03일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7870			
	KORD 번호	NN-7870-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	빨간 배추 생육 20일차 반복 2			
	용량	3.46 GB 6.58 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 03일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7871			
	KORD 번호	NN-7871-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	빨간 배추 생육 20일차 반복 3			
	용량	3.93 GB 7.47 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 03일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa subsp. pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7878			
	KORD 번호	NN-7878-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	빨간 배추 생육 30일차 반복 1			
	용량	4.0 GB 7.63 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 06일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7879			
	KORD 번호	NN-7879-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	빨간 배추 생육 30일차 반복 2			
	용량	3.9 GB 7.37 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 06일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	구체과제	사업단	구체과제	과제번호	8210260 31HD020
주관연구과제명	구분	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영분	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	인입연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등 록 자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등 록 내용	등록필증번호	NN-7880			
	KORD 번호	NN-7880-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	빨간 배추 생육 30일차 반복 3			
	용량	3.94 GB 7.43 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 06일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등 록 자 인격사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등 록 내용	등록필증번호	NN-7881			
	KORD 번호	NN-7881-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	빨간 배추 생육 40일차 반복 1			
	용량	3.5 GB 6.63 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 06일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업다	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	식 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa subsp. pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7876			
	KORD 번호	NN-7876-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	빨간 배추 생육 40일차 반복 2			
	용량	4.31 GB 8.21 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 06일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7882			
	KORD 번호	NN-7882-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	빨간 배추 생육 40일차 반복 3			
	용량	3.77 GB 7.08 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 06일</p> <p style="text-align: center;">국립농업과학원장 (지인) </p>					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7883			
	KORD 번호	NN-7883-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	빨간 배추 생육 50일차 반복 1			
	용량	3.34 GB 6.29 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 06일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	지 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인식사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	지 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/업종 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa subsp. pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7884			
	KORD 번호	NN-7884-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	빨간 배추 생육 50일차 반복 2			
	용량	3.63 GB 6.88 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 06일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	구분	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7877			
	KORD 번호	NN-7877-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	빨간 배추 생육 50일차 반복 3			
	용량	4.49 GB 8.57 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 06일</p> <p style="text-align: center;">국립농업과학원장 (지인) </p>					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7861			
	KORD 번호	NN-7861-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	일반 배추 생육 10일차 반복 1			
	용량	4.77 GB 9.4 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 03일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	구분	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 내분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7856			
	KORD 번호	NN-7856-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	일반 배추 생육 10일차 반복 2			
	용량	4.94 GB 9.66 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 03일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7854			
	KORD 번호	NN-7854-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	일반 배추 생육 10일차 반복 3			
	용량	5.04 GB 9.82 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 03일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7857			
	KORD 번호	NN-7857-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	일반 배추 생육 20일차 반복 1			
	용량	4.78 GB 9.33 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 03일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	구책과제	사업단	구책과제	과제번호	8210260 31HD020
주관연구과제명	구분	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	신입연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7858			
	KORD 번호	NN-7858-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	일반 배추 생육 20일차 반복 2			
	용량	4.82 GB 9.45 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 03일</p> <p style="text-align: center;">국립농업과학원장 (직인) </p>					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	지 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7859			
	KORD 번호	NN-7859-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	일반 배추 생육 20일차 반복 3			
	용량	5.44 GB 10.66 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 03일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa subsp. pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7862			
	KORD 번호	NN-7862-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	일반 배추 생육 30일차 반복 1			
	용량	4.34 GB 8.46 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 03일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7864			
	KORD 번호	NN-7864-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	일반 배추 생육 30일차 반복 2			
	용량	4.3 GB 8.42 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 03일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa subsp. pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7866			
	KORD 번호	NN-7866-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	일반 배추 생육 30일차 반복 3			
	용량	4.78 GB 9.41 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 03일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	지 위	선임연구원	
	성 명	(한글) 김홍민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/영종 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7867			
	KORD 번호	NN-7867-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	일반 배추 생육 40일차 반복 1			
	용량	3.53 GB 6.88 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 03일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

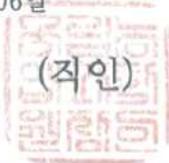
### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
ATIS 연계정보	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	생명정보명	<i>Brassica rapa subsp. pekinensis</i>			
등록 내용	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
	등록필증번호	NN-7868			
	KORD 번호	NN-7868-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용 용량	일반 배추 생육 40일차 반복 2 3.52 GB 6.86 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 03일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7872			
	KORD 번호	NN-7872-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	일반 배추 생육 40일차 반복 3			
	용량	3.78 GB 7.36 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 06일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업년	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7873			
	KORD 번호	NN-7873-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	일반 배추 생육 50일차 반복 1			
	용량	3.43 GB 6.67 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 06일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	선임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7874			
	KORD 번호	NN-7874-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	일반 배추 생육 50일차 반복 2			
	용량	3.47 GB 6.81 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 06일</p> <p style="text-align: center;">국립농업과학원장 (직인)</p> 					

### 생명정보 등록 확인서

사업구분	국책과제	사업단	국책과제	과제번호	8210260 31HD020
주관연구과제명	국문	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
	영문	Development of Artificial Intelligence-based Genome Analysis Programs and Digital Breeding Platform			
주관연구책임자	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	신임연구원	
	성 명	(한글) 김용민	(영문) Yong min Kim		
등록자 인적사항	소속 및 부서명	한국생명공학연구원 유전자교정연구센터	직 위	연구원	
	성 명	(한글) 홍성민	(영문) Seong Min Hong		
	연락처	[Redacted]			
	세부/협동 연구과제명	(세부) 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발			
ATIS 연계정보	생명정보명	<i>Brassica rapa</i> subsp. <i>pekinensis</i>			
	생명정보 대분류명	Transcriptome	생명정보 소분류명	RNA-Seq	
등록 내용	등록필증번호	NN-7875			
	KORD 번호	NN-7875-000001 (1 건)			
	생명정보형태	NGS SRA			
	등록내용	일반 배추 생육 50일차 반복 3			
	용량	3.71 GB 7.26 Gbp			
<p>“생명연구자원의 확보 관리 및 활용에 관한 법률” 제10조2항(개정 2018.12.24.)에 의거하여, 위와 같이 생명정보가 농업생명공학정보센터(NABIC)에 등록되었음을 확인합니다.</p> <p style="text-align: center;">2021년 12월 06일</p> <p style="text-align: center;">국립농업과학원장 (직인) </p>					

특허 실적 증빙 1

관인생략

## 출원번호통지서

출원일자 2021.12.22  
 특기사항 심사청구(유) 공개신청(무) 참조번호(KPA211390)  
 출원번호 10-2021-0185453 (접수번호 1-1-2021-1489319-80)  
 (DAS접근코드46BB)  
 출원인명칭 한국생명공학연구원(3-1999-034166-5) 외 1명  
 대리인성명 특허법인 한열(9-2020-100061-3)  
 발명자성명 김용민 홍성민 신아영 임용표 최수련  
 발명의명칭 배추의 종자순도 검정을 위한 KASP 마커 세트와 효율적인 마커 개발을 위한 방법

# 특 허 청 장

<< 안내 >>

1. 귀하의 출원은 위와 같이 정상적으로 접수되었으며, 이후의 심사 진행상황은 출원번호를 이용하여 특허로 홈페이지([www.patent.go.kr](http://www.patent.go.kr))에서 확인하실 수 있습니다.  
 2. 출원에 따른 수수료는 접수일로부터 다음날까지 동봉된 납입영수증에 성명, 납부자번호 등을 기재하여 가까운 은행 또는 우체국에 납부하여야 합니다.  
 ※ 납부자번호 : 0131(기관코드) + 접수번호  
 3. 귀하의 주소, 연락처 등의 변경사항이 있을 경우, 즉시 [특허고객번호 정보변경(경정), 정정신고서]를 제출하여야 출원 이후의 각종 통지서를 정상적으로 받을 수 있습니다.  
 4. 기타 심사 절차(제도)에 관한 사항은 특허청 홈페이지를 참고하시거나 특허고객상담센터(☎ 1544-8080)에 문의하여 주시기 바랍니다.  
 ※ 심사제도 안내 : <http://www.kipo.go.kr>-지식재산제도

