

Sequence Analysis of the Gene CYP79B3 in Cabbage (*Brassica Oleracea* Var. *Capitata*)

Bo Sun, Min Jiang, Qiao Yuan, Fen Zhang and Haoru Tang ^{a, *}

College of Horticulture, Sichuan Agricultural University, Chengdu 611130, China.

^{*, a} htang@sicau.edu.cn.

Keywords: Cabbage, glucosinolate, CYP79B3, sequence analysis.

Abstract. CYP79B3 is an important cytochrome P450 monooxygenases in glucosinolate biosynthesis. Here, the *Brassica oleracea* var. *capitata* CYP79B3 (BocCYP79B3) gene sequence was obtained from Brassica database (BRAD), and performed for sequence analysis. The BocCYP79B3 gene mapped to Scaffold000053, and contains an open reading frame of 1,632 bp that encodes a 543-amino acid protein with a calculated molecular mass of 61.54 kD and an isoelectric point (pI) of 8.59. Subcellular localization predicted the BocCYP79B3 gene was in the chloroplast. The conserved domain of the BocCYP79B3 protein is PLN02971. The CYP79B3 protein is most closely related to *B. napus*. The findings of the present study provide a molecular basis for the elucidation of CYP79B3 gene function in cabbage.

1. Introduction

Cabbage (*Brassica oleracea* var. *capitata*) is a member of the Brassicaceae family that is widely distributed in the world. In China, cabbage is an important vegetable crop, and consumed considerable every year. Cabbage is generally grown for its leafy head as common edible part, which are crispy, tender, and tasty [1]. Besides its good flavor, cabbage is also a rich source of nutrients, antioxidants, and anticarcinogenic compounds, including carbohydrates, vitamin C, carotenoids, and glucosinolates [1-2].

Glucosinolates are a group of sulfur- and nitrogen-containing secondary metabolites that are mainly found in the order of Brassicales and related groups of dicotyledonous angiosperms [3-4]. Glucosinolates and the hydrolytic myrosinase (β -thioglucoside glucohydrolase) are stored separately under normal situations, but they come into contact with each other when tissues are damaged, and then the glucosinolates are hydrolyzed into several degradation products, such as isothiocyanates and nitriles [5]. Glucosinolates and their degradation products have diverse biological functions, which contribute to human health, as well as the taste and odor of cruciferous crops. The anticancer activity of isothiocyanates has been widely studied, and the mechanism involved has been elucidated [6].

Glucosinolate metabolism in plants is modulated by numerous biotic and abiotic factors, and the regulatory network of glucosinolate metabolism has been well elucidated in *Arabidopsis* [6]. CYP79B3 belongs cytochrome P450 monooxygenases (cytochromes P450) of the CYP79 family. CYP79B3 together with the homolog CYP79B2 catalyze the conversion of tryptophan to indole-3-acetaldoxime (IAOx) [7]. The gene encoding the CYP79B3 protein has been isolated in *Arabidopsis thaliana* and Chinese cabbage [7-8]. To date, research studies on CYP79B3 in cabbage are limited. In the present study, the CYP79B3 gene sequence of cabbage was obtained from web database, and then sequence analysis of the CYP79B3 gene were analyzed. The present study aimed to establish the foundation for further studies on the molecular mechanism of CYP79B3 in cabbage.

2. Materials and Methods

2.1 Sequence Obtention of the Boccyp79b3 Gene.

The genomic DNA and mRNA sequences of CYP79B3 gene of cabbage were downloaded and obtained from The Brassica database (BRAD) (<http://brassicadb.org>), and then used to subsequent bioinformatic analysis.

2.2 Sequence Analysis of the BocCYP79b3 Gene.

The amino acid sequence, protein molecular weight, isoelectric point, stability index, and hydrophobicity of the BocCYP79B3 gene were analyzed and predicted by ExPASy (<http://web.expasy.org>) and NCBI (<https://www.ncbi.nlm.nih.gov/>). The prediction of protein secondary structure was done using DNASTar software. Subcellular localization was predicted by WoLF PSORT (<http://www.genscript.com/wolf-psort.html>). The conserved domain were predicted by NCBI (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>). Homology analysis of the CYP79B3 proteins was executed in DNAMAN.

3. Results

3.1 Analysis on Genomic Organization.

The Brassica database (BRAD) was used to analyze the chromosomal localization and genomic organization of BocCYP79B3. The gene ID in BRAD is Bol031784. The BocCYP79B3 gene was mapped to Scaffold000053 and has 2 exons and 1 intron (Fig. 1).

