

멸종위기 야생동물 II급 염주알다슬기 (*Koreoleptoxis nodifila*) 의 Aquaporin 유전자 발굴 및 생물정보학적 분석

홍찬의^{1,2}, 정준양^{2,3}, 신현준^{1,2}, 이혁^{1,2}, 이경원^{1,2}, 장이선^{1,2}, 상민규^{2,3}, 박지은^{2,3}, 송대권^{2,3}, 황희주⁴, 강세원⁵, 박소영⁶, 김기원⁷, 조성신¹, 박정한⁸, 박홍석⁹, 이준상¹⁰, 조용훈^{1,2}, 이응석^{1,2,3}

¹순천향대학교 자연과학대학 생명과학과, ²순천향대학교 한국자생동물자원활용 융복합연구소, ³순천향대학교 생명자원 바이오빅데이터 분석 및 활용 연구지원센터, ⁴순천향대학교 순천향의생명연구원, ⁵한국생명공학연구원 생물자원센터, ⁶국립낙동강생물자원관 다양성연구팀, ⁷순천향대학교 SW중심대학단 AI-SW교육원, ⁸한국기초과학지원연구원 국가연구시설장비진흥센터, ⁹주지앤시바이오, ¹⁰한국연체동물연구소

Identification and Bioinformatic Analysis of Aquaporin Gene Sequences in the Endangered Freshwater Snail *Koreoleptoxis nodifila*

Chan-Eui Hong^{1,2}, Jun Yang Jeong^{2,3}, Hyeon Jun Shin^{1,2}, Hyeok Lee^{1,2}, Kyoung Won Lee^{1,2}, Yi Seon Jang^{1,2}, Min Kyu Sang^{2,3}, Jie Eun Park^{2,3}, Dae Kwon Song^{2,3}, Hee Ju Hwang⁴, Se Won Kang⁵, So Young Park⁶, Ki Wan Kim⁷, Sung Sin Jo¹, Jung Han Park⁸, Hong Seog Park⁹, Jun Sang Lee¹⁰, Yong Hun Jo^{1,2} and Yong Seok Lee^{1,2,3}

¹Department of Biology, College of Natural Sciences, Soonchunhyang University, Asan, Chungnam 31538, Korea,

²Korea Native Animal Resources Utilization Convergence Research Institute (KNAR), Soonchunhyang University, Asan, Chungnam, Korea,

³Research Support Center for Bio-Bigdata Analysis and Utilization of Biological Resources, Soonchunhyang University, Asan, Chungnam, Korea,

⁴Soonchunhyang Institute of Medi-bio Science, Soonchunhyang University, Cheonan 31151, Korea.

⁵Biological Resource Center, Korea Research Institute of Bioscience and Biotechnology, Jeongeup, Jeonbuk 56212, Korea,

⁶Biodiversity Research Team, Animal & Plant Research Department, Nakdonggang National Institute of Biological Resources, Sangju, Gyeongbuk, 37242, Korea,

⁷Department of Artificial Intelligence and Software, National Center of Excellence in Software, Soonchunhyang University, Asan, Chungnam 31538, Korea

⁸National Research Facilities and Equipment Center, Korea Basic Science Institute (KBSI), Daejeon, South Korea

⁹Research Institute, GnC BIO Co., LTD., 621-6 Banseok-dong, Yuseong-gu, Daejeon, 34069, Korea,

¹⁰Korea Mollusca Research Institute, 238-87, Dongnae-ro, Chuncheon-si, Kangwon-do, 24399

ABSTRACT

Aquaporins (AQP) are membrane proteins that regulate cellular water balance and are widely expressed in both animals and plants. AQP contributes to maintaining cellular water homeostasis and body fluid balance under various environmental conditions. This is essential for mollusks, and AQP is known to be important factors involved in this process. The genetic information of *Koreoleptoxis nodifila*, an endangered gastropod from Korea, is poor, and there is no report on AQP. In this study, we analyzed transcriptome data obtained using the Illumina HiSeq 2000 platform to identify the AQP gene in *K. nodifila*. The *K. nodifila* AQP gene contains a 1,056-bp open reading frame (ORF) encoding a protein consisting of 352 amino acids. Conserved domain analysis revealed that this protein has six alpha-helical transmembrane structures, which suggests that it is classified as a membrane protein. InterPro analysis confirmed that this protein belongs to the major intrinsic protein (MIP) superfamily. Multiple sequence alignment using ClustalX2 identified a conserved asparagine-proline-alanine (NPA) motif associated with water transport functions. Also, endangered *K. nodifila* is difficult to verify through molecular biological experiments; therefore, an *in silico* analysis was performed. The results confirmed that *K. nodifila* AQP belongs to the Pleuroceridae lineage and is positioned within Gastropoda. This study is the first to elucidate the AQP gene sequence of *K. nodifila* and is expected to be used as basic data for securing genetic resources and functional studies.

Received: March 13, 2025; Revised: March 20, 2025;
Accepted: March 28, 2025

Corresponding author: Yong Seok Lee

Tel: +82 (041) 530-3040, e-mail: yslee@sch.ac.kr
1225-3480/24892

This is an Open Access Article distributed under the terms of the Creative Commons Attribution Non-Commercial License with permits unrestricted non-commercial use, distribution, and reproducibility in any medium, provided the original work is properly cited.

Keywords: *Koreoleptoxis nodifila*, aquaporin, bioinformatics analysis

서 론

Aquaporin (AQP) 은 동물계와 식물계 전반에 걸쳐 광범위하게 발현되며 세포내 또는 세포간 물의 흐름 조절에 관여하는 세포막 단백질로, AQP의 발현은 생체 내 주요 용매인 물이 인지질 이중층으로 이루어진 세포막을 이동하는 기작을 밝히는데 주요 인자로 작용하였다. AQP는 인간에서 최소 13 개 family가 확인되었으며, 효모, 박테리아 등의 경우에서도 AQP family가 확인되었다 (Campbell *et al.*, 2008; Verkman, 2013; Jung *et al.*, 2020). AQP는 고도로 보존된 Asn-Pro-Ala (NPA) 서열과 같은 특징적인 모티프를 포함하고 있으며, 관련 모티프의 넓은 분포와 높은 서열보존성은 AQP가 수행하는 세포의 수분 항상성과 체액 균형 유지가 생물 내에서 매우 중요함을 시사한다 (Krane and Goldstein, 2007; Ikeda *et al.*, 2011; Ibrahim and Soladoye, 2019). 또한 최근 연구에 따르면 AQP는 세포막의 물 수송 기능 뿐만 아니라, 가뭄, 염분 변화, 건조 등 환경적 스트레스에 반응하여 세포와 조직이 삼투압 변화에 적응할 수 있도록 돕는 것으로 나타났다 (Kapilan *et al.*, 2018; Lorente-Martínez *et al.*, 2023). 선행 연구된 사례들을 살펴보았을 때 식물에서는 AQP의 조절이 수분 부족 환경에서 수분 균형을 유지하는데 핵심적인 역할을 하며, 동물에서는 AQP가 암모니아 배출과 건조 저항성 같은 적응 과정에 기여한다고 확인되었다 (Kapilan *et al.*, 2018; Lorente-Martínez *et al.*, 2023).

