

PLANT & FOREST

Construction of a Full-length cDNA Library from *Cardamine manshurica* Nakai and Characterization of EST Dataset

Subin Im, Sung-Ho Lee, Yoon-Young Kim, Ju-Sang Kim, Dasom Kim, Yong Pyo Lim*

Department of Horticultural science, Chungnam National University, Daejeon 34134, Korea

 *Corresponding author: yplim@cnu.ac.kr

Abstract

Brassicaceae consists of important species that have significant amounts of metabolites, and many studies have been carried out in order to understand the mechanism that improves the content of these metabolites. In Brassicaceae, *Cardamine manshurica* Nakai is one of the important edible plants and is rich in oil, fiber, and various nutrients. In this study, we constructed cDNA library using leaves from 4 week-old plants and analyzed the ESTs of *C. manshurica* Nakai. One thousand thirty-nine ESTs were discovered which assembled to form 468 unigenes. The latter contained 116 contigs and 352 singletons. Similarity search of these ESTs with BLASTX revealed similarities with *Arabidopsis thaliana* 285 (31.9%), *Arabidopsis lyrata* 172 (19.3%), *Capsella rubella* 162 (18.1%), and *Eutrema salsugineum* 137 (15.3%). ESTs were functionally categorized into molecular function, biological process, and cellular component, and each category took 10.6%, 58.5%, and 30.9%, respectively. The functional analysis also found that 94.9% of ESTs showed at least one GO ID. Microsatellite analysis of 468 unigene sequences revealed 225 structures of which Di-, Tri-, Tetra-, Penta-repeats were 35.6% (80/225), 63.1% (142/225), 0.9% (2/225), and 0.4% (1/225), respectively. The results from our study can be a valuable resource for *Cardamine* research.

Keywords: brassicaceae, EST analysis, gene ontology, microsatellite, unigene

Introduction

배추과(Brassicaceae) 또는 십자화과(Cruciferae) 식물에는 휘발성이 강한 독특한 매운 향기를 내는 2차 대사산물을 함유하고 있으며, 이 물질들에는 생리활성작용 및 항암효과가 있어(Fenwick et al., 1983; Hwang et al., 2012), 식용뿐 아니라 약용적 가치가 점차 높아지고 있는 추세이다. 배추과는 전 세계적으로 약 338속 3700 여종이 분포하며(Al-Shehbaz et al., 2006a), 주로 온대지역에 생육하는 것으로 알려져 있다(Zhou et al., 2001). 황새냉이속(*Cardamine* L.)은 장각과가 편평하고 과피가 감기며 꽃잎은 흰색 또는 자주색인 특징으로 다른 속과 구분된다(Al-Shehbaz, 1988; Song, 2011). 배추과 내에서 꽃다지속(*Draba* L.), 다닥냉이속(*Lepidium* L.), 부지깽이나물속(*Erysimum* L.)



OPEN ACCESS

Citation: Im SB, Lee SH, Kim YY, Kim JS, Kim DS, Lim YP. 2016. Construction of a Full-length cDNA Library from *Cardamine manshurica* Nakai and Characterization of EST Dataset. Korean Journal of Agricultural Science 43:33-39.

DOI: <http://dx.doi.org/10.7744/kjoas.20160005>

Editor: Kee Woong Park, Chungnam National University, Korea

Received: September 30, 2015

Revised: December 4, 2015

Accepted: December 21, 2015

Copyright: ©2016 Korean Journal of Agricultural Science.

This is an Open-Access article distributed under the terms of the [Creative Commons Attribution Non-Commercial License](https://creativecommons.org/licenses/by/4.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

음으로 큰 황새냉이속은 아시아, 북아메리카, 유럽을 중심으로 약 200여종이 보고되어 있으며(Zhou et al., 2001; Warwick, 2006), 이중 특히 동북아시아 지역은 히말라야 지역과 더불어 종분화의 중심지로 인식되고 있다(Al-Shehbaz, 1988; Lihova et al., 2006). 아시아지역에는 극동러시아에 약 13종(Komarov, 1970), 몽고 7종(Zhou et al., 2001), 중국 50종(Zhou et al., 2001; Chen et al., 2008), 대만 5종(Ying, 1996), 일본에는 17종이 분포하는 것으로 알려져 있으며(Al-Shehbaz et al., 2006b), 한국에는 가장 최근 연구에서 15종이 자생하는 것으로 보고하고 있다(Song, 2011).

