

Next-generation sequencing-based transcriptome profiling analysis of *Pohlia nutans* reveals insight into the stress-relevant genes in Antarctic moss

Shenghao Liu · Nengfei Wang · Pengying Zhang ·
Bailin Cong · Xuezheng Lin · Shouqiang Wang ·
Guangmin Xia · Xiaohang Huang

Received: 1 August 2012 / Accepted: 31 January 2013 / Published online: 27 March 2013
© Springer Japan 2013

Abstract Genome-wide characterization of the *Pohlia nutans* transcriptome is essential for clarifying the role of stress-relevant genes in Antarctic moss adapting to the extreme polar environment. High-throughput Illumina sequencing was used to analyze the gene expression profile of *P. nutans* after cold treatment. A total of 93,488 unigenes, with an average length of 405 bp, were obtained. Gene annotation showed that 16,781 unigenes had significant similarity to known functional protein-coding genes, most of which were annotated using the GO, KOG and KEGG pathway databases. Global profiling of the differentially expressed genes revealed that 3,796 unigenes were significantly upregulated after cold treatment, while 1,405 unigenes were significantly downregulated. In addition, 816 receptor-like kinases and 1,309 transcription factors were identified from *P. nutans*. This overall survey of transcripts and stress-relevant genes can contribute to

understanding the stress-resistance mechanism of Antarctic moss and will accelerate the practical exploitation of the genetic resources for this organism.

Keywords Mosses · Transcriptome · Gene expression profiling · High-throughput RNA sequencing · Cold treatment

Introduction

Antarctic terrestrial ecosystems are cold, dry and low-nutrient environments, with drastic temperature fluctuations and paradoxically low levels of water availability (Pearce 2008). Due to this harsh climate, plant growth is largely limited to the ice-free areas around the coastal fringe of continental Antarctica, which are generally small and isolated. In these regions, cryptogams (mosses and lichens) and two vascular plant species constitute the predominant (though sparse) terrestrial vegetation (Skotnicki et al. 2005; Bokhorst et al. 2007). Recent studies have shown that these terrestrial plants are of ancient origins and have persisted in evolutionary isolation for tens of millions of years (Convey and Stevens 2007). To successfully colonize these low-temperature environments, they have evolved a number of strategies that range from molecular to whole cell, as well as to ecosystem levels. The process of genetic change that accumulates over a time scale of many generations in response to an organism's specific environmental niche is termed adaptation (Morgan-Kiss et al. 2006). In contrast to acclimation, intrinsic changes may occur among the functional genes of Antarctic organisms during adaptation. Therefore, these Antarctic plants offer exceptional opportunities for discovering new functional genes and gaining novel insights into the mechanisms of

Communicated by H. Atomi.

Electronic supplementary material The online version of this article (doi:10.1007/s00792-013-0528-6) contains supplementary material, which is available to authorized users.

S. Liu · N. Wang · B. Cong · X. Lin · S. Wang · X. Huang (✉)
Key Laboratory of Marine Bioactive Substance, The First
Institute of Oceanography, State Oceanic Administration,
6 Xian-Xia Ling Road, High-tech Park, Qingdao 266061,
Shandong, People's Republic of China
e-mail: xiaohanghuang@yahoo.ca; xiaohanghuang@fio.org.cn

S. Liu · G. Xia
School of Life Science, Shandong University,
Jinan 250100, People's Republic of China

P. Zhang
National Glycoengineering Research Center,
Shandong University, Jinan 250100, People's Republic of China

stress resistance, as well as genetic evolution, under extreme conditions.

Mosses have a dominant haploid vegetative stage with stems and leaves, can reproduce both sexually and asexually, and almost any part is capable of regeneration. The moss flora comprises relatively few species of wide ecological amplitude and is widespread across the Antarctic continent (Skotnicki et al. 2000). In the Antarctic environment, moss can be exposed to freezing temperatures (below $-7\text{ }^{\circ}\text{C}$) while in full sunlight (up to $2000\text{ }\mu\text{mol photons m}^{-2}\text{ s}^{-2}$), particularly in the late summer months when the snow cover has melted (Lovelock et al. 1995). The Chinese Great Wall Station is located in the Fildes Peninsula of King George Island, where the moss is overgrown and luxuriating. During China's 24th Antarctic expedition, we collected 12 moss samples in the vicinity of the Great Wall Station. These mosses were regularly cultured in an illumination incubator at temperatures below $16\text{ }^{\circ}\text{C}$ and could still flourish at temperatures down to $4\text{--}6\text{ }^{\circ}\text{C}$, suggesting that they have a strong ability to acclimate to low temperatures.

Previously, we constructed and analyzed the library of expressed sequence tags (ESTs) from Arctic moss to gain insight into the transcriptome involved in various cellular processes (Liu et al. 2010). The emergence of next-generation sequencing technology as a cutting edge approach for high-throughput sequence determination has dramatically improved the efficiency and speed of gene discovery (Schuster 2008). An example can be observed with the Illumina sequencing technology, which is able to generate over one billion bases of high-quality DNA sequence per run at less than 1 % of the cost of capillary-based methods (Huang and Marth 2008). Furthermore, this next-generation sequencing technique has also significantly accelerated and improved the sensitivity of gene expression profiling and is expected to boost collaborative and comparative genomics studies (Blow 2009). It has been demonstrated that by performing Illumina sequencing of the transcriptomes of organisms with completed genomes, the relatively short reads produced can be effectively assembled and used for gene discovery and the comparison of gene expression profiles (Rosenkranz et al. 2008). Despite its obvious potential, these next-generation sequencing methods have yet to be applied to Antarctic organisms.

Mosses are the dominant organisms in Antarctic terrestrial ecosystems. The ability of Antarctic mosses to adapt to the extreme polar environment is due to distinct functional genes and different gene expression profiles corresponding to stressed conditions. However, the genomic sequence resources available for Antarctic mosses are scarce (Peck et al. 2005). Therefore, in this study, the transcriptome analysis of the Antarctic moss, *Pohlia nutans*, in response to cold stress using the Illumina ultra high-throughput

sequencing technology for functional gene or gene family discoveries will offer us a better understanding of the mechanism of stress acclimation and provide us with a basis for effective engineering strategies to improve stress tolerance in agricultural crops.

Materials and methods

Plant materials and cold stress treatments

The Antarctic moss *Pohlia nutans* was collected from the terrene near the Great Wall Station of China in Antarctica (69.8S , 77.8E) during the 24th Antarctic expedition of China in 2008. The samples were placed in plastic containers and transported to China, where they were then cultivated on soil media in an illumination incubator at $16.5\text{ }^{\circ}\text{C}$ and 90 % relative humidity, at a light density of $80\text{ }\mu\text{mol photons m}^{-2}\text{ s}^{-1}$ with continuous light. The mosses cultivated under this condition were used as the experimental control. For the cold treatments, the mosses were transferred to $4\text{ }^{\circ}\text{C}$ and incubated for 6 h. The gametophyte tissues (moss leaves and stems) were frozen immediately with liquid nitrogen. The moss samples were ground in liquid nitrogen with a mortar and pestle, and the ground powder was transferred into 1.5-ml Eppendorf tubes and stored at $-80\text{ }^{\circ}\text{C}$.

RNA extraction and library preparation for transcriptome analysis

The total RNA was extracted from the liquid nitrogen ground moss powder using TRIzol reagent (Invitrogen, Carlsbad, CA, USA), following the manufacturer's standard protocol. The final total RNA was dissolved in $200\text{ }\mu\text{L}$ of RNase-free water. The concentration of the total RNA was determined by a NanoDrop spectrophotometer (Thermo Scientific, USA), and the RNA integrity value (RIN) was checked using the RNA 6000 Pico LabChip kit on an Agilent 2100 Bioanalyzer (Agilent, USA).

The total RNA was incubated with 10 U of DNase I (Ambion, Austin, TX, USA) at $37\text{ }^{\circ}\text{C}$ for 1 h, and then nuclease-free water was added to bring the sample volume to $250\text{ }\mu\text{L}$. The messenger RNA was further purified with the MicroPoly(A) PuristTM Kit (Ambion), following the manufacturer's manual. The mRNA was dissolved in $100\text{ }\mu\text{L}$ of RNA Storage Solution and the final concentration was determined using a NanoDrop spectrophotometer.

