INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
18–1194/2019/21–3–597–605
DOI: 10.17957/IJAB/15.0934
http://www.fspublishers.org

Full Length Article

Genome-Wide Analysis of the ATP-Binding Cassette A (ABCA) Gene Family and their Expression Profiles in High and Low Oil Cultivars of *Brassica napus*

Nannan Li[1,2*†], Zhongchun Xiao[2,3†], Mingwei Guan[2,3†], Shengting Li[1], Xiaohui Zhang[1] and Feng Luo[1]
[1]*College of Resources and Environment, Southwest University, Beibei, Chongqing 400715, People's Republic of China*
[2]*Academy of Agricultural Science, Southwest University, Beibei, Chongqing 400715, People's Republic of China*
[3]*College of Agronomy and Biotechnology, Southwest University, Beibei, Chongqing 400715, People's Republic of China*
[*]For correspondence: linannan2013@swu.edu.cn
[†]These authors contributed equally to this work

Abstract

ATP-binding cassette (ABC) proteins exist widely in plant cells and ABCA was a subfamily in ABC superfamily. AtABCA9 mediates acyl-CoA/Fatty acids import into Endoplasmic Reticulum (ER) in *Arabidopsis*. However, the characterization of *ABCA* genes in *Brassica napus* L remains unknown. In the present study, 32 members of *ABCA* gene (*BnABCAs*) families were identified in the *B. napus* through homology screening. We performed the analysis of gene structures, chromosomal distribution, phylogenetic relationships and motifs. Meanwhile, we also predicted the subcellular localization and 3D models of partial identified BnABCA members. Furthermore, the expression patterns of most identified *BnABCA* genes were examined in seeds of different stages of *B. napus* by RNA-Seq. The result revealed that *BnABCA9-1* and *BnABCA9-2* were mainly expressed in the seeds during seed oil accumulation stages. Furthermore, the expression levels of *BnABCA9-1 and BnABCA9-2* genes were analyzed in seeds of different stages of two high/low oil content cultivars of *B. napus* by qRT-PCR assays. The results imply that the BnABCA9 may be involved in the acyl-CoA export from ER or lipid degradation, which is different from the role of ABCA9 in *Arabidopsis*. © 2019 Friends Science Publishers

Keywords: *Brassica napus*; ABCA; Expression analysis; RNA-Seq; qRT-PCR

Introduction

ATP-binding cassette (ABC) proteins exist widely in plant cells (Higgins, 1992). The ABC protein superfamily contains over 120 proteins in *Arabidopsis thaliana* (Sanchez-Fernandez *et al.*, 2001) and majority of ABC proteins mainly are functional transporters. The ABC transporters in plants have three salient characteristics. First, it is energized by MgATP; second, transport is not related to the transmembrane H^+ electrochemical changes and finally, transport is sensitive to the vanadate difference (Rea, 2007).

In membrane biology, the research of intracellular lipid transfer is a fundamental aspect. ABC transporters have a large number of members in most species from humans to prokaryotes, and it is an essential gene family in transport system (Jones and George, 2004). The ABC transporters can make use of the energy of ATP to transport kinds of substances across cell membranes (Pighin *et al.*, 2004; Rea, 2007; Verrier *et al.*, 2008; Kuromori *et al.*, 2010; Le Hir *et al.*, 2013).

Until now, we know that the ABCAs and ABCGs may have the most likely probability in lipid transport. It is reported that lots of the member of ABCG is associated with lipid transform in *A. thaliana*. For example, two ABC transporters, named ABCG11 and ABCG12, mediated cuticular lipids secretion (Pighin *et al.*, 2004; Bird *et al.*, 2007; Luo *et al.*, 2007; Panikashvili *et al.*, 2007). Meanwhile, the ABCG13 was involved in the transport of flower cuticular lipids (Panikashvili *et al.*, 2011). Besides, the ABCG9/11/14 are important in lipid/sterol homeostasis regulation in *A. thaliana* (Le Hir *et al.*, 2013). Furthermore, the ABCG2 and ABCG20 are important in the suberin barrier synthesis in roots, while the ABCG1 and ABCG16 play important roles in synthesis of pollen wall (Yadav *et al.*, 2014). In rice, previous study revealed that OsABCG15 and OsABCG26 cooperatively regulated rice male reproduction. OsABCG15 mainly took charge of the transport of lipids from tapetal cells to anther locules for further pollen exine biosynthesis (Zhao *et al.*, 2015). As for ABCA subfamily, the majority of *ABCA* genes were mainly reported in mammalian, and there is evidence that some *ABCA* genes might contribute to brain lipid homeostasis, and lack of lipid transporting ABC proteins can cause many important genetic diseases (Kim *et al.*, 2008; van Meer *et*

al., 2008). Besides, the upregulated activity of ABCA proteins leaded to over accumulation of lipids in some specific animal tissues (Oram and Vaughan, 2006). However, the research of the ABCA gene family in plant remains unknown, except for ABCA9 in *A. thaliana*. ABCA9, an ABC transporter, which is localized to the endoplasmic reticulum (ER) and plays a vital role in fatty acid transport into the ER in *A. thaliana* (Kim *et al.*, 2013). It is known that the ABC proteins widespread exist and their functions are highly conserved (Higgins, 1992). Therefore, we suspect that the members of *ABCA* gene family may also play an important role in *B. napus*.

Oilseed rape (*B. napus*) is one of the major oil crops used for food, feed animal, source of biofuel etc., but some studies showed that the harvest index (HI) is only 0.2–0.3 of oilseed rape (Luo *et al.*, 2015). In China, the average oil content is about 40% of rapeseed varieties, which is 2–7% lower than the average oil content of exotic oilseed varieties. Thus, the research to improve the oil content and HI has become very important. In the present study, we hope that some important genes can be found by performing comprehensive analysis for the *ABCA* gene family, especially the *ABCA9* genes, which may contribute to promote the HI and oil content of the Oilseed rape. The BLASTP search program was used to query for *ABCA* family members in the genomes of three species, *B. rapa, B. oleracea* and *B. napus*, using AtABCA protein sequences as the query. 32 members of *ABCA* gene (*BnABCAs*) families were identified in the *B. napus.* through homology screening. We further performed analysis of gene structures, chromosomal distribution, phylogenetic relationships and motifs. Meanwhile, the subcellular localization and 3D models of partial identified BnABCA members were also predicted and analysized. Furthermore, the expression patterns of most identified *BnABCA* genes were examined in seeds of different stages of by the RNA-Seq data from our group, which was submitted to NCBI Gene Expression Omnibus (No. GSE77637) and to verify the reliability of RNA-Seq results. Subsequently, the expression of several *BnABCA9* genes of two high/low oil content cultivars of *B. napus.* (P474-HO and P525-LO) were further investigated using qRT-PCR. By our research, we hope to provide some guideline about the research of *ABCA* gene family in *B. napus.*

Materials and Methods

Plant Materials

Two different *B. napus* varieties (*P474-HO; P525-LO*) were grown in net house conditions in Chongqing, China. To gain into the expression patterns of some *BnABCA* members in different tissues of two rape lines, including stalk (St), leaf (Le), seeds after flowering 10 days (Se-10d), seeds after flowering 20 days (Se-20d), seeds after flowering 30 days (Se-30d), seeds after flowering 40 days (Se-40d) were collected from the two rape lines. All tissues were quickly frozen in liquid nitrogen and stored at -80 centigrade until use.

