BIODIVERSITAS
Volume 21, Number 11, November 2020
Pages: 5054-5064

ISSN: 1412-033X
E-ISSN: 2085-4722
DOI: 10.13057/biodiv/d211109

Isolation and characterization of chalcone synthase (*CHS*) gene from *Phalaenopsis* and *Doritaenopsis* orchids

DEWI SUKMA[1], ALINE SISI HANDINI[2], SUDARSONO[1,♥]

[1]PMB Laboratory, Department of Agronomy and Horticulture, Faculty of Agriculture, Institut Pertanian Bogor. Jl. Meranti, IPB University Campus, Darmaga, Bogor 16680, West Java, Indonesia. Tel.: +62-251-8629364, Fax.: +62-251-8629362, ♥email: sudarsono_agh@apps.ipb.ac.id
[2]Politeknik Kelapa Sawit Citra Widya Edukasi. Jl. Gapura No. 8, Rawa Banteng, Cibuntu, Cibitung, Bekasi 17520, West Java, Indonesia

Manuscript received: 29 August 2020. Revision accepted: 6 October 2020.

Abstract. *Sukma D, Handini AS, Sudarsono. 2020. Isolation and characterization of chalcone synthase (*CHS*) gene from* Phalaenopsis *and* Doritaenopsis *orchids. Biodiversitas 21: 5054-5064.* Chalcone synthase (*CHS*) is a key enzyme in flavonoid biosynthesis. The research aims to isolate and characterize the *CHS* gene nucleotide sequence diversity of nine *Phalaenopsis* and one *Doritaenopsis* genotypes. Genomic DNA was isolated from young leaves and used to PCR amplify the gene using *CHS* specific degenerate primers. Results of PCR amplification yielded DNA fragments of 700 bp. Upon sequencing and nucleotide sequence analysis, we found that the genomic fragments were partial *CHS* gene from *Phalaenopsis* and *Doritaenopsis* genotypes and deposited the sequences in the NCBI Genebank Database accession numbers of KR184089-KR184098. More analysis of the sequences confirmed they shared ranged from 81-99% sequence identities to known *CHS* genes from *Phalaenopsis* deposited in the NCBI Database. They also shared 84-100% amino acid residues identity to *CHS* polypeptides. Multiple sequence alignment (MSA) and phylogenetic analysis further revealed the determined *CHS* nucleotide sequences from *Phalaenopsis* and *Doritaenopsis* genotypes were closely related to *CHS* genes from other orchid species.

Keywords: *CHS* gene, DNA sequencing, genetic diversity analysis, multiple sequence alignment, phylogenetic analysis

INTRODUCTION

Phalaenopsis is one of the most important genera of the *Orchidaceae* family. The number of species is 47 species (Sweet 1980) and updated to 63 species distributed into five sub-genera (Tsai 2011). Breeding programs have generated thousands of newly breed *Phalaenopsis* varieties having a high variation of flower morphologies, sizes, color and quantities. A high number of *Phalenopsis* hybrids registered (an average of 100 new *Phalaenopsis* hybrids every two months from 1976 to 2008) in the Royal Horticulture Society (RHS) indicating the importance of novel *Phalaenopsis* hybrids for market demands (Chang et al. 2009).

The high economic value of *Phalaenopsis* is one of the reasons breeders promote *Phalaenopsis* breeding programs. Taiwan is the leader in *Phalaenopsis* breeding, and 50% of flower export values are from *Phalaenopsis* (Hsiao et al. 2015). Taiwan exported *Phalaenopsis* in either in vitro plantlets, young transplants, mature-ready to flower plants or flowering pot plants. Taiwan put the breeding and development of new varieties as one of the most important strategies to a leader *Phalaenopsis* producer in the world (Tang and Chen 2007).

The *Phalaenopsis* breeding program is mainly designed to improve flower size and color and other plant traits such as longevity, stalk length, leaf shape, ease of cultivation, and disease resistance (Tang and Chen 2007). Flower color is an essential visual character and *Phalaenopsis* species or hybrid are rich in color variation. Plant pigments include chlorophylls, anthocyanins, betalains, and carotenoids (Tanaka et al. 2008). Anthocyanin pigments are flavonoid type of pigments derived from phenylalanine. Characteristics of anthocyanins are water-soluble, synthesized in the cytosol, and localized in vacuoles.

The content of chlorophylls, anthocyanins, and carotenoids was varied in various organs of *Phalaenopsis* species or hybrids (Handini et al. 2016). The differences in anthocyanins, carotenoids, and betalains contents determined the color variations in *Phalaenopsis* flowers. Therefore, understanding how the genetic factors and the genes controlling flower colors will enable *Phalaenopsis* breeders to efficiently develop desirable flower colors.

Some reports of *CHS* from *Phalaenopsis* had been published. Five chalcone synthase (*CHS*) genes from the *Phalaenopsis* hybrid revealed gene duplications (Han et al. 2006). On the other hand, there are three main groups of *CHS* genes in the *Phal. aphrodite* genome based on their amino acid residues and the genes were in three different chromosomes (Kuo et al. 2019). Moreover, the existed *CHS* genes have the same structures, consisted of two exons and one intron. Since many *Phalaenopsis* species exist in Indonesia and they also contain anthocyanins as their flower pigment, it would be essential to evaluate and characterize the *CHS* gene in the *Phalaenopsis* from Indonesia. This research aims to isolate and characterize nucleotide sequences of the *CHS* gene from nine *Phalenopsis* and *Doritaenopsis* genotypes in Indonesia. Subsequently, we evaluated the determined nucleotide sequences to determine the genetic diversity of the orchids. Later on, we plan to use the nucleotide sequence diversities

of the *CHS* gene to develop markers for assisting the *Phalaenopsis* breeding program in the future.

MATERIALS AND METHODS

Plant materials

This research evaluated seven *Phalaenopsis* species, two hybrids, and one *Doritaenopsis* hybrid having flower color variations. The accessions being assessed include *Phal. amabilis (*PAB), *Phal. celebensis* (PCE), *Phal. amboinensis* (PAM*), Phal. gigantea* (PGG), *Phal.* 'Salu-Spot' (PSS), *Phal. cornucervi* (PCC), *Phal. bellina* (PBE), *Phal. violacea* (PVI), *Phal.* 'hybrid PH34' (PH34), and *Doritaenopsis* 'Kenneth Schubert' (DKS). Figure 1 presented color variations among the evaluated accessions, while Table 1 presented the taxonomic classification and pedigree of the orchid accessions used in this study.

Total genomic DNA isolation

The total genomic DNA was isolated from a fully open young leaf of the ten orchid accessions, following the basic CTAB protocols for DNA isolation, routinely used for extracting DNA from the fresh young leaf of *Phalaenopsis* orchids (Handini 2014; Elina et al. 2017; Sudarsono et al. 2017; Sukma et al. 2017; Raynalta et al. 2018; Humaira et al. 2019). Fresh young leaf tissues (~0.1 g) were ground to powder and transferred to pre-warmed extraction buffer (2% w/v CTAB, 1.4 M NaCl, 20 mM EDTA, 100 mM Tris-HCL, pH 8.0, and 2% v/v β-mercaptoethanol). After thorough homogenization and heating at 65ºC for 30 min, the suspension was extracted twice with an equal volume of chloroform: isoamyl alcohol (24: 1), centrifuged at 10.000 rpm 10 min, and the supernatant was transferred to a new microfuge tube. Total DNA precipitation from the supernatant was done by adding two volumes of pre-chilled (-20ºC) 95% ethanol and 0.1 volume of sodium acetate (0.3 M). After centrifugation at 10.000 rpm for 10 min, the DNA pellet was washed twice with 70 % ethanol, air-dried and dissolved in 250 µL of sterile TE buffer. The DNA concentration was estimated by electrophoresis using 1% agarose gel in SB buffer at 100 Volt for 60 minutes. The agarose gel was stained using gelred; the DNA was visualized under ultraviolet (UV) light and documented using a digital camera.

