

ORIGINAL ARTICLE

Open Access

# Genetic and Evolutionary Analysis of Purple Leaf Sheath in Rice



Han-shiuan Chin<sup>1†</sup>, Yong-pei Wu<sup>2†</sup>, Ai-ling Hour<sup>3</sup>, Chwan-yang Hong<sup>4</sup> and Yann-rong Lin<sup>1\*</sup>

## Abstract

**Background:** Anthocyanin accumulates in many plant tissues or organs, in rice for example leading to red, purple red and purple phenotypes for protection from damage by biotic and abiotic stresses and for reproduction. Purple leaf, leaf sheath, stigma, pericarp, and apiculus are common in wild rice and landraces and occasionally found in modern cultivars. No gene directly conferring anthocyanin deposited in a purple leaf sheath has yet been isolated by using natural variants. An F<sub>2</sub> population derived from *ssp. japonica* cv. Tainung 72 (TNG72) with purple leaf sheath (PSH) crossed with *ssp. indica* cv. Taichung Sen 17 (TCS17) with green leaf sheath (GSH) was utilized to isolate a gene conferring leaf sheath color.

**Results:** By positional cloning, 10-and 3-bp deletions in the R2R3 Myb domain of *OsC1* were uncovered in GSH varieties TCS17 and Nipponbare, respectively. Allelic diversity, rather than gene expression levels of *OsC1*, might be responsible for anthocyanin accumulation. Parsimony-based analysis of genetic diversity in 50 accessions, including cultivars, landraces, and A-genome wild rice, suggests that independent mutation occurred in Asian, African, South American, and Australian species, while *O. meridionalis* had a divergent sequence. *OsC1* was thought of as a domestication related gene, with up to 90 % reduction of genetic diversity in GSH; however, no values from three tests showed significant differences from neutral expectations, implying that *OsC1* had not been subjected to recent selection. Haplotype network analysis revealed that species from different continents formed unique haplotypes with no gene flow. Two major groups of haplotypes corresponding to 10-bp deletion and other sequences were formed in Asian rice, including *O. rufipogon*, *O. nivara* and *O. sativa*. Introgressions of *OsC1* between subspecies through natural and artificial hybridization were not rare. Because artificial and natural selection imposed admixture on rice germplasm in Taiwan, the genealogy of *OsC1* might not be congruent with the current distribution of alleles through lineage diversification.

**Conclusion:** *OsC1* is responsible for purple leaf sheath, and much new information about *OsC1* is provided e.g., new alleles, non-domestication syndrome, and incongruence of genealogy with geographic distribution.

**Keywords:** Genealogy, Genetic Diversity, *OsC1*, Purple Leaf Sheath

## Background

Plants accumulate diverse pigments in various tissues and organs related to photosynthesis, defense, and reproduction. Anthocyanins, belonging to the flavonoid class of pigment molecules, are important secondary metabolites in rice (Reddy et al. 1995). Anthocyanin accumulation in different tissues is sometimes involved in many physiological functions, such as modulation of hormone responses, protection from damage by ultra-

violet radiation, and defense responses to biotic and abiotic stresses (Reddy et al. 1994; Chalker-Scott 1999; Ithal and Reddy 2004). Purple stigma and apiculus attract insects for pollination and animals for seed dispersal, respectively. Different degrees of purple color are commonly found in the root, leaf sheath, leaf blade, stigma, and apiculus of wild rice species and landraces and occasionally seen in cultivars. Allelic diversity of genes conferring purple pigmentation is maintained in natural germplasm.