RNA Isolation and qRT-PCR Analysis

RNA extraction from 150–200 mg samples was performed using the EZ-10 RNA Mini-prep Kit (Sangon Biotech, Shanghai, China). DNase digestion was done using RNase-free (PremeScript^{TM} RT reagent Kit with gDNA Eraser, TAKARA). The cDNA was synthesized by the (PremeScript^{TM} RT reagent Kit with gDNA Eraser, TAKARA. The reaction system is 20 μL, containing 800 μg total RNA.

Identification of *ABCA* Gene Family in *Brassica* spp.

The ABCA amino acid, coding and the genomic sequences in *A. thaliana* were downloaded from *TAIR* (Lamesch *et al.*, 2012). BLASTP was performed according to (Altschul *et al.*, 1997) *BRAD database* (http://brassicadb.org/brad/index.php) and the *B. napus genome* (http://www.genoscope.cns.fr/brassicanapus/) at a cut-off value of < E−20 to find homologous *ABCA* genes in *B. rapa*, *B. oleracea* and *B. napus*, and the query sequences are the ABCA amino acid sequences in *A. thaliana.* According to the results of the BLASTP, download the CDSs and the genomic sequences from the *B. napus* genome and recorded the information of exon/intron.

Phylogenetic Analysis of ABCA Family in *B. napus*

To know the evolutionary relationships of ABCA family members in *B. napus*, multiple alignments of the ABCA proteins were performed in the four species *A. thaliana*, *B. rapa*, *B. oleracea* and *B. napus*. Dendrograms were generated by the MEGA 7.0 (Tamura *et al.*, 2013; Kumar *et al.*, 2016) program using the neighbor-joining (NJ) method (Lu *et al.*, 2015). Subsequently, the phylogenetic trees were adjusted by using the program FigTreev1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/).

Protein Properties Analyses

The ExPASy proteomics server database (http://expasy.org/) was used to predict the molecular weight and isoelectric points of BnABCAs (Gasteiger *et al.*, 2003). The exon/intron analysis was performed by using the Gene Structure Display Server (GSDS2.0, http://gsds.cbi.pku.edu.cn/index.php). The analysis of conserved motif was conducted by the *MEME* program (http://meme-suite.org/; (Bailey *et al.*, 2006). All identified motifs were annotated according to InterProScan (http://www.ebi.ac.uk/interpro/search/sequence-search; (Zdobnov and Apweiler, 2001).

Prediction about the Subcellular Localization

The subcellular localization was predicted by the *TargetP*

1.1 Server (http://www.cbs.dtu.dk/services/TargetP/; Emanuelsson *et al.*, 2007). At the same time, we also chose the WoLF PSORT (http://www.genscript.com/wolf-psort.html) as a verification tool (Horton *et al.*, 2007).

3D Model about ABCA9 Proteins

We chose 5 proteins (BnABCA9-1, BnABCA9-2, BnABCA9-3, BnABCA9-4, BnABCA9-5) and used the phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) to predict the 3D model of BnABCA9 proteins (Kelley and Sternberg, 2009).

Expression Patterns Analysis of *BnABCA* Genes

To investigate the spatial expression patterns of *BnABCA* genes, two RNA-seq data from our laboratory were used. One was submitted to *NCBI Gene Expression Omnibus* (accession number GSE77637). Real-time PCR was used to verify the consistency. We tested the expression level of *ABCA9* homologous genes in *B. napus* in two different varieties (P474-HO and P525-LO) as per result of e-*FP Browser* (http://bar.utoronto.ca/efp_arabidopsis/cgi-bin/efpWeb.cgi). Each sample was replicated three times.

Results

Identification and Sequence Analysis of ABCAs in *B. napus*

To identify all ABCA protein sequences in *B. napus, B. rapa and B. oleracea*, we used the BLASTP program and identified 32, 12, 14 ABCA candidate genes, respectively. According to the BLASTP program, we gave each gene an identifier from *BnABCA1-1* to *BnABCA12-4*, totally 32 numbers (Table 1). There were a maximum and minimum number of amino acids in addition to the BnABCA1-1(1860 aa) and BnABCA3-1(426 aa) proteins, the amino acids of all the other proteins ranged from 457 (BnABCA9-4) to 1853 (BnABCA1-2) and the average length was 903 aa. We could not get the molecular weight (MW) and theoretical pI of BnABCA12-1 protein due to some unknown amino acid sequences, but relative molecular weight of other BnABCA proteins varied from 47.72 kDa (*BnABCA3-1*) to 206.66 kDa (*BnABCA1-1*), meanwhile, the theoretical pI values ranged from 6.00 (*BnABCA3-1*) to 9.14 (*BnABCA6-3*), with 6 of their pIs were less than 7 (18.75%).

To study the evolutionary relationships of ABCA proteins in *B. oleracea, B. rapa* and *B. napus*, we used amino acid sequences of the ABCA family proteins from that of three species to create the unrooted neighbor joining phylogenetic tree. Based on the results of multiple sequence alignment and the phylogenetic analysis, *ABCA* family genes were divided into two groups. It was reported that the plant ABCA subfamily genes were divided to full-size and

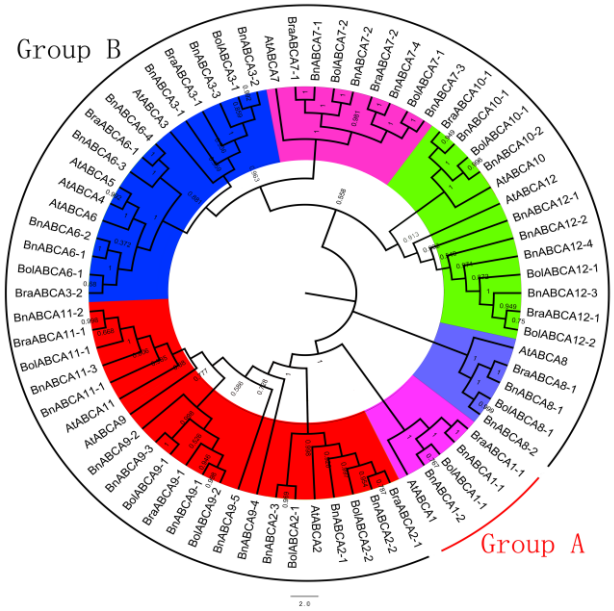

Fig. 1: The phylogenetic analysis of ABCA proteins from *A. thaliana, B. oleracea, B. rape* and *B. napus*
70 *ABCAs* were identified to construct the NJ tree with 1000 bootstraps based on the protein sequences. The ABCA proteins were grouped into two distinct types

half-size proteins in *Arabidopsis* genome, and there is only one full-size named *AtABCA1 in Arabidopsis*. Therefore, in group A, there are 5 proteins, including AtABCA1, BnABCA1-1, BnABCA1-2, BraABCA1-1 and BolABCA1-1 and others belong to Group B (Fig. 1).

