

9180130
44-010

농림축산식품 분야를 위한 데이터베이스 및 소프트웨어 개발

농림축산식품부

위한 데이터베이스 및 소프트웨어 개발

2021

농림식품기술기획평가원

보안 과제(), 일반 과제(O) / 공개(O), 비공개() 발간등록번호(O)
포스트게놈다부처유전체사업 2021년도 최종보고서

발간등록번호

11-1543000-004043-01

농림축산식품 분야를 위한 메타유전체의 통합 분석을 위한 데이터베이스 및 소프트웨어 개발

2022. 03. 29

주관연구기관 / 서울대학교 산학협력단

농림축산식품부
(전문기관) 농림식품기술기획평가원

제 출 문

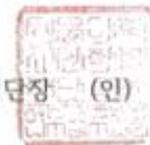
농림축산식품부 장관 귀하

본 보고서를 “농림축산식품 분야를 위한 메타유전체의 통합 분석을 위한 데이터베이스 및 소프트웨어 개발”(개발기간 : 2018. 04. 25 ~ 2021 . 12. 31)과제의 최종보고서로 제출합니다.

2022. 03. 29

주관연구기관명 :

서울대학교 산학협력단장 (인)



주관연구책임자 : 천 종 식

국가연구개발사업의 관리 등에 관한 규정 제18조에 따라 보고서 열람에 동의합니다.

최종보고서

보안등급
일민(), 보안()

중앙행정기관명				사업명		사업명	
전문기관명 (해당 시 작성)				사업명		내역사업명 (해당 시 작성)	
공고번호				총괄연구개발 식별번호 (해당 시 작성)			
				연구개발과제번호			
기술분류	국가과학기술 표준분류	LA0706	50%	LA0204	30%	LA0506	20%
	농림식품과학기술분류	CA0399	10%		0%		
총괄연구개발명 (해당 시 작성)		국문					
		영문					
연구개발과제명		국문		농림축산식품 분야를 위한 메타유전체의 통합 분석을 위한 데이터베이스 및 소프트웨어 개발			
		영문		Development of integrated database and bioinformatics software for metagenomics for agriculture			
주관연구개발기관		기관명		서울대학교 신학협력단		사업자등록번호	
		주소		(우) 08826 서울시 관악구 관악로1 서울대 연구 공원 942동 5층		119-82-03684	
연구책임자		성명		전충식		직위	
		연락처		직장전화		휴대전화	
		전자우편				국가연구자번호	
연구개발기간		전체		2018. 04. 25 - 2021. 12. 31(3년 9개월)			
연구개발비 (단위: 천원)		정부지원 연구개발비		기관부담 연구개발비		그 외 기관 등의 지원금 지방자치단체 기타()	
		현금		현금		현금	
		원금		원금		원금	
총계		750,000				750,000	
1년차		150,000				150,000	
2년차		200,000				200,000	
3년차		200,000				200,000	
4년차		200,000				200,000	
공동연구개발기관 등 (해당 시 작성)		기관명		책임자		직위	
		공동연구개발기관					
		위탁연구개발기관		서울대학교		윤성로	
		연구개발기관 외 기관					
연구개발담당자 실무담당자		성명		박세인		직위	
		연락처		직장전화		휴대전화	
				전자우편		국가연구자번호	

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

2022 년 02 월 10 일

연구책임자

전 충 식

주관연구개발기관의 장:

서울대학교 신학협력단장 (직인)

위탁연구개발기관의 장:

윤 성 로 (직인)

농림축산식품부장관·농림식품기술기획평가원장 귀하

최종보고서							보안등급			
							일반[], 보안[]			
중앙행정기관명				사업명	사업명					
전문기관명 (해당 시 작성)					내역사업명 (해당 시 작성)					
공고번호				총괄연구개발 식별번호 (해당 시 작성)						
				연구개발과제번호						
기술분류	국가과학기술 표준분류	LA0706	50%	LA0204	30%	LA0506	20%			
	농림식품과학기술분류	CA0399	10%		%		%			
총괄연구개발명 (해당 시 작성)		국문								
		영문								
연구개발과제명		국문	농림축산식품 분야를 위한 메타유전체의 통합 분석을 위한 데이터베이스 및 소프트웨어 개발							
		영문	Development of integrated database and bioinformatics software for metagenomics for agriculture							
주관연구개발기관		기관명	서울대학교 산학협력단		사업자등록번호		119-82-03684			
		주소	(우) 08826 서울시 관악구 관악로1 서울대 연구공원 942동 5층		법인등록번호					
연구책임자		성명	천종식		직위		교수			
		연락처	직장전화			휴대전화				
			전자우편			국가연구자번호				
연구개발기간	전체		2018. 04. 25 - 2021. 12. 31(3년 9개월)							
연구개발비 (단위: 천원)	정부지원 연구개발비	기관부담 연구개발비	그 외 기관 등의 지원금				합계			연구개발비 외 지원금
	현금	현금	현물	현금	현물	현금	현물	현금	현물	
총계	750,000							750,000		750,000
1년차	150,000							150,000		150,000
2년차	200,000							200,000		200,000
3년차	200,000							200,000		200,000
4년차	200,000							200,000		200,000
공동연구개발기관 등 (해당 시 작성)	기관명	책임자	직위	휴대전화	전자우편	비고				
	역할	기관유형								
공동연구개발기관										
위탁연구개발기관		서울대학교	윤성로	교수				위탁	대학	
연구개발기관 외 기관										
연구개발담당자 실무담당자		성명	박세인		직위		박사과정			
		연락처	직장전화			휴대전화				
			전자우편			국가연구자번호				

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

2022 년 03 월 29 일

연구책임자: 천종식 (인)

주관연구개발기관의 장: 서울대학교 산학협력단장 (직인)
위탁연구개발기관의 장: 윤성로 (직인)

< 요약 문 >

사업명				총괄연구개발 식별번호 (해당 시 작성)			
내역사업명 (해당 시 작성)				연구개발과제번호			
기술 분류	국가과학기술 표준분류	LA0706	50 %	LA0204	30 %	LA0506	20%
	농림식품 과학기술분류	CA0399	10 0%		%		%
총괄연구개발명		농림축산식품 미생물유전체전략연구사업단					
연구개발과제명		농림축산식품 분야를 위한 메타유전체의 통합 분석을 위한 데이터베이스 및 소프트웨어 개발					
전체 연구개발기간		2018. 04. 25 - 2021. 12. 31					
총 연구개발비		총 750,000 천원 (정부지원연구개발비: 750,000 천원, 기관부담연구개발비 : 0 천원, 지방자치단체: 0 천원, 그 외 지원금: 0천원)					
연구개발단계		기초[<input checked="" type="checkbox"/>] 응용[<input type="checkbox"/>] 개발[<input type="checkbox"/>] 기타(위 3가지에 해당되지 않는 경우)[<input type="checkbox"/>]		기술성숙도 (해당 시 기재)		착수시점 기준(<input type="checkbox"/>) 종료시점 목표(<input type="checkbox"/>)	
연구개발과제 유형 (해당 시 작성)							
연구개발과제 특성 (해당 시 작성)							
연구개발 목표 및 내용	최종 목표	<ul style="list-style-type: none"> ○ 농축산 관련 유용 미생물의 메타유전체 연구를 원활히 수행할 수 있도록 정보분석 파이프라인을 고도화 ○ 주요 농식품 미생물 유전체 기반 특화 데이터베이스를 구축하여, 관련 분야 연구자들의 신규 유전자 발굴 및 산업화 연구지원 					
	전체 내용	<ul style="list-style-type: none"> ○ 표준화된 NGS 데이터 유전체 및 마이크로바이옴 분석 파이프라인 구축 및 Deep-learning 등을 활용한 고도화 ○ 농식품 유용 미생물 종의 phylogenomics 파이프라인 구축 ○ 농식품 유용 미생물 종의 비교 유전체 및 pan/core genome database 구축 ○ 농식품 유용 미생물 종의 genomic island database 구축 ○ 표준화되고, big data 분석이 가능한 농식품 관련 대규모 마이크로바이옴 16S 및 shotgun database 구축 ○ 축산 (돼지,소,닭), 전통 발효 식품의 미생물 gene catalog database 구축 ○ 마이크로바이옴 네트워크 분석 알고리즘 및 소프트웨어 개발 					
연구개발성과	<ul style="list-style-type: none"> ○ 표준화되고, 고도화된 NGS 데이터 유전체 및 마이크로바이옴 분석 파이프라인 ○ 농축산 관련 유용 미생물 종에 대한 phylogenomics, 비교 유전체, pan/core genome, genomic island 등의 통합 데이터베이스 구축 ○ 농축산 관련 마이크로바이옴의 대규모 데이터베이스 구축 ○ 네트워크 분석을 비롯한 마이크로바이옴 관련 메타유전체 분석 파이프라인 개발 ○ 사업단 내 연구자를 위한 웹서버 및 파이프라인 분석 지원 						

연구개발성과 활용계획 및 기대 효과	<ul style="list-style-type: none"> ○ 농축산 관련 미생물 유전체 및 마이크로바이옴 데이터베이스의 구축 및 분석 소프트웨어의 제공을 통한 국내 관련 연구자 또는 기업의 연구역량 증가 ○ 학술적으로 높은 임팩트(인용 지수)의 논문 발표 ○ 농축산 관련 미생물 유전체 및 생명정보학 전문 인력의 교육 및 배출 												
연구개발성과의 비공개여부 및 사유													
연구개발성과의 등록·기탁 건수	논문	특허	보고서 원문	연구 시설 ·장비	기술 요약 정보	소프트 웨어	표준	생명자원		화합물	신품종		
								생명 정보	생물 자원		정보	실물	
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)	유전체			생명정보학		비교유전체		마이크로바이옴		데이터베이스			
영문핵심어 (5개 이내)	Genome			Bioinformatics		Comparative Genomics		Microbiome		Database			

< 목 차 >

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

별첨 자료 (참고 문헌 등)

1. 연구개발과제의 개요

○ 미생물의 90% 이상은 직접 배양이 어려움.

