

RESEARCH ARTICLE

Open Access



Genome-wide characterization and expression profiling of *SWEET* genes in cabbage (*Brassica oleracea* var. *capitata* L.) reveal their roles in chilling and clubroot disease responses

Wei Zhang¹, Shenyun Wang¹, Fangwei Yu¹, Jun Tang¹, Xi Shan², Kan Bao³, Li Yu¹, Hong Wang¹, Zhangjun Fei^{3*} and Jianbin Li^{1*}

Abstract

Background: The *SWEET* proteins are a group of sugar transporters that play a role in sugar efflux during a range of biological processes, including stress responses. However, there has been no comprehensive analysis of the *SWEET* family genes in *Brassica oleracea* (*BoSWEET*), and the evolutionary pattern, phylogenetic relationship, gene characteristics of *BoSWEET* genes and their expression patterns under biotic and abiotic stresses remain largely unexplored.

Results: A total of 30 *BoSWEET* genes were identified and divided into four clades in *B. oleracea*. Phylogenetic analysis of the *BoSWEET* proteins indicated that clade II formed first, followed by clade I, clade IV and clade III, successively. Clade III, the newest clade, shows signs of rapid expansion. The *Ks* values of the orthologous *SWEET* gene pairs between *B. oleracea* and *Arabidopsis thaliana* ranged from 0.30 to 0.45, which estimated that *B. oleracea* diverged from *A. thaliana* approximately 10 to 15 million years ago. Prediction of transmembrane regions showed that eight *BoSWEET* proteins contain one characteristic MtN3_slv domain, twenty-one contain two, and one has four. Quantitative reverse transcription-PCR (qRT-PCR) analysis revealed that five *BoSWEET* genes from clades III and IV exhibited reduced expression levels under chilling stress. Additionally, the expression levels of six *BoSWEET* genes were up-regulated in roots of a clubroot-susceptible cabbage cultivar (CS-JF1) at 7 days after inoculation with *Plasmodiophora brassicae* compared with uninoculated plants, indicating that these genes may play important roles in transporting sugars into sink roots associated with *P. brassicae* colonization in CS-JF1. Subcellular localization analysis of a subset of *BoSWEET* proteins indicated that they are localized in the plasma membrane.

Conclusions: This study provides important insights into the evolution of the *SWEET* gene family in *B. oleracea* and other species, and represents the first study to characterize phylogenetic relationship, gene structures and expression patterns of the *BoSWEET* genes. These findings provide new insights into the complex transcriptional regulation of *BoSWEET* genes, as well as potential candidate *BoSWEET* genes that promote sugar transport to enhance chilling tolerance and clubroot disease resistance in cabbage.

Keywords: Cabbage (*Brassica oleracea* var. *capitata* L.), Expression profile, RNA-Seq, Stress response, *SWEET*

* Correspondence: zf25@cornell.edu; jbli@jaas.ac.cn

³Boyce Thompson Institute, Cornell University, Ithaca, NY 14853, USA

¹Jiangsu Key Laboratory for Horticultural Crop Genetic Improvement, Institute of Vegetable Crops, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, People's Republic of China

Full list of author information is available at the end of the article



Background

Sugars as essential energy sources, are synthesized in leaves (source organs) and then translocated via phloem sap into sink organs, such as modified leaves, roots, seeds and fruits, and the amount of sugars influences plant development [1, 2]. In plants, sugar transport is mediated by proteins in the sucrose transporter (SUT) and monosaccharide transporter (MST) and SWEET families [3, 4]. SWEET, a novel class of sugar transporters, is a distinct transporter family which mediates influx or efflux of sugars from phloem parenchyma into the phloem apoplast [5–7]. In prokaryotes SWEET proteins contain only three transmembrane helices (TMHs), while in eukaryotes there are also examples with seven TMHs. The seven-TMHs has evolved with two tandemly repeated three-TMH units separated by a single TMH [8]. Phylogenetic analysis of SWEET proteins has shown that they can be divided into four clades [3], with clades I and II preferentially transporting hexoses, clade III transporting sucrose, and clade IV being responsible for the flux of fructose across the tonoplast [9–12].

With the rapid development of whole-genome sequencing, genome-wide identification of *SWEET* genes in various species have been reported, such as in *Arabidopsis thaliana* [3], rice (*Oryza sativa*) [13], sorghum (*Sorghum bicolor*) [14], soybean (*Glycine max*) [15], apple (*Malus domestica*) [16], grape (*Vitis vinifera*) [17], banana (*Musa acuminata*) [18], tomato (*Solanum lycopersicum*) [19] and rapeseed (*Brassica napus*) [20]. Biochemical and functional analyses have shown that *SWEET* genes play significant roles in various physiological processes, such as nectar secretion [21, 22], seed and pollen development [23, 24], leaf senescence [25], and responses to abiotic [12, 26–28] and biotic stresses [3, 17, 29–31]. For example, *AtSWEET8* is essential for pollen viability in *A. thaliana*, and decreased expression reduces starch content in pollen grains and causes male sterility [3]. Moreover, the *SWEET* genes are critical for carbon transport regulation in host-pathogen interactions and have been shown to be targets of extracellular pathogens [30]. Induction of *SWEET* genes upon pathogen infection has also been reported in alfalfa (*Medicago sativa*), rice, grape, *A. thaliana* and Chinese cabbage (*Brassica rapa*) [29–34]. For example, *MtN3*, the first identified member of the *SWEET* family, was found to participate in the host-*Rhizobium meliloti* interaction in alfalfa [32]. *OsSWEET11*, *OsSWEET13* and *OsSWEET14* were later shown to be associated with resistance to bacterial blight in rice [29, 31, 33]. In grape, the expression of *VvSWEET4* increased after *Botrytis cinerea* infection [17].

Clubroot disease is a soil-borne disease caused by the obligate biotrophic pathogen *P. brassicae* and is one of the most devastating diseases in Brassicaceae plants, reducing both crop quality and yields [35, 36]. The life cycle of *P. brassicae* consists of three stages: the survival stage of resting spores in the soil, the primary

infection (root hair infection) stage and the secondary infection (root cortex infection) stage [37–39]. The resting spores can survive in the soil for 6–12 years, making this disease hard to control once the soil has been contaminated [40]. In *A. thaliana* after *P. brassicae* infection the expression of *AtSWEET15* was strongly induced during gall formation, and the *atsweet11* mutant exhibited slower gall formation compared to wild-type plants [30, 34]. In Chinese cabbage, the expression of several *BrSWEET* genes from Clade I and III increased, as did glucose and fructose levels, in roots of a clubroot-susceptible line compared to a clubroot-resistant line following *P. brassicae* infection, suggesting a close relationship between *P. brassicae* growth, sugar translocation and the expression of *BrSWEET* genes [30].

Cold stress (CS), including chilling (< 20 °C) and freezing (< 0 °C), has a major impact on plant growth and development, limiting geographic distribution and productivity [41]. It has been long established that accumulation of soluble sugars can stabilize cellular components and membranes following CS [42, 43]. Overexpression of *AtSWEET4* has been shown to increase plant size and freezing tolerance in *A. thaliana* [28], while cold-stressed *AtSWEET16* overexpressing lines are unable to accumulate fructose and have increased tolerance to freezing stress [26]. The fructose-specific transporter *AtSWEET17* plays a primary role in fructose homeostasis following 1 week of 4 °C treatment [12]. Interestingly, the double mutant *atsweet11/atsweet12* was reported to release fewer electrolytes when the temperature was reduced to 4 °C, but exhibited greater freezing tolerance than both single mutants and wild-type *A. thaliana* [27].

Cabbage (*Brassica oleracea* var. *capitata* L.), belonging to the Brassicaceae family, is one of the most economically important leafy vegetable crops worldwide. Chilling and clubroot disease cause severe losses of yields and quality in this species, as well as in other Brassicaceae crops. Even though they have been associated with responses to chilling and clubroot disease in other species, little is known about the role of the *SWEET* sugar transporters in chilling and clubroot disease responses in cabbage. The objectives of this study were to conduct a genome-wide analysis of the *SWEET* gene family in *B. oleracea* and thirteen other species, and to develop a better understanding of the molecular evolution and function of the *SWEET* proteins in cabbage, while also providing a reference for other Brassicaceae species.

Methods

Identification of *SWEET* family genes in *B. oleracea* and thirteen other plant species

The *B. oleracea* whole-genome sequence used to identify the *BoSWEET* genes was downloaded from the *B.*

oleracea Genome Database (Bolbase, <http://ocri-genomics.org/bolbase/>) [44]. The amino acid sequences of the *A. thaliana* SWEET genes were retrieved from the TAIR database (<http://www.arabidopsis.org/>). The *Brassica rapa* whole-genome sequence was obtained from the BRAD database (<http://brassicadb.org/brad/>) [45]. The *Carica papaya*, *Populus trichocarpa*, *Vitis vinifera*, *Oryza sativa*, *Zea mays*, *Selaginella moellendorffii*, *Physcomitrella patens*, *Chlamydomonas reinhardtii*, *Volvox carteri*, *Ostreococcus lucimarinus* and *Ostreococcus tauri* gene information was downloaded from the Plaza v2.5 database (<http://bioinformatics.psb.ugent.be/plaza/news/index>) [46]. The Hidden Markov Model (HMM) corresponding to the MtN3/saliva (MtN3_slv) domain (PF03083) was retrieved from Pfam 31.0 (<http://pfam.xfam.org/>) and used to identify putative SWEET proteins with the “trusted cutoff” as the threshold [47, 48]. AtSWEET protein sequences were used as the seed sequences to carry out a BLASTP search in the sequences from the other species with an *E*-value threshold of $1e^{-10}$. The Pfam 31.0 database, the SMART database (<http://smart.embl-heidelberg.de/>) and the Conserved Domain Database (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi/>) were then used to further filter and analyze the potential SWEET protein sequences to validate the HMM and BLASTP search results [49, 50]. The SWEET protein sequences from *B. oleracea* and *B. rapa* were named by adding a suffix (a, b, c...etc.) based on sequence similarity to the corresponding AtSWEET proteins.

Phylogenetic analysis and characterization of BoSWEET proteins

The ProtParam tool (<https://web.expasy.org/protparam/>) was used to analyze the physical and chemical parameters of the BoSWEET proteins, including molecular weight and theoretical pI. The Gene Structure Display Server (GSDS, <http://gsds.cbi.pku.edu.cn/index.php>) was utilized to draw a schematic diagram of the gene structure according to the genomic sequences and the corresponding coding sequence of each *BoSWEET* gene [51]. The online MEME tool (<http://meme-suite.org/tools/meme>) was used to identify conserved protein motifs using default parameters [52]. The TMHs of BoSWEET proteins were predicted by the TMHMM Server v.2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>). The Conserved Domain Architecture Retrieval Tool (CDART) (<http://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi>) was used to draw the MtN3_slv domains [53]. The SignalP 4.1 server (<http://www.cbs.dtu.dk/services/SignalP/>) was used to predict the presence and location of signal peptide cleavage sites in the amino acid sequences. Multiple alignment of all the BoSWEET proteins was performed using ClustalW [54], and the

phylogenetic tree was constructed by MEGA7 with the bootstrap of 1000 replicates using the neighbor-joining (NJ) method [55].