본 속의 식물들은 어린순을 나물로 이용하며, 요도염, 방광염, 이질, 백대, 백일해 등에 효능이 있다고 기록되어 있다(Xiao, 1988; Bae, 2000). 본 속의 대표식물인 황새냉이는 수분 86.82%, 단백질 3.03%, 기름 0.47%, 무질소추출물 6.41%, 섬유소 2.15%, 재성분 1.33% 정도가 함유되어 있다(Kim, 2008). 본 속의 식물들은 어린순을 나물로 이용하며, 요도염, 방광염, 이질, 백대, 백일해 등에 효능이 있다고 기록되어 있다(Bae, 2000).

흔히 좁냉이(Koh, 1993; Korea national arboretum and the plant taxonomic society of Korea, 2007) 또는 말황새냉이(Oh, 2007; Song, 2011)로 불리는 *Cardamine parviflora* var. *manshurica* Kom.은 중국 길림성에서 채집된 표본에 기초하여 신변종으로 발표되었다(Komarov, 1903). 이후, 다수의 학자들이 유럽에 분포하는 *Cardamine parviflora* L.의 synonym으로 처리하여 왔으나(Kitagawa, 1939; Im, 1997; Zhou et al., 2001; Al-Shehbaz et al., 2006b; Warwick et al., 2006; Oh, 2007), 두 분류군은 형태적·지리적으로 명확하게 구별되는 것으로 밝혀졌다(Kim et al., 2015). Nakai(1914)는 *C. parviflora* var. *manshurica* Kom.와 형태적으로 동일한 분류군을 '*Cardamine manshurica* Nakai'라는 학명으로 기록하였으며, 이와 함께 간략한 도해와 일본어로 된 기재문을 수록하였다. 또한 그는 *C. manshurica* Nakai에 작고 귀여운 황새냉이라는 뜻을 가진 일본명(ひめたねつけばな)을 신청하였다. 이후, Lee (1996)는 이를 번역한 '애기황새냉이'를 국명으로 부여하였으며, Kim et al. (2015)은 Lee (1996)가 제시한 국명을 추천명으로 채택하였다.

애기황새냉이는 야생상태에서 1년에 3-5회 세대교번이 이루어져서 십자화과의 대표적 모델식물인 애기장대보다 세대교번이 짧다. 또한 발아율 및 결실율이 매우 높고, 건조하거나 다습한 환경에도 잘 자라며 자가수분이 가능하여 연구에 이용하기에 적합한 식물이다.

따라서, 본 논문에서는 애기황새냉이의 cDNA library 제작 및 EST 분석을 통해 유전적으로 애기황새냉이에 대해 연구하고, 유전자원으로 확보하고자 한다.

Materials and Methods

식물 재료 준비 및 RNA 추출

애기황새냉이의 시료는 충남대학교 실습답(36° 22' 3.64" N, 127° 21' 19.85" E)에서 채취한 자연집단의 성숙모본(Fig. 1)을 25°C, 장일조건(16 h light / 8 h dark)에서 배양하고 자가 수분하여 얻은 종자를 이용하였다. 수확한 종자를 음지에서 건조 후 동일 조건에서 4주간 생육시켜 잎을 실험 재료로 사용하였다. Total RNA의 추출은 액체질소로 동결시켜 곱게 파쇄한 시료 약 100 mg에 시료에 1 mL의 TRIzol® Reagent (Life technology, USA)를 이용하여 제조사의 지침에 따라 진행하였다. 최종 RNA pellet은 DEPC-treated D.W.에 녹여 NanoVue™ (GE Healthcare, USA)로 농도를 측정하였다. 추출한 total RNA로부터 PolyAtract® mRNA Isolation System (Promega, USA)을 이용하여 Poly (A+) RNA를 선별하였다.



Fig. 1. General shape of *C. manshurica* at research farm in Chungnam National University (A) and selected mature plants for self-pollination (B).