Table 1. The taxonomic classification and pedigree of the orchid accessions used in this study

Accessions	Remark	Section	Subgenera
Phal. amabilis	Species	*Phalaenopsis*	*Phalaenopsis*
Phal. celebensis	Species	*Stauroglottis*	*Phalaenopsis*
Phal. amboinensis	Species	*Amboinensis*	*Polychilos*
Phal. gigantea	Species	*Amboinensis*	*Polychilos*
Phal. 'Salu Spot'	Hybrid between *Phal.* 'Paifang Auckland'×*Phal.* 'Golden Amboin'	*Phalaenopsis*× *Amboinensis*	*Phalaenopsis*× *Polychilos*
Phal. cornucervi	Species	*Polychilos*	*Polychilos*
Phal. bellina	Species	*Amboinensis*	*Polychilos*
Phal. violacea	Species	*Amboinensis*	*Polychilos*
Phal. 'hybrid PH34'	Hybrid-unknown parent	*Unknown*	*Unknown*
Doritaenopsis 'Kenneth Schubert'	Hybrid between *Phal. pulcherrima*× *Phal. violacea*	*Esmeralda*× *Amboinensis*	*Phalenopsis*× *Polychillos*

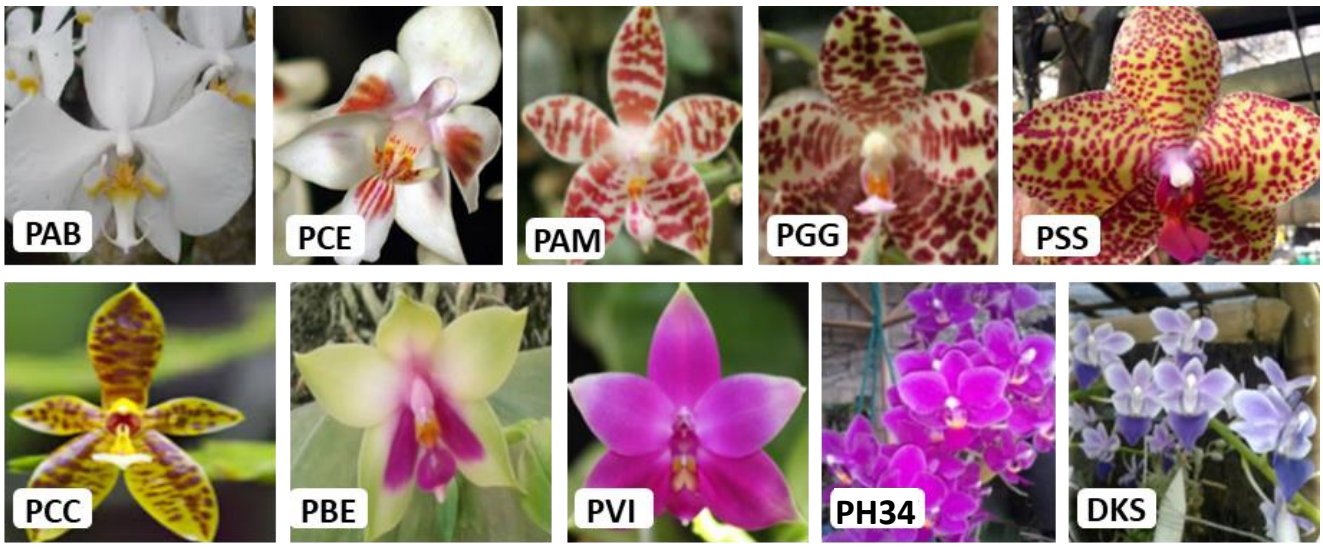


Figure 1. Flower color phenotypes of nine *Phalaenopsis* and a *Doritaenopsis* accessions used in this study. The accessions include PAB: *Phal* amabilis, PCE: *Phal. celebensis;* PAM: *Phal. amboinensis,* PGG*:* Phal *gigantea*, PSS: *Phal.* 'Salu Spot', PCC: *Phal. cornucervi*; PBE: *Phal. bellina*; PVI: *Phal. violacea*; PH34: *Phal.* 'hibrid PH34', DKS: *Doritaenopsis* 'Kenneth Schubert'

Table 2. Sequences of chalcone synthase (*CHS*) gene-specific primers to amplify the *CHS* gene from ten *Phalaenopsis* species and hybrids using total genomic fragments as templates

Primer Identity	Primer sequences (5'-3')	Tm* (°C)	Primer length (bases)	Estimated PCR product size (bp)
*CHS*1_F	CATTTTGGCCATCGGGAGAG	64.9	20	750
*CHS*1_R	GCTCATCAGCCGCCTCA	65.2	17	

Note: * Tm: Melting temperature

CHS specific degenerate primers and PCR amplification

One pair of *CHS* specific primers were designed using sequences of the *CHS* gene (GenBank accession No. U8877.1) available in the NCBI Gen-Bank DNA Database (https://www.ncbi.nlm.nih.gov/). The designed *CHS* specific, forward and reverse primer sequences were presented in Table 2. The PCR amplification was carried out in a 25 µL of a total volume containing KAPA2G™ Ready Mix PCR Kit (Kapa Biosystem Inc., USA), 1.0 µL of primers 10 µM (forward and reverse primer mix), and 30 ng of genomic DNA. The PCR amplification was performed in the BioRad T-100 GeneAmp® PCR System. Initial DNA template denaturation was carried at 95 °C for 3 minutes, followed by 35 cycles of denaturation at 95°C for 10 sec, primer annealing at 47°C to 62°C for 5 sec, and primer extension at 72°C for 1 sec, and a final extension at 72°C for 10 min. The amplicon was electrophoresed using 1% agarose gel in 1X SB buffer at 100 V for 30 min, the gel was stained with gelred, and the amplicon was visualized and photographed on a UV transilluminator.

DNA sequencing and BLAST analysis

The amplicon showing the clear and distinct band in agarose gel was sent to Base Asia (http://base-asia.com/, Malaysia) for DNA sequencing. Before DNA sequencing, the amplicon was electrophoresed in agarose gel, the target DNA fragment cut from the gel, and purified using GeneAid Gel/PCR DNA Fragments Extraction Kit (http: //www.geneaid.com/, Taiwan). The DNA sequencing was conducted using the BigDye® Terminator v3.1 cycle sequencing kit and the sequencing products were fractionated using automatic capillary electrophoresis on Applied Biosystems 310 DNA sequencer. Each genomic fragment of the *CHS* gene was sequenced from both sides using the forward and reversed primers, respectively.