Chromosomal localization of BoSWEET genes, and identification of orthologs and paralogs

The chromosomal localization of the *BoSWEET* genes was determined using MapChart 2.30 [56]. OrthoMCL (<http://orthomcl.org/orthomcl/>) was used to identify the orthologs and paralogs of the SWEET proteins in *B. oleracea* and *A. thaliana*. The relationships of orthologs and paralogs were plotted using the Circos software [57].

The occurrence of duplication events and divergence time of orthologous genes, as well as the selective pressure on duplicated genes, was estimated by calculating *Ks* (synonymous substitution rate) and *Ka* (nonsynonymous substitution rate) values using DnaSP 6 [58]. The *Ks* values of all the syntenic orthologs of the *SWEET* genes between *B. oleracea* and *A. thaliana* were then plotted as the density using an R package [59]. The divergence time was calculated using the formula, $T = Ks/2r$, with the value of *r* being 1.5×10^{-8} synonymous substitutions per site per year for dicotyledonous plants [60].

Chilling stress, sample collection and quantitative reverse transcription-PCR (qRT-PCR)

To investigate *BoSWEET* expression profiles in response to chilling stress, the advanced inbred cold-tolerance cabbage line 923 (CT-923) was used. Seeds were grown in sterilized soil in a growth chamber at 25 °C day /18 °C night, with a photoperiod of 14 h light/10 h dark. After 4 weeks, to induce chilling stress, the seedlings were maintained at 4 °C for 0, 3, 6, 12, 24 and 48 h. Next, samples were collected and immediately frozen in liquid nitrogen and stored at -80 °C till RNA extraction. For quantitative gene expression analysis, total RNA was extracted using the TaKaRa MiniBEST Plant RNA Extraction Kit according to the manufacturer's instructions (Takara Bio Inc., Dalian, China). The first-strand cDNA was synthesized using the PrimeScript™ RT reagent Kit with gDNA Eraser (TakaRa). The qRT-PCR reactions were performed using TB Green™ Premix Ex Taq™ II (TakaRa) and carried out on a Roche LightCycler® 480II PCR System. Gene specific primers were designed using Beacon Designer 7.7 (Premier Biosoft, CA, USA) and are listed in Additional file 1: Table S1. *BoActin2* was used as the reference gene [61]. All reactions were performed in triplicate, and the $2^{-\Delta\Delta CT}$ method was applied to calculate the relative expression [62]. Duncan's multiple range test at $P < 0.05$ was used to determine the significance level of the data, using the SPSS 21 software (SPSS Inc., USA).

RNA-Seq data analysis of *BoSWEET* genes

To analyze the *BoSWEET* expression profiles in different organs and in response to *P. brassicae* infection, RNA-Seq data from the GSE42891 (GEO database) and PRJNA453960 (BioProject accession) were downloaded from NCBI. The GSE42891 RNA-Seq data contained expression profiles of seven different organs/tissues (bud, callus, flower, leaf, root, silique and stem) of the cabbage homozygous line 02–12 [44]. The PRJNA453960 RNA-Seq data contained the root expression profiles from clubroot-resistant cabbage Xiangan336 (CR-XG336) and clubroot-susceptible cabbage Jingfeng No.1 (CS-JF1) at 7 (primary infection stage) and 28 (clubroot formation stage) days after inoculation (DAI) with *P. brassicae*. The corresponding root samples at 7 d and 28 d without inoculation were sampled as the mock control, and all the samples were collected in three biological replicates. Raw RNA-Seq reads were processed to trim the adapter and low-quality sequences using Trimmomatic [63]. The high-quality cleaned reads were aligned to the *B. oleracea* genome using HISAT [64] allowing up to 3 edit distances. Following alignments, raw counts for each gene were derived and normalized into FPKM (fragments per kilobase of exon model per million mapped reads). Raw count data was then fed to DESeq2 [65] to identify differentially expressed genes with a cut off of fold change > 2 and FDR < 0.05. Heatmap of *BoSWEET* gene expression profiles was generated using the pheatmap package (<https://cran.r-project.org/web/packages/pheatmap/>) based on the \log_2 transformed FPKM values.

Construction of *BoSWEET* transient expression vectors and subcellular localization studies in tobacco

To investigate the subcellular localization of the *BoSWEET* proteins, they were transiently expressed as translational GFP (green fluorescent protein) fusion proteins in tobacco (*Nicotiana benthamiana*) leaf epidermal cells. The full-length coding sequences of *BoSWEET11b*, *BoSWEET11c* and *BoSWEET12b* were amplified using a forward primer containing a *Kpn* I restriction site and a reverse primer containing a *Xba* I restriction site. The primers used are listed in Additional file 1: Table S1. The amplification products were digested with *Kpn* I and *Xba* I and ligated into the pCAMBIA1300-35S-GFP (35S-GFP) vector. The recombined plasmids were then transformed into *Agrobacterium tumefaciens* strain GV3101 [66]. *Agrobacterium* transient expression and infiltration was carried out according to previously published protocols [67]. Leaves transformed with the 35S-GFP vector alone were used as controls. Two days after infiltration, fluorescence and bright-light images of transiently infected tobacco leaves were obtained using a fluorescence microscope (BX41, Olympus, Rungis, France).

Results

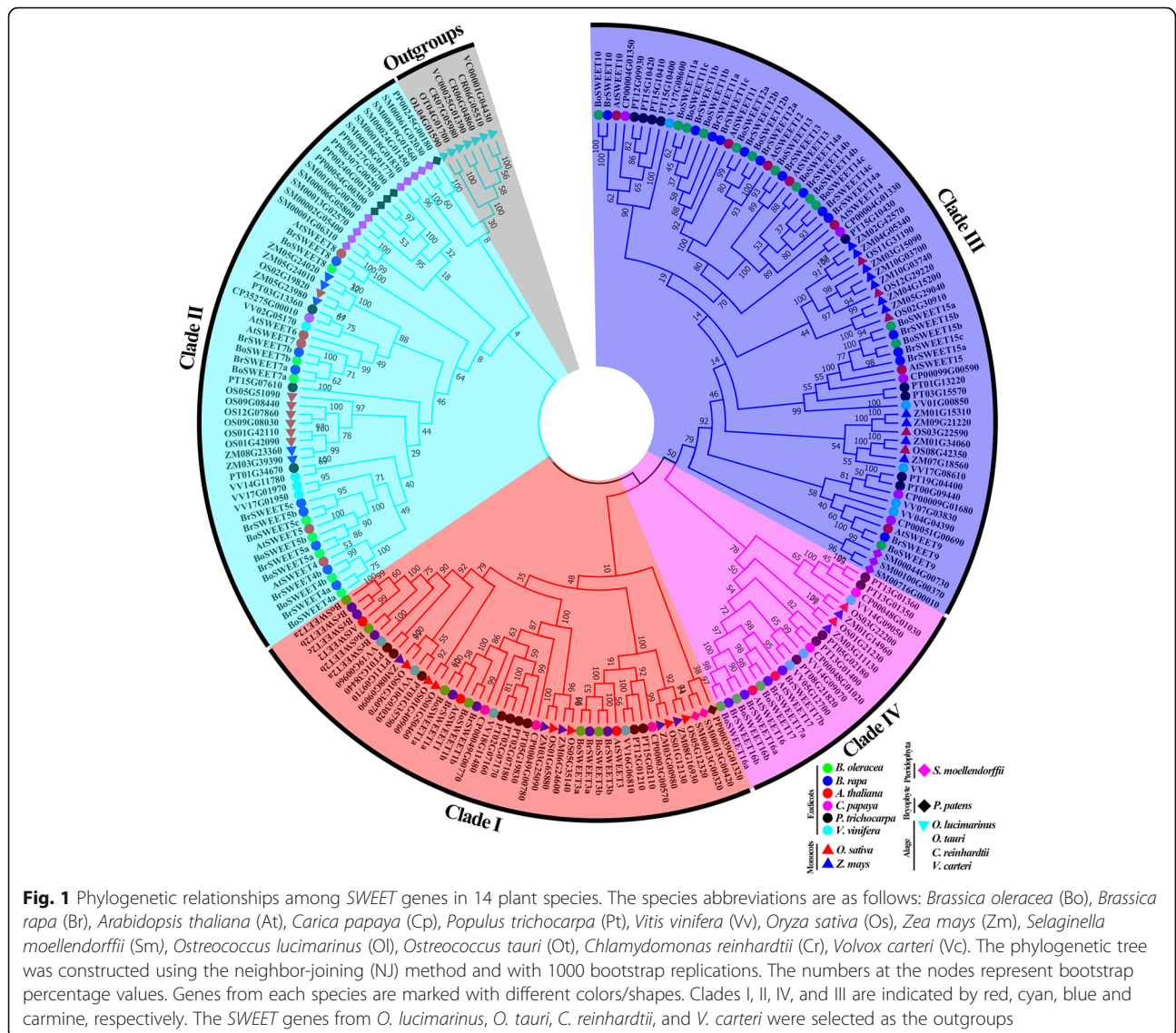
Evolutionary history of *SWEET* genes among plant species

All putative *SWEET* genes were identified in *B. oleracea* and thirteen other representative plant species (Additional file 2: Table S2). A phylogenetic tree of the identified 205 *SWEET* genes from the 14 species was constructed in order to investigate the evolutionary history of the family in the plant kingdom, with genes from the four algal species selected as the outgroups (Fig. 1, Additional file 2: Table S2). All other genes in the tree were clustered into four clades (clade I, clade II, clade III and clade IV) that were named according to the previously reported *A. thaliana* nomenclature [3]. According to the evolutionary distance with the outgroups, clade II formed first, followed by clade I, clade IV and clade III, with the latter expanding most rapidly (Fig. 1). The 30 *BoSWEET* and 33 *BrSWEET* genes identified in this study were named sequentially from *BoSWEET1* to *BoSWEET17* and *BrSWEET1* to *BrSWEET17*, respectively, according to their *A. thaliana* homologs.

The phylogenetic relationships between the genes in the four clades were investigated (Fig. 2). In total of 14 plant species included dicotyledonous species (*B. oleracea*, *B. rapa*, *A. thaliana*, *Carica papaya*, *P. trichocarpa* and *V. vinifera*), monocotyledonous species (*O. sativa* and *Z. mays*), a pteridophyte (*S. moellendorffii*), a bryophyte (*P. patens*) and algae (*C. reinhardtii*, *V. carteri*, *O. lucimarinus* and *O. tauri*). In an evolutionary context, examples of genes in Clade II were first observed in the four algal species (unicellular chlorophyta), which contained fewer *SWEET* members (1–3), and clade I and clade IV were predominantly observed in *P. patens* (bryophyte) and *S. moellendorffii* (pteridophyta), respectively. Clade III was first formed in monocots. In eudicots, all six selected species underwent the γ triplication event, and the larger number of *SWEET* genes in *P. trichocarpa* than in *C. papaya* and *V. vinifera* is likely due to the salicoid-specific genome duplication [68]. As consequence of the α and β duplication events that occurred after the divergence of Brassicales, *A. thaliana* also has more members than *C. papaya*. Finally, the number of *SWEET* genes in *B. oleracea* and *B. rapa* apparently doubled after the *Brassica*-specific whole-genome triplication (WGT) event, based on a comparison with *A. thaliana* [45].

