Sequence Analysis of the Geranylgeranyl Pyrophosphate Synthase Gene in Cabbage

Bo Sun[a], Min Jiang[b], Qiao Yuan[c], Fen Zhang[d], and Haoru Tang[*]

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

*Corresponding author e-mail: htang@sicau.edu.cn, [a]14099@sicau.edu.cn, [b]1754425081@qq.com, [c]1729759838@qq.com, [d]zhangf_12@163.com

Abstract. Geranylgeranyl pyrophosphate synthase (GGPPS) is an important enzyme in carotenoid biosynthesis. Here, the *Brassica oleracea* var. *capitata GGPPS* (*BocGGPPS*) gene sequences were obtained from *Brassica* database (BRAD), and preformed for 2.1. sequence analysis. The *BocGGPPS1, BocGGPPS2 and BocGGPPS3* genes mapped to chromosomes 2,3 and 9, and contains an open reading frame of 1,113 bp, 1,077 bp and 1, 116 bp that encodes a 370, 358, 371 amino acid protein, respectively. Subcellular localization predicted all *BocGGPPS* genes were in the chloroplast. The conserved domain of the BocGGPPS protein is Trans_IPPS_HT. Homology analysis indicates that the levels of identity among BocGGPPSs were all more than 55%, and the GGPPS protein is relatively conserved during plant evolution. The findings of the present study provide a molecular basis for the elucidation of *GGPPS* 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 that consumes a lot 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 ant carcinogenic compounds, including vitamin C, carbohydrates , carotenoids, and glucosinolates [1-2].

Carotenoids are 40-carben or 30-carben terpenoids composed of isoprene skeletons, which are a general term for important natural pigments and are widely found in plants and microorganisms [3-5]. Carotenoids participate in various plant physiological processes, including growth, development, and responses to multiple environmental factors. In green tissues, carotenoids act as antenna pigments in photosynthesis and transmit captured light energy to chlorophyll [3, 5]. In non-green tissues, carotenoids are also pigments that are the coloring factors of many flowers and fruits [5-6]. In addition, carotenoids are precursors to many volatile flavoring substances and phytohormones, such as abscisic acid and strigolactone [4]. Carotenoids, as precursors of vitamin A, are also essential compounds in the human diet [7]. At the same time, carotenoids have the functions of scavenging free radicals, delaying aging, inducing information transmission between cells, inhibiting cell proliferation and enhancing immunity [8-9].

The enzymes involved in the carotenoid biosynthetic pathway have been extensively studied in various model plants, including *Arabidopsis* [10], tomato [11], and citrus [12]. The first step in carotenoid synthesis is the formation of 40-carbon phytoene by two geranylgeranyl pyrophosphate

(GGPP) molecules catalyzed by phytoene synthase (PSY). [13-14]. Then, phytoene (achromatic carotenoid) is converted to lycopene (colored carotenoid) by desaturases and isomerases, including phytoene desaturase (PDS)[15], ζ-carotene desaturase (ZDS) [16], 15-*cis*-ζ-carotene isomerase (Z-ISO) [17], and carotenoid isomerase (CRTISO) [10]. Here after, bifurcation of the carotenoid biosynthetic pathway occurs, and the production of β-carotene and α-carotene is catalyzed by lycopene β-cyclase (β-LCY) and lycopene ε-cyclase (ε-LCY) [18-19].

Geranylgeranyl pyrophosphate (GGPP) is a central precursor for the synthesis of primary and secondary isoprenoid compounds such as chlorophylls, carotenoids and derivatives including the hormones abscisic acid (ABA) and strigolactones, gibberellins, plastoquinones, ubiquinones, phylloquinones, tocopherols, triterpenoids, polyprenols, dolichols, and prenylated proteins [20]. Geranylgeranyl pyrophosphate synthase (GGPPS) is an important enzyme for the biosynthesis of carotenoids, catalyzing the condensation of three molecules of isopentenyl pyrophosphate (IPP) and one molecule of idimethylallyl pyrophosphate (DMAPP) into Geranylgeranyl pyrophosphate GGPP [5]. The genes encoding the GGPPS protein have been isolated in various plant species, including *Arabidopsis* [20], *Nicotiana tabacum* [21], *Ginkgo biloba* [22], and *B. rapa* [14]. To date, research studies on *GGPPS* in cabbage are limited. In the present study, the *GGPPS* gene sequence of cabbage was obtained from web database, and then sequence analysis of the *GGPPS* gene were analyzed. The present study aimed to establish the foundation for further studies on the molecular mechanism of *GGPPS* in cabbage.

