Antarctic Science 31(1), 23–33 (2019) © Antarctic Science Ltd 2019 doi:10.1017/S0954102018000470

Identification, characterization and expression analysis of the chalcone synthase family in the Antarctic moss *Pohlia nutans*

XINGHAO YAO[1,3†], TAILIN WANG[1,4†], HUIJUAN WANG[1,3], HONGWEI LIU[1,3], SHENGHAO LIU[2], QINGANG ZHAO[1], KAOSHAN CHEN[1,3] and PENGYING ZHANG[1,3]

[1]*National Glycoengineering Research Center and School of Life Science, Shandong University, Jinan 250100, China*
[2]*Marine Ecology Research Center, The First Institute of Oceanography, State Oceanic Administration, Qingdao 266061, China*
[3]*Shandong Provincial Key Laboratory of Carbohydrate Chemistry and Glycobiology, Jinan 250100, China*
[4]*Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin 300308, China*
zhangpy80@sdu.edu.cn
†joint lead authorship

Abstract: Mosses have adapted to the Antarctic environment and are an ideal medium for studying plant resistance to abiotic stress. Chalcone synthase is the first committed enzyme in the flavonoid metabolic pathway, which plays an indispensable role in plant resistance to adversity. In this study, six genes (*Pn021*, *PnCHS088*, *Pn270*, *PnCHS444*, *PnCHS768* and *Pn847*) were identified in the Antarctic moss *Pohlia nutans* Lindberg transcriptome by reverse transcription polymerase chain reaction (RT-PCR) and rapid amplification of cDNA ends (RACE). Sequence alignment and three-dimensional structure analysis revealed the conserved amino acid residues of the enzymes of the chalcone synthase family, including three catalytic residues (Cys^{164}, His^{303} and Asn^{336}) and two substrate recognition residues (Phe^{215} and Phe^{265}). Phylogenetic analysis indicated that *PnCHS088*, *PnCHS444* and *PnCHS768* might be chalcone synthase but that *Pn021* is more like stilbenecarboxylate synthase. These genes were located at the transition between fungi and advanced plants in the phylogenetic tree. In addition, real-time PCR analysis revealed that the expression profiles of the six *P. nutans* genes were influenced by diverse abiotic stresses as well as by abscisic acid and methyl jasmonate. The results presented here contribute to the study of the *CHS* gene family in polar mosses and further reveal the mechanisms underlying the adaptation of mosses to extreme environments.

Received 2 May 2018, accepted 4 October 2018, first published online 11 January 2019

Key words: abiotic stress, extreme environment, flavonoids

Introduction

Drought, cold, high salt levels and ultraviolet (UV) radiation in the Antarctic limit plant growth. Mosses are the most widely distributed terrestrial plants in the Antarctic, and have adapted to the extreme environment with the aid of special functional genes and secondary metabolites. Flavonoids are the largest class of polyhydric phenylpropanoid secondary metabolites in plants and include flavonols, flavanones, anthocyanins and isoflavones. Many studies have shown that the accumulation of flavonoids is related to plant resistance to abiotic stress. Enhanced levels of flavonoids lead to increased salt (Oh *et al*. 2011), cold (Choi *et al*. 2009), drought (Castellarin *et al*. 2007) and UV-B (Landrey *et al*. 1995) tolerance in plants. The flavonoid metabolic pathway is an indispensable branch of the phenylpropane metabolic pathway. In a three-step enzymatic reaction, phenylalanine is converted into coumaryol coenzyme A. Then, upon catalysis by chalcone synthase (CHS) and chalcone isomerase (CHI), *p*-coumaroyl-CoA and malonyl-CoA form naringenin, the precursor of all flavonoid compounds.

CHS (EC 2.3.1.74) is the rate-limiting enzyme in the above pathway and catalyses the biosynthesis of naringenin chalcone through the condensation of three units of malonyl-CoA with one *p*-coumaroyl-CoA (Fig. 1). Its expression level affects the downstream metabolism of flavonoids in plants, thereby affecting the plants' ability to grow and resist external stresses. The downregulation of the *CHS* gene results in a decrease in flavonoid content during the development of yellow-seeded *Brassica napus* L. (Jiang *et al*. 2013). Li *et al*. (1993) found that CHS and CHI mutations in *Arabidopsis thaliana* (L.) Heynh. led to the emergence of individuals with UV-hypersensitive phenotypes. Cold stress (4°C) activates *CHS* expression and induces the accumulation of anthocyanin in the leaves and stems in *A. thaliana* (Leyva *et al*. 1995). The expression levels of most *NtCHS* genes are regulated during drought and salinity stresses, which suggests that tobacco *NtCHS* genes might play important roles in responding to these two stresses (Wang *et al*. 2017).

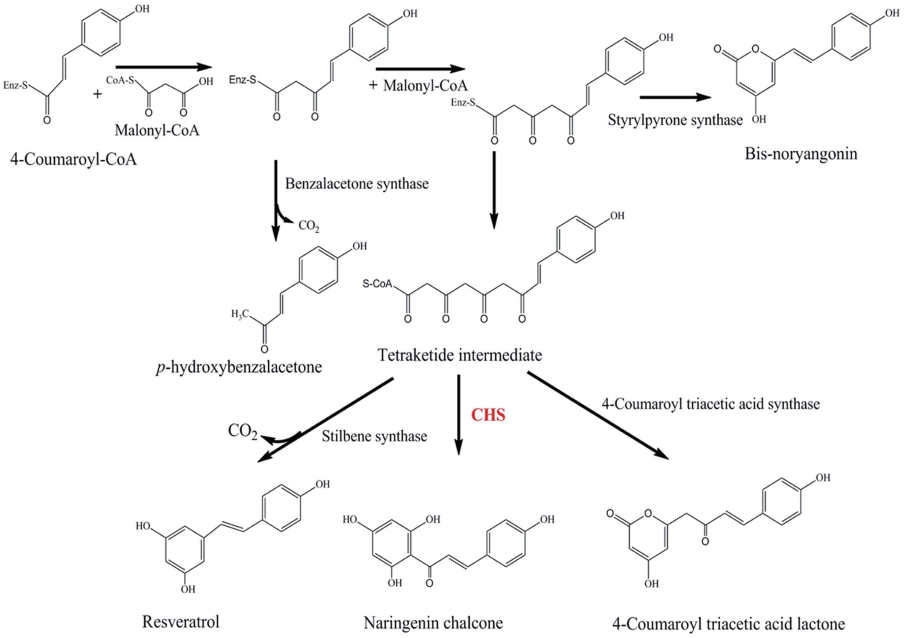

Fig. 1. Illustration of reactions catalysed by CHS and other enzymes of the CHS family.

