

Article

Isolation and Expression of Glucosinolate Synthesis Genes *CYP83A1* and *CYP83B1* in Pak Choi (*Brassica rapa* L. ssp. *chinensis* var. *communis* (N. Tsen & S.H. Lee) Hanelt)

Biao Zhu ¹, Zhizhou Wang ¹, Jing Yang ², ZhuJun Zhu ^{1,2,*} and Huasen Wang ²

¹ Department of Horticulture, Zhejiang University, Zijingang Campus, Yuhangtang Road 866, Hangzhou 310058, China; E-Mails: billzhu@zju.edu.cn (B.Z.); twzz@163.com (Z.W.)

² Department of Horticulture, School of Agricultural and Food Science, Zhejiang A & F University, Huan Cheng Bei Lu 88, Lin'an, Hangzhou 311300, China; E-Mails: yangjing@zafu.edu.cn (J.Y.); whsych66@163.com (H.W.)

* Author to whom correspondence should be addressed; E-Mail: zhjzhu@zju.edu.cn or zhuzj@zafu.edu.cn; Tel.: +86-571-6374-3001; Fax: +86-571-6374-1276.

Received: 6 February 2012; in revised form: 22 April 2012 / Accepted: 27 April 2012 /

Published: 15 May 2012

Abstract: *CYP83A1* and *CYP83B1* are two key synthesis genes in the glucosinolate biosynthesis pathway. *CYP83A1* mainly metabolizes the aliphatic oximes to form aliphatic glucosinolate and *CYP83B1* mostly catalyzes aromatic oximes to synthesis corresponding substrates for aromatic and indolic glucosinolates. In this study, two *CYP83A1* genes named *BcCYP83A1-1* (JQ289997), *BcCYP83A1-2* (JQ289996) respectively and one *CYP83B1* (*BcCYP83B1*, HM347235) gene were cloned from the leaves of pak choi (*Brassica rapa* L. ssp. *chinensis* var. *communis* (N. Tsen & S.H. Lee) Hanelt) “Hangzhou You Dong Er” cultivar. Their ORFs were 1506, 1509 and 1500 bp in length, encoding 501, 502 and 499 amino acids, respectively. The predicted amino acid sequences of *CYP83A1-1*, *CYP83A1-2* and *CYP83B1* shared high sequence identity of 87.65, 86.48 and 95.59% to the corresponding ones in *Arabidopsis*, and 98.80, 98.61 and 98.80% to the corresponding ones in *Brassica pekinensis* (Chinese cabbage), respectively. Quantitative real-time PCR analysis indicated that both *CYP83A1* and *CYP83B1* expressed in roots, leaves and petioles of pak choi, while the transcript abundances of *CYP83A1* were higher in leaves than in petioles and roots, whereas *CYP83B1* showed higher abundances in roots. The expression levels of glucosinolate biosynthetic genes were consistent with the glucosinolate profile accumulation in shoots of seven cultivars and three organs. The isolation and

characterization of the glucosinolate synthesis genes in pak choi would promote the way for further development of agronomic traits via genetic engineering.

Keywords: pak choi; glucosinolate; phylogenetic analysis; quantitative real-time PCR

1. Introduction

Pak choi (*Brassica rapa* L. ssp. *chinensis* var. *communis* (N. Tsen & S.H. Lee) Hanelt) is a very important vegetable crop in China and Northeast Asia because of its high yield and nutritional value [1]. It contains a number of bioactive components such as folates, vitamin C, carotenoids, polyphenols and glucosinolates which have anti-carcinogenic effects [2–4]. Glucosinolates are a group of plant secondary metabolites derived from different kinds of amino acids and can be hydrolyzed by myrosinase to produce a variety of compounds which can have beneficial effects on human health including anti-carcinogenic, cholesterol-reducing and other pharmacological effects [5]. The glucosinolates-myrosinase-system is also considered to be an efficient defense system for plant against herbivorous generalists and microorganisms [6–8].

To date, more than 132 individual glucosinolates have been detected, and they are grouped into aliphatic, aromatic and indole glucosinolates depending on the structure of their side-chain [9]. In recent years, significant advances have been made in our understanding of the biosynthetic pathway of glucosinolates. The formation of glucosinolates can be divided into three separate phases including the chain elongation, formation of the core structure and secondary modifications of side chain and glucose moiety of the parent glucosinolates [10]. Since the completion of the *Arabidopsis* genome sequencing, a significant progress on glucosinolate research has been made especially in the identification of the genes and regulators for glucosinolate biosynthesis. MAM (Methylthioalkylmalate synthase) genes are responsible for the chain-elongation. CYP79 enzymes catalyze the first committed step of glucosinolates core structure formation as the conversion of amino acids to aldoximes. CYP83 enzymes catalyze the key step for the core structure formation as the conversion of aromatic oximes, and the conversion of thiohydroximic acids to glucosinolates [11–18].

Of the two oxime-metabolizing enzymes CYP83A1 and CYP83B1, major progress has been made by using *Arabidopsis* as model plant. Aliphatic oximes derived from chain-elongated homologs of methionine are efficiently metabolized by CYP83A1 enzyme, whereas CYP83B1 enzyme metabolizes these substrates with very low efficiency. Aromatic oximes derived from phenylalanine, tryptophan, and tyrosine are metabolized by both enzymes, although CYP83B1 has higher affinity for these substrates than CYP83A1 [19,20]. However, most of these researches are carried out in the model plant *Arabidopsis*. As glucosinolates have been proved to have many beneficial effects on human health and *Arabidopsis* is not edible, it is necessary to focus on glucosinolates-containing crops which are important for human consumption. In the present study, two glucosinolates biosynthesis genes *BcCYP83A1* and *BcCYP83B1* were isolated from pak choi and characterized and their expression pattern in different organs and cultivars were also characterized.

