eJBio

Electronic Journal of Biology, 2010, Vol. 6(3): 73-79

Genome-wide identification of PHYTOCHELATIN and PHYTOCH_SYNTH domain-containing phytochelatin family from rice

Guo-ming Shen[1, 3], Cheng Zhu[2,*], Qi-zhen Du[3]

[1]*State Key Laboratory of Plant Physiology & Biochemistry, College of Life Sciences, Zhejiang University, Hangzhou, 310058 China.*

[2] *College of Life Sciences, China Jiliang University, Hangzhou, 310018 China.*

[3] *School of Food Science and Biotechnology, Zhejiang Gongshang University, 149 Jiaogong Road, Hangzhou, 310012 China.*

* Corresponding author. Tel: +86(0)571 8691 4510; fax: +86(0)571 8691 4510; Email: pzhch@cjlu.edu.cn

Abstract

Phytochelatin synthases (PCS) are cytosol proteins that may play vital roles in heavy metal detoxicity in bacteria, yeasts, plants, or worms. More than 10 PCS-encoded genes from different species have been reported during the last decade, and they were played critical roles in cadmium or arsine detoxfication. In this investigation, we identified 3 putative *PCS* and 11 putative *PCS*-like genes in rice genome by the bioinformatics analysis of recently completed rice genome data, classified them into 2 subfamilies on the basis of domain sequence similarity by constructing phylogenetic trees, and localized them in rice chromosomes. Reverse transcription polymerase chain reaction (RT-PCR) was performed to demonstrate that the expression of 3 putative PCS genes (*OsPCS3*, *OsPCS8*, and *OsPCS11*) could not be detected in the roots, stem base, nodes, internodes or petioles, but all the remaining 11 putative rice *PCS* genes were found to be expressed in the leaves and panicles, *OsPCS7* was induced by Hg^{2+} and Pb^{2+}, while *OsPCS9* was induced by Cd^{2+} and Zn^{2+} in roots, suggesting that *PCS* or *PCS*-like genes might play crucial roles in survival or in different heavy metal detoxification.

Keywords: Gene family; heavy metal detoxification; phylogenetic evolution; phytochelatin synthase; *Oryza sativa*.

1. Introduction

Phytochelatins (PCs) are a class of post-translationally synthesized peptides that may play a vital role in heavy metal detoxification and accumulation in plants [1-3]. PCs combine heavy metals such as cadmium (Cd) or arsine (As) with high affinity, and localize themselves bound with the heavy metal ions to the cell vacuoles, thus playing a role in detoxification [1,2]. Their synthesis is mediated by phytochelatin synthase (PCS) (γ-glutamylcysteine dipeptidyltrans-peptidase, EC 2.3.2.15) using glutathione (GSH) and related thiol tripeptides as substrates, transfering a γ-Glu-Cys unit from one thiol peptide to another or to pre-existing phytochelatin molecules [4,5]. PCS is activated by heavy metals, such as Cd [4].

PCS genes firstly were cloned by three independent laboratories with different methods [6-8]. In plants, there had been reported *PCS* genes in *Arabidopsis*, wheat (*Triticum aestivum* L.), *Thlaspi caerulescens* L. J. & C. Presl and *Thlaspi japonicum* H., soybean (*Glycine max* L. Merr.), *Nostoc* sp. PCC 7120 L., *Pteris vittata* L., lettuce (*Lactuca sativa* L.), lotus (*Lotus japonicus* L.), rice (*Oryza sativa* L.) [3,6-15] and so on. Overexpression of *Arabidopsis thaliana* phytochelatin synthase (*AtPCS1*) gene in *Escherichia coli* (T. Escherich) or *Mesorhizobium huakuii* (Chen) enhanced the accumulation of heavy metals in those bacteria [16,17], while, overspression of *AtPCS1* in zebrafish (*Danio rerio* Hamilton-Buchanan) or tobacco (*Nicotiana tobacum* L.) enhanced heavy metal tolerance [18,19]. Interestingly, overexpression of *AtPCS1* in *Arabidopsis* was paradoxically hypersensitive to cadmium [20,21]. Therefore, the functions of *PCS* genes and their homologues and their roles in heavy metal detoxification need to be further investigated in other plants.

The recent completion of the rice genome-sequencing project, together with its automated annotation process, enabled us for the first time to gauge the number of PCS family with a typical angiosperm [22]. In this report, we identified 3 *PCS* and 11 *PCS*-like genes in rice genome using bioinformatics methods, and analyzed their expression patterns by Reverse transcription polymerase chain reaction (RT-PCR).

2. Materials and Methods

Plant materials

Rice (*Oryza sativa* L. ssp *japonica* variety, Zhonghua 11) seeds were surface sterilized in 0.5%

sodium hypochlorite for 20 min, rinsed, and germinated in the dark on moistened filter paper at 30°C for 2 d, and then transferred the germinated seeds to 96-hole plastic floatings for 4 d. The uniformly germinated seedlings were transferred to black polyethylene barrels which contain 8 L rice culture solution [23] in growth chambers (MC1000 system; Snijders) at temperature regimes of 30/24°C (day/night) and 70% humidity under a 12-h photoperiod (photo flux density of 500 μM s^{-1} m^{-2}) during the growth period, the nutrient solution was replaced every five day. The roots, stembase, node, internode, petiole, leaves, and panicles were harvested and dipped in liquid nitrogen and then stored at -80°C until RNA extraction. For metal-induced gene expression experiment, 9-d old seedlings were transferred to half-strength rice culture solutions containing 100 mM $CdCl_2$, 100 mM $CuCl_2 \cdot 2H_2O$, 100 mM $ZnSO_4 \cdot 7H_2O$, 100 mM $HgCl_2$, 100 mM $MnCl_2$, 100 mM $Pb(NO_3)_2$, 100 mM $CoCl_2$ or 100 mM $AgNO_3$, respectively, for 1 d. The roots were harvested and dipped in liquid nitrogen and then stored at -80°C until RNA extraction.

cDNA synthesis and RT-PCR analysis

Total RNA was extracted from the roots, stembase, nodes, internodes, petioles, leaves and panicles in mature plants using Trizol reagent (Invitrogen, USA). cDNA synthesis was performed as described by Reale *et al* [24]. Reverse transcription polymerase chain reaction (RT-PCR) was conducted in a 25 μL reaction containing 20 pmole of *OsTUB16* (accession number X78143) primers (5'-CGCCT CTGCCATGTTCCGTGGAA-3' and 5'-GGCGGTAA TACGGTGATAATGTAA-3'), and gene-specific primers (see Supplemental Table 1), 10 mM dNTPs, 5 unit of Ex-Taq DNA polymerase (TakaRa, Japan), and 10×reaction buffer. RT-PCR was performed under optimal conditions for each gene, and the numbers of reaction cycles were 33~35. Afterward, 5 μL aliquots of the reaction mixtures were separated on 1.2% (w/v) agarose gels.

Sequence analysis and construction of the phylogenetic tree

Database searching, editing, and sequence alignments described as Cai and Lytton [25], sequence analysis of a variety in molecular evolutionary genetics analysis (MEGA) version 4.0 using the ClustalW program, building phylogenetic trees also use the MEGA 4.0 software package. Calculating the sequences of different base composition, the percentage of variation sites, codon usage, the percentage of conversions and transversions, or the percentage of transversions as the methods described by Tamura [26]. Two-parameter method to calculate the genetic distance between branches, using neighbor-joining, minimum evolution and maximum parsimony method for system reconstruction as described by Saitou and Nei [27]. Phylogenetic tree of the confidence level of the branch by re-sampling method (bootstrap) 1000 repeated testing, DNA sequence variation of the conversions and transversions given the same valuation were used in this investigation.

