Plant Gene and Trait 2020, Vol.11, No.3, 1-8
http://genbreedpublisher.com/index.php/pgt

Research Article Open Access

Cloning and Bioinformatics Analysis of Chlorophyll Degrading Gene *PPH* from *Curcuma alismatifolia Gagnep*

Ding Huaqiao [1], Mao Lihui [1], Hu Wei [1], Dong Qing [1], Liu Jianxin [1,2]

1 Xiaoshan Cotton and Bast Fiber Crops Research Institute of Zhejiang Province, Hangzhou, 311202, China

2 Yuanpei College, Shaoxing University, Shaoxing, 311200, China

Corresponding author email: liujianxin2000@aliyun.com

Plant Gene and Trait, 2020, Vol.11, No.3 doi: 10.5376/pgt.2020.11.0003

Received: 13 May, 2020

Accepted: 22 Jun., 2020

Published: 31 Aug., 2020

Preferred citation for this article:

Ding H.Q., Mao L.H., Hu W., Dong Q., and Liu J.X., 2020, Cloning and bioinformatics analysis of chlorophyll degrading gene *PPH* from *Curcuma alismatifolia Gagnep*, Plant Gene and Trait, 11(3): 1-8 (doi: 10.5376/pgt.2020.11.0003)

Abstract To obtain key enzyme gene *PPH* in chlorophyll degradation process of *Curcuma alismatifolia Gagnep*, on the basis of obtaining a large amount of transcriptome information by sequencing the full-length transcriptome, we had screened and analyzed these transcriptome information and obtained 2 *PPH* genes which named *PPH1* and *PPH2*. The *PPH1* gene (GenBank: MT077178) has a full-length cDNA sequence of 1 795 bp in length, an open reading frame of 1437 bp (from 138 to 1 574 bp), and encode a sequence with 478AA amino acid. The *PPH2* gene (GenBank: MT077179) has a full-length cDNA sequence of 1 393bp, an open reading frame of 1227bp (from 70 to 1296 bp), and encode a sequence with 408AA amino acid. Using Blast, Translate tool (ExPASy), Clustal Omega, Find Conserved Domains (NCBI), ProtParam, TMHMM Server, SOPMA, SWISS-MODEL, ClustalX (1.81), MEGA4.1 and so on. Their amino acid composition, physical and chemical properties, conserved domains, secondary structures, tertiary crystal structures, and molecular phylogeny were predicted and analyzed. The nucleotide and protein amino acid sequences of *PPH1* and *PPH2* have high homology with these *PPH* genes of other species, and both of them contain a conserved region PLN02578 with hydrolase characteristic. Molecular phylogenetic analysis showed that a small cluster of *PPH1* and *PPH2* were closest to Musa acuminate *PPH* (XP_018677219.1), but far away from dicotyledons. This study provided a molecular basis for improving color of *Curcuma alismatifolia Gagnep* sterile bracts by genetic transformation in the future.

Keywords *Curcuma alismatifolia Gagnep*; Chlorophyll; *PPH* (pheophytinase gene); Gene

Curcuma alismatifolia Gagnep, also called ‘Siam Tulip’ or ‘Tropical Tulip’, is a perennial bulbous herbaceous flower of *Curcuma* in Zingiberaceae, which is native to Chiang Mai and other places in Thailand. The flowering period of *Curcuma alismatifolia Gagnep* is more than 3 months. In East China, the flowering period is from late July to late October. *Curcuma alismatifolia Gagnep* is a spike, the upper parts are pink broad oval sterile bracts (the color of the bracts depends on the variety, here is the color of the variety ‘Chiang Mai pink’), and the lower parts are honeycomb green fertile bracts, containing small purple and white flowers, like gentle lotuses (Ding et al., 2013). However, the green tip phenomenon of chlorophyll deposition exists at the tip of sterile bracts of most *Curcuma alismatifolia Gagnep* varieties, and there is red pigment deposition below the green tip, which is called scorching phenomenon of ‘green tip with red bottom’ (Figure 1). It is closely related to the light intensity of the environment. The phenomenon is very obvious when it is in the open air all day, while it is relatively light when in the shade. The emergence of this phenomenon makes the bracts of *Curcuma alismatifolia Gagnep* show a scorching feeling, which seriously affects the beauty and reduces the ornamental value. The green part of chlorophyll plays a very important role in the formation of ‘green tip with red bottom’. Therefore, taking measures to reduce the content of chlorophyll in bracts is conducive to inhibit or reduce the ‘green tip with red bottom’ phenomenon.