The RNA was fragmented at $95\text{ }^{\circ}\text{C}$ for 5 min and annealed with Biotinylated Random Primers which contained the Illumina adapter sequencing primer. The RNA fragments were then captured on Streptavidin Magnetic Beads through interactions with the Biotinylated Random

Primers. An additional Illumina adapter sequencing primer was added to the 5' end of the cleaved mRNA fragments with T4 RNA ligase at 37 °C for 3 h. The cleaved RNA fragments were used for the first-strand cDNA synthesis using reverse transcriptase. The second-strand cDNA was synthesized using DNA polymerase I and RNase H. The products were enriched and purified via PCR to create the final Illumina sequencing cDNA library.

Transcriptome analysis

The Illumina sequencing was performed with the Illumina HiSeq 2000 platform at the Chinese National Human Genome Center (Shanghai), according to the manufacturer's instructions (Illumina, San Diego, CA, USA). Four fluorescently labeled nucleotides and a specialized polymerase were used to determine the clusters base by base in parallel. The size of the cDNA fragment was approximately 300 bp, and both ends of the cDNA fragment were sequenced. The image deconvolution and quality value calculations were performed using Genome Analyzer SCS V2.7 software. The clean data were obtained from the raw data using the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). Any short sequences (<100 bp) were removed. The resulting high-quality sequences were assembled into non-redundant consensus sequences using Velvet and Oases software (Zerbino and Birney 2008; Schulz et al. 2012). The original data sets are available at the NCBI Short Read Archive (SRA) with the accession number SRA051595. This Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the accession GACA00000000. The version described in this paper is the first version, GACA01000000.

The open reading frame was identified using an in-house developed program based on 'GetORF' from EMBOSS (Rice et al. 2000). Gene annotation was performed through a BLASTP search against the Swiss-Prot and GenBank databases (E value $<10^{-3}$), and the best sequence was chosen based on the gene annotation. Gene ontology analysis was performed using GoPipe (Chen et al. 2005), using BLASTP against the Swiss-Prot and TrEMBL databases (E value $<10^{-3}$). The metabolic pathway was constructed based on the KEGG database using the BBH (bi-directional best hit) method (Kanehisa et al. 2010). First, the analysis assigned a KO number to each protein, and then, the metabolic pathways were constructed based on the KO number.

Identification of the differentially expressed genes

Following the cDNA sequence assembly, the reads for the number of distinct sequences in the cold-treatment group

and the control group were derived and used as an approximate estimation of the gene expression levels in the gametophyte tissues of the Antarctic moss, following normalization to the RPKM (Reads Per Kilobase per Million reads) (Mortazavi et al. 2008). The differentially expressed genes between the two samples were identified by the DEGseq package, using the MARS (MA plot-based method with Random Sampling model) method (Wang et al. 2010a, b). A false-discovery rate (FDR) ≤ 0.001 and the absolute value of $\log_2(\text{Treat}/\text{Control}) \geq 1$ were used as the threshold to assess the significance of the differential gene expression. The differentially expressed genes sorted out according to KEGG category term and BLAST annotation.

Identification of receptor-like kinases and transcription factors

The protein sequences of the predicted receptor-like kinases for *Arabidopsis thaliana* were downloaded from the NCBI. The resulting sequences were saved in FASTA format, and the duplicate sequences were removed. A total of 994 receptor-like kinase sequences for *A. thaliana* were obtained and then used to BLAST against the normalized transcriptome database of *P. nutans* (E value $<10^{-3}$).

The protein sequences of the predicted transcription factors for *A. thaliana* were downloaded from the Plant Transcription Factor Database (<http://plantfdb.cbi.pku.edu.cn/>) (Zhang et al. 2011). For identification of the transcription factor-related unigenes from Antarctic moss, the protein sequence set of each predicted *A. thaliana* transcription factor family was BLASTed against the normalized transcriptome database of *P. nutans* (E value $<10^{-3}$).

RT-PCR analyze the expression of several cold stress-related genes

To validate the accumulation of mRNAs under cold stress conditions, total RNA extracts from control plants and plants cold-treated for 1, 3, 6, 12 and 24 h were used as templates and the concentration was measured by microvolume UV-Vis spectrophotometer. RT-PCRs were performed using the PrimeScript™ RT-PCR Kit (TaKaRa, Dalian, China). The ortholog(s) of CRLK1, MEKK1, MKK2, MPK4, CBF and ICE1 of *A. thaliana* in *P. nutans* were selected for RT-PCR analysis (Table S1). The moss β -tubulin gene was used to normalize the template. Gene-specific primers were used to amplify cDNA fragments of these selected genes with the program for 28–29 cycles (95 °C for 30 s; 56 °C for 30 s; 72 °C for 45 s) (Table S2). Amplification products were visualized on 1.2 % agarose (w/v) gels containing ethidium bromide, using a UV light

transilluminator (BioSpectrum® Imaging System, Ultra-Violet Products Ltd., California, USA).

Results

Illumina sequencing and assembly of the reads

To better understand the molecular mechanisms of the Antarctic moss *Pohlia nutans* in adapting to an extreme polar environment, we constructed two Illumina cDNA libraries from the cold-treatment samples and control samples. Approximately, 19,574,660 and 17,855,650 94-bp pair end (PE) raw reads from the cold-treatment tissue and control tissue, respectively, were generated using Illumina RNA sequencing analysis. Using the software Velvet and Oases for de novo assembly, all of the reliable reads from the cold-treatment samples and control samples were first assembled into contigs. The resulting contigs were pooled together and then further assembled, resulting in 145,538 scaffolds, including 49,054 scaffolds larger than 500 bp. The scaffold sequences were assembled into clusters that were then analyzed for consensus. Finally, a total of 93,487 non-redundant sequences (unigenes), ranging from 100 to 13,000 bp, were generated (Fig. S1). Following assembly, we analyzed the length distribution of all of the non-redundant sequences. As shown in Fig. S1, although 80 % of the transcripts were between 100 and 500 bp, we still identified 18,688 transcripts longer than 500 bp. These transcripts provide abundant data that will aid in producing a better reference of the stress-relevant genes in plants.

Annotation of the non-redundant sequences

“GetORF” from the EMBOSS software was used to analyze the reliable coding sequences (CDs) of the 93,488 non-redundant sequences and further translate them into functional proteins. Comparison with the GenBank nr database and Swiss-Prot database revealed that 16,781 non-redundant sequences had good comparability with known gene sequences in existing species (E value $<10^{-3}$). Within these known genes, the average encoded protein size was 275 amino acids. In addition, there were 7,783 non-redundant sequences (length >500 bp) that had no significant similarities to any other protein sequences or were annotated as an unknown protein or putative protein in the public databases.

Annotation of the 93,488 non-redundant sequences using the Gene Ontology (GO) database yielded better results, identifying approximately 19,439 sequences as cellular components, 20,294 sequences involved in molecular functions and 23,659 sequences involved in

biological process, and showed the genes to be distributed among more than 50 categories, including metabolism, growth, development, apoptosis, and immune defense (Fig. 1). Similarly, the clusters of eukaryotic orthologous groups of proteins (KOG)-annotated putative proteins were classified functionally into at least 25 molecular families, including cellular structure, biochemistry, metabolism, molecular processing, signal transduction, gene expression, and immune defense, which correspond to the categories observed in the GO analysis (Fig. S2). The results showed that out of the 93,488 non-redundant sequences, 29,828 transcripts had a KOG classification within the 25 categories. Within these categories, the transcripts with unknown function or putative function together (L + M) comprised the largest group of non-redundant sequences, while the transcripts with signal transduction mechanisms were the second largest group.

The metabolic pathway of the Antarctic moss non-redundant sequences was constructed based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database using the BBH (bi-directional best hit) method. Among the 93,488 non-redundant sequences, approximately 5,037 transcripts could be grouped into six categories comprising 219-known metabolic or signaling pathways, including cell growth and death, cell communication, development, environmental adaptation, signal transduction, folding, replication and repair (Fig. 2).