The Chromosomal Location and Sequences Analysis of *BnABCAs*

The chromosomal locations of *BnABCA* genes are shown in Fig. 2. Obviously, most of the *BnABCA* genes (87.5%) are located at the A or C genomes. According to the Table 1 and Fig. 2, it is evident that the chromosome C01 contains four genes, C03 contains six genes, C07 contains five genes and Cnn-random contains four genes. Besides, there are two genes on Chromosomes A01, five genes on A03, five genes on A06 and only one gene on A07. It is reported that the gene divergence and duplication were the major factors for evolutionary momentum (Vision *et al.*, 2000; Bowers *et al.*, 2003). To better understand the structural features of BnABCA proteins, we compared with all BnABCA exon/intron organizations. The most closely related *BnABCA* genes usually have the most similar gene structures. Except the homologous genes of *AtABCA1* in *B. napus L.* have 41 exons and the *BnABCA12-1* and *BnABCA3-1* have 7 and 8 exons, respectively, other genes have 14 to18 exons (Fig. 3).

To better understand the characteristics of the BnABCA proteins, the conserved motifs of BnABCA proteins were identified by using MEME 7.0 (Fig. 4).

Table 1: Main structural features of the BnABCA family members

Isoforms	Transcript name	At Orthologs	Location	gDNA size (bp)	Exon	CDS size (nts)	Peptide residues	Theoretical pI	Theoretical Mw (Da)
BnABCA1-1	GSBRNA2T00058390001(BnaA03g19570D)	At2g41700	chrA03:9270777..9282863	12087	41	5583	1860	6.21	206655.1
BnABCA1-2	GSBRNA2T00019127001(BnaC03g23460D)	At2g41700	chrC03:13069016..13081111	12096	41	5562	1853	6.13	205528.8
BnABCA2-1	GSBRNA2T00024765001(BnaC03g57180D)	At3g47730	chrC03:46434921..46439482	4562	17	2814	937	6.15	10323.17
BnABCA2-2	GSBRNA2T00063075001(BnaA06g17180D)	At3g47730	chrA06:9733452..9738132	4681	16	2898	965	6.68	106324.8
BnABCA2-3	GSBRNA2T00125859001(BnaC01g25120D)	At3g47730	chrC01:20568513..20572099	3587	14	2133	710	8.11	78756.4
BnABCA3-1	GSBRNA2T00125865001(BnaC01g25160D)	At3g47740	chrC01:20594996..20601540	6545	8	1281	426	6	47724.7
BnABCA3-2	GSBRNA2T00068088001(BnaC07g07140D)	At3g47740	chrC07:11412417..11417340	4924	16	2808	935	8.99	105441.3
BnABCA3-3	GSBRNA2T00128266001(BnaA07g05590D)	At3g47740	chrA07:5888346..5892526	4181	16	2808	935	9.02	105590.6
BnABCA6-1	GSBRNA2T00024763001(BnaC03g57200D)	At3g47770	chrC03:46455698..46460090	4393	16	2856	951	8.56	107505.9
BnABCA6-2	GSBRNA2T00063073001(BnaA06g17160D)	At3g47770	chrA06:9710170..9714869	4700	17	2796	931	8.75	105207.3
BnABCA6-3	GSBRNA2T00024764001(BnaC03g57190D)	At3g47770	chrC03:46450927..46455084	4158	15	2781	926	9.14	104232
BnABCA6-4	GSBRNA2T00063074001(BnaA06g17170D)	At3g47770	chrA06:9717169..9721301	4133	15	2661	886	8.95	100014.9
BnABCA7-1	GSBRNA2T00063072001(BnaA06g17150D)	At3g47780	chrA06:9704566..9709122	4557	16	2829	942	8.4	105265.5
BnABCA7-2	GSBRNA2T00024762001(BnaC03g57210D)	At3g47780	chrC03:46462394..46466914	4521	16	2829	942	8.71	105152.5
BnABCA7-3	GSBRNA2T00125871001(BnaC01g25190D)	At3g47780	chrC01:20621569..20626864	5296	16	2829	942	8.73	105302.8
BnABCA7-4	GSBRNA2T00138490001(BnaA01g20250D)	At3g47780	chrA01:11885644..11890217	4574	16	2829	942	8.59	105507
BnABCA8-1	GSBRNA2T00016337001(BnaA06g20660D)	At3g47790	chrA06:14207210..14211668	4459	18	2715	904	9.08	102093.2
BnABCA8-2	GSBRNA2T00129061001(BnaC03g53450D)	At3g47790	chrC03:38878588..38882655	4068	17	2490	829	8.99	94276.54
BnABCA9-1	GSBRNA2T00021504001(BnaC07g31470D)	At5g61730	chrC07:35545897..35550242	4346	16	2865	954	8.16	105839.7
BnABCA9-2	GSBRNA2T00113508001(BnaA03g40510D)	At5g61730	chrA03:20236924..20244851	7928	16	2523	840	8.36	92889.05
BnABCA9-3	GSBRNA2T00007894001(BnaCnng23540D)	At5g61730	chrCnn_random:21948910..21959469	10560	16	2637	878	8.27	97534.21
BnABCA9-4	GSBRNA2T00036430001(BnaCnng40470D)	At5g61730	chrCnn_random:39006476..39009023	2548	14	1374	457	6.84	50652.65
BnABCA9-5	GSBRNA2T00040962001(BnaCnng43390D)	At5g61730	chrCnn_random:42474205..42477521	3317	14	1884	627	8.66	69730.62
BnABCA10-1	GSBRNA2T00113506001(BnaA03g40530D)	At5g61740	chrA03:20247381..20251552	4172	17	2664	887	8.4	98142.77
BnABCA10-2	GSBRNA2T00021503001(BnaC07g31480D)	At5g61740	chrC07:35551653..35555815	4163	17	2667	888	8.69	98308.09
BnABCA11-1	GSBRNA2T00024195001(BnaA01g21510D)	At5g61690	chrA01:13675327..13679595	4296	16	2748	915	8.58	101436.7
BnABCA11-2	GSBRNA2T00113510001(BnaA03g40490D)	At5g61690	chrA03:20229240..20234407	5168	16	2988	995	8.82	110139.6
BnABCA11-3	GSBRNA2T00021506001(BnaC07g31450D)	At5g61690	chrC07:35533466..35537948	4483	15	2814	937	7.19	103589.1
BnABCA12-1	GSBRNA2T00021507001(BnaC07g31440D)	At5g61700	chrC07:35530885..35532991	2107	7	1539	512		
BnABCA12-2	GSBRNA2T00113512001(BnaA03g40470D)	At5g61700	chrA03:20224785..20228058	3274	15	1980	659	8.99	73735.19
BnABCA12-3	GSBRNA2T00036435001(BnaCnng40510D)	At5g61700	chrCnn_random:39028966..39033265	4300	18	2610	869	8.67	96082.4
BnABCA12-4	GSBRNA2T00050456001(BnaC01g27420D)	At5g61700	chrC01:25043591..25046858	3268	15	1983	660	8.99	73897.41

