

극지유전체 101 프로젝트: 극지생물 유전체
정보 분석 및 활용기반 구축

Polar Genome 101 Project: Genome analysis of polar
organisms and establishment of application platform



2020. 02. 28

한 국 해 양 과 학 기 술 원
부 설 극 지 연 구 소

제 출 문

극지연구소장 귀하

본 보고서를 “극지유전체 101 프로젝트: 극지생물 유전체 정보 분석 및 활용기반 구축” 과제의 최종 보고서로 제출합니다.



연구 책임자 : 김진형

참여 연구원 : 이준혁

〃 : 김한우

〃 : 이형석

〃 : 이정은

〃 : 강승현

〃 : 김보미

보고서 초록

과제관리번호		해당단계 연구기간	2017. 01. 01. ~ 2019. 12. 31.	단계 구분	(2단계) / (총3단계)
연구사업명	중 사업명	기관고유사업			
	세부사업명				
연구과제명	중 과제명				
	세부(단위)과제명	극지유전체 101 프로젝트: 극지생물 유전체 정보 분석 및 활용기반 구축			
연구책임자	김진형	해당단계 참여연구원수	총 : 44 명 내부 : 12 명 외부 : 32 명	해당단계 연구비	정부: 2,388,492 천원 기업: 천원 계: 2,388,492 천원
연구기관명 및 소속부서명	극지연구소 실용화연구사업단		참여기업명		
국제공동연구	상대국명 :		상대국연구기관명 :		
위탁연구	연구기관명 : 연세대학교		연구책임자 : 김우택		
	연구기관명 : 인천대학교		연구책임자 : 김연정		
	연구기관명 : 인천대학교		연구책임자 : 이재성		
	연구기관명 : 노스캐롤라이나대		연구책임자 : Joseph Covi		
요약(연구결과를 중심으로 개조식 500자 이내)					보고서 면수
<p>270</p> <ul style="list-style-type: none"> ○ 극지생물 고품질 유전체 해독 <ul style="list-style-type: none"> - 극지 고등생물 고품질 유전체 정보 획득 - 극지 미생물 고품질 유전체 정보 획득 - 극지 특이적 환경요인에 의한 단백질 발현지도 구축 ○ 유전체 정보 기반 극지 특이적 환경적응기작 규명 <ul style="list-style-type: none"> - 환경스트레스 처리에 따른 전사체 분석 - 발현 단백질 분석에 의한 기능성 단백질 발굴 - 발현 단백질내 단백질 상호조절 경로 및 기능 분석 - 이형 유전자 발현에 의한 후보 단백질의 기능 검증 ○ 유용 유전자원 후보 발굴 <ul style="list-style-type: none"> - 유전체 기반 유용활성 단백질, 펩타이드 발굴 - 유용 유전자 과다발현 형질전환체 제작 및 표현형분석 - 유용 유전자 형질전환 모델 생물 발굴 					
색인어 (각 5개 이상)	한글	유전체, 단백질, 전사체, 진화, 환경적응, 형질전환			
	영어	genome, proteome, transcriptome, evolution, environmental adaptation, transformation			

요 약 문

I. 제 목

- 극지유전체 101 프로젝트: 극지생물 유전체 정보 분석 및 활용기반 구축

II. 연구개발의 목적 및 필요성

1. 유전체 연구는 첨단생명공학기술을 활용한 생명자원의 확보와 활용에 관한 연구로서 미래 성장가능성이 높은 분야중 하나이며 세계적으로 연구 투자가 강화되는 추세임.
 - NGS(Next Generation Sequencing) 분석기술의 발달로 고효율, 저비용의 시퀀싱 분석이 비약적으로 발전 중임.
 - 전 세계에 걸쳐 생물의 유전체 및 유전자원 정보를 선점하기 위한 다양한 연구가 진행 중임.
2. 남극에는 약 7,200종의 동물과 300여종의 식물, 그리고 확인되지 않은 수만 종 이상의 미생물이 서식하는 등 극한환경에도 불구하고 높은 생물 다양성을 지니고 있으며, 고립된 극한환경에서 적응해온 극지생물은 다양한 산업분야로의 활용성이 높은 생명자원의 보고
 - 극지생물은 극한 환경에 적응하기 위한 독특한 적응관련 유전자, 단백질 및 다양한 이차대사물질을 보유하고 있어 미래자원으로서의 가치가 매우 높음.
 - 세계 각국은 극지생물을 중요한 미래자원으로 인식하고, 활용하기 위한 자원평가 및 기초 원천기술 개발에 주력하고 있음.
 - 생물다양성협약에서 생물자원에 대한 국가주권을 인정함에 따라 선진국 들은 체계적인 극지 및 해양생물정보은행을 구축하고 있으며, 유전체, 단백질 및 천연 대사물 탐색에 정부차원의 대규모 투자를 하고 있음.
3. 극지생물 유전체의 극지적응 관련 기작은 세계 공통의 관심사이자 풀리지 않은 난제로 남아 있음(Antarctic and Southern Ocean Science Horizon Scan Q.43).
 - 제1회 SCAR Antarctic and Southern Ocean Science Horizon Scan에서 남극 과학에서 풀어야 할 80개의 항목을 선정했으며, 이 중에서 Question 43은 남극 생물의 생존 원리에 관한 유전체기반의 물음을 담고 있음(Q.43. What is the genomic basis of adaptation in Antarctic and Southern Ocean organisms and

communities?).

4. 미국 NSF는 '남극생명체의 유전체 분석'을 국가주도연구의 전략적 우선순위로 선정하여 집중 투자하고 있음.
 - 미국의 분야별 R&D 투자 추이에서 Genomic Science는 주요 투자 대상임
 - 미국 에너지부 산하의 JGI (U.S. Department of Energy Joint Genome Institute)는 Non-human(해양동물, 식물, 미생물) 분야의 "Genome to Life" 프로젝트를 진행, 최근 극지 유전체 등 극한 생물에 대한 유전체 정보 수집 확대
5. 중국의 BGI에서 수행 중인 'Animal Genome Project'에 의해 전세계 주요 수산생물의 유전체 연구가 진행되고 있음.
6. 우리나라도 국가주도의 유전체 사업에 투자를 늘리고 있으나, 선진국 대비 기술 격차를 완전히 해소하지 못하고 있는 상황임.
 - 우리나라도 범 국가 차원에서 유전체 관련 기술 수준을 끌어올려 의료와 산업에 활용하고자 2014년부터 8년간 약 6천억 원이 투자됨.
 - 생물자원으로서 무한한 가치를 지닌 극지생물의 경우 유전체사업 대상에 포함 안 됨.
7. 극지연구소는 다년간 유전체 연구에 투자하여 해당 분야에서 높은 수준의 기술력을 확보하였고 연구결과의 양적 성장을 이루어 세계에서 독보적인 위치를 선점하였음(최근 5년 출판 논문 수 기준 BAS 대비 220% 수준).
 - 연구결과의 질적 수준은 BAS 대비 82% 수준(IF 기준)으로, 향후 수월성 향상에 집중한다면 세계 최고수준 달성 가능
8. 극지연구소 비전이 "극지연구 글로벌 선도기관" 도약을 위해서 수월성 연구에 집중 필요
9. 극지 생명과학 연구 분야에 있어서 유전체 연구는 타 분야 연구에 비해, 그 파급효과가 매우 높음. 2014년부터 2016년까지 연구 논문의 인용수에 있어서 유전체분야 연구는 기타 극지생명자연과학 분야보다 약 6배 많음(그림 2).
 - 국민 건강을 위협하는 기존 항생제 내성균의 등장에 따른 차세대 항생제 개발 요구
 - 기존의 항생제 개발 방법의 한계(배양되는 방선균에서만 연구, 반복적인 동일 물질 발견)를 극복할 수 있는 신종 방선균 또는 새로운 실험적 방법 요구

10. 선진국들은 이미 국가적 수준에서 전략적으로 새로운 항생제를 찾는 연구를 지원하고 있음. 미국은 2015년 오바마 대통령 주도하에 2020년 내 10개의 신규 항생제를 개발하자는 '10×20 프로그램' 진행 중, 유럽은 EU 주관의 "PharmaSea(2012.10.01.~2016.09.30)" 프로젝트를 수행하여 남극과 북극을 포함한 해양에서 새로운 항생제 및 신약 후보 물질 발굴 수행(13개국에서 24개의 기업 및 연구기관이 참여, 4년간 950만유로 이상의 투자)

11. 국내외 다른 연구자와의 극지생명자원 활용을 위한 중개연구 필요

- 극지 생물의 유전정보를 이용한 항생제 개발 분야에서의 학·연·산 공동연구 필요

12. 현재 알려진 천연 항생물질의 약 64% 정도가 방선균에 의해 생산되었음. 하지만 신규 항생물질 탐색과정에서 이미 밝혀진 물질들만 재차 분리되는 악순환이 계속되면서, 그 경제성과 효용성이 현격하게 떨어지고 있기 때문에 신규 항생물질의 탐색을 위해 극지 및 해양에서의 신규 방선균의 유전자 정보 활용 요구

13. 예산 미반영시 항생제 내성균 문제라는 시급한 국가현안에 대한 선제적 대응이 늦추어지고 기존의 모든 항생제에 내성을 가지는 슈퍼박테리아 출현이라는 국가적 재난 발생 가능성 증가(극지 신규 항생물질 연구개발 파이프라인 구축이 늦어지고, 그로 인한 해외로의 지식재산권 유출 예상)

Ⅲ. 연구개발의 내용 및 범위

1. 극지생물 고품질 유전체 해독

- 극지 고등생물 고품질 유전체 정보 획득
- 극지 미생물 고품질 유전체 정보 획득
- 유전체 정보 DB 구축

2. 유전체 정보 기반 극지 특이적 환경적응기작 규명

- 환경스트레스 처리에 따른 전사체 분석
- 극지 특이적 환경요인에 의한 단백질 발현지도 구축
- 발현 단백질내 단백질 상호조절 경로 및 기능 분석

3. 극지 특이적 유용 유전자원 후보 7종 이상 발굴

- 발현단백체 분석기반 기능성 단백질 발굴
- 이형 유전자 발현기반 후보 단백질 기능 검증
- 유전체 기반 유용활성 단백질, 펩타이드 발굴
- 유용 유전자 과다발현 형질전환체 제작 및 표현형분석
- 유용 유전자 형질전환 모델 생물 발굴

IV. 연구개발결과

1. 극지 고등생물 고품질 유전체 정보 획득

- 남극 빙어 (Icefish) 유전체 해독 (Nature Ecology & Evolution 논문 출판, 2019)
 - 남극빙어 유전체 해독을 위하여 3세대 염기서열 분석장비(Pacbio Sequel)를 활용하여 유전체 90Gb 서열분석후 1.06G 의 전장유전체 해독 완료
 - 분석한 유전자 서열을 이용하여 genome assembly 성공적으로 완료
 - 남극 빙어의 genetic map 완성
 - 남극빙어의 30,773개의 유전자를 확인하여 남극환경적응, 진화 기작의 유전체적 해석 완료
 - 남극의 저온 환경에 적응 과정 중 결빙방지단백질(AFGP), 난보호단백질 (zona pellucida), 항산화유전자의 확장 확인
 - 극단적인 광주기에 의한 생체시계관련 유전자의 손실 확인

2. 남극선대류의 고품질 유전체 분석

- 남극선대류 *Sanionia uncinata*의 유전체 고도화
 - 남극 바톤반도의 우점 선대류인 *Sanionia uncinata*의 극지 적응 및 진화적 특성을 파악하기 위해 유전체 분석 분야 최신 기술인 Sequel을 이용해 전체 21 Gb 데이터(예상 genome size 0.7 Gb의 30배 데이터) 확보 완료(2차년 결과)
 - *Sanionia*의 유전체가 4배체임을 고려하여 2개의 haplotype을 구분하는 작업을 진행하였고(v2), 다시 HiRise assembly 기술을 활용해서 염색체 수준의 assembly결과를 획득하였음(v3). v3는 v2에 비해 scaffold 개수는 60% 이상

줄었고 scaffold 길이는 6배 이상 늘어나, 유전체 품질의 고도화 달성

- 해독한 유전체의 완결성을 검증하기 위해 BUSCO 분석을 수행한 결과, 모델 이끼인 *Physcomitrella patens*와 비교했을 때 complete and single-copy BUSCO 비율은 모델 이끼종보다 더 높은 수치를 보였기 때문에 고품질 유전체로서의 기준을 충족하였음

3. 극지 미생물 고품질 유전체 정보 획득

○ 저온성 미생물 유전체 해독

- 저온성 미생물 17종의 유전체 해독을 통하여 저온활성효소, 저온특이적 유전자 선발
- 고품질의 미생물 유전체해독을 위하여 Long-read 서열 분석 방법을 통해 17종의 미생물 유전체 해독 완료
- 이들 남극 미생물 유전체로부터 저온특이적 유전자의 특성 분석을 통한 극지 특이적 환경적응기작 해석과 저온활성단백질 선발을 통한 활용성 검증

4. 극지 특이적 환경요인에 의한 단백질 지도 제작

○ 단백질 활성 지도제작

- 남극 저온박테리아의 저온적응유전자(Cold shock protein)가 도입된 형질전환 박테리아의 단백질 지도 구축
- cold shock protein이 knock-out 된 형질전환 박테리아에 해당 유전자를 도입 후 저온배양하여 단백질 지도 확인

5. 극지 환경에 의해 조절되는 생리변화 조절 경로 연구

- 남극 저온성 박테리아 *Psychrobacter sp.* PAMC 21119 유래 cold shock protein의 RNA chaperone의 기능을 검증하여 전사(transcription)를 용이하게 함으로써 다음 단계인 기능성 단백질 발현에 직접적으로 관여함을 규명
- Cold shock protein 유전자의 발현이 세포의 저온성장에 미치는 영향을 분석하기 위하여 저온 스트레스에 민감하게 반응하는 BX04 세포(CspA, B, E, G 결핍) 사용을 도입
- 형질 전환된 균주를 반복하여 얼리고 녹인 후 생존율에 영향을 끼치는지 확인 결과, 남극 저온성 박테리아 유래 Csp가 형질 전환된 박테리아가 positive control로 사용된 *E. coli* CspA와 비슷한 수준의 생존율을 확인

6. 전체 기반 유용활성 단백질, 펩타이드 기능 및 유효성 검증

○ 북극해 미생물

- 북극해에서 발견한 *Bacillus* sp. PAMC 23377 미생물 유래 cytochrome P450 (CYP) 효소인 BaCYP106A2 단백질을 이용하여 다양한 steroid 기질의 변형에 이용가능

○ 빙하 미생물

- 남극 빙하에서 발견된 스포어를 형성하는 호냉성 미생물인 *Paenisporosarcina* sp. TG-14 유래의 dihydrodipicolinate reductase (PaDHDPR) 효소 단백질의 구조-기능 연구수행, PaDHDPR 효소의 삼차구조 및 기질결합 부위 정보는 스포어를 형성하는 박테리아에 대한 항생제 신약개발 또는 합성 라이신 공정에 사용되는 균주개발에 이용 가능

7. 남극곰새풀 유용 유전자 과다발현 형질전환 벼의 표현형 분석

○ DaADF 과다발현 형질전환 벼 표현형 분석

- 남극곰새풀에 환경 스트레스를 처리했을 때 발현이 증가하는 유전자인 ADF (Actin Depolymerizing Factor)를 과다발현하는 형질전환 벼를 제작하여 표현형을 분석한 결과, 저온 내성이 야생형에 비해 4배 이상 증가하였음(2차년 결과)
- 저온 내성과 DaADF 생화학적 기능의 연관성을 확인하기 위해 저온 조건에서 남극곰새풀 세포내 미세섬유의 변화를 현미경으로 관찰한 결과, 대조군에 비해 저온 조건에서 남극곰새풀 뿌리 세포 내부 F-액틴이 감소하였음
- DaADF 과다발현 형질전환 벼의 뿌리 세포의 경우 야생형에서 비해 F-액틴이 줄어든 것을 확인하였으므로, actin 미세섬유의 변화가 저온 조건에서 세포 내부의 구조적 변화와 밀접하게 관련되어 있을 것이라 확인하였음

V. 연구개발결과의 활용계획

1. 과학분야: 남극 생물 특이적 유전자원 확보를 위한 교두보 마련

- 남극어류 유전체 해독은 극지연에서만 성공적 수행 및 세계적 선도
- 남극어류의 대표종인 icefish의 유전체를 해독하고 reference genome을 제공함으로써 남극어류 연구의 기반 제공

- 연구자들의 관심도가 높은 생물종을 우선 연구대상으로 선정하고 데이터를 선 공개함으로써 다학제 연구 활성화와 연구소 전체의 연구 수월성 제고
- 극지생명과학연구의 근간인 유전체 정보를 확보함으로써 극지생물의 적응과 진화 연구 토대 마련
- 극지 특이적인 생명현상 규명을 통한 지구 생명의 생성과 적응, 진화의 과학적 근거 제시
- 극지 지역을 중심으로 급속도로 진행되는 지구 온난화의 원인을 극지생물을 중심으로 기후 연구와 생물학 분야의 융합연구 연계점 제시
- 극지 환경 적응관련 핵심 유전자를 활용한 형질전환 생물 개발 및 기술력 선점
- 미지의 극지생명정보를 적극 확보하고 BT, IT, ET 등 융합학문을 촉진하여 극지생명공학 분야의 국가 경쟁력 제고

2. 사회/경제적 분야: 극지 분야 세계 바이오산업 시장 선도가능

- 극지 고유 생명정보 확보를 통한 극지연구의 저변 확대와 극지생명 연구 분야 세계 1위의 국제적 선도 기관으로 도약
- 극지 생명현상의 이해와 해석 정보를 바탕으로 새로운 생명자원을 개발하기 위한 4차 산업혁명의 핵심 생물자원 정보 제공
- 세계 바이오산업 시장 규모는 약 330조 원 (2013년 기준)로 연평균 10.8%의 고성장세를 보이고 있으며, 지속적으로 증가될 전망. 미래 글로벌 시장의 선점을 위해서 극지생물을 이용한 다양한 생명공학기술의 연구개발이 필요하며, 미래경제력 창조의 촉매 역할을 담당
- 극지생물의 환경적응, 진화 기작과 관련한 유전체를 확보하여 국가생명공학산업의 새로운 패러다임으로 다양한 신규 바이오소재의 활용가치를 제시함
- 극지 생물 신소재 개발을 위한 기초연구로 21세기 세계 생명공학 분야에서 선도적 역할 담당
- 극지 생물 특이적 유전자 분석 기술 전수를 통한 국내 연구 및 산업계 이익 창출
- 극지 생물의 극한 환경 적응관련 핵심 유전자를 활용한 의료, 미용 소재 개발을 통한 경제성 창출
- 극한 환경 적응관련 핵심 유전자 형질전환 활용 시 국가적 이익 창출과 더불어

반복되는 저수온, 고수온, 적조 등의 급격한 환경변화로 인한 수산생물의 피해
저감을 통한 국내 수산양식산업 전반의 사회 문제 해결 가능

3. 정책 분야: 정부시책에 부합된 연구 수행

- 나고야의정서 내 유전체 생물 주권에 있어 극지 생물 분야의 글로벌 선도 기관
으로 세계적 우위 선점 가능
- 극지 수산생물을 이용한 고부가 수산양식 기술 개발을 통한 정부시책에 부합
된 연구 수행 가능: 제1차 해양수산과학기술 육성 기본계획(과제 2-1 전략산
업 육성을 위한 상용화 기술 확보)
- 극지 해양생물 배양기술 확립을 통한 극지 해양자원 보존 등 국제사회 공동 현
안 해결가능: 제1차 해양수산과학기술 육성 기본계획(과제 2-2 국제 사회에
기여하는 해양수산과학기술 육성)



S U M M A R Y

I. Title

- Polar Genome 101 Project

II. Purpose and Necessity of R&D

1. Understanding of genomic properties and development of useful genomic resources from polar organisms

III. Contents and scope of R&D

1. High-quality genome data analysis from polar organisms
 - High-quality genome data collection from polar higher organisms
 - High-quality genome data collection from polar microorganisms
 - Proteom analysis of polar organisms affected by polar-specific environmental factors
2. Investigation of polar-specific environment adaptation mechanism from genome data analysis
 - Transcriptome analysis of polar organisms affected by the environmental stress treatment
 - Functional protein identification by proteome analysis
 - Analysis of protein signal transduction and biological function from proteome data
 - Functional verification of target gene using heterologous gene expression
3. Development of useful genetic resource candidate
 - Development of genome-based useful active protein and peptide
 - Production of target gene over-expressed transformants and phenotype analysis

- Development of model transformants with useful target gene oriented from polar organisms

IV. R&D Result

1. Verifying the possibility of antibiotic deformation using substrate flexibility of cryogenic protein derived from polar organisms

- Securing antibiotic modified enzyme genes from genetic information of various polar organisms
- Analyze the structure of six types of CYP enzymes and check the hydroxide reaction of antibiotics
- Various types of CYP enzymes have hydroxyl reactivity in different parts of antibiotics
- Verification of the activity of removing certain acetyl group of antibiotics in the PbAcE enzyme (Application for domestic patent)

2. Establishing a catalog by function of antibiotic strain enzymes

- Vector system for the expression of the re-combined antibiotic strain enzyme
- Establishment of an electro-magnetic protein refining system (second year, introduction of AVANT FPLC research equipment)
- Using glycosyl transferase, hydroxylase, methyltransferase, sulfotransferase, Isomerase, and Acyltransferase enzymes to find genes, re-combined protein mass production, biochemical properties analysis, and tertiary structural analysis.
- Once the catalog is completed for each function of low-temperature enzyme, it can be used not only for antibiotic candidate materials, but also for strain making of useful substances that cannot be chemically synthesized.

3. New antibiotic strain development succeeded.

- Development of a strain that combines sugar from Magnololol and Honokiol antibiotics
- Produce six types of antibiotic strains of quinolone

- Producing the Lafamycin antibiotic strain of the Macloide family in the process of an anti-bacterial activity test
 - Produce six types of antibiotic strains of quinolone
4. Establishing a research platform for developing new antibiotics
- Establishing a cooperative system with Gaffibio, a raw material company of a pharmaceutical company
 - Success in the production of strains of quinolon and steroid antibiotics seeking deformation in Gaffibio
 - Joint study on the antimicrobial activity analysis of marine life origin natural products from polar marine life with a research team from Chungnam National University's Pharmacology College
 - Joint research on antimicrobial activity analysis of antimicrobial materials developed with the research team of Chung-Ang University's Pharmacology College
 - Joint study on the structure of the strain of antibiotics developed with the research team of the pharmaceutical engineering department of Sunmoon University

V. Utilization Plan of R&D Results

1. Growing to become the world's top research institute in the field of polar genome research through 33 SCI journal publications and 4 patents application
2. Understanding of polar-specific environment adaptation mechanism and evolutionary evidence for polar organisms
3. Application study such as transformants development using the genetic information obtained from polar-specific living mechanisms
4. International polar research collaborations with genome and proteome data through the polar bioinformatic hub-network system

목 차

제 1 장 서론

제1절 연구의 배경 및 목적

제2절 연구개발의 필요성

1. 극지생물 유전체연구
2. 기술적 측면
3. 경제, 산업적 측면
4. 과학적 측면
5. 사회, 문화적 측면

제 2 장 국내외 기술개발 현황

제1절 국외 기술, 산업동향

1. 기술 및 특허동향 분석
2. 경제성 분석

제2절 국내 기술, 산업동향

1. 기술 및 특허동향 분석
2. 경제성 분석

제 3 장 연구개발 수행 내용 및 결과

제1절 연구목표

1. 연구 개발의 최종 목표
2. 연차별 성과목표

제2절 연구내용

1. 1차년도(2017년) 연구 결과
2. 2차년도(2018년) 연구결과
3. 3차년도(2019년) 연구결과

제 4 장 연구개발결과의 활용계획

1. 향후 연구방향
2. 성과 활용계획

제 5 장 참고문헌



C O N T E N T S

Chapter 1. Introduction

Section 1. The Background and Purpose of the Study

Section 2. the need for the study

1. Development of a useful substance using substrate flexibility of cryogenic enzyme in polar life
2. Technical aspects
3. Economic and industrial aspects
4. Scientific aspect
5. Social and cultural aspects

Chapter 2. Current Status of Domestic and Foreign Technology Development

Section 1. Development of a useful substance using substrate flexibility of cryogenic enzyme

1. Types of Antibiotics and Operation Techniques
2. New Antibiotics Excavation Cases
3. Technology and Patent Trend Analysis
4. Economic Analysis

Section 2. Development of Antibiotics for Merk Corporation, a major patent applicant

1. List of major applicants for antibiotics
2. An Analysis of Merk's Strategies for Developing Antibiotics

Chapter 3 Contents and Results of R&D Implementation

Section 1. Research Objectives

1. Final goal of research and development
2. Annual Discrimination Performance goal

Section 2. Research Contents

Chapter 4 Utilization Plan of R&D Results

1. Future Research Direction
2. Performance utilization plan

Chapter 5 References



제 1장 서론

제 1절 연구의 배경 및 목적

- (연구의 배경)
- (연구의 목적)

제 2절 연구의 필요성

1. 유전체 연구는 첨단생명공학기술을 활용한 생명자원의 확보와 활용에 관한 연구로서 미래 성장가능성이 높은 분야중 하나이며 세계적으로 연구 투자가 강화되는 추세임.
 - NGS(Next Generation Sequencing) 분석기술의 발달로 고효율, 저비용의 시퀀싱 분석이 비약적으로 발전 중임.
 - 전 세계에 걸쳐 생물의 유전체 및 유전자원 정보를 선점하기 위한 다양한 연구가 진행 중임.
2. 남극에는 약 7,200종의 동물과 300여종의 식물, 그리고 확인되지 않은 수만 종 이상의 미생물이 서식하는 등 극한환경에도 불구하고 높은 생물 다양성을 지니고 있으며, 고립된 극한환경에서 적응해온 극지생물은 다양한 산업분야로의 활용성이 높은 생명자원의 보고
 - 극지생물은 극한 환경에 적응하기 위한 독특한 적응관련 유전자, 단백질 및 다양한 이차대사물질을 보유하고 있어 미래자원으로서의 가치가 매우 높음.
 - 세계 각국은 극지생물을 중요한 미래자원으로 인식하고, 활용하기 위한 자원평가 및 기초 원천기술 개발에 주력하고 있음.
 - 생물다양성협약에서 생물자원에 대한 국가주권을 인정함에 따라 선진국 들은 체계적인 극지 및 해양생물정보은행을 구축하고 있으며, 유전체, 단백질 및 천연 대사물 탐색에 정부차원의 대규모 투자를 하고 있음.
3. 극지생물 유전체의 극지적응 관련 기작은 세계 공통의 관심사이자 풀리지 않은 난제로 남아 있음(Antarctic and Southern Ocean Science Horizon Scan Q.43).
 - 제1회 SCAR Antarctic and Southern Ocean Science Horizon Scan에서 남극 과학에서 풀어야 할 80개의 항목을 선정했으며, 이 중에서 Question 43은 남극

생물의 생존 원리에 관한 유전체기반의 물음을 담고 있음(Q.43. What is the genomic basis of adaptation in Antarctic and Southern Ocean organisms and communities?).

4. 미국 NSF는 '남극생명체의 유전체 분석'을 국가주도연구의 전략적 우선순위로 선정하여 집중 투자하고 있음.
 - 미국의 분야별 R&D 투자 추이에서 Genomic Science는 주요 투자 대상임(그림 1).



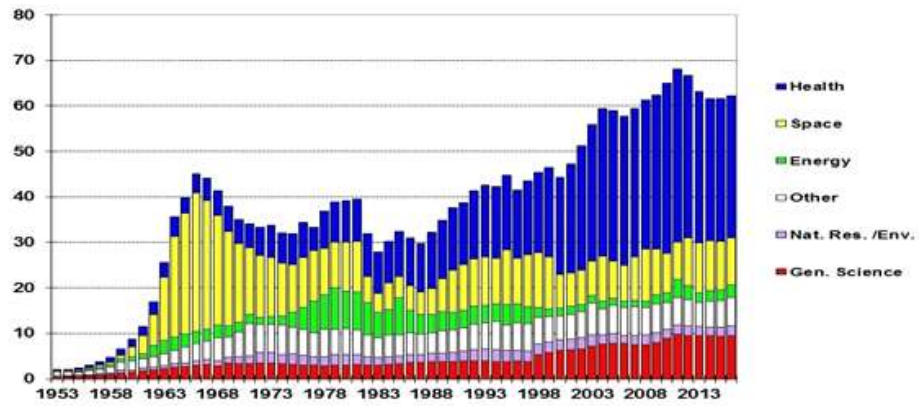


그림1. 미 연방정부 분야별 R&D 투자 추이(국방부 제외)



- 미국 에너지부 산하의 JGI (U.S. Department of Energy Joint Genome Institute)는 Non-human(해양동물, 식물, 미생물) 분야의 “Genome to Life” 프로젝트를 진행, 최근 극지 유전체 등 극한 생물에 대한 유전체 정보 수집 확대
5. 중국의 BGI에서 수행 중인 ‘Animal Genome Project’에 의해 전세계 주요 수산 생물의 유전체 연구가 진행되고 있음.
 6. 우리나라도 국가주도의 유전체 사업에 투자를 늘리고 있으나, 선진국 대비 기술 격차를 완전히 해소하지 못하고 있는 상황임.
 - 우리나라도 범 국가 차원에서 유전체 관련 기술 수준을 끌어올려 의료와 산업에 활용하고자 2014년부터 8년간 약 6천억 원이 투자됨.
 - 생물자원으로서 무한한 가치를 지닌 극지생물의 경우 유전체사업 대상에 포함 안 됨.
 7. 극지연구소는 다년간 유전체 연구에 투자하여 해당 분야에서 높은 수준의 기술력을 확보하였고 연구결과의 양적 성장을 이루어 세계에서 독보적인 위치를 선점하였음(최근 5년 출판 논문 수 기준 BAS 대비 220% 수준).
 - 연구결과의 질적 수준은 BAS 대비 82% 수준(IF 기준)으로, 향후 수월성 향상에 집중한다면 세계 최고수준 달성 가능
 8. 극지연구소 비전이 “극지연구 글로벌 선도기관” 도약을 위해서 수월성 연구에 집중 필요
 9. 극지 생명과학 연구 분야에 있어서 유전체 연구는 타 분야 연구에 비해, 그 파급효과가 매우 높음.
 - 2014년부터 2016년까지 연구 논문의 인용수에 있어서 유전체분야 연구는 기타 극지생명자연과학 분야보다 약 6배 많음(그림 2).

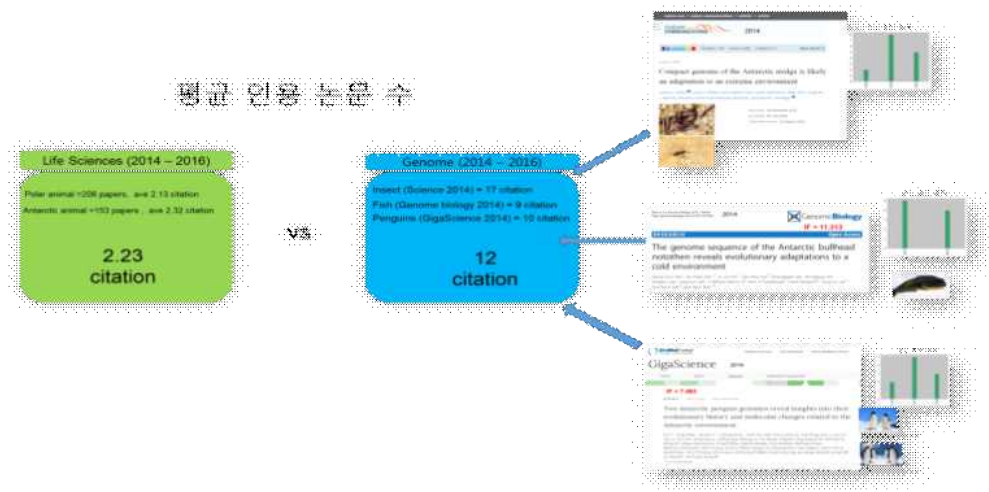


그림 2. 3년(2014-2016)간 출판된 논문의 평균 인용수 비교



제 2 장 국내외 기술개발 현황

제1절 국외기술, 산업동향

1. 미국·영국·프랑스 등의 선진국은 극지를 포함한 제3세계의 생명자원을 무차별적으로 수집, 분석하여 지적재산권을 확보하고 있을 뿐만 아니라 몬산토, 신젠타 등의 다국적 기업에서도 유전체를 기반으로 생명 경제시대를 주도하고자 생명정보를 경쟁적으로 확보하고 있음

○ 유전체 해독 기술의 발전은 유전체 연구를 근본적으로 변화시키고 있음. 2005년 NGS 기술이 개발된 이래 지속적인 기술개발이 이루어진 결과 NGS 유전체 해독 기기의 시장규모는 2008년 3.8억 달러로 확대되었고 2014년에는 NGS 매출이 36억 달러까지 증대될 것으로 예측 (Scientific Advisor)

○ 인간유전체 분야는 2009년 46억 달러에서 2014년 73억 달러로, 농업유전체 분야는 2009년 6억 달러에서 2014년 24억 달러로, 분자진단 분야는 2009년 5억 달러에서 2014년 8억 달러로, 기타 응용분야는 2009년 5억 달러에서 2014년 8억 달러로 확대될 것으로 전망하고 있음.

○ (미국) 전 세계 유전체 연구를 주도하고 있음. 21세기 해양청사진 등 해양관련 지속적 정책에 따라 해양에 대한 7개 중점기술 개발 분야를 선정하여 바이오 연구 투자 확대 및 연구 주도, 해양샘플링 수립 활동 및 해양생물종 유전체 연구 활발히 진행

- 미국의 유전체 연구는 보건성 산하의 국립보건원(NIH) 및 에너지성(DOE) 산하의 Joint Genome Institute (JGI)를 중심으로 수행되고 있음.

- JGI (Joint Genome Institute, Department of Energy, USA)의 경우 Non-human 분야 (해양동물, 식물, 미생물 분야)에 대한 집중적인 유전체 연구로 환경 문제 해결과 바이오에너지 개발, 2009년 4.1억 달러 투자하였으며, “Genome to Life” 프로그램을 통해 바이오에너지 생산과 환경에너지 정화를 위한 미생물 유전체를 중점적으로 지원하고 있음: Bioenergy, 해양생물유전체, Fungal 유전체, Microbial 유전체, Metagenome, 식물 유전체 연구를 대규모로 진행하고 있으며 최근 극지유전체 등 극한 생물에 대한 생명 정보 수집 확대

- J. Craig Venter Institute, BROAD institute, Institute for Genome Science, Washington Uni. of St. Louis 등에서도 각각 9%, 7%, 5%, 3%의 점유율을 보이는 등 미국의 연구기관 5곳이 46%의 비중을 보이며 주도하고 있음.

- (영국) 미국에 이어 세계 2위의 유전체 강국임. Wellcome Trust Sanger Institute, BRSC Genome Analysis Centre 등을 중심으로 암 유전체, 게놈연관 분석, 바이오에너지와 영양 및 농업 분야 등 경제적으로 중요한 생명체의 유전체 해독을 진행하고 있음.
- (일본) 이화학연구소, 동경대 의과대학, 국립유전학연구소를 중심으로 시스템 유전체 분야, 유전체 의과학 분야 및 환경유전체 분야에 대해 연구를 진행하고 있음. 산학연 협력 연구 활성화를 통한 해양유전체 연구 추진, 해양과학기술연구센터(JAMSTEC)와 해양생명공학연구소(MBI), 수산종합연구센터(FRA)를 중심으로 유전체 연구 투자
- (중국) Beijing Genome Institute(BGI)를 중심으로 유전체 정보 생산을 수행하고 있으며 전세계 동식물 1,000종의 주요 생물을 위주로 전체 게놈 해독 연구가 진행 중에 있어 이 분야 연구에서 가장 적극적임
- (캐나다) 연어 유전체 연구 추진, 수산해양부(DFP, Department of Fisheries and Oceans)를 중심으로 해양생명공학 분야 유전체학 및 생명공학 발전계획을 수립하였고, 대서양 연어 유전자지도 작성 등 연어 유전체 연구 수행
- 빅데이터 기술 및 long-read 시퀀싱 기술을 통해 생물 유전체 정보 축적은 더욱 가속화될 전망이며 이로부터 비롯된 생물정보들은 신약 개발 등 응용 기술에 대한 적용 범위를 확대할 예정

2. 국내 기술, 산업동향

가. 다부처 유전체 사업의 일환으로 2015년부터 해양수산부에서 “해양수산생물 유전체정보 분석 및 활용기반 구축” 사업에서 국내 해양생물(동물, 식물, 미생물) 100종의 유전체정보와 40종의 메타유전체 분석 및 유용기능 제시 진행

- 극지생물의 생명정보의 분석은 주로 극지연구소의 주도하에 진행되고 있음.
- 2014년 남극 고등생물 최초로 남극대구의 유전체 서열을 규명하고 저온적응기작을 해석함.
- 2013년 극지식물(남극좁새풀)의 전사체 분석을 통한 저온내성유전자 발굴, 기능규명하고, 이를 활용한 저온내성 작물 개발
- 극지연구소에서 극지 어류와 요각류 등 동,식물 20여 종의 전사체 정보 확보와 저온 적응 특이성 유전자 규명

3. 기술 및 특허동향 분석

가. 분석 배경 및 목적

- 본 특허 동향 분석에서는 『극지 생물 유전체 관련 특허』에 대한 특허 동향을 분석함으로써 주요국가의 특허출원동향 및 경쟁력 현황 등을 파악하고 각 분야별 핵심 특허 및 출원자 분석을 통한 분석대상기술의 R&D 전략 수립 및 IP 전략 수립에 대한 객관적인 타당성을 제공하고자 함

나. 특허동향 분석

(1) 극지 동물 유전체 관련 특허동향 분석

- 1999년 이후부터 현재까지 지속적인 특허 출원이 이루어지고 있으며, 미국의 경우 2010년, 유럽과 일본의 경우 2001년에 가장 많은 수의 특허를 출원한 것으로 나타남.
- 국가별 점유율은 미국특허가 34건으로 과반수가 넘는 점유율(59%)을 차지하고 있고, 다음으로 유럽(17%), 일본(14%), 한국(10%) 순으로 나타남.



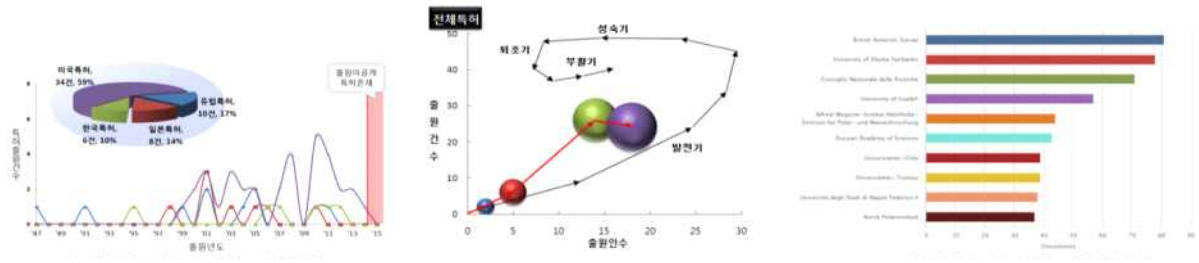


그림 3. 극지 동물 유전체 관련 분야 연도별 특허출원, 포트폴리오로 본 극지 동물 유전체 기술 분야의 위치, 극지 동물 유전체 관련 주요 연구기관 분석



- 극지 동물 유전체 관련 기술 분야에 대한 기술주기 변화를 살펴보면, '87-'93년에서 '94-'00년 구간을 거쳐 '01-'07년 구간까지 전체특허동향은 기술혁신의 주체인 출원인수와 기술혁신의 결과인 출원건수가 동시에 급격히 증가하였고, '08-'15년에 이르면서 출원건수는 다소 감소하였으나, 출원인수는 증가하고 있으며, 특허 출원 이후 1년 6개월이 지나야 공개되는 점을 감안하는 경우, 미공개된 출원건수가 더 존재할 것으로 예상되는 바, 해당 기술은 포트폴리오 기본 모델에서 발전기에 해당하는 것으로 분석됨.

○ 극지 식물 유전체 관련 특허동향 분석

- 극지 식물 유전체 관련 기술 분야의 전체적인 특허동향을 살펴보면, 1991년도 유럽 특허를 시작으로, 2000년대에 들어서서 미국이 타 국가 대비 많은 수의 출원을 하고 있음.
- 국가별 점유율은 미국특허가 가장 많은 특허출원 건수(21건)를 보였고, 75%의 점유율을 차지하고 있어 해당 기술 분야에서 미국이 기술 흐름을 리드하고 있는 것으로 나타났으며, 이 외에 한국(14%), 유럽(7%), 일본(4%) 순으로 나타남.



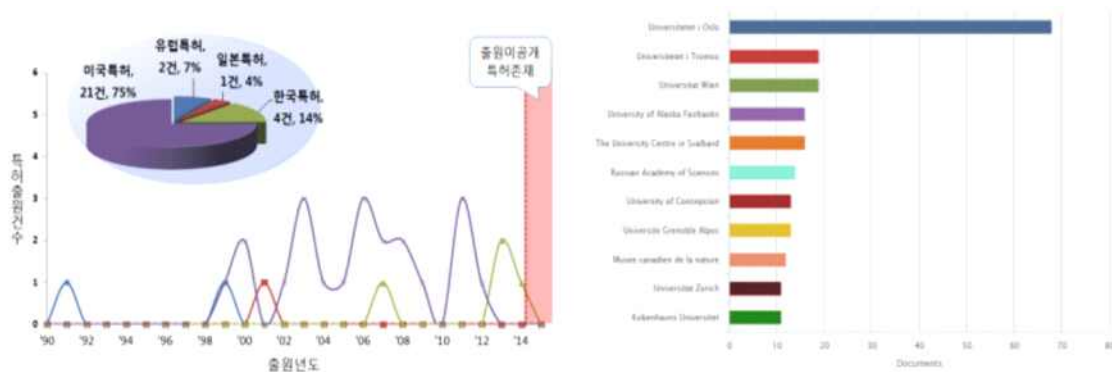


그림 4. 극지 식물 유전체 관련 분야 연도별 특허출원, 극지 식물 유전체 관련 주요 연구기관 분석



- 극지 식물 유전체 관련 기술에 대한 주요출원인으로는 VITROGEN SA(칠레), DANISCO AS(덴마크), 한국해양과학기술원(한국), AGRICULTURE VICTORIA SERVICES PTY LTD(호주) 등이 있음.
- 특이점은 한국은 자국기업에 의해 특허출원이 이루어진 반면, 미국, 유럽, 일본은 다국적 기업 중심의 출원활동을 보이고 있고, 주요 출원인 가운데 DANISCO AS의 경우 자국뿐만 아니라 미국에서도 특허를 출원하였음.
- 극지 식물 유전체 관련 발표 논문('81-'16년) 총 627건 중 Universitetet i Oslo(노르웨이)에서 68건으로 가장 많은 수의 논문을 발표하였고, 그 뒤를 이어 Universitetet i Tromso(노르웨이) 및 Universitat Wien(오스트리아)에서 각각 19건, University of Alaska Fairbanks(미국) 및 The University Centre in Svalbard(노르웨이)에서 각각 16건의 논문을 발표한 것으로 조사됨.

○ 극지 조류(algae) 유전체 관련 특허동향 분석

- 극지 조류 유전체 관련 기술 분야의 전체적인 특허동향을 살펴보면, 2004년도에 한국에서 처음으로 1건의 특허 출원이 이루어졌고, 2008년부터 현재까지 등락을 반복하며 소수의 특허 출원이 이루어지고 있으며, 미국의 경우 2008년에 급격히 많은 수의 출원이 이루어졌고, 2010년부터 꾸준한 출원 활동을 보이고 있음.
- 국가별 점유율은 미국특허가 가장 많은 특허출원 건수(24건)를 보였고, 59%의 점유율을 차지하고 있어 해당 기술 분야에서 미국이 기술 흐름을 리드하고 있는 것으로 나타났으며, 이 외에 한국(24%), 일본(12%), 유럽(5%) 순으로 나타남.
- 극지 조류 유전체 관련 기술에 대한 주요출원인으로는 SOLAZYME INC(미국), 삼성전자 주식회사(한국), 한양대학교 산학협력단(한국), UNIVERSITY OF STIRLING(영국), UNIVERSITY OF TEXAS(미국) 등이 있음.
- 특이점은 주요 출원인 가운데 SOLAZYME INC의 경우 자국 뿐 아니라 한국과 일본에서도 출원활동을 진행하였고, 삼성전자 주식회사 및 한양대학교 역시 미국과 유럽에도 특허를 출원하였음.

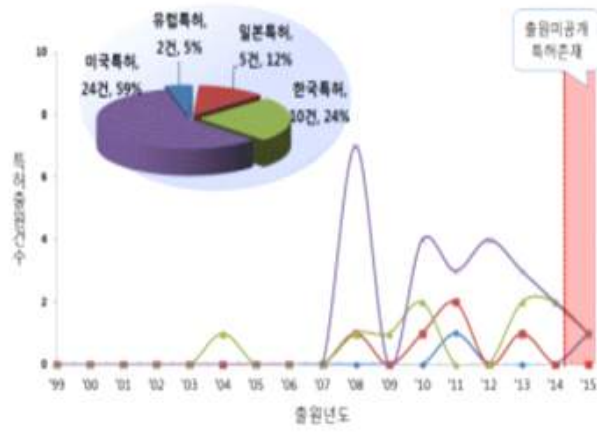


그림 5. 극지 조류 유전체 분야 연도별 특허출원



(가) 극지 미생물 유래 항생제 기술 분야는 미국을 제외한 국가의 유효 건수 및 출원인수가 매우 적었으므로 국가별 분석 수행에 어려움이 있어 전체 특허를 대상으로 주요출원인 분석을 수행하였음

- 또한, 극지 미생물로부터 유래된 항생제와 직접적으로 관련된 특허가 거의 없어, 본 기술은 주요 키워드를 포함하는 특허를 대상으로 분석을 진행하였기 때문에 해당 기술 분야를 주도하는 출원인 분석이 어려우므로, 각국 주요출원인에 대한 그래프는 생략함

(나) 본 기술 분야에 대한 주요출원인으로는 미국과 유럽 국적의 출원인이 주축을 이루고 있으며, 특히 미국 국적의 출원인이 과반수를 차지하고 있었고, FOAMIX(이스라엘), PROCTER & GAMBLE(미국), MONOSOLRX(미국) 등이 상위권에 속해 있음

① 특이점으로 주요출원인은 다국적 기업 외에도 대학교에서 많은 특허를 출원한 것으로 나타났으며, 여러 국가에 진출하여 활발한 출원활동을 진행한 출원인은 거의 없는 것으로 파악됨

② 다출원 1위 출원인은 제약회사인 FOAMIX(이스라엘)로 확인되었는데, 대상 기술과 관련된 특허를 보유하고 있지는 않았고, 주요출원인 중 CALIFORNIA 대학교(미국)는 다른 항생제 재발기술 분야에서도 상위권에 속한 출원인임

③ 한국 국적의 출원인은 주요출원인에 해당되지 않았고, 한국생명공학연구원(한국) 및 한국해양연구원(한국)에서 각각 1건씩 특허를 출원한 바 있으나, 항생기능을 갖는 재조합 단백질 제조방법과 남극에 서식하는 동물 유래 항생물질에 관한 특허인 것으로 확인됨

④ 본 기술 분야의 전체적인 출원 건수가 적은 것은 대상 기술이 아직 개발 초기 단계여서 본격적인 연구가 진행되지 않았거나 공개되지 않은 특허가 많이 존재하기 때문인 것으로 판단됨

제 3장 연구개발 수행 내용 및 결과

제 1절 연구목표

1. 연구개발의 최종목표

가. 극지생물 유전체 특성 규명과 극지 유용유전자원 활용을 통한 국제적 연구 선도

- 극지 대표생물 100종 이상의 고품질 유전체 정보 획득
- 극지생물의 환경적응과 진화 특성 규명
- 산업적 고부가가치 창출 가능한 극지생물 유전자원 15종 발굴

2. 연차별 성과목표

구분	년도	성과목표	연구내용	연구범위	연구비 (직접비)
1차년도	2017	극지생물 고품질 유전체 해독	- 극지 고등생물 고품질 유전체 정보 획득 - 극지 미생물 고품질 유전체 정보 획득	- 유전체 정보(12종; 고등생물 2종, 미생물 10종) - 유전체 특성분석 10건	450,000
		유전체 기반 극지 특이적 환경적응기작 규명	- 환경스트레스 처리에 따른 전사체 분석 - 극지 특이적 환경요인에 의한 단백질 발현지도 구축 - 발현 단백질내 단백질 상호조절 경로 및 기능 분석	- 전사체분석 4건 - 단백질 발현지도 1건 - 단백질 특성분석 10건 - 단백질 구조 및 기능 연구 2건 이상	400,000
		유용 유전자원 후보 발굴 극지 특이적 유용 유전자원 5종 이상 발굴	- 발현 단백질 분석에 의한 기능성 단백질 발굴 - 이형 유전자 발현에 의한 후보 단백질의 기능 검증 - 유전체 기반 유용활성 단백질, 펩타이드 발굴 - 유용 유전자 과다발현 형질전환체 제작 및 표현형분석 - 유용 유전자 형질전환 모델 생물 발굴	- 유용 유전자, 단백질 발굴 1건 이상 - 목적 단백질 발굴 10건 이상 - 단백질 기능 검증 5건 이상 - 형질전환체 3건	250,000
2차년도	2018	극지생물 고품질 유전체 해독	- 극지 고등생물 고품질 유전체 정보 획득 - 극지 미생물 고품질 유전체 정보 획득 - 극지 특이적 환경요인에 의한 단백질 발현지도 구축	- 유전체 정보(17종; 고등생물 2종, 미생물 15종) - 유전체, 단백질 특성분석 10건 - 단백질 발현지도 1건	450,000
		유전체 기반 극지	- 환경스트레스 처리에 따른 전사체 분석	- 유전체 기능분석 5건	400,000

		특이적 환경 응답기작 규명	<ul style="list-style-type: none"> - 모델종과의 비교유전체 분석을 통한 저온특이적 유전자군 선별 - 극지 환경에 의해 발현되는 단백질 지도 작성 및 조절 경로 연구 	<ul style="list-style-type: none"> - 단백질 구조 및 기능 연구 2건 이상 - 발현조절해석 2건 	250,000
		유용 유전자 후보 발굴	<ul style="list-style-type: none"> - 유전체 기반 유용활성 단백질, 펩타이드 발굴 및 기능검증 - 유용 유전자 과다발현 형질전환체 제작 및 표현형분석 	<ul style="list-style-type: none"> - 유용 단백질 발굴 2건 이상 - 형질전환 모델생물 1건 	
		극지 특이적 유용 유전자 후보 5종 이상 발굴	<ul style="list-style-type: none"> - 세포내 분자상호조절 기작에 근거한 후보 단백질 선정 - 유용 유전자 형질전환 모델 생물 발굴 		
3차년도	2019	극지 생물 고 품질 유전체 해독	<ul style="list-style-type: none"> - 극지 고등생물 고품질 유전체 정보 획득 - 고품질 미생물 유전체 정보 획득 	<ul style="list-style-type: none"> - 유전체정보(11종; 고등생물1종, 미생물 10종) - 유전체, 단백질 특성분석 10건 	350,000
		유전체 정보 극지 특이적 환경기작 규명	<ul style="list-style-type: none"> - 환경스트레스 처리에 따른 전사체 및 단백질 분석 - 복수 극지생물종의 발현 단백질 유사성 분석 - 극지 환경 특성에 의한 단백질 조절기작 연구 	<ul style="list-style-type: none"> - 유전체 기능분석 5건 - 단백질 구조 및 기능 연구 2건 이상 	400,000
		유용 유전자 후보 발굴	<ul style="list-style-type: none"> - 유전체 기반 유용활성 단백질, 펩타이드 기능 및 유효성 검증 - 유용 유전자 과다발현 형질전환체 표현형분석 	<ul style="list-style-type: none"> - 유용 단백질 발굴 2건 이상 - 형질전환 모델생물 1건 	350,000
		극지 특이적 유용 유전자 후보 5종 이상 발굴	<ul style="list-style-type: none"> - 복수 환경 요인에 의해 조절되는 단백질 후보 발굴 - 유용 유전자 형질전환 모델 생물 활용 		

제 2절 연구내용

- 남극 빙어의 유전체 분석을 통한 극한 환경으로의 적응 기작 규명 연구
- 남극 드래곤피쉬 *Parachaenichthys charcoti*의 게놈 분석을 통한 극한환경 적응 기작의 이해
- 남극개미자리의 현장 및 챔버 생육 샘플의 전사체 비교를 통한 환경적응기작 분석
- 남극좁새풀 유전자 DaGolS2 과발현에 의한 벼 형질전환체의 내냉성 증대 효과 분석
- 극지미세조류 KNF0032의 얼음결합 단백질 CmIBP의 분리 및 이들의 활성검증과 과다발현 형질전환체의 표현형 분석
- 남극 선대류 AP2 유전자 과다발현 형질전환체 제작과 저온 표현형 분석
- 남극 좁새풀 유래 Chalcone isomerase (DaCHI) 효소의 구조-기능 연구
- 극지 호냉성 박테리아(*Exiguobacterium antarcticum*) 유래 EaEST 효소의 구조-기능연구
- 바실러스 PAMC 23377 균주 유래 BaCYP106A2 효소의 구조와 기능적 특성 연구
- 북극 해양 미생물(*Colwellia psychrerythraea*) 유래 isoaspartyl dipeptidase 효소(CpsIadA) 의 삼차구조 및 기능적 특성 규명
- 남극 호냉성 미생물 유래의 저온성 esterase 효소의 삼차구조 해석
- 극지 특이적 환경요인에 의한 단백질 지도 제작
- 극지 환경에 의해 조절되는 생리변화 조절 경로 연구
- 극지 미생물에서 생산되는 유용 효소 단백질의 구조 분석 및 특성 연구

제 3절 연구결과

제 3-1절 남극 빙어의 유전체 분석을 통한 극한 환경으로의 적응 기작 규명 연구

1. 서론

Blackfin icefish (*Chaenocephalus aceratus*)는 남극암치아목(Notothenoid)내 남극빙어과(Channichthyidea)에 속하며 16종의 남극빙어과 내에서도 헤모글로빈과 심장 미오글로빈이 모두 손실된 5종의 남극빙어 중 한 종으로 적혈구 세포를 가지고 있지 않아 투명한 혈액을 가지고 있다[1,2]. 남극 빙어는 혈액 내 산소 운반에 있어 헤모글로빈을 사용하지 않기 때문에 남극암치아목 내 헤모글로빈을 갖는 어류 대비 혈액 용량의 10% 미만의 산소 운반 능력을 가지게 된다[1,3]. 남극 빙어는 산소결합단백질의 손실을 보상하기 위한 메커니즘을 진화 시켰는데 그 예시로 심박출량과 심장 크기의 증가, 강화된 혈관 시스템, 미토콘드리아 밀도와 형태의 변화 등을 들 수 있다[4-7]. 남극 대륙의 온도가 낮아지며 남극해의 온도가 -1.9도에 이르게 되고 남극암치아목 어류들은 동결방지단백질(Antifreeze glycoprotein)과 난보호단백질(ice-resistant egg chorion protein)등의 유전자의 확장 및 발현을 통해 남극해에서 생존, 적응, 진화에 성공할 수 있었다[8-11]. 남극암치아목은 연중 수온 변화가 없는 차가운 바다에 서식하며 온도 변화에도 열충격단백질(heat shock protein)의 발현이 증가 혹은 감소하는 전통적인 유전자 발현 패턴을 지니지 않는 것으로 알려져 있다[13,14]. 빙어를 포함한 8개의 남극암치아목 내 저서성 남극 어류는 부레가 퇴화된 것이 특징인데, 먹이 활동 시 부력을 증가시키기 위하여 뼈의 골밀도를 낮추고 비늘을 없애고 지질함량을 높이는 방향으로 진화하게 되었다 [15-20]. 이러한 남극 빙어 고유의 생리적 특징 및 극한 환경으로의 적응진화 과정을 유전체 정보를 기반으로 이해하기 위하여 남극 빙어의 전체 유전체를 해독하였다.

2. 국내외 기술개발 현황

남극 어류 유전체의 경우 2014년 남극 척추동물 최초로 남극대구의 유전체 서열을 해독하고 저온적응기작을 해석한 이래로 20여종의 다양한 남극 동·식물의 전사체정보 및 유전체 정보가 극지연구소 실용화사업단에 의해 해독 및 분석되어지고 있다. 남극암치아목 내에서도 가장 진화된 형태로 알려져 있는 남극빙어의 고품질의 유전체를 해독함으로써 극한 환경으로의 적응 기작을 규명할 수 있을 것으로 사료된다.

3. 연구 수행 결과

1) 유전 서열 해독 및 어셈블리

본 연구에서는 한 마리의 암컷 남극 빙어로부터 유전자를 추출하여 PacBio sequel system 을 사용하여 평균 길이 13Kbp 의 라이브러리를 제작, 유전체 크기의 90배 이상의 유전자 서열을 분석하였다(Supplementary Table1). 전장 유전체의 크기는 Jellyfish 프로그램으로 계산된 k-mer 값을 통해 1.1 Gbp 로 예상되었다(Supplementary Figure 3). FALCON-Unzip 어셈블을 통해 1.06 Gbp 의 유전자 서열이 3,852개의 contig로 어셈블 되었고, N50가 1.5 Mbp 로 고품질의 유전체를 구축하였다(Table 1). BUSCO를 기반으로 한 유전체 정보의 완성도 평가 결과 89%의 완성도를 보였다(Supplementary Table 2)[21]. 남극 빙어 유전체를 대상으로 유전자 예측과 상동성 검색 결과 30,773개의 단백질 코딩하는 유전자가 확인되었다. 5개의 조직을 이용한 small-RNA 전사체 분석 결과 334개의 고유 mature miRNA를 코딩하는 290개의 miRNA 유전자가 확인되었다(Table 1 및 Supplementary Table 5 및 6). 남극 빙어 게놈에는 약 50.4%의 반복 서열이 존재하였으며, 그 중 47.4%에 해당하는 대부분의 반복 서열은 transposable elements로 확인되었다(Supplementary Table 7 및 Supplementary Information 1). Kimura분석을 통해 transposable element copy들의 상대적 진화 연령을 계산하여 반복 요소의 진화양상을 추적하였으며, 다른 경골어류와의 비교 분석을 통해 비교적 최근 변형된 DNA transposons을 확인하였다(Supplementary Information 1)[22]. 이 결과는 영하의 온도로의 냉각 및 일련의 빙하 생성 사이클과 같은 강한 환경 변화에 생물이 노출될 시 transposable element가 유전체 내 대규모로 이동할 수 있다는 가설과 일치한다.

2) 유전학적 연쇄지도와 어셈블리 결과 상호 비교 분석

염색체 수준의 게놈 어셈블리를 만들기 위해, 남극 빙어의 유전자지도를 작성하였다. RAD-tag 시퀀싱을 이용해 암수 성체에 대해 각각 2천만개의 raw read를 확인하였으며, 83마리의 개별 자손 각각에 대해 평균 240만개의 raw read를 분석하였다[23-25]. Stacks 소프트웨어를 이용하여 60,038개의 RAD-tag를 발굴하였으며, 이 중 56,256(93.7%)개가 10마리 이상의 자손에 존재하는 것을 확인하였다. 12.JoinMap 4.1의 minimum logarithm of the odds (LOD) 기록을 통해 확보한 반수 염색체 기준 24개의 연쇄 그룹에 대한 마커(Supplementary Figure 4, accession SRP118539) 내 7,215개의 다형성 RAD-tag 중 4,952(55.8%)개가 83마리의 자손 중 60개 이상에 그리고 4,023 개가 수컷 지도, 암컷 지도 또는 암수 모두에 확률적으로 분포되는 것을 확인하였다. 유전학적 연쇄지도는 남극빙어 각 염색체가 송사리 각 염색체의 orthologue임을 나타내었기에 전체 연쇄지도의 송사리 게놈 대비 일치하는 연쇄지도 위치를 가진 남극 빙어

연쇄지도 그룹의 개수를 확인하였다. 남극 빙어 연쇄지도 그룹 6(Cac6)은 가장 많은 202개의 마커를 보유하고 있으며, Cac2는 가장 적은 131개의 마커를 보유한 것을 확인하였다. Cac21은 가장 길었으며(65.8 cM), 반면 Cac12는 가장 짧은 것으로 나타났다(46.7 cM). Chromonomer 소프트웨어(<http://catchenlab.life.illinois.edu/chromonomer>)를 이용해 전체 유전자 지도에 대한 contig를 정렬시킨 결과, 어셈블리 내 3,852개의 contig 중 1,065 Mb(77%) 어셈블리 총 길이 820 Mb에 대해 하나 이상의 마커로 유전자 맵 내에서 1,063(27%)개를 정렬시켰다. 이중 하나의 contig는 chimeric(Ice_000013)임을 확인하였으며, 한쪽 끝은 Cac7에 정렬되고 다른 쪽 끝은 Cac14에 정렬됨을 확인하였다(Supplementary Figure 4)[26].

Synolog를 이용해 보존된 syntenies를 확인하였으며, 남극 빙어 각 염색체는 송사리(*Oryzias latipes*, Ola; Figure 1a)와 European sea bass(*Dicentrarchus labrax*, Dla; Supplementary Figure 5a)에서 단일 염색체들에 대한 이중 상동성을 나타냈다[27]. 추가로 하나의 작은 내부 전좌를 발견했으나(Supplementary Figure 5d, e), 계통이 ~ 113 Ma 기준 분화 이후로는 빙어류, 농어 및 송사리 혈통에서 상동성을 나타내는 염색체 전좌가 발견되지 않았다[28]. 경골 어류의 염색체 진화 안정성은 포유동물 대비 큰 편으로 알려져 있다. 예를 들어 사슴의 경우 3종에서 40개의 반수체 염색체를 갖지만, 설치류의 경우 5종에서 51개의 반수체 염색체를 가진다. 남극 빙어는 경골어류 선조의 염색체 수를 유지하는 것으로 확인되었으나 많은 남극 nototheniid는 그렇지 않음을 확인하였다[29]. 예를 들어 *Notothenia* 속 다른 종들은 전체 선조 염색체의 중심 융합으로 인해 선조가 가진 24개의 염색체 대비 13, 12 또는 11개 염색체를 가지고 있는 경우도 있다[30].

남극 빙어, 농어 및 송사리 염색체의 이중 염색체는 전체적으로 비슷한 수준의 유전자 개수를 공유하지만 이들의 순서는 종종 잘 보존되지 않는 것을 확인하였다. 예를 들어 Cac12 및 Ola12의 경우 역위 및 전위에 의해 재배열된 다수의 보존된 syntenic block(도 1b, 블록 1 내지 8)을 공유하였다. 그러나 Ola12와 Dla12에서는 이러한 syntenic block의 순서가 농어 혈통에서 분리된 후 남극 빙어 혈통에서 재배열되었음을 확인하였다. 보존된 syntenic block '4'와 '5'는 농어와 stickleback 염색체(Figure 1c; Supplementary Figure 5b)에서 반대 순서를 나타내었으며, 농어와 송사리(Figure 1d) 또는 stickleback(Supplementary Figure 5c) 비교 결과는 송사리 계통 내 반전이 있었음을 확인하였다. Cac12 내 다른 계통 별 재배열된 현상이 확인되었으며, 추가 염색체 분석 결과 100 Myr 이상의 유전체 내 전위 부족에도 불구하고, 남극빙어 염색체 내 다양한 재배열이 농어목에서 계통이 분리된 후 발생했음을 확인하였다(Supplementary Figure 6).

3) 진화계통과 계놈 분화 분석

OrthoMCL 기반 계놈 서열 비교 분석 결과 남극 빙어가 13종의 경골어류에서 확인된 24,159개의 이중 상동 유전자 클러스터 중 18,636개를 보유하고 있음을 확인하였다. 13종 경골어류 계놈 전체에 걸쳐 보존되어 있는 3,718개의 orthologues 세트를 사용하여 전체 계통 발생도를 분석하였다(Supplementary Table 8 및 9). 진화 시간 보정 후 분석된 계통 발생도에 따르면, 3종의 남극 어류(*C. aceratus*, *Parachaenichthys charcoti* (Charcot 's dragonfish) 및 *Notothenia coriiceps* (bullhead notothen))의 공통 조상은 stickleback 계보 ~ 77 Ma에서 분기되었으며, 남극 빙어는 드래곤 피쉬 계보에서 ~ 7 Ma 전 분기되었음을 확인하였다(Figure 2a 및 Supplementary Figure 7). 유전자계 분석을 통해 9,647개의 유전자계 핵심 세트가 6종의 대표 물고기(3종의 남극 어류, stickleback, 송사리 및 지브라피쉬)에 공통으로 분포함을 확인하였으며, 445개의 남극 빙어 특이적 유전자계를 확인하였다(Supplementary Figure 8).

남극 빙어는 z score 기반 유전자계 확장 및 수축 분석 결과 13종의 경골어류 계놈 내에서 373개의 유전자계가 상당히 확장되었으며, 346개의 유전자계가 상당히 수축된 양상을 나타내었다(Supplementary Table 10 및 11). 남극 빙어 계통은 드래곤피쉬에서 분기된 후 13종의 경골어류 중 가장 큰 유전자계 진화를 겪었다(유전자 획득: 280개의 유전자; 유전자 손실: 6개의 유전자)(Figure 2a 및 Supplementary Table 12). 유전자계 상당수는 감각 인식(보조 참고 2), 산화 환원 효소 활성 및 이온 결합(Supplementary Table 13)에서 확장되었다. 드래곤 피쉬로부터 분기된 후 남극 빙어 계통에서 40개의 유전자가 특이적으로 선택진화를 겪은 것으로 나타났다(Supplementary Table 14). 선택진화된 유전자는 oxidoreductase 활성(헤모글로빈이 없는 특징과 관련이 있을 것으로 추정됨)과 지질 결합(저생 생활 변화에 대한 적응의 일환인 부력 증가와 관련이 있을 것으로 추정)로 나타났다(Supplementary Table 15).

4) 계놈 변이 및 집단 진화

남극 빙어 암컷 계놈 내 9,365,677개의 이형 접합성 단일 뉴클레오티드 다형성을 확인하였으며, 8.79×10^{-3} 의 시퀀싱 된 계놈에서 이형 접합 부위 빈도 다발생을 확인하였고, 이는 대서양 대구(2.09×10^{-3}) 및 바다 개복치 (0.78×10^{-3})와 같은 해양 어류 계놈 대비 매우 큰 것으로 확인되었다[31,32]. Pairwise sequentially Markovian coalescent (PSMC)모델 기반 남극 빙어 집단 통계학적 진화 양상에 있어 큰 두 개의 현상을 확인하였다[33]. 첫째, 남극 해수면 온도가 2.5°C로 내려가고, Plio-Pleistocene 냉각 사건 (3.0-0.9 Ma) 종료부터 ~ 1 Ma시까지 남극 빙어 집단이 최대 크기에 도달 한 것으로 나타났다[34]. Plio-Pleistocene의 느린 냉각기에 이루어진 적응이 남극 빙어의 최대 유효 집

단 크기를 만들었음을 유추할 수 있다. 둘째, 남극 빙어 집단은 mid-Pleistocene transition (~ 1.2 ~ 0.55 Ma)의 온도 변화기에서 감소하는 것으로 나타났으며(Figure 2b), 이 같은 현상은 열에 민감한 남극 빙어의 생리학적 부적응을 유도했을 것이라 여겨진다[35].

5) 항동결단백질과 난보호단백질의 확장

트립시노겐으로부터 유래한 것으로 알려진 항동결단백질(AFGP)은 남극 빙어 유전체 내에서 남극이빨고기(*Dissostichus mawsoni*)에서와 같이 연속하여 복제된 것이 확인되었다(Figure 3a)[36,37]. 남극 어류의 게놈 내에서 AFGP-trypsinogen locus 가 몇몇 농어목 어류의 게놈 내 mitochondrial ribosomal protein L (mrpl)과 E3 ubiquitin-protein ligase CBL (cbl) 사이에 존재하는 trypsinogen gene들과 일치하는 결과를 보였다(Figure 3a). 선례의 연구에서 Illumina sequencing 기반의 유전체 해독 결과 같은 종의 남극 빙어에서 4 개의 AFGP 유전자만이 발견되었다면, 이번 유전체 해독 결과에서는 11개의 AFGP 유전자, 10개의 trypsinogen 유전자, 2개의 trypsinogen-like protease 유전자가 각각 명명되었다[38].

남극 어류의 배아 및 치어에서는 AFGP 유전자의 발현이 나타나지 않는 것으로 알려져 있어 배아 시기에 결빙에 저항하는 방법에 대한 의문이 있었다[12,39]. Zona pellucida egg-coat 단백질은 배아가 성공적으로 수정될 수 있는 기능을 하며 포유류에서는 다정자수정을 막는 기능을 하는 것으로 알려져 있다[12,40]. 또한 어류의 알표면을 단단하게 하고 두껍게 만들어 배아를 보호하는 기능을 한다[41]. 선례 연구 결과에서 남극이빨고기의 *zona pellucida* 단백질이 얼음의 녹는점을 낮추었다는 보고가 있었다[12]. 남극빙어의 게놈에서 *zona pellucida* 유전자군이 확장되어 있는 것을 확인하였다. 총 131개의 *zona pellucida* 유전자 중에서 109개의 유전자가 20개의 contig 내에서 연속하여 복제된 것을 확인하였고(Supplementary Table 16), 유전자 서열 유사도를 기반으로 11개의 subfamily 로 명명하였다(Supplementary Table 17). 또 다른 남극 어류인 남극대구에서는 18개, 드래곤피쉬에서는 30개의 *zona pellucida* 유전자가 발견되었고, 7 종의 비극지어류에서는 16개~35개의 *zona pellucida* 유전자만이 명명되었다[42,43]. 한편 남극이빨고기의 발현유전체 결과에서는 18개의 *zona pellucida* 유전자가 발견되었다[12]. 남극 빙어의 게놈에서 zpax1, zpc1, zpc2, zpc5 유전자가 다른 어류들에 비하여 많은 수가 복제된 것으로 분석되었다(Supplementary Table 16). 예시로, 18개의 zpc5 유전자는 2개의 zpc3 유전자와 함께 1개의 contig (Ice__000281)에 연속하여 존재하는 것을 확인하였다(Figure 3b). 또 다른 contig (Ice__000114)의 경우 가시고기나 송사리와 같은 패턴의 synteny 에서 5개의 zpc3 유전자와 3개의 zpc5 유전자가 확장되어 있는 것을 확인하였다(Supplementary Figure 9).

척추동물에서 *zona pellucida* 유전자는 난소와 간에서 발현하는 것으로 알려져 있는 것과 마찬가지로 남극 빙어의 *zona pellucida* 유전자 역시 난소에서 높은 발현을 보였다(Supplementary Figure 10)[12,44]. 몇 개의 *zona pellucida* 유전자는 남극 빙어의 다른 조직에서도 발현값을 보였는데 난소 이외의 조직에서의 발현은 *zona pellucida* paralogues의 적응성, 즉 남극 빙어 계놈에서 확장된 유전자 군의 특이적 신기능화를 대변하는 가능성을 보여주는 결과라고 여겨진다[45].

6) 산소 결합 단백질

대부분의 경골어류의 유전체는 경골어류 특이적 계놈 복제 이벤트(Teleost genome duplication)를 거침으로 인해 *lcmt1*과 *aqp8*, *rhbdf1b* 유전자가 존재하는 LA cluster와 *mpg* 와 *npr13*, *kank2*를 가지는 MN cluster, 두개의 globin cluster 를 가지는 것으로 알려져 있다[46]. α -globin과 β -globin 을 각각 다른 염색체에서 가지는 포유류와는 다르게 경골어류의 경우, α -globin과 β -globin 이 cluster 내에서 교대로 존재한다[46]. Beta-globin 의 손실 및 위유전자화는 남극 빙어과 어류의 공동파생형질로써 16종의 남극 빙어과 어류 중 15종이 α -globin의 일부만을 가지는 것으로 알려져 있다[2].

남극빙어의 계놈에서 globin 유전자를 확인한 결과, LA cluster 내 α -globin의 흔적이 mapping 된 반면, MN cluster 에서는 globin 유전자 근방의 유전자들은 보존이 되어 있었지만 globin 유전자의 어떠한 흔적도 찾을 수 없었다(Supplementary Figure 11). Myoglobin, cytoglobin, neuroglobin 의 경우 남극 빙어 유전체 내에서 다른 경골어류와 같은 패턴으로 intact 하게 존재하는 것을 확인하였다(Supplementary Figure 12)[47-49]. 이 결과는 myoglobin 이 발현되지 않는 남극빙어의 계놈에 intact 한 유전자를 보유하고 있는 것을 확인한 바, 계놈 내에서 존재하지만 발현을 억제하는 myoglobin 분자 유전 메커니즘에 대한 이해가 필요함을 시사한다.

7) 산화적 스트레스

C. aceratus 를 포함하는 몇 종의 남극 빙어의 경우, 적혈구를 가지는 남극암치아목 어류들에 비하여 산화적 스트레스에 더욱 민감하다[50-52]. 골격 및 근육 세포 단위 부피당 다불포화지방산의 함량이 높은 미토콘드리아 부피가 붉은 피 남극 어류 대비 두배 높은 특성은 남극 빙어가 지질 과산화반응 및 항산화 스트레스 형성에 더욱 취약하게 하였다[52]. 나아가 남극 빙어는 심장 근육의 단백질 증가와 지질 손상에 의해 적혈구를 갖는 남극암치아목 어류에 비해 열적 내성이 낮은 것으로 예상할 수 있다[50]. 남극 빙어는 적혈구를 갖는 남극암치아목 어류에 비하여 미토콘드리아 내 호흡 기작에 의해 더 많은 항산화스

트레스를 생성하게 된다[51]. 마지막으로 항산화물질의 농도가 적혈구를 갖는 남극암치아목 어류에 비하여 낮다는 연구 결과가 있었다[52]. 이러한 결과들을 종합하였을 때 남극 빙어의 경우 항산화 방어 시스템을 강화하기 위한 선택압을 받았을 것이라 예상할 수 있었다[53].

Superoxide dismutase (SOD)와 NAD(P)H:quinone acceptor oxidoreductase (NQO)는 ROS 항상성과 관련된 유전자군으로써 *C. aceratus* 게놈 내에서 확장된 것을 확인하였다(Supplementary table 18). 농어목 내 어류에서는 3개의 SOD 유전자만을 갖는 것으로 알려져 있었으나 본 연구결과를 통해 남극빙어에서 *sod1*, *sod2*, 3개의 *sod3* 유전자를 확인함으로써 총 5개의 SOD 유전자를 명명하였다(Figure 3c, supplementary Table 19). Mutation rate 을 분석 한 결과, SOD3 유전자의 복제가 2백만년 전으로 비교적 최근에 복제되었다는 사실을 알 수 있었다(Figure 3d). SOD3 유전자는 세포 외질에서 단백질로 기능하는 것으로 알려져 있는데(실제로 세포 외 분비를 위한 signal peptide를 가지고 있는 것을 확인하였음) ROS 항상성을 조절하기 위한 세포 외질과 내에서의 SOD 유전자의 기능에 대한 추가 연구가 필요할 것으로 사료된다.

대부분의 어류 유전체 내에 2~10개의 *nqo1* 유전자를 포함하는 것으로 알려진 것에 비하여 남극 빙어의 유전체의 경우, 33개의 유전자가 명명되었다(supplementary Figure 13). 경골어류의 경우 *nqo1* 유전자가 2개의 loci 에 존재하게 되는데 (*vang-nqo1s-ackr3* and *il17-nqo1-gabarapl*) 남극 빙어 역시 2개의 보존된 synteny를 가지고 있었다. 추가적으로 *C. aceratus* 의 경우 3개의 contig 에 26개의 추가적인 *nqo1* 유전자를 가지는 것으로 확인되었다(Supplementary Figure 13). 또한 산화적 스트레스에 의해 손상된 DNA 교정하는 기능을 하는 것으로 알려져 있는 8-oxoguanine DNA glycosylase (*ogg1*) 유전자는 모든 경골어류에서 1개만이 알려져 있는 반면, 남극 빙어의 게놈에는 2개의 유전자가 분석되었다(Supplementary Fig 14).

8) 극단적인 광주기에 의한 생체리듬 변화

극지 생물들은 극단적인 일조시간을 갖는 환경에 서식하며 생체리듬과 관련된 유전자들의 기능 및 역할이 변화하였을 것이다. Cryptochrome 과 period 유전자는 각 유전자들의 발현에 의한 순환 피드백 시스템에 의해 생체 리듬을 조절하는 기능을 한다[54]. 경골어류의 경우 5개(가시고기, 메다카)부터 7개(제브라피쉬)까지 다양한 개수의 cryptochrome 유전자를 가지는 것으로 알려져 있다[55]. 남극빙어의 경우 cryptochrome과 period 유전자의 synteny는 보존되어 있음에도 불구하고 *cry1*, *cry2*, *per2a*, *per3* 유전자가 손실된 것을 확인하였다(Figure 4a and Supplementary Table 20). 남극 빙어의 게놈에는 어류 중 가장 적은 개수인 3개의 *cry* 유전자만이 존재하고 있었다. 남극 대구와 드래곤

피쉬 게놈의 경우 불완전한 어셈블 결과 때문에 synteny를 모두 확인하지 못하였지만, 그럼에도 불구하고 두 종의 남극 어류 게놈 내에서 남극 빙어와 같은 유전자가 손실된 것을 확인하였다. 이러한 결과는 겨울의 극야와 여름의 백야라는 극단적인 광주기가 남극 어류의 생체리듬 조절기의 효용성을 감소시켰으며 생체 리듬 조절과 관련된 유전자들의 변화를 유발했을 것이라는 가설에 대한 증거가 된다. 이 가설을 입증하기 위하여 남극 빙어를 포함한 남극암치아목 어류들의 행동학적 연구가 필요할 것이다.

4. 연구 개발 수행 방법

1) 샘플 채집 방법

체장 30cm 크기의 남극 빙어(Supplementary Fig. 1)는 2017년 1월 세종과학기지 근처 수심 20-30m의 마리아나소만에서 채집하였다 ($62^{\circ} 14' S$, $58^{\circ} 47' W$) (Supplementary Fig. 2). Genomic DNA 는 1 개체에서부터 추출하였다. Linkage map 을 작성하기 위한 추가 유전체 해독 및 miRNA 해독에 사용된 생물은 2012-2015년 수심 160-200m, Low Island (Antarctic Specially Protected Area 152, Western Bransfield Strait; latitudes $63^{\circ} 15' S - 63^{\circ} 30' S$; longitudes $62^{\circ} 00' W - 62^{\circ} 45' W$, bounded on the northeast by Low Island)) 에서 Antarctic Research and Supply Vessel Laurence M. Gould의 저인망 채집 시스템을 이용하여 채집하였다[56].

2) 유전체 시퀀싱 및 어셈블리

남극 빙어의 genomic DNA 는 한 마리의 암컷의 근육조직으로부터 추출하였다 (Supplementary Figure 1). Genomic DNA 라이브러리는 제조사의 지침에 따라 제작되었으며 PacBio Sequel System을 이용하여 시퀀싱 하였다. FALCON-Unzip 어셈블러를 이용하여 de novo 어셈블을 수행하였다[57]. 유전체의 완성도를 평가하기 위하여 BUSCO를 수행하였다[21]. 유전체 크기를 예측하기 위하여 Jellyfish를 사용하여 k-mer 분석을 수행하였다 (Supplementary Table 3)[58].

3) 조직 특이적 RNA 시퀀싱

뇌, 안구, 아가미, 심장, 장, 신장, 간, 근육, 난소, 피부, 비장, 위 총 12개의 서로 다른 조직에서 RNeasy mini kit(Qiagen)를 이용하여 RNA를 추출하였다 (Supplementary Table3). RNA 의 품질은 Agilent Bioanalyzer 를 이용하여 측정하였고 (Integrity value 8 이상)발현유전체 라이브러리 제작 및 시퀀싱은 Illumina Hiseq 2500 시스템을 사용하였다. 각각의 조직에서 발현된 발현유전체 정보는 TopHat를 이용하여 게놈에 매핑하였고 발현유전체 정보는

Cufflinks를 활용하여 분석되었다[59]

4) 유전자 명명

MAKER 프로그램을 이용하여 게놈 명명을 수행하였다[60]. RECON과 RepeatScout를 포함하는 RepeatModeler (version 1.0.3)을 활용하여 repeat library를 구축하였다[61-63]. simple repeats, satellites, low-complexity repeat은 Tandem Repeats Finder를 이용하여 예측하였다[63,64]. Repetitive element는 RepeaMasker를 이용하여 분석하였다[64]. 경골어류내 reference 단백질로는 비극지어류 10종(*Danio rerio*, *Gasterosteus aculeatus*, *Gadus morhua*, *Tetraodon nigroviridis*, *Takifugu rubripes*, *Astyanax mexicanus*, *Oryzias latipes*, *Poecilia formosa*, *Oreochromis niloticus* and *Xiphophorus maculatus*), 남극 어류 2종(*P. charcoti* and *N. coriiceps*)의 게놈이 활용되었다 (Supplementary Table3).