Differential gene expression analysis of the Antarctic moss after cold treatment

The MARS (MA plot-based method with Random Sampling model) method was used to identify the differentially expressed genes, and “an false-discovery rate (FDR) ≤ 0.001 and an absolute value of \log_2 Ratio ≥ 1 ” were used as the threshold to judge the significance of the differential gene expression. Among these genes, 3,796 non-redundant sequences were significantly upregulated, while 1,405 non-redundant sequences were significantly downregulated (Fig. 3). The distribution of the significant changes is illustrated in a volcano plot, where the statistical significance of each transcript is plotted against the fold change (Fig. S3). For each transcript, the ratio of the expression levels in the cold-treatment samples versus that of the control samples was plotted against the $-\log$ error rate. The horizontal green line indicates the significance threshold (FDR ≤ 0.001), and two vertical green lines indicate the twofold change threshold ($\log_2(\text{Treat}/\text{Control}) \geq 1$ or $\log_2(\text{Treat}/\text{Control}) \leq -1$). The results showed that several of the transcripts with the higher fold changes between the cold-treatment samples and control samples (far right and left of the plot, respectively) also had the smallest false discovery rate values ($-\log$ error rate).

Fig. 1 Histogram presentation of the Gene Ontology classification. The results were summarized into three main categories: biological processes, cellular components and molecular functions. The left y axis indicates the percentage of genes in each specific category

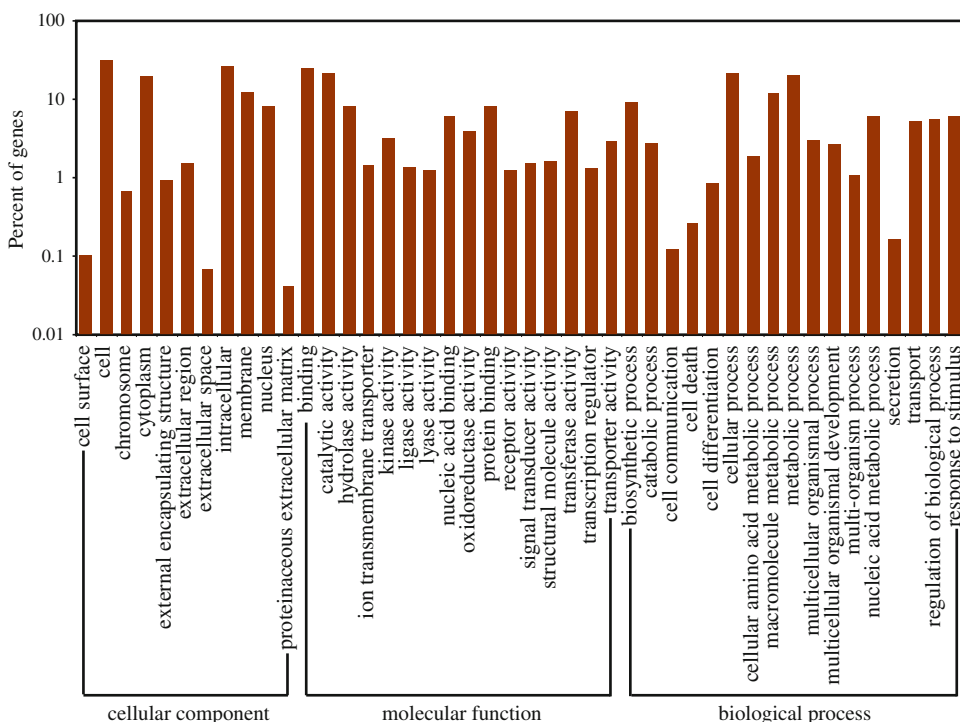
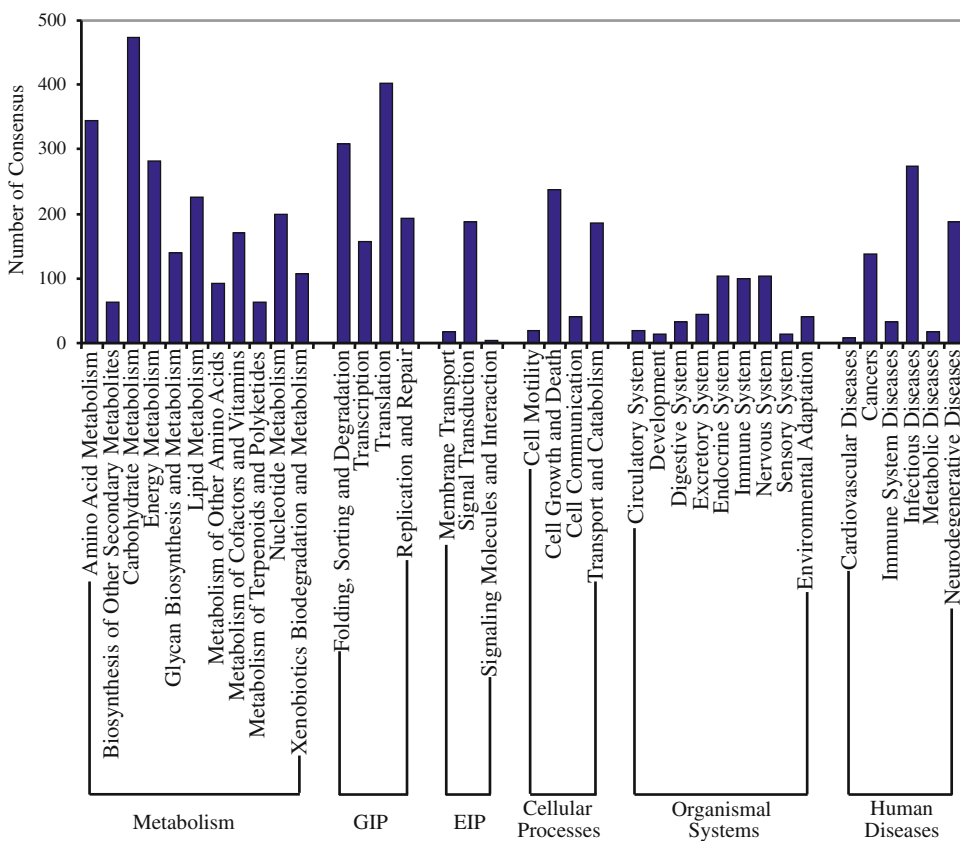


Fig. 2 The KEGG categories of the non-redundant consensus sequences. All of the non-redundant consensus sequences were annotated using the KEGG Automatic Annotation Server for pathway information, and approximately 5,037 consensus sequences were annotated. The categories GIP and EIP stand for genetic information processing and environmental information processing, respectively



Functional annotation showed that the differentially regulated genes were involved in many biochemical and physiological aspects such as signal transduction (receptor-like

kinase, mitogen-activated protein kinase family and phosphatase), transcription (AP2, NAC, bHLH, MYB and zinc finger transcription factor), transport (ABC transporter,

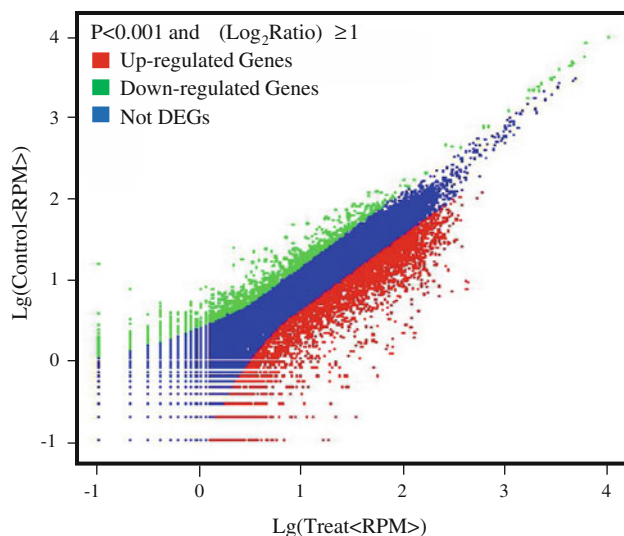


Fig. 3 Differential expression analysis of the identified genes. The expression level for each gene was included in the volcano plot. “Not DEGs” indicates “not-detected expression genes.” The x axis represents the Log_{10} of the transcripts per million of the treated sample, and the y axis indicates Log_{10} of the transcripts per million of the control sample. The limitations are based on $P < 0.001$ and the absolute value of $\log_2(\text{Treat}/\text{Control}) \geq 1$

vacuolar malate transmembrane transporter and tonoplast dicarboxylate transporter, sugar transporter), protein degradation (ubiquitin-protein ligase and F-box family protein), chemical modification (cytochrome P450 monooxygenase, cytochrome P450 reductase and glycosyltransferase). Furthermore, other function genes were also up- or down-regulated after cold stress including the genes of late embryogenesis abundant protein (LEA), S-adenosylmethionine decarboxylase and trehalose 6-phosphate synthase. Cold treatment usually makes the plants suffer oxidative stress and accumulate reactive oxygen species (ROS). Correspondingly, some antioxidant enzymes were up-regulated after cold stress such as thylakoid-bound ascorbate peroxidase, peroxidase and oxidoreductase. Chloroplast omega-6 fatty acid desaturase, microsomal delta-5 desaturase and delta9-fatty acid desaturase were up-regulated, which might be responsible for increasing the membrane unsaturated fatty acid content under cold stress. These representative cold-responsive genes of Antarctic moss *P. nutans* were sorted out and classified according to BLAST annotation as shown in Table 1.