특허 실적 증빙 1

2021-12-22

**【주소】** 대전광역시 유성구 대학로 99 충남대학교 농업생명과학대학  
원예학과

**【발명자】**

**【성명】** 최수련

**【성명의 영문표기】** Su Ryun Choi

**【주민등록번호】** [REDACTED]

**【우편번호】** 34134

**【주소】** 대전광역시 유성구 대학로 99 충남대학교 농업생명과학대학  
원예학과

**【출원언어】** 국어

**【심사청구】** 청구

**【핵산염기 서열목록 또는 아미노산 서열목록】**

**【서열개수】** 400

**【서열목록 전자파일】** 미첨부

**【이 발명을 지원한 국가연구개발사업】**

**【과제고유번호】** 1545023072

**【과제번호】** 821026031SB010

**【부처명】** 농림축산식품부

**【과제관리(전문)기관명】** 농림식품기술기획평가원

**【연구사업명】** 기술사업화지원(R&D)

**【연구과제명】** 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼  
개발

특허 실적 증빙 2

관인생략

## 출원번호통지서

출원일자 2023.09.01  
특기사항 심사청구(유) 공개신청(무) 참조번호(KPA221585)  
출원번호 10-2023-0116324 (접수번호 1-1-2023-0969470-37)  
(DAS접근코드FE1A)  
출원인명칭 한국생명공학연구원(3-1999-034166-5)  
대리인성명 특허법인 한얼(9-2020-100061-3)  
발명자성명 김용민 신아영 홍성민  
발명의명칭 배추의 계통군 검정을 위한 마커 및 이의 이용

## 특 허 청 장

<< 안내 >>

1. 귀하의 출원은 위와 같이 정상적으로 접수되었으며, 이후의 심사 진행상황은 출원번호를 이용하여 특허로 홈페이지([www.patent.go.kr](http://www.patent.go.kr))에서 확인하실 수 있습니다.  
2. 출원에 따른 수수료는 접수일로부터 다음날까지 동봉된 납입영수증에 성명, 납부자번호 등을 기재하여 가까운 은행 또는 우체국에 납부하여야 합니다.  
※ 납부자번호 : 0131(기관코드) + 접수번호  
3. 귀하의 주소, 연락처 등의 변경사항이 있을 경우, 즉시 [특허고객번호 정보변경(경정), 정정신고서]를 제출하여야 출원 이후의 각종 통지서를 정상적으로 받을 수 있습니다.  
4. 기타 심사 절차(제도)에 관한 사항은 특허청 홈페이지를 참고하시거나 특허고객상담센터(☎ 1544-8080)에 문의하여 주시기 바랍니다.  
※ 심사제도 안내 : <https://www.kipo.go.kr>-지식재산제도

2023-09-01

**【연구과제명】** 인공지능기술과 유전체 분석의 융합연구를 통한 무궁화 디지털 육종 플랫폼 개발

**【기여율】** 1/2

**【과제수행기관명】** 한국생명공학연구원

**【연구기간】** 2023.03.01 ~ 2024.02.29

**【이 발명을 지원한 국가연구개발사업】**

**【과제고유번호】** 1545026757

**【과제번호】** 821026033SB010

**【부처명】** 농림축산식품부

**【과제관리(전문)기관명】** 농림식품기술기획평가원

**【연구사업명】** 기술사업화지원

**【연구과제명】** 인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발

**【기여율】** 1/2

**【과제수행기관명】** 한국생명공학연구원

**【연구기간】** 2023.01.01 ~ 2023.12.31

**【취지】** 위와 같이 특허청장에게 제출합니다.

대리인 특허법인 한얼 (서명 또는 인)

**【수수료】**

**【출원료】** 0 면 46,000 원

**【가산출원료】** 68 면 0 원



## 기술이전계약서

(주)유클리드소프트(이하 “실시권자”라 한다)와 한국생명공학연구원(이하 “연구원”이라 한다)은 다음과 같이 기술이전계약을 체결한다.

### 제1조 (정의)

- ① 본 계약상 “기술”이란 2항의 저작권과 관련된 “전장유전체 데이터를 활용한 프라이머 제작 파이프라인 활용 기술”로서, 이와 관련하여 “연구원”이 보유하고 있는 연구·개발 결과 관련 발명, 노하우, 기술적 데이터 등 “기술”을 실시하기 위하여 필요한 모든 것을 말한다.
- ② 본 계약상의 “저작권이란 다음과 같다.

저작물 제호	등록번호	등록일
전장유전체 데이터를 활용한 프라이머 제작 파이프라인	C-2020-024215	2020.07.23
저작물의 종류	저작자	
컴퓨터프로그램저작물	한국생명공학연구원	



- ③ 본 계약상 “실시”란 특허법 제2조 제3호에서 정하는 실시를 말한다.
- ④ 본 계약상 “실시제품”이라 함은 “기술”의 전부 또는 일부를 사용하여 생산되는 제품 및 생산장치, 중간체나 원료, 공정 등 산출물 일체를 말한다.

### 제2조 (실시권의 내용)

- ① “연구원”은 “실시권자”에 대하여 본 계약에서 정하는 바에 따라 아래와 같은 내용으로 “기술”을 실시할 권한을 허락한다.

실시권의 유형	통상 실시권
실시 기간	계약체결일로부터 2032년 12월31일까지

- ② “실시권자”는 “연구원”의 사전 서면 동의 없이는 제3자로 하여금 “기술”을 실시하게 할 수 없다.

### 제3조 (실시 의무)

- ① “실시권자”는 본 계약체결일로부터 3년(이하, “실시개시기한”이라 한다) 이내에 “기술”을 실시하여야 한다.
- ② “실시권자”가 제1항에서 정하는 “실시개시기한” 내에 실시할 수 없는 경우, 그 사





유와 향후 실시 계획을 기재한 서면을 “연구원”에 제출해야 한다.

- ③ “연구원”은 제3항에서 정하는 바에 따라 “실시권자”가 제출한 서면을 검토하여 필요하다고 판단되면, “실시권자”와 별도 서면 합의를 통해 제1항의 “실시개시기한”을 연장할 수 있다.

**제4조 (기술료)**

- ① “실시권자”는 “연구원”으로부터 “기술”의 실시권을 받는 대가(이하, “기술료”라 한다)로 “연구원”에 다음과 같이 “기술료”(부가세 별도, 이하 같다)를 지급하기로 한다.

구분	지급기한	금액(부가세 별도)
정액기술료	계약체결일로부터 30일 이내	₩ 30,000,000원

- ② “실시권자”가 제1항에서 정하는 “기술료”를 납부 기한내에 “연구원”에 지급하지 않을 경우, “실시권자”는 미납 기술료에 대하여 연체이자(미납기술료(원) × (연체일수/365) × 5%)를 “연구원”에 지급하여야 한다.
- ③ “실시권자”가 “연구원”에 납부한 기술료는 어떤 경우에도 반환하지 아니한다.

**제5조 (개량기술)**

- ① “실시권자” 또는 “실시권자”의 임원 및 피용자가 “기술”의 성능이나 효과를 향상하기 위하여 “기술”을 개선, 확장, 대체 또는 추가발명 등의 방법으로 “기술”을 개량한 경우(이하 “개량기술”이라 한다), “실시권자”는 지체 없이 “연구원”에 그 사실을 서면으로 통보하여야 한다.
- ② “실시권자”가 “개량기술”에 관하여 특허, 저작권 등 지식재산권을 출원·등록하고자 하는 경우 반드시 사전에 “연구원”의 서면 동의를 받아야 한다.
- ③ “개량기술”에 관한 특허, 저작권 등 지식재산권에 관해서는 “실시권자”와 “연구원”이 공동으로 소유하기로 한다.
- ④ “개량기술”에 관한 특허, 저작권 등 지식재산권의 확보·유지에 필요한 제반 비용은 “실시권자”가 부담한다.

