Plant Molecular Biology Reporter (2019) 37:277–290
https://doi.org/10.1007/s11105-019-01156-3

ORIGINAL PAPER

# Morphological Characteristics and Gene Mapping of Purple Apiculus Formation in Rice

Yohannes Tsago[1,2] · Zhongkang Wang[1] · Jialin Liu[1] · Mustapha Sunusi[1] · Jamal Eshag[1] · Delara Akhter[1] · Chunhai Shi[1] · Xiaoli Jin[1]

Published online: 16 July 2019


## Abstract

Apiculus color of grain is an important trait which is used as a morphological marker in rice (*Oryza sativa*. L). In the present study, the purple apiculus mutant named as *Ospa* was developed from an *indica* cultivar using ethyl methanesulfonate mutagenesis. The *Ospa* mutant showed increased grain size, thousand-grain weight, and anthocyanin accumulation compared with the wild-type (WT). Histological analysis revealed that the size and number of cells in parenchyma layers of spikelet hulls was significantly higher in *Ospa* mutant than the WT. By map-based cloning, *OsPA* was located within the 60.74 kb region on the long arm of chromosome 1 where a T to C substitution was detected in the third exon of *LOC_Os01g39580*. The encoded polypeptide, predicted as anthocyanin regulatory *Lc* protein, contained the basic-helix-loop-helix (bHLH) domain. The mutation has changed cysteine to arginine in the amino acid sequence and modified the predicted secondary structure in the conserved bHLH domain of the encoded peptide. As the RT-qPCR analysis revealed, *OsPA* was specifically expressed in the young panicles of *Ospa* mutant. The *OsPA* gene might have regulated the grain anthocyanin content by up-regulating the expression of the anthocyanin regulatory MYB gene (*OsC*), and the anthocyanin biosynthesis genes (*OsF3'H*, *OsF3'5 'H*, *OsDFR*, and *OsANS*). Our study will provide new genetic material for a further study of the genetic mechanism of tissue-specific pigmentation, and the developed gene-based markers could be utilized in marker-assisted selection along with purple apiculus trait in rice.

**Keywords** Gene mapping · Phenotype analysis · Rice · Purple apiculus

## Introduction

Rice is grown on all the continents in the world except Antarctica and makes up the majority of the source of calories for half the world's population. Different anthocyanin pigments appear in photosynthetic and reproductive tissues and organs in rice. Diverse grades of purple pigmentation mostly appear in the apiculus, stigma, leaf blade, root, leaf sheath, and

**Electronic supplementary material** The online version of this article (https://doi.org/10.1007/s11105-019-01156-3) contains supplementary material, which is available to authorized users.

✉ Xiaoli Jin
jinxl@zju.edu.cn

Yohannes Tsago
yohannets@gmail.com

Zhongkang Wang
nwswangzhongkang@vip.qq.com

Jialin Liu
864400183@qq.com

Mustapha Sunusi
sunusimustapha10@yahoo.com

Jamal Eshag
jamaladam@zju.edu.cn

Delara Akhter
11516090@zju.edu.cn

Chunhai Shi
chhshi@zju.edu.cn

[1] Department of Agronomy, the Key Laboratory of Crop Germplasm Resource of Zhejiang Province, Zhejiang University, Hangzhou 310058, Zhejiang, China

[2] Department of Biology, Madda Walabu University, Bale Robe, Ethiopia

root. Anthocyanin accumulation in various tissues plays a key role in the stimulation of hormonal responses, protection from ultraviolet radiation, and biotic stress resistance and abiotic stresses tolerance (Chalker-Scott 1999; Ithal and Reddy 2004; Lin-Wang et al. 2010; Reddy et al. 1995; Saitoh 2004). Purple apiculus trait is commonly seen not only in wild species but also frequently observed in landraces and cultivars of rice. Purple apiculus attracts insects and animals for pollination and seed dispersal (Chin et al. 2016). It is also an important morphological marker for varietal selection and purity determination (Saitoh 2004).

Anthocyanins are pigment-producing secondary metabolite molecules classified as part of flavonoid (Reddy et al. 1995). The genetic basis of anthocyanin pigmentation has been widely studied in the past and several genes have been identified in rice, maize, Petunia, and model plants like Arabidopsis (Dooner et al. 1991; Brenda 2001; Koes et al. 2005). The expression and function of genes encoding enzymes that directly catalyze the conversion of intermediates in the anthocyanin biosynthesis pathway is modulated by numerous activators and transcription factors. The transcriptional activator genes *C1* and *Pl* that encode proteins with MYB DNA binding domains and *B* and *R* genes that code for proteins containing a basic helix-loop-helix (bHLH) domain are transcriptional activators and regulate downstream genes of the anthocyanin metabolism pathway in maize (Chandler et al. 1989). The *OsC*, *OsB1*, and *OsB2* genes were identified from rice based on sequence similarity with maize *C1* and *B* genes and regulated pigmentation pattern in different tissues (Reddy et al. 1995; Sakamoto et al. 2001). In addition to the *C* and *B* genes, the tissue-specific pigmentation patterns in rice were controlled by *R*, *B*, *Lc*, and *Sn* genes (Ludwig and Wessler 1990). Other genes like *Pl*, *Rc*, and *Rd* regulated purple leaf and red pericarp formation; whereas *kala4* is responsible for black rice trait (Furukawa et al. 2007; Gao et al. 2011; Liu et al. 2012; Maeda et al. 2014; Oikawa et al. 2015; Sakamoto et al. 2001). In addition to *Rc* and *Rd*, *Kala1*, and *Kala3* genes were responsible for the formation of the purple pericarp (Maeda et al. 2014; Reddy et al. 1995). A sequence rearrangement in the promoter of *Kala4* resulted in for the formation of the purple pericarp trait in rice (Oikawa et al. 2015). The red pericarp trait was regulated by *Rc* gene that encodes bHLH together with functional *Rd* gene that encodes dihydroflavonol reductase (Furukawa et al. 2007; Sweeney et al. 2006), also, the *pl* gene which was responsible for the purple coloration of the leaf encodes bHLH protein (Sakamoto et al. 2001). The bHLH domains of regulatory genes make protein–protein interaction with *OsC* that has an R2R3 MYB domain and activated the expression of the structural anthocyanin biosynthesis genes that directly participate in the anthocyanin production pathway (Koes et al. 2005; Sakamoto et al. 2001).

Basically, anthocyanin pigmentation required network of *C* (chromogen), *A* (activator), and *P* (purple, tissue-specific regulator), and other genes that determine tissue-specific distribution of coloration (Fan et al. 2007; Mikami et al. 2000; Saitoh 2004; Sakamoto et al. 2001; Sun et al. 2018; Zhao et al. 2016). Anthocyanin pigmentation in the grain of rice is dependent on a single-copy gene, *OsC* (chromogen in rice) in the presence of other functional activators, tissue-specific regulators, and structural anthocyanin biosynthesis genes (Fan et al. 2007; Mikami et al. 2000; Saitoh 2004; Zhao et al. 2016). Functional bHLH was required for activation of the structural anthocyanin biosynthesis gene expression together with *OsC* in rice (Sun et al. 2018). In the past, the genetic basis of apiculus coloration was widely investigated (Fan et al. 2007; Mikami et al. 2000; Saitoh 2004; Sakamoto et al. 2001; Sun et al. 2018; Takahashi 1957, 1982; Zhao et al. 2016). Apart from the basic gene *OsC*, the responsible activator and tissue-specific regulator genes for purple apiculus trait were not fully identified. In the present study, a new purple apiculus mutant, *Ospa*, was generated from wild-type (WT), Zhenong 34 (*O. sativa* L. ssp. *indica*) by ethyl methanesulfonate (EMS) mutagenesis. The objectives of the study were (1) to analyze the phenotypic differences between *Ospa* mutant and WT, (2) to map the candidate gene (*OsPA*) using the map-based cloning strategy, and (3) to develop molecular markers for *OsPA* that will assist molecular breeding in rice.

## Materials and Methods

### Plant Materials and Growth Condition

Rice (*Oryza sativa* L.) purple apiculus mutant was obtained from an $M_2$ population of the *indica* cultivar Zhenong 34 after EMS mutagenesis. For phenotypic segregation analysis and genetic mapping, $F_1$ plants were developed by crossing the *Ospa* mutant (female parent) and green apiculus Zhenongda 104 (male parent). Then, the $F_1$ plants were self-crossed to generate $F_2$ population. All the parents and 2302 $F_2$ individuals were cultivated in the paddy field at Zhejiang University in Hangzhou, China (N 30° 15′ 49″,E 120 7′15″) during the natural growing season.

### Measurement of Phenotypic Traits

The wild-type (Zhenong 34) and *Ospa* mutant were planted in the paddy field at the same growing season. Agronomic traits such as plant height (PH), panicle number per plant (PN), grain number per panicle (GN), seed setting rate (SR), thousand-grain weight (TGW), grain length (GL), grain width

(GW), grain thickness (GT), brown rice length (BRL), brown rice width (BRW), and brown rice thickness (BRT) were measured during maturity.

## Physiological Analysis

The grain anthocyanin content was recorded in the 5th, 10th, 15th, 20th, 25th, 30th, and 35th day after heading (DAH). Total grain anthocyanin contents were extracted from fresh matured hull samples following the procedures described by Yang et al. (2018) with minor modification. In brief, anthocyanin was isolated using 1 mL of 1% acidic methanol (HCl/methanol, 1:99) solution. 1 g of finely ground mature hull tissue was incubated at 25 °C in the dark. The extracted solution was centrifuged at 12,000×*g* rpm for 20 min. The absorbance values of the supernatant were measured against the reagent blank at 530 and 657 nm with spectrophotometer UV-1800 (Shanghai, China), and total anthocyanins were quantified based on the absorption of the extracts using the equation:$Q_{Anthocyanins} = (A530 - 0.25 \times A657) \times FW^{-1}$, where $Q_{Anthocyanins}$ is the quantity of total anthocyanins; A530 and A657 are the absorptions levels at 530 nm and 657 nm, respectively, and FW is the weight of hull tissue (g). Anthocyanins were quantified as absorption levels in three independent biological replicates.