Determination of conserved domain

Conservative functional domain analysis was carried out by using DNAMAN version 6.0 (Lynnon corporation, Vaudreuil-Dorion, Canada), Hidden Markov Model (HMM) [28] and on-line version of GlobPlot™ 2.3 software package (http://globplot.embl.de/) [29].

3. Results

3.1 Identification and structure of PCS and PCS-lick genes in rice

To identify the *PCS* family genes in rice, basic local alignment search tool (BLAST) [30] searches of the rice databases were performed using the PHYTOCHELATIN domain (pfam05023) of the *Arabidopsis* AtPCS1 protein as a query sequence. Three genes were identified to possibly encode PHYTOCHELATIN domain, namely, *OsPCS1*, *5* and *9* (Table 1). The search of "gene" using "phytochelatin synthase (*PCS*)" as a keyword in NCBI matched 7 homologues of *OsPCS 2-4, 6, 8, 10* and *12*. tBLASTn searches of the rice database were performed using the rice OsPCS2 protein as a query sequence, and other 4 genes were matched.
In order to predict the genetic classification of PCS in this study, we have built phylogenetic trees based on the full-length amino acid sequences [26], analysis showed that: the predicted 14 genes can be divided into two sub-group as showed in Figure 1a. In RiceGAAS (http:// ricegaas.dna.affrc.go.jp) web site on our forecast of 14 genes in exon-intron composition analysis results showed that 10 of the 14 OsPCS genes have 6 exons, 1 has 3 exons, and the other 3 have no intron (Figure1b).

We explored protein sequences for intrinsic protein disorder, domain and globularity prediction by software GlobPlot™ version 2.3. Unfortunately, PCS1, 5, 13 or 14 were not detected PHYTO-CHELATIN or PHYTOCHEL_SYNTH domain (see Supplemental Figure 1).

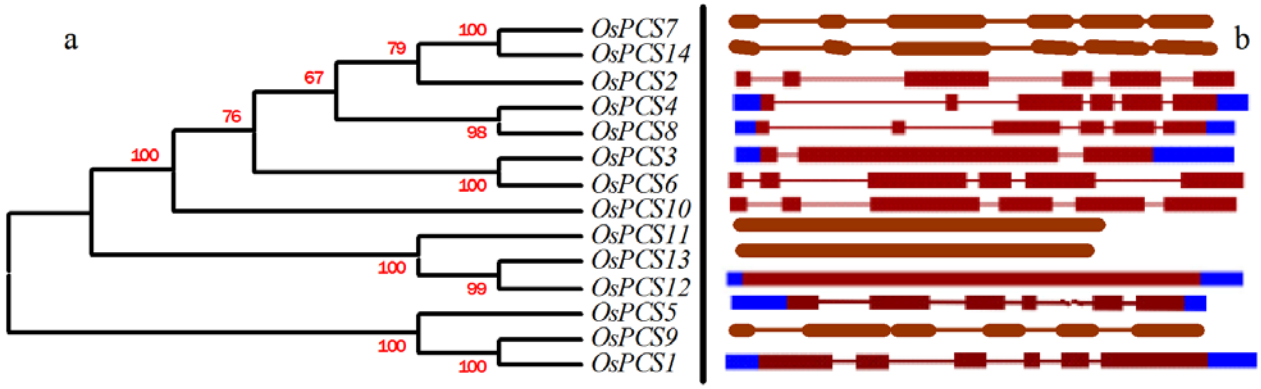

Figure 1. Phylogenetic analysis, exon-intron organi-zation of *PCS* genes in *Oryza sativa*. Neighbor-joining trees of full-length amino acid sequences encoded by rice *PCS* genes are shown. Bootstrap values (1,000 replicates) are placed at the nodes, and the scale bar corresponds to 0.1 estimated amino acid substitutions per site (a). The exon-intron organization of the corres-ponding *PCS* genes is shown for the rice (b) gene family. Exons are depicted as red boxes, introns as connecting thin lines, and untranslated region as blue boxes.

3.2 Chromosomal distribution

We used the data from Rice Genome Research Program (GRP) website (http:// rgp.dna.affrc.go.jp/) on the corresponding YAC or BAC cloning genetic map database to localize the map position of the predicated *PCS* genes [22]. Figure 2 showed that the distribution of the 14 *OsPCS* genes on different chromosomes in the rice genome. All of the predicted *OsPCS* sequences were distributed within chromosomes 3~7 and 10, and no homologues were located at other chromosomes. Although the 6 chromosomes had homologues, each differed in the number of *OsPCS* genes it contained chromo-somes 4 and 10 each contained only 1 *OsPCS* gene, chromosome 6 contained 2, chromosomes 5 and 7 each contained 3, and chromosome 3 had 4. *OsPCS 3-5* and *13* oriented the of short arm of rice chromosome, and the rest located in the long arm of rice chromosome, while *OsPCS 3*, *4*, *7* and *8* is located near the centromere, *OsPCS 5*, *6* and *12* located near the end of the long arm (from the end of the genetic map distance of less than 2 centi Morgan).

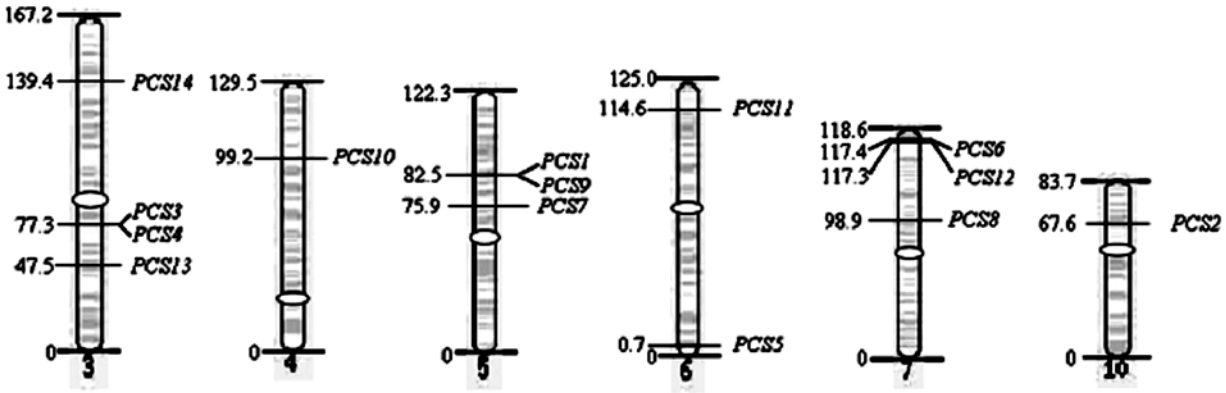

Figure 2. Chromosomal mapping of three *PHYTO-CHELATIN* domain *PCS* and eleven *PHYTOCHEL_ SYNTH* domain *PCS*-like genes from the complement rice genome sequence. Gene names were marked on the right side of chromosomes and genetic positions were on the left of chromosomes.