Identification and phylogenetic analysis of *SWEET* genes in *B. oleracea*

We constructed an unrooted phylogenetic tree (Additional file 3: Figure S1) of *SWEET* proteins in *B. oleracea*, *B. rapa* and *A. thaliana*. A total of 80 members clustered into four clades, containing 16, 21, 34, and 9 members, respectively. Based on the tree, eleven



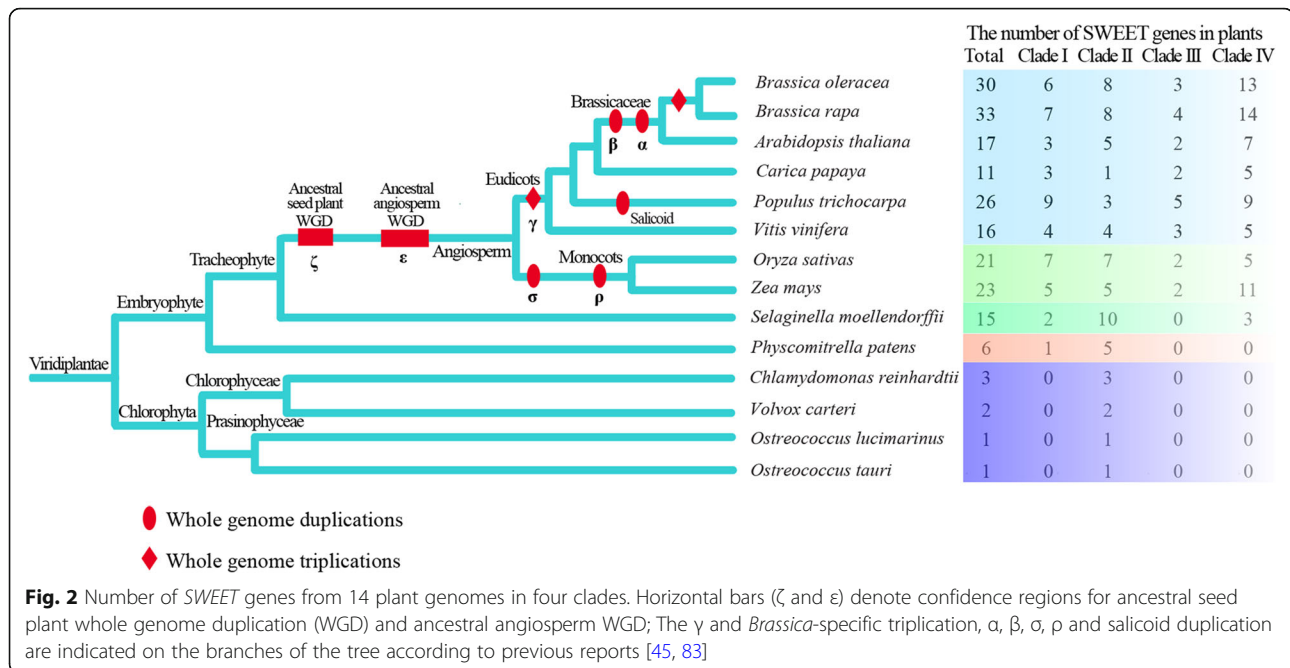
duplication clusters in *B. oleracea* were identified and are listed in Table 1.

The physical and chemical characteristics of the SWEET proteins in *B. oleracea* were predicted, and their sizes were found to range from 147 aa (BoSWEET7b) to 475 aa (BoSWEET5c), with the corresponding open reading frames (ORF) ranging from 441 bp to 1425 bp. The predicted molecular weights ranged from 16.20 kDa (BoSWEET7b) to 53.51 kDa (BoSWEET5c) (Table 1), and the theoretical isoelectric points (pI) from 6.28 (BoSWEET15b) to 9.56 (BoSWEET7a), with most, except for three, being higher than 7.60 (Table 1).

Chromosomal distribution and differential retention of SWEET genes in *B. oleracea*

The three subgenomes of *B. oleracea* (LF, the least fractionated blocks of *Brassica*; MF1, the most fractionated

blocks of *Brassica*; MF2, the medium fractionated blocks of *Brassica*) have been established to distinguish the degree of fractionation in genome evolution [44]. In this study, 29 *BoSWEET* genes (i.e., the whole family except *BoSWEET13*) were distributed among the three subgenomes (11 *BoSWEET* genes in LF, 9 in MF1, and 9 in MF2) (Table 1). Interestingly, all the *BoSWEET* genes were retained in *B. oleracea* after triplication and fractionation, except for the loss of *BoSWEET6*. Approximately half of the *BoSWEET* genes (8/17) were retained in two copies, while only five (*BoSWEET8*, -9, -10, -13 and -17) and three (*BoSWEET5*, *BoSWEET11* and *BoSWEET14*) were retained in one and three copies, respectively (Table 1). The retained copies have the same conserved collinear block. The physical positions of the *BoSWEET* genes on the *B. oleracea* chromosomes were identified (Fig. 3a), which revealed that 20 (66.7%) were



distributed across the nine chromosomes (C01–C09), with the exception of the C04 chromosome having no members, while the largest number on any single chromosome was five on chromosome 7 (Fig. 3a). Ten *SWEET* genes (*BoSWEET1a*, *-2a*, *-2b*, *-5a*, *-5c*, *-8*, *-10*, *-11a*, *-11c* and *-14a*) were not anchored on any of the *B. oleracea* chromosomes.

In eukaryotic genomes, a substantial proportion of protein-coding genes belong to multigene families, which have presumably evolved by the process of gene duplication. Duplicated genes from eukaryotic genomes have originated predominantly from inter-chromosomal duplications [69]. We observed that seven *BoSWEET* genes underwent segment duplication (7 duplications), leading to many homologs on different chromosomes, as indicated with red lines in Fig. 3a. To further highlight the specific retention of *BoSWEET* genes, their syntentic relationship with *AtSWEET* genes was determined using the Circos program (Fig. 3b).

Ks analysis of *B. oleracea* and *A. thaliana* *SWEET* genes

The *Ks* and *Ka* values of 30 syntenic *SWEET* orthologous pairs between *B. oleracea* and *A. thaliana* were analyzed (Table 2, Additional file 4: Table S3). The *Ka/Ks* ratios of all the *SWEET* syntenic orthologous pairs were far less than 1, indicative of purifying selection for retention. To estimate the divergence time between the two species, the *Ks* values of the orthologous *SWEET* genes were used and observed to range from 0.3 to 0.45, concentrating at approximately 0.35 (Fig. 4). According to the neutral substitution rate of 1.5×10^{-8} substitutions per site per year for dicotyledonous plants [60], this suggests the *SWEET* gene family of *B. oleracea* diverged

from *A. thaliana* approximately 10 to 15 million years ago (MYA). We concluded that the *SWEET* genes diverged concurrently with the Brassica-specific WGT event that occurred approximately 13–17 MYA [45, 70].

BoSWEET gene structures and associated conserved protein motifs

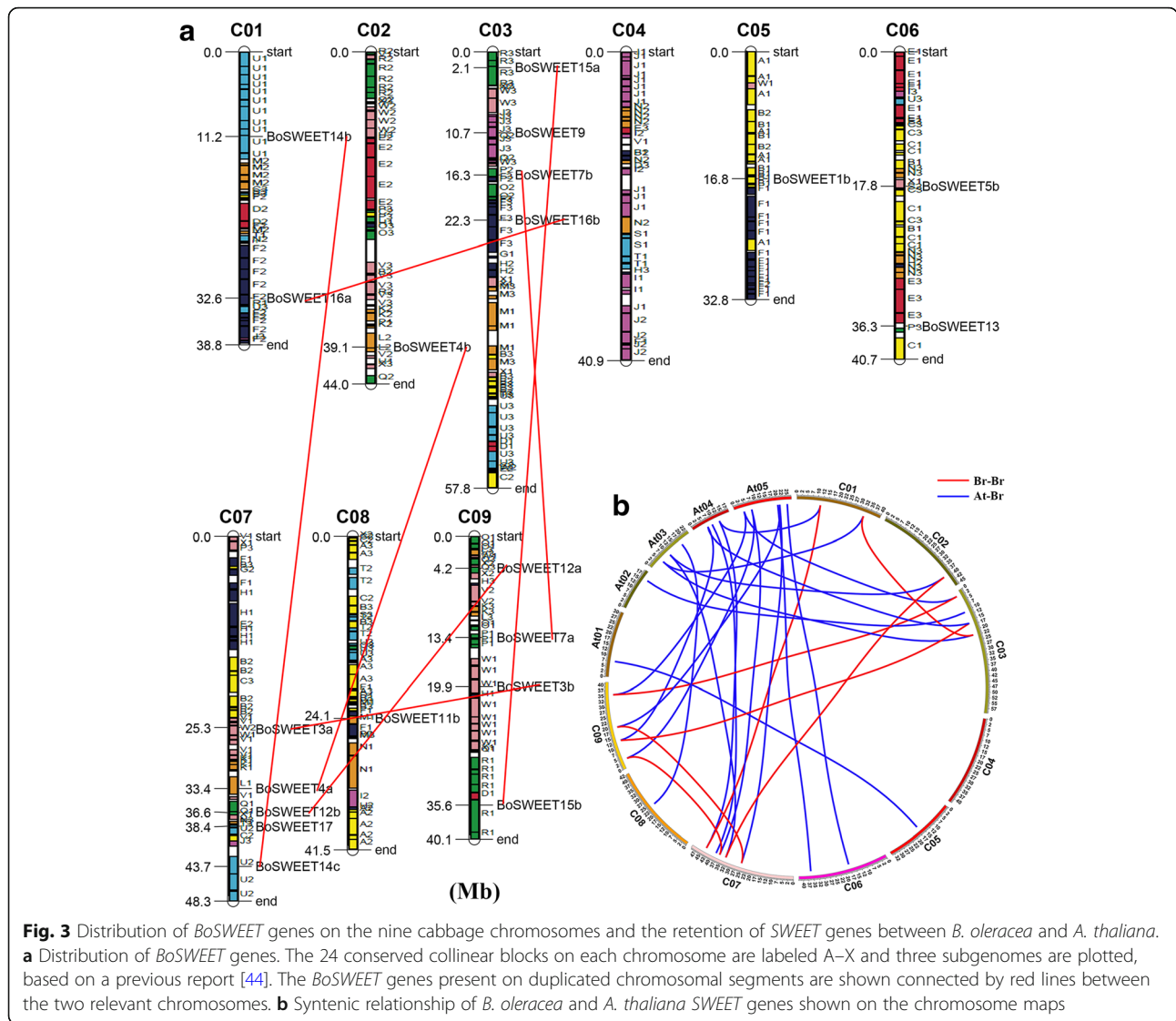
Gene structural diversity and conserved protein motif divergence played key roles in the evolution of the *SWEET* gene family [71]. The exon-intron organization of the cabbage *SWEET* genes was analyzed, and half (15 members) were found to have six exons, while seven members had four exons, three had five exons, three had three exons and *BoSWEET7b* and *BoSWEET5c* had 2 and 11 exons (Fig. 5), respectively. The exon lengths were similar, while the intron length varied, with eight *BoSWEETs* (*BoSWEET3b*, *-5b*, *-5c*, *-9*, *-11a*, *-11b*, *-11c* and *-17*) containing very long introns (Fig. 5). Six genes appeared to have exon-intron loss variations. For example, *BoSWEET2b* (four exons) lost the first two exons and two introns compared with *BoSWEET2a*; *BoSWEET5a* (four exons) lost the first exon and intron compared with *BoSWEET5b*; *BoSWEET11a* (three exons) lost the first three exons and introns compared with *BoSWEET11b*; and *BoSWEET14c* lost the third exon compared with *BoSWEET14b* (Additional file 5: Figure S2).