5) 비교유전학적 분석

Markov clustering algorithm을 적용한 OrthoMCL pipeline을 이용하여 orthologous gene cluster를 확인하였다[75]. 모든 단계에서 기본 옵션으로 13종의 게놈을 분석하였다(Supplementary Table 8). Protein coding gene은 Probabilistic Alignment Kit (PRANK)를 사용하여 정렬되었다[76]. 정렬이 되지 않은 gap 에 대해서는 Gblocks를 사용하여 제거하였다[77]. Bootstrap 1000회 옵션으로 RAxML을 이용하여 Maximum-likelihood 방법으로 계통도를 제작하였고 분지시점은 PAML 패키지내 MCMCtree를 이용하여 보정되었다 (median estimates of the pairwise divergence time for *D. rerio* and *G. morhua*: 230.4 Ma)[78,79].

7) 유전자군 분석

남극 빙어 특이적 형태, 생리활성, 생태와 관련된 유전자를 찾기 위하여 몇 종의 경골어류에서 밝혀진 선례 연구 및 유전체 서열을 이용하여 local blast 후 매 뉴얼하게 유전자를 확인하였다. 단백질 서열의 경우 Clustal Omega와 MUSCLE 프로그램을 이용하여 비교하였다[80,81]. FastTree 또는 RAxML을 이용하여 선정 유전자의 계통도를 작성하였다[82]. $\mu = D/2t = 3.28 \times 10^{-9}$ 의 mutation rate을 적용하여 유전자 복제 시기를 측정하였다.

5. 참고문헌

1. Ruud, J. T. Vertebrates without erythrocytes and blood pigment. Nature 173, 848 (1954).

2. Near, T. J., Parker, S. K. & Detrich, H. W., 3rd. A genomic fossil reveals key steps in hemoglobin loss by the Antarctic icefishes. *Mol. Biol. Evol.* 23, 2008-2016 (2006).
3. Egginton, S. Stress response in two Antarctic teleosts (*Notothenia coriiceps* Richardson and *Chaenocephalus aceratus* Lönnberg) following capture and surgery. *J. Comp. Physiol. B* 164, 482-491 (1994).
4. Sidell, B. D. & O'Brien, K. M. When bad things happen to good fish: the loss of hemoglobin and myoglobin expression in Antarctic icefishes. *J. Exp. Biol.* 209, 1791-1802 (2006).
5. Hemmingsen, E. A. in *Biology of Antarctic Fish* (ed G. di Prisco, Maresca, B., Tota, B.) 191 - 203 (Springer-Verlag, 1991).
6. Egginton, S. Blood rheology of Antarctic fishes: viscosity adaptations at very low temperatures. *J. Fish Biol.* 48, 513-521 (1996).
7. Kennett, J. P. Cenozoic evolution of Antarctic glaciation, the circum-Antarctic Ocean, and their impact on global paleoceanography. *J. Geophys. Res.* 82, 3843-3860 (1977).
8. Eastman, J. T. *Antarctic Fish Biology - Evolution in a Unique Environment.* (Academic Press, San Diego, 1993).
9. DeVries, A. The role of antifreeze glycopeptides and peptides in the freezing avoidance of Antarctic fishes. *Comp. Biochem. Physiol.* 90B, 611 - 621 (1988).
10. Cheng, C. H. & Detrich, H. W., 3rd. Molecular ecophysiology of Antarctic notothenioid fishes. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 362, 2215-2232 (2007).
11. Cao, L. et al. Neofunctionalization of zona pellucida proteins enhances freeze-prevention in the eggs of Antarctic notothenioids. *Nat. Commun.* 7, 12987 (2016).
12. Hofmann, G. E., Buckley, B. A., Airaksinen, S., Keen, J. E. & Somero, G. N. Heat-shock protein expression is absent in the Antarctic fish *Trematomus bernacchii* (family Nototheniidae). *J. Exp. Biol.* 203, 2331-2339 (2000).
13. Place, S. P. & Hofmann, G. Constitutive expression of a stress-inducible heat shock protein gene, hsp70, in phylogenetically distant Antarctic fish. *Polar*

- Biol. 28, 261-267 (2005).
14. Near, T. J., Jones, C. D. & Eastman, J. T. Geographic intraspecific variation in buoyancy within Antarctic notothenioid fishes. *Antarct. Sci.* 21, 123 - 129 (2009).
 15. Eastman, J. T., Witmer, L. M., Ridgely, R. C. & Kuhn, K. L. Divergence in skeletal mass and bone morphology in Antarctic notothenioid fishes. *J. Morphol.* 275, 841-861 (2014).
 16. Albertson, R. C. et al. Molecular pedomorphism underlies craniofacial skeletal evolution in Antarctic notothenioid fishes. *BMC Evol. Biol.* 10, 4 (2010).
 17. Hagen, W., Kattner, G. & Friedrich, C. The lipid compositions of high-Antarctic notothenioid fish species with different life strategies. *Polar Biol.* 23 (2000).
 18. Kock, K. H. Antarctic icefishes (Channichthyidae): a unique family of fishes. A review, Part I. *Polar Biol.* 28, 862 - 895 (2005).
 19. Hu, Y. et al. Evolution in an extreme environment: developmental biases and phenotypic integration in the adaptive radiation of Antarctic notothenioids. *BMC Evol. Biol.* 16, 142 (2016).
 20. Hemmingsen, E. A. *Biology of Antarctic Fish*, 191-203 (Springer Berlin Heidelberg, 1991).
 21. Auvinet, J. et al. Mobilization of retrotransposons as a cause of chromosomal diversification and rapid speciation: the case for the Antarctic teleost genus *Trematomus*. *BMC Genomics* 19, 339 (2018).
 22. Braasch, I. et al. A new model army: Emerging fish models to study the genomics of vertebrate Evo Devo. *J. Exp. Zool. B Mol. Dev. Evol.* 324, 316-341 (2015).
 23. Amores, A., Catchen, J., Ferrara, A., Fontenot, Q. & Postlethwait, J. H. Genome evolution and meiotic maps by massively parallel dna sequencing: spotted gar, an outgroup for the teleost genome duplication. *Genetics* 188, 799-808 (2011).
 24. Catchen, J. M., Amores, A., Hohenlohe, P. A., Cresko, W. A. &

- Postlethwait, J. H. Stacks: de novo genotype analyses from short-read sequence data. *G3 (Bethesda)* 1, 1-10 (2011).
25. Morescalchi, A. et al. A multiple sex-chromosome system in Antarctic ice-fishes. *Polar Biol.* 11, 655-661 (1992).
26. Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W. & Postlethwait, J. H. Stacks: building and genotyping loci de novo from short-read sequences. *G3 (Bethesda)* 1, 171-182 (2011).
27. Steinke, D., Salzburger, W. & Meyer, A. Novel relationships among ten fish model species revealed based on a phylogenomic analysis using ESTs. *J. Mol. Evol.* 62, 772-784 (2006).
28. Scherthan, H. in *eLS* (2012).
29. Amores, A., Wilson, C. A., Allard, C. A. H., Detrich, H. W. & Postlethwait, J. H. Cold fusion: massive karyotype evolution in the Antarctic bullhead notothen *Notothenia coriiceps*. *G3 (Bethesda)* 7, 2195-2207 (2017).
30. Star, B. et al. The genome sequence of Atlantic cod reveals a unique immune system. *Nature* 477, 207-210 (2011).
31. Pan, H. et al. The genome of the largest bony fish, ocean sunfish (*Mola mola*), provides insights into its fast growth rate. *GigaScience* 5, 36 (2016).
32. Li, H. & Durbin, R. Inference of human population history from individual whole-genome sequences. *Nature* 475, 493 (2011).
33. McKay, R. et al. Antarctic and Southern Ocean influences on Late Pliocene global cooling. *Proc. Natl. Acad. Sci. U.S.A.* 109, 6423-6428 (2012).
34. Hayward, B. W., Kawagata, S., Grenfell, H. R., Sabaa, A. T. & O'Neill, T. Last global extinction in the deep sea during the mid Pleistocene climate transition. *Paleoceanography* 22, PA3103 (2007).
35. Chen, L., DeVries, A. L. & Cheng, C.-H. C. Evolution of antifreeze glycoprotein gene from a trypsinogen gene in Antarctic notothenioid fish. *Proc. Natl. Acad. Sci. U.S.A.* 94, 3811-3816 (1997).
36. Nicodemus-Johnson, J., Silic, S., Ghigliotti, L., Pisano, E. & Cheng, C. H. C. Assembly of the antifreeze glycoprotein/trypsinogen-like protease genomic

- locus in the Antarctic toothfish *Dissostichus mawsoni* (Norman). *Genomics* 98, 194-201 (2011).
37. Wassarman, P. M. Zona Pellucida Glycoproteins. *J. Biol. Chem.* 283, 24285-24289 (2008).
38. Sano, K. et al. Comparison of egg envelope thickness in teleosts and its relationship to the sites of zp protein synthesis. *J. Exp. Zool. B Mol. Dev. Evol.* 328, 240-258 (2017).
39. Shin, S. C. et al. The genome sequence of the Antarctic bullhead notothen reveals evolutionary adaptations to a cold environment. *Genome Biol.* 15, 468 (2014).
40. Ahn, D.-H. et al. Draft genome of the Antarctic dragonfish, *Parachaenichthys charcoti*. *GigaScience* 6, 1-6 (2017).
41. Wu, T. et al. Bioinformatic analyses of zona pellucida genes in vertebrates and their expression in Nile tilapia. *Fish Physiol. Biochem.* 44, 435-449 (2018).
42. Opazo, J. C., Butts, G. T., Nery, M. F., Storz, J. F. & Hoffmann, F. G. Whole-genome duplication and the functional diversification of teleost fish hemoglobins. *Mol. Biol. Evol.* 30, 140-153 (2013).
43. Zhao, Y. et al. The major adult α -globin gene of Antarctic teleosts and its remnants in the hemoglobinless icefishes: calibration of the mutational clock for nuclear genes. *J. Biol. Chem.* 273, 14745-14752 (1998).
44. Mueller, I. A. et al. Exposure to critical thermal maxima increases oxidative stress in hearts of white- but not red-blooded Antarctic notothenioid fishes. *J. Exp. Biol.* 215, 3655-3664 (2012).
45. O'Brien, K. M. & Mueller, I. A. The unique mitochondrial form and function of Antarctic channichthyid icefishes. *Integr. Comp. Biol.* 50, 993-1008 (2010).
46. Klein, R. D. et al. Antioxidant defense system and oxidative status in Antarctic fishes: The sluggish rockcod *Notothenia coriiceps* versus the active marbled notothen *Notothenia rossii*. *J. Therm. Biol.* 68, 119-127 (2017).
47. Shearman, L. P. et al. Interacting molecular loops in the mammalian circadian clock. *Science* 288, 1013-1019 (2000).

48. Liu, C. et al. Molecular evolution and functional divergence of zebrafish (*Danio rerio*) cryptochrome genes. *Sci. Rep.* 5, 8113 (2015).
49. Lisiecki, L. E. & Raymo, M. E. A Pliocene Pleistocene stack of 57 globally distributed benthic $\delta^{18}\text{O}$ records. *Paleoceanography* 20, PA1003 (2005).
50. Jouzel, J. et al. Orbital and millennial Antarctic climate variability over the past 800,000 years. *Science* 317, 793-796 (2007).
51. Chin, C.-S. et al. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat. methods* 10, 563-569 (2013).
52. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210-3212 (2015).
53. Marçais, G. & Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* 27, 764-770 (2011).
54. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105-1111 (2009).
55. Holt, C. & Yandell, M. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC bioinformatics* 12, 491 (2011).
56. Bao, Z. & Eddy, S. R. Automated de novo identification of repeat sequence families in sequenced genomes. *Genome Res.* 12, 1269-1276 (2002).
57. Price, A. L., Jones, N. C. & Pevzner, P. A. De novo identification of repeat families in large genomes. *Bioinformatics* 21, i351-i358 (2005).
58. Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27, 573 (1999).
59. Smit, A., Hubley, R. & Green, P. RepeatMasker Open-3.0, <<http://www.RepeatMasker.org>> (2017).
60. Korf, I. Gene finding in novel genomes. *BMC bioinformatics* 5, 59 (2004).
61. Nawrocki, E. P., Kolbe, D. L. & Eddy, S. R. Infernal 1.0: inference of RNA alignments. *Bioinformatics* 25, 1335-1337 (2009).

62. Gardner, P. P. et al. Rfam: Wikipedia, clans and the “decimal” release. *Nucleic Acids Res.* 39, D141-D145 (2010).
63. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25, 955-964 (1997).
64. Desvignes, T., Detrich, H. W. & Postlethwait, J. H. Genomic conservation of erythropoietic microRNAs (erythromiRs) in white-blooded Antarctic icefish. *Mar. Genomics* 30, 27-34 (2016).
65. Li, L., Stoeckert, C. J. & Roos, D. S. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome res.* 13, 2178-2189 (2003).
66. Löytynoja, A. & Goldman, N. An algorithm for progressive multiple alignment of sequences with insertions. *Proc. Natl. Acad. Sci. U.S.A.* 102, 10557-10562 (2005).
67. Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17, 540-552 (2000).
68. Hedges, S. B., Dudley, J. & Kumar, S. TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics* 22, 2971-2972 (2006).
69. Yang, Z. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol. Biol. Evol.* 24, 1586-1591 (2007).
70. Han, M. V., Thomas, G. W., Lugo-Martinez, J. & Hahn, M. W. Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Mol. Biol. Evol.* 30, 1987-1997 (2013).
71. Zhang, G. et al. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* 346, 1311-1320 (2014).
72. Conesa, A. et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674-3676 (2005).
73. Zhang, G. et al. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* 346, 1311 - 1320 (2014).
74. Conesa, A. et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674 - 3676 (2005).

75. Li, C. et al. Two Antarctic penguin genomes reveal insights into their evolutionary history and molecular changes related to the Antarctic environment. *Gigascience* 3, 27 (2014).
76. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows - Wheeler transform. *Bioinformatics* 25, 1754 - 1760 (2009).
77. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792-1797 (2004).
78. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 - approximately maximum-likelihood trees for large alignments. *PloS one* 5, e9490 (2010).
79. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792 - 1797 (2004).
80. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5, e9490 (2010).
81. Lisiecki, L. E. & Raymo, M. E. A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records. *Paleoceanography* 20, PA1003 (2005).
82. Jouzel, J. et al. Orbital and millennial Antarctic climate variability over the past 800,000 years. *Science* 317, 793 - 796 (2007).

Table 1 | Icefish assembly and annotation statistics

Assembly	
Number of contigs	3,852
Total genome length from contigs (bp)	1,065,645,509
Longest contig (bp)	9,422,831
N50 contig length (bp)	1,500,626
Annotation	
Number of genes	30,773
Exon number	277,249
Total length of exons (bp)	50,279,998
Total length of repeats (bp)	523,290,133
G+C (%)	42.08



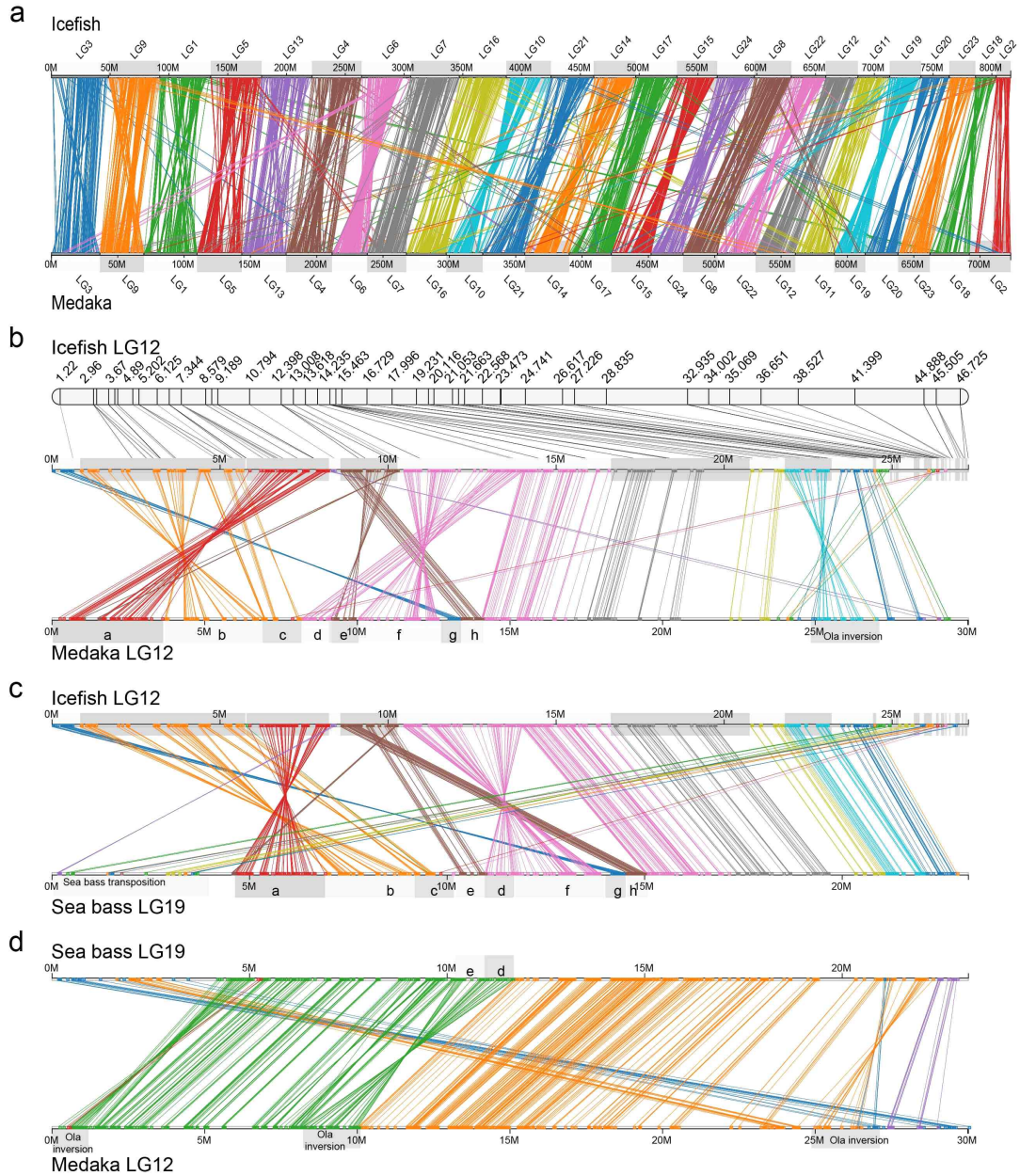


그림 1 Karyotype stability of icefish with respect to teleost outgroups. a. Gene content in icefish chromosomes supports a one-to-one correspondence between icefish and medaka chromosomes. Each line represents orthologous genes in icefish and medaka, color coded by icefish chromosome. The few lines that cross linkage groups likely represent incorrect orthology calls. b. A comparison of orthologous gene orders in icefish LG12 (Cac12) and medaka LG12 (Ola12) illustrates icefish-specific chromosome inversions and transpositions (see text). Each line represents orthologous genes in the icefish and medaka chromosome, color coded by icefish genomic scaffold. Conserved syntenic blocks are labelled a-h. c. Comparison of orthologous gene order in Cac12 and sea bass LG19. Conserved syntenic blocks are labelled a-h. d. Comparison of orthologous gene order between sea bass LG19 and medaka LG12 reveals that most chromosome rearrangements occurred after the divergence of the icefish lineage from the sea bass lineage.

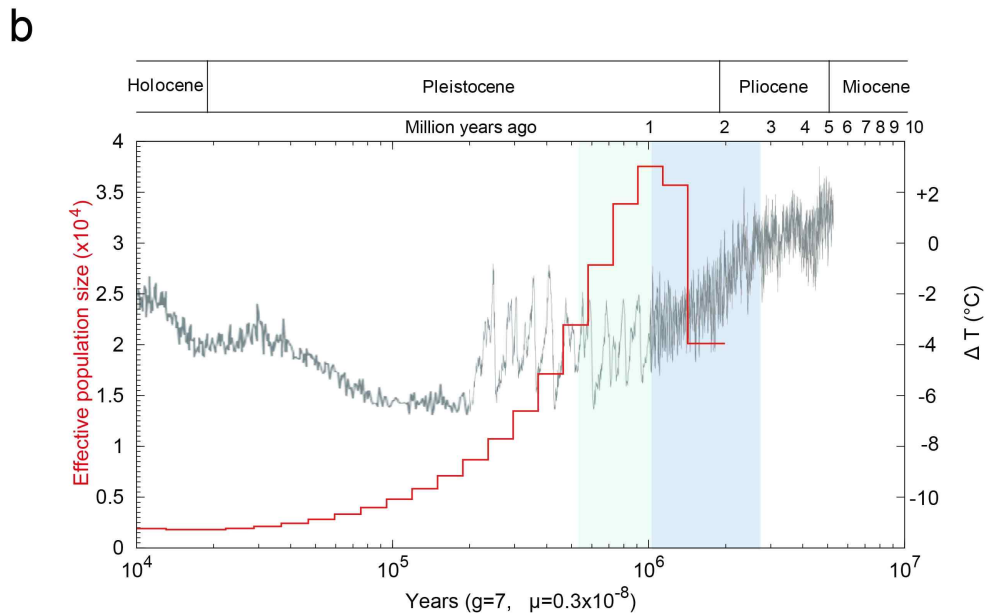
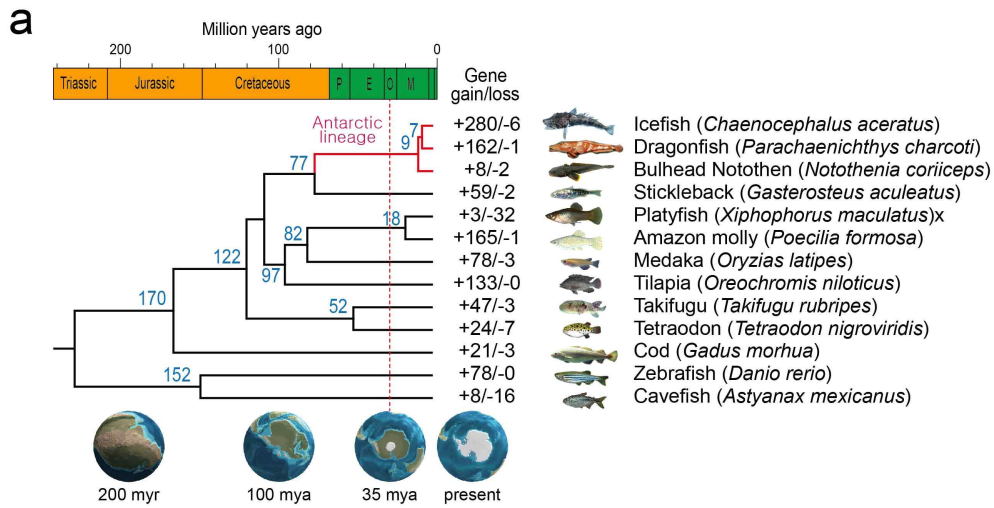


그림 2. Comparative analysis of the *C. aceratus* genome assembly. a. Phylogenetic tree and gene family gain-and-loss analysis including the number of gained gene families (+), lost gene families (-). Blue numbers specify divergence times between lineages. The red dotted line indicates the appearance of Antarctic ice sheets (35 Mya), which allowed the Circum-Antarctic Current to form after the opening of the Drake Passage. Subsequent cooling of the Southern Ocean drove the local extinction of most fish taxa and the adaptive radiation of Antarctic notothenioid suborder. b. Inferring icefish population history by pairwise sequential Markovian coalescent (PSMC) analysis. Left Y-axis represents the demographic history of *C. aceratus* (red line). During the Plio-Pleistocene period (3-0.9 Mya), which is shaded blue, Antarctic sea surface temperatures dropped by around 2.5°C , judged by a proxy for marine paleo-temperature changes based on oxygen isotope ratios^{49,50} (right Y-axis): concomitant decreases in marine temperatures (black line) probably allowed the cold-adapted *C. aceratus* populations to increase in size. Green shading represents the mid-Pleistocene transition (MPT), during which temperature fluctuations were large.



그림 3. Conserved synteny for expanded gene clusters identified in the blackfin icefish genome. a. The antifreeze glycoprotein (afgp) and trypsinogen (tryp) locus. b. The zona pellucida c5 (zpc5) locus. c. The superoxide dismutase 3 (sod3) gene cluster. Genomic neighborhoods are shown within representative sequenced teleost genomes. Each arrow indicates a complete gene oriented in the (5'→3') direction. The pink-shaded area in panel a indicates the trypsinogen gene locus. d Phylogenetic analysis of vertebrate sod3 genes. Divergence time was calculated applying mutation rate formula $\mu = D/2t = 3.28e-9$. Abbreviation: Ca, *C. aceratus*, icefish; Dr, *D. rerio*, zebrafish; Ga, *G. aculeatus*, stickleback; Nc, *N. coriiceps*, bullhead notothen; Ol, *O. latipes*, medaka; Pf, *P. Formosa*, Amazon molly; Tr, *T. rubripes*, fugu; Xm, *X. maculatus*, platyfish.



그림 4. Genomic evidence supporting gene loss events for blackfin icefish circadian rhythm-related genes. Genomic structures and syntenic comparisons of a. cryptochrome gene cry1ab and b. cry2. c. period gene per2 and d. per3 gene clusters are shown within representative sequenced teleost genomes.

Supplementary Table 1. Statistics for icefish genome sequencing using PacBio Sequel.

Library	No. of Cells	No. of Reads	Total data (Mbp)	N50 read length (bp)	Mean read length (bp)
Polymerase read	16	6,562,134	89,275	22,250	13,605
Subread		8,434,974	89,045	16,203	10,557



Supplementary Table 2. BUSCO scores for the completeness of the icefish genome sequence.
(<http://busco.ezlab.org>)

	Actinopterygii Gene Set	%
Complete BUSCOs (C)	4,123	89.9
Complete and single-copy BUSCOs (S)	3,930	85.7
Complete and duplicated BUSCOs (D)	193	4.2
Fragmented BUSCOs (F)	164	3.6
Missing BUSCOs (M)	297	6.5
Total BUSCO groups searched	4,584	



Supplementary Table 3. Sequencing statistics for icefish transcriptome analysis.

Organ	No. reads	Total data (bp)	% reads mapping*
Brain	42,517,602	4,294,277,802	94.76
Eye	48,799,334	4,928,732,734	95.45
Gill	51,072,374	5,158,309,774	94.35
Heart	56,781,088	5,734,889,888	92.23
Intestine	51,113,066	5,162,419,666	96.66
Kidney	53,863,870	5,440,250,870	90.35
Liver	78,598,880	7,938,486,880	86.68
Muscle	45,374,108	4,582,784,908	87.23
Ovary	49,080,678	4,957,148,478	91.09
Skin	45,497,130	4,595,210,130	89.49
Spleen	43,778,336	4,421,611,936	89.57
Stomach	57,809,718	5,838,781,518	86.44

*Data from each organ was mapped to assembled icefish genome using Tophat.



Supplementary Table 4. General statistics of the functional annotation.

		Number	Percent (%)
Total		30,773	
Annotated	Blastp	23,923	77.74
	Interpro	20,684	67.21
	GO	16,671	54.17
	KEGG	19,893	64.64



Supplementary Table 5. Sequencing statistics for icefish small-RNA transcriptome analysis.

Organ	Number of reads retained for analysis¹	Number of reads being annotated	% reads being annotated
Pronephric kidney	4,818,658	4,281,988	88.86
Pectoral girdle bone	9,459,845	7,820,112	82.67
Heart ventricle	596,018	545,265	91.48
Pectoral adductor muscle	2,715,548	2,343,132	86.29
Skeletal muscle	5,750,328	4,822,711	83.87

¹Passed sequence quality filter; passed size filter; have minimum of 5 counts



Supplementary Table 6. Information of miRNA sequence of blackfin icefish.

Detailed information is appended in Supplementary Table 6.xlsx

Supplementary Table 7. Statistics for the analysis of icefish repetitive elements.

	Repeats	Number	Length occupied	Percentage genome (%)		
Transposable element	LTR retrotransposon	LTR	1,512	154,453	0.01	
		LTR/Copia	1,423	1,254,557	0.12	
		LTR/DIRS	28,935	21,963,956	2.06	
		LTR/ERV1	14,347	9,558,534	0.90	
		LTR/ERVK	953	147,808	0.01	
		LTR/Gypsy	54,611	21,218,816	1.99	
		LTR/Ngaro	3,291	1,806,225	0.17	
		LTR/Pao	6,780	3,393,081	0.32	
		Total	111,852	59,497,430	5.58	
	non-LTR retrotransposon	SINE	15,410	2,555,151	0.24	
		SINE/5S	1,996	186,954	0.02	
		SINE/ID	563	47,443	0.00	
		SINE/MIR	19,288	3,038,263	0.29	
		SINE/tRNA	570	206,591	0.02	
		Total	37,827	6,034,402	0.57	
		LINE	10,421	5,399,187	0.51	
		LINE/CR1	3,384	1,119,014	0.11	
		LINE/I	8,488	2,563,237	0.24	
		LINE/Jockey	5,508	1,365,271	0.13	
		LINE/L1	8,227	3,347,285	0.31	
		LINE/L2	100,525	40,927,414	3.84	
		LINE/Penelope	1,301	284,035	0.03	
		LINE/Proto2	2,374	780,797	0.07	
		LINE/Rex	72,096	16,705,191	1.57	
		LINE/RTE	8,466	4,759,124	0.45	
		Total	220,790	77,250,555	7.25	
		DNA transposon	DNA	116,939	25,573,600	2.40
			DNA/Academ	3,694	896,244	0.08
			DNA/CMC	13,082	3,359,603	0.32
			DNA/Crypton	9,700	2,343,498	0.22
			DNA/Ginger	601	194,239	0.02
			DNA/Harbinger	2,760	1,440,879	0.14
			DNA/hAT	332,737	93,810,547	8.80
		DNA/Kolobok	8,025	2,297,447	0.22	
		DNA/Maverick	10,958	3,101,964	0.29	
		DNA/Merlin	291	111,556	0.01	
		DNA/MULE	250	19,659	0.00	
		DNA/Novosib	2,194	641,603	0.06	
		DNA/P	4,543	2,087,502	0.20	
		DNA/PIF	32,016	12,688,102	1.19	
		DNA/PiggyBac	7,135	2,199,829	0.21	
		DNA/Sola	484	201,769	0.02	
		DNA/TcMar	15,385	7,248,703	0.68	
		DNA/Zisupton	889	198,058	0.02	
		RC Helitron	2,519	2,163,670	0.20	
	Total	564,202	160,578,472	15.07		
	Unclassified	628,836	201,836,786	18.94		
Total Interspersed		1,563,507	505,197,645	47.41		
	Small RNA	1,074	454,551	0.04		
	Satellites	5,834	2,808,492	0.26		
	Simple repeats	324,248	26,602,004	2.50		
	Low complexity	26,275	1,647,308	0.15		
Total			536,710,000	50.36		

Supplementary Table 8. Genome assemblies/Gene models used in this study.

Common name	Species	Assembly/Gene model
Icefish	<i>Chaenocephalus aceratus</i>	This study
Dragonfish	<i>Parachaenichthys charcoti</i>	PRJNA330735
Bullhead notothen	<i>Notothenia coriiceps</i>	GCA 000735185.1
Stickleback	<i>Gasterosteus aculeatus</i>	BROAD S1
Platyfish	<i>Xiphophorus maculatus</i>	Xipmac4.4.2
Amazon molly	<i>Poecilia formosa</i>	Poecilia formosa-5.1.2
Medaka	<i>Oryzias latipes</i>	HdrR
Tilapia	<i>Oreochromis niloticus</i>	Orenil1.0
Takifugu	<i>Takifugu rubripes</i>	FUGU 5.0
Tetraodon	<i>Tetraodon nigroviridis</i>	TETRAODON 8.0
Cod	<i>Gadus morhua</i>	gadMor1
Zebrafish	<i>Danio rerio</i>	GRCz10
Cavefish	<i>Astyanax mexicanus</i>	AstMex102



Supplementary Table 9. Summary of orthologous gene clusters analyzed in 13 species.

Species name	No. of coding genes	No. of gene families	No. of genes in gene families	No. of orphan genes	No. of unique gene families	Ave. No. of genes in gene families
Icefish	30,773	18,636	24,159	6,614	581	1.3
Dragonfish	32,713	21,417	26,550	6,163	429	1.2
Bullhead notothen	32,635	20,153	23,387	9,248	34	1.2
Stickleback	20,786	16,131	19,601	1,185	48	1.2
Platyfish	20,379	16,966	19,748	631	6	1.2
Amazon molly	23,615	18,019	22,488	1,127	116	1.2
Medaka	19,699	14,981	18,136	1,563	132	1.2
Tilapia	21,437	16,282	20,541	896	139	1.3
Takifugu	18,523	14,767	17,747	776	47	1.2
Tetraodon	19,602	14,835	17,783	1,819	47	1.2
Cod	20,094	15,292	18,107	1,987	72	1.2
Zebrafish	25,397	16,334	22,815	2,582	320	1.4
Cavefish	23,042	16,014	19,522	3,520	124	1.2

Supplementary Table 10. Gene Ontology of expanded genes families in the icefish genome relative to the 13 species*.

Category	GO_acc	Term	p-value
P	GO:0043623	cellular protein complex assembly	6.40E-07
P	GO:0050906	detection of stimulus involved in sensory perception	2.00E-06
P	GO:0034622	cellular macromolecular complex assembly	2.20E-05
P	GO:0007606	sensory perception of chemical stimulus	3.00E-05
P	GO:0007600	sensory perception	2.90E-05
P	GO:0006461	protein complex assembly	4.80E-05
P	GO:0065003	macromolecular complex assembly	0.00023
P	GO:0071205	protein localization to juxtaparanode region of axon	0.029
P	GO:0007626	locomotory behavior	0.0092
P	GO:0034728	nucleosome organization	0.04
P	GO:0032507	maintenance of protein location in cell	0.04
P	GO:0042742	defense response to bacterium	0.027
P	GO:0065004	protein-DNA complex assembly	0.013
P	GO:0000050	urea cycle	0.004
P	GO:0019627	urea metabolic process	0.004
P	GO:0071466	cellular response to xenobiotic stimulus	0.0031
P	GO:0002460	adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	0.028
P	GO:0035113	embryonic appendage morphogenesis	0.044
P	GO:0045185	maintenance of protein location	0.04
P	GO:0006935	chemotaxis	0.024
P	GO:0043604	amide biosynthetic process	0.017
P	GO:0002449	lymphocyte mediated immunity	0.0048
F	GO:0019001	guanyl nucleotide binding	9.30E-07
F	GO:0004984	olfactory receptor activity	1.80E-06
F	GO:0017111	nucleoside-triphosphatase activity	0.001
F	GO:0004053	arginase activity	0.0011
F	GO:0004190	aspartic-type endopeptidase activity	0.0014
F	GO:0070001	aspartic-type peptidase activity	0.0014
F	GO:0016818	hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides	0.0019
F	GO:0016462	pyrophosphatase activity	0.0018
F	GO:0005506	iron ion binding	0.002
F	GO:0005093	Rab GDP-dissociation inhibitor activity	0.004
F	GO:0016813	hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds, in linear amidines	0.0091
F	GO:0008158	hedgehog receptor activity	0.01
F	GO:0004617	phosphoglycerate dehydrogenase activity	0.01

F	GO:0035004	phosphoinositide 3-kinase activity	0.04
F	GO:0004842	ubiquitin-protein ligase activity	0.012
F	GO:0003964	RNA-directed DNA polymerase activity	0.014
F	GO:0003796	lysozyme activity	0.017
F	GO:0020037	heme binding	0.027
F	GO:0003779	actin binding	0.033
F	GO:0019787	small conjugating protein ligase activity	0.022
F	GO:0032549	ribonucleoside binding	0.029
F	GO:0016881	acid-amino acid ligase activity	0.025
F	GO:0005540	hyaluronic acid binding	0.016
F	GO:0016303	1-phosphatidylinositol-3-kinase activity	0.04
F	GO:0004601	peroxidase activity	0.047
C	GO:0034702	ion channel complex	0.029
C	GO:0034703	cation channel complex	0.012
C	GO:0005581	collagen	0.014
C	GO:0043232	intracellular non-membrane-bounded organelle	0.01
C	GO:0005577	fibrinogen complex	0.0091
C	GO:0005604	basement membrane	0.024

* Species list are Supplementary Table 6.

극지연구소

Supplementary Table 11. Gene Ontology for contracted gene families in the icefish genome relative to the 13 species*.

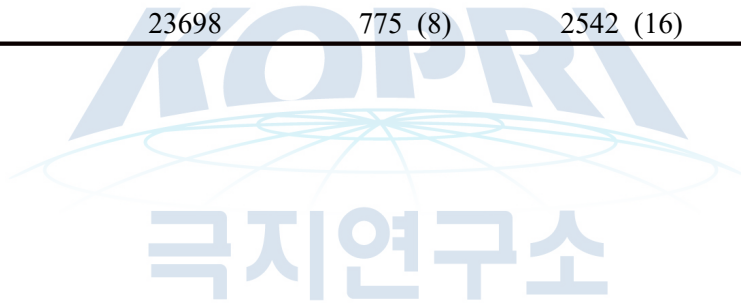
Category	GO acc	Term	p-value
P	GO:0046835	carbohydrate phosphorylation	1.90E-13
P	GO:0006003	fructose 2,6-bisphosphate metabolic process	1.30E-12
P	GO:0006000	fructose metabolic process	6.70E-10
P	GO:0044237	cellular metabolic process	4.60E-08
P	GO:0010467	gene expression	1.30E-07
P	GO:0006004	fucose metabolic process	4.00E-07
P	GO:0005996	monosaccharide metabolic process	5.20E-07
P	GO:0044260	cellular macromolecule metabolic process	5.20E-07
P	GO:0006066	alcohol metabolic process	1.40E-06
P	GO:0019318	hexose metabolic process	1.50E-06
P	GO:0044238	primary metabolic process	3.90E-06
P	GO:0031145	anaphase-promoting complex-dependent proteasomal ubiquitin-dependent protein catabolic process	4.20E-06
P	GO:0007200	activation of phospholipase C activity by G-protein coupled receptor protein signaling pathway coupled to IP3 second messenger	4.90E-06
P	GO:0007202	activation of phospholipase C activity	7.60E-06
P	GO:0010863	positive regulation of phospholipase C activity	9.40E-06
P	GO:0007067	mitosis	9.00E-06
P	GO:0022613	ribonucleoprotein complex biogenesis	9.60E-06
P	GO:0000280	nuclear division	1.10E-05
P	GO:0010518	positive regulation of phospholipase activity	1.40E-05
P	GO:0010517	regulation of phospholipase activity	1.40E-05
P	GO:0048285	organelle fission	1.70E-05
P	GO:0030071	regulation of mitotic metaphase/anaphase transition	2.10E-05
P	GO:0043170	macromolecule metabolic process	2.20E-05
P	GO:0007091	mitotic metaphase/anaphase transition	2.10E-05
P	GO:0060193	positive regulation of lipase activity	2.50E-05
P	GO:0007178	transmembrane receptor protein serine/threonine kinase signaling pathway	2.90E-05
P	GO:0060191	regulation of lipase activity	3.00E-05
P	GO:0000087	M phase of mitotic cell cycle	3.80E-05
P	GO:0008152	metabolic process	3.90E-05
P	GO:0042254	ribosome biogenesis	3.90E-05
P	GO:0016070	RNA metabolic process	4.20E-05
P	GO:0090092	regulation of transmembrane receptor protein serine/threonine kinase signaling pathway	4.30E-05
P	GO:0048015	phosphoinositide-mediated signaling	4.30E-05
P	GO:0006607	NLS-bearing substrate import into nucleus	7.90E-05
P	GO:0045333	cellular respiration	9.30E-05
P	GO:0007176	regulation of epidermal growth factor receptor activity	0.00011
P	GO:0007175	negative regulation of epidermal growth factor receptor activity	0.00011
P	GO:0018345	protein palmitoylation	0.00011
P	GO:0010469	regulation of receptor activity	0.00011
P	GO:0009987	cellular process	0.00011
P	GO:0006457	protein folding	0.00012
P	GO:0006396	RNA processing	0.00012
P	GO:0006364	rRNA processing	0.00014
P	GO:0055014	atrial cardiac muscle cell development	0.00014
F	GO:0003873	6-phosphofructo-2-kinase activity	3.80E-13
F	GO:0019200	carbohydrate kinase activity	8.80E-13
F	GO:0008443	phosphofructokinase activity	3.00E-11
F	GO:0003746	translation elongation factor activity	1.40E-05
F	GO:0031386	protein tag	2.70E-05
F	GO:0017124	SH3 domain binding	4.80E-05
F	GO:0051219	phosphoprotein binding	0.00011
F	GO:0001784	phosphotyrosine binding	0.00011
F	GO:0045309	protein phosphorylated amino acid binding	0.00011
C	GO:0030684	preribosome	2.20E-10
C	GO:0030529	ribonucleoprotein complex	1.70E-09
C	GO:0043231	intracellular membrane-bounded organelle	1.60E-08

C	GO:0043227	membrane-bounded organelle	1.80E-08
C	GO:0044424	intracellular part	2.90E-08
C	GO:0005634	nucleus	3.70E-08
C	GO:0005622	intracellular	3.40E-07
C	GO:0043229	intracellular organelle	7.80E-07
C	GO:0043226	organelle	9.10E-07
C	GO:0034457	Mpp10 complex	1.10E-06
C	GO:0005737	cytoplasm	2.60E-06
C	GO:0030686	90S preribosome	2.40E-06
C	GO:0045121	membrane raft	6.70E-06
C	GO:0030672	synaptic vesicle membrane	1.40E-05
C	GO:0032040	small-subunit processome	2.00E-05
C	GO:0030687	preribosome, large subunit precursor	2.10E-05
C	GO:0044428	nuclear part	5.80E-05
C	GO:0032991	macromolecular complex	6.10E-05
C	GO:0044455	mitochondrial membrane part	6.70E-05
C	GO:0030665	clathrin coated vesicle membrane	9.80E-05
C	GO:0016469	proton-transporting two-sector ATPase complex	0.0001
C	GO:0005689	U12-type spliceosomal complex	0.00013
C	GO:0044452	nucleolar part	0.00018
C	GO:0005680	anaphase-promoting complex	0.00018
C	GO:0005687	U4 snRNP	0.00022
C	GO:0031410	cytoplasmic vesicle	0.00029
C	GO:0031307	integral to mitochondrial outer membrane	0.0003
C	GO:0044444	cytoplasmic part	0.00028
C	GO:0000152	nuclear ubiquitin ligase complex	0.0003
C	GO:0005686	U2 snRNP	0.00041
C	GO:0071564	npBAF complex	0.00041
C	GO:0031982	vesicle	0.00039
C	GO:0016507	fatty acid beta-oxidation multienzyme complex	0.00041
C	GO:0000243	commitment complex	0.00042
C	GO:0030532	small nuclear ribonucleoprotein complex	0.00055
C	GO:0071011	precatalytic spliceosome	0.00058
C	GO:0000276	mitochondrial proton-transporting ATP synthase complex, coupling factor F(o)	0.00072
C	GO:0031306	intrinsic to mitochondrial outer membrane	0.0007
C	GO:0044446	intracellular organelle part	0.0009
C	GO:0044422	organelle part	0.00087
C	GO:0030688	preribosome, small subunit precursor	0.00089

* Species list are Supplementary Table 6.

Supplementary Table 12. CAFÉ1 gene family analysis results.

Species	No. of coding genes	No. of Expanded families	No. of Contracted families.	No. of Genes lost
Icefish	30,773	1793 (280)	6831 (6)	6869
Dragonfish	32,713	1174 (162)	4057 (1)	4080
Bullhead	32,661	1178 (8)	5534 (2)	5640
notothen				
Stickleback	27,576	465 (59)	10704 (2)	10732
Platyfish	20,779	184 (3)	1928 (32)	2002
Amazon molly	24354	1195 (165)	780 (1)	781
Medaka	20459	452 (78)	4953 (3)	5018
Tilapia	27088	800 (133)	4047 (0)	4057
Takifugu	18,523	390 (47)	1924 (3)	1935
Tetraodon	23,118	530 (24)	1895 (7)	1927
Cod	22,100	391 (21)	14427 (3)	14523
Zebrafish	25,403	1183 (78)	2205 (0)	2241
Cavefish	23698	775 (8)	2542 (16)	2753



Supplementary Table 13. CAFÉ results for significantly expanded genes in icefish.

Category	GO acc	Term	<i>p</i> -value
F	GO:0005488	binding	2.94E-03
F	GO:0005200	structural constituent of cytoskeleton	5.48E-03
F	GO:0030246	carbohydrate binding	5.98E-03
F	GO:0005506	iron ion binding	1.07E-02
F	GO:0003924	GTPase activity	2.12E-02
F	GO:0008199	ferric iron binding	1.79E-02
F	GO:0016702	oxidoreductase activity, acting on single donors with incorporation of molecular oxygen, incorporation of two atoms of oxygen	3.04E-02
F	GO:0016701	oxidoreductase activity, acting on single donors with incorporation of molecular oxygen	3.04E-02
F	GO:0004984	olfactory receptor activity	4.29E-02
P	GO:0051258	protein polymerization	1.42E-02
P	GO:0043623	cellular protein complex assembly	1.86E-02
P	GO:0007017	microtubule-based process	2.30E-02
P	GO:0034622	cellular macromolecular complex assembly	3.31E-02
P	GO:0006461	protein complex assembly	4.28E-02
P	GO:0070271	protein complex biogenesis	4.34E-02
P	GO:0006826	iron ion transport	2.04E-02
P	GO:0000041	transition metal ion transport	2.54E-02
P	GO:0006879	cellular iron ion homeostasis	2.79E-02
P	GO:0046916	cellular transition metal ion homeostasis	2.79E-02
P	GO:0055072	iron ion homeostasis	3.29E-02
P	GO:0055076	transition metal ion homeostasis	3.54E-02
P	GO:0050911	detection of chemical stimulus involved in sensory perception of smell	4.29E-02
P	GO:0050907	detection of chemical stimulus involved in sensory perception	4.53E-02
P	GO:0050906	detection of stimulus involved in sensory perception	4.53E-02
P	GO:0006875	cellular metal ion homeostasis	4.78E-02
P	GO:0009593	detection of chemical stimulus	4.78E-02
C	GO:0005874	microtubule	1.00E-02
C	GO:0015630	microtubule cytoskeleton	2.79E-02

Supplementary Table 14. Positively selected genes in the icefish genome. (Likelihood ratio test: $\alpha=0.05$).

Gene ID	Descriptions
Cac000489	synaptosomal-associated protein 47
Cac000662	septin-10-like isoform X2
Cac001617	ester hydrolase C11orf54 homolog
Cac001771	T-complex protein 1 subunit alpha
Cac002279	delta-type opioid receptor-like
Cac003235	protein FAM3C
Cac003886	tripeptidyl-peptidase 1
Cac004140	poly(ADP-ribose) glycohydrolase ARH3
Cac004885	nuclear pore complex protein Nup205
Cac004994	thioredoxin domain-containing protein 5
Cac005124	integrin alpha-6
Cac005514	high-affinity choline transporter 1-like
Cac006594	phosphatidylinositol 4-kinase type 2-alpha
Cac011177	neutrophil cytosol factor 1
Cac011177	neutrophil cytosol factor 1
Cac012351	lengsin
Cac013420	cerebral cavernous malformations 2 protein isoform X2
Cac013753	sonic hedgehog protein
Cac016183	translation initiation factor eIF-2B subunit beta
Cac016183	translation initiation factor eIF-2B subunit beta
Cac019931	pyruvate dehydrogenase phosphatase catalytic subunit 1
Cac019931	pyruvate dehydrogenase phosphatase catalytic subunit 1
Cac020370	zinc finger MYND domain-containing protein 10
Cac020655	beta-1:4 N-acetylgalactosaminyl transferase 1
Cac020945	mitotic checkpoint protein BUB3
Cac020991	F-box only protein 11 isoform X4
Cac021124	NADPH--cytochrome P450 reductase isoform X1
Cac021496	PX domain-containing protein 1
Cac021612	translocon-associated protein subunit alpha isoform X2
Cac022143	reticulon-4-interacting protein 1: mitochondrial isoform X1
Cac023068	UNC93-like protein MFSD11
Cac023070	mpv17-like protein
Cac025005	protein CYR61
Cac025225	rod cGMP-specific 3',5'-cyclic phosphodiesterase subunit beta
Cac025247	RNA 3'-terminal phosphate cyclase-like protein
Cac026046	AP-1 complex subunit gamma-1 isoform X2
Cac026882	sorting nexin-12
Cac027018	merlin-like
Cac027509	peripheral myelin protein 22-like
Cac027988	transcription cofactor vestigial-like protein 2 isoform X1
Cac028923	glomulin-like
Cac029538	FAST kinase domain-containing protein 1
Cac029831	general transcription factor 3C polypeptide 4

Supplementary Table 15. Gene ontology for positively selected genes in the icefish genome.

Category	GO accessory	Term	p-value
MF	GO:0035091	phosphatidylinositol binding	7.25E-05
MF	GO:0005543	phospholipid binding	2.48E-04
MF	GO:0008289	lipid binding	2.05E-03
MF	GO:0016651	oxidoreductase activity, acting on NAD(P)H	8.93E-04
MF	GO:0003958	NADPH-hemoprotein reductase activity	3.26E-03
MF	GO:0016175	superoxide-generating NADPH oxidase activity	3.26E-03
MF	GO:0050664	oxidoreductase activity, acting on NAD(P)H, oxygen as acceptor	3.26E-03
MF	GO:0016653	oxidoreductase activity, acting on NAD(P)H, heme protein as acceptor	6.50E-03
MF	GO:0010181	FMN binding	1.30E-02
MF	GO:0004722	protein serine/threonine phosphatase activity	3.21E-02
MF	GO:0004114	3',5'-cyclic-nucleotide phosphodiesterase activity	3.53E-02
MF	GO:0004112	cyclic-nucleotide phosphodiesterase activity	3.53E-02
BP	GO:0030908	protein splicing	6.50E-03
BP	GO:0016539	intein-mediated protein splicing	6.50E-03
BP	GO:0016485	protein processing	3.53E-02
BP	GO:0051604	protein maturation	3.53E-02
CC	GO:0000151	ubiquitin ligase complex	3.53E-02

Abbreviations: MF, Molecular function; BP, biological process; CC, Cellular component

극지연구소

Supplementary Table 16. Comparison of zona pellucida (zp) gene family members in icefish and other Antarctic fish.

	ZPAX1	ZPAX2	ZPC1	ZPC2	ZPC3	ZPC4	ZPC5	ZPB	ZPD	CGL	CGH	Total
<i>C. aceratus</i>	29	7	22	13	7	2	33	2	1	8	7	131
* <i>N. coriiceps</i>	1	2	2	1	1	2	1	3	1	4	0	18
* <i>P. charcoti</i>	4	1	2	2	2	1	3	6	1	7	1	30
** <i>D. mawsoni</i>	1	1	1	1	1	1	8	1	1	1	1	18
** <i>C. hamatus</i>	1	4	4	2	0	2	4	1	1	0	0	19
** <i>G. acuticeps</i>	3	1	3	2	2	2	4	2	1	0	0	20
** <i>T. bernacchii</i>	1	1	2	1	1	1	2	1	1	0	0	11
** <i>N. angustata</i>	0	0	1	0	0	0	1	0	0	0	0	2
** <i>D. eleginoides</i>	1	0	1	1	0	0	5	0	0	0	0	8
** <i>L. nudifrons</i>	1	1	1	1	1	1	1	1	1	0	0	9
** <i>E. maclovinus</i>	1	1	2	1	1	1	3	1	0	0	0	11

* The number of zp genes was annotated from our previous research^{2,3}.

** The number of zp genes of 8 species of Antarctic fish was derived from the Supplementary Table 1 in Cao et al., 2016⁴.



Supplementary Table 17. Genomic distribution of icefish zona pellucida (zp) family genes. The total number of annotated zp genes on each contig is indicated. Of the 131 icefish zp genes, 109 are distributed on 20 different scaffolds that have more than two genes, and 22 scaffolds have a single zp gene.

Contig	ZPAX1	ZPAX2	ZPC1	ZPC2	ZPC3	ZPC4	ZPC5	ZP3	ZP4	ZPB	ZPD	Total
Ice_000281					2		18					20
Ice_010079			7	3								10
Ice_000114					5		5					10
Ice_030120	7	3										10
Ice_040384	3	1					4					8
Ice_051004			4	4								8
Ice_060311	1	1	4									6
Ice_070981			5									5
Ice_080050								1	3			4
Ice_090470			2	2								4
Ice_100683							4					4
Ice_110055								3				3
Ice_120371	3										1	4
Ice_130009						1				1		2
Ice_140019						1				1		2
Ice_150293	2											2
Ice_160991	2											2
Ice_171379	1	1										2
Ice_182290	2											2
Ice_192443	2											2
Ice_200015									1			1
Ice_210036									1			1
Ice_220080								1				1
Ice_230090	1											1
Ice_240126									1			1
Ice_250182				1								1
Ice_260204							1					1
Ice_270271		1										1
Ice_280276				1								1
Ice_290336								1				1
Ice_300443	1											1
Ice_310473								1				1
Ice_320486								1				1
Ice_330519	1											1
Ice_340709							1					1
Ice_350845									1			1
Ice_361246	1											1
Ice_371821				1								1
Ice_382268	1											1
Ice_392717				1								1
Ice_402947	1											1
Total	29	7	22	13	7	2	33	8	7	2	1	131

Supplementary Table 18. Annotation of antioxidant defense system related genes in five teleosts. Genes were selected from the gene set of antioxidant activity (M15021) of Gene Set Enrichment Analysis5 (GSEA, <http://software.broadinstitute.org/gsea/index.jsp>).

	Icefish	Dragonfish	Antarctic bullhead	Takifugu	Medaka	Platyfish	Zebrafish
Superoxide dismutase	5	3	3	3	3	3	4
Catalase	1	1	1	1	1	1	1
NAD(P)H:quinone acceptor oxidoreductase	33	2	2	2	3	3	1
Glutathione peroxidase	4	7	4	9	6	6	9
Selenoprotein S	1	1	1	1	0	1	1
Eosinophil peroxidase	7	4	4	3	6	4	2
Microsomal glutathione S-transferase 3	2	2	2	2	2	2	2
Cytoglobin	2	2	2	2	2	2	2
8-oxoguanine DNA glycosylase	2	1	1	1	1	1	1
Thioredoxin reductase	3	3	2	2	2	3	2
Peroxiredoxin	6	6	6	6	6	6	6
Apolipoprotein A-IV	4	2	3	6	2	1	4
Total	70	34	31	38	34	33	35

Supplementary Table 19. sod genes identified in teleost genomes. The Genbank accession number is indicated for each sod gene.

	SOD1	SOD2	SOD3
Icefish	CacV3p_001772	CacV3p_016306	CacV3p_021817-1 CacV3p_021817-2 CacV3p_021817-3
Antarctic Bullhead	XP_010771234.1	XP_010770367.1	XP_010775098.1
Dragonfish	DFtranscrip-31740	DFtranscrip-29679	DFtranscrip-21872
Stickleback	ENSGACG00000020581	ENSGACG00000009000	ENSGACG00000017840
Takifugu	XP_003971372	XP_011610563	XP_011610687
Platyfish	XP_005807113	XP_005802672	XP_005807205
Amazon molly	XP_007555641	XP_007546844	XP_007551838
Medaka	XP_004076261	XP_004083519	XP_004084336 *
Zebrafish	NP_571369.1	NP_956270	NP_001092706 (chr7)* XP_001332758 (chr1)

Zebrafish has two sod3 genes (sod3a and sod3b); sod3a is located on chromosome 7 with similar synteny to that of other teleosts, whereas sod3b is located on chromosome 1.



Supplementary Table 20. Circadian rhythm-related genes identified in teleost genomes. The location of each gene in each chromosome or scaffold is noted for each gene. Chromosomal location was based on the updated Supplementary Material 13 of Toloza-Villalobos et al. (2015)6. Updated information is marked with an asterisk (*). A dash means that the gene was not found on a defined chromosome or scaffold.

	Icefish	Stickleback	Takifugu	Medaka	Platyfish	Zebrafish
<i>per1a</i>	-	-	-	-	-	Chr10:23010981-23028617
<i>per1b</i>	Cacp_005339 Ice_000086:2255858..2264187	groupIII:1517985-1529008	Scaffold_63:506313-512694	Chr18:15189752-15199550	JH556673.1:326107-343686*	Chr7:52277645-52301471
<i>per2a</i>	-	groupI:1855721-1863275*	Scaffold_344:165475-173866	Chr13:31788175-31800229	JH556784.1:457717-480890	-
<i>per2b</i>	Cacp_024627 Ice_000019:2927559..2939828	groupVII:3993749-4000926*	Scaffold_123:243442-251642	Chr17:25658720-25680938	JH556667.1:3479242-3504286	Chr2:48423772-48489514
<i>per3</i>	-	-	Chr19:11805147-11815446*	Chr5:29288111-29298254	JH556773.1:1356945-1376982	Chr11:41435220-41490992
<i>cry1aa</i>	Cacp_026503 Ice_000025:3445991..3456768	groupIV:32004462-32008670	Scaffold_4395:1-440	Chr23:19786254 - 19805276	JH556703.1:1996368-2009995	Chr4:12011692-12027887*
<i>cry1ab</i>	-	groupXIX:17896031-17900731:1*	Scaffold_21:12588914-1264794	Chr6:30902546-30911031	Chr2:17084212-17094297*	Chr18:15104200-15132014*
<i>cry1ba</i>	Cacp_006784 Ice_000125:1848029..1834983	groupXII:13140358-13147524*	Scaffold_278:81132-99203*	Chr7:28299032 - 28474995*	Chr1:29806902 - 29840394*	Chr8:21222300-21255421*
<i>cry1bb</i>	-	-	-	-	-	Chr22:783514-811211*
<i>cry2</i>	-	groupXIX:1002682-1016637	Scaffold_591:4558-10348*	Chr6:22492682-22503706	Chr2:4478878-4495387*	Chr25:13845092-13873484
<i>clocka</i>	Cacp_021314 Ice_000042:1091634..1109722	-	Scaffold_13:1587298-1594020	-	JH557215.1:159758-185684	Chr20:22202472-22246469
<i>clockb</i>	Cacp_021654 Ice_000015:997555..1012718	groupIX:489361-499374	Scaffold_563:28933-37732	Chr1:14996520-15007939	JH556907.1:802434-820177	Chr1:18174899-18208348

<i>npas2a</i>	Cacp_011192 Ice_000386:348502..366485	groupVII:17209951-1 7217908	Scaffold_6:1321121-13 31856	Chr14:17135493-171 47943	-	Chr5:24572014-246 14052
<i>arntl1a</i>	Cacp_017407 Ice_000301:749264..791055	groupXIX:1196030-12 06038	Scaffold_30:314818-32 2668	Scaffold212:822889- 852119	JH556669.1:1164407- 1191899	Chr25:18381074-18 406100
<i>arntl1b</i>	-	-	-	-	-	Chr7:67945182-680 02908
<i>arntl2a</i>	Cacp_009803 Ice_000337:138232..147037	-	Scaffold_105:333609-3 41010	Chr6:16473669-1648 4384	JH556735.1:1789641- 1801148	Chr18:15249987-15 289314
<i>arntl2b</i>	Cacp_026454 Ice_000025:673361..679455	groupIV:22530990-22 537106	Scaffold_657:8465-120 73	Chr23:2931241-2954 154	JH556873.1:690137-7 07607	-
<i>csnk1ea</i>	Cacp_012338 Ice_001973:12374..33989	groupV:5628628-5633 878	Scaffold_3:2837555-28 40657	Chr19:803568-81128 0	JH559539.1:29-8132	Chr12:20401738-20 414656
<i>csnk1eb</i>	Cacp_029385 Ice_000244:45..10972	groupXI:4260952-426 8829	Scaffold_253:334382-3 42511	Chr8:19677134-1968 4994	JH556752.1:1560221- 1567996	-
<i>csnk1da</i>	Cacp_014005 Ice_000322F:247084..274355	Scaffold54:477509-48 6247	Scaffold_41:895893-90 2348	Chr8:25232642-2523 9805	JH557172.1:17306-26 673	-
<i>csnk1db</i>	Cacp_006446 Ice_000123F:682442..704673	groupXI:1303431-131 2998	-	-	-	Chr12:34815412-34 831695
<i>csnk1dc</i>	-	-	-	-	-	Chr3:36172321-362 09462
<i>timeless</i>	Cacp_000716 Ice_000119F:944982..953940	Scaffold27:1143834-1 158386	Scaffold_47:1354118-1 362657	Chr5:27745117-2775 8015	JH556773.1: 50419-64006	Chr11:3195071-322 9802

Supplementary Table 21. The number of olfactory receptor genes in 6 teleost species.

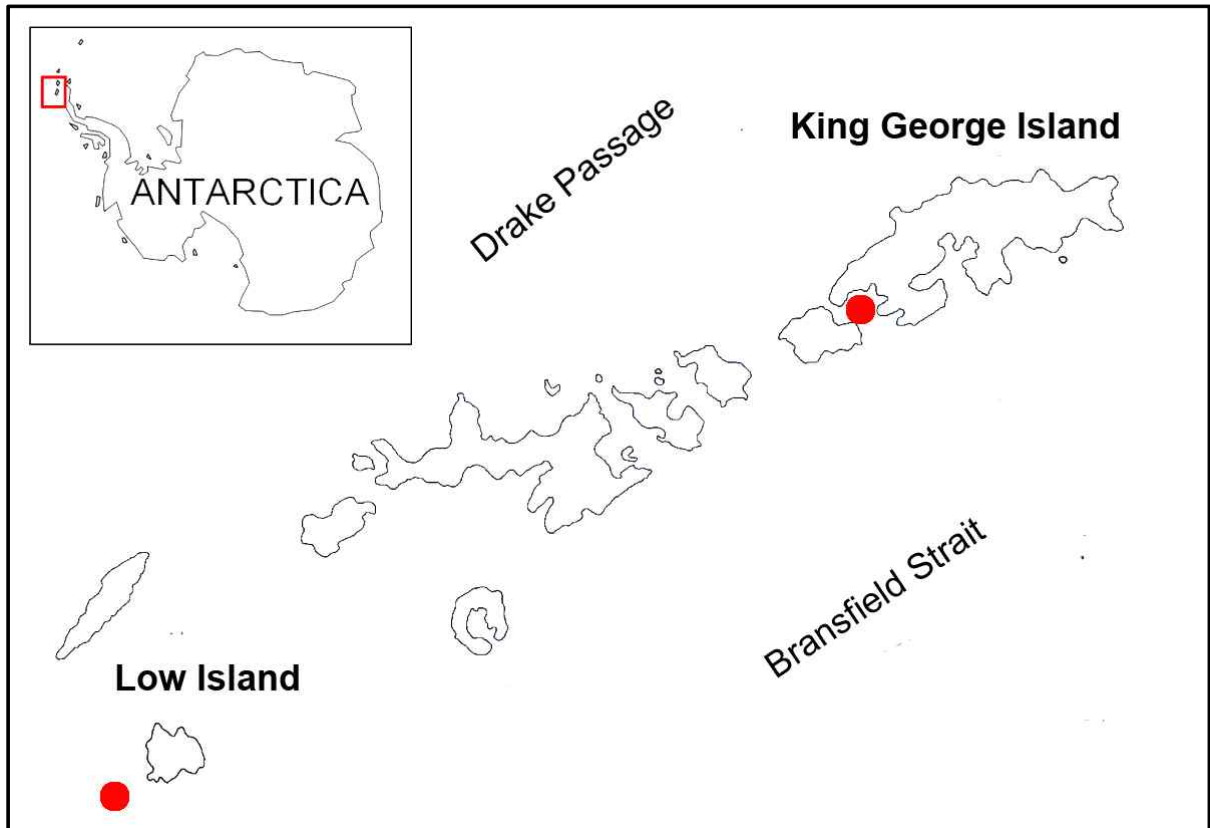
	Icefish	Medaka	Stickleback	Takifugu	Tetraodon	Zebrafish
alpha	0	0	0	0	0	0
beta	1	3	1	1	0	7
gamma	2	1	3	0	0	1
delta	161	44	112	55	14	69
epsilon	2	4	4	3	2	13
zeta	10	12	23	10	2	41
eta	35	30	12	31	12	45
theta	0	1	1	1	1	1
theta-2	0	0	0	0	0	1
kappa	1	1	3	1	1	2
lambda	0	1	1	0	0	1
Total	212	97	160	102	32	181



Supplementary Figure 1. Photograph of a female specimen of the Antarctic blackfin icefish, *Chaenocephalus aceratus*.

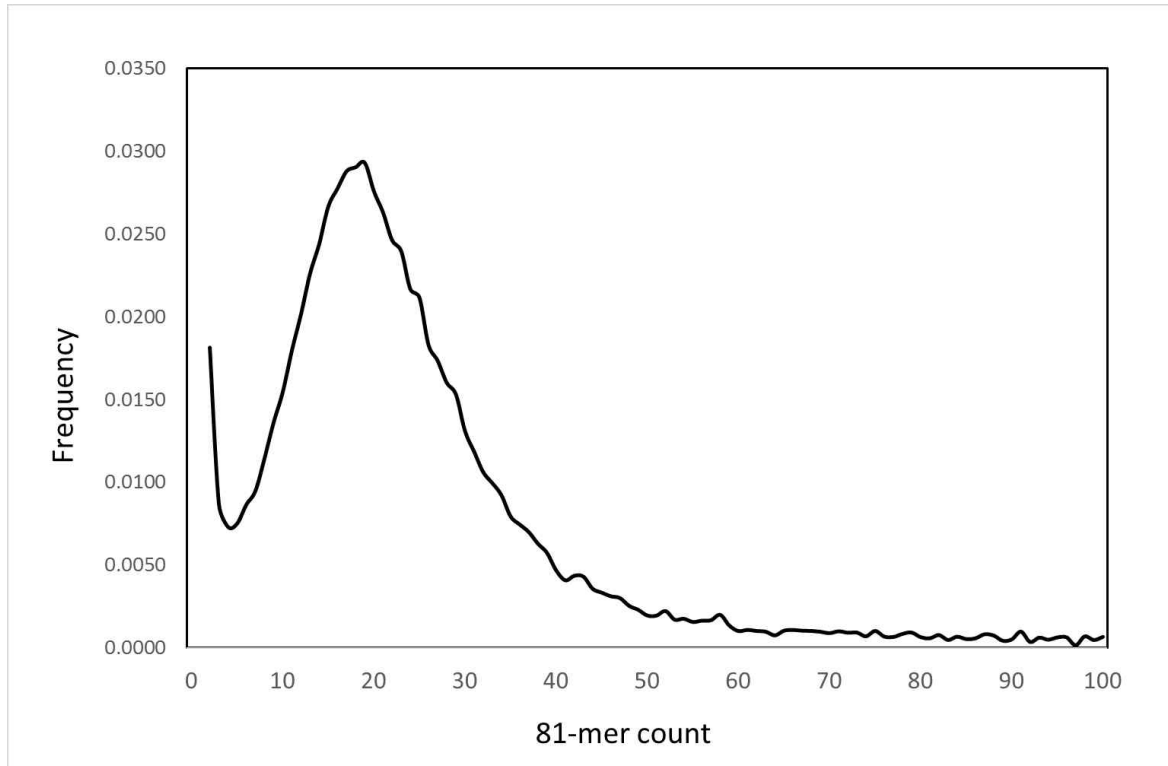


Supplementary Figure 2. Sampling sites for Antarctic blackfin icefish. Red dots represent sampling points.



극지연구소

Supplementary Figure 3. Estimation of genome size by K-mer analysis. Genome size was calculated as total error-corrected Pacbio sequence. The estimated genome size of *C. aceratus* is 1.1 Gb (81-mers).

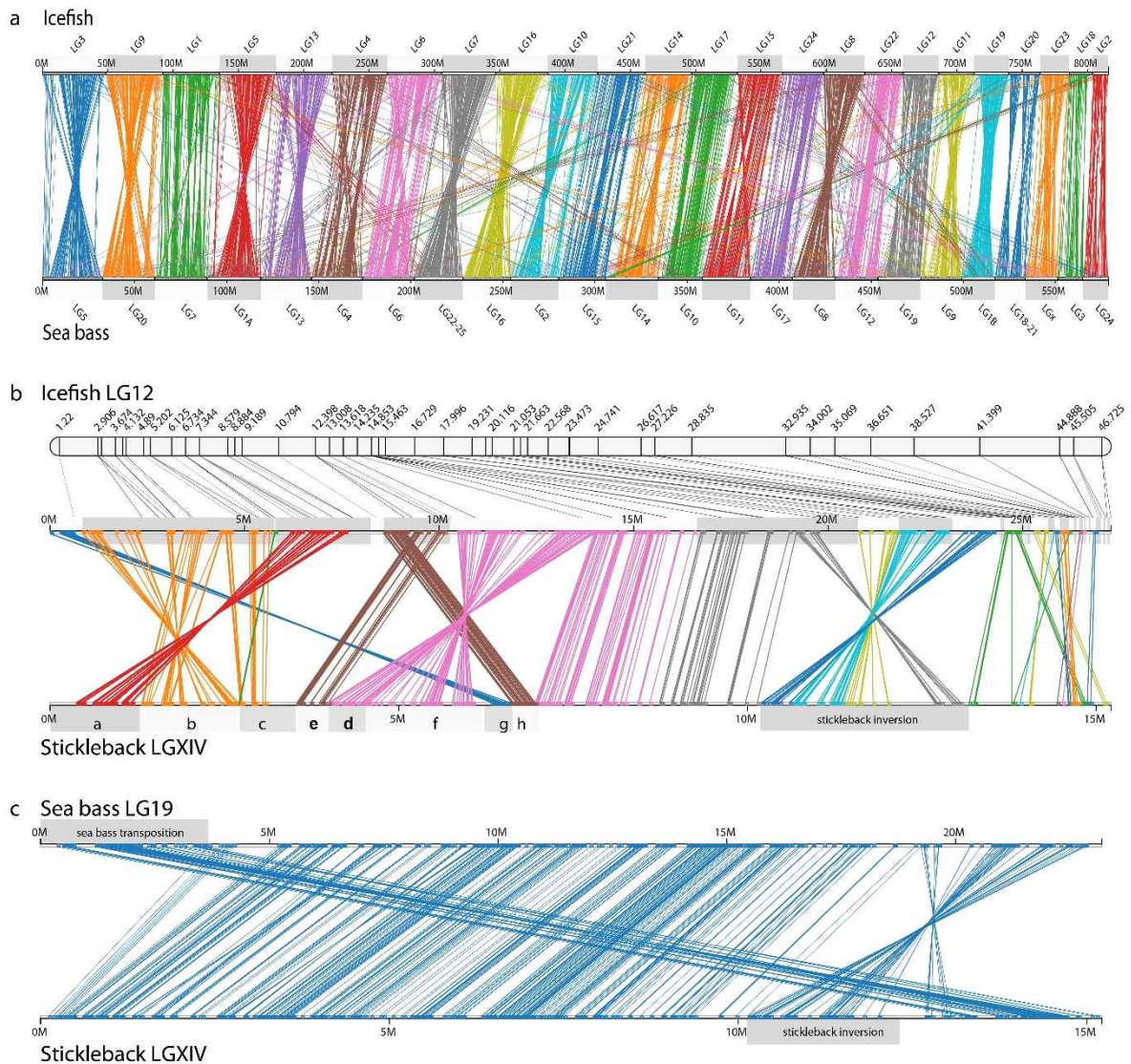


Supplementary Figure 4. Genetic linkage map for the blackfin icefish.

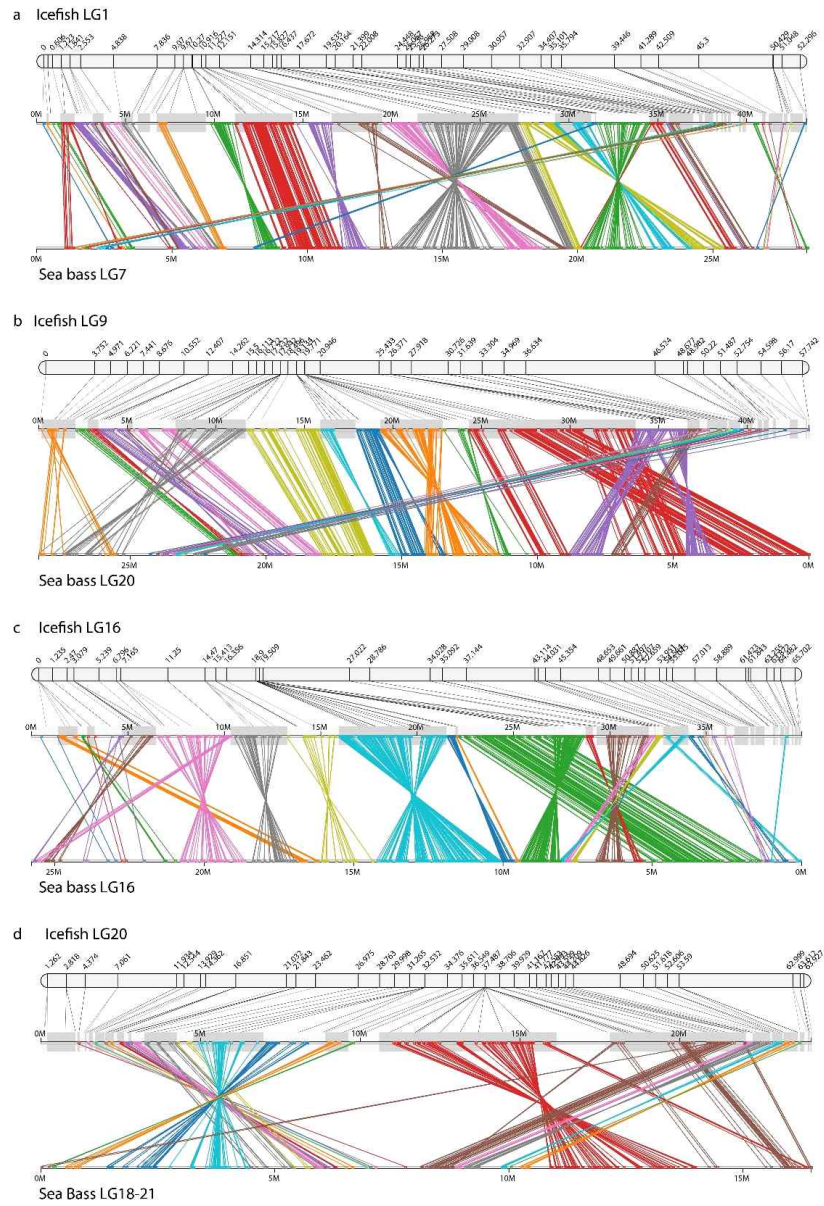


극지연구소

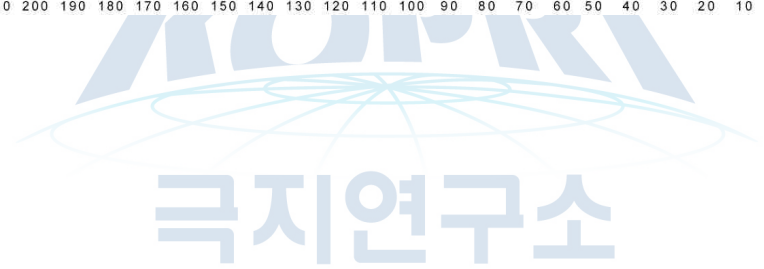
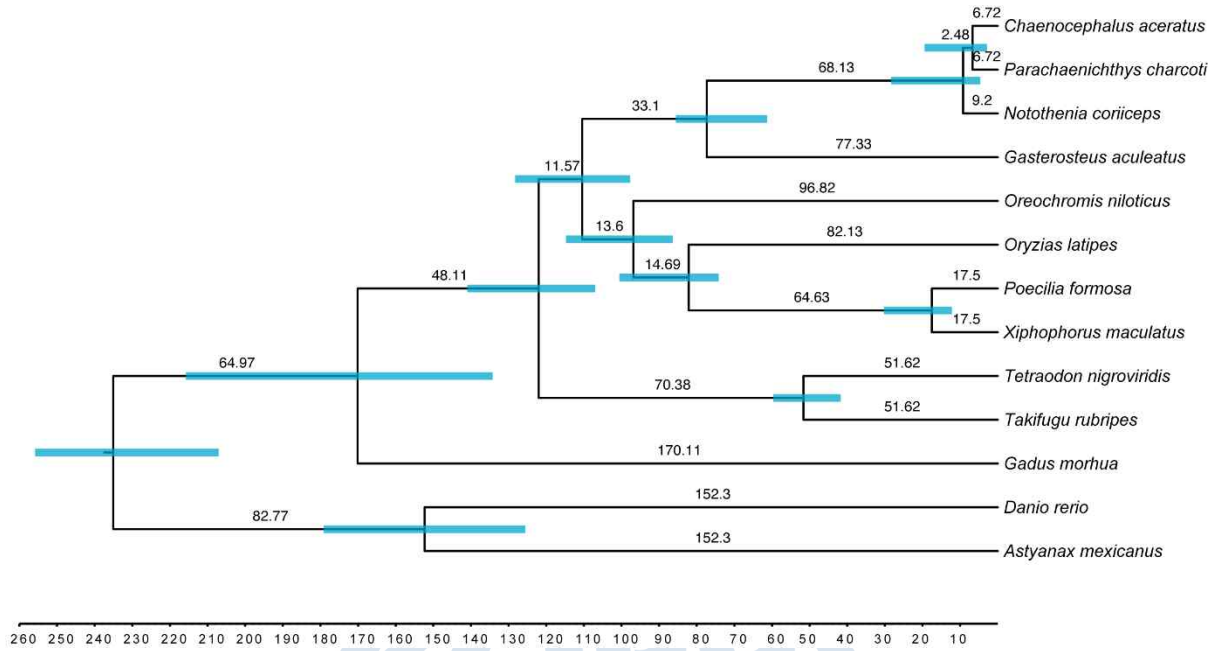
Supplementary Figure 5. Synteny of *C. aceratus* chromosomes with those of other teleosts. a. Analysis of gene content suggests a one-to-one correspondence between icefish and sea bass chromosomes. Icefish-specific color-coded lines connect predicted orthologous genes and show that no translocations occurred in the icefish lineage after it diverged from the sea bass lineage. b. Conservation of synteny between icefish LG12 and stickleback LGXIV. Lines connecting individual orthologous genes are color coded based on the icefish genomic contig. c. Conserved synteny between sea bass LG19 and stickleback LgXIV, which reveals that inversions and transpositions were frequent in the icefish lineage after it diverged from the sea bass lineage.



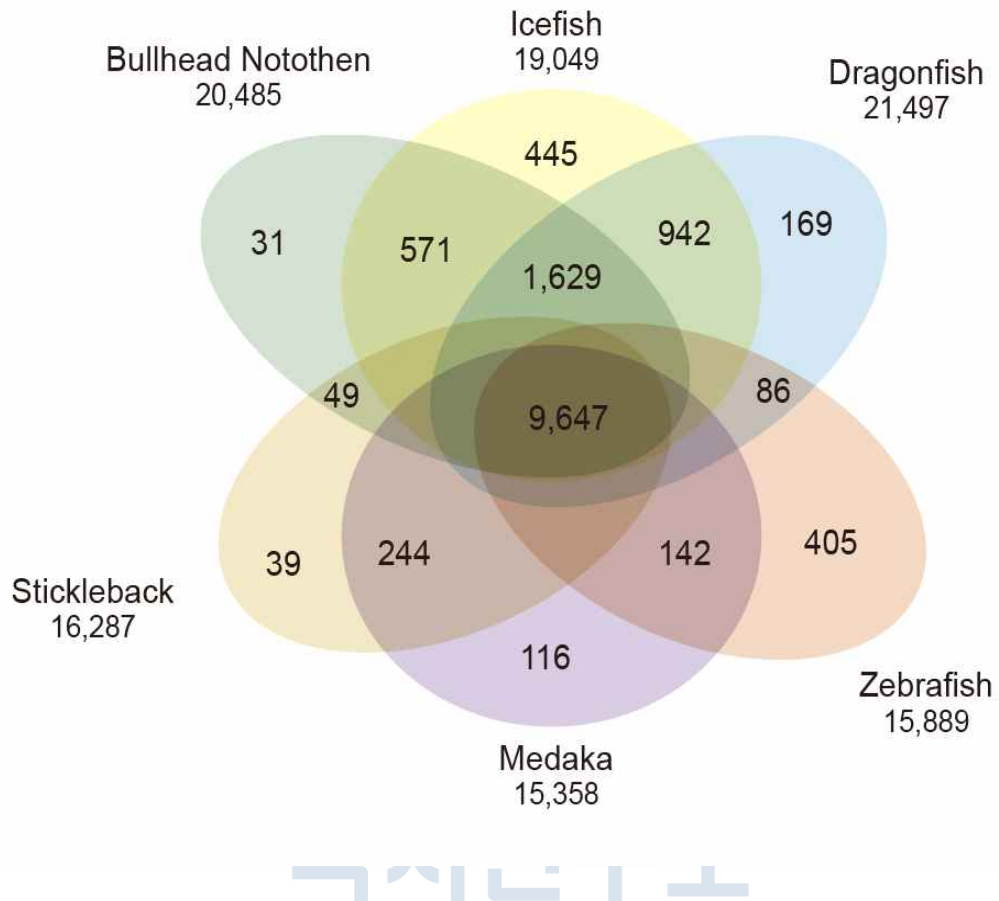
Supplementary Figure 6. Conservation of gene order between orthologous icefish and sea bass chromosomes. a, icefish LG1. b, icefish LG9. c, icefish LG16. d, icefish LG20. Lines connecting orthologous genes are color coded relative to the icefish genomic contig. Comparisons show many inversions and a few transpositions that distinguish these two species, whose lineages diverged more than 100 million years ago.