Anthocyanin metabolism is regulated by genes and influenced by environmental factors such as pH, ultra-violet radiation and temperature, which have been well

\* Correspondence: ylin@ntu.edu.tw

†Equal contributors

<sup>1</sup>Department of Agronomy, National Taiwan University, Taipei, Taiwan  
Full list of author information is available at the end of the article

studied in maize, *Petunia*, *Arabidopsis* and other model plants (Dooner et al. 1991; Brenda 2001; Koes et al. 2005). In addition to the genes encoding enzymes participating in the anthocyanin synthesis pathway, several regulatory genes including members of the *C1*, *Pl*, *B* and *R* gene families in maize have been identified and their functions elucidated. *C1* and *Pl* proteins containing Myb DNA binding domains regulate downstream genes of the flavonoid synthesis pathway; while *B* and *R* proteins containing a basic helix-loop-helix (bHLH) domain are transcriptional activators (Chandler et al. 1989). Based on sequence similarity, *OsC1*, *OsB1* and *OsB2*, which are homologous to maize *C1* and *B*, were isolated from rice (Reddy et al. 1998; Sakamoto et al. 2001). The N-terminal and bHLH domains of *OsB1* and *OsB2* take advantage of protein-protein interaction with *OsC1* that have an R2R3 Myb domain, and these proteins regulate directly the downstream genes of the flavonoid synthesis pathway (Sakamoto et al. 2001; Koes et al. 2005).

Asian rice, *Oryza sativa*, was domesticated from wild species, *O. rufipogon*, an estimated 10,000 years ago. The domestication syndromes are considered to include aspects of grain color, grain size, yield, and other desirable agronomic traits. Domestication-related genes have been classified in terms of crop domestication, improvement and diversification (Kovach et al. 2007; Larson et al. 2014). Shattering genes, *sh4* and *qSH1*, were deliberately modified to influence harvest efficiency as recently as the past 100 years, contributing directly to rice domestication (Zhang et al. 2009). Another key domestication-related gene, *sd-1*, resulting in semi-dwarf stature, was a key element of the 'Green Revolution' in the 1960s in Asia by virtue of dramatically increasing grain yield (Asano et al. 2007). The purple apiculus trait which is common in wild rice but rare in cultivars is controlled by a regulatory gene, *OsC1*. Allelic diversity of *OsC1* revealed signatures of selection in cultivated Asian rice but not in indigenous improved rice varieties in North India (Saitoh et al. 2004; Choudhury et al. 2014). Nevertheless, some crop improvement traits are regional preferences, such as grain texture and flavor regulated by *waxy* and *BADH2* which harbor allelic diversity in different landraces and cultivars (Olsen and Purugganan 2002; Kovach et al. 2009).

In rice, purple leaf sheath (PSH) as well as purple apiculus and stigma is common in wild species and landraces; however, green leaf sheath (GSH) is prevalent in modern cultivars. One or two major QTLs were suggested to confer leaf sheath color, and two QTLs mapped on chromosomes 1 and 6 could explain more than 50 % of phenotypic variation (Hadagal et al. 1980; Yue et al. 2006). A purple leaf sheath gene, *PSHI* (t), was narrowed down to an interval of 23.5 kb on chromosome 1 encompassing 6 candidate genes after high resolution of linkage mapping. No gene

directly conferring PSH has yet been isolated by using natural variants.

In this study, an  $F_2$  segregating population of *ssp. japonica* cv. Tainung 72 (TNG72) with purple leaf sheath  $\times$  *ssp. indica* cv. Taichung Sen 17 (TCS17) with green leaf sheath was utilized to isolate a gene conferring leaf sheath color by positional cloning. Four accessions with various levels of anthocyanin accumulation in the leaf sheath were used to investigate the relationship between gene expression and anthocyanin content. Analyses of haplotype and nucleotide diversity based on 50 rice accessions which included improved cultivars, landraces, and wild species were conducted to reveal if the gene has been responsive to artificial selection and the associated geographic distribution of its alleles.