After trimming and removing low-quality nucleotide sequences using Geneious Prime Software (Biomatters, USA), the remaining sequences were evaluated against all nucleotide accessions available at the NCBI GenBank DNA Database (http://ncbi.nlm.nih.gov) using online Basic Local Alignment Search Tool (BLAST) software version 2.8.0. (Altschul et al. 1997) at the BLAST website (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Both BLASTN for generating somewhat similar and MegaBLAST for highly similar sequences were conducted and both of the results were evaluated. The exon and intron parts of the putative *CHS* gene genomic fragment were determined by comparing the determined sequences against a reference full-length *CHS* gene sequence (GenBank Accession No. U8877.1). The identified coding sequences were translated into polypeptides and subsequently compared against all accessions of polypeptides at the NCBI GenBank DNA Database (http://ncbi.nlm.nih.gov) using BLASTP software version 2.8.0.+ (Altschul et al. 1997) at the BLAST website (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Several sample nucleotide and polypeptide accessions were downloaded from the NCBI DNA Database and used for subsequent multiple sequence alignment (MSA) analysis.

Multiple sequence alignment and phylogenetic analysis

Multiple sequence alignment and phylogenetic analysis followed the standard MSA and phylogenetic analysis (Bast 2013). Subsequently, basic sequence handling and preliminary MSA was done using Geneious Prime Software (Biomatters, USA; Genious.com). Final MSA and phylogenetic tree construction were conducted in MEGA X software (www.megasoftware.net/). The tree construction used the Neighbor-Joining method, while bootstrap analysis was done using 10.000 iterations (Tamura et al. 2013). The presence of insertion-deletions (INDELs) and single nucleotide substitutions mutations (SNPs) were evaluated based on the MSA outputs.

RESULTS AND DISCUSSION

PCR amplification of *CHS* genomic fragment

PCR amplification using the *CHS*1_F and *CHS*1_R primer pairs and the total DNA template from the evaluated *Phalaenopsis* accessions resulted in a single band with a fragment size of approximately 750 bp (Figure 2). The designed primers could amplify the target sequences from ten accessions evaluated based on PCR products' presence. The putative *CHS* genomic fragment amplified from ten orchid accessions were sent to Base Asia for direct sequencing of PCR products.

DNA sequencing and BLAST analysis

After trimming and removal of low-quality sequences at both the 5' and 3' end, the remaining high-quality sequences identified were only 733 bp. The identified sequences have been deposited and available in the NCBI GenBank DNA database under the accession number of KR184089.1-KR184098.1. The relative position of the amplicon to the *CHS* gene reference is presented in Figure 3. Based on the MSA analysis results, the amplified fragment is aligned to a partial exon 1 (112 bp), an intron (109 bp), and a partial exon 2 (512 bp) (Figure 3).

The BLAST analysis was used to facilitate the finding of sequence similarity between queries to the nucleotide and protein accessions in the database (Altschul et al. 1997). Results of BLAST analysis confirmed the identity of the amplified fragments as a partial *CHS* gene from *Phalaenopsis* and *Doritaenopsis* (Tables 3 and 4). Table 3 presented the BLASTN (megablast) analysis results of putative *CHS* sequences to all accessions in the GenBank Database. The query coverages of putative *CHS* sequences to the known *CHS* from *Phalaenopsis* were between 85-100%, *Oncidium* were 70-100%, and *Cymbidium was* 41%, while those to the known *BBS* from *Phalaenopsis* were 85% and *Dendrobium* was 69%. The putative *CHS* sequence identities to other *CHS* genes from *Phalaenopsis* were above 96%, while to *CHS* of other genera (*Oncidium* and *Cymbidium*) were between 81-86%. Similar sequence identity levels were also observed to the *BBS* from *Phalaenopsis* (96-100% identity), from *Cymbidium* (90% identity), and *Dendrobium* (87-88%), respectively (Table 3).

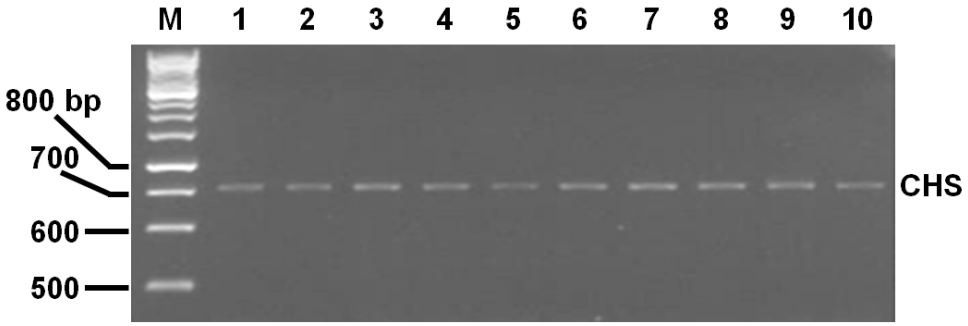
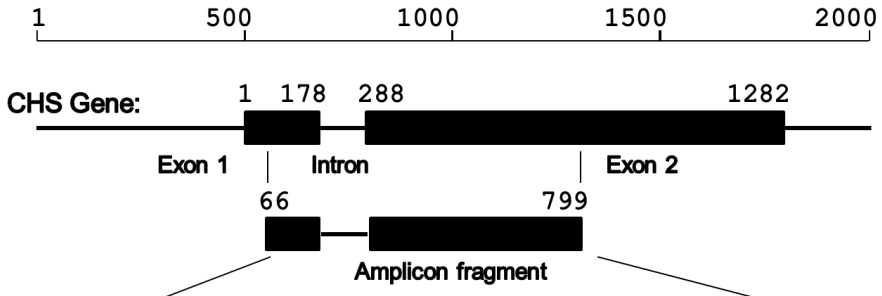


Figure 2. Gel photograph of PCR amplified DNA fragment generated using chalcone synthase (*CHS*) gene-specific primers (*CHS*1_F and *CHS*1_R primers) and genomic DNA templates of ten orchid accessions. The DNA fragments were approximately 750 bp. Lane M**:** DNA marker size (100 bp ladder), lane 1 to 10**:** amplicon of (1) *Phalaenopsis amabilis,* (2) *Phal. amboinensis,* (3) *Phal. bellina,* (4) *Phal. cornucervi,* (5) *Phal. violacea,* (6) *Phal.* 'hybrid PH34', (7) *Phal.* 'Salu Spot' *hybrid,* (8) *Phal. celebensis,* (9) *Phal. gigantea,* and (10) *Doritaenopsis* 'Kenneth Schubert'




```
GGGAGAGCGA ACCCAGACAA TATTATTGAA CAGAGCGCCT ACCCAGACTT CTACTTTCGT   60
GTCACCAATA GCGAGCACTT GGTCGACCTC AAAAAGAAAT TTCAACGCAT CTgtaagtca  120
tctcaattat ctaaagtctg ttctagtata atcgatctcc tttaatagca ttcctcatat  180
tataaatagg ctagagttaa taaatttata ataatcttca gGTGAGAAGA CGGCAATCAG  240
AAAGCGCCAC TTTGTCTGGA ACGAGGAATT TCTCACTGCA AACCCTTGCT TCAGCACATT  300
CATGGACAAA TCTTTAAACG TAAGGCAAGA GGTTGCTATA AGCGAGATAC CAAAGCTGGG  360
CGCGAAGGCG GCCACCAAGG CTATCGAGGA CTGGGGGCAG CCTAAATCGC GTATAACTCA  420
CCTAATCTTC TGCACCACGA GCGGCATGGA CTTACCTGGT GCTGATTACC AGCTCACCCA  480
AATCCTTGGC CTCAACCCAA ATGTTGAGCG TGTCATGCTC TATCAGCAGG GCTGTTTTGC  540
TGGCGGAACC ACGCTTCGCC TGGCCAAGTG TCTTGCTGAG AGCCGCAAAG GCGCACGAGT  600
TCTAGTGGTT TGTGCAGAGA CCACCACCGT GTTATTTCGT GCACCTTCTG AGGAGCATCA  660
GGATGACCTT GTGACCCAGG CTTTATTTGC CGATGGTGCC TCTGCAGTTA TAGTGGGCGC  720
CGATCCAGAT GAG
```