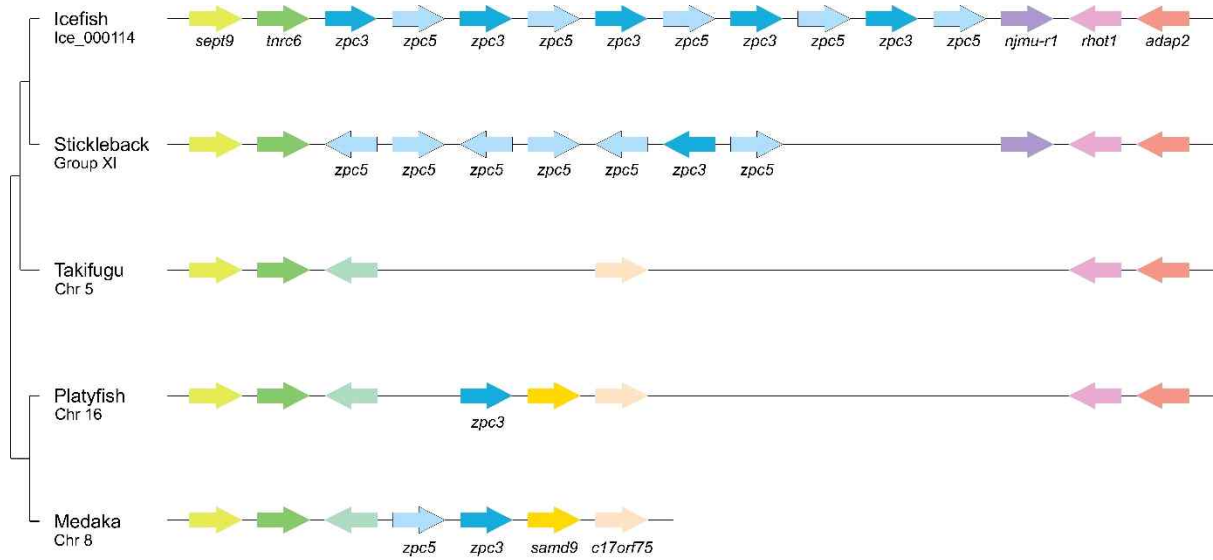
Supplementary Figure 7. Divergence time estimation of teleost fish. Blue bars depict 95% highest posterior density (HPD) time estimates (time scale: Mya).



Supplementary Figure 8. Venn diagram of orthologous gene clusters for five teleost fishes.



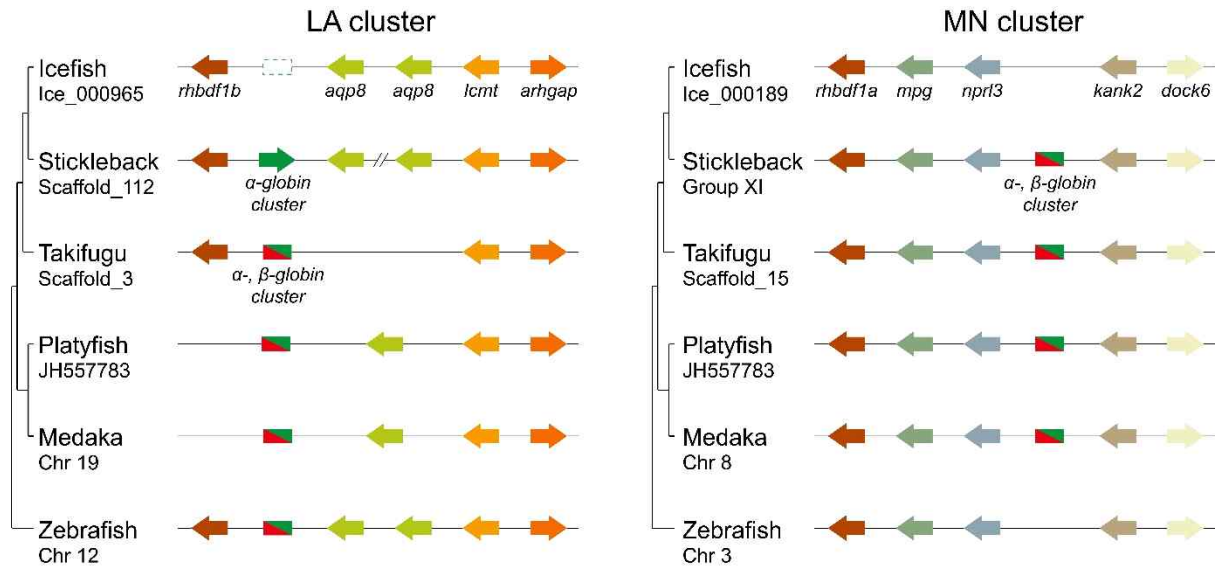
Supplementary Figure 9. Conserved synteny for *zpc3* and *zpc5* genes identified in the icefish genome. The icefish *zpc3* and *zpc5* are tandemly duplicated in a contig Ice_000114.



Supplementary Figure 10. Transcriptional profile of zona pellucida (zp) protein family genes in blackfin icefish organs. The transcriptional profile of zp genes was measured in twelve organs of icefish (i.e. brain, eye, gill, heart, intestine, kidney, liver, muscle, ovary, skin, spleen, and stomach). The level of expression for each mRNA is represented by color gradations in the heat map.

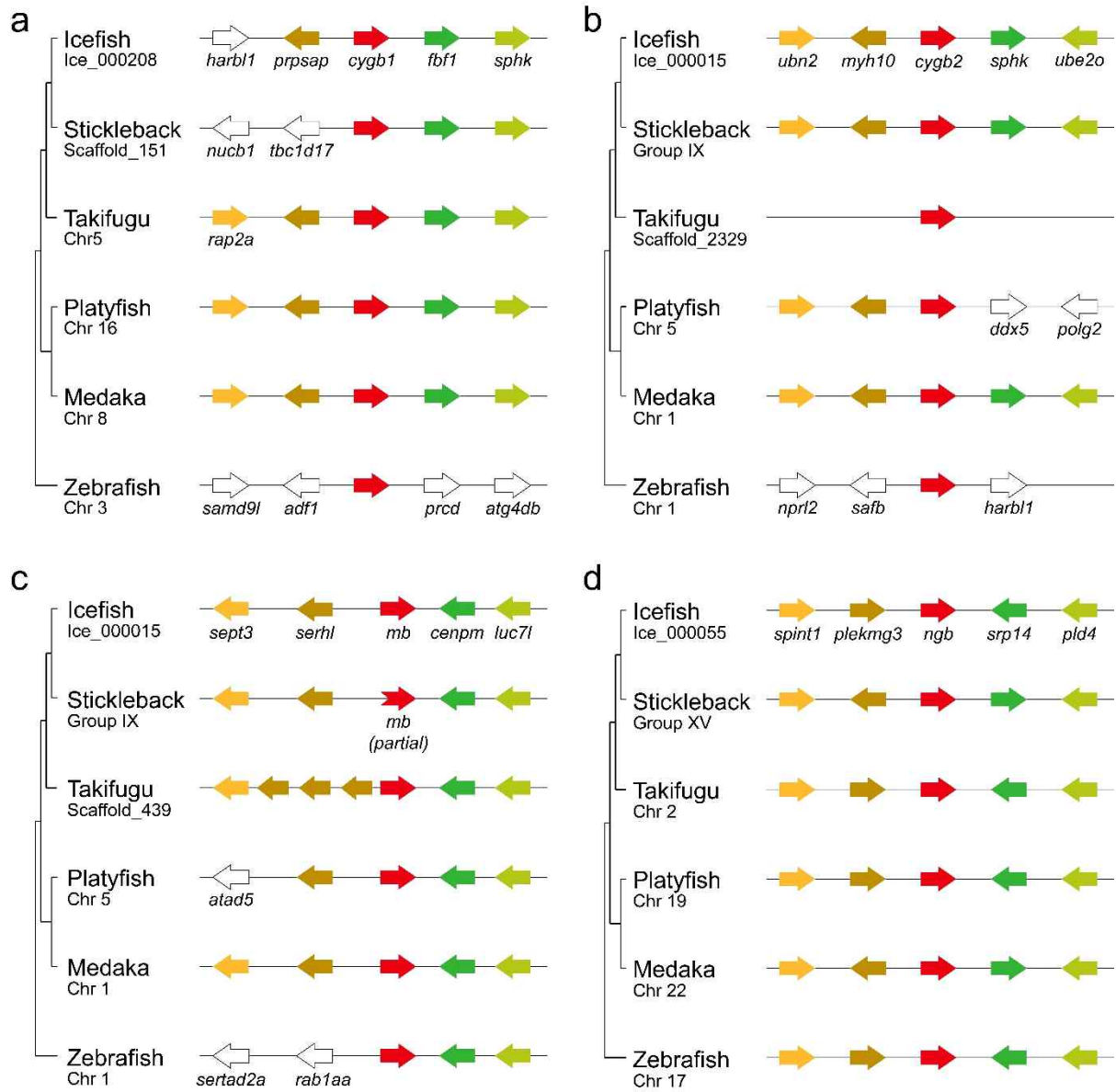


Supplementary Figure 11. Syntenic comparisons of icefish globin gene clusters. Teleosts have two globin gene clusters. The LA cluster, which encodes both hemoglobin α - and β -subunits, is adjacent to the *lcmt1* and aquaporin genes on one side and *rhbdf1b* on the other side. The MN cluster, which also encodes both hemoglobin alpha and beta subunits, has *rhbdf1a*, *mpg*, and *nrp13* on one side and *kank2* on the other side. Loss of complete copies of both α - and β -globin genes is observed for both the LA and MN clusters in the icefish genome. Each arrow indicates a gene in the 5'→3' direction. α - and β -globin clusters are shown by green and red boxes. The partial alpha globin gene containing only exon 3 is identified in the icefish LA cluster by a dashed box. This figure was updated from the figure 1 of Opazo et al.1.

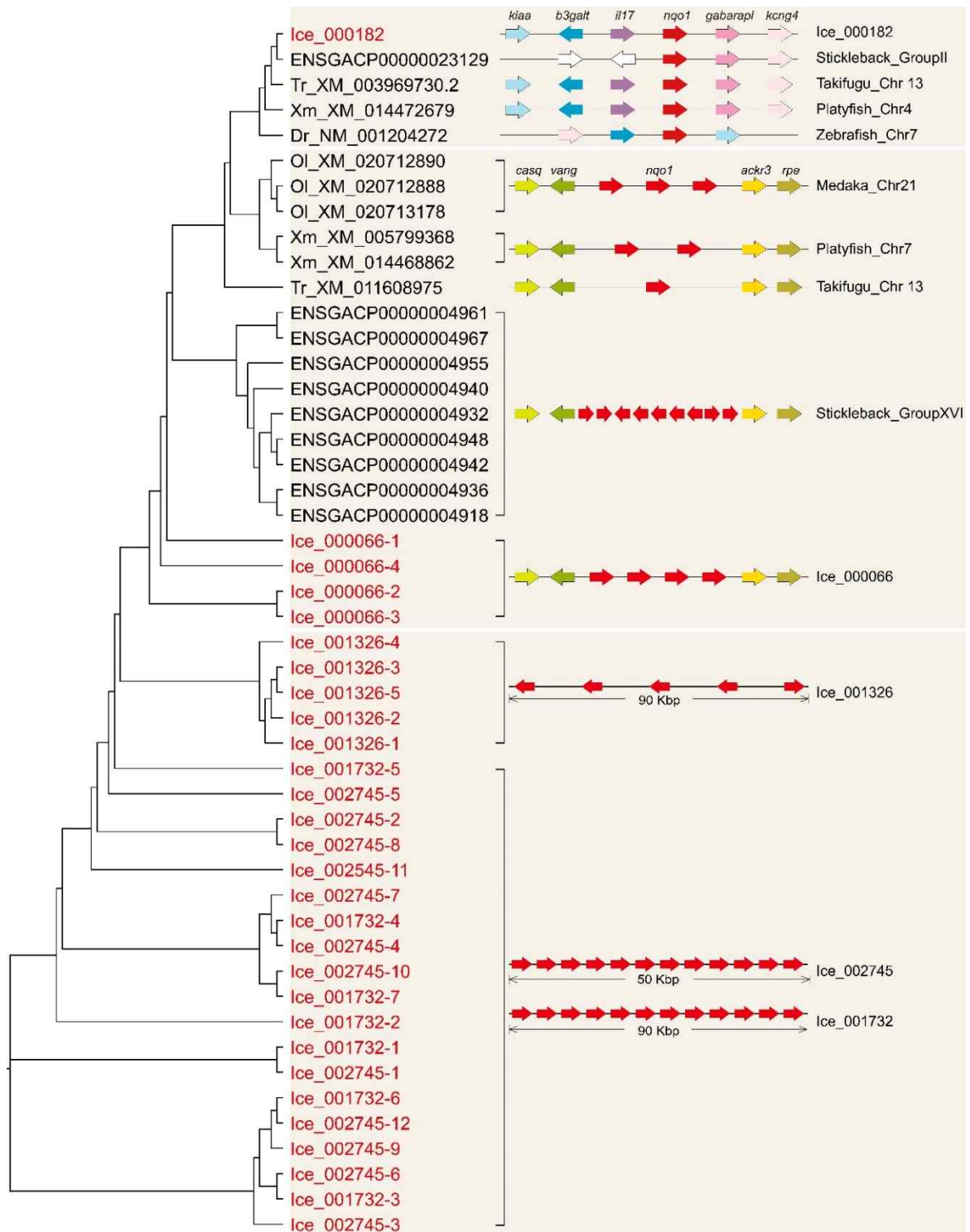


극지연구소

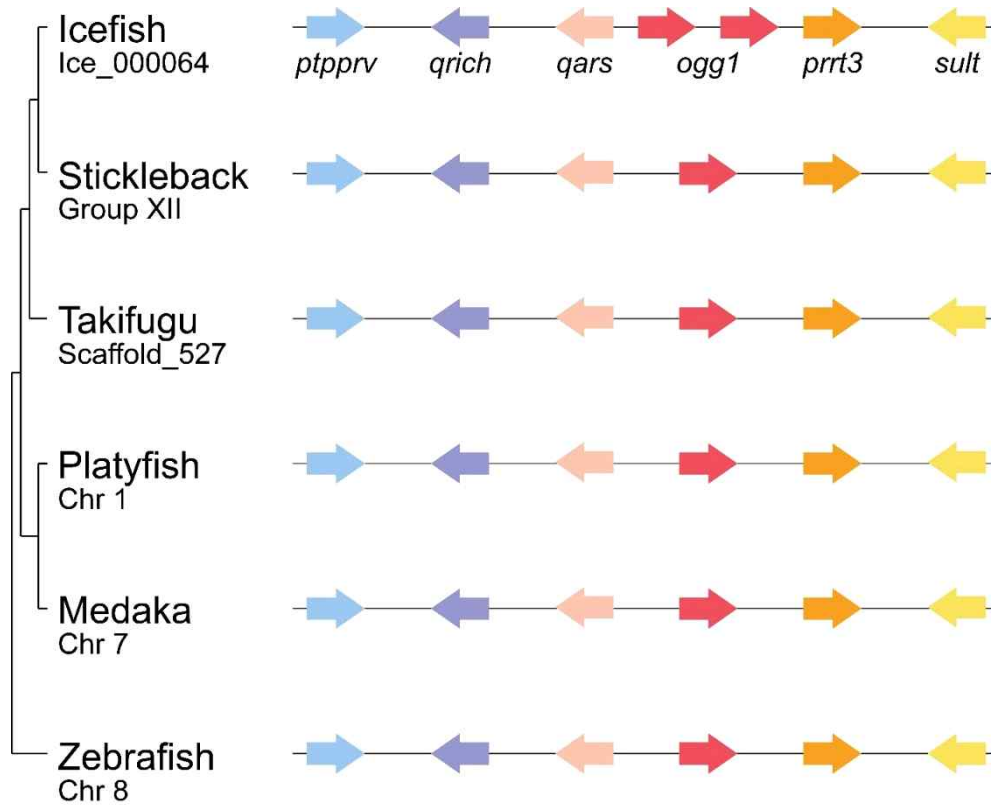
Supplementary Figure 12. Syntenic comparison of myoglobin and cytoglobin genes identified in icefish genome. Genomic structure and syntenic comparison of a. cytoglobin (*cygb1*) b. *cygb2*, c. myoglobin (*mb*) and d. neuroglobin (*ngb*) genes are shown within representative sequenced teleost genomes.



Supplementary Figure 13. Expansion of NAD(P)H quinone dehydrogenase 1 (*nqo1*) genes in the icefish genome. Comparison of genomic structure, synteny, and phylogenetic relationship of icefish *nqo1* genes shown within representative sequenced teleost genomes. Icefish *nqo1* genes are distributed on five different contigs.

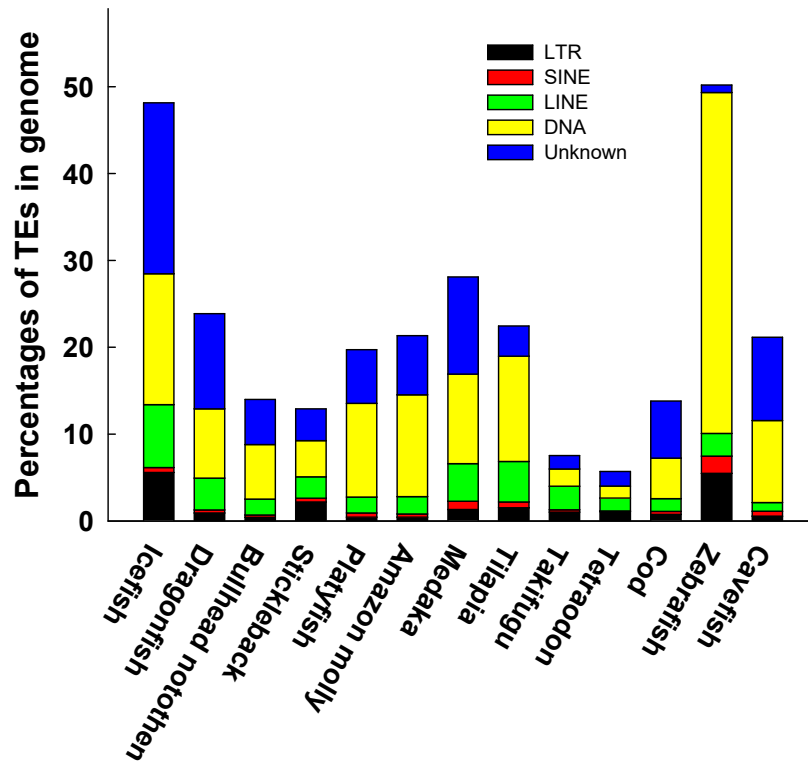


Supplementary Figure 14. The 8-oxoguanine DNA glycosylase (*ogg1*) gene is duplicated in the icefish genome. Other teleost genomes contain a single gene.

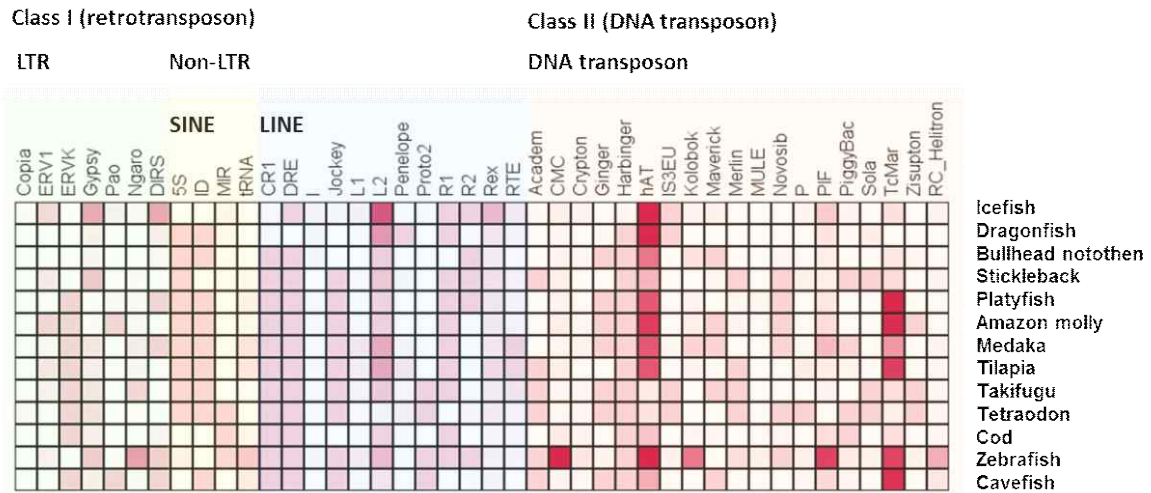


극지연구소

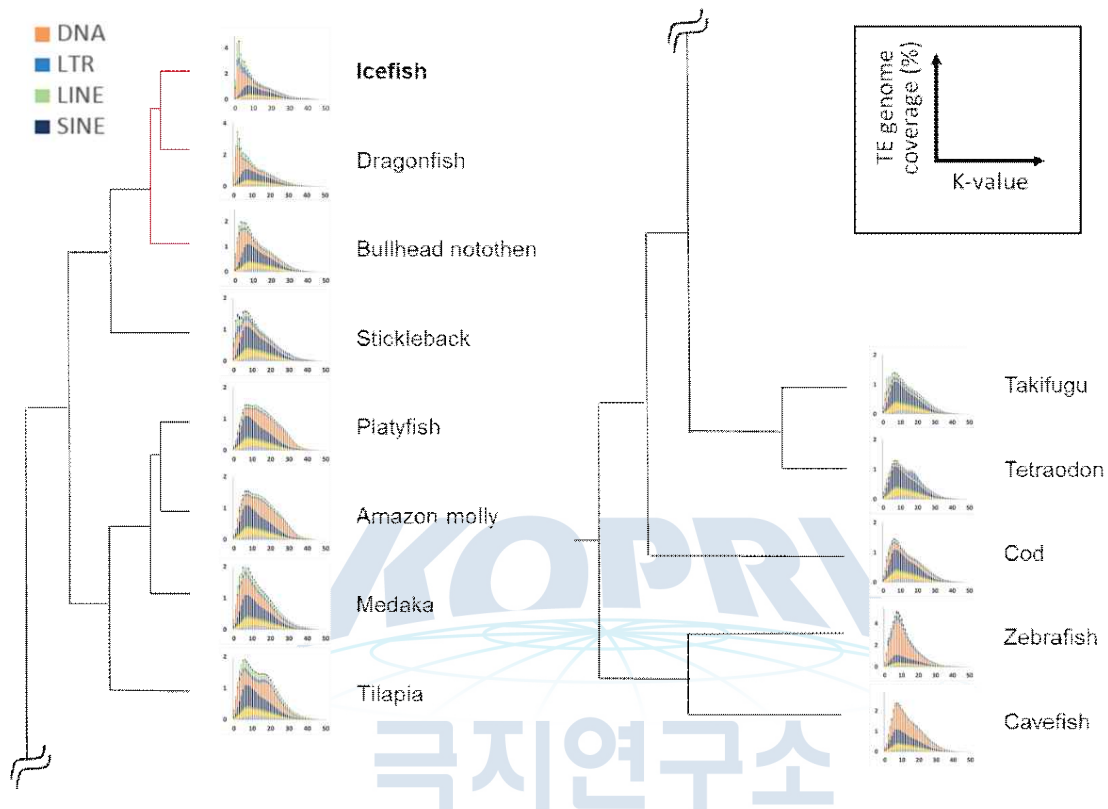
Supplementary Figure 15. Percentages of TEs, DNA transposons, LTR, LINE, and SINE retrotransposons, in different teleost genomes. The amounts of DNA transposons, LTR, LINE, and SINE retrotransposons, and unclassified elements (Unknown), as well as their respective proportions were estimated based on RepeatMasker outfiles.



Supplementary Figure 16. Diversity and abundance of TE superfamilies in teleosts. Presence or absence of TE superfamilies was determined using automatic annotation, manual verification, and literature information. The relative proportion of superfamilies is shown in squares according to the gradient scale (log) at the bottom.



Supplementary Figure 17. Kimura distance-based copy divergence analysis² of transposable elements in teleost genomes. Graphs represent genome coverage (Y-axis, %) for each type of TE (DNA transposons, SINE, LINE, and LTR retrotransposons) in the different genomes analyzed, clustered to their corresponding consensus sequence according to Kimura distances (X-axis, K-value from 0 to 50). Right inserted square shows axes for all graph.



제 3-2절 남극 드래곤피쉬 *Parachaenichthys charcoti*의 게놈 분석을 통한 극한환경 적응 기작의 이해

1. 서론

남극해의 어류 동물상은 132종과 8종으로 구성된 notothenioid teleost에 속하는 단일 혈통에 의해 점유되어 있다. 모든 남극 동물들은 극심한 계절적 먹이량 변화와 지속적인 저수온을 포함하는 극심한 남극해 환경에 적응하기 위해 진화해 왔다. Notothenioids는 남극해 어류 생물 총량의 90%를 차지하며, 추운 환경에서 살아남기 위해 분자와 생태학적 다양화를 겪어 왔다. 이 생물은 진화와 조상형의 유전체 연구에 있어서 매력적인 어류 모델로 여겨져 왔다. Bathydraconidae는 남극 teleost의 대부분을 차지하는 Notothenioidei에 속하는 과(family)이다 (Eastman et al. 1993; Eastman et al. 1998; Eastman et al. 2000; Eakin et al. 2009). 남극 bathydraconid 드래곤피쉬인 *Parachaenichthys charcoti*는 1906년 Vailant에 의해 처음 기록되었다 (Notothenioidei: Bathydraconidae; AphiaID: 234687; Fishbase ID: 7102). 이 생물은 남극의 포터 소만, 사우스 셔틀랜드 제도 등에서 발견된다. *P. charcoti*는 개체 발생 내내 내륙붕(inner shelves)에 머무는 것으로 알려져 있다(Casaux et al. 1990). 지금까지 생태학과 행동학에 대한 연구가 이루어져 왔으나, 아직까지 유전체 연구는 이루어진 바가 없다(Casaux et al. 1990; Barrera-Oro et al. 2002; Barrera-Oro et al. 2010; Eastman et al. 2002). 남극 teleost의 다양성과 진화를 이해하기 위한 유용 자료를 제공하기 위해서는 이 남극 어류 특이적인 특성 확인과 포괄적 유전자 연구가 필수적이다.

2. 연구개발 수행결과

1) 라이브러리 구성 및 DNA 염기서열 해독

P. charcoti(길이: ~45cm)는 2012년 1월 남극 대륙 북부의 세종기지 근처 Marian 소만 (62°14,58'47W)의 20~30m 깊이에서 낚시로 채집하였다(Figure 1). Gentra Puregene Blood Kit (Qiagen, Valencia, CA, USA)를 사용하여 *P. charcoti*로부터 고분자량의 genomic DNA를 추출하였다. DNA 염기서열 분석을 위해 300-, 400-, 450-bp 사이즈로 조각낸 DNA 조각으로부터 3가지 종류의 paired-end 라이브러리(PE300, PE400, and PE450)를 제작하였으며 표준 Illumina 라이브러리 준비 방법을 사용하였다. 스캐폴딩을 위해 3-kb, 5-kb, 8-kb, 20-kb 사이즈의 mate-pair 라이브러리(MP3K, MP5K, MP8K, MP20K)를 제작하였으며 제작사의 표준 준비 방법을 사용하였다(Illumina, San Diego, CA, USA). Draft genome의 유전자 명명을 위

해 expressed sequence tag이 필수적이기 때문에 Truseq Sample preparation v.2 (Illumina)를 사용해 전사체 라이브러리를 전체 RNA를 사용하여 제작하였다. 전체 RNA는 간조식으로 제작하였으며 RNeasy Mini Kit (Qiagen)와 RNase-Free DNaseI Kit (Qiagen)를 사용해 정제하였다. 추출한 RNA 샘플의 질과 농도 분석은 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA)를 사용하였다. 2µg의 전체 RNA로부터 mRNA를 분리하였으며 poly-A 분리방법을 사용하여 cDNA 라이브러리를 제작하였다. 전사체 염기서열 해독을 위해 cDNA를 500-bp 크기로 자른 후, Illumina 표준 방법을 사용하여 paired-end 라이브러리(PE500)를 제작하였다. 최종 전사체 라이브러리의 길이와 농도는 2100 Bioanalyzer를 사용하여 확인하였다. 전사체 염기서열 해독은 300 x 2 paired-end reads를 사용하였다 (Table 1). Illumina 염기서열 해독 결과는 FASTXToolkit (v. 0.0.11) (Gordon et al. 2017) 의 parameters -t 20, -l 70, -Q33 조건으로 트리밍 하였으며 paired 서열만을 다음 분석에 사용하였다. Table 1의 모든 염기서열 해독은 극지연구소에서 이루어졌다.

2) 유전체 어셈블리

DNA paired-end 라이브러리의 유전체 사이즈를 추정하기 위해 Jellyfish 2.2.5 (Jellyfish,

RRID:SCR_005491) (Marcais et al. 2011)를 사용하여 K-mer 분석을 하였다. 유전체의 예상 크기는 805 Mb이며, main peak의 크기로 보아 coverage depth는 x39이다 (figure 2). 트리밍한 paired-end 리드를 이용하여 초기 어셈블리를 진행하였으며 이때 Celera Assembler v. 8.3 (Celera Assembler, RRID:SCR_010750) (Myers et al. 2000)을 사용하였다. Celera Assembler 분석을 위해 paired-end 리드 데이터 Celera Assembler에 포함되어있는 GastqToCA를 이용하여 FRG포맷으로 변환하였다. Intel Xeon X7460 2.66 GHz 프로세서, 1Tb RAM, 80 프로세서 워크스테이션을 이용하여 어셈블리를 진행하였으며 사용한 파라미터는 다음과 같다: overlapper = ovl, unitigger = bogart, utgErrorRate = 0.03, utgErrorLimit = 2.5, utgGraphErrorRate = 0.030, utgGraphErrorLimit = 3.25, ovlErrorRate = 0.06, cnsErrorRate = 0.06, cgwErrorRate = 0.1, merSize = 28, doOverlapBasedTrimming = 1, merylMemory = 500 000, merylThreads = 40, ovlMemory = 8 Gb, ovlThreads = 2, ovlConcurrency = 40, ovlHashBlockLength = 300000 000, ovlRefBlockSize = 7 630 000, ovlHashBits = 24. 초기 어셈블리 크기는 709 Mb이었으며, N50 contig 길이는 5039 bp, N50 scaffold 길이는 6135 kb였으며, GC content는 40.66%이었다. Celera Assembler로 어셈블된 contig결과를 보면 contig coverage는 약 ×36.57 이었다. 초기 어셈블리 결과인 contig는 scaffolding에 사용되었으며 이 단계에서는 SSPACE v. 2.0 (SSPACE, RRID:SCR_005056)를 이

용하였으며 사용된 파라미터는 다음과 같다: -x 0, -k 3, -a 0.8, and -T 60(Boetzer et al. 2011). mate-pair 리드의 트리밍은 FASTX-Toolkit을 사용하였으며 scaffolding 단계에 활용하였다. scaffolding 후 scaffold의 개수는 153,398에서 12,381로 감소하였으며, N50 scaffold 길이는 6,135에서 166,726 bp로 증가하였다(Table 2). 최종 scaffold의 크기(~795 Mb)는 예상 유전체 크기(805 Mb)와 비슷하였다.

3) 유전자 명명

MAKER 유전자 명명 파이프라인(MAKER, RRID:SCR_005309)의 기본 설정으로 유전체 유전자 명명을 하였다(Cantarel et al. 2008). RepeatMasker v. 3.3.0 (RepeatMasker, RRID:SCR_012954)의 de novo repeat library를 사용하여 초기 repetitive elements를 구분하였다(Tarailo-Graovac et al. 2009). repeat library는 Repbase library (Ver. 20,140,131)를 기반으로 RepeatModeler v. 1.0.3 (Repeat-Modeler, RRID:SCR_015027) (Bao et al. 2002)를 사용하여 구축하였다. masked genome sequence로부터 ab initio 유전자 예측을 위해 SNAP gene finder (Korf et al. 2004)를 사용하였다. RNA 예측 증거로는 BLASTn을 사용한 전사체 어셈블리가 사용되었으며, 단백질 예측 증거로는 tBLASTx를 이용한 상동 단백질 정보가 사용되었다. PE500 라이브러리로부터 생성된 염기서열 리드를 CLC Genomics Workbench 8.0를 사용하여 전사체 어셈블리를 진행하였다. 총 6종의 단백질 서열이 분석에 사용되었으며 종 정보는 다음과 같다: *Notothenia coriiceps* (NCBI reference sequence NC_015653.1), *Danio rerio*, *Gasterosteus aculeatus*, *Takifugu rubripes*, *Tetraodon nigroviridis*, *Gadus morhua*(모든 데이터는 Ensembl release 69임). MAKER2는 유전자 명명의 품질 관리를 위해 annotation edit distance(AED) 계량방법이 적용되었다(Holt et al. 2011). AED 임계값은 0과 1사이의 값으로 한정되며, AED 임계값 0은 정렬된 증거와 주석을 단 유전자가 정확히 일치함을 보여주고 반대로 1의 값은 증거가 지지해주지 않음을 타낸다. 그러나 AED 임계값은 이러한 유전자 예측에 적용되지 않으나, AED 값은 유전자 주석에 표시되었으며, 상동유전자 분석과 유전자 획득 및 상실 분석 단계에서 표시 값이 고려되었다. 최종 유전자 명명 모델 단계에서 MAKER2를 이용하였으며, 모든 입력 데이터들이 사용되었으며, 그 결과 총 32,712개의 유전자가 예측되었다(Table 2). 명명된 유전자들은 평균 8개의 exon을 가지고 있으며, 평균 mRNA 길이는 1,412bp이고, 평균 coding DNA (CDS) 염기서열의 길이는 1,291bp이다. MAKER2를 사용한 repeat element 분석 결과 *P. charcoti* 유전체의 19.41%가 반복서열이었다.

유전체의 어셈블리와 유전자명명의 완성도를 추정하기 위해 Benchmarking Universal Single-Copy Orthologs (BUSCO) 분석(BUSCO, RRID:SCR_015008) (Simão et al. 2015)을 진행하였으며 이때 actinopterygii 계통 특이적인 프로파일 라이브러리를

사용하였다. BUSCO 분석 결과 88.6%의 진핵생물 전체 완성도를 보였으며 5.7%의 진핵생물 부분 완성도를 보이는 것을 확인하였다(Table 3). 32,712개의 유전자에 예상 기능을 할당하기 위해 Blast2GO v. 2.6.0 (Blast2GO, RRID:SCR_005828)를 사용하였다(Conesa et al. 2005). BLASTp 결과와 InterproScan(RRID:SCR_005829)을 사용하여 명명을 한 예상 유전자 18,455개(56.42%)의 기능 분류를 하였다. 유전자 ontology 명명 결과, “biological process” (20126, 61.52%), “molecular function” (20514, 62.71%), “cellular component” (15452, 47.23%)로 분류되었다. 총 3,846개의 단백질에 대해 효소 commission번호가 확인되었다.

4) 상동유전자 분석

Markov 클러스터링 알고리즘 기반하여 다중 진핵 유전체를 subgraph 분석하는 OrthoMCL(v. 2.0.5) (Li et al. 2003)을 사용하여 orthologous 그룹을 선별하였으며, 모든 분석단계에서 기본 파라미터인 percentMatchCutoff = 50 와 evalueExponentCutoff = -5를 적용하였다. 분석에는 총 7종의 어류 유전체 정보 (*D. rerio*, *G. aculeatus*, *T. rubripes*, *T. nigroviridis*, *G. morhua*, *N. coriiceps*, *P. charcoti*)가 사용되었다. 5종의 coding 염기서열은 Ensembl release 69를 사용하였으며, 한 유전자에 대한 여러 가지 타임의 단백질 중 한 개의 coding서열만이 분석에 사용되었다. *N. coriiceps*의 NCBI reference 서열 내의 coding 서열을 사용하여, *P. charcoti*는 MAKER 파이프라인에서 서로 다른 AED 값 (1, 0.75, 0.25)을 적용하여 생성한 세가지 그룹의 coding서열을 사용하였다. AED 임계값이 1인 경우 7종의 어류에서 공통으로 8951개의 orthologous 그룹이 확인되었다. *N. coriiceps* 유전자 32636 중 288개와 *P. charcoti* 유전자 32712 중 33개는 어떤 종에서도 발견되지 않았다. 2519개 유전자 그룹은 2종의 남극 물고기에서만 있는 것으로 확인되었다(Figure 3A). 유전자 예측에 대해 AED 임계값 0.25를 적용하자 *P. charcoti*에서 7568개의 orthologous 그룹이 확인되었다.

5) 유전자 획득 및 상실에 대한 가능성도 분석

진화과정을 통해 유의미한 크기 변화를 겪은 유전자 그룹을 식별하기 위해 orthologs의 크기 차이를 추정하였다(Hahn et al. 2005; Hahn et al. 2007). CAFE 3.0 프로그램(De et al. 2006)을 사용하여 AED 임계값이 서로 다른 *P. charcoti*의 3가지 유전자 그룹 결과에 대해 분석을 진행하였다. Newick 포맷의 bifurcating 계통수를 작성하기 위해 어류 7종의 coding 염기서열을 이용하여 계통유전학적 분석을 진행하였다. 총 8951개의 orthologous 유전자 세트가 상호 최고 BLASPTP hit 기준을 사용하여 선택되었으며, 그 과정에서 염기서열의 정렬은 PRANK(v. 130820) 프로그램(Loytynoja et al. 2005)의 codon model “-dna -codon”을 이용하였으며, 정렬이 잘

되지 않은 부분은 Gblock (v. 0.91) 프로그램의 codon model “-t=c” 설정을 이용해서 제거하였다(Castresana et al. 2000). 정렬이 잘 된 부분의 유전자들은 모두 연결하여 MEGA (v.7) 프로그램의(RRID:SCR_000667) (Tamura et al. 2013) neighbor-joining 방법을 이용하여 계통수 그렸다. CAFE 3.0 분석에 필요한 중분화 계통수 의 시간 단위는 TimeTree(Hedges et al. 2006) 데이터베이스의 중분화 시점 데이터를 사용하였다(Figure 3B). CAFE 3.0 프로그램은 $P < 0.05$ 를 적용하여 분석을 진행했으며, 출생률(λ)과 사망률(μ)은 프로그램 내의 “-s” 설정을 이용하는 LambdaMu를 이용하여 계산하였다. 분석 결과 *P. charcoti*는 937개의 유전자군을 획득하였으며 1916개의 유전자군에서 상실을 겪었다(Figure 3B).

남극 드래곤피쉬 *P. charcoti*는 빙어의 자매 혈통에 속하는 종으로, 헤모글로빈이 없는 유일한 척추동물이다. 드래곤피쉬(Bathydraconidae)와 빙어(Channichthyidae)는 일반적으로 일반적인 notothenioid의 조상으로부터 진화한 것으로 여겨졌는데, 헤마토크리트와 혈액 헤모글로빈 농도가 감소된 것이 특징이다(Bargelloni et al. 1998; Beers et al. 2010; D’avino et al. 1988; Di et al 1998; Kunzmann et al. 1991). 드래곤피쉬는 이러한 경향에서 적혈구 notothenioid 분류군 중 가장 유사한 양상을 보였으며, 드래곤피쉬의 글로빈 콤플렉스는 빙어의 조상형이 기능을 잃기 이전의 길이와 구성이 비슷했을 것이라는 가설이 세워졌다(Lau et al. 2012). 최근 발간된 *N. coriiceps* 유전체(Shin et al. 2014)와 함께 *P. charcoti*의 유전체는 남극 물고기가 영하의 온도에서 살아남기 위해 어떻게 진화해왔는지에 대한 우리의 이해를 넓힐 것이고 헤모글로빈 없는 남극 물고기와 이들의 뚜렷한 표현형(혈유량 증가, 저혈점도, 커다란 보어 모세관, 혈관의 비대와 혈관 증가, 심장비대, 고혈류 등)으로 진화하는 과정을 이해하는 중요한 단서를 제공할 것이다.

3. 참고문헌

1. Eastman JT , Pratt D, Winn W. Antarctic Fish Biology: Evolution in a Unique Environment. San Diego: Academic Press; 1993.
2. Eastman JT , Clarke A. A comparison of adaptive radiations of Antarctic fish with those of nonAntarctic fish. In: Fishes of Antarctica. Milan: Springer; 1998:3 - 26.
3. Eastman JT . Antarctic notothenioid fishes as subjects for research in evolutionary biology. Antarct Sci2000;12(03):276 - 87.
4. Eakin RR , Eastman JT, Near TJ. A new species and a molecular phylogenetic analysis of the Antarctic fish genus Pogonophryne (Notothenioidei:

- Artedidraconidae). *Copeia*2009;4(4):705 - 13.
5. Casaux RJ , Mazzotta AS, Barrera-Oro ER. Seasonal aspects of the biology and diet of nearshore nototheniid fish at Potter Cove, South Shetland Islands, Antarctica. *Polar Biol*1990;11(1):63 - 72.
 6. Barrera-Oro E . The role of fish in the Antarctic marine food web: differences between inshore and offshore waters in the southern Scotia Arc and west Antarctic Peninsula. *Ant Sci*2002;14(4):293 - 309.
 7. Barrera-Oro ER , Lager C. Egg-guarding behaviour in the Antarctic bathydraconid dragonfish *Parachaenichthys charcoti*. *Polar Biol*2010;33(11):1585 - 7.
 8. Eastman JT , Sidell BD. Measurements of buoyancy for some Antarctic notothenioid fishes from the South Shetland Islands. *Polar Biol*2002;25(10):753 - 60.
 9. Gordon A , Hannon GJ. Fast-toolkit. FASTQ/A short-reads pre-processing tools. http://hannonlab.cshl.edu/fastx_toolkit. Accessed 27 Feb 2017.
 10. Marçais G , Kingsford C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics*2011;27(6):764 - 70.
 11. Myers EW , Sutton GG, Delcher AL et al. A whole-genome assembly of *Drosophila*. *Science*2000;287(5461):2196 - 204.
 12. Boetzer M , Henkel CV, Jansen HJ et al. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics*2011;27(4):578 - 9.
 13. Cantarel BL , Korf I, Robb SMC et al. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res*2008;18(1):188 - 96.
 14. Tarailo Graovac M , Chen N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr Protocols Bioinform*2009:4 - 10.
 15. Bao Z . Automated de novo identification of repeat sequence families in sequenced genomes. *Genome Res*2002;12(8):1269 - 76.
 16. Korf I . Gene finding in novel genomes. *BMC Bioinformatics*2004;5(1):59.

17. Holt C , Yandell M. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics*2011;12(1):491.
18. Simão FA , Waterhouse RM, Ioannidis Pet al. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*2015;31(19):3210 - 2.
19. Conesa A , Gotz S, Garcia-Gomez JMet al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*2005;21(18):3674 - 6.
20. Li L . OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res*2003;13(9):2178 - 89.
21. Hahn MW , De Bie T, Stajich JEet al. Estimating the tempo and mode of gene family evolution from comparative genomic data. *Genome Res*2005;15(8):1153 - 60.
22. Hahn MW , Han MV, Han S. Gene family evolution across 12 drosophila genomes. *PLoS Genet*2007;3(11):e197.
23. De Bie T , Cristianini N, Demuth JPet al. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics*2006;22(10):1269 - 71.
24. Loytynoja A , Goldman N. From the cover: an algorithm for progressive multiple alignment of sequences with insertions. *Proc Natl Acad Sci U S A*2005;102(30):10557 - 62.
25. Castresana J . Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol*2000;17(4):540 - 52.
26. Tamura K , Stecher G, Peterson Det al. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol*2013;30(12):2725 - 9.
27. Hedges SB , Dudley J, Kumar S. TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics*2006;22(23):2971 - 2.
28. Balushkin A , Voskoboynikova O. Systematics and phylogeny of antarctic dragonfishes (Bathydraconidae, Notothenioidei, Perciformes). *J*

- Ichthyol1995:35(5):89 - 104.
29. Derome N , Chen W, Dettai Aet al. Phylogeny of Antarctic dragonfishes (Bathypagrus, Notothenioidei, Teleostei) and related families based on their anatomy and two mitochondrial genes. Mol Phylogenet Evol2002;24(1):139 - 52.
 30. Near TJ . A genomic fossil reveals key steps in hemoglobin loss by the Antarctic icefishes. Mol Biol Evol2006;23(11):2008 - 16.
 31. Bargelloni L , Marcato S, Patarnello T. Antarctic fish hemoglobins: evidence for adaptive evolution at subzero temperature. Proc Natl Acad Sci U S A1998;95(15):8670 - 5.
 32. 32. Beers JM , Borley KA, Sidell BD. Relationship among circulating hemoglobin, nitric oxide synthase activities and angiogenic poise in red- and white-blooded Antarctic notothenioid fishes. Comp Biochem Physiol A Mol Integr Physiol2010;156(4):422 - 9.
 33. 33. D'avino R , Di Prisco G. Antarctic fish hemoglobin: an outline of the molecular structure and oxygen binding properties—I. Molecular structure. Comp Biochem Physiol B Comp Biochem1988;90(3):579 - 84.
 34. Di Prisco G. Molecular adaptations of Antarctic fish hemoglobins. In: Fishes of Antarctica. Milan: Springer: 1998:339 - 53.
 35. Kunzmann A , Caruso C, Prisco GD. Haematological studies on a high-Antarctic fish: Bathypagrus marri Norman. J Exp Mar Biol Ecol1991;152(2):243 - 55.
 36. Lau Y , Parker SK, Near TJ et al. Evolution and function of the globin intergenic regulatory regions of the Antarctic dragonfishes (Notothenioidei: Bathypagrus). Mol Biol Evol2012;29(3):1071 - 80.
 37. Shin SC , Ahn DH, Kim SJ et al. The genome sequence of the Antarctic bullhead notothen reveals evolutionary adaptations to a cold environment. Genome Biol2014;15(9):468.
 38. Ahn DH , Shin SC, Kim BM et al. Supporting data for “Draft genome of the Antarctic dragonfish, *Parachaenichthys charcoti*.” GigaScience Database 2017.

Table 1. *P. charcoti* sequencing statistics

Library	Mode	Insert size (bp)	Library type	Trimmed reads	Trimmed sequence (bp)	Source
PE300	2 × 300	300	Paired-end	28,776,064	4,964,428,226	Genomic DNA
PE400	2 × 300	400	Paired-end	139,126,700	29,538,419,473	Genomic DNA
PE450	2 × 300	450	Paired-end	85,834,292	16,644,575,781	Genomic DNA
MP3K	2 × 300	3000	Mate-pair	70,517,546	4,925,657,177	Genomic DNA
MP5K	2 × 300	5000	Mate-pair	66,623,428	4,626,486,038	Genomic DNA
MP8K	2 × 300	8000	Mate-pair	61,240,982	4,212,744,363	Genomic DNA
MP20K	2 × 300	20000	Mate-pair	86,575,644	5,387,730,972	Genomic DNA
PE500	2 × 300	500	Paired-end	25,940,404	5,571,197,784	Liver RNA



Table 2. Global statistics of the *P. charcoti* genome assembly

		<i>P. charcoti</i>
Scaffold	Total scaffold length (bases)	794,596,176
	Gap size (bases)	86,840,902
	Scaffolds (n)	12,602
	N50 scaffold length (bases)	178,362
	Max scaffold length (bases)	1,318,127
Contig	Total contig length (bases)	709,540,340
	Contigs (n)	153,398
	N50 contig length (bases)	6145
	Max contig length (bases)	65,864
Annotation	Gene number (n)	32,712
	An average mRNA length (bases)	1412
	An average CDS length (bases)	1291
	An average of exons (n)	8
Repeat content (% of genome)		19.4

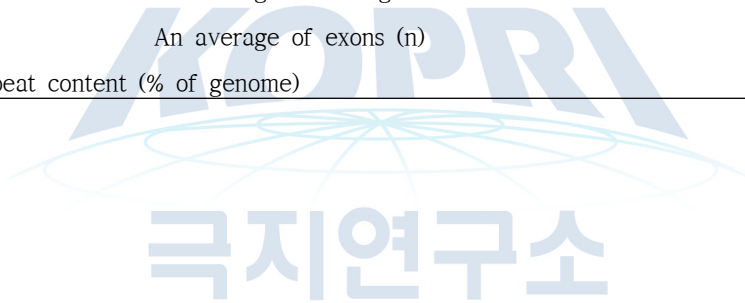


Table 3. Summarized benchmarks of the BUSCO assessment

	Actinopterygii (%)
Total BUSCO groups searched	4062 ^a
Complete BUSCOs	88.6
Complete and single-copy	86.3
Complete and duplicated	2.3
Partial	5.7
Missing	5.7

aNumber of total BUSCO groups searched.





그림 1. Photograph of Antarctic dragonfish, *P. charcoti*.



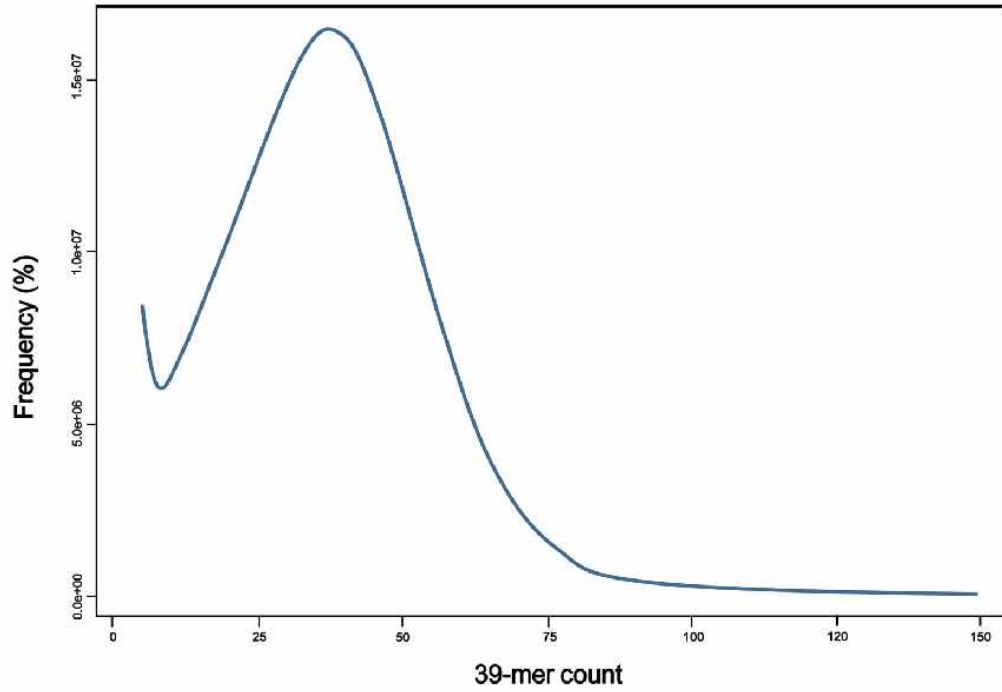
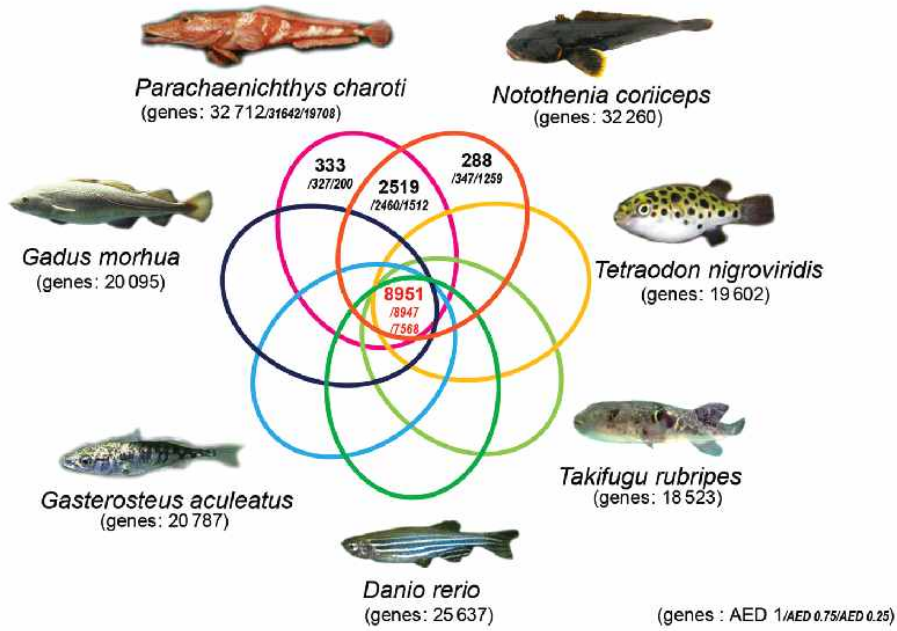


그림 2. Estimation of the *P. charcoti* genome size based on 39-mer analysis. X-axis represents the depth (peak at $\times 39$), and the y-axis represents the proportion. Genome size was estimated to be 805 Mb (total k-mer number/volume peak).

극지연구소

(A)



(B)

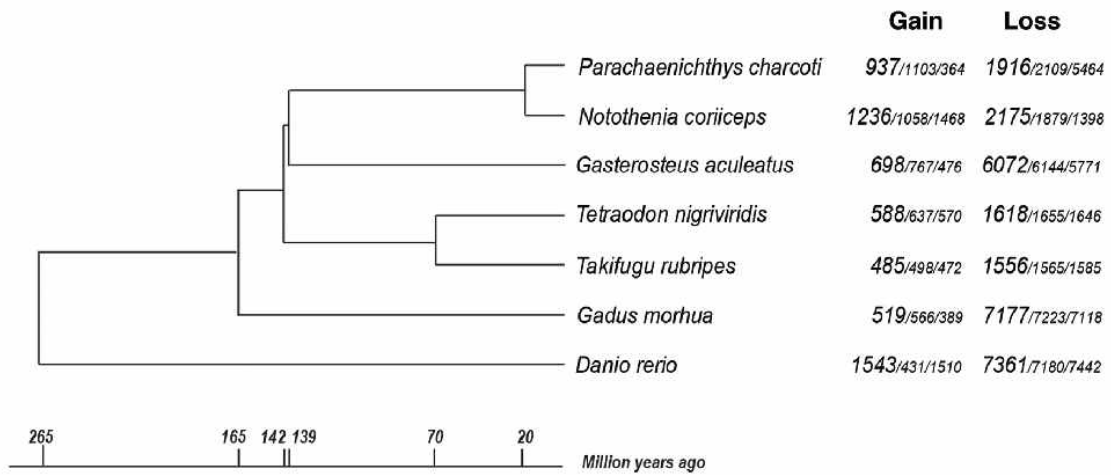


그림 3. Comparative genome analyses of the *P. charcoti* genome. (A) Venn diagram of orthologous gene clusters between 4 arthropod lineages. (B) Gene family gain and loss analysis. The number of gained gene families and lost gene families are indicated for each species. Time lines specify divergence times between the lineages.

제 3-3절 남극개미자리의 현장 및 챔버 생육 샘플의 전사체 비교를 통한 환경적응 기작 분석

Analysis on the environmental adaptation mechanism of an Antarctic flowering plant, *Colobanthus quitensis* (Kunth) Bartl., using comparative transcriptome analysis of field- and chambergrown samples.

요약

남극개미자리는 남극에 서식하는 관속식물 2종 중 하나이다. 남극개미자리는 남극의 혹독한 자연환경에서 생존해온 대표적인 극한 생물이다. 이 종은 자연 서식지에서는 쿠션이나 매트 형태로 자라며 일반적으로 고위도 및 고산지역에서 관찰된다. 이 종은 기후변화에 의한 지리분포 및 생리적 적응반응을 위한 지표종으로 수년간 연구되었으나 유전정보의 부재로 분자유전학적 연구는 매우 제한적으로 이루어져왔다. 본 과제에서는 남극개미자리의 전사체를 확보하고 이를 분석함으로써 극한지 식물의 환경적응기작을 규명하고자 하였다. 남극현장 및 배양챔버에서 자란 식물의 RNA library를 제작하고 HiSeq short-read sequencing을 수행하였으며 De novo assembly 및 blast 분석을 통해 47,070개의 blast-hit이 존재하는 contig를 확인하였다. DEG (Differentially expressed gene) 분석 결과, 광합성 경로 유전자군 및 환경스트레스 반응 유전자군 등이 남극현장샘플에서 발현이 현저하게 증가하는 것을 확인하였다. 특히 열형태형성 조절유전자로 알려진 PIF (Phytochrome interacting factor) 유전자군이 남극현장샘플과 배양챔버샘플 사이에 발현 차이가 나타나며, 이들의 발현량이 온도에 따라 다르다는 것을 확인하였다. 따라서, 이들 유전자의 차등적 발현이 극지 환경에서의 남극개미자리의 생장을 조절하여, 추운 환경에서 열 손실을 최소화하는 쿠션 모양의 성장형태를 유도하리라 여겨진다.

Abstract

Colobanthus quitensis is one of the two vascular plants inhabiting the Antarctic. In natural habitats, it grows in the form of a cushion or mats, commonly observed in high latitudes or alpine vegetation. Although this species has been investigated over many years to study its geographical distribution and physiological adaptations to climate change, very limited genetic information is available. The high-throughput sequencing with a de novo assembly analysis yielded 47,070 contigs with blast-hits. Through the functional classification and enrichment

analysis, we identified that photosynthesis and phenylpropanoid pathway genes show differential expression depending on the habitat environment. We found that the known 'plant core environmental stress response (PCESR)' genes were abundantly expressed in Antarctic samples, and confirmed that their expression is mainly induced by lowtemperature. In addition, we suggest that differential expression of thermomorphogenesis-related genes may contribute to phenotypic plasticity of the plant, for instance, displaying a cushion-like phenotype to adapt to harsh environments.

1. 서론

극한의 온도, 가뭄, 높은 염분 및 높은 자외선은 세포 구조를 손상시키고 생리 기능을 손상시킬 수 있는 전형적인 환경 스트레스 요인으로 식물의 광합성을 억제하고 성장을 지연시킨다. 이러한 환경 스트레스 요인에 대처하기 위해, 식물은 스트레스 신호 전달을 조절하여 세포막 유동성 증가 및 세포보호물질의 생산 등, 세포 내 생리학적 및 대사 변화를 포함하는 스트레스 저항 전략을 개발했다. 한편, 식물은 환경의 변화에 따라 생리학 및 형태학적 변화를 겪게 되는데 이를 표현형가소성 (phenotypic plasticity)라 일컫는다.

“Cushion plant”라 불리는 종들은 높은 위도 또는 고산 식물에서 나타나는데, 일반적으로 둥근 반구형 혹은 매트 형태로 나타난다. 극지생태계에서 이러한 성장 형태는 식물의 생존에 도움이 될 뿐 아니라 열과 수분의 손실을 최소화하여, 미생물, 곤충 및 유식물이 생존하기에 적합한 서식처로서 생태적 기능을 하는 것으로 생각된다. 이러한 식물의 성장 형태는 다양한 과와 속에 걸쳐 나타나는데, *Silene acaulis* (Caryophyllaceae), *Azorella compacta* (Apiaceae), *Androsace helvetica* (Primulaceae), *Raoulia eximia* (Asteraceae) 등이 대표적이며, 이는 극한 환경에 대처하기 위한 표현형 가소성과 관련된 수렴진화의 대표적 예라 할 수 있다.

남극개미자리(*Colobanthus quitensis*, Caryophyllaceae)는 쿠션이나 매트 형태로 자라는 자가수정 종으로서, 멕시코 (17°N)에서 남극 반도 (68°S)까지 넓은 지역에 분포하며, 해발 0~4200 미터의 고도에 분포하고 서식지에 따라 상당한 형태학적 차이를 보이는 것으로 알려져 있다. 집단에 따른 형태학적 비교 분석에 따르면, 위도가 높은 곳에 서식하는 집단은 mesophyll 층이 두껍고, 전체 표면적이 상대적으로 감소하는, 즉, 더 작고 두꺼운 잎으로 구성되며, 잎 조직의 색소가 상대적으로 증가되어 있음이 밝혀졌다. 남극개미자리의 광합성 효율은 주로 온도에 의해 조절되었으며, 특히, OTC (Open Top Chamber) 현장연구에 따르면, OTC로 인한 온도상승효과는 식물 형태학적 특성 외에도 세포 화학 성분 및 광합성 변수에 영향을 미침으로써 식물 생

장에 영향을 미치는 것으로 나타났다.

남극개미자리에 관한 많은 연구결과에도 불구하고, 자연 서식지에서 자라는 식물을 이용한 유전자 조절기작에 대한 연구결과는 전무하였다. 따라서 본 과제에서는, 남극 현장에서 남극개미자리의 유전자조절 네트워크를 규명하기 위해, de novo 전사체를 구축하였고, 남극현장샘플에서 많이 발현되는 유전자그룹을 선별하였다. 이들에 대한 GO·KEGG 분석 및 ortholog 분석을 통해, 유전자 서열 유사성에 따른 생물학적 기능을 추론하고 그에 따른 분류를 진행하였다. 그 결과, 광합성경로 유전자군 및 환경스트레스 반응 유전자군, 열형태형성관련 유전자군 등이 남극현장샘플에서 발현이 현저하게 증가하는 것을 확인하였다.

2. 재료 및 방법

가. 연구지역 및 샘플링 방법

2013년 1월 킹조지 섬의 바톤 반도에 있는 세종과학기지(62°14'29" S; 58°44'18" W) 부근에서 자라는 현장샘플들을 채집하였다. 샘플은 현장에서 즉시 TRIzol 시약 (Invitrogen, 미국)에 넣어졌고, RNA 추출을 위해 1시간 안에 기지 내 실험실로 옮겨졌으며 이 샘플을 ANT 시료라 지정하였다. 기온과 광량은 온도 및 PAR 센서와 연결된 데이터 로거 CR800 (Campbell Scientific, 미국)으로 측정되었다. 1월중 바톤 반도의 미소 기후 조건은 그림 1a와 같다. 현장의 공기 온도는 토양 표면 위 5 cm에 설치된 온도 센서에서 측정되었다. 월 평균기온은 약 2.7 °C, 최고 기온은 14°C이었다. 최대 PAR 강도는 맑은 날 ca. 1,845 $\mu\text{mol m}^{-1} \text{s}^{-1}$ (2013년 1월 19일)부터 흐린 날 ca. 629 $\mu\text{mol m}^{-1} \text{s}^{-1}$ (2013년 1월 15일)까지 다양했다. 광도인 PAR 값은 자정부터 오전 04시까지 0에 가까웠다. 현지 시간 기준 12:00-14:00 사이에 ANT 샘플을 채취했다. 한편, Barton 반도의 일부 식물은 한국 극지 연구소의 실험실로 옮겨졌으며 장일조건 (20 : 4 h, 150 $\mu\text{mol m}^{-1} \text{s}^{-1}$)의 16 °C 기후 챔버에서 2 % sucrose 함유된 0.5x Murasige Skoog (MS) 배지에서 재배되었다. 해당조건에서 3 주 동안 식물을 재배한 후, 비슷한 발달 단계에서 엽경이 1.5 ~ 2 cm 인 잎을 수집했다. 식물 하부의 오래된 잎을 제거하고 새롭게 생긴 상부의 잎을 RNA 추출을 위해 샘플링하였으며, 이를 "LAB" 샘플이라 지정하였다. 저온 스트레스 처리를 위해, 16°C에서 자란 식물을 2°C로 옮기고 수확할 때까지 배양하였다. 염 스트레스 처리를 위해, 식물을 0.5 x MS 액체배지로 옮기고 7일 동안 성장시킨 다음 150 mM NaCl이 보충된 0.5 x MS 액체배지로 옮겼다. 건조 스트레스 처리를 위해, 식물을 3mm 여과지로 옮기고 30 분 동안 클린벤치에서 건조시켰다. 이어서, 여과지 상의 식물을 다시 빈생장 박스로 옮기고 샘플링 시점까지 키웠다. 각 조건 하에서 RNA 추출을 위해 0, 1 및 7 일 후에 x3 반복 샘플로 샘플링하였다. 모든 실험실 배양 샘플은 ZT8에 샘

플링하였다.

나. 잎길이 측정

온도에 따른 형태학적 차이를 비교하기 위해, 16°C에서 자란 잎의 길이가 1.5 ~ 2cm 인 식물을 2°C와 8°C의 생장 챔버로 옮기고 4주 동안 배양한 후 잎의 길이의 변화를 측정했다. 개별 잎의 길이는 ImageJ 프로그램(<https://imagej.nih.gov/ij/>)에 의해 측정되었다. 통계 분석은 t-검정 ($p < 0.05$)으로 수행되었다.

다. RNA 추출 및 라이브러리 제작

RNA를 TRIzol 시약을 사용하여 추출하고 DNase I (QIAGEN, 독일)로 처리하여 contaminated DNA를 제거한 다음, 제조업체의 프로토콜에 따라 RNeasy Mini kit (QIAGEN, 독일)를 사용하여 정제하였다. 각 샘플은 3 반복수로 진행되었다. RNA 무결성 및 농도는 Bioanalyzer (RNA 6) (Agilent Technologies, 독일) 및 Qubit®RNA kit (Thermo Fisher Scientific, 미국)를 사용하여 결정되었다. RNA-Seq 라이브러리를 구축하기 위해, 1.5µg의 RNA를 TruSeq RNA kit v2 (Illumina, 미국)에 적용하여 라이브러리를 제작하였다. 라이브러리는 Illumina 권장 방법에 따라 만들어졌으며, 제작 후 Bioanalyzer 및 라이브러리 qPCR 정량화 방법을 사용하여 정량화되었다. 정량 후, 이들은 동일한 비율로 혼합하여 Illumina HiSeq Rapid SBS 키트 v2의 단일 플로우 셀(2 × 100개 실행)에 로딩하였다. 시퀀싱은 HiSeq 2500 플랫폼(Illumina)에서 수행되었으며, 총 16.2 Gb (160M paired-end reads)의 시퀀싱 데이터가 생성되었다(Q30 > 93 %).

라. De novo assembly, annotation 및 DEG analysis

De novo assembly는 CLC Genomics Workbench v7.5 소프트웨어(QIAGEN, 독일)를 사용하여 수행되었다. 어댑터 시퀀스를 트리밍한 후, 낮은 품질의 시퀀스 ($Q < 0.001$, ambiguity < 2bps)와 너무 길이가 짧거나 (50 bps 이하) 및 중복 서열을 제거하는 방식으로 리드 필터링을 수행하였다. 생성된 관독 값은 다음과 같은 파라미터 (wordsize = 20 및 bubble size = 50, length > 200 bps)로 어셈블된 후, 이들 contig 들은 CD-HIT을 사용하여 클러스터링을 진행하였다. 총 95,010 개의 assembled contig를 query로 NR 데이터베이스에 대한 BLASTX 검색을 수행하였다(cutoff value E-value < 1×10^{-10}). Gene ontology mapping 및 annotation은 Blast2GO를 사용하여 수행되었다(cutoff value E-value < 1×10^{-10}). GO enrichment 분석은 Fisher's exact test (cutoff FDR < 0.05)와 함께 AgriGO 오픈리소스를 사용하여 수행되었다. full-length로 추정되는 cDNA 서열은 웹 기반 ORF 예측 도구인 Full-lengther를 사용하여 예측되었다. 중간 ortholog는 reciprocal best-hit (RBH)

분석방법을 택했으며, Arabidopsis의 단백질서열에 대한 상호 blastP 검색에 soft masking 및 Smith-Waterman 정렬 옵션(-soft_masking true -use_sw_tback)을 적용하여 가장 낮은 E-value 및 query coverage >50, protein identity >50을 갖는 서열을 ortholog로 간주하였다. KEGG 분석은 KEGG (Kyoto Encyclopedia of Genes and Genomes) 데이터베이스의 각 pathway에 blast 검색을 통해 mapping되었다. KEGG enrichment 분석은 KOBAS web server에 수행되었다. 각 contig의 발현량은 FPKM 으로 결정하였으며, 서열상의 발현차이는 two-group 비교 방법을 사용하였으며 통계적 유의성은 Baggerley's tests 및 t-tests 로 검정되었다.(corrected p-value of false discovery rate <0.05 and difference ≠ 0).

마. Quantitative PCR 분석

cDNA는 superscript III (Invitrogen)를 사용하여 합성되었으며 유전자별 프라이머는 표 3에 표시되었다. Internal control 유전자 선별을 위하여 LAB/ANT 전사체 결과를 이용하여 다음의 7개의 유전자를 선별하였다: 18S rRNA (contig32901), UBC28, ubiquitin-conjugating enzyme E2 28-like (contig3602), RPB6A, DNA-directed RNA polymerase subunit (contig9755), TIM, triosephosphate chloroplast-like (contig19814), CHC, clathrin heavy chain (contig6535), GLX2-4, lactoylglutathione chloroplast-like (contig18505), RPL3, 50S ribosomal protein L3 (contig39479). 이들 유전자들의 적합도를 검정하기 위하여 RefFinder (<http://leonxie.esy.es/RefFinder/>) 방법을 사용하였으며, 증폭 효율이 95 - 105% ($R^2 > 0.97$) 인 경우 좋은 control 로 간주되었다. 그 결과, TIM > 18S > GLX2-4 > UBC28 > RPL3 > RPB6A > CHC 순으로 적합도가 나타났으며, 한편, geNorm 알고리즘은 TIM 과 18S 유전자의 조합 사용을 추천하였다. 이들 두 예측 프로그램의 결과를 바탕으로 우리는 18S와 TIM을 control로 사용하였다. qPCR은 SYBR® Premix Ex Taq™ DNA polymerase (Takara, 일본)를 사용하였으며 실험에 사용된 프라이머 정보는 표3에 기재하였다.

3. 결과

가. 남극개미자리 ANT 및 LAB 샘플의 형태적 비교

남극 *C. quitensis* 남극개미자리 식물은 지리적 분포 및 미소기후의 차이에 의해 형태적, 생리적 반응의 차이를 보인다. 극한지식물의 필드 전사체 분석은 주어진 환경 스트레스에 대해 그들이 어떤 방식으로 유전자 조절을 하는지에 대한 정보를 제공한다. 따라서 우리는 극지환경 스트레스 저항 및 표현형 가소성과 관련된 분자유전학적 기작을 규명하기 위하여 남극반도의 극한 환경에서 자란 식물(ANT, 2013년 바톤반도에서 채집)과 실험실 챔버에서 배양된 식물(LAB)의 전사체 비교 분석을 수

행하였다. 챔버의 온도는 남극개미자리의 최적광합성 온도 14~18 °C 사이인 16 °로 설정되었다. 잎의 길이는 생장 기간 동안 관찰되었는데, LAB 샘플은 잎이 길고 하배측이 상대적으로 길었으며, ANT 샘플은 작고 두꺼운 잎으로 구성되어 잎간의 밀도가 훨씬 조밀했다(그림 1b - d).

나. 남극개미자리 전사체 시퀀싱, 어셈블리 및 기능 분류

전체 160,264,674 reads 가 Illumina HiSeq. 2500 시퀀싱을 통해 생성되었다. Trimming 후, 143,972,486 high-quality reads ($Q > 30$, 13.7 Gb)를 CLC Genome Workbench를 이용하여, 101,690 contigs (averaging 764 bp, N50 value=1,058 bp)로 어셈블리를 하였으며, 이후 CD-HIT-EST를 이용하여, 95,010 contig를 제작하였다.

이들에 대해 GenBank nr 또는 UniProt 데이터베이스에 대해 BLASTX를 수행한 결과, nr 기준, 47,070 contigs가 최소 하나이상의 hit를 가지며 (E-value of 1×10^{-5}), 36,387 contig 가 맵핑되었다(표1). 또한 UniProt 데이터베이스 기준, 45,037 유전자 (47.4%)가 기존의 단백질과 유사성을 보였으며, 이중 10,456 유전자(11%)가 “complete” 로 예측되었으며, 11,837 유전자 (12.5%)를 “new“, 38,116 유전자 (40.1%)를 “unknown“으로 예측되었다(표1). Blast된 서열들의 유사 서열에 대한 종 분포를 확인했을 때 *A. thaliana*, *Vitis vinifera* 가 각각 42%, 11%로 확인되었다.

남극개미자리의 예측된 유전자들의 기능을 분류하기 위해 각 contig에 GO terms을 부여한 결과, 26,346 contigs 가 하나이상의 GO term이 부여되었고, 이들은 PlantGO slim terms에 의해 분류되었다(그림 2). 또한, 유전자의 기능을 유추하기 위해 KEGG databases (organism code: ath, E-value $< 10^{-10}$) 맵핑한 결과, 11,950 contig 가 5개의 기능카테고리로 분류되었다 (그림 3).

다. Simple-sequence repeats 검색

Simple-sequence repeats (SSRs) 은 집단유전학에서 매우 잘 확립된 마커이다. 앞서 기술했듯이, 남극개미자리는 지리적 분포에 따른 생태적 정보를 파악하기 위해서는 이들의 유전적 다형성을 확인하는 것이 매우 필요하다. 따라서 이들의 유전적 다형성을 확인할 수 있는 분자 마커를 개발하는 것은 매우 의미가 있다. 우리는 이를 위해 EST-SSR을 검색하였다. 그 결과, 전체 7,749 contig에서 1~6 bp 차이를 보이는 서열부위를 찾았다. 이들 중 대략 60.5%가 mononucleotides, 28.3%가 trinucleotides, 9.3%가 dinucleotides, 나머지 ~2%는 tetra-, penta-, hexanucleotides였다. 가장 빈번히 나타나는 dinucleotide SSR는 AG/CT (382/4.43%)였으며, trinucleotide SSRs의 경우, AAC/GTT, AAT/ATT, ACC/GGT, AAG/CTT, ATC/ATG는 4.2 -

4.9%의 빈도를 보였다 (표 2).

라. ANT 및 LAB 샘플에서 DEG (differentially expressed genes) 분석

ANT 및 LAB 샘플의 read를 95,010 contigs에 re-mapping하였고, 각각의 contig에 mapping 된 read 수가 계산되었다. Beta-binomial test를 통해, 3,902 transcripts가 differentially expressed genes (DEGs)로 확인되었는데, 이중 2,127 transcripts가 ANT 샘플에서 up-regulated, 1,775 transcripts가 down-regulated 된 것을 확인하였다 (cutoff of FDR corrected p-value < 0.05 (Supplementary 표s S5 and S6). Comparative GO enrichment 분석을 3,902 DEGs에 실시한 결과, 'response to stress (GO: 0006950)' 및 'photosynthesis (GO: 0015979)' 카테고리 DEG 그룹에서 상대적으로 강화되어 나타나는 것을 보였다. 이 중, 'response to stress', 'response to (abiotic, biotic or external) stimulus', 'extracellular region'는 up-regulated 및 down-regulated 유전자군에서 강화되었다. 한편, 'secondary metabolic process', 'response to endogenous stimulus', 'plasma membrane', 'vacuole', 'ribosome' and 'nucleolus' 는 up-regulated 유전자군에서 강화되었으며, 'photosynthesis', 'carbohydrate metabolic process', 'cellular homeostasis', 'plastid' 는 down-regulated 유전자군에서 강화되어 나타났다(그림 4)

마. 대사경로 비교분석

ANT 샘플에서 변화된 대사경로를 확인하기 위해 DEG를 대상으로 KEGG enrichment 분석을 수행하였다 (FDR < 0.01, Fisher's exact test). 그 중, 'ribosome', 'phenylpropanoid biosynthesis', 'biosynthesis of secondary metabolites', 'oxidative phosphorylation', 'flavonoid biosynthesis', 'glutathione metabolism', 등이 ANT-upregulated 유전자에서 강화된 경로로 확인되었다(그림 5).

DEG에 대한 KEGG 와 GO enrichment 분석은 다양한 경로에 관련된 유전자들의 차등 발현을 통하여, 주어진 환경에 대한 세포대사능력이 달라진다는 것을 암시한다. 특히 'photosynthesis' 경로와 관련한 많은 유전자들이 ANT 샘플에서 감소하나, 광보호와 관련된 PGR5 (PROTON GRADIENT REGULATION 5, Contig18861) 유전자가 ANT 샘플에서 증가한 점은 주목할 만하다.

또한, 'phenylpropanoid biosynthesis' (KEGG ID: map00940) 경로 유전자들이 ANT-up-regulated gene에서 현저히 많은 비율로 존재하였다(p-value: 1.7×10^{-19}), (그림 4, 5). Phenylpropanoids들은 페닐알라닌으로부터 유도되는 이차대사산물들이며 이들은 식물세포에서 구조형성물질 및 신호전달물질로 매우 중요한 역할을 한다. 'Phenylpropanoid biosynthesis' 의 주요하위경로는 'monolignol biosynthesis'

(KEGG ID: M0039) 와 'flavanone biosynthesis' (KEGG ID: M00137)로서 각각 페닐알라닌을 리닌과 narigenin으로 변환시키는 경로이다. 이들 경로에 속하는 유전자들은 Phenylalanine ammonia-lyase (PAL; Contig20542), cinnamic acid 4-hydroxylase (C4H; Contig49114, Contig70533) 와 4-hydroxycinnamoyl-CoA ligase (4CL; Contig3499), hydroxycinnamoyl-CoA shikimate hydroxycinnamoyl transferase (HCT; Contig13431, Contig29449, Contig 40563), hydroxycinnamoyl-CoA reductase (CCR; Contig33804), cinnamyl alcohol dehydrogenase (CAD; Contig24021), chalcone synthase (CHS; Contig2685), chalcone isomerase (CHI; Contig16136) 등이었다(그림 6).

바. 식물핵심스트레스반응 유전자의 발현

식물체에 비생물적 스트레스 환경을 처리하면 스트레스 반응 초기에 공통적으로 일련의 유전자들이 유도되는데, 이를 식물 코어 환경스트레스 반응(Plant Core Environmental Stress Response)이라 정의하였다. 이러한 현상은 모델식물인 애기장대에서 잘 연구되어 있으며, 다른 식물에서도 공통적으로 나타나는 것으로 보아 환경스트레스에 대한 보편적 식물생리반응으로 여겨져왔다. 우리는 남극환경에서 자라는 식물들도 극한 환경을 견디내기 위하여 이반응과 관련된 유전자(PCESR genes)들이 높은 수준의 발현을 보일 것이라 가정하고, 애기장대와 남극개미자리의 ortholog 유전자 중 이들 유전자들의 발현양상을 확인하였다. 그 결과, 8530개의 RBH ortholog 중, 35개의 PCESR 유전자가 있었고 이중 40%인 17개의 유전자가 ANT 샘플에서 발현량이 현저히 증가 ($p < 0.05$) 되어 있는 것을 확인하였다.

우리는 이들 유전자들이 저온, 고염, 건조와 같은 비생물학적 스트레스 요인에 어떻게 반응하는지를 확인하기 위해, 각 스트레스를 24시간 처리 후 처리 전후의 발현양상을 quantitative RT-PCR (qPCR)을 통해 확인하였다. 그중, 14/17 유전자가 ANT 샘플에서 높은 발현을 보였다. 이들 유전자는 다음과 같은 다양한 생물학적 기능에 예측되는 유전자들을 포함하였다(그림 7): APETALATA2 (AP2)/ethylene response factor (ERF) protein (Contig14494), NAC27-like (Contig38763), MYB44-like protein (Contig16997); C2H2 type zinc finger proteins ZAT10 (Contig22074), ZAT12 (Contig28814); CCR4-associated factor (Contig45609); calcium-binding protein (Contig38763); DUF246-containing protein (Contig7995); sugar transporter 13 (STP13)-like gene (Contig21564); PUB23 E3 ligase family protein (Contig99760); major facilitator superfamily gene (Contig15740); ring finger proteins (Contig34911, Contig20878); polygalacturonase inhibitor protein (Contig42834). 한편, 이들 유전자들은 다양한 환경요인 중 저온에 특화되어 반응하는 양상을 보였는데, 이는 남극환경에서 온도가 스트레스 신호전달

반응에서 가장 중요한 요소임을 암시한다.