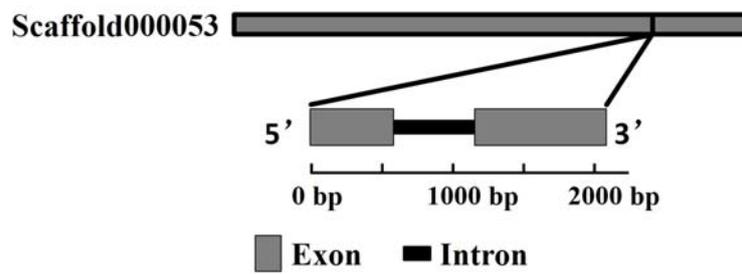


Fig. 1. Chromosomal location and genomic structure of BocCYP79B3.

3.2 Protein Physical and Chemical Properties Analysis.

Sequence analysis indicated that the BocCYP79B3 gene contained a 1,632-bp open reading frame (ORF), which encoded a 543-amino acids protein with a calculated molecular mass of 61.54 kD and an isoelectric point (pI) of 8.59. The amino acid types and proportions of the BocCYP79B3 gene was shown in Figure 2, the highest number of amino acid is Leucine (Leu), whereas the lowest number is Tryptophan (Trp). Its predicted formula was C₂₇₆₅H₄₃₉₁N₇₄₁O₇₇₇S₃₄. Its total average hydrophilicity index was -0.162, liposoluble index was 88.91, and instability index in solution was 40.6. There are 21 α -helices, 18 β -sheets, and 30 β -sheets. Most amino acids are hydrophilic amino acids.

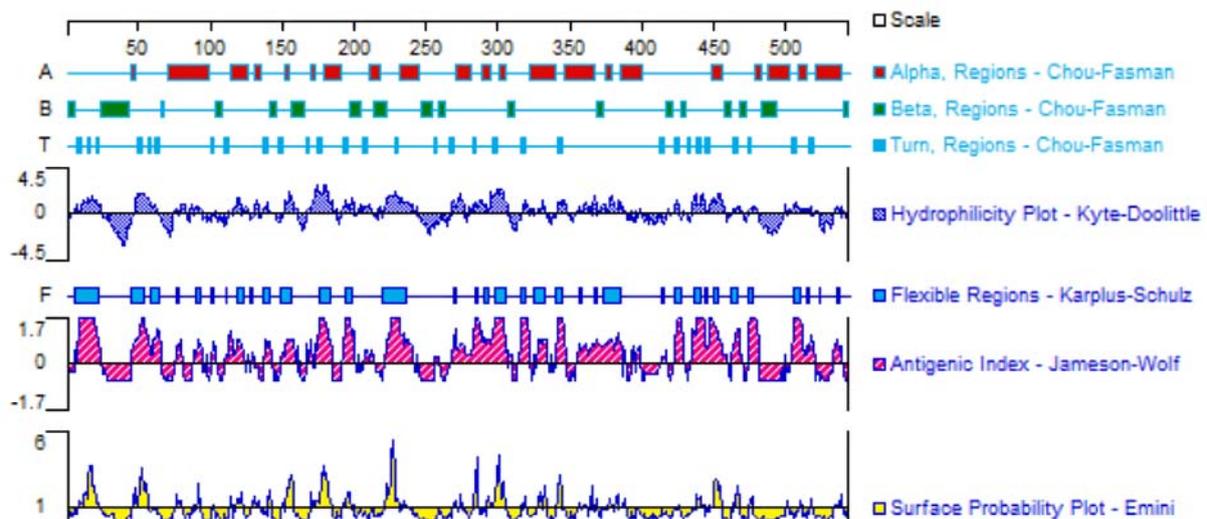


Fig. 2. Prediction of protein secondary structure of BocCYP79B3

3.3 Subcellular Localization and Conserved Domain Analysis.

Subcellular localization of the BocCYP79B3 gene was predicted by WoLF PSORT to be in the chloroplast. The analysis using Conserved Domain Database (CDD) demonstrated that the amino

acid sequence of the BocCYP79B3 protein belonged the p450 superfamily and has the conserved domain PLN02971 (Fig. 3).



Fig. 3. Conserved domains analysis of BocCYP79B3

3.4 Homology Analysis.

Homology analysis demonstrated that the amino acid sequence of the BocCYP79B3 protein shared high homology with those of 18 other higher plant species (Table 1). Table 1 shows that the BocCYP79B3 had the highest identities (> 90%) with several CYP79B3 proteins of Cruciferae such as *Brassica napus*, *Raphanus sativus*, *Arabidopsis thaliana*, and the BocCYP79B3 protein is more closely related to *B. napus*. All of the levels of identity were > 50% with other species cited in our study, indicating that the CYP79B3 protein is highly conserved among different species.