Identification of receptor-like kinases and transcription factors

BLAST software was used to search the normalized the Antarctic moss transcriptome database and a total of 816 transcripts encoding putative receptor-like kinases from the transcriptome of *P. nutans* were characterized, using the *A. thaliana* receptor-like kinase sequences as a reference

(E value $< 10^{-3}$). Further annotation showed that these proteins contain subfamilies that include leucine-rich repeat receptor-like kinase, proline-rich receptor kinase, lectin-rich receptor kinase, mitogen-activated protein kinase, and Ca^{2+} and calmodulin-dependent protein kinase. Differential gene expression analysis showed that there were 42-upregulated and 33-downregulated putative receptor-like kinases in *P. nutans* after cold treatment (Table 1). These differentially expressed genes accounted for 9.2 % of the total isolated receptor-like kinases.

In the *P. nutans* transcriptome, 1,309 transcripts encoding putative transcription factors were identified using *A. thaliana* transcription factor as queries, comprising 1.4 % (1,309/93,487) of the total non-redundant sequences (Fig. S3). The target transcripts were further classified into 52 corresponding transcription factors families. Differential gene expression analysis showed that 106 transcription factors were significantly upregulated, including those in the AP2/ERF (20 transcripts), WRKY (21 transcripts), C2H2 (7 transcripts), NAC (6 transcripts) and bHLH (6 transcripts) families, while 63 transcription factors were significantly downregulated, including the WRKY (13 transcripts), NAC (6 transcripts), C2H2 (6 transcripts) and bHLH (6 transcripts) families (Table 1). These differentially expressed genes accounted for 12.9 % of the total isolated transcription factors.

RT-PCR analyze the expression of several cold stress-related genes in *P. nutans*

According to the previously published reports, we selected orthologs of CRLK1, MEKK1, MKK2, MPK4 and CBF transcription factors in *P. nutans* for analyzing their expression levels under cold stress and validating the reliability of Solexa/Illumina sequencing data. The expression levels of these genes in *P. nutans* transcriptome sequencing and the percent identity with orthologs in *A. thaliana* were listed in Table S1. RT-PCR results showed that the expression levels of one MEKK1 ortholog (MFC03145), one MKK2 ortholog (MFC88317), three CBF orthologs (MFC40570, MFC14588, MFC00537) and one ICE1 ortholog (MFC41806) were significantly increased after cold stress (Fig. 4). However, the selected CRLK1 and MPK4 orthologs in *P. nutans* were not increased after cold stress (Fig. S5). These results showed that expressions of these selected fifteen genes were basically consistent between the RT-PCR and the transcriptome differential gene expression analysis.

Discussion

Antarctic mosses are capable of growing under extreme climate conditions, which include extreme cold, aridity and

Table 1 Representative cold-responsive genes of Antarctic moss *P. nutans*

Gene ID	Cold	Control	Fold	Functional annotation	Gene family
MFC74230	1607	109	4.00	Leucine-rich receptor-like protein kinase [<i>Arabidopsis thaliana</i> , NP_173166.2]	Receptor-like protein kinase
MFC17337	1239	508	1.40	Receptor-like kinase [<i>Marchantia polymorpha</i> , BAF79934.1]	
MFC04784	759	315	1.38	Receptor-like kinase [<i>Marchantia polymorpha</i> , BAF79947.1]	
MFC35418	1059	438	1.39	Receptor protein kinase, putative [<i>Ricinus communis</i> , XP_002527538.1]	
MFC86560	548	270	1.13	Receptor protein kinase, putative [<i>Ricinus communis</i> , XP_002520456.1]	
MFC04716	1576	122	3.81	Protein kinase-like protein [<i>Arabidopsis thaliana</i> , NP_177036.1]	Protein kinase
MFC71588	791	175	2.29	AKIN10 (Arabidopsis SNF1 kinase homolog 10) [<i>Arabidopsis thaliana</i> , NP_566130.1]	
MFC57702	1084	283	2.05	Serine-threonine protein kinase, plant-type, putative [<i>Ricinus communis</i> , XP_002525575.1]	
MFC82023	907	240	2.03	NF-1-like serine/threonine protein kinase [<i>Glycine max</i> , AAD23582.1]	
MFC12774	1647	480	1.89	Lectin protein kinase family protein [<i>Arabidopsis lyrata</i> subsp. <i>Lyrata</i> , XP_002869284.1]	
MFC02941	1036	344	1.70	Serine/threonine protein kinase AtPK19, putative [<i>Solanum lycopersicum</i> , ABI34273.1]	
MFC04999	452	170	1.52	Mitogen-activated protein kinase, putative [<i>Ricinus communis</i> , XP_002524836.1]	
MFC03145	362	144	1.44	Protein kinase-like protein [<i>Arabidopsis thaliana</i> , NP_177036.1]	
MFC03367	613	104	2.67	Putative protein phosphatase 2C [<i>Arabidopsis thaliana</i> , NP_174731.1]	Phosphatase
MFC00537	331	18	4.31	AP2 domain transcription factor [<i>Arabidopsis thaliana</i> , AAM64613.1]	AP2 transcription factor
MFC14588	2608	181	3.96	Transcriptional factor DBF1 [<i>Physcomitrella patens</i> , ABA43687.2]	
MFC23320	1385	167	3.17	CBF2-like protein [<i>Cissus Antarctica</i> , ABU55671.1]	
MFC77984	201	23	3.24	Dehydration-responsive element-binding protein 2A [<i>Oryza sativa Indica Group</i> , A2WL19.2]	
MFC19911	1810	308	2.67	Dehydration-responsive element-binding protein 2C [<i>Ricinus communis</i> , XP_002520794.1]	
MFC50482	1169	249	2.34	DNA binding protein, putative [<i>Ricinus communis</i> , XP_002511590.1]	
MFC34744	621	1375	−1.03	Putative AP2/EREBP transcription factor [<i>Arabidopsis thaliana</i> , AAT44951.1]	
MFC07853	186	93	1.11	NAC domain-containing protein, putative [<i>Ricinus communis</i> , XP_002532571.1]	NAC transcription factor
MFC05080	1698	882	1.06	NAC domain protein NAC3 [<i>Glycine max</i> , NP_001238234.1]	
MFC20077	330	196	0.87	Transcription factor, putative [<i>Ricinus communis</i> , XP_002527580.1]	
MFC79899	924	191	2.39	bHLH transcription factor [<i>Musa acuminata AAA Group</i> , ADQ38902.1]	bHLH transcription factor
MFC25128	742	161	2.32	Transcription factor BIM1, putative [<i>Ricinus communis</i> , XP_002514597.1]	
MFC65542	756	357	1.20	Basic helix-loop-helix family protein [<i>Arabidopsis lyrata</i> subsp. <i>Lyrata</i> , XP_002891184.1]	
MFC00217	189	80	1.35	MYB transcription factor MYB173 [<i>Glycine max</i> , NP_001236032.1]	MYB transcription factor
MFC06881	1585	833	1.04	MYBR domain class transcription factor [<i>Malus x domestica</i> , ADL36775.1]	
MFC26808	464	1286	−1.36	R2R3-MYB transcription factor MYB3 [<i>Pinus taeda</i> , ABD60282.1]	
MFC93478	761	202	2.03	Transcription factor viviparous 1 [<i>Picea abies</i> , AAG22585.1]	B3 transcription factor
MFC09871	602	167	1.96	Transcription factor ABI3-like [<i>Physcomitrella patens</i> , BAE80314.1]	
MFC34485	1256	224	2.60	Dof transcription factor [<i>Physcomitrella patens</i> , BAL46027.1]	Other transcription factor
MFC30140	2678	947	1.61	GATA transcription factor 19 [<i>Zea mays</i> , NP_001150675.1]	
MFC66468	2410	344	2.92	PLATZ transcription factor family protein [<i>Arabidopsis thaliana</i> , NP_001078404.1]	