사. 남극개미자리의 환경에 따른 형태적 변화

남극개미자리는 대표적인 cushion plant로, 짧은 잎으로 구성된 뾰족한 반구를 형성하는 표현형을 가지며, 고위도 및 고산지대에 서식한다. 16 °C의 생장 챔버에서 필드 식물을 옮겨 재배할 때 남극개미자리의 컴팩트한 구조가 느슨해지고 잎이 길어짐을 확인하였다(그림 1c,d). 최근 보고에 따르면 일부 PHYTOCHROME INTERACTORS(PIFs)는 열형태형성과정(Thermomorphogenesis)과 관련이 있는 것으로 나타났다. 특히 애기장대 PIF4(AT2G43010)의 경우 온도가 상승할수록 유전자 발현이 증가하는 반면, 온도가 감소할수록 유전자 발현이 감소한다고 알려져 있다. 이와 함께, 형질전환식물 연구에서, hypocotyl 및 petiole의 길이는 이 유전자들의 발현량과 역상관관계가 있다는 것이 보고되었으며, 이러한 결과들은 이 유전자들이 온도변화에 의한 잎의 표현형에 참여한다는 것을 암시한다. 이와 관련하여, 우리는 잎이 작은 ANT 샘플에서 PIF4 및 관련 하위 생장 촉진 유전자의 발현이 감소되어 있을 것이라는 가설을 세웠다. 그 가설을 증명하기 위해, 먼저, 우리는 남극개미자리의 PIF 유전자가 필드 혹은 챔버 조건에서 표현형이 변화하는지 확인했다. AtPIF4를 쿼리로 사용한 TBLASTN 검색 결과, bHLH 도메인이 포함된 5개의 contig (Contig8088, Contig41295, Contig35032, Contig6363, Contig64547)가 검색되었다 (E-value <10⁻¹⁰). 예상대로, 이들의 발현 수준은 ANT 샘플에서 유의미한 감소를 보였다(FDR corrected p-value <0.05). PIF4는 HFR, IAA29, IAA19, ATHB, SAUR23 등의 성장촉진 유전자의 표현을 활성화한다고 알려져 있다. 따라서 우리는 PIF4의 타겟으로 알려진 이들 유전자의 ortholog의 발현이 환경조건에 따라 달라지는지 확인했다. 그 결과, ATHB2(Contig12609), HFR(Contig41293), SAUR23(Contig10897)의 표현값이 LAB 표본보다 ANT에서 더 낮았다(그림 8).

이러한 결과는 저온생장실험에서 재현할 수 있었다. PIF4와 그 하위단계 유전자인 ATHB2와 HFR의 16 °C 샘플에 비해, 상대적 발현량은 8 °C, 2 °C 조건의 샘플 모두에서 상당히 감소하였으며(그림 8c), 잎 길이는 저온 재배 샘플에서 현저히 줄어든 것으로 나타났다(그림 8d). 또한, 저온조건보다 16 °C 이상에서 뿌리 및 shoot 생장이 증가하는 것이 관찰되었다. 이러한 결과들은, ANT, 즉 현장 식물의 특징적인 표현형은 PIF4와 그 하위단계 유전자의 하향조절과 연관 있음을 시사한다.

4. 논의

온도 변화로 인한 식물의 생장 변화는 많은 연구에서 보고되었다. 북극과 남극 식물의 경우, 표현형 가소성에 대한 몇 가지 보고가 있었다. 저온은 기관의 초기발생구조와 mesophyll의 palisade cell에 영향을 미치는 것으로 나타났다. 남극현장에서 자라는

식물과 저온처리를 한 식물은 모두 다중적 돌출부위와 인접한 접촉면의 면적이 증가된 변형된 엽록체 모양을 가지고 있다. 표현형 가소성은 다른 현화식물 중 하나인 남극좁새풀에서도 뚜렷이 나타난다. 최근 남극개미자리의 OTC를 이용한 온도변화 실험에서 광합성 효율이 잎의 해부학적 특성에 따른 CO₂ 전달효율에 의해 조절된다는 것이 밝혀졌는데, 이는 mesophyll의 CO₂ 전도율과 탄소 동화를 향상시켜 결과적으로 잎 탄소 증가와 식물 성장을 촉진한다. 이것은 온도에 의한 표현형 가소성이 식물의 성장과 적응에 큰 역할을 한다는 것을 암시한다.

본 연구에서 우리는 온도가 높은 생장 챔버에서 자랄 때 잎이 길어지고 잎 사이의 밀도가 느슨해진다는 것을 관찰했다. 따라서, 우리는 온도와 함께 잎의 해부학적 변화를 일으키는 발달조절에 PIF 전사인자가 역할을 한다고 제안한다. PIF는 고온과 관련된 열변형성 생성을 수반하는 메커니즘으로는 잘 알려져 있으나, 저온 조절 역할에서 파생된 생태학적 이익에 대한 논의는 많지 않았다. 본 연구 결과에 의하면 PIF4 전사인자의 발현이 남극샘플에서 낮은 것으로 밝혀졌으며, 하위 타겟으로 알려진 HFR, IAA19, ATHB, SAUR23 등의 유전자 발현도 LAB 표본보다 ANT 표본에서 낮은 것으로 확인되었으며 이는 기존의 연구 결과 및 가설을 지지하였다. 이들 유전자의 발현 차이에 의한 형태변화는 궁극적으로 작은 잎을 만들고, 잎과 잎 사이의 공간이 매우 촘촘하게 만들어 보온을 유지하는 것으로 생각된다.

한편 phenylpropanoid 생합성 경로는 스트레스와 관련된 다양한 종류의 이차대사산물을 생성한다. 이 연구에서 'phenylpropanoid biosynthesis' 경로에 관여하는 거의 모든 유전자는 ANT 시료에서 강하게 상향 조절되었으며, 이는 이 경로에서 생성된 부산 물질이 비생물적 및 생물학적 스트레스로부터 식물 세포를 보호하는 데 중요한 역할을 할 수 있음을 암시한다. *Arabidopsis*, *Zea mays*, *Eucalyptus globulus*, *Lotus japonicus* 등에서의 연구 결과는 PAL, CCR, CAD, CHS 및 CHI와 같은 phenylpropanoid 경로의 주요 효소가 저온, 가뭄, UV 또는 병원체 감염에 의해 강하게 유도되는 것으로 나타났으며 이들의 유전자 발현 패턴은 조직 및 종에 따라 상이하다. 이와 관련하여, ANT 샘플에서 phenylpropanoid biosynthesis 유전자의 발현증가는 저온, 가뭄 및 병원체와 같은 다양한 환경적 요인으로 반응이 축적된 결과일 수 있다. 또한, 이러한 결과는 기존의 OTC 테스트에서 관찰되었던, 'OTC 외부, 즉, 낮은 온도의 개방된 외부환경에서 성장한 식물이 hemicellulose, cellulose, lignin 등의 함량이 OTC 내부의 샘플보다 더 높다'라는 연구 결과와 상통한다.

5. 참고문헌

1. Krasensky, J. & Jonak, C. Drought, salt, and temperature stress-induced metabolic rearrangements and regulatory networks. *Journal of Experimental Botany* 63, 1593-1608, doi:10.1093/jxb/err460 (2012).

2. Thomashow, M. F. PLANT COLD ACCLIMATION: Freezing Tolerance Genes and Regulatory Mechanisms. *Annual Review of Plant Physiology and Plant Molecular Biology* 50, 571-599, doi:0.1146/annurev.arplant.50.1.571 (1999).
3. Xiong, L., Schumaker, K. S. & Zhu, J.-K. Cell Signaling during Cold, Drought, and Salt Stress. *Plant Cell* 14, S165-S183, doi:10.1105/tpc.000596 (2002).
4. Guy, C. L. Cold Acclimation and Freezing Stress Tolerance: Role of Protein Metabolism. *Annual Review of Plant Physiology and Plant Molecular Biology* 41, 187-223, doi:10.1146/annurev.pp.41.060190.001155 (1990).
5. Schlichting, C. D. The Evolution of Phenotypic Plasticity in Plants. *Annual Review of Ecology and Systematics* 17, 667-693, doi:10.1146/annurev.es.17.110186.003315 (1986).
6. Gratani, L. Plant Phenotypic Plasticity in Response to Environmental Factors. *Advances in Botany 2014*, 17, doi:10.1155/2014/208747 (2014).
7. Aubert, S., Boucher, F., Lavergne, S., Renaud, J. & Choler, P. 1914 - 2014: A revised worldwide catalogue of cushion plants 100 years after Hauri and Schröter. *Alpine Botany* 124, 59-70, doi:10.1007/s00035-014-0127-x (2014).
8. Antonsson, H., Björk, R. G. & Molau, U. Nurse plant effect of the cushion plant *Silene acaulis* (L.) Jacq. in an alpine environment in the subarctic Scandes, Sweden. *Plant Ecology & Diversity* 2, 17-25, doi:10.1080/17550870902926504 (2009).
9. Molenda, O., Reid, A. & Lortie, C. J. The Alpine Cushion Plant *Silene acaulis* as Foundation Species: A Bug's-Eye View to Facilitation and Microclimate. *PLOS ONE* 7, e37223, doi:10.1371/journal.pone.0037223 (2012).
10. Greene, D. & Holtom, A. Studies in *Colobanthus quitensis* (Kunth) Bartl. and *Deschampsia antarctica* Desv. III. Distribution habitats and performance in Antarctic Botanical Zone. *British Antarctic Survey Bulletin* 26, 1-29 (1971).
11. Moore, D. Studies in *Colobanthus quitensis* (Kunth) Bartl. and *Deschampsia antarctica* Desv. II. Taxonomy, distribution and relationships. *British Antarctic Survey Bulletin* 23, 633 - 680 (1970).
12. Smith, R. & Poncet, S. *Deschampsia antarctica* and *Colobanthus quitensis* in the

Terra Firma Islands, British Antarctic Survey Bulletin 74, 31-35 (1987).

13. Sáez, P. L. et al. Photosynthetic limitations in two Antarctic vascular plants: importance of leaf anatomical traits and Rubisco kinetic parameters. *Journal of Experimental Botany* 68, 2871-2883, doi:10.1093/jxb/erx148 (2017).
14. Bascunan-Godoy, L., Garcia-Plazaola, J. I., Bravo, L. A. & Corcuera, L. J. Leaf functional and micro-morphological photoprotective attributes in two ecotypes of *Colobanthus quitensis* from the Andes and Maritime Antarctic. *Polar Biology* 33, 885-896, doi:10.1007/s00300-010-0765-4 (2010).
15. Bascunan-Godoy, L. et al. Cold-acclimation limits low temperature induced photoinhibition by promoting a higher photochemical quantum yield and a more effective PSII restoration in darkness in the Antarctic rather than the Andean ecotype of *Colobanthus quitensis* Kunt Bartl (Cariophyllaceae). *BMC Plant Biology* 12, 114 (2012).
16. Bravo, L. et al. Effect of cold acclimation on the photosynthetic performance of two ecotypes of *Colobanthus quitensis* (Kunth) Bartl. *Journal of Experimental Botany* 58, 3581 - 3590, doi:10.1093/jxb/erm206 (2007).
17. Xiong, F. S., Ruhland, C. T. & Day, T. A. Photosynthetic temperature response of the Antarctic vascular plants *Colobanthus quitensis* and *Deschampsia antarctica*. *Physiologia Plantarum* 106, 276-286, doi:10.1034/j.1399-3054.1999.106304.x (1999).
18. Giełwanowska, I., Pastorczyk, M., Lisowska, M., Węgrzyn, M. & Górecki, R. in *Polish Polar Research* Vol. 35 627 (2014).
19. Lütz, C. Cell physiology of plants growing in cold environments. *Protoplasma* 244, 53-73, doi:10.1007/s00709-010-0161-5 (2010).
20. Giełwanowska, I. et al. in *Polish Polar Research* Vol. 32 139 (2011).
21. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* 28, 27-30, doi:10.1093/nar/28.1.27 (2000).
22. Ellis, J. R. & Burke, J. M. EST-SSRs as a resource for population genetic analyses. *Heredity* 99, 125-132 (2007).

23. Acuña-Rodríguez, I. S., Oses, R., Cortés-Vasquez, J., Torres-Díaz, C. & Molina-Montenegro, M. A. Genetic diversity of *Colobanthus quitensis* across the Drake Passage. *Plant Genetic Resources* 12, 147-150, doi:10.1017/S1479262113000270 (2014).
24. DalCorso, G. et al. A Complex Containing PGRL1 and PGR5 Is Involved in the Switch between Linear and Cyclic Electron Flow in Arabidopsis. *Cell* 132, 273-285, doi:10.1016/j.cell.2007.12.028 (2008).
25. Munekage, Y. et al. PGR5 is involved in cyclic electron flow around photosystem I and is essential for photoprotection in Arabidopsis. *Cell* 110, 361-371, doi:S009286740200867X [pii] (2002).
26. Vogt, T. Phenylpropanoid Biosynthesis. *Molecular Plant* 3, 2-20, doi: <http://dx.doi.org/10.1093/mp/ssp106> (2010).
27. Hahn, A. et al. Plant core environmental stress response genes are systemically coordinated during abiotic stresses. *Int J Mol Sci* 14, 7617-7641, doi:10.3390/ijms14047617 (2013).
28. Kilian, J. et al. The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *Plant Journal* 50, 347-363, doi:10.1111/j.1365-313X.2007.03052.x (2007).
29. Calzadilla, P. I., Maiale, S. J., Ruiz, O. A. & Escaray, F. J. Transcriptome Response Mediated by Cold Stress in *Lotus japonicus*. *Frontiers in Plant Science* 7, doi:10.3389/fpls.2016.00374 (2016).
30. Liu, Z. et al. Temporal transcriptome profiling reveals expression partitioning of homeologous genes contributing to heat and drought acclimation in wheat (*Triticum aestivum* L.). *BMC Plant Biology* 15, 152, doi:10.1186/s12870-015-0511-8 (2015).
31. Zhou, J. et al. Global genome expression analysis of rice in response to drought and high-salinity stresses in shoot, flag leaf, and panicle. *Plant Molecular Biology* 63, 591-608, doi:10.1007/s11103-006-9111-1 (2007).
32. Kumar, S. V. et al. Transcription factor PIF4 controls the thermosensory activation of flowering. *Nature* 484, 242-245, doi:10.1038/nature10928 (2012).

33. Leivar, P. & Monte, E. PIFs: Systems Integrators in Plant Development. *Plant Cell* 26, 56-78 (2014).
34. Leivar, P. et al. The Arabidopsis Phytochrome-Interacting Factor PIF7, Together with PIF3 and PIF4, Regulates Responses to Prolonged Red Light by Modulating phyB Levels. *Plant Cell* 20, 337-352, doi:10.1105/tpc.107.052142 (2008).
35. Zhang, Y. et al. A Quartet of PIF bHLH Factors Provides a Transcriptionally Centered Signaling Hub That Regulates Seedling Morphogenesis through Differential Expression-Patterning of Shared Target Genes in Arabidopsis. *PLOS Genetics* 9, e1003244, doi:10.1371/journal.pgen.1003244 (2013).
36. de Lucas, M. et al. A molecular framework for light and gibberellin control of cell elongation. *Nature* 451, 480-484, doi:10.1038/nature06520 (2008).
37. Champigny, M. J. et al. RNA-Seq effectively monitors gene expression in *Eutrema salsugineum* plants growing in an extreme natural habitat and in controlled growth cabinet conditions. *BMC Genomics* 14, 578, doi:10.1186/1471-2164-14-578 (2013).
38. Mishra, Y. et al. Arabidopsis plants grown in the field and climate chambers significantly differ in leaf morphology and photosystem components. *BMC Plant Biology* 12, 6-6, doi:10.1186/1471-2229-12-6 (2012).
39. Suorsa, M. Cyclic electron flow provides acclimatory plasticity for the photosynthetic machinery under various environmental conditions and developmental stages. *Frontiers in Plant Science* 6, 800, doi:10.3389/fpls.2015.00800 (2015).
40. Ivanov, A. G. et al. Implications of alternative electron sinks in increased resistance of PSII and PSI photochemistry to high light stress in cold-acclimated *Arabidopsis thaliana*. *Photosynth Res* 113, 191-206, doi:10.1007/s11120-012-9769-y (2012).
41. Ifuku, K., Endo, T., Shikanai, T. & Aro, E. M. Structure of the chloroplast NADH dehydrogenase-like complex: nomenclature for nuclear-encoded subunits. *Plant Cell Physiol* 52, 1560-1568, doi:10.1093/pcp/pcr098 (2011).
42. Ishikawa, N. et al. NDH-Mediated Cyclic Electron Flow Around Photosystem I is

- Crucial for C4 Photosynthesis. *Plant Cell Physiol* 57, 2020-2028, doi:10.1093/pcp/pcw127 (2016).
43. Munekage, Y. et al. Cyclic electron flow around photosystem I is essential for photosynthesis. *Nature* 429, 579-582, doi:10.1038/nature02598 (2004).
 44. Lauvergeat, V. et al. Two cinnamoyl-CoA reductase (CCR) genes from *Arabidopsis thaliana* are differentially expressed during development and in response to infection with pathogenic bacteria. *Phytochemistry* 57, 1187-1195, doi: http://dx.doi.org/10.1016/S0031-9422(01)00053-X (2001).
 45. Fornale, S. et al. ZmMYB31 directly represses maize lignin genes and redirects the phenylpropanoid metabolic flux. *Plant J* 64, 633-644, doi:10.1111/j.1365-313X.2010.04363.x (2010).
 46. Zhao, S. et al. Identification of phenylpropanoid biosynthetic genes and phenylpropanoid accumulation by transcriptome analysis of *Lycium chinense*. *BMC Genomics* 14, 802, doi:10.1186/1471-2164-14-802 (2013).
 47. Li, J., Ou-Lee, T. M., Raba, R., Amundson, R. G. & Last, R. L. Arabidopsis Flavonoid Mutants Are Hypersensitive to UV-B Irradiation. *Plant Cell* 5, 171-179, doi:10.1105/tpc.5.2.171 (1993).
 48. Jean-Marc Domon, L. B., Sébastien Acket, Elodie Caudeville, Stéphanie Arnoult, Hélène Zub, Françoise Gillet, Isabelle Lejeune-Hénaut, Maryse Brancourt-Hulmel, Jérôme Pelloux, Catherine Rayon. Cell wall compositional modifications of *Miscanthus* ecotypes in response to cold acclimation. *Phytochemistry* 85, 51-61, doi:http://dx.doi.org/10.1016/j.phytochem.2012.09.001 (2013).
 49. Tomotaka Shinya, K. H., Shinichi Onogi, Akiyoshi Kawaoka. Transcript Level Analysis of Lignin and Flavonoid Biosynthesis Related Genes in *Eucalyptus globulus*. *American Journal of Plant Sciences* 5, 2764-2772, doi:10.4236/ajps.2014.518293 (2014).
 50. Apel, K. & Hirt, H. REACTIVE OXYGEN SPECIES: Metabolism, Oxidative Stress, and Signal Transduction. *Annual Review of Plant Biology* 55, 373-399, doi:10.1146/annurev.arplant.55.031903.141701 (2004).
 51. Davletova, S., Schlauch, K., Coutu, J. & Mittler, R. The Zinc-Finger Protein Zat12 Plays a Central Role in Reactive Oxygen and Abiotic Stress Signaling in

- Arabidopsis. *Plant Physiology* 139, 847-856, doi:10.1104/pp.105.068254 (2005).
52. Mittler, R. et al. Gain- and loss-of-function mutations in *Zat10* enhance the tolerance of plants to abiotic stress. *FEBS Letters* 580, 6537-6542, doi:10.1016/j.febslet.2006.11.002 (2006).
 53. Yao, C. et al. Gene encoding polygalacturonase inhibitor in apple fruit is developmentally regulated and activated by wounding and fungal infection. *Plant Molecular Biology* 39, 1231-1241, doi:10.1023/a:1006155723059 (1999).
 54. John, U. P. et al. Ice recrystallization inhibition proteins (IRIPs) and freeze tolerance in the cryophilic Antarctic hair grass *Deschampsia antarctica* E. Desv. *Plant, Cell & Environment* 32, 336-348, doi:10.1111/j.1365-3040.2009.01925.x (2009).
 55. Cho, S. K., Ryu, M. Y., Song, C., Kwak, J. M. & Kim, W. T. Arabidopsis PUB22 and PUB23 Are Homologous U-Box E3 Ubiquitin Ligases That Play Combinatory Roles in Response to Drought Stress. *Plant Cell* 20, 1899-1914, doi:10.1105/tpc.108.060699 (2008).
 56. Giełwanowska, I., Pastorczyk, M., Kellmann-Sopyła, W., Górniak, D. & Górecki, R. J. Morphological and Ultrastructural Changes of Organelles in Leaf Mesophyll Cells of the Arctic and Antarctic Plants of Poaceae Family Under Cold Influence. *Arctic, Antarctic, and Alpine Research* 47, 17-25, doi:10.1657/aaar0014-019 (2015).
 57. Lee, H. et al. Expressed Sequence Tag Analysis of Antarctic Hairgrass *Deschampsia antarctica* from King George Island, Antarctica. *Mol. Cells* 25, 258-264 (2008).
 58. Gorsuch, P. A., Pandey, S. & Atkin, O. K. Temporal heterogeneity of cold acclimation phenotypes in Arabidopsis leaves. *Plant, Cell & Environment* 33, 244-258, doi:10.1111/j.1365-3040.2009.02074.x (2010).
 59. Hurry, V. M. & Huner, N. P. A. Low Growth Temperature Effects a Differential Inhibition of Photosynthesis in Spring and Winter Wheat. *Plant Physiology* 96, 491-497, doi:10.1104/pp.96.2.491 (1991).
 60. Sebastian, N., Erika, H. & Christian, K. Critically low soil temperatures for root growth and root morphology in three alpine plant species. *Alpine Botany* 126,

- 11-21, doi:10.1007/s00035-015-0153-3 (2016).
61. Stefanowska, M., Kuraś, M., Kubacka-Zebalska, M. & Kacperska, A. Low Temperature Affects Pattern of Leaf Growth and Structure of Cell Walls in Winter Oilseed Rape (*Brassica napus* L., var. *oleifera* L.). *Annals of Botany* 84, 313-319, doi:10.1006/anbo.1999.0924 (1999).
62. Quint, M. et al. Molecular and genetic control of plant thermomorphogenesis. *Nature Plants* 2, 15190, doi:10.1038/nplants.2015.190 (2016).
63. Jiang, Y. & Deyholos, M. K. Comprehensive transcriptional profiling of NaCl-stressed *Arabidopsis* roots reveals novel classes of responsive genes. *BMC Plant Biology* 6, 25, doi:10.1186/1471-2229-6-25 (2006).
64. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150-3152, doi:10.1093/bioinformatics/bts565 (2012).
65. Conesa, A. et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, doi:10.1093/bioinformatics/bti610 (2005).
66. Du, Z., Zhou, X., Ling, Y., Zhang, Z. & Su, Z. agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Research* 38, W64-W70, doi:10.1093/nar/gkq310 (2010).
67. Lara, A. et al. in *Innovations in Hybrid Intelligent Systems Vol. 44 Advances in Soft Computing* (eds Emilio Corchado, JuanM Corchado, & Ajith Abraham) Ch. 47, 361-368 (Springer Berlin Heidelberg, 2007).
68. Moreno-Hagelsieb, G. & Latimer, K. Choosing BLAST options for better detection of orthologs as reciprocal best hits. *Bioinformatics* 24, 319-324, doi:10.1093/bioinformatics/btm585 (2008).
69. Xie, C. et al. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Research* 39, W316-W322, doi:10.1093/nar/gkr483 (2011).

표 1. Summary of sequencing, assembly, and annotation

Total reads		
ANT sample		86,457,992
LAB sample		73,806,682
Total		160,264,674
Assembly		
Total Unigenes		95,010
N50 (bp)		1,058
Average size (bp)		764
BLASTX (nr DB)	Counts	Ratio
Without Blast Hits	47,940	0.505
With Blast Hits	47,070	0.495
With Mapping Results	36,387	0.383
Annotated Sequences	26,347	0.277
Total	95,010	1
ORF prediction (Uniprot DB)		
Unigenes with orthologue in DBs	45,037	0.474
Putative New Genes	11,837	0.125
Unknown	38,116	0.401
Total	95,010	1

극지연구소

표 2. Statistics of SSR markers identified from *C. quitensis* transcriptome analysis

SSR mining	
Total number of sequences examined	95,010
Total size of examined sequences (bp)	72,865,140
Total number of identified SSRs	8,619
Number of SSR containing sequences	7,749
Number of sequences containing more than 1 SSR	749
Number of SSRs present in compound formation	246
Frequency of SSR	one per 8.5kb
Distribution to different repeat type classes	
Mononucleotide	5,214 (60.5%)
Dinucleotide	805 (9.3%)
Trinucleotide	2,438 (28.3%)
Tetranucleotide	91 (1.1%)
Pentanucleotide	52 (0.6%)
Hexanucleotide	19 (0.2%)



표 3. Primer sequences used for qPCR analysis

Gene	Contig	Direction	Sequences
Stress response genes			
ZAT10	Contig22074	F	ATGGCACTAGAAGCGTGGAGCTC
		R	CATGAGGAGGCAGAAAAGCGAGATAC
ZAT12	Contig28814	F	ATTAGCGAGGGGAATAGTCACG
		R	CCTCGTCATCAACCTCTTCT
CML37	Contig38763	F	ATTTCTCAAAGAACATGGAAAACGA
		R	CGTTGTTTTTCGTCTATGTATTTCG
EFR1A	Contig14494	F	GATATTTTCGGACCCGCGGGA
		R	GCCGTCTCAAACGTCCCCAA
NAC27	Contig35858	F	AAAAATAACAATAACCTTCCTCCGG
		R	GGGTATTTCTATCTCTTGGTG
MYB44	Contig16997	F	TATCGTCGGAAAATGGATCAAAATC
		R	CGAACTTATCATGAGCCTTGATGA
CCR4-Not component	Contig45609	F	ATGGATACTGAGTTTCTGGTGT
		R	GAAAATCACAACCCATATCAGG
DUF246	Contig7995	F	GACGATAATGGGGAAAGAGATAA
		R	ACAAGATCAACATTCGGATGAAG
STP13 like	Contig21564	F	TTGAAAATAAGATGTTAAGGCCCTCCA
		R	GCAACCTTTTTTGCATCACATACAA
E3 ligase PUB23	Contig99760	F	GGAAATGCCTTGAGTTGTTTAG
		R	CACGGGATATATGTAATGAGCTAA
Major-facilitator superfamily	Contig15740	F	ATCGCGCTAGGATGCAAAGTC
		R	GCTTCAAGTGCCCGATCAAATAA
Ring finger	Contig34911	F	GGACGAGTTAACTCTGAGTCAACT
		R	GGGCCTGACTCATCAGAGTT
Ring finger	Contig20878	F	CATGCTCAACCCGACTCTTC
		R	CGTAGTCGAAAAGTAGGGAAGC
PGIP	Contig42834	F	GCAGAGCTATGCAACCCAAG
		R	GGCAATGTCACCGAAAATA
Internal reference genes			
18S	Contig32901	F	TCAGTTGGTGGAGCGATTTTGTC
		R	GGAAGGCCATAGTCCCTCTA
UBC28	Contig3602	F	AGTGGGCGAAGACATGTTTC
		R	TTGGGGGCTTGAAAGGATAGTC
TIM	Contig19814	F	ATGCTCTGAGTGAAGGTCTTGG
		R	ACAACGTTTTCCAGCTTGG
RPL3	Contig39479	F	TGTGACTCAGGTGAAAACCG
		R	ATTCGCCTTCGTCAAATGCC
RPB6A	Contig9755	F	TTGGAACACGTGCTCTTCAG
		R	AATGGTATCTTGCGCTCACG
GLX2-4	Contig18505	F	GGTTATGGCCCTGAAGATTCTC
		R	ACATCTTCCACTGCAATGCC
CHC	Contig6535	F	AAAGAGTGCTTTGCCGCATG
		R	ACTTTGCCTGTGTACTCACG
PIF4 and growth promoting genes			
PIF4	Contig8088	F	TGGGAATTGGTTTTGGGCATG
		R	ACCCGCATTTCCTTCAAATG
ATHB2	Contig12609	F	AAGAAGGCGGTGAGAATTTCG
		R	GCCTCAATCCAAGCCGTTTC
HFR	Contig41293	F	ACCAGTCTGAAAGGAGAAGGAG
		R	GCTGCTTTAGTACTCTATCACC
SAUR23	Contig10897	F	AACACTAGTCTCTAGCGCGAAG
		R	AAGCGGAACCTATGCATCTGG

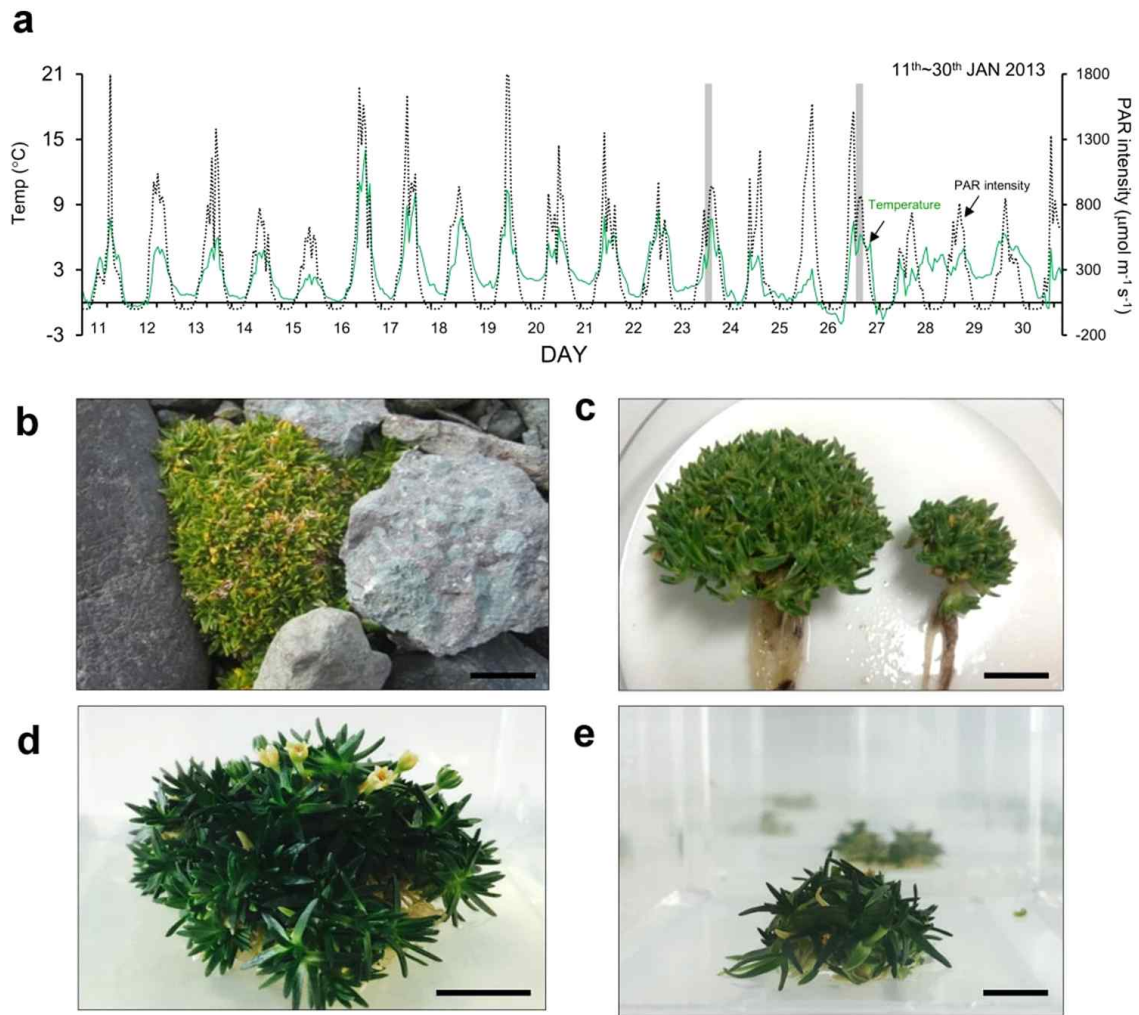


그림 1. Microclimates of the sampling site in Antarctica and morphology of *C. quitensis* plants (a) The temperature (green solid line) and light intensity (black dotted line) of the plant sampling sites had been measured during summer research activities in January 2013. The temperature of the site was measured from a temperature sensor installed 5 cm above the soil surface. During the period, the average temperature was 4.3 °C and 1.8 °C for day and night, respectively. Luminosity was close to zero from midnight to 04:00 AM. For ANT samples, plants were harvested between 11:00 and 13:00 during mid-to-late January (shaded). (b,c) Wild plants from the natural habitat of the Antarctic, and (d,e) a plant grown in a climate chamber. Side-bars indicate 1 cm. The *C. quitensis* found on the coast of the Baton Peninsula are often found around rocks and their overall shape is hemispherical or similar to moss carpets. When young plants with a diameter of 2 cm or less were transferred to a climate chamber maintained at a temperature of 16 °C and grown in nutrient medium, the density of leaves became much looser than those of similar sized Antarctic plants.

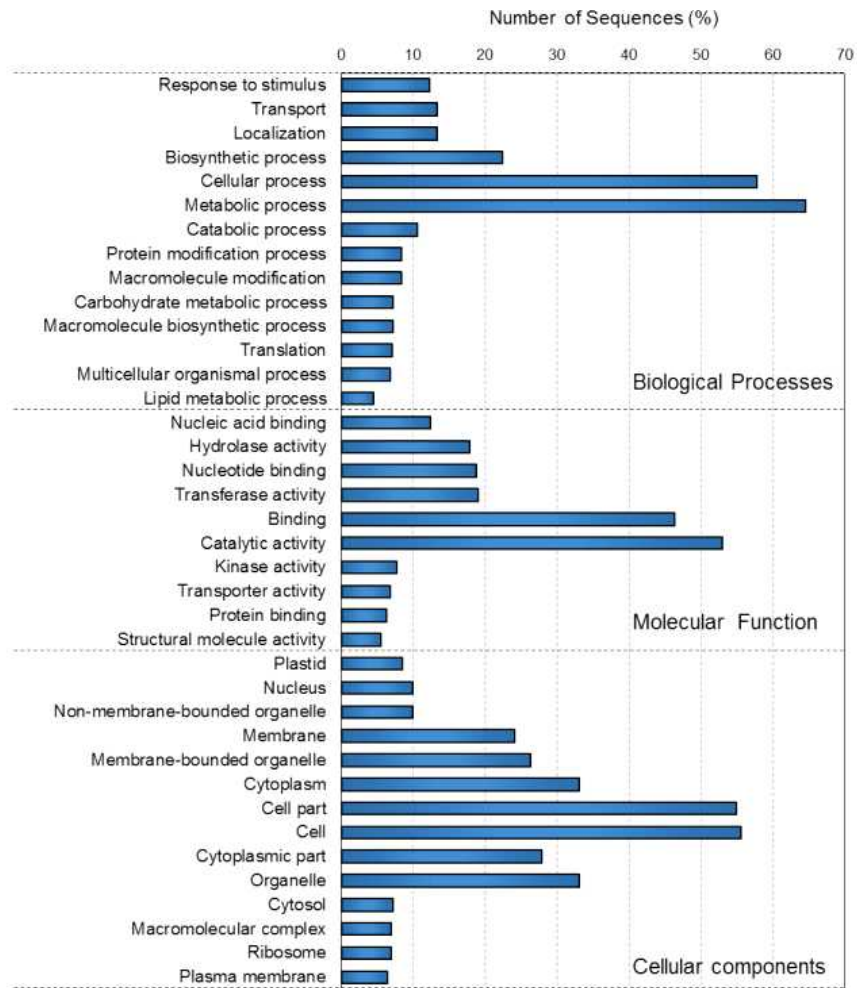


그림 2. Functional GO classification of the *C. quitensis* transcriptome. Sequences with BLASTX matches were assigned GO terms and classified into different functional categories (biological process, molecular function, and cellular component).

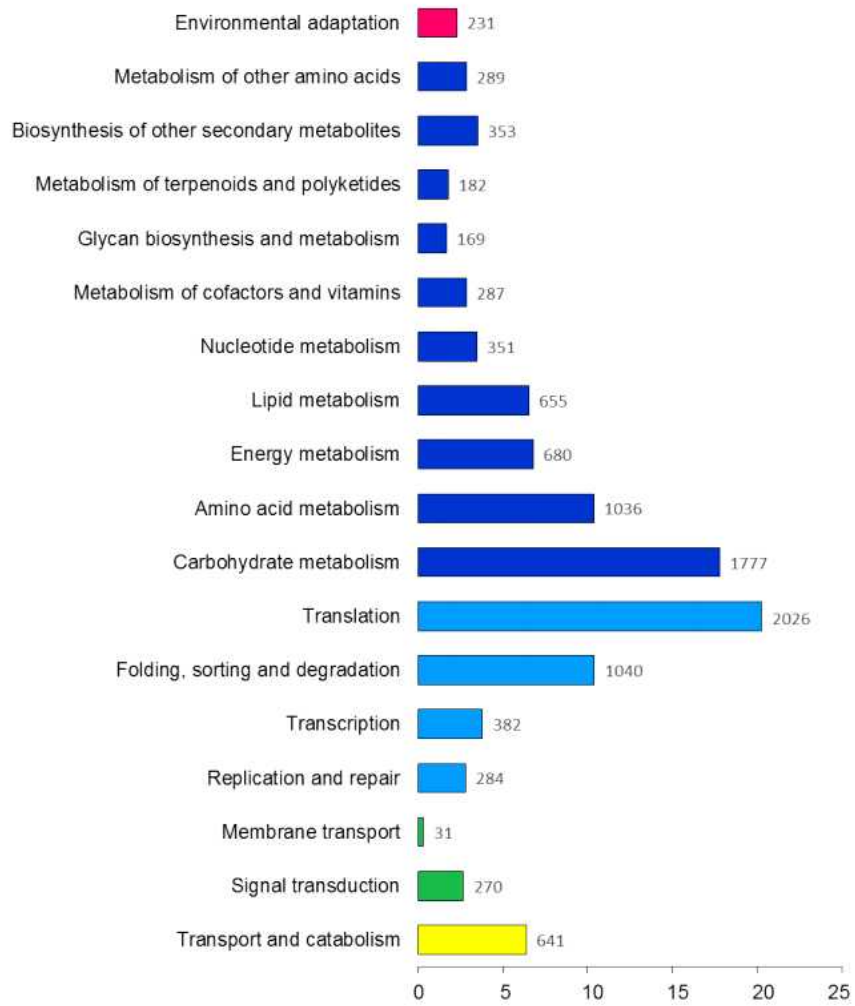


그림 3. Pathway assignment based on KEGG. Pathways were assigned into five categories. OS: Organismal systems, M: Metabolism, GIP: Genetic information processing, EIP: Environmental information processing, CP: Cellular processes

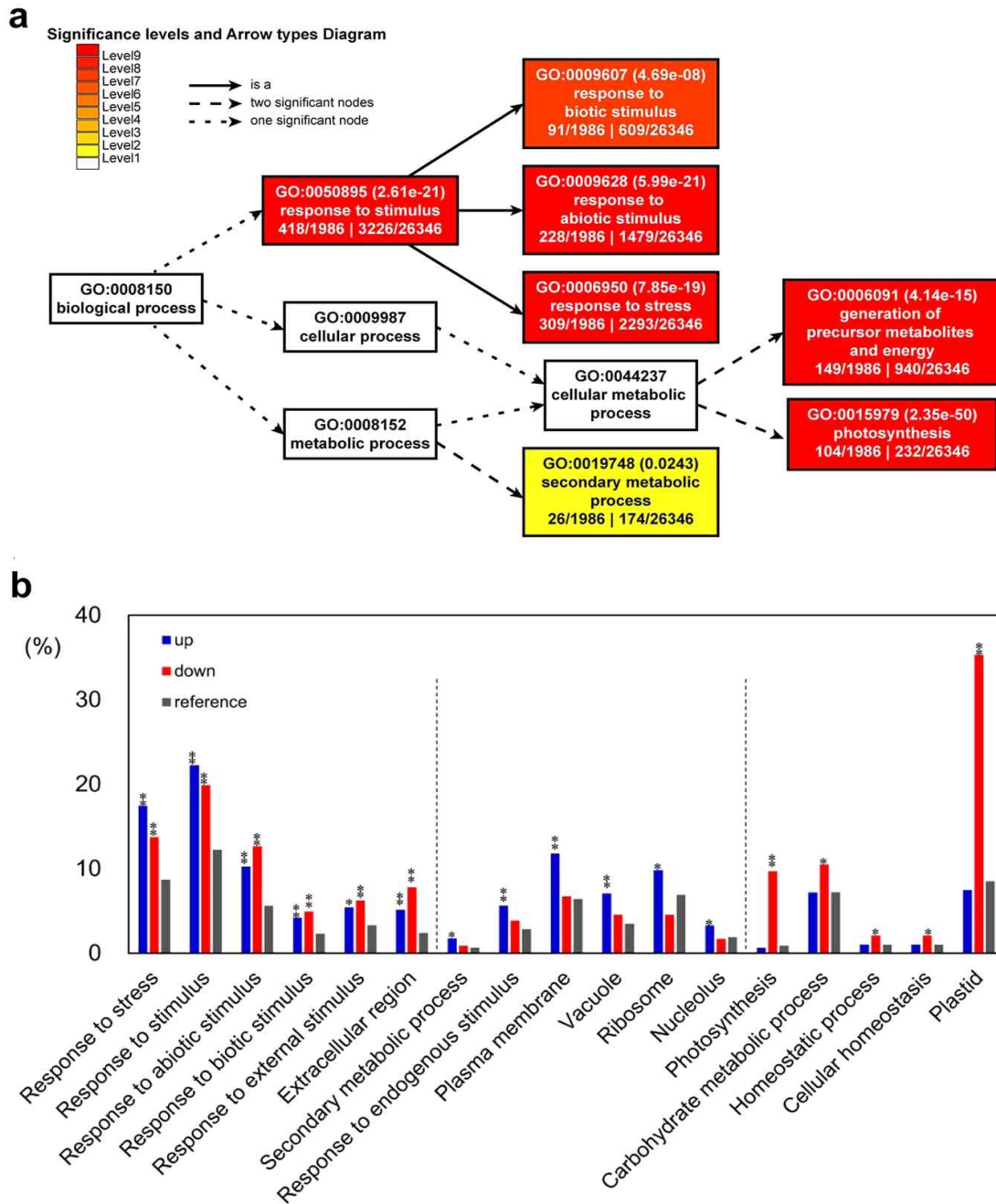


그림 4. Functional classification of DEGs by plant GO slim terms. (a) GO terms of all DEGs with significance ($FDR < 0.05$, Fisher's exact test) are presented with box colors when compared with those of a total of 26,346 GO-annotated contigs. The 그림 was drawn by AgriGO analysis tool65 (b) The percentage of enriched GO terms of upregulated- and downregulated gene groups were compared with those of a total of contigs. The blue, red, and grey bars indicate the percentage of GO terms of the upregulated, downregulated and a total of contigs, respectively. Statistical significance is indicated with asterisks (* $FDR < 0.05$, ** $FDR < 0.001$).

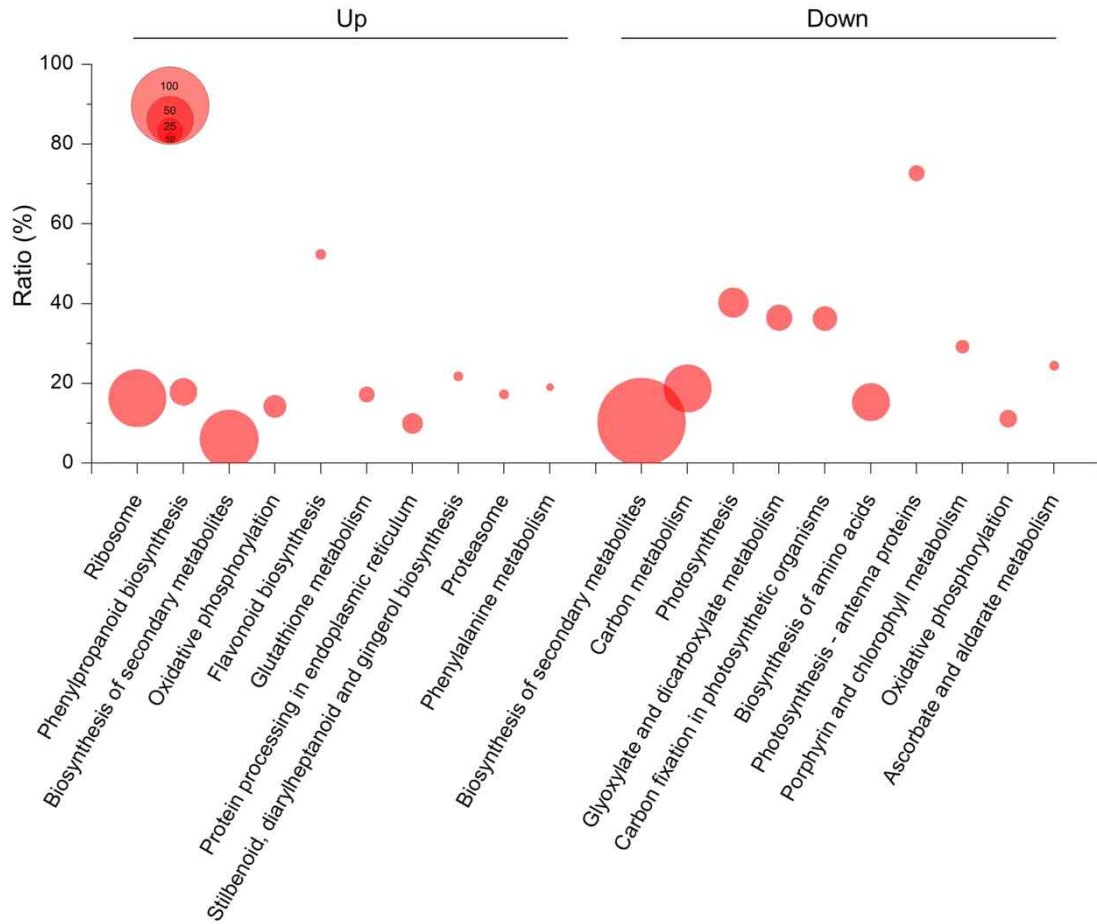


그림 5. A scatterplot of enriched KEGG pathways in upregulated- and downregulated gene groups. The significantly enriched KEGG pathways were indicated as dots (FDR < 0.05, Fisher's exact test) and the dot sizes represent the number of the genes included in each cluster. The horizontal axis represents the ratio of the number of differentially expressed genes and the number of all genes in the pathway.

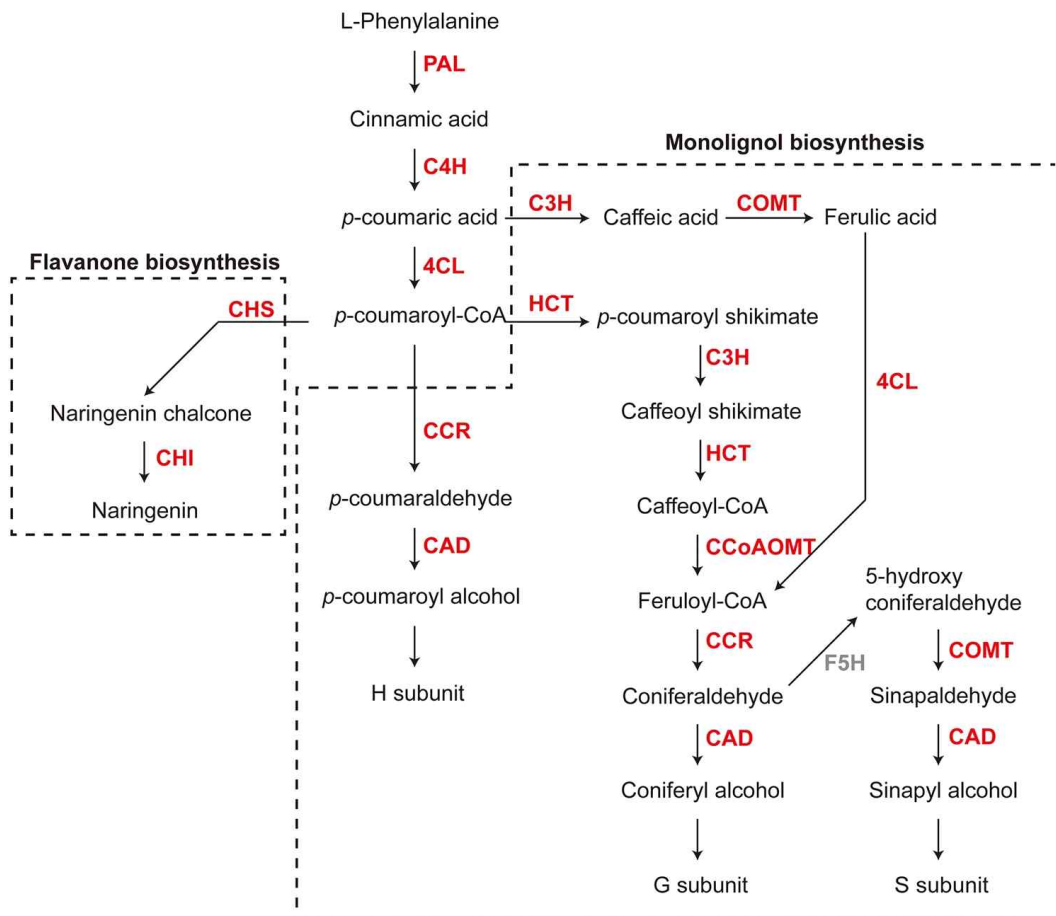


그림 6. The genes of 'phenylpropanoid biosynthesis' pathway were upregulated in ANT samples. The 'phenylpropanoid biosynthesis' pathway consists of the 'monolignol biosynthesis' (KEGG ID: M0039) and the 'flavanone biosynthesis' (KEGG ID: M00137) pathways, which convert phenylalanine to lignin and naringenin, respectively. The major enzymes in these pathways tended to be upregulated in ANT samples, and these are marked in red

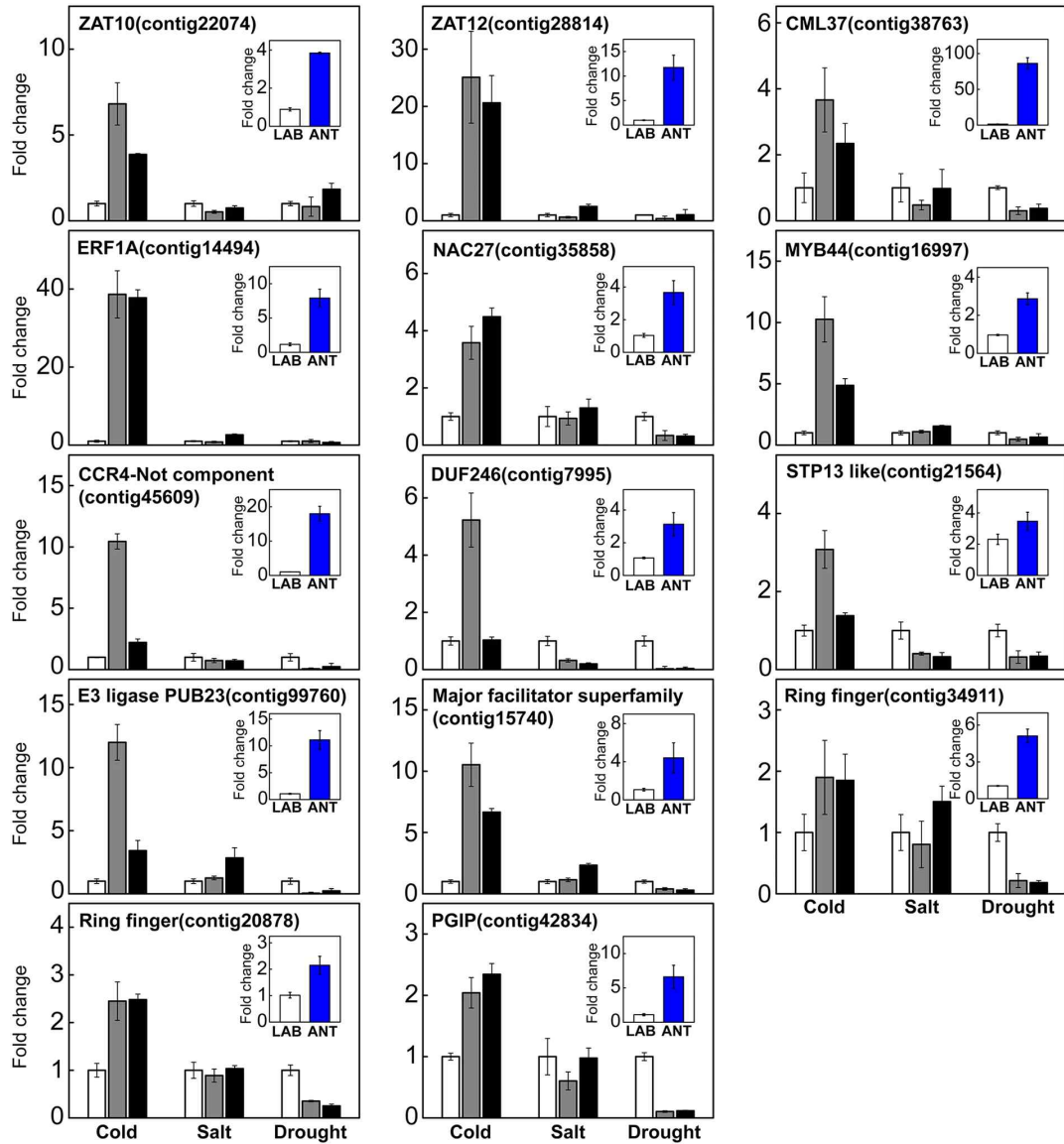


그림 7. qPCR analysis of PCESR - orthologs from *C. quitensis* in LAB- vs. ANT samples and the plants exposed to different abiotic stressors. The main graphs are the results of the abiotic stress treated samples and inset graphs are the results of the LAB- vs. ANT samples. The genes were selected from the 14 PCESR orthologs which have higher expression values in ANT samples from RNA-Seq results. For stress treatment, plants at similar developmental stages were treated with low-temperature (2 °C), high salt (150 mM NaCl) or dehydration for 1 and 7 days. The left vertical axis indicates the relative ratio of transcript abundance of selected genes compared to 18S rRNA and TIM (Contig19814) which were predicted as the best internal controls by reference gene prediction tools^{71,72}. Mean and standard deviation are presented (n = 3) and white, grey and black bars display the 0, 1 and 7 days of stress treatments, respectively.

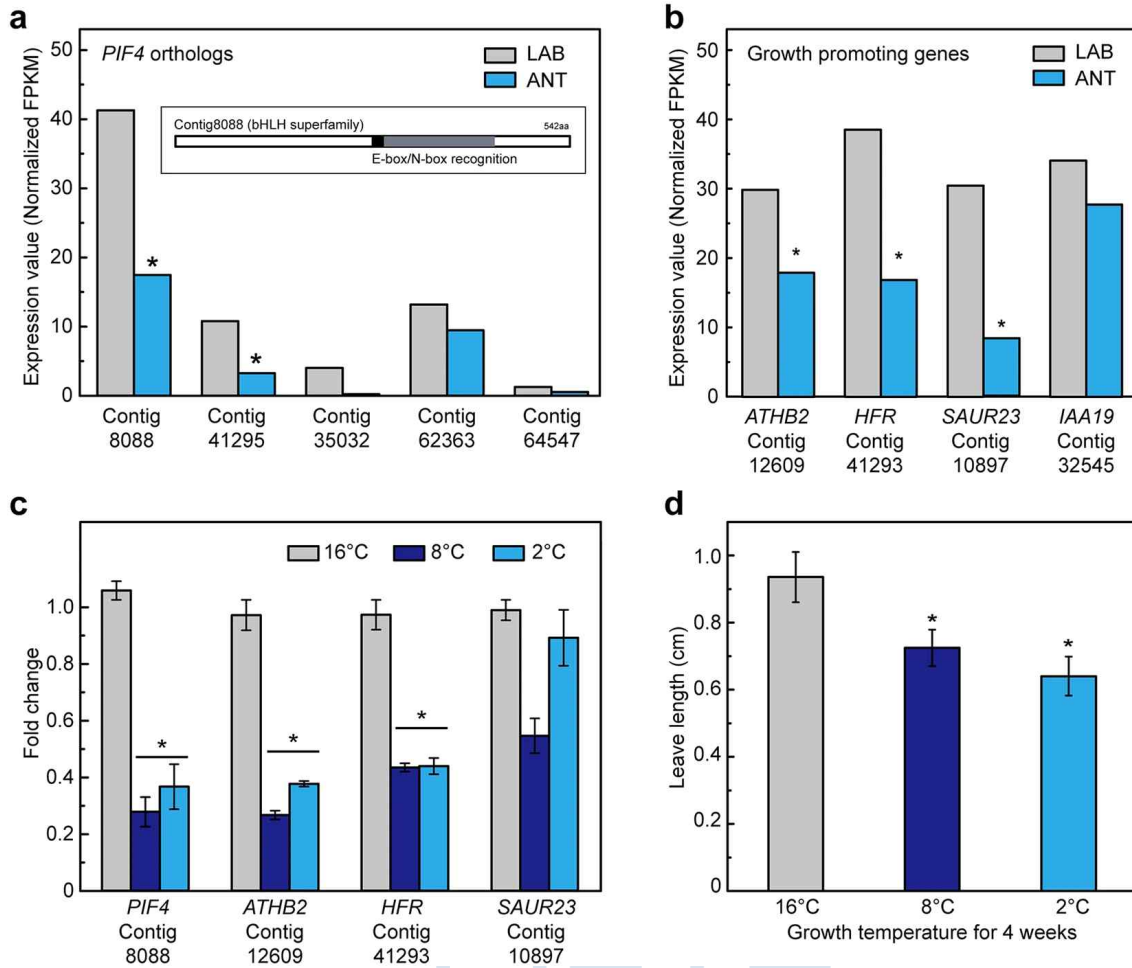


그림 8. The relative gene expression of the orthologs of *PIF4* and growth promoting genes as *PIF4*-targets. (a) The relative gene expression of *PIF4*-like orthologs which were found by TBLASTN search using *PIF4* as a query. In an inset, it is shown that Contig8088 contains the E-box/N-box recognition site. (b) The expression values of growth-promoting genes, *HFR*, *SAUR23*, *ATHB2*, and *IAA19* as *PIF*-targets. The grey and light-blue bars indicate the expression values of LAB and ANT samples, respectively. The left axis indicates the expression values (normalized FPKM). (c) Relative expression of *PIF4* and growth promoting genes, *ATHB2*, *HFR* and *SAUR23* from *C. quitensis* plants treated 2 °C, 8 °C and 16 °C for 4 weeks. Mean and standard deviations are shown (n = 3). (d) The leaf length of *C. quitensis* plants treated 2 °C, 8 °C and 16 °C for 4 weeks were measured by ImageJ program. Mean and standard deviations are displayed (n = 5). The significances were represented by p-values (*p < 0.05, t-test).

제 3-4절 남극좁새풀 유전자 DaGolS2 과발현에 의한 벼 형질전환체의 내냉성 증대 효과 분석

Poaceae type II galactinol synthase 2 improves cold and drought tolerance by accumulation of RFOs in Antarctic flowering plant *Deschampsia antarctica*

요약

남극좁새풀(*Deschampsia antarctica*)은 남극에 적응해서 서식하고 있는 벼과식물이다. 남극좁새풀이 남극의 외부 환경과 실험실 조건에서 저온 및 건조 스트레스를 받았을 때, 공통적으로 갈락티놀로 대표되는 라피노오스 계열 올리고당(RFO)이 고도로 축적되었고, 이는 갈락티놀 합성효소(GolS) 전사체의 증가를 동반하였다. 단자엽인 벼과식물은 소수의 GolS 유전자를 가지고 있어, I형과 II형의 두 가지 그룹으로 나뉜다. II형 GolS는 저온에 적응한 단자엽 식물에서 고도로 확장되었다. 남극좁새풀과 벼에서 유래한 II형 GolS2(DaGolS2, OsGolS2)가 과발현된 형질전환 벼 식물은 야생형과 비교하여 저온 및 건조 스트레스에 대해 현저하게 증가된 내성을 보였다. RFO 함량 및 GolS 효소 활성은 정상 및 스트레스 조건 모두에서 야생형 식물보다 DaGolS2 및 OsGolS2 과발현체에서 더 높았다. DaGolS2 및 OsGolS2 과발현체는 저온 및 건조 처리 후 야생형 식물에 비해 감소된 수준의 ROS를 함유하였다. 전반적으로, 이러한 결과는 벼과식물의 II GolS2가 남극좁새풀과 벼에서 RFO의 축적을 유도하고 ROS 수준을 감소시킴으로써 가뭄과 냉해 스트레스에 대한 저항성을 증가시키는 역할을 한다는 것을 시사한다.

Abstract

Deschampsia antarctica is a Poaceae grass that has adapted to and colonized Antarctica. When *D. antarctica* plants were subjected to cold and dehydration stress both in the Antarctic field and in laboratory experiments, galactinol, a precursor of raffinose family oligosaccharides (RFOs), and raffinose were highly accumulated, which was accompanied by upregulation of galactinol synthase (GolS). The Poaceae monocots have a small family of GolS genes, which are divided into two distinct groups called types I and II. Type II GolSs are highly expanded in cold-adapted monocot plants. Transgenic rice plants, in which type II *D. antarctica* GolS2 (DaGolS2) and rice GolS2 (OsGolS2) were constitutively expressed, were markedly tolerant to cold and drought stress as compared to the

wild-type rice plants. The RFO contents and GolS enzyme activities were higher in the DaGolS2- and OsGolS2-overexpressing progeny than in the wild-type plants under both normal and stress conditions. DaGolS2 and OsGolS2 overexpressors contained reduced levels of ROS relative to the wild-type plants after cold and drought treatments. Overall, these results suggest that Poaceae type II GolS2s play a conserved role in *D. antarctica* and rice in response to drought and cold stress by inducing the accumulation of RFO and decreasing ROS levels.

1. 서론

남극은 저온, 제한된 물 가용성, 높은 토양 염도, 연간 태양광량의 급격한 변동 및 짧은 낮 길이로 인해 식물 성장과 발달을 심각하게 제한하는 가혹한 환경이다. 이처럼 불리한 조건에도 불구하고, 남극에는 두 가지의 현화식물이 있으며, 이들 중 하나인 남극좁새풀(*Deschampsia antarctica* Desv.)는 남극의 유일한 벼과식물이다(Bravo et al., 2001). 남극좁새풀은 광합성 특성과 이에 영향을 미치는 잎의 해부학적 구조 및 엽록체의 미세구조 등에서 남극 환경에 특화된 형태학적 특성을 보인다(Gielwanowska et al., 2005; Sáez et al., 2019). 또한, 다양한 유전자의 발현 조절을 수반하는 생리적 및 생화학적 변화는 남극 환경에 대한 남극좁새풀의 소성 반응과 관련이 있는 것으로 알려져 있다(Lee et al., 2010; Lee et al., 2013). 예를 들어, 결빙 조건에 대한 반응으로 얼음 재결정화 억제 단백질(DaIRIP) 발현 및 재결정화 억제(RI) 활성의 증가는 남극좁새풀의 결빙 내성에 기여할 수 있다(John et al., 2009). 또한, 남극좁새풀은 폐놀형 분자를 합성할 뿐만 아니라 효소 및 비효소적 항산화 시스템을 활성화하여 UVB 유발 산화 스트레스에 대한 내성을 나타냈다(Köhler et al., 2017). 여름철 성장기(1월 - 2월) 동안 남극좁새풀은 비구조적 탄수화물을 많이 축적하여 식물의 대사 적응 능력을 높인다(Zuñiga et al., 1996). 이러한 특성은 가혹한 환경에서 생존하는 능력과 관련이 있기 때문에, 남극좁새풀은 다양한 비생물적 스트레스에 대한 식물 적응을 담당하는 유전자 및 대사 기작의 연구를 위한 모델로 사용될 수 있다.

2. 재료 및 방법

가. 식물 재료 및 성장 조건

남극좁새풀 식물체는 2007년 1월 King George Island의 Barton Peninsula에 있는 King Sejong Antarctic Station(62°14'29"S; 58°44'18"W) 근처에서 채집하였다. 식물체는 시험관 내에서 16시간/8시간 광주기에서 1/2 MS 배지[Murashige and Skoog (MS) (Duchefa Biochemie, Haarlem, The Netherlands), 2% sucrose, 0.8% phytoagar, pH 5.7], 15°C에서 150 $\mu\text{mol}/\text{m}^2/\text{s}$ 의 광량에서 재배하였고 3주마다 새로

운 배지로 옮겼다. 벼(*Oryza sativa* L.)의 경우 자포니카 품종 '동진'의 건조 종자를 70% 에탄올로 세척한 다음 0.4% NaClO 용액으로 멸균했다. 멸균된 종자를 발아시키고 10-12일 동안 3% sucrose 및 0.75% phytoagar, pH 5.7의 MS 배지에서 재배시켰다. 발아된 유식물을 토양에 이식하고 온실에서 장일조건, 25-30°C에서 성장시켰다.

나. 스트레스 처리 및 qRT-PCR 반응

저온 처리를 위해, 15°C에서 키운 남극곰새풀 식물체를 다양한 시간(1시간-7일) 동안 4°C 챔버에서 배양했다. 탈수 처리를 위해, 식물체를 여과지로 옮기고, 15°C에서 건조시키고, 1, 2, 4 시간 후에 샘플링하였다. 일주기 리듬에 의한 변화를 피하기 위해 발현 분석을 위한 모든 샘플링을 동시에 수행하였다. RNeasy Plant Mini Kit (Qiagen, Hilden, Germany)를 사용하여 남극곰새풀 및 벼 식물의 성숙한 잎으로부터 total RNA를 분리하였다. RNA의 양과 품질은 ND-1000 분광 광도계(NanoDrop Technologies, Wilmington, DE, USA)를 사용하여 결정하였다. TOPscript 역전사 효소(Enzymomics, South Korea) 및 올리고 (dT) primer를 이용하여 2µg의 total RNA로부터 cDNA를 합성하였다. qRT-PCR 분석은 20µL 반응 혼합물[1µL의 1:10 희석된 cDNA 주형, 2µM의 각 프라이머, 10µL의 TB Green Premix ExTaq (TaKaRa, Japan)]에서 수행하였다. 증폭 절차는 95°C에서 5분의 변성 및 효소 활성화, 95°C에서 10초, 55°C에서 10, 72°C에서 15초, 40 사이클을 수행하였다. DaEF1a 유전자를 내부 대조군으로 사용하였다. PCR 증폭에 사용된 프라이머의 DNA 서열은 Table 1에 열거하였다.

다. 갈락티놀과 라피노스 함량 측정

남극곰새풀과 벼 식물체의 성숙한 잎(100 mg)을 액체 질소를 이용하여 미세한 분말로 분쇄한 다음, 1mL의 80% 메탄올로 균질화시켰다. 90°C에서 10분 동안 끓인 다음 10,000g에서 5분 동안 원심분리하였다. 상등액은 질소 플러시를 사용하여 실온에서 밤새 건조시켰다. 1mL의 HPLC 등급 증류수에 재현탁시키고 0.2µm 필터로 여과하였다. 당 분석은 Suga-pak(Waters, USA) 컬럼 및 RI 검출기(Shodex, RI-101, Japan)를 사용하여 HPLC(Ultimate 3000, Dionex, USA)로 분석하였다. 갈락티놀 및 라피노스의 피크는 표준체와 비교하여 계산하였다.

라. 세포 내 위치 분석

DaGolS1 및 DaGolS2 코딩 영역의 3' 말단에 sGFP를 붙여 35S CaMV 프로모터를 포함하는 pBI221 벡터에 삽입하였다. 35S:DaGolS1-sGFP, 35S:DaGolS2-sGFP 및 35S:OsGolS2-sGFP construct는 PEG 방법을 사용하여 남극곰새풀 또는 벼의 원형

질체에 발현시켰다. 형광 신호는 형광 현미경(BX51, Olympus, Japan)으로 관찰하였다. sGFP를 세포질 마커로 사용하였다.

마. 계통학적 발생 분석

단자엽 벼과식물과 쌍자엽식물을 대상으로, DaGolS1과 DaGolS2의 상동체의 아미노산 서열을 GenBank 데이터베이스에서 확보했다. MEGA7 소프트웨어를 사용하여 JTT 매트릭스 기반 모델로 계통수를 구축하였다(Kumar et al., 2016). 분석에 사용한 남극초새풀, *Oryzae*, *Leersia perrieri*, *Triticeae*, *Aegilops tauschii*, *Hordeum vulgare*, *Zea mays*, *Sorghum bicolor*, *foxtail millet*, *Brachypodium distachyon*, *Musa acuminata*, *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, *Selaginella moellendorffii*의 GenBank accession 번호는 Table 2에 제시하였다.

바. DaGolS2와 OsGolS2 과발현 형질전환 벼 식물체 제작

남극초새풀 DaGolS2 (GenBank MK286465)와 벼의 OsGolS2 (GenBank XP_015645380.1)의 전장 유전자 영역을 옥수수 유비퀴틴 프로모터를 포함하는 pGA2897 벡터 플라스미드에 삽입하였다. Ubi:DaGolS2 및 Ubi:OsGolS2 재조합 플라스미드는 전기천공법으로 아그로박테리움 균주 LBA4404로 형질전환되었다(Park et al. 2016). 벼 형질전환의 모든 과정은 최근에 확립된 프로토콜에 따라 수행하였다(Cui et al. 2018). 생성된 형질전환체 T0 식물은 온실 조건에서 토양에 이식하고 논에서 추가로 번식시켰다. 개별 식물체에서 수확한 형질전환 종자를 하이그로마이신 B 40 mg/L가 포함된 MS 배지에서 발아시켜 동형 접합 T2 식물체를 선별하였다. 동형 접합 T4 DaGolS2 과발현체와 T3 OsGolS2 과발현체가 이후 표현형 분석에 사용되었다.

사. 야생형과 형질전환 벼의 스트레스 처리

저온 스트레스 처리를 위해, 장일조건 28°C에서 5주간 재배한 벼 식물체를 4°C의 냉장실로 옮겼다. 8일간 저온 처리 후, 식물을 28°C로 옮겨 50일 동안 키운 다음 성장 패턴을 모니터링 하였다(Cui et al., 2018). 건강한 잎이 계속 자라는 식물은 생존한 것으로 간주되었고 생존율은 회복 후 50 일에 결정하였다. 6개 이상의 생물학적 독립 실험으로부터 데이터를 얻었다. 4°C에서 저온 처리(0, 5 및 10d)의 다른 시점에서 전해질 누출 분석을 수행하였다. 야생형 및 형질전환 벼 식물의 저온 스트레스 처리한 유식물을 진탕기(200rpm) 상에 35mL의 증류수를 함유하는 시험관에 실온에서 밤새 침지시켰다. 각 샘플의 전해질 전도도는 전도도 측정기(Orion Star A212, Thermo Scientific, USA)를 사용하여 평균 전후에 결정하였다(Min et al. 2016). 건조 처리를 위해, 5주간 키운 야생형과, T4 Ubi:DaGolS2, T3 Ubi:OsGolS2 식물체를

잎이 시들 때까지 9-10일 동안 물 공급을 중단했다. 그 다음 식물에 물을 다시 공급하고 스트레스 회복 시점에서 성장을 모니터링했다. 건강한 잎으로 자라는 식물은 생존한 것으로 간주하였다. 생존율과 잎 엽록소 함량은 회복 3-4주에 결정하였다. 6개 이상의 생물학적 독립 실험으로부터 데이터를 얻었다. 잎의 수분 손실률을 측정하기 위해, 6주된 야생형 및 형질전환 벼 식물로부터 분리된 잎을 여과지에 놓고 여러 시간 간격(15, 30, 60, 90, 120, 180, 240)에 정량하였다. 수분 손실률은 초기 신선한 중량의 백분율로 계산하였다. 메틸 바이올로겐(MV) 처리를 위해, 야생형 및 T4 Ubi:DaGolS2 및 T3 Ubi:OsGolS2 형질전환체 종자를 MS 배지에서 발아시켰다. 발아된 유식물을 0, 2 및 5 μ M MV(Sigma-Aldrich, Missouri, USA)가 포함된 MS 배지로 옮기고 동일한 조건에서 7일 동안 배양하였다.

3. 결과

가. 남극종새풀의 갈락티놀과 라피노스 함량은 남극 환경과, 저온 및 건조 스트레스 조건에서 증가

RFO는 광범위한 비생물적 스트레스에 반응하여 다양한 식물의 조직에 축적된다고 알려져 있다(Sengupta et al., 2015). 따라서 본 연구에서는 남극종새풀이 해양성 남극 지역에 적응하는 과정에서 식물체 내 RFO가 결정적인 역할을 할 것이라 가정했다. 이를 검증하기 위해, 바톤반도의 세종기지 주변(62°04'29"S; 58°44'18"W) 현장에 서식하는 남극종새풀을 채집하고 실험실로 옮겨온 후 15°C, 장일 조건(16h 낮/8h 밤)에서 6일 동안 배양하였다. 이후 식물체를 다시 남극 현장으로 옮기고 다양한 시점(8 시간, 3 일 및 5 일)에서 수확하였다. 이 과정에서 남극종새풀의 잘 알려진 저온 유도 유전자인 DaIRIP의 발현은 식물체가 남극 현장에서 실험실로 옮겨 졌을 때 기저 수준으로 떨어졌다가, 남극 현장으로 다시 이동한 후 점차적으로 증가 하였다(Figure 1A).

대조구로 사용한 남극 현장 최초 샘플에서, 남극종새풀 잎의 갈락티놀과 라피노스 함량은 각각 $1.10 \pm 0.20 \text{mg/gFW}$ 및 $12.90 \pm 1.20 \text{mg/gFW}$ 였다(Figure 1B). 실험실 조건에서 6일간 배양했을 때(L6d), 갈락티놀 및 라피노스 함량은 각각 $0.002 \pm 0.001 \text{mg/gFW}$ 및 $1.19 \pm 0.04 \text{mg/gFW}$ 로 현저하게 감소하였다. 갈락티놀 및 라피노스의 함량은 식물을 남극 현장으로 옮겨심은 후에 점차적으로 증가했다. 8시간(F8h), 3일(F3d) 및 5일(F5d)에서, 갈락티놀 함량은 $0.84 \pm 0.18 \text{mg/gFW}$, $1.89 \pm 0.37 \text{mg/gFW}$ 및 $5.86 \pm 1.30 \text{mg/gFW}$ 이었고, 동일한 시점에 라피노스 함량은 각각 $2.44 \pm 0.17 \text{mg/gFW}$, $9.42 \pm 1.96 \text{mg/gFW}$ 및 $38.67 \pm 2.20 \text{mg/gFW}$ 로 증가하였다(Figure 1B). 갈락티놀은 라피노스보다 더 빠르게 증가했는데, 이는 변화하는 환경에서 두 가지 RFO가 단계적으로 합성됨을 시사한다. 대조적으로, 자당, 포도당 및

과당의 수준은 RFO만큼 크게 변하지 않았다(Figure 1C-D).

다음에는 실험실에서 배양한 남극곰새풀 식물을 대상으로 거온과 건조 스트레스에 반응하는 RFO 함량을 측정했다. 저온 스트레스 처리를 위해, 15°C에서 성장한 남극곰새풀 식물을 4°C에서 10일 동안 배양했다. 탈수 처리를 위해, 남극곰새풀 식물을 15°C에서 여과지 위에서 4일 동안 배양했다. 그 결과 저온 및 건조 처리 후 갈락티놀과 라피노스의 수준이 크게 증가했다. 처리 전 대조군의 갈락티놀 ($0.59 \pm 0.39 \text{mg/gFW}$)은 저온 처리 후 $4.32 \pm 1.08 \text{mg/gFW}$, 건조 처리 후 $2.05 \pm 0.20 \text{mg/gFW}$ 로 증가했다(Figure 1E-G). 라피노스는 대조군에서 $0.53 \pm 0.53 \text{mg/gFW}$, 저온 처리 후 $21.12 \pm 0.80 \text{mg/gFW}$ 로, 건조 처리 후 $5.62 \pm 0.73 \text{mg/gFW}$ 로 증가했다(Figure 1E-G). 대조적으로, 자당, 포도당 및 과당의 상대 수치는 건조와 저온 처리 조건에서 변하지 않았다(Figure 1G). 따라서, RFO의 축적은 비생물 스트레스 환경에서 남극곰새풀이 적응하는 과정에 중요한 역할을 할 것으로 보인다. 이러한 결과는 RFO의 생합성과 축적이 가혹한 남극 환경에 대처하기 위해 남극곰새풀이 발달시킨 정교한 적응기작일 가능성이 높다.

나. 남극곰새풀 DaGolS1과 DaGolS2 유전자 특성 분석

갈락티놀 합성효소(GolS)는 RFO 생합성 경로의 초기 단계를 촉매하는 핵심 효소다. 남극곰새풀은 두 개의 상동성 GolS 유전자인 DaGolS1(GenBank MK286464)과 DaGolS2(GenBank MK286465)를 가지고 있다. DaGolS1과 DaGolS2는 서로 72% 상동성을 갖는 336개 및 326개의 아미노산으로 구성된 서열을 포함한다(Figure 2A). 두 단백질 모두 1개의 글리코실 트랜스퍼라제 도메인과 보존된 DxD 모티프를 포함한다(Sengupta et al., 2012)(Figure 2A). 실시간 qRT-PCR 분석 결과 남극곰새풀이 남극 현장에서 실험실로 옮겨질 때 DaGolS1과 DaGolS2의 발현 수준이 변하지 않았다(Figure 2B). 그러나 DaGolS1과 DaGolS2의 전사 수준은 식물을 실험실에서 다시 남극 현장으로 이동 한 후 8 시간에 빠르게 증가했다가, 장기간 노출(3일에서 5일)되었을 경우 기저 수준으로 다시 복귀되었다(Figure 2B). DaGolS1 및 DaGolS2의 이러한 발현 양상은 초기 저온 반응 유전자의 서브 세트를 활성화시키는 전사 인자인 DaCBF4 및 DaCBF7의 발현 양상과 매우 유사했다(Figure 2B; Byun et al., 2015; Byun et al., 2018). 이러한 결과는 DaGolS1과 DaGolS2가 남극곰새풀에서 저온 스트레스 반응의 초기 단계에서 기능한다는 것을 의미한다. 이러한 개념은 실험실에서 저온 조건(4°C)에서 8시간(최대 25배 및 8배) 및 건조 조건에서 4일(최대 6배 및 11배)에 남극곰새풀에서 DaGolS1 및 DaGolS2 전사체가 빠르게 증가했다는 결과에 의해 뒷받침되었다. 따라서, RFO의 증가(Figure 1)는 스트레스 조건에서 DaGolS1 및 DaGolS2(Figure 2B-D)의 신속한 전사 유도와의 비례관계를 보였다.

DaGolS의 세포 내 위치를 원형질체 시스템을 통해 조사하였다. 원형질체를 남극종새풀의 성숙한 잎으로부터 제작한 후, 35S:DaGolS1-sGFP와 35S:DaGolS2-sGFP 플라스미드를 형질전환하였다. 발현된 단백질은 형광 현미경으로 관찰하였다. 그 결과 DaGolS1-sGFP 및 DaGolS2-sGFP에 의한 형광 신호는 주로 세포질 분획에 위치하고 있었으며 이는 DaGolS1 및 DaGolS2가 세포질 효소임을 시사한다(Figure 2E).

다. DaGolS2는 벼과 II형 GolS에 속한다

단자엽 벼과식물 유래의 GolS 서열로 계통발생 분석을 수행한 결과, 이들은 소수의 GolS 유전자군을 가지고 있으며, 이는 I형과 II형의 두 가지 그룹으로 구분되었다(Figure 3). 예를 들어, 남극종새풀과 벼는 두 개 유전자(유형 I:II = 1:1)를 가지고 있고 보리는 세 개의 유전자(유형 I:II = 1:2)를 가지며 옥수수에는 네 개 유전자가 있다(유형 I:II = 3:1). 대표적인 동결 내성 작물이면서 6배체인 겨울 밀 (*Triticum aestivum* L.)은 B 계통에 6개 유전자를 포함하며, 대부분 II형 GolS 유전자였다(유형 I:II = 1:5)(Figure 3). 밀의 II형 유전자 중 하나인 TaGolS1의 기능이 확인된 바 있는데, TaGolS1을 과발현하는 형질전환 벼는 저온 스트레스에 대한 내성을 증가시켰다(Shimosaka et al., 2015). DaGolS1은 I형 GolS이고 DaGolS2는 II형 효소이기 때문에(Figure 3), 우리는 저온 내성과 연관이 있을 것으로 예상되는 DaGolS2를 선택하여 남극종새풀 II형 GolS의 식물 내 역할을 조사했다.

