

Arginine Kinase를 이용한 대추귀고둥 (*Ellobium chinense*) 의 분자계통학적 분석

박지은¹, 상민규¹, 황희주², 송대권¹, 정준양¹, 박소영³, 강세원⁴, 박홍석⁵, 한연수⁶, 이준상², 이용석¹

¹순천향대학교 자연과학대학 생명시스템학과, ²순천향대학교 기초과학연구소, ³국립낙동강생물자원관 다양성연구팀, ⁴한국생명공학연구원 생물자원센터, ⁵주지앤시바이오, ⁶전남대학교 농업생명과학대학 식물생명공학부

Molecular Phylogenetic Analysis of *Ellobium chinense* Using Arginine Kinase

Jie Eun Park¹, Min Kyu Sang¹, Hee-Ju Hwang², Dae Kwon Song¹, Jun Yang Jeong¹, So Young Park³, Se Won Kang⁴, Hong seog Park⁵, Yeon soo Han⁶, Jun Sang Lee² and Yong Seok Lee¹

¹Department of Life Science and Biotechnology, College of Natural Sciences, Soonchunhyang University, Asan, Chungnam, 31538, Korea,

²Institute for Basic Sciences, College of Natural Sciences, Soonchunhyang University, Asan, Chungnam, 31538, Korea,

³Biodiversity Research team, Nakdonggang National Institute of Biological Resources, Sangju, Gyeongbuk, 37242, Korea,

⁴Biological Resource Center; Korea Research Institute of Bioscience and Biotechnology, Jeongeup, Jeonbuk 56212, Korea,

⁵Research Institute, GnC BIO Co., LTD., 621-6 Banseok-dong, Yuseong-gu, Daejeon, 34069, Korea

⁶Department of Applied Biology, Institute of Environmentally-Friendly Agriculture (IEFA), College of Agriculture and Life Sciences, Chonnam National University, Gwangju 61186, Korea

ABSTRACT

In Korea, *Ellobium chinense* is one of the well-known small, air-breathing snails, and was registered as the endangered species in 2005. The number of *E. chinense* populations has rapidly declined in recent years. In this context, genetic information of *E. chinense* is needed for species conservation in the future. However, the registered genetic information of *E. chinense* in NCBI is only 55 nucleotides and 53 proteins. Thus, we sequenced an *E. chinense* cDNA library using the Illumina platform, and selected arginine kinase (AK) gene which has been used as a molecular phylogenetic marker. AK sequence of *E. chinense* was analyzed through bioinformatic programs, and the biological importance of *E. chinense* was discussed in conjunction with molecular phylogenetic trees.

Key Words : Arginine kinase (AK), *Ellobium chinense*, Illumina Hiseq 2500, phylogenetic analysis

INTRODUCTION

Arginine kinase (AK) 는 Phosphagen (guanidino)

Kinases 그룹에 속한다 (Suzuki *et al.*, 1997). Phosphagen Kinases는 자연적으로 발생하는 guanidine 화합물로 고에너지 인산염 화합물 group의 가역적 전달을 촉진하는 역할을 하는 것으로 알려져 있으며, Hypotaurocyamine Kinase (HTK), Glycocyamine Kinase (GK), Thalessemine Kinase (ThK), Opheline Kinase (OP), Lumbricine Kinase (LK), Taurocyamine Kinase (TK), Arginine kinase (AK) 등을 포함하고 있다 (Uda *et al.*, 2006; Mahon and Neigel, 2008; Jarilla *et al.*, 2014). 이들 효소는 ATP의 환원율이 높은 것으로 알려져 있으며, 다양한 생물군에서 에너지 대사에 주요한 역할을 하는 것으로 밝혀졌다 (Wyss *et al.*, 1992; Ellington, 2001; McLeish and Kenyon, 2005;

Received: September 21, 2020; Revised: September 25, 2020;
Accepted: September 30, 2020

Corresponding author: Yong Seok Lee

Tel: +82 (41) 530-3040, e-mail: yslee@sch.ac.kr
1225-3480/24769

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.

Ellington and Suzuki, 2006).

이들 효소 group은 크게 AK와 HTK를 포함하는 AK group과 CK, GK, LK 및 TK가 포함되는 CK group으로 분리된다 (Uda *et al.*, 2008). 이 중 AK가 무척추동물 및 일부 척추동물에서 존재하는 것과는 달리 CK는 척추동물에서만 발견되는 것으로 알려져 있다 (Wyss *et al.*, 1992; Ellington, 2001; McLeish *et al.*, 2005; Ellington *et al.*, 2006). PK group에 속한 효소 중 AK는 다른 PK group 효소와 달리 변형되지 않은 아미노산을 사용하며, 전형적인 기능성 단량체로 알려져 있다 (Uda *et al.*, 2005; Jarilla *et al.*, 2014). 또한 원생동물, 선형동물, 절지동물, 연체동물, 중생동물 등 다양한 연체동물 분류군의 근육에서 분리, 확인된 바 있다 (Suzuki *et al.*, 2000). 연체동물 전반에서 확인되고 동시에 PK group에 속하는 gene으로만 lineage를 작성한 선행논문이 있는 등 (Suzuki and Furukohri, 1994; Uda *et al.*, 2005; Uda and Suzuki, 2007) AK가 연체동물의 분자계통학적 위치를 찾기 용이한 marker인 것으로 판단하여 본 연구과제를 진행하였다.