연체동물은 해양, 담수, 육상 등 다양한 환경에 서식하며, 지구상에서 종다양성이 두번째로 큰 동물 문 (Phylum) 에 속한다 (Haszprunar, 2020). 체표면에 분포된 점액을 통해 호흡을 하는 연체동물의 생리학적 특성상, 다양한 환경에 적응하며 생존하기 위해서는 체내 수분 평형과 삼투압 조절이 필수적이다 (Williams, 1855; Liu *et al.*, 2025). 국내 중·북부 하천에 분포하는 담수산 복족류인 *Koreoleptoxis nodifila* (염주알다슬기) 는 수질이 양호하고 수심이 깊은 하천 상류지역에 서식하는 것으로 알려져 있다. 하지만 지속적인 서식지의 오염과 생태계 파괴 그리고 남획으로 인해 출현지역 및 개체수의 감소로 인하여 2017년도에 멸종위기 야생동물 II급으로 지정되었다. 종 절멸 위기에서 보호되어야 하는 현황에도 불구하고 *K. nodifila* 관련 연구는 개체의 생물학적 특성 및 분포 현황만이 연구되었으며, 번식 현황, 서식지 동향, 분류학적 특성 등에 대한 기초자료는 부족한 상황이다 (Lee and Min, 2002; Lee *et al.*, 2005; Song *et al.*, 2022). 또한 이 종에 대해 NCBI에 등록되어 있는 *K. nodifila*의 유전정보는 Nucleotide 44건, Protein 42건, SRA 데이터 2건에 불과하며, 연체동물의 환경적응에 주요한 작용을 할 것으로 예측되는 AQP 유전자에 대한 유전 정보 또한 전무한 실정이다.

따라서 이 연구는 *K. nodifila* 전사체 데이터에서 확보한 AQP 유전자를 생물정보학적으로 분석하여 정확한 서열을 탐색하기 위해 수행되었다. 도출된 결과들은 유전자원의 확보 및 기능적 연구를 위한 기초 데이터로 활용될 것으로 사료되며 멸종위기종 보호 및 보전 전략에 대한 기초 자료로서 활용될 수 있을 것으로 기대된다.

재료 및 방법

K. nodifila (염주알다슬기) 는 Cerithioidea (짜부락고등상과), Pleuroceridae (다슬기과) 에 속하는 복족류로 원주시 방항경청의 허가를 받아 2014년 7월 강원도 영월군에서 채집을 진행하였다.

1. cDNA library 구축

*K. nodifila*의 total RNA 추출은 냉동 상태의 내장낭 조직 300 mg을 분쇄한 샘플을 사용하여 Trizol Reagent (3 mL, Thermo Fisher Scientific, USA) 를 이용해 제조사 프로토콜에 따라 수행하였다. 추출된 total RNA의 농도와 품질은 NanoDrop 분광광도계와 Agilent 2100 BioAnalyzer (Agilent Technologies) 의 RNA Nano Chip을 사용하여 평가하였다. 이후, Absolutely mRNA Purification Kit (Stratagene, CA, USA) 를 이용하여 total RNA로부터 mRNA를 정제한 후, cDNA 합성을 수행하였다. 정제된 mRNA는 cDNA Library Construction Kit (Stratagene, CA, USA) 를 사용하여 cDNA 라이브러리로 구축하였으며, Illumina HiSeq 2000 시퀀서를 이용한 전사체 분석을 진행하였다.

2. *K. nodifila* Unigenes 확보

Illumina HiSeq 2000 플랫폼에서 획득된 raw reads는 Cutadapt (v1.11) (Martin, 2011) 와 Sickle (v1.33) (Joshi and Fass, 2011) 프로그램을 사용하여 어댑터 서열과 저품질 서열을 제거함으로써 clean reads를 확보하였다. 확보된 clean reads는 Trinity (v2.0.4) (Grabherr *et al.*, 2011) 를 사용하여 *De novo* 어셈블리 (Grabherr *et al.*, 2011; Kang *et al.*, 2017) 를 수행하여 contig 서열을 생성하였다. 생성된 contigs는 TransDecoder (Haas and Papanicolaou, 2016) 를 사용하여 Open Reading Frame (ORF) 을 예측하였으며, TGICL (TIGR Gene Indices Clustering Tools v2.1) (Pertea *et al.*, 2003) 을 활용하여 ORF 서열을 클러스터링하여 Unigene을 구축하였다.