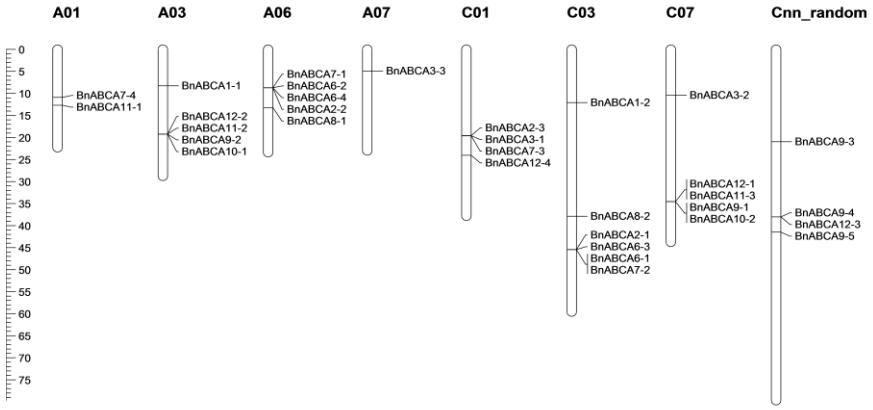

Fig. 2: The distribution of the *BnABCA* genes in the *B. napus* genome

The chromosomal position of each *BnABCA* gene was mapped according to the *B. napus* genome. The chromosome number is indicated at the top of each chromosome

The result showed that there were four (motif 3, 5, 8, 10) motifs highly conserved except for BnABCA3-1, BnABCA12-1, BnABCA2-3, BnABCA9-4. Except for the four genes mentioned above and the BnABCA9-5, the motif 2 was found in the rest, and if the BnABCA12-2, BnABCA12-4 and BnABCA9-2 were removed, the motif 1 was found in the rest.

Prediction about the Subcellular Localization

We performed prediction of BnABCAs subcellular localization by two different tools, TargetP 1.1 Server and WoLFPSORT. However, the results of the two tools are different from each other (Table 2).

3D Models of BnABCA9-1 to BnABCA9-5 Proteins

From all the given predicted models provided by phyre2, we chose the models that had the highest scores. The result showed that the 3D structures of BnABCA9-1, BnABCA9-2, BnABCA9-5 were different from the BnABCA9-3 and BnABCA9-4 (Fig. 5). Therefore, they

Table 2: Prediction about the subcellular localization of all BnABCA proteins

Name	cTP	mTP	SP	other	Loc	RC
BnABCA1-1	0.018	0.939	0.004	0.049	M	1
BnABCA1-2	0.012	0.93	0.011	0.04	M	1
BnABCA2-1	0.004	0.94	0.014	0.112	M	1
BnABCA2-2	0.005	0.929	0.012	0.119	M	1
BnABCA2-3	0.008	0.056	0.613	0.724	_	5
BnABCA3-1	0	0.017	0.982	0.334	S	2
BnABCA3-2	0.096	0.133	0.102	0.821	_	2
BnABCA3-3	0.081	0.135	0.102	0.829	_	2
BnABCA6-1	0.122	0.345	0.031	0.436	_	5
BnABCA6-2	0.077	0.378	0.036	0.473	_	5
BnABCA6-3	0.217	0.313	0.076	0.263	M	5
BnABCA6-4	0.186	0.304	0.078	0.281	M	5
BnABCA7-1	0.131	0.394	0.031	0.363	M	5
BnABCA7-2	0.131	0.404	0.034	0.353	M	5
BnABCA7-3	0.024	0.464	0.022	0.642	_	5
BnABCA7-4	0.064	0.451	0.026	0.48	_	5
BnABCA8-1	0.063	0.234	0.079	0.758	_	3
BnABCA8-2	0.058	0.23	0.083	0.754	_	3
BnABCA9-1	0.015	0.234	0.076	0.687	_	3
BnABCA9-2	0.112	0.073	0.158	0.606	_	3
BnABCA9-3	0.181	0.055	0.044	0.866	_	2
BnABCA9-4	0.047	0.088	0.256	0.875	_	2
BnABCA9-5	0.002	0.272	0.215	0.808	_	3
BnABCA10-1	0.02	0.587	0.043	0.555	M	5
BnABCA10-2	0.02	0.594	0.039	0.561	M	5
BnABCA11-1	0.019	0.563	0.094	0.231	M	4
BnABCA11-2	0.019	0.641	0.086	0.184	M	3
BnABCA11-3	0.019	0.612	0.08	0.191	M	3
BnABCA12-1	0.026	0.532	0.113	0.531	M	5
BnABCA12-2	0.044	0.308	0.094	0.832	_	3
BnABCA12-3	0.042	0.299	0.074	0.713	_	3
BnABCA12-4	0.044	0.308	0.094	0.832	_	3

were divided into two groups. We further focused on the expression regularity differences between them especially in lipid transport in seeds.

Expression Patterns of *BnABCAs* in Seed

Until now, only the characterization of *ABCA9* and the *ABCA2* in *A. thaliana* has been reported to connect with lipid transport and kin recognition, respectively. Because the HI has relationship with lipid transport, we focused on the expression of BnABCAs in seeds. We showed the expression patterns of *BnABCAs* by the RNA-seq data, which is from 3 different sources about different stage of seeds and leaves. To verify the accuracy of transcriptome sequencing, the *BnABCA9-1* and *BnABCA9-2* were selected to perform the qRT-PCR analysis in two different lines (P474-HO and P525-LO).