Among the 30 *BoSWEET* proteins, we found that the most *BoSWEET* (21) contained two MtN3_slv domains, a characteristic eukaryote *SWEET* protein domain, whereas eight contained only one MtN3_slv domain (Table 1). The number of TMHs ranged from three to eleven and concentrated in seven (Additional file 6: Figure

Table 1 SWEET family genes in *B. oleracea* and corresponding orthologs in *A. thaliana*

Group	tPCK Chr	<i>Arabidopsis thaliana</i>		<i>Brassica oleracea</i>		Subgenome	Chromosome	Gene ID	Gene name	Gene ID	Chromosome	ORF (bp)	PL (aa)	MW (kD)	pI	Number of MtN3_slv
		Gene name	Gene ID	ORF (bp)	PL (aa)											
Clade I	tPCK1	B	AtSWEET1	AT1G21460	744	247	BoSWEET1a	Bol012880	Scaffold000195_P1	MF1	741	246	27.805	9.2	2	
	tPCK2	F	AtSWEET2	AT3G14770	711	236	BoSWEET1b	Bol0009655	C05	LF	756	251	27.111	9.38	2	
							BoSWEET2a	Bol002799	Scaffold000374	LF	711	236	26.683	8.95	2	
Clade II	tPCK5	W	AtSWEET3	AT5G53190	792	263	BoSWEET2b	Bol011848	Scaffold000206	MF1	594	197	22.330	8.58	2	
							BoSWEET3a	Bol017165	C07	MF1	780	259	29.239	8.6	2	
							BoSWEET3b	Bol030288	C09	LF	648	215	24.056	7.64	1	
Clade III	tPCK7	L	AtSWEET4	AT3G28007	756	251	BoSWEET4a	Bol042944	C07	LF	747	248	27.479	9.07	2	
							BoSWEET4b	Bol031203	C02	MF1	738	245	27.323	8.78	2	
							BoSWEET5a	Bol019179	Scaffold000133	MF1	636	211	24.316	8.92	2	
Clade IV	tPCK6	E	AtSWEET5	AT5G62850	723	240	BoSWEET5b	Bol019426	C06	LF	738	245	27.918	9.3	2	
							BoSWEET5c	Bol020702	Scaffold000121_P2	MF2	1428	475	53.510	8.36	4	
							/	/	/	/	/	/	/	/	/	/
Clade V	tPCK5	P	AtSWEET6	AT4G10850	777	258	BoSWEET7a	Bol006816	C09	LF	750	249	27.277	9.56	2	
							BoSWEET7b	Bol025689	C03	MF1	444	147	16.200	7.78	1	
							BoSWEET8	Bol012888	Scaffold000192	LF	717	238	26.670	8.9	2	
Clade VI	tPCK3	J	AtSWEET9	AT2G39060	777	258	BoSWEET9	Bol020406	C03	MF2	531	176	19.551	7.85	1	
							BoSWEET10	Bol001620	Scaffold000431	MF2	870	289	33.058	9.2	2	
							BoSWEET11a	Bol003199	Scaffold000361	MF1	534	177	19.181	8.68	1	
Clade VII	tPCK6	M	AtSWEET11	AT3G48740	870	289	BoSWEET11b	Bol037413	C08	LF	870	289	31.945	9.32	2	
							BoSWEET11c	Bol002048	Scaffold000407	MF2	534	177	19.208	8.98	1	
							BoSWEET12a	Bol036230	C09	MF2	834	277	30.623	9.14	2	
Clade VIII	tPCK7	Q	AtSWEET12	AT5G23660	858	285	BoSWEET12b	Bol017126	C07	LF	867	288	31.782	9.07	2	
							BoSWEET13	Bol045233	C06	/	600	199	22.122	8.49	1	
							BoSWEET14a	Bol002776	Scaffold000375	MF2	819	272	29.971	9.18	2	
Clade IX	tPCK5	R	AtSWEET15	AT5G13170	879	292	BoSWEET14b	Bol039481	C01	LF	822	273	30.114	9.2	2	
							BoSWEET14c	Bol042202	C07	MF1	612	203	22.429	9.12	1	
							BoSWEET15a	Bol034222	C03	MF2	705	234	26.232	6.58	2	
Clade X	tPCK2	F	AtSWEET16	AT3G16690	693	230	BoSWEET15b	Bol043455	C09	LF	540	179	19.812	6.28	1	
							BoSWEET16a	Bol034791	C01	MF1	696	231	25.759	8.69	2	
							BoSWEET16b	Bol022976	C03	MF2	552	183	20.362	9.35	2	
Clade XI	tPCK4	T	AtSWEET17	AT4G15920	726	241	BoSWEET17	Bol037218	C07	MF2	723	240	26.513	6.73	2	

Note: tPCK Chr Chromosome of translocation Proto-Calepineae Karyotype, ancestral genome of *Brassica* species, CCB conserved collinear block, LF the least fractionated blocks of *Brassica*, MF1 the medium fractionated blocks of *Brassica*, MF2 the most fractionated blocks of *Brassica*, PL protein length, MW molecular weight.



S3). However, we identified the BoSWEET5c protein with four MtN3_slv domains that has a similar predicted protein architecture to VV14G09070 from *V. vinifera*, and that together constitute a novel sub-type, which were named extra-SWEET (Additional file 7: Figure S4). To further investigate the structural diversity, the conserved motif structure of all the BoSWEET proteins was analyzed. In total, ten conserved motifs (motifs 1–10) were identified, among which motifs 7, 2, 4, and motifs 1, 6, 3 were annotated as MtN3_slv domains, respectively. Motif 5 or motifs 9, 5 connects the two MtN3_slv domains (Fig. 5). In addition, none of the 30 BoSWEET proteins were predicted to contain signal peptide sequences.

Characterization of *BoSWEET* expression in response to chilling stress

A total of eight *BoSWEET* genes were selected for expression pattern analysis using the chilling-tolerant line

CT-923. The eight selected *BoSWEET* genes represented all four clades, with *BoSWEET2b* and *BoSWEET4a* belonging to clade I and clade II, respectively; *BoSWEET11b*, *BoSWEET11c*, *BoSWEET12b* and *BoSWEET15b* belonging to clade III, and *BoSWEET16a* and *BoSWEET17* belonging to clade IV. We found that expression of *BoSWEET2b*, *BoSWEET4a* and *BoSWEET15b* were induced after chilling stress (Fig. 6), with *BoSWEET2b* expression increasing to maximal levels at 6 h following chilling stress. *BoSWEET4a* and *BoSWEET15b* expression peaked at 24 h after chilling stress. In addition, while the expression of *BoSWEET11b*, *BoSWEET11c*, *BoSWEET12b*, *BoSWEET16a* and *BoSWEET17* decreased in response to chilling, the expression patterns were different. The expression of *BoSWEET11b* and *BoSWEET12b* declined to a minimum at 12 h, then rose to reach nearly half of the expression of the control. The expression of *BoSWEET11c* declined

Table 2 Non-synonymous (K_a) and synonymous substitution rate (K_s) between orthologous *SWEET* gene pairs in *B. oleracea* and *A. thaliana*

Orthologous gene pairs	K_a	K_s	K_a/K_s	Duplication date (MYA)
<i>AtSWEET1</i>				
<i>BoSWEET1a</i>	0.0393	0.4481	0.0877	14.9
<i>BoSWEET1b</i>	0.0404	0.3674	0.1100	12.2
<i>AtSWEET2</i>				
<i>BoSWEET2a</i>	0.0458	0.2509	0.1825	8.4
<i>BoSWEET2b</i>	0.0756	0.368	0.2054	12.3
<i>AtSWEET3</i>				
<i>BoSWEET3a</i>	0.1024	0.3765	0.2720	12.6
<i>BoSWEET3b</i>	0.1163	0.39	0.2982	13.0
<i>AtSWEET4</i>				
<i>BoSWEET4a</i>	0.0491	0.3639	0.1349	12.1
<i>BoSWEET4b</i>	0.0624	0.3623	0.1722	12.1
<i>AtSWEET5</i>				
<i>BoSWEET5a</i>	0.0709	0.3544	0.2001	11.8
<i>BoSWEET5b</i>	0.0509	0.3204	0.1589	10.7
<i>BoSWEET5c</i>	0.0853	0.3782	0.2255	12.6
<i>AtSWEET7</i>				
<i>BoSWEET7a</i>	0.0863	0.4894	0.1763	16.3
<i>BoSWEET7b</i>	0.1344	0.6556	0.2050	21.9
<i>AtSWEET8</i>				
<i>BoSWEET8</i>	0.2009	1.3091	0.1535	43.6
<i>AtSWEET9</i>				
<i>BoSWEET9</i>	0.0642	0.2567	0.2501	8.6
<i>AtSWEET10</i>				
<i>BoSWEET10</i>	0.101	0.4938	0.2045	16.5
<i>AtSWEET11</i>				
<i>BoSWEET11a</i>	0.0434	0.2373	0.1829	7.9
<i>BoSWEET11b</i>	0.0214	0.2179	0.0982	7.3
<i>BoSWEET11c</i>	0.0287	0.179	0.1603	6.0
<i>AtSWEET12</i>				
<i>BoSWEET12a</i>	0.0599	0.3024	0.1981	10.1
<i>BoSWEET12b</i>	0.0476	0.3442	0.1383	11.5
<i>AtSWEET13</i>				
<i>BoSWEET13</i>	0.0961	0.2736	0.3512	9.1
<i>AtSWEET14</i>				
<i>BoSWEET14a</i>	0.0712	0.312	0.2282	10.4
<i>BoSWEET14b</i>	0.0652	0.3597	0.1813	12.0
<i>BoSWEET14c</i>	0.0967	0.3643	0.2654	12.1
<i>AtSWEET15</i>				
<i>BoSWEET15a</i>	0.1207	0.3539	0.3411	11.8
<i>BoSWEET15b</i>	0.1114	0.344	0.3238	11.5
<i>AtSWEET16</i>				
<i>BoSWEET16a</i>	0.0912	0.3653	0.2497	12.2
<i>BoSWEET16b</i>	0.1011	0.4047	0.2498	13.5
<i>AtSWEET17</i>				
<i>BoSWEET17</i>	0.0457	0.3157	0.1448	10.5