연구대상으로 선정된 *Ellobium chinense*는 Ellobiidae과의 육상 폐산 복족류 연체 동물인 작은 공기 호흡 달팽이 종이다 (Min, 2004). 이 종은 대한민국 환경부에서 멸종 위기에 처한 것으로 판단하여 2005년 멸종위기 야생생물 II급으로 지정 후 지역 법으로 보호받고 있다 (<https://species.nibr.go.kr/Endangeredspecies/home/main.jsp>) (Lee and Son, 2012). *E. chinense*는 한반도의 서 남해안의 염습지 혹은 해안선 근처 민물 갯벌의 돌과 *Zoysia sinica* (갯잔디) 와 같은 염생식물 등의 서식지에서 발견되는 것으로 알려져 있다 (Min, 2004; Li *et al.*, 2020). 2011년 멸종 위기 종 목록집에 따르면 *E. chinense*의 점유 면적 (AOO), 발생 범위 (EOO) 및 서식지 품질 감소로 개체수가 급격하게 감소한 것으로 보고되었다 (Köhler and Rintelen, 2011; 서재화 *et al.*, 2018). 지속적인 국토의 해안 개발, 갯벌 매립 및 염습지의 파괴로 인해 *E. chinense*의 종 보존을 위한 적극적인 노력이 필요하다. 하지만 2020년 09월 기준으로 NCBI를 확인한 결과 해당종에 대한 유전정보는 본 연구진들이 선행연구를 통해 update한 유전정보 (Kang, 2016; Kang *et al.*, 2018) 외 추가적인 데이터 update는 없는 것으로 확인되었다.

이번 연구에서는 *E. chinense*에서 확인된 AK 유전자를 이용하여 해당 종의 계통분류학적 위치를 확인하고, 이를 통해 향후 연체동물문의 세밀한 분자계통학적 연구에 AK 유전자가 활용될 가능성이 있는지에 대하여 알아보고자 한다.

MATERIALS AND METHODS

1. Materials

멸종위기 야생생물 II급으로 지정된 *E. chinense*의 채집을 위해 영산강유역 환경청의 허가를 받아 (허가번호 2014-11) 진행하였다. 2014년도 7월 전라남도 장흥군 안양면 수문리 일대에서 채집을 진행하였으며, 4개의 개체를 획득하였다. 이를 해부하여 내장낭을 적출한 후 분석에 이용하였다.

2. Methods

1) cDNA library 구축 및 염기서열 분석

*E. chinense*에서 적출된 내장낭 조직을 액체질소를 통해 동결하여 sample의 균질화를 진행하였다. 균질화 된 시료를 이용하여 total RNA 추출을 진행하였으며, 이 과정에서 50-100 mg의 시료 당 1 ml의 Trizol reagent (Thermo Fisher Scientific, USA) 를 사용하였다. 추출된 RNA를 Spectrophotometer 및 Agilent Technologies 2100 Bioanalyzer를 사용하여 quality 및 농도를 확인한 뒤, Illumina TruSeq RNA Sample Preparation Kit (Illumina, San Diego, CA, USA) 를 사용하여 mRNA로 정제 후 cDNA를 합성하였다. 합성된 cDNA를 통해 cDNA library를 구축하였으며, Illumina HiSeq 2500 (NGS) 플랫폼을 이용하여 염기서열 분석을 수행하였다.

2) *E. chinense*의 AK sequence 확보를 위한 분석

Illumina HiSeq 2500을 통해 생산된 raw reads를 가공하고 clean reads를 얻기 위해 cutadapt (v1.11) (Martin, 2011), Sickle (v1.33) (Joshi and Fass, 2011) 및 Fastq_filter를 사용하였다. 이를 통해 raw reads에 있던 adapt 서열 및 low quality sequence가 제거된 clean reads를 확보하였다. 확보된 clean reads를 Trinity (v2.0.4) 를 이용하여 *de novo* assembly를 진행하였다 (Grabherr *et al.*, 2011; Kang *et al.*, 2017). *de novo* assembly를 통해 확보된 contigs로 TransDecoder (Haas and Papanicolaou, 2016) 를 이용하여 Open Reading Frame (ORF) sequence를 확보하였다. ORF 서열은 TGICL (Pertea *et al.*, 2003) 프로그램을 이용하여 clustering 한 후 분석에 사용될 unigenes을 확보하였다. 분자계통분석에 사용하고자 하는 AK gene을 포함한 unigene 추출을 위하여 PANM DB (v3.0) (Kang *et al.*, 2019) 와 BLAST program을 활용하였다. 그 결과 도출된 unigene은 NCBI의 nr (non-redundant) 데이터베이스에 BLAST하여 서열을 검증하였다. 검증된 서열은 EMBOSS package sixpack (https://www.ebi.ac.uk/Tools/st/emboss_sixpack/) (Rice *et al.*, 2000; Madeira *et al.*, 2019) 을 이용하여 AK 아미노산 서열을 확보하였다.