- 현재 배양되어 있는 미생물들은 미생물 전체의 1% 에 지나지 않는다는 보고도 있음(Cho J. et al., 2004).

- 미생물 생태에서 존재하는 미생물들의 전체를 파악하고 서로간의 네트워크 분석을 해야 하는 상황에서 전체 1%정도밖에 분석할 수 없다면, 그 연구는 무의미 하다고 볼 수 있음.

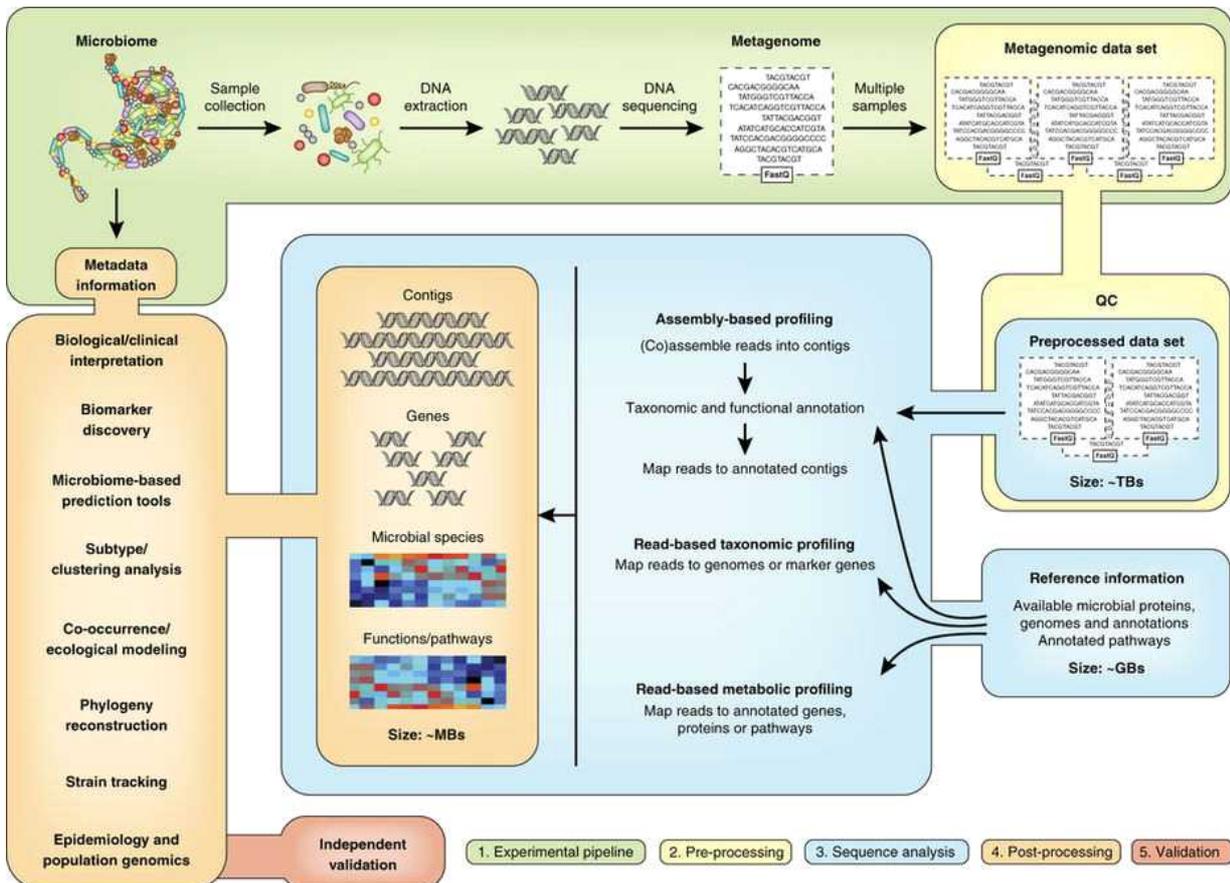
○ 메타지노믹스

- 1998년 이래로 환경시료로부터 시료에 존재하는 모든('Meta-') 유전자를 시퀀싱하여 해당 시료에 존재하는 모든 미생물군을 조사하려는 학문인 메타지노믹스분야가 등장했음(Jo Handelsman et al., 1998).

- 메타지노믹스는 환경 미생물학자들에게 환경 시료내의 미생물 군집에 존재하는 유전적 다양성을 직접적으로 탐구할 수 있는 실험 방법.

- 메타지노믹스 방법이 시행 된 이래로 메타지노믹스라 하면 특정 유전자위를 증폭시켜 얻는 방법이 주를 이루었음. (high-throughput 16S rRNA gene sequencing), 앰플리콘 메타지노믹스.

- 앰플리콘 메타지노믹스 환경 시료 내의 미생물들이 갖고 있는 표지 유전자인 16S rRNA 유전자를 증폭함으로써 환경 시료 속에 있는 개체의 정보를 얻는 방법임. 지금까지도 많은 환경미생물학 분야에서는 이 방법으로 미생물 생태를 연구하고 있음.



○ 샷건 메타지노믹스

- 기존 메타지노믹스의 한계 : 표지 유전자의 미생물 개체 내의 개수가 복수인 것에서부터 증폭이 되지 않는 것까지 다양한 단점이 발견되어 새로운 실험적 방법이 요구됨.

- 샷건(무작위 대상) 메타지노믹스 방법을 고안하게 되었음(JunJie Qin et al., 2010).

- 샷건 메타지노믹스방법은 환경 시료 내에 있는 미생물들의 유전자를 무작위로 쪼개어 해당 유전자들을 시퀀싱한 다음 생물정보학적 기법들로 정보를 얻어 내는 방법

- 장점: 기존의 한 유전자 단위 증폭 메타지노믹스에 비해 추가적인 정보들을 다량 얻을 수 있음.
 - 단점: 정보 해독에 필요한 생물정보학적 기법들이 어려움.
- 생물정보학적 기법들을 제대로 수행할 수 있는 연구진이 있다면 기존의 방법보다 다양하고 정확한 정보들을 얻을 수 있기에 최근 들어 각광을 받고 있는 분야임.

[농식품과 관련된 미생물 연구를 위한 환경 분석]

- 환경 시료와 미생물 군집의 상호 작용
 - 거의 모든 농식품의 시료는 미생물 군집과 상호작용함.
 - 예) 김치는 발효되는 시시각각 미생물의 작용에 의해 발효가 됨(Jeon C. et al., 2011).
- 환경 시료 분석에서 메타지노믹스의 필요성
 - 변화하는 미생물 군집의 환경 분석은 과거에 어떤 미생물들이 존재했는지에만 관심이 있었음.
 - 미생물 환경 학자들의 연구하려는 목적과 필요한 정보들이 고도화 됨.
 - 샷건 메타지노믹스로 얻을 수 있는 정보가 미생물 학자들의 연구 목적을 충족할 수 있음.
- 생물정보학의 중요성
 - 샷건 메타지노믹스로 얻을 수 있는 정보가 기하급수적으로 늘어남.
 - 정보가 늘어남에 따라 정보를 올바르게 정확히 처리할 수 있는 능력을 필요로 하게 됨.
 - 생물정보학자들의 연구 활동이 중요시 됨.
- 생물정보학에서 다룰 수 있는 샷건 메타지노믹스 분야
 - 통합 유전체 분석
 - 후성유전체와 유전체와의 통합 연계 분석
 - 대사산물 총체 분석
 - 통합단백질분석
 - 다중오믹스 분석
 - 숙주 미생물 상호 대사 네트워크 분석
 - 미생물 유전체 데이터베이스 구축
 - 미생물 군집 메타지노믹스 데이터베이스 구축

2. 연구개발과제의 수행 과정 및 수행 내용

- 연차별 세부 과제들의 연구 목표 및 해당 연구 수행 내용

구분 (연도)	세부과제명	세부연구목표	연구개발 수행내용
1차 년도 (2018)	대규모 마이크로바이 옴 데이터베이스 구축	공공 데이터저장소로부터 농식품 관련 마이크로바 이옴 NGS 데이터를 확보 처리 방식 및 메타데이터 정보가 상이한 파편화된 데이터로부터 통일된 포 맷의 데이터베이스 구축	NCBI SRA 및 MG-RAST에서 시퀀스를 다운로드 받고 파편화된 데이터 양식을 편집 받은 데이터 필터링 및 샘플 단위에서 clustering을 거침
	phylogenomi cs 파이프라인	계통수 작성을 위한 bacterial core gene 기반	세균 및 고세균에 대해서 core gene을 선별하고 그 기반의 분류 파이프라인 구축

구축	의 파이프라인을 구축	
	농식품 관련 속과 종에 대해서 계통수 작성	phylogenomic 파이프라인을 활용한 계통도 작성
농식품 관련 종에 대한 genomic island 데이터베이스 구축	농식품 분야 중요 종에서 phylogenomics와 연동하여 genomic island를 예측할 수 있는 알고리즘 및 파이프라인을 개발하고 이를 데이터베이스로 구축	축산 농가에 피해를 끼치는 <i>Vibrio cholerae</i> 에 대해서 computational approach를 활용하여 새로운 genomic island를 찾아냄
	종 내 균주 간 계통과 genomic island의 이동 경로 및 이의 데이터베이스를 구축	genomic island 데이터를 활용하여 데이터베이스 구축
기존 개발된 알고리즘 선행조사 및 기술 확장	해당 연구에 관련된 선행 연구를 결과에 대해 좀 더 조사하고 이를 토대로 참고 및 보완할 점을 살핌	선행조사를 통하여 향후 연구 방향을 가다듬는데 활용
기존 연구 선행 조사 및 마이크로바이옴 서열 분석을 위한 본 연구 기관내의 기 확보 기술 확장 (위탁)	선행 연구 조사	마이크로바이옴 서열 군집화 알고리즘 선행연구 조사
	기 확보 기술 hc-OTU 적용 가능성 및 한계점 파악	각 마이크로바이옴 서열 군집화 알고리즘 한계 파악
	기 확보 기술 적용 가능성 파악	기 확보 기술(hc-OTU)의 선행연구 대비 우수성 파악
	기 확보 기술 적용 한계점 파악	기 확보 기술(hc-OTU)의 마이크로바이옴 데이터 적용의 한계점 파악
	기 확보 개선 및 개발 계획 확립	기 확보 기술의 마이크로바이옴 데이터 적용을 위해 한계점 개선 계획 수립