The *CHS* gene has been widely cloned, and its role in resistance has been studied in multiple plant species; this gene is a single-copy gene in species such as *Arabidopsis*, *Petroselinum crispum* (Mill.) Fuss, and yellow snapdragon. Moreover, there are also multicopy *CHS* genes in *Physcomitrella patens* (Hedw.) Bruch & Schimp., morning glory, chrysanthemum and other plants, which shows that *CHS* genes have gone through varying degrees of duplication and differentiation and have formed a multigene family during the evolutionary process. A genome-wide investigation of *CHS* genes in rice (*Oryza sativa* L.) was performed by Han (2017), and the results revealed that the *CHS* genes exhibit both diversity and conservation in many aspects (Han *et al*. 2017). Chalcone synthase genes exhibit various expression patterns in maize (*Zea mays* L.), indicating functional diversification (Han *et al*. 2016). Surprisingly, although it was the earliest land plant, *Physcomitrella patens*. contains seventeen *CHS* superfamily genes, and duplication and diversification of the *CHS* superfamily genes have occurred in the moss lineage (Koduri *et al*. 2010). The special environment of Antarctica has created unique species. Mosses are the dominant vegetation in the coastal regions of the Antarctic Peninsula and represent the earliest existing land plants; they have been reproducing on land for 470 million years (Roads *et al*. 2014). Recent studies have also shown that Antarctic mosses were the origin of all land plants and have continued to evolve in isolation for tens of millions of years (Convey & Stevens 2007). Another study speculated that in the absence of sexual reproduction in extreme Antarctic conditions, any change in the nucleotide sequence of moss is most likely caused by immigration or mutation *in situ* and could be involved in the evolution of new species (Seppelt *et al*. 1992, Skotnicki *et al*. 2005). Although *CHS* genes have been well studied in advanced plants, no studies have investigated the *CHS* family of genes in the Antarctic moss *Pohlia nutans* Lindberg. In the present study, six gene transcripts were cloned from *P. nutans* and a series of analyses were performed including multiple sequence alignment, structure simulation, subcellular localization, phylogenetic analysis and expression analysis.

Materials and methods

Plant material and stress treatment

The Antarctic moss *P. nutans* was obtained from a scientific investigation of the 24th Chinese National Antarctic Research Expedition. It was inoculated in sterile soil for culture, and placed in an illumination incubator at 16°C with 70% relative humidity and continuous light (109 μmol photons m^{-2} s^{-1}). Mosses cultured under the above conditions were later used as experimental controls (Liu *et al*. 2014). *P. nutans* samples were subjected to different treatments, such as cold (cultured at 4°C for three hours), drought (transferred to a 20% (w/v) Polyethylene Glycol (PEG) 6000 solution (Dingguo Corp., Beijing, China) for six hours), salinity (sprayed once with 200 mM NaCl; the stems and leaves were taken after six hours) and UV-B radiation (exposed under an 8 W UV-B lamp, 302 nm, reaching 80 μW cm^{-2} irradiance for six hours). In addition, the samples were also treated with the plant hormones abscisic acid (ABA)

https://doi.org/10.1017/S0954102018000470 Published online by Cambridge University Press

and methyl jasmonate (MeJA); they were sprayed with 50 μM ABA or 50 μM MeJA and incubated at 16°C for one hour. The stems and leaves of the moss plants were cut and immediately placed in a liquid nitrogen tank (-197°C) to be used as RNA extraction material.

Extraction of total RNA and rapid amplification of cDNA ends

Total RNA was obtained from the stems and leaves of *P. nutans* according to the modified cetyltrimethylammonium bromide method (Gambino *et al*. 2008). Both 5'- and 3'- rapid amplification of cDNA ends (RACE) were performed with a SMART™ RACE cDNA Amplification Kit (Clontech). The above procedures were carried out according to the manufacturer's instructions. Gene-specific primers for 5'- and 3'-RACE reactions were designed from the core cDNA sequence of chalcone synthase. Reverse transcription-polymerase chain reaction (RT-PCR) analysis was performed using total RNA as the template. First, a cDNA library was constructed with 5'/3'-RACE CDS Primer A and PrimeScript™ Reverse Transcriptase (37°C for 60 min). Then, six *P. nutans* genes were amplified with gene-specific primers (see Supplemental Table SI at https://doi.org/10.1017/S0954102018000470), Nested Universal Primer A and TransStart™ FastPfu DNA Polymerase (TransGenBiotech, Beijing). Using the Tiangen gel purification kit (Tiangen, Beijing, China), the 5'- or 3'-RACE products were purified from the gel, ligated into a pEASY-Blunt cloning vector (TransGen Biotech, Beijing), and transferred into competent cells (*Escherichia coli* DH5α). The positive clones were further sequenced, and the generated sequence files were compiled. The full-length cDNA sequences of the six *P. nutans* genes were identified and assembled by using the SeqMan program (DNAStar 7.1, DNASTAR Inc., USA).

Gene information acquisition, multiple sequence alignment and three-dimensional structure construction

Sequencing of the transcriptome of *P. nutans* was completed by using Illumina ultra-high throughput sequencing technology. Data were retrieved from the NCBI Short Read Archive (SRA) (accession number SRA051595). Chal_sti_synt_N (PF02797) and Chal_sti_synt_C (PF00195) domain models built with the hidden Markov model were downloaded from the Sanger Centre (http://pfam.sanger.ac.uk/). First, using the HMMER package according to the provided instructions, homologous protein sequences were obtained with the program hmmsearch. Then the genes with the highest scores above the threshold (with the highest identity) were chosen as target genes. Finally, the six full-length gene sequences of *P. nutans* (*Pn021*, *PnCHS088*, *Pn270*, *PnCHS444*, *PnCHS768* and *Pn847*) were obtained and then annotated further. The GC (guanine-cytosine) content of each nucleic acid sequence was analysed with DNASTAR software. The molecular weight and isoelectric point were calculated with the Swiss Biological Information Institute (SIB) ExPASy Compute PI/Mw online server, and functional annotation was performed based on protein sequence from NCBI basic local alignment search tool (BLAST) searches and related literature. Multiple sequence alignments were performed with the BioEdit Clustal-W program, and GeneDoc was used to show the active site. Finally, MEGA (5.0) was used to perform phylogenetic analysis. All parameters used were the default parameters for the programs.

For the construction of the three-dimensional structure, the crystal structures of *Marchantia polymorpha* L. stilbenecarboxylate synthase 2 (STCS2) and polyketide synthase 1 of *Huperzia serrata* (Thunb. ex Murray) Trevis. were used as templates, and homology modelling was performed using the SWISS-MODEL website. AutoDock software was used to predict the interactions with the substrate molecules. Finally, PyMOL software was used for visual observation and analysis. The key active amino acids were marked with a stick model.

Subcellular localization

The subcellular localization of PnCHS444 was examined with an N-terminal green fluorescent protein (GFP) tag. The gene was subcloned into a modified pBI221-GFP vector and the recombinant plasmid was transformed into *Escherichia coli* (Migula) Castellani and Chalmers (Zhang *et al*. 2000). The gene-specific PCR primers are listed in Table SI (at https://doi.org/10.1017/S0954102018000470). Using a Biomiga Plasmid Maxprep Kit (www.biomiga.com), the recombinant plasmid was prepared. Next, the plasmid was transferred into the protoplasts of *Arabidopsis*. Transformed tissues were cultured in a dark room at 23°C and collected after ten hours. They were observed with a confocal laser microscope. (Zeiss LSM 700 Inverted Confocal Microscope).