Pheophytinase (*PPH*) is a key enzyme for chlorophyll degradation and metabolism discovered recently (Eckardt, 2009; Schelbert et al., 2009; Asumi et al., 2010). *PPH* specifically cleaves phytol of pheophytin a, prompting pheophytin a to remove phytol and become phaeophorbide a. And its acting substrate is pheophytin a (Chen et al., 2017). *PPH* is a major chlorophyll dephytase during aging (Schelbert et al., 2009). In the process of leaf aging, the

expression of *PPH* gene is enhanced, which accelerates the conversion of pheophytin a to phaeophorbide a, thus accelerating leaf aging (Liu et al., 2016).

Figure 1 Scorching phenomenon of 'green tip with red bottom' of *Curcuma alismatifolia Gagnep*

PPH is commonly found in higher plant genomes, and has been obtained by cloning *Cucumis sativus* (Wang et al., 2011), *Arabidopsis thaliana* (AT5G13800), *Lolium perenne* (KT345726.1), *Guzmania* (KP723523) (Liu et al., 2016), *Camellia sinensis* (MK986828.1), *Brassica oleracea* (OL386), *Amaranthus tricolor* (KY353111.1), *Brassica rapa* (AC189212.2) and other plants. At present, there is no report about the key enzyme gene *PPH* in chlorophyll degradation and metabolism of *Curcuma alismatifolia Gagnep*. The purpose of this study is to obtain the key enzyme gene *PPH* information of chlorophyll degradation, provide a basis for improving the color of *Curcuma alismatifolia Gagnep* sterile bracts in the future, and pave the way for further enriching and exploring the theory of chlorophyll degradation and metabolism.

1 Results and Analysis

1.1 Acquisition of *PPH1* and *PPH2*

From the database, 13 pieces of information annotated as pheophytinase were screened out, of which 11 pieces were 1-2K and 2 pieces were 2-3K. Multiple sequence alignment (Clustal Omega) of nucleotides and putative amino acids was performed to eliminate repeat sequences, and then BlastN and BlastP online alignment of nucleotides and putative amino acids was performed. Finally, two high quality pheophytinase gene information were obtained, named *PPH1* and *PPH2* respectively (Figure 2).

PPH1, GenBank accession number was MT077178, the full-length cDNA sequence was 1 795 bp, ORF was 138~1 574 bp, the inferred number of protein amino acids was 478 AA, the starting codon was ATG and the ending codon was TAG. *PPH2*, GenBank accession number was MT077179, the full-length cDNA sequence was 1 393 bp, ORF was 70~1 296 bp, the inferred number of protein amino acids was 408 AA, the starting codon and ending codon were the same as *PPH1*.

The Clustal Omega comparison of the amino acid sequences of *PPH1* and *PPH2* showed that except that *PPH1* had 71 more amino acids at the starting position, most of the other amino acid sequences of the two were the same, and only 6 amino acids were different (Figure 3).