2. Materials and methods

2.1. Sequence obtain of the Boggs's gene

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

2.2. Sequence Analysis of the Boggs's Gene

ExPASy (http://web.expasy.org) and NCBI (https://www.ncbi.nlm.nih.gov/) online software were used to analyze and predict the amino acid sequence, protein molecular weight, isoelectric point, stability index and hydrophobicity of the *Boggs's* gene. The WoLF PSORT (http:www.genscript.com/wolf-psort.htm1) online software was used to predict the Subcellular localization. The NCBI (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) was used to predict the conserved domain. We used the NCBI and DNAMAN to download the amino acid sequence of GGPPS from eight additional species and subjected to multiple sequence alignment, respectively. Phylogenetic tree analysis of the GGPPS proteins was executed in MEGA 6.0 was used to execute the phylogenetic tree analysis of the GGPPS proteins by the neighbour-joining (NJ) method.

3. Results

3.1. Analysis on genomic organization

The *Brassica* database (BRAD) was used to analyse the chromosomal localization and genomic organization of *Boggs's*. There are three genes of *GGPPS* in cabbage chromosomes, *BocGGPPS1*, *BocGGPPS2* and *BocGGPPS3*, and the gene IDs in BRAD are Bol028967, Bol025714 and Bol045796, respectively. The *BocGGPPS1* gene was mapped to chromosomes 1 and has 1 exon and 0 intron, the *BocGGPPS2* gene was mapped to chromosomes 3 and has 2 exons and 1 intron, and the *BocGGPPS3* gene was mapped to chromosomes 8 and has 1 exon and 0 intron (Fig. 1).

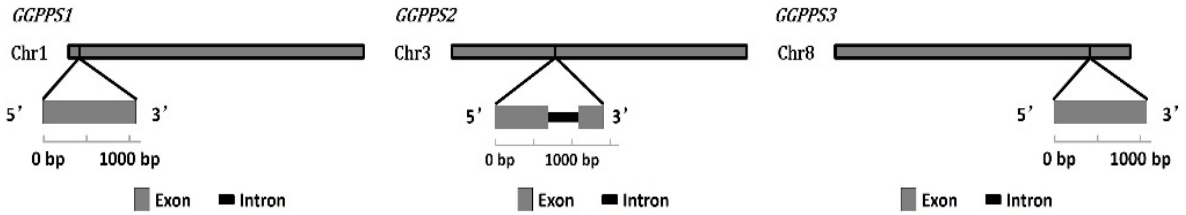

Figure 1. Chromosomal location and genomic structure of *BocGGPPSs*.

3.2. Protein physical and chemical properties analysis

Sequence analysis indicated that the *BocGGPPS1*, *BocGGPPS2* and *BocGGPPS3* gene contained open reading frame (ORF) of 1,113 bp, 1,077 bp and 1,116 bp, encoding a 370, 358, 371 amino acids protein with a calculated molecular mass of 39.81 kD, 38.43 kD and 40.48 kD, and an isoelectric point (pI) of 6.07, 6.27 and 5.28. The amino acid types and proportions of the *BocGGPPSs* gene was shown in Figure 2, the highest number of amino acid in each gene is Leucine (Leu), whereas the lowest number is Tryptophan (Trp). The predicted formula *BocGGPPS1*, *BocGGPPS2* and *BocGGPPS3* were $C_{1742}H_{2847}N_{491}O_{543}S_{14}$, $C_{1684}H_{2775}N_{469}O_{516}S_{18}$ and $C_{1794}H_{2912}N_{484}O_{544}S_{16}$ respectively. Their total average hydrophilicity index were -0.05, 0.082 and 0.025, lip soluble index were 99.68, 103.85 and 103.13, and instability index in solution were 44.59, 44.65 and 42.57, respectively.