3. *K. nodifila* aquaporin (AQP) 서열 확보

361	AGG	CGG	GAG	TGT	CGC	GAT	CAA	AAC	CAT	GTC	TTC	GCA	AGT	GCA	CCC	CGC	TAG	CAC	GGC	GCT	420
121	R	E	Q	T	L	Q	S	Q	A	L	S	E	I	R	N	P	Y	P	A	M	140
421	CAG	AGA	ACA	GAC	CCT	GCA	GTC	TCA	GGC	ATT	GTC	CGA	AAT	ACG	AAA	CCC	ATA	TCC	CGC	CAT	140
141	E	G	L	Y	V	Y	T	A	N	K	L	E	A	I	R	T	G	T	D	E	160
481	GGA	GGG	GCT	ATA	CGT	TTA	CAC	TGC	TAA	TAA	GCT	GGG	AGC	TAT	TCG	CAC	CGG	CAC	TGA	CGA	160
161	K	K	S	Y	K	I	F	D	V	R	E	I	K	S	I	S	F	W	R	180	
541	GAA	AAA	ATC	GTC	TTA	CAA	AAAT	CTT	CGA	TGT	CCG	AGA	AAAT	ATC	CAT	CAG	TTT	CTG	GAG	180	
181	A	V	F	A	E	F	L	A	Q	I	L	F	V	F	L	G	A	S	A	200	
601	GGC	AGT	TTT	TGC	CGA	GTT	TTT	GGC	GCA	GAT	ACT	TTT	TGT	GTT	CCT	CGG	TGG	TGC	ATC	GGC	200
201	K	F	T	K	E	T	E	D	A	Q	Y	N	A	Q	I	K	I	A	L	G	220
661	TAA	GTT	CAC	CAA	AGA	AAC	GGA	GGA	TGC	TTA	CMA	CGC	ACA	GAT	AAAT	CAA	GAT	CGC	CCT	GGG	220
221	F	G	L	S	I	M	A	L	Q	M	I	G	H	V	S	G	H	I	S	I	240
721	TTT	CGG	CCT	TTC	AAT	CAT	GGC	GCT	CAT	CCA	GAT	GAT	TGG	TCA	CGT	GAG	CGG	CGG	CCA	TAT	240
241	N	P	A	V	T	I	A	M	A	V	V	M	N	I	S	V	V	R	A	V	260
781	TAA	CCC	TGC	CGT	TAC	CAT	CGC	GAT	GGC	GGT	GGT	CAT	GAA	CAT	TTC	TGT	CGT	CAG	AGC	GGT	260
261	F	Y	V	I	A	Q	C	L	G	A	I	V	G	G	Y	I	L	L	G	L	280
841	TTT	CTA	CGT	CAT	CGC	CCA	GTG	CCT	TGG	AGC	TAT	CGT	TGG	TGG	CTA	CAT	TTT	ACT	AGG	ATT	280
281	T	P	H	S	F	N	D	D	L	A	V	T	F	I	H	K	D	L	T	V	300
901	GAC	GCC	TCA	CAG	TTT	CAA	CGA	CGA	CCT	GGC	CGT	GAC	GTT	CAT	CCA	CAA	AGA	CCT	GAC	AGT	300
301	G	Q	G	V	G	V	E	I	I	L	T	F	V	L	V	F	V	I	F	G	320
961	GGG	GCA	GGG	TGT	CGG	GGT	GGA	GAT	CAT	CCT	TAC	CTT	TGT	CCT	CGT	CTT	CGT	CAT	CTT	CGG	1020
321	T	T	D	P	N	R	P	S	F	G	S	P	S	L	L	I	G	L	T	V	340
1021	CAC	CAC	CGA	TCC	CAA	CCG	CCC	GTC	CTT	CGG	CAG	CCC	TTC	GCT	GCT	CAT	CGG	GCT	GAC	AGT	1080
341	T	L	L	L	L	S	G	V	S	F	T	G	A	S	M	N	P	S	R	S	360
1081	CAC	CCT	CCT	CCA	TCT	GTC	CGG	GGT	GAG	CTT	CAC	AGG	TGC	CAG	CAT	GAA	TCC	GTC	TGC	GTC	1140
361	L	G	S	A	V	A	S	N	R	F	S	N	H	W	V	Y	W	V	G	P	380
1141	GCT	GGG	TTT	TGC	CGT	GGC	GTC	CAA	CAG	GTT	TAG	TAA	TCA	TTG	GGT	GTA	CTG	GGT	GGG	GCC	1200
381	I	L	G	G	I	L	A	A	G	T	Y	K	L	I	I	N	P	Y	R	K	400
1201	CAT	TCT	GGG	AGG	CAT	CCT	GGC	AGC	AGG	CAC	GTA	CAA	GCT	GAT	CAT	CAA	CCC	CTA	CAG	GAA	1260
401	A	L	T	M	D	E	A	I	Q	K	M	Y	S	D	A	P	I	F	P	L	420
1261	AGC	ACT	CAC	CAT	GGA	CGA	AGC	CAT	ACA	GAA	AAAT	GTA	TTC	CGA	CGC	GCC	CAT	CTT	TCC	ACT	1320
421	Q	D	E	N	F	V	A	I	P	R	D	Y	F	K	A	S	D	G	K	R	440
1321	GCA	AGA	CGA	GAA	CTT	TGT	GGC	CAT	ACC	ACG	TGA	CTA	CTT	CAA	GGC	TTC	CGA	CGG	AAA	ACG	1380
441	A	N	G	K	E	V	T	E	S	S	N	F	*	T	*	S	A	I	E	Y	460

Fig. 1. The complete nucleotidesequence of Aquaporin and deduced amino acid sequence of *K. nodifila*. (A) The red box area represents the ORF of AQP, (B) Red letter : Characteristic part of MIP protein.

*K. nodifila*의 Unigene 서열은 BLASTx 프로그램과 PANM-DB (ver. 5.1) (Kang *et al.*, 2015; Song *et al.*, 2023) 를 활용하여 기능적 주석 (annotation) 을 수행하였다. 주석 결과, AQP로 확인된 Unigenes 중 full-length 서열을 확보하기 위해 추가 분석을 진행하였다. 이를 위해 NCBI nr (non-redundant) 데이터베이스에서 BLAST 분석을 실시하여 서열을 확인하였다. 확보한 AQP full-length 서열은 EMBOSS package의 sixpack tool (Rice *et al.*, 2000) 을 활용하여 분석하였다. 이를 통해 *K. nodifila*의 AQP 아미노산 서열을 확보하였다.

4. *K. nodifila*의 Aquaporin 단백질의 2차 구조 및 3차 구조 예측

NCBI에 등록된 서열 데이터를 활용하여 BLAST 분석을 수행하였으며, 유사한 서열에서 conserved domain을 확인하였다. 또한, TMHMM Server v.2.0 (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>) 을 이용하여 transmembrane helices를 예측하였으며, *K. nodifila*의 motif 분석은 InterProScan (<https://www.ebi.ac.uk/interpro/search/sequence>) 을 이용하여 확인하였다. PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred>) 로 단백질의 2차 구조를 예측하였으며, Phyre2 (<http://bioinf.cs.ucl.ac.uk/psipred>) 소프트웨어를 활용하여 3차 구조를 예측하고, 시각화 하였다.

5. Multiple Sequence Alignment 및 Molecular Phylogenetic Analysis

이전에 수행된 *Charonia lampas* (나팔고둥) (Jung *et al.*,

2020) 의 AQP 서열을 포함한 NCBI nr (Non-redundant) 데이터베이스에서 BLAST 분석을 수행한 뒤, 근연종의 aquaporin complete 서열을 FASTA 형식으로 확보하였다. 이후 ClustalX2 (v2.1) 프로그램을 이용하여 multiple sequence alignment를 수행하였다 (Larkin *et al.*, 2007). Phylogenetic 분석은 MEGA-X 프로그램을 활용하여 Maximum-Likelihood (ML) 방법으로 수행하여, 확보한 *K. nodifila* AQP 서열과 다른 연체동물 AQP 서열 간의 계통적 분지를 서열의 유사도에 맞춰 확인하였다 (Kumar *et al.*, 2018).