구분 (연도)	세부과제명	세부연구목표	연구개발 수행내용
2차 년도 (2019)	농식품 중요 종의 pan/core-genome 데이터베이스 구축 및 마이크로바이옴 Gene Catalogue 데이터베이스 구축	농식품 관련 중요 종의 core-genome 데이터베이스 구축, QC 기준 정립	비교유전체를 통한 분류학적 검증은 거친 유전체와 분류체계를 활용한 합리적인 QC기준 제시
		농식품 관련 중요 종에 대한 pan-genome 데이터베이스 구축	농식품 관련 중요 종에 대한 pan-genome 데이터베이스를 만들고, gene catalogue 데이터베이스에 활용할 수 있도록 구축
		농식품 관련 마이크로바이옴 Non-redundant Gene Catalogue 구축	shotgun data와 pan-genome의 유전자를 통합하여, 대규모의 유전자 카탈로그 구축
		작물 병해충과 관련된 <i>Bacillus</i> 의 계통유전체적 관계 재정립	<i>B. cereus</i> 및 <i>B. thuringiensis</i> 유전체에 대한 비교유전체 분석을 통해 생물정보학적 기법으로 유전체 스크리닝 수행
		타 연구과제와의 연계를 위하여 데이터베이스 및 웹서비스 구축	대규모 농축산 마이크로바이옴 데이터베이스 및 웹서비스 제공에 앞서 먼저 쥐 데이터를 사용하여 데이터베이스 및 웹서비스를 구축
		Information theory 및 deep learning	예상 데이터 규모 및 목표 성능치 파악
		정보이론 기반 군집화 기	메타지놈 기 확보 기술구현 및 한계점 파악

기본 메타지놈 분석 알고리즘 연구개발 (위탁)	술 구현	
	정보이론 기반 군집 내 맵핑(분류) 기술 구현	대용량 데이터 적용을 위한 가속기 개발
	맵핑 결과 해석 및 피드백 반영	데이터 확보 및 초기적용
	알고리즘 최종 최적화 및 모듈화	초기 적용 결과 개선 및 최적화

구분 (연도)	세부과제명	세부연구목표	연구개발 수행내용
3차 년도 (2020)	농식품 마이크로바이옴 분석을 위한 메타유전체 및 메타전사체 통합 분석 파이프라인 구축	농식품관련 중요 종 및 중요 유전자 선별 및 pan-genome 데이터베이스 구축	대규모의 gene clustering으로 pan-genome 데이터베이스 구축
		농축산 microbiome data의 Non-redundant Gene Catalogue 구축	Pan-genome을 활용하여 농축산 data에 대한 gene catalogue 구축
		농축산 메타유전체의 정확하고 신속한 분석을 위한 classifier에 적용 가능한 data 확보	농축산 data를 이용한 classifier training
		전체 농축산 microbiome data에 대한 gene profile database 구축	Gene Catalogue로부터 현재까지 공개되고 수집된 농축산 microbiome data 샘플에 대한 gene profile을 생성
		농축산 microbiome data로부터 mycobium profile의 확보	진균의 데이터를 이용해 mycobium의 profile 확보
		16S rRNA 유전자 기반의 농축산 microbiome data에 대한 taxonomic profile database 구축	쥐 microbiome data로 부터 농축산 microbiome data로에 대한 taxonomic 확장
	Information theory 및 deep learning 기반 메타지놈 분석 알고리즘 연구개발 (위탁)	메타지놈 서열 내 잡음 정도 파악 및 목표 성능치 선정	유전체 데이터 확보 및 분석
		오류 정정 이론 인경신경망 함수화	인공신경망 연구 및 적용 방향 모색
		딥러닝 환경 구축	딥러닝 라이브러리 선택 및 환경 설정
		오류 정정 이론 딥러닝 기술 도입	알고리즘 개발 및 적용
	알고리즘 최종 최적화 및 모듈화	주관기관 데이터 확보 및 초기 적용	

구분 (연도)	세부과제명	세부연구목표	연구개발 수행내용
4차 년도 (2021)	메타유전체 통합오믹스 파이프라인 구축	1, 2, 3년 차에 개발된 파이프라인의 고도화를 통한 효율성 증진 및 웹 서비스 제공	한정된 서버 자원을 효율적으로 활용하면서도 정확한 분석을 제공하도록 파이프라인을 개선

		농축산 microbiome의 raw data 보충 및 분류학적, 기능적 profile의 데이터베이스 구축	기존에 수집한 샘플에 더해 추가적으로 데이터 수집하여 profile 생성
		Multiple set의 Profile로부터 바이오마커를 효율적으로 발굴할 수 있도록 하는 데이터베이스 보완	EggNOG, COG 등의 functional annotation 정보를 추가하여 바이오마커 발굴에 활용
		메타유전체/메타전사체/메타단백체/메타대사체 통합 오믹스 파이프라인 구축	메타유전체 데이터와 다른 오믹스 데이터를 비교하여 통합 파이프라인을 구축하고자 하는 계획
		농식품 microbiome의 유전체 비교분석을 위한 플랫폼 프로그램 및 reference dataset 제작	Average amino acid identity 계산 도구, bacterial core gene set 제작
마이크로바이옴 네트워크 분석 파이프라인 구축 및 고도화 (위탁)	확보된 기술 모듈화	Sequence-to-model 기반 알고리즘 모듈화 및 구현	
	초고용량 메타지놈 네트워크 도출	고용량 데이터 확보 및 적용	
	최종 분석 파이프라인 구현	추가 분석 알고리즘 구축	
	협동 연구자 피드백 반영 및 파이프라인 최적화	주관기관 피드백 반영	

3. 연구개발과제의 수행 결과 및 목표 달성 정도

1) 연구수행 결과

(1) 정성적 연구개발성과

○ 연차별 세부 과제들의 연구 목표 및 해당 연구 수행 결과

구분 (연도)	세부과제명	세부연구목표	연구결과
1차년도 (2018)	대규모 마이크로바이옴 데이터베이스 구축	공공 데이터저장소로부터 농식품 관련 마이크로바이옴 NGS 데이터를 확보	마이크로바이옴 염기서열 데이터 및 편집된 환경 데이터를 storage 서버에 저장
		처리 방식 및 메타데이터 정보가 상이한 파편화된 데이터로부터 통일된 포맷의 데이터베이스 구축	메타유전체 서열을 샘플 단위로 합한 후 군집화 작업을 거쳐 분석 수행. 다수의 메타유전체 샘플을 비교분석한 결과를 얻음
	phylogenomics 파이프라인 구축	계통수 작성을 위한 bacterial core gene 기반의 파이프라인을 구축 농식품 관련 속과 종에 대해서 계통수 작성	개선된 파이프라인을 제작하여 프로그램 등록 예정 core gene에 기반한 계통도를 작성
	농식품 관련 종에 대한	농식품 분야 중요 종에서 phylogenomics와 연동하	새롭게 찾아낸 genomic island를 기존에 알려진 genomic island와 함께 반영하여 데이터베이스 구축

genomic island 데이터베이스 구축	여 genomic island를 예측할 수 있는 알고리즘 및 파이프라인을 개발하고 이를 데이터베이스로 구축	
	종 내 균주 간 계통과 genomic island의 이동 경로 및 이의 데이터베이스를 구축	웹사이트 형태로 genomic island 정보 공개
기존 개발된 알고리즘 선행조사 및 기술 확장	해당 연구에 관련된 선행 연구를 결과에 대해 좀 더 조사하고 이를 토대로 참고 및 보완할 점을 살핌	선행조사 결과를 연구에 참여하는 연구원 간 공유하여 향후 연구에 활용 예정
기존 연구 선행 조사 및 마이크로바이 옴 서열 분석을 위한 본 연구 기관내의 기확보 기술 확장 (위탁)	선행 연구 조사	EZPRIT, CD-HIT, UCLUST등의 선행연구 확인
	기확보 기술 hc-OTU 적용 가능성 및 한계점 파악	많은 수의 변수를 필요로 하는 한계점 확인
	기확보 기술 적용 가능성 파악	hc-OTU는 성능 및 속도에서 큰 우위를 가짐
	기확보 기술 적용 한계점 파악	Seq-to-Seq 기반으로 집단의 전체적 특성을 반영하지 못함
	기확보 개선 및 개발 계획 확립	각 마이크로바이옴 집단에 대해 각각 대표 모델을 학습하여, 각 집단의 전체적 특성을 반영할 수 있도록 함

구분 (연도)	세부과제명	세부연구목표	연구결과
2차 년도 (2019)	농식품 중요 종의 pan/core-ge nome 데이터베이스 구축 및 마이크로 바이옴 Gene Catalogue 데이터베이스 구축	농식품 관련 중요 종의 core-genome 데이터베이스 구축, QC 기준 정립	유전체 데이터베이스에서 모아진 유전체들을 completeness와 contamination 등을 고려해 퀄리티 좋은 유전체에 대한 분석 진행
		농식품 관련 중요 종에 대한 pan-genome 데이터베이스 구축	EzBioCloud reference 데이터베이스와 내부에서 확보한 genome data를 이용해 pan-genome 구축하기 위한 파이프라인 작성
		농식품 관련 마이크로바이옴 Non-redundant Gene Catalogue 구축	위탁기관에서 개발한 알고리즘으로 효율성 및 정확성 검증할 예정
		작물 병해충과 관련된 <i>Bacillus</i> 의 계통유전체적 관계 재정립	Bt toxin을 가진 유전체는 다계통적이며 플라스미드를 통해 빈번히 이동함을 확인
	타 연구과제와의 연계를 위하여 데이터베이스 및 웹서비스 구축	NCBI의 Short Read Archive (SRA)로부터 건강한 쥐의 소화기관이나 대변에서 얻은 박테리아 16S rDNA 앰플리콘 데이터를 사용하여 Murine Microbiome Database (MMDB)를 구축하여 웹서비스를 제공	
	Information theory 및 deep learning 기반 메타지놈 분석 알고리즘 연구개발	예상 데이터 규모 및 목표 성능치 파악	예상 데이터 규모 및 연구 진행 계획 수립
		정보이론 기반 군집화 기술 구현	Sequence-to-sequence 접근의 hc-OTU 기술 한계점 파악
정보이론 기반 군집 내 맵핑(분류) 기술 구현		서열 데이터 효율적 분석을 위한 FPGA 기반 가속기 개발	
	맵핑 결과 해석 및 피드백	주관기관과 협의하여 차년도 진행	