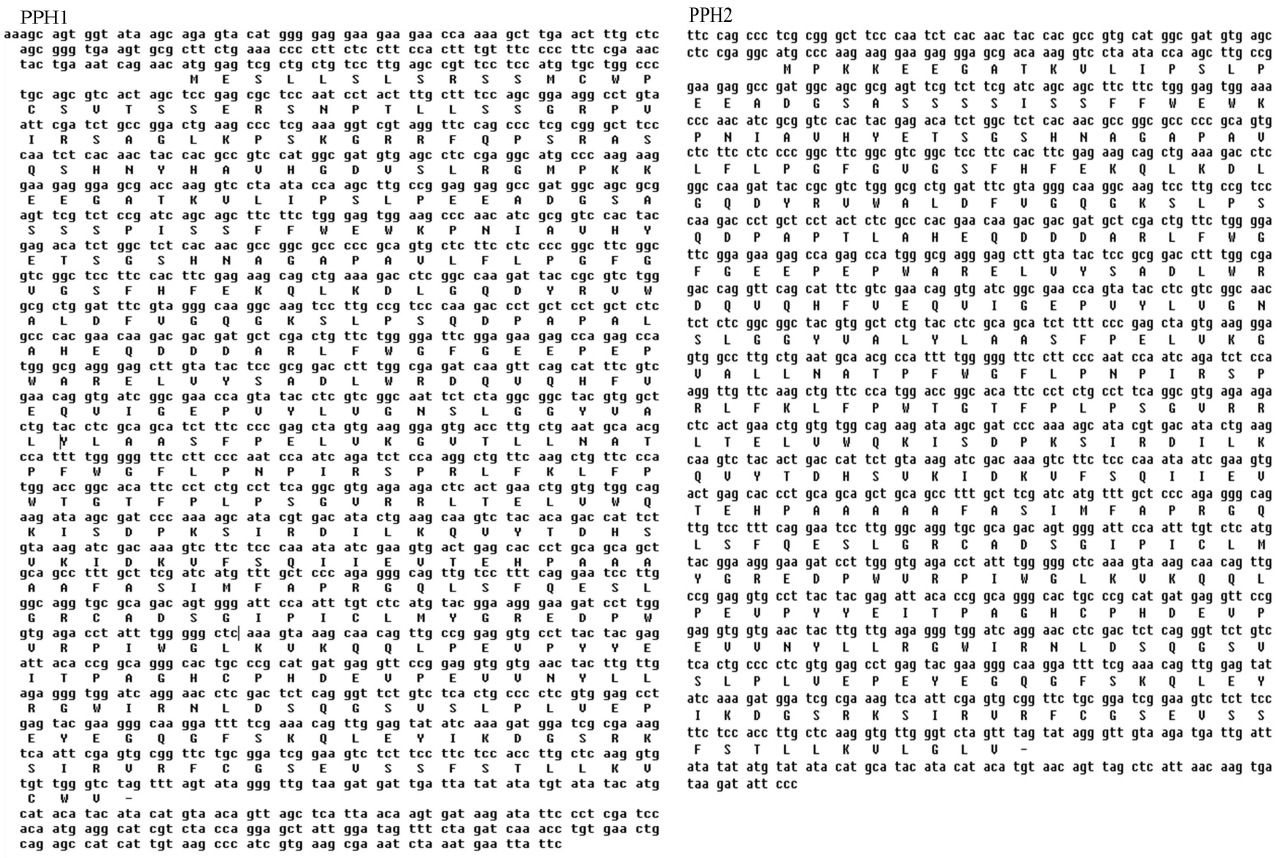

Figure 2 cDNA sequence and amino acid sequence of *PPH1* and *PPH2*

Figure 3 Alignment of amino acid sequences of *PPH1* and *PPH2*

1.2 Blast analysis of full-length cDNA and amino acid sequence

Through the BlastN alignment of full-length cDNA sequence of *PPH1* gene, it was found that *PPH1* gene had high homology with the *PPH* genes in *Musa acuminata subsp. Malaccensis*, *Phoenix dactylifera*, *Elaeis guineensis*, *Ananas comosus*, *Phalaenopsis equestris*, *Asparagus officinalis*, and *Dendrobium catenatum*. Among them, it had the highest homology with *Musa acuminata subsp. Malaccensis* (XM_018821674.1) and *Phoenix dactylifera* (XM_008790358.2, XM_008790357.2), which were 74.46% and 71.96% respectively. Then, the protein amino acid sequence of *PPH1* gene was analyzed by BlastP online comparison. It can be seen that it had the highest consistency with *Musa acuminata subsp. Malaccensis* |XP_018677219.1| and *Nelumbo nucifera* Gaertn. |XP_010269706.1|, which were 71.06% and 69.09% respectively.