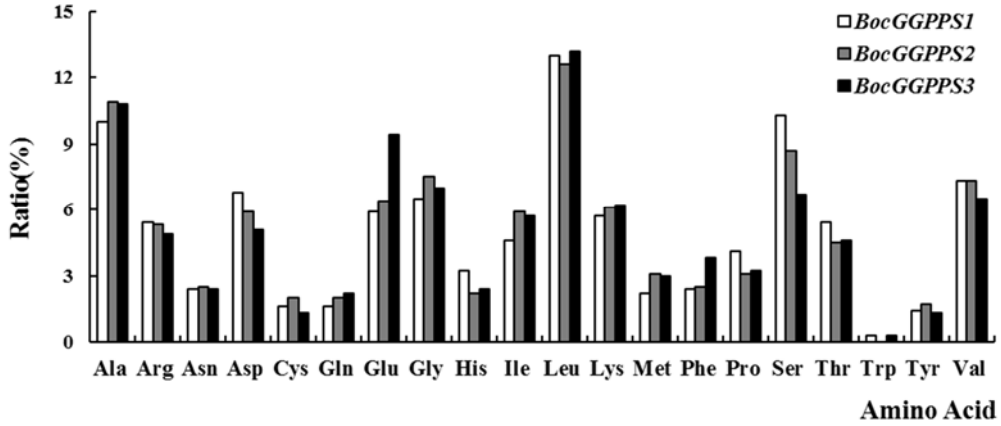

Figure 2. Amino acid composition of BocGGPPSs.

3.3. Subcellular localization and conserved domain analysis

We used the WoLF PSORT to predict that Subcellular localization of the *BocGGPPS1*, *BocGGPPS2* and *BocGGPPS3* are in the chloroplast. The analysis of Conserved Domain Database (CDD) demonstrated that the amino acid sequence of all Boggs's proteins have one conserved domain Trans_IPPS_HT and one Isoprenoid_Biosyn_C1 superfamily (Fig. 3).

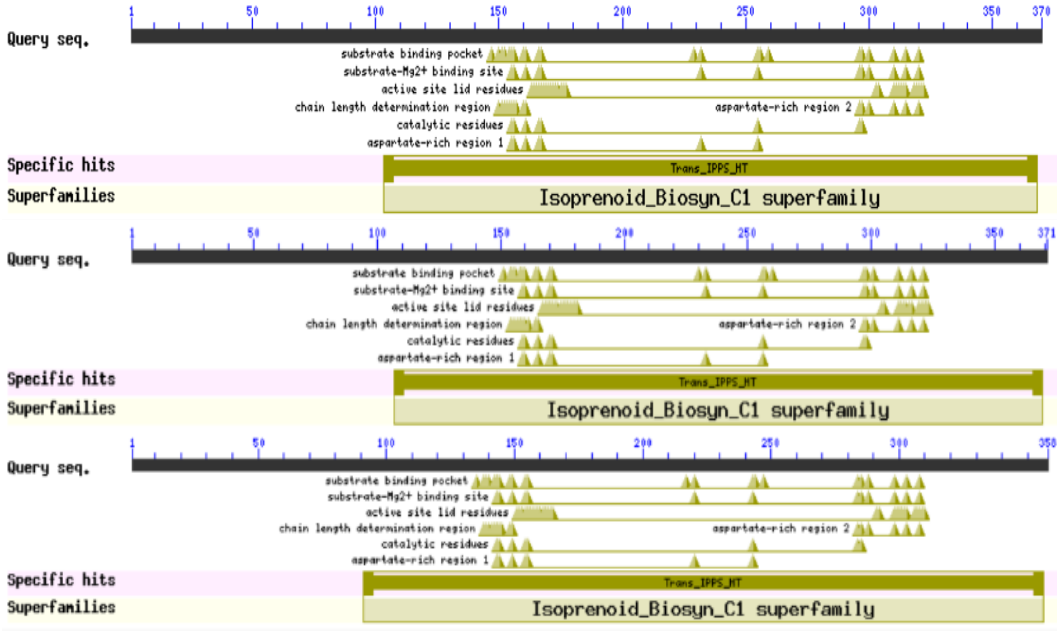

Figure 3. Conserved domains analysis of BocGGPPSs.

3.4. Homology and phylogenetic tree analysis

Homology analysis demonstrated that the amino acid sequence of the Boggs's protein shared moderate homology with those of other higher plant species. The levels of identity among BocGGPPSs were all more than 55%. Figure 4 shows that all BocGGPPSs had the highest identities with several GGPPS proteins of Cruciferae and all of the levels of identity were > 59%, such as *B. rapa*, *B. napus*, *Arabidopsis thaliana*. BocGGPPSs showed > 51% identity with other species we studied, indicating that GGPPS proteins are relatively conservative in different species. Other than this, significant differences were found near the N-termini of GGPPS proteins of various plant species (Fig. 4).