Total flavonoids were measured in the 5th, 10th, 15th, 20th, 25th, 30th and 35th DAH by the aluminum method previously reported by Shen et al. (2009) with slight modification. Mature hulls (1 g) ground with liquid nitrogen were mixed with 25 mL of methanol containing 1% HCl at 25 °C in the dark. The methanolic extracts were centrifuged at 4,000×*g* rpm for 5 min, and the supernatant was separated. Aliquots (0.5 mL) of extracts or standard solutions were transferred into 15 mL conical centrifuge tubes, and 2 mL double distilled $H_2O$ and 0.15 mL 5% $NaNO_2$ were added. Then, 0.15 mL 10% $AlCl_3.6H_2O$ and 1 mL 1 M NaOH were added in 5 min interval respectively. After 15 min, total flavonoids content was recorded in the form of absorbance values by spectrophotometer UV-1800 (Shanghai, China) at the wavelength of 415 nm, and the measurement was repeated three times. Total flavonoid content was calculated using the standard quercetin curve prepared by serial dilution of quercetin and expressed as μg quercetin equivalent (μg QE) per g of fresh weight.

## Histological Analysis

The cross-sectional analysis was carried out to investigate the cellular anatomy of the spikelet hulls. The double fixation was carried out by excising tissue samples from the young spikelet hulls. The samples were fixed with 2.5% glutaraldehyde in phosphate buffer (0.1 M, pH 7.0) for 4 h and washed three times in the phosphate buffer for 15 min at each step followed by post-fixation with 1% $OsO_4$ in phosphate buffer for 2 h. Consecutively, the specimen was first dehydrated with 30%, 50%, 70%, 80%, 90%, 95%, and 100% ethanol for 20 min at each step and transferred to absolute acetone for 20 min. The infiltration process was carried out by placing the specimen in 1:1 and 1:3 mixture of absolute acetone and Spurr resin mixture for 1 h and 3 h respectively at 25 °C. Then, the samples were placed in Spurr resin mixture overnight. Embedding was done by placing the specimen in Eppendorf contained Spurr resin and heating at 70 °C for 9 h. The specimen was sectioned in LEICA EM UC 7 Ultratome and stained by uranyl acetate alkaline lead citrate for 5 and 10 min respectively. The cross section was observed in transmission electron microscope Hitachi Model H-7650 TEM (Hitachi, Japan). The cell number in the outer parenchyma layer was analyzed using the multipoint tool of the ImageJ program version 1.51(National Institute of Health, USA).

For investigating the inner parenchyma layers cells and tubercles in the outer layer, the tissue samples from the young spikelet hull were fixed with 2.5% glutaraldehyde in phosphate buffer (0.1 M, pH 7.0) for 4 h and washed three times in the phosphate buffer for 15 min at each step. Subsequently, the samples were post-fixed with 1% $OsO_4$ in the phosphate buffer for 2 h and washed three times in the phosphate buffer for 15 min at each step. Following post-fixation, the samples were dried by 30%, 50%, 70%, 80%, 90%, 95%, and 100% ethanol for 20 min at each step and transferred to the mixture of alcohol and iso-amyl acetate (1:1, *v*:*v*) for 30 min followed by pure iso-amyl acetate overnight. Finally, the samples were dried in Model HCP-2 critical point dryer (Hitachi, Japan) with liquid $CO_2$, coated with gold-palladium and observed in scanning electron microscope TM-1000 SEM (Hitachi, Japan). Cell area in the inner parenchyma layer was analyzed using the ImageJ program version 1.51.

## Mapping and Cloning of the *OsPA* Gene

The DNA isolation was carried out from fresh rice leaves (Murray and Thompson 1980). The candidate gene for *Ospa* mutant was located by fine mapping after bulked segregant analysis (BSA) (Michelmore et al. 1991). BSA was carried out using the DNA isolated from the female parent (*Ospa* mutant) and male parent (Zhenongda 104 with WT trait) and DNA pooled from 20 individuals with purple apiculus and 20 individuals with green apiculus respectively. A total of 400 SSR markers covering all the 12 chromosomes of rice were employed for mapping the candidate region. The candidate markers were selected based on the polymorphism of the parents, $F_2$ mutant pool, and $F_2$ wild-type pool taking in to account the similarity of $F_2$ mutant pool to the female parent (*Ospa* mutant). For genetic mapping, 10 InDel markers were developed by comparing the diversity of *indica* and *japonica* rice DNA sequences from the gramene database (Gupta et al.

2016; Tello-Ruiz et al. 2018). Analyzing recombination events in the respective positions of the newly designed InDel markers (Table S1) between the WT and *Ospa* mutant individuals of $F_2$ together with the parental lines, the mapping region was further delimited. For each marker designed, the candidate region was amplified with PCR cycle of denaturation at 95 °C for 5 min, followed by 35 cycles of 94 °C for 30 s; annealing at 55 °C for 30 s, followed by 72 °C for 30 s, and elongation step at 72 °C for 10 min. The PCR results were analyzed by polyacrylamide gel (8%) electrophoresis.

Analyzing the predicted function of the candidate ORFs in the delimited region (Table S2) using the Rice Genome Annotation Project (Kawahara et al. 2013), primers were designed for the candidate gene PCR amplification and sequencing (Table S1). The candidate gene was PCR amplified with cycles of denaturation at 95 °C for 5 min, followed by 35 cycles of 98 °C for 15 s; annealing at 55 °C for 30 s followed by 68 °C for 30 s, and elongation at 72 °C for 10 min. Agarose gel (1%) was used to visualize amplification result.

### Structural and Functional Prediction of OsPA Protein Sequence

The secondary structure of the encoded protein was predicted based on the ResQ approach (Yang et al. 2016), and the function prediction of the OsPA protein was carried out based COFACTOR (Zhang et al. 2017) and COACH prediction (Yang et al. 2013a, b). The peptide model for the conserved domain where the mutation occurred was produced using the SWISS-MODEL program (Waterhouse et al. 2018). For sequence alignment, the protein sequence of OsPA was obtained from the rice genome annotation project database (Kawahara et al. 2013), and orthologous proteins were obtained using the protein basic local alignment search tool (Altschul et al. 1990). Protein family database (El-Gebali et al. 2019) and the conserved domain database (Sayers et al. 2019) programs were used to perform a domain analysis. The obtained sequences were compared with rice OsPA protein sequence using the Clustal Omega program (Sievers et al. 2011) with default settings and were viewed in the Jalview program version 2 (Waterhouse et al. 2009).

### Phylogenetic Analysis

To analyze the evolutionary relationship and functional association between orthologous proteins of OsPA, additional sequences were obtained by the protein basic local alignment search tool (Altschul et al. 1990). The alignment of protein sequences was carried out and the phylogenetic tree was created using MEGA X program (Kumar et al. 2018). Neighbor-joining (NJ) method and a bootstrap test were performed with 1000 replicates to evaluate the support of clusters and nodes in the phylogenetic tree.

### RNA isolation and real-time quantitative PCR analysis.

For analyzing the expression profile, total RNA was isolated from seedlings, culm, and root of *Ospa* mutant and WT. In addition, total RNA was extracted from the panicles at 2, 7, 12, 17, and 22 days after booting (DAB) for studying the time-course of gene expression. To evaluate the expression level of structural anthocyanin biosynthesis genes, total RNA was isolated from the panicle of *Ospa* mutant and WT at 17 DAB. Total RNA isolation was carried out using Trizol (Invitrogen, USA) following the manufacturer's instruction. The Prime Script RT reagent Kit with gDNA Eraser and SYBR premix ExTaq™ (Takara, Japan) was used for quantitative real-time PCR (qRT-PCR) (Takara, Japan). One microgram RNase-free DNase I treated total RNA was utilized for the first-strand cDNA synthesis in a 20 μL reaction volume. Quantitative RT-PCR was performed using the SYBR Premix ExTaq™ (Tli RNaseH Plus). The *OsACTIN1* was used as an internal control for the RT-PCR analysis. The RT-PCR was carried out with activation at 95 °C for 30 s, followed by 40 cycles of denaturation at 95 °C for 5 s, and annealing and extension at 60 °C for 30 s. The primers used in the qRT-PCR analysis were listed in Table S3.

## Results

### Phenotypic Characterization of *Ospa* Mutant

The morphological and agronomic traits of WT and *Ospa* mutant were measured at maturity under field conditions. The *Ospa* mutant plants exhibited purple coloration in the apiculus, while the WT has green apiculus (Fig. 1a, b). Traits difference between WT and *Ospa* mutant was observed only in the grains and a significant increase in GL, GW, and TGW in *Ospa* mutants compared with those of WT (Table 1). There was no difference in PH, PL, PN, GN, GT, BRL, BRW, BRT, and SR between WT and *Ospa* mutants (Table 1).