Phylogenetic analysis

After obtaining the full-length sequences of the genes, the codons were used to deduce protein sequences and conduct a cophylogenetic analysis with 40 representative CHS protein sequences. These sequences were from the protein database of NCBI and included proteins from plants, fungi and bacteria. Finally, a phylogenetic tree including the six proteins and chalcone synthase family

members from other plants was constructed using MEGA 6 software (MEGA, Raynham, MA, USA) with the neighbour-joining method.

Real-time PCR was used to analyse mRNA expression levels

The cDNA sequences were constructed with total RNA (500 ng) and a mixture of oligo(dT)18 (Biasini *et al.* 2014) and random hexamer primers. Gene-specific primers were designed with Primer 5 (http://www.premierbiosoft.com). Real-time PCR assays were conducted using Bestar SybrGreen qPCR Master Mix (DBI Bioscience, Shanghai, China). The PCR program was 95°C for 20 s, 58°C for 20 s and 72°C for 20 s. To check the PCR specificity, the program ended with the generation of a melt curve by ramping from 65°C to 95°C in 2 s intervals. All reactions were completed in triplicate. The comparative expression levels of the selected genes were normalized to the expression of the tubulin gene. Ct (threshold cycle) values were used to quantify the mRNA expression of the six *P. nutans* genes.

Results

The retrieval, identification and molecular cloning of chalcone synthase family genes from the Antarctic moss Pohlia nutans

To obtain the *CHS* family genes of the Antarctic moss *P. nutans*, a homologous gene mode was used to screen the transcriptome database. Using HMMER software, homologous protein sequences were obtained with the program hmmsearch. After comparing their annotations in the NCBI database, the *CHS* family genes of *P. nutans* were filtered and manually identified. The target sequence was subjected to a BLAST search on the NCBI website and annotated. However, only *Pn021* had a full-length nucleotide sequence. For the other five genes, full-length gene amplification was performed under the guidance of RACE technology and the full-length cDNAs of the *CHS* genes were further sequenced. The sequences have been deposited in the NCBI database, and their accession numbers are given in Table I.

The Swiss Institute of Bioinformatics (SIB) online server was used for analysis; the nucleic acid GC content of the six genes ranged from 52.0–65.0%. In addition, the lengths of the six amino acid sequences ranged from 355–423 residues. The molecular weights of the six proteins ranged from 38.4–45.5 KDa, and the range of isoelectric points (pI) was 6.28–6.87 (Table I).

Multiple sequence alignment of the chalcone synthase superfamily

To compare the sequences with several known CHSs from other species, the *CHS* homologous genes of other species were found in the public NCBI EST and genomic databases (Fig. 2). The CHS homologous protein sequences of six species, including *Medicago sativa* L., *A. thaliana*, *Pinus sylvestris* L., *Psilotum nudum* (L.) P.Beauv., *Marchantia paleacea* L. and *P. patens* were compared to those of the six *P. nutans* genes. The identity of the six *P. nutans* genes at the amino acid level with the *CHS* genes of other species was 30.0–65.0%. The conserved amino acids in the six *P. nutans* genes are consistent with those conserved in the type III polyketide synthases (PKSs). *PnCHS088*, *PnCHS444* and *PnCHS768* contain two typical polypeptide sequences of the CHS family (GCFAGGTVLR and GFGPGL), the catalytic triad (Cys^{164}, His^{303} and Asn^{336}) and phenylalanine residues (Phe^{221} and Phe^{271}) associated with substrate-specific binding. All of the above amino acid residues are important for the cyclization reactions in CHS/STS-type PKSs, and these amino acid residues are important for starter substrate selectivity. However, the Cys^{177} catalytic site of Pn021 was replaced by Phe^{177}, and ALGPG was substituted with GLGPG. GFGPG in Pn847 was replaced with SFGPG, which may have caused differences in the starter molecules. Additionally, Pn270 and Pn847 appeared as 2'-oxoalkylresorcinol synthase (ORS) orthologues in BLAST searches.

Three-dimensional structure simulation for PnCHS444

To characterize the spatial structural properties of PnCHS444, a 3D homology model was developed based on the crystal structures of *M. polymorpha*. STCS2 and

Table I. Sequence information of the six *P. nutans* genes.

Gene ID	Protein length (AA)	Isolectric point (pI)	Molecular weight (MW)	(C + G) %Content	Results of blast annotation	GeneBank Accession Number
Pn021	420	6.282	45482.62	60.65	stilbenecarboxylate synthase	MH715894
Pn088	400	6.611	42742.38	64.75	chalcone synthase	MH715895
Pn270	355	6.874	38465.4	52.25	alkylresorcinol synthase	MH715896
Pn444	423	6.858	45314.23	65.09	chalcone synthase	MH715897
Pn768	393	6.815	42100.73	63.54	chalcone synthase	MH715898
Pn847	404	6.654	43707.88	59.09	alkylresorcinol synthase	MH715899

https://doi.org/10.1017/S0954102018000470 Published online by Cambridge University Press