## Result

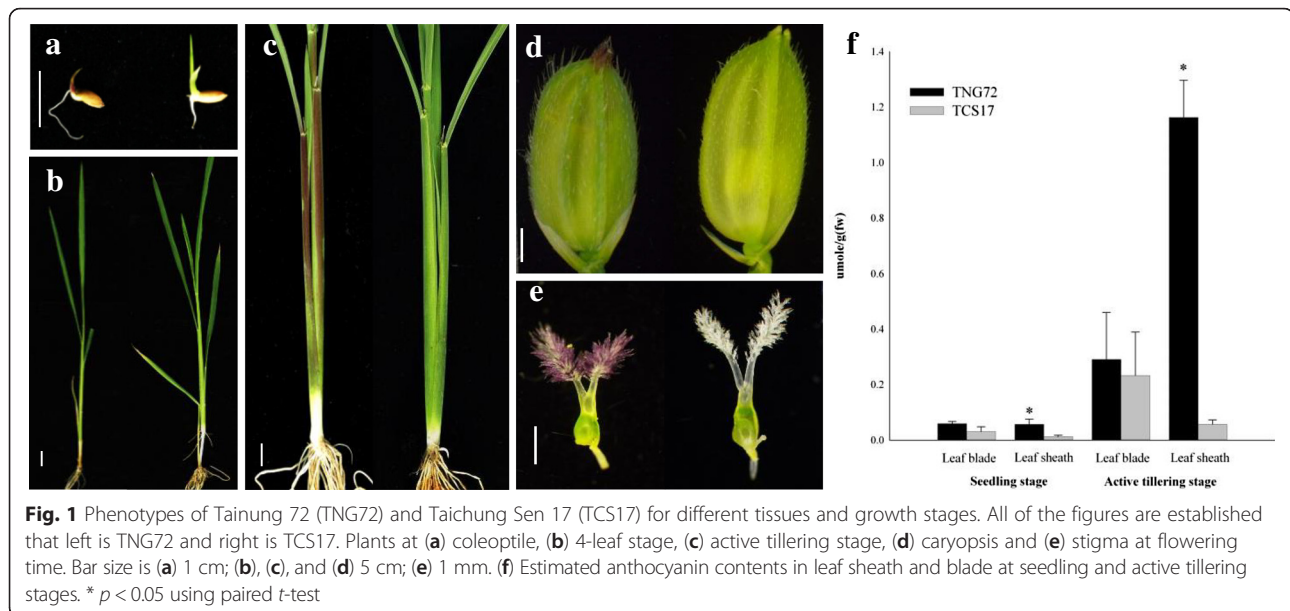
### Accumulation of Anthocyanins in Tissues of TNG72 and TCS17 at Various Growth Stages

TNG72 possessed purple coleoptile, leaf sheath, apiculus, and stigma, and the degrees of purple color varied at different growth stages (Fig. 1a-e). The leaf sheath of TNG72 was green at the 4-leaf stage but gradually turned purple until the active tillering stage, remaining purple until maturity. TNG72 accumulated various amounts of anthocyanin with different localization in rice tissues. However, TCS17 remained green for the entire life cycle, representing the most common phenotype of cultivated rice (Fig. 1a-e).

Both TNG72 and TCS17 possess green leaf blades for their entire life. The anthocyanin contents of TNG72 and TCS17 leaf blades did not differ significantly although TNG72 was somewhat higher at both seedling and tillering stages (Fig. 1f) and may have already had anthocyanins at a level too low to be discerned visually. On the other hand, TNG72 had significantly higher ( $p < 0.05$ ) anthocyanin content in leaf sheaths than TCS17. Leaf sheath anthocyanin contents were estimated as 0.06  $\mu\text{mol/g}$  and 0.01  $\mu\text{mol/g}$  at the 4-leaf seedling stage; and 1.16  $\mu\text{mol/g}$  and 0.05  $\mu\text{mol/g}$  at active tillering for TNG72 and TCS17, respectively (Fig. 1f).