Table 1. The homology comparison among amino acid sequences of CYP79B3 from plant species

Plant species	Protein name	GenBank accession No.	Identity with BocCYP79B3 (%)
<i>Brassica napus</i>	BnCYP79B3	XP_013743573.1	99
<i>Raphanus sativus</i>	RsCYP79B3	XP_018480772.1	98
<i>Eutrema salsugineum</i>	EsCYP79B3	XP_006404690.1	94
<i>Arabidopsis thaliana</i>	AtCYP79B3	BAF01576.1	92
<i>Capsella rubella</i>	CrCYP79B3	XP_006296223.1	91
<i>Leavenworthia alabamica</i>	LaCYP79B3	AGN12785.1	87
<i>Arabis alpina</i>	AaCYP79B3	KFK30571.1	85
<i>Isatis tinctoria</i>	ItCYP79B3	ARB51597.1	85
<i>Camelina sativa</i>	CsCYP79B3	XP_010429154.2	85
<i>Tarenaya hassleriana</i>	ThCYP79B3	XP_019057431.1	81
<i>Carica papaya</i>	CpCYP79B3	XP_021887085.1	61
<i>Bambusa ventricosa</i>	BvCYP79B3	ABD84027.1	59
<i>Sorghum bicolor</i>	SbCYP79B3	XP_002466099.1	58
<i>Eucalyptus grandis</i>	EgCYP79B3	KCW62254.1	58
<i>Panicum hallii</i>	PhCYP79B3	PAN23637.1	57
<i>Populus nigra</i>	PnCYP79B3	AHI88992.1	56
<i>Populus trichocarpa</i>	PtCYP79B3	PNT08538.1	55
<i>Populus euphratica</i>	PeCYP79B3	XP_011005474.1	54

4. Summary

The present study analyzed the BocCYP79B3 gene of cabbage. Zang et al. (2008) transformed the *Arabidopsis* CYP79B3 and CYP83B1 into Chinese cabbage to modulate the indole glucosinolate pathway flux. It was found that overexpression of single CYP79B3 did not affect the profiles of indole glucosinolates. However, co-expressing CYP79B3 with CYP83B1 significantly increased the contents of glucobrassicin, 4-hydroxy glucobrassicin and 4-methoxy glucobrassicin [8]. Previous studies have shown that the CYP79B3 protein is relatively conserved in plants [7]. The findings of the present study show that CYP79B3 from cabbage is highly conserved in plants, similar to that observed in earlier reports. The findings of the present study may serve as a foundation for future studies on the functions of CYP79B3 in glucosinolate metabolism in cabbage.

Acknowledgements

This work was supported by National Natural Science Foundation of China (31500247), key project of Department of Education of Sichuan Province (14ZA0016), and National Student Innovation Training Program (201710626030).

References

- [1]. M. Wennberg, J. Ekvall, K. Olsson, and M. Nyman, Changes in carbohydrate and glucosinolate composition in white cabbage (*Brassica oleracea* var. *capitata*) during blanching and treatment with acetic acid, *Food Chem.* 95 (2006) 226-236.
- [2]. S. Rokayya, C.J. Li, Y. Zhao, Y. Li, and C.H. Sun, Cabbage (*Brassica oleracea* L. var. *capitata*) phytochemicals with antioxidant and anti-inflammatory potential, *Asian Pac. J. Cancer Prev.* 14 (2014) 6657-6662.
- [3]. B.G. Hansen, R.E. Kerwin, J.A. Ober, V.M. Lambrix, T. Mitchell-Olds, J. Gershenzon, B.A. Halkier, and D.J. Kliebenstein, A novel 2-oxoacid-dependent dioxygenase involved in the formation of the goitrogenic 2-hydroxybut-3-enyl glucosinolate and generalist insect resistance in *Arabidopsis*, *Plant Physiol.* 148 (2008) 2096-2108.
- [4]. X. Yan, and S. Chen. Regulation of plant glucosinolate metabolism, *Planta* 226 (2007) 1343-1352.
- [5]. E. Andréasson, L.B. Jørgensen, A.S. Höglund, L. Rask, and J. Meijer, Different myrosinase and idioblast distribution in *Arabidopsis* and *Brassica napus*, *Plant Physiol.* 127 (2001) 1750-1763.
- [6]. N. Benkeblia, *Phytonutritional Improvement of Crops*, John Wiley & Sons, Inc., Hoboken, 2017, pp. 407-733.
- [7]. A.K. Hull, R. Vij, and J.L. Celenza, *Arabidopsis* cytochrome P450s that catalyze the first step of tryptophan-dependent indole-3-acetic acid biosynthesis, *P. Natl. Acad. Sci. USA* 97 (2000) 2379-2384.
- [8]. Y.X. Zang, M.H. Lim, B.S. Park, S.B. Hong, and D.H. Kim, Metabolic engineering of indole glucosinolates in Chinese cabbage plants by expression of *Arabidopsis* CYP79B2, CYP79B3, and CYP83B1, *Mol. Cells* 25 (2008) 231-241.