2. Results and Discussion

2.1. Identification and Sequence Analysis of CYP83A1 and CYP83B1

Two aliphatic oxime-metabolizing enzymes namely *BcCYP83A1-1*, *BcCYP83A1-2*, and one aromatic oxime-metabolizing enzyme *BcCYP83B1* were isolated from the leaves of pak choi “Hangzhou You Dong Er” cultivar. The ORF of the genes were 1506, 1509 and 1500 bp in length, encoding 501, 502 and 499 amino acids, respectively (Figures 1 and 2). A BLAST search revealed that the primary structure of *CYP83A1-1*, *CYP83A1-2* and *CYP83B1* shared high sequence identity of 86.81, 87.43 and 90.13% to the corresponding ones in *Arabidopsis* (NM117451, NM119299), respectively, and shared high sequence identity of 99.14, 99.07 and 99.47% to the corresponding ones in *Brassica pekinensis* (*Brassica rapa* L. ssp. *pekinensis* var. *communis*, Chinese cabbage, FJ376049, FJ376050, FJ376051) respectively. The amino acid sequence deduced from *CYP83A1-1* and *CYP83A1-2* showed 87.65 and 86.48% identity to that of the *Arabidopsis CYP83A1*, and 98.80 and 98.61 % identity to *Brassica pekinensis CYP83A1-1* and *CYP83A1-2*, respectively. The deduced amino acid sequence of *CYP83B1* showed 95.59% and 98.80% identity to that of the *Arabidopsis* and *Brassica pekinensis*, respectively. The sequence similarity between *CYP83A1-1* and *CYP83A1-2* were 85.88% in this study, identical to that of 85.29% in *Brassica pekinensis*.

Figure 1. Multiple alignment of deduced amino acid sequences of *CYP83A1*. Black shaded and other shaded boxes show identical and similar amino acids, respectively. GenBank accession number of the *CYP83A1*s: *Arabidopsis thaliana* (NM117451), *Brassica pekinensis* (FJ376049, FJ376050) and *Brassica chinensis* (JQ289997, JQ289996).

Chinensis 83A1-1	MEDIIHQGVVALAAVLLFFLSQKKKTKRYKLP PGPVIGLNPVILGNLHQLRFFVNPQRFF YGWAK	60
Pekinensis 83A1-1	MEDIIHQGVVALAAVLLFFLSQKKKTKRYKLP PGPVIGLNPVILGNLHQLRFFVNPQRFF YGWAK	60
Arabidopsis 83A1	MEDIIHQGVVALAAVLLFFLPQKKKTKRYKLP PGPVIGLNPVILGNLHQLRFFVNPQRFF YGWAK	60
Chinensis 83A1-2	MEDIIHQGVVALAAVLLFFRLKQRRKTKRYKLP PGPVIGLNPVILGNLHQLRFFVNPQRFF YGWAK	60
Pekinensis 83A1-2	MEDVILFGVVAAVLLFFLHLKQRRKTKRYKLP PGPVIGLNPVILGNLHQLRFFVNPQRFF YGWAK	60
Chinensis 83A1-1	KYSPITFFSYKIGSMTMVVIESAELTKELLKTKQDINFAANRPEHHRGHEFIMAYGRYDMA NHYT	120
Pekinensis 83A1-1	KYGPITFFSYKIGSMTMVVIESAELTKELLKTKQDINFAANRPEHHRGHEFIMYGRHMDMA NHYT	120
Arabidopsis 83A1	KYGPILLSYRIGSMTMVVIESAELTKELLKTKQDVNFADRPEHHRGHEFIMSYGRRDMA NHYT	120
Chinensis 83A1-2	KYGPITFFSYKIGSMTMVVIESAELTKELLKTKQDVNFHIDRPEHHRGQEFMSYGRRDMA HHYT	120
Pekinensis 83A1-2	KYGPITFFSYKIGSMTMVVIESAELTKELLKTKQDVNFHIDRPEHHRGQEFMSYGRRDMA HHYT	120
Chinensis 83A1-1	PYYREMRKMGMNHLFSPTRVATFKHVREEEARRMMCKVAKAAAEPEEDVDIISELMLTFTNS	180
Pekinensis 83A1-1	PYYREMRKMGMNHLFSPTRVATFKHVREEEARRMMCKVAKAAAEPEEDVDIISELMLTFTNS	180
Arabidopsis 83A1	PYYREIRRKMGMNHLFSPTRVATFKHVREEEARRMMDKAAKAAADVSESVDIISELMLTFTNS	180
Chinensis 83A1-2	PYYRDIRRKMGMNHLFSPTRVATFKHVREEEARRMMDKAAADAADA'SINSVDIISELMLTFTNS	180
Pekinensis 83A1-2	PYYRDIRRKMGMNHLFSPTRVATFKHVREEEARRMMDKAAADAADA'SINSVDIISELMLTFTNS	180
Chinensis 83A1-1	VVCRQAFGKKYNE DGEEMKRFUUKICLYGTQSVLGGKFFFSDFFPFWGFLLDLWGLTAYMKE	239
Pekinensis 83A1-1	VVCRQAFGKKYNE DGEEMKRFUUKICLYGTQSVLGGKFFFSDFFPFWGFLLDLWGLTAYMKE	239
Arabidopsis 83A1	VVCRQAFGKKYNE DGEEMKRFUUKICLYGTQSVLGGKFFFSDFFPYLGFLLDLWGLTAYMKE	239
Chinensis 83A1-2	VVCRQAFGKKYNE DGEEMKRFUUKICLYGTQSVVGGKFFFSDFFPFDGYRLDDLTKLTAYMKE	240
Pekinensis 83A1-2	VVCRQAFGKKYNE DGEEMKRFUUKICLYGTQSVVGGKFFFSDFFPFDGYRLDDLTKLTAYMKE	240
Chinensis 83A1-1	CFERQDITYQEDITDETLDPNREKPEITESMDLLMEIYKDPFASVFTVIENVKGVILNI VV	299
Pekinensis 83A1-1	CFERQDITYQEDITDETLDPNREKPEITESMDLLMEIYKDPFASVFTVIENVKGVILNI VV	299
Arabidopsis 83A1	CFERQDITYQEVVNETLDPNRVKPEITESMDLLMGDIYKEQPFASEFTAENVKAVLDI VV	299
Chinensis 83A1-2	CFERQDITYLQEIIVDETLDPNRVKPEITDSMDLLMEIYRDQPFASEFTAENVKAVVLDI VV	300
Pekinensis 83A1-2	CFERQDITYLQEIIVDETLDPNRVKPEITDSMDLLMEIYRDQPFASEFTAENVKAVVLDI VV	300

Figure 1. Cont.