`

Figure 3. The relative position of the amplified fragment to the *CHS* gene reference. The amplicon covered from the 66th nucleotide position of exon 1 to the 799th of exon 2 of the *CHS* gene reference. The underlined sequences indicated the partial sequences of *CHS*1_F (forward-) and *CHS*1_R (reverse-) primers in the amplicon sequences. The capital letters represented the exon sequences, while the smaller letters represented the intron

Table 3. Representative results of BLASTN (MegaBLAST) analysis using the putative chalcone synthase (*CHS*) nucleotide sequences from this study to all nucleotide accessions in the NCBI GenBank DNA Database

NCBI Accession	Organism	Gene	Query cover	E value	Identity (%)
AY825502.1	*Phalaenopsis* hybrid	*CHS*	100	0	99
AY282575.1	*Phal.* x *Doritaenopsis* hybrid	*CHS*	100	0	99
AY825503.1	*Phal.* hybrid	*CHS*	100	0	98
EF139435.1	*Phal. gigantea*	*CHS*2	97	0	96
U88077.1	*Phal.* sp. 'True Lady'	*CHS*	85	0	98
DQ118024.1	*Oncidium* 'Gower Ramsey'	*CHS*2	100	0	83
DQ118022.1	*O.* 'Gower Ramsey'	*CHS*3	100	2E-161	81
DQ118023.2	*O.* hybrid	*CHS*1	70	3E-149	86
EU045579.1	*Cymbidium floribundum*	*CHS*	41	5E-87	86
X79903.1	*Phal.* sp. 'Sport1'	*BBS*	85	0	100
X79904.1	*Phal.* sp. 'Sport1'	*BBS*	85	0	99
XM020716367.1	*Phal. equestris*	*BBS*	85	0	98
XM020716366.1	*Phal. equestris*	*BBS*	85	0	96
KM186175.1	*C.* hybrid	*BBS*	69	0	90
XM020834440.1	*D. catenatum*	*BBS*	69	1E-172	88
XM020834438.1	*D. catenatum*	*BBS*	69	1E-167	88
XM020848439.1	*D. catenatum*	*BBS*	69	3E-163	87
XM020848443.1	*D. catenatum*	*BBS*	69	2E-156	87

Note: *CHS*: Chalcone synthase, *BBS*: Bibensyl synthase

Table 4. Representative results of BLASTP (Protein BLAST) analysis using the translated amino acid residues from putative chalcone synthase (*CHS*) sequences from this study to all polypeptide accessions in the NCBI GenBank Protein Database

GenBank Acc. No.	Protein	Organism	Query cover-age (%)	E-Value	Identity (%)
ALJ57231.1	*CHS*	*Phalaenopsis*	100	7E-155	100
ALJ57233.1	*CHS*	*Phalaenopsis*	99	1E-153	100
AAV70116.1	*CHS*	*Phalaenopsis*	100	2E-151	99
AAX54693.1	*CHS*	*Phalaenopsis*	100	1-135	88
AAZ32095.1	*CHS*	*Oncidium*	100	1E-137	89
AAZ32094.1	*CHS*	*Oncidium*	100	1E-136	89
AAZ32093.1	*CHS*	*Oncidium*	100	1E-129	85
AFU07715.1	*CHS*	*Paphiopedilum*	100	1E-130	85
AFU07712.1	*CHS*	*Paphiopedilum*	100	2E-129	84
AFU07709.1	*CHS*	*Paphiopedilum*	100	1E-115	76
AFU07717.1	*CHS*	*Paphiopedilum*	100	5E-108	70
CAA56277.1	*BBS*	*Phalaenopsis*	100	1E-152	100
XP_020572026.1	*BBS*	*Phalaenopsis*	100	2E-151	99
XP_020572025.1	*BBS*	*Phalaenopsis*	100	4E-147	96
XP_020572030.1	*BBS*	*Phalaenopsis*	100	2E-134	88
XP_020591420.1	*BBS*	*Phalaenopsis*	99	6E-112	74
AIM58716.1	*BBS*	*Cymbidium*	100	6E-140	91
XP_020704098.1	*BBS*	*Dendrobium*	100	6E-142	91
PKU72005.1	*BBS*	*Dendrobium*	100	1E-139	90
XP_020690097.1	*BBS*	*Dendrobium*	100	4E-137	92
XP_020690099.1	*BBS*	*Dendrobium*	100	1E-136	91
XP_020704102.1	*BBS*	*Dendrobium*	100	3E-134	87

Note: *CHS*: Chalcone synthase, *BBS*: Bibensyl synthase

The BLASTP analysis of translated products from putative *CHS* fragments resulted in 88 accessions having high amino acid identities to the queried polypeptides. Seventy-two outputs of the BLASTP analysis results were the *CHS* proteins, and 16 were *BBS*. Out of the 88, four were *CHS*, and five were *BBS* from *Phalaenopsis*, while 11 were *CHS*, and the other 11 were *BBS* from other Orchid species, respectively. Table 4 showed the representative outputs of BLASTP analysis using translated polypeptides of the identified putative *CHS* gene. The putative *CHS* amino acid identities to other *CHS* polypeptides from *Phalaenopsis* were 88-100%, while *CHS* of other genera (*Oncidium* and *Paphiopedilum*) were between 70-89% (Table 3). We also observed similar levels of amino acid residue identity to the *BBS* polypeptide from *Phalaenopsis* (74-100% identity) and other orchid species (*Cymbidium* and *Dendrobium*, 87-92% identity), respectively (Table 4).

Multiple sequence alignment and phylogenetic analysis

The multiple sequence alignment (MSA) analysis for the polypeptides translated from the identified *CHS* fragment revealed only a few amino acid variations. Figure 4 presented the results of the MSA analysis using *CHS* amino acid residues. The MSA results also revealed four *CHS* polypeptide variants based on the amino acid residues at the 185th-190th positions. The majority of the *CHS* polypeptides contain the FADGAS amino acid residues at those positions (25 accessions), followed by the LPMVFL residues (9 accessions), the LPMCLC (1 accession) or FGDGAS (1 accession). The evaluated *Phalaenopsis* and *Doritaenopsis* diploid accessions are identified as heterozygous. The loci of *CHS* encode *CHS* polypeptide with either the FADGAS and LPMVFL, FADGAS and LPMCLC, or FADGAS and LPMCLC residues (the 185th-190th amino acid positions, Figure 4). Although there are a few more amino acid variations in the *CHS* coding sequences; However, the variations only exist among *CHS* and *BBS* accessions from NCBI DNA Database (Figure 4). In those variable regions, the *CHS* sequences characterized in this study were all the same (Figure 4).