We constructed a phylogenetic tree to illustrate the relationship between the cabbage GGPPS proteins and 37 other higher plant species (Fig. 5). A total of four major clusters were identified, BocGGPPS2 belongs the second cluster, and BocGGPPS1 and BocGGPPS3 belong the fourth cluster. Sequence alignment indicated that the BocGGPPS2 protein sequence was more consistent with Lepidus ape alum and Camellia sinensis, the BocGGPPS3 protein sequence was more consistent with that of Upland cotton, and the BocGGPPS1 protein was a single small cluster.

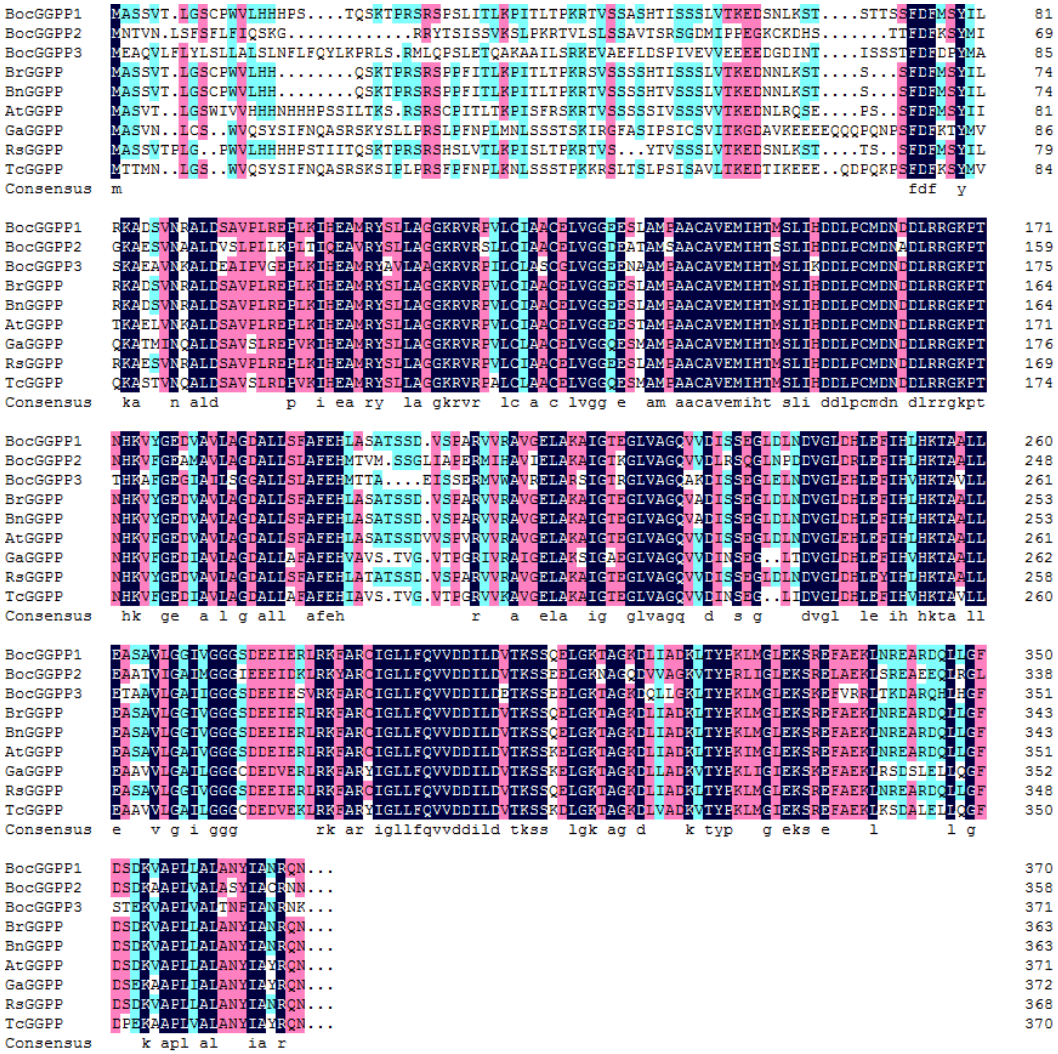

Figure 4. Amino acid sequence alignment of BocGGPPSs and the GGPPS protein of other species.

Boc: *Brassica oleracea* var. *capitata*; Br: *Brassica rapa* (XP_009141710.1); Bn: *Brassica napus* (XP_013738757.1); At: *Arabidopsis thaliana* (NP_195399.1); Ga: *Gossipier arboretum* (XP_017638158.1); Rs: *Raphanus sativus* (XP_018467471.1); Tc: *Theobroma cacao* (EOX91186.1).