272	CCT	AAG	TTC	TTC	CTC	AAA	CTC	AAA	ATG	GCC	GCA	GAT	GCA	GGA	AAG	316
91	P	K	F	F	L	K	L	K	M	A	A	D	A	G	K	105
317	TTG	TTC	GAT	CTT	CTG	AAG	GCT	GAC	AAA	AAC	TGC	AAG	TCT	CTC	CTT	361
106	L	F	D	L	L	K	A	D	K	N	C	K	S	L	L	120
362	GCC	AAA	AAT	CTT	ACG	AAG	GAG	CGC	TTC	GAT	GCC	CTG	AAG	GAC	AAG	406
121	A	K	N	L	T	K	E	R	F	D	A	L	K	D	K	135
407	AAA	ACC	AAA	TTT	GGT	GGC	GGC	CTC	GCC	GAC	TGC	ATC	CGC	TCA	GGA	451
136	K	T	K	F	G	G	G	L	A	D	C	I	R	S	G	150
452	TGC	GAA	AAC	ACA	GAC	AGC	GGT	GTG	GGC	GTG	TAC	GCC	TCC	GAC	CCC	496
151	C	E	N	T	D	S	G	V	G	V	Y	A	S	D	P	165
497	GAT	GCA	TAC	ACC	GTC	TTC	GCT	CCC	CTG	CTG	GAC	GCT	A	V	ATC	541
166	D	A	Y	T	V	F	A	P	L	L	D	A	W	I	K	180
542	GAC	TAT	CAT	AAG	GTC	ACC	GAG	CTG	AAC	CAC	CCC	AAC	CCT	GAC	TTT	586
181	D	Y	H	K	V	T	E	L	N	H	P	N	P	D	F	195
587	GGC	GAC	ATC	GAC	AAG	CTC	GAC	TTT	GGT	GAC	CTT	GAT	CCC	TCT	GGC	631
196	G	D	I	D	K	L	D	F	G	D	L	D	P	S	G	210
632	AAC	ATG	ATC	GTG	TCC	ACC	CGT	GTG	CGC	GTG	GGC	CGC	AGC	CAC	GAC	676
211	N	M	I	V	S	T	R	V	R	V	G	R	S	H	D	225
677	AGC	TAT	GGC	TTC	CCC	CCG	GTG	CTG	AAG	TGT	GAG	GAG	CGT	GTT	GAG	721
226	S	Y	G	F	P	P	V	L	K	C	E	E	R	V	E	240
722	ATG	GAG	AAG	AAG	ACC	TG	GAT	GCC	CTG	AAG	AAG	CTG	GAG	GGT	GAA	766
241	M	E	K	K	T	V	D	A	L	K	K	L	E	G	E	255
767	CTG	AAG	GGA	ACA	TAC	CAC	CCA	CTG	ACT	GGC	ATG	AGC	AAG	GAG	ACA	811
256	L	K	G	T	Y	H	P	L	T	G	M	S	K	E	T	270
812	CAG	AAA	CAG	CTG	ACT	GAG	GAC	CAC	TTC	CTC	TTC	AAT	GAC	AGT	GAC	856
271	Q	K	Q	L	T	E	D	H	F	L	F	N	D	S	D	285
857	AGG	TTC	TTG	AGG	GCT	GCT	GGT	GGT	TAC	AAC	GAC	TGG	CCC	ACC	GGC	901
286	R	F	L	R	A	A	G	G	Y	N	D	Trp	P	T	G	300
902	CGT	GGA	ATC	TAC	TTC	AAC	GAC	AAC	AAA	ACC	TTC	CTT	GTC	TGG	ATC	946
301	R	G	I	Y	F	N	D	N	K	T	F	L	V	Trp	I	315
947	AAC	GAG	GAG	GAT	CAT	CTC	CGC	TTC	ATC	TCT	ATG	CAG	AAG	GGT	GGC	991
316	N	E	E	D	H	L	R	F	I	S	M	Q	K	G	G	330
992	AAC	CTC	AAC	GAG	GTT	TAC	TCC	AGA	CTT	GTT	AAG	GCC	ATC	AGA	GCC	1036
331	N	L	N	E	V	Y	S	R	L	V	K	A	I	R	A	345
1037	CTT	GAA	GGC	AGT	GGT	CTC	TCC	TTC	GCC	AAG	CGT	GAT	GGT	TTG	GGC	1081
346	L	E	G	S	T	L	S	F	A	K	R	D	G	G	G	360
1082	TAC	CTG	ACC	TTC	TGC	CCC	TCA	AAC	TTG	GGC	ACC	ACC	CTT	CGT	GCC	1126
361	Y	L	T	F	C	P	S	N	L	G	T	T	L	R	A	375
1127	TCT	GTG	CAT	ATC	AAG	ATC	CCC	AAG	CTG	TCT	GCC	AGC	CCC	GAG	TTC	1171
376	S	V	H	I	K	I	P	K	L	S	A	S	P	E	F	390
1172	AAG	GCA	TTC	TGT	GAC	AAG	TAC	AAT	ATC	CAG	GCC	AGA	GGC	ATT	CAT	1216
391	K	A	C	C	D	K	Y	N	I	Q	A	R	G	I	H	405
1217	GGA	GAG	CAC	ACA	GAA	TCT	GTA	GGA	GGC	GTG	TAT	GAT	ATC	TCC	AAC	1261
406	G	E	H	T	E	S	V	G	G	V	Y	D	I	S	N	420
1262	AAG	AGG	AGA	CTG	GGT	CTC	ACA	GAA	ATT	CAG	GCC	ATC	CAA	GAA	ATG	1306
421	K	R	R	L	G	L	T	E	I	Q	A	I	Q	E	M	435
1307	CGC	AGG	GGT	GTT	GAG	GCA	GTC	ATT	GCA	GAG	GAG	AAG	AAG	TTG	GGC	1351
436	R	R	G	V	E	A	V	I	A	E	E	K	K	L	G	450
1352	GGT	GGA	TCT	TAA	ATT	AAA	AAT	AAA	AAA	ACA	GCA	ACA	ACA	AAC	GAC	1396
451	G	G	S	End	I	K	N	K	K	T	A	T	T	N	D	465

Fig. 1. The complete nucleotide sequence of Arginine kinase and deduced amino acid sequence of *E. chinense*.