Figure 5 presented the phylogenetic tree constructed using the *CHS* and *BBS* nucleotide sequences from *Phalaenopsis* and other orchid species. Based on *CHS* sequences, the evaluated *Phalaenopsis* belonged to two main groups. The first group (Group I) consisted of six *CHS* genes from *Phal. equestris* (Acc. No. KF769460.1) and *Phal. hybrids* (Acc. No. DQ089652.1 and AY825503.1), *CHS* from *Oncidium* (Acc. No. DQ118024.1 and DQ118023.2) and a *BBS* gene from *Cymbidium* (Acc. No. EU045579.1). All of the *CHS* isolated in this study belonged to the second group (Group II). In Group II (Figure 5) also include *CHS* from *Phalaenopsis* sp. 'True Lady' (Acc. No. U88077.1), *CHS* from *Bletilla striata* (Acc. No. KF812890.1, *BBS* from *Dendrobium* (Acc. No. XM020834440.1, XM020834438.1, XM020848439.1, XM020848443.1), *BBS Phal. equestris* (Acc. No. XM020716367.1 and XM020716366.1) and *BBS Phal.* sp. 'Sport1' (Acc. No. X79903.1).

Discussion

In this study, we have used *CHS* specific primers (*CHS*1_F and *CHS*1_R) to PCR amplify the gene encoding *CHS* from the genomic DNA of nine *Phalaenopsis* and a *Doritaenopsis* accessions as the template DNA. The PCR amplification resulted in a single, of an approximately 750 bp DNA fragment. Upon sequencing of the DNA fragment, 733 bp of nucleotide sequences were determined. Further analysis of the sequences showed high sequence identities to either *CHS* or *BBS* genes for various orchids species.

However, multiple sequence alignment analysis results among known *CHS* and *BBS* genes from both *Phalaenopsis* and other orchid species indicated that the isolated fragments in this study were *CHS* from the nine *Phalaenopsis* and a *Doritaenopsis* accessions. Although both *CHS* and *BBS* are members of the PKS type III protein superfamily, the *BBS* genes are involved in the biosynthesis of dihydro-m-hydroxy-resveratrol from m-hydroxy phenyl propyl dihydro-m-coumaroyl-CoA and malonyl-CoA (Wang et al. 2014), while *CHS* genes were involved in the flavonoid biosynthesis (Wang et al. 2014; Zhao and Tao 2015).

Determined sequences encompass nucleotide positions of the 66th to the 799th of *CHS* coding sequences and represent part of the exon 1, the intron, and part of the exon 2. Moreover, the identified exon sequences represent 726 bp (62%) out of 1173 bp of the *CHS* coding sequences. Hence, the identified fragment encodes 242 amino acid residues (62%) out of 391 residues of the *CHS* polypeptides.

Three partial *CHS* genes from *Phal. aphrodite* (*Pa*TC127065, *Pa*TC147749, *Pa*TC151596), consisted of 1039, 981, and 1237 bp nucleotide sequences (Chao et al. 2017). The largest *Pa*TC151596 encompassed up to 99% of the complete *CHS* coding sequences. In another report, the presence of five *CHS* loci from *Phal. aphrodite* genome sequences (*PaCHS*1, *PaCHS*2, *PaCHS*3, *PaCHS*4, and *PaCHS*5), which encode polypeptides having 388 (*PaCHS*2) to 395 (*PaCHS*1) amino acids residues were confirmed (Kuo et al. 2018). Several researchers have isolated and characterized the *CHS* gene from various plant species, such as *Freesia* hybrid and *Grewia asiatica* (Sun et al. 2015; Wani et al. 2017). The *FhCHS*1 coding sequences from *F.* hybrid consisted of 1079 bp and translated into 389 amino acids (Sun et al. 2015), while the *GaCHS*1 coding sequences from *G. asiatica* consisted of 1170 to 1176 bp and encoding 390 to 392 amino acids (Wani et al. 2017).

There is a high expression of the *PhCHS*5 gene during the petal and lip development of *Phalaenopsis* flowers (Han et al. 2006). Several of the accession in this research varied for anthocyanins and carotenoid content in their flowers except for PH34, PSY, DKS, and PVI did not contain carotenoids (Handini et al. 2016). The *Phalaenopsis* PH34 hybrid with a deep pink flower contains the highest anthocyanin (Handini et al. 2016). Our study has isolated the *CHS* gene from *Phalaenopsis* and *Doritaenopsis* orchids having a wide color variation. Although there is variation in their anthocyanin content, the evaluated orchids have the same *CHS* nucleotide sequences. Therefore, the *CHS* gene may not be the only factor affecting flower color in either *Phalaenopsis* or *Doritaenopsis* orchids.

The *CHS* is a key enzyme for anthocyanin biosynthesis (Dao et al. 2011). The *CHS* enzyme belongs to a plant-specific polyketide synthase (PKS) type III protein (Schroder et al. 1997), which catalyzes the first step in the phenylpropanoid pathway for flavonoid biosynthesis (Zhao and Tao 2015). Flavonoid has an essential role in forming flower colors, such as pink, red, and purple (Tanaka et al. 2008). A group of secondary metabolites consisting of subgroups includes flavonols, flavanones, flavanonols, flavanols or catechins, anthocyanins and chalcones. Anthocyanins include cyanidin, delphinidin, malvidin, pelargonidin and peonidin, which are pigments responsible for colors in plants, flowers and fruits (Panche et al. 2016).