전장 cDNA Library 제작 및 염기 서열 분석

선별된 Poly (A)⁺ RNA로부터 전장 cDNA Library를 제작하였다. cDNA library의 제작은 In-Fusion® SMARTer® Directional cDNA Library Construction Kit (Takara, USA)을 이용하였다. 제조사의 지침에 따라 1 ug의 Poly (A)⁺ RNA를 재료로 Library 제작을 진행하였다. Library Clone들은 X-gal을 이용한 Blue/White selection 방법을 통해 cDNA를 포함하는 Clone들만 선별하여 Library로 확립하였다. 제작된 cDNA Library 복사본을 바이오니아(Korea)사에 의뢰하여 염기서열 분석을 수행하였다.

EST 염기서열의 분석

분석으로 얻은 염기서열은 CodonCode Aligner Ver. 3.7.1을 사용하여 전처리 및 contig assembly 과정을 수행하였다. 이를 통해 vector 서열 및 100 bp 이하 서열을 제거하였고, contig와 singlet 염기서열로 정렬하여 unigene 염기서열을 얻었다. 분석된 각각의 서열에 대한 상동성 탐색과 Gene ontology의 결정은 Blast2GO (<http://blast2go.bioinfo.cipf.es/>) 프로그램을 사용하여 수행하였다(Conesa et al., 2005). 분석에 필요한 parameter는 BLASTX cut off 10^{-5} , mapping E-value 10^{-5} , annotation cut off 45, GO weight 10으로 설정하여 진행하였다. Microsatellite는 SSR Finder (<http://www.csufresno.edu/ssrfinder>)를 사용하여 분석하였다. Di nucleotide는 5회 반복 이상, Tri nucleotide 이상의 경우 4회 이상 반복된 것을 microsatellite로 분류하였다.

Results and Discussion

애기황새냉이 전장 cDNA Library의 제작

애기황새냉이에서 추출한 total RNA를 NanoVue™로 측정된 결과, UV 흡광도는 A260에서 12.71, A280에서 7.73으로 산출되었으며, 농도는 420.8 ng/ul로 나타났다. 전장 cDNA Library의 제작은 선별된 Poly (A)⁺ RNA를 3M Sodium acetate를 이용한 침전방법을 통해 농축하고, 0.5 ug/ul로 희석하여 총 1 ug의 RNA를 이용하여 수행하였다. 합성된 cDNA는 CHROMA SPIN™+TE-1000 column (Takara, USA)을 사용하여 1000 bp 이하의 cDNA와 salts, proteins 등을 제거하였다. 선별된 cDNA는 pSMART2IF 벡터에 재조합하고, *E. coli* strain DH5 α 에 Ice cold-shock transformation 방법을 통해 형질전환 시켰다.

염기서열 해독 및 분석

얻어진 2,235개의 EST clone중 총 1,054개에 대하여 염기서열을 분석하였다. EST 염기서열의 평균 길이는 약 1,094 bp이며 전체 분석 서열 길이는 1.04×10^6 bp 이었다. EST 염기서열은 CodonCode Aligner Ver.

3.7.1을 사용한 전처리 과정을 통해 Vector 염기서열과 Low quality 분석 부분을 제거한 후 50 bp이하의 서열은 분석에서 제외되어 총 1,036개의 염기서열이 연계분석에 사용되었다. 684개의 염기서열이 116개의 contig로 통합되었고, singleton은 352개로 총 468개의 Unigene을 얻었다(Fig. 2).

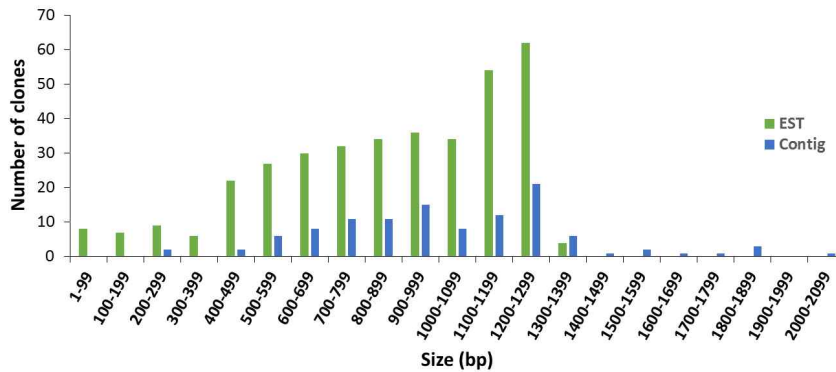


Fig. 2. Distribution of the sequence length of ESTs and contigs.