결과 및 고찰

K. nodifila AQP 유전자의 ORF는 총 352개의 아미노산, 즉 1,056 bp로 구성되어 있음을 확인하였다. 해당 서열은 246개의 adenine (A), 191개의 thymine (T), 298개의 guanine (G), 321개의 cytosine (C) 로 이루어져 있으며, GC contents는 58.6%로 확인되었다 (Fig. 1). Sixpack 프로그램을 이용하여 *K. nodifila*의 AQP 아미노산 서열을 예측한 결과, 보존된 영역 (conserved domain) 과 두 개의 주요 NPA (N-Asparagine, P-Proline, A-Alanine) 모티프가 확인되었다. 첫 번째 NPA 모티프는 241-243번째 아미노산 위치에서, 두 번째 NPA 모티프는 1080-1082번째 아미노산 위치에서 발견되었다. 이러한 특징은 AQP 단백질의 특성을 반영하며, 막 단백질로서의 기능과 관련이 있는 것으로 판단된다 (Fig. 1). 복족류의 AQP 단백질 구조적 특성을 비교하기 위해, 이전에 수행한 *C. lampas* (Jung *et al.*, 2020) 와 *K. nodifila*의

멸종위기 야생동물 II급 염주알다슬기의 Aquaporin 유전자 발굴 및 생물정보학적 분석

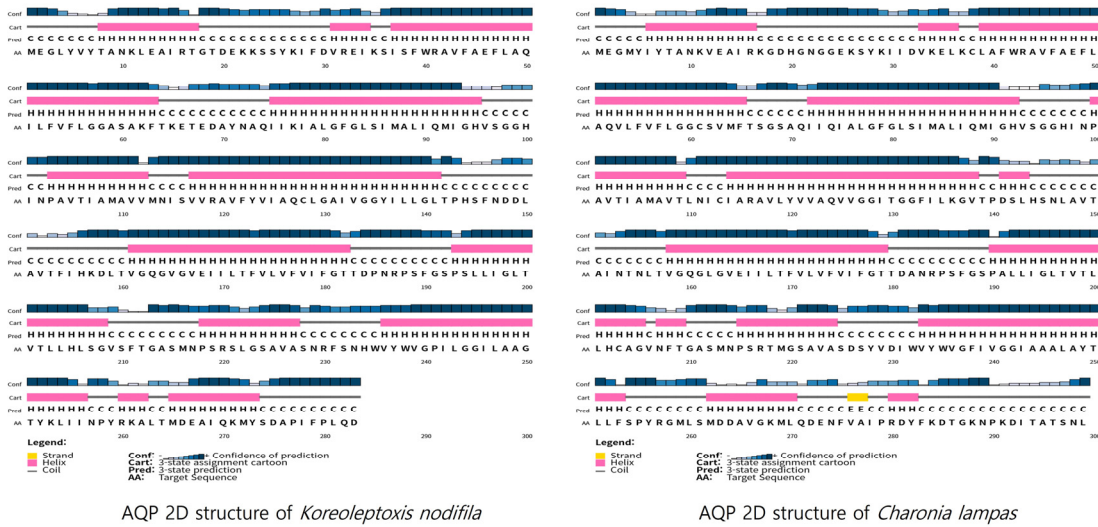


Fig. 2. The predicted 2D structure of aquaporin sequence by PSIPRED with *C. sauliae* and *K. nodifila*.

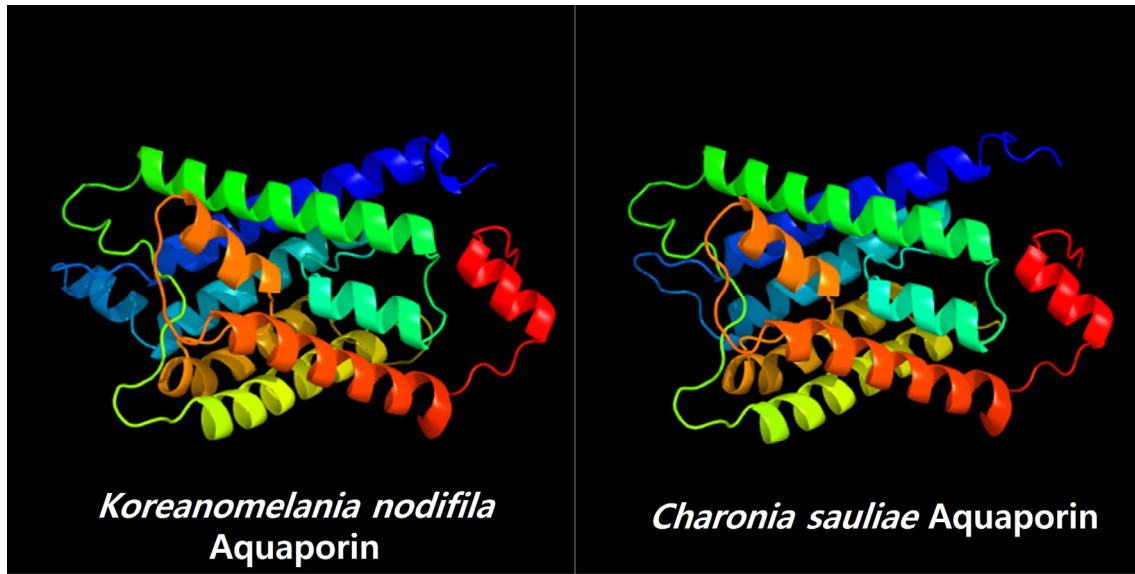


Fig. 3. The predicted 3D structure of aquaporin sequence by Phyre2 with *K. nodifila* and *C. sauliae*.