1 80
CHS-PABa GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-PABb GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-PAMa GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-PAMb GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-PBEa WRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-PBEb WRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-PCCa WRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-PCCb WRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-PCEa WRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-PCEb WRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-PGGa GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-PGGb GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-PH47a WRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-PH47b WRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-PSSa GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-PSSb GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-DKSa GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-DKSb GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-PVIa GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-PVIb GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-AY282575.1 GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-AY825502.1 GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
CHS-KF812890.1 GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTSNPSFSTFMDKSLYVRQEVAIREI
CHS-U88077.1 GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTSNPCFSTFMDKSLNVRQEVAIREI
CHS1-DQ118023.2 GRANPEHFIEQSSYPDFTFRVTNSEHLVALKKKFQRICDKTAIRKRHFAWNEELLTTNSCFQTFMGNSLNVRQEFAIREI
CHS2-DQ118024.1 GRANPENFVEQNAYPDLEFRATNSKHLVNLKKKFQRICDKTAIRKRHFAWNEEFITANPCLQTFMDNSLNVRQEFAITYI
CHS3-DQ118022 GSANPENIIEQSTYPDFTFHLTNSEHLVDLKNKFQRICDKTAIRKRHFAWNEELLTANPCLRTFMDNSLNVRQEFVIREI
BBS-KM186175.1 GRANPENFIEQSTYPDFTFRVTNSEHLVNLKKKFQRICDKTAIRKRHFVWNEELLNANPCLGTFMDNSLNVRQEFAIREI
BBS-X79903.1 GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
BBS-X79904.1 GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAISEI
BBS-XM020716366 GRANPENFIEQSAYPDFTFRVTNSEHLVDLKKKFQRICDKTAIRKRHFVWNEEFLTANPCFSTFMDKSLNVRQEVAIREI
BBS-XM020716367 GRANPDNIIEQSAYPDFYFRVTNSEHLVDLKKKFQRICEKTAIRKRHFVWNEEFLTSNPCFSTFMDKSLNVRQEVAIREI
BBS-XM020834438 GRANPENFIEQSTYPDFTFRITNSEHLVDLKKKFQRICDKTAIRKRHFVWNEEFTTTNPCLHTFMDKSLDVRQEVAIREI
BBS-XM020834440 GRANPENFIEQSTYPDFTFRITNSEHLVDLKKKFQRICDKTAIRKRHFVWNEEFTTTNPCLHTFMDKSLDVRQEVAIREI
BBS-XM020848439 GRANPENFIEQSAYPDLEFRITNSEHLVDLKNKFKRICDKTAIRKRHFVWIEEFITANPCFSTFMDKSLNIRQEVAIREI
BBS-XM020848443 GRANPENFIEQSAYPDLEFRITKSEHLVDLKNKFKRICDKTAIRKRHFVWTEEFTTANPCFSTFMEKSLNTRQEVAIREI
81 160
CHS-PABa PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-PABb PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-PAMa PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-PAMb PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-PBEa PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-PBEb PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-PCCa PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-PCCb PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-PCEa PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-PGGa PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-PGGb PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-PH47a PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-PH47b PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-PSSa PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-PSSb PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-DKSa PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-DKSb PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-PVIa PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-PVIb PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-AY282575.1 PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-AY825502.1 PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMALPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS-KF812890.1 PKLGAKAATKAIEDWGQPKSRISHLIFCTTSGMDLPGADYQLTQILGLNPNVERLMLYQQGCFAGGTTLRLAKCLAESRK
CHS-U88077.1 PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS1-DQ118023.2 PKLGAQAANKAIQEWGQPKSRITHLIFCTTSGMDLPGADYQLTEILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS2-DQ118024.1 PKLGAEAATKAIQEWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
CHS3-DQ118022 PKLGAQAATKAIQEWGQSKSRITHLIFSTTSGMDLPGADYQLTQILGLNPNIERVMLYHQGCFAGGTTLRLTKSLAESRK
BBS-KM186175.1 PKLGAEAATKAIQEWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNIERVMLYQQGCFAGGTTLRLAKCLAESRK
BBS-X79903.1 PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
BBS-X79904.1 PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
BBS-XM020716366 PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
BBS-XM020716367 PKLGAKAATKAIEDWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
BBS-XM020834438 PKLGAKAAAKAIQEWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
BBS-XM020834440 PKLGAKAAAKAIQEWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESRK
BBS-XM020848439 PKLGAEAATKAIQEWGQPKSRITHLIFCTTSGMDLPGADYQLTQILGLNPNVERVMLYQQGCFAGGTTIRLAKCLAESRK
BBS-XM020848443 PKLGAEAAAKAIQEWGQPKSRITHLIFCTRSGMSLPGPDYQLTQILGLNPNVERVMLYQQGCFAGGTTLRLAKCLAESHK

Figure 4. Multiple sequence alignment (MSA) analysis results among 190 amino acid residues of the chalcone synthase (*CHS*) and bibenzyl synthase (*BBS*) originated from *Phalaenopsis* and other orchid species. The first 20 polypeptides were from this study, and the rest were from the NCBI GenBank Protein Database. The shaded letters represented the position of variant amino acid residues

The *CHS* is members of the PKS type III protein superfamily containing highly conserved amino acid residues, and they have a conserved catalytic structure consisting of four amino acid residues (Cys-His-Asn-Phe) (Ferer et al. 1999). In this study, the *CHS* genes isolated from *Phalaenopsis* and *Doritaenopsis* orchids may also contain both the conserved amino acid residues and the Cys-His-Asn-Phe catalytic structure of the PKS type III protein superfamily. However, since the isolated fragment was only partial *CHS* gene, the findings can only confirm that the amplified fragment was the *CHS* from *Phalaenopsis* and *Doritaenopsis*. This study can not confirm the conserved catalytic structure since it is located beyond the identified sequences.

The *BBS* is also a member of the PKS type III protein superfamily. In Cymbidium orchid, the *BBS* gene has been reported not associated with pigment accumulation (Wang et al. 2014). The type III polyketides are the simple homodimeric structured enzyme composed of 40-45 kDa subunits, each having homology sequences (Taura et al. 2016). The conserved Cys-His-Asn catalytic triad of the PKS type III enzyme functions in sequential condensations of C2 units from malonyl-CoA (Abe and Morita et al. 2010). At least 30 SNPs were found among the *CHS* and *BBS* sequences from the evaluated *Phalaenopsis* species. Moreover, the *CHS* and *BBS* amino acid residues from *Phalaenopsis* are also closely related to CHA and *BBS* from other orchid genera (*Cymbidium* and *Oncidium*).

The *CHS* gene structure was highly conserved among different plant taxa (Conrad et al. 2016), and most of the reported *CHS* gene consisted of two exons and an intron (Ferrer et al. 1999). The *CHS* gene from *Clivia miniata* (Conrad et al. 2016) and fourteen *CHS* genes from maize also consisted of two exons and one intron (Han et al. 2016). The *FhCHS*1 from red *Freesia* hybrid also contained one intron with an 1170 bp open reading frame encoding 390 amino acid residues (Sun et al. 2015) while that of *Antirrhinum* contained two introns and two exons (Sommer and Saedler 1988). In this study, the isolated *CHS* gene from nine *Phalaenopsis* and a *Doritaenopsis* confirmed the typical *CHS* gene structure with the two exons and an intron.