Similarly, through the BlastN alignment of full-length cDNA sequence of *PPH2* gene, it was found that it had high homology with the *PPH* genes in *Musa acuminata subsp. Malaccensis, Phoenix dactylifera, Ananas comosus, Elaeis guineensis, Phalaenopsis equestris, Asparagus officinalis, Setaria italica* and *Dendrobium catenatum*. Among them, it had the highest homology with *Musa acuminata subsp. Malaccensis* (XM_018821674.1) and *Elaeis guineensis* (XM_010923945.3), which were 74.33% and 72.46% respectively. Then, the protein amino acid sequence of *PPH2* gene was analyzed by BlastP online comparison. It can be seen that it had the highest consistency with *Musa acuminata subsp. Malaccensis* |XP_018677219.1| and *Ananas comosus* |XP_020114200.1|, which were 71.88% and 70.65% respectively.

1.3 Characteristic analysis of putative protein amino acids

Using online ProtParam to predict and analyze the physical and chemical properties of the putative *PPH1* amino acid sequence, its relative molecular mass was 53.25 kD, the positively and negatively charged amino acid residues were 48 (Arg + Lys) and 51 (Asp+Glu), and the isoelectric point pI was 6.48. TMHMM Server v. 2.0 was used for online transmembrane prediction, which showed that there was no transmembrane structure and it was an outside protein. Using SOPMA to predict its secondary structure online, it can be seen that it was mainly composed of 16.11% extend strand, 32.43% alpha helix, 48.33% random coil and 12.19% beta turn.

Using online ProtParam to predict and analyze the physical and chemical properties of the putative *PPH2* amino acid sequence, its relative molecular mass was 45.52 kD, the positively and negatively charged amino acid residues were 38 (Arg+Lys) and 48 (Asp+Glu), and the isoelectric point pI was 5.46. TMHMM Server v. 2.0 was used for online transmembrane prediction, which showed that there was no transmembrane structure and it was an outside protein. Using SOPMA to predict its secondary structure online, it can be seen that it was mainly composed of 13.48% extend strand, 37.50% alpha helix, 41.91% random coil and 7.11% beta turn. The conserved regions of *PPH1* and *PPH2* were analyzed on NCBI by using Specialized Blast tool (Lu et al., 2020). It was found that both of them contained conserved region PLN02578 (Aron et al., 2017), which was with hydrolase characteristic (Lu et al., 2020). The protein amino acid sequence of *PPH1* was submitted to SWISS-MODEL (https://swissmodel.expasy.org/) and the template was searched in the protein structure library. Finally, it was found that the consistency of the 96th~422nd residues in protein sequence of *PPH1* and the A-chain sequence of human soluble bifunctional cyclohydrolase C-terminal domain 6i5 e.1 had reached 21.13%. Then, the A-chain of 6i5e.1 was used as a template to predict the three-dimensional model of PPH1 protein (Abis et al., 2019) (Figure 4A). *PPH2* was subjected to the same treatment, and it was found that the consistency of the 27th~352nd residues in protein sequence of *PPH2* and the A-chain sequence of soluble epoxide hydrolase 3wk4.1 had reached 21.58% (Amano et al., 2014). Then, it was used as a template to predict the three-dimensional model of PPH2 protein (Figure 4B).

1.4 Molecular phylogenetic analysis

The obtained amino acid sequences of *PPH1*, *PPH2* and *PPH* of other species were compared with multiple amino acid sequences by ClustalX(1.81) software, and then the molecular system tree (MEGA4.1 software) was constructed (Figure 5). The results showed that *PPH1* and *PPH2* were clustered into a small group, and then they had the closest genetic relationship with *Musa acuminata subsp. Malaccensis* (XP_018677219.1), clustering into one group. And then they clustered into a big group with other monocotyledons such as *Brachypodium distachyon*

(XP_010227493.1), *Elaeis guineensis* (XP_010931560.1, XP_010922247.1), *Panicum miliaceum* (RLM73077.1), *Setaria italica* (XP_004977026.1), *Asparagus officinalis* (XP_020254622.1), *Ananas comosus* (OAY84910.1, XP_020114200.1), *Phalaenopsis equestris* (XP_020582283.1), *Dendrobium catenatum* (XP_020672292.1), *Musa acuminata subsp. Malaccensis* (XP_018677219.1) and *Phoenix dactylifera* (XP_008788579.1). But they were distinguished from dicotyledons such as *Nelumbo nucifera* Gaertn. (XP_010269706.1, XP_010279530.1) and *Cinnamomum micranthum* (RWR81325.1).