NCBI (National Center for Biotechnology Information) nr database의 BLASTX검색은 Blast2GO (<http://www.blast2go.de/>) 프로그램을 사용하여 최소 E-value 값을 10^{-4} 로 설정하여 상동성을 검색하였다(Fig. 3). 전체 1,036개의 서열에 대한 BLASTX 분석 결과 이 중 No hit는 44개(4.2%), Unknown gene은 8개(0.8%) 그리고 Known gene은 984개(95%)로 확인되었다(Fig. 3A). 984개의 Known gene의 종별 분포는 *Arabidopsis thaliana* 285개(31.9%), *Arabidopsis lyrata* 172개(19.3%), *Capsella rubella* 162개(18.1%), *Eutrema salsugineum* 137개(15.3%) 순으로 나타났으며, 이중 특히 *A. thaliana* (애기장대)와 *A. lyrata* (뫇장대)는 높은 상동성을 보여주었다(Fig. 3B). 이는 *A. thaliana*의 경우 오랜 기간 유전체 연구가 이뤄진 모델식물로서 많은 분석 자료가 축적되어 이러한 높은 상동성이 나타난 걸로 생각된다. 또한 *A. lyrata*는 *A. thaliana*와 매우 가까운 근연종으로 높은 상동성을 나타낸 것으로 보인다.

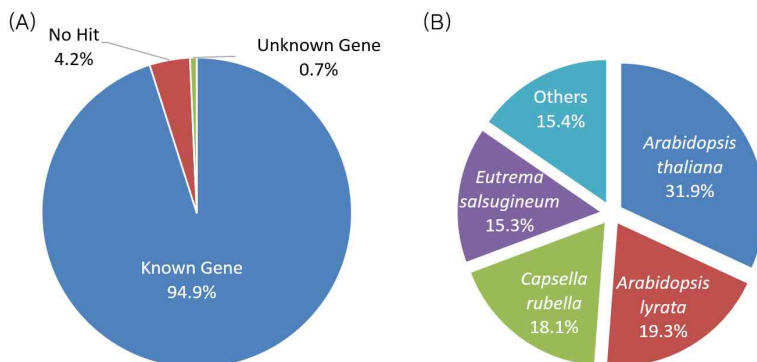


Fig. 3. Graphical representation of functional annotation and distribution of BLASTN top-hit species. (A) classification of all ESTs with BLASTN. (B) BLASTN top-hit species showing high homology with known genome sequences.

분석이 수행된 1,036 EST 염기서열에 대한 gene ontology는 Blast2GO software (<https://www.blast2go.com/>) 의 mapping 기능을 사용하여 Functional category에 따라 분류하고 Plant Go Slim

기능을 사용하여 중복되는 GO ID를 최소화 시켰다. 각각의 비율은 Molecular function 10.6%, Biological process 58.5%, Cellular component 30.9%로 나타났다(Fig. 4). 전체 EST 염기서열 중 적어도 한 개 이상의 GO ID가 부여된 경우는 94.9% (984/1036)로 나타났으며, GO ID는 EST 염기서열마다 평균 10.5개, 전체 10,967개로 확인됐다. 기능에 따른 분류에 있어 molecular function은 'protein binding', Biological process는 'response to abiotic stimulus', Cellular component는 'plastid'가 가장 높은 비율로 나왔다.

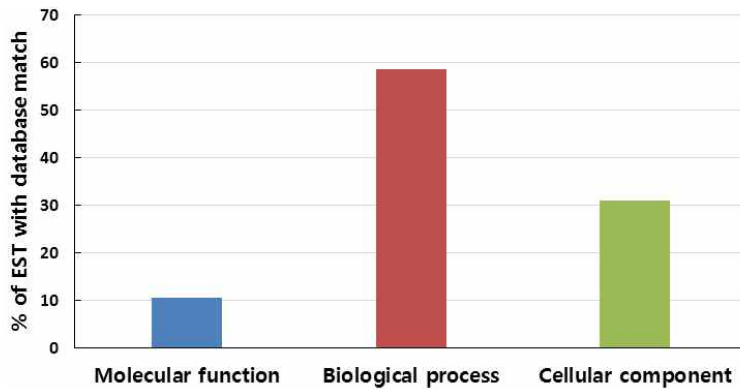


Fig. 4. Graphical representation of three functional categories of EST sequences.