라. DaGolS2와 OsGolS2를 과발현하는 형질전환 벼 제작

다양한 식물 종을 대상으로 스트레스 내성 유전자의 식물 내에서의 역할을 탐색하기 위해 유전자 형질전환 기법이 널리 사용되어 왔지만, 남극종새풀을 대상으로 안정적인 유전자 변형 기법은 아직 확립되지 않았다. 남극종새풀은 단자엽 식물의 모델인 벼와 같은 벼과에 속한다. 따라서, DaGolS2의 역할을 조사하기 위해 옥수수 유비퀴틴 프로모터 (Ubi)의 제어하에 벼에서 DaGolS2를 과발현하는 형질전환 벼를 제작하였다(Figure 4A). 정상적인 조건에서, DaGolS2를 과발현하는 T4 식물 (Ubi:DaGolS2)은 야생형과 비교하여 형태적 차이를 보이지 않았다(Figure 4B). 계통 서던 결과를 토대로, 2 개의 독립된 Ubi:DaGolS2 라인이 선택되었고, 여기서 DaGolS2 전사체를 RT-PCR로 검출하였다(Figure 4C, D). 더불어, GolS 활성화도 야생형에 비해 Ubi:DaGolS2 라인에서 증가하였다(Figure 4E).

벼과 식물에서 보편적인 GolS2의 생물학적 기능을 알아보기 위해, 우리는 또한 DaGolS2의 벼 상동체인 OsGolS2를 분리하였다. OsGolS2는 DaGolS2와 79%의 상동성을 갖는 328개의 아미노산을 포함한다(Figure 4F). OsGolS2는 건조와 고염 스트레스에 의해 유전자 발현이 빠르게 유도되었다. 그러나 DaGolS2와 달리 OsGolS2는 0.5-2일간의 저온 스트레스 조건에서는 유도되지 않았고, 5일 저온 처리 후에야 증

가하였다(Figure 4G). 벼에서 OsGolS2의 이러한 저온 유도 패턴은 DaGolS1이나 DaGolS2와는 달랐는데, 남극종새풀에서는서 저온처리 8시간만에 발현이 증가하였기 때문이다(Figure 2C). OsGolS2 과발현 형질전환 벼를 제작하여 표현형 분석을 진행한 결과, 야생형과 비교하여 Ubi:OsGolS2 자손에서 증가된 GolS 활성이 확인되었다(Figure 4L).

마. DaGolS2와 OsGolS2를 과발현 형질전환 벼는 저온 스트레스에 강하다

DaGolS2 및 OsGolS2 과발현 식물체의 내냉성 표현형을 조사하기 위해, 야생형과 T4 Ubi:DaGolS2 및 T3 Ubi:OsGolS2 식물체를 5주 동안 28°C에서 키운 후, 4°C의 냉장실로 옮기고 8일 동안 처리한 후, 다시 28°C 성장실로 옮겨 50일 동안 회복시켰다. 대부분의 야생형 벼는 저온 스트레스로부터 회복된 후 잎이 변색되면서 생존하지 못했다($16.8 \pm 2.5 \sim 20.9 \pm 3.8\%$) (Figure 5A). 대조적으로, DaGolS2와 OsGolS2 과발현체는 보다 건강하게 성장을 재개하였다(각각 $64.0 \pm 9.2 \sim 84.4 \pm 4.0\%$ 및 $61.0 \pm 4.8 \sim 63.8 \pm 4.4\%$) (Figure 5A-D).

잎의 엽록소 함량을 측정하기 위해, 저온 처리 전후에 각 유전자형의 식물에서 성숙한 잎을 분리하였다. 저온 처리 전에, 야생형과 Ubi:DaGolS2 및 Ubi:OsGolS2 식물의 엽록소 함량은 차이가 없었다(Figure 5E, F). 그러나, DaGolS2 및 OsGolS2 과발현체는 저온에 반응하여 야생형보다 더 많은 양의 엽록소를 함유하였다. 저온 처리(4°C)에서 1개월 회복 후, 야생형 잎의 엽록소 함량은 $1.5 \pm 0.7 \sim 2.0 \pm 0.4$ mg / gDW인 반면, Ubi:DaGolS2 및 Ubi:OsGolS2 식물의 엽록소 함량은 각각 $10.4 \pm 1.3 \sim 10.8 \pm 0.6$ 및 $8.5 \pm 0.9 \sim 10.3 \pm 1.5$ mg/gDW였다(Figure 5E, F).

저온 스트레스에 대한 세포 반응을 정량화하기 위해, 8일된 유식물을 사용하여 전해질 누출 분석을 수행하였다. 야생형, Ubi:DaGolS2 및 Ubi:OsGolS2의 유식물을 4°C에서 0, 5 및 10일 동안 배양하고, 모종을 25 ml에서 밤새 35 ml의 증류수에 침지시킨 후 전도도 측정기로 전해질 누출을 측정하였다. Figure 5G에서 볼 수 있듯이 Ubi:DaGolS2 및 Ubi:OsGolS2 유식물은 5일째 각각 $12.4 \pm 0.7 \sim 12.5 \pm 0.7\%$ 및 $12.5 \pm 0.7 \sim 12.7 \pm 0.7\%$, 10일째 각각 $15.4 \pm 1.0 \sim 15.8 \pm 1.2\%$ 및 $14.4 \pm 2.4 \sim 16.5 \pm 2.0\%$ 였으며, 야생형은 5일째 $14.7 \pm 1.1\%$, 10일째 $25.6 \pm 1.7\%$ 였다. 이러한 결과는 DaGolS2 및 OsGolS2의 과발현이 저온 처리에 대한 내성을 증가시켰으며, 이는 DaGolS2 및 OsGolS2가 벼의 저온 스트레스 반응에서 긍정적인 역할을 한다는 것을 시사한다.

바. DaGolS2와 OsGolS2를 과발현 형질전환 벼는 건조 스트레스에 강하다

야생형, T4 Ubi:DaGolS2 및 T3 Ubi:OsGolS2 식물을 28°C에서 5주 동안 성장시킨 다

음 9-10일 동안 물을 주지 않고 식물에 건조 스트레스를 주었다. 그 후 식물에 물을 다시 주고 3-4주 동안 자라도록 하여 회복을 관찰했다. 건조 스트레스로부터 회복한 후 Ubi:DaGolS2 식물체 중 $28.0 \pm 4.0 - 32.0 \pm 2.8\%$, Ubi:OsGolS2 식물체의 $38.1 \pm 5.7 - 51.0 \pm 8.1\%$ 가 성장을 재개했다(Figure 6A-D). 대조적으로, 야생형의 대부분은 잎이 변색되어 심하게 시들었고 생존율은 $8.8 \pm 2.0 - 11.0 \pm 2.0\%$ 로 낮았다(Figure 6A-D).

건조 처리 후 DaGolS2와 OsGolS2 과발현 형질전환체는 야생형보다 훨씬 많은 양의 엽록소를 포함하고 있었다(각각 $9.6 \pm 2.5 - 10.7 \pm 1.7$ 및 $11.3 \pm 2.1 - 11.6 \pm 1.1$ mg/gDW)(Figure 6E, F). 또한, 형질전환체의 잎은 야생형보다 수분 손실 속도가 더 느렸다. 실온에서 5시간 동안 둔 후, 야생형 잎은 생중량의 $48.9 \pm 4.9\%$ 를 유지하였다. 대조적으로, Ubi:DaGolS2 및 Ubi:OsGolS2 잎은 생중량의 $61.7 \pm 1.8 - 65.1 \pm 3.9\%$ 및 $62.8 \pm 2.1 - 63.0 \pm 2.1\%$ 를 여전히 각각 유지하였다(Figure 6G). 따라서, DaGolS2 및 OsGolS2의 과발현은 벼의 건조 스트레스에 대한 내성을 향상시켰다. 이는 DaGolS2와 OsGolS2가 저온과 건조 스트레스에 대한 반응에서 긍정적인 요인임을 나타낸다.

4. 논의

고등 식물은 다수의 환경 요인에 의해 야기되는 불리한 성장 조건에 대처하기 위해 복잡한 생리적 및 생화학적 적응법을 발전시켰다(Seki et al. 2007). 남극종새풀은 식물들이 가혹한 남극 기후를 견딜 수 있도록 하는 형태학적 및 생리적 적응과 관련한 다양한 특성을 가지고 있다(Cavieres et al. 2016). RFO는 식물의 비구조 탄수화물이며 광범위한 비생물적 스트레스에 반응하여 보호 역할을 한다(Sengupta et al. 2015). 남극종새풀이 성장기 및 종자 성숙 과정에서 탄수화물을 많이 축적하는 것으로 알려져 있지만, 탄수화물 중 특히 RFO의 세포 역할과 비생물적 스트레스 내성 반응과의 연관성은 명확하지 않다(Zuñiga et al. 1996, Zúñiga-Feest et al. 2003).

이 연구에서 우리는 먼저 남극의 자연 조건과 실험실에서 인위적으로 조성된 저온 및 건조에서 남극종새풀의 RFOs의 변화를 관찰했다. 그 결과 자연 조건과 실험실 조건 모두에서 식물체 내 갈락티놀과 라피노스 함량이 현저히 증가했다(Figure 1). 식물에서 RFO 생합성의 첫 번째 단계를 촉매하는 효소인 갈락티놀 합성효소를 암호화하는 두 개의 유전자 DaGolS1와 DaGolS2를 분리하고 이들의 유전자 발현양상을 조사한 결과, 두 유전자는 저온과 건조 조건에서, 그리고 식물이 남극 현장으로 옮겨졌을 때 빠르게 발현이 증가하였다(Figure 2). 이러한 DaGolS의 초기 유도 양상은 저온 반응의 핵심 전사 인자인 DaCBF4와 DaCBF7을 연상시킨다(Byun et al. 2015, Byun et al. 2018). 기존 보고에서, AtGolS 유전자는 10시간 저온처리에 의해 신속하게 유도되었다(Fowler and Thomashow 2002, Taji et al. 2002). 벼과식물에서는,

겨울종인 *Festuca arundinacea* 유래의 6개 GolS 유전자, 밀과 *Lolium perenne*의 3개 GolS 유전자와 TaGolS3가 저온 적응 반응의 일부로서 조기 유도되는 특성을 보였다(Wang et al. 2016a, Wang et al. 2016b). 이러한 연구결과는 GolS가 저온 적응종에서 식물의 내한성에 중요한 기능을 할 수 있음을 시사한다. 그러나 GolS 유전자의 전사 패턴은 작물에 따라 차이를 보인다. 벼의 OsGolS2는 건조 조건에서는 1시간째에 유도되었지만, 저온 조건에서는 5일만에 발현이 증가하였다 (Figure 4G). 건조 스트레스에 대한 GolS의 이러한 두드러진 반응은 옥수수에서도 알려졌다. 옥수수의 ZmGolS1, ZmGolS2 및 ZmGolS3의 전사체는 건조와 고온 조건에서 발현이 증가했지만 저온 조건에서는 별다른 반응을 보이지 않았다(Zhao et al. 2004). 저온과 건조 처리 조건에서 Ubi:DaGolS2와 Ubi:OsGolS2 식물들 간 유의한 표현형 차이가 없었기 때문에, 우리는 GolS의 전사 조절 시스템이 식물종의 적응 전략으로서의 기후에 따라 다양한 방식으로 차별화되었다고 가정하였다.

글리코실 트랜스퍼라제 8(GT8) 유전자군에 속하는 GolS는 고등 식물 특이적 대사물인 RFO의 생합성을 관장하는 주요 효소이다(Yin et al. 2010). GT8 유전자군에 속하는 많은 구성원 중에서 GolS는 다양한 비생물 스트레스에 반응하는 유일한 그룹이다(Sengupta et al. 2012). 식물에서 GolS 유전자 수가 증가하고 개별 GolS 유전자의 발현 양상이 차별화된 것은 불리한 환경을 효과적으로 극복하도록 진화하는 과정에서 RFO의 필요성을 시사한다(Taji et al. 2002, Downie et al. 2003, Gu et al. 2016). GolS 유사 단백질들의 계통도에서, 단자엽과 쌍자엽 간의 뚜렷한 분리가 확인되며, 이는 GolS가 단자엽과 쌍자엽 식물의 다양화 이전에 존재했음을 시사한다(Yin et al. 2010, Sengupta et al. 2012). 단자엽 식물의 GolS는 I형과 II형 효소로 분류된다(Figure 3). II형 GolS의 뚜렷한 특징은 저온 스트레스 내성 작물로 알려진 Triticeae 하위 식물군에서 그 수가 크게 늘어난 것이다. 겨울밀(*T. aestivum* L.)은 I형보다 II형에 속하는 GolS의 수가 더 많으며, 저온에 약한 옥수수는 I형보다 II형 GolS의 수가 더 적었다. 각 유형별 GolS 개수는 다양한 환경 및 생리적 상황에서 식물의 성장과 발달을 유지하기 위한 식물 유전자의 기능적 분기와 관련될 수 있다.

GolS 단백질의 GT8 도메인은 고도로 보존되어 있지만, 단자엽과 쌍자엽 GolS의 DxD 모티프는 뚜렷하게 구분되는 특성이 있다(Sengupta et al. 2012). 이는 단자엽 I형 GolS가 II형 GolS보다 쌍자엽 GolS와 더 인접한 상관 관계가 있음을 나타낸다. 예측된 인산화 부위 Ser324가 II형 DaGolS2에서는 잘 보존되어 있지만 I형 DaGolS1에서는 잘 보존되지 않다는 사실도, 남극좀새풀 DaGolS 효소의 독특한 특성을 야기할 가능성이 있다. 인산화 부위의 이러한 변화는 특정 단백질 변형과 관련될 수 있으며, 순차적으로 다른 발달 및 환경 신호에 대한 반응으로써 GolS의 세포내 기능을 미세하게 조정 가능하다(Zhang et al. 2018). 전반적으로 이러한 결과는 각 GolS 타입이

식물에서 명확하고 다양한 역할을 수행하여 불안정성을 극복하고 지속적으로 변동하는 환경에서 생존하기 위해 기능적으로 진화했음을 암시한다.

결론적으로, 본 연구결과는 벼과의 II형 GolS인 DaGolS2 및 OsGolS2의 식물 내 과발현이 RFO의 축적을 유도하고 ROS를 감소시킴으로써 형질전환 벼의 저온 및 건조 스트레스에 대한 내성을 증가시켰음을 보여 주었다. 따라서 극한 환경 조건 하에서 세포 단백질과 막을 안정화하고 ROS 축적을 막기 위해 RFO를 효율적으로 생산하기 위해 DaGolS 유전자를 유도하는 것은 해양성 남극에서 남극종새풀이 살아남기 위해 발전시킨 효과적인 생존 전략이라 할 수 있다.

5. 참고문헌

1. Bartels, D. and Sunkar, R. (2005) Drought and salt tolerance in plants. *Crit. Rev. Plant Sci.* 24: 23-58.
2. Blokhina, O., Virolainen, E. and Fagerstedt, K.V. (2003) Antioxidants, oxidative damage and oxygen deprivation stress: a review. *Ann. Bot.* 91: 179-194.
3. Bravo, L.A., Ulloa, N., Zuniga, G.E., Casanova, A., Corcuera, L.J. and Alberdi, M. ~ (2001) Cold resistance in Antarctic angiosperms. *Physiol. Plant.* 111: 55-65.
4. Byun, M.Y., Lee, J., Cui, L.H., Kang, Y., Oh, T.K. and Park, H. (2015) Constitutive expression of DaCBF7, an Antarctic vascular plant *Deschampsia antarctica* CBF homolog, resulted in improved cold tolerance in transgenic rice plants. *Plant Sci.* 236: 61-74.
5. Byun, M.Y., Cui, L.H., Lee, J., Park, H., Lee, A., Kim, W.T., et al. (2018) Identification of rice genes associated with enhanced cold tolerance by comparative transcriptome analysis with two transgenic rice plants overexpressing DaCBF4 or DaCBF7, isolated from Antarctic flowering plant *Deschampsia antarctica*. *Front. Plant Sci.* 9: 601.
6. Cavieres, L.A., Saez, P., Sanhueza, C., Sierra-Almeida, A., Rabert, C., Corcuera, L.J., et al. (2016) Ecophysiological traits of Antarctic vascular plants: their importance in the responses to climate change. *Plant Ecol.* 217: 343-358.
7. Cui, L.H., Min, H.J., Byun, M.Y., Oh, H.G. and Kim, W.T. (2018) OsDIRP1, a putative RING E3 ligase, plays an opposite role in drought and cold stress responses as a negative and positive factor, respectively, in rice (*Oryza sativa* L.). *Front. Plant Sci.* 9: 1797.

8. Downie, B., Gurusinghe, S., Dahal, P., Thacker, R.R., Snyder, J.C., Nonogaki, H., et al. (2003) Expression of a GALACTINOL SYNTHASE gene in tomato seeds is up-regulated before maturation desiccation and again after imbibition whenever radicle protrusion is prevented. *Plant Physiol.* 131: 1347–1359.
9. ElSayed, A.I., Rafudeen, M.S. and Golldack, D. (2014) Physiological aspects of raffinose family oligosaccharides in plants: protection against abiotic stress. *Plant Biol. J.* 16: 1–8.
10. Fiske, C.H. and Subbarow, Y. (1925) The colorimetric determination of phosphorus. *J. Biol. Chem.* 66: 375–400.
11. Fowler, S. and Thomashow, M.F. (2002) Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. *Plant Cell* 14: 1675–1690.
12. Gidekel, M., Destefano-Beltran, L., Garcia, P., Mujica, L., Leal, P., Cuba, M., et al. (2003) Identification and characterization of three novel cold acclimation-responsive genes from the extremophile hair grass *Deschampsia antarctica* Desv. *Extremophiles* 7: 459–469.
13. Gielwanowska, I., Szczuka, E., Bednara, J. and Gorecki, R. (2005) Anatomical features and ultrastructure of *Deschampsia antarctica* (Poaceae) leaves from different growing habitats. *Ann. Bot.* 96: 1109–1119.
14. Gu, L., Zhang, Y., Zhang, M., Li, T., Dirk, L.M.A., Downie, B., et al. (2016) ZmGOLS2, a target of transcription factor ZmDREB2A, offers similar protection against abiotic stress as ZmDREB2A. *Plant Mol. Biol.* 90: 157–170.
15. Heath, R.L. and Packer, L. (1968) Photoperoxidation in isolated chloroplasts: I. Kinetics and stoichiometry of fatty acid peroxidation. *Arch. Biochem. Biophys.* 125: 189–198.
16. Hue, N.T., Tran, H.T.T., Phan, T., Nakamura, J., Iwata, T., Harano, K., et al. (2013) Hsp90 and reactive oxygen species regulate thermotolerance of rice seedlings via induction of heat shock factor A2 (OsHSFA2) and galactinol synthase 1. *Agric. Sci.* 04: 154–164.
17. John, U.P., Polotnianka, R.M., Sivakumaran, K.A., Chew, O., Mackin, L., Kuiper,

- M.J., et al. (2009) Ice recrystallization inhibition proteins (IRIPs) and freeze tolerance in the cryophilic Antarctic hair grass *Deschampsia antarctica* E. Desv. *Plant Cell Environ.* 32: 336–348.
18. Köhler, H., Contreras, R.A., Pizarro, M., Cortés-Antiquera, R. and Zúñiga, G.E. (2017) Antioxidant responses induced by UVB radiation in *Deschampsia antarctica* Desv. *Front. Plant Sci.* 8: 921.
19. Kumar, S., Stecher, G. and Tamura, K. (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33: 1870–1874.
20. Lee, H., Kim, J.H., Park, M., Kim, I.C., Yim, J.H. and Lee, H.K. (2010) Reference genes validation for qPCR normalization in *Deschampsia antarctica* during abiotic stresses. *Antarct. Sci.* 22: 477–484.
21. Lee, J., Noh, E.K., Choi, H.S., Shin, S.C., Park, H. and Lee, H. (2013) Transcriptome sequencing of the Antarctic vascular plant *Deschampsia antarctica* Desv. under abiotic stress. *Planta* 237: 823–836.
22. Lehle, L. and Tanner, W. (1973) The function of myo-inositol in the biosynthesis of raffinose. Purification and characterization of galactinol: sucrose 6-galactosyltransferase from *Vicia faba* seeds. *Eur. J. Biochem.* 38: 103–110.
23. Loewus, F.A. and Murthy, P.P.N. (2000) myo-Inositol metabolism in plants. *Plant Sci.* 150: 1–19.
24. Min, H.J., Jung, Y.J., Kang, B.G. and Kim, W.T. (2016) CaPUB1, a hot pepper u-box E3 ubiquitin ligase, confers enhanced cold stress tolerance and decreased drought stress tolerance in transgenic rice (*Oryza sativa* L.). *Mol. Cells* 39: 250–257.
25. Nishizawa, A., Yabuta, Y. and Shigeoka, S. (2008) Galactinol and raffinose constitute a novel function to protect plants from oxidative damage. *Plant Physiol.* 147: 1251–1263.
26. Nishizawa-Yokoi, A., Yabuta, Y. and Shigeoka, S. (2008) The contribution of carbohydrates including raffinose family oligosaccharides and sugar alcohols to protection of plant cells from oxidative damage. *Plant Signal. Behav.* 3: 1016–1018.

27. Park, K.Y., Kim, E.Y., Seo, Y.S. and Kim, W.T. (2016) Constitutive expression of CaPLA1 conferred enhanced growth and grain yield in transgenic rice plants. *Plant Mol. Biol.* 90: 517–532.
28. Peshev, D., Vergauwen, R., Moglia, A., Hideg, E. and Van den Ende, W. (2013) Towards understanding vacuolar antioxidant mechanisms: a role for fructans- *J. Exp. Bot.* 64: 1025–1038.
29. Sraez, P.L., Rivera, B.K., Ramirez, C.F., Vallejos, V., Cavieres, L.A., Corcuera, L.J., et al. (2019) Effects of temperature and water availability on light energy utilization in photosynthetic processes of *Deschampsia antarctica*. *Physiol. Plant.* 165: 511–523.
30. Salvi, P., Kamble, N.U. and Majee, M. (2018) Stress-inducible galactinol synthase of chickpea (CaGolS) is implicated in heat and oxidative stress tolerance through reducing stress-induced excessive reactive oxygen species accumulation. *Plant Cell Physiol.* 59: 155–166.
31. Selvaraj, M.G., Ishizaki, T., Valencia, M., Ogawa, S., Dedicova, B., Ogata, T., et al. (2017) Overexpression of an *Arabidopsis thaliana* galactinol synthase gene improves drought tolerance in transgenic rice and increased grain yield in the field. *Plant Biotechnol. J.* 15: 1465–1477.
32. Seki, M., Umezawa, T., Urano, K. and Shinozaki, K. (2007) Regulatory metabolic networks in drought stress responses. *Curr. Opin. Plant Biol.* 10: 296–302.
33. Sengupta, S., Mukherjee, S., Parween, S. and Majumder, A.L. (2012) Galactinol synthase across evolutionary diverse taxa: functional preference for higher plants- *FEBS Lett.* 586: 1488–1496.
34. Sengupta, S., Mukherjee, S., Basak, P. and Majumder, A.L. (2015) Significance of galactinol and raffinose family oligosaccharide synthesis in plants. *Front. Plant Sci.* 6: 656.
35. Seo, D.H., Ahn, M.Y., Park, K.Y., Kim, E.Y. and Kim, W.T. (2016) The N-terminal UND motif of Arabidopsis U-box E3 ligase PUB18 is critical for the negative regulation of ABA-mediated stomatal movement and determination of its ubiquitination specificity to Exo70B1. *Plant Cell* 28: 2952–2973.

36. Shimosaka, E. and Ozawa, K. (2015) Overexpression of cold-inducible wheat galactinol synthase confers tolerance to chilling stress in transgenic rice. *Breed. Sci.* 65: 363-371.
37. Stevenson, J.M., Perera, I.Y., Heilmann, I., Persson, S. and Boss, W.F. (2000) Inositol signaling and plant growth. *Trends Plant Sci.* 5: 252-258.
38. Taji, T., Ohsumi, C., Iuchi, S., Seki, M., Kasuga, M., Kobayashi, M., et al. (2002) Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in *Arabidopsis thaliana*. *Plant J.* 29: 417-426.
39. Valluru, R. and Van den Ende, W. (2011) Myo-inositol and beyond-emerging networks under stress. *Plant Sci.* 181: 387-400.
40. Van den Ende, W. and Valluru, R. (2009) Sucrose, sucrosyl oligosaccharides, and oxidative stress: scavenging and salvaging. *J. Exp. Bot.* 60: 9-18.
41. Wang, Y., Dai, Y., Tao, X., Wang, J.Z., Cheng, H.Y., Yang, H., et al. (2016a) Heat shock factor genes of tall fescue and perennial ryegrass in response to temperature stress by RNA-seq analysis. *Front. Plant Sci.* 6: 1226.
42. Wang, Y., Liu, H., Wang, S., Li, H. and Xin, Q. (2016b) Overexpression of a common wheat gene GALACTINOL SYNTHASE3 enhances tolerance to zinc in *Arabidopsis* and rice through the modulation of reactive oxygen species production. *Plant Mol. Biol. Rep.* 34: 794-806.
43. Xiong, H., Yu, J., Miao, J., Li, J., Zhang, H., Wang, X., et al. (2018) Natural variation in OsLG3 increases drought tolerance in rice by inducing ROS scavenging. *Plant Physiol.* 178: 451-467.
44. Xue, H., Chen, X. and Li, G. (2007) Involvement of phospholipid signaling in plant growth and hormone effects. *Curr. Opin. Plant Biol.* 10: 483-489.
45. Yin, Y., Chen, H., Hahn, M.G., Mohnen, D. and Xu, Y. (2010) Evolution and function of the plant cell wall synthesis-related glycosyltransferase family 8. *Plant Physiol.* 153: 1729-1746.
46. Zhang, M., Su, J., Zhang, Y., Xu, J. and Zhang, S. (2018) Conveying endogenous and exogenous signals: MAPK cascades in plant growth and defense. *Curr. Opin. Plant Biol.* 45: 1-10.

47. Zhao, T.-Y., Martin, D., Meeley, R.B. and Downie, B. (2004) Expression of the maize GALACTINOL SYNTHASE gene family: (II) Kernel abscission, environmental stress and myo-inositol influences accumulation of transcript in developing seeds and callus cells. *Physiol. Plant.* 121: 647-655.
48. Zuniga, G.E., Alberdi, M. and Corcuera, L.J. (1996) Non-structural carbohydrates in *Deschampsia antarctica* Desv. from South Shetland Islands, Maritime Antarctic. *Environ. Exp. Bot.* 36: 393-399.
49. Zuniga-Feest, A., Inostroza, P., Vega, M., Bravo, L.A. and Corcuera, L.J. (2003) Sugars and enzyme activity in the grass *Deschampsia antarctica*. *Antartic. Sci.* 15: 483-491.
50. Zuther, E., Buchel, K., Hundertmark, M., Stitt, M., Hinch, D.K. and Heyer, A. G. (2004) The role of raffinose in the cold acclimation of *Arabidopsis thaliana*. *FEBS Lett.* 576: 169-173.



Table 1. Information of primers used in this study.

Oligonucleotide name	Sequence	Purpose
<i>DaGolS1</i> CDS F	ATGGCTCCCGAGCTGGCCGGCAA	gene cloning
<i>DaGolS1</i> CDS R	CGCGGCCAATGGCGCAGTTATGAAC	gene cloning
<i>DaGolS2</i> CDS F	ATGGCCCCGATGGCGCTCAATG	gene cloning
<i>DaGolS2</i> CDS R	CGCGGCGGAGGGCGCGGGGAAG	gene cloning
<i>DaGolS1</i> qPCR F	AGATACCCTGTCGCTCACCT	qRT-PCR
<i>DaGolS1</i> qPCR R	TGGACTCGTTTCGACCAAACC	qRT-PCR
<i>DaGolS2</i> qPCR F	GACGTGTACAAGCCCATCCC	qRT-PCR
<i>DaGolS2</i> qPCR R	TGAACGACCTTCACCTTGCC	qRT-PCR
<i>DaEF1a</i> qPCR F	TTTGTCCACTGCTACACTCGTGGT	qRT-PCR
<i>DaEF1a</i> qPCR R	TCGAAGGCTGACGGACATAACCAA	qRT-PCR
<i>DaIRIP</i> qPCR F	AAGGGACCAACAACAATGTC	qRT-PCR
<i>DaIRIP</i> qPCR R	GACACCCAGTTACGATAGTG	qRT-PCR
<i>DaDhn</i> qPCR F	GCATGCACAGTGTAATGGGCTGAA	qRT-PCR
<i>DaDhn</i> qPCR R	AAGGCGGAATACATGACGCTGACA	qRT-PCR
<i>SODec2</i> qPCR F	TGGGCAACTCGCAGATCGCC	qRT-PCR
<i>SODec2</i> qPCR R	GCCCTGGCTTGAGCCCAGAG	qRT-PCR
<i>CATB</i> qPCR F	GCTTGCTTTCTGCCAGCGATAAT	qRT-PCR
<i>CATB</i> qPCR R	AAATAGTTTGGGCAAGACGGTGC	qRT-PCR
<i>OsGolS2</i> CDS F	ATGATGGGGCCGAACGTG	gene cloning
<i>OsGolS2</i> CDS R	CTACGCGGCGGAGGGCG	gene cloning
<i>OsGolS2</i> RT F	GTTCGAAGCCATGGAGGTTCA	RT-PCR
<i>OsGolS2</i> RT R	AGCTATCAAGACTTGACGGCG	RT-PCR
<i>OsRab16b</i> RT F	ACAAGGGCAACAACCACCAG	RT-PCR
<i>OsRab16b</i> RT R	GCTTGCAATGGCATCACAAAG	RT-PCR
<i>OsDREB1A</i> RT F	AGG CCG TCG AGG ACT TCT T	RT-PCR
<i>OsDREB1A</i> RT R	TAGTAGCTCCAGAGTGGAGT	RT-PCR
<i>OsUbiquitin</i> F	ATGCAGATCTTTGTGAAGACATTG	RT-PCR
<i>OsUbiquitin</i> R	TTACTGACCACCACGGAGGC	RT-PCR
<i>Hph</i> probe F	ATGAAAAAGCCTGAACTCACC	Southern blot probe
<i>Hph</i> probe R	CTATTCCTTTGCCCTCGG	Southern blot probe

Table 2. The accession numbers of GolS homologs used in phylogenetic analysis.

Clade	Family (Subfamily)	Species	Gene ID	Accession number	
Lycophyte	Selaginellaceae	<i>Selaginella moellendorffii</i>	SmGolS1	XP_002979905	
			SmGolS2	XP_002965247	
			SmGolS3	XP_002986570	
Eudicot	Brassicaceae	<i>Arabidopsis thaliana</i>	AtGolS1	NP_182240	
			AtGolS2	NP_176053	
			AtGolS3	NP_172406	
			AtGolS4	NP_176250	
			AtGolS5	NP_197768	
			AtGolS6	NP_567741	
			AtGolS7	NP_176248	
			AtGolS10	OAO91933	
			<i>Brassica napus</i>	BnGolS1	XP_013691918
				BnGolS2	XP_009113368
	BnGolS4	XP_013725304			
	BnGolS5	XP_022547051			
	BnGolS6	XP_013701350			
	Fabaceae	<i>Glycine max</i>	GmGolS1	XP_003554564	
			GmGolS2A	NP_001341795	
GmGolS2B			NP_001343305		
Monocot	Musaceae	<i>Musa acuminata</i>	MaGolS1	XP_009411607	
			MaGolS1L1	XP_009383943	
	Poaceae (Pooideae)	<i>Deschampsia antarctica</i>	DaGolS1	MK286464	
			DaGolS2	MK286465	
		<i>Triticum aestivum</i>	TraesCS4B02G179300	A0A1D5XM82	
			TraesCS4D02G180800	A0A1D5XYD6	
			TraesCS4A02G125300	W5DZ91	
			TaGolS1	BAF51565	
			TaGolS2	BAF51566	
			TraesCS2A02G095800	-	
			TraesCS7B02G493200	--	
			TraesCS2A02G094500	A0A1D5TQ22	
			TraesCS2B02G109700	W5BPL7	
			TraesCS2B02G109800	W5BND2	
			TraesCS2B02G109400		
			TraesCS2D02G093000	A0A1D5TR42	
		TraesCS2A02G094700	A0A1D5TR42		
		<i>Triticum dicoccoides</i>	TRIDC4BG034310	-	
			TRIDC4AG017380	-	
			TRIDC2AG011270	-	
			TRIDC2AG010850	-	
			TRIDC2BG013270	-	
	TRIDC2AG010870		-		
	<i>Triticum urartu</i>	TRIUR3_29787	M8A600		
		TRIUR3_15601	T1MD70		
		TRIUR3_24943	M7Z717		
		TRIUR3_24942	M7ZGF6		

(Continue) Table 2. The accession numbers of GolS homologs used in phylogenetic analysis.

			TRIUR3_33493	M7Z2E0
		<i>Aegilops tauschii</i>	F775_28615	M8CK84
			F775_09025	M8AXM8
			F775_25352	-
			F775_29160	NIQRD9
			F775_25855	M8C9J6
			F775_25473	M8BEI8
		<i>Hordeum vulgare</i>	HORVU4Hr1G052450	M0UZW7
			HORVU2Hr1G015720	A0A287H5W7
			HORVU2Hr1G015700	A0A287H5V9
		<i>Brachypodium distachyon</i>	BdGolS1	XP_003558046
			BdGolS2	XP_003559767
	<i>Poaceae</i> (<i>Panicoideae</i>)	<i>Zea mays</i>	ZmGolS1	NP_001105748
			ZmGolS2	NP_001105750
			ZmGolS4	ONM60967
		<i>Sorghum bicolor</i>	SbGolS2	XP_002467954
		<i>Setaria italic</i>	SiGolS1	XP_004984579
			SiGolS2	XM_004958687
	<i>Poaceae</i> (<i>Oryzeae</i>)	<i>Oryza brachyantha</i>	OB07G32460	J3MPB4
			OB03G25060	XP_006649987
		<i>Leersia perrieri</i>	LPERR07G23740	A0A0D9X346
			ORGLA03G0145400	I1PAQ8
		<i>Oryza punctata</i>	OPUNC07G25320	A0A0E0LQ01
			OPUNC03G14690	A0A0E0KCY9
		<i>Oryza meridionalis</i>	OMERI07G23550	A0A0E0EGE1
		<i>Oryza glumipatula</i>	OGLUM07G26920	A0A0E0APH9
			OGLUM03G15340	A0A0D9Z6F4
		<i>Oryza sativa Indica Group</i>	BGIOSGA023687	A2YQ42
			BGIOSGA012500	A2XG08
		<i>Oryza nivara</i>	ONIVA07G26690	A0A0E0I5X4
			ONIVA03G16170	A0A0E0GLL2
		<i>Oryza sativa Japonica Group</i>	OsGolS1	XP_015628833
			OsGolS2	XP_015645380
		<i>Oryza rufipogon</i>	ORUFI07G27890	A0A0E0QCZ6
			ORUFI03G15800	A0A0E0NU89
		<i>Oryza barthii</i>	OBART07G26950	A0A0D3GV71
			OBART03G15180	A0A0D3FHS2
		<i>Oryza glaberrima</i>	ORGLA03G0145400	I1PAQ8
	ORGLA07G0213400		I1QDE2	

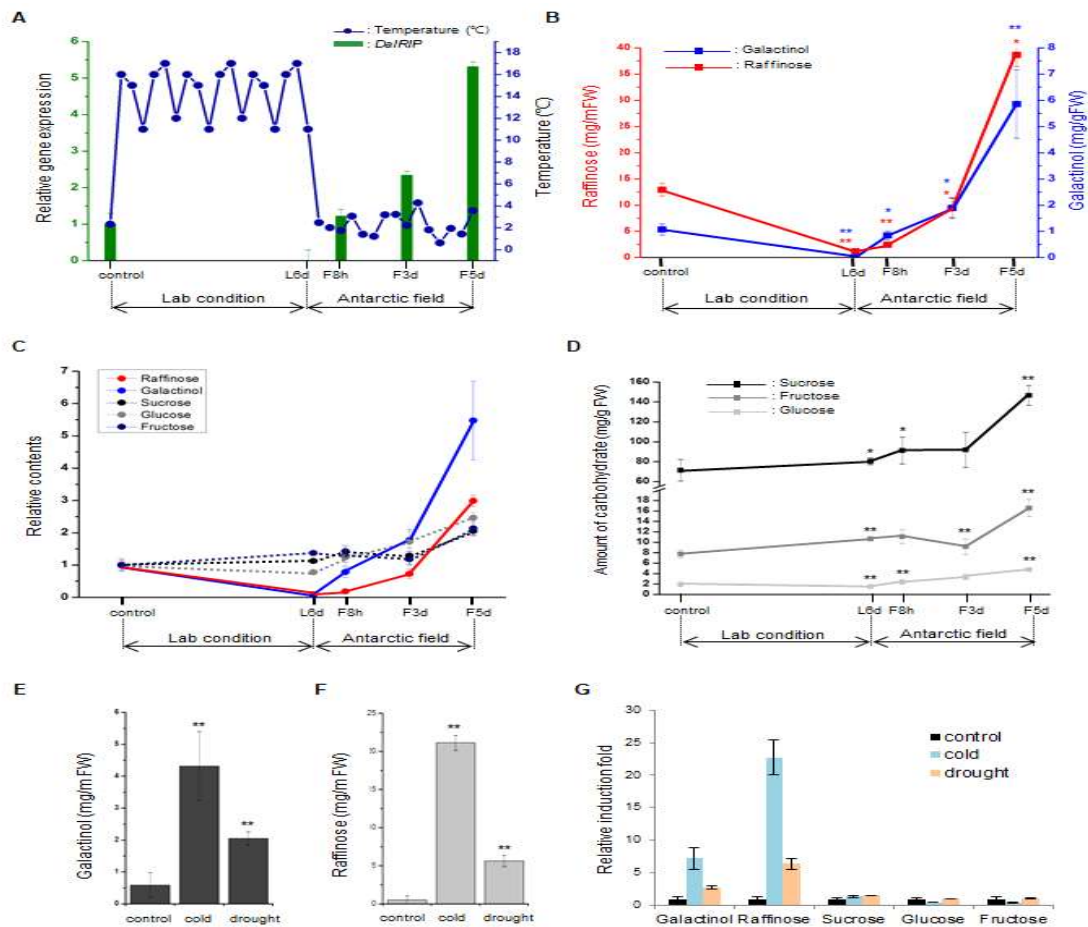


그림 1. Amounts of galactinol and raffinose were increased in response to cold and drought stress in *Deschampsia antarctica*. (A) Expression patterns of DaIRIP in response to different growth conditions. *Deschampsia antarctica* plants were collected from Antarctica field (control) and incubated in the laboratory conditions for 6 d (L6d), after which the plants were transferred back to Antarctic field for 8 h (F8h), 3 d (F3d) and 5 d (F5d). Total RNAs were isolated from each sample and used for real-time qRT-PCR analysis with a gene-specific primer set (Table 1). The relative expression of DaIRIP was normalized to that of DaEF1a that was served as an internal control. Data are means \pm SD (* P < 0.05, ** P < 0.01, Student's t-test) from three biologically independent experiments. Changes in the temperatures in the Antarctic field and laboratory conditions are indicated. (B) Galactinol and raffinose content in *D. antarctica* plants grown in the Antarctic field and laboratory conditions. Data are means \pm SD (* P < 0.05, ** P < 0.01, Student's t-test) from three biologically independent experiments. (C) Relative levels of galactinol, raffinose, sucrose, glucose and fructose in response to different growth conditions in *D. antarctica* plants. Relative sugar levels indicate the measured values of each growth point divided by the value of field control. (D) Amounts of sucrose, glucose and fructose in *D. antarctica* plants grown in the Antarctic field and laboratory conditions. Data are means \pm SD (* P < 0.05, ** P < 0.01, Student's t-test) from three biologically independent experiments. (E, F) Galactinol and raffinose content in the laboratory-cultured *D. antarctica* plants before (control) and after cold (4°C for 10 d) and drought (air dried on the filter paper at 15°C for 4 d) treatments. (G) Relative carbohydrate levels in the laboratory-cultured *D. antarctica* under stress conditions. Relative levels of each carbohydrate indicate the measured values of each stress treatment divided by the value of normal condition.

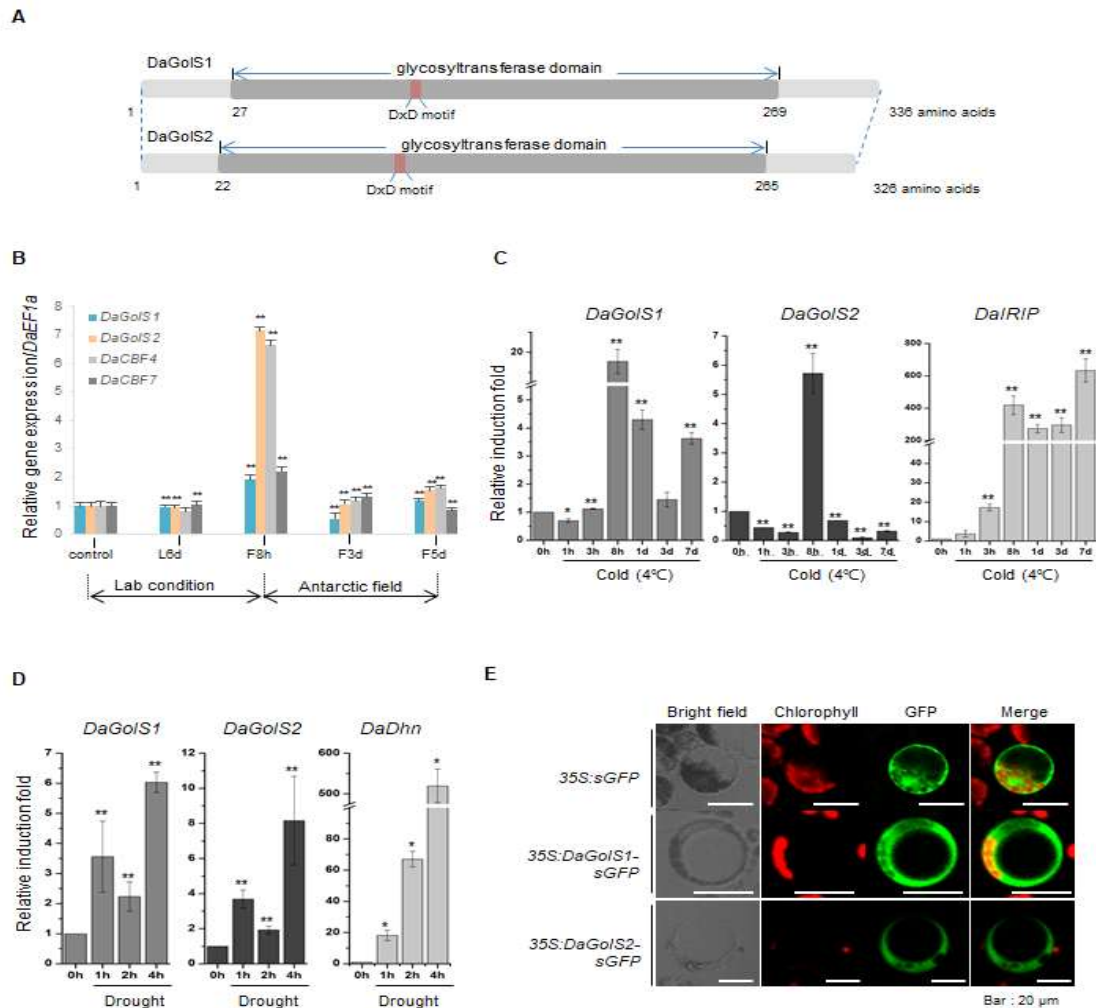


그림 2. Identification and characterization of DaGolS1 and DaGolS2 in *Deschampsia antarctica*. (A) Schematic structures of predicted DaGolS1 and DaGolS2. The glycosyltransferase domain is shown as a dark gray bar and the DxD catalytic motif is marked as a red bar. (B) Expression patterns of DaGolS1 and DaGolS2 in response to the Antarctica field and laboratory conditions in *D. antarctica* plants. Real-time qRT-PCR analysis was performed as described in Fig. 1A. DaCBF4 and DaCBF7 were used as positive controls for cold stress. Relative expression level of each gene was normalized to that of DaEF1a. Data are means \pm SD (* P < 0.05, ** P < 0.01, Student's t-test) of three biologically independent experiments. (C, D) Expression profiles of DaGolS1 and DaGolS2 in response to cold and drought stress. Laboratory-cultured 3-week-old *D. antarctica* seedlings were subjected to cold (4°C for 10 d) and dehydration (air dried on the filter paper at 15°C for 4 d) treatments, and total RNAs prepared from the treated tissues were analyzed by qRT-PCR using gene-specific primer sets (Table 1). DaIRIP and DaDhn were used as positive controls for cold and drought stress, respectively. Relative expression level of each gene was normalized to that of DaEF1a. Data are means \pm SD (* P < 0.05, ** P < 0.01, Student's t-test) of three biologically independent experiments. (E) Subcellular localization of DaGolS1 and DaGolS2. The 35S:DaGolS1-sGFP and 35S:DaGolS2-sGFP fusion constructs were transfected into the protoplasts prepared from mature leaves of *D. antarctica*. The fluorescent signals of the expressed proteins were visualized by fluorescence microscopy. sGFP was used as a cytosolic marker protein. Bars = 20 μ m.

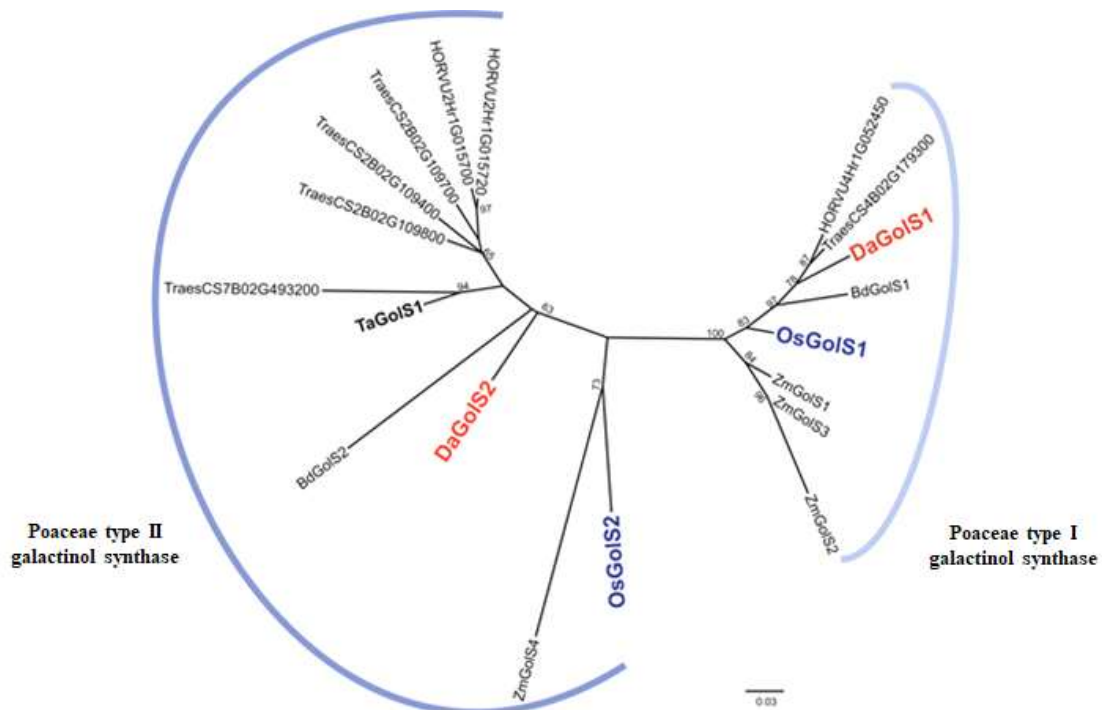


그림 3. Phylogenetic analysis of GolS proteins. The full-length amino acid sequences of DaGolS1, DaGolS2 and GolS homologs from Poaceae monocot crops were retrieved from the GenBank database and proofread. Phylogenetic trees were constructed from the data sets by the neighbor-joining method based on the Jones-Taylor-Thornton (JTT) matrix-based model using MEGA7 software. The tree is drawn to scale with branch lengths in the same units, as those of the evolutionary distances were used to infer the phylogenetic tree. The evolutionary distances were computed using the JTT matrix-based method. The Poaceae monocot GolS members are divided into two different groups, type I and type II.

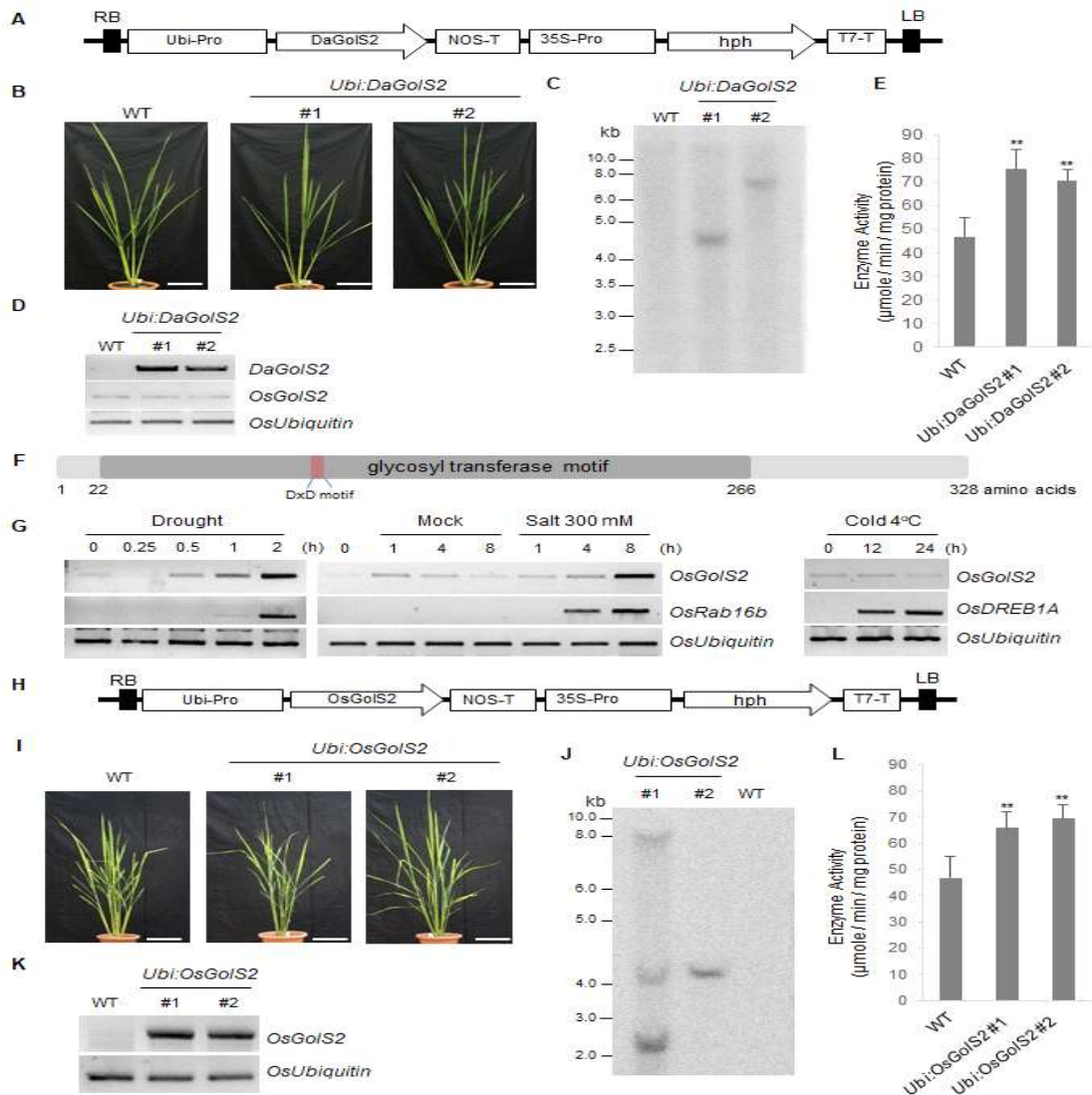


그림 4. Molecular characterizations of DaGolS2- and OsGolS2-overexpressing transgenic rice plants. (A) Schematic representation of a DaGolS2-overexpressing binary vector construct. RB, right border; pUbi, maize ubiquitin promoter; NOS-T, NOS terminator; 35S-Pro, CaMV 35S promoter; hph, hygromycin phosphotransferase; T7-T, T7 terminator; LB, left border. (B) Morphology of 2-month-old wild-type (WT) and T4 Ubi:DaGolS2 (independent lines #1 and #2) transgenic rice plants grown under long-day condition (16 h light and 8 h dark). Bars = 20 cm. (C) Genomic Southern blot analysis. Total leaf genomic DNA was extracted from the wild-type and T4 Ubi:DaGolS2 (lines #1 and #2) rice plants. The DNA was digested with HindIII and hybridized to a ³²P-labeled hygromycin B phosphotransferase (hph) probe. (D) RT-PCR analysis of the wild-type and T4 Ubi:DaGolS2 (lines #1 and #2) rice plants to examine the overexpression of DaGolS2 and endogenous transcript level of OsGolS2. OsUbiquitin was used as a loading control. (E) GolS enzyme activity assay. Crude extracts (50 μg protein) of the wild-type and Ubi:DaGolS2 (lines #1 and #2) rice seedlings were used in each assay. Activity was calculated as μmol Pi released per mg protein per min. Data are means ± SD (**P < 0.01, Student's t-test) of three biologically independent experiments. (F) Schematic structure of deduced OsGolS2. The glycosyltransferase and DxD motifs are indicated as dark gray and red bars, respectively.

(G) Expression patterns of OsGolS2 in response to drought (0, 0.25, 0.5, 1 and 2 h), high salinity (300 mM for 0, 1, 4 and 8 h) and low temperature (4°C for 0, 0.5, 1, 2 and 5 d) stress in rice plants. OsRab16b was used as a positive control for drought and salt treatments, whereas OsDREB1A was used as a positive control for cold stress. (H) Schematic representation of an OsGolS2-overexpressing binary vector construct. RB, right border; pUbi, maize ubiquitin promoter; NOS-T, NOS terminator; 35S-Pro, CaMV 35S promoter; hph, hygromycin phosphotransferase; T7-T, T7 terminator; LB, left border. (I) Morphology of 2-month-old WT and T3 Ubi:OsGolS2 (independent lines #1 and #2) rice plants grown under long-day condition. Bars = 20 cm. (J) Genomic Southern blot analysis. Total leaf genomic DNA was extracted from the wild-type and T3 Ubi:OsGolS2 (lines #1 and #2) rice plants, digested with HindIII, and hybridized to a 32P-labeled hygromycin B phosphotransferase (hph) probe. (K) RT-PCR analysis of the wild-type and T3 Ubi:OsGolS2 (lines #1 and #2) rice plants to examine the expression of OsGolS2 transcript. OsUbiquitin was used as a loading control. (L) GolS enzyme activity assay. Crude extracts (50 µg protein) of the wild-type and Ubi:OsGolS2 (lines #1 and #2) rice seedlings were used in each assay. Activity was calculated as µmol Pi released per mg protein per min. Data are means ± SD (**P < 0.01, Student's t-test) of three biologically independent experiments.



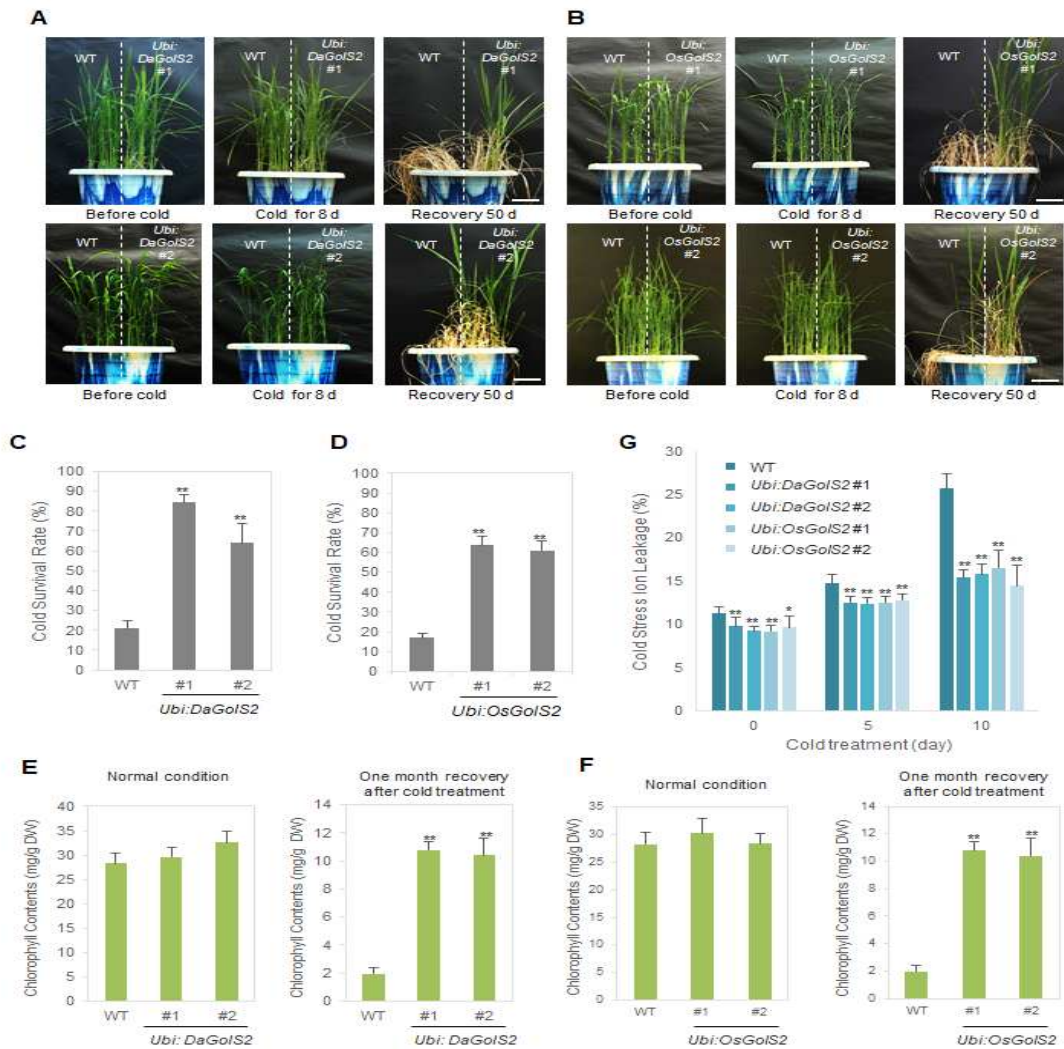


그림 5. Increased tolerance of DaGolS2- and OsGolS2-overexpressing transgenic rice plants in response to cold stress. (A, B) Cold stress phenotypes of the wild-type, Ubi:DaGolS2 and Ubi:OsGolS2 transgenic rice plants. Light-grown, 5-week-old wild-type, T4 Ubi:DaGolS2 (lines #1 and #2) and T3 Ubi:OsGolS2 (lines #1 and #2) rice plants were transferred to a cold room at 4°C for 8 d, after which the plants were recovered at 28°C for 50 d. Bars = 15 cm. (C, D) Survival rates of the wild-type, Ubi:DaGolS2 and Ubi:OsGolS2 plants in response to cold stress. Data are means \pm SE ($n \geq 6$ biologically independent experiments; >30 plants were used in each assay, ** $P < 0.01$, Student's t-test). (E, F) Total leaf chlorophyll content of the wild-type, Ubi:DaGolS2 and Ubi:OsGolS2 plants before and after cold treatments. The amounts of leaf chlorophyll (chlorophyll a + chlorophyll b) of mock- (before cold) and cold-treated plants were determined 1 month after recovery from cold stress. Data are means \pm SE ($n \geq 3$ biologically independent experiments; >10 plants were used in each assay, ** $P < 0.01$, Student's t-test). (G). Electrolyte leakage analysis of the wild-type, Ubi:DaGolS2 and Ubi:OsGolS2 rice plants in response to cold stress. Electrolyte leakage analysis was conducted using 8-day-old wild-type, T4 Ubi:DaGolS2 (lines #1 and #2) and T3 Ubi:OsGolS2 (lines #1 and #2) seedlings at different time points before and after cold (4°C) treatment (0, 5 and 10 d). Data are means \pm SD ($n = 3$ biologically independent experiments; >12 plants of each genotype were used in each experiment, ** $P < 0.01$, Student's t-test).

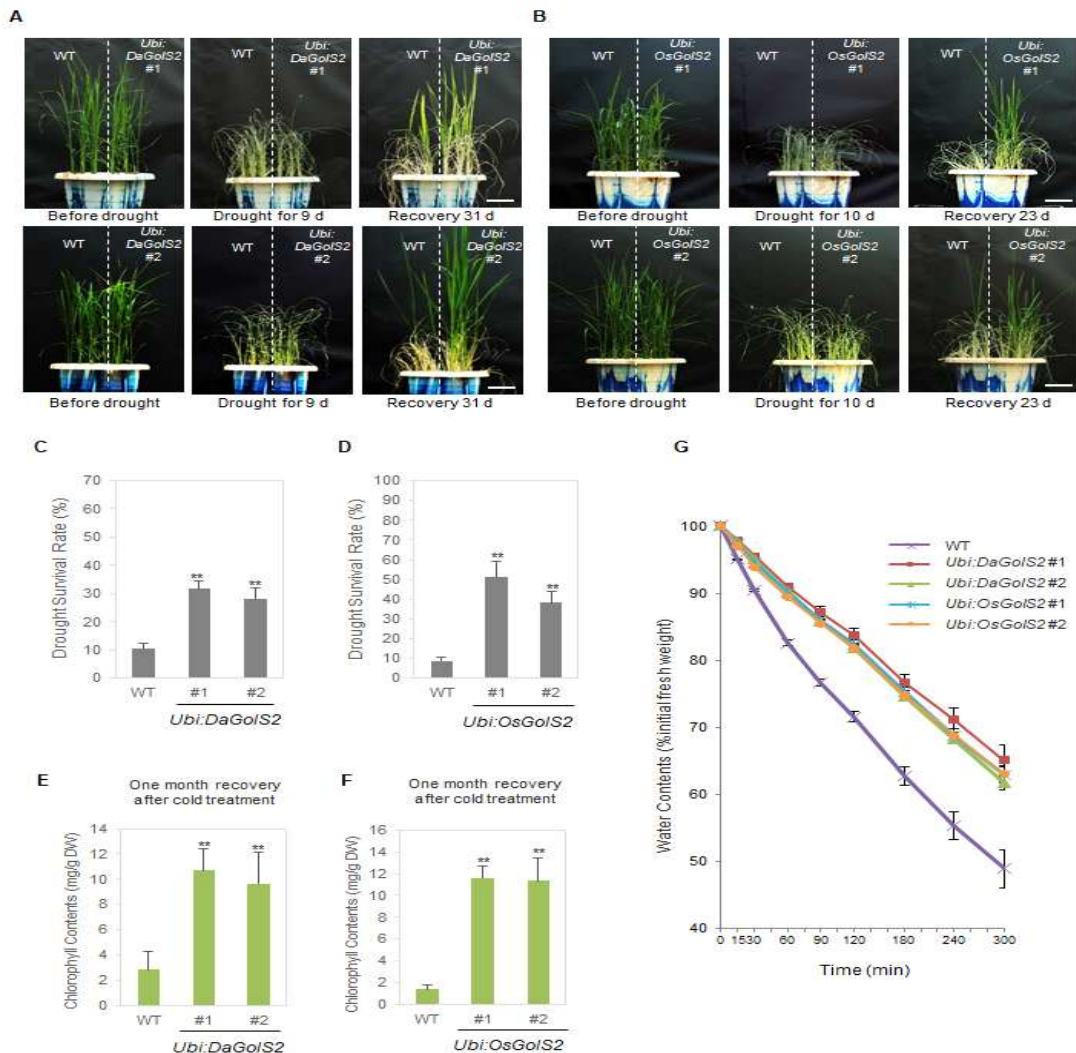


그림 6. Increased tolerance of DaGolS2- and OsGolS2-overexpressing transgenic rice plants in response to drought stress. (A, B) Drought stress phenotypes of the wild-type, Ubi:DaGolS2 and Ubi:OsGolS2 transgenic rice plants. Light-grown, 5-week-old wild-type, T4 Ubi:DaGolS2 (lines #1 and #2) and T3 Ubi:OsGolS2 (lines #1 and #2) rice plants were grown at 28°C without watering for 9/10 d. Dehydration-treated plants were re-watered and their growth patterns were monitored for 23 d after re-watering. Bars = 15 cm. (C, D) Survival rates of the wild-type, Ubi:DaGolS2 and Ubi:OsGolS2 rice plants in response to drought stress. Data are means \pm SE ($n \geq 6$ biologically independent experiments; >30 plants were used in each assay, ** $P < 0.01$, Student's t-test). (E, F) Total leaf chlorophyll content of the wild-type, Ubi:DaGolS2 and Ubi:OsGolS2 plants after drought treatment. The amounts of leaf chlorophyll (chlorophyll a + chlorophyll b) of drought-treated plants were determined 1 month after recovery from the stress. Data are means \pm SE ($n \geq 3$ biologically independent experiments; >10 plants were used in each assay, ** $P < 0.01$, Student's t-test). (G) Water loss rates of detached leaves. The leaves of 5-week-old wild-type, T4 Ubi:DaGolS2 (lines #1 and #2) and T3 Ubi:OsGolS2 (lines #1 and #2) rice plants were detached, and their fresh weights were measured at the indicated time points. The rate of water loss was calculated as the percentage of initial fresh weight of the detached leaves. Data are means \pm SD ($n = 3$ biologically independent experiments; >6 plants of each genotype were used in each experiment).

제 3-5절 극지미세조류 KNF0032의 얼음결합 단백질 CmIBP의 분리 및 이들의 활성검증과 과다발현 형질전환체의 표현형 분석

Type II Ice-Binding Proteins Isolated from an Arctic Microalga Are Similar to Adhesin-Like Proteins and Increase Freezing Tolerance in Transgenic Plants

요약

극지방의 미세 조류 얼음 결합 단백질(IBP)은 저온 적응에 중요하지만 게놈 전체 수준에서 잘 연구되어 있지 않았다. 북극 녹조류 *Chloromonas* sp. KNF0032의 전사체 연구를 통해 남극 빙설조류 *Chloromonas* sp. CCMP681에서만 발견된 바 있는 IBP와 상동의 6개 *Chloromonas* IBP 유전자 (CmIBPs)를 발견하였다. 이 유전자는 다수의 엑손/인트론 구조 및 프로모터 지역의 저온-반응성 시스-요소로 구성되었고 저온에서 풍부하게 발현되었다. CmIBP 단백질의 구조는 평행한 b-시트 및 반복된 TXT 모티프를 갖는 b-솔레노이드 형태를 갖는 것으로 예측되었다. 이러한 특징은 곤충 및 식물의 얼음결합 단백질과 유사했다. 3개의 대표적인 CmIBP (CmIBP1, CmIBP2 및 CmIBP3)의 생물학적 기능을 시험관 내 분석 및 형질전환 식물시스템을 사용하여 시험 하였다. CmIBP1은 시험관 내 및 형질전환 식물 모두에서 가장 효과적인 얼음 재결정화 억제(IRI) 활성을 가지며, CmIBP2 및 CmIBP3가 뒤 따랐다. 순화되지 않은 조건에서 자란 모든 형질전환 식물은 동결내성이 있었고, 특히 35S::CmIBP1 식물이 가장 효과적이었다. 저온 적응 후, 35S::CmIBP2 식물만이 약간 증가된 동결 내성을 나타냈다. 본 연구에서 검증한 형질전환 식물체 내의 동결내성 향상은 미세조류의 얼음결합 단백질이 기능을 보인 첫 사례에 해당한다.

Abstract

Microalgal ice-binding proteins (IBPs) in the polar region are poorly understood at the genome-wide level, although they are important for cold adaptation. Through the transcriptome study with the Arctic green alga *Chloromonas* sp. KNF0032, we identified six *Chloromonas* IBP genes (CmIBPs), homologous with the previously reported IBPs from Antarctic snow alga CCMP681 and Antarctic *Chloromonas* sp. They were organized with multiple exon/ intron structures and low-temperature-responsive cis-elements in their promoters and abundantly expressed at low temperature. Structurally, the CmIBPs were predicted to have b-solenoid forms with parallel b-sheets and repeated TXT motifs. The biological

functions of three representative CmIBPs (CmIBP1, CmIBP2 and CmIBP3) were tested using in vitro analysis and transgenic plant system. CmIBP1 had the most effective ice recrystallization inhibition (IRI) activities in both in vitro and transgenic plants, and CmIBP2 and CmIBP3 had followed. All transgenic plants grown under nonacclimated condition were freezing tolerant, and especially 35S::CmIBP1 plants were most effective. After cold acclimation, only 35S::CmIBP2 plants showed slightly increased freezing tolerance. This is the first report to improve the freezing tolerance of plants by heterologously expressed microalgal IBP proteins.

서론

호냉성 미세 조류는 0-4°C에서 견딜 수 있고 15°C 미만의 온도에서 최적의 성장 속도를 보이는 종으로 설원, 빙하 및 극지방 담수호수에서 생존한다(Eddy 1960; Morita 1975; Morgan-Kiss et al. 2006). 어느점 근처의 온도에서 얼음 결정의 성장은 세포를 둘러싼 물 분자의 반복적인 동결-해동주기와 그에 따른 탈수로 인해 세포에 심각한 물리적 손상을 일으킨다(Davies 2014; Bar Dolev et al. 2016). 이러한 손상으로부터 세포를 보호하기 위해, 호냉성 미세 조류는 얼음 결정의 평면에 결합하여 성장을 억제하는 얼음 결합 단백질(IBP)을 생성한다(Raymond et al. 2009; Raymond and Morgan-Kiss 2013; Raymond 2014; Raymond and Morgan-Kiss 2017). IBP는 다양한 호냉성 및 저온기피 생물체에서 발견되며 주요 활동에 따라 부동액 단백질(AFP), 부동액 당단백질(AFGP) 또는 얼음 재결정 억제 단백질(IRIP)이라고 불린다(Davies 2014; Bar Dolev et al. 2016). 남극 어류(Gonzalez-Aguero et al. 2013) 및 곤충(Graether and Sykes 2004; Duman 2015; Schrodinger 2015)의 AFP 또는 AFGP에 대한 연구에서 주로 설명된 열 히스테리시스(TH) 활동은 녹는점을 어느점의 아래로 낮추는 기능을 하여, 더 낮은 온도에서 얼음 결정이 형성 되도록 한다(Davies 2014; Bar Dolev et al. 2016). 일반적으로 동결내성 식물(Verbruggen et al. 2009; Middleton et al. 2012; Bredow et al. 2017) 및 녹조류(Raymond et al. 2009; Morgan-Kiss 2013; Raymond 2014; Raymond and Morgan-Kiss 2017)의 IRIP는 얼음재결정 억제(IRI) 활성을 보인다. 얼음재결정 억제활성은 0°C 미만의 온도에서 얼음 결정의 성장을 억제하여 세포 외 또는 세포 내 얼음 결정을 작게 유지하는 능력을 의미한다(Davies 2014; Bar Dolev et al. 2016).

극지방 미세 조류의 IBP는 단백질 서열 상동성에 근거하여 2가지 유형으로 나뉜다(Raymond and Morgan-Kiss 2013). 제 1형 IBP는 *Chlamydomonas raudensis* UWO241(Raymond and Morgan-Kiss 2013), *Chloromonas brevispina*(Raymond 2014) 및 *Chlamydomonas* ICE-MDV(Raymond and Morgan-Kiss 2017)의 세 가지

남극 녹조에서 발견되었다. X-선 결정학 연구에 따르면 제 1형 IBP에는 알 수 없는 기능의 도메인 3494(DUF3494)가 포함되어 있으며 α -나선과 나란히 β -솔레노이드 접힘을 형성한다. 제 1형 IBP는 박테리아와 곰팡이에서 프로테오 박테리아, 규조류 및 코페포드로 수평 유전자 전달(HGT) 이벤트를 통해 확산된다는 가설이 인정되고 있다(Sorhannus 2011). 제 1형 IBP와는 달리, 제 2형 IBP는 남극 빙설조류 CCMP681 및 남극 *Chloromonas sp.*(Raymond et al. 2009; Jung et al. 2016)에서 발견되었다. 이들 단백질은 TH 활성보다 IRI 활성에서 더 효과적인 것으로 밝혀졌으며, 단백질 구조 예측법을 통해 반복된 Thr 잔기로 구성된 얼음 결합 부위의 평행한 평면을 갖는 β -솔레노이드 구조를 가짐을 보여 주었다(Raymond et al. 2009; Jung et al. 2016). 제 2형 IBP 단백질은 상동성의 단백질이 단백질 데이터베이스에서 발견되지 않았지만, β -솔레노이드의 표면 구조는 다년생 식물(Middleton et al. 2012), 동결 저항성 곤충(Graether and Sykes 2004), 또는 남극 해양 박테리아, *Marinomonas primoryensis*(Guo et al. 2017)에서 유래한 다른 IBP에서 발견되었다. 이 연구에서 북극 담수로부터 분리된 호냉성 녹색 조류 *Chloromonas sp.* KNF0032의 IBP 단백질의 in vivo 기능과 형질전환 식물체 내의 동결내성 기능을 검증하였다.

2. 재료 및 방법

가. 변형 배양 조건 및 온도 처리

본 연구에 사용된 모든 KNF 및 KSF 균주는 북극 다산 과학기지과 남극 세조과학기지 근처에서 수집된 것으로 KCCPM(KOPRI Culture Collection)에서 얻은 것이다. 균주는 연속 광($30 \mu\text{mol m}^{-2} \text{s}^{-1}$) 조건의 8°C 성장 챔버에서 BBM 배지에서 유지되었다. 온도 반응 실험을 위해 *Chloromonas sp.* KNF0032 배양체를 지수 단계에서 $4, 8, 12$ 및 16°C 로 설정된 성장 챔버로 옮기고, 온도에 대한 그들의 반응을 평가하기 위해 일주일 동안 동일한 광 조건 하에서 성장시켰다.

나. RNA 추출 및 RNA 시퀀싱 라이브러리 구축

온도 처리 후, 조류 세포를 원심 분리를 통해 수확하고 액체 질소로 동결하여 막자사발에서 세포를 깬다. 총 RNA는 RNeasy Mini Kit(Qiagen, Hilden, Germany)를 사용하여 이전에 설명한 제조사의 프로토콜에 따라 RNase free DNase Set(Qiagen)과 함께 추출하였다(Cho et al. 2018). 간략하게, 2 개의 생물학적 복제물을 제조하였다. RNA의 완전성 및 농도는 Bioanalyzer system(RIN > 6; Agilent Technologies, Palo Alto, CA, USA) 및 Qubit RNA Broad-range Assay Kit(Life Technologies, Carlsbad, CA, USA)을 사용하여 측정하였다. RNA-Seq 라이브러리를 구축하기 위해, 각 샘플의 $1.5 \mu\text{g}$ 총 RNA를 TruSeq RNA Sample Prep Kit v2 (Illumina, San

Diego, CA, USA)에 인풋으로 사용하였다. 라이브러리는 Bioanalyzer와 Illumina 가이드 라인에 따라 라이브러리 정량적 PCR 방법을 사용하여 정량화되었다. 정량 후, 이들은 동일한 비율로 다중화되었다. 시퀀싱은 MiSeq 시퀀서 시스템(Illumina)에서 수행되었고, 총 4.98Gb(28M paired-end reads)의 시퀀싱 데이터가 생성되었다. 시퀀싱 데이터는 NCBI 시퀀스 아카이브에서 수탁 번호 SRP139417로 제공된다.

다. 드 노보 어셈블리, 주석 및 차등 발현 유전자 분석

De novo 어셈블리는 CLC Genomics Workbench v7.5 소프트웨어(Qiagen)를 사용하여 수행되었다. 낮은 품질의 시퀀스(quality score < 0.001, ambiguity < 2 bp)를 제외하고 어댑터, 길이가 짧고(length > 50 bp), 중복되는 시퀀스를 제거하여 판독 값을 필터링하였다. 결과 판독 값은 다음 파라미터(word size = 20, bubble size = 50, length > 200 bp)를 사용하여 조립하였다. 총 29,631 개의 조립된 contig(Average: 894 bp, N50: 1,246 bp)에 임계값이 1e-3인 non-redundant protein database를 이용하여 BLASTX 검색을 수행했다(Altschul et al. 1990). Gene ontology mapping 및 annotation (E-value < 1x10⁻¹⁰)은 Blast2GO 플랫폼을 사용하여 수행되었다(Conesa and Gotz 2008). 발현 값을 전사 수준에서 FPKM(Fragments Per Kilobase of transcript per Million)으로 추정하였다. CLC Genomics Workbench에서 Baggerley의 테스트로 쌍별 비교를 수행하고, 차등 발현된 유전자를 컷오프 값(corrected P-value FDR < 0.05, difference ≠ 0)을 사용하여 확인하였다.

라. 유전자 구조 및 실리코 프로모터 분석

Chloromonas sp. KNF0032 균주의 genomic DNA는 DNeasy Plant Mini Kit(Qiagen)를 사용하여 추출하였다. 6 개의 CmIBP 유전자의 코딩 영역은 유전자-특이적 프라이머(Table 1)를 사용하여 PCR에 의해 증폭되었고, 서열은 Sanger-sequencing에 의해 확인되었다. 6 개의 CmIBP 유전자의 upstream 1 kb는 KNF0032의 부분 게놈 어셈블리 데이터로부터 얻은 CmIBP 게놈 contig의 서열 정보에 기초하여 예측되었다. 게놈 contig는 query로서 CmIBP 서열을 사용하는 Blast 검색에 의해 발견되었고, 발견된 서열은 Sanger 시퀀싱에 의해 확인되었다. 확인된 게놈 DNA 서열을 사용하여 TRANSFAC에서 시스 조절 요소 분석을 확인하였다(Matys et al. 2003). CmIBP1, CmIBP2 및 CmIBP3의 엑손-인트론 경계는 게놈 DNA 서열(GenBank ID : MN011069 - MN011071)과 코딩 서열 영역을 비교하여 결정하였다.

마. 정량적 실시간 역전사 PCR 분석

Quantitative realtime Reverse Transcriptase PCR (qRT-PCR) 분석을 위해 *Chloromonas* sp. KNF0032 균주를 8°C에서 지수 성장 단계로 배양한 다음, 0°C 챔버

에서 1, 2, 3 및 7 일 동안 배양하였다. 주어진 조건 하에서 수확된 세포로부터 총 RNA를 추출하고 RNeasy Plant Mini Kit(Qiagen)를 사용하여 정제하였다. cDNA는 2 μ g의 총 RNA로부터 합성되었다. 유전자-특이적 프라이머는 Table 1에 열거되어 있다. 내부 대조군 유전자를 선택하기 위해, 모든 전사체 데이터에서 평균 발현값 (> 30) 및 변동 계수(CV < 0.1)에 기초하여 10 개의 후보 유전자를 선택 하였다 (Table 2; Gonzalez-Aguero et al. 2013). 10 개의 후보 유전자의 증폭 효율 테스트를 수행하여 (n = 3) 발현 안정성을 테스트하고, 임계주기(Ct) 값을 RefFinder 분석 도구와 함께 사용 하였다(Xie et al. 2012). RefFinder는 geNorm, NormFinder, BestKeeper 및 delta CT 방법의 결과를 종합적으로 비교하여 contig5016 (methyltransferase-like protein)과 contig3 9(elongation factor-like protein)라는 두 가지 참조 유전자를 제안하였다. 생물학적 3 반복체로 분석을 수행하고, 평균 발현 값을 표준편차로 플롯팅하였다.