Table 1 continued

Gene ID	Cold	Control	Fold	Functional annotation	Gene family
MFC54755	434	80	2.55	Zinc-finger transcription factor [<i>Ananas comosus</i> , ACN43326.1]	Zinc finger protein or transcription factor
MFC35044	1512	183	3.16	FYVE zinc finger protein [<i>Ectocarpus siliculosus</i> , CBN74398.1]	
MFC00019	873	171	2.47	Zinc finger family-like protein [<i>Picea sitchensis</i> , ADM77691.1]	
MFC67365	465	101	2.32	Zinc finger transcription factor [<i>Ananas comosus</i> , ACN43326.1]	
MFC45050	1107	269	2.15	Protein binding/zinc ion binding [<i>Arabidopsis thaliana</i> , NP_565180.1]	
MFC25637	1071	276	2.07	Zinc finger family protein [<i>Arabidopsis lyrata</i> subsp. <i>Lyrata</i> , XP_002888862.1]	
MFC44263	1191	355	1.86	EPF-type Cis2-His2 zinc finger transcription factor [<i>Selaginella moellendorffii</i> , XP_002984343.1]	
MFC66915	921	416	1.26	Zinc finger protein, putative [<i>Ricinus communis</i> , XP_002527000.1]	
MFC20764	79	734	−3.10	Zinc-finger transcription factor [<i>Ananas comosus</i> , ACN43326.1]	
MFC44921	1399	217	2.80	Putative plastidic glucose transporter 2 [<i>Arabidopsis thaliana</i> , NP_849855.2]	Sugar transporter
MFC31537	1358	215	2.77	Putative sugar transporter [<i>Oryza sativa Japonica Group</i> , BAD21807.1]	
MFC79082	1193	290	2.15	Sugar transporter, putative [<i>Ricinus communis</i> , XP_002520608.1]	
MFC79788	917	164	2.60	Putative sugar transporter [<i>Prunus armeniaca</i> , AAB88879.1]	
MFC05968	1000	242	2.16	Putative tonoplastic monosaccharide transporter [<i>Vitis vinifera</i> , ADP37149.1]	
MFC79082	1193	290	2.15	Sugar transporter, putative [<i>Ricinus communis</i> , XP_002520608.1]	
MFC22612	1990	296	2.86	Beta-amylase, putative [<i>Ricinus communis</i> , XP_002511858.1]	Amylase
MFC31313	1199	198	2.71	Beta-amylase/catalytic/cation binding [<i>Arabidopsis thaliana</i> , NP_182112.2]	
MFC05390	381	71	2.54	Glycosyltransferase [<i>Lycium barbarum</i> , BAG80542.1]	Glycosyltransferase
MFC35722	626	314	1.11	Glycosyltransferase-like protein [<i>Selaginella moellendorffii</i> , XP_002975186.1]	
MFC57138	294	15	4.41	delta9-fatty acid desaturase [<i>Phanerochaete chrysosporium</i> , BAJ04706.1]	Fatty acid desaturase
MFC03373	882	379	1.33	Chloroplast omega-6 fatty acid desaturase [<i>Jatropha curcas</i> , ABU96742.1]	
MFC20159	1045	400	1.50	Microsomal delta-5 desaturase [<i>Thalassiosira pseudonana</i> CCMP1335, XP_002296094.1]	
MFC00874	1618	696	1.33	Microsomal delta-5 desaturase [<i>Salpingoeca</i> . ATCC 50818, EGD79088.1]	
MFC46529	1085	467	1.33	Omega-6 fatty acid desaturase [<i>Oryza sativa Japonica Group</i> , BAD09897.1]	
MFC43853	1343	643	1.18	Cinnamoyl CoA reductase [<i>Picea abies</i> , CAK18610.1]	Oxidoreductase
MFC03124	369	154	1.37	Quinone oxidoreductase [<i>Brucella</i> . BO2, ZP_07474467.1]	
MFC47891	1056	412	1.47	Oxidoreductase/transcriptional repressor [<i>Arabidopsis lyrata</i> subsp. <i>Lyrata</i> , XP_002890185.1]	
MFC27619	769	398	1.06	Oxidoreductase [<i>Aeromicrobium marinum</i> , ZP_07717703.1]	
MFC09572	405	216	1.02	FAD oxidoreductase, putative [<i>Ricinus communis</i> , XP_002527265.1]	
MFC10326	393	123	1.79	Thylakoid-bound ascorbate peroxidase [<i>Brassica oleracea</i> , BAD14931.1]	
MFC02813	721	326	1.26	Peroxidase-like protein [<i>Oryza sativa Japonica Group</i> , BAD82471.1]	
MFC78170	1023	219	2.34	Late embryogenesis abundant hydroxyproline-rich glycoprotein [<i>Arabidopsis thaliana</i> , NP_193063.2]	LEA protein
MFC01171	398	92	2.23	Late embryogenesis abundant, putative [<i>Ricinus communis</i> , XP_002528631.1]	
MFC07113	3503	829	2.19	Late embryogenesis abundant protein/LEA protein [<i>Arabidopsis thaliana</i> , NP_188888.1]	
MFC11804	933	234	2.11	Late embryogenesis abundant protein [<i>Brassica napus</i> , BAB88877.1]	
MFC57734	747	208	1.96	LEA protein group 3 [<i>Marchantia polymorpha</i> , BAJ78281.1]	