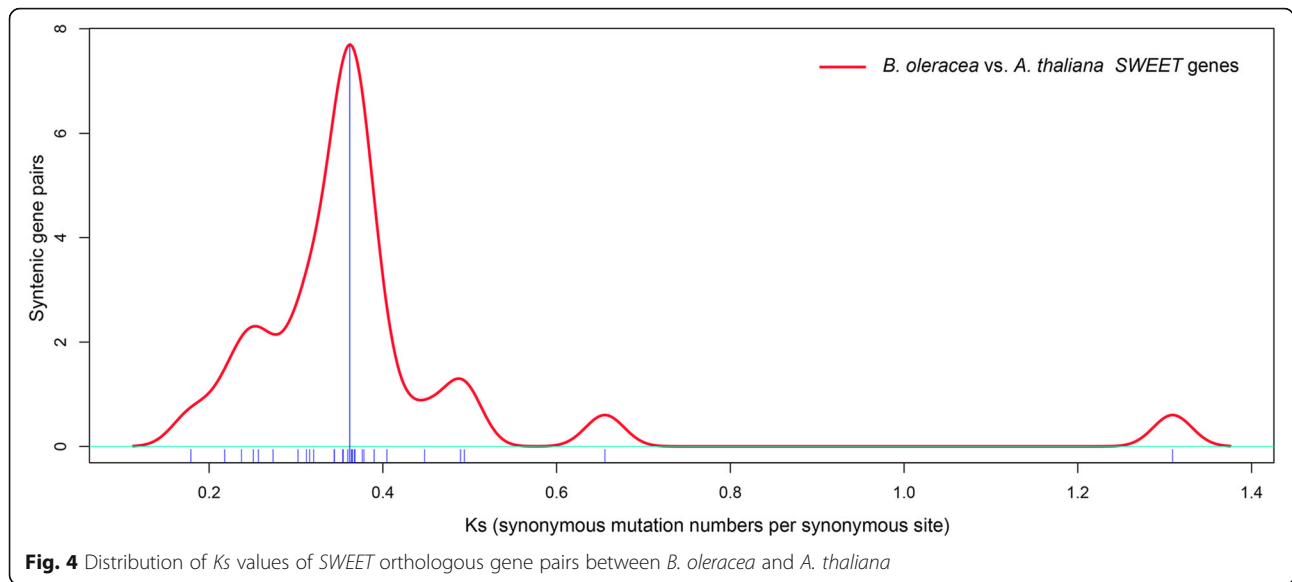
sharply to a minimum at 3 h after treatment, then increased to a maximum at 48 h. *BoSWEET16a* and *BoSWEET17* had a similar expression pattern after chilling exposure, with expression declining rapidly and remaining at a low level even after 12 h (Fig. 6).

Expression of *BoSWEET* genes in different organs and in response to *P. brassicae* infection

We examined the RNA-Seq data set (GSE42891) to determine transcript levels of the *BoSWEET* genes in the cabbage bud, callus, flower, leaf, root, silique and stem. The expression of most of the *BoSWEET* genes exhibited different patterns (Fig. 7a). Seven were expressed in all organs, whereas the expression of three was not detected. Several genes were expressed in only one or two

organ types, such as *BoSWEET3a* in buds, *BoSWEET3b* and *BoSWEET5c* in buds and flowers, *BoSWEET14c* in roots and stems, and *BoSWEET16a* in calli and roots (Fig. 7a). This diversity of expression patterns suggested a broad range of biological functions of *BoSWEET* genes during growth and development of cabbage.

We also examined the expression patterns of *BoSWEET* genes in CR-XG336, a clubroot-resistant line, and CS-JF1, a clubroot-susceptible line, infected by *P. brassicae* at two infection stages, and observed variation between the two different cultivars and at different infection stages (Fig. 7b). Several genes were nearly undetectable in both cultivars both before *P. brassicae* infection and at two infection stages. Six (*BoSWEET4a*, *-11c*, *-12a*, *-12b*, *-14b* and *-15b*) and three

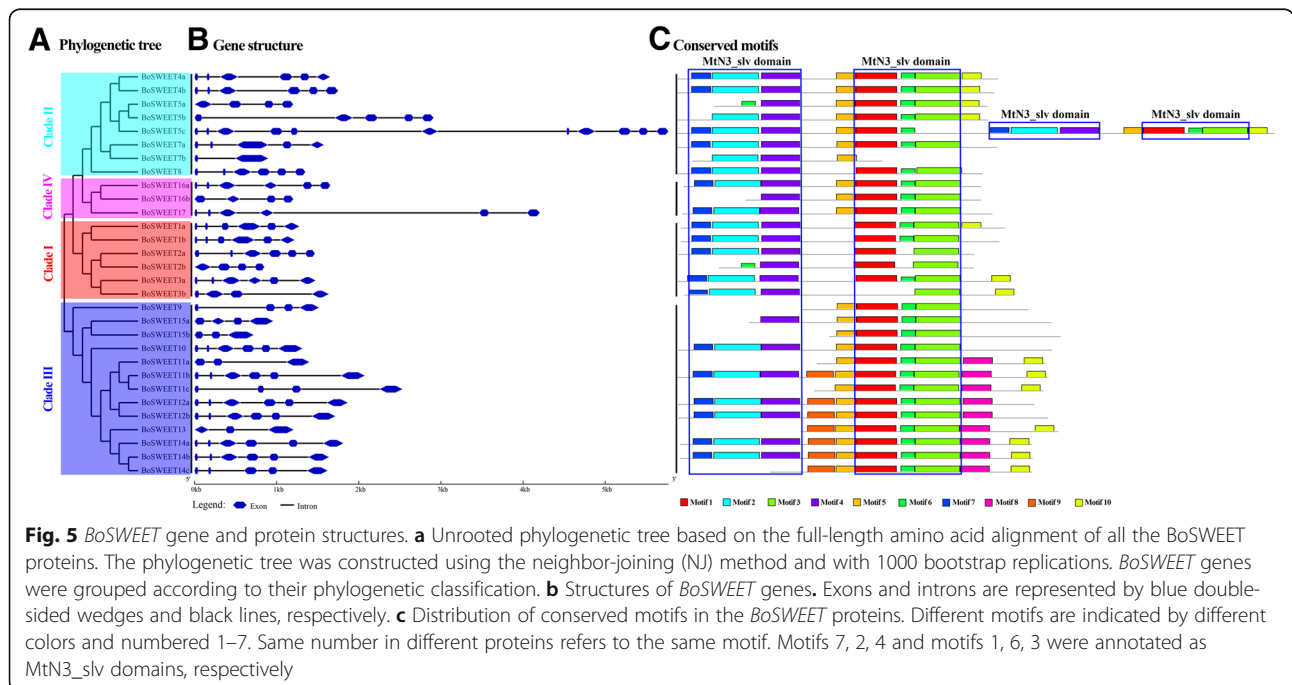


(*BoSWEET14c*, *BoSWEET16a* and *BoSWEET17*) *BoSWEET* genes were significantly up- and down-regulated, respectively, in CS-JF1 at 7 DAI (primary infection stage) compared with mock-treated plants. No *BoSWEET* genes were significantly up- or down-regulated in CR-XG336 at 7 DAI compared with mock-treated plants (Fig. 7b, Additional file 8: Table S4). Moreover, at 28 DAI (club-root formation stage), six *BoSWEET* genes (*BoSWEET12b*, -13, -14a, -14b, -14c and -17) were significantly down-regulated in CS-JF1 after infection, whereas two (*BoSWEET1a* and *BoSWEET16a*) and three

(*BoSWEET13*, *BoSWEET14a* and *BoSWEET14b*) were up- and down-regulated in CR-XG336, respectively (Fig. 7b, Additional file 8: Table S4).

Subcellular localization analysis of BoSWEET proteins following heterologous expression in tobacco

To study the subcellular localization of *BoSWEET* proteins, *BoSWEET11b*, *BoSWEET11c* and *BoSWEET12b* were heterologously and transiently expressed in tobacco leaf epidermal cells as translational GFP fusion proteins. All three proteins were found to be localized to the



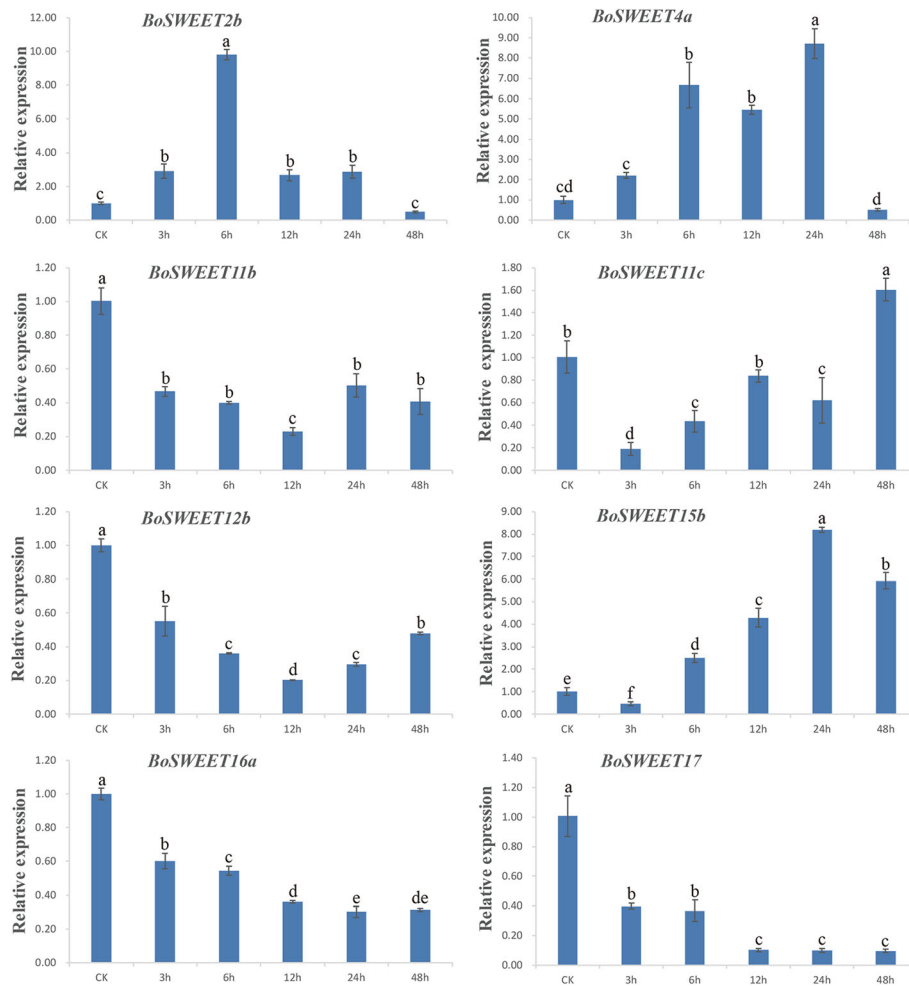


Fig. 6 qRT-PCR analysis of eight *BoSWEET* genes in cabbage leaves following chilling treatment. Data are presented as means \pm standard deviations of three technical replicates derived from one bulked biological replicate. A Duncan's multiple range test was used to calculate the significance level of the data at $P < 0.05$

