

# 기수갈고둥 (*Clithon retropictus*) 의 Arginine kinase 유전자를 이용한 분자계통학적 연구

홍찬의<sup>1</sup>, 상민규<sup>1,3</sup>, 송대권<sup>1,3</sup>, 박지은<sup>1,3</sup>, 정준양<sup>1,3</sup>, 김용태<sup>1</sup>, 신현준<sup>2</sup>, 류자미<sup>1</sup>, 황희주<sup>3</sup>, 박소영<sup>4</sup>, 강세원<sup>5</sup>, 이준상<sup>6</sup>, 정기윤<sup>1</sup>, 이용석<sup>1,3</sup>

<sup>1</sup>순천향대학교 자연과학대학 생명과학과, <sup>2</sup>대구대학교 생명과학과, <sup>3</sup>순천향대학교 한국자생동물자원활용 융복합연구소, <sup>4</sup>국립낙동강생물자원관 전략기획실 혁신성과부, <sup>5</sup>한국생명공학연구원 생물자원센터, <sup>6</sup>순천향대학교 자연과학대학 기초과학연구소

## Molecular Phylogenetic Studies of *Clithon retropictus* Using Arginine Kinase Gene Sequence

Chan Eui Hong<sup>1</sup>, Min Kyu Sang<sup>1,3</sup>, Dae Kown Song<sup>1,3</sup>, Jie Eun Park<sup>1,3</sup>, Yong Tae Kim<sup>1</sup>, Hyun Jun Sin<sup>2</sup>, Liu Ziwei<sup>1</sup>, Hee Ju Hwang<sup>3</sup>, So Young Park<sup>4</sup>, Se Won Kang<sup>5</sup>, Jun Sang Lee<sup>6</sup>, Ki Yoon Jung<sup>1</sup> and Yong Seok Lee<sup>1,3</sup>

<sup>1</sup>Department of Biology, College of Natural Sciences, Soonchunhyang University, Asan, Chungnam 31538, Korea

<sup>2</sup>Department of Biological science, Daegu University, Gyeongsan 38453, Korea

<sup>3</sup>Korea Native Animal Resources Utilization Convergence Research Institute

<sup>4</sup>Performance Management Division, Strategic Planning Department, Nakdonggang National Institute of Biological Resources

<sup>5</sup>Biological Resource Center, Korea Research Institute of Bioscience and Biotechnology, Jeongeup, Jeonbuk 56212, Korea

<sup>6</sup>Institute for Basic Sciences, College of Natural Sciences, Soonchunhyang University, Asan, Chungnam, 31538, Korea

### ABSTRACT

*Clithon retropictus* is a small gastropod belonging to the Neritidae, and inhabits gravel areas in brackish waters. It has been designated as an endangered species as class II, because of its habitat conditions and reckless development of coastlines. In September 2021, *C. retropictus* registered in NCBI has only 80 nucleotides, 52 proteins, and 74 genes. In this context, we sequenced the *C. retropictus* cDNA library using the Illumina Hiseq 2500 platform. Arginine Kinase (AK) sequence was obtained by using BLAST program. To investigate the molecular phylogenetic analysis of *C. retropictus*, a dendrogram was generated using AK gene as a marker. Based on the phylogenetic tree, it is classified into three classes (Cephalopods, Gastropods, and Bivalvia). In addition, analysis of AK 2D structure using 38 species shows that no striking similarity was found. In conclusion, AK gene of *C. retropictus* consisted of 1,053 bp (351 amino acids) and could be utilized for molecular phylogenetic classification.

**Keyword:** *Clithon retropictus*, Arginine Kinase, phylogenetic tree

### INTRODUCTION

Received: September 15, 2021; Revised: September 26, 2021;  
Accepted: September 30, 2021

Corresponding author: Yong Seok Lee

Tel: +82 (41) 530-3040, e-mail: yslee@sch.ac.kr  
1225-3480/24799

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.

Arginine Kinase (AK) 는 Phosphagen Kinase group (PK) 에 속하는 유전자로 거의 모든 진핵생물의 에너지 항상성 조절에 관여한다. PK group은 ATP에 의존적인 인산화를 촉매하는 완충 시스템으로 척추동물에서는 phosphocreatine 이 유일한 PK group에 해당한다. 하지만 무척추동물에서는 phosphocreatine을 제외하고도 최소 6개의 고유한 PK group과 각각에 해당하는 kinase가 존재한다. 여기에는 phosphoglycocyanine (glycocyanine kinase, GK), phosphotaurocyanine (taurocyanine kinase, TK),