3) 연체동물 서열 확보를 통한 Phylogenetic 분석

NCBI (<https://www.ncbi.nlm.nih.gov/>) 에 검색 key word를 “Gastropoda” “arginine kinase”, “Bivalvia” “arginine kinase”, “Cephalopoda” “arginine kinase”로 하여 도출 된 서열들 중 full-length 아미노산 서열을 사용해 phylogenetic 분석에 사용할 유전자 pool을 구성하였다. 이 때 연구진들이 선행하여 분석했던 *Satsuma myomphala* (거

제외줄달팽이) (Min Kyu Sang *et al.*, 2020)와 *Mirus junensis* (두타산입술대고둥아재피) (Min *et al.*, 2018) 의 AK 서열을 추가하였다. MEGA X program (Kumar *et al.*, 2018) 을 사용하여 phylogenetic 분석을 시행하였으며 (Jones *et al.*, 1992; Kumar *et al.*, 2018), Neighbor-Joining algorithm을 사용하여 dendrogram을 완성하였다 (Felsenstein, 1985; Saitou and Nei, 1987).

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

Class	Species	Accession number
Bivalvia	<i>Scapharca broughtonii</i>	BAD11949
	<i>Pholas orientalis</i>	ACP43447.1
	<i>Crassostrea gigas</i>	EKC24881
	<i>Azumapecten farreri</i>	AEX08673
	<i>Mizuhopecten yessoensis</i>	OWF45839
	<i>Hyriopsis schlegelii</i>	AEO94538
	<i>Corbicula japonica</i>	BAB91357.1
	<i>Archivesica packardana</i>	AXE71657.1
	<i>Ensis directus</i>	AAM90698.1
	<i>Paphia undulata</i>	ACP43446.1
	<i>Meretrix lyrata</i>	ACP43445.1
	<i>Meretrix meretrix</i>	ACP43444.1
	<i>Calyptogena kaikoi</i>	BAE16974.1
	<i>Solen strictus</i>	BAB91358.1
Cephalopoda	<i>Nautilus pompilius</i>	BAA95594
	<i>Amphioctopus fangsiao</i>	AEK65120
	<i>Octopus vulgaris</i>	BAA95609
	<i>Sepiella maindroni</i>	AEK26855
	<i>Sepia pharaonis</i>	AKS26488
	<i>Sepioteuthis lessoniana</i>	BAA95610
Gastropoda	<i>Aplysia kurodai</i>	BAB41095
	<i>Aplysia californica</i>	XP_005099408
	<i>Pomacea canaliculata</i>	AYH91743
	<i>Haliotis diversicolor supertexta</i>	AJW60778
	<i>Haliotis madaka</i>	BAA05100.1
	<i>Conus anemone novaehollandiae</i>	ADK73590
	<i>Conus monile</i>	ANC48006
	<i>Conus araneosus</i>	AQM52449
	<i>Conus miles</i>	AQQ10876
	<i>Conus litteratus</i>	ARS01451
	<i>Conus frigidus</i>	ARU12142
	<i>Conus ebraeus</i>	ASF90538
	<i>Conus lividus</i>	ATG85037
	<i>Nesiohelix samarangae</i>	AHC02701
	<i>Mirus junensis</i>	KJM 34_2
	<i>Satsuma myomphala</i>	KJM 36_2
	<i>Ellobium chinense</i>	-
<i>Turbo cornutus</i>	BAA22870.1	
<i>Biomphalaria glabrata</i>	ADH59421	
<i>Semisulcospira libertina</i>	AGN95434	
<i>Cellana grata</i>	BAB41096	

4) *E. chinense*의 AK protein 구조 예측

확보된 *E. chinense*의 AK gene에서 전사되는 단백질의 구조를 예측하기 위해 PSIPRED protein structure prediction server (v4.0) (<http://globin.bio.warwick.ac.uk/psipred/>) 를 이용하여 단백질의 2차 구조 예측을 진행하였다 (McGuffin *et al.*, 2000; Buchan and Jones, 2019).