**제6조(비밀유지)**

- ① “실시권자”는 “연구원”이 제공한 정보를 “특허”의 실시를 위해서만 사용할 수 있으며, “연구원”의 사전 동의 없이 그 정보를 제3자에게 제공하거나 누설해서는 안 된다.
- ② “실시권자”는 그 임원 및 피용자 또는 그 승계인이 본조에서 정한 의무를 충실히 이행하도록 필요한 조치를 취하여야 한다.





- ③ 본 조에서 정한 의무는 본 계약이 만료된 이후나 해제 또는 해지 등으로 종료되었을 경우에도 계속 유효하다.

**제7조 (보증면제)**

“연구원”은 “기술”의 유효성(등록가능성을 포함한다)이나 사업성, 상업성에 대해 어떠한 보증도 하지 아니하며, “기술”의 실시가 제3자의 권리를 침해하지 않는다거나 “기술”의 실시로 인하여 “실시권자”에게 발생한 어떠한 손해나 손실(제3자에 대한 실시료 지불을 포함)에 대해서도 책임을 지지 아니한다.

**제8조 (계약의 해지)**

- ① “연구원”은 다음 각 호의 경우에 30일의 기한을 두고 “실시권자”에게 그 이행을 서면으로 최고하고, “실시권자”가 최고를 받은 날로부터 30일이 경과할 때까지 이를 이행하지 아니하면 본 계약을 해지할 수 있다.
  - 1. “실시권자”가 제4조에서 정한 기술료의 전부 또는 일부를 납부하지 않는 경우
  - 2. “실시권자”가 기타 본 계약에서 정한 의무를 이행하지 아니하는 경우
- ② “연구원”은 다음 각 호의 경우 별도의 서면 통보 없이 본 계약을 즉시 해지할 수 있다.
  - 1. “실시권자”가 부도, 파산, 청산, 해산된 경우
  - 2. “실시권자”가 본 계약에서 정한 의무 이행을 명시적으로 거절한 경우
  - 3. “연구원”이 우편이나 이메일 발송, 전화통화 등 통상적인 방법에 의하여 “실시권자”에게 연락을 취했음에도 불구하고, 연속하여 3개월 이상 연락이 두절되거나 “실시권자”의 소재가 불분명한 경우
- ③ 본 계약이 해지되는 경우 “실시권자”는 다음 각 호에서 정하는 사항을 즉시 이행하여야 한다.
  - 1. “기술”의 실시 중단 (제3자가 “특허”를 실시하고 있을 경우, 제3자로 하여금 “기술”의 실시를 중지하게 할 의무를 포함한다)
  - 2. “기술” 관련 자료의 반환 또는 폐기
  - 3. 계약 해지일까지 미납된 기술료 및 연체이자의 지급

**제9조 (명칭 사용)**

- ① “실시권자”는 광고, 판매촉진, 기타 선전이나 쟁송의 목적으로 “연구원”의 명칭을 암시하거나 사용해서는 아니 된다.
- ② “실시권자”가 “연구원”의 명칭을 사용할 필요가 있는 경우 “연구원”으로부터 각 사전 서면 동의를 받아야 한다.





**제10조 (변경 사항)**

- ① “실시권자”는 본 계약 체결 이후 다음 각 호에서 정하는 사실이 변경되었거나, 변경될 예정인 경우 지체없이 “연구원”에 서면으로 통보하여야 한다.
  - 1. 상호 또는 법인의 명칭
  - 2. 대표자 또는 기술이전 담당자
  - 3. 대표 주소 및 연락처
  - 4. 기타 본 계약에 관한 주요 사항
- ② “실시권자”가 전항에서 정한 의무를 위반하여 발생한 “연구원”의 착오는 “실시권자”의 모든 항변으로부터 면책된다.

**제11조 (계약의 효력 등)**

- ① 본 계약은 “실시권자”와 “연구원”이 모두 서명 또는 날인한 날로부터 효력을 발생하며, 본 계약 체결 이전에 있었던 의견교환, 구두 또는 문서상의 합의에 우선한다.
- ② 본 계약은 그 일부가 무효, 부적법 또는 집행 불능인 경우 및 “특허”가 거절, 무효, 취소되는 경우에도, 본 계약의 나머지 규정은 완전한 효력을 가지고 존속한다.
- ③ 본 계약의 내용은 “실시권자”와 “연구원”의 별도 서면 합의에 의해서만 유효하게 변경할 수 있다.

**제12조 (분쟁의 해결)**

본 계약의 해석이나 이행과 관련하여 당사자간에 분쟁이 발생하는 경우, “연구원”의 소재지를 관할하는 법원을 제1심 법원으로 하여 해결하기로 한다.

본 계약서는 2통을 작성하여 당사자가 기명날인한 후 각자 1통씩 보관하기로 한다.

2022. 07. 18.

“실시권자”

대전 서구 대덕대로317번길 20, 458호(월  
평동)

(주)유클리드소프트

대표이사 채 은 경



“연구원”

대전광역시 유성구 과학로 125

한국생명공학연구원

원 장 김 장 성





## 비밀준수약정서

한국생명공학연구원(이하 "갑"이라 한다)과 주식회사 유클리드소프트(이하 "을"이라 한다)(이하 "갑"과 "을"을 각각 "당사자", 통칭하여 "당사자들"이라 한다)은 "갑"의 "전장유전체 데이터를 활용한 프라이머 제작 파이프라인 (표 1의 프로그램)"에 대한 검토(이하 "본 검토"라 한다)와 관련하여 다음과 같이 비밀을 유지할 것을 동의하여 상호 약정(이하 "본 계약"이라 한다) 한다.

표 1. 특허 리스트

연번	등록번호	저작물 제호	주창작자
1	C-2020-024215	전장유전체 데이터를 활용한 프라이머 제작 파이프라인	김용민 책임연구원

### 제 1 조 (목적)

"본 계약"은 "본 검토"를 수행하는 데 있어서 "갑"이 "을"에게 제공하거나 "을"이 "갑"에게 제공하는 모든 "기술정보", "물질", "취득정보"에 대한 비밀을 준수하고 관련한 당사자들의 권리와 의무를 정함을 목적으로 한다.

### 제 2 조 (용어 정의)

- ① "본 계약"에서 사용하는 용어의 뜻은 다음 각 호와 같다.
- "기술정보" : "본 검토"를 위하여 "갑"이 "을"에게 제공하거나 "을"이 "갑"에게 제공하는 서면 기술자료 뿐만 아니라 그와 관련된 일체의 지식 및 정보 등을 총칭한다.
  - "물질" : "본 검토"를 위하여 "갑"과 "을" 상호간에 제공하는 물질 뿐만 아니라, 이러한 물질 또는 "기술정보"를 이용하여 생산한 모든 제품을 의미한다.
  - "취득정보" : "갑" 또는 "을"이 "기술정보" 및 "물질"을 활용하여 "본 검토"를 통해 취득한 정보를 의미한다. 단, "갑" 또는 "을"이 자사의 제품개발을 위해 기 확보해 놓은 기술 또는 정보임을 증명할 수 있는 것은 "취득정보"에 포함되지 아니한다.
  - "종사원 및 관계인" : 해당 당사자의 임원, 사원, 대표자 및 대리인을 의미한다.





- ② "본 계약"에서 정보제공자는 "기술정보", "물질", "취득정보"를 제공하는 자를 의미한다. 본 계약에서 정보수령자는 정보제공자로부터 "기술정보", "물질", "취득정보"를 수령하는 자를 말한다.
- ③ 정보제공자가 서면이나 서류 형태로 "기술정보", "물질", "취득정보"를 제공하는 경우 서류상에 비밀임을 알리는 문구("비밀", "비밀정보", "기밀정보")를 명백히 표시해야 한다. 정보제공자가 서면이나 서류 이외의 형태 즉 구두나 영상 또는 당사자의 시설, 장비 샘플 기타 품목들을 관찰하거나 조사함으로써 제공할 경우에는 공개시 정보수령자에게 본 정보가 비밀에 속한다는 것을 알리고 또한 공개일로부터 [10]일 이내에 서면으로 비밀임을 알리는 문구를 표시하여 정보수령자에게 확인시켜야 한다. 비밀임을 알리는 표시(고지)하지 않은 정보는 본 계약상의 비밀을 유지해야 하는 "기술정보", "물질", "취득정보"에 해당하는 것으로 보지 아니한다.