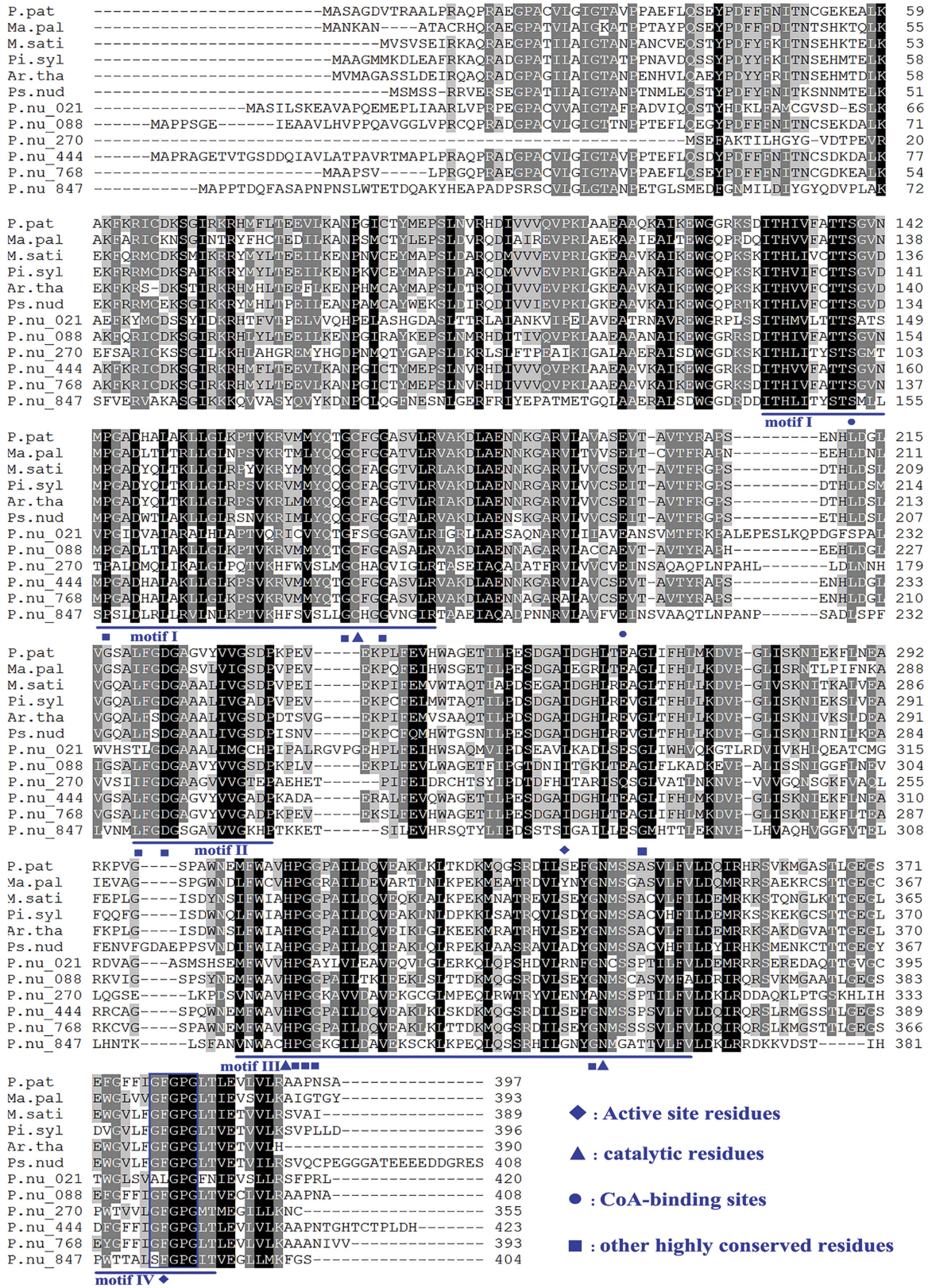

Fig. 2. Amino acid alignment showing the functional residues in the six *P. nutans* proteins. Identical amino acid residues are shaded in black; similar residues are shaded in grey. The active site residues, catalytic site, CoA binding site and other highly conserved amino acid sequences are marked with blue symbols, and the signature GFGPG loop is shown in the blue box. Reference sequences: *Medicago sativa*, P30074; *Arabidopsis thaliana*, CAI30418; *Pinus sylvestris*, CAA43166; *Psilotum nudum*, BAA87922; *Marchantia paleacea*, BAD42328; and *Physcomitrella patens*, CHS13, ABU87504.

https://doi.org/10.1017/S0954102018000470 Published online by Cambridge University Press

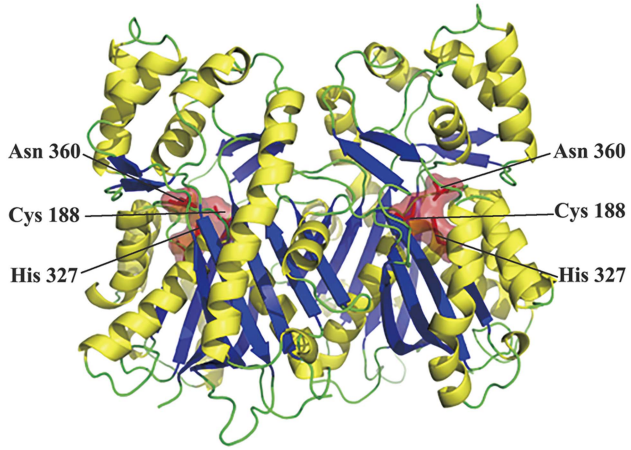

Fig. 3. 3D structure modelling of PnCHS444. α-helices are in yellow, β-sheets are in blue, and loops are in light green. The superimposed catalytic site residues for the 3D structure are in red and are annotated individually.

polyketide synthase 1 from *Huperzia serrata*. The three-dimensional structure of PnCHS444 was predicted with the SWISS-MODEL server.

As shown in Fig. 3, the 3D structure homology of PnCHS444 exhibits an alpha beta 3-layer (αβα) sandwich architecture and possesses the typical catalytic site containing the conserved residues of Cys^{188}, His^{327} and Asn^{360}. The 3D model results of the other five proteins also demonstrated similar structures, and so the structure of PnCHS444 was used as the representative structure. In general, the 3D structures of homologous proteins are more conserved than their amino acid sequences. There was no significant difference in the core conserved sites. In the 3D structure of CHS, the CoA binding tunnel binds *p*-coumaroyl-CoA, which provides access to the different catalytic site residues (marked as red in Fig. 3), catalysing the loading of the *p*-coumaroyl starter unit and the decarboxylative condensation reactions to extend the polyketide.

Subcellular localization of PnCHS444 in Arabidopsis mesophyll protoplasts

The distribution of PnCHS444 in the cells provides an important reference. A subcellular localization vector for PnCHS444 was constructed, and its production was significantly increased upon UV-B irradiation. The data revealed that GFP fluorescence was spread over the entire protoplast in the p35S::GFP/pBI221-carrying (i.e. vector control) protoplasts (Fig. 4a). However, for p35S::PnCHS444::GFP containing protoplasts, the fluorescence of the PnCHS444-GFP fusion protein was only distributed in the cell membrane and in the membranes of the endothelial organelles (Fig. 4e).

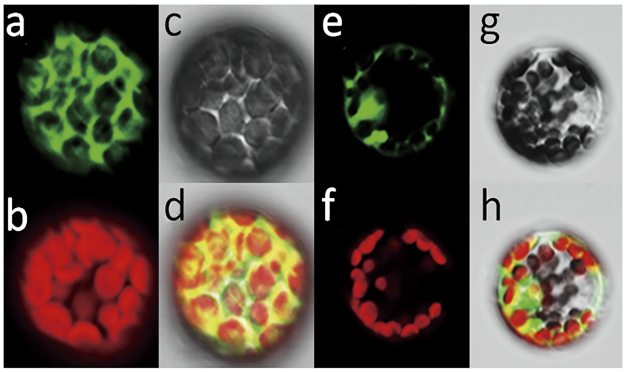

Fig. 4. Subcellular localization of *PnCHS444* in *Arabidopsis* protoplasts. **a**, **b**, **c & d.** The control p35S::GFP/pBI221 vector. **e**, **f**, **g & h.** The constructed p35S::*PnCHS444*::GFP/pBI221 vector. **a.** The fluorescence of the control vector was spread over the entire cell. **e.** The fluorescence of the fusion protein was only distributed in the cell membrane and the membranes of the endothelial organelles. **b & f.** Fluorescence of the chloroplast. **d & h.** The combined fluorescence of GFP and the chloroplast. **c & g.** Bright field of the protoplast. The confocal images were taken ten hours after protoplast transfection.

Phylogenetic analysis of the chalcone synthase superfamily

Phylogenetic analysis was performed with deduced amino acid sequences of the six *P. nutans* proteins of the CHS family, as well as for a range of catalytically verified and putative chalcone synthases from other species, including advanced plants, ferns, bryophytes, algae and fungi. Most species have multiple *CHS* family genes. These multiple genes are prone to mutation or ectopic exchange with other related gene copies, which is important in plant evolution.