분리된 468개의 unigene에서 microsatellite서열을 분리하여 구조를 분석하였다. 각각의 microsatellite는 반복 서열의 구조에 따라 분류하였고, 전체 총 225개의 microsatellite를 확인하였다. Di-, Tri-, Tetra-, Penta-repeat의 빈도는 각각 35.6% (80/225), 63.1% (142/225), 0.9% (2/225), 0.4% (1/225)로 나타났다. 전체 microsatellite 중 Tri repeat인 (AAG)_n이 27.6% (62/225)로 다른 microsatellite에 비해 가장 높은 비율이 확인되었다(Table 1).

Table 1. Characterization of the repeat structures for the identified microsatellite sequences.

Repeat unit	Repeat sequence	Synonym	Observed No.	Total No.
Di-nt	AT	TA	18	80
	AC	CA, GT, TG	16	
	AG	GA, CT, TC	45	
	CG	GC	1	
Tri-nt	AAC	ACA, CAA, GTT, TTG, TGT	12	142
	AAG	AGA, GAA, CTT, TTC, TCT	62	
	AAT	ATA, TAA, ATT, TTA, TAT	9	
	ACC	CCA, CAC, GGT, GTG, TGG	10	
	AGC	GCA, CAG, GCT, CTG, TGC	8	
	ACG	CGA, GAC, CGT, GTC, TCG	1	
	AGG	GGA, GAG, CCT, CTC, TCC	9	
	ACT	CTA, TAC, AGT, GTA, TAG	3	
	ATC	TCA, CAT, GAT, ATG, TGA	26	
	GCC	CCG, CGC, GGC, GCG, CGG	2	
Tetra-nt	TTTC	TTCT, TCTT, CTTT, GAAA, AAAG, AAGA, AGAA	2	2
Hexata-nt	GGAAGA	GGAAGA	1	1
Total				225

본 연구에서는 애기황새냉이(*C. manshurica*)의 EST library를 제작하여 전체 1,054개의 EST clone을 분석하였다. 전체 116개의 contig와 352개의 singleton을 분리하여 총 468개의 unigene을 분리해 냈으며 EST 서열분석을 통해 총 225개의 microsatellite를 분리해 냈다. 이중 BLASTX를 통해 상동성이 확인되지 않은 4.2%의 no hit 유전자는 애기황새냉이의 특이적 유전자일 가능성이 높을 것으로 생각된다. 또한 항균 단백질로 알려진 Defensin과 관련된 유전자들이 몇몇의 clone에서 발견되었으며, 근연관계인 *A. thaliana*에 존재하는 13개 defensin gene (Penninckx et al., 1996)의 비교 연구 등도 가능할 것이다.

본 연구를 통해 얻어진 애기황새냉이 유전자원과 유전 정보들은 다양한 분야의 기초자료로 사용될 것으로 사료된다.

Acknowledgements

본 연구는 “충남대학교 2014 자체연구과제” 지원에 의해 수행되었으며, 저자 중 임수빈 · 이성호는 논문작성에 동일한 기여를 하였습니다.