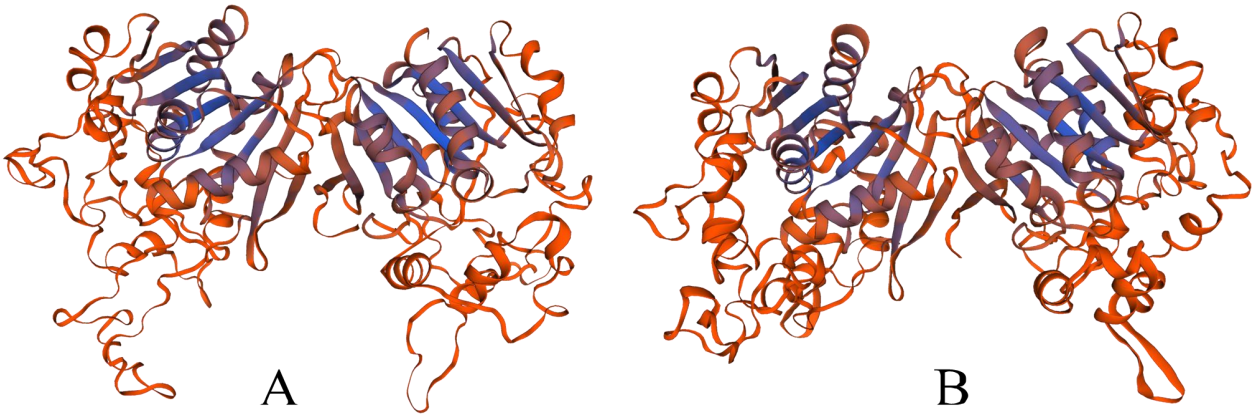

Figure 4 Three-dimensional model of PPH protein
Note: A: PPH1; B: PPH2

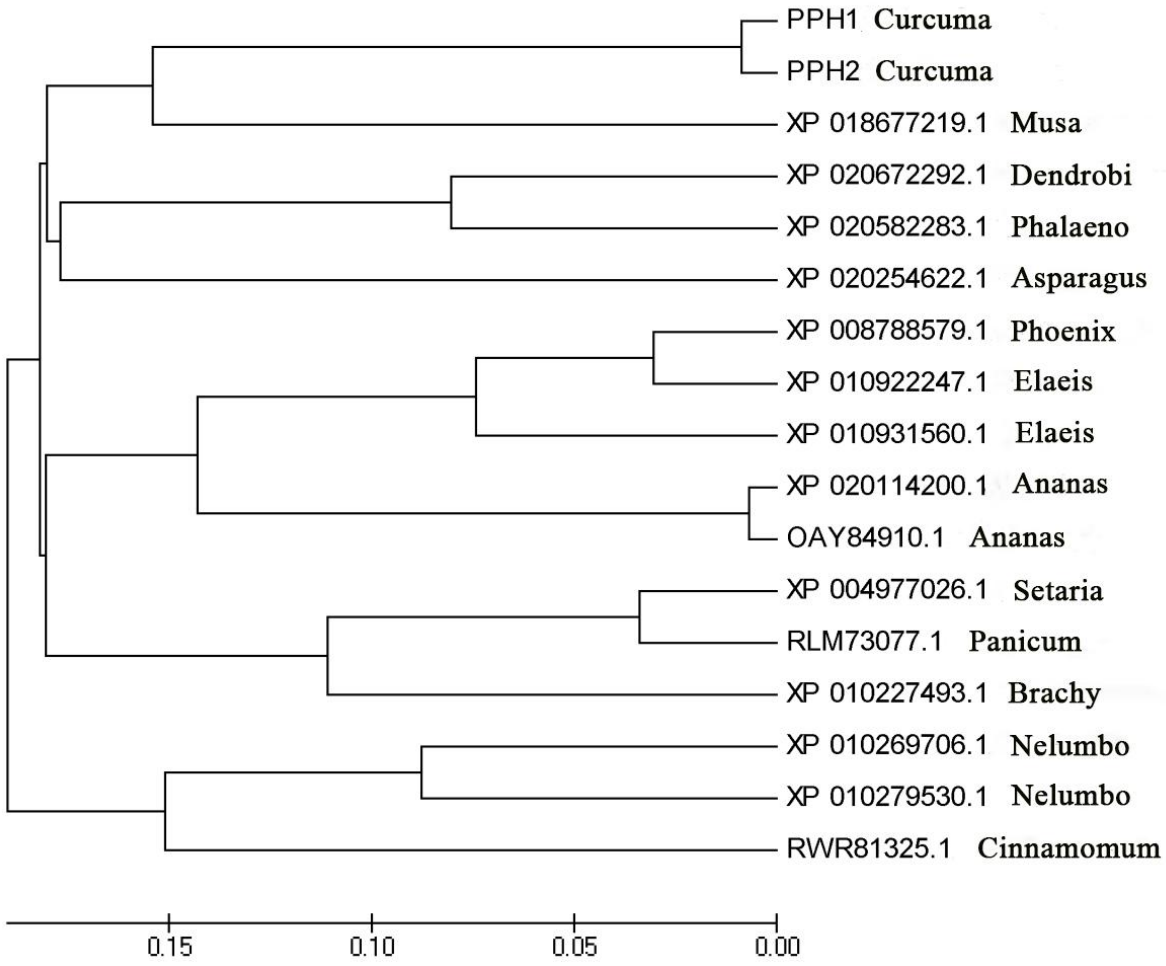

Figure 5 Molecular evolution analysis of *PPH* gene

2 Discussion

Pheophytinase is a key enzyme for chlorophyll degradation and metabolism newly discovered in *Arabidopsis thaliana* by Schelbert et al. (2009), Zhang (2017). *PPH* is a candidate enzyme involved in the porphyrin-phytol hydrolysis process. It is very active for pheophytin, but has no effect on chlorophyll. The role of *PPH* enzyme is to convert pheophytin a to phaeophorbide a. Its discovery rewrote the process of chlorophyll degradation. Previously, it was always believed that phytol was removed first and then magnesium was removed in the process of chlorophyll degradation, but recently with the deepening of research on pheophytinase, it has been believed that magnesium may be removed first and then phytol was removed (Tian et al., 2010). The substrate of the deplantation reaction was corrected from chlorophyll a to pheophytin a, that is, magnesium was removed earlier than phytol (Eckardt, 2009; Tian et al., 2010; Tang and Mao, 2011), while chlorophyllase was not necessary (Schelbert et al., 2009).

There are few reports on *PPH* gene expression, only in *Arabidopsis thaliana*, *Brassica oleracea*, *Cucumis sativus* and *Guzmania lingulata* (Liu, 2016). According to BLASTP search (NCBI) for homologous proteins, PPH proteins were widely distributed in algae and land plants and