Chinensis 83A1-1	ACTDTAAAAVVWGMTYLMKYPQVMKKAQAEVREYAREKKGSTFIETEDDVKNLPYFRALVKE	359
Pekinensis 83A1-1	ACTDTAAAAVVWGMTYLMKYPQVMKKAQAEVREYAREKKGSTFIETEDDVKNLPYFRALVKE	359
Arabidopsis 83A1	AGTDTAAAAVVWGMTYLMKYPQVLKKAQAEVREYMKKKGSTFVETEDDVKNLPYFRALVKE	359
Chinensis 83A1-2	ACTDTAAAAVVWGMTYLMKYPHVMKKAQAEVREYMRERCITFVETEDDVKNLPYFRALVKE	360
Pekinensis 83A1-2	ACTDTAAAAVVWGMTYLMKYPHVMKKAQAEVREYMRERCITFVETEDDVKNLPYFRALVKE	360
Chinensis 83A1-1	TLRBEPVLP LLIPRACIIQDTKDAGYDIIPAGTTVNVNAWAVSRDEKEWGPINPDEFKPERFL	419
Pekinensis 83A1-1	TLRBEPVLP LLIPRACIIQDTKDAGYDIIPAGTTVNVNAWAVSRDEKEWGPINPDEFKPERFL	419
Arabidopsis 83A1	TLRBEPVLP LLIPRACIIQDTKDAGYDIIPAGTTVNVNAWAVSRDEKEWGPINPDEFKPERFL	419
Chinensis 83A1-2	TLRBEPVLP LLVPRACIQNTKIIAGYDIIPAGTTVNVNAWAVSRDEKEWGPINPDEFKPERFL	420
Pekinensis 83A1-2	TLRBEPVLP LLVPRACIQNTKIIAGYDIIPAGTTVNVNAWAVSRDEKEWGPINPDEFKPERFL	420
Chinensis 83A1-1	EKDVDFKGT DYEFIFPGSGRRMCPGMRLGAAGLENPANLLKIPNFKLPNGIKPEEI NMD	479
Pekinensis 83A1-1	EKDVDFKGT DYEFIFPGSGRRMCPGMRLGAAGLENPANLLKIPNFKLPNGIKPEEI NMD	479
Arabidopsis 83A1	EKDVDFKGT DYEFIFPGSGRRMCPGMRLGAAGLENPANLLKIPNFKLPNGMKPDDI NMD	479
Chinensis 83A1-2	EKDVDFKGT DYEFIFPGSGRRMCPGMRLGLAMLEVPYANLLFKIFDFKLPNGMKPDEI NMD	480
Pekinensis 83A1-2	EKDVDFKGT DYEFIFPGSGRRMCPGMRLGLAMLEVPYANLLFKIFDFKLPNGMKPDEI NMD	480
Chinensis 83A1-1	VMTGLAMHKSIDHLKL VPEKVS K	501
Pekinensis 83A1-1	VMTGLAMHKSIDHLKL VPEKVS K	501
Arabidopsis 83A1	VMTGLAMHKSQHLKL VPEKVKK	501
Chinensis 83A1-2	VMTGLAMHKSIDHLKL VPEKVS	502
Pekinensis 83A1-2	VMTGLAMHKAHMLML VPEKVV	502

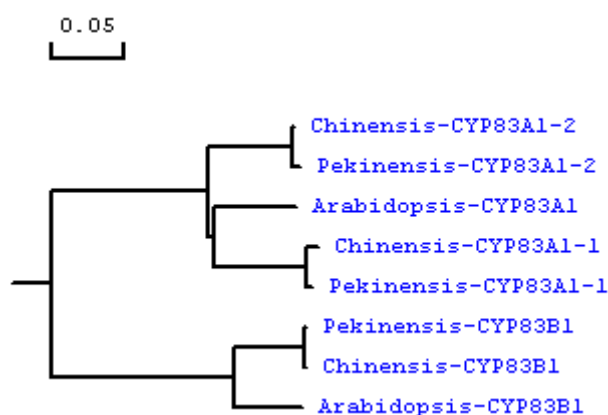
Figure 2. Multiple alignment of deduced amino acid sequences of *CYP83B1*. Black shaded and other shaded boxes show identical and similar amino acids, respectively. GenBank accession number of the *CYP83B1*s: *Arabidopsis thaliana* (NM119299), *Brassica pekinensis* (FJ376051) and *Brassica chinensis* (HM347235).