### Isolation of the Gene Conferring the Leaf Sheath Color

The  $F_2$  population of TNG72  $\times$  TCS17 was used to identify the gene conferring leaf sheath color by positional cloning. A total of 632  $F_2$  progenies included 473 and 159 individuals exhibited PSH and GSH, respectively. The segregation ratio of 2.9:1 (PSH: GSH) followed single-gene Mendelian inheritance, indicating that GSH was recessive. A randomly selected 46  $F_2$  plants with green leaf sheath were genotyped with 117 polymorphic markers distributed over the 12 rice chromosomes, permitting the gene to be coarsely mapped between CH0639 and RM276 on the short arm of chromosome 6



(Fig. 2). An additional 113  $F_2$  plants with GSH were genotyped with 3 more markers for fine mapping. The target interval was 162.04 kb, between CH0611 and RM253 encompassed by 3 BACs, OsJNBa00161019, OsJNBb0015B15, and P0529B09. By retrieving the Rice Genome Annotation Project database (RGAP, <http://rice.plantbiology.msu.edu>), it was found that one of 25 candidate genes, *Os06g10350*, annotated as a MYB family transcription factor, was the most likely gene. *Os06g10350* corresponded to *Os06g0205100*, annotated as transcription factor MYB6, *O. sativa* C1, a rice homolog of maize C1 in the Rice Annotation Project database (RAP, <http://rapdb.dna.affrc.go.jp/>).

TNG72 with PSH possessed a full-length *OsC1* DNA sequence encoding 272 amino acids. TCS17 with GSH had a 10-bp deletion in the R3 Myb domain, which was considered a core domain in exon 3 of *OsC1*. The 10-bp deletion resulted in a frame shift, translating only 207 amino acids and leading to a premature stop codon. In the reference genome sequence of Nipponbare on RGAP, a 3-bp deletion of part of the R3 Myb domain in exon 2 was noted, for which Nipponbare has GSH. The results revealed that an aberrant R3 Myb domain of anthocyanin regulatory C1 protein caused a deficit in anthocyanin synthesis in TCS17 and Nipponbare.

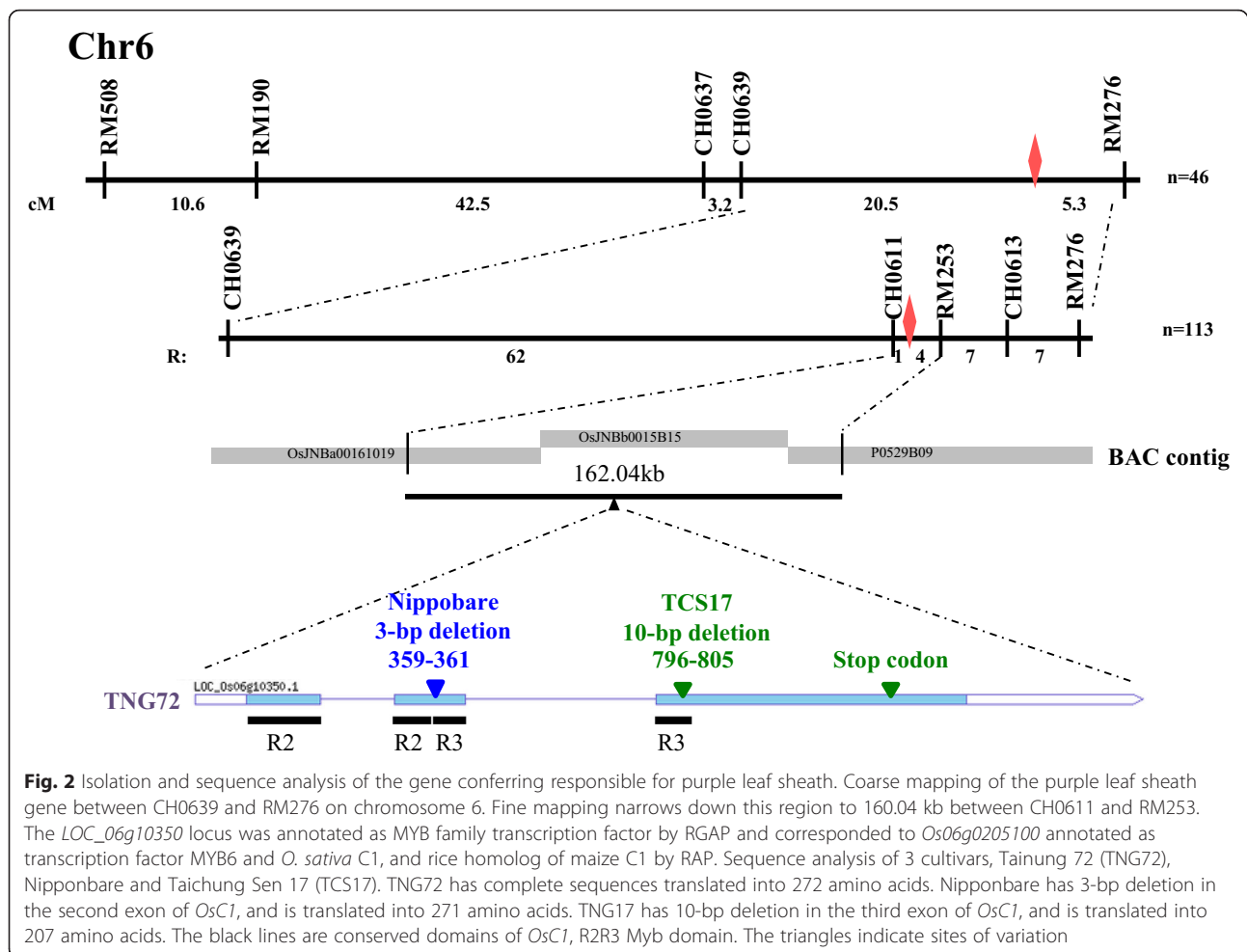
#### Natural Variation in Leaf Sheath Colors