기수갈고둥 (*Clithon retropictus*) 의 Arginine kinase 유전자를 이용한 분자계통학적 연구

720	CGA	GAC	ACT	AAA	ACT	ACC	ACC	ATG	CCT	TCC	AAG	CAG	ACT	TGG	TCC	764
1	R	D	T	K	T	T	T	M	P	S	K	Q	T	W	S	15
765	GAC	AAG	GAT	ATG	CGT	CTG	AAT	GTG	GAC	GAG	CTG	TGG	AAG	AAA	CTG	809
16	D	K	D	M	R	L	N	V	D	E	L	W	K	K	L	30
810	GAA	ACT	GCC	ACC	AAC	TGC	CAG	TGC	CTC	CTC	AAA	AAA	CAC	CTG	ACC	854
51	E	T	A	T	N	C	Q	S	L	L	K	K	H	L	T	45
855	AAG	CCA	CTT	TTC	AAT	CAG	CTG	AAG	GGA	CTA	AAA	ACA	AAG	TTC	GGG	899
46	K	P	L	F	N	Q	L	K	G	L	K	T	K	F	G	60
900	GGA	ACA	CTC	GCC	GAC	TGT	ATC	CGA	TCA	GGA	TGC	CTG	AAC	CCC	GAC	944
61	G	T	L	A	D	C	I	R	S	G	C	L	N	P	D	75
945	AGC	GGT	GTG	GGC	ATC	TAC	GCC	TGG	GAC	CCC	GAG	GCC	TAC	ACG	GTG	989
76	S	G	V	G	I	Y	A	S	D	P	E	A	Y	T	V	90
990	TTC	GCG	CCC	GTG	CTG	GAC	GCC	GTG	ATC	AAG	GAC	TAC	CAC	AAG	GTG	1034
91	F	A	P	V	L	D	A	V	I	K	D	Y	H	K	V	105
1035	GAC	AAG	CTG	AAC	CAC	CCC	GAG	CCC	AGC	TTC	GGT	GAC	CTG	GAT	AAC	1079
106	D	K	L	N	H	P	E	P	S	F	G	D	L	D	N	120
1080	CTG	GGC	TTC	GGC	GAC	CTG	GAC	CCC	TCG	GGC	AAC	ATG	ATC	GTC	TCC	1124
121	L	G	F	G	D	L	D	P	S	G	N	M	I	V	S	135
1125	ACC	CGC	GTG	CGT	GTG	GGC	CGC	AGC	CAC	GAC	AGC	TAC	GGC	TTC	CCC	1169
136	T	R	V	R	V	G	R	S	H	D	S	Y	G	F	F	150
1170	CCC	GTG	CTG	AAG	AAG	GAG	GAC	CGC	CTG	GAC	ATG	GAG	AAG	AAG	ACC	1214
151	P	V	L	K	K	E	D	R	L	D	M	E	K	K	T	165
1215	GTG	GAG	GCG	CTG	AAG	AAG	CTG	GAC	GGG	GAC	CTG	GCT	GGC	GAC	TAC	1259
166	V	E	A	L	K	K	L	D	G	D	L	A	G	D	Y	180
1260	TTC	CCC	CTG	ACC	GGC	ATG	ACC	GCC	GAC	ACG	CAG	AAG	CAG	CTG	ATG	1304
181	F	P	L	T	G	M	T	A	D	T	Q	K	Q	L	M	195
1305	GAG	GAC	CAC	TTC	CTC	TTC	AAC	GAC	AGC	GAC	CGC	TTC	CTC	AAG	GCG	1349
196	E	D	H	F	L	F	N	D	S	D	R	F	L	K	A	210
1350	GCT	GGC	GGC	TAC	GAC	GAC	TGG	CCC	CTG	GGC	CGT	GGC	ATC	TAC	TTC	1394
211	A	G	G	Y	D	D	W	P	L	G	R	G	I	Y	F	225
1395	AAC	AAG	GCC	AAG	ACC	TTC	CTG	GTC	TGG	GTC	AAC	GAG	GAG	GAC	CAC	1439
226	N	K	A	K	T	F	L	V	W	V	N	E	E	D	H	240
1440	CTC	CGC	TTC	ATC	TCC	ATG	CAG	AAG	GGA	GGC	GAC	CTC	GCA	GAG	GTC	1484
241	L	R	F	I	S	M	Q	K	G	G	D	L	A	E	V	255
1485	TAC	AAG	CGC	CTG	GTC	AAG	GCC	ATC	CGC	GCC	CTG	GAG	AGC	TCT	GGA	1529
256	Y	K	R	L	V	K	A	I	R	A	L	E	S	S	G	270
1530	CTC	AGC	TTC	GCC	AAG	CGC	GAG	GGC	CTG	GGC	TAC	CTG	ACC	TTC	TGC	1574
271	L	S	F	A	K	R	E	G	L	G	Y	L	T	F	C	285
1575	CCC	TCC	AAC	CTG	GGC	ACC	ACG	CTG	CGT	GCC	TCC	GTG	CAC	ATC	AAG	1619
286	P	S	N	L	G	T	T	L	R	A	S	V	H	I	K	300
1620	ATC	CCC	AAG	CTG	GCC	GCC	ACC	CCT	GAG	TTC	AAG	GAG	TTC	TGC	GAC	1664
301	I	P	K	L	A	A	T	P	E	F	K	E	F	C	D	315
1665	AAA	TAC	AAC	ATC	CAG	GCC	CGA	GGC	ATC	CAC	GGG	GAG	CAC	ACC	GAG	1709
316	K	Y	N	I	Q	A	R	G	I	H	G	E	H	T	E	330
1710	TCG	GTG	GGC	GGC	GTG	TAC	GAC	ATC	TCC	AAC	AAG	CGC	CGT	CTG	GGC	1754
331	S	V	G	G	V	Y	D	I	S	N	K	R	R	L	G	345
1755	CTG	ACC	GAG	ATC	GAG	GCT	CTG	CAG	GAG	ATG	CGC	AGA	GGC	GTG	GAG	1799
346	L	T	E	I	E	A	L	Q	E	M	R	R	G	Y	E	360
1800	GCG	GTG	ATC	GCG	AAG	GAG	AAG	AGC	CTG	TAG	ACA	CCC	GAA	CTC	TCC	1844
361	A	V	I	A	K	E	K	S	L	*	T	P	E	L	S	375