played a very important role (Schelbert et al., 2009; Agustin et al., 2010). All PPH proteins had a conserved region with hydrolase characteristic, so did *PPH1* and *PPH2* in this study, indicating that PPH proteins among species were highly conserved. As for the prediction of protein tertiary crystal structure, different from Luzia et al. (2018) using Protein Homology/analogY Recognition Engine V 2.0 (Phyre2) method and using RCSB PDB as a template to construct the tertiary crystal structure of *PPH* from *Arabidopsis thaliana*, we used the A-chain of cyclohydrolase 6i5e.1 and the A-chain of 3wk4.1 as templates to structure the tertiary crystal structures of *PPH1* and *PPH2* from *Curcuma alismatifolia Gagnep* by SWISS-MODEL method. There were some differences in the morphological structure between the model constructed in this study and that constructed by Luzia et al. (2018). Similarly, Luzia et al. (2018) also conducted molecular system analysis of *PPH* between *Arabidopsis thaliana* and other species. The results showed that it was only closely related with three PPH-like proteins of *Arabidopsis thaliana*, and the consistency with other PPH proteins was low, ranging from 25.1% to 31.1%. In this study, *PPH1* and *PPH2* from *Curcuma alismatifolia Gagnep* were clustered into a small group. Among them, *PPH1* and *PPH2* had the highest genetic relationship with *PPH* from *Musa acuminata subsp. Malaccensis* (XP_018677219.1) (71.06% and 71.88% respectively), but far away from dicotyledons.