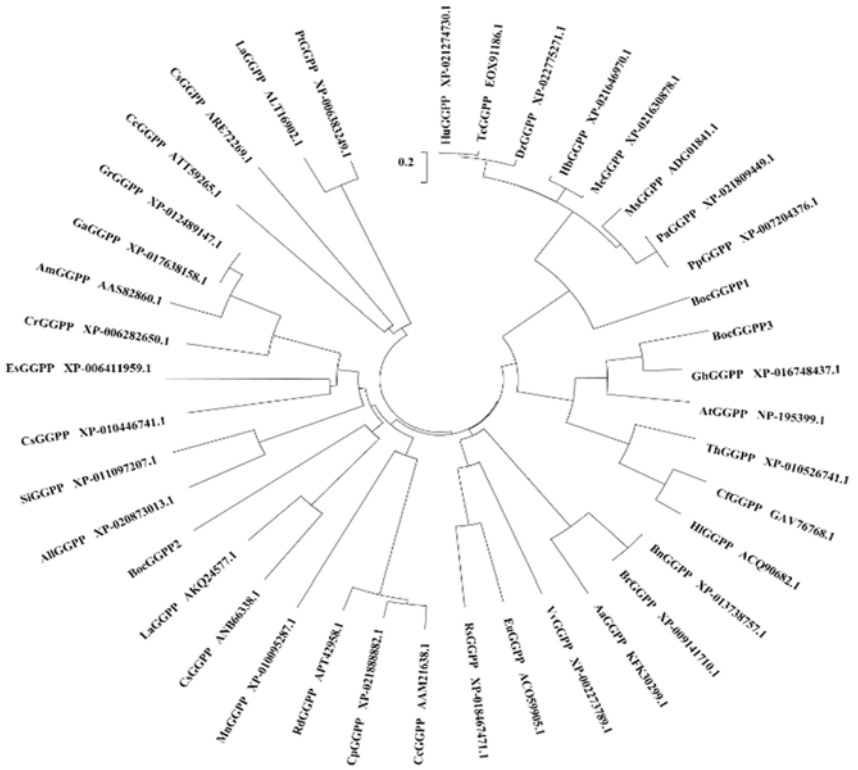

Figure 5. Phylogenetic tree analysis of BocGGPPSs with GGPPS proteins of other species

4. Conclusion

The present study analysed the *Boggs's* gene of cabbage. GGPPS enzyme is encoded by 12-copy genes in *Arabidopsis*, the genes are expressed in specific tissues and specific biological developmental stages, indicating their important role in the isoprene synthesis pathway. [20]. However, the *GGPPS* genes occurred as three copies in cabbage, suggesting that GGPPS enzymes may have experienced different evolutionary patterns to make them have different functions. Previous studies have shown that the GGPPS protein is relatively conserved in plants [21-22]. The GGPPS protein of *Ginkgo biloba* is similar to the GGPPS protein of *Taxus Canadensis* and *Picea abies*, showing 73% and 73% homology, respectively [22], and GGPPS3 of *N. tabacum* exhibits 86% homology that of tomato [21]. The findings of the present study show that GGPPS from cabbage is relatively conserved; in the Cruciferae, these proteins show >59% homology, similar to that observed in earlier reports. The results of this study can lay the foundation for future research on the function of GGPPS in the metabolism of carotenoids in cabbage.

Acknowledgments

This work was supported by National Natural Science Foundation of China (31500247), Sichuan Science and Technology Program (2018NZ0081), and National Student Innovation Training Program (201710626030).