Table 1 continued

Gene ID	Cold	Control	Fold	Functional annotation	Gene family
MFC00982	1085	286	2.04	S-Adenosylmethionine decarboxylase [<i>Olea europaea</i> , ACZ73828.1]	speD
MFC56053	2105	504	2.18	S-Adenosylmethionine decarboxylase proenzyme [<i>Vitis vinifera</i> , CAD98785.2]	
MFC06651	921	332	1.59	Trehalose-phosphate synthase [<i>Selaginella moellendorffii</i> , XP_002989716.1]	Trehalose-phosphate synthase
MFC24377	80	30	1.53	Trehalose-6-phosphate synthase [<i>Selaginella pulvinata</i> , ADD71944.1]	
MFC47000	171	65	1.51	Trehalose-6-phosphate synthase, putative [<i>Musa balbisiana</i> , ABF70084.1]	
MFC09274	460	180	1.47	Trehalose 6-phosphate synthase [<i>Nicotiana tabacum</i> , BAI99252.1]	
MFC38284	1599	638	1.44	Truncated trehalose-phosphate synthase 1 [<i>Physcomitrella patens</i> subsp. <i>patens</i> , ABO61743.1]	
MFC14059	1072	463	1.32	Trehalose 6-phosphate synthase [<i>Nicotiana tabacum</i> , BAI99252.1]	
MFC07511	1780	412	2.22	Flavonoid 3'-hydroxylase [<i>Epimedium sagittatum</i> , ADE80941.1]	P450 monooxygenase and reductase
MFC24430	1069	275	2.07	Medium chain fatty acid hydroxylase [<i>Selaginella moellendorffii</i> , XP_002979685.1]	
MFC28853	164	433	-1.29	Flavonoid 3'-hydroxylase [<i>Epimedium sagittatum</i> , ADE80941.1]	
MFC43413	764	314	1.40	NADPH-cytochrome P450 reductase [<i>Pseudotsuga menziesii</i> , CAA89837.3]	
MFC35837	2080	899	1.32	Cytochrome P450 reductase [<i>Taxus wallichiana</i> var. <i>chinensis</i> , AAX59902.1]	
MFC35878	1433	133	3.54	Ubiquitin-conjugating enzyme variant, putative [<i>Ricinus communis</i> , XP_002522972.1]	Ubiquitin-protein ligase
MFC10223	442	53	3.17	Ubiquitin-protein ligase, putative [<i>Ricinus communis</i> , XP_002525648.1]	
MFC57746	354	56	2.77	Ubiquitin-protein ligase, putative [<i>Ricinus communis</i> , XP_002525648.1]	
MFC84560	1561	587	1.52	F-box family protein [<i>Populus trichocarpa</i> , XP_002309467.1]	F-box family protein
MFC07176	443	56	3.10	F-box/kelch-repeat protein SKIP25 [<i>Medicago truncatula</i> , XP_003612576]	
MFC51793	669	287	1.33	F-box family member [<i>Zea mays</i> , NP_001132560.1]	
MFC02228	256	66	2.07	F-box family protein [<i>Populus trichocarpa</i> , XP_002304679.1]	
MFC05072	605	147	2.15	ATP-binding cassette transporter [<i>Selaginella moellendorffii</i> , XP_002985478.1]	ABC transporter family
MFC32178	1976	542	1.98	ATP-binding cassette transporter [<i>Selaginella moellendorffii</i> , XP_002985478.1]	
MFC35626	654	180	1.98	ATP-binding cassette transporter F9 [<i>Physcomitrella patens</i> subsp. <i>patens</i> , XP_001757662.1]	
MFC07819	640	200	1.79	ATP-binding cassette transporter F9 [<i>Physcomitrella patens</i> subsp. <i>patens</i> , XP_001757662.1]	
MFC51373	1037	353	1.67	ATP-binding cassette transporter F1 [<i>Physcomitrella patens</i> subsp. <i>patens</i> , XP_001779695.1]	
MFC26997	50	143	-1.40	ATP-binding cassette transporter G5 [<i>Physcomitrella patens</i> subsp. <i>patens</i> , XP_001781566.1]	
MFC03002	966	309	1.76	Unknown [<i>Picea sitchensis</i> , ABK24775.1]	Moss-specific genes
MFC49704	1345	266	2.45	Predicted protein [<i>Physcomitrella patens</i> subsp. <i>patens</i> , XP_001774012.1]	
MFC77682	927	363	1.47	Predicted protein [<i>Physcomitrella patens</i> subsp. <i>patens</i> , XP_001783573.1]	

The Gene ID MFC14179 is the abbreviation of MFC14179; Cold: Gene raw reads from cold-treatment sample; Control: Gene raw reads from control treatment sample; Fold: $\log_2(\text{Treat}/\text{Control})$ normalized; Annotation: blastp hit in nr database. All these genes can be retrieved from Transcriptome Shotgun Assembly Sequence Database of NCBI website

a photoperiod that ranges from continuous sunlight in the summer to continuous darkness in the winter (Skotnicki et al. 2000). These plants offer exceptional opportunities for gaining novel insights into the mechanisms of plant survival under extreme conditions (Peck et al. 2005). The transcriptome is the complete repertoire of expressed RNA transcripts in a cell, and its characterization is essential in deciphering the functional complexity of the genome and in obtaining a better understanding of the cellular activities in organisms, including growth, development, immune defense and stress response (Xiang et al. 2010). The Solexa/Illumina RNA-seq deep sequencing approaches have overcome many of the inherent barriers of the traditional EST approach, making the detection of alternative splicing events and low-abundance transcripts possible (Xiang et al. 2010; Wang et al. 2010a, b). Recently, these platforms have been applied to several species, such as the Monterey pine (*Pinus radiata*), the cucumber (*Cucumis sativus* L.), sesame (*Sesamum indicum* L.) and the moss *P. patens*, for different purposes (Li et al. 2011; Wei et al. 2011; Qi et al. 2012; Xiao et al. 2011).

In this study, transcriptome profile sequencing analysis of cold shock-treated Antarctic mosses was conducted using the Illumina platform. A large data set of 93,488 non-redundant consensus sequences, including 18,688 sequences longer than 500 bp, with complete or partial lengths of encoding regions was generated (Fig. S1). Although a high number of transcripts were short-length reads, which can result in several assembled contigs and singletons for each gene (Hsiao et al. 2011), this isolated consensus sequence dataset will reflect the sophisticated organization of Antarctic moss for favoring growth, reproduction and adaptation to the polar environment. Homology searches revealed that 16,781 consensus sequences had significant similarities with known gene sequences in existing species. In addition, there were 7,783 consensus sequences (length >500 bp) that had no significant similarities to any other protein sequences or were annotated as “unknown protein” or “putative protein” in the public databases. Together, these non-redundant consensus sequences nearly equaled the number of clusters found in the model plant *P. patens* subsp. *patens* (Rensing et al. 2008). Subsequently, when we set a higher threshold (sequence length >1000 bp and raw reads >100) to screen out the genes annotated as unknown protein or putative protein, 1,793 genes matched these conditions in the Antarctic moss transcriptome. These novel moss-specific genes may perform specific roles, such as ecological strategies for environmental adaptation, and may be quite divergent from those of other plant species.

Comparing the sequencing abundance (raw reads) of the functional genes between the treatment and control samples provided support for the expression levels of the different genes responding to stress. The results showed that

the cold treatment resulted in large alterations of the transcriptome profile, including significant upregulation of 3,796 transcripts and downregulation of 1,405 transcripts (Fig. 3). Our data were consistent with the number of differentially expressed genes in moss model organism *P. patens* that, 4,129 transcripts were identified as differentially expressed or as developmental-stage specific genes by the Illumina digital gene expression tag profiling technique (Xiao et al. 2011). Previously, the proteomic analysis (2-DE and mass spectroscopy) were used to study the cold stress response in the moss, *P. patens* and 45 changed proteins were identified involved in energy metabolism, signaling, cytoskeleton, and defense proteins. Our results also showed that the cold-responses in Antarctic moss are also multi-aspects including the processes of cold sensing and signal transduction, alterations of metabolism, accumulation of defense proteins (Table S1) (Wang et al. 2009).

Receptor-like kinases play many critical roles in plant development and physiology, including developmental patterning, hormone signaling, disease resistance and stress responses (Lehti-Shiu et al. 2009; Gish and Clark 2011). In Antarctic moss *P. nutans* transcriptome, differential gene expression analysis showed that nearly 42 receptor-like kinases were upregulated, and 33 receptor-like kinases were downregulated after the cold treatment, including leucine-rich repeat receptor-like kinases, mitogen-activated protein kinases, and Ca^{2+} and calmodulin-dependent protein kinases (Table 1). Furthermore, a recent study showed that a novel calcium/calmodulin-regulated receptor-like kinase (CRLK1) mainly localized in the plasma membrane might confer cold sensing and signal transduction in *A. thaliana* (Yang et al. 2010). Using CRLK1 (NP_568809.2) BLASTP search against *P. nutans* transcriptome protein data, we identified a full-length CRLK1 ortholog in *P. nutans* (MFC60113). The amino acid sequence deduced from MFC60113 was 57.9 % identity with CRLK1 of *A. thaliana* and 81.1 %, 83.1 identity with CRLK1 ortholog of *P. patens* subsp. *patens* (XP_001753295.1 and XP_001755574.), respectively. However, the expression level of CRLK1 ortholog in *P. nutans* was not increased after cold stress treatment (Fig. S5, Table S1). In plants, mitogen-activated protein kinase signaling under abiotic stress had been comprehensively studied. Overexpressing MEKK1, MKK2, MPK4 exhibited enhanced tolerance to cold and salt stress (Rodriguez et al. 2010). In our RT-PCR analysis, one MEKK1 (MFC03145) and one MKK2 (MFC88317) genes were significantly up-regulated after cold stress in *P. nutans*. However, the other selected MEKK1, MKK2 and MPK4 were not changed (Fig. 4; Fig. S5, Table S1).