바. 단백질의 계통 발생학적 분석 및 구조적 예측

균주의 계통 발생적 위치를 명확히 하기 위해, 18S rDNA 서열을 5 개의 균주 Chloromonas sp.에서 JO2 및 SS17HR(Verbruggen et al. 2009)을 사용하여 증폭하고 이들 서열을 NCBI(GenBank ID : MH400028-MH400032)에 등록하였다. Chloromonas, Chlamydomonas 및 Chlorella 종의 추가 서열을 NCBI에서 다운로드하였다. MAFFT ver.7(Katoh et al. 2017)을 사용하여 다중 정렬을 수행하였고, 계통 발생학적 분석은 MEGA7의 Neighbor-Joining 방법에 의해 수행되었다. 부트스트랩 분석은 1,000 번의 반복 실험으로 수행되었으며 support value(> 50)는 지점에 표시 하였다. Chloromonas KNF0032로부터의 6 개의 CmIBP 유전자의 코딩 서열을 RNA-Seq 분석을 위해 준비된 cDNA로부터 증폭시키고, 이들 서열을 NCBI 데이터베이스에 등록하였다(GenBank ID : MH400035-MH400040). Chloromonas의 나머지 4개 균주에서 상동성 IBP 단백질을 스크리닝하기 위해, 상기한 바와 같이 1 주일 동안 2 $^{\circ}$ C에서 배양된 샘플의 cDNA를 제조하였다. CDS의 전체 서열을 증폭시키기 위한 유전자-특이적 프라이머를 사용하여 RT-PCR을 수행하였다(Table 1). PCR 증폭은 다음 조건으로 수행되었다 : 98 $^{\circ}$ C에서 30초, 98 $^{\circ}$ C에서 10초, 53 $^{\circ}$ C에서 30초, 72 $^{\circ}$ C에서 50초 30번 반복, 72 $^{\circ}$ C 7분의 최종 연장.

이 연구에 사용된 다른 IBP 서열은 NCBI 데이터베이스에서 다운로드하였다. 웹 서버에서 PhyML 3.0을 사용하여 최대 가능성 트리를 구성하고(Lefort et al. 2017), IBP 및 Adhesin 계통 구축을 위한 substitution model은 Akaike Criterion 기준에 따라 WAG + G + F 모델에서 선택되었습니다. 부트스트래핑은 1,000 번 반복되었고 계통발생수는 FigTree v1.4.3(Rambaut 2012)으로 시각화하였다.

추정 신호 펩티드는 SignalP 4.1(Petersen et al. 2011)로 예측되었고 도메인 구조는 SMART(Letunic and Bork 2018)를 사용하여 예측되었습니다. 이황화 결합은 SCRATCH 단백질 예측 변수 (<http://scratch.proteomics.ics.uci.edu/index.html>, 2019년 8월 26일에 마지막으로 액세스)에서 DIpro에 의해 예측되었습니다. Phyre2 웹서버(Kelley et al. 2015)로 3차 구조를 예측하고 PyMol 프로그램 (Schrodinger 2015)으로 그래픽뷰를 조작했다.

사. 재조합 CmIBP의 클로닝 및 단백질 정제

CmIBP 재조합 단백질을 구축하기 위해, 신호 펩티드 및 막 횡단 서열을 제외한 CmIBP1, CmIBP2 및 CmIBP3의 코딩 영역을 유전자 특이적 프라이머로 증폭하고, pET22b 발현 벡터(Novagen, Gibbstown, NJ, USA)에 클로닝하였다. 서열을 확인한 후, BL21 (DE3) 세포 (Novagen)로 형질 전환하였다. 단백질을 정제를 위해, 형질 전환체 대장균 세포를 흡광도(600nm)에서 0.6이 될 때까지 37°C에서 키우고, 0.5 mM IPTG를 첨가하고 15°C에서 48시간 동안 단백질을 유도하고, 원심분리하여 세포를 수확하였다. 펠렛을 cOMplete, EDTA-free protease inhibitor cocktail (Roche, Darmstadt, Germany)이 첨가된 용해 완충액 (50mM Tris-HCl, 300mM NaCl, 20mM Imidazole, 10 % glycerol 및 0.1 % NP40, pH 8.0)에 재현탁시켰다. 얼음에서 35% 진폭으로 10분 동안 초음파 처리(Vibra-Cell; Sonics, Newtown, CT, USA)에 의해 세포벽을 파괴하고, 용해물을 4°C에서 30분 동안 21,130g에서 원심분리하였다. 가용성 단백질을 Ni-NTA agarose column (Qiagen)을 사용하여 정제하였다. 단백질을 용리 완충액(50mM Tris-HCl, 300mM NaCl 및 250mM Imidazole, pH 8.0)으로 용출시킨 후, Amicon Ultracel-10K 원심분리 필터(Merck Millipore, Cork, Ireland)를 사용하여 50mM Tris-HCl(pH 8.0)과 교환하였다. 정제된 단백질의 용해도는 anti-His6 항체(# sc-53073; Santa Cruz Biotechnology)를 사용하여 웨스턴 블롯 분석으로 검증되었고, 정제된 단백질은 후속 분석에 즉시 사용되었다.

아. CmIBP를 과발현하는 형질전환 애기장대 식물의 생성

CmIBP의 open reading frame 서열은 Myc epitope, cauliflower mosaic virus(CaMV) 35S promoter, octopine synthase gene terminator를 갖는 pEarleygate203 벡터(Earley et al. 2006)에 클로닝 되었으며, 확인된 구축물은 *Agrobacterium tumefaciens* GV3101 균주로 형질전환한 뒤, 애기장대(Col-0)에 floral-dip 방법을 사용하여 형질전환을 수행하였다(Zhang et al. 2006). Hygromycin (50 mg ml⁻¹)을 함유하는 0.5 Murashige 및 Skoog 배지에서 형질전환된 식물을 선택하였다. 각 유전자마다 10개의 독립적인 동형 접합체 라인(T3)이 생성되었고, 이 중 두 라인을 이용하여 데이터를 생성하였다. 이종 CmIBP 유전자의 과발현은 RT-PCR에 의해 확인

되었다. 비저온적응(NA, non-cold acclimated) transgenic plants는 $120 \mu\text{mol m}^{-2} \text{s}^{-1}$ 의 빛 강도로 16 : 8 h (밝은 : 어두운)주기에서 3-4주 동안 배양하고 저온적응(CA, cold acclimate) 식물체는 4°C 로 옮기고 2일 더 성장시켜, IRI 및 이온 누출 분석에 활용하였다(Bredow et al. 2017).

자. 애기장대 형질전환 식물체에서 단백질 추출

NA 및 CA 조건 하에서 성장한 2주차 형질전환 식물체를 proteasome inhibitor MG132(0.1mM)가 첨가된 액체배지에 2일간 처리하여 단백질 추출에 사용하였다. 준비된 유식물체(30mg FW)를 막자와 막자사발을 사용하여 액체 질소에서 분쇄하고 cOmplete EDTA-free protease inhibitor cocktail(Roche, Basel, Swiss)이 보충된 $200\mu\text{l}$ 의 native protein extraction buffer($10\text{mM Tris-HCl pH } 7.5$ 및 25mM NaCl)를 넣었다. 모든 추출 단계는 4°C 또는 얼음에서 수행되었다. 샘플을 충분히 섞어준 뒤 10분 동안 얼음 위에 유지시켰다. $13,523 \text{ g}$ 에서 10분 동안 원심분리한 후, 상층액을 새로운 튜브로 옮기고 DC 단백질 분석법(Bio-Rad, Hercules, CA, USA)을 사용하여 총 단백질 농도를 계산 하였다.

아포플라스트(Apoplast) 단백질 추출을 위해, NA 조건에서 4주간 성장시킨 야생형(Col-0) 및 $35\text{S}::\text{CmIBP1 \#6}$ 식물을 이용하였다(Haslam et al. 2003). 야생형 및 $35\text{S}::\text{CmIBP1 \#6}$ 식물의 30개의 신선한 잎을 수확하고, Tween20의 0.05%로 헹구고 탈이온수(DW)로 2회 헹구었다. 세척된 잎을 cOmplete EDTA-free protease inhibitor cocktail을 포함하는 $200\mu\text{l}$ native protein extraction buffer($10\text{mM Tris-HCl pH } 7.5$ 및 25mM NaCl)에 완전히 담그고, desiccator에서 70kPa 에서 3분 동안 진공-침투시켰다. 남은 완충액을 제거한 후, 파라필름으로 말아서 침윤된 잎을 20ml 주사기에 넣고 주사기를 원심분리를 위해 50ml 튜브에 넣었다. 아포플라스트 단백질을 원심분리(15분, 55g)에 의해 수확하고, 세포질 단백질을 상기 기재된 총 단백질을 추출하는 방식으로 나머지 잎으로부터 추출하였다.

아포플라스트 및 세포질 단백질을 농축시키고, 이들의 완충액을 Amicon Ultracel-10K 원심분리 필터(Merck Millipore, Cork, Ireland)를 사용하여 $50\text{mM Tris-HCl(pH } 8.0)$ 로 교환하였다. 단백질 농도는 DC 단백질 분석에 의해 측정되었다. 야생형 및 CmIBP1 \#6 형질전환 식물체($2.18\text{-}2.29 \text{ g FW}$)의 총 아포플라스트 단백질을 각각 $144 \mu\text{g}(4.8 \mu\text{g 잎})$ 및 $126 \text{ mg}(4.2 \mu\text{g 잎})$ 으로 수득하였다. 야생형 및 CmIBP1 \#6 형질전환 식물체의 세포질 단백질은 $620 \mu\text{g}$ 및 $930 \mu\text{g}$ 을 각각 수득하였으며, 따라서 아포플라스트 단백질은 각각 총 잎 단백질의 0.77% 및 0.45%로 측정되었다. 제조된 단백질을 웨스턴 블롯팅 및 IRI 분석에 즉시 사용하였다.

차. 웨스턴 블롯팅 분석

총 단백질 샘플(50 ug)을 12% SDS-PAGE 겔에 로딩하고 제조업체의 프로토콜에 따라 PVDF막에 블롯팅하였다. 형질전환 식물에서 CmIBP 단백질을 검출하기 위해 anti-cMyc(#sc-40; Santa Cruz Biotechnology)를 사용하였고, alkaline phosphatase(AP)-conjugated anti-mouse(#sc-2008; Santa Cruz Biotechnology) 항체를 이차항체로 하여 BCIP/NBT 기질(#sc-24981; Santa Cruz Biotechnology)의 비색 분석에 사용하였다. 아포플라스트 및 세포질 단백질 샘플(50 ug)을 12% SDS-PAGE 겔에 로딩하고 PVDF 막에 블롯팅하였다. anti-tubulin(T5168, Sigma, USA) 및 anti-PR2(pathogenesis related protein-2; AS12 2366, Agrisera, Sweden)를 사용하여 세포질 및 아포플라스트 단백질 분획을 각각 검출하였다. AP-conjugated anti-mouse(#sc-2008; Santa Cruz Biotechnology) 항체가 이차 항체로 사용되었다. BCIP/NBT 기질(#sc-24981; Santa Cruz Biotechnology)을 비색 분석에 사용하였다.

카. IRI 분석

제조된 단백질을 동결 저장 없이 IRI 분석에 사용 하였다. 단백질 샘플 및 대조군으로서 BSA 0.1 mg/ml의 농도를 사용하였다. IRI 분석은 다음 조건 하에서 BX2000 광학 현미경(Olympus, Tokyo, Japan)에 장착된 LINKAM THMS600 heating-freezing stage(Linkam Scientific, Tadworth, UK)를 사용하여 수행하였다 : -20°C 5분, -10°C 5분, -6°C 30분. 각 샘플에서 얼음 입자의 크기(면적 3.2 mm²)는 ImageJ(Schneider et al. 2012)를 사용하여 25-75 %에 해당하는 얼음 입자 크기의 평균값으로 측정되었습니다. 각각의 샘플에서 얼음 입자의 크기는 완충액 대조군의 크기와 관련하여 표현 하였다. 모든 실험을 3회 반복하였다. 차이는 Student 's t-test (P <0.05)로 확인하였다.

타. 전해질 누출 분석

야생형과 형질전환 식물체를 4주 동안 키운 뒤 전해질 누출을 분석하였다. 식물체당 2개의 성숙한 로제트 잎(7번 및 8번 잎)을 절단하고 250ul의 DW를 함유하는 유리 튜브에 넣었다. 각 식물의 10개의 잎을 제조하고 냉동 처리(-6°C)를 다음과 같이 수행 하였다. 유리 튜브를 순환 수조(EYELA, Tokyo, Japan)에 -1°C에서 1시간 동안 놓은 다음 얼음핵을 첨가하였다. 물이 얼기 시작한 후, 온도를 1°C/h 속도로 감소시키고, 온도가 -6°C에서 1시간 처리한 뒤, 순환 수조에서 튜브를 제거하고, 튜브를 어둠 속에서 밤새 4°C에서 해동시켰다. 모든 잎 및 DW를 추가 10ml의 DW를 갖는 50ml 유리 튜브로 옮기고 실온에서 밤새 교반하였다. 초기 전도도(Ci)를 측정 한 뒤, 멸균 처리 후 최종 전도도(Cf)를 측정하였다. Ci/Cf 비율은 백분율로 표시했다. 상대 이온 누출은 각 조건에서 야생형의 이온 누출에 대한 백분율로 표시했다. 이 실험

험은 4번 수행되었으며 차이는 Student's t-검정 ($P < 0.05$)에 의해 검증하였다.

3. 결과

가. 호냉성 조류 *Chloromonas* sp. KNF0032의 저온 반응 유전자

Chloromonas sp.의 저온적응 메커니즘을 이해하기 위해 균주를 최적 성장 온도 (8°C), 저온 (4°C) 및 고온 (12°C 및 16°C)에서 전사체 분석을 하여, 저온-반응성 유전자를 조사하였다. 전사체 분석을 통해, 29,631 개의 contig가 de novo assembly 되었고, 이 중 5,153 개의 contig가 정상 성장 온도(8°C)와 비교하여 차별적인 유전자 발현을 나타냈다. 저온 적응에 관여하는 유전자를 검색하기 위해, 4°C 에서 발현이 증가된 62개의 annotated contig(4°C vs 8°C , FDR < 0.05 , 차이 > 0)를 선택하였다. 표적 유전자 그룹을 더욱 좁히기 위해, 우리는 4 , 8 , 12 및 16°C 에서 FPKM 값에 기초하여 계층적 군집화를 수행하였고, 최종적으로 저온에서 구체적으로 발현된 29 개의 contig를 선택하였다. 이러한 저온 특이 발현 contig 중 6개(21%)가 IBP로 annotation 되었다(Figure 1a). 6개의 contig의 게놈 단편은 Sanger Sequencing에 의해 확인되었다. 그들은 2.7-4.6 kb 길이의 게놈 영역에 걸쳐 9-12 개의 엑손을 가졌다. 따라서 이 연구에서 CmIBP1 ~ CmIBP6이라고 명명하였다 (Table 3). CmIBP1-CmIBP6의 FPKM 수준은 저온에서 증가했다. 특히, CmIBP1, CmIBP2 및 CmIBP3는 저온에서 현저하게 향상된 발현을 나타냈다. 균주 KNF0032를 0°C 로 옮기고 1주일 동안 노출시키면, 3개의 CmIBP 유전자 모두 8°C 에서보다 더 많이 유도되었고, 이들의 발현은 2일째에 최대에 도달하였다(Figure 1b, c). 또한, CmIBP1, CmIBP2 및 CmIBP3는, CBF(C-repeat/DRE binding factor), ABI (ABA insensitive) 및 WRKY와 같은 식물 전사인자에 대한 프로모터에서 결합 모티프를 가지고 있다(Figure 2; Table 3). 이 결과는 CmIBPs가 동결 온도 적응 반응에 참여하고, 이러한 유전자의 전사는 현화식물의 신호전달 과정과 유사하게 cis-element 결합으로 유도되는 positive regulator으로써 cold stress signaling pathway에 관여할 수 있다는 것으로 추정된다. 이들 IBP가 다른 *Chloromonas* 균주에도 존재함을 확인하기 위해, 제 2형 IBP 특이적 프라이머를 사용하여 8개의 추가 KCCPM 균주에 대해 RT-PCR 수행하였다. 결과적으로, KNF0032의 동일한 클레이드의 균주는 CmIBP 상동성 유전자를 가졌고 이것이 저온에서 강하게 유도되었음을 확인했다(Figure 3).

나. Thr이 풍부한 CmIBP는 곤충 AFP와 구조가 유사

CmIBP 유전자는 각각 분자량이 32-39kDa인 단백질로 340-415개 아미노산(aa)으로 구성된다. 모든 CmIBP는 SignalP 4.0 서버에 따라 N terminus에 분비 신호 펩티드 (18-31 aa)가 있는 것으로 예측되었고, CmIBP3 및 CmIBP5는 추가적인 막 통과 도메인을 갖는 것으로 예측되었다(Figure 4). CmIBP 서열의 BLAST 검색은

Chlamydomonas sp. CCMP681(이하, CCMP681 IBP; Raymond et al. 2009) 및 Antarctic *Chloromonas* sp.(이하 ChloroIBP; Jung et al. 2016; Figure 4)의 2형 IBP와 매우 높은 유사성을 나타냈다. 계통수는 CmIBP, 4개의 CCMP681 IBP 및 ChloroIBP가 박테리아, 곰팡이, prasinophytes, 녹조류 및 규조류에서 DUF3494 도메인을 포함하는 이전에 보고된 제 1형 IBP와 구별되는 클레이드를 형성했음을 분명히 보여주었다(Figure 5). CmIBP는 50-53% 소수성 아미노산(V, I, L, MF, W, C, GP 및 A)으로 구성되었지만 가장 풍부한 아미노산은 친수성 아미노산 Thr(T; 12-15%)이다. CmIBP는 곤충 AFP 및 MpIBP와 비교하여 서열 유사성이 적지만, 둘 다 12-15%의 높은 비율로 나타나는 친수성 Thr 잔기를 가졌다(Table 1). 특히, Cys 잔기는 CmIBP의 약 5%를 점유하였고, Cys가 없는 MpIBP와 비교하여 현저하게 더 큰 부분을 나타낸다(Table 4). CmIBP1은 7개의 이황화 결합을 갖고, CmIBP2 및 CmIBP3은 각각 6 및 5개의 이황화 결합을 가졌다. CmIBP 및 CCMP681 IBP의 다중 정렬은 이전에 TXT 모티프(Raymond et al. 2009)로 지정된 7개의 반복된 Thr-X-Thr 잔기(Figure 6의 모티프 I-VII)의 존재를 나타내었다. 모든 TXT 모티프는 일정한 간격으로 배치되었지만 첫 번째 및 두 번째 TXT 모티프는 세 번째에서 일곱 번째 모티프(25-29 잔기)와 약간 떨어져 있었다(40-43 잔기). CmIBP3, CmIBP4, CmIBP5 및 CmIBP6의 TXT 모티프 서열의 Thr 잔기는 완전히 보존되었지만, CmIBP1 및 CmIBP2 TXT 모티프의 첫 번째 또는 세 번째 잔기에 해당하는 Thr 잔기는 종종 Lys (K), Asp (D), Ile (I) 또는 Ser (S)으로 치환되었다(Figure 6). CmIBP의 3 차 단백질 구조는 β -솔레노이드 형태(템플릿 단백질 ID : c5gkd, 리아제)를 갖는 것으로 예측되었다. N-말단의 부분 서열은 예측할 수 없었지만 일부 모티프 (III-VII)는 평행한 β -시트를 포함하는 평평한 평면을 형성하는 것으로 성공적으로 입증되었다(Figure 6). 다중의 CmIBP TXT 모티프로 구성된 평행한 β -시트는 곤충 AFP, 식물 IRIP 및 남극 박테리아 IBP (MpIBP)의 것과 유사하지만, 서열 유사성 및 크기가 낮다(Bar Dolev et al. 2016).

다. 재조합 CmIBP1은 5배 더 높은 IRI 활성을 나타낸다

저온에서 6개의 CmIBP의 향상된 유전자 발현에 기초하여, 추가 기능 연구의 표적으로서 3개의 대표적인 단백질, CmIBP1, CmIBP2 및 CmIBP3을 선택하였다. 그들의 결빙 결합 활성을 조사하기 위해, 재조합 CmIBP1, CmIBP2 및 CmIBP3을 제조하고(Figure 7) TH 및 IRI 활성을 시험하였다. 3개의 재조합 단백질의 TH 활성은 무시할 만했다(TH <0.1 °C, 0.1 mg/ml). 그러나, 더 작은 얼음 결정의 형성을 열역학적으로 선호하는 억제된 얼음 재결정화의 파라미터인 IRI 활성(Knight and DeVries 2009)은 완충제(50mM Tris-HCl, pH 8.0) 및 BSA 0.1 mg/ml 대조군에 비해 CmIBP 재조합 단백질에서 매우 높았다(Figure 8). IRI 활동의 강도는 주어진 온도

와 지속 시간에서 얼음 입자가 얼마나 작게 유지되는지에 의해 비교될 수 있으므로, CmIBP의 상대적인 얼음 입자 크기는 얼음 입자의 평균 크기에서 베퍼에 이르기까지 계산되었습니다. 재조합 단백질 샘플을 6°C에서 30분 동안 처리할 때, 얼음 입자의 성장이 0.2, 0.1 및 0.05 mg/ml의 CmIBP1에서 강하게 억제되었고, 큰 입자는 관찰되지 않았다. 0.1 mg/ml 농도에서 동일한 처리 조건 하에서 CmIBP2는 지연된 얼음 재결정으로 중간 정도의 IRI 효과를 보인 반면, CmIBP3는 효과가 가장 낮았으며, 다각형 경계를 갖는 더 큰 얼음 입자를 생성했다(Figure 8a). CmIBP1과 CmIBP2의 상대적인 얼음 입자 크기는 각각 베퍼 컨트롤보다 5 배와 2.5 배 더 강했지만 CmIBP3의 효과는 더 낮았다(Figure 8b). 또한, 0.2 mg/ml에서의 재조합 CmIBP1의 IRI 효과는 얼음 재결정화 과정을 거의 완전히 억제하였다(Figure 8). 이러한 결과는 CmIBP의 IRI 강도가 상이하며, 이는 얼음 결합에 대한 CmIBP의 친화도의 차이와 관련될 수 있음을 시사한다.

라. CmIBP는 형질전환 애기장대의 동결 내성을 향상시킨다

재조합 CmIBP는 IRI 활성을 보였으므로, CmIBP가 식물 발현 시스템에서 동결 손상으로부터 세포를 보호함으로써 동결 내성을 개선할 것이라는 가설을 세웠다. 이 가설을 시험하기 위해, CmIBP1, CmIBP2 및 CmIBP3을 과발현하는 형질전환 애기장대 식물(이하, 35S::CmIBP1, 35S::CmIBP2 및 35S::CmIBP3)을 생성하였다. 정상적인 성장 조건에서 형질전환 식물과 야생형 식물(Col-0) 사이에는 유의한 표현형 또는 발달 차이가 관찰되지 않았다(Figure 9). 특히, 전이 유전자의 RNA 전사체는 저온 순응에 상관없이 꾸준히 축적되었지만, 단백질 발현 수준은 각 유전자에서 상이하였다(Figure 7). CmIBP1의 경우, 단백질 발현은 MG132 및 저온 적응 처리에 의해 덜 영향을 받지만, CmIBP2 및 CmIBP3 단백질은 영향을 미치지 않았다. CmIBP2의 단백질 밴드의 강도는 차가운 순응 후에 더 강한 것으로 보이지만, 식물 세포에서 이중 단백질이 어떻게 조절되는지는 여전히 명확하지 않다.

저온 적응의 내인성 효과를 제거하기 위해, 비적응(NA) 형질전환 식물을 IRI 및 전해질 누출 분석에 사용하였다. 35S::CmIBP 형질전환 식물의 단백질 추출물에 대한 IRI 실험은 야생형과 비교하여 모든 시험된 형질전환 식물에서 얼음 입자의 크기가 현저히 감소하거나 얼음 결정화가 상당히 지연됨을 보여 주었으며(Figure 9a), 이것은 35S::CmIBP 식물체의 세포추출물이 IRI 활성을 갖는다는 것을 보여준다. IRI 활성은 아포플라스트 및 세포질 영역 모두에서 나타났다(Figure 9b).

35S::CmIBP 형질전환 식물이 동결 조건 하에서 세포 보호 능력을 향상시켰는지 여부를 조사하기 위해, 이들 식물에 대해 전해질 누출 분석을 수행하였다(Figure 9c). NA 조건 하에서 재배된 모든 형질전환 식물에서 전해질 누출 수준은 감소했으

며 특히 35S::CmIBP1 식물의 전해질 누출 수준은 다른 식물보다 현저히 낮았습니다. 저온 순응 후, 35S::CmIBP2 식물의 전해질 누출은 야생형 식물보다 더 감소하였으나, 다른 형질전환 식물의 효과는 유사하였다. 모든 결과는 미세 조류 CmIBP가 아포플라스트로 분비되고 형질전환 식물체에서 얼음 결정의 성장을 억제함으로써 동결 내성의 증가에 기여한다는 것을 시사한다.

4. 논의

제 2형 IBP는 CCMP681, 남극 녹조류(Raymond et al. 2009)에 의해 처음 보고 되었으며, 이 연구는 제 2형 IBP에 대한 중요한 생물학적 정보를 제공한다. 이 연구에서, RNA-Seq의 de novo 어셈블리를 통해, 우리는 제 2형 IBP 유전자가 게놈에 존재하고 이들 중 다수가 온도-의존적 방식으로 조절됨을 입증하였다. 우리는 약 10 IBPs의 코딩 서열을 annotation하고 그들의 유전자 및 프로모터 지역을 분석했다. IBP 이소형의 활성을 비교하여 IRI 활성이 다르다는 것을 발견했으며, 식물에서 발현될 때 동결 내성을 개선하여 농업 응용 분야에서 가능한 역할을 제안할 수 있음을 실험적으로 입증했다. 따라서 이것은 지금까지 제 2형 IBP에 대한 가장 포괄적이고 진보된 연구이다.

가. 2형 IBP의 희귀성

녹조류는 얼음이 많은 환경에서 자주 발견되지만 미세 조류 유형 1 및 2 IBP는 자연적으로 고르지 않게 분포되어 있는 것으로 보인다. 유형 1과 달리 유형 2 IBP는 몇 가지 계통으로 구성된 특정 분류군에서 발견된다. 북극 및 남극 해빙 군집내 발현 분석에 의하면 IBP 전사체의 대략 90%가 제 1형 IBP로 분류되고 DUF3494 도메인을 가졌다(Uhlig et al. 2015). 또한, 제 2형 IBP의 *Chloromonas* 종은 현재까지 북극에서 보고된 바가 없다. 최근, amplicon deep sequencing을 이용한 홍설 조류의 양극성 분산에 관한 연구에 따르면 KNF0032에 속하는 *Stephanospaherinia* clade에 해당하는 ITS2 OTU는 매우 드물게 발견되었다(전체 관독 값의 0.1%, 남극 관독 값의 0.066%; Segawa et al. 2019). 따라서, 그들의 동결권 바이오매스 비율은 낮기 때문에, 배양 농축을 통해서만 단일 균주로서 분리될 수 있는 것으로 여겨진다. 또한, 이전의 연구는 다른 눈 조류 그룹이 지배하는 홍설 조류에 중점을 두었기 때문에 샘플링 중에 이를 제외했을 수 있다. 이 연구에서 우리는 북극 및 남극 미세 조류 균주 모음인 KCCPM에서 얻은 *Chloromonas* 균주가 단일 클레이드를 형성하고 모두 북극과 남극 중 모두에서 2형 IBP를 함유하고 있음을 발견했다 (Figure 3). 따라서 다양한 미세 조류의 생태학적 상태와 서식지 분포를 정의하기 위해서는 샘플링 환경을 다양화하고 극지 육상 및 담수 조류에 대한 더 많은 데이터를 수집해야 한다.

나. KNF0032 게놈 내 다수의 CmIBP 존재 및 이들의 생리적 역할

다수의 IBP 유전자가 호냉성 조류에서 발견되며 이것은 자연에서 생물학적 중요성을 나타낸다. 현재까지 최소 9개 이상의 IBP 단백질이 *C. brevispina*(Raymond 2014), *Chlamydomonas raudensis* UWO241(Raymond and Morgan-Kiss 2013) 및 *Chlamydomonas* sp. ICE-MDV(Raymond and Morgan-Kiss 2017)에서 발견되었고, 이것의 발현은 영하의 온도에서 증가했다. *Chloromonas* sp. KNF0032 계놈은 최소 6개의 CmIBP 유전자가 존재하며, 이것의 프로모터에는 저온반응성 cis-element를 가지고 있어 이것이 저온반응 조절기작에 관여할 것이라 추정된다. CmIBP는 온도에 대한 전사 반응이 상이하였고, 각각의 유전자는 상이한 재결정화 억제 활성을 가졌다. 예를 들어, 6개의 CmIBP 중 CmIBP1, 2 및 3은 저온에 대해 현저히 높은 전사 반응을 나타냈다. 세 단백질 모두 RI 활성을 보였지만 그 정도는 달랐으며 CmIBP1은 정제된 재조합 단백질과 형질전환 식물의 세포 추출물 모두에서 가장 강한 RI 활성을 나타냈다. 따라서, 3개의 CmIBP는 *Chloromonas* sp. KNF0032는 자연 환경에서 저온에 노출될 때, 전체적인 IRI 활성을 향상시키는데 관여할 것이다. 유사하게 온대성의 겨울을 나는 다년생 식물인 *Lolium perenne*은 4개의 IRI 단백질과 1개의 AFP를 가지고 있다(Lauersen et al. 2011). 이들 중 LpIRI2 또는 LpIRI3는 형질전환 애기장대의 동결 생존을 가장 잘 개선시킨다. 이들 2개의 LpIRI 유전자가 LpAFP와 함께 발현될 때, 형질전환 식물의 동결 내성이 더욱 향상되었으며(Bredow et al. 2017). 이들 유전자는 동결 조건 하에서 상이한 세포 활성 및 부가 효과를 갖는다는 것을 암시한다. 많은 극단적인 유기체가 계놈에 IBP 유전자를 가지고 있다는 것은 잘 알려져 있지만, IBP의 유전자 조절 기능에 대해 명확히 해야 할 것이 많이 있다. 따라서, IBP의 세포 기능에 대한 추가의 유전자 분석은 IBP가 극한의 생존을 위해 필수적인지 여부에 대한 답이 필요할 수 있다.

다. CmIBP-과발현 형질 전환 식물의 IRI 활성 및 동결 내성 표현형

CmIBP1, CmIBP2 및 CmIBP3은 분비 신호를 가지며, CmIBP3은 N terminus에 추가적으로 막 횡단 도메인을 갖는다. 따라서, 이들은 세포 외 공간으로부터 분비되거나 세포 주위의 얼음 결정의 성장을 억제하기 위해 막-고정된 것으로 추정된다. 유형 1 또는 유형 2 IBP를 갖는 많은 미세 조류에서 배양 배지에서 IRI 활성을 나타냈다. 특히, 유형 2 IBP는 CCMP681 및 남극 *Chloromonas* sp.의 배양 배지에서 검출되었다(Raymond et al. 2009, Jung et al. 2016). 최근의 연구에서 CCMP681 유형 2 IBP의 신호 펩티드와 mCherry 형광 단백질을 결합한 융합 단백질은 *Chlamydomonas reinhardtii* 발현 시스템의 세포에서 성공적으로 분비되었다(Molino et al. 2018). 그러나, 아포플라스트는 일반적으로 세포질보다 높은 동결점을 갖고 아포플라스트에서 얼음핵 생성이 시작되기 때문에 CmIBP가 세포 밖으로 분비되고 형질전환 식물의 아포플라스트 영역에서 기능하는지 여부를 확인해야 했다. 애기장대에서 분비 신

호가 없는 LpAFP의 과다발현은 동결 내성이 개선되지 않는다는 사실(Bredow et al. 2017)은 얼음 결합 단백질이 세포 외 영역에서 또는 세포질 영역에서 발현되는지가 식물의 동결 내성에 중요하다는 가설을 지지한다 (Bredow and Walker 2017). 우리의 분석에서, IRI 활성은 형질 전환 식물의 세포 형성 및 세포질 분획 둘 모두에서 관찰되었는데, 이는 CmIBP가 35S::CmIBP1 식물의 세포질 및 아포플라스트 영역 모두에 존재함을 시사한다. 이는 CmIBP의 분비 신호가 식물의 분비 신호와 다르고 분비 효율이 식물의 분비 효율보다 낮을 수 있기 때문일 수 있다. 이러한 분비 신호 펩티드가 고등 식물과 미세 조류간에 상호 교환 가능한지 여부를 확인하기 위해 더 많은 연구가 필요하다. 따라서, 형광 리포터와 CmIBP의 분비 신호 펩티드의 조합과 같은 실험 방법을 사용하여 세포 하 위치를 확인할 필요가 있다. 우리의 결과는 이중 CmIBP 단백질이 저온 적응 없이 유전자 변형 식물에서 동결 내성을 향상시킬 수 있다는 것을 암시한다. 또한 식물 세포에서 이중 단백질은 번역 후 수준에서 조절되는 것으로 보이며(Figure 7), 단백질 안정성, 단백질(또는 DNA)-단백질 상호 작용 및 세포 내 위치 등을 확인해야 한다. 그러나, 이중 CmIBPs 발현 시스템을 적용함으로써, 우리는 유전자 이식 식물 및 작물이 넓은 온도 범위에서 재배될 수 있고 서리에 대한 저항을 증가시킬 수 있다고 제안한다. IBP의 기능을 향상시키려면 세포에서 더 높은 IRI 활성으로 단백질 안정성을 높이는 방법을 개발해야 한다. 예를 들어, 애기장대에서 곤충 AFP가 발현 될 때 코딩 서열의 최적화는 식물 동결 온도의 감소에 긍정적인 영향을 미쳤으며(Huang et al. 2002), 얼음 결합 표면을 넓히기 위한 단백질 크기의 다양한 조작 방법이 IBP 활동 증가에 있어서 효과적이었다(Bar Dolev et al. 2016). 식물에서 세포 외 분비를 유도하는 신호 펩티드의 융합이 효과적일 것으로 예상된다. 따라서, CmIBP에 대해 최적화 된 유전자 공학적 방법을 사용함으로써 보다 개선된 동결 내성을 갖는 형질 전환 식물을 개발할 수 있을 것으로 기대된다.

5. 참고문헌

- 1) Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* 215: 403-410.
- 2) Bar Dolev, M., Braslavsky, I. and Davies, P.L. (2016) Ice-Binding Proteins and Their Function. *Annu Rev Biochem* 85: 515-542.
- 3) Bredow, M., Vanderbeld, B. and Walker, V.K. (2017) Ice-binding proteins confer freezing tolerance in transgenic *Arabidopsis thaliana*. *Plant Biotechnol J* 15: 68-81.
- 4) Bredow, M. and Walker, V.K. (2017) Ice-Binding Proteins in Plants. *Front Plant*

- 5) Cho, S.M., Lee, H., Jo, H., Lee, H., Kang, Y., Park, H., et al. (2018) Comparative transcriptome analysis of field- and chamber-grown samples of *Colobanthus quitensis* (Kunth) Bartl, an Antarctic flowering plant. *Sci Rep* 8: 11049.
- 6) Conesa, A. and Gotz, S. (2008) Blast2GO: A comprehensive suite for functional analysis in plant genomics. *Int J Plant Genomics* 2008: 619832.
- 7) Davies, P.L. (2014) Ice-binding proteins: a remarkable diversity of structures for stopping and starting ice growth. *Trends Biochem Sci* 39: 548-555.
- 8) Duman, J.G. (2015) Animal ice-binding (antifreeze) proteins and glycolipids: an overview with emphasis on physiological function. *J Exp Biol* 218: 1846-1855.
- 9) Earley, K.W., Haag, J.R., Pontes, O., Opper, K., Juehne, T., Song, K., et al. (2006) Gateway compatible vectors for plant functional genomics and proteomics. *The Plant Journal* 45: 616-629.
- 10) Eddy, B. (1960) The use and meaning of the term 'psychrophilic'. *Journal of Applied Microbiology* 23: 189-190.
- 11) Gonzalez-Aguero, M., Garcia-Rojas, M., Di Genova, A., Correa, J., Maass, A., Orellana, A., et al. (2013) Identification of two putative reference genes from grapevine suitable for gene expression analysis in berry and related tissues derived from RNA-Seq data. *BMC Genomics* 14: 878.
- 12) Graether, S.P. and Sykes, B.D. (2004) Cold survival in freeze-intolerant insects: the structure and function of beta-helical antifreeze proteins. *Eur J Biochem* 271: 3285-3296.
- 13) Guo, S., Stevens, C.A., Vance, T.D.R., Olijve, L.L.C., Graham, L.A., Campbell, R.L., et al. (2017) Structure of a 1.5-MDa adhesin that binds its Antarctic bacterium to diatoms and ice. *Sci Adv* 3: e1701440.
- 14) Haslam, B.R.P., Downie, A.L., Raveton, M., Gallardo, K., Job, D., Pallett, K.E., et al. (2003) The assessment of enriched apoplastic extracts using proteomic approaches. *Annals of Applied Biology* 143: 81-91.
- 15) Huang, T., Nicodemus, J., Zarka, D. G., Thomashow, M. F., Wisniewski, M,

- Duman, J. G. (2002) Expression of an insect (*Dendroides canadensis*) antifreeze protein in *Arabidopsis thaliana* results in a decrease in plant freezing temperature. *Plant Molecular Biology* 50(3): 333-344.
- 16) JJung, W., Campbell, R.L., Gwak, Y., Kim, J.I., Davies, P.L. and Jin, E. (2016) New Cysteine-Rich Ice-Binding Protein Secreted from Antarctic Microalga, *Chloromonas* sp. *PLoS One* 11: e0154056.
- 17) Katoh, K., Rozewicki, J. and Yamada, K.D. (2017) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in bioinformatics* bbx108, <https://doi.org/10.1093/bib/bbx108>.
- 18) Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. and Sternberg, M.J.E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* 10: 845.
- 19) Knight, C.A. and DeVries, A.L. (2009) Ice growth in supercooled solutions of a biological "antifreeze", AFGP 1-5: an explanation in terms of adsorption rate for the concentration dependence of the freezing point. *Phys Chem Chem Phys* 11: 5749-5761.
- 20) Lauersen, K.J., Brown, A., Middleton, A., Davies, P.L. and Walker, V.K. (2011) Expression and characterization of an antifreeze protein from the perennial rye grass, *Lolium perenne*. *Cryobiology* 62: 194-201.
- 21) Lefort, V., Longueville, J.-E. and Gascuel, O. (2017) SMS: Smart Model Selection in PhyML. *Molecular Biology and Evolution* 34: 2422-2424.
- 22) Letunic, I. and Bork, P. (2018) 20 years of the SMART protein domain annotation resource. *Nucleic Acids Research* 46: D493-D496.
- 23) Matys, V., Fricke, E., Geffers, R., Gößling, E., Haubrock, M., Hehl, R., et al. (2003) TRANSFAC®: transcriptional regulation, from patterns to profiles. *Nucleic Acids Research* 31: 374-378.
- 24) Middleton, A.J., Marshall, C.B., Faucher, F., Bar-Dolev, M., Braslavsky, I., Campbell, R.L., et al. (2012) Antifreeze protein from freeze-tolerant grass has a beta-roll fold with an irregularly structured ice-binding site. *J Mol Biol* 416: 713-724.

- 25) Molino, J.V.D., Carvalho, J.C.M., and Mayfield, S.P. (2018) Comparison of secretory signal peptides for heterologous protein expression in microalgae: Expanding the secretion portfolio for *Chlamydomonas reinhardtii*. Plos One. 13(2):e0192433
- 26) Petersen, T.N., Brunak, S., von Heijne, G. and Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nature methods 8: 785.
- 27) Rambaut, A. (2012) <http://tree.bio.ed.ac.uk/software/figtree/>.
- 28) Raymond, J.A. (2014) The ice-binding proteins of a snow alga, *Chloromonas brevispina*: probable acquisition by horizontal gene transfer. Extremophiles 18: 987-994.
- 29) Raymond, J.A., Janech, M.G. and Fritsen, C.H. (2009) Novel Ice-Binding Proteins from a Psychrophilic Antarctic Alga (Chlamydomonadaceae, Chlorophyceae)(1). J Phycol 45: 130-136.
- 30) Raymond, J.A. and Morgan-Kiss, R. (2013) Separate origins of ice-binding proteins in antarctic chlamydomonas species. PLoS One 8: e59186.
- 31) Raymond, J.A. and Morgan-Kiss, R. (2017) Multiple ice-binding proteins of probable prokaryotic origin in an Antarctic lake alga, *Chlamydomonas sp.* ICE-MDV (Chlorophyceae). J Phycol 53: 848-854.
- 32) Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. Nature methods 9: 671.
- 33) Schrodinger, LLC (2015) The PyMOL Molecular Graphics System, Version 1.8.
- 34) Segawa, T., Matsuzaki, R., Takeuchi, N., Akiyoshi, A., Navarro, F., Sugiyama, S., et al. (2019) Bipolar dispersal of red-snow algae. Nat. Communication. 9:3094
- 35) Sorhannus, U. (2011) Evolution of antifreeze protein genes in the diatom genus *fragilariopsis*: evidence for horizontal gene transfer, gene duplication and episodic diversifying selection. Evol Bioinform Online 7: 279-289.
- 36) Uhlig, C., Kilpert, F., Frickenhaus, S., Kegel, J.U., Krell, A., Mock, T., et al.

- (2015) In situ expression of eukaryotic ice-binding proteins in microbial communities of Arctic and Antarctic sea ice. *ISME J* 9: 2537-2540.
- 37) Verbruggen, H., Ashworth, M., LoDuca, S.T., Vlaeminck, C., Cocquyt, E., Sauvage, T., et al. (2009) A multi-locus time-calibrated phylogeny of the siphonous green algae. *Mol Phylogenet Evol* 50: 642-653.
- 38) Xie, F., Xiao, P., Chen, D., Xu, L. and Zhang, B. (2012) miRDeepFinder: a miRNA analysis tool for deep sequencing of plant small RNAs. *Plant Mol Biol* 80: 75-84.
- 39) Zhang, X., Henriques, R., Lin, S.-S., Niu, Q.-W. and Chua, N.-H. (2006) Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nature Protocols* 1: 641.



Table 1. Information of primers used in this study.

Name	Sequences (5' to 3')
For qPCR analyses	
CmIBP1 qPCR F	ACT CTT TGA ATG CGC GAG TG
CmIBP1 qPCR R	TGA CCA AAG CTG CAT TGC TG
CmIBP2 qPCR F	AAC ATT CAC GGG CAC TTT GG
CmIBP2 qPCR R	TCA CAG CCG TGT TTG AAG TG
CmIBP3 qPCR F	TGA TTG GAC AGT GTG TGT GG
CmIBP3 qPCR R	ATC ATG AGC CTT GCA CCA AG
KNF032-5016 F	CGT GTG CAT GCA ATT ACC AC
KNF032-5016_R	AAA GTG CGT GCA TCA CCA AC
KNF032-39 F	TGG CCT TTT GCA TTG TCA CC
KNF032-39 R	AAA GTG AAG CCG ACC ATT GC
KNF032-15619_F	ACC GCG CAC GAA ACA TTT AC
KNF032-15619 R	TCT CGA ACG TGA CAA ATC GC
KNF032-646 F	ACC GTT TCG TCA TTG TGC AC
KNF032-646_R	TGC ATG CCT TGT TGG GAA TG
KNF032-8467 F	TGC GCT TCA AGA ATG AGC TG
KNF032-8467 R	TGC TGG CCT TTC ACA TGA TG
KNF032-433 F	TCC ATG GAC AAC ACA ATG CG
KNF032-433 R	AGC TGC ATC GAA GCA AAA GC
KNF032-662 F	TTT TGC CGT GGT TGC TTG AC
KNF032-662 R	TGT CAC TTG AAT CGC TGA GC
KNF032-1423 F	ACA TTG CTT TGG CGT GGA TG
KNF032-1423 R	ATG AGG GTT GGT GGT GTT TG
KNF032-1534 F	TTT TGA AAG TGC GCG ACT GG
KNF032-1534 R	AGC GCA GTA AAC TGG TGA AC
KNF032-11906 F	TGA CAC GCA CGT TCC AAA AG
KNF032-11906 R	ACG TGT GCC CAA AGT TGA AG
For cloning into recombinant protein expression vector	
CmIBP1 F BamHI	CGG GAT CCG ATG GCC GAG GTC ACG TGC
CmIBP1 R XhoI	CCG CTC GAG CTG GGC CGG GCA GAG
CmIBP2 F BamHI	CGG GAT CCG ATG ATC GTT GTG TGC AAG ATG
CmIBP2 R XhoI	CCG CTC GAG GTA GCA CTC CCT GGG
CmIBP3 F BamHI	CGG GAT CCG ATG GTG ATT GTT GTT TGC AAG ATG G
CmIBP3(-TM) F BamHI	CGG GAT CCG ATG CCT GGA GCG GGC CGT CAG
CmIBP3(-TM) R XhoI	CCG CTC GAG ATA GCA ATA ACG AGG CCC
For constructing transgenic plants	
CmIBP1 ORF F	ATG GCC CCA GCA GCT GGC ACA
CmIBP1 ORF R	CTA CTG GGC CGG GCA GAG TG
CmIBP2_ORF_F	ATG CCT AGCT CTTC AAT GAA GCT GTT
CmIBP2 ORF R	CTA GTA GCA CTC CCT GGG
CmIBP3 ORF F	ATG TCA ACA ACA ACC ACT ACC ATG AAG
CmIBP3 ORF R	CTA ATA GCA ATA ACG AGG CCC TTC C

Table 2. Amplification efficiency test of selected 10 internal control genes.

Contig	Description	FPKM value			Amplification test	
		Avr.	Std.	CV [†]	Efficiency	R ²
Contig5016	methyltransferase-like protein 1- partial	31.7	0.656	0.021	1.09	0.996
Contig39	elongation factor-like protein	56.4	1.172	0.021	0.98	0.997
Contig15619	TPR-like protein	32.8	0.735	0.022	0.94	0.966
Contig646	eukaryotic initiation factor	51.3	1.634	0.032	1.00	0.953
Contig8467	eukaryotic translation initiation factor 2 gamma	95.9	3.066	0.032	1.14	0.980
Contig433	wd40 repeat-like protein	52.4	1.731	0.033	0.99	0.965
Contig662	exostosin-like glycosyltransferase	83.8	2.826	0.034	0.95	0.993
Contig1423	l-aspartate oxidase	108.1	3.741	0.035	0.85	0.937
Contig1534	mitochondrial pyruvate dehydrogenase kinase	38.8	1.374	0.035	1.12	0.995 [†]
Contig11906	protein phosphatase 2a catalytic subunit	49.5	1.755	0.035	1.01	0.960

CV value was calculated from dividing standard deviation (Std.) by average of FPKM value (Avr.).

극지연구소

Table 3. Summary of genetic features of six CmIBP genes. Gene sizes and number of exons were confirmed by PCR-based sequencings of genomic DNA, and the plant type transcription factor binding sites were found within -1kb upstream sequences of six CmIBP genes.

Gene	Gene structure		Promoter region			
	Size (bp)	No. Exon	Size (bp) †	CBF	ABI3	WRKY 1,2,3,18
<i>CmIBP1</i>	4,003	11	2,880	2	2	1
<i>CmIBP2</i>	3,206	12	8,940	1	1	2
<i>CmIBP3</i>	3,298	11	660	2	1	1
<i>CmIBP4</i>	3,242	9	2,040	1	1	1
<i>CmIBP5</i>	2,749	9	333	1	n.d.	n.d.
<i>CmIBP6</i>	4,574	11	1,579	3	1	n.d.

† Size of the PCR-confirmed upstream region from the start codon



Table 4. Comparison of the protein composition of CmIBPs and other Thr-rich AFPs[†]

	Length (aa)	Thr		Motif	Cys		Disulfid e bonds	UniProtKB
		No.	(%)		No.	(%)		
CmIBP1	341	41	12.0	TXT	17	5.0	7	In this study
CmIBP2	323	39	12.1	TXT	14	4.3	6	In this study
CmIBP3	293	39	13.3	TXT	13	4.4	5	In this study
TmAFP	84	17	20.2	TCT	16	19.0	8	O16119
sbwAFP	121	20	16.5	TXT	10	8.3	5	Q9GSA6
MpIBP-RIV	326	20	6.1	TXN	0	0	n.a.	A1YIY3

[†] The signal peptide regions were excluded from the calculation.



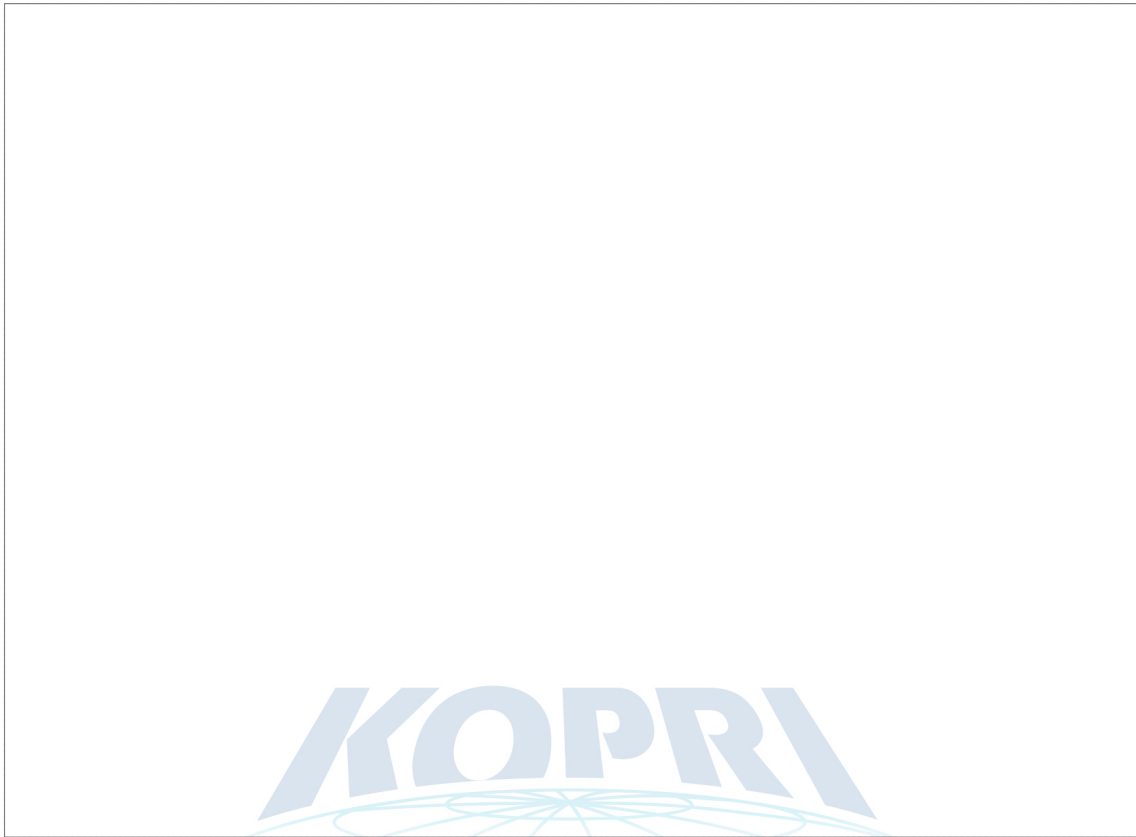


그림 1. Transcriptional response of IBPs from the psychrophilic green alga *Chloromonas sp.* KNF0032. (a) Heat map displaying the FPKM levels of 29 annotated genes that increased in response to low temperature. Among those, six genes were identified as CmIBP (Chloromonas IBPs). (b) Transcript levels of the CmIBP gene results (CmIBP1 - CmIBP6) obtained from RNA-Seq analysis. (c) Transcriptional changes in CmIBP1, CmIBP2, and CmIBP3 in response to low temperature treatment quantified by qRT-PCR relative to the expression values of the internal controls contig5016 (methyltransferase-like protein) and contig39 (elongation factor-like protein). Internal control genes were selected and evaluated based on transcriptome data (detailed in the Materials and methods section). Mean relative expression values are plotted with standard deviation ($n = 3$).

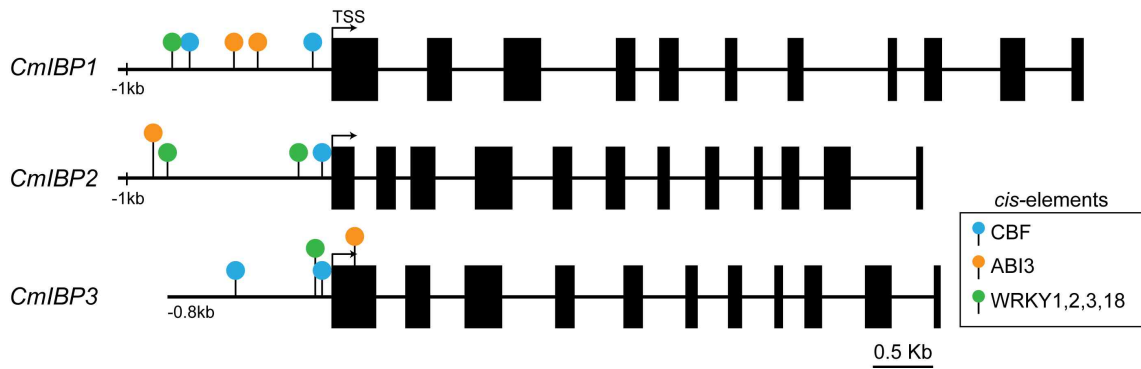


그림 2. Prediction of exon - intron structures and putative transcription factor binding sites in upstream regions of CmIBP1, CmIBP2, and CmIBP3. CmIBPs have 11 - 12 exons and cis-elements that are known to be bound by plant transcription factors. The black boxes indicate exon regions, and TSS indicates the transcription start sites of the genes. Putative transcription factor binding sites, predicted by the TRANSFAC database (Matys et al. 2003), and are labeled with arrows and circle heads.



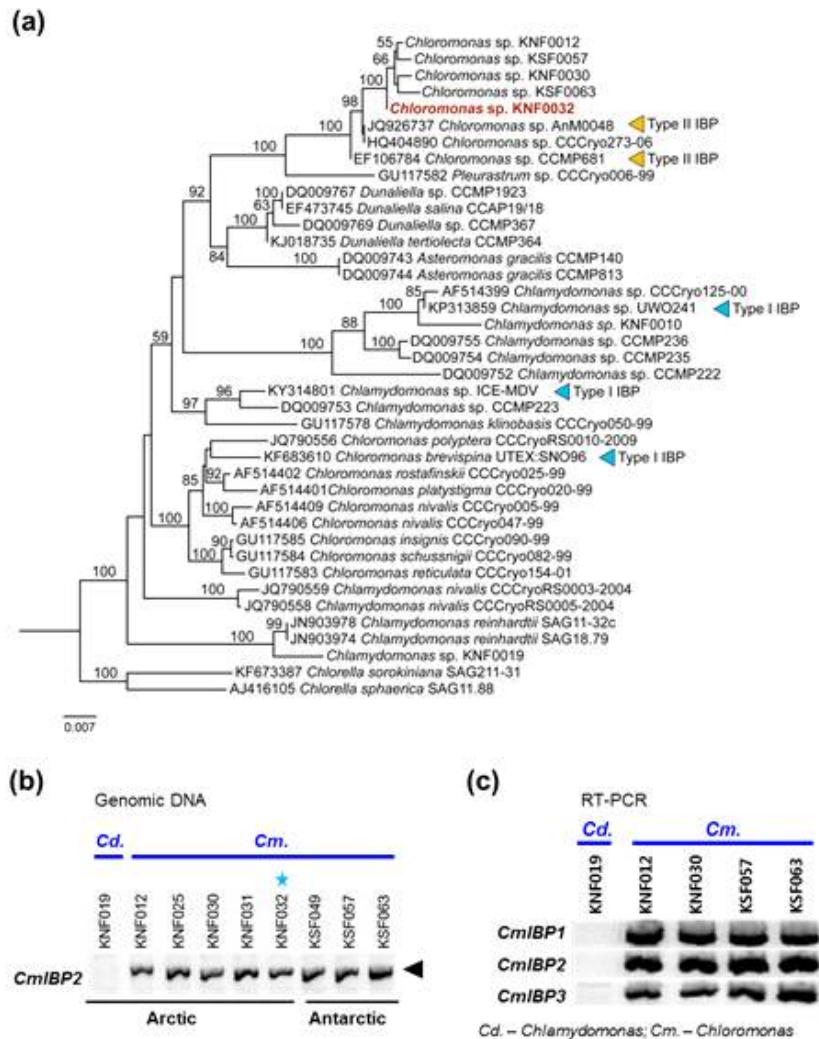


그림 3. Phylogenetic analysis of 18S rDNA sequences and screening of CmIBP homologous genes at genomic and transcriptional level (a) For constructing a phylogenetic tree, the 18S rDNA region of five *Chloromonas* strains (KNF0012, KNF0030, KNF0032, KSF0057 and KSF0063) and *Chlamydomonas* strain (KNF0010) were amplified using gene specific primers JO2 and SS17HR (Verbruggen et al., 2009) and those sequences were deposited in GenBank (MH400028-MH400033). Additional 34 sequences of *Chloromonas*, *Chlamydomonas* and *Chlorella* species were downloaded from NCBI. Multiple alignments were performed and a phylogenetic tree was constructed for 1,154 nucleotides regions in 18S rDNA sequences by Neighbor-Joining methods with a Kimura-2-parameter in MEGA7. Bootstrap analyses were performed by 1,000 replicates and the value is shown on the branches. *Chloromonas* sp. KNF0032 used in this study is indicated in red. Blue and yellow arrow heads mark the species which have reported the type I and type II IBP proteins, respectively. (b) Amplification of CmIBP2 CDS region from 9 KCCPM strains, 5 Arctic and 3 Antarctic *Chloromonas* and an Arctic *Chlamydomonas* strains using gene specific primer (Table 1). Arrow head indicated in ~3.2kb which is the actual size of CmIBP2 gene. (c) RT-PCR was performed from the selective four *Chloromonas* and one *Chlamydomonas* strains using gene specific primers (Table 1). Total RNA was extracted from the samples were cultivated at 2°C for a week.

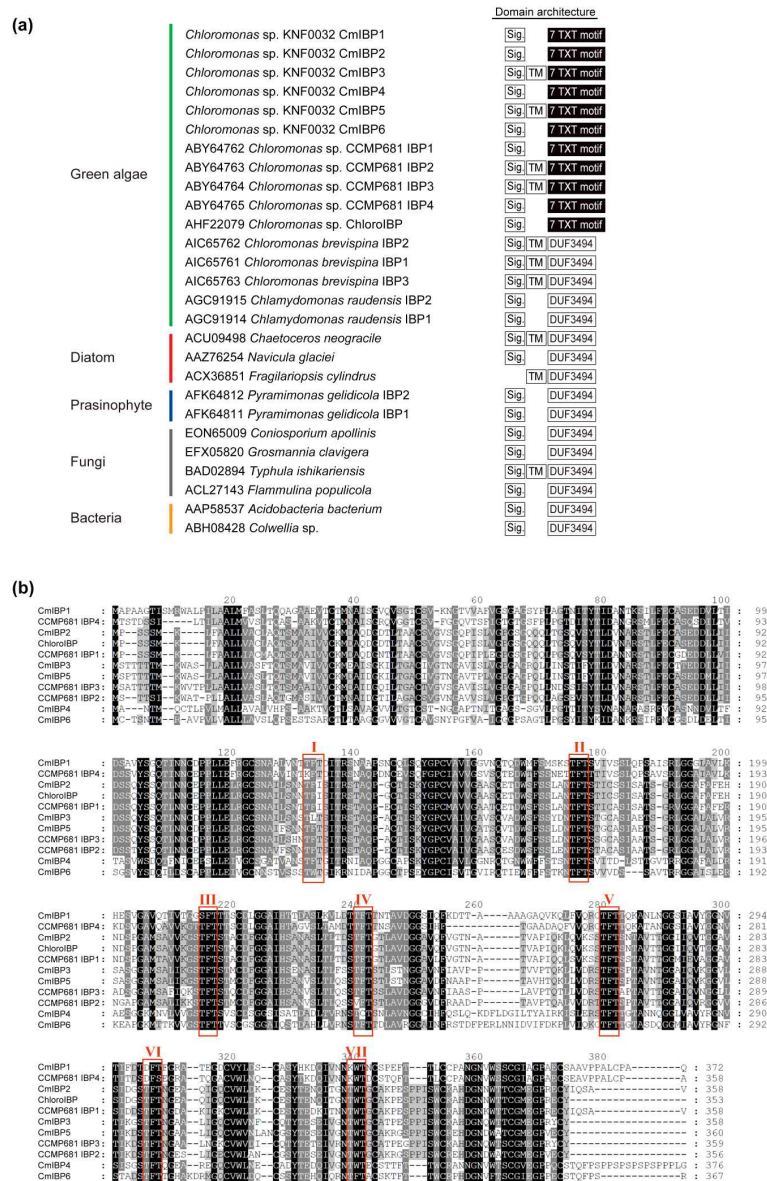


그림 4. Domain architecture and multiple alignments of six CmIBPs, four CCMP681 IBPs and ChloroIBP amino acid sequences. (a) Domain architecture and signal peptide of six CmIBP proteins with previously studied 21 IBP proteins were predicted by SMART and SignalIP webserver, respectively. Domain name was abbreviated as follow: Sig, Signal peptide; TM, transmembrane; 7 TXT motif, seven repeated TXT motifs; DUF3494, domain of unknown function 3494. (b) Multiple alignments of six CmIBPs and other type II IBPs were performed by the G-INS-i methods in MAFFT version 7. Homologous sequences are shown in black and gray background color. The seven conserved TXT motifs (I~VII) are displayed in red boxes

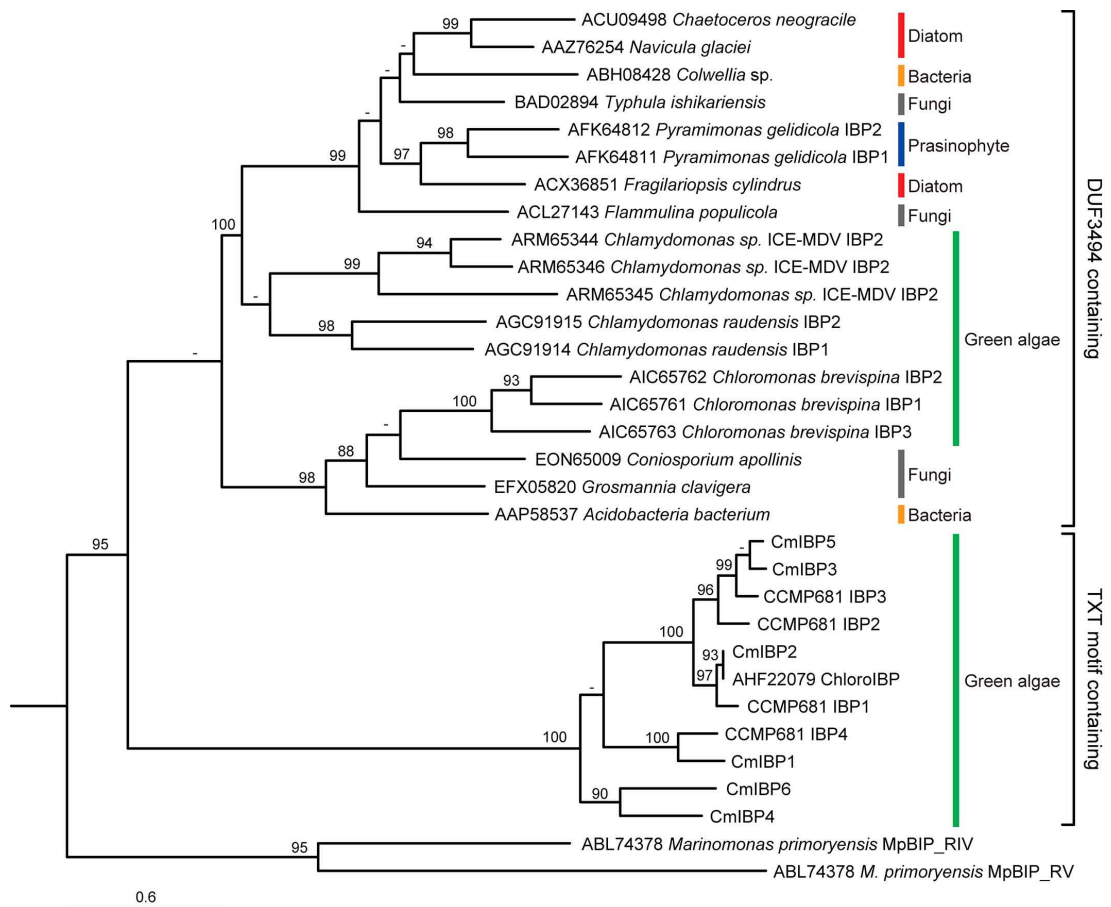


그림 5. Maximum likelihood tree of CmIBPs with previously reported IBP proteins from fungi, bacteria, diatoms, prasinophytes, and green algae. The TXT motif-containing IBP group including CmIBPs, CCMP681 IBPs, and ChloroIBP formed a monophyletic clade clearly separated from type I IBPs containing the DUF3494 domain. The phylogenetic tree was constructed with PhyML using the WAG+G+F substitution model. Bootstrap analysis was performed with 1,000 replicates and only supporting values > 80 are presented on the branches.

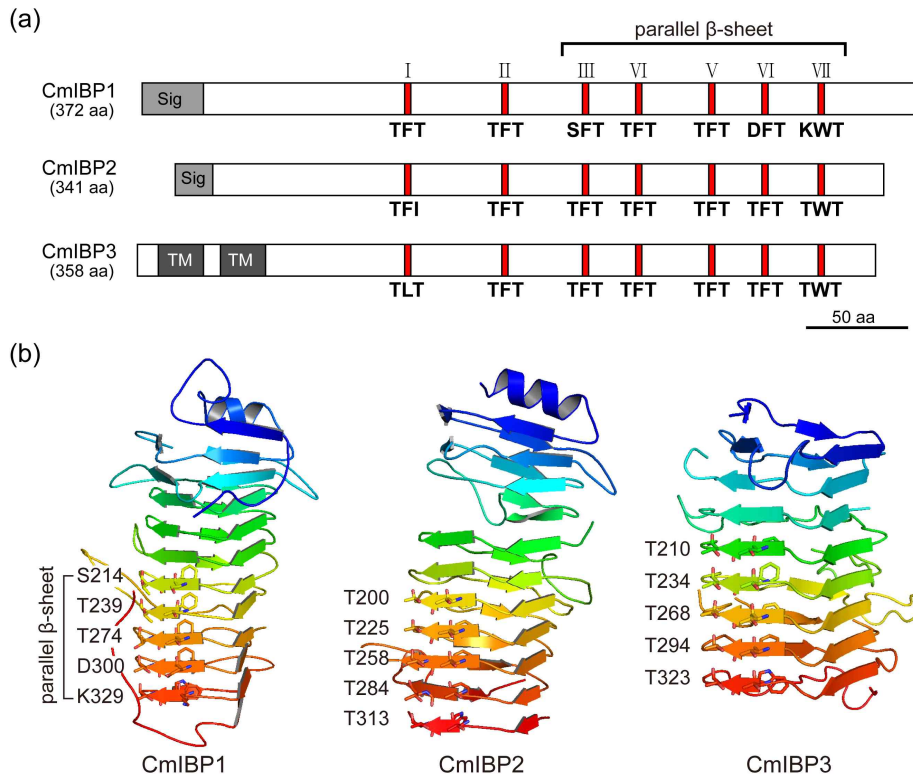


그림 6. Prediction of the domains and tertiary structures of the CmIBP1, CmIBP2, and CmIBP3 proteins. (a) Common TXT motifs in CmIBP1, CmIBP2, and CmIBP3 are marked in red and roman numerals (I - VII), and the predicted signal peptides (Sig.) and transmembrane domain are presented as gray boxes. (b) All three proteins were predicted to have a β -solenoid form by the Phyre2 webserver. The third through seventh TXT motifs spaced at regular intervals on protein sequences were arranged on one side of the parallel β -sheets. The aligned sequence regions (CmIBP1: 150 - 356 aa, CmIBP2: 89 - 337 aa, CmIBP3: 117 - 341 aa) to the template (PDB ID: c5gkdA) were used for prediction

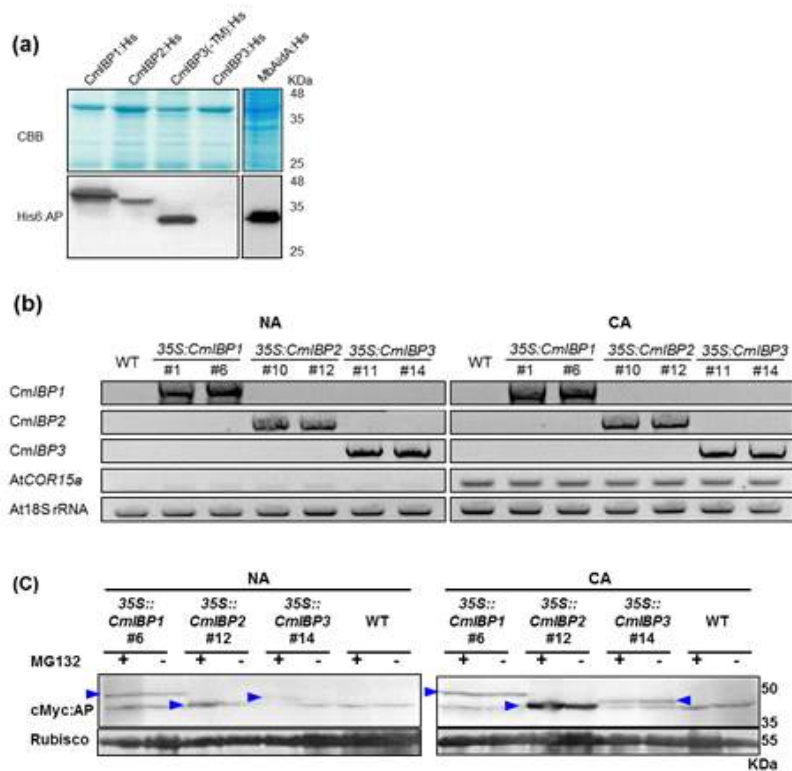


그림 7. Expression of recombinant CmIBP proteins in *E. coli*, and the RNA and protein expression of heterologous CmIBP genes in Arabidopsis transgenic plants. (a) Western blot assay using an anti-His6 antibody, to detect CmIBP proteins in each soluble fractions of recombinant CmIBP1, CmIBP2 and CmIBP3(-TM) and CmIBP3 protein. Recombinant CmIBP1 and CmIBP2 were successfully detected in soluble fraction, but in the case of CmIBP3, the signal was detected only when the transmembrane regions (1-53 amino acids) was excluded. (b) RT-PCR was performed using cDNA from six transgenic Arabidopsis lines and Col-0 wildtype using gene specific primers (Table 1), At 18S rRNA (forward- ATA CGT GCA ACA AAC CCC GA and reverse- CAT CGA AAG TTG ATA GGG CA) and AtCOR15a (forward- TGA TCC ATA TCC TCT CTC TT and reverse-TAA AGA ATG TGA CGG TGA CT) primers. Two-week-old seedlings grown under non-acclimated (NA) and cold-acclimated at 4°C (CA) for 2 days were harvested. PCR amplification was performed with following conditions: 98°C for 30 s followed by 28 cycles of 98°C for 10 s, 53°C for 10 s and 72°C for 30 s, with a final extension of 7 min at 72°C. 18S rRNA sequences were used for internal control. (d) Comparison of protein levels under the NA and CA conditions of CmIBP1:cMyc (40 KDa), CmIBP2:cMyc (37 KDa), and CmIBP3:cMyc (38 KDa), which were constitutively expressed by the CaMV 35S promoter. Ponceau staining of the ribulose-1,5-bis-phosphate carboxylase/oxygenase (Rubisco) large subunit served as a loading control.

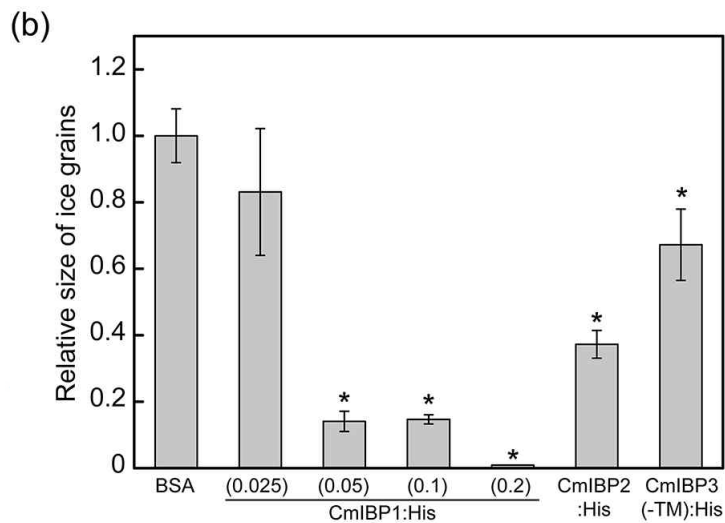
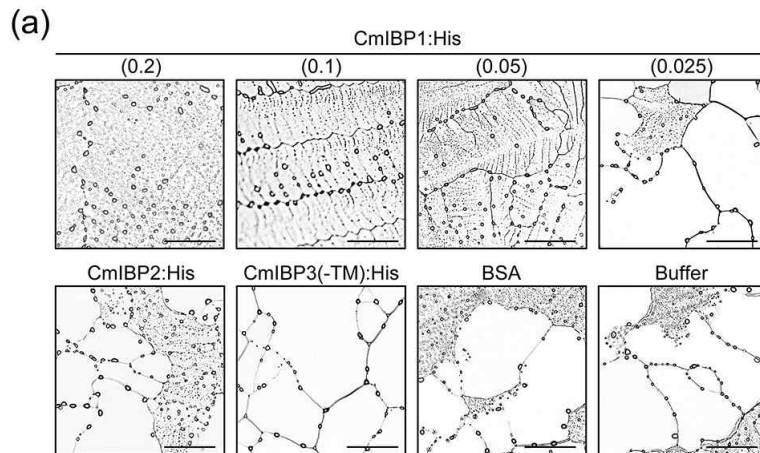


그림 8. IRI activity assay of recombinant CmIBPs. (a) Images of ice grains after freezing for 30 min at -6°C in IRI assay. (b) The size of individual ice grains in each sample (3.2 mm^2 in the area) was measured by ImageJ program. Protein concentration was 0.1 mg ml^{-1} as a default, and the number of parentheses showed the concentration of CmIBP1. The relative size of ice grains were calculated from the average size of each samples divided by that of the BSA control. The experiments were repeated twice, and the asterisks represent statistical significance ($p < 0.05$) determined by the t-test, compared to the control. The scale bar shows 0.25 mm

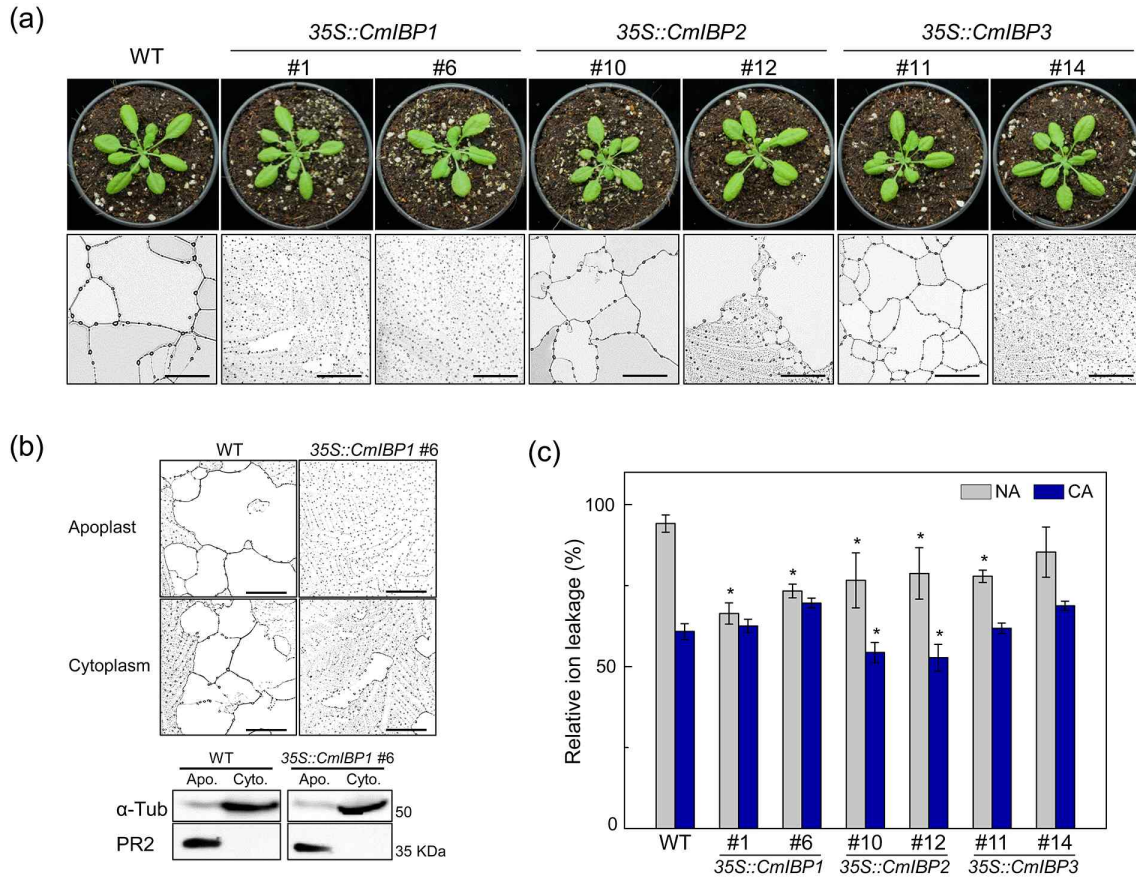


그림 9. Transgenic Arabidopsis plants overexpressing CmIBPs enhance freezing tolerance due to IRI activity. (a) Images of transgenic Arabidopsis plants under normal growing conditions at 22°C (upper panel) and the ice grain images of the cell extracts from each plant after freezing for 30 min at -6°C in IRI assay (lower panel). Scale bar shows 0.25 mm. (b) IRI assay and western blotting results using apoplastic and cytoplasmic protein fractions extracted from wild-type and 35S::CmIBP1 #6 plant leaves. IRI assay was performed as previously described. In western blotting, anti-PR-2, pathogen related protein-2, and anti- α -tubulin used as the apoplastic and cytoplasmic marker protein antibody, respectively. (c) Electrolyte leakage measurements after freezing treatment at -6°C of the transgenic plants grown under NA and CA conditions. Relative ion leakages of the samples were compared to that of wild-type plants grown under NA condition. The experiments were repeated four times and the asterisks represent statistical significance ($p < 0.05$) determined by the t-test compared to the wild-type control.

제 3-6절 남극 선타류 AP2 유전자 과다발현 형질전환체 제작과 저온 표현형 분석

Antarctic moss AP2 genes can improve cold tolerance in model moss
Physcomitrella patens

요약

낮깃털이끼(*Sanionia uncinata*)는 남극 지역에서 가장 우점하는 이끼종이다. 남극의 극한 생태계에서 생산자로서의 중요성에도 불구하고, 이들의 저온 내성과 관련된 분자 기작은 아직 아무것도 밝혀지지 않았다. 본 연구에서는 낮깃털이끼의 저온 내성에 관여하는 주요 전사인자를 발굴하고 그들의 생체내 기능을 알아보기로 전사체 정보를 기반으로 21개 AP2 유전자를 찾아냈고, 이들을 과발현하는 식물체를 제작하여 저온과 관련된 표현형을 분석하였다. 21개 유전자들 중 50% 정도는 저온과 건조 조건에서 발현이 증가하였고, 또 이들 중 절반은 스트레스 호르몬인 ABA에 의해서도 발현이 증가하였다. 모델 이끼 과발현체의 표현형 분석 결과, SuAPL 3, 4, 11, 13, 14, 21 과발현체는 야생형에 비해 증가된 저온 저항성을 나타내어 4°C에도 꾸준히 성장하여 초록색을 띄며 높은 광합성능을 유지하고 있었다. 이들 중 특히 저항성이 두드러지는 SuAPL 3/14 과발현체의 경우 저온 외에 고염과 ABA 처리 조건에서 야생형에 비해 저항성이 증가했으며 특히 mannitol을 처리한 삼투 스트레스 조건에서는 저항성이 두드러지게 증가했다. 결론적으로, 본 연구결과는 낮깃털이끼의 AP2 유전자의 과발현이 모델 이끼의 저온을 비롯한 여러 가지 환경스트레스에 대한 내성을 증가시켰음을 보여 주었다. 이러한 결과는 남극과 같은 극한의 환경에서 차별적으로 진화한 AP2 유전자는 낮깃털이끼의 극지 적응을 위한 독특한 생존 전략임을 시사한다.