The phylogenetic tree showed that the CHS family of advanced plants forms a sister clade to that of ferns, and the CHS family of mosses is in the transition zone of the phylogenetic tree. The CHS family of algae and fungi is located at the bottom of the phylogenetic tree. Moss enzymes clustered together and showed high sequence similarities. The six genes of *P. nutans* clustered in three different small branches (Fig. 5). The Pn021 branch (marked with a blue triangle) was longer than others and was separate from the rest, suggesting a larger change during the evolutionary process. However, the branch length for Pn270 was short, indicating that few nucleotide changes occurred. The branches of Pn270 and Pn847 shared the same internal nodes, indicating that Pn270 and Pn847 have closely related lineages. The PnCHS768, PnCHS444 and PnCHS088 branches shared the same node as PpCHS, a bona fide CHS from *P. patens*, indicating that those proteins share the same ancestor genes. The CHSs of bryophytes are clustered between

https://doi.org/10.1017/S0954102018000470 Published online by Cambridge University Press

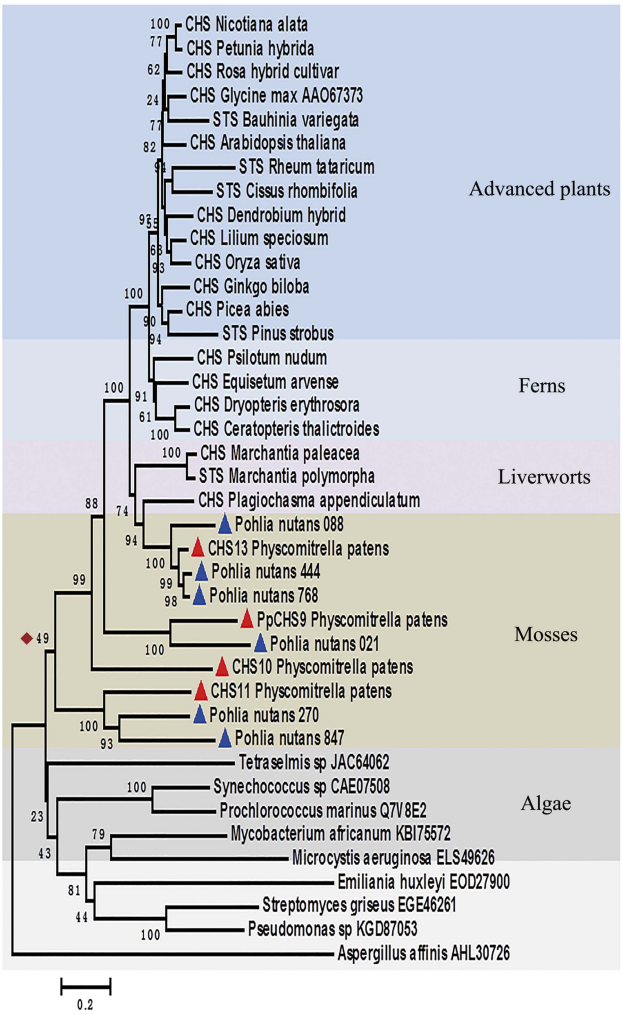

Fig. 5. Phylogenetic tree of the CHS family. The numbers in the branches indicate the reliability level. The accession numbers of the enzymes included in the analysis are shown below. *Physcomitrella patens* (PpCHS9, A9RSI0), *P. patens* (CHS13, ABB84527), *P. patens* (CHS10, XP001781520), *P. patens* (CHS11, ABU87504), *Plagiochasma appendiculatum* Lehm. & Lindenb. (CHS, AIV42295), *Arabidopsis thaliana* (CHS, AED91961), *Glycine max* (L.) Merr. (CHS, AAO67373), *Equisetum arvense* L. (CHS, BAA89501), *Ginkgo biloba* L. (CHS, AAY52458), *Picea abies* (L.) H. Karst. (CHS, ABD24228), *Rosa hybrid cultivar* (CHS, BAC66467), *Lilium speciosum* Thunb. (CHS, BAE79201), *Dendrobium hybrid cultivar* (CHS, AAU93767), *Oryza sativa Japonica Group* (CHS, BAA19186), *Psilotum nudum* (L.) P.Beauv. (CHS, Q9SLY0), *Dryopteris erythrosora* (D.C.Eaton) Kuntze. (CHS, AHU87077), *Ceratopteris thalictroides* (L.) Brongn. (CHS, AFN02448), *Nicotiana alata* Link & Otto. (CHS,C5IWS6), *Petunia* × *hybrida* hort. ex E.Vilm. (CHS, P08894), *Marchantia polymorpha* L. (STCS, AAW30009), *Pinus strobus* L. (STS, CAA87013), *Nicotiana sylvestris* Speg. & Comes. (CHSL, CAA74847), *Rheum tataricum* L.f. (STS, AAP13782), *Cissus rhombifolia* Vahl. (STS, AAM21772), *Bauhinia variegate* (L.) Benth. (STS, ABF59517), *Streptomyces griseus* (Krainsky) Waksman and Henrici. (EGE46261), *Pseudomonas* Migula. (KGD87053), *Aspergillus affinis* Davolos, Persiani, Pietrangeli & Maggi. (AHL30726), *Mycobacterium tuberculosis variant africanum* Castets *et al.*, ATCC 25420 (KBI75572), *Tetraselmis* F.Stein. (JAC64062), *Emiliania huxleyi* (Lohm.) Hay and Mohler. (EOD27900), *Microcystis aeruginosa* Kützing. (ELS49626), *Synechococcus* Nägeli. (CAE07508), *Prochlorococcus marinus* Chisholm. (Q7V8E2).

those of fungi and higher plants. Even in the inner family of bryophytes, the separation of genes was not clear; the same situation was observed for the algae. However, the origin node of these branches or the root of the cluster was the same (marked with a red diamond), and it can be inferred that the higher plant chalcone genes originated from a common ancestor gene.

Expression analysis of six genes of the Antarctic moss Pohlia nutans

To investigate the responses of the six *P. nutans* genes to abiotic stress, their expression levels were analysed by real-time PCR. As shown in Fig. 6, the expression levels of the six *P. nutans* genes were markedly different.

https://doi.org/10.1017/S0954102018000470 Published online by Cambridge University Press

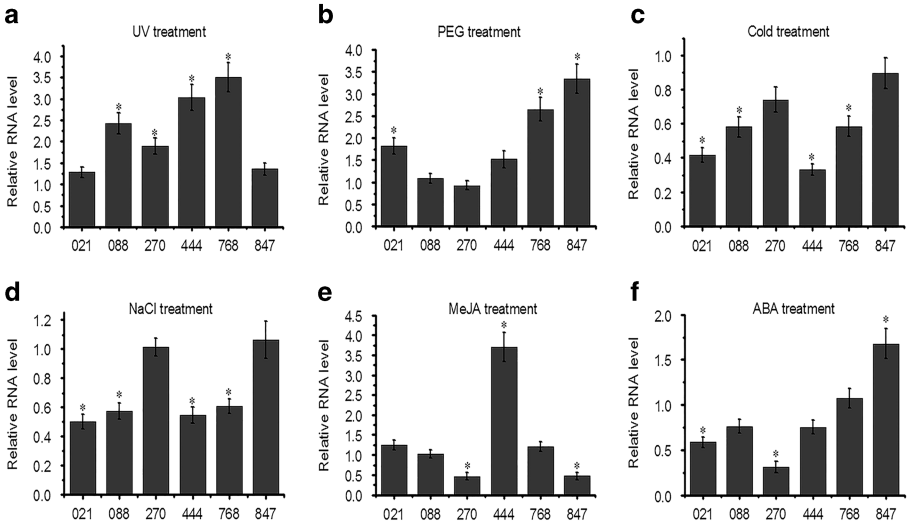

Fig. 6. Expression levels of the six *P. nutans* family genes after treatment of **a.** UV for 6 h, **b.** 20% PEG for 6 h, **c.** 4°C for 3 h, **d.** 200 mM NaCl for 6 h, **e.** 50 μM MeJA for 1 h, or **f.** 50 μM ABA for 1 h. The vertical bars indicate the means ± SD ($n = 3$). The asterisks (*) signify differences in means between the control group and the treatment group (Student's *t*-test, $P \leq 0.05$).