Fig. 1. The complete nucleotide sequence of Arginine kinase and deduced amino acid sequence of *C. retropictus*.

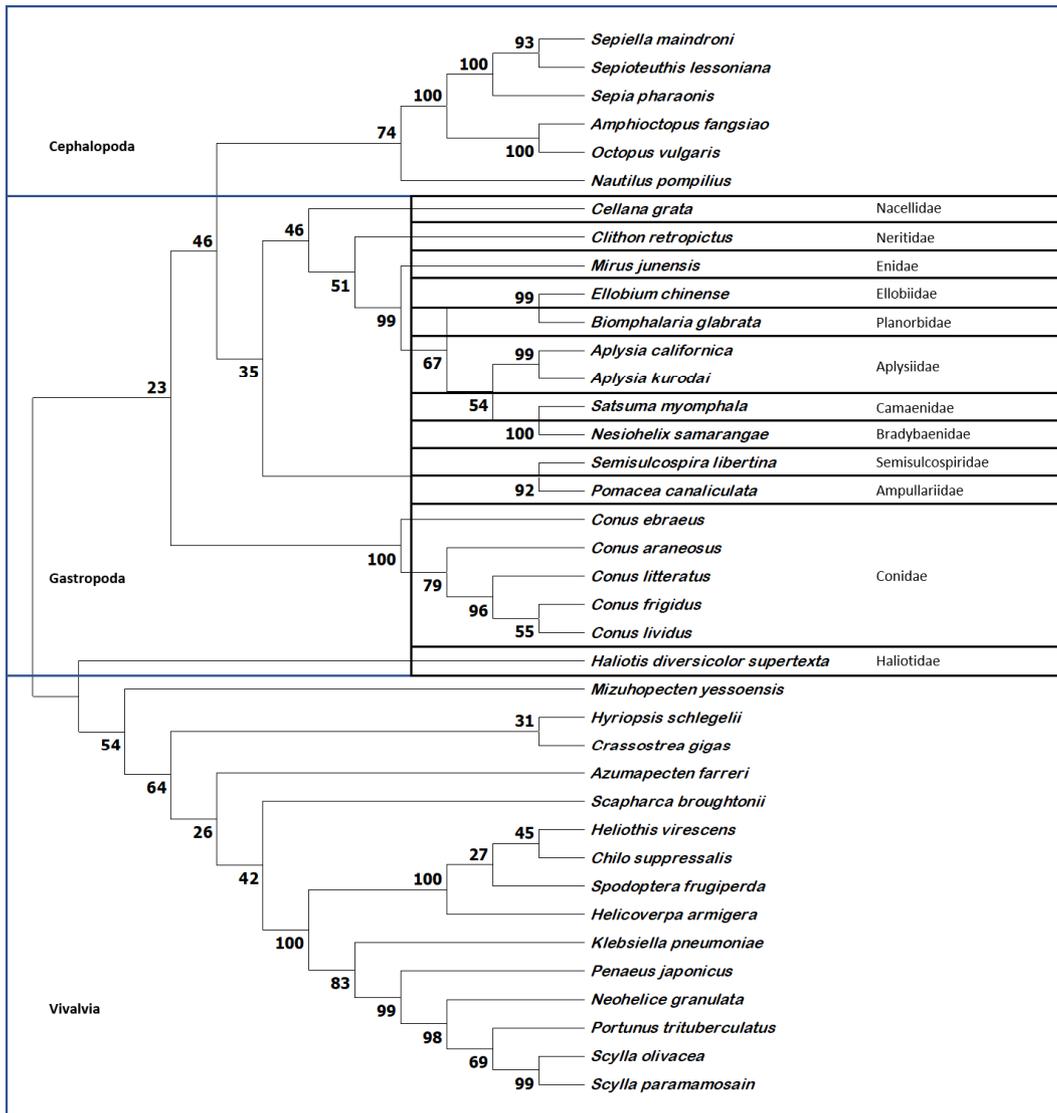
phospholombricine (lombricine kinase, LK), phosphoopheline (opheline kinase, OP), phosphohypotaurocyamine (hypotaurocyamine kinase, HTK), 그리고 phosphoarginine (arginine kinase, AK) 가 포함된다 (Tanaka *et al.*, 2007; Voncken *et al.*, 2013).