3.3 Expression pattern of PCS and PCS-like genes

In order to verify whether the predicted genes can be truly expressed and to know their expression patterns, we have extracted RNA in different organs from the rice. Before reverse transcription, RNAs were treated with RNase-free Dnase to avoid genomic DNA contamination for the reliability of the experiment. Primers were designed based on coding open reading frame regions of *PCS* genes and are shown in Supplemental Table 1. The RT-PCR results show that not all predicted genes were expressed in tissues tested under normal growth conditions. *OsPCS3, OsPCS8, OsPCS11* genes could not be detected even if we modified the RT-PCR conditions. The other 11 *OsPCS* genes expressed almost in all the leaves and panicles, and different expression patterns were observed in the roots, stem base, nodes, internodes, and petioles (Figure 3a). To investigate metal inducing gene expression pattern, we selected *OsPCS7* standing

for the *PCS*-like subfamily and *OsPCS9* standing for the *PCS* family. *OsPCS7* was induced by Hg^{2+} and Pb^{2+}, while *OsPCS9* was induced by Cd^{2+} and Zn^{2+} (Figure 3b).

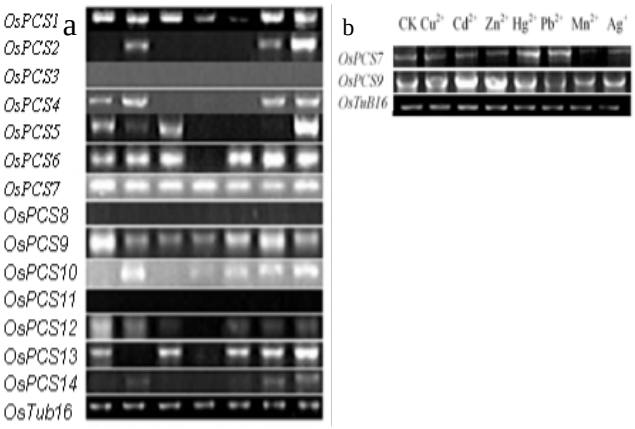

Figure 3. RT-PCR analysis of *PCS* genes. *OsPCS* gene family using mRNA extracted from roots(R), stembase (SB), node (N), inter-node (IN), petiole (Pe), leaves (L) and panicle (Pa) from mature plant (a). *OsPCS7 and OsPCS9* gene using mRNA extracted from seedling roots (10d) (b).

4. Discussions

4.1 Family relationships of putative rice PCS/ PCS-like sequences

We investigated the relationship among these putative rice PCS/PCS-like sequences by generating an alignment of 14 identified PCS amino acid sequences followed by generation of a neighbor-joining phylogenetic tree (Figure 1a). The combined phylogeny PCS sequences revealed two subfamilies of putative orthologous genes, namely, *OsPCS2*, *3*, *4*, *6*, *7*, *8*, *10*, *11*, *12*, *13* and *14*; and *OsPCS1*, *5*, and *9*, suggesting that *PCS* and *PCS*-like genes were divergent long years ago. *OsPCS2*, *OsPCS 7* and *OsPCS 14* have similar sequences to *OsPCS11*, *OsPCS12,* and *OsPCS13,* but are distributed in different 3 chromosomes. Other genes, *OsPCS6* and *OsPCS8*, *OsPCS3* and *OsPCS4,* are distributed in the same chromosome but with long evolutionary distance, and *OsPCS1* and *OsPCS9* are located in the same chromosome and are more closely related to each other (Figure 2), indicating that the divergence of genes was not relative to chromosome distribution.

4.2 The PCS genes might be divided into two subfamilies

Based on the phylogenetic tree of *OsPCS* genes, we inferred that phytochelatin synthase genes in eukaryotic organisms might be divided into two subfamilies, the PHYTOCHELATIN (PCS) subfamily serving as heavy metal chelators as formerly reported [31,32], and the unknown function gene family, namely, the PHYTOCHEL_SYNTH (PCS-like) subfamily.

To validate our hypothesis, firstly, we searched in NCBI (www.ncbi.nlm.nih.gov) using the keywords “phytochelatin synthase,” and there were 94 items matched in protein database, 22 in gene database and 115 in nucleotide database. Secondly, we searched in UniProtKB/Swiss-Prot (http://www.ebi.ac.uk/swiss-prot/) also using the keyword “phytochelatin synthase,” and there were 48 entries in protein database, and 157 in nucleotide database. Some retrieved sequences were discarded on the basis of the criteria described in Materials and Methods. All the remaining 87 sequences with information on species, sizes, and their accession numbers are available in Table 1. For the remote relationship between those family members, divergences take place not only in the peptide sequences but also in the nucleotide sequences, so it is very difficulty to put those subfamily members into a phylogenetic tree. We selected eukaryotic phytochelatin synthases to assess the evolutionary relationship between those subfamily members. Our results clear demonstrate that those family members can be classified into two subfamilies (Figure 4). To confirm our inference, we used BLAST to search all of the 87 sequences in NCBI, and found that 2 mono-domains, namely, the “PHYTO-CHELAYIN” and the “PYTOCHEL_SYNTH,” were matched. All of the “PHYTO- CHELAYIN” domain family members have conserved amino acids of Q, C, G, GH, P, and L in all organism species (see Supplemental Figure 2a), while the “PYTOCHEL_ SYNTH” domain family is a C-rich family [28] with high probability in site of 49, 69, 95, and 149 in the full domain (see Supplemental Figure 2b). Thus, phylogenetic relationships analysis of all predicted *PCS* genes from NCBI database clearly supports our hypothesis.