After UV treatment, the expression levels of the six *P. nutans* genes increased by different magnitudes. The expression levels of *PnCHS088* and *PnCHS444* increased by factors of 2.4 and 3.0, respectively (Fig. 6a). After simulated drought stress using PEG6000, the expression levels of *PnCHS088* and *Pn270* remained essentially unchanged; those of the other four genes were upregulated. The expression levels of *Pn021*, *PnCHS444* and *Pn847* increased by factors of 1.8, 1.5 and 3.3, respectively (Fig. 6b). In addition, the expression levels of the six *P. nutans* genes were all downregulated under low-temperature conditions (Fig. 6c). With the exception of *Pn270/847*, the other four *P. nutans* genes were all downregulated under high-salinity treatment. The expression levels of *Pn021*, *PnCHS088*, *PnCHS444* and *PnCHS768* were reduced to 0.50, 0.57, 0.54 and 0.62 times their previous expression levels, respectively (Fig. 6d).

As molecular signals, phytohormones effectively control the response of stress-related proteins and affect plant resistance. The flavonoid metabolic pathway was regulated by *Myb*. Moreover, *Myb* was affected by abscisic acid and methyl jasmonate. The expression levels of *PnCHS444* and *Pn847* increased by factors of 3.6 and 1.7, respectively, after ABA and MeJA treatment (Fig. 6e & f).

Discussion

Flavonoids are important secondary metabolites that are synthesized by different enzymes and take part in many physiological processes in plants, including disease resistance, auxin transport and abiotic stress responses. As CHS is the rate-limiting enzyme in the flavonoid metabolic pathway, the expression levels of CHS affect the metabolism of flavonoids and the ability of plants to grow and resist external stresses. Low temperature induces *CHS* gene expression, resulting in anthocyanin accumulation in petunia flowers (Shvarts *et al.* 1997). Overexpression of the *EaCHS1* (*Eupatorium adenophorum* Spreng. *CHS1*) gene in tobacco enhances its tolerance to salinity stress (Chen *et al.* 2015). The *CHS* gene is actively expressed in parsley, which accumulates flavonoids in leaf epidermal cells to protect against harmful UV-B radiation (Schmelzer *et al.* 1988). The ability of Antarctic mosses to adapt to the extreme polar environment is due to distinct functional genes and different gene expression profiles that correspond to stressful conditions. In this study, six *P. nutans* genes were isolated, and these genes possessed a relatively conserved coding region of approximately 1.2 kb (Fig. 1 & Table I), which is approximately the same length as the coding regions of *SoCHS1* (*Syringa oblata* Lindl, 1.17 kb), *LrCHS* (*Lamiophlomis rotata* (Bentham ex J.D. Hooker) Kudô, 1.17 kb), and *PpCHS* (*Physcomitrella patens*, 1.19 kb) and consistent with *CHS* genes identified in other plant lineages. *PnCHS088*, *PnCHS444* and *PnCHS768* contain two typical polypeptide sequences (GCFAGGTVLR and GFGPGL), a catalytic triad (Cys^{164}, His^{303} and Asn^{336}) and substrate-specific binding sites (Phe^{221} and Phe^{271}) (Jez *et al.* 2002). However, *Pn021*, *Pn270* and *PnCHS847* are not completely conserved, showing evidence of base pair substitution or deletion. The Cys^{177} catalytic site of Pn021 was replaced by Phe^{177} (see Fig. 2), a phenomenon which is rare in the type III PKSs. Pn021 might not exhibit catalytic activity similar to typical type III PKSs, although the annotation of Pn021 upon BLAST search

https://doi.org/10.1017/S0954102018000470 Published online by Cambridge University Press

showed it to be one of the stilbenecarboxylate synthase (STCS) family members. Genetic mutations in *P. nutans* may be caused by its prolonged exposure to extreme conditions. This is consistent with Dale's assumption that genetic variation occurring in *Hennediella heimii* (Hedw.) R.H. Zander is possibly due to increased exposure to UV-B irradiation (Dale *et al*. 1999). Through the construction of 3D models for the PnCHS proteins, the higher-order structures and catalytic sites of the proteins were predicted more thoroughly. For the catalytic site, the typical Cys-His-Asn catalytic triad is conserved in the CHS proteins of *P. nutans*, with an exception of PnCHS021 (Fig. 3). Studies have shown that four amino acids, Cys^{164}, Phe^{215}, His^{303} and Asn^{336} situated at the intersection of the CoA-binding tunnel and the active site cavity play an essential and distinct role during the decarboxylation of malonyl-CoA and the formation of chalcone. Cys^{164} plays a role as the active-site nucleophile in polyketide formation. The importance of His^{303} and Asn^{336} in the malonyl-CoA decarboxylation reaction has also been demonstrated. Phe^{215} may help orient substrates at the active site during elongation of the polyketide intermediate (Jez *et al*. 2000). The distribution of CHS in cells varies among species. Subcellular localization analysis showed that the distribution of PnCHS444 is similar to that of the CHS of buckwheat (Hrazdina *et al*. 1987), which is distributed in the cell membrane and in the membranes of endothelial organelles (Fig. 4). However, the CHS of grape berry is localized in the rough endoplasmic reticulum (ER) and the cytoplasm of the skin cells, with low localization in the cell wall (Tian *et al*. 2008). The CHS of Arabidopsis is predominantly distributed in the cytosol (Zhang *et al*. 2017), and was a cytoplasmic and nuclear-localized enzyme for the local synthesis of flavonoids (Saslowsky & Winkel-Shirley 2001).