2D 및 3D 구조 분석을 수행하였다 (Fig. 2, Fig. 3). 분석 결과, *K. nodifila*는 15개의 alpha-helix 구조를 가지는 것으로 예측되었으며, *C. lampas*는 14개의 alpha-helix 구조와 1개의 beta-strand 구조를 가지는 것으로 예측되었다. 막 관통 alpha-helix는 15-30개의 아미노산을 포함한다는 점을 기반으로 확인하였을 때 두 종 모두 막 관통 alpha-helix는 6개로 예측되었다 (Arkin and Brunger, 1998). 이러한 구조적 특징은 AQP 단백질이 생체막에서 선택적으로 물을 수송하는 역할을 수행하는데 필수적인 요소로 알려져 있다 (Liu *et al.*,

2005). 또한 alpha-helix의 배열이 두 종에서 보존되어 있다는 점은, AQP 단백질이 복족류 내에서 기본적인 기능을 유지하고 있음으로 사료된다. TMHMM Server v.2.0 을 사용하여 분석한 결과, *K. nodifila*의 AQP 단백질은 6개의 transmembrane helices를 포함하고 있음을 확인하였다 (Fig. 4). 기존 연구에서 보고된 AQP 단백질의 6개 막관통 도메인 및 세포 내·외 위치 특성과 일치하며, 이는 AQP 단백질이 30kDa 단량체로 구성된 4량체 구조를 형성하며, 각각 독립적인 물 수송 채널로 기능한다는 점을 뒷받침한다 (Feher,

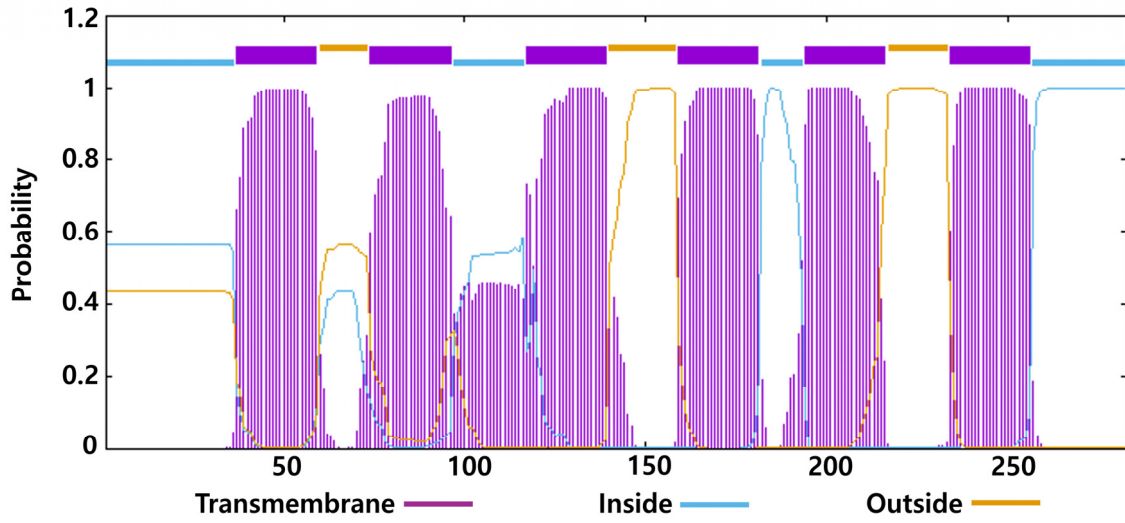


Fig. 4. Prediction of transmembrane helices in *K. nodifila* aquaporin using TMHMM-2.0.

361	AGG	CGG	GAG	TGT	CGC	GAT	CAA	AAC	CAT	GTC	TTC	GCA	AGT	GCA	CCC	CGC	TAG	CAC	GGC	GCT	420
121	R	E	Q	T	L	Q	S	Q	A	L	S	E	I	R	N	P	Y	P	A	M	140
421	CAG	AGA	ACA	GAC	CCT	GCA	GTC	TCA	GGC	ATT	GTC	CGA	AAT	ACG	AAA	CCC	ATA	TCC	CGC	CAT	480
141	E	G	L	Y	V	Y	T	A	N	K	L	E	A	I	R	T	G	T	D	E	160
481	GGA	GGG	GCT	ATA	CGT	TTA	CAC	TGC	TAA	TAA	GCT	AGG	AGC	TAT	TCG	CAC	CGG	CAC	CGA	540	
161	K	K	S	S	Y	K	I	F	D	V	R	E	I	K	S	I	S	F	W	R	180
541	GAA	AAA	ATC	GTC	TTA	CAA	AAT	CTT	CGA	TGT	CCG	AGA	AAT	CAA	ATC	CAT	CAG	TTT	CTG	GAG	600
181	A	V	F	A	E	F	L	A	Q	I	L	F	V	F	L	G	G	A	S	A	200
601	GGC	AGT	TTT	TGC	CGA	GTT	TTT	GGC	GCA	GAT	ACT	TTT	TGT	GTT	CCT	CGG	TGG	TGC	ATC	GGC	660
201	K	F	T	K	E	T	E	D	A	Y	N	A	Q	I	I	K	I	A	L	G	720
661	TAA	GTT	CAC	CAA	AGA	AAC	GGA	GGA	TGC	TTA	CAA	CGC	ACA	GAT	AAT	CAA	GAT	CGC	CCT	GGG	720
221	F	G	L	S	I	M	A	L	I	Q	M	I	G	H	V	S	G	G	H	I	240
721	TTT	CGG	CCT	TTC	AAT	CAT	GGC	GCT	CAT	CCA	GAT	TGG	TCA	CGT	GAG	CGG	CGG	CCA	TAT	780	
241	N	P	A	V	T	I	A	M	A	V	V	M	N	I	S	V	V	R	A	V	260
781	TAA	CCC	TGC	CGT	TAC	CAT	CGC	GAT	GGC	GGT	GGT	CAT	GAA	CAT	TTC	TGT	CGT	CAG	AGC	GGT	840
261	F	Y	V	I	A	Q	C	L	G	A	I	V	G	Y	I	L	L	G	I	L	280
841	TTT	CTA	CGT	CAT	CGC	CCA	GTG	CCT	TGG	AGC	TAT	CGT	TGG	TGG	CTA	CAT	TTT	ACT	AGG	ATT	900
281	T	P	H	S	F	N	D	D	L	A	V	T	F	I	H	K	D	L	T	V	300
901	GAC	GCC	TCA	CAG	TTT	CAA	CGA	CGA	CCT	GGC	CGT	GAC	GTT	CAT	CCA	CAA	AGA	CCT	GAC	AGT	960
301	G	Q	G	V	G	V	E	I	I	L	T	F	V	L	V	I	F	G	I	F	320
961	GGG	GCA	GGG	TGT	CGG	GGT	GGA	GAT	CAT	CCT	TAC	CTT	TGT	CCT	CGT	CTT	CGT	CAT	CIT	CGG	1020
321	T	T	D	P	N	R	P	S	F	S	P	S	L	I	G	L	I	T	V	S	340
1021	CAC	CAC	CGA	TCC	CAA	CCG	CCC	GTC	CTT	CGG	CAG	CCC	TTC	GCT	GCT	CAT	CGG	GCT	GAC	AGT	1080
341	T	L	L	H	L	S	G	V	S	F	T	G	A	S	M	N	P	S	R	S	360
1081	CAC	CCT	CCT	CCA	TCT	GTC	CGG	GGT	GAG	CIT	CAC	AGG	TGC	CAG	CAT	GAA	TCC	GTC	TCG	GTC	1140
361	L	G	S	A	V	A	S	N	R	F	S	N	H	W	V	Y	W	V	G	P	380
1141	GCT	GGG	TTC	TGC	CGT	GGC	GTC	CAA	CAG	GTT	TAG	TAA	TCA	TTG	GGT	GTA	CTG	GGT	GGG	GCC	1200
381	I	L	G	G	I	L	A	A	G	T	Y	K	L	I	I	N	P	Y	R	K	400
1201	CAT	TCT	GGG	AGG	CAT	CCT	GGC	AGC	AGG	CAC	GTA	CAA	GCT	GAT	CAT	CAA	CCC	CTA	CAG	GAA	1260
401	A	L	T	M	D	E	A	I	Q	K	M	Y	S	D	A	P	I	F	P	L	420
1261	AGC	ACT	CAC	CAT	GGA	CGA	AGC	CAT	ACA	GAA	AAT	GTA	TTC	CGA	CGC	GCC	CAT	CIT	TCC	ACT	1320
421	Q	D	E	N	F	V	A	I	P	R	D	Y	F	K	A	S	D	G	K	R	440
1321	GCA	AGA	CGA	GAA	CTT	TGT	GGC	CAT	ACC	ACG	TGA	CTA	CTT	CAA	GGC	TTC	CGA	CGG	AAA	ACG	1380
441	A	N	G	K	E	V	T	E	S	S	N	F	*	T	*	S	A	I	E	Y	460