Abstract

Sanionia uncinata is the most dominant moss species in Antarctica. Despite their importance as producers in the extreme ecosystems of Antarctica, the molecular mechanisms associated with enhanced cold tolerance have yet to be identified. In this study, we found 21 AP2 genes based on transcriptome to identify key transcription factors involved in the cold tolerance and to examine their in vivo functions in *S. uncinata*. Transcription of about 50% of the 21 genes was increased at cold and dry conditions, and again half of which were also increased by the stress hormone ABA. As a result of analyzing phenotypes of model moss plants overexpressing SuAPL 3, 4, 11, 13, 14, and 21, showed increased cold tolerance

compared to wild type, and continued to grow at 4°C showing continuous photosynthesis. Among them, SuAPL 3/14 overexpressing plants, which are particularly resistant to cold, were more resistant than wild-type under high salt and ABA treatment. In conclusion, the results showed that overexpression of the AP2 gene of *S. uncinata* increased resistance to various environmental stresses in model moss, including low temperature. These results suggest that the AP2 gene, which has evolved differentially in extreme environments such as Antarctica, is a unique survival strategy for polar adaptation of *S. uncinata*.

1. 서론

극심한 건조와 영하 온도라는 극한의 조건 때문에 남극 육상의 식생은 항상 생존의 생리학적 한계를 겪는다. 때문에 환경 요인의 약간의 변화가 육상 생태계에 큰 영향을 미치기도 한다(Robinson et al. 2003). 따라서 남극의 이끼는 기후 변화에 의한 영향에 매우 취약할 수 있고, 개체 수분보다는 종의 분포와 생태생리적 적응의 맥락에서 그들의 반응을 이해하는 것이 중요하다. 낮깃털이끼(*Sanionia uncinata*)는 남극 반도에서 가장 우점하는 육상식물 중 하나이다(Nakatsubo 2002; Putzke et al. 2015). 기존 연구에서는 자외선 및 금속 오염과 같은 환경 요인의 영향을 연구하기 위한 대상으로 활용되었다(Fernandes et al. 2017; Lud et al. 2002; Samecka-Cymerman et al. 2011). 또한 수분 손실을 방지하는 데 도움이 되는 카펫 모양의 개체군을 형성하여 장기적으로 수분을 유지함으로써 탈수에 대한 내성을 갖는 것으로 보인다(Zúñiga-González et al. 2016).

이처럼 낮깃털이끼의 남극의 저온 생태계에서 생산자로서의 역할의 중요성에도 불구하고, 낮깃털이끼의 저온 내성에 대한 분자 기작은 아직 아무것도 밝혀지지 않았다. 본 연구에서는 이 종의 저온 내성을 정량화하고 분자 수준에서 지구온난화로 인한 환경 변화에 대한 반응을 예측하기 위해, 저온 내성에 관여하는 주요 전사인자를 발굴하고 그들의 생체내 기능을 알아보려고 하였다. 이를 위해 낮깃털이끼 전사체 정보를 기반으로 21개 AP2 유전자를 찾아냈고, 이들을 과발현하는 식물체를 제작하여 저온과 관련된 표현형을 분석하였다.

2. 재료 및 방법

가. 식물 재료 및 성장 조건

남극에서 채집한 낮깃털이끼는 멸균된 배우체의 형태로 실험실에서 유지하였다. 배양은 Knop-ME 배지에서 진행하였고(Schween et al. 2003), 3주마다 T10 basic ULTRA-TURRAX (IKA)로 분쇄한 후 계대하였다. 배양실은 장일 조건, 20-25°C

를 유지하였다.

나. 스트레스 처리 및 qRT-PCR 반응

저온 처리를 위해, 25°C에서 키운 낮깃털이끼 배양체를 다양한 시간(1시간-7일) 동안 4°C 챔버에서 배양했다. 탈수 처리를 위해, 식물을 여과지로 옮기고, 15°C에서 건조시키고, 24시간 이내 다양한 시간에 샘플링하였다. ABA 처리를 위해, 식물을 100 μ M ABA를 함유하고 있는 배지로 옮기고, 24시간 이내 다양한 시간에 샘플링하였다. 일주기 리듬에 의한 변화를 피하기 위해 발현 분석을 위한 모든 샘플링을 동시에 수행하였다. RNeasy Plant Mini Kit (Qiagen, Hilden, Germany)를 사용하여 남극곰새풀 및 벼 식물의 성숙한 잎으로부터 total RNA를 분리하였다. RNA의 양과 품질은 ND-1000 분광 광도계(NanoDrop Technologies, Wilmington, DE, USA)를 사용하여 결정하였다. TOPscript 역전사 효소(Enzymomics, South Korea) 및 올리고(dT) primer를 이용하여 2 μ g의 total RNA로부터 cDNA를 합성하였다. qRT-PCR 분석은 20 μ L 반응 혼합물[1 μ L의 1:10 희석된 cDNA 주형, 2 μ M의 각 프라이머, 10 μ L의 TB Green Premix ExTaq (TaKaRa, Japan)]에서 수행하였다. 증폭 절차는 95°C에서 5분의 변성 및 효소 활성화, 95°C에서 10초, 55°C에서 10, 72°C에서 15초, 40 사이클을 수행하였다. SuTub 유전자를 내부 대조군으로 사용하였다.

다. 이끼 형질전환체 제작

원형질체 형질전환 과정은 기존의 보고를 참고하여 수행하였다(Strepp et al, 1998). 분리한 원형질체를 계수한 다음 1.2x10⁶개mL⁻¹의 농도로 3M 배지(5mM MgCl₂, 0.1% MES, 0.48M mannitol, pH5.6)에 재현탁시켰다. Plasmid DNA purification kit(Macherey-nagel, Germany)를 이용해서 분리한 pAct:SuAPL 플라스미드는 AseI 제한효소(NEB, USA)로 처리하여 선형화시켰다. 잘린 DNA는 NucleoSpin® Gel and PCR Clean-up Kit(Macherey-nagel)로 정제한 다음 500 ng μ L⁻¹의 농도로 0.1M Ca(NO₃)₂ 용액에 재현탁시켰다. 이 용액을 멸균된 등근바닥 튜브(SPL Life Science, South Korea)로 옮긴 다음 250 μ L의 원형질체-현탁액을 튜브에 첨가하고 부드럽게 혼합한 다음 350 μ L의 PEG 용액(3M 배지에 녹인 4% w/v PEG4000)을 첨가했다. 혼합물을 가끔씩 부드럽게 섞어주면서 30분 간 실온에 둔 후, 3 분마다 3M 배지 1, 2, 3, 4 ml를 첨가해 주었다. 750 rpm에서 10분간 원심분리해서 원형질체를 가라앉히니 후, 재생 배지(250 mg L⁻¹ KH₂PO₄, 250 mg L⁻¹ MgSO₄ × 7H₂O, 250 mg L⁻¹ KCl, 1 g L⁻¹ Ca(NO₃)₂ × 4H₂O, 12.5 mg L⁻¹ FeSO₄ × 7H₂O, 50 g L⁻¹ glucose, 30 g L⁻¹ mannitol, pH 5.8, 0.22 μ m 필터로 멸균)에 현탁한 후, 6구 웰플레이트로 옮겨 어둠 속에서 24시간 동안 배양했다. 그 후, 정상 성장 조건에서 10일 동안 배양하고, 재생된 원형질체를 셀로판을 깬 Knop 배

지로 옮겼다. 9일 후, 셀로판 채로 15 mg L⁻¹ 하이그로마이신(AG Scientific, USA)을 포함하는 배지로 옮겨, 항생제 내성을 보이는 콜로니를 선택하여 분리했다.

라. 이끼 형질전환체 검증을 위한 RT-PCR

제조사의 프로토콜에 따라 MiniBEST Plant RNA Extraction Kit(TAKARA)를 사용하여 이끼의 사상체로부터 total RNA를 추출하였다. RNA의 양과 품질은 NanoDrop Lite Spectrophotometer(Thermo Scientific, USA)를 사용해서 결정했다. TOPscript 역전사 효소(Enzynomics, South Korea)와 올리고 (dT) primer를 이용하여 2 μ g의 total RNA로부터 1차 cDNA를 합성하였다. RT-PCR 반응은 20 μ L 반응 혼합물[1 μ L cDNA template, 2 μ M의 각 프라이머, 10 μ L TB Green Premix Ex Taq]에서 수행하였다. 증폭 절차는 다음과 같았다. 95 $^{\circ}$ C에서 5분의 변성 및 효소 활성화, 95 $^{\circ}$ C에서 30초, 55 $^{\circ}$ C에서 30초, 72 $^{\circ}$ C에서 30초의 28 사이클을 수행하였다. Tubulin을 내부 대조군으로 사용했다(Gao et al. 2015).

마. 엽록소 형광 측정

엽록소 형광 측정은 엽록소 형광측정기(MINI-PAMII, Heinz Walz, Germany)를 사용하여 진행하였다. 측정하기 전에 *P. patens* 식물체를 20분 동안 암처리했다. Fv/Fm 값은 10개 이상 개체를 측정하여 평균값을 사용하였다.

3. 결과 및 토의

낮깃털이끼는 남극 킹조지섬에서 우점하는 선대류로서, 극저온, 부족한 수분, 높은 자외선량 등의 남극환경에 높은 적응력을 보유하고 있는 것으로 알려져 있다. AP2 단백질은 식물에서 대사 작용, 성장과 발달, 환경 인자에 대한 반응 등 다양한 생물학적 반응을 조절하는 대표적인 전사인자로서, 남극 선대류의 AP2 유전자들 중 저온 환경 적응에 기능하는 주요 유전자를 선별하고 특성을 파악한다면 식물의 환경 적응 기작을 분자 수준에서 이해하는 데 크게 기여할 수 있을 것이다. 이를 위해 선행 연구를 통해 저온 스트레스에 의해 발현이 증가하는 21개의 AP2 유전자를 선별하였다.

이들 유전자를 대상으로 다양한 환경 스트레스 조건에서 유전자 발현양상을 조사하였다. 먼저 남극의 대표적인 환경 스트레스인 저온 조건에서의 발현분석 결과, 대부분의 AP2 유전자들의 발현이 증가했고 특히 SuAPL 1, 2, 3, 4, 7, 8, 13, 14, 17, 23 등의 유전자들은 발현의 증가폭이 두드러지고 스트레스 처리 초기인 1시간 이후부터 빠른 증가를 보였다(Figure 1). 또한 저온과 건조는 둘 다 수분과 관련해서 탈수를 방지하는 쪽으로 식물이 적응을 했기 때문에 분자 수준에서 기작을 공유하는 경우가 많다. AP2 유전자들이 건조 조건에서도 발현양상이 변화하는지를 확인한 결과,

SuAPL 1, 2, 3, 4, 7, 8, 12, 14, 15, 16, 17, 23 등의 유전자들의 발현이 증가했다 (Figure 2). 특히 SuAPL1의 경우 저온과 건조 모두에서 매우 강한 수준으로 발현이 증가했다. 또한 ABA는 흔히 스트레스 호르몬으로 알려져 있을 정도로 환경 스트레스 반응과 관련이 높은 식물 호르몬이다. AP2 유전자의 작용이 ABA 호르몬의 영향을 받는지를 분석하기 위해 100 μ M ABA를 처리하여 발현양상을 분석한 결과, SuAPL 1, 3, 8, 16, 17 등의 유전자들의 발현이 8배 이상 증가하였다(Figure 3). 이들 중 SuAPL 1, 3, 8, 16, 17 등 5개 유전자는 저온과 건조 조건에서도 발현이 증가했기 때문에 이들은 ABA에 의해 반응이 매개되는 유전자들일 것으로 추측되며, 나머지 유전자들은 ABA와는 독립적으로 작용할 것으로 판단된다.

낮깃털이끼 AP2 유전자들의 생체 내 기능을 알아보기 위해 이들 유전자들을 과발현하는 모델 이끼(*Physcomitrella patens*) 형질전환체를 제작하였다. 낮깃털이끼를 대상으로 하는 형질전환 기법은 아직 구축되어 있지 않기 때문에 모델 이끼를 사용하였다. 21개 유전자들을 *P. patens*의 액틴 유전자 프로모터와 결합시킨 플라스미드를 제작한 후 PEG를 활용한 원형질체 형질전환 기법으로 형질전환체를 제작하였다. Hygromycin을 포함하는 배지에서 생존하는 콜로니들을 선별하여 외래 플라스미드의 삽입 여부를 PCR로 확인한 결과, 대부분의 콜로니에서 플라스미드가 삽입된 것을 확인할 수 있었다(Figure 4). 또한 삽입된 플라스미드에서 SuAPL 유전자들의 전사체가 정상적으로 만들어지는지를 RT-PCR로 확인한 결과, 유전자별로 적게는 1개 라인, 많게는 5개 라인으로 개수의 차이는 있었지만 모든 유전자들에 대해 과발현 형질전환체를 성공적으로 확보하였다(Figure 5).

실험에 사용한 21개 AP2 유전자들이 모두 저온 조건에서 발현이 증가하는 특성에 기반하여 선별되었기 때문에 AP2 과발현체들의 저온 저항성에 어떤 변화가 생겼는지를 분석했다. 야생형과 AP2 과발현체들을 정상 온도 조건인 20°C와 저온인 4°C에서 배양하여 이들 엽상체의 성장 모습을 관찰하였다(Figure 6). 일부 과발현체에서는 정상 온도에서도 야생형과 상당히 다른 형태를 보였다. 특히 SuAPL 3, 4, 11, 12, 13, 14 과발현체는 야생형에 비해 배우체의 발달이 잘 이루어지지 못하고 사상체 위주의 성장을 보였다. 최종적으로 생성되는 배우체의 개수도 대폭 줄었는데, 이는 AP2 유전자가 식물의 발달과정에서 중요한 역할을 한다는 기존의 보고와도 일치하는 결과이다. 저온 조건에서는 몇몇 과발현체들이 야생형에 비해 뚜렷한 저항성을 보였다. 야생형은 저온 조건에서 정상적인 성장을 하지 못하고 갈색으로 죽어가는 반면, SuAPL 3, 4, 11, 13, 14, 21 과발현체는 초록색을 유지하면서 상대적으로 높은 수준의 성장률을 보였다. 가지적인 성장 외에도 광합성 지표인 Fv/Fm을 측정하여 저온 스트레스에 대한 내성을 정량화하였다. 20°C와 4°C 조건에서 Fv/Fm을 측정하면 이들 사이의 상대값을 그래프로 표시하였다(Figure 7). 그 결과 SuAPL 3, 4,

13, 14 과발현체에서 눈에 띄게 증가한 값이 관측되었는데, 이는 이들 과발현체가 저온 조건에서도 상대적으로 광합성능의 저하가 일어나지 않는다는 것을 의미한다. SuAPL 4, 13은 AP2 도메인을 2개 갖는 AP2 유전자군이고, SuAPL 3, 14는 환경 스트레스와 기능적으로 연관성이 높은 ERF 유전자군에 속하기 때문에 최종적으로 SuAPL 3, 14를 선정하여 추후 연구에 활용하였다.

SuAPL3/14 유전자 과발현체들을 대상으로, 라인간 오차가 있는지를 확인하기 위해 2개 이상의 독립된 라인들을 대상으로 저온 저항성을 분석하였다(Figure 8). 그 결과, 저온 조건에서의 성장과 광합성 특성 모두에서 2개 이상의 라인들이 공통적으로 야생형에 비해 높은 수준의 저항성을 보였다. 따라서 증가된 저항성이 우연한 효과가 아니라 과다발현되는 유전인 SuAPL3/14에 의한 것으로 결론낼 수 있다. 또한 저온 스트레스는 건조, 고염, ABA 등 다른 종류의 환경 스트레스와도 직접적으로 깊은 연관성을 보이는 경우가 많아서, 저온 내성이 증가된 식물체의 경우 다른 종류의 환경 스트레스에도 높은 내성을 보이는 경우가 많다. 따라서 SuAPL3/14 과다발현체를 대상으로 고염, 삼투, ABA 처리에 따른 성장특성을 분석하였다(Figure 9). 그 결과, 조사한 거의 모든 조건에서 이들 과다발현체는 증가된 내성을 보였으며, 특히 mannitol을 처리한 삼투 스트레스 조건에서 야생형에 비해 두드러지는 저항성을 나타내었다.

결론적으로, 본 연구결과는 낮기털이끼의 AP2 유전자의 과발현이 모델 이끼의 저온을 비롯한 여러 가지 환경스트레스에 대한 내성을 증가시켰음을 보여 주었다. 이러한 결과는 남극과 같은 극한의 환경에서 차별적으로 진화한 AP2 유전자는 낮기털이끼의 극지 적응을 위한 독특한 생존 전략임을 시사한다.

4. 참고문헌

- 1) Fernandes A, Mazzei J, Oliveira C, Evangelista H, Marques M, Ferraz E, Felzenszwalb I (2017) Protection against UV-induced toxicity and lack of mutagenicity of Antarctic *Sanionia uncinata*. *Toxicology* 376:126-136
- 2) Gao B, Zhang D, Li X, Yang H, Zhang Y, Wood AJ. (2015) De novo transcriptome characterization and gene expression profiling of the desiccation tolerant moss *Bryum argenteum* following rehydration. *BMC Genomics* 16:416
- 3) Lud D, Moerdijk T, Van de Poll W, Buma A, Huiskes A (2002) DNA damage and photosynthesis in Antarctic and Arctic *Sanionia uncinata* (Hedw.) Loeske under ambient and enhanced levels of UV B radiation. *Plant Cell Environ* 25:1579-1589

- 4) Nakatsubo T (2002) Predicting the impact of climatic warming on the carbon balance of the moss *Sanionia uncinata* on a maritime Antarctic island. *J Plant Res* 115:99-106
- 5) Putzke J, Athanasio CG, de Albuquerque MP, Victoria FC, Pereira AB (2015) Comparative study of moss diversity in South Shetland Islands and in the Antarctic Peninsula. *Rev Chil His Nat* 88:6
- 6) Robinson SA, Wasley J, Tobin AK (2003) Living on the edge - plants and global change in continental and maritime Antarctica. *Glob Chang Biol* 9:1681-1717
- 7) Samecka-Cymerman A, Wojtuń B, Kolon K, Kempers A (2011) *Sanionia uncinata* (Hedw.) loeske as bioindicator of metal pollution in polar regions. *Polar Biol* 34:381-388
- 8) Schween G, Hohe A, Koprivova A, Reski R. (2003) Effects of nutrients, cell density and culture techniques on protoplast regeneration and early protonema development in a moss, *Physcomitrella patens*. *J Plant Physiol* 160:209-212
- 9) Strepp R, Scholz S, Kruse S, Speth V, Reski R. (1998) Plant nuclear gene knockout reveals a role in plastid division for the homolog of the bacterial cell division protein FtsZ, an ancestral tubulin. *Proc Natl Acad Sci U S A* 95:4368-4373
- 10) Zuniga-Gonzalez P, Zuniga GE, Pizarro M, Casanova-Katny A (2016) Soluble carbohydrate content variation in *Sanionia uncinata* and *Polytrichastrum alpinum*, two Antarctic mosses with contrasting desiccation capacities. *Biol Res* 49:6

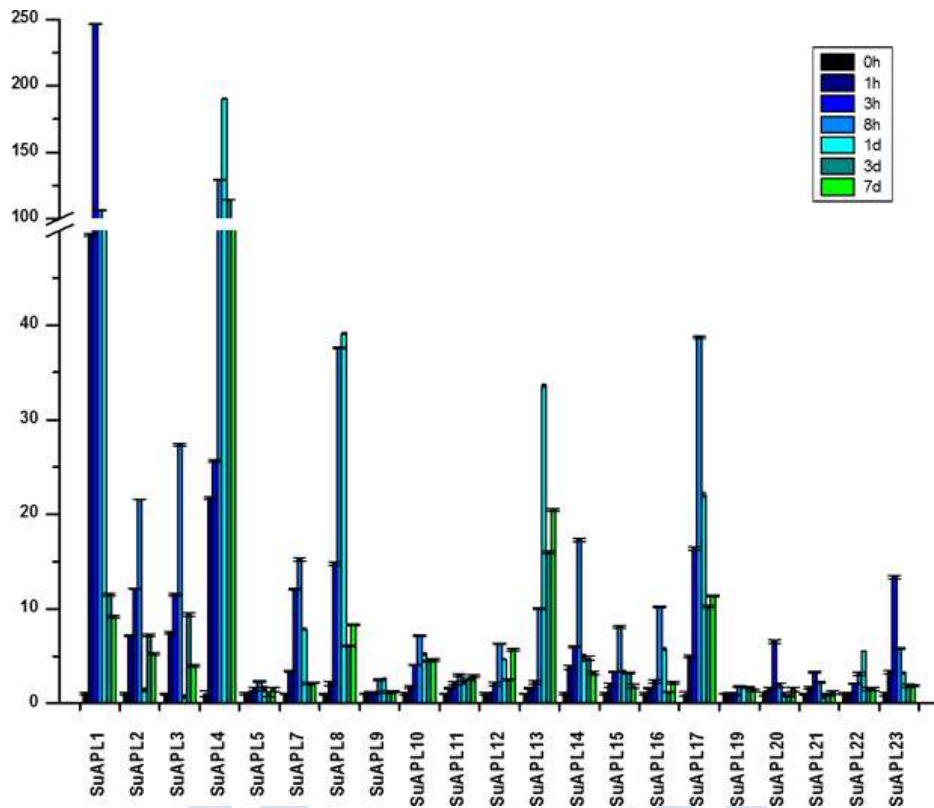


그림 1. RT-qPCR to check the expression of SuAPL genes in cold stressed *Sanionia uncinata*. Color codes indicate the duration time of cold treatment. SuTub was used as the internal reference. The data were shown as the average \pm SD of three biological replicates.

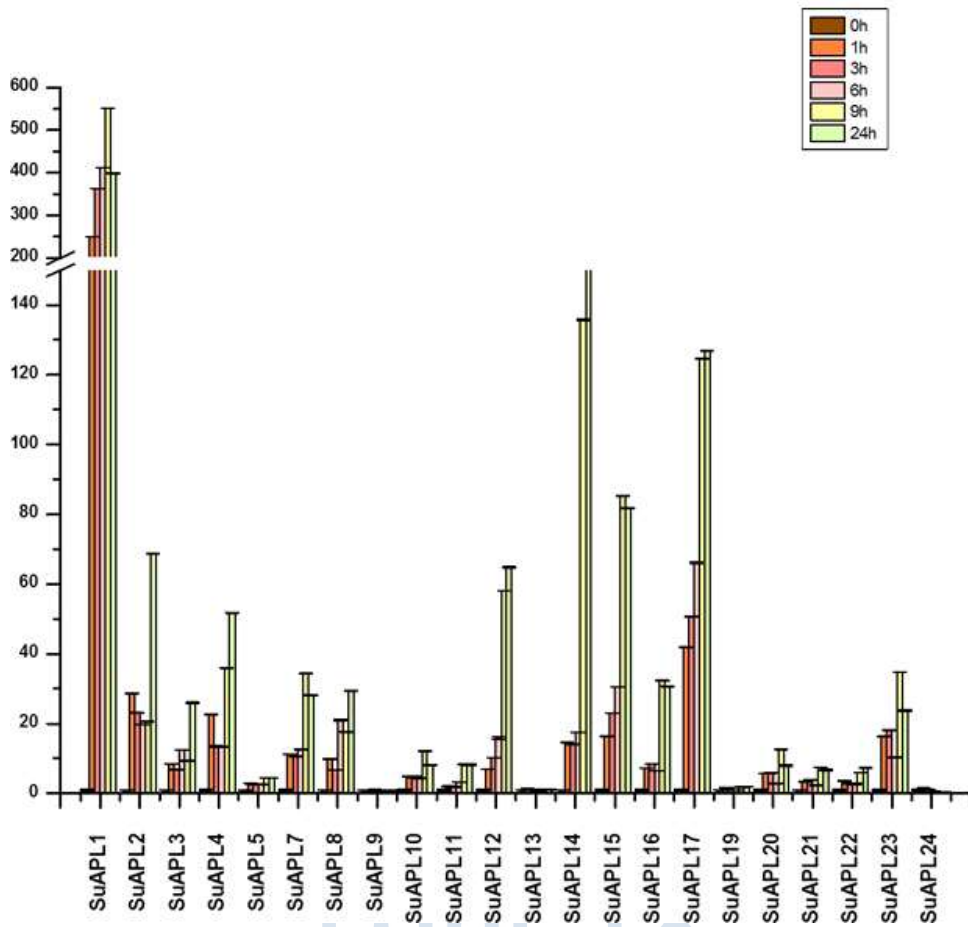


그림 2. RT-qPCR to check the expression of SuAPL genes in dehydration stressed *Sanionia uncinata*. Color codes indicate the duration time of dehydration treatment. SuTub was used as the internal reference. The data were shown as the average \pm SD of three biological replicates.

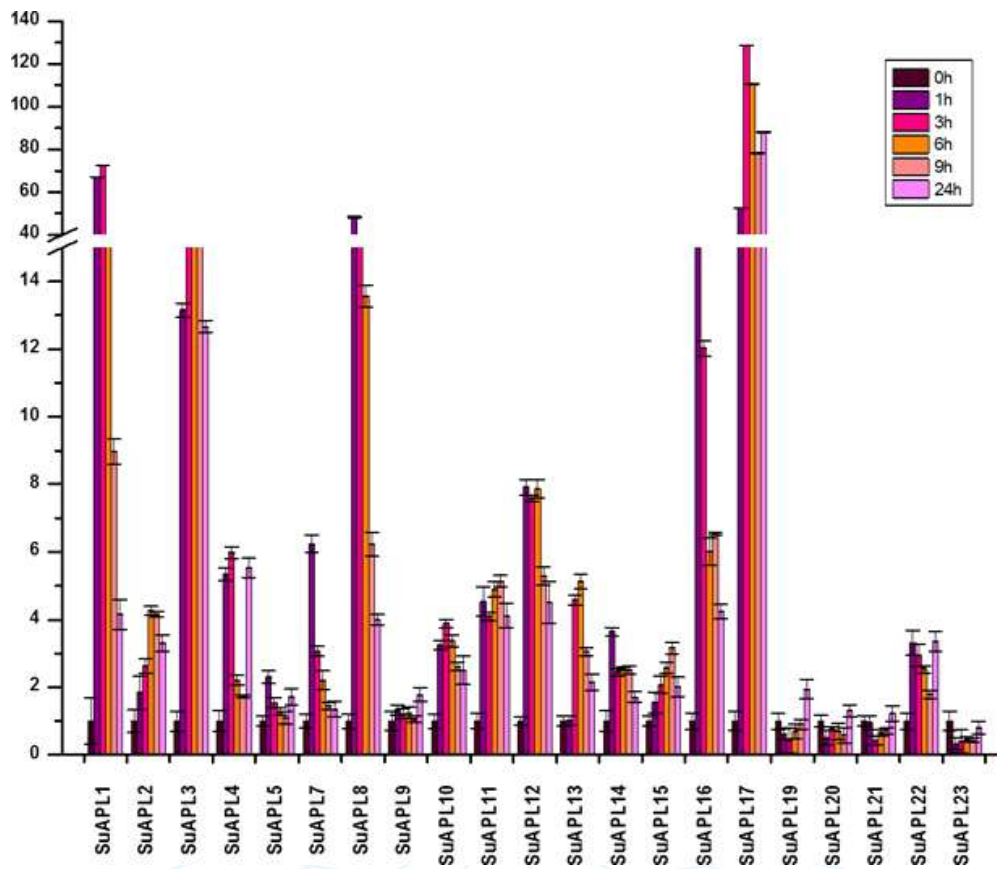


그림 3. RT-qPCR to check the expression of SuAPL genes in 100 μ M ABA treated *Sanionia uncinata*. Color codes indicate the duration time of 100 μ M ABA treatment. SuTub was used as the internal reference. The data were shown as the average \pm SD of three biological replicates.

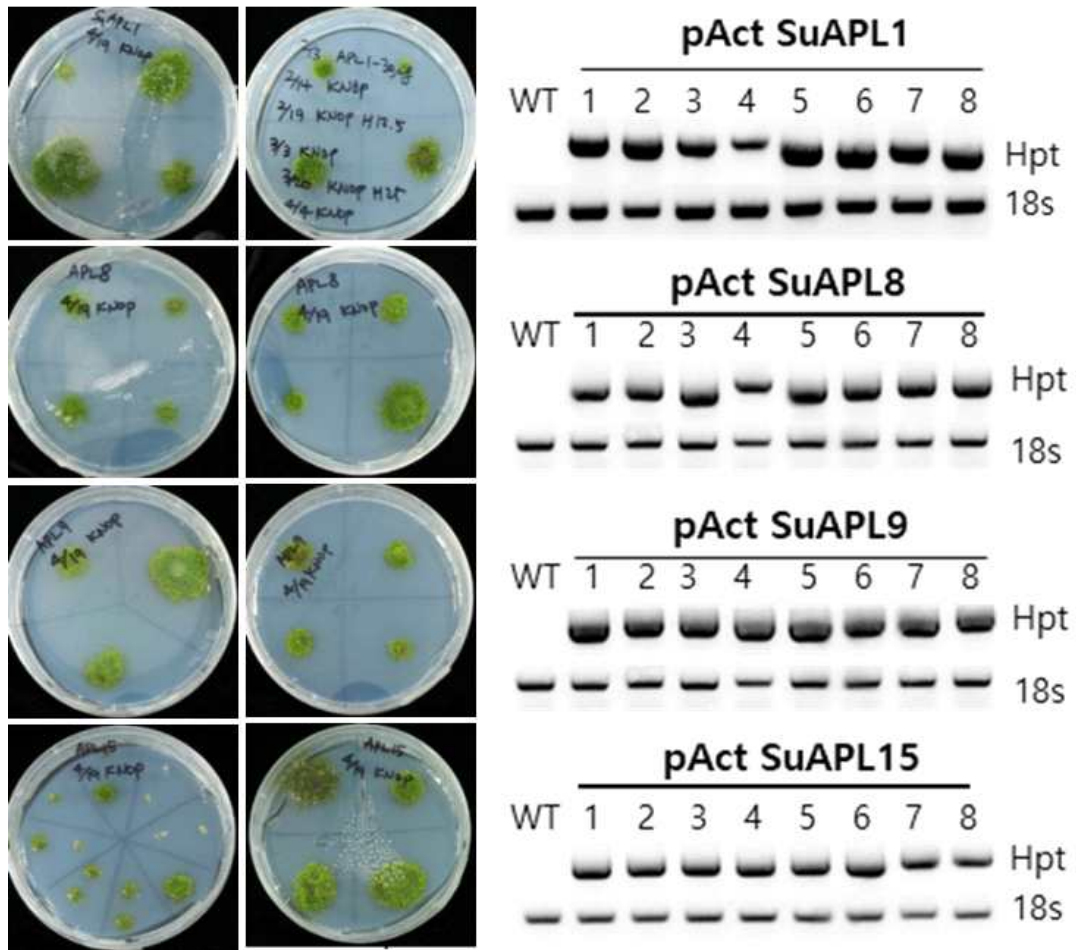


그림 4. Generation of pAct:SuAPL transgenic *Physcomitrella patens* plants. Genomic Polymerase chain reaction (PCR) analysis of transgenic lines and wild type (WT) moss. Genomic DNA was extracted from protonema cells and used for PCR to amplify hygromycin phosphotransferase (hpt).

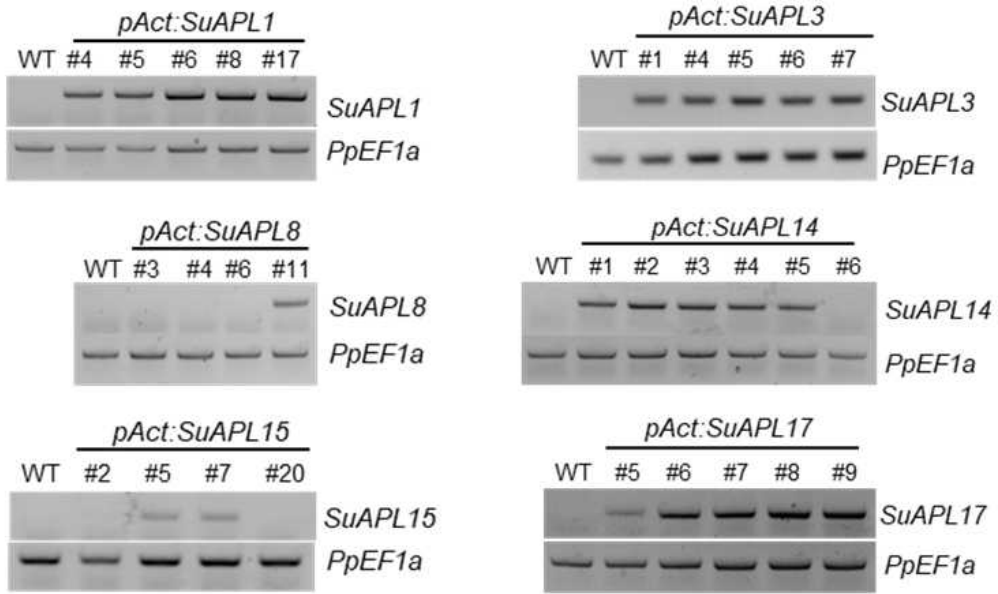


그림 5. Molecular analysis of pAct:SuAPL transgenic *Physcomitrella patens* plants. Reverse transcription (RT)-PCR analysis of SuAPLs expression in six transgenic plants. PpEF1a was used internal control.



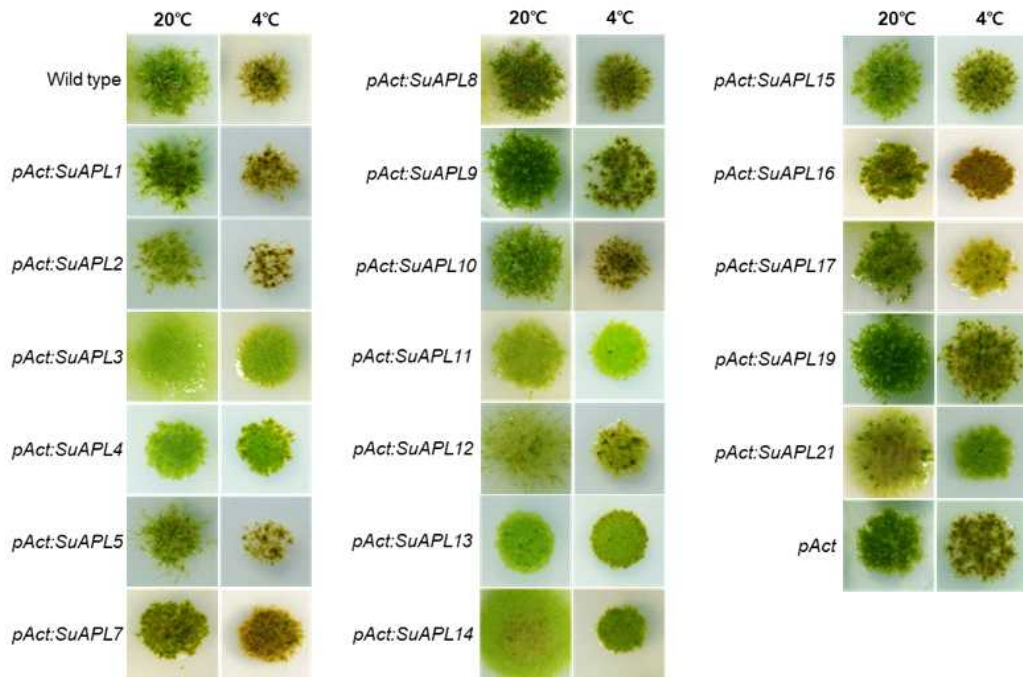
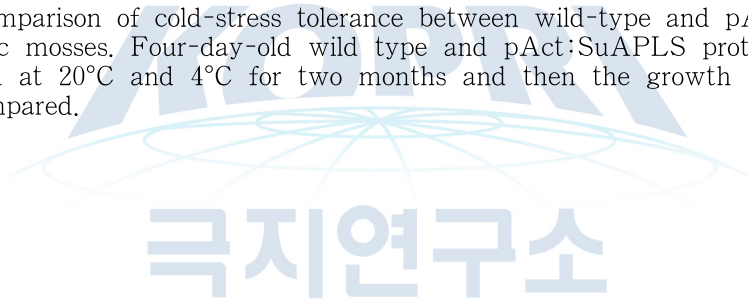


그림 6. Comparison of cold-stress tolerance between wild-type and pAct:SuAPLs transgenic mosses. Four-day-old wild type and pAct:SuAPLs protonema were incubated at 20°C and 4°C for two months and then the growth morphologies were compared.



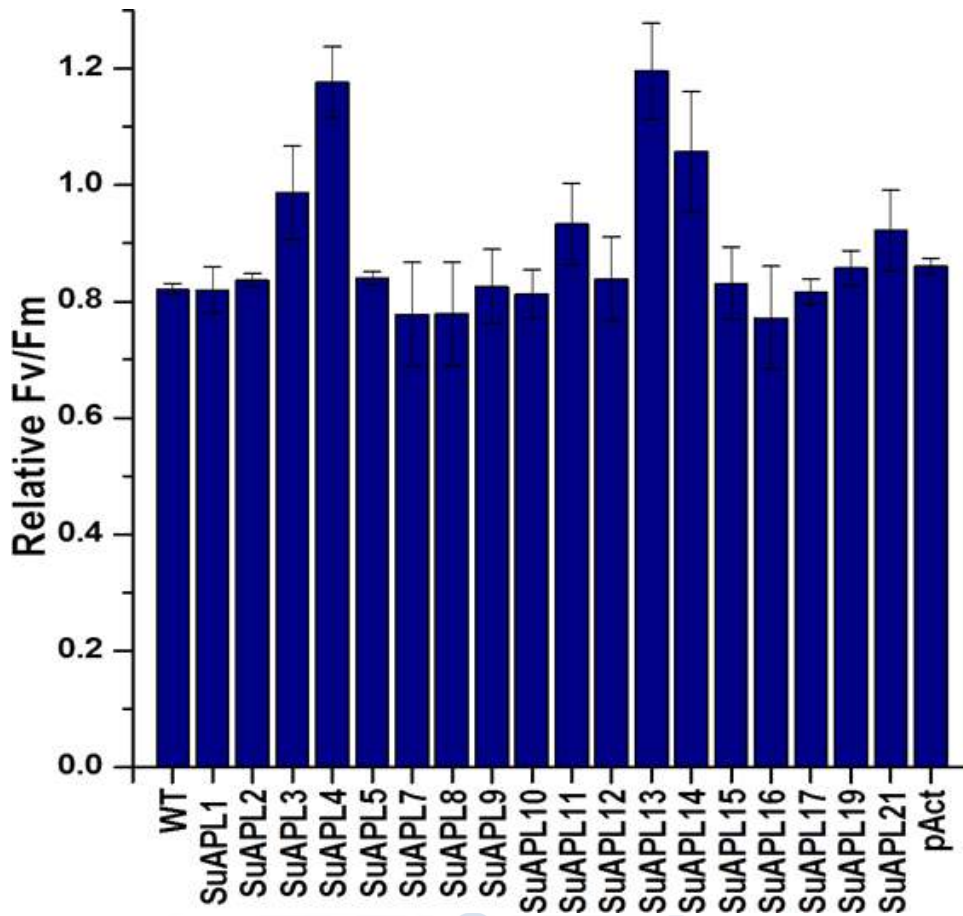


그림 7. Comparison of maximum quantum efficiency of PSII photochemistry (Fv/Fm) between wild-type and pAct:SuAPLs transgenic mosses. The ten biological replicates were used for each line (n = 10). Results are the means with \pm standard deviation shown by vertical bars.

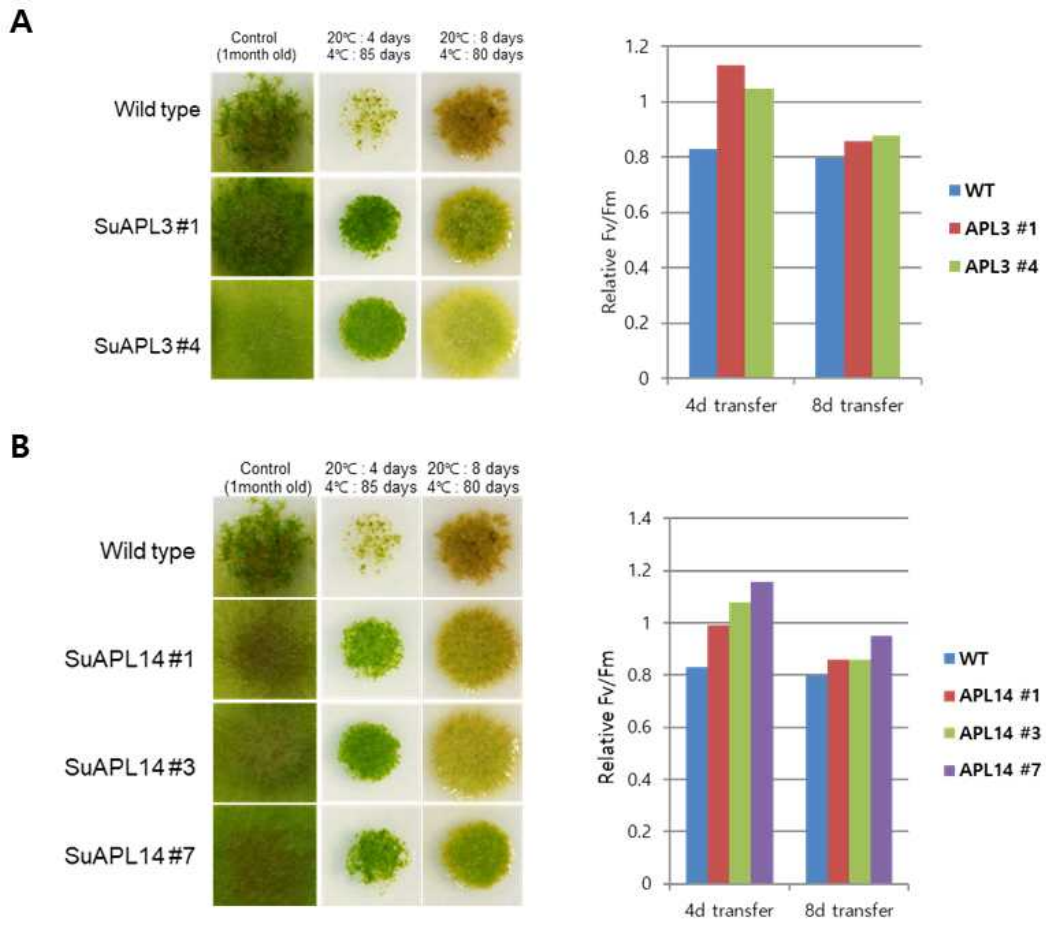


그림 8. Comparison of cold-stress tolerance morphologies and maximum quantum efficiency of PSII photochemistry (F_v/F_m) between wild-type, pAct:SuAPL3 and pAct:SuAPL14 transgenic mosses.

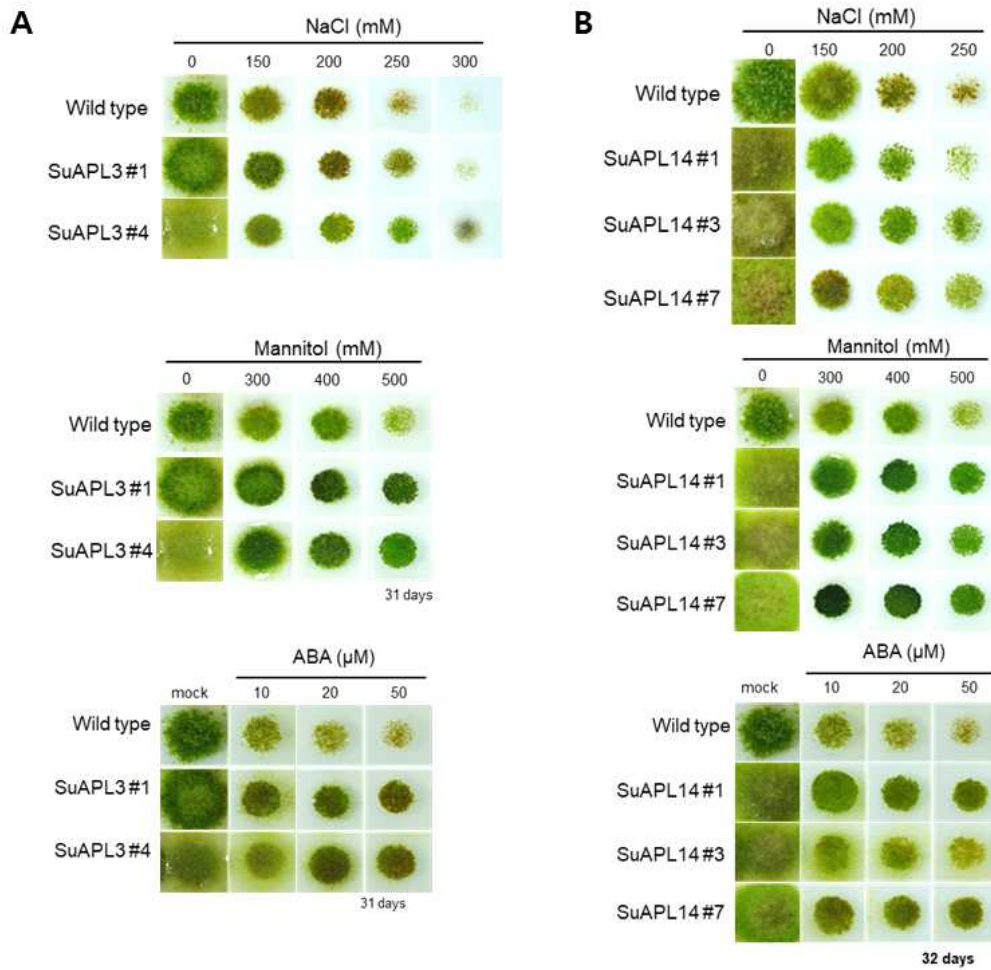


그림 9. Comparison of cold-stress tolerance between wild-type and pAct:SuAPLs transgenic mosses. Four-day-old wild type and pAct:SuAPLs protonema were incubated at 20°C and 4°C for two months and then the growth morphologies were compared.

제 3-7절 남극 쯔새풀 유래 Chalcone isomerase (DaCHI) 효소의 구조-기능 연구

- DaCHI 효소의 삼차구조 및 생화학적 특성 분석
- 저온 및 강한 UV 조건에서 DaCHI 효소 발현 증가 확인
- 남극 쯔새풀의 극지 환경적응에 DaCHI 효소의 역할 규명



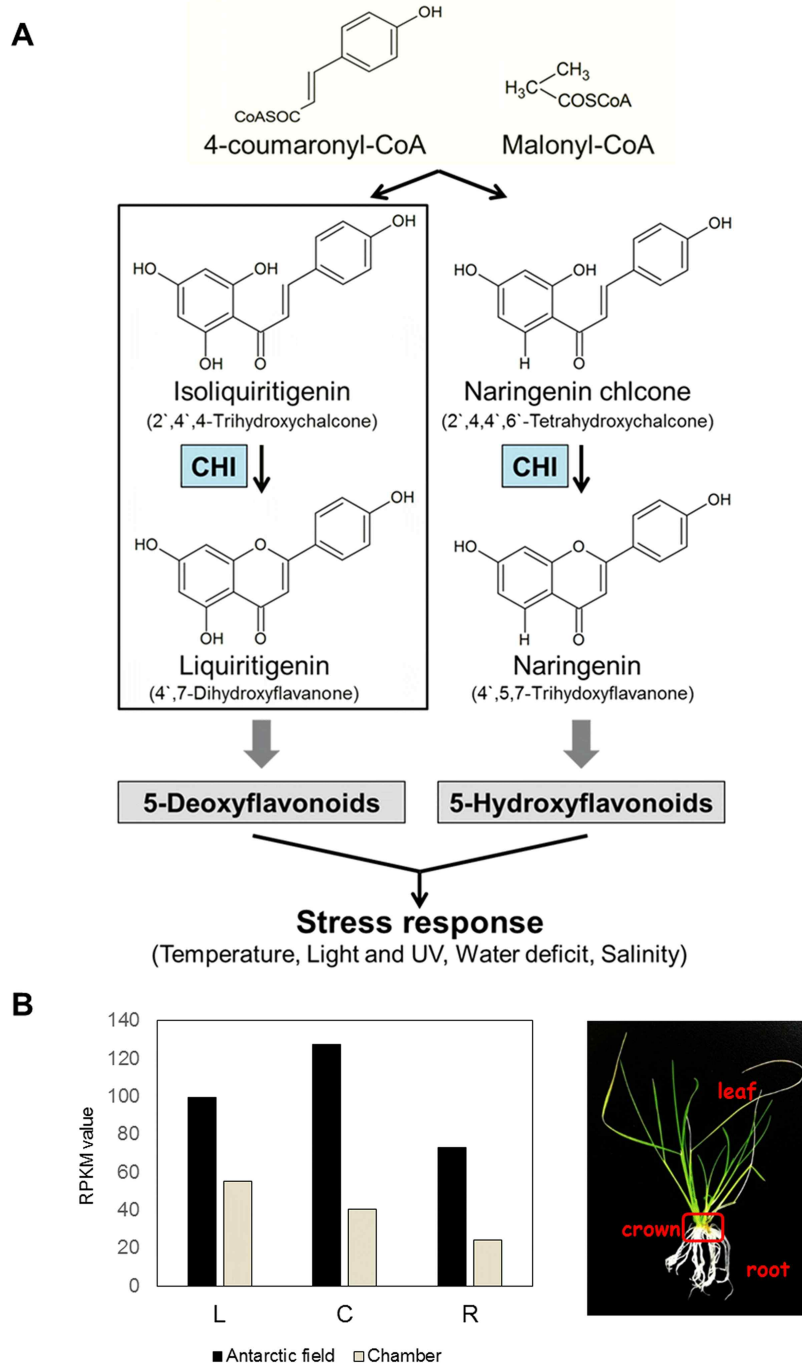


그림 1. (A) 남극 쉼새풀이 극지 환경스트레스에 적응할 수 있게 항산화 물질 합성에 관여하는 CHI 효소 기작 (B) 남극 쉼새풀의 조직별 CHI의 발현 패턴 분석

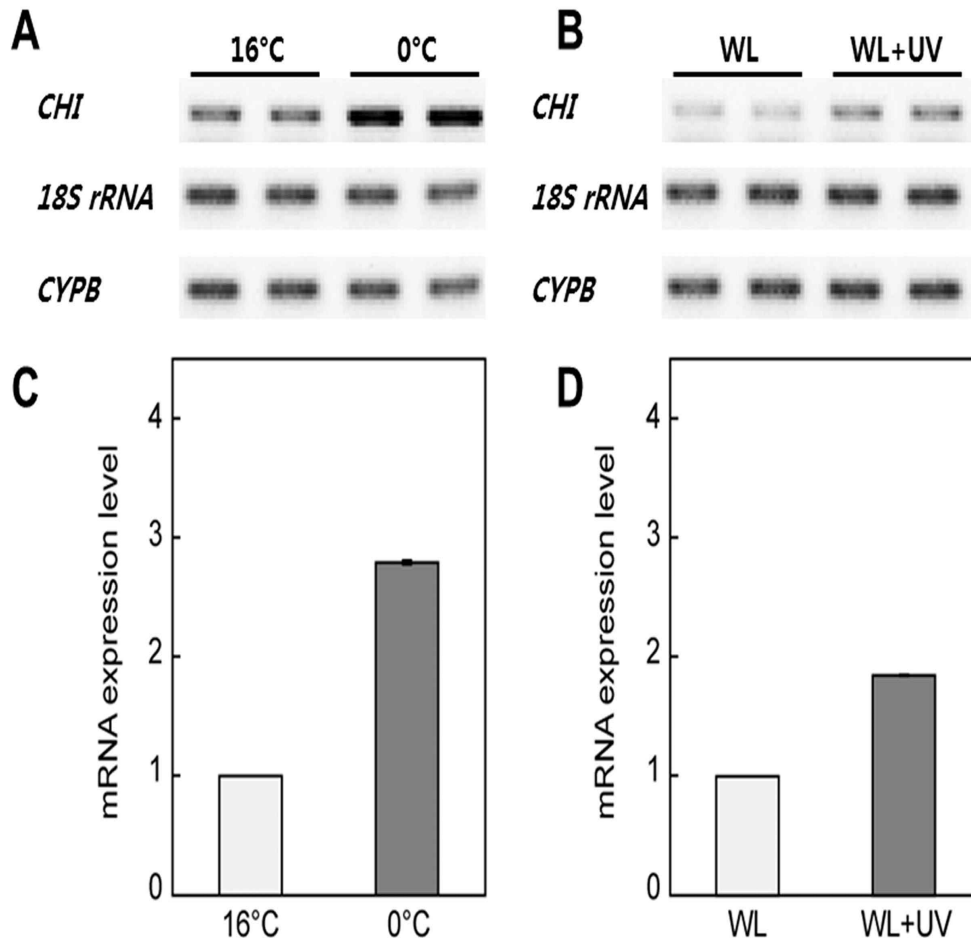


그림 2. 저온 (A and C) 및 강한 UV환경 (B and D) 에서 DaCHI의 발현 변화

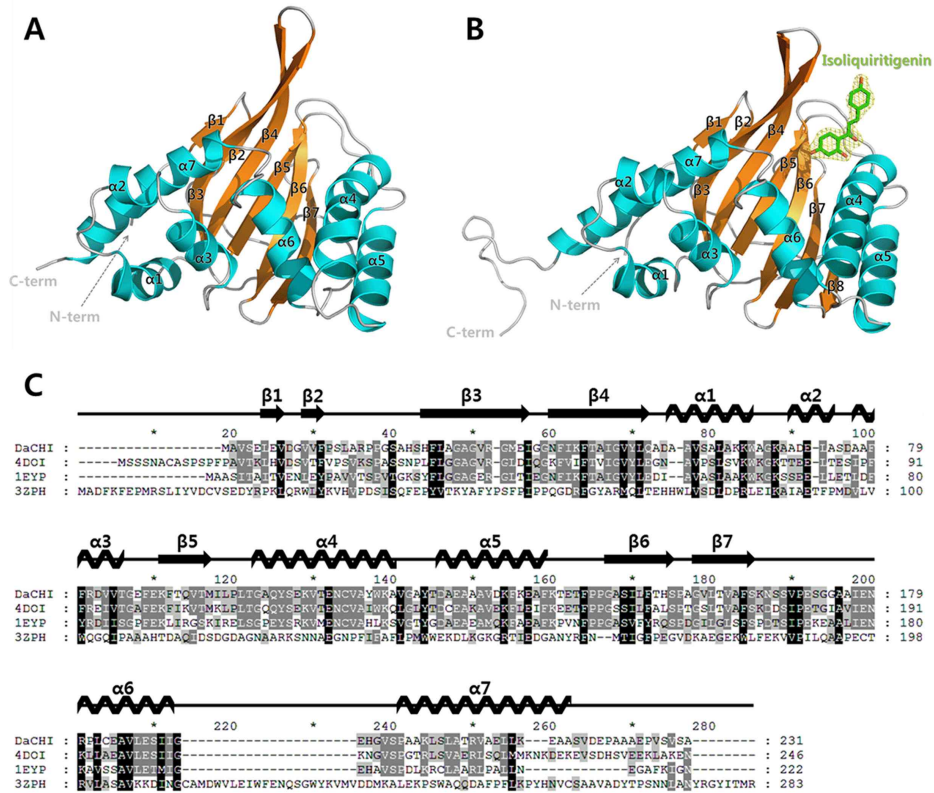


그림 3. (A and B) DaCHI 효소의 기질 (isoliquiritigenin) 복합체 삼차구조 (C) DaCHI 효소의 아미노산 서열과 기존에 알려진 다른 유사 CHI 효소와의 아미노산 비교

극지연구소

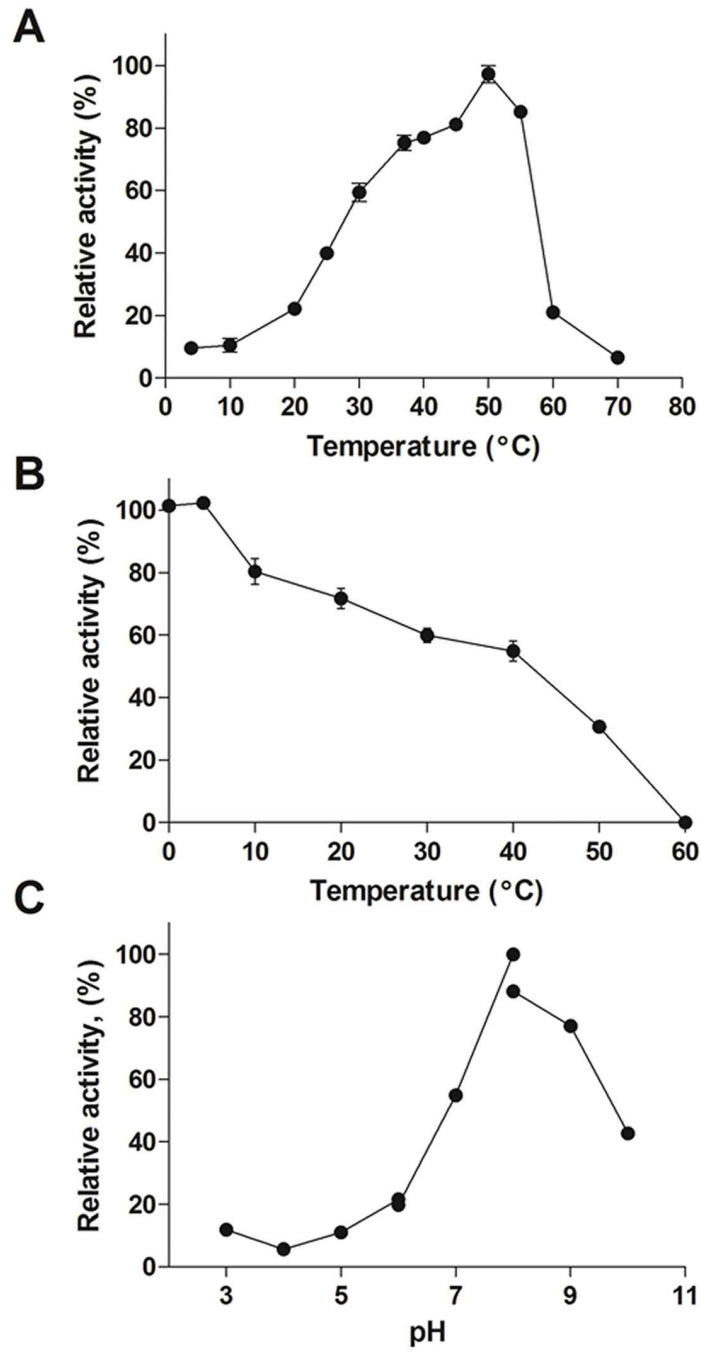


그림 4. DaCHI 효소의 생화학적 특성 연구 (A) 온도별 DaCHI 효소 활성 비교 (B) DaCHI의 온도에 따른 효소 활성 안정성 변화 (C) pH 조건에 따른 DaCHI 효소 활성 비교

제 3-8절 극지 호냉성 박테리아(*Exiguobacterium antarcticum*) 유래
EaEST 효소의 구조-기능연구

- 극지 미생물 유래의 저온성 esterase 효소의 삼차구조 및 특성 분석
- 저온성 esterase 효소는 정밀 화합물, 약품 산업의 생촉매제로 활용가능



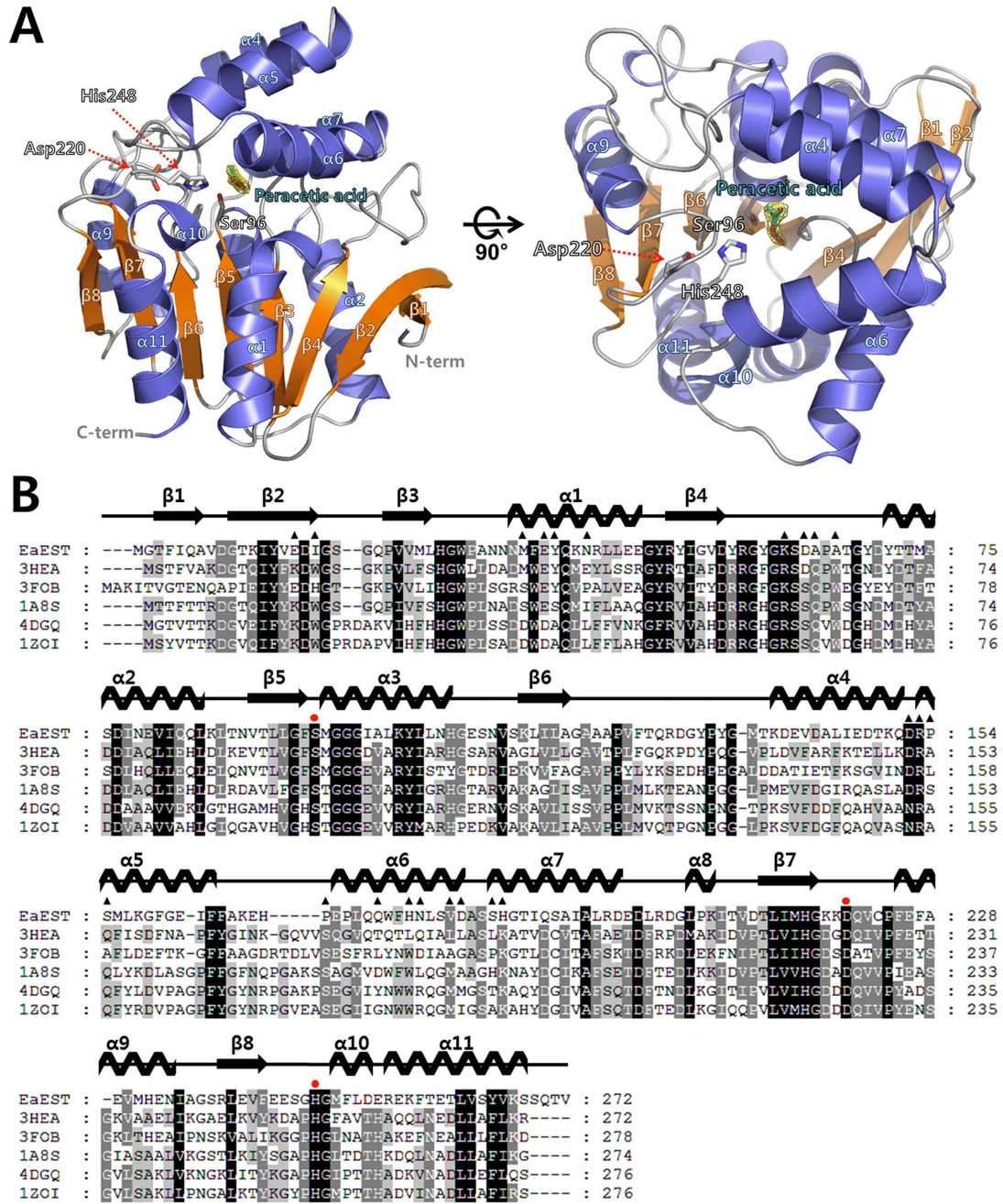


그림 1. EaEST 효소의 삼차구조와 활성잔기 위치 (A) 활성부위에 peracetate 가 결합된 EaEST 효소의 삼차구조 모습 (B) 저온성 EaEST 효소의 아미노산 서열과 기존에 알려진 다른 미생물에 존재하는 상동성이 높은 esterase 효소의 아미노산 비교, 붉은 동그라미로 표시된 잔기가 EaEST 효소의 catalytic triad 잔기

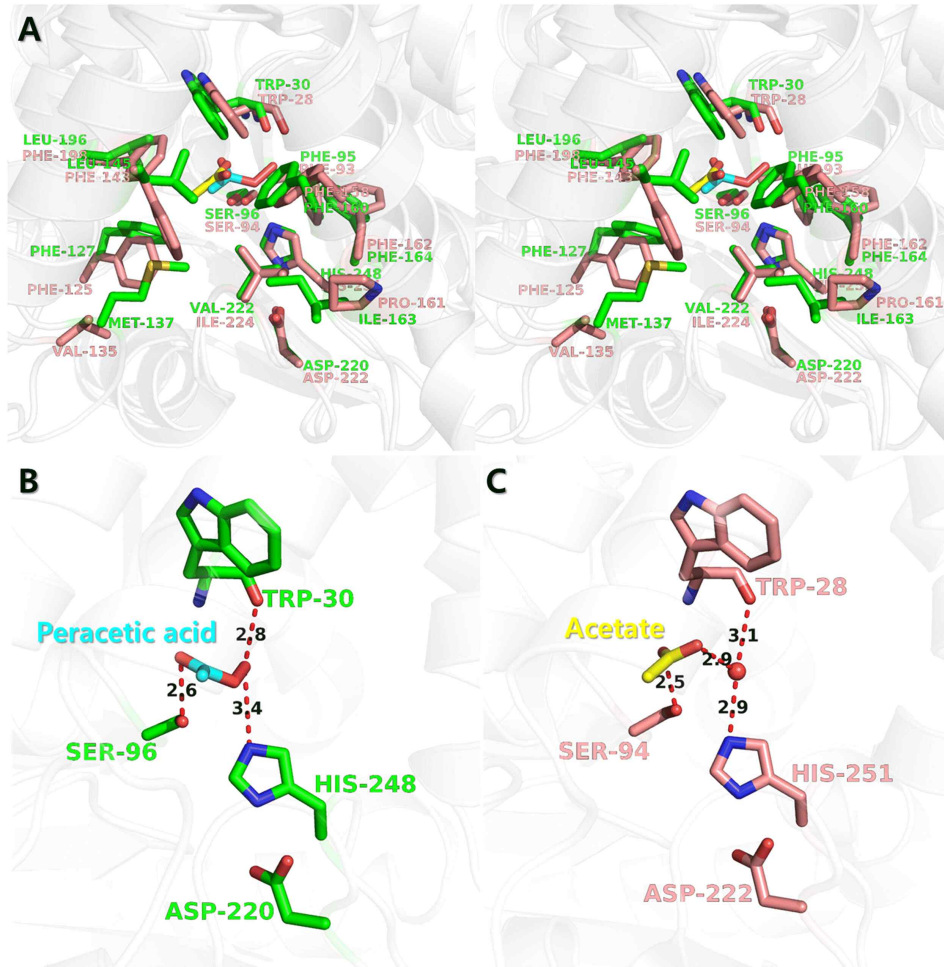


그림 2. EaEST 효소의 기질결합부위 (A) EaEST 효소의 기질결합 부위와 기존에 알려진 PfEST (*Pseudomonas fluorescens* esterase) 기질 결합부위 구조 비교, EaEST 효소의 기질결합 부위에 peracetic acid가 결합되어 있음. 이는 EaEST가 perhydrolase activity를 가진다는 사실을 나타냄 (B) 저온성 EaEST 효소에서 Peracetic acid 결합에 중요한 잔기를 나타내는 그림 (C) PfEST 효소에서 Acetate 결합에 중요한 잔기를 나타내는 그림. 가운데 붉은 색의 구형태인 structural water 가 Acetate 결합에 중요한 역할 수행

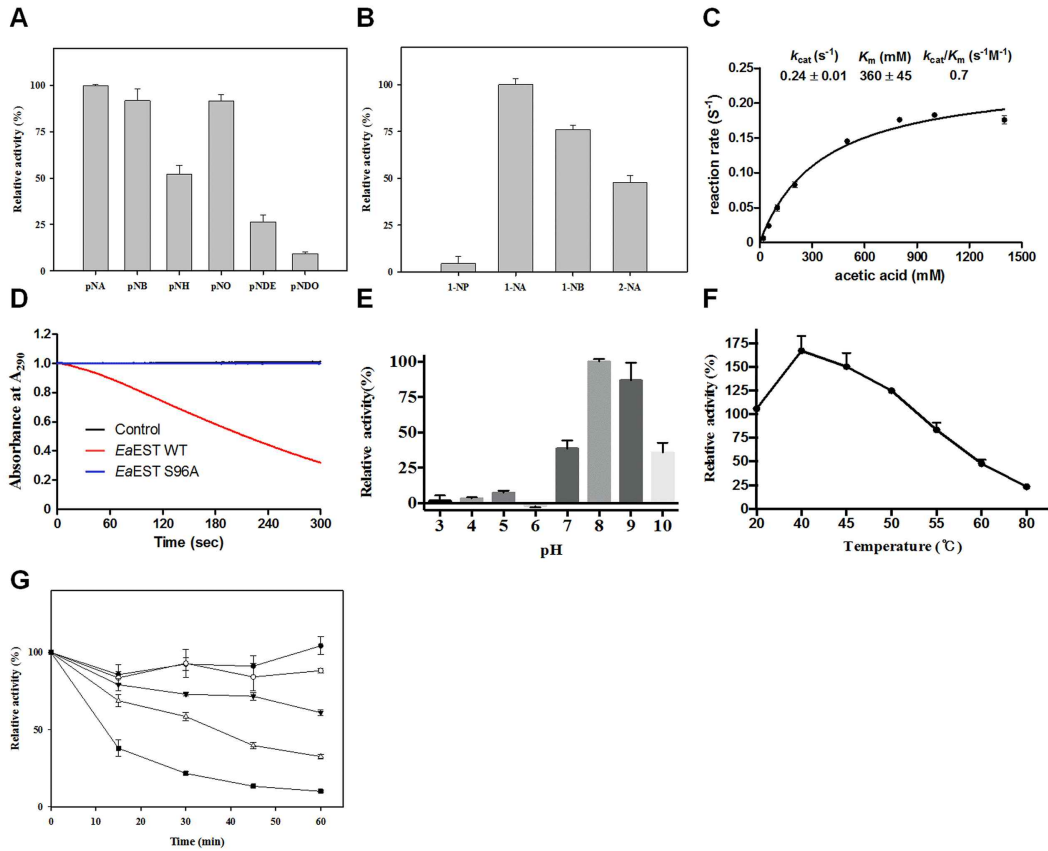


그림 3. EaEST 효소의 활성특성 규명 (A) EaEST 효소의 기질특이성 (B) EaEST 효소의 naphthyl derivatives에 대한 기질특이성 (C) EaEST 효소의 Perhydrolase 활성 (D) EaEST 효소 wild-type 과 S96A 돌연변이체의 활성비교. S96A 돌연변이체는 Perhydrolase 활성을 잃어버림 (E) pH 조건별 EaEST 효소 활성 비교 (F) 온도별 EaEST 효소 활성 비교 (G) EaEST 효소의 열안정성 실험

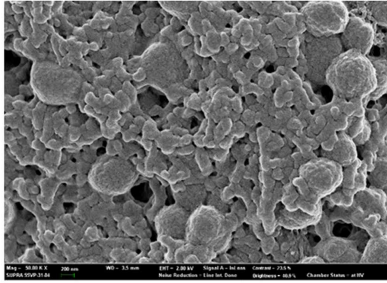
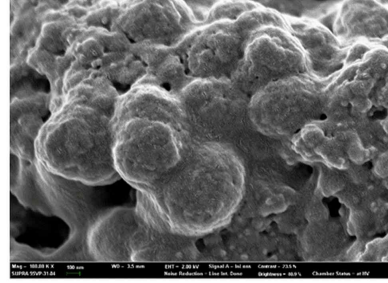
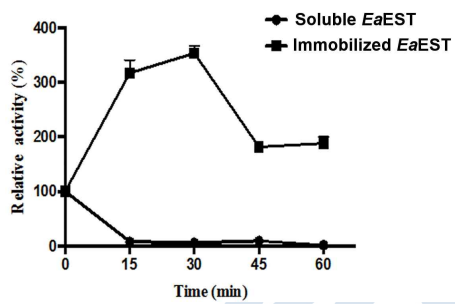
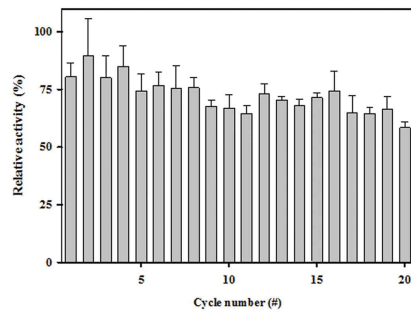
A**B****C****D**

그림 4. EaEST 효소의 고정화 연구 (A) Cross-linking에 의해 고정화된 EaEST 효소의 전자현미경 사진 (B) 고정화된 EaEST 효소의 고배율 전자 현미경 사진 (C) 용액상과 고정화된 상태에서의 EaEST 효소의 활성유지 시간 비교 (D) 고정화된 EaEST 효소는 20번의 효소반응 후에도 약 60% 이상의 효소 활성을 유지

극지연구소

제 3-9절 바실러스 PAMC 23377 균주 유래 BaCYP106A2 효소의 구조와
기능적 특성연구

- Steroid 수산화 효소는 의약품 개발에 이용 될 수 있어 최근 많은 관심을 받고 있음
- 북극 Kara Sea에서 분리한 Bacillus sp. PAMC 23377 미생물로 부터 신규 CYP steroid 수산화 효소 유전자를 발굴하고, 효소활성 특성 및 삼차구조를 규명함



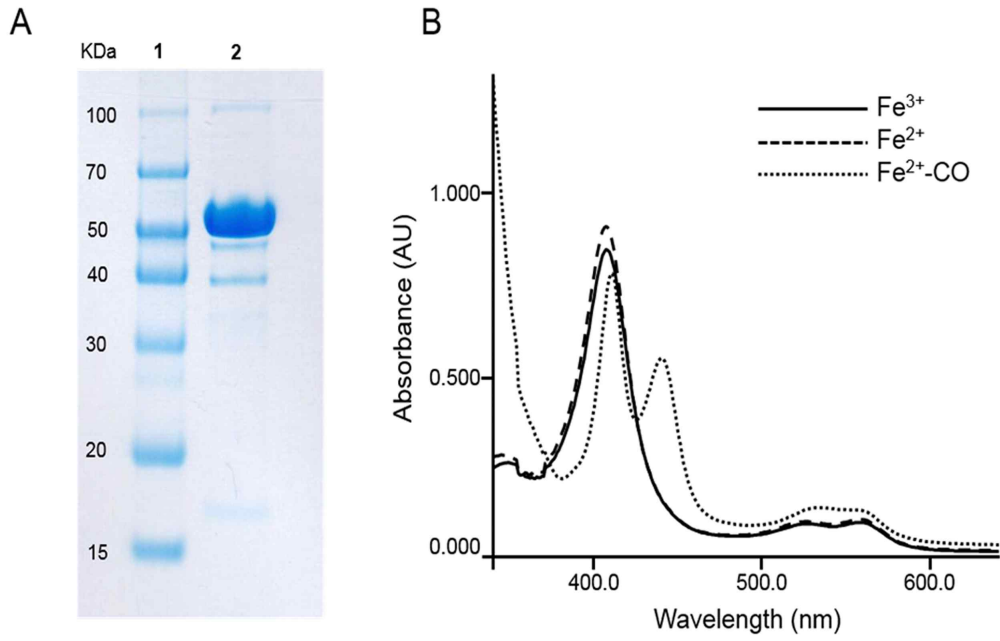
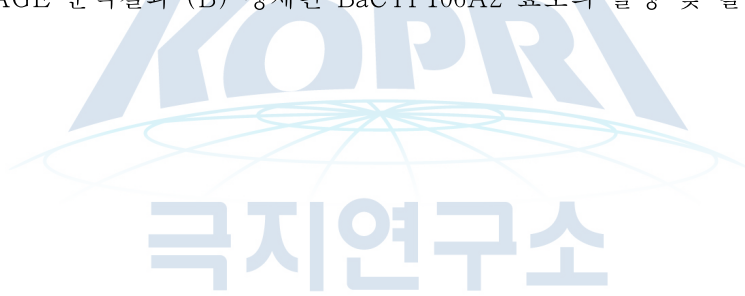


그림 1. 재조합 BaCYP106A2 효소 생산 (A) 정제된 BaCYP106A2 효소의 SDS-PAGE 분석결과 (B) 정제된 BaCYP106A2 효소의 폴딩 및 활성 확인



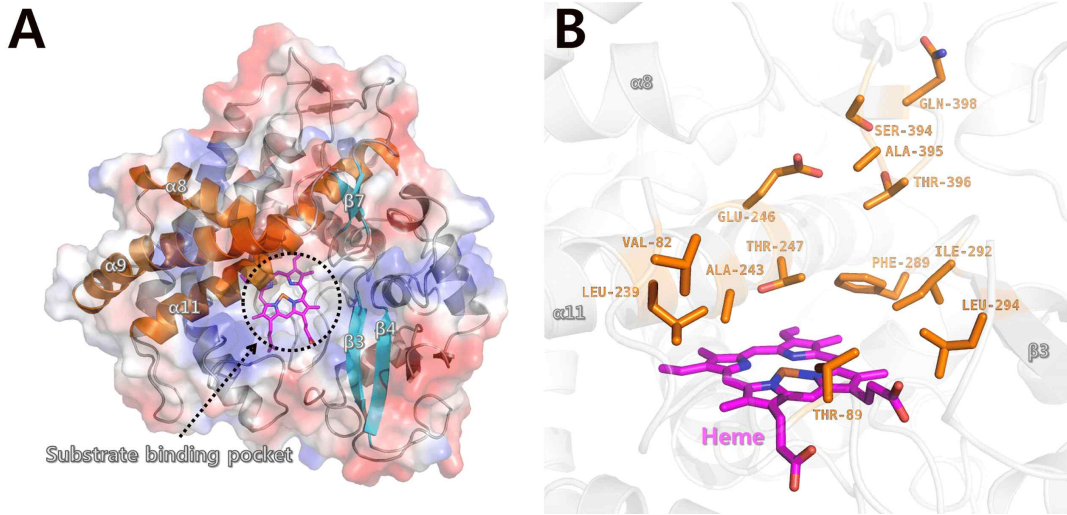
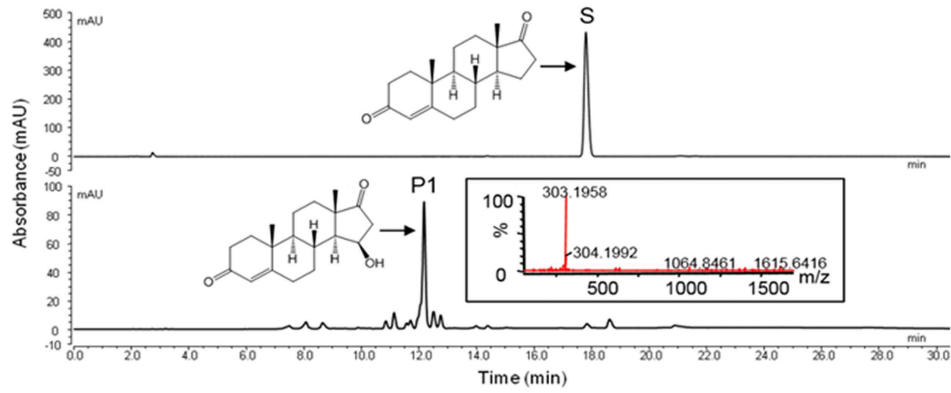


그림 2. BaCYP106A2 효소의 삼차구조 모습 (A) BaCYP106A2 효소에 결합되어 있는 heme 분자와 예상되어지는 기질 결합부위 (B) 기질결합에 관여할 것으로 생각되어지는 소수성 잔기들의 모습



A



B

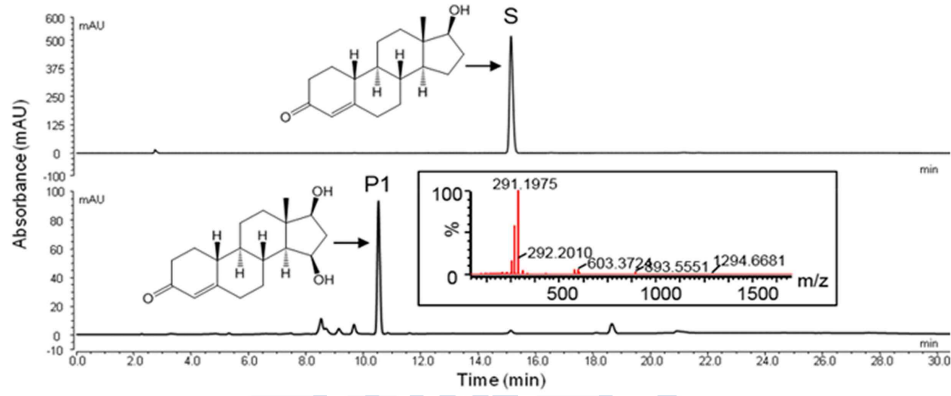


그림 3. HPLC 와 LC-MS 분석을 통한 BaCYP106A2 효소의 기질특이적 활성 확인 (A) 4-Androstenedione 기질에 대한 BaCYP106A2 효소의 수산화 부위 확인 (B) Nandrolone substrate에 대한 BaCYP106A2 효소의 수산화 부위 확인

제 3-10절 북극 해양 미생물(*Colwellia psychrerythraea*) 유래 isoaspartyl dipeptidase 효소(CpsIadA)의 삼차구조 및 기능적 특성 규명

- CpsIadA의 효소 활성부위 잔기 구성에서 기존에 알려진 중온성 IadA 효소와의 차이점을 밝히고, 이 차이점 때문에 isoaspartyl dipeptidase 효소가 두 개의 Type으로 나누어진다는 사실 규명
- 저온성 isoaspartyl dipeptidase 효소는 정밀 화합물 합성 산업의 생촉매제로 활용 가능



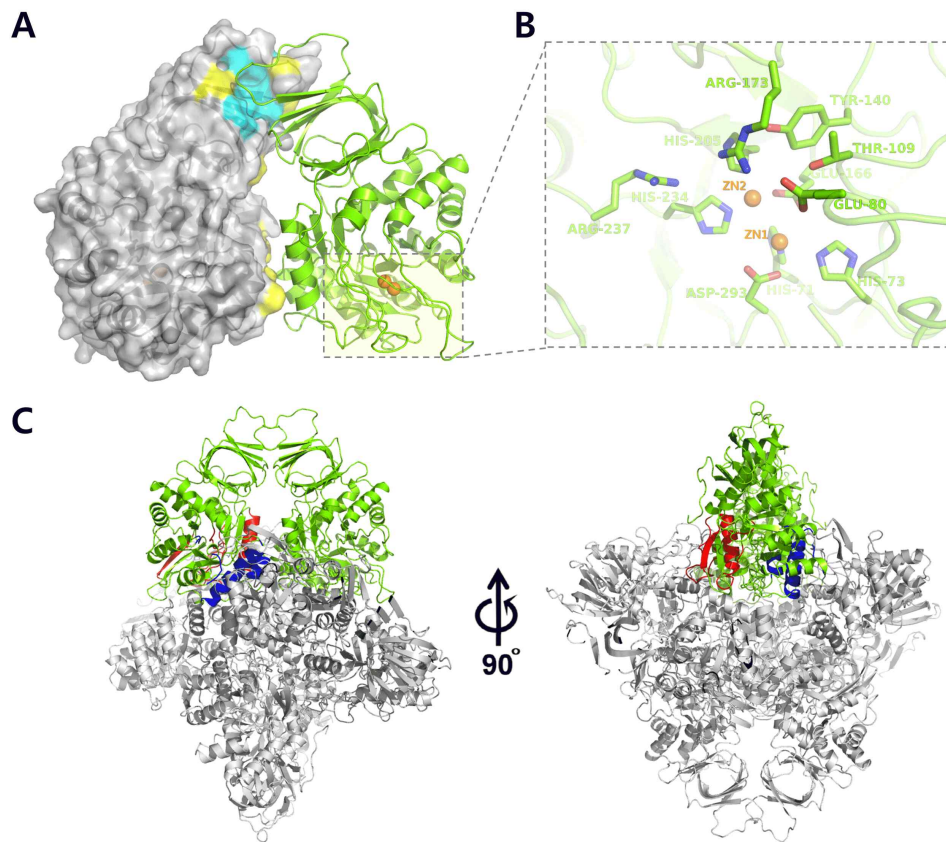


그림 1. CpsIadA 효소의 삼차구조 모습 (A) CpsIadA 효소의 dimer 형성 부위 (B) CpsIadA 효소의 활성 부위에 존재하는 Zn metal ion의 모습 (C) CpsIadA 효소의 octamer 구조 모습. 용액상에서 CpsIadA 효소는 octamer 형태로 존재

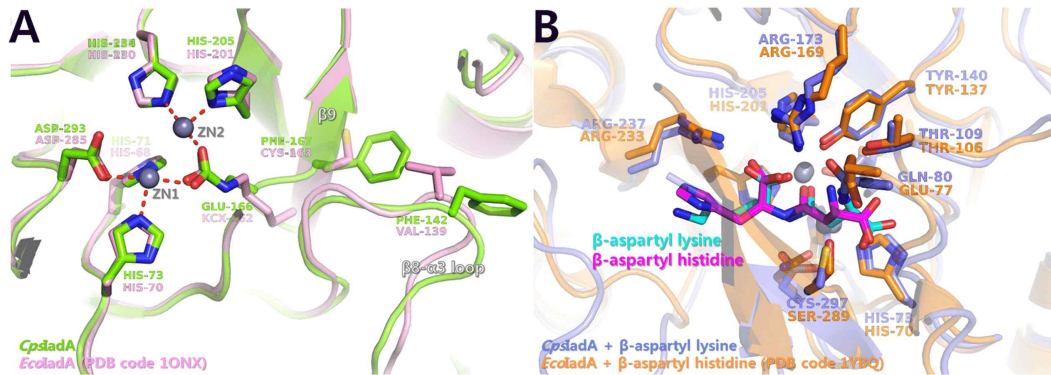


그림 2. 저온성 CpsIadA 효소의 구조와 기존에 알려진 중온성 EcoladA 효소의 구조 비교 (A) 저온성 CpsIadA 효소의 활성부위 구조와 EcoladA 효소의 활성부위 비교. EcoladA 효소는 post-translationally 수식화된 carbamylated-lysine 162 잔기를 가지는데 비해 CpsIadA 효소는 같은 자리에 Glutamate 166 잔기를 가짐 (B) CpsIadA 효소의 기질결합 부위에 Beta-Aspartyl lysine이 결합된 구조와 기존의 EcoladAdp 결합된 Beta-Aspartyl histidine 복합체 구조와의 비교 그림. CpsIadA 의 삼차구조 정보로부터 CpsIadA 의 활성부위는 기존에 알려진 중온성 미생물인 대장균 유래 IadA 효소 (EcoladA) 와 다른 효소 활성잔기를 가진다는 사실 규명. EcoladA 효소는 post-translationally 수식화된 carbamylated-lysine 162 잔기를 가지는데 비해 CpsIadA 효소는 같은 자리에 Glutamate 166 잔기를 가짐으로써 EcoladA 와는 다른 활성 기작을 가짐. 이런 차이점 때문에 isoaspartyl dipeptidase 효소가 두 개의 Type 으로 나누어진다는 사실 규명. 즉, 효소 활성 부위에 Glutamate 잔기를 가지는 Type I IadA 그룹과 그자리에 carbamylated-lysine 잔기를 가지는 Type II IadA 그룹으로 구분가능. 본 연구에서는 최초의 Type I IadA 구조를 규명한 것으로 정밀 화합물 제조에 이 효소를 활용하고자 할때에는 CpsIadA 는 carbamylation 이라는 효소 활성 과정이 필요없다는 장점을 가짐.