The importance of gene duplication in evolution is well documented (Zhang 2003). *CHS9*, *CHS10*, *CHS11* and *CHS13* from *P. patens* have been characterized. Gene duplications gave rise to the progenitor of *PpCHS11* and the ancestor of all other *CHS* superfamily genes. *Pn270* and *Pn847* are in the same evolutionary position as *PpCHS11* (Fig. 5). *CHS11* is a 2'-oxoalkylresorcinol synthase (*PpORS*) (Kim *et al*. 2013). *Pn270* and *Pn847* may be (oxo)alkylresorcinol synthases similar to *PpORS*, and they would be the first enzymes belonging to the same basal clade as *PpORS*. *PpORS* and its two paralogues have thus far been the only members of the basal clade of phylogenetic trees of plant type III PKSs. According to previous studies, fungal type III polyketide synthases (Kim *et al*. 2013) and *PpCHS11* have similar catalytic activity. Although flavonoids have been detected in microalgae (Goiris *et al*. 2014), CHS-like enzymes (CHSL) extracted from the algae cannot catalyse 4-coumaroyl-CoA and malonyl-CoA to form naringin chalcone. Compared with PpORS, CHSL is only similar in terms of substrate selectivity (Baharum *et al*. 2011), which might provide an important reference for the evolution of the *CHS* family genes and the initial formation of the flavonoid pathway. *CHS13* is a bona fide *CHS* (Jiang *et al*. 2006). *PnCHS088*, *PnCHS444* and *PnCHS768* are near *P. patens* CHS13 in the phylogenetic tree (Fig. 5). Although biochemical characterization is necessary, *PnCHS088*, *PnCHS444* and *PnCHS768* are probably bona fide *CHS* family genes. *Pn021* might encode a non-CHS enzyme since it is near PpCHS9 (Koduri *et al*. 2010) in the tree, which is more similar to STCS, although this has yet to be demonstrated. The tree suggests that diversification of the type III PKSs, including the appearance of bona fide *CHS* genes, might have preceded the divergence of the moss species.

In nature, plants are exposed to a variety of abiotic stresses, such as drought, cold, high salt, and UV radiation. *CHS* is very commonly induced in many plant species under abiotic stress, resulting in changes in expression levels. The enhanced expression of *CHS* enhances the accumulation of flavonoids and isoflavonoids (Dao *et al*. 2011). Keeping *Arabidopsis* under high-intensity light conditions for 24 hours causes a fiftyfold increase in CHS activity and an accumulation of anthocyanin pigments (Feinbaum & Ausubel 1988). *PpORS* was recently shown to confer dehydration resistance (Li *et al*. 2018). *Pn847* (a putative *PpORS* orthologue) also showed a similar phenomenon under PEG treatment. In this study, abiotic stress promoted the upregulation of several *CHS* family genes in Antarctic moss. In addition, several CHS family genes were also downregulated after stress treatment (Fig. 6). The *CHSA* and *CHSJ* paralogues of *Nicotiana tabacum* L. have different expression patterns under abiotic stresses (Chen *et al*. 2015). Phytohormones are a class of active substances that are induced by specific environmental signals in plant cells and can regulate the physiological responses of plants. ABA and MeJA are the main determinants of enhanced plant tolerance to abiotic stress (Zhu 2002), and they influence the biosynthesis of flavonoids (Xu *et al*. 2015). Tobacco plants accumulate more flavonoids under ABA and MeJA treatment than under control conditions (Chen *et al*. 2017). As shown in Fig. 6e & f, the mRNA production levels of six *Pohlia nutans* genes under ABA and MeJA treatment were different from control conditions. These results were similar to those for the family of flavonoid 3'-hydroxylase and flavonoid 3', 5'-hydroxylase genes from *P. nutans* (Liu *et al*. 2014).

Based on the results of this study, it is concluded that the expression profiles of these genes were influenced by diverse abiotic stresses. This study also provides an important reference regarding the adaptations of Antarctic mosses to the polar environment.

https://doi.org/10.1017/S0954102018000470 Published online by Cambridge University Press

Conclusions

Pn088, *Pn444* and *Pn768* were identified and analysed as putative CHS genes. *Pn021* was found to be similar to STCS, and the other two genes (*Pn270* and *Pn847*) were suggested to be *ORS* orthologues. They are all involved in regulating abiotic stress tolerance in *P. nutans*. Such changes in expression are closely related to adaptation to extreme environments.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (41206176 and 41476174); the Natural Science Foundation of Shandong Province (ZR2017MD013 and ZR2014DQ012); the Basic Scientific Fund for National Public Research Institutes of China (2014T04); the Excellent Creative Team Fund of Jinan; and the Fund for Transformation of Agricultural Scientific and Technological Achievements of Shandong Province. Thanks to the School of Life Science of Shandong University for providing the advanced platform, and to the reviewers for their helpful comments that significantly improved the paper. Thanks also to Xu Li for helping with the construction of 3D models.

Author contribution

Xinghao Yao and Tailin Wang conceived and, together with Kaoshan Chen and Pengying Zhang, designed the experiments. Huijuan Wang and Hongwei Liu performed the experiments and analysed the data. Shenghao Liu and Qingang Zhao contributed materials and analysis tools, and Xinghao Yao wrote the paper.

Supplemental material

Table SI. showing primers for RACE, plasmid construction and quantitative real-time PCR analysis will be found at http://dx.doi.org/10.1017/S0954102018000470

References

Baharum, H., Morita, H., Tomitsuka, A., Lee, F.C., Ng, K.Y., Rahim, R.A., *et al.* 2011. Molecular cloning, modeling, and site-directed mutagenesis of type III polyketide synthase from *Sargassum binderi* (Phaeophyta). *Marine Biotechnology*, **13**, 845–856.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., *et al.* 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research*, **42**(Web Server issue), W252–W258.

Castellarin, S.D., Matthews, M.A., Gaspero, G.D. & Gambetta, G.A. 2007. Water deficits accelerate ripening and induce changes in gene expression regulating flavonoid biosynthesis in grape berries. *Planta*, **227**, 101–112.

Chen, L.J., Guo, H.M., Lin, Y. & Cheng, H.M. 2015. Chalcone synthase EaCHS1 from *Eupatorium adenophorum* functions in salt stress tolerance in tobacco. *Plant Cell Reports*, **34**, 885–894.

Chen, S., Pan, X.H., Li, Y.T., Cui, L.J., Zhang, Y.C., Zhang, Z.M. *et al.* 2017. Identification and characterization of chalcone synthase gene family members in *Nicotiana tabacum*. *Journal of Plant Growth Regulation*, **36**, 374–384.

Choi, S., Kwon, Y.R., Hossain, M.A., Hong, S.W., Lee, B.H. & Lee, H. 2009. A mutation in ELA1, an age-dependent negative regulator of PAP1/MYB75, causes UV- and cold stress-tolerance in *Arabidopsis thaliana* seedlings. *Plant Science*, **176**, 678–686.

Convey, P. & Stevens, M.I. 2007. Antarctic biodiversity. *Science*, **317**, 1877–1878.

Dale, T.M., Skotnicki, M.L., Adam, K.D. & Selkirk, P.M. 1999. Genetic diversity in the moss *Hennediella heimii* in Miers Valley, Southern Victoria Land, Antarctica. *Polar Biology*, **21**, 228–233.

Dao, T.T.H., Linthorst, H.J.M. & Verpoorte, R. 2011. Chalcone synthase and its functions in plant resistance. *Phytochemistry Reviews*, **10**, 397–412.

Feinbaum, R.L. & Ausubel, F.M. 1988. Transcriptional regulation of the *Arabidopsis thaliana* chalcone synthase gene. *Molecular and Cellular Biology*, **8**, 1985–1992.

Gambino, G., Perrone, I. & Gribaudo, I. 2008. A rapid and effective method for RNA extraction from different tissues of grapevine and other woody plants. *Phytochemical Analysis*, **19**, 520–525.