According to the RNA-sequence data, we found that the expression of *BnABCA1-1*, *BnABCA1-2*, *BnABCA2-1* and *BnABCA2-2* had high expression levels in the first week after flowering (1wp) and had low expressions after 2wp. The expressions *BnABCA1-1* and *BnABCA1-2* were upregulated after 2wp and reached the highest levels in 30 to 35 days after flowering (Fig. 6 and 7). The results showed that the expression levels of *BnABCA2-1* and *BnABCA2-2*

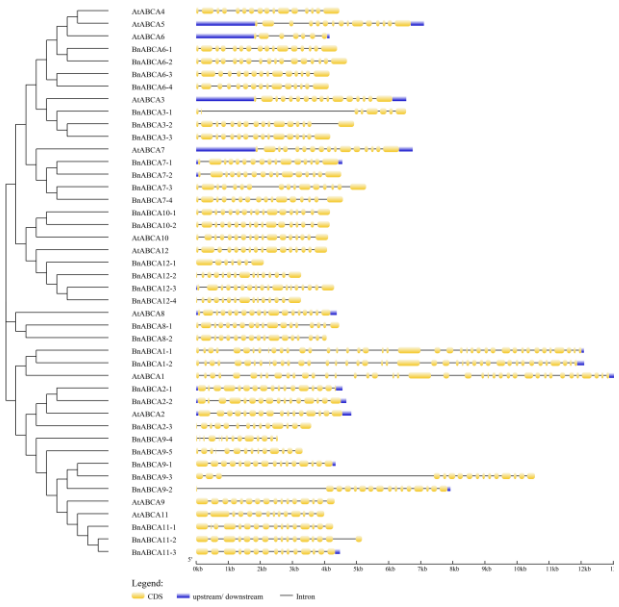

Fig. 3: The exon-intron structure of the *BnABCA* genes according to their phylogenetic relationships

An unrooted phylogenetic tree was constructed with 1000 bootstraps based on the full-length sequences of *BnABCA*. Exon-intron structure analyses of the *BnABCA* genes were performed using the online tool GSDS. The lengths of the exons and introns of each *BnABCA* gene are proportional. 11 segmental duplicates are highlighted by the red box

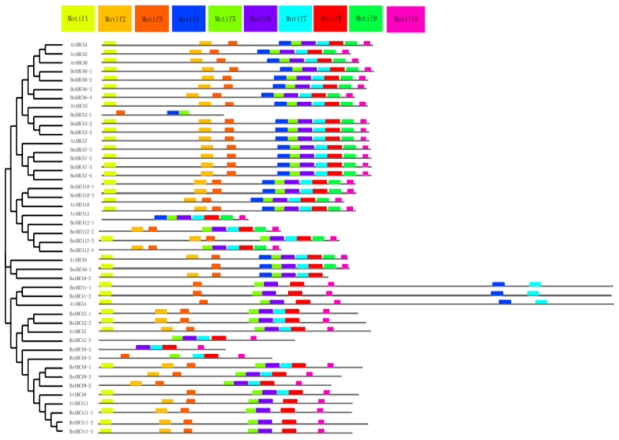

Fig. 4: The conserved motifs of the BnABCA proteins according to their phylogenetic relationships

The conserved motifs of the BnABCA proteins were identified by MEME. Gray lines represent the non-conserved sequences, and each motif is indicated by a colored box numbered at the top. The lengths of motifs in each protein are proportional

were low but sustainable upregulated after 1wp. The expression of *BnABCA6-3*, *BnABCA6-4*, *BnABCA7-1*, *BnABCA8-1*, *BnABCA8-2* showed that these five genes most likely played important roles in the last several weeks after flowering. However, the *BnABCA10-1* and *BnABCA10-2* mostly affected the early stage in the process

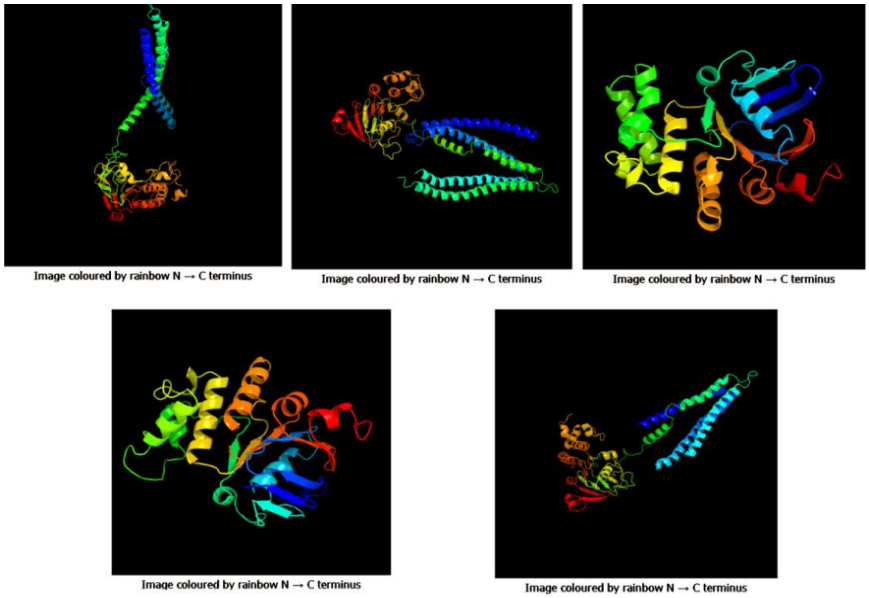

Fig. 5: 3D models of five selected BnABCA9 transporter proteins
Model were generated by Phyre2 server at normal mode. Figures were arranged in an order from BnABCA9-1 to BnABCA9-5

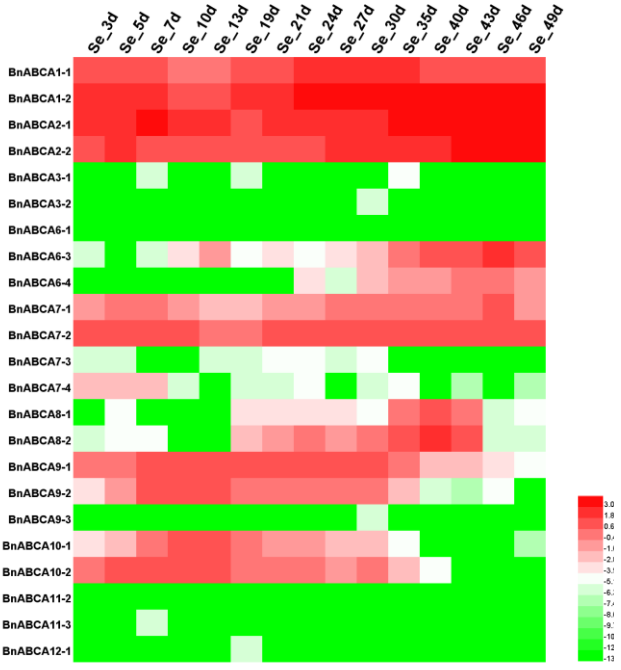

Fig. 6: The tissue-specific expression levels of *BnABCA* genes were obtained from unpublished RNA-Seq data from our laboratory
There were 15 different stage of seed after flowering time (Se_3d means the seeds of 3 days after flower and so on). Red indicates high expression and white indicates low expression

of seed development. The expression of *BnABCA7-2* was stable in the development of the seed and then was decreased after the 2wp and 3wp. Both *BnABCA9-1* and *BnABCA9-2* showed high expression levels after flowering 7 to 30 days, exactly the stage of lipid accumulation. Moreover, other genes showed no or very low expression level in seeds.