Chinensis 83B1	MDLFLILIAAMVAVA AAFLLRSSTKKSRLP PGP KGLPIQNGNLHQMEKFN PQHFLFRLSKL	60
Pekinensis 83B1	MDLFLILIAAMVAVA AAFLLRSSTKKSRLP PGP KGLPIQNGNLHQMEKFN PQHFLFRLSKL	60
Arabidopsis 83B1	MDLFLILIAAGLVAA AAFLLRSSTKKSRLP PGP KGLPIQNGNLHQMEKFN PQHFLFRLSKL	60
Chinensis 83B1	YGPFFFTMKIGGRRLAVI LSSAELAKE LLKTQDLNFTARP LLKGQQTMSYQGRELGFGQYTA	120
Pekinensis 83B1	YGPFFFTMKIGGRRLAVI LSSAELAKE LLKTQDLNFTARP LLKGQQTMSYQGRELGFGQYTA	120
Arabidopsis 83B1	YGPFFFTMKIGGRRLAVI LSSAELAKE LLKTQDLNFTARP LLKGQQTMSYQGRELGFGQYTA	120
Chinensis 83B1	YYREMRKMMVNLFSPNRVASFRPVREEECQRMMDKIYKAADQSGTVDLSE LLSFTNCV	180
Pekinensis 83B1	YYREMRKMMVNLFSPNRVASFRPVREEECQRMMDKIYKAADQSGTVDLSE LLSFTNCV	180
Arabidopsis 83B1	YYREMRKMMVNLFSPNRVASFRPVREEECQRMMDKIYKAADQSGTVDLSE LLSFTNCV	180
Chinensis 83B1	VCRQAFGKRYNEYGT EMKRFIINELYETQALLGTLFFSDLEFPYFGFLDNL TGLNARLKRAF	240
Pekinensis 83B1	VCRQAFGKRYNEYGT EMKRFIINELYETQALLGTLFFSDLEFPYFGFLDNL TGLNARLKRAF	240
Arabidopsis 83B1	VCRQAFGKRYNEYGT EMKRFIINELYETQALLGTLFFSDLEFPYFGFLDNL TGLNARLKRAF	240
Chinensis 83B1	KELDTYLQELLDETLDP SRPKPETE SEINDL LMQIDYKDQPF SIKFTHE NVKAMGLDI VVPG	300
Pekinensis 83B1	KELDTYLQELLDETLDP SRPKPETE SEINDL LMQIDYKDQPF SIKFTHE NVKAMGLDI VVPG	300
Arabidopsis 83B1	KELDTYLQELLDETLDP RRPKPETE SEINDL LMQIDYKDQPF SIKFTHE NVKAMGLDI VVPG	300
Chinensis 83B1	TDTAAAVVVWAMTYL IIKYPEAMKKAQDEVRNVVGDKGYSSEEDIPNLPYLKAVIKESLRL	360
Pekinensis 83B1	TDTAAAVVVWAMTYL IIKYPEAMKKAQDEVRNVVGDKGYSSEEDIPNLPYLKAVIKESLRL	360
Arabidopsis 83B1	TDTAAAVVVWAMTYL IIKYPEAMKKAQDEVRNVVGDKGYSSEEDIPNLPYLKAVIKESLRL	360
Chinensis 83B1	EPVVPPIILLHRETIADAKI GGYDIPAKTINQVNAWAVSRDTAAAGDNPNEFIPERFMNEQK	420
Pekinensis 83B1	EPVVPPIILLHRETIADAKI GGYDIPAKTINQVNAWAVSRDTAAAGDNPNEFIPERFMNEQK	420
Arabidopsis 83B1	EPVVPPIILLHRETIADAKI GGYDIPAKTINQVNAWAVSRDTAAAGDNPNEFIPERFMNEHK	420
Chinensis 83B1	GVDFKQDFELLPFGSGRRMCPAMHLGVAMVEIPFANLLYRFDWSLPKGIKKPEDIKMDVM	480
Pekinensis 83B1	GVDFKQDFELLPFGSGRRMCPAMHLGVAMVEIPFANLLYRFDWSLPKGIKKPEDIKMDVM	480
Arabidopsis 83B1	GVDFKQDFELLPFGSGRRMCPAMHLGVAMVEIPFANLLYRFDWSLPKGIKKPEDIKMDVM	480
Chinensis 83B1	TGLAMHKKDHLVLAPRTHI	499
Pekinensis 83B1	TGLAMHKRDHLVLAPRTHI	499
Arabidopsis 83B1	TGLAMHKKEHLVLAPRTHI	499

2.2. Phylogenetic Analysis of CYP83A1 and CYP83B1

We compared the isolated sequences with the corresponding ones in *Arabidopsis thaliana* (NM117451, NM119299) and *Brassica pekinensis* (FJ376049, FJ376050, FJ376051) by multiple alignments, and reconstructed the phylogenetic tree by ClustalW analysis program. The results showed that the homologs of *CYP83A1* and *CYP83B1* from the *Brassicaceae* family are highly conserved, with low variation in DNA sequence. As shown in the phylogenetic tree, the glucosinolate synthesis genes of *Brassica chinensis* and *Brassica pekinensis* clustered first with each other and then with the genes of *Arabidopsis* (Figure 3), which demonstrated the highly congruent relationship among different sources of glucosinolate synthesis genes.

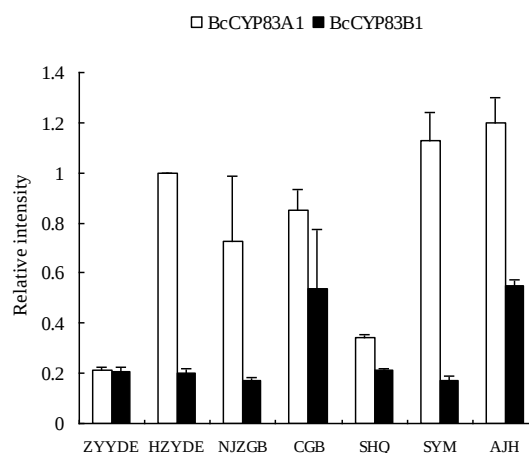
Figure 3. Phylogenetic tree of the cDNA sequences of the plant glycosyltransferase superfamily. GenBank accession numbers are: *Arabidopsis thaliana* CYP83A1 (NM117451), *Arabidopsis thaliana* CYP83B1 (NM119299), *Brassica pekinensis* CYP83A1-1 and CYP83A1-2 (FJ376049, FJ376050), *Brassica pekinensis* CYP83B1 (FJ376051), *Brassica chinensis* CYP83A1-1 and CYP83A1-2 (JQ289997, JQ289996), *Brassica chinensis* CYP83B1 (HM347235). The horizontal scale shows the number of differences per 100 residues derived from the ClustalW alignment.