3 Materials and Methods

3.1 Test materials

The full-length transcriptome database of flowering plants of *Curcuma alismatifolia Gagnep* variety 'Chiang Mai pink' was used as test materials. 'Chiang Mai pink' was the main *Curcuma alismatifolia Gagnep* variety for garden application and cut flower application in the market. The materials were taken from the experimental base of Xiaoshan Cotton and Bast Fiber Crops Research Institute. Corresponding author Liu Jianxin took the stems, leaves, sterile bracts, fertile bracts and small flower bracts of 'Chiang Mai pink' as mixed materials in 2015 and sequenced the full-length transcriptome (PBIsoSeq) based on Pacbio method. A total of 64 471 transcriptome information of *Curcuma alismatifolia Gagnep* were obtained and annotated. The relevant data were published in another paper.

3.2 Acquisition of *PPH1* and *PPH2*

According to the annotation information of transcriptome database, the information annotated as pheophytinase was found, and then sequence analysis, comparison, screening and PCR verification were carried out, and finally *PPH1* and *PPH2* genes were obtained.

3.3 Bioinformatics analysis of *PPH1* and *PPH2* gene

Using bioinformatics methods, the obtained cDNA nucleotide sequences, the primary structure, secondary structure and tertiary structure of the protein amino acids of *PPH1* and *PPH2* genes were studied and analyzed. Specifically, nucleotide alignment (BlastN), multiple sequence alignment (Clustal Omega), protein amino acid alignment (BlastP), protein amino acid sequence prediction (ExPASy's Translate tool), physical and chemical

characteristic analysis (ProtParam), conserved region search (NCBI's Find conserved domains), protein transmembrane structure prediction (TMHMM Server v. 2.0), secondary structure prediction (SOPMA) (Combet et al., 2000), and tertiary crystal structure model prediction (SWISS-MODEL) (Lambert et al., 2002) were carried out. Finally, the molecular evolution analysis was performed by using BlastP, MEGA4.1 and ClustalX_(1.81) programs. That is, firstly, the protein amino acid sequence information of *PPH* of other species with high consistency was obtained by BlastP online alignment, then the obtained sequences were performed multiple sequence alignment, and finally the genetic relationship was analyzed by generating molecular system tree.

Authors' contributions

DHQ was responsible for writing this manuscript; MLH, HW and DQ participated in the design and revision of the manuscript; LJX completed the previous transcriptome sequencing and guided the design of the manuscript. All authors read and approved the final manuscript.

Acknowledgments

This study was funded by Provincial Key Research and Development Program of Zhejiang (2019C02025).

References

Abis G., Charles R.L., Kopec J., Yue W.W., Atkinson R.A, Bui T.T., Lynham S., Popova S., Sun Y.B., Fraternali F., Eaton P., and Conte M.R., 2019, 15-deoxy-Delta12,14-Prostaglandin J2inhibits human soluble epoxide hydrolase by a dual orthosteric and allosteric mechanism, Commun Biol., 2(188): 1-14
https://doi.org/10.1038/s42003-019-0426-2
PMid:31123712 PMCid:PMC6525171

Agustin M.B., Pedro M.C., Gustavo A., and Martinez., 2010, Chlorophyllase versus pheophytinase ascandidates for chlorophyll dephytilation during senescence of broccoli, J Plant Physiol., 7: 1-7

Amano Y., Yamaguchi T., Tanabe E., 2014, Structural insights into binding of inhibitors to soluble epoxide hydrolase gained by fragment screening and X-ray crystallography, Bioorg.Med.Chem., 22(8): 2427-34
https://doi.org/10.1016/j.bmc.2014.03.001
PMid:24656800

Aron M.B., Yu B., Lianyi H., Jane H., Lanczycki C.J., Shennan L., Farideh C., Myra K.D., Renata C.G., Noreen R.G., Marc G., David I.H., Fu L., Gabriele H.M., James S.S., Narmada T., Zhouxi W., Roxanne A.Y., Dachuan Z., Chanjuan Z., Lewis Y.G., and Stephen H.B., 2017, "CDD/SPARCLE: functional classification of proteins via subfamily domain architectures.", Nucleic Acids Res., 45(D): 200-203
https://doi.org/10.1093/nar/gkw1129
PMid:27899674 PMCid:PMC5210587