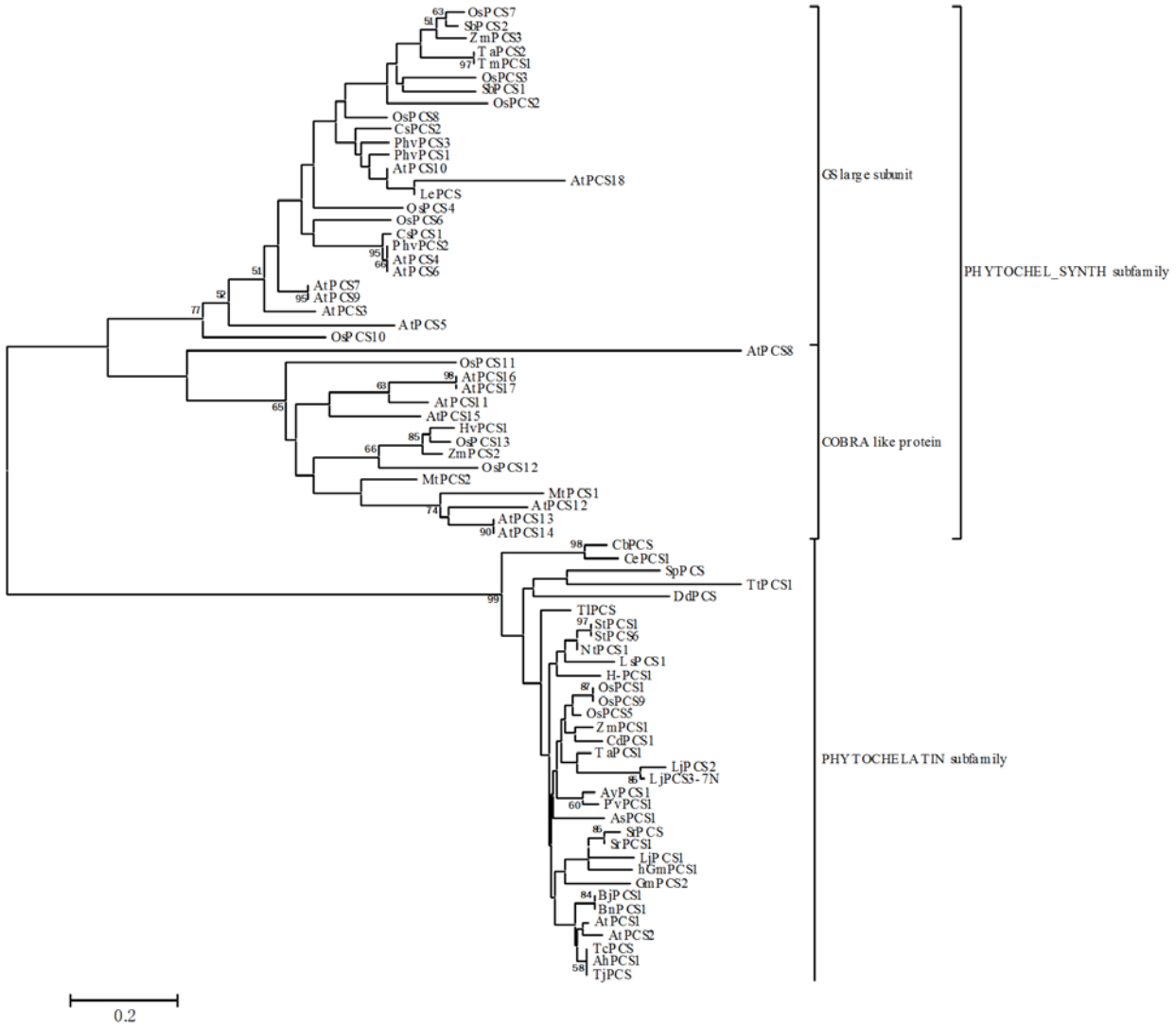

Figure 4. Phylogenetic relationships of the PHYTOCHELATIN and PHYTOCHEL_SYNTH domains encoded by *PCS* genes from NCBI database. The unrooted tree was constructed from the alignment using Clustal W and the neighbor-joining method with MEGA 4.0. Numbers on branches indicate the percentage of 1000 bootstrap replicates that support the adjacent node; low bootstrap support (<50%) was not reported. The scale bar corresponds to 0.1 estimated amino acid substitutions per site.

Table1. A list of *PCS* and *PCS*-like genes.

Gene Name	Species	Taxonomy	Common Name	Protein NCBI Acc	Sequence GI	Protein Length
*AhPCS1**	***Arabidopsis halleri***	Magnoliophyta	Cress	AAS45236	42742267	485
*AsPCS1**	*Allium sativum*	Magnoliophyta	Garlic	AAO13809	27448224	506
*AtPCS1**	*Arabidopsis thaliana*	Magnoliophyta	Mouse ear cress	AAD29446	4768281	502
AtPCS10	*Arabidopsis thaliana*	Magnoliophyta	Mouse ear cress	AAM19781	20453072	456
AtPCS11	*Arabidopsis thaliana*	Magnoliophyta	Mouse ear cress	BAB03074	11994745	653
AtPCS12	*Arabidopsis thaliana*	Magnoliophyta	Mouse ear cress	AAO64887	29029016	672
AtPCS13	*Arabidopsis thaliana*	Magnoliophyta	Mouse ear cress	BAD43466	51969548	668
AtPCS14	*Arabidopsis thaliana*	Magnoliophyta	Mouse ear cress	CAB79566	7269564	717
AtPCS15	*Arabidopsis thaliana*	Magnoliophyta	Mouse ear cress	BAB10345	10177156	663
AtPCS16	*Arabidopsis thaliana*	Magnoliophyta	Mouse ear cress	AAO64838	29028918	515
AtPCS17	*Arabidopsis thaliana*	Magnoliophyta	Mouse ear cress	CAB78654	7268361	661
AtPCS18	*Arabidopsis thaliana*	Magnoliophyta	Mouse ear cress	BAB10644	10177318	204
*AtPCS2**	*Arabidopsis thaliana*	Magnoliophyta	Mouse ear cress	AAK94671	16519291	452
AtPCS3	*Arabidopsis thaliana*	Magnoliophyta	Mouse ear cress	AAO63437	28951027	452
AtPCS4	*Arabidopsis thaliana*	Magnoliophyta	Mouse ear cress	CAC01762	9755608	395
AtPCS5	*Arabidopsis thaliana*	Magnoliophyta	Mouse ear cress	AAB60732	2160169	454
AtPCS6	*Arabidopsis thaliana*	Magnoliophyta	Mouse ear cress	AAT44976	48525349	431
AtPCS7	*Arabidopsis thaliana*	Magnoliophyta	Mouse ear cress	AAG12670	10092257	404