생물체의 대사활동에 필수적인 항상성과 관련한 AK 유전자는 세대가 반복되는 과정에서 큰 변동없이 보존되어 후손에게 전달되는 대표적 유전자로 알려져 있다. 이러한 AK 유전자는 환형동물, 연체동물, 절지동물 등의 다양한 무척추동물 전반에서 발견된다. 대부분의 AK는 단량체이며, 40 kDa 의 분자량을 가진다. 연체동물의 경우 선천성 면역 시스템이나 근육운동에 필요한 ATP의 보충 역할 또한 AK 유전자가 관여한다고 알려져 있다 (Suzuki *et al.*, 1998; Uda *et al.*, 2006; Jeong

and Lee, 2013).

기수갈고둥 (*Clithon retropictus*) 은 갈고둥과에 속하는 멸종위기 야생생물 II급인 생물로, 우리나라 남부지방과 제주도 등에 주로 분포하며 해외에서는 일본, 대만, 중국 남부에서 확인된 바 있으며, 남획의 우려는 없으나 서식 분포가 협소하여 하수구나 도랑 등에서 배출되는 소량의 오염원으로도 큰 피해를 입을 수 있어 주의가 필요하다고 알려져 있으며, 서식지 근처의 보나 제방시설에 의한 해수유입 차단, 무분별한 해안선 개발 등 인간의 활동으로 인해 멸종위기에 처해있다 (Lee, 2012).

2012년 멸종위기 야생생물 II급으로 지정된 생물임에도 해당 종에 대한 인지도는 다소 낮은 편이다. 2021년 9월 기준, NCBI (National Center for Biotechnology Information)



**Fig. 2.** Results of phylogenetic analysis of evolutionary relationships using arginine kinase. The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein J, 1985). The evolutionary distances were computed using the Poisson correction method (Zuckerkanl E and Pauling L, 1965) and are in the units of the number of amino acid substitutions per site. This analysis involved 38 amino acid sequences. All ambiguous positions were removed for each sequence pair (pairwise deletion option). There were a total of 383 positions in the final dataset. Evolutionary analyses were conducted in MEGA X (Kumar *et al.*, 2018).

에 등록된 기수갈고둥 (*C. retropictus*) 의 nucleotide 수는 80건, gene의 수는 74건, protein 52건, SRA Experiments 는 1건으로 해당 종 관련 정보는 매우 부족한 것을 확인할 수 있다. 뿐만 아니라 기수갈고둥 (*C. retropictus*) 의 genome은 1건으로 전체 서열이 아닌 mitochondria DNA 서열만 분석된 것을 확인할 수 있었다.

본 연구에서는 기수갈고둥의 NGS 전사체 데이터를 활용하

여 AK 유전자를 동정하고 NCBI에 등록되어있는 두타산입술 대고둥아재비 (*Mirus junensis*), 대추귀고둥 (*Ellobium chinense*) 및 거제외줄달팽이 (*Satsuma myomphala*) 의 AK 유전자 서열을 포함한 PANM 데이터베이스를 활용하여 분자계통분류학적 위치를 확인하고자 수행되었다 (Min *et al.*, 2018; Kang *et al.*, 2019; Park *et al.*, 2020; Sang *et al.*, 2020).