(위탁)	반영	주관기관과 협의하여 차년도 진행
	알고리즘 최종 최적화 및 모듈화	

구분 (연도)	세부과제명	세부연구목표	연구결과
3차 년도 (2020)	농식품 마이크로바이 옴 분석을 위한 메타유전체 및 메타전사체 통합 분석 파이프라인 구축	농식품관련 중요 종 및 중요 유전자 선별 및 pan-genome 데이터베이스 구축	총 5억여 개의 gene데이터를 처리하여 pan-genome 데이터베이스 구축
		농축산 microbiome data의 Non-redundant Gene Catalogue 구축	정확한 annotation 정보 전달 가능 및 clustering한 gene 개수 줄일 수 있는 정확하고 효율적인 Gene catalogue 제작
		농축산 메타유전체의 정확한 분석을 위한 classifier에 적용 가능한 data 확보	다양한 species들로부터 short read로 구성된 raw data를 마련
		전체 농축산 microbiome data에 대한 gene profile database 구축	농축산 microbiome data를 분석하여 하나로 모아 데이터베이스로 구축
		농축산 microbiome data로부터 mycobium profile의 확보	진균의 core-genome을 선별하여 mycobium profile을 확보하여 여러 분석 수행
		16S rRNA 유전자 기반의 농축산 microbiome data에 대한 taxonomic profile database 구축	농축산 microbiome data를 16S rRNA유전자 기반으로 세균종의 분포 분석
	Information theory 및 deep learning 기반 메타지놈 분석 알고리즘 연구개발 (위탁)	메타지놈 서열 내 잡음 정도 파악 및 목표 성능치 선정	유전체 데이터 연구 방향 확립
		오류 정정 이론 인경신경망 함수화	오류 및 숨겨진 메시지 검출 알고리즘 개발
		딥러닝 환경 구축	PyTorch 라이브러리 사용 구현
		오류 정정 이론 딥러닝 기술 도입	딥러닝 기반 Malicious DNA 검출 알고리즘 개발
		알고리즘 최종 최적화 및 모듈화	Sequence-to-model기반의 메타지놈 초기 분석

구분 (연도)	세부과제명	세부연구목표	연구결과
4차 년도 (2021)	메타유전체 통합오믹스 파이프라인 구축	1, 2, 3년 차에 개발된 파이프라인의 고도화를 통한 효율성 증진 및 웹 서비스 제공	Pan-genome 데이터베이스를 접목 딥러닝 모델 성능 실험 Gene catalogue 시각화 웹페이지
		농축산 microbiome의 raw data 보충 및 분류학적, 기능적 profile의 데	283개의 샘플을 추가하여 총 1,320개의 raw data 각 샘플 별 taxonomic profile 생성 및 전체 샘플에 대한 functional gene catalogue 구축

		이터베이스 구축	
		Multiple set의 Profile로부터 바이오마커를 효율적으로 발굴할 수 있도록 하는 데이터베이스 보완	Gene catalogue의 각 gene에 EggNOG annotation 및 COG category 정보를 덧붙인 tree 구조의 데이터베이스 생성
		메타유전체/메타전사체/메타단백체/메타대사체 통합 오믹스 파이프라인 구축	메타유전체와 타 오믹스 데이터의 불균형으로 인해 메타유전체에 집중하는 방향으로 변경
		농식품 microbiome의 유전체 비교분석을 위한 플랫폼 프로그램 및 reference dataset 제작	AAI 계산 프로그램 EzAAI 및 개선된 bacterial core gene set인 UBCG2 제작
마이크로바이옴 네트워크 분석 파이프라인 구축 및 고도화 (위탁)	확보된 기술 모듈화	Model building과 Classification 두 단계의 알고리즘 구현	
	초고용량 메타지놈 네트워크 도출	다양한 실험을 통해 알고리즘 성능 향상 확인	
	최종 분석 파이프라인 구현	딥러닝 기반 분석 알고리즘 연구 및 개발	
	협동 연구자 피드백 반영 및 파이프라인 최적화	Species-level 추가 분석	

(2) 정량적 연구개발성과

○ 가. 미생물 유전체사업의 성과목표

성과목표	전략 미생물 해독	유용 유전 자원 확보	표준 유전체 해독	메타 유전체 분석	유전체 분석 기술 개발	NABIC 등록	병원성 미생물 진단마커개발	병원성 미생물 정보 완성	미생물 병발생 기작 규명
최종목표					10			4	
1차년도	목표				1			1	
	실적				0			0	
2차년도	목표			0	3	0		1	
	실적			20	4	22		2	
3차년도	목표				3	0		1	
	실적				3	1		1	
4차년도	목표				3	0		1	
	실적				3	2		2	
계	목표			0	10	0		4	
	실적			20	10	25		5	

○ 나. 기타 성과목표

성과 목표	사업화지표				연구기반지표				
	지식	기술	사업화	기	학술성과	교 인	정책	기	

	재산권			실시 (이전)							술 인 증	논문			학 술 발 표	육 지 도	력 양 성	활용 홍 보		타 (타 연 구 활 용 등)	
	특 허 출 원	특 허 등 록	프 로 그 램 등 록	건 수	기 술 료	제 품 화	매 출 액	수 출 액	고 용 창 출	투 자 유 치		SC I	비 SC I	논 문 평 균 IF				정 책 활 용	홍 보 전 시		
																					건
가중치																					
최종목표		1	6									4	4	2.0	8	4	4				
1차년도	목표														1	1	1				
	실적														8	0	0				
2차년도	목표		2									2	2	2.0	2	1	1				
	실적		4									3	1	3.9	4	2	3				
3차년도	목표		2									1	1	2.0	2	1	1				
	실적		3									3	0	3.2	4	1	1				
4차년도	목표	1	2									1	1	2.0	3	1	1				
	실적	0	3									8	0	6.3	0	2	2				
소계	목표	1	6									4	4	2.0	8	4	4				
	실적	0	10									14	1	5.1	16	5	6				
종료 1차년도			0									0	0		8	0	0				
종료 2차년도			4									3	1		12	2	3				
종료 3차년도			7									6	1		16	3	1				
종료 4차년도			10									14	1		16	5	2				
종료 5차년도																					
소계			10									14	1		16	5	6				
합계			10									14	1		16	5	6				

(3) 세부 정량적 연구개발성과

[미생물유전체사업 성과]

□ 메타유전체 분석

번호	분석대상 (유전체, 유전자원 명칭)	분석내용	등록일자	등록번호	생산량 (GB)
1	도드람 30kg 돼지 분변샘플1	돼지 분변, blood Sample Sequencing	2019.12.16	NG-1255	0.168
2	도드람 30kg 돼지 분변샘플2	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1258	0.19
3	도드람 30kg 돼지 분변샘플3	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1259	0.283
4	도드람 30kg 돼지 분변샘플4	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1260	0.187
5	도드람 30kg 돼지 분변샘플5	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1261	0.173
6	도드람 30kg 돼지 분변샘플6	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1262	0.171
7	도드람 30kg 돼지 분변샘플7	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1263	0.181
8	도드람 30kg 돼지 분변샘플8	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1264	0.165
9	도드람 30kg 돼지 분변샘플9	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1265	0.111
10	도드람 30kg 돼지 분변샘플10	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1266	0.265
11	도드람 90kg 돼지 분변샘플1	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1267	0.128
12	도드람 90kg 돼지 분변샘플2	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1268	0.21
13	도드람 90kg 돼지 분변샘플3	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1269	0.223
14	도드람 90kg 돼지 분변샘플4	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1270	0.195
15	도드람 90kg 돼지 분변샘플5	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1271	0.216
16	도드람 90kg 돼지 분변샘플6	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1272	0.179
17	도드람 90kg 돼지 분변샘플7	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1273	0.203
18	도드람 90kg 돼지 분변샘플8	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1274	0.287
19	도드람 90kg 돼지 분변샘플9	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1275	0.225
20	도드람 90kg 돼지 분변샘플10	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1276	0.265

□ 유전체 분석기술 개발

번호	분석대상 (유전체, 유전자원 명칭)	분석내용	등록일자	등록번호	생산량 (GB)
1	Bacterial genome	단쇄지방산 생성 관련 유전자를 유전체 내에서 탐지 후에 염기서열 추출.	2019.12.02	C-2019-036764	0.0051
2	Bacterial genome	대규모의 전장유전체로부터 genomic island를 찾아낸 후 데이터 베이스에 등록 후 유사한 genomic island를 clustering 함.	2019.12.02	C-2019-036765	0.0225
3	Fungal genome	균류 유전체에서 균류의 core 유전자 포함 여부 및 유전자 서열을 추출.	2019.12.02	C-2019-036766	0.0001