극지연구소

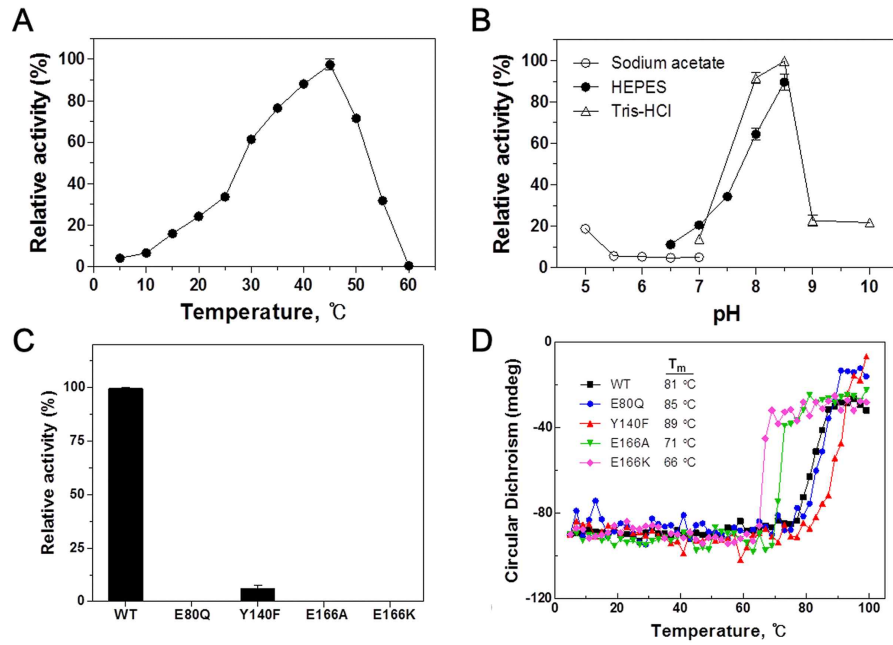


그림 3. CpsIadA 효소의 생화학적 특성 분석 (A) 온도별 CpsIadA 활성 분석 (B) pH별 CpsIadA 활성 분석 (C) CpsIadA 효소 활성부위 잔기들의 mutation에 의한 효소 활성 변화 (D) CpsIadA 효소 돌연변이체에 대한 열안정성 분석

극지연구소

제 3-11절 남극 호냉성 미생물 유래의 저온성 esterase 효소의 삼차구조 해석

- peracetate 가 결합된 고해상도 EaEST 삼차구조 해석
- Analytical ultracentrifugation 실험으로 용액상에서 EaEST가 trimer를 형성한다는 사실 규명
- EaEST 구조에서 trimer 형성에 중요한 잔기 확인
- peracetate 결합에 의한 EaEST 구조 변화 관찰

2) EaEST 효소의 생화학적 특성 규명

- 효소 활성실험으로 EaEST가 esterase 활성 뿐만아니라 perhydrolase 활성을 가진다는 사실 규명
- 구조비교를 통해 효소 활성에 중요한 catalytic triad 와 기질결합에 중요한 잔기 확인
- EaEST효소의 조건별 (온도, pH, Urea) 효소 활성 비교

3) 성과의 우수성 및 의의

- 세계 최초의 호냉성 미생물 유래 esterase 구조 및 생화학적 특성 규명
- EaEST 효소의 고정화 연구를 통해 Cross-linking에 의해 고정화된 EaEST 효소의 전자현미경 사진 확보, 고정화된 EaEST 효소는 20번의 효소반응 후에도 약 60% 이상의 효소 활성을 유지한다는 사실 발견
- 저온성 EaEST 효소의 산업적 활용 가능성 확인

(2) 북극 해양미생물 유래 oligoribonuclease (CpsORN) 효소의 구조분석 및 small RNA 분해기작 규명, 논문 발표, 2019, Scientific Reports (IF: 4.122, mrnIF 82.540)

1) 북극 해양미생물 유래 oligoribonuclease (CpsORN) 효소의 구조분석 및 small RNA 분해기작 규명

- 대표적 저온성 미생물인 *Colwellia psychrerythraea*에서 유래한 oligoribonuclease (CpsORN) 단백질의 삼차구조연구
- 미생물이 신속하게 외부환경 변화에 반응하기 위해서는 RNA의 조절이 중요함

- RNA 분해기작에는 다양한 ribonuclease 작용함
- CpsORN 단백질은 작은 크기의 (≤ 5) RNA만을 분해함
- RNA와의 결합구조 규명을 통해 CpsORN의 RNA 분해 기작 설명

2) 성과의 우수성 및 의의

- 최초의 저온성 ORN 단백질에 대한 구조적 특성 분석 결과임
- 최초의 저온성 ORN 단백질과 RNA와의 결합 구조 규명
- 저온성 미생물은 외부 환경에 반응하기 위해 RNA 조절을 하며, CpsORN은 RNA 분해 과정에 필수적임
- CpsORN 구조 정보를 이용해 phosphodiester 결합의 분해과정을 이해하고, 이를 이용하여 화합물변형 디자인이 가능함
- 우수 논문 발표, Scientific Reports (IF: 4.122, mrnIF 82.540)



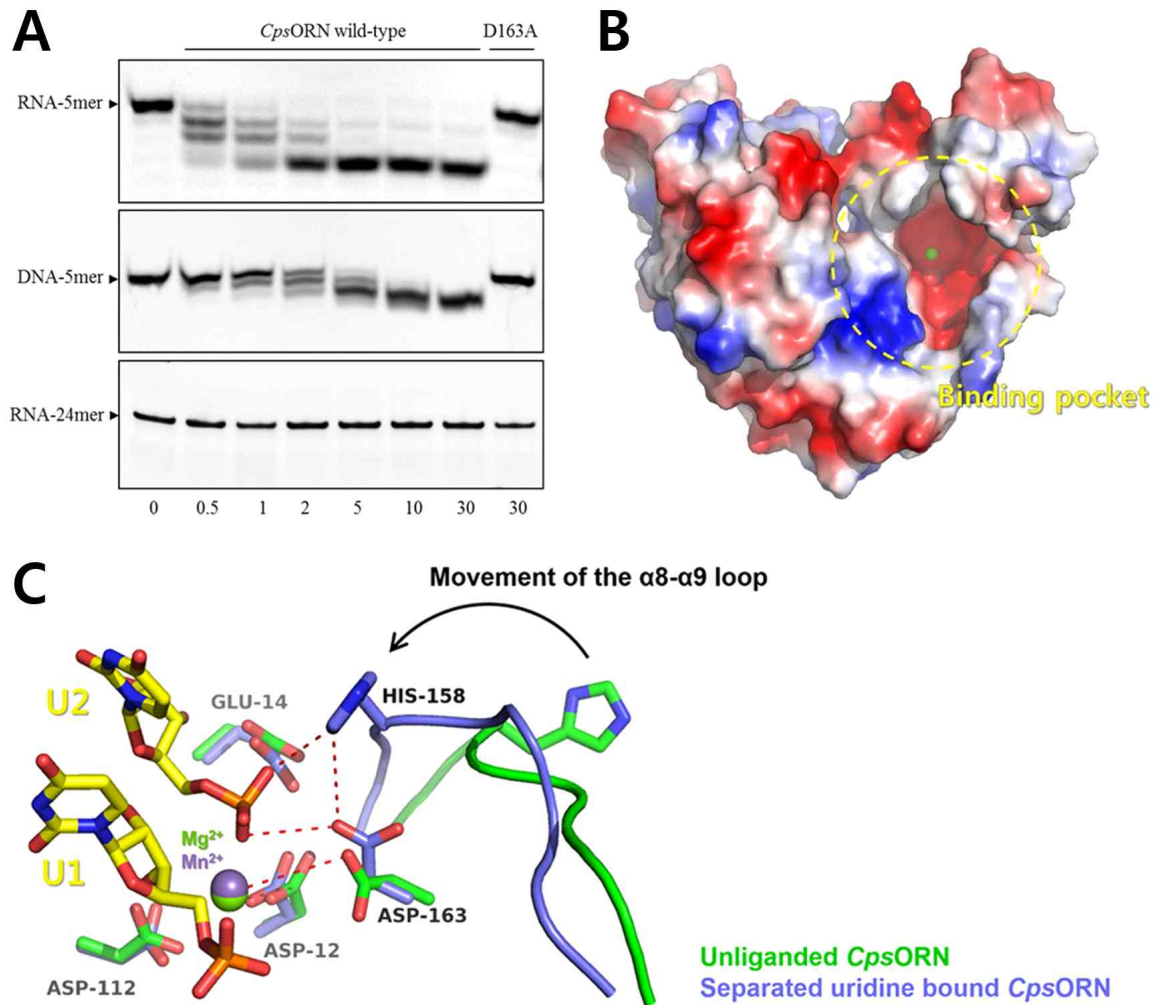


그림 1. (A) 작은 크기의 (≤ 5) RNA만을 분해하는 *Cps*ORN (B) 이량체를 이루고 있으며 RNA가 결합하는 틈을 지니고 있는 *Cps*ORN의 구조 (C) RNA의 결합 여부에 따라 α 8- α 9 고리 부분의 움직임을 보이며, 이 부분이 분해 활성화에 중요한 역할을 함

제 3-12절 극지 특이적 환경요인에 의한 단백질 지도 제작

가. 남극 저온성 박테리아 유래 cold shock protein이 도입된 박테리아의 표준 단백질 지도

1) 서론

남극 저온성 박테리아 *Psychrobacter* sp. PAMC 21119는 남극 동토로부터 분리되었다(Kim, Shin et al. 2012). 이 균주는 영하의 온도(-5도)에서도 성장하였고 영하의 온도 조건에서 높은 발현율을 보인 cold shock protein (ppCspC, ppCspA)을 확인하였다. cold shock protein (ppCspC, ppCspA) 유전자를 저온 스트레스에 민감하게 반응하는 BX04 세포(CspA, B, E, G 결핍)에 형질 전환하고 박테리아의 단백질 지도를 구축하였다.

2) 연구수행내용

- 샘플 준비

저온 스트레스에 민감하게 반응하는 BX04 세포(CspA, B, E, G 결핍)에 cold shock protein을 형질 전환한다. 이때 사용하는 세포는 액체 배양 후 세포수를 계산하여 동량의 균질한 세포액을 카나마이신 배지에 접종한다. BX04 세포와 형질 전환된 세포를 15도 인큐베이터에서 배양한다.

- 단백질 추출

콜로니를 회수 한 후 2-DE lysis solution (7 M urea, 2 M thiourea, 1% (w/v) dithiothreitol (DTT), 4% (w/v) 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate hydrate (CHAPS), 2% (v/v) ampholyte, and 1 mM benzamidine-HCl)을 넣어주고 갈아준다. 14도에서 14,000rpm으로 30분간 원심 분리 한 후 상층액의 단백질만 따로 새 튜브에 옮긴다. 단백질의 농도는 Bradford method를 이용하여 측정한다 (Bradford, 1976).

- 2차 전기영동

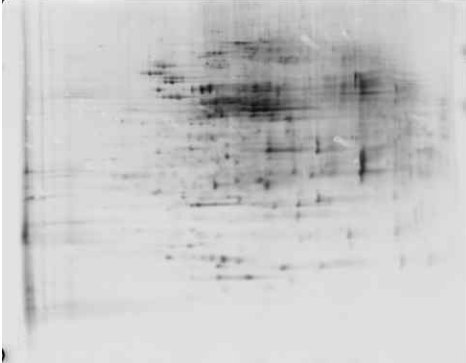
농도 측정을 끝낸 단백질 샘플은 cup-loading method를 이용하여 Immobiline DryStrip gels에 등전점을 기준으로 하여 1차 전기영동을 수행한다. 1차 전기영동을 마친 strip은 SDS-PAGE (10 - 15% gradient, 26 × 20 cm)에 옮겨 2차 전기영동을 실시한다. 단백질 스팟은 은염색법을 이용하여 확인한다 (Mykytczuk et al., 2011).

참고문헌

1. Kim, S. J., S. C. Shin, S. G. Hong, Y. M. Lee, I.-G. Choi and H. Park (2012). Genome sequence of a novel member of the genus *Psychrobacter* isolated from Antarctic soil, *Am Soc Microbiol*.
2. Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72: 248-254.
3. Mykytczuk, N.C., Trevors, J.T., Foote, S.J., Leduc, L.G., Ferroni, G.D., and Twine, S.M. (2011) Proteomic insights into cold adaptation of psychrotrophic and mesophilic *Acidithiobacillus ferrooxidans* strains. *Antonie van Leeuwenhoek* 100: 259-277.



15°C Bacteria



15°C Bacteria + cold shock protein

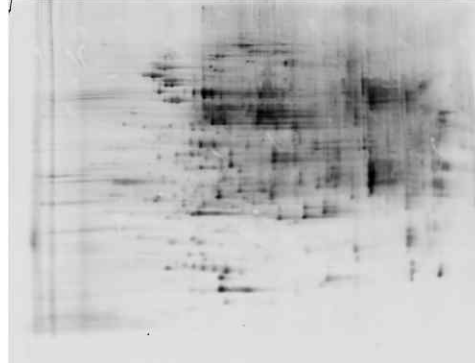


그림 1. 남극 저온성 박테리아 유래 cold shock protein이 도입된 박테리아의 표준 단백질 지도



나. 남극빙하 소실에 의해 유도되는 해수의 염도감소에 따른 주요 해양생물인 남극 해수 갑각류 *Gondogeneia antarctica*의 단백질 분석

1) 서론

지구온난화는 오늘날 세계가 직면하고 있는 가장 중요한 환경문제 중 하나이다. 지구온난화가 계속되면 해수면 상승은 점점 가속도가 붙게 된다. IPCC 보고서에 따르면 그린란드와 남극 빙상이 녹으면서 해수면 상승 속도가 빨라지고 있다(Stocker, Plattner et al. 2014). 해수면 상승은 저지대 해안에 사는 사람뿐만 아니라 연안생태계에도 영향을 미칠 것으로 생각된다. 세종기지 인근 빙하소실로 인한 담수도입에 의해 기지 앞바다의 염도가 25ppt 수준으로 수차례 감소한 기록을 확인하였다. 남극의 주요 해양생물인 곤도지니아를 이용하여 해양 염도 변화에 의해 유도되는 단백질 발현의 변화를 단백질 지도로 구축한다.

2) 연구수행내용

- 치사율 실험

세종기지 주변에서 채집한 10마리의 곤도지니아를 해수가 담긴 비이커에 준비한다. 기존 염도의 해수에 세종기지 주변 염도 31.5ppt에서 시작하여 5~10%씩 한시간 간격으로 변화를 주며 관찰하였다. 염도를 감소시키는데 사용한 담수는 기지 주변의 눈을 녹여 사용하였다.

- 시료 준비

기지 앞바다에서 과거 25ppt 수준으로 수차례 감소한 기록을 토대로 3포인트 염분 조건 (26, 30, 34 ppt)을 가지고 실험 준비를 하였다. 3개의 비이커를 준비하여 각각의 염도 조건을 맞추고 채집한 곤도지니아를 넣어 24시간, 48시간 후 샘플링 하였다.

염도별로 샘플링한 곤도지니아는 2-DE lysis solution (7 M urea, 2 M thiourea, 1% (w/v) dithiothreitol (DTT), 4% (w/v) 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate hydrate (CHAPS), 2% (v/v) ampholyte, and 1 mM benzamidine-HCl)을 넣어주고 갈아준다. 14도에서 14,000rpm으로 30분간 원심 분리 한 후 상층액의 단백질만 따로 새 튜브에 옮긴다. 단백질의 농도는 Bradford method를 이용하여 측정한다 (Bradford, 1976).

- 2차 전기 영동

농도 측정을 끝낸 단백질 샘플은 cup-loading method를 이용하여 Immobiline

DryStrip gels에 등전점을 기준으로 하여 1차 전기영동을 수행한다. 1차 전기영동을 마친 strip은 SDS-PAGE (10 - 15% gradient, 26 × 20 cm)에 옮겨 2차 전기영동을 실시한다. 단백질 스팟은 은염색법을 이용하여 확인한다. 치사율 실험 결과, 23.2ppt부터 죽는 개체를 확인하였다. 100% 치사율을 보인 염도는 18.2ppt임을 확인하였다. 남극 해양생물인 곤도지니아가 염도의 농도에 따라 변화를 보인 단백질 지도를 구축하였다.

참고문헌

1. Stocker, T., G.-K. Plattner and Q. Dahe (2014). IPCC climatechange 2013: the physical science basis-findings and lessons learned. EGU General Assembly Conference Abstracts.
2. Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72: 248-254.



〈표. 염분 스트레스에 따른 *Gondogeneia antarctica*의 치사율〉

처리 시간	염도	비율	치사개체수	관찰 시간
PM 7 : 37	31.5ppt	5%		PM 8 : 37
PM 8 : 37	31ppt	5%		PM 9 : 37
PM 9 : 37	27.4ppt	10%		PM 10 : 37
PM 10 : 37	26.1ppt	5%		PM 11 : 37
PM 11 : 37	24.3ppt	5%		AM12 : 37
AM12 : 37	23.2ppt	5%	1마리(0.8cm)	AM 1 : 37
AM 1 : 37	22.1ppt	5%		AM 2 : 37
AM 2 : 37	21.2ppt	5%	1마리(1cm)	AM 3 : 37
AM 3 : 37	19.7ppt	5%	2마리(0.8cm, 0.9cm)	AM 4 : 37
AM 4 : 37	18.2ppt	5%	6마리(0.9cm(2), 0.8cm(2), 0.7cm(2))	AM 5 : 37



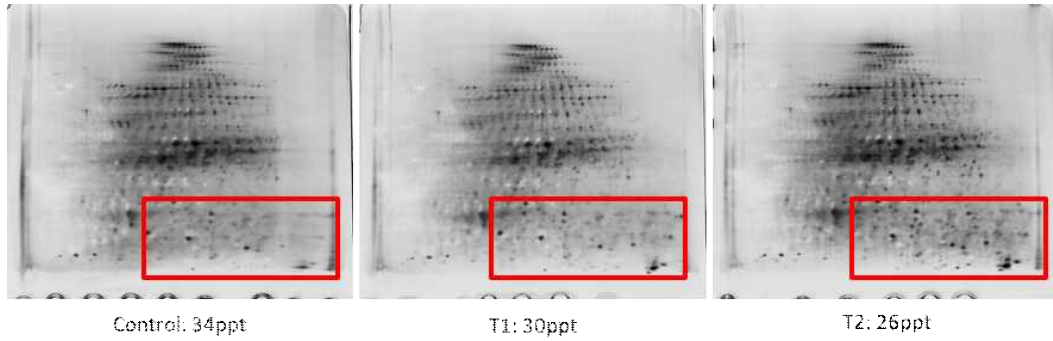


그림 1. 염도감소 스트레스에 의해 유도되는 *Gondogeneia antarctica*의 단백질 지도 변화



제 3-13절 극지 환경에 의해 조절되는 생리변화 조절 경로 연구

가. 남극 저온성 박테리아 *Psychrobacter* sp. PAMC 21119 유래 cold shock protein의 RNA chaperone의 기능

1) 서론

남극 저온성 박테리아 *Psychrobacter* sp. PAMC 21119는 남극 동토로부터 분리되었고 이 균주는 영하의 온도(-5도)에서도 성장하였다(Kim, Shin et al. 2012). 저온 스트레스 실험결과, 온도 조건별로 다른 발현율을 나타낸 두 종류의 Csp isoform (ppCspC, ppCspA) 단백질을 확인하였다.

저온 스트레스가 박테리아에 미치는 영향은 막 유동성 감소, mRNA 번역 및 전사의 효율 저하에 따른 핵산 구조 안정화, 리보솜 기능 방해, 비효율적인 단백질 접힘 등의 결과를 가져온다. 이러한 저온 스트레스의 부정적 영향에 대처하기 위해 Csp 단백질이 유도된다(Phadtare and Severinov 2010). Csp 단백질은 불포화지방산 형성을 증가시키고 세포막유동성의 재확립으로 세포막으로의 결합을 유도하며, mRNA 번역과 전사, 리보솜 기능의 회복, 세포 단백질의 적절한 접힘 등을 유도한다(Phadtare and Severinov 2010, Jones 2012). cold shock protein의 RNA chaperone의 기능을 검증하여 전사(transcription)를 용이하게 함으로써 다음 단계인 기능성 단백질 발현에 직접적으로 관여함을 규명하였다.

2) 연구수행내용

- 재조합 단백질 제작 및 정제

pET-28a(+) vector에 ppCspC/ppCspA/CspA를 클로닝한다. BL21(DE3)에 형질전환 후 kanamycin이 포함된 10ml LB배지에 접종하고, 37도에서 배양한다. kanamycin이 포함된 1L LB배지에 10ml을 접종하고 OD 0.6이 되는 시점에서 25도로 배양기 온도를 낮춘 후 0.5M IPTG를 첨가하여 24시간 배양한다. 균체를 회수하고 초음파로 파쇄한다. 원심분리하여 상층액(가용성 단백질 분획)과 침전(불용성 단백질 분획)을 회수하고, 각각을 SDS-PAGE로 분석하여 단백질의 발현을 조사한다. 원심 분리한 상층액 중 가용성 단백질을 10ml의 Ni-bead에 흡착시키고, 300mM NaCl, 10mM imidazol을 포함한 50mM 인산완충액(pH8.0)으로 세정한 후, 300mM imidazol의 농도에 따라 재조합 단백질을 정제하였다.

cold shock test

저온 스트레스에 민감하게 반응하는 BX04 세포(CspA, B, E, G 결핍)에 cold

shock protein을 형질 전환한다. 이때 사용하는 세포는 액체 배양 후 세포수를 계산하여 동량의 균질한 세포액을 kanamycin이 포함된 LB배지에 접종한다. BX04 세포와 형질 전환된 세포를 37/15/10도 인큐베이터에서 배양한 후 세포 성장을 관찰한다(Xia, Ke et al. 2001).

freezing and thawing assay

ppCspC/ppCspA/CspA를 형질전환하여 액체 배양하고 과발현 시킨 균주를 1ml 씩 튜브에 분주한다. -20도에서 2시간동안 냉동 후 1시간동안 얼음위에서 녹이는 과정을 한 사이클로 하여 2번 3번 반복한다. 각 사이클의 균주는 LB 플레이트에 분주한 후 37도 인큐베이터에서 24시간 배양한다. 배양한 균주는 CFU (colony forming units)로 카운팅하며 데이터는 총 5번의 실험으로 평균을 계산하였다. 실험의 대조군으로는 공vector를 형질 전환한 균주를 사용하였다 (Jung, Yi et al. 2010).

transcription antitermination assay

자체 transcription anti-termination 기능이 삭제된 RL211 세포주를 사용하여 외부 유전자 도입에 의해 유도되는 transcription anti-termination activity 활성을 측정하였다. pINIII 플라스미드에 클로닝된 ppCspC, ppCapA, CspA를 E.coli RL211에 형질 전환한다. 형질전환된 균주는 ampicillin이 포함된 LB 액체배지에서 밤새 배양한 후 새로운 배지에 1/100로 희석한다. OD600에서 1.0이 될 때까지 배양 한 후 ampicillin 배지에 chloramphenicol (30ug/ml)이 포함되거나 포함되지 않은 배지에 5ul씩 떨어뜨린다. 인큐베이터에서 2~3일간 배양한다. CspA는 positive control, pINIII vector는 negative control로 사용한다 (Landick, R. 1990).

DNA binding assay

정제된 재조합 단백질을 농도별로 완충액(10mM Tris-HCl, pH7.5)과 dsDNA (M13mp18RFIDNA, NEB), ssDNA (M13mp18, NEB), lucmRNA (FLuc-mRNA, TriLink Biotechnologies Co.) 100ng을 혼합한 후 얼음 위에서 30분간 반응시켰다. 반응 후 단백질과 DNA, RNA의 결합을 확인하고자 1% agarose gel로 전기영동 하였다. GelRed (Biotium, USA)로 염색하여 확인하였다(Yang, Wang et al. 2012).

FITC assay

RNA chaperone의 기능 중 하나인 stem-loop structure melting activity 측정을 위하여 FITC-BHQ (molecular beacon)을 이용하고 Csp 결합 시 방출되는 형광을 측정하였다. 두 개의 oligonucleotides는 각각 FITC와 BHQ1 라벨을 붙였

다. molecular beacon substrates는 FITC:BHQ1=1:20으로 섞은 후 95도에서 2분간 변성하고 20분간 얼음위에 보관한다. annealed DNA는 분석용 표지 기질로 사용한다. 10mM Tris-HCl(pH7.5) 완충액에 1pmol의 분자 기질과 60ug 정제된 단백질을 넣고 얼음위에서 반응시킨다. 반응 후 555/575nm에서 형광을 측정한다(Melencion, Chi et al. 2017).



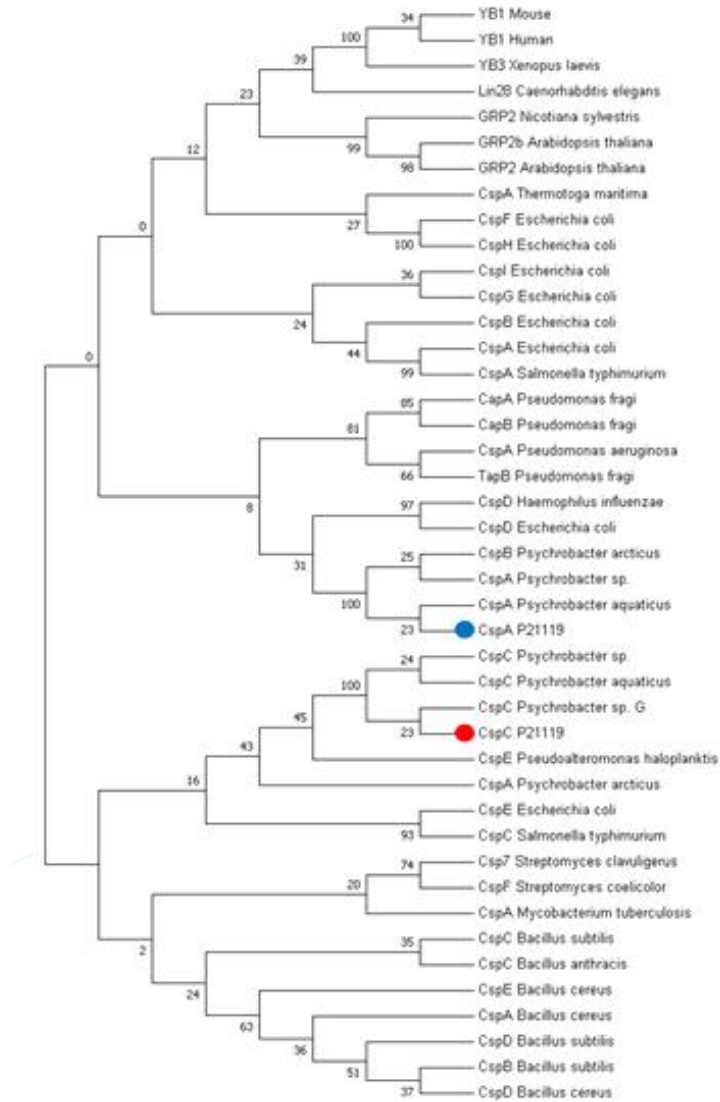


그림 1. 남극 저온박테리아 *Psychrobacter sp.* PAMC 21119의 두가지 cold shock protein의 계통도

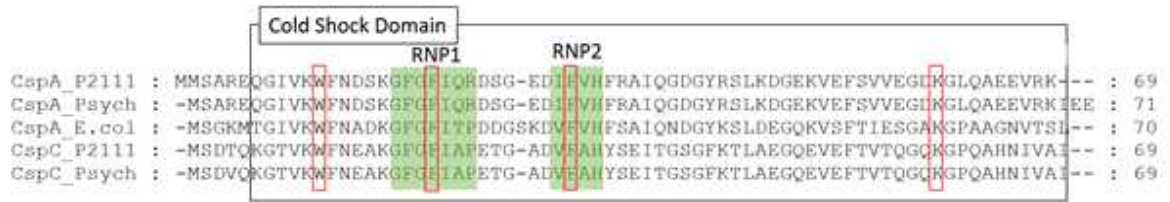


그림 2. Psychrobacter PAMC 21119 유래 Csp의 아미노산 서열 분석. RNP1, RNP2 (RNA binding domain): red box, DNA binding site



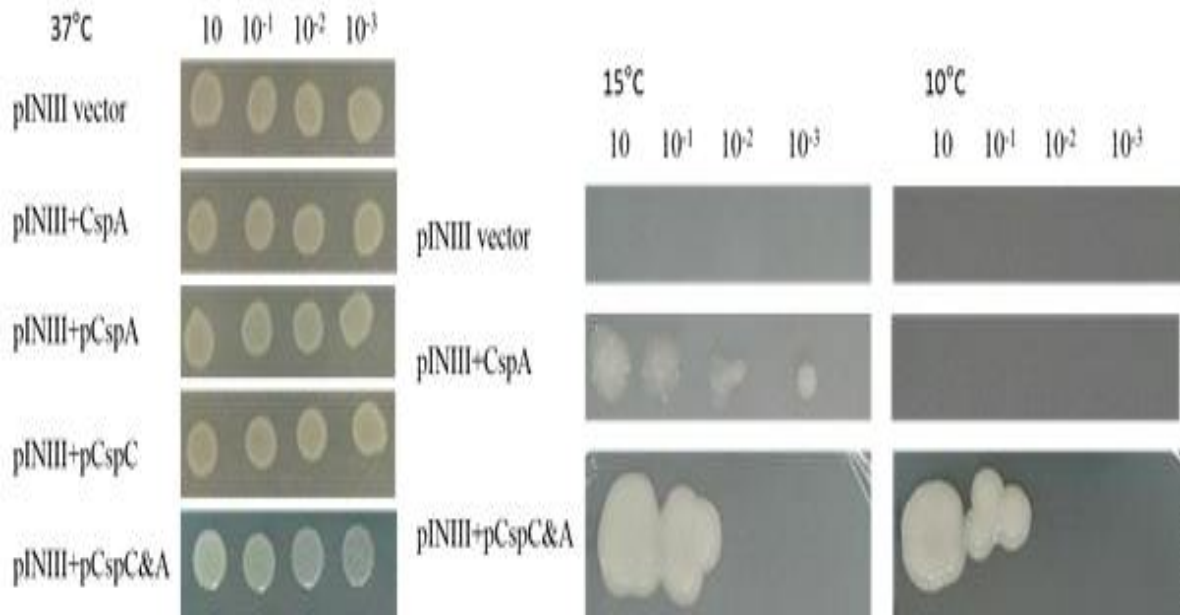


그림 3. 저온조건에서의 이형 *csp* 유전자 발현 균주의 세포 성장



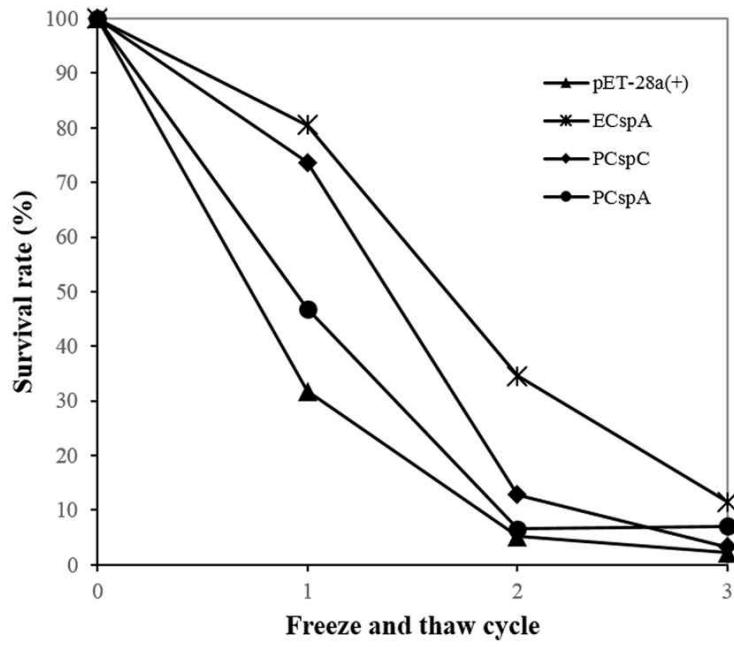


그림 4. 형질 전환된 박테리아의 생존율 곡선



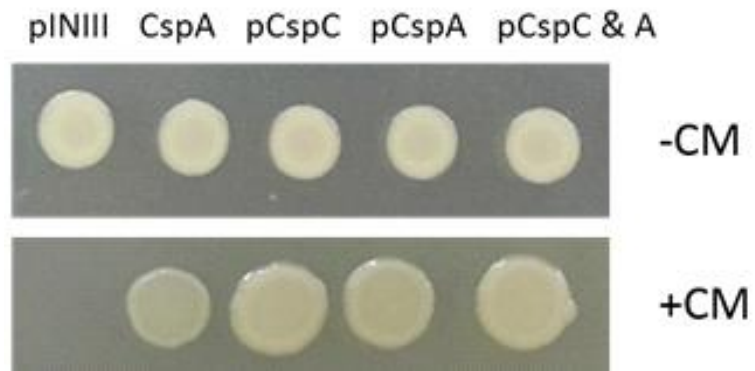


그림 5. transcription anti-termination 활성



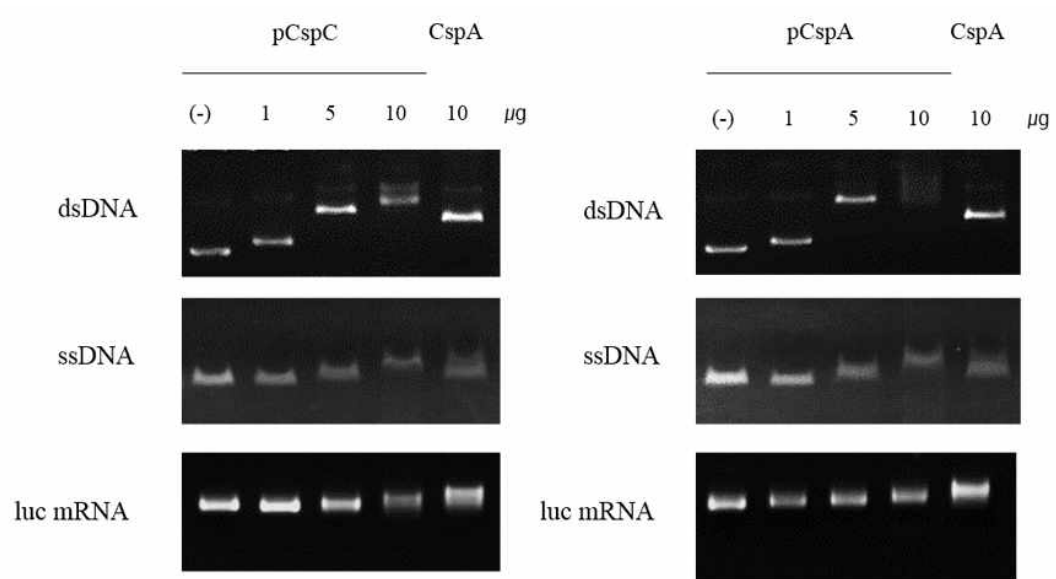


그림 6. 남극 저온 박테리아 유래 Csp의 DNA와 RNA binding activity 확인



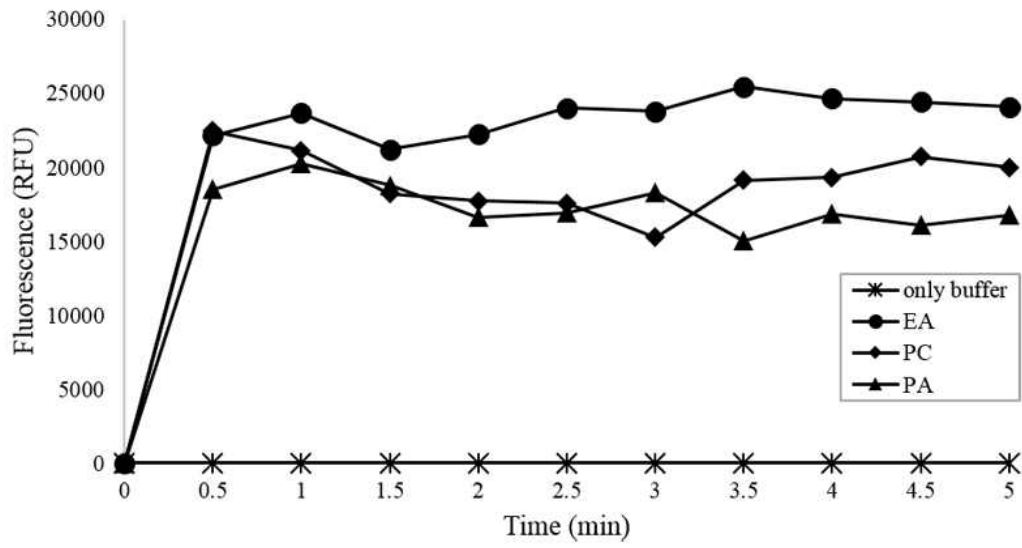


그림 7. FITC-BHQ를 이용한 남극 유래 Csp protein의 nucleic acid melting activity 확인



참고문헌

- 1) Kim, S. J., S. C. Shin, S. G. Hong, Y. M. Lee, I.-G. Choi and H. Park (2012). Genome sequence of a novel member of the genus *Psychrobacter* isolated from Antarctic soil, *Am Soc Microbiol*.
- 2) Phadtare, S. and K. J. R. b. Severinov (2010). "RNA remodeling and gene regulation by cold shock proteins." *7(6): 788-795*.
- 3) Jones, T. J. S. r. o. f. m. (2012). "Response of *Escherichia coli* to environmental stress." *293-330*.
- 4) Xia, B., H. Ke and M. Inouye (2001). "Acquirement of cold sensitivity by quadruple deletion of the *cspA* family and its suppression by PNPase S1 domain in *Escherichia coli*." *Molecularmicrobiology40(1): 179-188*.
- 5) Landick, R., J. Stewart and D. Lee (1990). "Amino acid changes in conserved regions of the beta-subunit of *Escherichia coli* RNA polymerase alter transcription pausing and termination." *Genes&Development4(9): 1623-1636*.
- 6) Jung, Y. H., J.-Y. Yi, H. J. Jung, Y. K. Lee, H. K. Lee, M. C. Naicker, J.-h. Uh, I. S. Jo, E. J. Jung and H. J. T. p. j. Im (2010). "Overexpression of cold shock protein A of *Psychromonas arctica* KOPRI 22215 confers cold-resistance." *29(2): 136-142*.
- 7) Yang, C., L. Wang, V. S. Siva, X. Shi, Q. Jiang, J. Wang, H. Zhang and L. J. P. o. Song (2012). "A novel cold-regulated cold shock domain containing protein from scallop *Chlamys farreri* with nucleic acid-binding activity." *7(2)*.
- 8) Melencion, S. M. B., Y. H. Chi, T. T. Pham, S. K. Paeng, S. D. Wi, C. Lee, S. W. Ryu, S. S. Koo and S. Y. J. I. j. o. m. s. Lee (2017). "RNA chaperone function of a universal stress protein in *Arabidopsis* confers enhanced cold stress tolerance in plants." *18(12): 2546*.

나. 남극 저온성 박테리아 *Psychrobacter* sp. PAMC 21119 유래 ppRidA의 ribonuclease의 기능

1) 서론

남극 저온성 박테리아 *Psychrobacter* sp. PAMC 21119를 저온 처리하였을 때 발현된 ppRidA의 ribonuclease 기능을 확인하기 위하여 실험을 진행하였다. ribonuclease는 RNA metabolism에 필수적이며 저온 스트레스의해 전사 과정이 제대로 이루어지지 않을 시 부정확한 정보를 지닌 mRNA의 분해에 필요한 효소이다(Morishita, Kawagoshi et al. 1999).

2) 연구수행내용

재조합 단백질 제작 및 단백질 정제

pET-28a(+) vector에 ppRidA를 클로닝한다. BL21(DE3)에 형질전환 후 kanamycin이 포함된 10ml LB배지에 접종하고, 37도에서 배양한다. kanamycin이 포함된 1L LB배지에 10ml을 접종하고 OD 0.6이 되는 시점에서 25도로 배양기 온도를 낮춘 후 0.5M IPTG를 첨가하여 24시간 배양한다. 균체를 회수하고 초음파로 파쇄한다. 원심분리하여 상층액(가용성 단백질 분획)과 침전(불용성 단백질 분획)을 회수하고, 각각을 SDS-PAGE로 분석하여 단백질의 발현을 조사한다. 원심 분리한 상층액 중 가용성 단백질을 10ml의 Ni-bead에 흡착시키고, 300mM NaCl, 10mM imidazol을 포함한 50mM 인산완충액(pH8.0)으로 세정한 후, 300mM imidazol의 농도에 따라 재조합 단백질을 정제하였다.

- string DB

ribonuclease 기능을 확인하기 위해 STRING 알고리즘을 통해 단백질-단백질 상호 작용을 예측하였다(Szklarczyk, Franceschini et al. 2015).

- mRNA 분해 실험

ppRidA의 ribonuclease 기능을 분석하기 위해 정제된 재조합 단백질을 농도별로 완충액(10mM Tris-HCl, pH7.5)과 lucmRNA (FLuc-mRNA, TriLink Biotechnologies Co.) 100ng을 혼합한 후 저온(on ice)과 37°C에서 mRNA 분해 실험 수행하였다. 반응 후 단백질의 mRNA 분해능을 확인하고자 1% agarose gel로 전기영동 하였다. GelRed (Biotium, USA)로 염색하여 확인하였다(Villalobos-Osnaya, Garza-Ramos et al. 2018).

두 가지 cold shock protein의 일차 아미노산 서열을 분석하여 cold shock domain 내에 2군데의 RNA binding domain과 4군데의 DNA binding site가 있음을 확인하였다.

STRING 알고리즘 결과는 Psychrobacter sp. 1501 균주의 endoribonuclease와 exoribonuclease로 예측된 단백질들과 상화 작용한다는 것을 보여주었다. ppRidA의 ribonuclease 기능은 농도에 따라 mRNA의 분해가 이루어지는 것을 확인하였고 저온 (on ice)에서도 psyRidA가 ribonuclease 활성을 가지는 것으로 확인하였다.



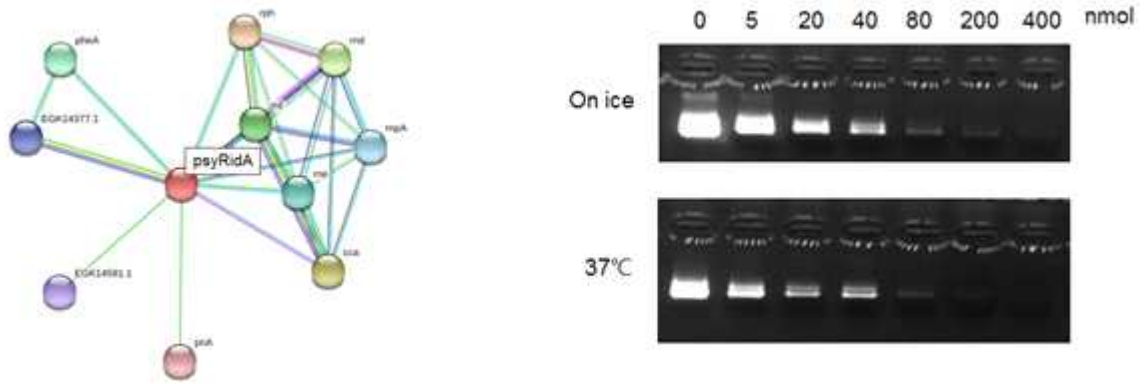


그림 1. ppRidA의 단백질-단백질 상호작용 예측 및 mRNA 분해 활성 측정 확인



참고문헌

- 1) Morishita, R., A. Kawagoshi, T. Sawasaki, K. Madin, T. Ogasawara, T. Oka and Y. J. J. o. B. C. Endo (1999). "Ribonuclease activity of rat liver perchloric acid-soluble protein, a potent inhibitor of protein synthesis." 274(29): 20688-20692.
- 2) Szklarczyk, D., A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos and K. P. J. N. a. r. Tsafou (2015). "STRING v10: protein - protein interaction networks, integrated over the tree of life." 43(D1): D447-D452.
- 3) Villalobos-Osnaya, A., G. Garza-Ramos, I. N. Serratos, C. Millán-Pacheco, A. González-Robles, R. Arroyo, L. I. Quintas-Granados and M. E. J. P. r. Alvarez-Sanchez (2018). "Identification of a perchloric acid-soluble protein (PSP)-like ribonuclease from *Trichomonas vaginalis*." 117(11): 3639-3652.



제 3-14절 극지 미생물에서 생산되는 유용 효소 단백질의 구조 분석 및 특성 연구

○ 난분해성 물질 분해 관련 효소의 효율적 단백질 생산 시스템 구축

- 저온성 미생물의 단백질은 저온에서의 고활성을 가지기 위하여 유연성이 많은 분자 구조로 진화되어 왔으며, 이러한 생물리학적 특성은 재조합 단백질 생산에 있어서 가용성 혹은 활성을 가지는 단백질 생산에 어려움이 있음.
- 유연성이 많은 구조를 가진 단백질의 효율적 생산을 위하여 저온성 샤페론을 공발현하여 생산 효율성을 향상 시킴.

○ 효율적인 단백질 결정화와 회절 정보를 활용한 구조 분석

- 다양한 융합 단백질을 구축하여 단백질 결정을 위한 스크리닝에 활용하고 구조 분석에 활용 가능한 시료를 다양하게 확보.
- 효소의 활성 기작을 규명하기 위하여 활성부위 혹은 기질 결합 부위의 돌연변이체를 제작하여, 기질-단백질 복합체의 구조를 분석하고, 효소와 다양한 기질의 반응 정보를 확보함.

○ 난분해성 물질 나프탈렌의 생분해 관련 효소 NahB 의 구조와 활성기작 연구

- 두 개 이상의 벤젠 링으로 구성된 polycyclic aromatic hydrocarbons (PAHs)는 전세계적으로 널리 알려진 유기 환경 오염물질로서, 석유 및 석탄의 불완전 연소로 배출되어지는 발암 독성 물질이며, 100여 종의 난분해성 PAH가 보고되어짐.
- PAH의 기본 단위체로 알려진 나프탈렌 생분해에 관련된 효소군의 유전정보가 극지유래 *Pseudomonas* sp. MC1 균주에서 발현 및 활성에 관한 정보가 확보되었으며, 관련 pathway의 효소군 중에서 활성기작에 관하여 잘 알려지지 않은 cis-dihydrodiol naphthalene dehydrogenase (NahB) 단백질 효소의 구조를 통하여 활성기작에 관한 기능성을 검증함.
- PAH 분해 관련 효소의 단백질 구조를 분자 수준에서 해석하였으며, 기질 및 리간드의 복합체의 분자 구조를 해석하였음. 단백질의 분자 구조는 *Pandora* sp. *pnomenus*의 BphB 단백질과 같은 기본 골격 구조를 가지고 있었으며, 기질인 PAH의 결합 부위에서 매우 유연한 loop구조를 가지고 있는 것을 확인하였음. 생분해 활성기작에 중요한 리간드 결합과 이와 관련된 구조적 변화 과정을 분자 구조 분석한 결과, 초기 기질 결합 과정에서 다양한 구조의 기질과 반응할 수 있다는 것을 검증함

- 나프탈렌 분해 효소 NahX의 생화학적 활성 특성 및 단백질 구조 분석
 - 지금까지 정확한 기능이 밝혀지지 않은 NahX의 활성 기작을 규명하기 위하여 단백질 구조를 분석하였으며, 특이적 활성을 추정함
 - 단백질 구조의 유사도 조사를 통하여 기능을 예측 분석하였으며, 유사 단백질의 효소 활성도를 조사함
 - 철이온 결합 활성능과 활성간의 연관성을 확인하고, 관련 아미노산 서열의 단백질 구조에서의 분석으로 확인하였음.
- 신규 저온성 esterase 의 특성 및 생산 전략에 관한 연구
 - 생물 및 화학 산업에 있어서 esterase는 에스테르 물질을 카르복시산 또는 알코올로 분해할 수 있는 매우 유용하게 생촉매제로 알려져 있음. 특히 저온에서의 생화학 반응이 필요한 분야에서는 저온성 esterase의 요구가 높은 것으로 알려져 있음.
 - 극지 유래 미생물에서 esterase 효소 PsEst3를 대장균에서 발현하였으나, 발현 생산량이 매우 낮게 생산됨에 따라, 본 연구팀이 개발한 저온성 샤페론의 공발현 시스템을 도입하여 생산량을 증가시킴.
 - 가용성 효소 PsEst3의 발현량을 증가시킴에 따라 성공적인 단백질 분리 및 정제가 가능해짐. 본 효소의 온도에 따른 활성 특성을 비교하였으며, 반응 온도의 특성에 의해 전형적인 저온성 효소로 확인이 되어짐.
- 극지 환경 특이적 단백질 발현, 정제, 삼차구조해석 및 기능 분석
- 극지식물 유래 항산화 활성 효소 monodehydroascorbate reductase, MDHAR, 의 구조와 효소 활성기작 연구
 - ascorbic acid, AsA 는 식물에서 매우 풍부한 항산화 물질로 알려져 있으며, AsA의 항산화 활성을 유지하기 위한 재생을 위해서는 환원효소 MDHAR의 활성 조절이 요구되며, 저온, 건조, 강한 자외선과 같은 극한 환경에서의 산화 스트레스에 의한 활성 산소종의 증가와 MDHAR의 기능성에 대한 연관성 규명함.
 - 극지식물 Antarctic hair grass (*Deschampsia antarctica*) 유래의 MDHAR의 생화학적 특성 분석
 - 효소 단백질 DaMDHAR의 활성 검증 및 단백질 구조 분석을 통하여 중온성 MDHAR과의 구조적 유사도를 분석함.

- DaMDHAR의 활성 부위에 대한 단백질 구조를 분석하여 효소 활성 기작에 대하여 규명하기 위하여, 조효소 NADH와 결합된 복합체에 대한 분자 구조를 해석하였음. 단백질의 기본 구조는 중온성 효소와 같은 기본 골격 구조를 가지고 있었으나, 단백질 표면에는 non-conserved 구조를 보여줌.
- DaMDHAR 발현 효모에서의 스트레스 반응을 조사를 위하여, freezing and thawing (FT) 실험을 통하여 세포 생존율을 조사함.
- 중온성 MDHAR과 비교하여 DaMDHAR 발현 효모에서 FT와 같은 환경 스트레스에 대하여 강한 내성을 보여 줌.



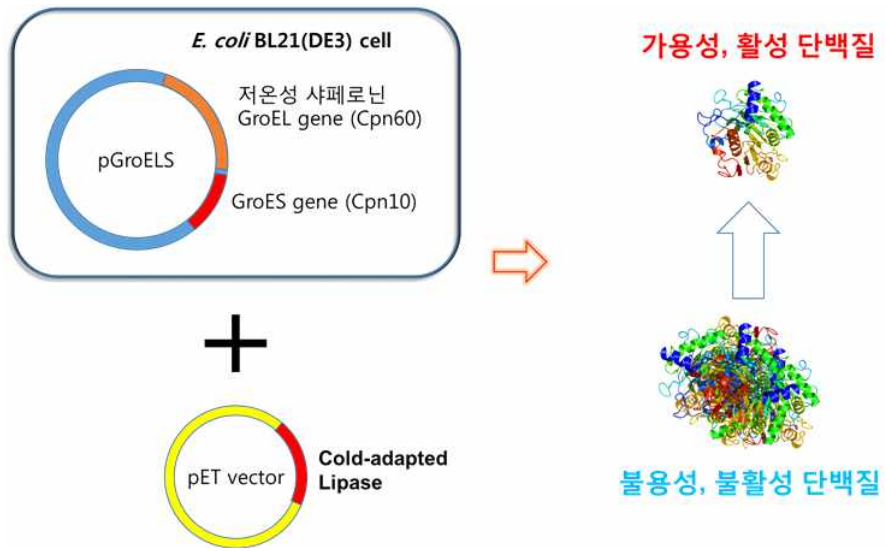


그림 1. 대장균에서 외래 단백질의 효율적 생산 시스템 예시



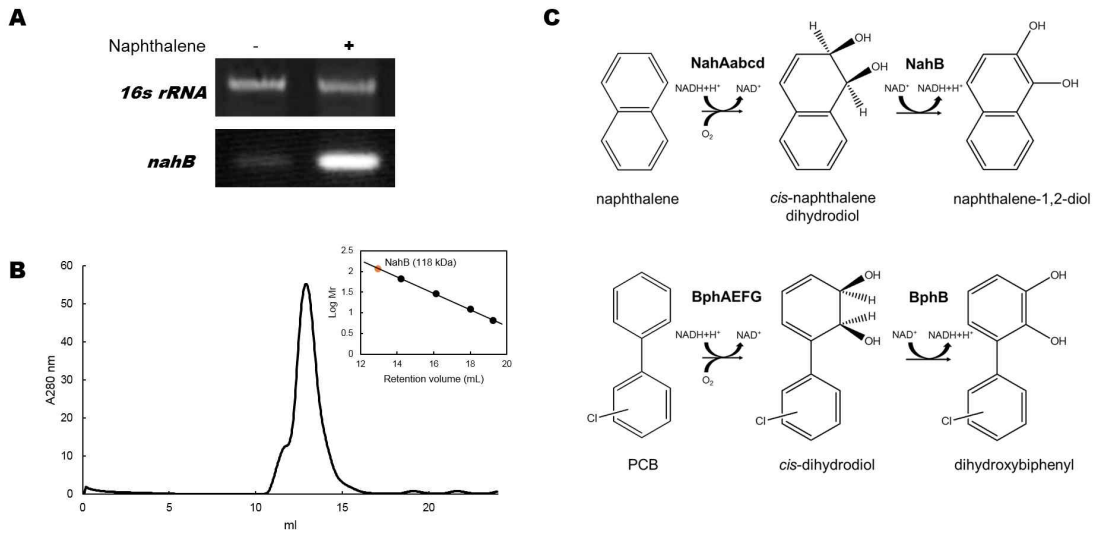


그림 2. nahB 단백질의 발현 분석. A. 기질인 나프탈렌 존재하에서 nahB의 RNA 발현 분석, B. 정제된 NahB 단백질 효소의 size-exclusion chromatography 분석, C. naphthalene과 polychlorinated biphenyl (PCB)의 생분해 도식



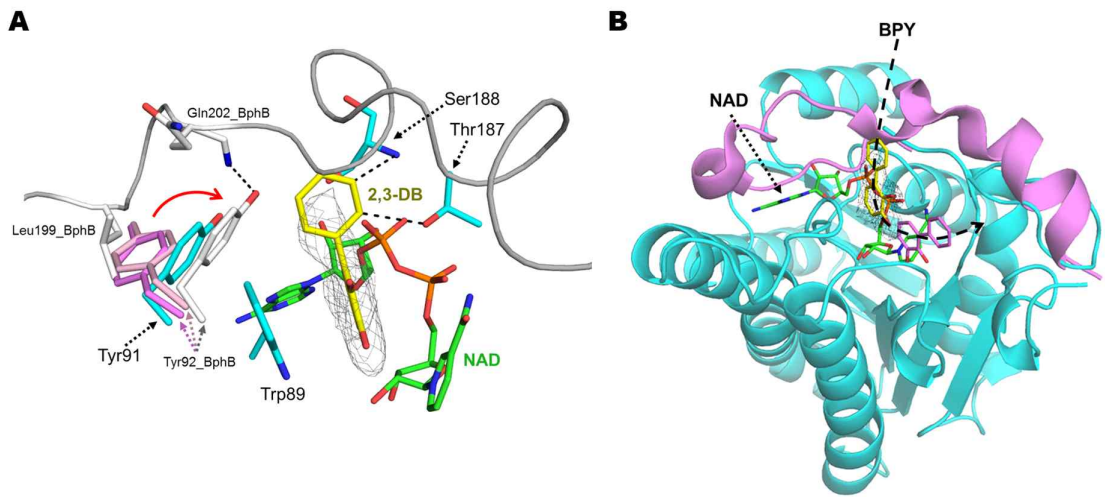


그림 3. 효소 단백질 NahB의 기질 결합 부위를 분석한 구조. A. 조효소 NAD와 결합된 효소 NahB의 구조, B. 리간드 2,3-dihydroxybiphenyl (2,3-DB)와 결합된 효소-리간드 복합체의 구조



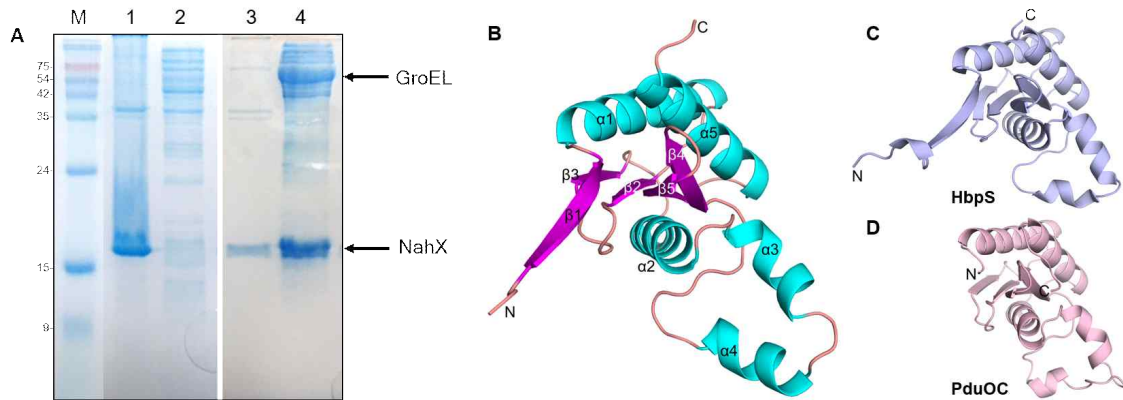


그림 4. 저온성 GroELS에 의한 재조합 NahX의 발현량 증가와 단백질 구조



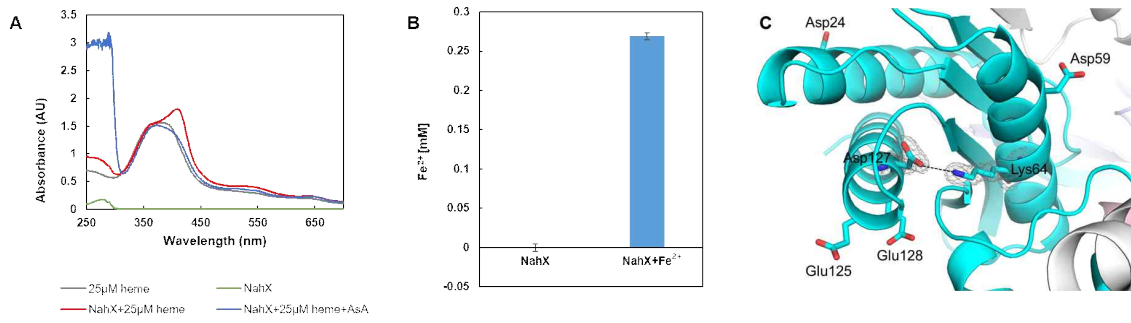


그림 5. NahX에 의한 heme의 결합과 분해에 대한 측정 비교



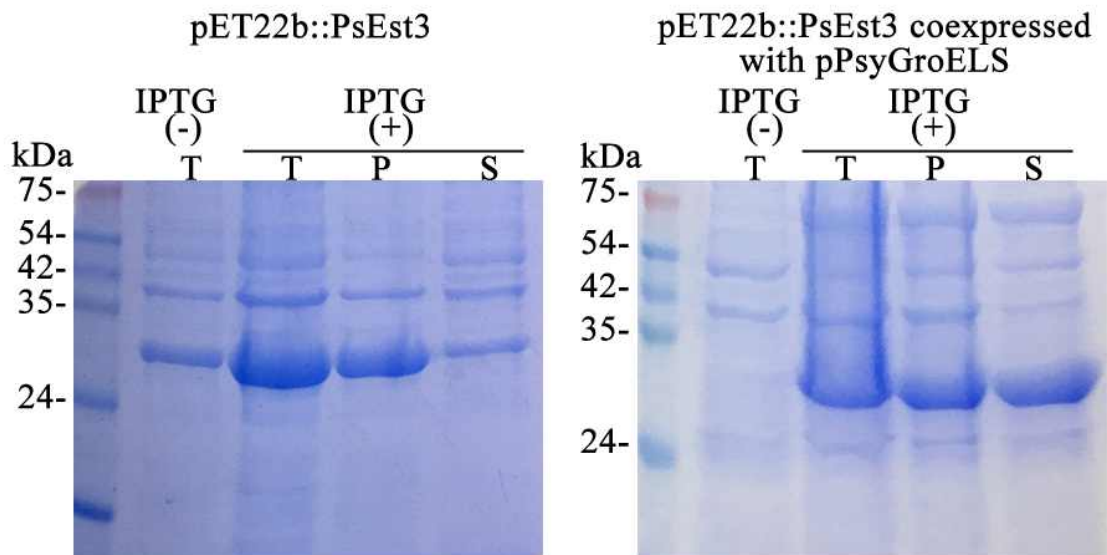


그림 6. 저온성 샤페론 PsyGroELS 공발현을 이용한 효소 PsEst3의 가용성 향상



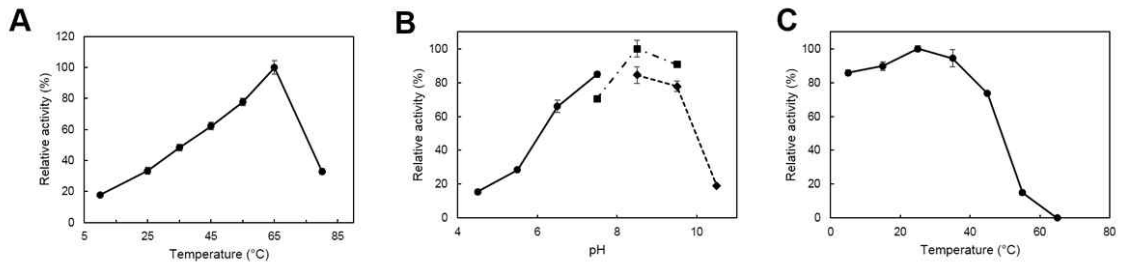


그림 7. DaMDHAR의 생화학적 효소 반응 특성 조사



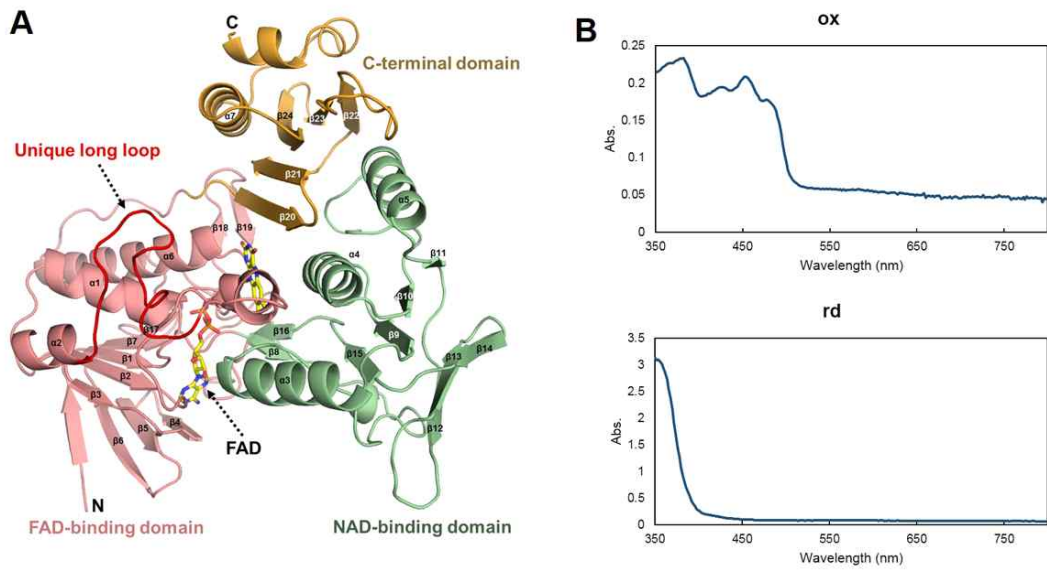


그림 8. DaMDHAR의 활성상태 검정 및 단백질 구조 분석



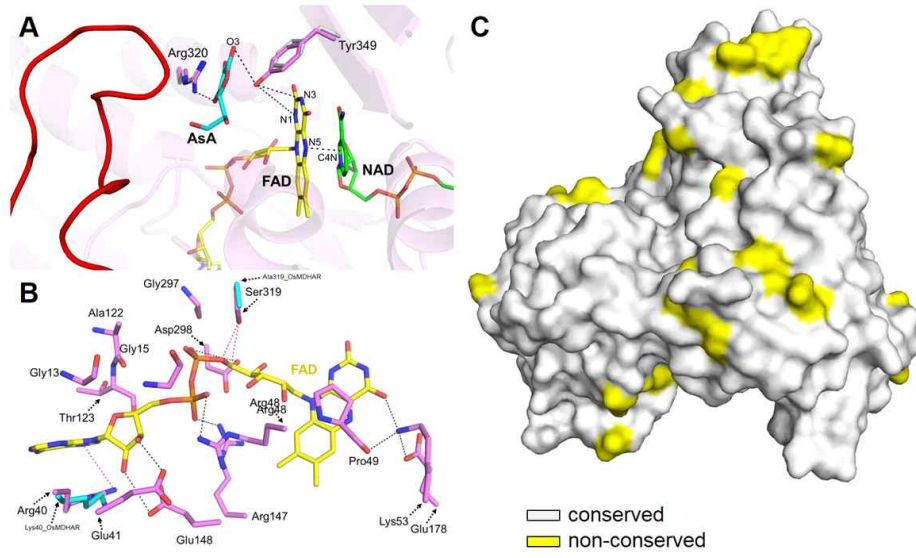


그림 9. DaMDHAR의 활성부위 구조 및 표면 구조의 보존성



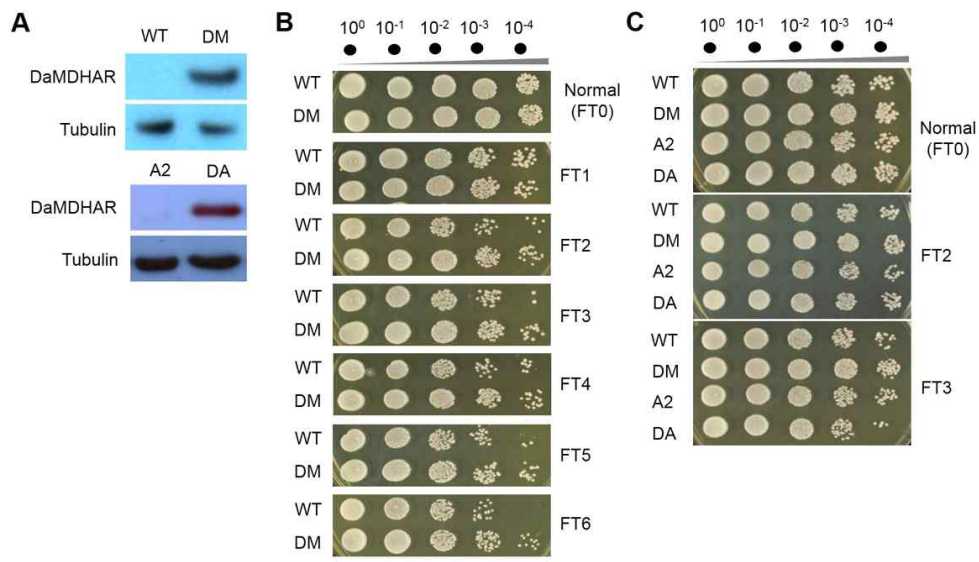


그림 10. DaMDHAR 발현 효모의 환경 스트레스 내성 활성



제 4장 연구개발결과의 활용계획

1. 향후 연구방향

가. 극지 생물 고품질 유전체 정보 획득 및 극지 생명정보 통합 관리 체계 구축

- 선도 유전체 정보 확보 및 지도 구축
- 비교유전체 연구기반 개체군 유전체 연구 파이프라인 구축
- 포스트 유전체 활용 극지 집단 군집/ 환경 특이적 유용 자원 발굴

나. 통합 유전체 정보망 및 활용기반 구축

- 핵심 대상 생물군 선정
- 극지생물 유전체 데이터망을 활용한 기후 및 환경요소 영향 평가
- 환경 인자별 유전체 정보 분석을 통한 환경 적응 관련 기능유전자 발굴
- 극지생물 유래 형질전환 시스템 확보
- 극지생물 배양 및 유지기술 확립
- 극지생물 유전체 활용 모델시스템 구축

다. 극지 통합생명정보 활용

전략 수립 및 산업적 응용성 검증

- 극지 생물자원 유전체 정보의 체계적 활용 전략 수립을 통한 고유/희귀 생물 자원의 산업적 가치와 활용성 증대
- 극지 유전체 전략 생물군 정립 및 초정밀 유전체 활용
- 극지 생물 특이형질 활용성 검증

라. 극지 유전체기반 전략 모델 실용화 달성

- 극지 어류 유래 결빙방지단백질 활용 환경내성 수산자원 개발
- 극지생물 유래의 저온 특이적 소재 발굴
- 극지 유전자원을 활용한 작물 개량

2. 성과 활용계획

가. 과학분야: 남극 생물 특이적 유전자원 확보를 위한 교두보 마련

- 남극어류 유전체 해독은 극지연에서만 성공적 수행 및 세계적 선도
- 남극어류의 대표종인 icefish의 유전체를 해독하고 reference genome을 제공함으로써 남극어류 연구의 기반 제공
- 연구자들의 관심도가 높은 생물종을 우선 연구대상으로 선정하고 데이터를 선공개함으로써 다학제 연구 활성화와 연구소 전체의 연구 수월성 제고
- 극지생명과학연구의 근간인 유전체 정보를 확보함으로써 극지생물의 적응과 진화 연구 토대 마련
- 극지 특이적인 생명현상 규명을 통한 지구 생명의 생성과 적응, 진화의 과학적 근거 제시
- 극지 지역을 중심으로 급속도로 진행되는 지구 온난화의 원인을 극지생물을 중심으로 기후 연구와 생물학 분야의 융합연구 연계점 제시
- 극지 환경 적응관련 핵심 유전자를 활용한 형질전환 생물 개발 및 기술력 선점
- 미지의 극지생명정보를 적극 확보하고 BT, IT, ET 등 융합학문을 촉진하여 극지생명공학 분야의 국가 경쟁력 제고

나. 사회/경제적 분야: 극지 분야 세계 바이오산업 시장 선도가능

- 극지 고유 생명정보 확보를 통한 극지연구의 저변 확대와 극지생명 연구 분야 세계 1위의 국제적 선도 기관으로 도약
- 극지 생명현상의 이해와 해석 정보를 바탕으로 새로운 생명자원을 개발하기 위한 4차 산업혁명의 핵심 생물자원 정보 제공
- 세계 바이오산업 시장 규모는 약 330조 원 (2013년 기준)로 연평균 10.8%의 고성장세를 보이고 있으며, 지속적으로 증가될 전망. 미래 글로벌 시장의 선점을 위해서 극지생물을 이용한 다양한 생명공학기술의 연구개발이 필요하며, 미래경제력 창조의 촉매 역할을 담당
- 극지생물의 환경적응, 진화 기작과 관련한 유전체를 확보하여 국가생명공학산업의 새로운 패러다임으로 다양한 신규 바이오소재의 활용가치를 제시함
- 극지 생물 신소재 개발을 위한 기초연구로 21세기 세계 생명공학 분야에서 선

도적 역할 담당

- 극지 생물 특이적 유전자 분석 기술 전수를 통한 국내 연구 및 산업계 이익 창출
- 극지 생물의 극한 환경 적응관련 핵심 유전자를 활용한 의료, 미용 소재 개발을 통한 경제성 창출
- 극한 환경 적응관련 핵심 유전자 형질전환 활용 시 국가적 이익 창출과 더불어 반복되는 저수온, 고수온, 적조 등의 급격한 환경변화로 인한 수산생물의 피해 저감을 통한 국내 수산양식산업 전반의 사회 문제 해결 가능

다. 정책 분야: 정부시책에 부합된 연구 수행

- 나고야의정서 내 유전체 생물 주권에 있어 극지 생물 분야의 글로벌 선도 기관으로 세계적 우위 선점 가능
- 극지 수산생물을 이용한 고부가 수산양식 기술 개발을 통한 정부시책에 부합된 연구 수행 가능: 제1차 해양수산과학기술 육성 기본계획(과제 2-1 전략산업 육성을 위한 상용화 기술 확보)
- 극지 해양생물 배양기술 확립을 통한 극지 해양자원 보존 등 국제사회 공동 현안 해결가능: 제1차 해양수산과학기술 육성 기본계획(과제 2-2 국제 사회에 기여하는 해양수산과학기술 육성)

제 5장 참고문헌

1. Zhang G, Li Y, Fang L, Pfeifer BA. Tailoring pathway modularity in the biosynthesis of erythromycin analogs heterologously engineered in *E. coli*. *Sci Adv*. 2015 May 29;1(4):e1500077. doi: 10.1126/sciadv.1500077. eCollection 2015 May. PubMed PMID: 26601183; PubMed Central PMCID: PMC4640655.
2. Vicente M, Hodgson J, Massidda O, Tonjum T, Henriques-Normark B, Ron EZ. The fallacies of hope: will we discover new antibiotics to combat pathogenic bacteria in time? *FEMS Microbiol Rev*. 2006 Nov;30(6):841-52. Review. PubMed PMID:17064283.
3. Kummerer K. Antibiotics in the aquatic environment--a review--part I. *Chemosphere*. 2009 Apr;75(4):417-34. doi: 10.1016/j.chemosphere.2008.11.086. Epub 2009 Jan 30. Review. PubMed PMID: 19185900.
4. Kummerer K. Antibiotics in the aquatic environment--a review--part II. *Chemosphere*. 2009 Apr;75(4):435-41. doi: 10.1016/j.chemosphere.2008.12.006. Review. PubMed PMID: 19178931.
5. Singh RK, Tiwari SP, Rai AK, Mohapatra TM. Cyanobacteria: an emerging source for drug discovery. *J Antibiot (Tokyo)*. 2011 Jun;64(6):401-12. doi:10.1038/ja.2011.21. Epub 2011 Apr 6. Review. PubMed PMID: 21468079
6. Clardy J, Fischbach MA, Walsh CT. New antibiotics from bacterial natural products. *Nat Biotechnol*. 2006 Dec;24(12):1541-50. Review. PubMed PMID: 17160060.
7. Hong W, Zeng J, Xie J. Antibiotic drugs targeting bacterial RNAs. *Acta Pharm Sin B*. 2014 Aug;4(4):258-65. doi: 10.1016/j.apsb.2014.06.012. Epub 2014 Jul 31. Review. PubMed PMID: 26579393; PubMed Central PMCID: PMC4629089.