Goiris, K., Muylaert, K., Voorspoels, S., Noten, B., Paepe, D.D., Baart, G.J.E. & Cooman, L.D. 2014. Detection of flavonoids in microalgae from different evolutionary lineages. *Journal of Phycology*, **50**, 483–492.

Han, Y.H., Cao, Y.P., Jiang, H.Y. & Ding, T. 2017. Genome-wide dissection of the chalcone synthase gene family in *Oryza sativa*. *Molecular Breeding*, **37**, 119.

Han, Y.H, Ding, T., Su, B. & Jiang, H.Y. 2016. Genome-wide identification, characterization and expression analysis of the chalcone synthase family in maize. *International Journal of Molecular Sciences*, **17**, 161.

Hrazdina, G., Zobel, A.M. & Hoch, H.C. 1987. Biochemical, immunological, and immunocytochemical evidence for the association of chalcone synthase with endoplasmic reticulum membranes. *Proceedings of the National Academy of Sciences of the United States of America*, **84**, 8966–8970.

Jez, J.M., Bowman, M.E. & Noel, J.P. 2002. Expanding the biosynthetic repertoire of plant type III polyketide synthases by altering starter molecule specificity. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 5319–5324.

Jez, J.M., Ferrer, J.L., Bowman, M.E., Dixon, R.A. & Noel, J.P. 2000. Dissection of malonyl-coenzyme A decarboxylation from polyketide formation in the reaction mechanism of a plant polyketide synthase. *Biochemistry*, **39**, 890–902.

Jiang, C.G., Schommer, C.K., Kim, S.Y. & Suh, D.Y. 2006. Cloning and characterization of chalcone synthase from the moss, *Physcomitrella patens. Phytochemistry*, **67**, 2531–2540.

Jiang, J.J., Shao, Y.L., Li, A.M., Lu, C.L., Zhang, Y.T. & Wang, Y.P. 2013. Phenolic composition analysis and gene expression in developing seeds of yellow- and black-seeded *Brassica napus*. *Journal of Integrative Plant Biology*, **55**, 537–551.

Kim, S.Y., Che, C.C., Wiedemann, G., Jepson, C., Rahimi, M., Rothwell, J.R., *et al.* 2013. *Physcomitrella* PpORS, basal to plant type III polyketide synthases in phylogenetic trees, is a very long chain 2'-oxoalkylresorcinol synthase. *Journal of Biological Chemistry*, **288**, 2767–2777.

Koduri, P.K.H., Gordon, G.S., Barker, E.I., Colpitts, C.C., Ashton, N.W. & Suh, D.Y. 2010. Genome-wide analysis of the chalcone synthase superfamily genes of *Physcomitrella patens*. *Plant Molecular Biology*, **72**, 247–263.

Landrey, L.G., Chapple, C.C.S. & Last, R.L. 1995. *Arabidopsis* mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-B injury and oxidative damage. *Plant Physiology*, **109**, 1159–1166.

Leyva, A., Jarillo, J.A., Salinas, J. & Martinez-Zapater, J.M. 1995. Low temperature induces the accumulation of phenylalanine ammonia-lyase and chalcone synthase mRNAs of *Arabidopsis thaliana* in a light-dependent manner. *Plant Physiology*, **108**, 39–46.

Li, L., Aslam, M., Rabbi, F., Vanderwel, C., Ashton, N.W. & Suh, D.Y. 2018. PpORS, an ancient type III polyketide synthase, is required for integrity of leaf cuticle and resistance to dehydration in the moss, *Physcomitrella patens. Planta*, **247**, 527–541.

Li, J., Ou-Lee, T.M., Raba, R., Amundson, R.G. & Last, R.L. 1993. Arabidopsis flavonoid mutants are hypersensitive to UV-B irradiation. *Plant Cell*, **5**, 171–179.

Liu, S.H., Ju, J.F. & Xia, G.M. 2014. Identification of the flavonoid 3′-hydroxylase and flavonoid 3′,5′-hydroxylase genes from Antarctic moss and their regulation during abiotic stress. *Gene*, **543**, 145–152.

Oh, J.E., Kim, Y.H., Kim, J.H., Kwon, Y.R. & Lee, H.J. 2011. Enhanced level of anthocyanin leads to increased salt tolerance in Arabidopsis *PAP1-D* plants upon sucrose treatment. *Journal of the Korean Society for Applied Biological Chemistry*, **54**, 79–88.

Roads, E., Longton, R.E. & Convey, P. 2014. Millennial timescale regeneration in a moss from Antarctica. *Current Biology*, **24**, 222–223.

Saslowsky, D. & Winkel-Shirley, B. 2001. Localization of flavonoid enzymes in Arabidopsis roots. *Plant Journal*, **27**, 37–48.

Schmelzer, E., Jahnen, W. & Hahlbrock, K. 1988. *In situ* localization of light-induced chalcone synthase mRNA, chalcone synthase, and flavonoid end products in epidermal cells of parsley leaves. *Proceedings of the National Academy of Sciences of the United States of America*, **85**, 2989–2993.

Seppelt, R.D., Green, T.G.A., Schwartz, A.M.J. & Frost, A. 1992. Extreme southern locations for moss sporophytes in Antarctica. *Antarctic Science*, **4**, 37–39.

Shvarts, M., Borochov, A. & Weiss, D. 1997. Low temperature enhances petunia flower pigmentation and induces chalcone synthase gene expression. *Physiologia Plantarum*, **99**, 67–72.

Skotnicki, M.L., Mackenzie, A.M., Clements, M.A. & Selkirk, P.M. 2005. DNA sequencing and genetic diversity of the 18S–26S nuclear ribosomal internal transcribed spacers (ITS) in nine Antarctic moss species. *Antarctic Science*, **17**, 377–384.

Tian, L., Wan, S.B., Pan, Q.H., Zheng, Y.J. & Huang, W.D. 2008. A novel plastid localization of chalcone synthase in developing grape berry. *Plant Science*, **175**, 431–436.

Wang, S.S., Xie, X.D., Zhang, L., Lin, F.C., Luo, Z.P., Li, F., *et al*. 2017. Comparative analysis of chalcone synthase gene family among *Nicotiana tabacum* L. and its diploid progenitors. *Tobacco Science & Technology*, **13**, 1–14.

Xu, W.J., Dubos, C. & Lepiniec, L. 2015. Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. *Trends in Plant Science*, **20**, 176–185.

Zhang, J.Z. 2003. Evolution by gene duplication: an update. *Trends in Ecology & Evolution*, **18**, 292–298.

Zhang, X.B., Abrahan, C., Colquhoun, T.A., Liu, C.J. 2017. A proteolytic regulator controlling chalcone synthase stability and flavonoid biosynthesis in Arabidopsis. *Plant Cell*, **29**, 1157–1174.

Zhang, Y.M., Muyrers, J.P., Testa, G. & Stewart, A.F. 2000. DNA cloning by homologous recombination in *Escherichia coli*. *Nature Biotechnology*, **18**, 1314–1317.

Zhu, J.K. 2002. Salt and drought stress signal transduction in plants. *Annual Review of Plant Biology*, **53**, 247–273.

https://doi.org/10.1017/S0954102018000470 Published online by Cambridge University Press