Asumi F., Yasuo S., Hirofumi T., and Naoki Y., 2010, Effects of postharvest ethanol vapor treatment on activities and gene expression of chlorophyll catabolic enzymes in broccoli florets, Postharvest Biology and Technology, 2010, 55: 97-102
https://doi.org/10.1016/j.postharvbio.2009.08.010

Chen J.Y., Zhu X.Y., Ren J., Qiu K., Li Z.P., Z.K., Gao J., Zhou X., and Kuai B.K., 2017, Suppressor of overexpression of CO 1 negatively regulates dark-induced leaf degreening and senescence by directly repressing pheophytinase and other senescence-associated genes in Arabidopsis, Plant Physiol., 173(3): 1881-1891
https://doi.org/10.1104/pp.16.01457
PMid:28096189 PMCid:PMC5338665

Combet C., Blanchet C., Geourjon C., and Deléage G., 2000, NPS@: network protein sequence analysis, TIBS 25: 147-150
https://doi.org/10.1016/S0968-0004(99)01540-6

Ding H.Q., Liu J.X., Wang W.Y., and Zou Q.C.,2013, Dwarfing effect of different plant growth retarders on *Curcuma alismatifolia*, Zhejiang Nongye Kexue (Zhejiang Agricultural Sciences), 5: 559-562

Eckardt N.A., 2009, A new chlorophyll degradation pathway, Plant Cell, 21: 700
https://doi.org/10.1105/tpc.109.210313
PMid:19304932 PMCid:PMC2671716

Lambert C., Leonard N., De Bolle X., and Depiereux E., 2002, ESyPred3D: Prediction of proteins 3D structures, Bioinformatics, 18: 1250-1256
https://doi.org/10.1093/bioinformatics/18.9.1250
PMid:12217917

Liu J.X., Ding H.Q., Tian D.Q., Wang W.Y., and Liu H.C., 2016, Isolation of Chlorophyll metabolism key genes and molecular mechanism of green fade in guzmania bracts discoloration process, Zhongguo Nongye Kexue (Scientia Agricultura Sinica), 49(13): 2593-2602

Lu S., Wang J., Chitsaz F., Derbyshire M.K., Geer R.G., Gonzales N.R., Gwadz M., Hurwitz D.I., Marchler G.H., Song J.S., Thanki N., Yamashita R.A., Yang M., Zhang D., Zheng C., Lanczycki C.J., Marchler-Bauer A., 2020, CDD/SPARCLE: the conserved domain database in 2020, Nucleic Acids Research, 48(D1):265-268
https://doi.org/10.1093/nar/gkz991
PMid:31777944 PMCid:PMC6943070

Luzia G., Kathrin S., Undine K., and Stefan H., 2018, Catalytic and structural properties of pheophytinase, the phytol esterase involved in chlorophyll breakdown, J Exp Bot., 69(4): 879-889
https://doi.org/10.1093/jxb/erx326
PMid:29036670 PMCid:PMC5853334

Schelbert S., Aubry S., Burla B., Agne B., Kessler F., and Krupinska K., 2009, Pheophytin pheophorbide hydrolase (pheophytinase) is involved in chlorophyll breakdown during leaf senescence in *Arabidopsis*, Plant Cell, 21: 767-785
https://doi.org/10.1105/tpc.108.064089
PMid:19304936 PMCid:PMC2671698

Tang L., and Mao Z.G., 2011, Degradation Pathway of Plant Chlorophyll and Its Molecular Regulation. Zhiwu Shengli Xuebao (Plant Physiology Journal), 47 (10): 936-942

Tian F.X., Hui Z., Wang G.K., Fan Z.Y., and Wang W., 2010, Chlorophyll degradation and stay-green mutant in plant, Zhiwu Shenglixue Tongxun (Plant Physiology Communications), 46(5): 505-511

Wang W., Xu Y.J., and Wan Z.J., 2011, Cloning and expression analysis of key genes *PPH* and PAO for chlorophyll degradation in cucumber, Yuanyi Xuebao (Acta Horticulturae Sinica), 38(6): 1104-1110

Zhang W., 2017, The relationship between chlorophyll metabolism and tocopherol biosynthesis in *Arabidopsis*, Dissertation for PhD, Huazhong Agricultural University, Supervisor: Zhang C.Y., pp.21