References

- Al-Shehbaz IA, Aria K, Ohba H. 2006b. Cruciferae. In Flora of Japan. Vol. 2a. Iwatsuki K, Boufford DE, Ohba H (eds.). Kodansha Ltd., Tokyo. pp. 454-511.
- Al-Shehbaz IA. 1988. The genera of Arabideae (Cruciferae; Brassicaceae) in the Southeastern United States, Journal of the Arnold Arboretum 69:85-166.
- Al-Shehbaz IA, Beilstein MA, Kellogg EA. 2006a. Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview. Plant Systematics and Evolution 259:89-120.
- Bae GH. 2000. The Medicinal Plants of Korea. Kyohaksa, Seoul pp. 187-191.
- Bleeker W, Franzke A, Pollmann K, Brown AHD, Hurka H. 2002. Phylogeny and biogeography of Southern Hemisphere high-mountain Cardamine species (Brassicaceae). Australian Systematic Botany 15:575-581.
- Chen LB, Liu KM, Liu LH. 2008. A New Species of Cardamine (Brassicaceae) from Hunan, China. A Journal for Botanical Nomenclature 18:135-137.
- Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674-3676.
- Fenwick GR, Griffiths NM, Heaney RK. 1983. Bitterness in Brussels sprouts (*Brassica oleracea* L. var. *gemmifera*): the role of 126 glucosinolates and their breakdown products. Journal of the Science of Food and Agriculture 34:73-80.
- Hwang ES, Hong EY, Kim GH. 2012. Determination of Bioactive Compounds and Anti-cancer Effect from Extracts of Korean Cabbage and Cabbage. The Korean Journal of Food And Nutrition 25:259-265.
- Im RJ. 1997. Brassicaceae. In Im RJ, Kim HS, Gwak JS, La EC, Lee YJ, Lee GP, Han GS (eds.), Flora Coreana, Vol. 3. The Science and Technology Publishing House, Pyongyang. pp. 32-91.
- Kim YY, Ji SJ, Oh BU. 2015. Taxonomic review of *Cardamine manshurica* (Kom.) Nakai (Brassicaceae). Korean Journal of Plant Taxonomy 45:136-144.
- Kitagawa M. 1939. Lineamenta Florae Manshuricae. Report of the Institute of Scientific Research, Manchoukuo 3:1-487.
- Koh GS. 1993. An observation of Korean wild plants in color. Woosung, Seoul.
- Komarov HL. 1970. Flora of the U.S.S.R.. Vol. 8. Keter Press. Washington, D.C..
- Komarov VL. 1903. Flora of Manshuriae t. 2. pars 1. Trudy Imp. S.-Peterburgsk. Bot. Sada 22:1-452.

- Korea National Arboretum and The Plant Taxonomic Society of Korea. 2007. A Synonymic List of Vascular Plants in Korea. Gyeonggi.
- Kudoh H, Marhold K, Lihová J. 2006. Note on *Cardamine impatiens* L., *C. flexuosa* With., *C. hirsuta* L. and *C. manshurica* L. in Japan. *Bunrui*. 6:41-49 (in Japanese with English summary)
- Kudoh H, Ishiguri Y, Kawano S. 1992. *Cardamine hirsuta* L., a new ruderal species introduced into Japan. *Journal of Phytogeography and Taxonomy* 40:85-89.
- Kudoh H, Ishiguri Y, Kawano S. 1993. Phenotypic variability in life history traits and phenology of field populations of *Cardamine flexuosa* and *C. fallax* (Cruciferae) in Honshu, Japan. *Plant Species Biology* 8:7-20.
- Lee WT. 1996. *Lineamenta Florae Koreae*. Academy publishing Co. Ltd, Seoul.
- Lihova J, Marhold K, Kudoh H, Koch MA. 2006. Worldwide phylogeny and biogeography of *Cardamine flexuosa* and ITS Relatives. *American Journal of Botany* 93:1206-1221.
- Nakai T. 1914. Cruciferae. In *Tyosen Syokubutsu*. Seibido Publishig Co., Tokyo. Pp. 103-118.
- Oh BU. 2007. Brassicaceae. In *Flora of Korea* Editorial Committee (eds.), *The Genera of Vascular Plants of Korea*. Academic Publishing Co., Seoul. pp. 427-461.
- Okayama H, Berg P. 1982. High efficiency cloning of full-length cDNA. *Molecular and Cellular Biology* 2:161-170.
- Peninckx IA, Eggermont K, Teras FR, Thomma BP, De Samblanx GW, Buchala A, Metraux JP, Maners JM, Brodkaset WF. 1996. Pathogen-induced systemic activation of a plant defensin gene in *Arabidopsis* follow a salicylic acid-independent pathway. *The Plant Cell* 8:2309-2323.
- Song IG. 2011. A molecular phylogenetic study of the genus *Cardamine* L. (Brassicaceae) in Northeast Asia. PhD Diss., Yeungnam University, Kyeongsan.
- Warwick SI, Francis A, Al-Shehbaz IA. 2006. Species checklist and database on CD-Rom. *Plant Systematics and Evo* 259:249-258.
- Ying SS. 1996. *Flora of Taiwan*. Vol. 2. T. C. Huang. Editorial Committee of the Flora of Taiwan. Tah Jinn Printing Company, Ltd., Taipei. pp. 745-769.
- Zhou T, Lu L, Yang G, Al-Shehbaz IA. 2001. Brassicaceae (Cruciferae). In *Flora of China*. Vol. 8. Wu ZY, Raven PH, Hong DY (eds.). Science and Missouri Botanical Garden Press, Beijing and St Louis. pp. 1-193.