### Physiological Changes of *Ospa* Mutant

Related to the purple color formation in the apiculus, grain anthocyanin content was measured in WT and *Ospa* mutant. Significantly higher anthocyanin content was recorded in the *Ospa* mutant hulls than the WT ($\alpha = 0.01$) (Fig. 1c). However, a significant difference ($\alpha = 0.05$) was not seen for the grain flavonoid content between the *Ospa* mutant and WT (Fig. 1d). It was notable that the differences between WT and *Ospa* mutant might be limited only to the grains and grain composition. The anthocyanin and grain flavonoid contents were also analyzed during the grain development stages. Anthocyanin contents increased dramatically in the *Ospa*

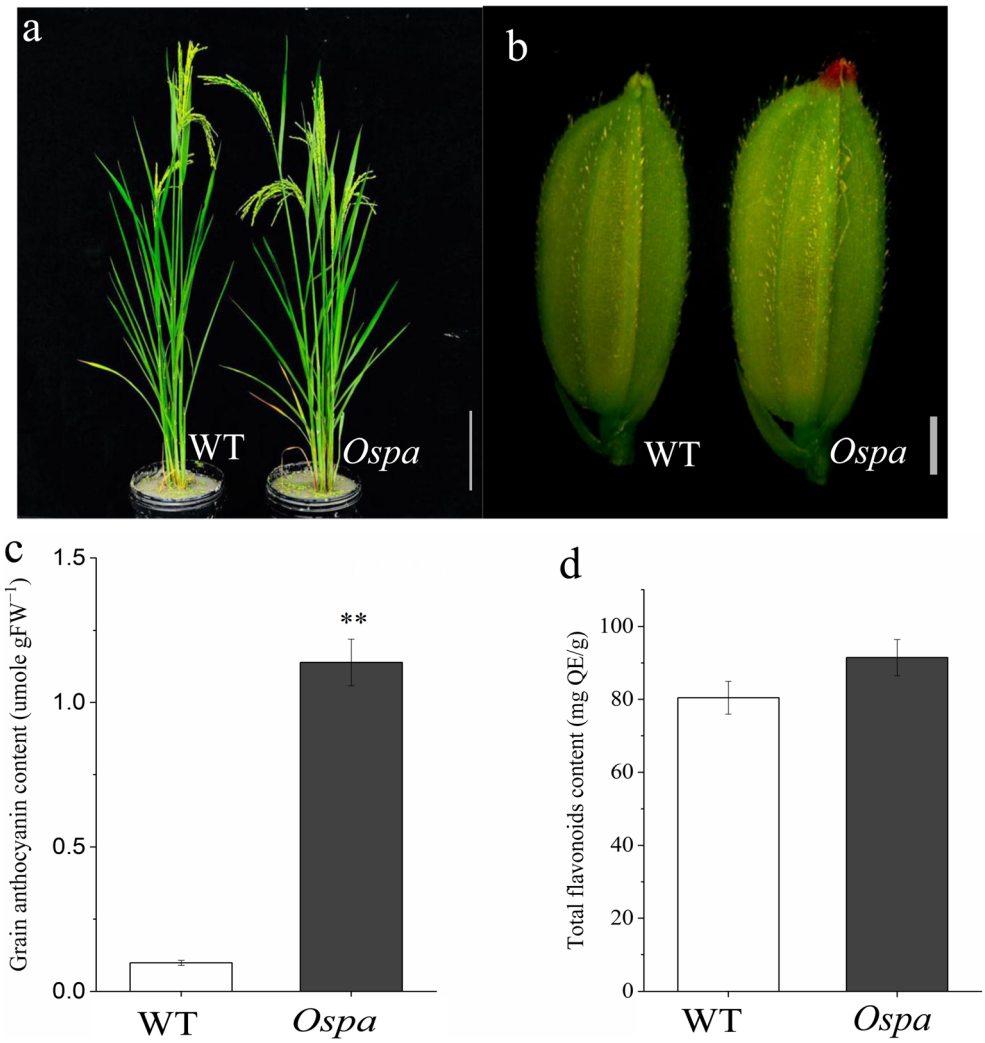


**Fig. 1** Comparison of different traits between WT and *Ospa* mutant. **a** Phenotype of WT and *Ospa*. **b** Spikelet of WT and *Ospa* mutant. Scale bar, 5 cm in (**a**) and 2 mm in (**b**). **c** Grain anthocyanin content of WT and *Ospa* mutant. **d** Total flavonoids content of WT and *Ospa* mutant. All data represented as mean ± SD. **represents significant difference at the 0.01 level by the Student *t* test ($n = 3$)

mutant from 5 to 15 days after heading (DAH) and gradually declined in the later stages, whereas there was no significant change in the hulls of WT (Fig. S1a). The grain flavonoids content increased from 5 to 20 DAH and maintained the same level in the later stages, but values were not significantly different between the Os*pa* mutant and WT (Fig. S1b). These results showed that the purple apiculus trait might be due to the increased anthocyanin accumulation in the *Ospa* mutant.

## Changes in Cellular Anatomy of *Ospa* Mutant

In addition to pigmentation and increased anthocyanin content, increased GL, GW, and TGW were measured in the *Ospa* mutant. To find out the cellular anatomy in relation to the increased grain size, the spikelet hulls were investigated by transmission electron microscopy (TEM) in WT and *Ospa* mutant. The anthocyanin accumulation was evident in the purple color observed in the outer parenchyma layer of the spikelet hulls in the *Ospa* mutant (Fig. 2a–d). The number of parenchyma cells in the outer layer of the spikelet hull of *Ospa* mutant was significantly higher than that in WT (Fig. 2c–e). Likewise, while the cell area in the inner surface of the spikelet hulls was larger in *Ospa* mutant than that in WT (Fig. 2f–h). On the contrary, the number of tubercles per mm$^2$ was significantly lower in *Ospa* mutant than the WT (Fig. 2i–k). Therefore, the cells in the mutant were bigger in size and higher in number in *Ospa* mutant than WT. These findings clearly showed that the increase in grain length and width might be from the expansion and proliferation of the spikelet hull cells.

## Genetic Analysis and Map-Based Cloning of *OsPA* Gene

The *Ospa* mutant showed purple coloration at the apiculus, while the WT exhibited green apiculus. All the $F_1$ plants from the crossing of the *Ospa* mutant (female parent) and green apiculus Zhenongda104 (male parent) showed the green apiculus, and the phenotypic segregation between WT (green

**Table 1** Agronomic traits of WT and *Ospa*

| Traits | WT | *Ospa* |
|---|---|---|
| PH (cm) | 82.87 ± 1.30 | 82.33 ± 0.98 |
| PL (cm) | 22.07 ± 1.03 | 22.00 ± 1.13 |
| PN (no.) | 9.13 ± 1.19 | 8.87 ± 1.25 |
| GNP (no.) | 207.60 ± 1.30 | 207.20 ± 1.32 |
| SR (%) | 91.46 ± 1.05 | 91.03 ± 1.56 |
| GL (mm) | 7.30 ± 0.59 | 9.23 ± 0.29** |
| GW (mm) | 3.13 ± 0.23 | 3.40 ± 0.21* |
| GT (mm) | 2.13 ± 0.23 | 2.20 ± 0.25 |
| BRL (mm) | 5.30 ± 0.59 | 5.70 ± 0.68 |
| BRW (mm) | 2.13 ± 0.13 | 2.20 ± 0.25 |
| BRT (mm) | 1.67 ± 0.24 | 1.78 ± 0.26 |
| TGW (g) | 26.23 ± 0.20 | 26.73 ± 0.29** |

PH, plant height; PL, panicle length; PN, panicle number per plant; GN, grain number per panicle; SR, seed setting rate; GL, grain length; GW, grain width; GT, grain thickness; BRL, brown rice length; BRW, brown rice width; BRT, brown rice thickness; TGW, 1000-grain weight. PL, GNP, and SR were measured using the largest panicle. **and *represent a highly significant and significant difference at the 0.01 and 0.05 level by the Student *t* test, respectively. All data represented as mean ± SD

apiculus) and *Ospa* phenotypes (purple apiculus) was observed in the $F_2$ population. There were 1,839 green apiculus and 576 purple apiculus phenotypes which fitted the 3:1 ratio of the Mendelian model ($\chi^2 = 3.40 < \chi^2_{0.05} = 3.84$), indicating that the purple apiculus of *Ospa* mutant was regulated by a single recessive gene.

The bulked segregant analysis (BSA) was carried out using SSR and InDel markers spanning the 12 chromosomes in rice. Polymorphic marker between the parents and the $F_2$ population derived from *Ospa* mutant/Zhenongda 104 was located on the long arm of chromosome 1. The *OsPA* gene was first delimited to a region between R1M30 and 1MYH01 (Fig. 3a). Subsequently, InDel markers were developed by comparing *indica* and *japonica* rice genomic sequences and used for fine mapping (Fig. 3b). Analyzing recombination events in the candidate region in 576 $F_2$ individual plants with the newly designed InDel markers, the mapping region was delimited to a 60.74 kb between markers 1MYH03 and 1MYH06.

The target region contained seven open reading frames (ORFs) based on the functional prediction by Genome Annotation Project (Kawahara et al. 2013) (Fig. 3c). After cloning and sequencing the viable candidate genes, a single nucleotide substitution was detected only in *LOC_01g39580*, encoding anthocyanin regulatory *Lc* protein. In the *Ospa* mutant, T to C substitution of the +3730th nucleotide in the third exon was detected. The nucleotide substitution resulted in the replacement of cysteine with arginine in the encoded protein of *OsPA* gene (Fig. 3d, e). Hence, we thought the difference in the gene and the resultant protein was the cause of the observed phenotypic, cellular, and physiological changes in the *Ospa* mutant.

## Predicted Structure of OsPA and Its Orthologous Proteins

The coding DNA sequence (CDS) of *OsPA* gene is 1,113-bp long with four exons and three introns. The CDS was predicted to encode a polypeptide of 371 amino acids with the calculated molecular mass of 40.29 kDa. The structural prediction showed the secondary structure of the encoded protein. The cysteine to arginine change was located on the helix motif of the peptide, which resulted in the formation of additional helix structure in the *Ospa* mutant (Fig. 4a, b; Fig. S2). The functional prediction of the protein was identified nucleic acid and peptide as ligands with several corresponding positions of binding sites, while there was no enzyme active site predicted that the gene might function as a transcription factor (TF) (Table S4).

Orthologous proteins were identified for structural similarity analysis. The protein basic alignment showed that the orthologous protein sequences from monocot plants such as *Zea mays*, *Triticum urartu*, *Sorghum bicolor*, *Setaria italica*, and *Panicum hallii* had high sequence similarity. The OsPA protein was predicted as TF, containing basic helix-loop-helix (bHLH) domain, in which the amino acid substitution was detected (Fig. S3). The results suggested that the mutation might have altered the peptide binding domain of the TF, which was conserved in the basic helix-loop-helix (bHLH) domain (Fig. 4b; Fig. S2 and S3).