RESULT AND DISCUSSION

Transcriptome 분석을 통해 얻은 *E. chinense*의 227,032 개 unigene을 PANM DB에 BLAST 하여 12개의 AK 추정 서열을 확보하였다. 확보된 서열의 정확도 확인을 위해 NCBI BLAST를 수행 후 Genscan 프로그램을 통해 orf 서열을 추출하였다. 그 결과 *E. chinense*의 AK 서열은 356개의 아미노

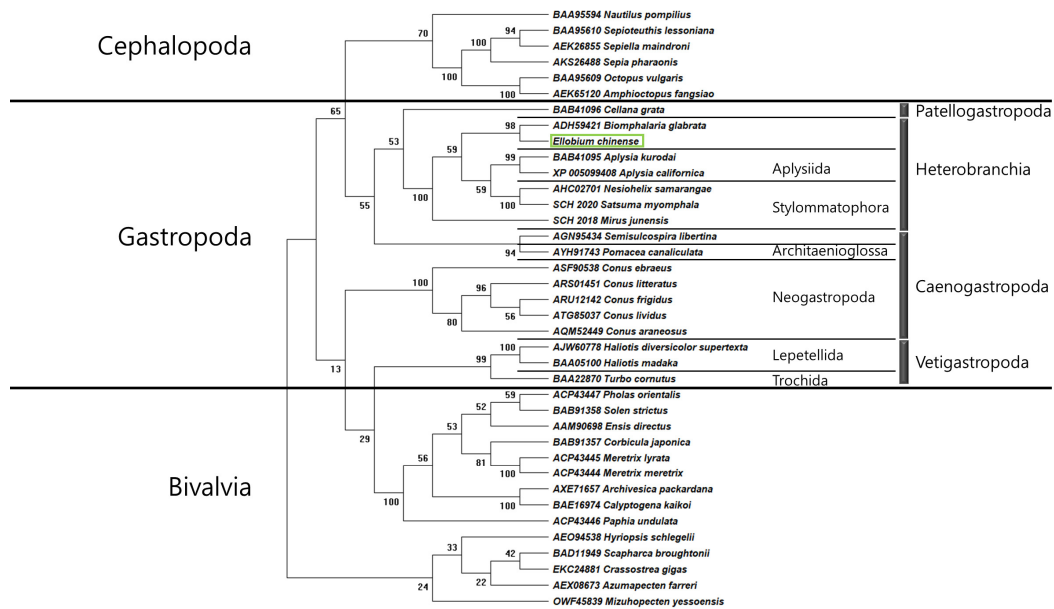


Fig. 2. Phylogenetic analysis result using Neighbor-joining algorithm. The AK gene was used as a marker and was performed in MEGA X using 41 amino acid sequences.

산, 즉 1,068개의 염기로 구성되어 있었다. 해당 서열은 269개의 adenine, 222개의 thymine, 288개의 guanine, 289개의 cytosine으로 이루어져 있었으며 전체 서열의 GC contents는 54.03%로 확인되었다 (Fig. 1).

연체동물문 내 *E. chinense*의 분자계통학적 위치를 확인하기 위한 marker로 AK gene을 사용하기 위하여 phylogenetic analysis를 진행하였다. 이 과정에서 사용하기 위한 gene pool의 구성을 위해 NCBI에 등록 되어있는 Gastropoda (복족류), Cephalopoda (두족류), Bivalvia (이매패류)의 AK 서열 중 full-length를 가지는 정보를 사용하였다. Phylogenetic analysis를 위한 gene pool 구성에 본 연구진이 선행하여 연구한 바 있는 *M. junensis*와 *S. myomphala*의 AK 서열을 추가하였다. 그 결과 총 31 속, 41 종의 AK 서열로 구성된 gene pool을 확보하였다 (Table. 1).

확보된 gene pool을 MEGA X program을 이용하여 진화적 거리가 가까운 상호 종관계를 가장 잘 유추할 수 있는 Neighbor-joining algorithm으로 분석하였다 (Saitou *et al.*, 1987). 그 결과 선행된 연구와 같이 41종의 생물이 같은 Phylum끼리 분지되어 grouping 된 것을 확인하였다 (Fig. 2). *E. chinense*가 포함 되어있는 Gastropoda의 경우 Patellogastropoda, Heterobranchia, Caenogasteropoda 및 Vetigastropoda로 분지되어 4개의 subclass로 grouping 되었으며, 각각의 subclass에서 분지된 group은 Aplysiida, Stylommatophora, Architaenioglossa, Neogastropoda, Lepetellida, Trochida로 묶을 수 있었다 (Fig. 2).

Neighbor-joining을 통해 분석된 phylogenetic tree를 확인한 결과 연구 대상종인 *E. chinense*와 유연관계가 가까운 group은 Heterobranchia로 볼 수 있었으며, 그 결과 본 연구진은 Heterobranchia에 속하는 *Biomphalaria glabrata*, *Aplysia kurodai*, *Nesiohelix samarangae*, *Satsuma myomphala*, *Mirus junensis*의 AK gene과 *E. chinense*의 AK gene에서 전사된 단백질의 2D 구조를 비교하여 보았다.

그 결과 6종의 gastropoda 중 육상생활을 하며 활발한 움직임 보이는 *S. myomphala*와 *N. samarangae*는 16개의 α -helix와 11-12개의 β -pleated sheet 구조를 가진 것으로 예측되었던 반면 수생생활을 하거나 습윤한 환경에서 생활하는 gastropoda는 13-14개의 α -helix와 11-12개의 β -pleated sheet 구조를 가진 것으로 예측되었다. *N. samarangae*, *S. myomphala*, *M. junensis*는 같은 order인 Stylommatophora에 속하지만 생육지의 특성에 따라 AK prediction 결과의 차이가 나타나는 것을 확인할 수 있었다. 이를 토대로 생각해보았을 때 습윤한 수생생활사를 가지는 종보다 건조한 육상생활사를 보내는 종 간의 AK 구조 차이점에 대한 추가적인 연구가 더 필요할 것으로 생각된다 (Table. 2 and Fig. 3).