To further investigate the role of BnABCA9 in lipid biosynthesis, we performed qRT-PCR analysis on the *BnABCA9-1* and *BnABCA9-2* in the seeds of two different cultivars P474-HO (higher oil content in seeds) and P525-LO (lower oil content in seeds) under three different stages. The result showed that *BnABCA9.1* was obviously expressed highly in the 30-day seeds and 40-day seeds of P525-LO than the seeds of P474-HO. *BnABCA9.2* were also obviously higher expressed in all three stage seeds of P525-LO than P474-HO (Fig. 8). In *Arabidopsis*, AtABCA9 was known as the acyl-CoA transporter, which imported acyl-CoA into ER for TAG biosynthesis. Therefore, higher expression of *AtABCA9* caused higher TAG content in *Arabidopsis*. The result of *BnABCA9* in *B. napus* revealed that BnABCA9 may be involved in the lipid export from ER for further lipid degradation in peroxisome.

Discussion

Fatty acids are very important elements for lives, serving as storage lipids and precursors of signaling molecules and they are essential for the biosynthesis of cell membranes. Fatty acids are synthesized in plastids, and are further imported into ER for phospholipids and neutral lipids biosynthesis. Therefore, the import of fatty acids or acyl-CoA into ER is very necessary (Li-Beisson *et al.*, 2013). Lipid transporters play a vital role in oil accumulation. The members of ABC gene families have a highly conserved

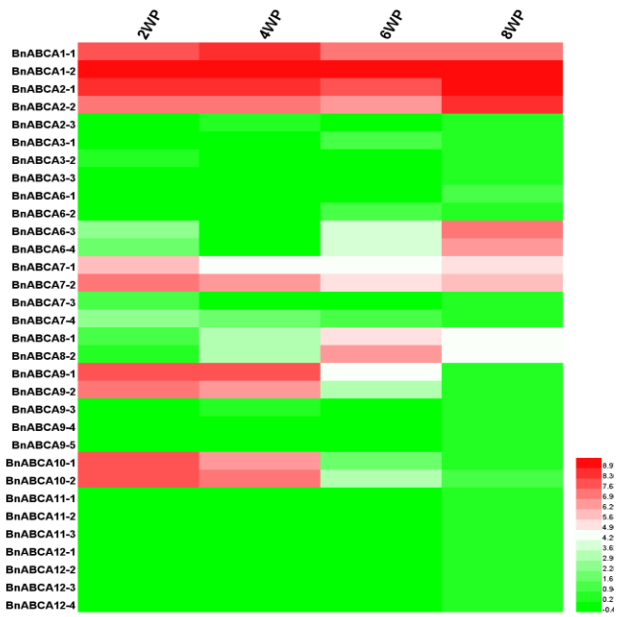

Fig. 7: The tissue-specific expression levels of *BnABCA* genes were obtained from the RNA-seq data

The tissue-specific expression levels of BnABCA genes in seeds of 4 different stages after pollination were examined (2WP, 4WP, 6WP and 8WP means that the seeds were 2, 4, 6 and 8 weeks after pollination, respectively). Red indicates high expression and white indicates low expression

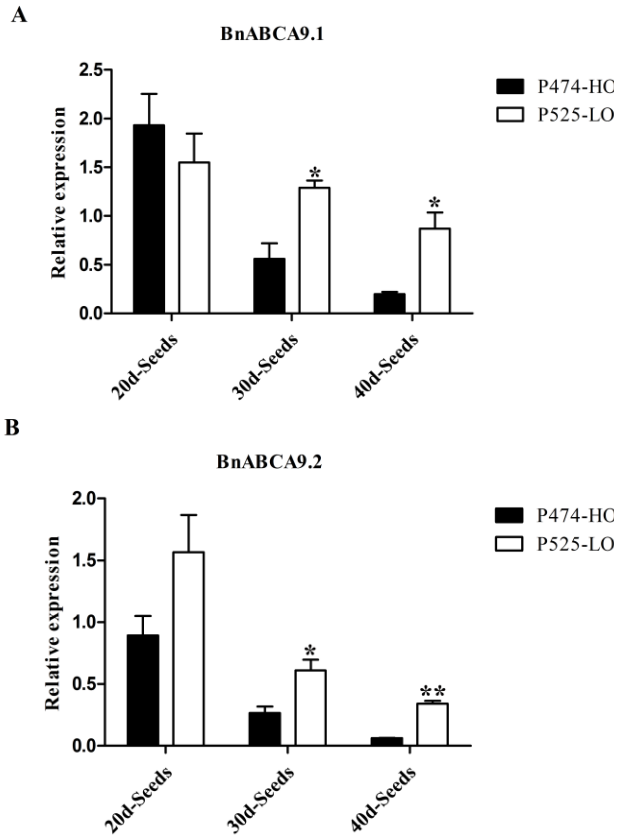

Fig. 8: The expression patterns of the *BnABCA9-1* and *BnABCA9-2* in 3 tissues of "P474-HO" and "P525-LO" were analyzed by qRT-PCR

20d-Seeds: the seeds after flowering 20 days; 30d-Seeds: the seeds after flowering 30 days; 40d-Seeds: the seeds after flowering 40 days

structures and functions, and the function of *AtABCA9* has been confirmed to be related with fatty acid and/or acyl-CoA transport into ER. There, the present study for the characterization of the *ABCA* gene family in *B. napus* is very important to investigate the role of ABCA proteins during lipid transport in *B. napus*. We hope that our result may be helpful to find the correlation between the ABCA genes and the lipid transport, and provide assistance for molecular breeding of high oil content.

As regards evolution of the BnABCA genes, most of homologous of *Arabidopsis* genes generally exist three copies in *B. napus* (Lysak *et al.*, 2005, 2007). However, due to the genome shrinkage and gene loss, the *B.rapa* and the *B. napus* only contain 1.5–2 and 2–6 copies of each homolog gene, respectively (Mun *et al.*, 2009; Chalhoub *et al.*, 2014). There are two different types of ABCs in plant, one is full-size (Group A) and the other is half-size (Group B). Molecular characterization also showed a large difference between group A and group B. Furthermore, the full-size members only have ABCA1 in plants. The structure of exon-intron also is important in evolution (Rogozin *et al.*, 2005; Xu *et al.*, 2012). In the present study, we find that most of the motifs of *BnABCA* members in group B were also appeared in group A, which may suggest that they have similar origin or the evolutionary patterns in different species. The study also showed that the member in group A had much more exons than the member in group B, which may indicate that the members of group A have more functions than group B. These data (including the analysis of proteins) showed the evolution of the *BnABCA* genes.