For understanding the phylogenetic relationship and functional associations of related proteins, additional protein sequences were identified alignment of orthologs were performed. Only the orthologs from monocot plants were found to be and possessed bHLH domain. A phylogenetic tree was constructed for 87 orthologs based on the NJ method. The phylogenetic analysis showed that 87 orthologs were grouped into nine clades with effective bootstrap support (Fig. S4). The orthologs obtained from *O. sativa* ssp. *japonica*, *O. sativa* ssp. *indica*, *O. brachntha*, *O. longistaminata*, and *A. tauschii* were in the same clade. Therefore, it is noteworthy that the orthologs conserved in other monocots might be functionally related.

## The Expression of Anthocyanin and Grain Size-Related Genes

The *Ospa* mutant was characterized and the candidate gene which might have resulted in the observed phenotypic,

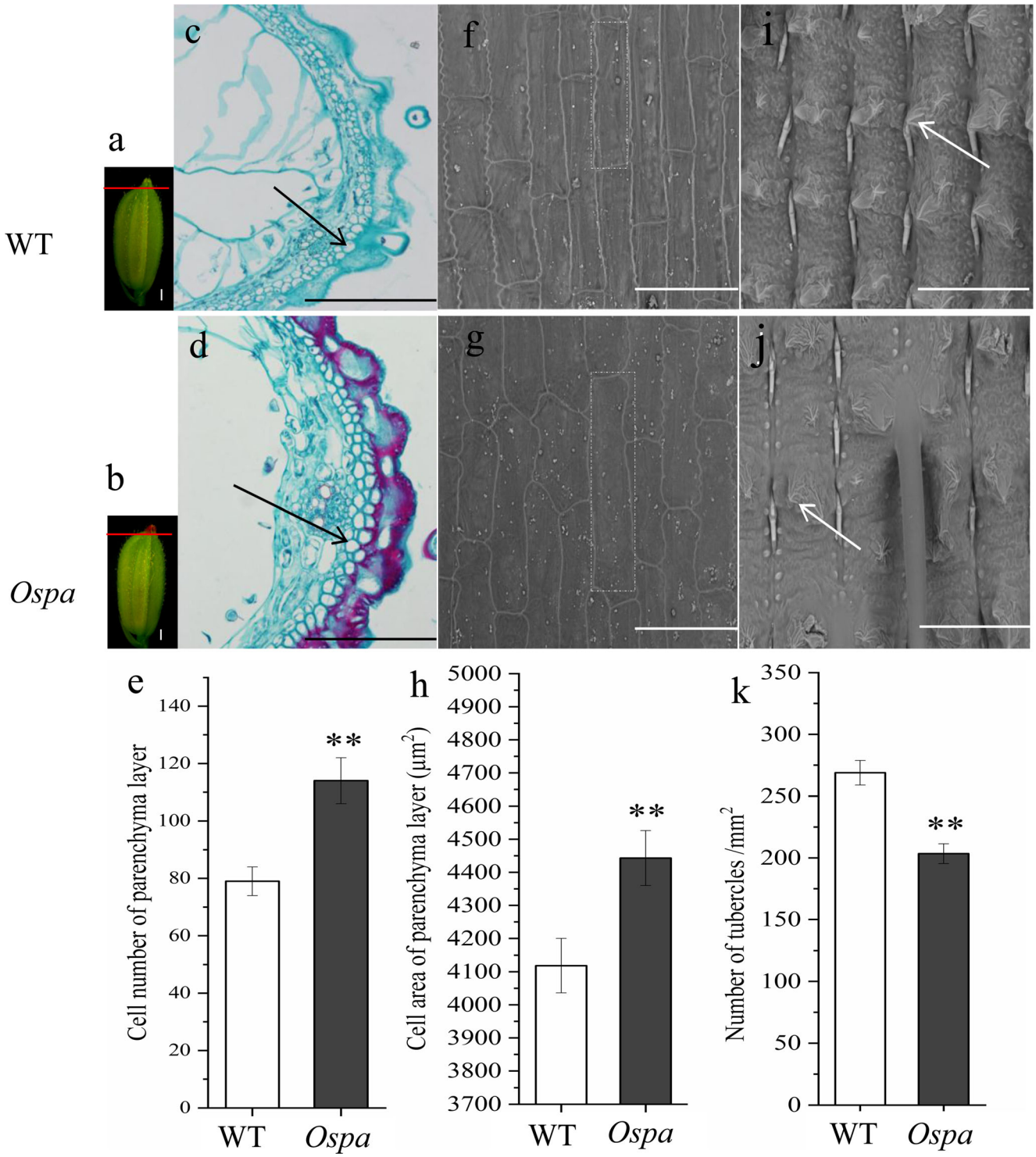


**Fig. 2** Histological analysis of spikelet hulls. **a, b** Spikelets of WT and *Ospa* mutant. Scale bars, 2 mm. The red lines indicated the positions of tissue analyzed. **c, d** TEM analysis of parenchyma layers of spikelet in WT and *Ospa* mutant respectively. Scale bars, 100 μm. **f, g** SEM analysis of inner and outer layers of spikelet hulls in WT and *Ospa* mutant respectively. Scale bars, 100 μm. **i, j** SEM analysis of outer surface of spikelet hulls in the WT and *Ospa* mutant respectively. Scale bars, 100 μm. **e** Comparison of cell number of inner parenchyma layer in WT and *Ospa* mutant. **h** Comparison of the cell area of inner surfaces of spikelet hulls in WT and *Ospa* mutant. **k** Comparison of the number of tubercles/mm$^2$ in the outer surfaces of spikelet hulls in WT and *Ospa* mutant. All data represent mean ± SD ($n = 3$). **represents significant difference at the 0.01 level by the Student *t* test

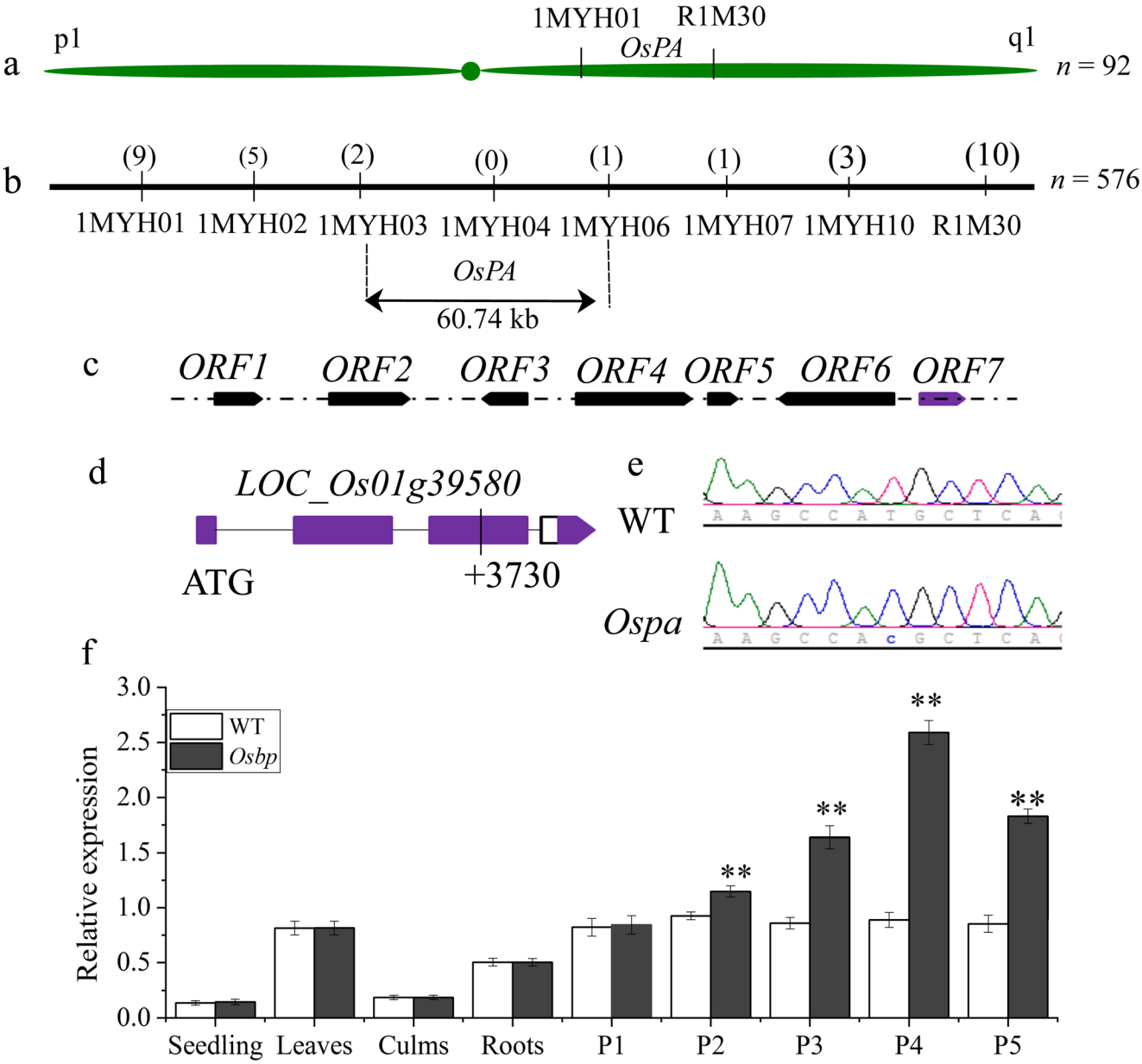


**Fig. 3** Schematic diagram of map-based cloning and expression profile of *OsPA* gene. **a** Bulked segregant analysis (BSA) located *OsPA* between markers R1M30 and 1MYH01 using 40 $F_2$ individuals with WT and *Ospa* mutant trait. **b** *OsPA* was fine mapped to an interval of 60.74 kb region by analyzing the recombination events in 576 $F_2$ individuals between markers 1MYH03 and 1MYH06. **c** Seven ORFs were annotated in the 60.74 kb region. **d** Schematic representation of *LOC_01g39580*. Filled boxes represent exons. **e** Site of mutation in *LOC_01g39580* as revealed by genomic sequencing of WT and *Ospa* mutant. **f** Relative expression profile of *OsPA* in different tissues. *P1*, *P2*, *P3*, *P4*, and *P5*; expression level in the panicle in the 2, 7, 12, 17, and 22 DAB. All data represent mean ± SD ($n = 3$). **represents significant difference at the 0.01 level by the Student *t* test

cellular, and biochemical changes was identified based on the map-based cloning strategy. The sequence analysis confirmed the genetic difference between the WT and *Ospa* mutant. Then, the mRNA transcript of the *OsPA* gene was analyzed in the different plant tissues and panicle during the growth stages using RT-qPCR. The results showed that the *OsPA* gene highly expressed in the young panicles, while lower expression was exhibited in leaf, stem, and root. When the expression pattern of *OsPA* gene was studied during the panicle growth stages in the WT and *Ospa* mutant, dramatic increase in expression was observed in the mutant, starting at 7 days after booting (DAB) and reaching the maximum expression level at the 17 DAB, while less variation of the expression level was found in the WT (Fig. 3f). Therefore, the results indicated that the tissue-specific significant increase of the *OsPA* mRNA transcript in the panicle of the *Ospa* mutant and could be linked to apiculus coloration and the increased accumulation of anthocyanin in the grains.