4	Bacterial metagenome	메타유전체 데이터의 QC를 한 후 짧은 NGS서열에 대한 유전체를 assembly하여 단백질로 발현될 가능성이 높은 유전자를 예측	2019.12.02	C-2019-036767	0.0338
5	Bacterial genome	평균 아미노산 유사도 고속 계산 파이프라인 (AAIRE, Average Amino acid Identity Rapid Evaluation pipeline)	2020.10.30	2020-043655	3.75*10 ⁻²
6	Bacterial genome	보존단백질의 비율 계산 파이프라인 (POPCAL, Percentage of Conserved Protein Calculator)	2020.10.30	2020-043656	3.84*e ⁻⁶
7	Bacterial & archaeal genome	UUCGp (Up-to-date Core Genes for Prokaryotes) - 원핵 생물의 계통 분석 도구	2020.10.30	2020-043657	1.57*10 ⁻²
8	Bacterial metagenome	Raw2Diversity (메타지놈 다양성 분석 파이프라인)	2021.09.06	C-2021-035706	3.73*10 ⁻²
9	Bacterial metagenome	GeneCat (유전자 카탈로그 제작 파이프라인)	2021.09.06	C-2021-035707	3.73*10 ⁻²
10	Bacterial metagenome	PanBuilder (범유전체 데이터베이스 제작 파이프라인)	2021.09.06	C-2021-035708	0.2*e ⁻⁴

□ NABIC 등록

번호	분석대상 (유전체, 유전자원 명칭)	분석내용	등록일자	등록번호	생산량 (GB)
1	도드람 30kg 돼지 분변샘플1	돼지 분변, blood Sample Sequencing	2019.12.16	NG-1255	0.168
2	도드람 30kg 돼지 분변샘플2	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1258	0.19
3	도드람 30kg 돼지 분변샘플3	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1259	0.283
4	도드람 30kg 돼지 분변샘플4	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1260	0.187
5	도드람 30kg 돼지 분변샘플5	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1261	0.173
6	도드람 30kg 돼지 분변샘플6	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1262	0.171
7	도드람 30kg 돼지 분변샘플7	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1263	0.181
8	도드람 30kg 돼지 분변샘플8	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1264	0.165
9	도드람 30kg 돼지 분변샘플9	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1265	0.111
10	도드람 30kg 돼지 분변샘플10	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1266	0.265
11	도드람 90kg 돼지 분변샘플1	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1267	0.128
12	도드람 90kg 돼지 분변샘플2	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1268	0.21
13	도드람 90kg 돼지 분변샘플3	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1269	0.223
14	도드람 90kg 돼지 분변샘플4	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1270	0.195
15	도드람 90kg 돼지 분변샘플5	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1271	0.216

16	도드람 90kg 돼지 분변샘플6	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1272	0.179
17	도드람 90kg 돼지 분변샘플7	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1273	0.203
18	도드람 90kg 돼지 분변샘플8	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1274	0.287
19	도드람 90kg 돼지 분변샘플9	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1275	0.225
20	도드람 90kg 돼지 분변샘플10	돼지 분변, blood Sample Sequencing	2019.12.24	NG-1276	0.265
21	Mycobacterium avium NTM-127	해당 균주에 대한 Whole Genome Sequencing	2019.11.07	NG-1228	1.42*10 ⁻⁴
22	Mycobacterium avium NTM-151	해당 균주에 대한 Whole Genome Sequencing	2019.11.07	NG-1229	1.07*10 ⁻⁴
23	Mycolicibacterium sp. SMC_7	해당 균주에 대한 Whole Genome Sequence	2020.10.27	NG-1321	5.91*10 ⁻³
24	폐렴막대균 3293의 전장유전체	해당 균주에 대한 Whole Genome Sequence	2021.11.23	NN-7843	2.74
25	폐렴막대균 R1의 전장유전체	해당 균주에 대한 Whole Genome Sequence	2021.11.23	NN-7845	3

병원성미생물 정보 완성

번호	분석대상 (유전체, 유전자원 명칭)	분석내용	등록일자	등록번호	생산량 (GB)
1	Mycobacterium avium NTM-127	해당 균주에 대한 Whole Genome Sequencing	2019.11.07	NG-1228	1.42*10 ⁻⁴
2	Mycobacterium avium NTM-151	해당 균주에 대한 Whole Genome Sequencing	2019.11.07	NG-1229	1.07*10 ⁻⁴
3	Mycolicibacterium sp. SMC_7	해당 균주에 대한 Whole Genome Sequence	2020.10.27	NG-1321	5.91*10 ⁻³
4	폐렴막대균 3293의 전장유전체	해당 균주에 대한 Whole Genome Sequence	2021.11.23	NN-7843	2.74
5	폐렴막대균 R1의 전장유전체	해당 균주에 대한 Whole Genome Sequence	2021.11.23	NN-7845	3

[과학적 성과]

논문(국내외 전문 학술지) 게재

번호	논문명	학술지명	주저자명	호	국명	발행기관	SCIE 여부 (SCIE/BIOSIS/CAS/ISI)	게재일	등록번호 (ISSN)
----	-----	------	------	---	----	------	----------------------------------	-----	----------------

1	Comparative genomic and phylogenomic analyses clarify relationships within and between Bacillus cereus and Bacillus thuringiensis: proposal for the recognition of two Bacillus thuringiensis genomovars	Frontiers in Microbiology	Inwoo Baek	10	Switzerland	Frontiers	SCI	2019.08	10.3389/fmicb.2019.01978
2	Introducing Murine Microbiome Database (MMDB): A Curated Database with Taxonomic Profiling of the Healthy Mouse Gastrointestinal Microbiome	Microorganisms	Junwon Yang	7(11)	Switzerland	MDPI	SCI	2019.10	10.3390/microorganisms7110480
3	Energy-Efficient Inference Accelerator for Memory-Augmented Neural Networks on an FPGA	2019 Design, Automation & Test in Europe Conference & Exhibition (DATE)	Seongsik Park		Europe	Design, Automation & Test in Europe Conference & Exhibi	SCI	2019.05	10.23919/DATE.2019.8715013
4	VCGIDB: a database and web resource for the genomic islands from Vibrio cholerae	Pathogens	Youngjae Hur	8(4)	Switzerland	MDPI	SCI	2019.11	10.3390/pathogens8040261
5	AnomiGAN: Generative adversarial networks for anonymizing private medical data	Pacific Symposium on Biocomputing (PSB)	Ho Bae	25	USA	Biocomputing	SCI	2020.01	10.1142/9789811215636_0050

6	DNA Privacy: Analyzing Malicious DNA Sequences using Deep Neural Networks	IEEE Transactions on Computational Biology and Bioinformatics	Ho Bae	5	USA	Gene	SCI	2020.08	10.1109/TCBB.2020.3017191
7	Metagenomic association analysis of gut symbiont <i>Limosilactobacillus reuteri</i> without host-specific genome isolation	Frontiers in Microbiology-Microbial Symbioses	Sein Park	11	Switzerland	Frontiers	SCI	2020.11	10.3389/fmicb.2020.585622
8	Taxonomic composition and variation in the gut microbiota of laboratory mice	Mammalian Genome	Junwon Yang	32	USA	Springer	SCI	2021.04	10.1007/s00335-021-09871-7
9	Introducing EzAAI: a pipeline for high throughput calculations of prokaryotic average amino acid identity	Journal of Microbiology	Dongwook Kim, Sein Park	59	Korea	Springer	SCI	2021.04	10.1007/s12275-021-1154-0
10	UBCG2: Up-to-date bacterial core genes and pipeline for phylogenomic analysis	Journal of Microbiology	Jihyeon Kim, Seong-In Na	59	Korea	Springer	SCI	2021.05	10.1007/s12275-021-1231-4
11	Protein transfer learning improves identification of heat shock protein families	PLOS ONE	Seonwo Min	16	USA	PLoS	SCI	2021.05	10.1371/journal.pone.0251865
12	VicPred: A <i>Vibrio cholerae</i> Genotype Prediction Tool	Frontiers in Microbiology	Imchang Lee	12	Switzerland	Frontiers	SCI	2021.08	10.3389/fmicb.2021.691895
13	Pre-training of deep bidirectional protein sequence representations with structural information	IEEE Access	Seonwo Min	9	USA	IEEE	SCI	2021.09	10.1109/ACCESS.2021.3110269

14	TargetNet: functional microRNA target prediction with deep neural networks	Bioinformatics	Seonwo Min	In Press	UK	Oxford University Press	SCI	2021. 10	10.1093/bioinformatics/btab733
15	AMDB: a database of animal gut microbial communities with manually curated metadata	Nucleic Acid Research	Junwon Yang		United Kingdom	Oxford Academic	SCI	2021. 11	10.1093/nar/gkab1009

□ 국내 및 국제 학술회의 발표

번호	회의명칭	발표자	발표일시	장소	국명
1	ASM microbe 2018	허영재	2018.06.10	Atlanta	미국
2	2018 한국실험동물학회 국제학술대회	천종식	2018.07.18	부산 BEXCO 컨벤션홀 Hall A	한국
3	2018년도 한국미생물학회연합 국제학술대회	천종식	2018.10.11	더케이호텔서울	한국
4	2018 한국영양학회 추계학술대회	천종식	2018.10.18	알펜시아 컨벤션센터 (강원도 평창)	한국
5	2018 대한생물정신의학회 추계학술대회	천종식	2018.10.26	백범김구기념관	한국
6	한국식품영양과학회 정기학술대회	천종식	2018.10.31	부산 BEXCO	한국
7	2018 대한치매학회 추계학술대회	천종식	2018.11.03	백범김구기념관	한국
8	Understanding human gut microbiome and diseases	천종식	2018.11.16	코엑스	한국
9	한국미생물생명공학회 동계심포지엄	백인우	2019.01.15	용평리조트	한국
10	ASM microbe 2019	김지현	2019.06.21	San Francisco	미국
11	ASM microbe 2019	백인우	2019.06.23	San Francisco	미국
12	Proceedings of Design, Automation and Test in Europe (DATE)	박성식	2019.03.28	Florence	이탈리아
13	Pacific Symposium on Biocomputing	배호	January 2020.	Hawaii	USA
14	한국미생물 생명공학회	양준원	2020.01.13	용평리조트 드래곤밸리호텔	대한민국
15	ACKSS2020	천종식	2020.11.05	서울 홍은동 스투아 호텔	대한민국
16	KDDW2020	천종식	2020.11.20	Online Conference	대한민국

[기술적 성과]

□ 저작권(소프트웨어, 서적 등)