According to previous research, AtABCA9 protein mediates the transport of fatty acids and/or acyl-CoA into the ER and plays a vital role in the process of TAG biosynthesis (Kim *et al.*, 2013). We identified five homologous genes of *AtABCA9* in *B. napus* through evolutionary analysis. Besides, because of the widespread existence and the highly conserved function or structure of *ABC* family genes, we conjecture that the *BnABCA9s* have similar characterization with *AtABCA9* gene. From the expression patterns of *BnABCA9s* and related qRT-PCR results, we found that the stage of the highest expression of *BnABCA9-1* and *BnABCA9-2* in seeds was exactly the stage of the oil fastest accumulation in seed. Furthermore, the expressions of *BnABCA9-1* and *BnABCA9-2* appeared in seeds, not leaf and stalk. Therefore, we believe that the *BnABCA9-1* and the *BnABCA9-2* have a close connection with fatty acids and/or acyl-CoA transport for lipid biosynthesis in seeds of *B. napus*. Interestingly, *BnABCA9-1* and *BnABCA9-2* showed higher expression in the lower-oil content cultivar P525-LO than the higher oil-content cultivar P474-HO. This opposite phenotype with the expression of *AtABCA9* in *Arabidopsis*

revealed the BnABCA9 may be involved in the acyl-CoA export from ER or lipid degradation in *B. napus*. Besides, our result also revealed that the expressions of *BnABCA1-1*, *BnABCA1-2*, *BnABCA2-1*, *BnABCA2-2*, *BnABCA7-1*, *BnABCA7-2*, *BnABCA10-1* and *BnABCA10-2* may have connection with fatty acids and/or lipid accumulation in seeds based on the RNA-seq data. These results may contribute to our understanding of fatty acid and/or lipid transport in *B. napus*.

Conclusion

In the present study, we identified 32 *ABCA* genes (*BnABCAs*) from *B. napus*. These genes were further used for the analysis of gene structures, chromosomal distribution, phylogenetic relationships and motifs analysis. And the subcellular localization and 3D models of partial identified *BnABCA* members were also predicted. Furthermore, the expression patterns of most *BnABCA* genes in seeds of different stages were investigated by RNA-Seq. The result shows that *BnABCA9-1* and *BnABCA9-2* were the two genes mainly expressed in the seeds during seed oil accumulation stages. The expression levels of these two genes in seeds of different stages were then verified by qRT-PCR assays. The results reveal that BnABCA9 may be involved in the acyl-CoA export from ER or the lipid degradation process, which is different from the role of ABCA9 in *Arabidopsis*. The present study is important to further characterize the function of *BnABCA* gene family during seed oil biosynthesis of *B. napus*.

Acknowledgements

This work was supported by National Key R & D Program of China (2018YFD0200903), National Natural Science Foundation of China (31400063; 31500038) and Fundamental Research Funds for the Central Universities (XDJK2017B030; SWU116021; XDJK2018C095; SWU118114; SWU115018), Research Funds of Scientific Platform and Base Construction (cstc2014pt-sy0017) and The Recruitment Program for Foreign Experts (No. WQ20125500073).

Reference

Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D.J. Lipman, 1997. Gapped blast and psi-blast: A new generation of protein database search programs. *Nucl. Acid. Res.*, 25: 3389–3402

Bailey, T.L., N. Williams, C. Misleh and W.W. Li, 2006. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucl. Acid. Res.*, 34(Web Server issue)**,** W369–W373

Bird, D., F. Beisson, A. Brigham, J. Shin, S. Greer, R. Jetter, L. Kunst, X.W. Wu, A. Yephremov and L. Samuels, 2007. Characterization of *Arabidopsis* abcg11/wbc11, an atp binding cassette (abc) transporter that is required for cuticular lipid secretion. *Plant J.*, 52: 485–498

Bowers, J.E., B.A. Chapman, J. Rong and A.H. Paterson, 2003. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature*, 422: 433–438

Chalhoub, B., F. Denoeud, S. Liu, I.A. Parkin, H. Tang, X. Wang, J. Chiquet, H. Belcram, C. Tong, B. Samans, M. Correa, C. Da Silva, J. Just, C. Falentin, C.S. Koh, I. Le Clainche, M. Bernard, P. Bento, B. Noel, K. Labadie, A. Alberti, M. Charles, D. Arnaud, H. Guo, C. Daviaud, S. Alamery, K. Jabbari, M. Zhao, P.P. Edger, H. Chelaifa, D. Tack, G. Lassalle, I. Mestiri, N. Schnel, M.C. Le Paslier, G. Fan, V. Renault, P.E. Bayer, A.A. Golicz, S. Manoli, T.H. Lee, V.H. Thi, S. Chalabi, Q. Hu, C. Fan, R. Tollenaere, Y. Lu, C. Battail, J. Shen, C.H. Sidebottom, X. Wang, A. Canaguier, A. Chauveau, A. Berard, G. Deniot, M. Guan, Z. Liu, F. Sun, Y.P. Lim, E. Lyons, C.D. Town, I. Bancroft, X. Wang, J. Meng, J. Ma, J.C. Pires, G.J. King, D. Brunel, R. Delourme, M. Renard, J.M. Aury, K.L. Adams, J. Batley, R.J. Snowdon, J. Tost, D. Edwards, Y. Zhou, W. Hua, A.G. Sharpe, A.H. Paterson, C. Guan and P. Wincker, 2014. Plant genetics. Early allopolyploid evolution in the post-neolithic brassica napus oilseed genome. *Science*, 345: 950–953

Emanuelsson, O., S. Brunak, G. von Heijne and H. Nielsen, 2007. Locating proteins in the cell using targetp, signalp and related tools. *Nat. Protoc.*, 2: 953–971

Gasteiger, E., A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel and A. Bairoch, 2003. Expasy: The proteomics server for in-depth protein knowledge and analysis. *Nucl. Acid. Res.*, 31: 3784–3788

Higgins, C.F., 1992. Abc transporters: From microorganisms to man. *Annu. Rev. Cell Biol.*, 8: 67–113

Horton, P., K.J. Park, T. Obayashi, N. Fujita, H. Harada, C.J. Adams-Collier and K. Nakai, 2007. Wolf psort: Protein localization predictor. *Nucl. Acid. Res.*, 35: W585–W587

Jones, P.M. and A.M. George, 2004. The abc transporter structure and mechanism: Perspectives on recent research. *Cell Mol. Life Sci.*, 61: 682–699

Kelley, L.A. and M.J. Sternberg, 2009. Protein structure prediction on the web: A case study using the phyre server. *Nat. Protoc.*, 4: 363–371

Kim, S., Y. Yamaoka, H. Ono, H. Kim, D. Shim, M. Maeshima, E. Martinoia, E.B. Cahoon, I. Nishida and Y. Lee, 2013. Atabca9 transporter supplies fatty acids for lipid synthesis to the endoplasmic reticulum. *Proc. Natl. Acad. Sci. USA*, 110: 773–778

Kim, W.S., C.S. Weickert and B. Garner, 2008. Role of atp-binding cassette transporters in brain lipid transport and neurological disease. *J. Neurochem.*, 104: 1145–1166

Kumar, S., G. Stecher and K. Tamura, 2016. Mega7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evolution*, 33: 1870–1874

Kuromori, T., T. Miyaji, H. Yabuuchi, H. Shimizu, E. Sugimoto, A. Kamiya, Y. Moriyama and K. Shinozaki, 2010. Abc transporter atabcg25 is involved in abscisic acid transport and responses. *Proc. Natl. Acad. Sci. USA*, 107: 2361–2366