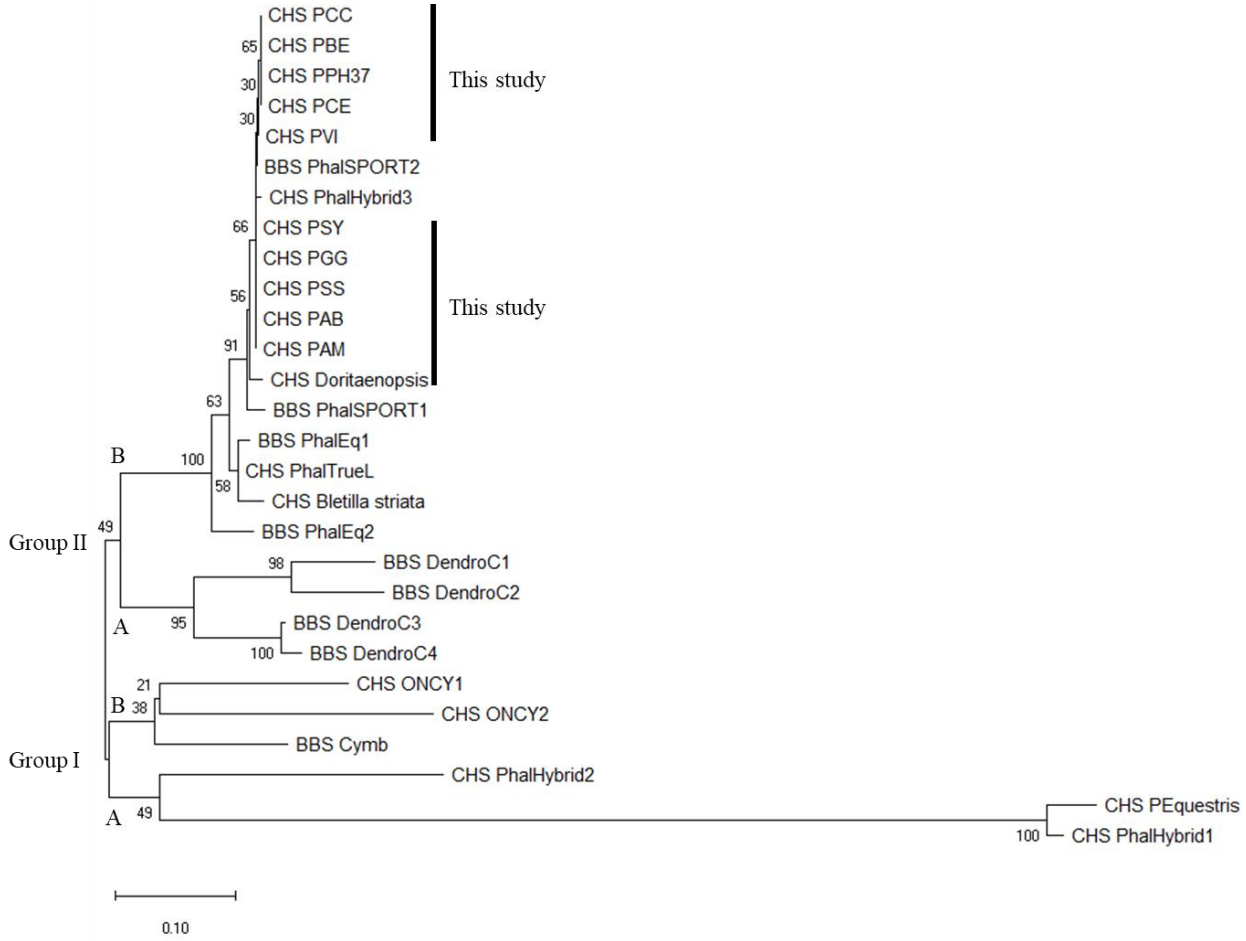


Figure 5. Phylogenetic of *Phalaenopsis CHS* nucleotide sequences was inferred using the Neighbor-Joining method. The numbers in the figure indicated the output values of the bootstrap test using 1000 iterations. The phylogenetic tree was conducted in MEGA X

Some of the *Phalaenopsis* species used in this study (*Phal. amboinensis, Phal. violacea, Phal. bellina, Phal. gigantea, Phal. cornucervi*, and *Phal. amabilis*) belong to the Polychilos subgenus. Meanwhile, *Phal. celebensis* belong to the *Phalaenopsis* subgenus. Two of the accessions were intraspecific hybrid (*Phal.* 'Salu Spot' and *Phal.* 'hibrid PH34'), and one was an interspecific hybrid (*Doritaenopsis*). Although taxonomically, some of the evaluated species belong to different subgenera or different sections, they all have similar *CHS* sequences. Moreover, both the intraspecific *Phalaenopsis* hybrids and the interspecific hybrid also contain similar *CHS* sequences, as shown in this study. Therefore, the nine *Phalaenopsis* and the *Doritaenopsis* accessions belonged to the same group II.B. in the phylogenetic analysis.

The orchid accessions belong to two subgenera (*Polychilos* and *Phalaenopsis*). The *Polychilos* subgenus is divided into sections Amboinensis and Polychilos, while the *Phalaenopsis* subgenus is divided into sections *Amabilis* and *Stauroglottis*. In this study, the evaluated species from section *Amboinensis* include *Phal. amboinensis, Phal.violacea, Phal. bellina,* and *Phal. gigantea;* while that from section *Polychilos* is *Phal. cornucervi*. The accession from section *Stauroglottis* is *Phal. celebensis,* while those from section *Phalaenopsis* is *Phal. amabilis. Phal.* 'Salu Spot', *an* advanced hybrid between *Phal.* 'Paifang Auckland' (section *Phalaenopsis*) and *Phal.* 'Golden Amboin' (section *Phalaenopsis* x *Amboinensis*). The *Phal.* 'hybrid PH34' is unknown, but the suggested progenitor parent was from the *Phalenopsis* subgenus, section *Esmeralda* and *Phalaenopsis*.

Phylogenetic analysis using the *CHS* sequences placed all of the evaluated accessions to the same closely related group (Group II.B). Although some of the evaluated accessions belong to different subgenera and sections, and two are interspecific hybrids, they showed similar *CHS* sequences. Such results suggest either the evaluated accessions are probably closely related or the *CHS* sequences are very conserved among the evaluated accessions. More data are probably needed to support either of these hypotheses.

The nucleotide diversity of specific genes is useful for evaluating genetic diversity among plant species and developing the gene-specific markers for an association between genetic variation and plant phenotypes. Three markers groups can be used to evaluate allelic variations within plant genomes, such as simple sequence repeats (SSRs), insertion-deletions (InDels), and single nucleotide polymorphisms (SNPs) (Mammadov et al. 2012). In this research, only limited numbers of SNPs existed among

nucleotide sequences of the evaluated *CHS* gene of nine *Phalaenopsis* and one *Doritaenopsis* orchids. However, the full-length *CHS* isolation and characterization may identify more SNPs in the *CHS* coding regions. These SNPs may be developed into SNP-based *CHS* gene-specific markers, which will be useful for supporting orchid breeding for flower color variations.

Researchers have used SNP markers for association analysis to certain plant traits such as pungency in *Capsicum* (Garces-Claver et al. 2007), hairy root phenotype in *Brassica* (Zhang et al. 2018), the dwarf character in *Lagerstroemia* (Ye et al. 2016), clustering and disease resistance trait in sugar beet (Simko et al. 2012; Bakooie et al. 2015). The SNP position is a critical point in its application to plant breeding, and the SNPs that existed in the coding regions are the most useful for genetic analysis (Zang et al. 2018). Therefore, identifying SNP presence in the *CHS* coding region is an essential step toward association analysis among SNPs and flower color in *Phalaenopsis*.

In conclusion, we have isolated and characterized the partial *CHS* nucleotide sequences from nine *Phalaenopsis* and one *Doritaenopsis* orchids. The *CHS* nucleotide sequences from those 10 orchids were quite conserved, except for a few regions. The determined *CHS* sequences may probably be essential for the development of the gene-specific markers, such as the single nucleotide amplified polymorphism (SNAP) markers for flower color prediction in *Phalaenopsis* sp. and *Doritaenopsis* sp. Availabilities of the developed SNAP markers would be beneficial for supporting *Phalaenopsis* and *Doritaenopsis* breeding in the future.

ACKNOWLEDGEMENTS

Funding for part of this research was from National Strategic Research Grant, Entitled 'Perakitan Varietas Anggrek *Phalaenopsis* Berbasis Spesies Asli dan Introduksi dan Penggunaan Teknologi Marka Molekuler untuk Percepatan Perolehan Varietas Unggul Baru' Project No. 40/IT3.11/LT/2014, the Directorate General of Higher Education, Ministry of National Education, Republic of Indonesia, under the coordination of DS.

REFERENCES

Abe I, Morita H. 2010. Structure and function of the chalcone synthase superfamily of plant type III polyketide synthases. Nat Prod Rep 27: 809-838. DOI: 10.1039/b909988n.

Bakooie M, Pourjam E, Mahmoudi SB, Safaie N, Naderpour M. 2015. Development of an SNP marker for sugar beet resistance/susceptible genotyping to root-knot nematode. J Agr Sci Tech 17: 443-454.

Bast F. 2015. Sequence similarity search, multiple sequence alignment, model selection, distance matrix and phylogeny reconstruction. Nature Protocols · July 2013. DOI: 10.1038/protex.2013.065.

Chang YK, Veilleux RE, Iqbal MJ. 2009. Analysis of genetic variability among *Phalaenopsis* species and hybrids using Amplified Fragment Length Polymorphism. J Amer Soc Hort Sci 134(1): 58-66. 2009.