AtPCS8	*Arabidopsis thaliana*	Magnoliophyta	Mouse ear cress	NP_683544	22330948	170
AtPCS9	*Arabidopsis thaliana*	Magnoliophyta	Mouse ear cress	AAM67179	21618129	441
*AvPCS**	*Anabaena variabilis*	Cyanophyceae	Anabaena flos-aquae	ABA22569	75702893	239
*AyPCS1**	*Athyrium yokoscense*	Filicophyta	Athyriaceae	BAB64932	15617416	488
*BjPCS1**	*Brassica juncea*	Magnoliophyta	Leaf mustard	CAC37692	13928024	485
*BmPCS**	*Burkholderia mallei*	Proteobacteria	Proteobacteria	YP_103177	53723721	243
*BnPCS1**	*Brassica napus*	Magnoliophyta	Oilseed rape	CAK24968	100801748	486
*BpPCS**	*Burkholderia pseudomallei*	Proteobacteria	Proteobacteria	YP_333954	76810443	245
*BtPCS**	*Burkholderia thailandensis*	Proteobacteria	Proteobacteria	YP_442569	83720815	243
*BxPCS**	*Burkholderia xenovorans*	Proteobacteria	Proteobacteria	ABE34681	91691483	247
*CbPCS**	*Caenorhabditis briggsae*	Nematoda	Nematode*	CAE59471	39597243	372
*CdPCS1**	*Cynodon dactylon*	Magnoliophyta	Bermudagrass	AAO13810	33187146	504
*CePCS1**	*Caenorhabditis elegans*	Nematoda	Nematode*	AAK62992	14485520	371
CsPCS1	*Cleome spinosa*	Magnoliophyta	Spiny spiderflower	ABD96834	90657534	440
CsPCS2	*Cleome spinosa*	Magnoliophyta	Spiny spiderflower	ABD96927	90657629	436
*DdPCS**	*Dictyostelium discoideum*	Mycetozoa	Amoeba	XP_635353	66803020	626
GmPCS2	*Glycine max*	Magnoliophyta	Soybean	AAX03071	59773938	499
*hGmPCS1**	*Glycine max*	Magnoliophyta	Soybean	AAL78384	18699092	498
*H-PCS1**	*Helianthus sp.*	Magnoliophyta	Sunflower	AAX03068	59773932	486
*HvPCS1**	*Hordeum vulgare*	Magnoliophyta	Barley	CAD42639	38174821	379
LePCS	*Lycopersicon esculentum*	Magnoliophyta	Tomato	BT013422(n)	4710483 7(n)	444
*LjPCS1**	*Lotus japonicus*	Magnoliophyta	Japanese mallotus	AAQ01752	33286859	501
*LjPCS3-7N**	*Lotus japonicus*	Magnoliophyta	Japanese mallotus	AAY81941	67773368	479
*LsPCS1**	*Lactuca sativa*	Magnoliophyta	Lettuce	AAU93349	53760453	490
MtPCS1	*Medicago truncatula*	Magnoliophyta	Barrel medic	ABE91155	92892137	667
MtPCS2	*Medicago truncatula*	Magnoliophyta	Barrel medic	ABE84746	92877766	697
*N-PCS**	*Nostoc sp.*	Cyanobacteria	Nostoc	BAD10973	42491328	242
*NtPCS1**	*Nicotiana tabacum*	Magnoliophyta	Common tobacco	AAO74500	29470177	501
*OsPCS1**	*Oryza sativa*	Magnoliophyta	Rice	AAO13349	48448412	473
OsPCS10	*Oryza sativa*	Magnoliophyta	Rice	CAE02786	38344980	439
OsPCS11	*Oryza sativa*	Magnoliophyta	Rice	BAD45565	52076665	683
OsPCS12	*Oryza sativa*	Magnoliophyta	Rice	BAC16043	22831184	675
OsPCS13	*Oryza sativa*	Magnoliophyta	Rice	ABF95484	108707689	671
OsPCS14	*Oryza sativa*	Magnoliophyta	Rice	NONE	NONE	457
OsPCS2	*Oryza sativa*	Magnoliophyta	Rice	AAL58276	18071417	425
OsPCS3	*Oryza sativa*	Magnoliophyta	Rice	AAS07098	41469147	468
OsPCS4	*Oryza sativa*	Magnoliophyta	Rice	AAS07075	41469124	458
*OsPCS5**	*Oryza sativa*	Magnoliophyta	Rice	AAX03070	59773936	502
OsPCS6	*Oryza sativa*	Magnoliophyta	Rice	BAC83873	34394570	451
OsPCS7	*Oryza sativa*	Magnoliophyta	Rice	AAV31332	54287588	457
OsPCS8	*Oryza sativa*	Magnoliophyta	Rice	BAC83872	34394569	446
*OsPCS9**	*Oryza sativa*	Magnoliophyta	Rice	AAV32132	54291763	543
PhvPCS1	*Phaseolus vulgaris*	Magnoliophyta	Garden bean	AAR13303	38194915	414
PhvPCS2	*Phaseolus vulgaris*	Magnoliophyta	Garden bean	AAR13305	38194917	463
PhvPCS3	*Phaseolus vulgaris*	Magnoliophyta	Garden bean	AAR13304	38194916	448
*PmPCS**	*Prochlorococcus marinus*	Cyanobacteria	Cyanobacteria	CAE21188	33640733	249
*PvPCS1**	*Pteris vittata*	Filicophyta	Chinese brake	AAT11885	47155943	476
SbPCS1	*Sorghum bicolor*	Magnoliophyta	Sorghum	AAO17705	30090025	425
SbPCS2	*Sorghum bicolor*	Magnoliophyta	Sorghum	AAO17706	30090026	449
*S-PCS**	*Synechococcus sp.*	Cyanobacteria	Cyanobacteria	ZP_01086227	87303439	271
*SpPCS**	*Schizosaccharomyces pombe*	Fungi	Fission yeast	Q10075	1351693	414
*SrPCS**	*Sesbania rostrata*	Magnoliophyta	Sesbania	AAY83876	67944509	465
*SrPCS1**	*Sesbania rostrata*	Magnoliophyta	Sesbania	AAY82881	67906844	233
*s-SpPCS**	*Strongylocentrotus purpuratus*	Metazoa	Purple urchin*	XP_780594	72012505	233
*StPCS1**	*Solanum tuberosum*	Magnoliophyta	Potato	CAD68110	28569706	503
*StPCS6**	*Solanum tuberosum*	Magnoliophyta	Potato	CAD68107	28569700	467
*TaPCS1**	*Triticum aestivum*	Magnoliophyta	Wheat	AAD50592	5757804	500
TaPCS2	*Triticum aestivum*	Magnoliophyta	Wheat	AAR14311	38230578	456
*TcPCS**	*Thlaspi caerulescens*	Magnoliophyta	Alpine penny-cress	BAB93120	21104518	485
*TjPCS**	*Thlaspi japonicum*	Magnoliophyta	Japanese cress	BAB93119	21104516	485
*TlPCS**	*Typha latifolia*	Magnoliophyta	Common cattail	AAG22095	24963931	421
*TmPCS1**	*Triticum monococcum*	Magnoliophyta	One-grained wheat	AAO86520	30090032	457

*TtPCS1**	*Tetrahymena thermophila*	Alveolata	Tetrahymena *	AAY68362	76096940	446
*ZmPCS1**	*Zea mays*	Magnoliophyta	Maize	AAX03063	59773922	476
ZmPCS2	*Zea mays*	Magnoliophyta	Maize	AAX03069	59773934	507
ZmPCS3	*Zea mays*	Magnoliophyta	Maize	AAQ81633	34733385	667