Lamesch, P., T.Z. Berardini, D. Li, D. Swarbreck, C. Wilks, R. Sasidharan, R. Muller, K. Dreher, D.L. Alexander, M. Garcia-Hernandez, A.S. Karthikeyan, C.H. Lee, W.D. Nelson, L. Ploetz, S. Singh, A. Wensel and E. Huala, 2012. The *Arabidopsis* information resource (tair): Improved gene annotation and new tools. *Nucl. Acid. Res.*, 40 (Database issue): D1202–D1210

Le Hir, R., C. Sorin, D. Chakraborti, T. Moritz, H. Schaller, F. Tellier, S. Robert, H. Morin, L. Bako and C. Bellini, 2013. Abcg9, abcg11 and abcg14 abc transporters are required for vascular development in *Arabidopsis*. *Plant J.*, 76: 811–824

Li-Beisson, Y., B. Shorrosh, F. Beisson, M.X. Andersson, V. Arondel, P.D. Bates, S. Baud, D. Bird, A. Debono, T.P. Durrett, R.B. Franke, I.A. Graham, K. Katayama, A.A. Kelly, T. Larson, J.E. Markham, M. Miquel, I. Molina, I. Nishida, O. Rowland, L. Samuels, K.M. Schmid, H. Wada, R. Welti, C. Xu, R. Zallot and J. Ohlrogge, 2013. Acyl-lipid metabolism. *The Arabidopsis Book*, 11: e0161

Lu, K., W. Guo, J. Lu, H. Yu, C. Qu, Z. Tang, J. Li, Y. Chai and Y. Liang, 2015. Genome-wide survey and expression profile analysis of the mitogen-activated protein kinase (mapk) gene family in brassica rapa. *PloS One*, 10: e0132051

Luo, B., X.Y. Xue, W.L. Hu, L.J. Wang and X.Y. Chen, 2007. An abc transporter gene of *Arabidopsis* thaliana, atwbc11, is involved in cuticle development and prevention of organ fusion. *Plant Cell Physiol.*, 48: 1790–1802

Luo, X., C. Ma, Y. Yue, K. Hu, Y. Li, Z. Duan, M. Wu, J. Tu, J. Shen, B. Yi and T. Fu, 2015. Unravelling the complex trait of harvest index in rapeseed (*Brassica napus* l.) with association mapping. *BMC Genom.*, 16: 379

Lysak, M.A., K. Cheung, M. Kitschke and P. Bures, 2007. Ancestral chromosomal blocks are triplicated in brassiceae species with varying chromosome number and genome size. *Plant Physiol.*, 145: 402–410

Lysak, M.A., M.A. Koch, A. Pecinka and I. Schubert, 2005. Chromosome triplication found across the tribe brassiceae. *Genom. Res.*, 15: 516–525

Mun, J.H., S.J. Kwon, T.J. Yang, Y.J. Seol, M. Jin, J.A. Kim, M.H. Lim, J.S. Kim, S. Baek, B.S. Choi, H.J. Yu, D.S. Kim, N. Kim, K.B. Lim, S.I. Lee, J.H. Hahn, Y.P. Lim, I. Bancroft and B.S. Park, 2009. Genome-wide comparative analysis of the *Brassica rapa* gene space reveals genome shrinkage and differential loss of duplicated genes after whole genome triplication. *Genom. Biol.*, 10: R111

Oram, J.F. and A.M. Vaughan, 2006. Atp-binding cassette cholesterol transporters and cardiovascular disease. *Circul. Res.*, 99: 1031–1043

Panikashvili, D., S. Savaldi-Goldstein, T. Mandel, T. Yifhar, R.B. Franke, R. Hofer, L. Schreiber, J. Chory and A. Aharoni, 2007. The *Arabidopsis* desperado/atwbc11 transporter is required for cutin and wax secretion. *Plant Physiol.*, 145: 1345–1360

Panikashvili, D., J.X. Shi, L. Schreiber and A. Aharoni, 2011. The *Arabidopsis* abcg13 transporter is required for flower cuticle secretion and patterning of the petal epidermis. *New Phytol.*, 190: 113–124

Pighin, J.A., H.Q. Zheng, L.J. Balakshin, I.P. Goodman, T.L. Western, R. Jetter, L. Kunst and A.L. Samuels, 2004. Plant cuticular lipid export requires an abc transporter. *Science*, 306: 702–704

Rea, P.A., 2007. Plant atp-binding cassette transporters. *Annu. Rev. Plant Biol.*, 58: 347–375

Rogozin, I.B., A.V. Sverdlov, V.N. Babenko and E.V. Koonin, 2005. Analysis of evolution of exon-intron structure of eukaryotic genes. *Brief Bioinform.*, 6: 118–134

Sanchez-Fernandez, R., T.G. Davies, J.O. Coleman and P.A. Rea, 2001. The *Arabidopsis thaliana* abc protein superfamily, a complete inventory. *J. Biol. Chem.*, 276: 30231–30244

Tamura, K., G. Stecher, D. Peterson, A. Filipski and S. Kumar, 2013. Mega6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evolut.* 30: 2725–2729

van Meer, G., D.R. Voelker and G.W. Feigenson, 2008. Membrane lipids: Where they are and how they behave. *Nat. Rev. Mol. Cell. Biol.*, 9: 112–124

Verrier, P.J., D. Bird, B. Buria, E. Dassa, C. Forestier, M. Geisler, M. Klein, U. Kolukisaoglu, Y. Lee, E. Martinoia, A. Murphy, P.A. Rea, L. Samuels, B. Schulz, E.P. Spalding, K. Yazaki and F.L. Theodoulou, 2008. Plant abc proteins - a unified nomenclature and updated inventory. *Trends Plant Sci.*, 13: 151–159

Vision, T.J., D.G. Brown and S.D. Tanksley, 2000. The origins of genomic duplications in *Arabidopsis. Science*, 290: 2114–2117

Xu, G., C. Guo, H. Shan and H. Kong, 2012. Divergence of duplicate genes in exon-intron structure. *Proc. Natl. Acad. Sci. USA*, 109: 1187–1192

Yadav, V., I. Molina, K. Ranathunge, I.Q. Castillo, S.J. Rothstein and J.W. Reed, 2014. Abcg transporters are required for suberin and pollen wall extracellular barriers in *Arabidopsis*. *Plant Cell*, 26: 3569–3588

Zdobnov, E.M. and R. Apweiler, 2001. Interproscan - an integration platform for the signature-recognition methods in interpro. *Bioinformatics*, 17: 847–848

Zhao, G.C., J.X. Shi, W.Q. Liang, F.Y. Xue, Q. Luo, L. Zhu, G.R. Qu, M.J. Chen, L. Schreiber and D.B. Zhang, 2015. Two atp binding cassette g transporters, rice atp binding cassette g26 and atp binding cassette g15, collaboratively regulate rice male reproduction. *Plant Physiol.*, 169: 2064–2079

(Received 22 August 2018; Accepted 20 October 2018)