No	저작권명	창작일	저작자명	등록일	등록번호	저작권자명	기여율
1	유전체 내 단쇄지방산 생성 관련 유전자 탐지 프로그램	2019.11. 08	백인우, 천종식	2019.12.02	C-2019-0367 64	서울대학교 산학협력단	100%
2	GI-Scanner - 대규모 유전체 데이터베이스기반 genomic island(유전체 섬) 발굴 프로그램	2019.11. 08	허영재, 천종식	2019.12.02	C-2019-0367 65	서울대학교 산학협력단	100%
3	UUCGf: 균류 유전체 내 core-gene(핵심유전 자) 탐색 파이프라인	2019.11. 08	김동욱, 천종식	2019.12.02	C-2019-0367 66	서울대학교 산학협력단	100%
4	메타유전체 유전자 분석 도구	2019.11. 08	박세인, 천종식	2019.12.02	C-2019-0367 67	서울대학교 산학협력단	100%
5	평균 아미노산 유사도 고속 계산 파이프라인 (AAIRE, Average Amino acid Identity Rapid Evaluation pipeline)	2020.10. 30	김동욱, 천종식	2020.10.30	C-2020-0432 52	서울대학교 산학협력단	100%
6	보존단백질의 비율 계산 파이프라인 (POPCAL, Percentage of Conserved Protein Calculator)	2020.10. 30	김지현, 천종식	2020.10.30	C-2020-0432 53	서울대학교 산학협력단	100%
7	UUCGp (Up-to-date Universal Core Genes for Prokaryotes) - 원핵 생물의 계통 분석 도구	2020.10. 30	나성인, 천종식	2020.10.30	C-2020-0432 54	서울대학교 산학협력단	100%
8	Raw2Diversity (메타 지놈 다양성 분석 파 이프라인)	2021.08. 27	양준원, 천종식	2021.09.06	C-2021-0357 06	서울대학교 산학협력단	100%
9	GeneCat (유전자 카 탈로그 제작 파이프 라인)	2021.08. 27	박세인, 천종식	2021.09.06	C-2021-0357 07	서울대학교 산학협력단	100%
10	PanBuilder (범유전 체 데이터베이스 제 작 파이프라인)	2021.08. 27	박세인, 천종식	2021.09.06	C-2021-0357 08	서울대학교 산학협력단	100%

[사회적 성과]

전문 연구 인력 양성

번호	분류	기준 연도	현 황											
			학위별				성별		지역별					
			박사	석사	학사	기타	남	여	수도권	충청권	영남권	호남권	기타	
1	이임창	2019	1				1		1					
2	백인우	2019	1				1					1		
3	한재흥	2019		1			1		1					
4	허영재	2020	1				1		1					
5	민선우	2021	1				1		1					
6	조정희	2021	1					1	1					

[그 밖의 성과]

□ 교육 지도

No	교육 및 컨설팅명	참석대상	인원	교육기간
1	Next Generation Sequencing: New Diagnostic Tool for Pediatric Infectious Diseases	연구실 학생	7	2019.05.08
2	Effective intestinal microbiome strategy for obesity treatment	연구실 학생	8	2019.08.02
3	그린바이오(마이크로바이옴 분야) 포럼	농림부 식품산업정책과, 농진청, 관련 연구자, 기업 등	20	2020.10.16
4	서울대학교 행정대학원 공공리더십과정	공무원/공공기관 직원	30	2021.10.05
5	2021 서울대학교 푸드테크 최고책임자과정	최고책임자과정	20	2021.10.21

2) 목표 달성 수준

추진 목표	달성 내용	달성도(%)
○ Information theory 및 deep learning 기반 메타지놈 분석을 위한 기반기술 고안 및 고도화	○ 기확보 기술의 마이크로바이옴 데이터 적용을 위해 한계점 개선	○ 100
○ 농식품 유용 미생물 종의 phylogenomics 파이프라인 구축	○ 세균 및 고세균에 대해서 core gene을 선별하고 그 기반의 분류 파이프라인 구축	○ 100
○ 농식품 유용 미생물 종의 비교 유전체 및 pan/core genome database 구축	○ 비교유전체를 통한 분류학적 검증을 거친 유전체와 분류체계를 활용한 합리적인 QC기준 제시 및 농식품 관련 중요 종에 대한 pan-genome 데이터베이스를 만듦	○ 100
○ 농식품 유용 미생물 종의 genomic island database 구축	○ 오염된 축산 식품 등에서 발견되어 피해를 끼치는 <i>Vibrio cholerae</i> 에 대해서 computational approach를 활용하여 새로운 genomic island를 찾아냄	○ 100
○ 표준화되고, big data 분석이 가능한 농식품 관련 대규모 마이크로바이옴 16S 및 shotgun database 구축	○ 쥐뿐만 아니라 다른 가축들로부터 16S rRNA 유전자 기반의 농축산 microbiome data에 대한 taxonomic profile database 구축하여 세균종의 분포 분석.	○ 100
○ 축산 (돼지, 소, 닭), 전통 발효 식품의 미생물 gene catalog	○ http://leb.snu.ac.kr/amdb 에서 검색 가능.	

<p>database 구축</p> <ul style="list-style-type: none"> ○ 메타지놈 데이터 분석 ○ 마이크로바이옴 네트워크 분석 알고리즘 및 소프트웨어 개발 	<ul style="list-style-type: none"> ○ shotgun data와 pan-genome의 유전자를 통합하여, 정확한 annotation 정보 전달 및 clustering 가능한 효율적인 대규모의 유전자 카탈로그 구축. ○ 메타지놈 데이터를 통해 Genus-level 분류를 진행했으며 CDS sequence를 Pan-genome DB에 맵핑하기 위해 Species-level 분류를 추가로 진행함. ○ Model building과 Classification 두 단계의 알고리즘 구현 및 알고리즘 성능향상 확인 	
--	--	--

4. 목표 미달 시 원인분석

1) 목표 미달 원인(사유) 자체분석 내용

본 과제의 세부 목표 중 달성하지 못한 사항들에 대해 다음과 같이 분석함.

○ 특허 등록에 대한 건

- 특허 등록 일정과 본 과제 주요 결과물의 산출 시기를 함께 고려하지 못한 것이 가장 주된 사유로 분석됨.

- 본 보고서 3-(2) 정량적 연구개발성과에서 알 수 있듯 3차년도 후반기 ~ 4차년도에 걸쳐 대부분의 성과물이 나왔는데, 일반적으로 특허 등록에 1년 반에서 2년 이상의 시간이 필요하다는 점에서 결과물들을 과제 목표에 맞게 등록까지 마치기 곤란하였음.

- 또한 본 연구의 결과물들이 새로운 기술의 개발보다는 기존에 공개된 다양한 농축산 관련 데이터를 수집 및 가공하여 데이터베이스를 만들고 기존의 분석 도구들을 조합하여 필요한 정보들을 효율적으로 서비스하는 것에 더 초점을 맞추고 있었으므로, 이를 특허 등록의 형태보다는 논문 저술과 같은 형태로 목표를 설정했으면 더 적절했을 것으로 생각됨.

○ 기타 연구 목표에 대한 건

- 목표 설정 당시 멀티오믹스 데이터를 다방면으로 수집해 네트워크 분석을 하고자 하였으나, 이 때 기대와는 달리 시간이 지나도 메타유전체 데이터에 비해 메타전사체 등 다른 종류의 데이터의 양이 부족하여 이를 동등하게 같은 선에 두고 분석하기는 오히려 정확도가 떨어질 것으로 생각되었음.

- 이에 따라 유전체 및 메타유전체 데이터에 집중해 더욱 세밀한 분석이 가능한 데이터베이스를 제작하기로 하였음.

2) 자체 보완활동

본 과제가 최초로 설정한 목표에 도달하지 못한 위 사항들에 대해 다음과 같이 보완하였음.

- 관련 결과물로서 나온 데이터베이스 및 이를 활용한 분석에 대하여 논문 발표를 통해 공개하였음

- 지식재산권 중 프로그램 등록을 통해 본 과제가 목표로 한 분석 과정을 실행할 수 있는 프로그램을 추가적으로 등록하였음

3) 연구개발 과정의 성실성

- 정량 목표의 추가 달성

- 미생물 유전체사업의 성과목표 중 하나인 “병원성 미생물 정보 완성”에 대해 목표 4건, 실적 5건의 추가 실적을 내었으며, 성과목표에 포함되지 않았던 항목인 “메타유전체 분석” 및 “NABIC 등록”에 대해서도 각각 20건, 25건의 실적을 달성함.
- 기타 성과목표에서도 “프로그램 등록” 목표 6건, 실적 10건, “논문” 목표 SCI/비SCI 총 8건, 실적 SCI 14건을 비롯하여 총 15건 등 미달 항목인 “특허 등록”을 제외한 모든 부분에서 높은 추가 달성율을 기록함.