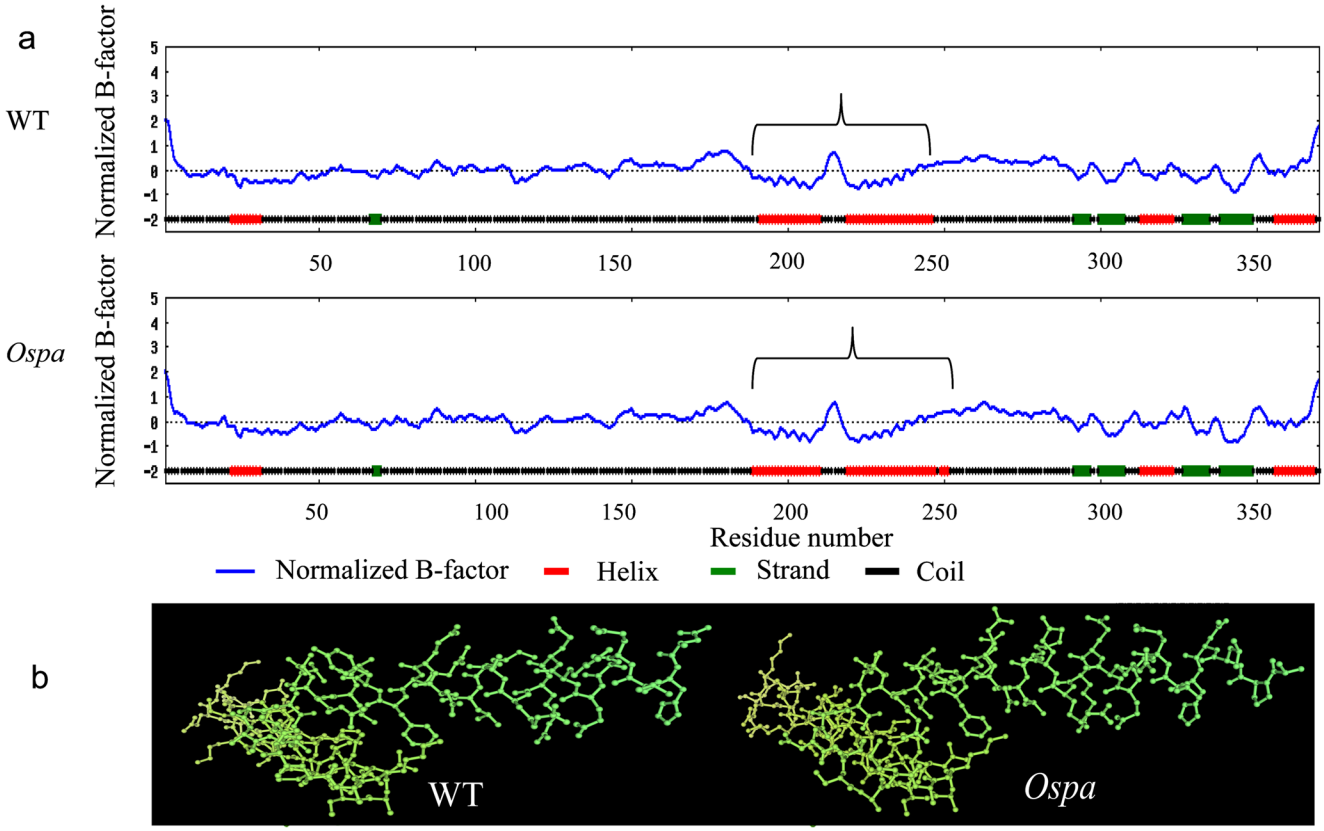


**Fig. 4** Structure prediction of the WT and *Ospa* mutant amino acid sequence. **a** Secondary structure prediction of the WT and *Ospa* amino acid sequence. Normalized B-factor represents the level of accuracy in the prediction. **b** Ball and stick model prediction of the WT and *Ospa* mutant protein structure for the region indicated with the brace signs in (**a**) which represents HLH conserved domains. Doted-box represents region in the *Ospa* mutant indicates the extended helix motif

As the apiculus coloration phenotype is determined by metabolic pathways and their regulatory cascades, the expression profiling of related genes underlying purple apiculus were conducted. The expression level of the structural anthocyanin biosynthesis genes, such as *OsF3'H*, *OsF3'5'H*, *OsDFR*, and *OsANS* (Aizza and Dornelas 2011; Shih et al. 2008), significantly increased in the panicle of the *Ospa* mutant (Fig. 5a). In addition to structural enzymatic genes, the expression of color producing gene, MYB gene (*OsC*), was significantly higher in the *Ospa* mutant than WT. On the other hand, the expression of *OsPAL*, *OsCHS*, and *OsCHI*, which direct conversion of upstream intermediates in the biosynthesis pathway (Aizza and Dornelas 2011; Hichri et al. 2011a, b; Hossain et al. 2018; Shih et al. 2008), remained not different between the WT and *Ospa* mutant. In a similar fashion to the *OsPA* gene, *OsC* showed tissue-specific expression in the panicle of *Ospa* mutant (Fig. 5a-b). Despite polymorphic markers were obtained for the mapped loci and DNA, and polypeptide sequence changes were found only in the *Ospa* gene, the gene expression levels of *OsF3'H*, *OsF3'5'H*, *OsDFR*, and *OsANS* increased along with *OsPA* gene in the mutant. The results showed that the expression levels of the regulatory and structural anthocyanin biosynthesis genes were upregulated in the mutant which might have activated anthocyanin accumulation in apiculus of the *Ospa* mutant grains.

In addition to the apiculus coloration, the *Ospa* mutant exhibited a significant increase in the grain size and thousand-grain weight. In relation to this change, the mRNA transcript levels of genes regulating grain size and weight were analyzed between the WT and *Ospa* mutant. The results exhibited that the expression level of *OsAGSW1* and *SG3*, which regulated grain size and weight by modulating cell size and proliferation in the spikelet hulls (Li et al. 2015; Wang et al. 2018) significantly increased in the *Ospa* mutant. The results indicated the mutation had an effect on grain size and weight as well (Fig. 5c). The other genes, such as *GS3*, *GS5*, *GW5*, *GW8*, *qGL3*, *SMG1*, and *TGW6*, which showed similar function (Duan et al. 2014; Fan et al. 2006; Ishimaru et al. 2013; Kang et al. 2018; Li et al. 2011; Liu et al. 2017; Zhang et al. 2016) did not show a significant difference in the expression level between the WT and *Ospa* mutant ($\alpha = 0.05$) (Fig. 5c). Also, *WTG1*, *D11*, *SLG*, *OsBUL1*, *PGLI*, *APG*, *SG1*, *SMOS1*, *GLW7*, and *GS2*, which regulated grain size and weight by modulating only cell expansion (Aya et al. 2014; Feng et al. 2016; Heang and Sassa 2012a, b; Huang et al. 2017; Jang et al. 2017; Si et al. 2016; Tong et al. 2018;