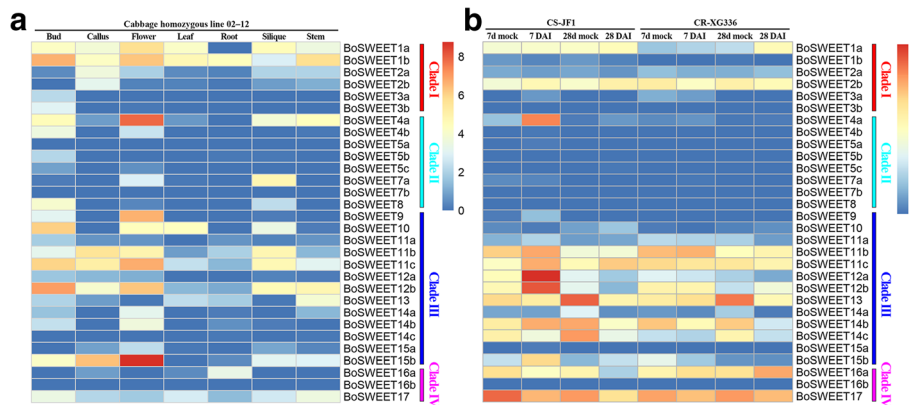
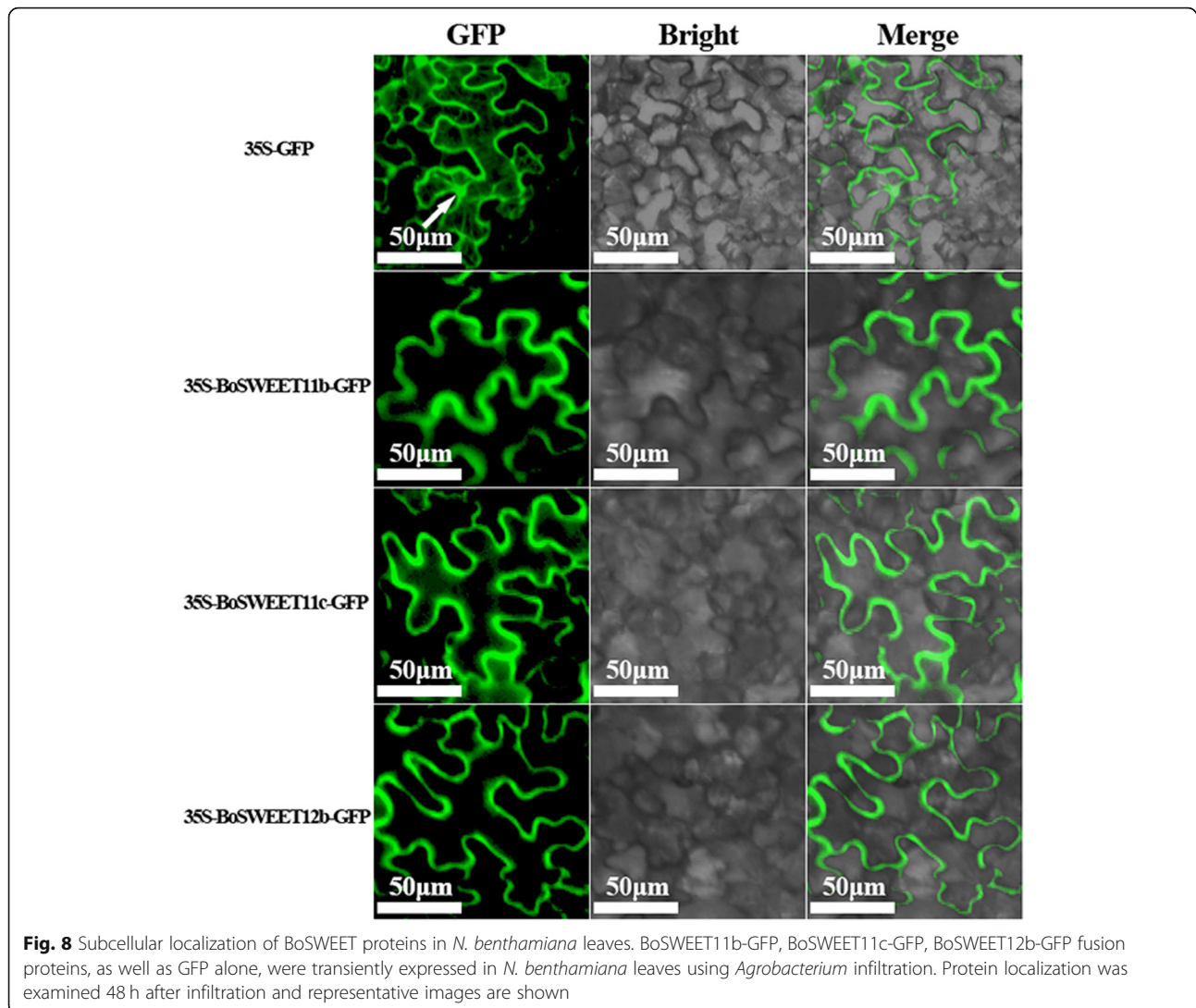


Fig. 7 Expression patterns of *BoSWEET* genes analyzed by RNA-Seq. **a** Heatmap showing *BoSWEET* gene expression in different organs. Bud, callus, flower, leaf, root, silique and stem are represented. **b** Heatmap showing expression dynamics of *BoSWEET* genes in cabbage roots following *P. brassicae* infection. Expression levels of the *BoSWEET* genes are shown as the Log₂ transformed FPKM values obtained from the RNA-Seq data. DAI, days after inoculation



plasma membrane (PM), whereas the control 35S-GFP (GFP alone) was detected in the PM, cytoplasm and nucleus (Fig. 8). These results suggested that all three proteins are PM-localized (Fig. 5, Additional file 5: Figure S2).

Discussion

Evolutionary history of *SWEET* genes among fourteen plant species

Gene family evolution is characterized by gene duplication via whole-genome duplication (WGD), tandem gene duplication and segmental duplication events [72]. Following duplication, duplicated gene pairs can undergo different fates, including neo-functionalization (functional diversification), sub-functionalization (partitioning of the function between daughter copies) and non-functionalization [73, 74]. In plants, genome duplication has been shown to contribute their ability to adapt to diverse environments, including drought, pathogen attack, and extreme

temperatures, as well as reproductive development [75]. In this study, the number of *SWEET* genes was found to vary considerably among fourteen plant species. For example, there were only 1–3 members in algae, with the salt water species *O. lucimarius* and *O. tauri* having only one copy and the fresh water species *V. carteri* and *C. reinhardtii* having three. The evolution from unicellular (algae) to multicellular plants led to further expansion of the *SWEET* gene family. The bryophyte *P. patens* is an early diverging land plant with only one primitive protophloem and has six *SWEET* gene family members, while the pteridophyte *S. moellendorffii* evolved to have phloem and has 15 *SWEET* genes, which is similar to the number in many angiosperms (Fig. 2). These results suggested that the expansion of *SWEET* genes might play an important role in land plant adaptations to terrestrial conditions.

Among Brassicaceae species, *B. oleracea*, *B. rapa* and *A. thaliana* have experienced γ , β and α WGD events. The number of *SWEET* genes in *B. oleracea* (30) and *B. rapa* (33) nearly doubled when compared with the number in *A. thaliana* (17) after the *Brassica*-specific WGT event, during which genes had experienced triplication, differentiation and fractionation (loss) [45, 70]. However, the gene balance hypothesis assumes that genes participate in macromolecular complexes, signaling and transcriptional networks are preferentially retained, thus avoiding the instability and unfitness of network caused by the imbalance associated with loss of one member of a complex [76–78]. Similarly, we found evidence of *SWEET* gene fractionation in *B. oleracea* after the split with *A. thaliana* from the recent common ancestor, with *BoSWEET6* being lost, five *BoSWEETs* (*BoSWEET8*, -9 , -10 , -13 and -17) and eight *BoSWEETs* (*BoSWEET1*, -2 , -3 , -4 , -7 , -12 , -15 , -16) having one and two separate orthologous genes in *A. thaliana*, respectively (Table 1; Additional file 3: Figure S1). WGD or WGT generates the multiplicity in gene copy numbers, and is the main source of genetic redundancy. We firmly believe that changes in *SWEET* genes during the WGT event played critical roles in the adaptation and expansion of rich morphotypes of *Brassica* crops.

Structure and evolution of *BoSWEET* genes

The *SWEET* proteins have evolved from an internal repeat duplication of three-TMH unit and fusion with an insertion of TMH4 [8]. In this study, we searched for ten conserved motifs among the 30 *BoSWEET* proteins and found the MtN3_slv domains were located in duplication regions in almost all *BoSWEET* proteins (Fig. 5). The extra-*SWEET* protein (*BoSWEET5c*, 11-TMHs) may have originated from two internal *SWEET* (7-TMHs, two MtN3_slv domains) duplications, similar to the duplication of semi-*SWEET* (3-TMHs, one MtN3_slv domain) and subsequently evolved into the *SWEET* [8, 79]. In *V. vinifera*, berries accumulate high levels of sugars, and the *SWEET* protein, VV14G09070, might play a role in mediating elevated levels of sucrose transport. To this end, it would be very interesting to investigate the spatial expression pattern of VV14G09070 and its potential function in long distance sugar transport during flower or berry development [80].

Expression patterns of *BoSWEET* genes in different organs and in response to chilling stress

Transcriptome sequencing revealed that more *BoSWEET* genes were highly expressed in flowers and buds than in the other organs investigated. Of the highly expressed *BoSWEET* genes, most were in clade III (Fig. 7a), which is consistent with studies of rapeseed (*Brassica napus*) [20]. The expression of *SWEET* genes has been shown to

change in response to abiotic stress in several plant species [19, 26, 81]. For example, in tomato *SISWEET10a*, $-10b$, $-10c$, $-11a$, $-11b$, $-11c$, $-11d$, $-12a$ and $-12c$ were reported to have similar expression patterns and were up-regulated several fold in leaves but substantially down-regulated in roots in response to sugar treatment, and salt, heat and cold stresses [19]. We observed that the expression of *BoSWEET11b*, $-11c$, $-12b$, $-16a$ and -17 was down-regulated by chilling stress (Fig. 6), which may contribute to the accumulation of glucose and fructose in leaves, leading to increased chilling tolerance. Heterologously expressed *BoSWEET11b*, *BoSWEET11c* and *BoSWEET12b* proteins fused to a GFP marker were observed to be accumulated in the PM, consistent with a role in mediating sucrose efflux from phloem parenchyma cells into the sieve element-companion cell complex [9]. The expression of *BoSWEET16a* and *BoSWEET17* declined rapidly after chilling stress and remained at low levels from 12 h to 48 h. The orthologous genes from *A. thaliana*, *AtSWEET16* and *AtSWEET17*, have been reported to export fructose from the vacuole and contribute to cytosolic fructose homeostasis [11, 12, 26].