**제 3 조 (약정 내용)**

- ① "갑"과 "을"은 협의하여 "본 검토" 관련 연구 방향을 설정하고, 연구 설계 및 수행을 하며, 연구 결과에 대하여 상호 협의 공유한다.
- ② "갑"과 "을"은 "기술정보", "물질", "취득정보"를 상대방의 사전 서면동의 없이 본 계약 목적 이외 타 목적으로 이용하거나 공개하지 않는다.
- ③ "갑" 또는 "을"은 "본 검토"를 통해 취득하거나 상대방이 제공한 "기술정보", "물질", "취득정보"를 이용하여 특허 등 지식재산권을 출원·등록하고자 하는 경우 상대방에게 서면 동의를 받아 별도의 공동출원협약을 체결하여야 하고, 제품을 실시하고자 할 경우 상대방과 별도의 기술실시계약을 체결하여야 한다. 여기서 실시는 특허법 제2조제3호에 따른 실시를 의미한다.
- ④ "갑"과 "을"이 "본 검토"를 통해 취득하거나 상대방이 제공한 "기술정보", "물질", "취득정보"에 근거하여 국내외에서 사업화하고자 하는 경우, 당사자들은 상호 성실과 신의를 바탕으로 사업 추진 조건 등에 대한 합의에 기초한 사업계획서를 작성하여 생산 및 판매를 하여야 한다.

**제 4 조 (검토기간 등)**





- ① "갑"과 "을"은 "본 검토"를 계약체결일로부터 [1]년 이내(검토기간)에 종결하기로 한다. 다만, 당사자들의 합의에 의하여 검토기간을 연장할 수 있다.
- ③ "갑" 또는 "을"은 "본 검토" 종료 후 상대방의 요청에 따라 제공한 "기술정보"와 "물질"을 "갑" 또는 "을"에게 반환 및 반환할 수 없는 경우 폐기해야 한다.

**제 5 조 (책임의 범위)**

- ① "갑"과 "을"은 소속 "종사원 및 관계인"의 직접 또는 간접행위에 의해 제3조(약정 내용)를 위반한 비밀누설 행위에 대해 사용자로서의 책임을 져야 한다.
- ② 본 조 제1항에서 정한 "종사원 및 관계인"이 "본 검토"가 종료되기 전에 퇴직 또는 전보 등으로 비밀에 접근할 수 있는 범위에 들지 않게 된 경우에도 본 계약의 비밀유지 의무기간 동안 이들에 의하여 직접 또는 간접적으로 본 계약상의 제반 비밀사항이 누설되는 때에는 각 당사자는 여전히 사용자로서의 본 조 제1항의 책임을 부담한다.
- ③ "갑" 또는 "을"이 연구결과를 실험하거나 성능검사를 하기 위해 "종사원 및 관계인"의 범위에 들지 아니하는 자에게 의뢰하거나 장비를 임대하여야 할 경우 상대방의 사전 서면동의를 얻어야 한다.

**제 6 조 (면책)**

"갑" 또는 "을"은 "본 검토"와 관련하여 상대방에게 제공하는 "기술정보" 또는 "물질"과 관련하여 발생하는 여하한 손실 또는 제3자에게 발생한 손해 등 어떠한 사항에도 일체의 책임을 지지 아니한다.

**제 7 조 (과실의 추정)**

"갑"과 "을"은 제3조(약정 내용)에서 정한 계약내용을 위반하였거나 위반할 우려가 있다고 어느 일방이 합리적인 이유 하에 추정하는 경우 "갑"과 "을"은 공동으로 이를 조사하고 이 경우 "갑"과 "을"은 상호에게 적극 협조해야 한다.

**제 8 조 (손해배상)**

- ① "갑" 및 "을"이 본 계약을 위반하였을 때에는 본 계약의 위반으로 인한 손해에 대하여 상대방에게 배상책임이 있다.





- ② 각 당사자는 천재지변, 폭동, 전쟁 등과 같은 불가항력으로 인한 계약상 의무의 불이행이나 이행 지연에 대하여는 책임을 지지 아니하며, 불가항력적인 요인이 제거된 후에는 지속적으로 의무를 이행하여야 한다.

**제 9 조 (약정위반에 해당되지 않는 사항)**

“갑” 또는 “을”의 다음 각 호의 1에 해당하는 행위 또는 사항은 약정위반으로 간주하지 않는다.

1. 누설된 상대방의 “기술정보”, “물질”, “취득정보”의 내용이 누설된 시점에서 공개된 타인의 특허내용과 동일한 경우
2. 누설된 상대방의 “기술정보”, “물질”, “취득정보”의 내용이 누설된 시점에서 이미 공지의 사항이거나 제3자에 의해 간행물에 기재된 내용인 경우
3. 당사자 일방이 상대방의 “기술정보”, “물질”, “취득정보”에 의존하지 않고 독자적으로 취득한 정보임을 증명할 수 있는 경우
4. 법령에 의해 개시가 의무인 정보
5. 그 외 공지의 사실

**제 10 조 (계약기간)**

- ① “본 계약”은 “당사자들”이 서명(또는 기명 날인) 한 날로부터 효력을 발생한다. “본 계약”의 계약기간(본 계약에 있어 “계약기간”이라 한다)은 계약체결일로부터 [5]년간으로 한다.
- ② “본 계약”의 제3조(약정내용)에 규정된 비밀준수의무는 본 계약이 종료되거나 중도에 해지 또는 해제에도 불구하고 “본 계약” 체결일로부터 [5]년간(본 계약에 있어 “비밀유지의무기간”이라 한다) 유효하다.

**제 11 조 (신의성실 및 해석)**

- ① 본 계약서에 명기되지 아니한 사항 및 해석상 이의가 있을 때에는 상호 신의 성실의 원칙에 의한 합의에 따라 결정한다.
- ② 본 조 제1항의 규정에도 불구하고 본 계약과 관련된 분쟁은 서울중앙지방법원을 제1심 전속적 관할법원으로 한다.





**제 12 조 (계약해지)**

- ① “당사자들” 중 일방이 본 계약의 조건을 위반하여 일방이 상대방에게 [30]일간의 기간을 정하여 이의 시정을 최고하였음에도 불구하고 동 기간 내에 이를 시정하지 않을 경우 시정을 최고한 일방은 즉시 본 계약을 해지할 수 있다.
- ② 다음 각 호에 해당하는 경우 계약기간 중이라도 본 계약을 즉시 해지할 수 있다.
  - 1. 강제집행 등으로 본 계약상의 의무이행이 어려울 것으로 판단되는 경우 또는 파산 또는 회생의 신청이 이루어졌을 경우
  - 2. 천재지변 등으로 인해 본 계약을 더 이상 이행하기 불가능 할 경우

**제 13 조 (통지의무)**

“갑”과 “을”은 본 계약 체결 당시에 알고 있는 상호, 대표자, 소재지, 업종, 주소가 변동되거나 합병, 영업양도, 부도, 회생신청, 파산신청 등 신용상태에 변동이 있거나 변경될 우려가 있는 경우 이를 지체 없이 상대방에게 통지하여야 한다.

**제 14 조 (계약변경)**

“본 계약”의 일부 또는 전부를 변경할 필요가 있는 경우에는 “갑”과 “을”의 서면 합의에 의하여 이를 변경하고, 그 변경 내용은 변경한 날 그 다음날부터 효력을 가진다.