**Table 1.** List of species and accession number of gene, used in phylogenetic analysis

Class	Species	Accession number
Gastropoda	<i>Nesiohelix samarangae</i>	AHC02701.1
	<i>Biomphalaria glabrata</i>	ADH59421.1
	<i>Aplysia kurodai</i>	BAB41095.1
	<i>Semisulcospira libertina</i>	AGN95434.1
	<i>Pomacea canaliculata</i>	AYH91743.1
	<i>Cellana grata</i>	BAB41096.1
	<i>Conus ebraeus</i>	ASF90538.1
	<i>Conus litteratus</i>	ARS01451.1
	<i>Haliotis diversicolor supertexta</i>	AJW60778.1
	<i>Conus frigidus</i>	ARU12142.1
	<i>Conus araneosus</i>	AQM52449.1
	<i>Conus lividus</i>	ATG85037.1
	<i>Aplysia californica</i>	XP_5099408.1
	<i>Mirus junensis</i>	SCH 2018
<i>Clithon retropictus</i>	SCH	
<i>Ellobium chinense</i>	SCH	
<i>Cellana grata</i>	BAB41096.1	
Bivalvia	<i>Hyriopsis schlegelii</i>	AEO94538.1
	<i>Mizuhopecten yessoensis</i>	OWF45839.1
	<i>Scapharca broughtonii</i>	BAD11949.1
	<i>Azumapecten farreri</i>	AEX08673.1
	<i>Heliothis virescens</i>	ADE27964.1
	<i>Neohelice granulata</i>	AAF43438.1
	<i>Klebsiella pneumoniae</i>	WP_151490152.1
	<i>Penaeus japonicus</i>	APU53308.1
	<i>Chilo suppressalis</i>	QEQ50054.1
	<i>Helicoverpa armigera</i>	ABU98622.1
	<i>Azumapecten farreri</i>	AEX08673.1
	<i>Crassostrea gigas</i>	EKC24881.1
	<i>Portunus trituberculatus</i>	ADO22718.1
	<i>Scylla olivacea</i>	ACP43443.1
<i>Scylla paramamosain</i>	AEY84969.1	
Cephalopoda	<i>Sepioteuthis lessoniana</i>	BAA95610.1
	<i>Sepiella maindroni</i>	AEK26855.1
	<i>Sepia pharaonis</i>	AKS26488.1
	<i>Nautilus pompilius</i>	BAA95594.1
	<i>Octopus vulgaris</i>	BAA95609.1
	<i>Amphioctopus fangsiao</i>	AEK65120.1

## MATERIALS AND METHODS

### 1. Materials

이번 연구에 사용된 종은 기수갈고둥 (*C. retropictus*) 으로, 연체동물문 복족강 고리갈고둥목 갈고둥과에 속한다. 대상종은 유속이 일정하게 유지되고 있는 기수지역 자갈지대에 서식하

고 있으며, 영산강유역환경청의 허가를 받아 2014년 경북 울진군에서 채집했다.

### 2. Methods

#### 1) cDNA library 구축과 염기서열 분석

샘플채집 후 액체질소에서 급속냉동 후 -80℃ 초저온냉동

**Table 2.** Branched into Class, Subclass, and Family within Gastropoda

Species	Class	Subclass	Family
<i>Cellana grata</i>	Gastropoda	Patellogastropoda	Nacellidae
<i>Clithon retropictus</i>	Gastropoda	Neritimorpha	Neritidae
<i>Biomphalaria glabrata</i>	Gastropoda	Heterobranchia	Planorbidae
<i>Aplysia californica</i>	Gastropoda	Heterobranchia	Aplysiidae
<i>Aplysia kurodai</i>	Gastropoda	Heterobranchia	Aplysiidae
<i>Satsuma myomphala</i>	Gastropoda	Heterobranchia	Camaenidae
<i>Nesiohelix samarangae</i>	Gastropoda	Heterobranchia	Bradybaenidae
<i>Mirus junensis</i>	Gastropoda	Heterobranchia	Enidae
<i>Semisulcospira libertina</i>	Gastropoda	Caenogastropoda	Semisulcospiridae
<i>Conus ebraeus</i>	Gastropoda	Caenogastropoda	Conidae
<i>Conus litteratus</i>	Gastropoda	Caenogastropoda	Conidae
<i>Conus lividus</i>	Gastropoda	Caenogastropoda	Conidae
<i>Conus araneosus</i>	Gastropoda	Caenogastropoda	Conidae
<i>Conus frigidus</i>	Gastropoda	Caenogastropoda	Conidae
<i>Haliotis diversicolor supertexta</i>	Gastropoda	Vetigastropoda	Haliotidae
<i>Pomacea canaliculata</i>	Gastropoda	Caenogastropoda	Ampullariidae
<i>Ellobium chinense</i>	Gastropoda	Heterobranchia	Ellobiidae

고 (Deep-freezer)에서 보관한 기수갈고둥 (*C. retropictus*)의 내장낭을 추출한 뒤 Trizol (Thermo Fisher Scientific, USA) 기반 RNA extraction 방법으로 total RNA의 density와 quality를 확인하였다. 추출된 total RNA에서 mRNA를 정제한 후, cDNA library 구축을 위해 TruSeq Standard mRNA Library Prep KIT (Illumina, USA)를 사용하여 cDNA를 합성하였다. cDNA library에서 fragment를 추출하여 Illumina HiSeq 2500 Sequencer를 이용하여 전사체 분석을 수행하였다.