Involvement of *BoSWEET* genes in the response to *P. brassicae* infection

Previous studies revealed that *P. brassicae* obtains sugars from hosts to complete its life cycle, involving the formation of galls, which act as an additional sink. In *A. thaliana*, it was found that sucrose accumulated in uninfected leaves, but not in *P. brassicae* infected leaves because sucrose was exported from leaves into the clubroot galls [82]. The expression of sugar transporter genes may therefore influence plant-*P. brassicae* interactions. In this study, the expression of six *BoSWEET* genes (*BoSWEET4a*, $-11c$, $-12a$, $-12b$, $-14b$ and $-15b$) was up-regulated in roots in the susceptible CS-JF1 upon *P. brassicae* infection at 7 DAI compared with mock plants. In contrast, no *BoSWEET* genes were up-regulated in the resistant CR-XG336 at 7 DAI (Fig. 6; Additional file 8: Table S4). We infer from these results that these six *BoSWEET* genes could be responsible for transporting sugars to the sink roots associated with *P. brassicae* colonization in CS-JF1. Clade III *SWEET* proteins have been shown to function as sucrose transporters involved in long distance sugar transport in *A. thaliana* [3, 9], consistent with our observation that all the up-regulated *BoSWEET* genes involved in *P. brassicae* infection in CS-JF1, except *BoSWEET4a*, belong to clade III. Clade III *SWEET* proteins have also been shown to be involved in sugar transport in clubroot disease establishment [30]. We also noted that the expression of the *BoSWEET1a* and *BoSWEET16a* genes was approximately 4-fold and 2-fold higher, respectively, in CR-XG336 at 28 DAI than in the uninfected control (Additional file 8: Table S4). Whether the increased expressions of the two *BoSWEET* genes

contribute to CR-XG336 resistance to *P. brassicae* infection stress remains to be established.

Conclusions

In this study, 30 *BoSWEET* genes were identified in the *B. oleracea* genome and further clustered into four clades based on a phylogenetic tree of 205 *SWEET* homologs from fourteen representative plant species. Clade II was evolutionarily the most ancient, while clade I, clade IV and clade III were formed successively. The ORF lengths of the *BoSWEET* genes ranged from 441 bp to 1425 bp, and the *Ks* values of the orthologous *SWEET* genes from *B. oleracea* and *A. thaliana* ranged from 0.30 to 0.45, meaning that the estimated time of divergence of the two species was approximately 10 to 15 MYA. Eight *BoSWEET* proteins were predicted to contain a single MtN3_slv domain, twenty-one to contain two, and one (*BoSWEET5c*) to have four. qRT-PCR analysis showed that the expression of five *BoSWEET* genes decreased when plants were exposed to chilling stress. The expression of six *BoSWEET* genes was up-regulated in CS-JF1 following *P. brassicae* infection at 7 DAI compared with mock controls, and we hypothesize that they might be responsible for transporting sugars to the sink roots during *P. brassicae* colonization. Overall, these findings facilitate unraveling the potential candidate *BoSWEET* genes involved in the response to chilling and clubroot disease, and provide valuable information to facilitate the breeding of chilling tolerant and clubroot disease-resistant cultivars in cabbage.

Additional files

Additional file 1: Table S1. Primers used in this study. (XLSX 10 kb)

Additional file 2: Table S2. Amino acid sequences of 205 *SWEET* proteins from 14 plant species. (XLSX 35 kb)

Additional file 3: Figure S1. Phylogenetic tree of *B. oleracea*, *B. rapa* and *A. thaliana* *SWEET* proteins. Phylogenetic analysis of 80 *SWEET* proteins from *B. oleracea* (30), *B. rapa* (33) and *A. thaliana* (17) showing similar groups in all species. Four clades were marked with different colors. (TIF 1595 kb)

Additional file 4: Table S3. Nucleotide sequences of 47 *SWEET* genes from *B. oleracea* and *A. thaliana*. (XLSX 19 kb)

Additional file 5: Figure S2. Gene organization of *BoSWEET* genes. (TIF 290 kb)

Additional file 6: Figure S3. Protein structure of *BoSWEET*s in *B. oleracea*. Red rectangles signify the TMHs, and blue and carmine lines indicate the intracellular and extracellular regions, respectively. (TIF 545 kb)

Additional file 7: Figure S4. Conserved domain architecture of the *BoSWEET5c* and VV14G09070 proteins. (TIF 177 kb)

Additional file 8: Table S4. Expression of *BoSWEET* genes in roots of clubroot-resistant and susceptible cabbage lines at different *P. brassicae* infection stages. (XLSX 18 kb)

Abbreviations

CDART: Conserved Domain Architecture Retrieval Tool; GSDS: Gene Structure Display Server; HMM: Hidden Markov model; *Ka*: Nonsynonymous substitution rate; *Ks*: Synonymous substitution rate; MEME: Multiple EM for Motif Elicitation; qRT-PCR: Quantitative reverse transcription-PCR;

SMART: Simple Modular Architecture Research Tool; TMH: Transmembrane helix; WGD: Whole-genome duplication; WGT: Whole-genome triplication

Acknowledgements

The authors gratefully acknowledge Dr. Liang Xu of Nanjing Agricultural University for his helpful advice and critical reading of the manuscript.

Funding

This work was in part supported by the China Postdoctoral Science Foundation funded project (2017 M621669), National Key Research and Development Program of China (2016YFD0101702), National Natural Science Foundation of China (31701773), Key Research and Development Program of Jiangsu Province (BE2017379), Jiangsu Agriculture Science and Technology Innovation Fund (CX(18)2006) and National Modern Agriculture Industry Technology System (CARS-25). The funding body played no role in study design and data collection, analysis, interpretation of data and in writing the manuscript.

Availability of data and materials

Plant materials are available under request to the respective owner institutions. The datasets supporting the results of this article are included within the article and its additional files.

Authors' contributions

JL and ZF conceived the and supervised the work. WZ did bioinformatics analyses and drafted the manuscript. FY, LY and HW performed validation of experiments. SW and XS performed the qRT-PCR analysis. KB and JT processed the transcriptome data. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Jiangsu Key Laboratory for Horticultural Crop Genetic Improvement, Institute of Vegetable Crops, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, People's Republic of China. ²Zhenjiang Agricultural Research Institute, Jurong, Jiangsu 212400, People's Republic of China. ³Boyce Thompson Institute, Cornell University, Ithaca, NY 14853, USA.

Received: 25 September 2018 Accepted: 14 January 2019

Published online: 29 January 2019

References

- Lemoine R, La Camera S, Atanassova R, Dedaldechamp F, Allario T, Pourtau N, et al. Source-to-sink transport of sugar and regulation by environmental factors. *Front Plant Sci.* 2013;4:272.
- Rennie EA, Turgeon R. A comprehensive picture of phloem loading strategies. *Proc Natl Acad Sci U S A.* 2009;106(33):14162–7.
- Chen LQ, Hou BH, Lalonde S, Takanaga H, Hartung ML, Qu XQ, et al. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature.* 2010;468(7323):527–32.
- Doidy J, Grace E, Kuhn C, Simonplais F, Casieri L, Wipf D. Sugar transporters in plants and in their interactions with fungi. *Trends Plant Sci.* 2012;17(7):413–22.
- Slewinski TL. Diverse functional roles of monosaccharide transporters and their homologs in vascular plants: a physiological perspective. *Mol Plant.* 2011;4(4):641–62.
- Braun DM. SWEET! The pathway is complete. *Science.* 2012;335(6065):173–4.
- Chen LQ. SWEET sugar transporters for phloem transport and pathogen nutrition. *New Phytol.* 2014;201(4):1150–5.