**제 15 조 (권리의무 승계)**

“본 계약”상의 모든 권리와 의무는 “갑” 또는 “을”의 합병, 영업양도, 경영 위임 등의 경우에도 “갑” 또는 “을”의 합병회사, 영업양수인, 경영수임인 등에게 승계되며, “갑” 또는 “을”은 그들로 하여금 “본 계약”상의 권리와 의무를 승계하는 것에 동의하도록 할 의무를 진다.

**제 16 조 (권리양도금지)**

“갑”과 “을”은 상대방의 사전 서면동의 없이 본 계약상의 일체의 권리 및 의무 등을 제3자에게 양도, 증여, 대물변제, 대여하거나 담보로 제공할 수 없으며, 위탁, 하도급 할 수 없다.





**제 17 조 (권리의 귀속 등)**

- ① "본 계약"에 따라 제공되는 "물질" 또는 "기술정보" 또는 "취득정보"에 대한 모든 권리는 이를 제공한 "당사자"에 속한다.
- ② "본 계약"은 어떠한 경우에도 정보수령자에게 "물질" 또는 "기술정보" 또는 "취득정보"에 관한 어떠한 권리나 사용권을 부여하는 것으로 해석되지 않는다.

본 계약 체결의 증거로서 본 계약서 [2]통을 작성하여 "갑"과 "을"이 기명 날인한 후에 각각 [1]통씩 보유한다.

계약체결일 : [2022]년 [06]월 [27]일

<p>"갑"</p> <p>대전광역시 유성구 과학로 125(어은동) 한국생명공학연구원 (법인등록번호 : 160171-0001632) 대표자 원장 김장성 (인)</p> 	<p>"을"</p> <p>대전 서구 대덕대로317번길 20, 458호(월평동) (주)유클리드소프트 (법인등록번호 : 160111-0371065) 대표자 채은경(인)</p> 
---	--



## 확 인 서

### ■ 기술실시계약

실시계약명	기술이전계약서	
(예정) 실시기업명	㈜유클리드소프트	
관련 연구과제	연구과제명 (부처/계정번호)	인공지능 기반 유전체 분석 프로그램 및 디지털 육종 플랫폼 개발
	연구과제 수행기간	2021.04.01. - 2023.12.31.
	연구책임자	김용민
	기여율	
실시 계약 대상 기술	순번	출원번호(국가)
	1	등록번호 C-2020-024215
	2	
		기술의 명칭 또는 내용
		전장유전체 데이터를 활용한 프라이머 제작 파이프라인

위 기술실시계약과 관련하여 연구책임자 김용민, (예정)실시기업의 대표(또는 기술이전 총괄책임자) 박주한 은/는 각자의 독자적인 검토 및 책임 하에 다음과 같은 절차를 진행하였음을 확인합니다.

- 다 음 -

#### <기술실시계약 연구책임자>

1. 위 연구책임자는 실시계약 대상기술에 관하여, (예정)실시기업이 실시계약에서 정한 바에 따라 이전받아 실시할 수 있는 수준의 기술을 개발하였으며, 이를 실시계약 체결 이전에 확인하였음.
2. 위 연구책임자는 (예정)실시기업의 대표(또는 기술이전 총괄책임자)에게 위 실시계약 대상기술의 내용, 범위, 실시방법 등에 관하여 충분히 설명하였음.

#### <(예정)실시기업>

1. 위 연구책임자로부터 위 실시계약 대상기술의 기술의 내용, 범위, 완성도, 실시방법 등에 관하여 충분한 설명을 받았고, 실시계약 대상기술의 실시가능성에 관하여 필요한 기술검토를 완료하였음.
2. 위 (예정)실시기업은 위와 같은 기술검토에 따라 실시계약 대상기술의 실시가능성을 인정하였는바, 이에 스스로의 책임 하에 한국생명공학연구원과 사이에 위

기술실시(이전) 증빙 1

실시계약을 체결하고자 위 기술실시계약을 신청함.

- 위 (예정)실시기업은 본 확인서 제출 이후 실시계약 대상기술의 실시가능성에 대해서 일체의 민형사상 이의를 제기하지 아니할 것을 약속함.

2022 년 5 월 30 일

연구책임자

김 용 만



(예정) 실시기업

(주) 유클리드소프트

대표자

대표이사 박 주 환



한국생명공학연구원장 귀하

## 기술실시계약 체결 관련자료

### 1. 기술의 내용

#### 1)기술개요

계통별 SNP calling 기술.

- ① paired-end fastq 파일을 이용하여 quality check, quality control 등 데이터의 전처리.
- ② 표준유전체에 mapping.
- ③ 계통별 표준유전체에 대한 변이(SNP) 산출.
- ④ 계통별 변이 정보를 활용하여 집단의 변이 데이터로 통합.

#### 2)기술 우수성(유사한 타기술과의 차이점 위주)

- ① 특정 디렉토리 상에 업로드된 작물의 paired-end fastq 파일을 자동으로 인식.
- ② 데이터의 quality check, quality control, mapping 및 변이 calling의 자동화.
- ③ 데이터의 quality check, quality control, mapping 및 변이 calling의 pipeline을 일원화.

#### 3)기대 효과

이전 대상 기술을 활용하여 유전체 데이터 분석을 위한 생산 공정에 도입, 향후 개체 및 계통 특이적 변이 및 마커 선정에 활용할 수 있음.

### 2. 기술활용결과의 형태(재료생산,완제품생산,일부공정대체,생산성향상 등)

유전체 분석 공정 내 raw 데이터 처리 **공정 대체**

### 3. 기술관련 특허 및 디자인 현황(출원, 등록 등)

실시 계약 대상 기술	출원번호(국가)	기술의 명칭
	등록번호 C-2020-024215	전장유전체 데이터를 활용한 프라이머 제작 파이프라인

### 4. 기술관련 노하우(Know-How)

- ① quality check, quality control, mapping 및 변이 calling 세부 공정 내 작물 데이터를 다루기에 최적, 최신의 프로그램 선택
- ② 문헌 연구 및 경험적 연구에 따른 최적의 option 선택

### 5. 연구종료일 이후 생산개시까지 당원(연구실)이 제공할 내용 :

기술실시(이전) 증빙 1

형태	구분
1. 보고서등 관련서류 제공	여·부
2. 실시자측 요원에 대한 기술교육	여·부
3. 전문가 지원(PILOT PLANT TEST지원)	여·부
4. 관련 장비이관(기자재,시설등)	여·부

### 청렴유지 서약서

한국생명공학연구원 기술이전책임자 김용민 (이하 "기술이전책임자"라 함)와 기술을 이전받는 (주) 유클리트소프트의 실무책임자 박주한 (이하 "실시기업 실무책임자"라 함)는(은) 기술이전계약 체결 및 동 계약의 이행이 공정하고 투명하게 이루어질 수 있도록 다음과 같이 서약한다.

1. "기술이전책임자"는 "실시기업 실무책임자"와의 기술이전계약 및 이의 이행에 있어,
  - 첫째, 기술이전 계약을 준수하고 "한국생명공학연구원"의 관련규정이 정하는 절차에 따라 신속하고 공정하게 처리하겠습니다.
  - 둘째, 기술이전과 관련하여 선물, 향응, 접대 등을 요구하거나 제공하지 않겠으며, 부당한 강요 및 협박 등 어떠한 불법적인 행위도 행사하지 않을 것이며, 이를 위반할 시에는 관계법령에 의거 엄중히 책임지도록 하겠습니다.
  - 셋째, 직무수행 중 알게 된 정보를 이용하여 주식 등 유가증권·부동산 등과 관련된 재산상 거래 또는 투자를 하거나 타인에게 그러한 정보를 제공하여 재산상 거래 또는 투자를 돕는 행위를 일체 하지 않겠습니다.
  - 넷째, 상기 청렴유지계약의 위반사항이 신고되었을 경우 신고인과 신고내용 등에 관하여 비밀을 보장할 것이며, "실시기업 실무책임자" 또는 신고인이 신고에 따른 어떠한 차별이나 불이익도 당하지 않도록 하겠습니다.
2. "실시기업 실무책임자"는 "기술이전책임자"와의 기술이전계약 및 이의 이행에 있어,
  - 첫째, 기술이전과 관련한 일체의 불공정 행위를 하지 않겠습니다.
  - 둘째, 기술이전과 관련하여 기술이전책임자 및 관련자에게 직·간접적으로 금품·향응 등의 부당한 이익을 제공하거나 받지 않겠습니다.
  - 셋째, 이전기술을 사업화 등 기술실시 이외의 목적에 사용하거나 한국생명공학연구원에 불이익이 초래되는 방법으로 활용하지 않을 것입니다.
  - 넷째, 본 기술이전계약을 위반하는 경우 국가연구개발사업 참여 및 기술이전 제한, 계약해지 등 일체의 제재조치와 관련하여 한국생명공학연구원을 상대로 손해배상 청구 기타 민·형사상 어떠한 이의 제기도 하지 않겠습니다.