**PCS* marked asterisk.

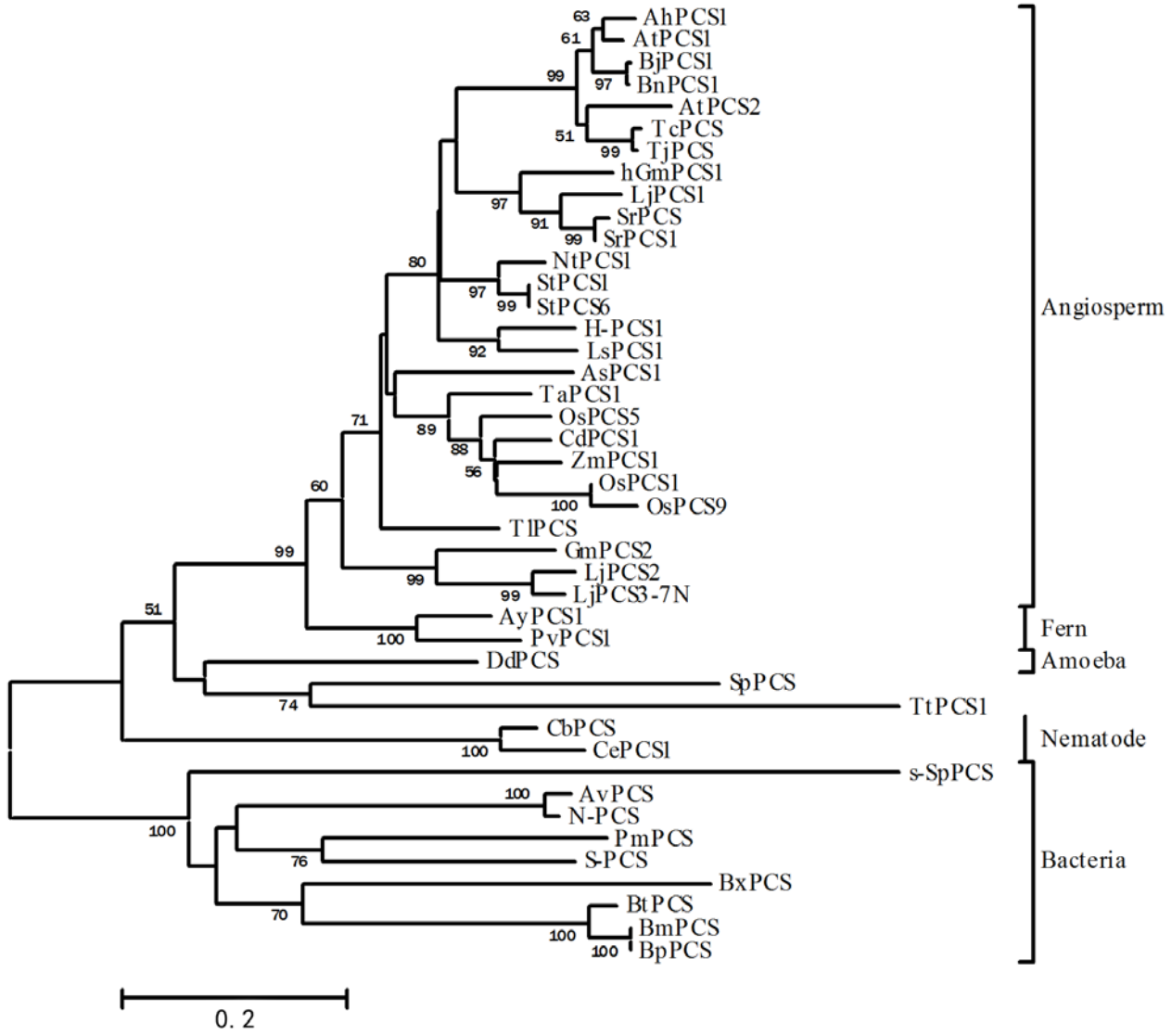

Figure 5. Phylogenetic relationships of the *PHYTOCHELATIN* domain encoded by *PCS* genes from NCBI database. The unrooted tree was constructed from the alignment using MEGA 4.0 and the neighbor-joining method. Numbers on branches indicate the percentage of 1000 bootstrap replicates that support the adjacent node; low bootstrap support (<50%) was not reported. The classification clades annotated at right.

4.3 The PHYTOCHELATIN domain family diversity monogene to multi-gene family

We used the mono-domain of Arabidopsis AtPCS1 to tBlastn and searched for "Phytochelatin" subfamily members in model organisms including all 30 species displayed in NCBI, and we also used tblastn to search in TIGR (www.tigr.org). We found that only 6 species (*Dictyostelium discoideum*, *Schizosaccharomyces pombe*, *Canorhabditis elegans*, *Arabidopsis thaliana*, *Oryza sativa* and *Zea mays*) contain the "PYTOCHEL_SYNTH" family genes. Interestingly, our result indicates that the genesis of "PYTOCHELATIN" gene subfamily might come from their mono-gene ancient ancestor, for in *Dictyostelium discoideum*, *Schizosaccharomyces pombe* and *Canorhabditis elegans* genomes only one gene was matched, while in *Arabidopsis thaliana*, *Oryza sativa* and *Zea mays* genomes more than one gene were matched (Figure 5). And the "PYTOCHEL_SYNTH" family might have evolved from the ancient ancestor of "PYTOCHELATIN" family.

4.4 The bacteria PCS chelate heavy metal ions might not through Cys amino acids

Multi-alignment analysis of the functional domain of "PYTOCHELATIN" in bacteria showed that there was only one Cys in the conserved domain (Supplemental Figure 2a), indicating that bacterial chelate heavy metal ions might have different mechanisms compared to eukaryotic ones.

5. Conclusions

Phytoremediation is an inovation area which interested more investigators who are concerned on heavy metal contamination. PCS involved critical roles in heavy metal remediation, predicating new PCS genes will inforce phytoremediation by transgenetic methods. And our findinds will interest those who are get in with phytoremediation.

Furthermore, the PCS-like family containing the cell wall protein COBRA-like and glutathione synthase (GS) large subunit (data not shown), will uncover the roles of COBRA-like and GS large subunit in heavy metal detoxification in the future.

Acknowledgments

This work was supported by grants from the Project of National Key Basic Research and Development of China (2007CB109305) and the National Natural Science Foundation of China (NO. 30671255).

References

[1] Cobbett C.S. (2000a) Phytochelatin biosynthesis and function in heavy-metal detoxification. *Curr. Opin. Plant Biol.*, **3**: 211-216.

[2] Cobbett C.S. (2000b) Phytochelatins and their roles in heavy metal detoxification. *Plant Physiol.*, **123**: 825-832.

[3] Li J.C., Guo J.B., Xu W.Z., Ma M. (2007) RNA interference-mediated silencing of phytochelatin synthase gene reduce cadmium accumulation in rice seeds. *J. Integr. Plant Biol.*, **49**(7): 1032-1037.

[4] Vatamaniuk O.K., Mari S., Lu Y.P., Rea P.A. (2000) Mechanism of heavy metal ion activation of phytochelatin (PC) synthase: blocked thiols are sufficient for PC synthase-catalyzed transpeptidation of glutathione and related thiol peptides. *J. Biol. Chem.*, **275**: 31451-31459.

[5] Piechalak A., Tomaszewska B., Barałkiewicz D. (2003) Enhancing phytoremediative ability of *Pisum sativum* by EDTA application. *Phytochemistry*, **64**(7): 1239-1251.

[6] Clemens S., Kim E.J., Neumann D., Schroeder J.I. (1999) Tolerance to toxic metals by a gene family of phytochelatin synthases from plants and yeast. *EMBO J.*, **18**(12): 3325-3333.

[7] Ha S.B., Smith A.P., Howden R., Dietrich W.M., Bugg S., O'Connell M.J., Goldsbrough P.B., Cobbett C.S. (1999) Phytochelatin synthase genes from *Arabidopsis* and the yeast *Schizosaccharomyces pombe*. *Plant Cell,* **11**: 1153-1164.

[8] Vatamaniuk O.K., Mari S., Lu Y.P., Rea P.A. (1999) *AtPCS1*, a phytochelatin synthase from Arabidopsis: isolation and *in vitro* reconstitution. *Proc. Natl. Acad. Sci. USA.*, **96**: 7110-7115.

[9] Cazale A.C., Clemens S. (2001) *Arabidopsis thaliana* expresses a second functional phytochelatin synthase. *FEBS Lett.*, **507**(2001): 215-219.

[10] Mizuno T., Sonoda T., Horie K., Senoo K., Tanaka A., Mizuno N., Obata, H. (2003) Cloning and characterization of phytochelatin synthase from a nickel hyperaccumulator *Thlaspi japonicum* and its expression in yeast. *Soil Sci. Plant Nutr.*, **49**: 285-290.

[11] Oven M., Page J.E., Zenk M.H., Kutchan T.M. (2002) Molecular characterization of the homo-phytochelatin synthase of soybean *Glycine max*: relation to phytochelatin synthase. *J. Biol. Chem.,* **277**: 4747-4754.

[12] Tsuji N. , Nishikori S., Iwabe O., Shiraki K.,Miyasaka H., Takagi M., Hirata K., Miyamoto K. (2004) Characterization of phytochelatin synthase-like protein encoded by alr0975 from a prokaryote, *Nostoc sp*. PCC 7120. *Biochem. Biophys. Res. Comm*., **315**(3): 751-755.

[13] Dong R. (2005) Molecular cloning and characterization of a phytochelatin synthase gene, *PvPCS1*, from *Pteris vittata* L. *J. Ind. Microbiol. Biot.*, **32**(11-12): 527-533.

[14] He Z.Y., Lia J.C., Zhang H.Y., Ma M. (2005) Different effects of calcium and lanthanum on the expression of phytochelatin synthase gene and cadmium absorption in *Lactuca sativa*. Plant Sci., **168**(2): 309-318.