PK group에 포함된 효소들 간의 lineage 분석과 AK gene을 이용한 계통발생학적 연구는 지속적으로 이루어져 왔다 (Suzuki *et al.*, 2000; Tanaka *et al.*, 2007; Uda *et al.*, 2010). PK group의 lineage를 계통학적으로 접근한 논문을 살펴보면 극피동물에 속하는 *Stichopus* (해삼류)와 환형동물

Table 2. Present AK 2D prediction result as number of α -helix and β -pleated sheet

Species	Number of α -helix	Number of β -pleated sheet
<i>Ellobium chinense</i>	13	11
<i>Mirus junensis</i>	14	11
<i>Aplysia kurodai</i>	14	11
<i>Biomphalaria glabrata</i>	14	12
<i>Satsuma myomphala</i>	16	11
<i>Nesiohelix samarangae</i>	16	12

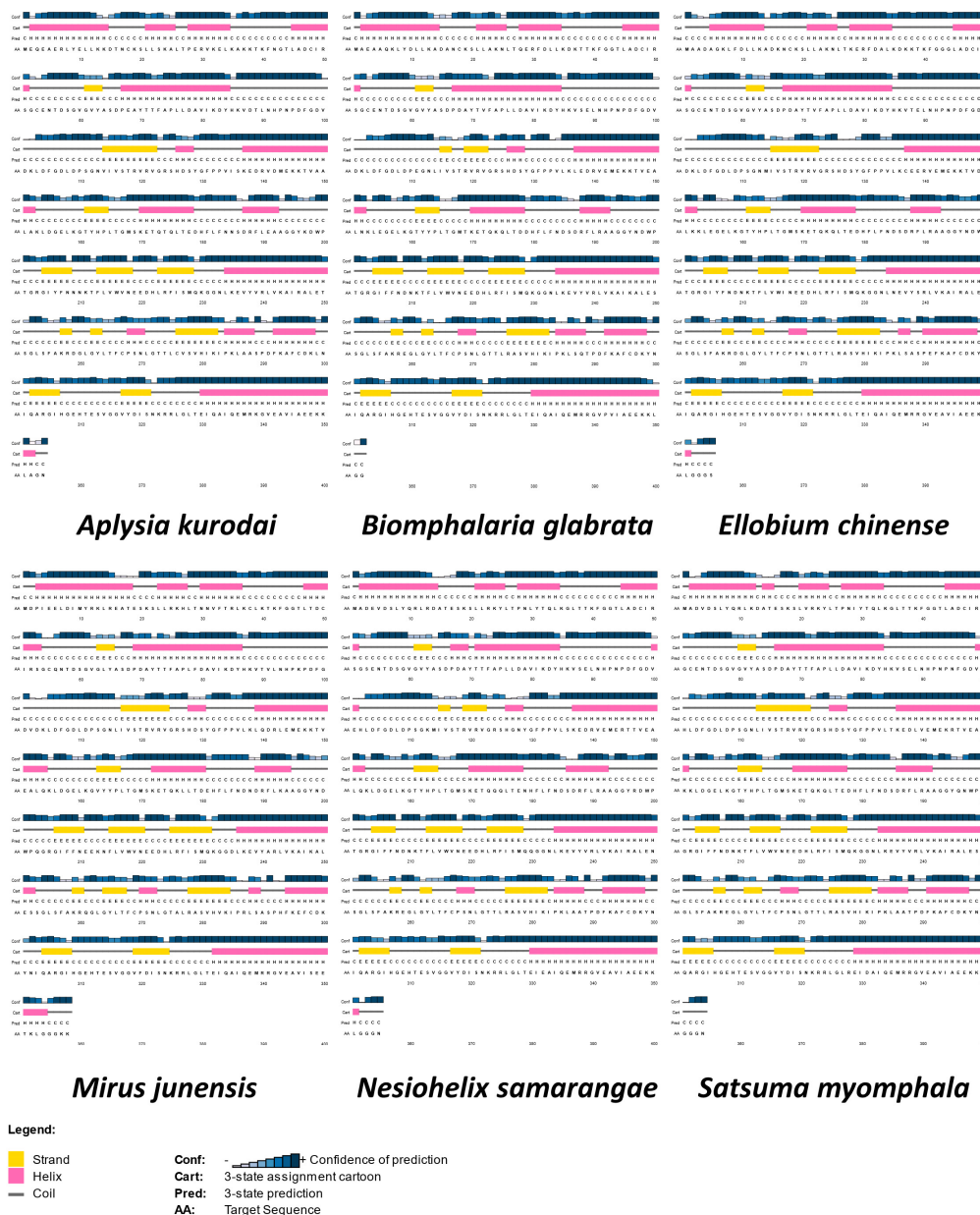


Fig. 3. The predicted 2D structure of arginine kinase by PSIPRED 4.0. Comparative analysis between 5 species and *E. chinense*. All 6 species has close Evolutionary distance.