2022 년 5 월 30 일

한국생명공학연구원  
 주 소 : 대전광역시 유성구 과학로 125  
 기술이전책임자 : 김 용 민 (인)  
 연 락 처 : [Redacted]

(주) 유클리트소프트  
 주 소 : 대전광역시 서구 대동매문 317번길 20, 4층  
 실시기업 실무책임자 : 박 주 한 (인)  
 연 락 처 : [Redacted]

제 C-2020-024215 호



# 저작권 등록증

- 1. 저작물의 제호(명칭)      전장유전체 데이터를 활용한 프라이머 제작 파이프라인
- 2. 저작물의 종류            컴퓨터프로그램저작물>응용프로그램>과학기술
- 3. 저작자 성명(법인명)      한국생명공학연구원  
대전광역시 유성구 과학로
- 4. 생년월일(법인등록번호)    160171-0001632
- 5. 창작연월일                2020년03월23일
- 6. 공표연월일                2020년07월21일
- 7. 등록연월일                2020년07월23일
- 8. 등록사항                  저작자 : 한국생명공학연구원,  
창작 : 2020.03.23, 공표 : 2020.07.21

「저작권법」 제53조에 따라 위와 같이 등록되었음을 증명합니다.

2020년 07월 27일

한국저작권위원회



신규채용 증빙 1 (유클리드소프트, 2021년, 신하승, 장주영)

출력일시: 2021.12.27 14:03

4대 사회보험 사업장 가입자 명부						
발급번호	20211227227146	발급일시	2021-12-27 14:02	사업장 관리번호 31486593830		
구분	국민연금	건강보험	산재보험	고용보험		
사업자등록번호	314-86-59383	314-86-59383	314-86-59383	314-86-59383		
사업장 명칭	주식회사유클리드소프트	주식회사유클리드소프트	주식회사유클리드소프트	주식회사유클리드소프트		
■ 가입 내역(발급일자 현재기준)				1 / 8		
연번	주민(외국인) 등록번호	성명	자격 취득일			
			국민연금	건강보험	산재보험	고용보험
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
<div style="display: flex; justify-content: space-between;"> </div> <p>▷ 위 사업장 가입자 명부는 4대사회보험 정보연계시스템이 국민연금공단, 국민건강보험공단, 근로복지공단의 가입자 정보를 실시간 연계받아 제공하는 것이며, 발급사실 여부는 발급일로부터 90일까지 4대사회보험 포털사이트(<a href="http://www.4insure.or.kr">www.4insure.or.kr</a>)의 [발급사실확인] 메뉴에서 확인 가능합니다. "청렴한 정보연계서비스, 4대 사회보험이 함께 합니다."</p>						

신규채용 증빙 1 (유클리드소프트, 2021년, 신하승, 장주영)

■ 가입 내역(발급일자 현재기준)		발급번호: 20211227227146		출력일시: 2021.12.27 14:03		2 / 8	
연번	주민(외국인) 등록번호	성명	자격취득일				
			국민연금	건강보험	산재보험	고용보험	
21							
22							
23							
24							
25							
26							
27							
28							
29							
30							
31							
32							
33							
34							
35							
36							
37							
38							
39							
40							
41							
42							
43							
44							
45							
46							
47							
48							
49							
50							
51							

> 위 사업장 가입자 명부는 4대사회보험 정보연계시스템이 국민연금공단 국민건강보험공단 근로복지공단의  
 가입자 정보를 실시간 연계받아 제공하는 것이며, 발급사실 여부는 발급일로부터 90일까지 4대사회보험  
 포털사이트(www.4insure.or.kr)의 [발급사실확인] 메뉴에서 확인 가능합니다.  
 "청렴한 정보연계서비스, 4대 사회보험이 함께 합니다."

신규채용 증빙 1 (유클리드소프트, 2021년, 신하승, 장주영)

■ 가입 내역(발급일자 현재기준)		발급번호: 20211227227146		출력일시: 2021.12.27 14:03		3 / 8	
연번	주민(외국인) 등록번호	성명	자격취득일				
			국민연금	건강보험	산재보험	고용보험	
52							
53							
54							
55							
56							
57							
58							
59							
60							
61							
62							
63							
64							
65							
66							
67							
68							
69							
70							
71							
72							
73							
74							
75							
76							
77							
78							
79							
80							
81							
82							

> 위 사업장 가입자 명부는 4대사회보험 정보연계시스템이 국민연금공단, 국민건강보험공단, 근로복지공단의 가입자 정보를 실시간 연계받아 제공하는 것이며, 발급사실 여부는 발급일로부터 90일까지 4대사회보험 포털사이트(www.4insure.or.kr)의 [발급사실확인] 메뉴에서 확인 가능합니다.  
"정령한 정보연계서비스, 4대 사회보험이 함께 합니다."

신규채용 증빙 1 (유클리드소프트, 2021년, 신하승, 장주영)

■ 가입 내역(발급일자 현재기준)		발급번호: 20211227227146		출력일시: 2021.12.27 14:03		4 / 8	
연번	주민(외국인) 등록번호	성명	자격취득일				
			국민연금	건강보험	산재보험	고용보험	
83							
84							
85							
86							
87							
88							
89							
90							
91							
92							
93							
94							
95							
96							
97							
98							
99							
100							
101							
102							
103							
104							
105							
106							
107							
108							
109							
110							
111							
112							
113							

위 사업장 가입자 명부는 4대사회보험 정보연계시스템이 국민연금공단, 국민건강보험공단, 근로복지공단의 가입자 정보를 실시간 연계받아 제공하는 것이며, 발급사실 여부는 발급일로부터 90일까지 4대사회보험 포털사이트(www.4insure.or.kr)의 [발급사실확인] 메뉴에서 확인 가능합니다.  
"정려한 정보연계서비스, 4대 사회보험이 함께 합니다."

신규채용 증빙 1 (유클리드소프트, 2021년, 신하승, 장주영)

■ 가입 내역(발급일자 현재기준)		발급번호: 20211227227146		출력일시: 2021.12.27 14:03		5 / 8	
연번	주민(외국인) 등록번호	성명	자격취득일				
			국민연금	건강보험	산재보험	고용보험	
114							
115							
116							
117							
118							
119							
120							
121							
122							
123							
124							
125							
126							
127							
128							
129							
130							
131							
132							
133							
134							
135							
136							
137							
138							
139							
140							
141							
142							
143							
144							

위 사업장 가입자 명부는 4대사회보험 정보연계시스템이 국민연금공단, 국민건강보험공단, 근로복지공단의 가입자 정보를 실시간 연계받아 제공하는 것이며, 발급사실 여부는 발급일로부터 90일까지 4대사회보험 포털사이트(www.4insure.or.kr)의 [발급사실확인] 메뉴에서 확인 가능합니다.  
"청렴한 정보연계서비스, 4대 사회보험이 함께 합니다."