[15] Loscos J., Naya L., Ramos J., Clemente M.R., Matamoros M.A., Becana M. (2006) A reassessment of substrate specificity and activation of phytochelatin synthases from model plants by physiologically relevant metals. *Plant Physiol.*, **140**: 1213-1221.

[16] Sauge-Merle S., Cuine S., Carrier P., Lecomte-Pradines C., Luu D.T., Peltier G. (2003) Enhanced toxic metal accumulation in engineered bacterial cells expressing *Arabidopsis thaliana* phytochelatin synthase. *Appl. Environ. Microbiol.*, **69**(1): 490-494.

[17] Sriprang R., Hayashi M., Ono H., Takagi M., Hirata K., Murooka Y. (2003) Enhanced accumulation of Cd^{2+} by a *Mesorhizobium* sp. transformed with a gene from *Arabidopsis thaliana* coding for phytochelatin synthase. *Appl. Environ. Microbiol.*, **69**(3): 1791-1796.

[18] Konishi T., Matsumoto S., Tsuruwaka Y., Shiraki K., Hirata K., Tamaru Y., Takagi M. (2006) Enhancing the tolerance of zebrafish (*Danio rerio*) to heavy metal toxicity by the expression of plant phytochelatin synthase. *J.Biotech.*, **122**(3): 316-325.

[19] Pomponi M., Censi V., Di Girolamo V., De Paolis A., Di Toppi L.S., Aromolo R., Costantino P., Cardarelli M. (2006) Overexpression of *Arabidopsis* phytochelatin synthase in tobacco plants enhances Cd^{2+} tolerance and accumulation but not translocation to the shoot. *Planta*, **223**(2): 180-190.

[20] Lee S., Moon J.S., Ko T.S., Petros D., Goldsbrough P.B., Korban S.S. (2003) Overexpression of *Arabidopsis* phytochelatin synthase paradoxically leads to hypersensitivity to cadmium stress. *Plant Physiol.*, **131**: 656-663.

[21] Li Y., Dhankher O.P., Carreira L., Lee D., Chen A., Schroeder J.I., Balish R.S., Meagher R.B. (2004) Overexpression of phytochelatin synthase in *Arabidopsis* leads to enhanced arsenic tolerance and cadmium hypersensitivity. *Plant Cell Physiol.*, **45**(12): 1787-1797.

[22] Jiang S.Y., Jasmin P.X., Ting Y.Y., Ramachandran S. (2005) Genome-wide identification and molecular characterization of Ole_e_I, Allerg_1 and Allerg_2 domain-containing Pollen-Allergen-like genes in *Oryza sativa*. *DNA Res.*, **12**(3): 167-179.

[23] Yoshida S., Forno D.A., Cock J.H. (1971) Laboratory manual for physiological studies of rice. Manila, The Philippines: International Rice Research Institute.

[24] Reale S., Doveri S.D., Angiolillo, A., Lucentini L., Pilla F., Martín A., Donini P., Lee D. (2006) "SNP-based markers for discriminating olive (*Olea europaea* L.) cultivars. (single- ucleotide polymorph", *Genome*, **49**(9): 1193-1205.

[25] Cai X.J., Lytton J. (2004) The cation/Ca^{2+} exchanger superfamily: phylogenetic analysis and structural implications. *Mol. Biol. Evol.*, **21**(9): 1692-1703.

[26] Tamura K., Dudley J., Nei M., Kumar S. (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.,* **24**(8): 1596-1599.

[27] Saitou N., Nei M. (1987) The neighbor-joining method: a new method for reconstruction of phylogenetic trees. *Mol. Biol. Evol.*, **4**: 406-425.

[28] Schuster-Böckler B., Schultz J., Rahmann S. (2004) HMM Logos for visualization of protein families. *BMC Bioinformatics*, **5**: 7.

[29] Linding R., Russell R.B., Neduva V., Gibson T.J. (2003) GlobPlot: exploring protein sequences for globularity and disorder. *Nucl. Acids Res.*, **31**(13): 3701-3708.

[30] Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller, W., Lipman D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.*, **25**(17): 3389-3402.

[31] Cobbett C.S., Goldsbrough P. (2002) Phytochelatins and metallothioneins: roles in heavy metal detoxification and homeostasis. *Ann. Rev. Plant Biol.*, **53**: 159-182.

[32] Rea P.A., Vatamaniuk O.K., Rigden D.J. (2004) Weeds, worms, and more. Papain's long-lost cousin, phytochelatin synthase. *Plant Physiol.*, **136**: 2463-2474.

Supplemental Files

Table 1. Primers used for RT-PCR analyses.

Gene name	Primer sequence (Forward)	Primer sequence (Reverse)
PCS1	5'- CAGCCTCGTCTCCGTCTTCC-3'	5'- GCAATGCGGTTGCTTCACAG-3'
PCS2	5'-TATAGATGGCGATTGGTGTT-3'	5'-TGCGGTAGTTCAAGTTAGTG-3'
PCS 3	5'-GGCGACTGCTCCAAGTTCAA-3'	5'- TCGCATCAGCACCTCCGACT-3'
PCS4	5'-ATGGCTATGTTGCTGTTGTT-3'	5'- TACTAAGTTCCACTGCGAGT-3'
PCS5	5'- CTGATGACGCAACTGGTCTA-3'	5'- TCGAAACAAAGGTAATACGG-3'
PCS6	5'- GGCGACTGCTCCCGCTTCAA-3'	5'-ACGCCCATCCCTGCCTGAAC-3'
PCS7	5'-CTGGGACTACCAACAAGACA-3'	5'-CAACCGTGGAACTGAATGAA-3'
PCS8	5'-GCAGTGCCAAGGTTGTGAGA-3'	5'-ACGACGGTTTATGGTTAGGT-3'
PCS9	5'-ATGGGGGCGGAGGTCCATGA-3'	5'-TCAATGCAAGGTTCTAGGAGTGA-3'
PCS10	5'-CTGTCGTCGCTCACGCAGAA-3'	5'-TGAATGGCAACAACAAGCAC-3'
PCS11	5'-AACACCACCACCCGCTACCT-3'	5'-AACCGCCACAACCATCACAA-3'
PCS12	5'- CGCTATGCCGTCTTGTTTGC-3'	5'-GGTCGCAGTTGAGCACCTTG-3'
PCS13	5' GCGACCTCGTCATCACCTAC-3'	5'-ACGCCACAGTAATCACCACA-3'
PCS14	5'- AGAAGGAGGTCATTTGGTCG-3'	5'- AAAGGTGAAAGCCATAGAAT-3'