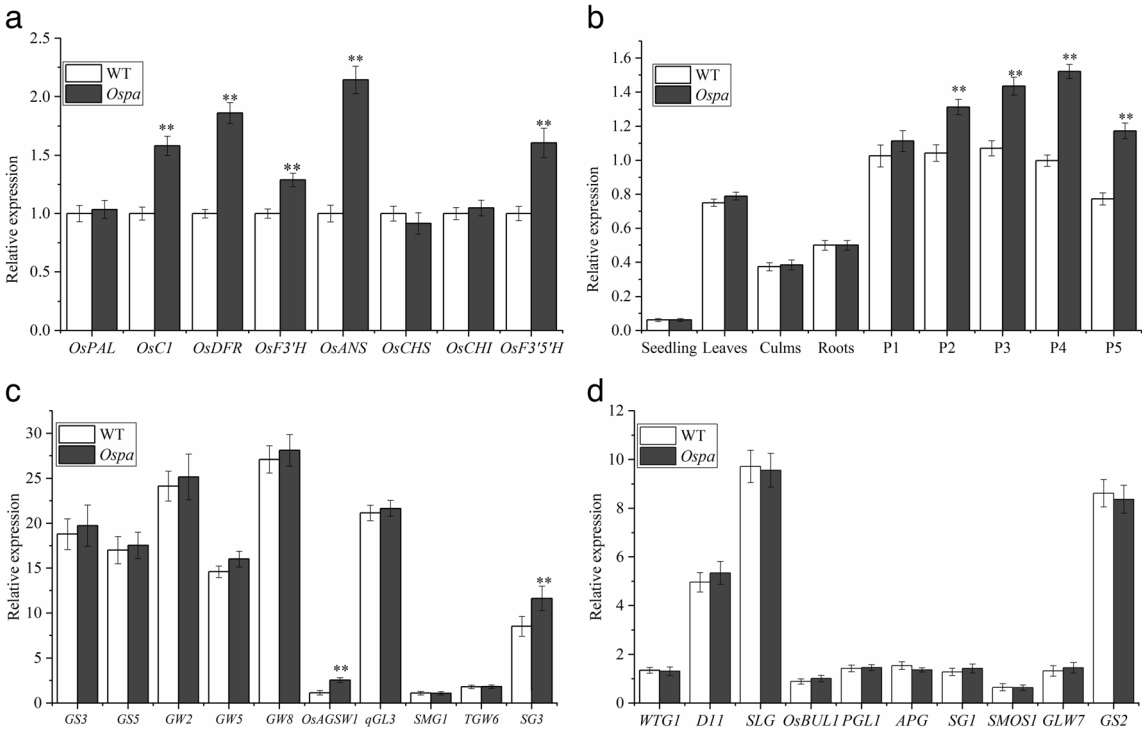


**Fig. 5** Gene expression analysis. **a** Relative expression of structural and regulatory genes of anthocyanin biosynthesis in spikelet hulls at 17 DAB (*OsC*, *LOC_Os06g10350*; *OsANS*, *LOC_Os06g42130*; *OsPAL*, *LOC_Os02g41630*; *OsCHS*, *LOC_Os11g32650*; *OsDFR*, *LOC_Os01g44260*; *OsF3'H*, *LOC_Os04g56700*; *OsCHI*, *LOC_Os03g60509*; *OsF3'5'H*, *LOC_Os10g17260*). **b** Relative expression profile of *OsC* in different tissues (*P1*, *P2*, *P3*, *P4*, and *P5*; expression level in the panicle in the 2, 7, 12, 17, and 22 DAB (**d**)). **c** Relative expression of cell size and proliferation related genes controlling grain size and weight in spikelet hulls at 17 DAB (*GS3*, *Os03g0407400*; *GS5*, *Os05g0158500*; *GW2*, *Os02g0244100*; *GW5*, *Os05g0187500*; *GW8*, *Os08g0531600*; OsAGSW1; *Os05g0323800*; *qGL3*, *Os03g0646900*; *SMG1*, *Os02g0787300*; *TGW6*, *Os06g0623700*; SG3, *Os03g0388800*) **d** Relative expression of cell expansion related genes controlling grain size and weight in spikelet hulls at 17 DAB (*WTG1*, *Os08g0537800*; *D11*, *Os04g0469800*; *SLG*, *Os08g0562500*; *OsBUL1*, *Os02g0747900*; *PGL1*, *Os03g0171300*; *APG*, *Os05g0139100*; *SG1*, *Os09g0459200*; *SMOS1*, *Os05g0389000*; *GLW7*, *Os07g0505200*; *GS2*, *Os02g0701300*). *OsACTIN1*, *LOC_Os03g50885* was used as internal control. All data represent mean ± SD ($n = 3$). **represents significant difference at the 0.01 level by the Student *t* test

Wuhan et al. 2013; You et al. 2007) did not show a significant difference in the expression level between WT and *Ospa* mutant ($\alpha = 0.05$) (Fig. 5d).

## Discussion

Apiculus color is one of the easily observed traits and regarded as a vital morphological marker in rice breeding. In the past, apiculus color has attracted the interest of a lot of researchers (Fan et al. 2007; Mikami et al. 2000; Reddy et al. 1995; Saitoh 2004; Sun et al. 2018; Zhao et al. 2016). Genes coding for bHLH proteins were involved in activation OsC expression which is considered as the basic gene for color production and regulated expression of structural anthocyanin biosynthesis genes to determine pigment distribution in the in rice (Sun et al. 2018). The bHLH domains make protein–protein interaction with the R2R3 MYB domain and activate downstream genes of the structural anthocyanin biosynthesis pathway (Koes et al. 2005; Sakamoto et al. 2001). Mutations altering the HLH domain can alter protein–protein interaction between HLH protein and any other protein and enhance or diminish the activities of the bHLH proteins. For instance, a single amino acid substitution in a bHLH transcription factor led to alteration of the interaction with partner proteins in Arabidopsis (Zhao et al. 2012). In the present study, a purple apiculus mutant, Ospa, was developed from wild-type with a green apiculus with EMS treatment and the phenotypic differences between the WT and Ospa mutant were observed in the grains. OsPA was identified by map-based cloning on chromosome 1. The OsPA protein was predicted as TF containing coil, strand, and helix motifs; however, the mutation was located in the helix motif which resulted in formation continuation of the helix motif in the predicted OsPA protein which is similar to the annotation by rice genome project which predicted as anthocyanin regulatory Lc protein (Kawahara et al. 2013) and annotation by gramene database as transcriptional activator protein (Gupta et al. 2016). The conserved domain analysis showed that OsPA protein was predicted to possess HLH domain which was highly conserved in the analyzed monocot species. The T to C nucleotide substitution that changed cysteine to arginine in the OsPA polypeptide was

identified in the HLH domain. The differences in the predicted peptide structure and mRNA transcript levels between the WT and Ospa mutant have implied that the mutation could have enhanced the expression of the OsPA and related genes to form purple apiculus trait and increased the grain size.

*Lc*-like proteins have been shown to affect the growth and development in addition to the anthocyanin biosynthesis. Overexpression of *Lc* gene in an Arabidopsis *ttg1* mutant complemented growth-related traits beyond anthocyanins content, indicating that genes coding for bHLH proteins with similarity to *Lc* can be involved in the regulation of epidermal cell growth (Lloyd et al. 1994). The bHLH *AtPIF4* played a key role in regulating morphogenesis (Leivar and Quail 2011). Consistent with these observations, the *OsPA* gene possessed the bHLH domain and analysis of cellular structure between WT and *Ospa* mutant revealed increased cell number and size in the parenchyma layers of spikelet hulls which, in turn, might have resulted in increased grain size and weight in our study. The significant increase of grain size and weight was supported by the mRNA transcript level analysis of genes related to grain size using qRT-PCR. Expression of *OsAGSW1* and *SG3* were significantly upregulated in the *Ospa* mutant. Previously, *OsAGSW1* increased grain size and weight by regulating the number of external parenchyma cells in the spikelet hulls of *OsAGSW1*-overexpressed rice and RNAi lines (Li et al. 2015). Similarly, EMS-generated sg3 mutant showed an increase in grain size and TGW by increasing cell proliferation in the spikelet hulls (Wang et al. 2018). In our study, the *Ospa* mutant showed upregulated expression of *OsAGSW1* and *SG3*, and it might have regulated the size and number of spikelet hull cells. Grain size in rice is highly controlled by the size of the spikelet hull (Li et al. 2011; Song et al. 2007). The results suggested that the *OsPA* gene might have affected cell size in spikelet hulls in addition to its effect on apiculus coloration synergistically with *OsAGSW1* and *SG3*.

Anthocyanins formation was catalyzed by enzymes encoded by structural genes, which utilized phenylalanine as a common substrate (Winkel-Shirley 2001). Chalcone synthase, chalcone isomerase, and flavanone 3-hydroxylase are encoded by the early biosynthesis genes are and are involved in the production of common precursors, dihydroflavonols, in the anthocyanin biosynthesis pathway whereas dihydroflavonol 4-reductase and anthocyanidin synthase are encoded by late biosynthesis genes (Ma et al. 2009). The bHLH proteins were involved in the regulation of the anthocyanin biosynthesis synergistically with MYB-protein, and WD-repeat proteins and many other proteins (Goff et al. 1992; Grotewold 2006; Hernandez et al. 2007; Nair and Burley 2000; Ramsay and Glover 2005; Sun et al. 2018). The bHLH protein *Lc* was specifically involved in tissue-specific anthocyanin pigmentation (Ludwig and Wessler 1990). This is a common theme in other plant species, in which an *Lc*-like bHLH protein requires an MYB protein partner to activate tissue-specific anthocyanin biosynthesis (Bernhardt et al. 2003; Quattrocchio et al. 1998; Spelt et al. 2000; Sun et al. 2018; Zhang et al. 2003). In the present study, the gene expression analysis showed that the *OsPA* gene expressed specifically in the panicle of the *Ospa* mutant and the expression dramatically increased during panicle developmental stages. In the past, the *OsC*, which is and MYB TF, and different from *OsPA* locus, was linked with purple apiculus trait (Liu et al. 2012). Recently, *OsC* was reported as a regulator of the red apiculus trait indicating that it is the basic gene for color formation in rice (Zhao et al. 2016). In the present study, the expression of *OsC* significantly increased in the immature panicles of *Ospa* mutant along *OsPA* gene. Therefore, *OsPA* and *OsC* might have regulated purple apiculus trait formation synergistically. Furthermore, significantly upregulated the expression of *OsF3'H*, *OsF3'5'H*, *OsDFR*, and *OsANS* specifically in the *Ospa* mutant was observed in our study. The enzyme product of *OsF3'5'H* directs the conversion of dihydrokaempferol to dihydromyricetin whereas the enzyme product of *OsDFR* directs conversion of dihydromyricetin, dihydroquercetin, and dihydrokaempferol into leucoanthocyanidins. Similarly, the enzyme product of *OsANS* directs the conversion of leucoanthocyanidins into anthocyanidins (Aizza and Dornelas 2011; Hichri et al. 2011a). The upregulated expression of *OsPA* and *OsC* along structural anthocyanin biosynthesis genes such as *OsF3'H*, *OsF3'5'H*, *OsDFR*, and *OsANS* suggested that the *OsPA* gene might have involved in positive regulation of the conversion of unique intermediates in the anthocyanin pathway that produce purple color in the apiculus of *Ospa* mutant. The observed changes in the grain anthocyanin content; expression of *OsC*, *OsF3'H*, *OsF3'5'H*, *OsDFR*, and *OsANS*; the size of parenchyma cells of spikelet hulls, size, and weight of grains could be linked to the single nucleotide change and the upregulated *OsPA* gene expression detected in the *Ospa* mutant.