에 속하는 Sabellastarte (갯지렁이류) 에서 발견된 AK가 CK group에 속하는 kinase sequence와 클러스터 된 것을 확인할 수 있으며, 이를 통해 AK group이 CK group으로부터 분지되어 2차적으로 진화하였음을 유추한 선행 연구결과가 있다 (Suzuki *et al.*, 1994; Uda *et al.*, 2005; Uda *et al.*, 2007). 또한 최근 연구자료에 식용 가능한 이매패류, 복족류 등 패각을 가진 연체동물들을 대상으로 allergen을 확인한 결과 대부분의 식용 연체동물에서 AK가 높은 농도로 발현됨을 확인한 연구 결과가 있다 (Lopata *et al.*, 2016). 이러한 선행 연구들의 결과를 확인하였을 때 AK gene 및 PK super family group에 속하는 gene들은 대다수의 연체동물에서 안정적으로 확인할 수 있으며 이를 이용한 분자계통학적 분석이 용이할 것으로 생각해 볼 수 있다. 또한 앞선 결과분석을 통해 AK 유전자의 2D 구조 예측 시 식생활환경에 따라 α -helix의 수가 다르게 나타난 것을 확인해 볼 수 있었다. 이를 토대로 AK gene에 대한 지속적인 연구는 향후 PK super family group에 속하는 gene을 이용한 Mollusca의 분자계통분류 분석 및 연구대상종의 생활환경 유추에 유용한 기초자료가 될 것으로 생각된다.

CONCLUSIONS

*E. chinense*의 AK 유전자는 1,068개의 염기서열로 이루어져 있으며, 54.03%의 GC contents를 가진다. Bioinformatic 분석을 통해 선정된 38종의 AK 유전자 서열 및 선행연구를 통해 발표된 *S. myomphala* 와 *M. junensis*의 서열로 구성된 gene pool을 이용하여 phylogenetic analysis를 진행하였다. 그 결과 선행된 연구와 같이 3개의 Class (Gastropoda, Cephalopoda, Bivalvia) 로 분류되는 것을 확인할 수 있었으며, *E. chinense*가 포함된 Gastropoda group 또한 Subclass 및 Order 단위의 분류군으로 묶이는 것을 확인할 수 있었다. Phylogenetic analysis 결과 *E. chinense*와 유연 관계가 가까운 분류군이 Heterobranchia인 것으로 확인되었다. 해당 group에 속한 종의 AK 단백질과 *E. chinense*의 AK 단백질의 2D 구조를 비교 분석 한 결과 생물종의 생활환경에 따라 α -helix의 수가 다르게 나타나는 구조적 차이점이 있는 것을 추정 할 수 있었다. 이러한 연구결과는 향후 AK gene을 이용한 지속적인 연구를 통해 PK super family에 속하는 유전자가 Mollusca의 분자계통학 분석 및 대상종의 생활환경 유추에 유용한 지표가 될 수 있을것으로 사료된다.

ACKNOWLEDGMENT

본 연구는 교육부 (한국연구재단, NRF-2017R1D1A3B06034971)

의 연구비 지원 및 순천향대학 학술연구비의 지원을 받아 수행되었습니다.

REFERENCE

- Buchan, D.W.A., and Jones, D.T. (2019) The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids. Res.*, **47**: W402-W407.
- Ellington, W., and Suzuki, T. (2006) Evolution and divergence of creatine kinase genes. *Molecular anatomy and physiology of proteins: creatine kinase. Nova Science, New York*: 1-27.
- Ellington, W.R. (2001) Evolution and physiological roles of phosphagen systems. *Annual review of physiology*, **63**: 289-325.
- Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *evolution*, **39**: 783-791.
- Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., and Regev, A. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.*, **29**: 644-652.
- Haas, B., and Papanicolaou, A. (2016) TransDecoder (find coding regions within transcripts). Google Scholar.
- Jarilla, B.R., Uda, K., Suzuki, T., Acosta, L.P., Urabe, M., and Agatsuma, T. (2014) Characterization of arginine kinase from the caenogastropod *Semisulcospira libertina*, an intermediate host of *Paragonimus westermani*. *Journal of Molluscan Studies*, **80**: 444-451.
- Jones, D.T., Taylor, W.R., and Thornton, J.M. (1992) The rapid generation of mutation data matrices from protein sequences. *Bioinformatics*, **8**: 275-282.
- Joshi, N., and Fass, J. (2011) Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33)[Software]
- Kang, S.W. (2016) Genome and transcriptome characterization of the four Korean land snails using NGS and bioinformatics databases. *Department of Biology, Graduate School of Soonchunhyang University*.
- Kang, S.W., Park, S.Y., Hwang, H.J., Chung, J.M., Sang, M.K., Min, H.R., Park, J.E., Cho, H.C., Patnaik, B.B., and Lee, Y.S. (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.
- 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., Kim, S., Park, H.S., Park, S.H., Park, Y.S., Han, Y.S., Lee, J.S., and Lee, Y.S. (2017) Sequencing