Purple leaf sheath is a common phenotype in wild rice. In this study, among the 7 wild rices utilized in the *Oryza* Map Alignment Project (OMAP), all except *O. nivara* (103813) exhibit PHS. One accession of *O. rufipogon* collected from Taiwan also has purple leaf sheath. Nevertheless, the majority of cultivated rice has GSH.

Among the 43 accessions of *O. sativa* selected to study genetic diversity of *OsC1*, 17 *indica* and 12 *japonica* accessions possess green leaf sheath, while 5 *indica* and 9 *japonica* accessions possess purple leaf sheath (Table 1). Because of the intention to study allelic variation of *OsC1*, more *japonica* landraces exhibiting various degrees of purple color on the leaf sheath were included.

Various degrees of purple color were obvious among different rice accessions. Fourteen accessions, including 2 cultivars and 3 landraces of *indica* rice, and 1 cultivar and 8 landraces of *japonica* rice, ranged from light to dark purple. According to the RHS (Royal Horticultural Society) Color Charts 5th Edition, the leaf sheath of 6, 3, 3, and 2 accessions showed tyran rose, pansy purple, red purple, and blackish purple, respectively (Fig. 3a). The estimated anthocyanin content of the GSH cultivar, Tainung 67 (TNG67), was 0.03  $\mu\text{mol/g}$ . The anthocyanin contents of 14 PSH ranged from 1.04–42.77  $\mu\text{mol/g}$ , for which Chuan4 with tyran rose had the least while Kun Shan Wu Siang Keng (KSWSK) with blackish purple had the most. The 14 accessions could be divided into 7 groups with significant differences (LSD,  $p \leq 0.05$ ), in which color degrees were in proportion to anthocyanin contents (Fig. 3a).

One *japonica* GSH accession, TNG67, and 3 *japonica* PSH accessions, TNG72, Shang Chi Tsao Tao (SCTT) and KSWSK, exhibiting different degrees of purple color in the leaf sheath were used to evaluate gene expression of *OsC1*. At the active tillering stage, the estimated anthocyanin contents of these four accessions were significantly different (Fig. 3a). The relative gene expression of 3 accessions with purple leaf sheath was 18-fold higher than TNG67, a highly significant difference (Fig. 3b).



Nevertheless, there were no significant differences among the 3 PSH accessions by using LSD analysis.