신규채용 증빙 1 (유클리드소프트, 2021년, 신하승, 장주영)

■ 가입 내역(발급일자 현재기준)		발급번호: 20211227227146		출력일시: 2021.12.27 14:03 6 / 8		
연번	주민(외국인) 등록번호	성명	자격취득일			
			국민연금	건강보험	산재보험	고용보험
145						
146						
147						
148						
149						
150						
151						
152						
153						
154						
155						
156						
157						
158						
159						
160						
161						
162						
163						
164						
165						
166						
167						
168						
169						
170						
171		신하승				
172						
173						
174						
175						

위 사업장 가입자 명부는 4대사회보험 정보연계시스템이 국민연금공단, 국민건강보험공단, 근로복지공단의 가입자 정보를 실시간 연계받아 제공하는 것이며, 발급사실 여부는 발급일로부터 90일까지 4대사회보험 포털사이트(www.4insure.or.kr)의 [발급사실확인] 메뉴에서 확인 가능합니다.  
"청렴한 정보연계서비스, 4대 사회보험이 함께 합니다."

신규채용 증빙 1 (유클리드소프트, 2021년, 신하승, 장주영)

■ 가입 내역(발급일자 현재기준)		발급번호: 20211227227146		출력일시: 2021.12.27 14:03 7 / 8		
연번	주민(외국인) 등록번호	성명	자격 취득일			
			국민연금	건강보험	산재보험	고용보험
176						
177						
178						
179						
180						
181						
182						
183		장주영				
184						
185						
186						
187						
188						
189						
190						
191						
192						
193						
194						
195						
196						
197						
198						



▷ 위 사업장 가입자 명부는 4대사회보험 정보연계시스템이 국민연금공단, 국민건강보험공단, 근로복지공단의 가입자 정보를 실시간 연계받아 제공하는 것이며, 발급사실 여부는 발급일로부터 90일까지 4대사회보험 포털사이트([www.4insure.or.kr](http://www.4insure.or.kr))의 [발급사실확인] 메뉴에서 확인 가능합니다.  
"정렬한 정보연계서비스, 4대 사회보험이 함께 합니다."

신규채용 증빙 1 (유클리드소프트, 2021년, 신하승, 장주영)

■ 가입 내역(발급일자 현재기준)      발급번호: 20211227227146      출력일시: 2021.12.27 14:03      8 / 8

연번	주민(외국인) 등록번호	성명	자격 취득일			
			국민연금	건강보험	산재보험	고용보험
이 하 여 백						
<p>▷ 위 사업장 가입자 명부는 [확인용]으로 신청·발급된 것임을 알려드립니다.</p> <p>- [확인용]은 4대 사회보험의 업무목적용 위해서만 제공하는 것이므로 재직증명용, 경력증명용, 대출용 등 다른 용도로 사용시에는 발급 기관에 법적 책임이 없다는 점을 알려드립니다.</p> <p>- 타 기관 제출을 위한 용도로 발급을 원하시는 경우에는 각 공단 지사 창구로 신청하시기 바랍니다.</p> <p>▷ 위 사업장 가입자 명부는 국민연금공단, 국민건강보험공단, 근로복지공단의 가입자 정보를 실시간 연계 받아 제공하는 것입니다. (문의전화: 국민연금 1355, 건강보험 1577-1000, 산재·고용보험 1588-0075)</p> <p>- 사업장 가입자 명부의 내용이 사실과 다를 경우에는 해당 공단으로 문의하시기 바랍니다.</p> <p>- 과거 가입내역은 해당 보험별 각 공단에 문의하여 발급받으시기 바랍니다.</p> <p>▷ [산재보험]의 경우, '자격취득일'은 근로자 고용일을 뜻하며, 건설업 및 벌목업 등 '자진신고 사업장'은 근로자 고용정보 신고 대상이 아니므로 '자격취득일(고용일)'은 표기되지 않습니다.</p> <p>▷ 위 사업장 가입자 명부는 [사업장 관리번호]를 기준으로 작성되었습니다.</p>						
위와 같이 국민연금 가입내역을 확인합니다.		위와 같이 건강보험 가입내역을 확인합니다.		위와 같이 산재보험 가입내역을 확인합니다.		
						
						

▷ 위 사업장 가입자 명부는 4대사회보험 정보연계시스템이 국민연금공단, 국민건강보험공단, 근로복지공단의 가입자 정보를 실시간 연계받아 제공하는 것이며, 발급사실 여부는 발급일로부터 90일까지 4대사회보험 포털사이트(www.4insure.or.kr)의 [발급사실확인] 메뉴에서 확인 가능합니다.  
"청림한 정보연계서비스, 4대 사회보험이 함께 합니다."

출력일시 : 2022.11.11 08:10

## 4대 사회보험 사업장 가입자 명부

발급번호	20221111484807	발급일시	2022-11-11 08:08	사업장 관리번호	31486593830
구분	국민연금	건강보험	산재보험	고용보험	
사업자등록번호	314-86-59383	314-86-59383	314-86-59383	314-86-59383	
사업장 명칭	주식회사유클리드소프트	주식회사유클리드소프트	주식회사유클리드소프트	주식회사유클리드소프트	

■ 가입 내역(발급일자 현재기준) 1 / 1

연번	주민(외국인) 등록번호	성명	자격 취득 일			
			국민연금	건강보험	산재보험	고용보험
1		윤봉구				
2		윤영권				
3		이종석				
4		장인지				
5		최열				

이 하 여 백

- ▷ 위 사업장 가입자 명부는 [확인용]으로 신청·발급된 것임을 알려드립니다.
- [확인용]은 4대 사회보험의 업무목적을 위해서만 제공하는 것이므로 재직증명용, 경력증명용, 대출용 등 다른 용도로 사용시에는 발급 기관에 법적 책임이 없다는 점을 알려드립니다.
- 타 기관 제출을 위한 용도로 발급을 원하시는 경우에는 각 공단 지사 창구로 신청하시기 바랍니다.
- ▷ 위 사업장 가입자 명부는 국민연금공단, 국민건강보험공단, 근로복지공단의 가입자 정보를 실시간 연계 받아 제공하는 것입니다.(문의전화: 국민연금 1355, 건강보험 1577-1000, 산재·고용보험 1588-0075)
- 사업장 가입자 명부의 내용이 사실과 다를 경우에는 해당 공단으로 문의하시기 바랍니다.
- 과거 가입내역은 해당 보험별 각 공단에 문의하여 발급받으시기 바랍니다.
- ▷ [산재보험]의 경우, '자격취득일'은 근로자 고용일을 뜻하며, 건설업 및 벌목업 등 '자진신고 사업장'은 근로자 고용정보 신고 대상이 아니므로 '자격취득일(고용일)'은 표기되지 않습니다.
- ▷ 위 사업장 가입자 명부는 [사업장 관리번호]를 기준으로 작성되었습니다.(일부인원에 대한 가입자명부)

위와 같이 국민연금 가입내역을 확인합니다.  <b>국민연금공단</b> 이사장	위와 같이 건강보험 가입내역을 확인합니다.  <b>국민건강보험공단</b> 이사장	위와 같이 산재보험 가입내역을 확인합니다.  <b>근로복지공단</b> 대전지역본부장	위와 같이 고용보험 가입내역을 확인합니다.  <b>근로복지공단</b> 대전지역본부장
			



▷ 위 사업장 가입자 명부는 4대사회보험 정보연계시스템이 국민연금공단, 국민건강보험공단, 근로복지공단의 가입자 정보를 실시간 연계받아 제공하는 것이며, 발급사실 여부는 발급일로부터 90일까지 4대사회보험 포털사이트(www.4insure.or.kr)의 [발급사실확인] 메뉴에서 확인 가능합니다.  
"정렬한 정보연계서비스, 4대 사회보험이 함께 합니다."