2.3. Expression of CYP83A1 and CYP83B1 in Different Cultivars and Various Organs

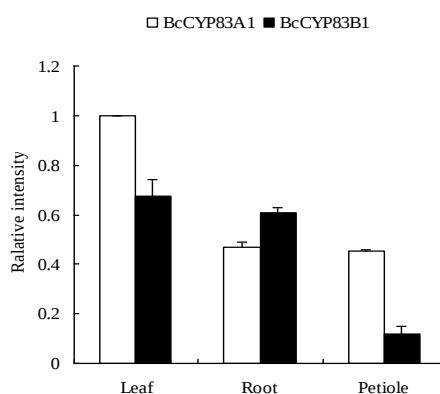
Quantitative real-time PCR (qRT-PCR) was employed to confirm the expression patterns of the two genes in seven cultivars, as well as in different organs. Actin was used as an internal reference control for total RNA input. In our analysis, the gene-specific primers used for amplification of *CYP83A1* recognized and amplified both *CYP83A1-1* and *CYP83A1-2* genes. Both *CYP83A1* and *CYP83B1* were detected in the leaves of all the seven cultivars and had the same expression pattern except for cultivar “ZYYDE”. The expression patterns of the two genes in leaves of seven cultivars are illustrated in Figure 4. *CYP83A1* was presented in higher abundance compared with *CYP83B1* in most of the cultivars, whereas the *CYP83A1* and *CYP83B1* transcripts in “ZYYDE” cultivar were nearly same.

Figure 4. Relative expression analysis of *CYP83A1* and *CYP83B1* in leaves of seven cultivars of pak choi. Error bars represent the standard deviation (SD) from three biological repeats.



Both the aliphatic oxime-metabolizing enzyme and aromatic oxime-metabolizing enzyme were identified in the three organs (roots, leaves and petioles) of pak choi (cv. HZYDE). Though *CYP83A1* and *CYP83B1* exist in all those organs, they displayed spatial expression patterns (Figure 5). The expression level of *CYP83A1* was higher in leaves than in roots or petioles, whereas the *CYP83B1* showed higher mRNA abundance in roots compared with another two.

Figure 5. Relative expression analysis of *CYP83A1* and *CYP83B1* in leaf, petiole and root of pak choi cv. HZYDE. Error bars represent the standard deviation (SD) from three biological repeats.



2.4. Glucosinolate Profile Concentrations in Different Cultivars and Various Organs

The concentrations of aliphatic, indole and aromatic glucosinolates in different pak choi cultivars and in different organs of “Hangzhou You Dong Er” cultivar were analyzed. As shown in Table 1, the aliphatic glucosinolate concentrations in leaves of most of the cultivars were higher than indole and aromatic glucosinolates. The glucosinolate profiles in different organs varied differently. As leaf and petiole had particularly higher aliphatic glucosinolates, however, root had much higher indole and aromatic glucosinolates.

Table 1. Aliphatic, indole and aromatic glucosinolates content (μmol) in shoots of different cultivars and in different organs of “Hangzhou You Dong Er” cultivar. Data are the means of three biological repeats $\pm$ standard deviation.

	ZYYDE	HZYDE	NJZGB	CGB	SHQ	SYM	AJH	Leaf	Root	Petiole
Aliphatic GS	0.43 $\pm$ 0.07	0.94 $\pm$ 0.07	0.49 $\pm$ 0.07	0.54 $\pm$ 0.08	0.98 $\pm$ 0.04	2.90 $\pm$ 0.34	0.87 $\pm$ 0.10	1.89 $\pm$ 0.29	0.12 $\pm$ 0.02	0.48 $\pm$ 0.05
Indole GS	0.68 $\pm$ 0.03	0.76 $\pm$ 0.03	0.30 $\pm$ 0.01	0.45 $\pm$ 0.02	0.52 $\pm$ 0.02	0.59 $\pm$ 0.03	0.75 $\pm$ 0.10	0.25 $\pm$ 0.06	0.94 $\pm$ 0.10	0.24 $\pm$ 0.04
Aromatic GS	0.38 $\pm$ 0.02	0.51 $\pm$ 0.02	0.31 $\pm$ 0.03	0.59 $\pm$ 0.07	0.63 $\pm$ 0.04	0.46 $\pm$ 0.07	0.61 $\pm$ 0.05	0.28 $\pm$ 0.06	0.78 $\pm$ 0.07	0.25 $\pm$ 0.03
Total GS	1.49 $\pm$ 0.07	2.21 $\pm$ 0.09	1.10 $\pm$ 0.09	1.58 $\pm$ 0.11	2.13 $\pm$ 0.05	3.95 $\pm$ 0.41	2.23 $\pm$ 0.04	2.43 $\pm$ 0.24	1.84 $\pm$ 0.17	0.97 $\pm$ 0.05

2.5. Discussion

In this study, we isolated two *CYP83A1* genes and one *CYP83B1* gene from pak choi “Hangzhou You Dong Er” cultivar. Amino acid sequence analysis and sequence alignment with their homologs from *Brassica pekinensis* and *Arabidopsis* showed the two key glucosinolate synthetic genes *CYP83A1* and *CYP83B1* within *Brassicaceae* family are highly conserved. As shown in Figures 1 and 2, the genes isolated from pak choi have both high DNA and amino acid sequence conservation with their *Brassica pekinensis* counterparts. The identity of DNA sequence and deduced amino acid sequence of *CYP83B1* between the two species was 99.47% and 98.80%. The two copies of *CYP83A1* namely *BcCYP83A1-1* and *BcCYP83A1-2* from pak choi were also highly similar to the corresponding ones of *Brassica pekinensis*, whereas there was only one copy of *CYP83A1* in *Arabidopsis*. This is consistent with the previous finding by Zang *et al.* that identities within *Brassica rapa* were usually higher than those with *Arabidopsis* [21]. They isolated and analysed the glucosinolate synthesis genes in *Brassica pekinensis* by genome-wide identification [21]. The result was also accordant to the notion that *Brassica* species evolved and were triplicated after divergence from *Arabidopsis* lineage [22–24]. The observed glucosinolate synthesis sequence similarities among the *Brassicaceae* (mostly >90%) were higher than those between *Brassica* and *Arabidopsis* (mostly 80–90%) in this study also supported the presumptive evolution order.