#### Nucleotide Diversity and Haplotype Network of *OsC1*

Fifty accessions collected from 9 countries, mainly from Taiwan, were subjected to analysis of *OsC1* nucleotide variation and its haplotype network (Table 1). These 50 accessions could be classified into 5 groups, with Groups I-IV sharing high similarity of DNA sequence while Group V including 6 A-genome wild rices exhibited more allelic diversity (Fig. 4). For Group I, all 14 accessions shared high sequence similarity with TNG72. The majority of Group I belonged to *ssp. japonica* except 1 *indica* cultivar, Tainung Sen 20. The leaf sheaths of three accessions were tyran rose or pansy purple. Nevertheless, 8 accessions had GSH although *OsC1* function was predicted to be normal. Group II, consisting of 4 *indica* and 2 *japonica* accessions, shared the same sequence as SCTT. There were 5 SNPs that differentiated the group from the *OsC1* sequence of TNG72, and one SNP located at position 1,192 bp in exon 3 conferred nonsynonymous substitution of valine to alanine.

All 6 accessions had PSH with colors of tyran rose, pansy purple, and red purple. Group III, including 5 accessions, had one SNP in common at position 918 bp on exon 3, resulting in nonsynonymous substitution from proline to glutamine. All 5 accessions shared high sequence similarity to KSWSK. Three accessions, including 2 *japonica* landraces and *O. rufipogon* Taiwan Type 1, had blackish purple leaf sheath; while the other 2 had green leaf sheath. Group IV, consisting of 16 *indica* and 3 *japonica* accessions, had the same 10-bp deletion at position 796–805 bp as TCS17, causing frameshift mutation. All 19 accessions had green leaf sheath. Group V consisted of all A-genome wild rices except *O. rufipogon* Taiwan Type 1. The nucleotide sequences of the 6 wild rices were highly diverse, especially that of *O. meridionalis*. Five of the 6 wild rices had PSH with various degrees of purple color but *O. nivara* (103813) had GSH (Fig. 4).

Haplotype network analysis of *OsC1* gene sequences of these 50 accessions revealed 17 haplotypes with four major distinct groups comprised of Asian, African, Australian, and South American rice, in agreement with genealogical lineages. The 43 accessions of Asian cultivated

**Table 1** A total of 50 accessions for nucleotide diversity analysis

Taxon	Accession	Origin	Color of leaf sheath	
<i>O. sativa</i> ssp. <i>indica</i>				
cultivars	Tainung Sen 20	Taiwan	Purple	
	Jianungyu 892229	Taiwan	Green	
	Jianungyu 892234	Taiwan	Green	
	Taichung Sen 17	Taiwan	Green	
	Taichung Native 1	Taiwan	Green	
	IR 13525-118-3-2-2-2	The Philippines	Purple	
	IR64	The Philippines	Green	
	Hua Keng Sen 7	China	Green	
	Hu Han 15	China	Green	
	landrace	Cheng Ching Yu	Taiwan	Green
		Cheng Wu Chan	Taiwan	Green
		Chu Tzu	Taiwan	Green
		Dee Geo Woo Gen	Taiwan	Green
		Hua Lou	Taiwan	Green
Midon		Taiwan	Green	
Jinya-149		Taiwan	Purple	
G 124		India	Purple	
Shui Pai Tiao		China	Purple	
Fu P1 Sen		China	Green	
Hsia Men Chung		China	Green	
Liu Shih Jih Tsao		China	Green	
Yin Yu Tzu		China	Green	
<i>O. sativa</i> ssp. <i>japonica</i>				
cultivars	Tainung 72	Taiwan	Purple	
	Taichung 65	Taiwan	Green	
	Tainung 67	Taiwan	Green	
	Shinriki	Japan	Green	
	Asamurasaki	Japan	Purple	
	Kameji	Japan	Green	
	Nipponbare	Japan	Green	
	Nohrin 1	Japan	Green	
	landrace	Chuan4	Taiwan	Purple
		Kun Shan Wu Siang Keng	Taiwan	Purple
Tongsisai		Taiwan	Purple	
Warisanmochi 2		Taiwan	Purple	
Baridon		Taiwan	Green	
Chien Tzu Chu		Taiwan	Green	
Ch'ih K'o		Taiwan	Green	
Nobohai		Taiwan	Green	
Munagurusu		Taiwan	Green	