- and *de novo* assembly of visceral mass transcriptome of the critically endangered land snail *Satsuma myomphala*: Annotation and SSR discovery. *Comp. Biochem. Physiol. Part D. Genomics Proteomics*, **21**: 77-89.
- Kang, S.W., Patnaik, B.B., Park, S.Y., Hwang, H.J., Chung, J.M., Sang, M.K., Min, H.R., Park, J.E., Seong, J., Jo, Y.H., Noh, M.Y., Lee, J.D., Jung, K.Y., Park, H.S., Han, Y.S., Lee, J.S., and Lee, Y.S. (2018) Transcriptome analysis of the threatened snail *Ellobium chinense* reveals candidate genes for adaptation and identifies SSRs for conservation genetics. *Genes Genomics*, **40**: 333-347.
- Köhler, F., and Rintelen. (2011) *Ellobium chinense*. The IUCN Red List of Threatened Species 2011. The IUCN Red List of Threatened Species.
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.*, **35**: 1547-1549.
- Lee, J.S., and Son, M.-H. (2012) Red Data Book of Endangered Mollusks in Korea.
- Li, R., Zhang, S., Zhang, L., Yu, K., Wang, S., and Wang, Y. (2020) Field study of the microplastic pollution in sea snails (*Ellobium chinense*) from mangrove forest and their relationships with microplastics in water/sediment located on the north of Beibu Gulf. *Environ. Pollut.*, **263**: 114-368.
- Lopata, A.L., Kleine-Tebbe, J., and Kamath, S.D. (2016) Allergens and molecular diagnostics of shellfish allergy. *Allergo Journal International*, **25**: 210-218.
- Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D., and Lopez, R. (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids. Res.*, **47**: W636-W641.
- Mahon, B.C., and Neigel, J.E. (2008) Utility of arginine kinase for resolution of phylogenetic relationships among Brachyuran genera and families. *Mol. Phylogenet. Evol.*, **48**: 718-727.
- Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011*, **17**:(3).
- McGuffin, L.J., Bryson, K., and Jones, D.T. (2000) The PSIPRED protein structure prediction server. *Bioinformatics*, **16**: 404-405.
- McLeish, M.J., and Kenyon, G.L. (2005) Relating structure to mechanism in creatine kinase. *Critical reviews in biochemistry and molecular biology*, **40**: 1-20.
- Min, D.-K. (2004) Mollusks in Korea pp. Min Molluscan Research Institute.
- Min, H.R., Hwang, H.J., Chung, J.M., Sang, M.K., Cho, H.C., Park, J.E., Jung, K.Y., Park, H.S., Han, Y.S., and Lee, Y.S. (2018) Molecular Phylogenetic studies of *Mirus junensis* using Arginine kinase gene sequence. *The Korean Journal of Malacology*, **34**: 107-114.
- Min Kyu Sang, Hee-Ju Hwang, Jong Min Chung, Jie Eun Park, Dae Kwon Song, Jun Yang Jeong, So Young Park, Hong seog Park, Yong Hun Jo, Jong dae Lee, Jun Sang Lee, Yong Seok Lee, and Kang, S.W. (2020) Phylogenetic analysis of endangered wild animal Class II *Satsuma myomphala* using the arginine kinase gene. *The Korean Journal of Malacology*, **36**: 97-104.
- 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. Elsevier current trends.
- Saitou, N., and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular biology and evolution*, **4**: 406-425.
- Suzuki, T., Fukuta, H., Nagato, H., and Umekawa, M. (2000) Arginine kinase from *Nautilus pompilius*, a living fossil. Site-directed mutagenesis studies on the role of amino acid residues in the Guanidino specificity region. *J. Biol. Chem.*, **275**: 23884-23890.
- Suzuki, T., and Furukohri, T. (1994) Evolution of phosphagen kinase: primary structure of glycoamine kinase and arginine kinase from invertebrates. *Journal of molecular biology*, **237**: 353-357.
- Suzuki, T., Inoue, N., Higashi, T., Mizobuchi, R., Sugimura, N., Yokouchi, K., and Furukohri, T. (2000) Gastropod arginine kinases from *Cellana grata* and *Aplysia kurodai*. Isolation and cDNA-derived amino acid sequences. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, **127**: 505-512.
- Suzuki, T., Kawasaki, Y., and Furukohri, T. (1997) Evolution of phosphagen kinase1: Isolation, characterization and cDNA-derived amino acid sequence of two-domain arginine kinase from the sea anemone *Anthopleura japonicus*. *Biochemical Journal.*, **328**: 301-306.
- 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 Biochem. Mol. Biol.*, **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. *Comp. Biochem. Physiol. Part D Genomics Proteomics*, **1**: 209-218.
- Uda, K., Ishida, M., Matsui, T., and Suzuki, T. (2010) Arginine Kinase from the Tardigrade, *Macrobiotus occidentalis*: Molecular Cloning, Phylogenetic Analysis and Enzymatic Properties. *Zoological Science*, **27**: 796-803.
- Uda, K., Iwai, A., and Suzuki, T. (2005) Hypotaurocyamine kinase evolved from a gene for

- arginine kinase. *FEBS Lett.*, **579**: 6756-6762.
- Uda, K., and Suzuki, T. (2007) A novel arginine kinase with substrate specificity towards D-arginine. *The protein journal.*, **26**: 281-291.
- Uda, K., Yamamoto, K., Iwasaki, N., Iwai, M., Fujikura, K., Ellington, W.R., and Suzuki, T. (2008) Two-domain arginine kinase from the deep-sea clam *Calyptogena kaikoi*-evidence of two active domains. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.*, **151**: 176-182.
- Wyss, M., Smeitink, J., Wevers, R.A., and Wallimann, T. (1992) Mitochondrial creatine kinase: a key enzyme of aerobic energy metabolism. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, **1102**: 119-166.
- 서재화, 권선만, 이수연, 박창득, 이소희, 이정연, 유정선, and 서민환 (2018) 한눈에 보는 멸종위기 야생생물 (2017년 개정)