Fig. 5. Identification of the MIP superfamily domain in *K. nodifila* aquaporin based on InterPro analysis. The regions highlighted in red indicate the conserved Major Intrinsic Protein (MIP) superfamily domain, as predicted by InterPro.

2017). 또한, AQP는 세포질 내에 위치하는 소수성 말단 (amino 및 carboxyl 그룹) 을 포함하고 있으며, 막 단백질로서의 기능을 수행하는 데 필수적인 구조적 특징을 가지고 있다 (Bai *et al.*, 1996). InterPro (<https://www.ebi.ac.uk/interpro/>) 분석을 수행한 결과, *K. nodifila*의 AQP 단백질은 세포막을 통한 물과 용질의 투과성을 촉진하는 MIP (Major Intrinsic Protein) superfamily에 속하는 것으로 확인되었으며, 두 개의 주요 보존된 영역 (conserved domains) 이 존재함을 확인하였다 (Fig. 5). MIP superfamily의 단백질들은

세포 내외 및 세포 간 물 이동을 조절하는 막 통합 단백질 (membrane-integrated protein) 로 기능하며, 일부 단백질은 작은 중성 분자, 이온을 선택적으로 운반하는 것으로 알려져 있다 (Verma *et al.*, 2015).

동일한 연구팀에 의해 선행 수행된 *C. lampas*의 연구 결과를 포함하여 연체동물의 AQP 아미노산 서열 12종을 대상으로 다중 서열 정렬 (multiple sequence alignment) 을 수행한 결과, 연체동물문 내에서 특이적인 서열 패턴의 유사성이 확인되었다 (Fig. 6). 분석 결과, 대부분의 서열에서

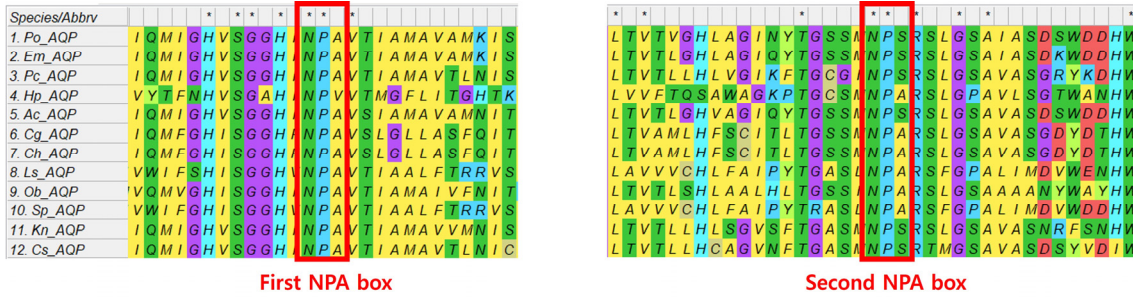


Fig. 6. Multiple sequence alignment results of *K. nodifilia* and 11 AQP protein reference sequences. There are First NPA box and second NPA box. The analysis was performed by clustalX2 using representative amino acid sequences from *Plakobranthus ocellatus* (GFO22746.1), *Elysia marginata* (GFS14283.1), *Pomacea canaliculata* (AYH91719.1), *Helix pomatia* (AOS89752.1), *Aplysia californica* (XP_005111061.2), *Crassostrea gigas* (XP_011446842.1), *Crassostrea hongkongensis* (AJF12150.1), *Lymnaea stagnalis* (AHL29026.1), *Octopus bimaculoides* (XP_014767491.1), *Stagnicola palustris* (AHL29028.1), *Charonia sauliae* (SCH).