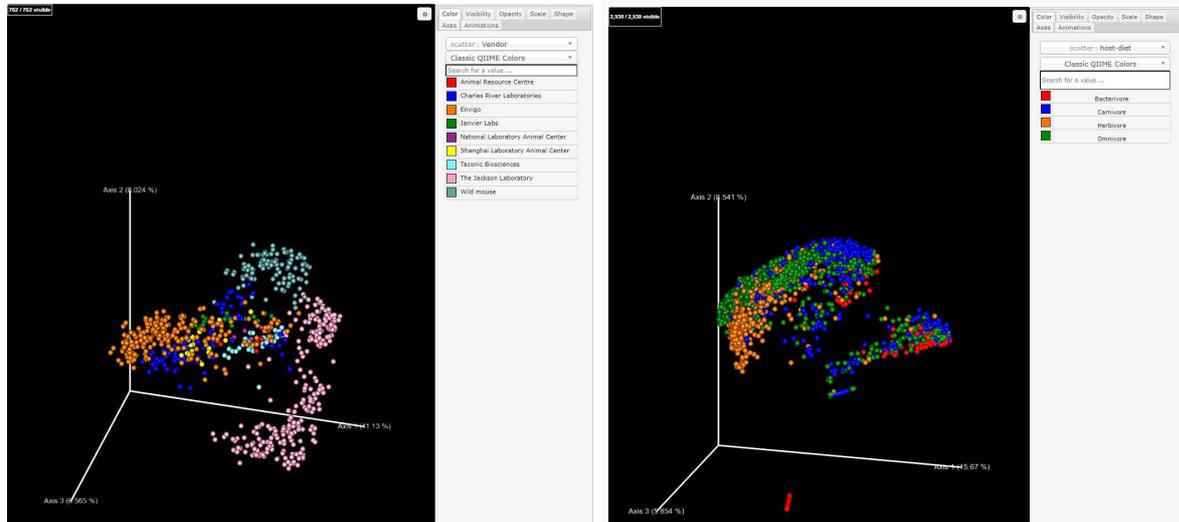
- 빅데이터 수집 및 분석을 통한 다양한 데이터베이스 개발

- 정량 목표가 아니더라도 본 연구에서 데이터베이스들의 개발 과정에서 메타유전체에 집중하여 대규모의 데이터 수집 및 분석이 이루어졌음.
- 예를 들어 농축산 마이크로바이옴 shotgun 메타유전체 데이터의 경우 총 1,320개의 샘플을 수집하여 약 2억 8천만 개의 유전자 서열을 얻었으며 이로부터 4천만 여 개의 유전자 클러스터를 추출하여 gene catalog를 구성함.
- 위 gene catalog를 만드는데 보조적 역할을 한 pan-genome 데이터베이스는 20만 여 개의 미생물 유전체 데이터를 사용하여 5억 여 개의 유전자 데이터를 가지고 있음.
- 또한 세부적으로 메타데이터를 정리하여 만든 데이터베이스인 AMDB의 경우 농축산 관련 동물을 포함한 총 467 종의 동물로부터 2,530 개의 장내 미생물 데이터 샘플을 수집하였고 이를 웹서비스의 형태로 공개함.

5. 연구개발성과의 관련 분야에 대한 기여 정도

- 다양한 데이터베이스 및 관련 정보에 대한 웹서비스

- 본 과제의 연구 결과로서 논문이 발표된 데이터베이스는 모두 웹에서 접근 가능함.
- 대표적으로 아래의 두 그림의 경우 각각 MMDB와 AMDB의 웹 페이지에서 제공하는 일부 정보를 보여주며, 이렇게 데이터베이스가 가지고 있는 마이크로바이옴 자료의 메타데이터 뿐 아니라 이를 이용해 분석한 결과들을 시각적으로 보여줌으로써 연구자들이 쉽게 이해하고 사용할 수 있게 함.



○ 미생물 유전체 분석 도구 공개

- 본 과제의 결과물 중 유전체 분석 도구와 reference 자료의 다수는 논문과 웹사이트에 공개되어있음.
- 종 단위 이상에서 세균 유전체의 분류군 간 비교를 위한 도구인 EzAAI와 세균 유전체의 core gene set인 UBCG2가 대표적 사례임.
- UBCG2의 경우 전체 박테리아 종에서 95% 이상 single copy gene으로 가지고 있는 유전자들을 core gene set에 포함 시켰음.
- EzAAI의 경우 아래와 같이 전 세계에서 꾸준히 사용되고 있음.



6. 연구개발성과의 관리 및 활용 계획

- 위에서 설명한 웹에 공개된 데이터베이스와 다양한 분석 도구 및 자료들은 차후에도 접근 가능하도록 유지하여 연구자들이 사용할 수 있도록 할 계획임.

< 별첨 자료 >

중앙행정기관 요구사항	별첨 자료
1.	1) 자체평가의견서 2) 연구성과 활용계획서
2.	1) 2)

[뒷면지]

주 의

1. 이 보고서는 농림축산식품부에서 시행한 ○○연구개발사업 ○○연구개발과제 최종보고서이다.
2. 이 연구개발내용을 대외적으로 발표할 때에는 반드시 농림축산식품부에서 시행한 ○○연구개발사업의 결과임을 밝혀야 한다.
3. 국가과학기술 기밀 유지에 필요한 내용은 대외적으로 발표 또는 공개하여서는 안 된다.

자체평가의견서

1. 과제현황

		과제번호		918013-04	
사업구분	포스트게놈 다부처 유전체사업				
연구분야				과제구분	단위 [O]
사업명	포스트게놈 다부처 유전체사업				주관
총괄과제	기재하지 않음			총괄책임자	기재하지 않음
과제명	농림축산식품 분야를 위한 메타유전체의 통합 분석을 위한 데이터베이스 및 소프트웨어 개발			과제유형	기초
연구개발기관	서울대학교 산학협력단			연구책임자	천종식
연구기간 연구개발비 (천원)	연차	기간	정부	민간	계
	1차년도	2018. 04. 25 - 2018. 12. 31	150,000		150,000
	2차년도	2019. 01. 01 - 2019. 12. 31	200,000		200,000
	3차년도	2020. 01. 01 - 2020. 12. 31	200,000		200,000
	4차년도	2021. 01. 01 - 2021. 12. 31	200,000		200,000
	계		750,000		750,000
	참여기업				
상대국		상대국연구개발기관			

2. 평가일 : 2022. 02. 07

3. 평가자(연구책임자) :

소속	직위	성명
서울대학교	교수	천종식

4. 평가자(연구책임자) 확인 :

본인은 평가대상 과제에 대한 연구결과에 대하여 객관적으로 기술하였으며, 공정하게 평가하였음을 확약하며, 본 자료가 전문가 및 전문기관 평가 시에 기초자료로 활용되기를 바랍니다.

확 약	
-----	---

자체평가의견서

1. 과제현황

		과제번호		918013-04	
사업구분	포스트게놈 다부처 유전체사업				
연구분야				과제구분	단위 [O]
사업명	포스트게놈 다부처 유전체사업				주관
총괄과제	기재하지 않음			총괄책임자	기재하지 않음
과제명	농림축산식품 분야를 위한 메타유전체의 통합 분석을 위한 데이터베이스 및 소프트웨어 개발			과제유형	기초
연구개발기관	서울대학교 산학협력단			연구책임자	천종식
연구기간 연구개발비 (천원)	연차	기간	정부	민간	계
	1차년도	2018. 04. 25 - 2018. 12. 31	150,000		150,000
	2차년도	2019. 01. 01 - 2019. 12. 31	200,000		200,000
	3차년도	2020. 01. 01 - 2020. 12. 31	200,000		200,000
	4차년도	2021. 01. 01 - 2021. 12. 31	200,000		200,000
	계		750,000		750,000
	참여기업				
상대국			상대국연구개발기관		

2. 평가일 : 2022. 02. 07

3. 평가자(연구책임자) :

소속	직위	성명
서울대학교	교수	천종식

4. 평가자(연구책임자) 확인 :

본인은 평가대상 과제에 대한 연구결과에 대하여 객관적으로 기술하였으며, 공정하게 평가하였음을 확약하며, 본 자료가 전문가 및 전문기관 평가 시에 기초자료로 활용되기를 바랍니다.

확약	
----	--

I. 연구개발실적

※ 다음 각 평가항목에 따라 자체평가한 등급 및 실적을 간략하게 기술(200자 이내)

1. 연구개발결과의 우수성/창의성

■ 등급 : 우수

마이크로바이옴 분석을 위한 다양한 데이터베이스 및 프로그램, 그리고 분석 결과를 논문 등을 통해 공개하였음. 데이터베이스의 경우 쥐의 장내에 존재하는 다양한 미생물 종의 특징을 상세하게 정리한 MMDB, 이를 확장하여 467 종의 동물에서 각각 나타나는 미생물을 정리한 AMDB가 있으며, 세균 유전체 간 차이를 측정하기 위해 평균 아미노산의 유사도를 계산하는 프로그램인 EzAAI와 전체 세균 종에서 공통적으로 나타나 계통 분석에 활용할 수 있는 유전자 세트인 UBCG2도 본 연구의 결과로 발표하였음.

2. 연구개발결과의 파급효과

■ 등급 : 보통

논문을 통해 발표된 데이터베이스 및 프로그램은 현재 웹에서 접근 가능하며, 이들은 마이크로바이옴 분석에 유용하게 쓰일 수 있음. 예를 들어 EzAAI의 경우 한국 뿐 아니라 중국, 미국, 인도 등 다양한 국가의 학자들이 다운로드 받아 세균 유전체 연구에 활용하고 있음.

3. 연구개발결과에 대한 활용가능성

■ 등급 : 우수

상술하였듯 본 연구과제의 주요 결과물들이 웹 상에 공개되어 있기 때문에 목적에 맞는 데이터베이스와 프로그램을 자유롭게 활용할 수 있음. 예를 들어 AMDB의 경우 다양한 동물 종의 마이크로바이옴을 상세히 다루고 있기 때문에 하나의 동물 종에 대해 연구하던 학자들이 비슷한 식성을 가진 다른 동물 종과 비교하는 방식으로 활용될 수 있음.

4. 연구개발 수행노력의 성실도

■ 등급 : 보통

본 연구과제에서 계획하였던 정량적 실적은 특허 등록 1건을 제외하면 모두 달성하였으며, 이 중에서도 논문 등록이나 학술 발표, 프로그램 등록 등 주요 실적은 초과 달성하였음. 또한 메타유전체 분석이나 NABIC 등록과 같이 처음에 계획하지 않았던 실적들도 각각 20, 22건 달성함으로써 부족함이 다소 있던 부분을 보완하고자 하였음.

5. 공개발표된 연구개발성과(논문, 지적소유권, 발표회 개최 등)

■ 등급 : 우수

논문의 경우 SCI 4건, 비 SCI 4건 발표를 계획하였으나, SCI 논문만 14건을 게재하여 총 15건을 달성하였음. 프로그램 등록 또한 계획한 6건을 초과하는 10건을 등록하였고, 학술 발표도 8회를 계획하였으나 총 16회의 발표를 진행함.