8. Xuan YH, Hu YB, Chen LQ, Sosso D, Ducat DC, Hou BH, et al. Functional role of oligomerization for bacterial and plant SWEET sugar transporter family. *Proc Natl Acad Sci U S A*. 2013;110(39):E3685–94.
9. Chen LQ, Qu XQ, Hou BH, Sosso D, Osorio S, Fernie AR, et al. Sucrose efflux mediated by SWEET proteins as a key step for phloem transport. *Science*. 2012;335(6065):207–11.
10. Zhou Y, Liu L, Huang W, Yuan M, Zhou F, Li X, et al. Overexpression of *OsSWEET5* in rice causes growth retardation and precocious senescence. *PLoS One*. 2014;9(4):e94210.
11. Chardon F, Bedu M, Calenge F, Klemens PAW, Spinner L, Clement G, et al. Leaf fructose content is controlled by the vacuolar transporter SWEET17 in *Arabidopsis*. *Curr Biol*. 2013;23(8):697–702.
12. Guo WJ, Nagy R, Chen HY, Pfrunder S, Yu YC, Santelia D, et al. SWEET17, a facilitative transporter, mediates fructose transport across the tonoplast of *Arabidopsis* roots and leaves. *Plant Physiol*. 2014;164(2):777–89.
13. Yuan M, Wang SP. Rice *MtN3/Saliva/SWEET* family genes and their homologs in cellular organisms. *Mol Plant*. 2013;6(3):665–74.
14. Mizuno H, Kasuga S, Kawahigashi H. The sorghum *SWEET* gene family: stem sucrose accumulation as revealed through transcriptome profiling. *Biotechnol Biofuels*. 2016;9(1):127.
15. Patil G, Valliyodan B, Deshmukh RK, Prince SJ, Nicander B, Zhao M, et al. Soybean (*Glycine max*) *SWEET* gene family: insights through comparative genomics, transcriptome profiling and whole genome re-sequencing analysis. *BMC Genomics*. 2015;16(1):520.
16. Wei X, Liu F, Chen C, Ma F, Li M. The *Malus domestica* sugar transporter gene family: identifications based on genome and expression profiling related to the accumulation of fruit sugars. *Front Plant Sci*. 2014;5:569.
17. Chong J, Piron M, Meyer S, Merdinoglu D, Bertsch C, Mestre P. The *SWEET* family of sugar transporters in grapevine: VvSWEET4 is involved in the interaction with *Botrytis cinerea*. *J Exp Bot*. 2014;65(22):6589–601.
18. Miao H, Sun P, Liu Q, Miao Y, Liu J, Zhang K, et al. Genome-wide analyses of *SWEET* family proteins reveal involvement in fruit development and abiotic/biotic stress responses in banana. *Sci Rep*. 2017;7(1):3536.
19. Feng C, Han J, Han X, Jiang J. Genome-wide identification, phylogeny, and expression analysis of the *SWEET* gene family in tomato. *Gene*. 2015;573(2):261–72.
20. Jian H, Lu K, Yang B, Wang T, Zhang L, Zhang A, et al. Genome-wide analysis and expression profiling of the *SUC* and *SWEET* gene families of sucrose transporters in oilseed rape (*Brassica napus* L.). *Front Plant Sci*. 2016;7:1464.
21. Ge Y, Angenent GC, Wittich PE, Peters J, Franken J, Busscher M, et al. *NECT1*, a novel gene, highly expressed in nectary tissue of *Petunia hybrida*. *Plant J*. 2000;24(6):725–34.
22. Lin IW, Sosso D, Chen LQ, Gase K, Kim S, Kessler D, et al. Nectar secretion requires sucrose phosphate synthases and the sugar transporter SWEET9. *Nature*. 2014;508(7497):546–9.
23. Chen LQ, Cheung LS, Feng L, Tanner W, Frommer WB. Transport of sugars. *Annu Rev Biochem*. 2015;84(1):865–94.
24. Sosso D, Luo D, Li Q, Sasse J, Yang J, Gendrot G, et al. Seed filling in domesticated maize and rice depends on SWEET-mediated hexose transport. *Nat Genet*. 2015;47(12):1489–93.
25. Quirino BF, Normanly J, Amasino RM. Diverse range of gene activity during *Arabidopsis thaliana* leaf senescence includes pathogen-independent induction of defense-related genes. *Plant Mol Biol*. 1999;40(2):267–78.
26. Klemens PAW, Patzke K, Deitmer JW, Spinner L, Hir RL, Bellini C, et al. Overexpression of the vacuolar sugar carrier AtSWEET16 modifies germination, growth, and stress tolerance in *Arabidopsis*. *Plant Physiol*. 2013;163(3):1338–52.
27. Le Hir R, Spinner L, Klemens PAW, Chakraborti D, De Marco F, Vilaine F, et al. Disruption of the sugar transporters AtSWEET11 and AtSWEET12 affects vascular development and freezing tolerance in *Arabidopsis*. *Mol Plant*. 2015;8(11):1687–90.
28. Liu X, Zhang Y, Yang C, Tian Z, Li J. AtSWEET4, a hexose facilitator, mediates sugar transport to axial sinks and affects plant development. *Sci Rep*. 2016;6(1):24563.
29. Antony G, Zhou J, Huang S, Li T, Liu B, White FF, et al. Rice *xa13* recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene *Os-11N3*. *Plant Cell*. 2010;22(11):3864–76.
30. Li H, Li X, Xuan YH, Jiang J, Wei Y, Piao Z. Genome wide identification and expression profiling of *SWEET* genes family reveals its role during *Plasmodiophora brassicae*-induced formation of clubroot in *Brassica rapa*. *Front Plant Sci*. 2018;9:207.
31. Liu Q, Yuan M, Zhou Y, Li X, Xiao J, Wang S. A paralog of the *MtN3/saliva* family recessively confers race-specific resistance to *Xanthomonas oryzae* in rice. *Plant Cell Environ*. 2011;34(11):1958–69.
32. Gamas P, C NF, Lescure N, Cullimore JV. Use of a subtractive hybridization approach to identify new *Medicago truncatula* genes induced during root nodule development. *Mol Plant-Microbe Interact*. 1996;9(4):233–42.
33. Hutin M, Sabot F, Ghesquiere A, Koebnik R, Szurek B. A knowledge-based molecular screen uncovers a broad-spectrum *OsSWEET14* resistance allele to bacterial blight from wild rice. *Plant J*. 2015;84(4):694–703.
34. Siemens J, Keller I, Sarx J, Kunz S, Schuller A, Nagel W, et al. Transcriptome analysis of *Arabidopsis* clubroots indicate a key role for cytokinins in disease development. *Mol Plant-Microbe Interact*. 2006;19(5):480–94.
35. Dixon GR. The occurrence and economic impact of *Plasmodiophora brassicae* and clubroot disease. *J Plant Growth Regul*. 2009;28(3):194–202.
36. Ludwigmuller J, Prinsen E, Rolfe SA, Scholes JD. Metabolism and plant hormone action during clubroot disease. *J Plant Growth Regul*. 2009;28(3):229–44.
37. Dixon GR. Clubroot (*Plasmodiophora brassicae* Woronin) – an agricultural and biological challenge worldwide. *Can J Plant Pathol*. 2014;36:5–18.
38. Gravot A, Deleu C, Wagner G, Lariagon C, Lukan R, Todd CD, et al. Arginase induction represses gall development during clubroot infection in *Arabidopsis*. *Plant Cell Physiol*. 2012;53(5):901–11.
39. Zhang H, Feng J, Zhang S, Zhang S, Li F, Strelkov SE, et al. Resistance to *Plasmodiophora brassicae* in *Brassica rapa* and *Brassica juncea* genotypes from China. *Plant Dis*. 2014;99(6):776–9.
40. Wallenhammar AC. Prevalence of *Plasmodiophora brassicae* in a spring oilseed rape growing area in Central Sweden and factors influencing soil infestation levels. *Plant Pathol*. 1996;45(4):710–9.
41. Chinnusamy V, Zhu J, Zhu J. Cold stress regulation of gene expression in plants. *Trends Plant Sci*. 2007;12(10):444–51.
42. Levitt J. Frost, drought, and heat resistance. *Annu Rev Plant Biol*. 1951;2(1):1–85.
43. Santarius KA. The protective effect of sugars on chloroplast membranes during temperature and water stress and its relationship to frost, desiccation and heat resistance. *Planta*. 1973;113(2):105–14.
44. Liu S, Liu Y, Yang X, Tong C, Edwards D, Parkin IAP, et al. The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nat Commun*. 2014;5(3930):3930.
45. Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, et al. The genome of the mesopolyploid crop species *Brassica rapa*. *Nat Genet*. 2011;43(10):1035–9.
46. Van Bel M, Proost S, Wischnitzki E, Movahedi S, Scheerlinck C, De Peer YV, et al. Dissecting plant genomes with the PLAZA comparative genomics platform. *Plant Physiol*. 2012;158(2):590–600.
47. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res*. 2011;39:29–37.
48. Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res*. 2016;44:279–85.
49. Letunic I, Bork P. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res*. 2018;46(D1):D493–6.
50. Marchlerbauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, et al. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res*. 2017;45(D1):D200–3.
51. Hu B, Jin J, Guo A, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics*. 2015;31(8):1296–7.
52. Bailey TL, Boden M, Buske FA, Frith MC, Grant CE, Clementi L, et al. MEME suite: tools for motif discovery and searching. *Nucleic Acids Res*. 2009;37:202–8.
53. Geer LY, Domrachev M, Lipman DJ, Bryant SH. CDART: protein homology by domain architecture. *Genome Res*. 2002;12(10):1619–23.
54. Larkin MA, Blackshields G, Brown NP, Chenna R, Mcgettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007;23(21):2947–8.
55. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol*. 2016;33(7):1870–4.
56. Voorrips RE. MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered*. 2002;93(1):77–8.
57. Krzywinski M, Schein JE, Birol I, Connors JM, Gascoyne RD, Horsman D, et al. Circos: an information aesthetic for comparative genomics. *Genome Res*. 2009;19(9):1639–45.
58. Rozas J, Ferrer-mata A, Sanchez-delbarrio JC, Guirao-rico S, Librado P, Ramosonsins SE, et al. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol Biol Evol*. 2017;34(12):3299–302.

59. Ihaka R, Gentleman R. R: a language for data analysis and graphics. *J Comput Graph Stat.* 1996;5(3):299–314.
60. Koch MA, Haubold B, Mitchell-Olds T. Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (Brassicaceae). *Mol Biol Evol.* 2000;17(10):1483–98.
61. Lee SS, Jung WY, Park HJ, Lee A, Kwon S, Kim H, et al. Genome-wide analysis of alternative splicing in an inbred cabbage (*Brassica oleracea* L.) line 'HO' in response to heat stress. *Curr Genomics.* 2018;19(1):12–20.
62. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods.* 2001;25(4):402–8.
63. Bolger A, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114–20.
64. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods.* 2015;12(4):357–60.
65. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.
66. Hellens RP, Edwards EA, Leyland N, Bean SJ, Mullineaux PM. pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. *Plant Mol Biol.* 2000;42(6):819–32.
67. Sparkes I, Runions J, Kearns A, Hawes C. Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat Protoc.* 2006;1(4):2019–25.
68. Tuskan GA, Difazio SP, Jansson S, Bohlmann J, Grigoriev IV, Hellsten U, et al. The genome of black cottonwood, *Populus trichocarpa* (torr. & gray). *Science.* 2006;313(5793):1596–604.
69. Friedman R, Hughes AL. Gene duplication and the structure of eukaryotic genomes. *Genome Res.* 2001;11(3):373–81.
70. Cheng F, Mandakova T, Wu J, Xie Q, Lysak MA, Wang X. Deciphering the diploid ancestral genome of the mesohexaploid *Brassica rapa*. *Plant Cell.* 2013;25(5):1541–54.
71. Xu G, Guo C, Shan H, Kong H. Divergence of duplicate genes in exon–intron structure. *Proc Natl Acad Sci U S A.* 2012;109(4):1187–92.
72. Maere S, De Bodt S, Raes J, Casneuf T, Van Montagu M, Kuiper M, et al. Modeling gene and genome duplications in eukaryotes. *Proc Natl Acad Sci U S A.* 2005;102(15):5454–9.
73. Cusack BP, Wolfe KH. When gene marriages don't work out: divorce by subfunctionalization. *Trends Genet.* 2007;23(6):270–2.
74. Blanc G, Wolfe KH. Functional divergence of duplicated genes formed by polyploidy during arabidopsis evolution. *Plant Cell.* 2004;16(7):1679–91.
75. Cheng F, Wu J, Fang L, Sun S, Liu B, Lin K, et al. Biased gene fractionation and dominant gene expression among the subgenomes of *Brassica rapa*. *PLoS One.* 2012;7(5):e36442.
76. Birchler JA, Veitia RA. The gene balance hypothesis: from classical genetics to modern genomics. *Plant Cell.* 2007;19(2):395–402.
77. Freeling M. The evolutionary position of subfunctionalization, downgraded. In: *Plant genomes*, vol. 4. Basel: Karger Publishers; 2008. p. 25–40.
78. Freeling M. Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition. *Annu Rev Plant Biol.* 2009;60:433–53.
79. Bowers JE, Chapman B, Rong J, Paterson AH. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature.* 2003;422(6930):433–8.
80. Afoufbastien D, Medici A, Jeauffre J, Coutosthevenot P, Lemoine R, Atanassova R, et al. The *Vitis vinifera* sugar transporter gene family: phylogenetic overview and macroarray expression profiling. *BMC Plant Biol.* 2010;10(1):245.
81. Seo PJ, Park J, Kang SK, Kim S, Park C. An *Arabidopsis* senescence-associated protein SAG29 regulates cell viability under high salinity. *Planta.* 2011;233(1):189–200.
82. Evans JL, Scholes JD. How does clubroot alter the regulation of carbon metabolism in its host? *Asp Appl Biol.* 1995;42:125–32.
83. Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, et al. Ancestral polyploidy in seed plants and angiosperms. *Nature.* 2011;473(7345):97–100.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