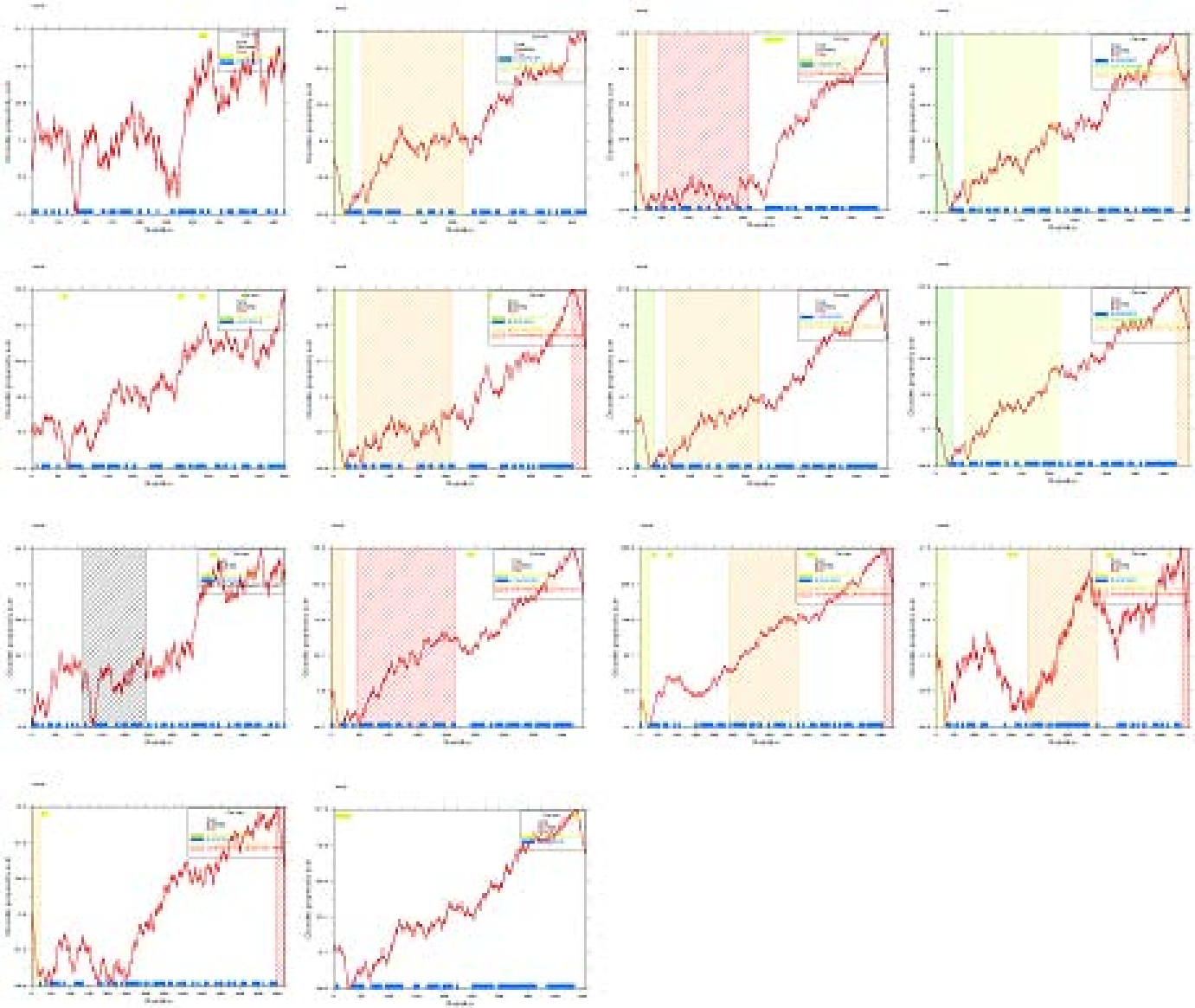

Figure 1. Exploring protein sequences for globularity and disorder in the rice *PCS* genes (line one from left to right, PCS1-5; line two from left to right, PCS6-10; and line three from left to right, PCS10-14. PCS1, 5, 13 and 14 have not been dictated *PHYTOCHELATIN* or *PHYTOCHEL_SYNTH* domain. The *PHYTOCHEL_SYNTH* domain colored yellow, pistachio or red in the broadest column, the *PHYTOCHELATIN* domain colored black.

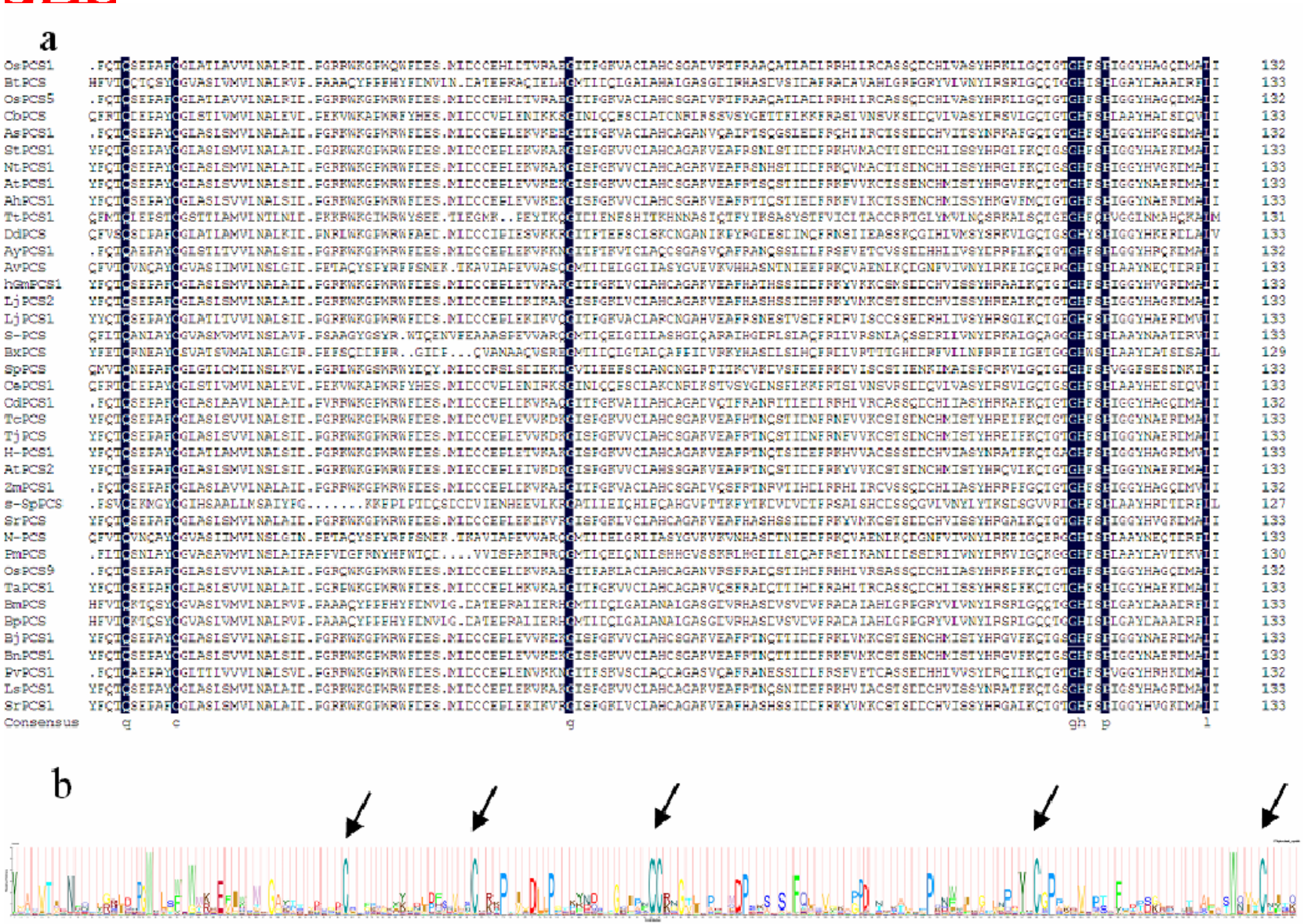

Figure 2. Functional domain of PCS proteins. The Phytochelatin domain multi-alignment by DNAMAN 6.0 (a). The Phytochel_synth domain are Cys-rich family multi-alignment by HMM(b), the Cys frequent site highlight by arrows.