## Conclusion

Apiculus color of grain is one of the easily observed traits in rice. In this study, we generated a new purple apiculus mutant, *Ospa*, using ethyl methanesulfonate (EMS) mutagenesis. We analyzed the phenotypic differences and isolated the candidate gene, *OsPA*, employing the map-based cloning strategy. In addition, we have shown that *OsPA* has regulated anthocyanin level quantifying the anthocyanin content in the grains and studying the expression of *OsPA* and structural anthocyanin biosynthesis genes using RT-qPCR. Investigating the cellular anatomy using histological analysis, we have demonstrated that parenchyma cells of spikelet hulls were more in number and enlarged size in the *Ospa* mutant which might implicate the pleiotropic nature of the *OsPA* gene. Since the gene has not

been reported before, it could provide new insight into the network of genes involved in tissue-specific pigmentation and the formation of purple apiculus trait as well as expansion and proliferation of spikelet hull parenchyma cells in rice. The *Ospa* mutant developed for this study can be utilized in rice selective breeding, and the developed gene-based markers would be applied to rice molecular breeding. Studying the function of *OsPA* by gene transformation in focus on its sub-cellular expression and interaction with anthocyanin regulatory and structural biosynthesis genes as well as genes related to grain size and weight will be worthwhile in the future.

**Authors' Contributions** X. L. Jin conceived the experiments and supervising the work. C. H. Shi found and affirmed the *Ospa* mutant. C. H. Shi and X. L. Jin developed the $F_2$ populations for the phenotypic segregation analysis and the gene mapping. Y. Tsago carried out the experiments, performed the statistical analysis, and wrote the manuscript. Z. K. Wang offered support in primer designing, DNA sequence assembly, and analysis skills. J. L. Liu offered support in the hybridization work and histological analysis skills. M. Sunusi, J. Eshag, and D. Akhter participated in manuscript editing. All authors read and approved the final manuscript.

**Funding Information** This work was supported by the Science and Technology Office of Zhejiang Province, China (Grant Nos. 2016C02050-6 and 2012C12901-2) and the National Key Research and Development of China (Grant No. 2017YFD0100300-5).

## References

Aizza LCB, Dornelas MC (2011) A genomic approach to study anthocyanin synthesis and flower pigmentation in passionflowers. J Nucleic Acids 2011:371517

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Aya K, Hobo T, Sato-Izawa K, Ueguchi-Tanaka M, Kitano H, Matsuoka M (2014) A novel *AP2*-type transcription factor, *SMALL ORGAN SIZE1*, controls organ size downstream of an auxin signaling pathway. Plant Cell Physiol 55:897–912

Brenda WS (2001) Flavonoid biosynthesis. a colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiol 126: 485–493

Bernhardt C, Lee MM, Gonzalez A, Zhang F, Lloyd A, Schiefelbein J (2003) The bHLH genes *GLABRA3* (*GL3*) and *ENHANCER OF GLABRA3* (*EGL3*) specify epidermal cell fate in the Arabidopsis root. Development 130:6431–6439

Chalker-Scott L (1999) Environmental significance of anthocyanins in plant stress responses. Photochem Photobiol 70:1–9

Chandler VL, Radicella JP, Robbins TP, Chen J, Turks D (1989). Two regulatory genes of the maize anthocyanin pathway are homologous: Isolation of B utilizing R genomic sequences. Plant Cell 1: 1175–1183

Chin HS, Wu YP, Hour AL, Hong CY, Lin YR (2016) Genetic and evolutionary analysis of purple leaf sheath in rice. Rice 9:14

Dooner HK, Robbins TP, and Jorgensen RA (1991) Genetic and developmental control of anthocyanin biosynthesis. Annu Rev Genet 25: 173–199

Duan PG, Rao YC, Zeng DL, Yang YL, Xu R, Zhang BL, Dong GJ, Qian Q, Li YH (2014) *SMALL GRAIN 1*, which encodes a mitogen-activated protein kinase kinase 4, influences grain size in rice. Plant J 77:547–557

El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, Sonnhammer ELL, Hirsh L, Paladin L, Piovesan D, Tosatto SCE, Finn RD (2019) The Pfam protein families database in 2019. Nucleic Acids Res 47:D427–D432

Fan CH, Xing YZ, Mao HL, Lu TT, Han B, Xu CG, Li XH, Zhang QF (2006) *GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. Theor Appl Genet 112:1164–1171

Fan FJ, Fan YV, Du JH, Zhuang JY (2007) Fine mapping of *C* (chromogen for anthocyanin) gene in rice. Zhongguo Shuidao Kexue 21: 454–458

Feng ZM, Wu CY, Wang CM, Roh J, Zhang L, Chen J, Zhang SZ, Zhang H, Yang CY, Hu JL, You XM, Liu X, Yang XM, Guo XP, Zhang X, Wu FQ, Terzaghi W, Kim SK, Jiang L, Wan JM (2016) *SLG* controls grain size and leaf angle by modulating brassinosteroid homeostasis in rice. J Exp Bot 67:4241–4253

Furukawa T, Maekawa M, Oki T, Suda I, Iida S, Shimada H, Takamure I, Kadowaki K-i (2007) The *Rc* and *Rd* genes are involved in proanthocyanidin synthesis in rice pericarp. Plant J 49:91–102

Gao DY, He B, Zhou YH, Sun LH (2011) Genetic and molecular analysis of a purple sheath somaclonal mutant in japonica rice. Plant Cell Rep 30:901–911

Goff SA, Cone KC, Chandler VL (1992) Functional analysis of the transcriptional activator encoded by the maize *B* gene: evidence for a direct functional interaction between 2 classes of regulatory proteins. Genes Dev 6:864–875

Grotewold E (2006) The genetics and biochemistry of floral pigments. In: Annual review of plant biology, vol 57. Annual Reviews, Palo Alto, pp 761–780

Gupta P, Naithani S, Tello-Ruiz MK, Chougule K, D'Eustachio P, Fabregat A, Jiao Y, Keays M, Lee YK, Kumari S, Mulvaney J, Olson A, Preece J, Stein J, Wei S, Weiser J, Huerta L, Petryszak R, Kersey P, Stein LD, Ware D, Jaiswal P (2016) Gramene database: navigating plant comparative genomics resources. Curr Plant Biol 7-8:10–15

Heang D, Sassa H (2012a) Antagonistic actions of HLH/bHLH proteins are involved in grain length and weight in rice. PLoS One 7:11

Heang D, Sassa H (2012b) A typical bHLH protein encoded by *POSITIVE REGULATOR OF GRAIN LENGTH 2* is involved in controlling grain length and weight of rice through interaction with a typical bHLH protein APG. Breed Sci 62:133–141

Hernandez JM, Feller A, Morohashi K, Frame K, Grotewold E (2007) The basic helix-loop-helix domain of maize *R* links transcriptional regulation and histone modifications by recruitment of an EMSY-related factor. Proc Natl Acad Sci U S A 104:17222–17227

Hichri I, Barrieu F, Bogs J, Kappel C, Delrot S, Lauvergeat V (2011a) Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. J Exp Bot 62:2465–2483

Hichri I, Deluc L, Barrieu F, Bogs J, Mahjoub A, Regad F, Gallois B, Granier T, Trossat-Magnin C, Gomès E, Lauvergeat V (2011b) A single amino acid change within the R2 domain of the VvMYB5b transcription factor modulates affinity for protein partners and target promoters selectivity. BMC Plant Biol 11:117

Hossain MR, Kim H-T, Shanmugam A, Nath UK, Goswami G, Song J-Y, Park J-I, Nou I-S (2018) Expression profiling of regulatory and biosynthetic genes in contrastingly anthocyanin-rich strawberry (*Fragaria ananassa*) cultivars reveals key genetic determinants of fruit color. Int J Mol Sci 19:656

Huang K, Wang DK, Duan PG, Zhang BL, Xu R, Li N, Li YH (2017) *WIDE AND THICK GRAIN 1*, which encodes an otubain-like protease with deubiquitination activity, influences grain size and shape in rice. Plant J 91:849–860

Ishimaru K, Hirotsu N, Madoka Y, Murakami N, Hara N, Onodera H, Kashiwagi T, Ujiie K, Shimizu B, Onishi A, Miyagawa H, Katoh E (2013) Loss of function of the IAA-glucose hydrolase gene *TGW6* enhances rice grain weight and increases yield. Nat Genet 45:707–713

Ithal N, Reddy AR (2004) Rice flavonoid pathway genes, *OsDFR* and *OsANS*, are induced by dehydration, high salt and ABA, and contain stress responsive promoter elements that interact with the transcription activator, *OsC1*-MYB. Plant Sci 166:1505–1513

Jang S, An G, Li HY (2017) Rice leaf angle and grain size are affected by the *OsBUL1* transcriptional activator complex. Plant Physiol 173: 688–702

Kang YJ, Shim KC, Lee HS, Jeon YA, Kim SH, Kang JW, Yun YT, Park IK, Ahn SN (2018) Fine mapping and candidate gene analysis of the quantitative trait locus *gw8.1* associated with grain length in rice. Genes Genomics 40:389–397

Kawahara Y, de la Bastide M, Hamilton JP, Kanamori H, McCombie WR, Ouyang S, Schwartz DC, Tanaka T, Wu JZ, Zhou SG, Childs KL, Davidson RM, Lin HN, Quesada-Ocampo L, Vaillancourt B, Sakai H, Lee SS, Kim J, Numa H, Itoh T, Buell CR, Matsumoto T (2013) Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. Rice 6: 10

Koes R, Verweij W, Quattrocchio F (2005) Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. Trends Plant Sci 10:236–242

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549

Leivar P, Quail PH (2011) PIFs: pivotal components in a cellular signaling hub. Trends Plant Sci 16:19–28

Li YB, Fan CC, Xing YZ, Jiang YH, Luo LJ, Sun L, Shao D, Xu CJ, Li XH, Xiao JH, He YQ, Zhang QF (2011) Natural variation in *GS5* plays an important role in regulating grain size and yield in rice. Nat Genet 43:1266–1270

Li T, Jiang JM, Zhang SC, Shu HR, Wang YQ, Lai JB, Du JJ, Yang CW (2015) *OsAGSW1*, an ABC1-like kinase gene, is involved in the regulation of grain size and weight in rice. J Exp Bot 66:5691–5701