II. 연구목표 달성도

세부연구목표 (연구계획서상의 목표)	비중 (%)	달성도 (%)	자체평가
Information theory 및 deep learning 기반 메타지놈 분석을 위한 기반기술 고안 및 고도화	15	100	기확보 기술의 마이크로바이옴 데이터 적용을 위해 한계점 개선
농식품 유용 미생물 종의 phylogenomics 파이프라인 구축	15	100	세균 및 고세균에 대해서 core gene을 선별하고 그 기반의 분류 파이프라인 구축
농식품 유용 미생물 종의 비교 유전체 및 pan/core genome database 구축	15	100	비교유전체를 통한 분류학적 검증을 거친 유전체와 분류체계를 활용한 합리적인 QC기준 제시 및 농식품 관련 중요 종에 대한 pan-genome 데이터 베이스를 만들
농식품 유용 미생물 종의 genomic island database 구축	15	100	오염된 축산 식품 등에서 발견되어 피해를 끼치는 <i>Vibrio cholerae</i> 에 대해서 computational approach를 활용하여 새로운 genomic island를 찾아냄
표준화되고, big data 분석이 가능한 농식품 관련 대규모 마이크로바이옴 16S 및 shotgun database 구축	15	100	쥐뿐만 아니라 다른 가축들로부터 16S rRNA유전자 기반의 농축산 microbiome data에 대한 taxonomic profile database 구축하여 세균종의 분포 분석
축산 (돼지,소,닭), 전통 발효 식품의 미생물 gene catalog database 구축	15	100	Shotgun data와 pan-genome의 유전자를 통합하여, 정확한 annotation 정보 전달 및 clustering 가능한 효율적인 대규모의 유전자 카탈로그 구축
마이크로바이옴 네트워크 분석 알고리즘 및 소프트웨어 개발	10	100	Model building과 Classification 두 단계의 알고리즘 구현 및 알고리즘 성능향상 확인
합계	100점		

III. 종합의견

1. 연구개발결과에 대한 종합의견

본 연구과제에서 계획했던 사항들을 성실히 수행하여 논문, 프로그램 등록, 학술 발표 등의 방식으로 다양한 결과물이 공개되었음. 정량적으로 일부 미흡한 부분이 있지만 이를 보완하기 위해 다른 부분에서 추가적으로 결과물을 만들어 내었으며, 정성적인 측면에서도 연차별 세부적인 과제 목표들을 각각 적절한 방식으로 수행하였음.

2. 평가시 고려할 사항 또는 요구사항

3. 연구결과의 활용방안 및 향후조치에 대한 의견

데이터베이스, 프로그램과 같이 본 연구과제의 주요 결과물들은 모두 논문에 기재된 대로 웹을 통해 공개되어있기 때문에 전 세계의 관련 분야 학자들이 접근하여 활용할 수 있음.

IV. 보안성 검토

o 연구책임자의 보안성 검토의견, 연구개발기관 자체의 보안성 검토결과를 기재함

※ 보안성이 필요하다고 판단되는 경우 작성함.

1. 연구책임자의 의견

해당 사항 없음

2. 연구개발기관 자체의 검토결과

연구성과 활용계획서

1. 연구과제 개요

사업추진형태	<input type="checkbox"/> 자유응모과제 <input type="checkbox"/> 지정공모과제	분 야		
연구과제명	농림축산식품 분야를 위한 메타유전체의 통합 분석을 위한 데이터베이스 및 소프트웨어 개발			
주관연구개발기관	서울대학교 산학협력단		주관연구책임자	천종식
연구개발비	정부지원 연구개발비	기관부담연구개발비	기타	총연구개발비
	750,000			750,000
연구개발기간	2018. 04. 25 - 2021. 12. 31			
주요활용유형	<input type="checkbox"/> 산업체이전 <input type="checkbox"/> 교육 및 지도 <input type="checkbox"/> 정책자료 <input checked="" type="checkbox"/> 기타(학술적 활용)			

2. 연구목표 대비 결과

당초목표	당초연구목표 대비 연구결과
Information theory 및 deep learning 기반 메타지놈 분석을 위한 기반기술 고안 및 고도화	기확보 기술의 마이크로바이옴 데이터 적용을 위해 한계점 개선
농식품 유용 미생물 종의 phylogenomics 파이프라인 구축	세균 및 고세균에 대해서 core gene을 선별하고 그 기반의 분류 파이프라인 구축
농식품 유용 미생물 종의 비교 유전체 및 pan/core genome database 구축	비교유전체를 통한 분류학적 검증을 거친 유전체와 분류체계를 활용한 합리적인 QC기준 제시 및 농식품 관련 중요 종에 대한 pan-genome 데이터베이스를 만들
농식품 유용 미생물 종의 genomic island database 구축	오염된 축산 식품 등에서 발견되어 피해를 끼치는 <i>Vibrio cholerae</i> 에 대해서 computational approach를 활용하여 새로운 genomic island를 찾아냄
표준화되고, big data 분석이 가능한 농식품 관련 대규모 마이크로바이옴 16S 및 shotgun database 구축	쥐뿐만 아니라 다른 가축들로부터 16S rRNA유전자 기반의 농축산 microbiome data에 대한 taxonomic profile database 구축하여 세균종의 분포 분석
축산 (돼지,소,닭), 전통 발효 식품의 미생물 gene catalog database 구축	Shotgun data와 pan-genome의 유전자를 통합하여, 정확한 annotation 정보 전달 및 clustering 가능한 효율적인 대규모의 유전자 카탈로그 구축
마이크로바이옴 네트워크 분석 알고리즘 및 소프트웨어 개발	Model building과 Classification 두 단계의 알고리즘 구현 및 알고리즘 성능향상 확인

3. 연구목표 대비 성과

성과목표		전략 미생물 해독	유용 유전 자원 확보	표준 유전체 해독	메타 유전체 분석	유전체 분석 기술 개발	NABIC 등록	병원성 미생물 진단마 커개발	병원성 미생물 정보 완성	미생물 병발생 기작 규명
최종목표						10			4	
1차 년도	목표					1			1	
	실적					0			0	
2차 년도	목표					3			1	
	실적				20	4	22		2	
3차 년도	목표					3			1	
	실적					3	1		1	
4차 년도	목표					3			1	
	실적					3	2		2	
계	목표					10			4	
	실적				20	10	25		5	

(단위 : 건수, 백만원, 명)

성과 목표	사업화지표										연구기반지표									
	지식 재산권			기술 실시 (이전)		사업화					기술 인증	학술성과			교육 지도	인력 양성	정책 활용 홍 보		기 타 (타 연 구 활 용 등)	
	특 허 출 원	특 허 등 록	프 로 그 램 등 록	건 수	기 술 료	제 품 화	매 출 액	수 출 액	고 용 창 출	투 자 유 치		기 술 인 증	논문				학 술 발 표	정 책 활 용		홍 보 전 시
													SC I	비 SC I						
단위	건	건	건	건	백 만 원	백 만 원	백 만 원	백 만 원	명	백 만 원	건	건	건	건	명	건	건			
가중치																				
최종목 표		1	6									4	4	2.0	8	4	4			
1 차 년 도	목 표														1	1	1			
	실 적														8	0	0			
2 차 년 도	목 표			2								2	2	2.0	2	1	1			
	실 적			4								3	1	3.9	4	2	3			
3 차 년 도	목 표			2								1	1	2.0	2	1	1			
	실 적			3								3	0	3.2	4	1	1			
4	목		1	2								1	1	2.	3	1	1			

차 년 도	표 실 적	0	3								8	0	0	6.	0	2	2		
	0	3												3	0	2	2		
소 계	목 표	1	6								4	4	2.	8	4	4			
	실 적	0	10								14	1	5.	16	5	6			
													1						

4. 핵심기술

구분	핵심기술명
①	
②	
③	

5. 연구결과별 기술적 수준

구분	핵심기술 수준					기술의 활용유형(복수표기 가능)				
	세계 최초	국내 최초	외국기술 복 제	외국기술 소화·흡수	외국기술 개선·개발	특허 출원	산업체이전 (상품화)	현장애로 해 결	정책 자료	기타
①의 기술										
②의 기술										
③의 기술										
·										
·										

* 각 해당란에 v 표시

6. 각 연구결과별 구체적 활용계획

핵심기술명	핵심기술별 연구결과활용계획 및 기대효과
①의 기술	
②의 기술	
③의 기술	

7. 연구종료 후 성과창출 계획

성과목표	전략 미생물 해독	유용 유전 자원 확보	표준 유전체 해독	메타 유전체 분석	유전체 분석 기술 개발	NABIC 등록	병원성 미생물 진단마 커개발	병원성 미생물 정보 완성	미생물 병발생 기작 규명
최종목표					10			4	
연구기간내 달성실적				20	10	25		5	
연구종료후 성과창출 계획									

(단위 : 건수, 백만원, 명)

성과 목표	사업화지표											연구기반지표								
	지식 재산권				기술 실시 (이전)		사업화					기술 인증	학술성과			교육 지도	인력 양성	정책 활용·홍보		기타 (타 연구 활용 등)
	특허 출원	특허 등록	품종 등록	S M A R T	건수	기술료	제품화	매출액	수출액	고용 창출	투자유치		논문 SCI	비SCI	논문 평균 IF			학술 발표	정책 활용	
단위	건	건	건	건	건	백만원	건	백만원	백만원	명	백만원	건	건	건	건	명	건	건		
가중치																				
최종목표																				
연구기간 내 달성실적																				
연구종료 후 성과장출 계획																				

성과 목표	사업화지표											연구기반지표								
	지식 재산권				기술 실시 (이전)		사업화					기술 인증	학술성과			교육 지도	인력 양성	정책 활용·홍보		기타 (타 연구 활용 등)
	특허 출원	특허 등록	프로그램 등록	건수	기술료	제품화	매출액	수출액	고용 창출	투자유치	논문 SCI		비SCI	논문 평균 IF	학술 발표			정책 활용	홍보 전시	
단위	건	건	건	건	백만원	백만원	백만원	백만원	명	백만원	건	건	건	건	명	건	건			
가중치																				
최종목표		1	6									4	4	2.0	8	4	4			
연구기간 내 달성실적		0	10									14	1	5.1	16	5	6			
연구종료 후 성과장출 계획																				

8. 연구결과의 기술이전조건(산업체이전 및 상품화연구결과에 한함)

주 의

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