Transcription factors exhibit sequence-specific DNA binding and are capable of activating or repressing the transcription of downstream target genes. In this study, we

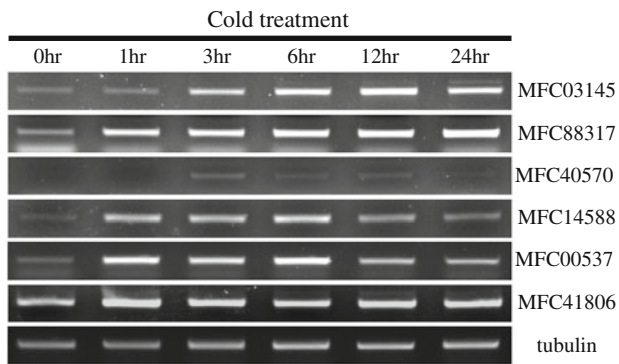


Fig. 4 RT-PCR analysis of MEKK1, MKK2, CBF and ICE1 orthologs in Antarctic moss *P. nutans* after cold stress. MFC03145 is MEKK1 ortholog; MFC88317 is MKK2 ortholog; MFC40570, MFC14588 and MFC00537 are CBF orthologs; MFC41806 is ICE1 ortholog. β -Tubulin was used to normalize the template

identified 1,309 unigenes encoding putative transcription factors, corresponding to 52 subfamilies and comprising 1.4 % (1,309/93,487) of the unigenes of the *P. nutans* transcriptome (Fig. S3). Compared to the 5.7 % of plant genes that have been shown to be transcription factor genes (Libault et al. 2009), the 1.4 % of genes associated with transcription factors in *P. nutans* is low. However, in the genome of the model plant *P. patens* subsp. *patens*, 1,188 genes were found to encode transcription factors and were classified into 53 subfamilies (Zhang et al. 2011). This finding suggests that the underestimation of transcription factor genes in *P. nutans* could have been attributed to the relatively short reads per sequencing reaction, therefore, causing the assembled unigenes to be unable to cover the full length of the transcripts or to represent untranslated regions. Nevertheless, the number of transcription factors in *P. nutans* (1,309) seems to be similar level with those in *P. patens* (1,188) (Fig. S3). The transcription factors in the CBF family, existing in multiple plant species, are the key regulators of the cold-responsive genes. Constitutive overexpression of either CBF1, CBF2, or CBF3 in *A. thaliana* results in expression of the CBF regulon and brings about an increase in freezing tolerance without a cold stimulus, indicating that the CBF regulon has a fundamental role in cold acclimation (Zhou et al. 2011). The upstream regulators of the CBF named inducer of CBF expression (ICE) acts as a positive regulator of CBFs. In Antarctic moss *P. nutans*, three CBF orthologs (MFC40570, MFC14588, MFC00537) and one ICE1 ortholog (MFC41806) were significantly increased after cold stress in RT-PCR analysis (Fig. 4). Therefore, the cold acclimation in Antarctic moss *P. nutans* may also be mainly mediated by CBF transcription factor-dependent manner.

It was demonstrated that factors such as the accumulation of soluble sugars, unsaturated fatty acid and

antioxidants also play important roles in alleviating freezing-induced cellular damage (Sun et al. 2007; Bhyan et al. 2012). Recent study showed that soluble sugars are involved in responses to stress, and act as signaling molecules that activate specific or hormone cross-talk transduction pathways (Ramel et al. 2009). They are distributed through the plant via sugar transporters, which are involved not only in sugar long-distance transport via the loading and the unloading of the conducting complex, but also in sugar allocation into source and sink cells (Afoufa-Bastien et al. 2010). In *P. nutans* transcriptome, we identified a total of 63 sugar transporter candidates including monosaccharide transporter, glucose transporter and hexose transporter. Differential gene expression analysis showed that nearly 16 sugar transporters were upregulated after cold stress (Table 1). Meanwhile, it has been demonstrated that there is a general tendency in the cells to increase the amount of unsaturated fatty acid lipids to enhance membrane fluidity. Acclimation to low-temperature stress via an increase in expression of desaturases has been documented in bacteria, algae, plants, and animals (Morgan-Kiss et al. 2006). In the present study, a total of 16 unigenes in the Antarctic moss transcriptome were identified as possible members of the fatty acid desaturase family, including omega-3, omega-6, delta 9 and delta-12 desaturase. Differential gene expression analysis showed that 6 fatty acid desaturases were significantly upregulated after cold stress (Table 1). Together, these differentially expressed transcripts of sugar transporter, fatty acid desaturase, oxidoreductase may indicate that sugars, unsaturated fatty acid and antioxidant should also have function in *P. nutans* cold acclimation.

LEA proteins are accumulated in vegetative plant tissues after exposure to environmental stresses, and these proteins are thought to play a role in freezing tolerance (Hundertmark and Hincha 2008). Cold acclimation in bryophytes is also accompanied by accumulation of several transcripts for LEA proteins (Minami et al. 2005). We identified a total of 21 LEA proteins in *P. nutans* transcriptome and 9 of them were significantly upregulated after cold stress (Table 1). Polyamines are present in all living organism and implicate in a wide range of cellular physiological processes. Overexpression of carnation S-adenosylmethionine decarboxylase gene generates a broad-spectrum tolerance to abiotic stresses in transgenic tobacco plants (Wi et al. 2006). We also identified two S-adenosylmethionine decarboxylase genes in *P. nutans* transcriptome and both of them were significantly upregulated after cold stress (Table 1). Moreover, the disaccharide trehalose is considered as a universal stress molecule, protecting cells and biomolecules from injuries imposed by many abiotic stress including freezing, salinity, oxidation and desiccation (Tibbett et al. 2002; Reina-Bueno et al. 2012). Similarly, in

P. patens transcriptome, we identified a total of 20 trehalose-phosphate synthase genes and more than 11 of them were upregulated after cold stress, suggesting the important role of trehalose in Antarctic moss acclimatizing the extreme environment (Table 1). Furthermore, cytochrome P450 monooxygenase (hydroxylase) together with cytochrome P450 reductase participate in a variety of biochemical pathways to produce primary and secondary metabolites such as phenylpropanoids, alkaloids, terpenoids, lipids, and glucosinolates, as well as plant hormones (Mizutani and Sato 2011). In model plant *P. patens*, proteomic analysis found that cytochrome P450 protein and NADH-cytochrome P450 reductase were up-regulated after cold stress. In Antarctic moss *P. nutans*, we also found that cytochrome P450 protein and NADH-cytochrome P450 reductase are regulated by cold stress (Table 1).

Taken together, these data obtained from annotated or known genes suggested that Antarctic moss adopted the similar biochemical and genetic processes as higher plant in response to cold stress. However, Antarctic moss transcriptome comprised a relatively large proportion of genes annotated as unknown protein or putative protein, and several of these genes were also significantly regulated by cold stress. Therefore, cold acclimation in Antarctic moss displays distinct differences with higher plants, which suggests a significant alteration during the evolution of land plants. Moreover, five up-regulated moss-specific genes in Antarctic moss transcriptome have been further confirmed by Real-Time PCR approach and are being verified by transgenic *Arabidopsis*. We believe that these results will contribute to clarify the genetic mechanism of Antarctic moss in the processes of acclimatizing and adapting the rigorous climate in Antarctic continent and accelerate practical exploitation of Antarctic moss gene resources in improving the stress tolerance in agricultural crops.

Acknowledgments We thank the Chinese National Human Genome Center (Shanghai) for their collaboration and support in the de novo assembly and annotation of the *P. nutans* transcriptome. We are very grateful for the critical comments and suggestions from the anonymous reviewers. This work was financially supported by the China National Natural Science Foundation (40906103 and 41206176) and Basic Scientific Research Foundation of The First Institute of Oceanography (SOA).