Chao YT, Yen SH, Yeh JH, Chen WC, Shih MC. 2017. Orchidstra 2.0 A transcriptomics resource for the orchid family. Plant Cell Physiol 58 (1): e9. DOI: 10.1093/pcp/pcw220.

Dao T, Linthorst H, Verpoorte R. 2011. Chalcone synthase and its functions in plant resistance. Phytochem Rev 10 (3): 397. DOI: 10.1007/s11101-011-9211-7P.

Conrad AC, Mathabatha MF. 2016. Characterization and expression analyses of chalcone synthase (*CHS*) and anthocyanidin synthase (ANS) genes in *Clivia miniata*. Transcriptomics 4: 136. DOI: 10.4172/2329-8936.1000136.

Han Y, Ding T, Su B, Jiang H. 2016. Genome-wide identification, characterization and expression analysis of the chalcone synthase family in maize. Intl J Mol Sci 17 (2): 161. DOI: 10.3390/ijms17020161PMID: 26828478.

Han YY, Ming F, Wang W, Wang JW, Ye MM, Shen DL. 2006. Molecular evolution and functional specialization of chalcone synthase superfamily from *Phalaenopsis* orchid. Genetica 128: 429-438. DOI: 10.1007/s10709-006-7668-x.

Handini AS. 2014. Morphological and biochemical analysis on *Phalaenopsis* and genetic diversity by SNAP marker. [Thesis]. Bogor Agricultural University, Bogor. [Indonesian]

Handini AS, Sukma D, Sudarsono. 2016. Analisis keragaman morfologi dan biokimia pada anggrek *Phalaenopsis* (*Orchidaceae*) J. Agron Indonesia 44 (1) : 62-67. DOI: 10.24831/jai.v44i1.12502.

Hsiao HI, Chang CC, Wu CH. 2015. Decision-making factors for country-of-origin agriculture branding in international markets: the case of orchids in Taiwan. J Horticulture 2: 134. DOI: 10.4172/2376-0354.1000134.

Humaira M, Purwito A, Sudarsono, Sukma D. 2020. Multiplikasi tunas in vitro anggrek *Phalaenopsis* dan analisis keragaman Genetik dengan marka SNAP. J Agronomi Indonesia 48 (1): 59-67. DOI: 10.24831/jai.v48i1.29149.

Kuo YT, Chao YT, Chen WC, Shih MC, Chang SB. 2019. Segmental and tandem chromosome duplications led to divergent evolution of the chalcone synthase gene family in *Phalaenopsis* orchids. Annals of Botany 123: 69-77. DOI: 10.1093/aob/mcy136.

Mammadov J, Aggarwal R, Buyyarapu R, Kumpatla S. 2012. SNP markers and their impact on plant breeding. Int J Plant Genomics 2012: 728398. DOI: 10.1155/2012/728398.

Panche AN, Diwan AD, Chandra SR. 2016. Flavonoids: an overview. J Nutr Sci 2016;5: e47. DOI: 10.1017/jns.2016.41.

Raynalta E, Elina J, Sudarsono, Sukma D. 2018. Clonal fidelity of micro-propagated *Phalaenopsis* plantlets based on assessment using eighteen *Ph-Pto* SNAP marker loci. Agrivita J Agric Sci 40: 390-402. DOI: 10.17503/agrivita.v40i3.1493.

Simko I, Eujayl IA, van Hintum T. 2012. Empirical evaluation of DArT, SNP, and SSR marker-systems for genotyping, clustering, and assigning sugar beet hybrid varieties into populations. Plant Science: 154-162. DOI: 10.1016/j.plantsci.2011.12.009.

Sommer H, Bonas U, Saedler H. 1988. Transposon-induced alterations in the promoter region affect transcription of the chalcone synthase gene of *Antirrhinum majus*. Mol Gen Genet 211: 49-55. 24. DOI: 10.1007/BF00338392.

Sudarsono, Haristianita MD, Sukma D. 2017. Molecular marker development based on diversity of genes associated with pigment biosynthetic pathways to support breeding for novel colors in *Phalaenopsis*. Acta Hort 1167: 237-244. DOI: 10.17660/ActaHortic.2017.1167.4.4

Sukma D, Elina J, Giyanto, Sudarsono. 2017. Disease resistance breeding of *Phalaenopsis* spp. for tropical environment and molecular marker development for plant selection. Acta Hort 1167: 237-244. DOI: 10.17660/ActaHortic.2017.1167.36.

Sun W, Meng X, Liang L, Jiang W, Huang Y, He J, Hu H, Almqvist J, Gao X, Wang L. 2015. Molecular and biochemical analysis of chalcone synthase from Freesia hybrid in flavonoid biosynthetic pathway. PLoS ONE 10(3): e0119054. DOI: 10.1371/journal.pone.0119054.

Tamura K, Stecher G, Peterson D, Filipski V, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30(12): 2725-2729. DOI: 10.1093/molbev/mst197 PMID: 24132122.

Tanaka Y, Sasaki N, Ohmiya A. 2008. Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. Plant J 54 (4): 733-749. DOI: 10.1111/j.1365-313X.2008.03447.

Tang CY, Chen WH. 2007. Breeding and development of new varieties in *Phalaenopsis*. pp 1-22. In: W.H. Chen and H.H. Chen HH (eds.). Orchid Biotechnology. World Scientific, Singapore.

Taura F, Iijima M, Yamanaka E, Takahashi H, Kenmoku H, Saeki H, Morimoto S, Asakawa Y, Kurosaki F, Morita H. 2016. A Novel Class of Plant Type III Polyketide Synthase Involved in Orsellinic Acid

Biosynthesis from *Rhododendron dauricum*. Front Plant Sci 7: 1452. DOI: 10.3389/fpls.2016.01452.

Tsai CC. 2011. Molecular phylogeny and biogeography of *Phalaenopsis* species. pp 1-24. In: W.H. Chen and H.H. Chen (eds.). Orchid Biotechnology II. Word Scientific. Singapore.

Wang L, Albert NW, Zhang H, Arathoon S, Boase MR, Ngo H, Schwinn KE, Davies KM, Lewis DH. 2014. Temporal and spatial regulation of anthocyanin biosynthesis provides diverse flower color intensities and patterning in Cymbidium orchid. Planta 240: 983-1002.

Wani TA, Pandith SA, Gupta AP, Chandra S, Sharma N, Lattoo SK. 2017. Molecular and functional characterization of two isoforms of chalcone synthase and their expression analysis in relation to flavonoid constituents in *Grewia asiatica* L. PLoS ONE 12(6): e0179155. DOI: 10.1371/journal.pone.0179155

Ye Y, Cai M, Ju Y, Jiao Y, Feng L, Pan H, Cheng T, Zhang Q. 2016. Identification and validation of SNP markers linked to dwarf traits using SLAF-Seq technology in *Lagerstroemia*. PLoS ONE 11(7): e0158970. DOI: 10.1371/journal.pone.0158970

Zhang W, Mirlohi S, Li X, He Y. 2018. Identification of functional single-nucleotide polymorphisms affecting leaf hair number in *Brassica rapa*. Plant Physiol 177: 491-503. DOI: 10.1104/pp.18.00025.

Zhao D, Tao J. 2015. Recent advances on the development and regulation of flower color in ornamental plants. Front Plant Sci 6: 261. DOI: 10.3389/fpls.2015.00261