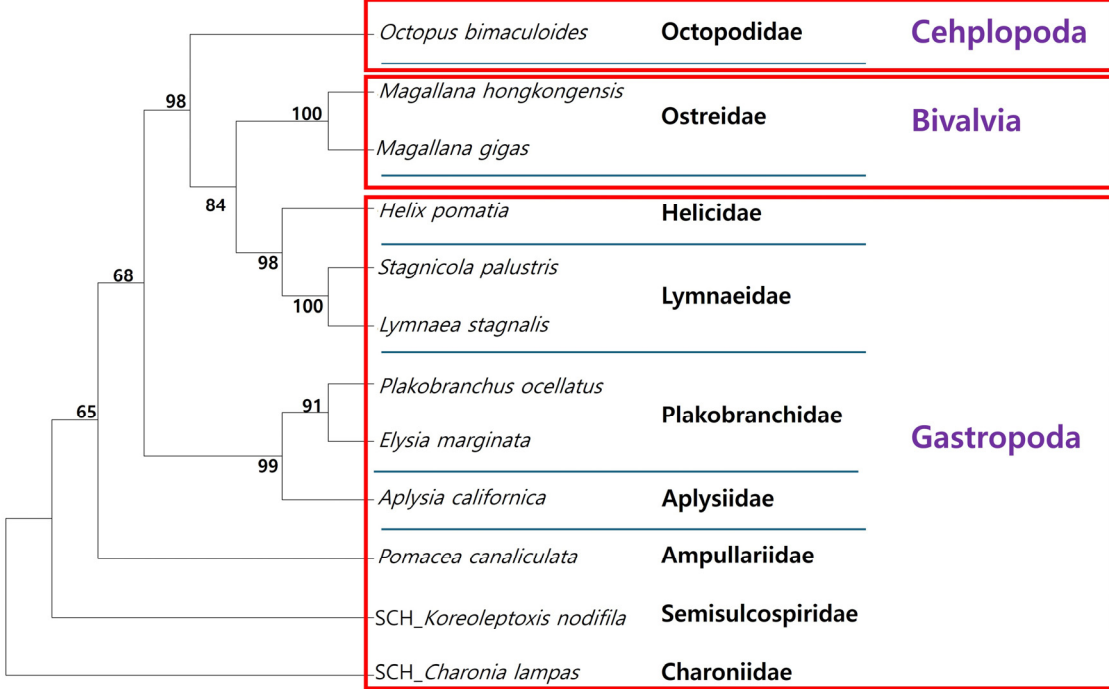


Fig. 7. The result of phylogenetic analysis on evolutionary relationships. It was divided into big three groups (Cephlopoda, Gastropoda, Bivalvia). In Gastropoda, there are 7 families (Helicidae, Lymnaeidae, Aplysiidae, Plakobranthidae, Ampullariidae, Pleuroceridae, Ranellidae). In Bivalvia, there are 1 family (Ostreidae). In Cephlopoda, there are 1 family (Octopodidae).

NPA motif (NPA box) 가 보존되어 있는 것으로 확인되었다. 그러나 일부 종에서는 First NPA box 및 Second NPA box에서 서열 불일치가 관찰되었으며, 이는 아미노산 alanine과 serine 사이의 mistranslation에 의해 발생한 것으로 사료된다 (Guo *et al.*, 2009). 이러한 변이는 특정 환경 적응이나 기능적 차이와 관련이 있을 가능성이 있으며, 이에 대한 추가적인 분자생물학적 실증 연구가 필요할 것으로 판단

된다. 다만 이번 연구 대상 종인 *K. nodifilia*는 멸종위기종으로, 분자생물학적 기법을 통한 검증에 한계가 있어, 확보한 *K. nodifilia* AQP 서열의 신뢰성을 평가하기 위해 *in silico* 분석을 수행하였다. MEGA-X 프로그램을 이용하여 Maximum Likelihood (ML) 방법으로 계통발생학적 분석 (phylogenetic analysis) 을 수행한 결과, 분석된 12종의 AQP 서열은 크게 3

개의 주요 그룹과 9개의 family로 분류됨이 확인되었다 (Fig. 7). Gastropoda 내에서 *C. lampas*, *K. nodifila*, *Pomacea canaliculata* 및 *Helix pomatia* 는 독립적인 그룹을 형성하였으며, Aplysiidae 및 Plakobanchidae는 각각 *Aplysia californica* (Aplysiidae) 및 *Elysia marginata*, *Plakobanchus ocellatus* (Plakobanchidae) 그룹으로 구분된다. *Stagnicola palustris*와 *Lymnaea stagnalis*는 Lymnaeidae로 묶이며 높은 bootstrap 점수를 확인할 수 있었다. Bivalvia의 Ostreidae에서는 *Magallana hongkongensis*와 *Magallana gigas*가 포함되었다. 해당 그룹은 monophyletic clade를 형성하였으며, 높은 bootstrap 점수를 기반으로 Bivalvia AQP의 보존성과 진화적 유사성이 확인되었다. *Octopus bimaculoides*는 Cephalopoda로 독립적인 그룹을 형성하였다. 이러한 결과는 확보된 *K. nodifila* AQP 서열이 높은 신뢰성을 가지며, Gastropoda에 속하는 다른 종의 AQP와 높은 유사성을 보임을 시사한다.

K. nodifila AQP 서열을 최초로 확보한 이번 연구 결과는, *K. nodifila* 에 대한 생물학적 이해를 증진시키는데 주요한 기초자료로써 활용 될 수 있으며, 향후 연구에서는 다양한 연체동물을 통해 확보된 AQP 서열을 활용하여, AQP가 연체동물의 환경 적응 과정에 미치는 명확한 생리학적 역할을 규명할 필요가 있을 것으로 판단된다. 또한, 지속적인 데이터 축적과 체계적인 분석을 통해 연체동물 AQP 연구의 기초 자료를 확충하고, 나아가 종 보호 및 보전에 기여할 수 있을 것으로 사료된다.

요 약

Koreanomelania nodifila (염주알다슬기) 는 멸종위기 II 급인 생물로 적극적인 유전자원확보가 필요한 상황이다. 이번 연구에서는 *K. nodifila*의 전사체 분석을 기반으로 aquaporin (AQP) 유전자를 확보하였다. *K. nodifila* AQP 유전자는 1,056 bp의 ORF를 가지며, 352개의 아미노산으로 구성되어 있음을 확인하였다. Conserved domain 분석 결과 6개의 transmembrane helix를 포함하는 전형적인 막단백질 구조를 보였다. 또한, InterPro 분석을 통해 MIP (Major Intrinsic Protein) superfamily에 속함을 확인하였다. 멸종위기종인 *K. nodifila*는 분자생물학적 실험을 통한 AQP 검증이 어려운 관계로, *in silico* 분석을 수행한 결과, *K. nodifila*의 AQP는 Pleuroceridae 계통에 속하며, Gastropoda내에 위치하는 것을 확인하였다. *K. nodifila*의 AQP 서열은 지속적인 데이터 축적과 체계적인 분석을 통해 연체동물 AQP 연구와 멸종위기종 보존과 보전에 대한 기초 자료로 활용될 수 있을 것으로 사료된다.

사 사

이 논문은 2021년도 정부 (교육부) 의 재원으로 한국연구재단의 (NRF-2017R1D1A3B06034971 / NRF-2021R1A6A1A03039503), 한국 기초과학지원연구원 국가연구시설장비센터 (2022R1A6C101B794), 순천향대학교 학술연구비의 지원을 받아 수행하였습니다.