References

- Afoufa-Bastien D, Medici A, Jeauffre J, Coutos-Thévenot P, Lemoine R, Atanassova R, Laloi M (2010) The *Vitis vinifera* sugar transporter gene family: phylogenetic overview and microarray expression profiling. *BMC Plant Biol* 10:245
- Bhyan SB, Minami A, Kaneko Y, Suzuki S, Arakawa K, Sakata Y, Takezawa D (2012) Cold acclimation in the moss *Physcomitrella patens* involves abscisic acid-dependent signaling. *J Plant Physiol* 169(2):137–145
- Blow N (2009) Transcriptomics: the digital generation. *Nature* 458:239–242
- Bokhorst S, Huiskes A, Convey P, Aerts R (2007) The effect of environmental change on vascular plant and cryptogam communities from the Falkland Islands and the Maritime Antarctic. *BMC Ecol* 7:15
- Chen ZZ, Xue CH, Zhu S, Zhou FF, Ling XB, Liu GP, Chen LB (2005) GoPipe: streamlined gene ontology annotation for batch anonymous sequences with statistics. *Prog Biochem Biophys* 32:187–191
- Convey P, Stevens MI (2007) Antarctic Biodiversity. *Science* 317:1877–1878
- Gish LA, Clark SE (2011) The RLK/Pelle family of kinases. *Plant J* 66:117–127
- Hsiao YY, Chen YW, Huang SC, Pan ZJ, Fu CH, Chen WH, Tsai WC, Chen HH (2011) Gene discovery using next-generation pyrosequencing to develop ESTs for *Phalaenopsis orchids*. *BMC Genomics* 12:360
- Huang W, Marth G (2008) EagleView: a genome assembly viewer for next-generation sequencing technologies. *Genome Res* 18:1538–1543
- Hundertmark M, Hinch DK (2008) LEA (late embryogenesis abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC Genomics* 9:118
- Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M (2010) KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res* 38:355–360
- Lehti-Shiu MD, Zou C, Hanada K, Shiu SH (2009) Evolutionary history and stress regulation of plant receptor-like kinase/pelle genes. *Plant Physiol* 150:12–26
- Li XG, Wu HX, Southerton SG (2011) Transcriptome profiling of *Pinus radiata* juvenile wood with contrasting stiffness identifies putative candidate genes involved in microfibril orientation and cell wall mechanics. *BMC Genomics* 12:480
- Libault M, Joshi T, Benedito VA, Xu D, Udvardi MK, Stacey G (2009) Legume transcription factor genes: what makes legumes so special? *Plant Physiol* 151:991–1001
- Liu SH, Lee HS, Kang PS, Huang XH, Yim JH, Lee HK, Kim II-C (2010) Complementary DNA library construction and expressed sequence tag analysis of an Arctic moss, *Aulacomnium turgidum*. *Polar Biol* 33:617–626
- Lovelock CE, Jackson AE, Melick DR, Seppelt RD (1995) Reversible photoinhibition in Antarctic moss during freezing and thawing. *Plant Physiol* 109:955–961
- Minami A, Nagao M, Ikegami K, Koshiba T, Arakawa K, Fujikawa S, Takezawa D (2005) Cold acclimation in bryophytes: low-temperature-induced freezing tolerance in *Physcomitrella patens* is associated with increases in expression levels of stress-related genes but not with increase in level of endogenous abscisic acid. *Planta* 220(3):414–423
- Mizutani M, Sato F (2011) Unusual P450 reactions in plant secondary metabolism. *Arch Biochem Biophys* 507(1):194–203
- Morgan-Kiss RM, Prisco JC, Pocock T, Gudynaite-Savitch L, Huner NP (2006) Adaptation and acclimation of photosynthetic microorganisms to permanently cold environments. *Microbiol Mol Biol Rev* 70:222–252
- Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5:621–628
- Pearce DA (2008) Climate change and the microbiology of the Antarctic Peninsula region. *Sci Prog* 91:203–217
- Peck LS, Clark MS, Clarke A et al (2005) Genomics: applications to Antarctic ecosystems. *Polar Biol* 28:351–365

- Qi XH, Xu XW, Lin XJ, Zhang WJ, Chen XH (2012) Identification of differentially expressed genes in cucumber (*Cucumis sativus* L.) root under waterlogging stress by digital gene expression profile. *Genomics* 99:160–168
- Ramel F, Sulmon C, Gouesbet G, Couée I (2009) Natural variation reveals relationships between pre-stress carbohydrate nutritional status and subsequent responses to xenobiotic and oxidative stress in *Arabidopsis thaliana*. *Ann Bot* 104(7):1323–1337
- Reina-Bueno M, Argandoña M, Salvador M, Rodríguez-Moya J, Iglesias-Guerra F et al (2012) Role of trehalose in salinity and temperature tolerance in the model halophilic bacterium *Chromohalobacter salexigens*. *PLoS ONE* 7(3):e33587
- Rensing SA, Lang D, Zimmer AD et al (2008) The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* 319:64–69
- Rice P, Longden I, Bleasby A (2000) EMBOSS: the European molecular biology open software suite. *Trends Genet* 16:276–277
- Rodríguez MC, Petersen M, Mundy J (2010) Mitogen-activated protein kinase signaling in plants. *Annu Rev Plant Biol* 61:621–649
- Rosenkranz R, Borodina T, Lehrach H, Himmelbauer H (2008) Characterizing the mouse ES cell transcriptome with Illumina sequencing. *Genomics* 92:187–194
- Schulz MH, Zerbino DR, Vingron M, Birney E (2012) Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* 28:1086–1092
- Schuster SC (2008) Next-generation sequencing transforms today's biology. *Nat Methods* 5:16–18
- Skotnicki ML, Ninham JA, Selkirk PM (2000) Genetic diversity, mutagenesis and dispersal of Antarctic mosses—a review of progress with molecular studies. *Antarct Sci* 12:363–373
- Skotnicki ML, Mackenzie AM, Clements MA, Selkirk PM (2005) DNA sequencing and genetic diversity of the 18S–26S nuclear ribosomal internal transcribed spacers (ITS) in nine Antarctic moss species. *Antarctic Sci* 17:377–384
- Sun MM, Li LH, Xie H, Ma RC, He YK (2007) Differentially expressed genes under cold acclimation in *Physcomitrella patens*. *J Biochem Mol Biol* 40(6):986–1001
- Tibbett M, Sanders F, Cairney J (2002) Low-temperature-induced changes in trehalose, mannitol and arabinol associated with enhanced tolerance to freezing in ectomycorrhizal basidiomycetes (*Hebeloma* spp.). *Mycorrhiza* 12(5):249–255
- Wang X, Yang P, Zhang X, Xu Y, Kuang T, Shen S, He Y (2009) Proteomic analysis of the cold stress response in the moss, *Physcomitrella patens*. *Proteomics* 9(19):4529–4538
- Wang L, Feng Z, Wang X, Wang X, Zhang X (2010a) DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* 26:136–138
- Wang L, Li PH, Brutnell TP (2010b) Exploring plant transcriptomes using ultra high-throughput sequencing. *Brief Funct Genomics* 9:118–128
- Wei WL, Qi XQ, Wang LH, Zhang YX, Hua W, Li DH, Lv HX, Zhang XR (2011) Characterization of the sesame (*Sesamum indicum* L.) global transcriptome using Illumina paired-end sequencing and development of EST-SSR markers. *BMC Genomics* 12:451
- Wi SJ, Kim WT, Park KY (2006) Overexpression of carnation S-adenosylmethionine decarboxylase gene generates a broad-spectrum tolerance to abiotic stresses in transgenic tobacco plants. *Plant Cell Rep* 25:1111–1121
- Xiang LX, He D, Dong WR, Zhang YW, Shao JZ (2010) Deep sequencing-based transcriptome profiling analysis of bacteria-challenged *Lateolabrax japonicus* reveals insight into the immune relevant genes in marine fish. *BMC Genomics* 11:472
- Xiao L, Wang H, Wan P, Kuang T, He Y (2011) Genome-wide transcriptome analysis of gametophyte development in *Physcomitrella patens*. *BMC Plant Biol* 11:177
- Yang T, Chaudhuri S, Yang L, Du L, Poovaiah BW (2010) A calcium/calmodulin-regulated member of the receptor-like kinase family confers cold tolerance in plants. *J Biol Chem* 285(10):7119–7126
- Zerbino DR, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821–829
- Zhang H, Jin JP, Tang L, Zhao Y, Gu XC, Gao G, Luo JC (2011) PlantTFDB 2.0: update and improvement of the comprehensive plant transcription factor database. *Nucleic Acids Res* 39: D1114–D1117
- Zhou MQ, Shen C, Wu LH, Tang KX, Lin J (2011) CBF-dependent signaling pathway: a key responder to low temperature stress in plants. *Crit Rev Biotechnol* 31(2):186–192