Consistent with the results of previous studies in *Arabidopsis*, both the *CYP83B1* and *CYP83A1* transcripts were expressed in all tested organs (roots, leaves and petioles) of pak choi [25,26]. Mizutani reported that *CYP83A1* expression levels in *Arabidopsis* was highest in leaves and significantly higher in roots and stems than in flowers and siliques, while *CYP83B1* was the most highly expressed in roots [25]. In our study, the expression level of *CYP83A1* was also highest in leaves, and the *CYP83B1* expressed higher in leaves and roots than in petioles. These results were consistent with the previous finding that *CYP83A1* expressed highest in the leaves and *CYP83B1* were preferentially expressed in the roots [25,26]. The gene expression pattern in the organs was consistent with the glucosinolate profile accumulations as aliphatic glucosinolates were higher in leaves and petioles while indole and aromatic glucosinolates were higher in roots. The results were similar to Brown’s finding that aliphatic glucosinolates are typically higher in the leaves compared to the roots while indole glucosinolates are abundant in the roots [27]. When compared within an appointed organ, we found that the expression patterns of the two genes were different. *CYP83A1* transcripts were almost two-folds of *CYP83B1* transcripts in leaves and petioles, in contrast, the expression levels of the *CYP83A1* were relatively lower than *CYP83B1* in roots. Our finding was also accordant to the conclusion that the *CYP83A1*

catalyzed aliphatic glucosinolates biosynthesis and accumulation in leaves and the CYP83B1 catalyzed aromatic and indole glucosinolates mostly found in roots [15].

The expression patterns of the two genes in leaves of seven cultivars in our study were similar except “ZYYDE” cultivar and the gene expression levels were consistent with the glucosinolate profile accumulation. The expression levels of *CYP83A1* in leaves of six of seven cultivars were higher than *CYP83B1* and the aliphatic glucosinolates of the six *CYP83A1* higher expression cultivars were higher than another two glucosinolate profiles, whereas *CYP83A1* and *CYP83B1* expressed nearly identical in “ZYYDE” cultivar. While focused on the specific gene expression and glucosinolate accumulation in individual cultivar, “SYM” cultivar had high *CYP83A1/CYP83B1* transcript level ratio and high aliphatic glucosinolates/indole glucosinolates ratio. And “CGB” cultivar had the higher *CYP83B1* transcript level and relative higher aromatic glucosinolates accumulation.

3. Experimental Section

3.1. Plant Materials

The seeds of seven pak choi cultivars (*Brassica rapa* L. ssp. *chinensis* var. *communis* (N. Tsen & S.H. Lee) Hanelt) with different characters were sown in plastic pot with vermiculite in a growth chamber with 26/22 °C (12 h light/12 h dark) for 2 weeks. The seven commonly plant and consumed pak choi cultivars in southern China used are “Hangzhou You Dong Er” (HZYDE), “Zhou Ye You Dong Er” (ZYYDE), “Si Yue Man” (SYM), “Shanghai Qing” (SHQ), “Nanjing Zhong Gan Bai” (NJZGB), “Chang Geng Bai” (CGB) and “Ai Jiao Huang” (AJH). The young leaves from two-week old plants of seven pak choi cultivars were collected as experimental materials for quantitative real-time PCR (qRT-PCR) analysis. And three organs including roots, petioles and leaves of “Hangzhou You Dong Er” cultivar were collected for the expression analysis.

3.2. Isolation of Total RNA and Synthesis of Full-Length cDNA Sequence

Total RNA was extracted from samples using TRIzol reagent RNAiso™ Plus (Takara, D9108A) following the manufacturer’s protocols. Approximately 0.1 g of tissue was extracted in 1 mL of extraction buffer. After melted in RNA-free water, RNA was quantified by UV spectroscopy and its integrity was visually assessed on ethidium bromide stained agarose gels. The cDNA was synthesized from 500 ng of DNA-free RNA with a PrimeScript RT reagent Kit (Takara, DRR037A) following the manufacturer’s protocol.

Degenerate primers allowing amplification were designed based on sequences corresponding to highly conserved peptide regions of *CYP83A1* and *CYP83B1* of *Arabidopsis* and *Brassica pekinensis*. The cDNAs were amplified by PCR using the degenerate primers, then cloned into the pMD18-T Easy vector (Takara, Code D101A) and sequenced (Invitrogen). Rapid amplification of cDNA ends (RACE) was performed to obtain the 3’ and 5’ ends of the genes with the SMART RACE cDNA Amplification Kit (Clontech) according to the manufacturer’s instructions. Sequence analysis of open reading fragment (ORF) and encoded amino acid sequence of genes were deduced by BioXM 2.6 (Might, Guangzhou, China, 2007). Physicochemical properties of the deduced protein were predicted by Protparam [28].

3.3. Phylogenetic Analysis

The nucleotide and deduced amino acid sequences were used for multiple alignments and phylogenetic tree analysis. The tree was obtained by ClustalW analysis program [29].

3.4. Quantitative Real-Time RT-PCR

The oligonucleotide primers for quantitative real-time PCR (qRT-PCR) analysis were designed with primer3 (version 0.4.0) based on the gene sequences (JQ289997, JQ289996, HM347235 and EX087730) [30]. For *CYP83A1*, we designed a specific primer which can amplify both *CYP83A1-1* and *CYP83A1-2* for qRT-PCR according to the highly conserved peptide regions compared with *Brassica chinensis*, *Brassica pekinensis* and *Arabidopsis*. The qRT-PCR primers for *CYP83B1* and actin gene were designed the same way. The size of all qPCR products was limited to 100–300 nucleotides. PCR product size was analyzed on 1% agarose gels stained with ethidium bromide and the specificity of qRT-PCR products was then confirmed by sequencing. The qPCR primers of *CYP83A1*, *CYP83B1* and Actin were as follows respectively:

CYP83A1

(F) 5'-TCCTCTCCTTATCCCTCGTGCTTGC-3'

(R) 5'-CCAAGACGCATTCCAGGGCACATTC-3'

CYP83B1

(F) 5'-AGACTCTTGACCCTAGCCGTCCTA-3'

(R) 5'-CCTTTGTCACCGACCACATTCCT-3'

Actin

(F) 5'-CGCTTAACCCGAAAGCTAAC-3'

(R) 5'-TACGCCCCTAGCGTAAAGA-3'