References

[1] 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] L. Pizarro, C. Stange, Cien, Light-dependent regulation of carotenoid biosynthesis in plants, Inv. Agr. 36 (2009) 143 - 162.
[4] L.H. Liu, Z.Y. Shao, M. Zhang, Q.M. Wang, Regulation of carotenoid metabolism in tomato. Mol. Plant. 8 (2015) 28 - 39.
[5] N. Nisar, L. Li, S. Lu, N.C. Khin, and B.J. Pogson, Carotenoid Metabolism in Plants, Mol. Plant. 8 (2015) 68 - 82.
[6] P.A. Tuan, J.K. Kim, J. Lee, W.T. Park, D.Y. Kwon, Y.B. Kim, H.H. Kim, H.R. Kim, S.U. Park, Analysis of carotenoid accumulation and expression of carotenoid biosynthesis genes in different organs of Chinese cabbage (*Brassica rapa* subsp. *pekinensis*), Excli J. 11 (2012) 508 - 516.
[7] N.I. Krinsky, and E.J. Johnson, Carotenoid actions and their relation to health and disease, Mol. Aspects Med. 26 (2005) 459 - 516.
[8] S.T. Mayne, Beta-carotene, carotenoids, and disease prevention in humans, FASEB J. 10 (1996) 690 - 701.
[9] F. Giovannucci, and J. Natl, Tomatoes, tomato-based products, lycopene, and cancer: review of the epidemiologic literature, Cancer Inst. 91 (1999) 317 - 331.
[10] H. Park, S.S. Kreunen, A.J. Cuttriss, D. Della Penna, B.J. Pogson, Identification of carotenoid isomerase provides insight into carotenoid biosynthesis, prolamellar body formation, and photomorphogenesis, Plant Cell 14 (2002) 321 - 332.
[11] T. Isaacson, G. Ronen, D. Zamir, and J. Hirschberg, Cloning of tangerine from tomato reveals a carotenoid isomerase essential for the production of beta-carotene and xanthophylls in plants, Plant Cell. 14 (2002) 333 - 342.
[12] M. Kato, Y. Ikoma, H. Matsumoto, M. Sugiura, H. Hyodo, and M. Yano, Accumulation of Carotenoids and Expression of Carotenoid Biosynthetic Genes during Maturation in Citrus Fruit, Plant Physiol. 134 (2004) 824 - 837.
[13] P.A. Scolnik, and G.E. Bartley, Nucleotide Sequence of an Arabidopsis cDNA for Phytoene Synthase, Plant Physiol. 104 (1994) 1471 - 1472.
[14] P.R. Li, S.J. Zhang, S.F. Zhang, F. Li, H. Zhang, F. Cheng, J. Wu, X.W. Wang, and R.F. Sun, Carotenoid biosynthetic genes in *Brassica rapa*: comparative genomic analysis, phylogenetic analysis, and expression profiling, BMC Genomics. 16 (2015) 492.
[15] P.A. Scolnik, and G.E. Bartley, Phytoene desaturase from Arabidopsis, Plant Physiol. 103 (1993) 1475.
[16] P.A. Scolnik, and G.E. Bartley, Nucleotide sequence of a putative geranylgeranyl pyrophosphate synthase from Arabidopsis, Plant Physiol. 109 (1995) 1499.
[17] Y. Chen, F. Li, and E.T. Wurtzel, Isolation and Characterization of the Z-ISO Gene Encoding a Missing Component of Carotenoid Biosynthesis in Plants, Plant Physiol. 153 (2010) 66 - 79.
[18] B. Pogson, K.A. McDonald, M. Truong, G. Britton, and D. Della Penna, Arabidopsis Carotenoid Mutants Demonstrate That Lutein Is Not Essential for Photosynthesis in Higher Plants, Plant Cell. 8 (1996) 1627 - 1639.
[19] Y.M. Shi, R. Wang, Z.P. Luo, L.F. Jin, P.P. Liu, Q.S. Chen, Z.F. Li, F. Li, C.Y. Wei, M.Z. Wu, P. Wei, H. Xie, L.B. Qu, F.C. Lin, and J. Yang, Molecular cloning and functional characterization of the lycopene ε-cyclase gene via virus-induced gene silencing and its expression pattern in *Nicotiana tabacum*, Int. J. Mol. Sci. 15 (2014) 14766 - 14785.
[20] G. Beck, D. Coman, E. Herren, M.A. Ruiz-Sola, M. Rodríguez-Concepción, W. Gruissem, and E. Vranová, Characterization of the GGPP synthase gene family in *Arabidopsis thaliana*, Plant Mol. Biol. 82 (2013) 393 - 416.
[21] F. Li, M. Li, L.F. Jin, Z.P. Luo, J. Yang, and F.C. Lin, Cloning and Characterization of a New Gene Encoding Geranylgeranyl Pyrophosphate Synthase from *Nicotiana tabacum*, Tob. Sci. Technol. 5 (2012) 62 - 66.
[22] H. J. Zhang, B.Y. Tan, and F.L. Cao, Cloning and characterization of the geranylgeranyl pyrophosphate synthase from *Ginkgo biloba* Linn., J. Nanjing Forestry Univ. 37 (2013) 8 - 12.