REFERENCES

- Arkin, I.T., and Brunger, A.T. (1998) Statistical analysis of predicted transmembrane α -helices. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology*, **1429**: 113-128.
- Bai, L., Fushimi, K., Sasaki, S., and Marumo, F. (1996) Structure of Aquaporin-2 Vasopressin Water Channel (*). *Journal of Biological Chemistry*, **271**: 5171-5176.
- Campbell, E.M., Ball, A., Hoppler, S., and Bowman, A.S. (2008) Invertebrate aquaporins: a review. *Journal of Comparative Physiology B*, **178**: 935-955.
- Feher, J. (2017) 2.6 - Active Transport: Pumps and Exchangers. *In*: Feher J (ed) Quantitative Human Physiology (Second Edition). Academic Press, Boston, pp 170-181.
- Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., and Zeng, Q. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology*, **29**: 644-652.
- Guo, M., Chong, Y.E., Shapiro, R., Beebe, K., Yang, X.-L., and Schimmel, P. (2009) Paradox of mistranslation of serine for alanine caused by AlaRS recognition dilemma. *Nature*, **462**: 808-812.
- Haas, B., and Papanicolaou, A. (2016) TransDecoder (find coding regions within transcripts). Google Scholar.
- Haszprunar, G. (2020) Mollusca (Molluscs). *eLS*, **1**: 565-571.
- Ibrahim, O.R., and Soladoye, O. (2019) The role of aquaporins in the regulation of body fluids homeostasis. *Journal of Clinical Medicine of Kazakhstan*, : 15-20.
- Ikeda, M., Andoo, A., Shimono, M., Takamatsu, N., Taki, A., Muta, K., Matsushita, W., Uechi, T., Matsuzaki, T., and Kenmochi, N. (2011) The NPC motif of aquaporin-11, unlike the NPA motif of known aquaporins, is essential for full expression of molecular function. *Journal of Biological Chemistry*, **286**: 3342-3350.
- Joshi, N., and Fass, J. (2011) Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33)[Software].
- Jung, J.M., Hwang, H.J., Sang, M.K., Park, J.E., Song, D.K., Jung, J.Y., Jo, Y.H., Han, Y.S., Park, H.S., and Lee, J.S. (2020) 멸종위기야생동물 I 급 나팔고둥

- (*Charonia sauliae*) 의 Aquaporin 유전자를 활용한 분자 계통 분류학적 연구. *Korean J. Malacol* 36: 35-43.
- Kang, S.W., Park, S.Y., Patnaik, B.B., Hwang, H.J., Kim, C., Kim, S., Lee, J.S., Han, Y.S., and Lee, Y.S. (2015) Construction of PANM Database (Protostome DB) for rapid annotation of NGS data in Mollusks. *Korean Journal of Malacology*, **31**: 243-247.
- Kang, S.W., Patnaik, B.B., Hwang, H.-J., Park, S.Y., Chung, J.M., Song, D.K., Patnaik, H.H., Lee, J.B., Kim, C., and Kim, S. (2017) Sequencing and de novo assembly of visceral mass transcriptome of the critically endangered land snail *Satsuma myomphala*: annotation and SSR discovery. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, **21**: 77-89.
- Kapilan, R., Vaziri, M., and Zwiazek, J.J. (2018) Regulation of aquaporins in plants under stress. *Biological Research*, **51**: 4.
- Krane, C.M., and Goldstein, D.L. (2007) Comparative functional analysis of aquaporins/glyceroporins in mammals and anurans. *Mammalian Genome*, **18**: 452-462.
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution*, **35**: 1547-1549.
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., and Lopez, R. (2007) Clustal W and Clustal X version 2.0. *bioinformatics*, **23**: 2947-2948.
- Lee, J.S., and Min, D.G. (2002) 한국산 연체동물 목록. *The Korean Journal of Malacology* 18: 93-217.
- Lee, J.S., Min, D.G., and Min, K.A. (2005) 우렁이와 달팽이 pp. 민 패류연구소. 서울.
- Liu, K., Kozono, D., Kato, Y., Agre, P., Hazama, A., and Yasui, M. (2005) Conversion of aquaporin 6 from an anion channel to a water-selective channel by a single amino acid substitution. *Proceedings of the National Academy of Sciences*, **102**: 2192-2197.
- Liu, T., Nie, H., Huo, Z., and Yan, X. (2025) Genome-wide identification of aquaporin and their potential role in osmotic pressure regulation in *Ruditapes philippinarum*. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*: 101436.
- Lorente-Martínez, H., Agorreta, A., Irisarri, I., Zardoya, R., Edwards, S.V., and San Mauro, D. (2023) Multiple instances of adaptive evolution in aquaporins of amphibious fishes. *Biology*, **12**: 846.
- Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal*, **17**: 10-12.
- Pertea, G., Huang, X., Liang, F., Antonescu, V., Sultana, R., Karamycheva, S., Lee, Y., White, J., Cheung, F., and Parvizi, B. (2003) TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics*, **19**: 651-652.
- Rice, P., Longden, I., and Bleasby, A. (2000) EMBOSS: the European molecular biology open software suite. *Trends in genetics*, **16**: 276-277.
- Song, D.K., Lee, Y.S., Sang, M.K., Park, J.E., Jung, J.Y., Hong, C.E., Kim, Y.T., Lyu, Z., and Lee, H. (2023) PANM 데이터베이스 업데이트 (Version 5.1): 연체동물 전 사체 서열 분석의 정확도 증진을 위한 오염되어진 곰팡이 유전자 서열 검색기능 추가.
- Song, M.Y., Kim, D.H., Kim, K.H., and Kim, D.H. (2022) 한탄강 수계의 다슬기류 분포 특성. *The Korean Journal of Malacology* 38: 47-55.
- Verkman, A. (2013) Aquaporins. *Current Biology*, **23**: R52-R55.
- Verma, R.K., Gupta, A.B., and Sankararamakrishnan, R. (2015) Chapter Twenty-Three - Major Intrinsic Protein Superfamily: Channels with Unique Structural Features and Diverse Selectivity Filters. *In*: Shukla AK (ed) *Methods in Enzymology*. Academic Press, pp 485-520.
- Williams, T. (1855) XXIX.—On the mechanism of aquatic respiration and on the structure of the organs of breathing in invertebrate animals. *Annals and Magazine of Natural History*, **16**: 315-329.