For gene expression analysis, first-stranded cDNAs from the leaves of the seven cultivars, and first-stranded cDNAs from the leaves, the petioles and the roots of “HZYDE” cultivar were used as templates. RT-PCR was performed on Applied Biosystems 7300 Real-Time PCR system with the SYBR Green qPCR kit (Takara, DRR081A). Ten-fold diluted cDNA was used as the template in real-time PCR. 25 µL total reaction volume comprised 12.5 µL 2× SYBR Green PCR Master Mix (Takara), 1 µL of each primer (10 µM), 2 µL of cDNA, and 8.5 µL of PCR-grade water. The PCR conditions were as the follows: initiated by 30 s at 95 °C, then followed by 40 cycles of 95 °C for 5 s, 60 °C for 31 s, and completed with a melting curve analysis program. No-template controls and melting curve analyses were included in every reaction. *Brassica actin* (EX087730) was used as the housekeeping gene. A melting curve of denatured double-stranded cDNA was established to test the purity of the products, and the products were re-sequenced to confirm the specificity. Triplicate quantitative PCR experiments were performed for each sample, and the expression values obtained were normalized against actin. Analysis of the relative gene expression data was done using the $2^{-\Delta\Delta Ct}$ method [31,32].

3.5. Glucosinolate Analysis

The glucosinolate extraction and analysis procedure was performed according to the method of Krumbein with slight modification [33]. Briefly, 0.25 g sample powder were boiled with 10 mL 70% methanol to inactive myrosinase, then the supernatants were loaded onto a 1 mL mini-column (JT Baker, Phillipsburg, NJ, USA) containing 500 μ L activated DEAE SephadexTM A25 (Amersham Biosciences, Uppsala, Sweden) to de-sulphate overnight with aryl sulfatase (Sigma-Aldrich Co., St. Louis, MO, USA). The resultant desulpho (ds)-glucosinolates were eluted with ultra pure water and stored at -20°C prior to separation by high performance liquid chromatography (HPLC).

For HPLC, 20 μ L samples were analyzed in an Agilent 1200 HPLC system (G1311A quaternary pump, G1322A vacuum solvent delivery degasser, G1316A thermostatted column compartment, Agilent Technologies, Inc., Palo Alto, CA, USA) consisting of a G1329A auto injector, a prontosil ODS2 column ($250 \times 4 \mu\text{m}$, $5 \mu\text{m}$, Bischoff, Leonberg, Germany) and a G1315B diode array detector (DAD) set at 229 nm. The mobile phase was ultra pure water (A) and acetonitrile (Tedia, Fairfield, OH, USA) (B) in a linear gradient from 0% to 20% B in 32 min, then constant 20% B for 6 min, following by 100% B and 0% B prior to the injection of the next sample. The flow rate was 1.3 mL min^{-1} .

4. Conclusions

In the present study, we isolated and analyzed the glucosinolate biosynthetic genes in pak choi. This is the first report about glucosinolate genes in pak choi although they have been studied in *Arabidopsis*. The gene expression levels and glucosinolate profile accumulation were also compared. The isolation and further characterization of the genes responsible for the synthesis and regulation of glucosinolates in pak choi are important for the development of vegetables that are beneficial to health.

Acknowledgments

This research was financially supported by NSFC (30871718), ZJNSF (R3080360, Y3110340) and the fund for Zhejiang Higher School Innovative Research Team (T200916).

References

1. Tay, D.C.S.; Toxopeus, H. *Brassica rapa* L. cv. Group Pak Choi. In *Plant Resources of South-East Asia, 8, Vegetables*; Siemonsma, J.S., Piluek, K., Eds.; Pudoc Scientific Publishers: Wageningen, The Netherlands, 1993; pp. 130–134.
2. He, H.J.; Fingerling, G.; Schnitzler, W.H. Seasonal variations of the glucosinolate amounts in *Brassica campestris* L. spp. *chinensis*. *J. Appl. Bot.* **2000**, *74*, 198–202.
3. Lewis, L.; Fenwick, G.R. Glucosinolate content of brassica vegetables: Analysis of twenty-four cultivars of calabrese (green sprouting broccoli, *Brassica oleracea* L. var. *botrytis* subvar. *cymosa* Lam.). *Food Chem.* **1987**, *25*, 259–268.
4. Kurilich, A.C.; Tsau, G.J.; Brown, A.; Howard, L.; Klein, B.P.; Jeffery, E.H.; Kushad, M.; Wallig, M.A.; Juvik, J.A. Carotene, tocopherol and ascorbate contents in subspecies of *Brassica oleracea*. *J. Agric. Food Chem.* **1999**, *47*, 1576–1581.