Lin-Wang K, Bolitho K, Grafton K, Kortstee A, Karunairetnam S, McGhie TK, Espley RV, Hellens RP, Allan AC (2010) An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in Rosaceae. BMC Plant Biol 10:50

Liu X, Sun X, Wang WY, Ding HF, Liu W, Li GX, Jiang MS, Zhu CX, Yao FY (2012) Fine mapping of *Pa-6* gene for purple apiculus in rice. J Plant Biol 55:218–225

Liu JF, Chen J, Zheng XM, Wu FQ, Lin QB, Heng YQ, Tian P, Cheng ZJ, Yu XW, Zhou KN, Zhang X, Guo XP, Wang JL, Wang HY, Wan JM (2017) *GW5* acts in the brassinosteroid signalling pathway to regulate grain width and weight in rice. Nat Plants 3(5):17043

Lloyd AM, Schena M, Walbot V, Davis RW (1994) Epidermal-cell fate determination in Arabidopsis: patterns defined by a steroid-inducible regulator. Science 266:436–439

Ludwig SR, Wessler SR (1990) Maize *R* gene family tissue-specific helix-loop-helix proteins. Cell 62:849–851

Ma H, Pooler M, Griesbach R (2009) Anthocyanin regulatory/structural gene expression in Phalaenopsis. J Am Soc Hortic Sci 134:88–96

Maeda H, Yamaguchi T, Omoteno M, Takarada T, Fujita K, Murata K, Iyama Y, Kojima Y, Morikawa M, Ozaki H, Mukaino N, Kidani Y, Ebitani T (2014) Genetic dissection of black grain rice by the development of a near-isogenic line. Breed Sci 64:134–141

Michelmore RW, Paran I, and Kesseli, R. V. (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proceedings of the National Academy of Sciences of the United States of America 88:9828–9832

Mikami AT, Khin-Thidar, Sano Y (2000) A candidate for *C* (Chromogen for anthocyanin) gene. Rice Genet Newsl 17:54–56

Murray MG, Thompson WF (1980) Rapid isolation of high molecular-weight plant DNA. Nucleic Acids Res 8:4321–4325

Nair SK, Burley SK (2000) Recognizing DNA in the library. Nature 404: 715–717

Oikawa T, Maeda H, Oguchi T, Yamaguchi T, Tanabe N, Ebana K, Yano M, Ebitani T, Izawa T (2015) The birth of a black rice gene and its local spread by introgression. Plant Cell 27:2401–2414

Quattrocchio F, Wing JF, van der Woude K, Mol JNM, Koes R (1998) Analysis of bHLH and MYB domain proteins: species-specific regulatory differences are caused by divergent evolution of target anthocyanin genes. Plant J 13:475–488

Ramsay NA, Glover BJ (2005) MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. Trends Plant Sci 10:63–70

Reddy VS, Dash S, Reddy AR (1995) Anthocyanin pathway in rice (*Oryza sativa* L.): identification of a mutant showing dominant inhibition of anthocyanins in leaf and accumulation of proanthocyanidins in pericarp. Theor Appl Genet 91:301–312

Saitoh K (2004) Allelic diversification at the *C* (*OsC1*) locus of wild and cultivated rice: nucleotide changes associated with phenotypes. Genetics 168:997–1007

Sakamoto W, Ohmori T, Kageyama K, Miyazaki C, Saito A, Murata M, Noda K, Maekawa M (2001) The purple leaf (*Pl*) locus of rice: the *Pl*(*w*) allele has a complex organization and includes two genes encoding basic helix-loop-helix proteins involved in anthocyanin biosynthesis. Plant Cell Physiol 42:982–991

Sayers EW, Agarwala R, Bolton EE, Brister JR, Canese K, Clark K, Connor R, Fiorini N, Funk K, Hefferon T, Holmes JB, Kim S, Kimchi A, Kitts PA, Lathrop S, Lu ZY, Madden TL, Marchler-Bauer A, Phan L, Schneider VA, Schoch CL, Pruitt KD, Ostell J (2019) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 47:D23–D28

Shen Y, Jin L, Xiao P, Lu Y, Bao JS (2009) Total phenolics, flavonoids, antioxidant capacity in rice grain and their relations to grain color, size and weight. J Cereal Sci 49:106–111

Shih CH, Chu H, Tang LK, Sakamoto W, Maekawa M, Chu IK, Wang M, Lo C (2008) Functional characterization of key structural genes in rice flavonoid biosynthesis. Planta 228:1043–1054

Si LZ, Chen JY, Huang XH, Gong H, Luo JH, Hou QQ, Zhou TY, Lu TT, Zhu JJ, Shangguan YY, Chen EW, Gong CX, Zhao Q, Jing YF, Zhao Y, Li Y, Cui LL, Fan DL, Lu YQ, Weng QJ, Wang YC, Zhan QL, Liu KY, Wei XH, An K, An G, Han B (2016) *OsSPL13* controls grain size in cultivated rice. Nat Genet 48:447–457

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soeding J, Thompson JD, Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. Mol Syst Biol 7:539

Song XJ, Huang W, Shi M, Zhu MZ, Lin HX (2007) A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nat Genet 39:623–630

Spelt C, Quattrocchio F, Mol JNM, Koes R (2000) *Anthocyanin1* of Petunia encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. Plant Cell 12: 1619–1631

Sun XM, Zhang ZY, Chen C, Wu W, Ren NN, Jiang CH, Yu JP, Zhao Y, Zheng XM, Yang QW, Zhang HL, Li JJ, Li ZC (2018) The C-S-A gene system regulates hull pigmentation and reveals evolution of anthocyanin biosynthesis pathway in rice. J Exp Bot 69:1485–1498

Sweeney MT, Thomson MJ, Pfeil BE, McCouch S (2006) Caught red-handed: *Rc* encodes a basic helix-loop-helix protein conditioning red pericarp in rice. Plant Cell 18:283–294

Takahashi (1957) Analysis on apiculus color genes essential to anthocyanin coloration rice. Journal of the Faculty of Agriculture, Hokkaido University 50(3):266–362

Takahashi (1982) Genetic studies on rice plants. Journal of the Faculty of Agriculture, Hokkaido University 61, 91–42

Tello-Ruiz MK, Naithani S, Stein JC, Gupta P, Campbell M, Olson A, Wei S, Preece J, Geniza MJ, Jiao Y, Lee YK, Wang B, Mulvaney J, Chougule K, Elser J, Al-Bader N, Kumari S, Thomason J, Kumar V, Bolser DM, Naamati G, Tapanari E, Fonseca N, Huerta L, Iqbal H, Keays M, Fuentes AM-P, Tang A, Fabregat A, D'Eustachio P, Weiser J, Stein LD, Petryszak R, Papatheodorou I, Kersey PJ, Lockhart P, Taylor C, Jaiswal P, Ware D (2018) Gramene 2018: unifying comparative genomics and pathway resources for plant research. Nucleic Acids Res 46:D1181–D1189

Tong XH, Wang YF, Sun AQ, Bello BK, Ni S, Zhang J (2018) *Notched belly grain 4*, a novel allele of *dwarf 11*, regulates grain shape and seed germination in rice (*Oryza sativa* L.). Int J Mol Sci 19:9

Wang ZK, Zeng DD, Qin R, Liu JL, Shi CH, Jin XL (2018) A novel and pleiotropic factor *SLENDER GRAIN3* is involved in regulating grain size in rice. Rice Sci 25:132–141

Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ (2009) Jalview version 2: a multiple sequence alignment editor and analysis workbench. Bioinformatics 25:1189–1191

Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T (2018) SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46:W296–W303

Winkel-Shirley B (2001) Flavonoid biosynthesis: a colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiol 126:485–493

Wuhan Z, Pingyong S, Qiang H, Fu S, Jie W, Huafeng D (2013) Fine mapping of *GS2*, a dominant gene for big grain rice. Crop J 1:160–165

Yang J, Roy A, Zhang Y (2013a) BioLiP: a semi-manually curated database for biologically relevant ligand-protein interactions. Nucleic Acids Res 41:D1096–D1103

Yang J, Roy A, Zhang Y (2013b) Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. Bioinformatics 29:2588–2595

Yang J, Wang Y, Zhang Y (2016) ResQ: an approach to unified estimation of B-factor and residue-specific error in protein structure prediction. J Mol Biol 428:693–701

Yang CC, Zeng DD, Qin R, Alamin M, Jin XL, Shi CH (2018) Rice gene, *BBH*/*Lsi1*, regulates the color of rice hull by reducing the absorption and deposition of silicon and accumulating excess flavonoid. Plant Growth Regul 85:133–142

You SC, Cho SH, Zhang H, Paik HC, Lee CH, Li J, Yoo JH, Lee BW, Koh HJ, Seo HS, Paek NC (2007) Quantitative trait loci associated with functional stay-green *SNU-SG1* in rice. Mol Cell 24:83–94

Zhang F, Gonzalez A, Zhao MZ, Payne CT, Lloyd A (2003) A network of redundant bHLH proteins functions in all *TTG1*-dependent pathways of *Arabidopsis*. Development 130:4859–4869

Zhang YD, Zhu Z, Zhao QY, Chen T, Yao S, Zhou LH, Zhao L, Zhao CF, Wang CL (2016) Haplotypes of *qGL3* and their roles in grain size regulation with *GS3* alleles in rice. Genet Mol Res 15:10

Zhang C, Freddolino PL, Zhang Y (2017) COFACTOR: improved protein function prediction by combining structure, sequence and protein-protein interaction information. Nucleic Acids Res 45: W291–W299

Zhao H, Wang X, Zhu D, Cui S, Li X, Cao Y, Ma L (2012) A single amino acid substitution in IIIf subfamily of basic helix-loop-helix transcription factor *AtMYC1* leads to trichome and root hair patterning defects by abolishing its interaction with partner proteins in Arabidopsis. J Biol Chem 287:14109–14121

Zhao SS, Wang CH, Ma J, Wang S, Tian P, Wang JL, Cheng ZJ, Zhang X, Guo XP, Lei CL (2016) Map-based cloning and functional analysis of the chromogen gene C in rice (*Oryza sativa* L.). J Plant Biol 59: 496–505