출력일시 : 2024.02.20 11:04

<b>4대 사회보험 사업장 가입자 명부</b>					
<b>확인용</b>					
발급번호	20240220587772	발급일시	2024-02-20 11:03	사업장 관리번호	31486593830
구분	국민연금	건강보험	산재보험	고용보험	
사업자등록번호	314-86-59383	314-86-59383	314-86-59383	314-86-59383	
사업장 명칭	주식회사유클리드소프트	주식회사유클리드소프트	주식회사유클리드소프트	주식회사유클리드소프트	
■ 가입 내역(발급일자 현재기준)					1 / 1
연번	주민(외국인) 등록번호	성명	자격취득일		
			국민연금	건강보험	산재보험
1		최부겸			
이 하 여 백					
<p>▷ 위 사업장 가입자 명부는 [확인용]으로 신청·발급된 것임을 알려드립니다.</p> <p>- [확인용]은 4대 사회보험의 업무목적을 위해서만 제공하는 것이므로 재직증명용, 경력증명용, 대출용 등 다른 용도로 사용시에는 발급 기관에 법적 책임이 없다는 점을 알려드립니다.</p> <p>- 타 기관 제출을 위한 용도로 발급을 원하시는 경우에는 각 공단 지사 창구로 신청하시기 바랍니다.</p> <p>▷ 위 사업장 가입자 명부는 국민연금공단, 국민건강보험공단, 근로복지공단의 가입자 정보를 실시간 연계 받아 제공하는 것입니다. (문의전화: 국민연금 1355, 건강보험 1577-1000, 산재·고용보험 1588-0075)</p> <p>- 사업장 가입자 명부의 내용이 사실과 다를 경우에는 해당 공단으로 문의하시기 바랍니다.</p> <p>- 과거 가입내역은 해당 보험별 각 공단에 문의하여 발급받으시기 바랍니다.</p> <p>▷ [산재보험]의 경우, '자격취득일'은 근로자 고용일을 뜻하며, 건설업 및 발목업 등 '자진신고 사업장'은 근로자 고용정보 신고 대상이 아니므로 '자격취득일(고용일)'은 표기되지 않습니다.</p> <p>▷ 위 사업장 가입자 명부는 [사업장 관리번호]를 기준으로 작성되었습니다.</p>					
위와 같이 국민연금 가입내역을 확인합니다. <b>국민연금 이 사</b> 		위와 같이 건강보험 가입내역을 확인합니다. <b>국민건강보 이 사</b> 		위와 같이 산재보험 가입내역을 확인합니다. <b>근로복지 대전지역</b> 	
					

<b>본</b>	  
<p>▷ 위 사업장 가입자 명부는 4대사회보험 정보연계시스템이 국민연금공단, 국민건강보험공단, 근로복지공단의 가입자 정보를 실시간 연계받아 제공하는 것이며, 발급사실 여부는 발급일로부터 90일까지 4대사회보험 포털사이트(www.4insure.or.kr)의 [발급사실확인] 메뉴에서 확인 가능합니다.</p> <p>"청렴한 정보연계서비스, 4대 사회보험이 함께 합니다."</p>	

신규채용 증빙 3 (유클리드소프트, 2023년, 신예빈, 김연주, 김유상)

출력일시 : 2023.10.26 16:42

### 4대 사회보험 사업장 가입자 명부

발급번호	20231026375181	발급일시	2023-10-26 16:42	사업장 관리번호	31486593830
구분	국민연금	건강보험	산재보험	고용보험	
사업자등록번호	314-86-59383	314-86-59383	314-86-59383	314-86-59383	
사업장 명칭	주식회사유클리드소프트	주식회사유클리드소프트	주식회사유클리드소프트	주식회사유클리드소프트	

■ 가입 내역(발급일자 현재기준) 1 / 1

연번	주민(외국인) 등록번호	성명	자격 취득 일			
			국민연금	건강보험	산재보험	고용보험
1		신예빈				
2		김연주				
3		김유상				

이 하 여 백

▷ 위 사업장 가입자 명부는 [확인용]으로 신청·발급된 것임을 알려드립니다.  
 - [확인용]은 4대 사회보험의 업무목적용을 위해서만 제공하는 것이므로 재직증명용, 경력증명용, 대출용 등 다른 용도로 사용시에는 발급 기관에 법적 책임이 없다는 점을 알려드립니다.  
 - 타 기관 제출을 위한 용도로 발급을 원하시는 경우에는 각 공단 지사 창구로 신청하시기 바랍니다.  
 ▷ 위 사업장 가입자 명부는 국민연금공단, 국민건강보험공단, 근로복지공단의 가입자 정보를 실시간 연계 받아 제공하는 것입니다. (문의전화: 국민연금 1355, 건강보험 1577-1000, 산재·고용보험 1588-0075)  
 - 사업장 가입자 명부의 내용이 사실과 다를 경우에는 해당 공단으로 문의하시기 바랍니다.  
 - 과거 가입내역은 해당 보험별 각 공단에 문의하여 발급받으시기 바랍니다.  
 ▷ [산재보험]의 경우, '자격취득일'은 근로자 고용일을 뜻하며, 건설업 및 벌목업 등 '자진신고 사업장'은 근로자 고용정보 신고 대상이 아니므로 '자격취득일(고용일)'은 표기되지 않습니다.  
 ▷ 위 사업장 가입자 명부는 [사업장 관리번호]를 기준으로 작성되었습니다. (일부인원에 대한 가입자명부)

위와 같이 국민연금 가입내역을 확인합니다.  <b>국민연금공단</b> 이 사 장	위와 같이 건강보험 가입내역을 확인합니다.  <b>국민건강보험공단</b> 이 사 장	위와 같이 산재보험 가입내역을 확인합니다.  <b>근로복지공단</b> 대전지역본부장	위와 같이 고용보험 가입내역을 확인합니다.  <b>근로복지공단</b> 대전지역본부장
--	--	---	--





▷ 위 사업장 가입자 명부는 4대사회보험 정보연계시스템이 국민연금공단, 국민건강보험공단, 근로복지공단의 가입자 정보를 실시간 연계받아 제공하는 것이며, 발급사실 여부는 발급일로부터 90일까지 4대사회보험 포털사이트(www.4insure.or.kr)의 [발급사실확인] 메뉴에서 확인 가능합니다.  
 "청결한 정보연계서비스, 4대 사회보험이 함께 합니다."



# 우수논문발표상

소 속: 한국생명공학연구원  
이 름: 고상진

논문제목: Construction of high-contiguity reference genome sequences of two Hibiscus syriacus accessions (Gangneung and Baekdansim)

저 자: Sangjin Go, Seongmin Hong, Ah-Young Shin and Yong-Min Kim\*

귀하는 2022년 8월 11일~13일 라마다프라자 제주 호텔에서 개최된 '2022년 한국식물생명공학회 정기학술발표회'에서 우수논문발표상으로 선정되었기에 이 상을 드립니다.

2022년 8월 12일

(사) 한국식물생명공학회 회장 이효연



< 연구개발성과 활용계획표 >

구분(정량 및 정성적 성과 항목)		연구개발 종료 후 5년 이내	
국외논문	SCIE	2	
	비SCIE		
	계	2	
국내논문	SCIE		
	비SCIE		
	계		
특허출원	국내		
	국외		
	계		
특허등록	국내	2	
	국외		
	계	2	
인력양성	학사		
	석사		
	박사		
	계		
사업화	상품출시		
	기술이전		
	공정개발		
제품개발	시제품개발		
비임상시험 실시			
임상시험 실시 (IND 승인)	의약품	1상	
		2상	
		3상	
	의료기기		
진료지침개발			
신의료기술개발			
성과홍보			
포상 및 수상실적			
정성적 성과 주요 내용			

< 별첨 자료 >

중앙행정기관 요구사항	별첨 자료
1. 공통 요구자료	1) 자체평가의견서
	2) 연구성과 활용계획서
	3) 연구부정행위 예방 확인서
2.	1)
	2)

주 의

1. 이 보고서는 농림축산식품부에서 시행한 기술사업화지원사업의 연구보고서입니다.
2. 이 보고서 내용을 발표하는 때에는 반드시 농림축산식품부에서 시행한 기술사업화지원사업의 연구결과임을 밝혀야 합니다.
3. 국가과학기술 기밀 유지에 필요한 내용은 대외적으로 발표 또는 공개하여서는 안 됩니다.