## 2) 기수갈고둥 (*C. retropictus*)의 Arginine Kinase 서열 확보

Illumina HiSeq 2500 platform에서 생산된 raw reads에서 adapter sequence와 low quality sequence를 제거하기 위해 Cutadapt (v1.11) (Martin, 2011)와 Sickle (v1.33) (Joshi and Fass, 2011)을 사용했다. 그 결과로서 확보된 clean reads를 Trinity 프로그램을 사용하여 *De novo* assembly (Grabherr *et al.*, 2011)를 진행하여 contigs로부터 ORF 서열을 확보했다. 그 후 clustering을 하기 위해 TGICL (TIGR Gene Indices Clustering tools. Ver 2.1) (Pertea *et al.*, 2003)를 이용하여 전사체 분석에 사용할 수 있는 unigenes를 확보했다. 확보된 unigenes 서열을 PANM DB (ver. 3.0) (Kang *et al.*, 2019)와 BLAST 프로그램을 이용하여 annotation 한 결과, AK로 annotation 된 서열 중

full-length 서열을 확보하였다. 확보된 서열을 NCBI의 non-redundant 데이터베이스에서 BLAST를 진행하여 arginine kinase 서열을 검증하였으며, EMBOSS package의 sixpack 프로그램을 이용하여 arginine kinase 아미노산 서열을 확보했다 (Rice *et al.*, 2000; Madeira *et al.*, 2019).

## 3) 연체동물서열 확보 및 Phylogenetic 분석

NCBI (<https://www.ncbi.nlm.nih.gov/>)에서 AK로 검색한 결과 중 Gastropoda, Bivalvia, Cephalopoda에 속하는 각각의 full-length 아미노산 서열만을 선정하였다. 또한 선행하여 수행된 *M. junensis*, *E. chinense*, *S. myomphala* 및 *C. retropictus*의 AK 유전자 서열을 포함시킨 후 MEGA X 프로그램의 Neighbor-joining 알고리즘 (1,000 bootstrap replication)으로 phylogenetic 분석을 수행하여 dendrogram을 확보하였다 (Kumar *et al.*, 2018; Min *et al.*, 2018; Park *et al.*, 2020; Sang *et al.*, 2020).

## 4) Arginine kinase의 단백질 2차 구조 예측

확보된 *C. retropictus*와 37종의 AK 유전자 단백질 2D 구조를 예측, 비교하기 위하여 EMBOSS package의 Garnier 프로그램 (Garnier *et al.*, 1978)을 이용하여  $\alpha$ -helix,  $\beta$ -sheet, turns, coli 수를 확인하였다.

**Table 3.** Prediction of AK 2D structures using Garnier program

Species	Class	Helix	Sheet	Turns	Coil
<i>Mizuhopecten yessoensis</i>		167	51	86	47
<i>Hyriopsis schlegelii</i>		159	69	88	34
<i>Crassostrea gigas</i>		161	47	92	51
<i>Azumapecten farreri</i>		193	54	72	47
<i>Scapharca broughtonii</i>		146	67	83	52
<i>Heliothis virescens</i>		172	58	69	56
<i>Chilo suppressalis</i>		176	57	68	54
<i>Spodoptera frugiperda</i>		173	58	69	55
<i>Helicoverpa armigera</i>	Bivalvia	173	57	69	56
<i>Klebsiella pneumoniae</i>		169	64	66	58
<i>Penaeus japonicus</i>		168	67	64	57
<i>Neohelice granulata</i>		169	68	66	54
<i>Portunus trituberculatus</i>		162	72	73	50
<i>Scylla olivacea</i>		161	76	70	50
<i>Scylla paramamosain</i>		161	76	70	50
<b>Average</b>		167	63	74	51
<i>Cellana grata</i>		155	65	64	63
<i>Clithon retropictus</i>		163	55	80	52
<i>Mirus junensis</i>		156	62	82	58
<i>Ellobium chinense</i>		153	71	84	55
<i>Biomphalaria glabrata</i>		141	68	94	49
<i>Aplysia californica</i>		159	60	79	57
<i>Aplysia kurodai</i>		157	70	75	52
<i>Satsuma myomphala</i>		142	74	87	51
<i>Nesiohelix samarangae</i>		131	75	87	62
<i>Semisulcospira libertina</i>	Gastropoda	168	45	80	58
<i>Pomacea canaliculata</i>		126	59	95	52
<i>Conus ebraeus</i>		169	63	69	47
<i>Conus araneosus</i>		164	57	83	54
<i>Conus litteratus</i>		165	60	76	54
<i>Conus frigidus</i>		166	63	77	49
<i>Conus lividus</i>		155	62	85	52
<i>Haliotis diversicolor supertexta</i>		191	48	73	42
<b>Average</b>		157	62	81	53
<i>Sepiella maindroni</i>		164	54	80	51
<i>Sepioteuthis lessoniana</i>		148	56	92	52
<i>Sepia pharaonis</i>		155	59	80	55
<i>Amphioctopus fangsiao</i>	Cephalopoda	161	56	83	48
<i>Octopus vulgaris</i>		147	72	87	41
<i>Nautilus pompilius</i>		157	64	80	49
<b>Average</b>		164	62	76	51