5. Mithen, R.F.; Dekker, M.; Verkerk, R.; Rabot, S.; Johnson, I.T. The nutritional significance, biosynthesis and bioavailability of glucosinolates in human foods. *J. Sci. Food Agric.* **2000**, *80*, 967–984.
6. Manici, L.M.; Lazzeri, L.; Palmieri, S. *In vitro* fungitoxic activity of some glucosinolates and their enzyme-derived products towards plant pathogenic fungi. *J. Agric. Food Chem.* **1997**, *45*, 2768–2773.
7. Wittstock, U.; Gershenzon, J. Constitutive plant toxins and their role in defense against herbivores and pathogens. *Curr. Opin. Plant Biol.* **2002**, *5*, 300–307.
8. Wittstock, U.; Kliebenstein, D.J.; Lambrix, V.; Reichelt, M.; Gershenzon, J. Glucosinolate Hydrolysis and Its Impact on Generalist and Specialist Insect Herbivores. In *Integrative Phytochemistry: From Ethnobotany to Molecular Ecology*; Romeo, J.T., Ed.; Pergamon: Amsterdam, The Netherlands, 2003; pp. 101–125.
9. Agerbirk, N.; Olsen, C.E. Glucosinolate structures in evolution. *Phytochemistry* **2012**, *77*, 16–45.
10. Wittstock, U.; Halkier, B.A. Glucosinolate research in the *Arabidopsis* era. *Trends Plant Sci.* **2002**, *7*, 263–270.
11. Bak, S.; Feyereisen, R. The involvement of two P450 enzymes, CYP83B1 and CYP83A1, in auxin homeostasis and glucosinolate biosynthesis. *Plant Physiol.* **2001**, *127*, 108–118.
12. Barlier, I.; Kowalczyk, M.; Marchant, A.; Ljung, K.; Bhalarao, R.; Bennett, M.; Sandberg, G.; Bellini, C. The *SUR2* gene of *Arabidopsis thaliana* encodes the cytochrome P450 CYP83B1, a modulator of auxin homeostasis. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 14819–14824.
13. Delarue, M.; Prinsen, E.; van Onckelen, H.; Caboche, M.; Bellini, C. *Sur2* mutations of *Arabidopsis thaliana* define a new locus involved in the control of auxin homeostasis. *Plant J.* **1998**, *14*, 603–611.
14. Grubb, C.D.; Abel, S. Glucosinolate metabolism and its control. *Trends Plant Sci.* **2006**, *11*, 89–100.
15. Halkier, B.A.; Gershenzon, J. Biology and biochemistry of glucosinolates. *Ann. Rev. Plant Biol.* **2006**, *57*, 303–333.
16. Hansen, C.H.; Du, L.C.; Naur, P.; Olsen, C.E.; Axelsen, K.B.; Hick, A.J.; Pickett, J.A.; Halkier, B.A. CYP83B1 is the oxime-metabolizing enzyme in the glucosinolate pathway in *Arabidopsis*. *J. Biol. Chem.* **2001**, *276*, 24790–24796.
17. Piotrowski, M.; Schemenewitz, A.; Lopukhinat, A.; Mueller, A.; Janowitz, T.; Weiler, E.W.; Oecking, C. Desulfoglucosinolate sulfotransferases from *Arabidopsis thaliana* catalyze the final step in the biosynthesis of the glucosinolate core structure. *J. Biol. Chem.* **2004**, *279*, 50717–50725.
18. Smolen, G.; Bender, J. *Arabidopsis* cytochrome p450 *cyp83B1* mutations activate the tryptophan biosynthetic pathway. *Genetics* **2002**, *160*, 323–332.
19. Bak, S.; Tax, F.E.; Feldmann, K.A.; Galbraith, D.W.; Feyereisen, R. CYP83B1, a cytochrome P450 at the metabolic branch point in auxin and indole glucosinolate biosynthesis in *Arabidopsis*. *Plant Cell* **2001**, *13*, 101–111.
20. Naur, P.; Petersen, B.L.; Mikkelsen, M.D.; Bak, S.; Rasmussen, H.; Olsen, C.E.; Halkier, B.A. CYP83A1 and CYP83B1, two nonredundant cytochrome P450 enzymes metabolizing oximes in the biosynthesis of glucosinolates in *Arabidopsis*. *Plant Physiol.* **2003**, *133*, 67–72.

21. Zang, Y.X.; Kim, H.U.; Kim, J.A.; Lim, M.H.; Jin, M.; Lee, S.C.; Kwon, S.J.; Lee, S.I.; Hong, J.K.; Park, T.H.; *et al.* Genome-wide identification of glucosinolate synthesis genes in *Brassica rapa*. *FEBS J.* **2009**, *276*, 3559–3574.
22. Bowers, J.E.; Chapman, B.A.; Rong, J.K.; Paterson, A.H. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* **2003**, *422*, 433–438.
23. Lysak, M.A.; Koch, M.; Pecinka, A.; Schubert, I. Chromosome triplication found across the tribe Brassiceae. *Genome Res.* **2005**, *15*, 516–525.
24. Yang, Y.W.; Lai, K.N.; Tai, P.Y.; Li, W.H. Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. *J. Mol. Evol.* **1999**, *48*, 597–604.
25. Mizutani, M.; Ward, E.; Ohta, E. Cytochrome p450 superfamily in *Arabidopsis thaliana*: Isolation of cDNAs, differential expression, and RFLP mapping of multiple cytochromes P450. *Plant Mol. Biol.* **1998**, *37*, 39–52.
26. Xu, W.; Bak, S.; Decker, A.; Paquette, S.M.; Feyereisen, R.; Galbraith, D.W. Microarray-based analysis of gene expression in very large gene families: The cytochrome P450 gene superfamily of *Arabidopsis thaliana*. *Gene* **2001**, *272*, 61–74.
27. Brown, P.D.; Tokuhiya, J.G.; Reichelt, M.; Gershenzon, J. Variation of glucosinolate accumulation among different organs and developmental stages of *Arabidopsis thaliana*. *Phytochemistry* **2003**, *62*, 471–481.
28. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*; Walker, J.M., Ed.; Humana Press Inc.: Totowa, NJ, USA, 2005; pp. 571–607.
29. Thompson, J.D.; Higgins, D.G.; Gibson, T.J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **1994**, *22*, 4673–4680.
30. Rozen, S.; Skaletsky, H.J. Primer3 on the WWW for General Users and for Biologist Programmers. In *Bioinformatics Methods and Protocols: Methods in Molecular Biology*; Krawetz, S., Misener, S., Eds.; Humana Press Inc.: Totowa, NJ, USA, 2000; pp. 365–386.
31. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_t}$ method. *Methods* **2001**, *25*, 402–408.
32. Yin, X.R.; Chen, K.S.; Allan, A.C.; Wu, R.M.; Zhang, B.; Lallu, N.; Ferguson, I.B. Ethylene-induced modulation of genes associated with the ethylene signaling pathway in ripening kiwifruit. *J. Exp. Bot.* **2008**, *59*, 2097–2108.
33. Krumbein, A.; Schonhof, I.; Schreiner, M. Composition and contents of phytochemicals (glucosinolates, carotenoids and chlorophylls) and ascorbic acid in selected *Brassica* species (*B. juncea*, *B. rapa* subsp. *nipposinica* var. *chinoleifera*, *B. rapa* subsp. *chinensis* and *B. rapa* subsp. *rapa*). *J. Appl. Bot. Food Qual.* **2005**, *79*, 168–174.