**RESULTS AND DISCUSSION**

기수갈고둥 (*C. retropictus*) 의 AK 유전자 coding region 은 총 1053 bp 로 351개의 아미노산으로 확인되었으며, GC 함유량은 62.2%로 확인되었다. 해당 서열의 염기구성은 234 개의 adenine, 164개의 thymine, 313개의 guanine, 342개의 cytosine으로 이루어져 있었다 (Fig. 1).

연체동물문 내에서 기수갈고둥 (*C. retropictus*) 의 분자계

통학적 위치 확인을 위해 marker로서 AK gene을 사용하여 molecular phylogenetic analysis를 진행했다. NCBI에 등록된 Gastropoda, Cephalopoda, Bivalvia의 full-length AK 유전자 서열 및 본 연구진에 의해 선행 연구된 *M. junensis*와 *S. myomphala*, *M. junensis* 그리고 대상종인 *C. retropictus*의 AK 서열을 추가하여 총 38종의 AK 유전자 서열로 구성된 서열을 모아 분석을 수행하였다. 38종의 AK 유전자 서열은 MEGA X 프로그램을 이용하여

Neighbor-joining 알고리즘으로 (Saitou and Nei, 1987) 분석한 결과 기존에 알려진 Class에 포함되어 Clade가 형성되는 것을 확인하였다. 기수갈고둥 (*C. retropictus*) 이 포함된 Gastropoda에서 Ampullariidae, Semisulcospiridae, Nacellidae, Neritidae, Planorbidae, Aplysiidae, Camaenidae, Bradybaenidae, Conidae, Enidae, Ellobiidae 및 Haliotidae로 분지되어 12개의 Family가 확인되었는데, 본 연구의 대상종인 *C. retropictus*와 같은 Family에 속하는 생물의 AK 유전자 서열이 아직 밝혀지지 않아 분석에 포함할 수 없었다. (Fig. 2), (Table 1), (Table 2).

38 종의 AK 단백질  $\alpha$ -helix,  $\beta$ -sheet, turns, coil 수를 Garnier 프로그램을 이용하여 예측하였다. Bivalvia에 속하는 서열들은 평균 167개의  $\alpha$ -helix, 63개의  $\beta$ -sheet, 74개의 turns, 51개의 coil로 예측되었으며, Cephalopoda는 각각 평균 164개, 62개, 76개, 51개, Gastropoda의 경우는 각각 157개, 62개, 81개, 53개로 예측되었다. 각 Class들의 평균을 비교해보았을 때, 큰 차이가 없는 것으로 확인되었으며, 또한 각 Class 내에서도 어떠한 상관관계를 찾기 어려웠다 (Table 3).

AK 유전자를 이용하여 Mollusca 내에서 Class, Subclass 또는 Order 수준에서 계통분류를 확인하는 연구가 선행되어져 왔다. 본 연구의 결과는 Family를 확인함으로써 다시 한번 더 AK 유전자가 계통분류에 marker로서 사용될 수 있고 Family 수준까지의 분류가 가능하다는 것을 확인하였다 (Sang *et al.*, 2020).

## CONCLUSIONS

*C. retropictus*의 AK 유전자는 1053bp의 염기서열로 이루어져 있으며, GC 함유량은 62.2%였다. 생물정보학적 분석을 통해 선정된 종에 AK 유전자 서열을 선행되어진 연구를 통해 발표된 *E. chinense*와 *M. junensis*, *S. myomphala*에 대상종인 *C. retropictus*의 서열을 포함한 총 38종의 연체동물을 대상으로 하여 phylogenetic analysis를 진행하였다. 그 결과 Gastropoda, Cephalopoda, Bivalvia의 3개 Class로 나뉘어졌으며, *C. retropictus*가 포함된 Gastropoda 내에선 Family 단위로 분류군이 묶이는 것을 확인할 수 있다. Garnier 프로그램을 이용하여 38종의 AK 단백질의 2D 구조를 예측하여  $\alpha$ -helix,  $\beta$ -sheet, turns, coil을 이루는 아미노산 숫자의 평균을 확인한 결과, 각 Class들의 평균은 큰 차이가 없는 것으로 확인되며, 각각의 종들에서도 상호관계를 찾기 어려웠다. 이러한 결과들을 통해, AK 유전자는 Mollusca의 분자계통학적 분석에서 유용한 지표가 될 수 있다는 것을 재확인할 수 있었으며, AK 단백질 2차 구조에 대한 추가적인 연구

가 더 필요할 것으로 판단된다.

## ACKNOWLEDGEMENT

본 논문은 교육부에서 지원하는 지역대학 우수과학자 (한국연구재단, NRF-2017R1D1A3B06034971) 및 중점연구소 (NRF-2021R1A6A1A03039503), 2016학년도 순천향대학교 교수 연구년제에 의하여 수행되었습니다.

## REFERENCE

- Downs, M., Johnson, P., and Zeece, M. (2016) Insects and their connection to food allergy Insects as sustainable food ingredients. Elsevier, pp 255-272.
- 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.
- Ji Eun Jeong, and Lee, Y.S. (2013) Identification, sequence characterization and expression analysis of the arginine kinase gene in response to laminarin challenge from the Oriental land snail, *Nesiohelix samarangae*. *Department of Life Science, Soochunhyang University*, **29**: 171-179.
- Joshi, N., and Fass, J. (2011) Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software] Available at <https://github.com/najoshi/sickle>.
- Kang, S.W., Park, S., Hwang, H., Chung, J., Sang, M., Min, H., Park, J., Cho, H., Patnaik, B., and Lee, Y. (2019) PANM DB ver 3.0: an update of the bioinformatics database for annotation of large datasets from sequencing of species under Protostomia clade. *The Korean Journal of Malacology*, **35**: 73-75.
- 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.
- Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R., Potter, S.C., and Finn, R.D. (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic acids research*, **47**: 636-641.
- Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal*. **17**: 10-12.
- McGuffin, L.J., Bryson, K., and Jones, D.T. (2000) The PSIPRED protein structure prediction server. *Bioinformatics*, **16**: 404-405.
- Min, R., Hee Ju Hwang, Jong Min Chung, Min Kyu

- Sang, Hang Chul Cho, Jie Eun Park, Ki Yoon Jung, Hong Seog Park<sup>1</sup>, Han<sup>2</sup>, Y.S., and Lee, a.Y.S. (2018) Molecular Phylogenetic studies of *Mirus junensis* using Arginine kinase gene sequence. *The Korean Journal of Malacology*, **34**: 107-114.
- Park, J., Min Kyu Sang, Hee-Ju Hwang, Dae Kwon Song, Jun Yang Jeong, So Young Park, Se Won Kang, Hong Seog Park, Yeon Soo Han, Lee, J.S., and Lee, a.Y.S. (2020) Molecular Phylogenetic Analysis of *Ellobium chinense* Using Arginine Kinase. *The Korean Journal of Malacology*, **36**(3): 133-141.
- Pertea, G., Huang, X., Liang, F., Antonescu, V., Sultana, R., Karamycheva, S., Lee, Y., White, J., Cheung, F., Parvizi, B., Tsai, J., and Quackenbush, J. (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.
- Saitou, N., and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**: 406-425.
- Sang, M., Hwang, H., Chung, J., Park, J., Song, D., Jeong, J., Park, S., Park, H., Jo, Y., and Lee, J. (2020) Phylogenetic analysis of endangered wild animal class II *Satsuma myomphala* using the arginine kinase gene. *The Korean Journal of Malacology*, **36**: 97-104.
- Tanaka, K., Ichinari, S., Iwanami, K., Yoshimatsu, S., and Suzuki, T. (2007) Arginine kinase from the beetle *Cissites cephalotes* (Olivier). Molecular cloning, phylogenetic analysis and enzymatic properties. *Insect Biochemistry and Molecular Biology*, **37**: 338-345.
- Uda, K., Fujimoto, N., Akiyama, Y., Mizuta, K., Tanaka, K., Ellington, W.R., and Suzuki, T. (2006) Evolution of the arginine kinase gene family. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, **1**: 209-218.
- Voncken, F., Gao, F., Wadforth, C., Harley, M., and Colasante, C. (2013) The phosphoarginine energy-buffering system of *trypanosoma brucei* involves multiple arginine kinase isoforms with different subcellular locations. *PLoS One*, **8**: e65908.