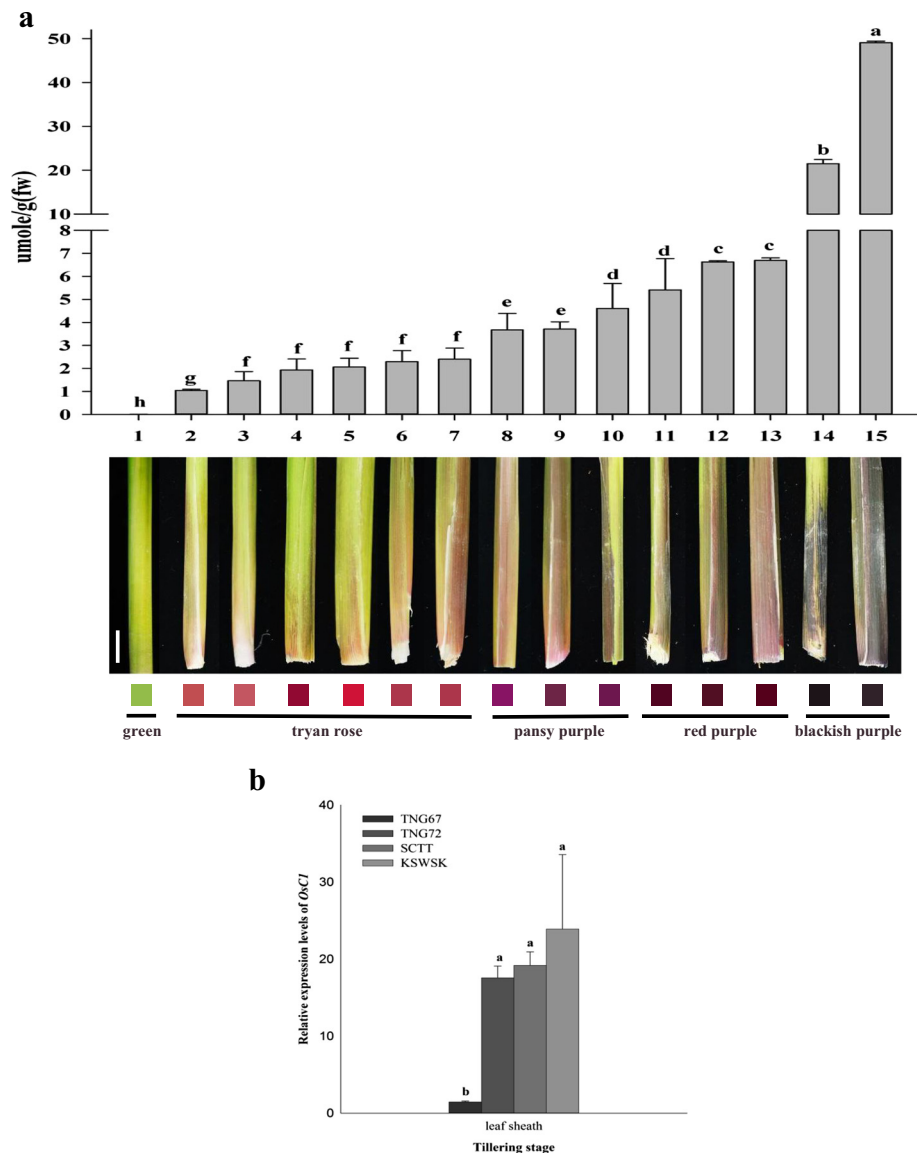
**Table 1** A total of 50 accessions for nucleotide diversity analysis (Continued)

	Ssall-Bye	Taiwan	Purple
	4233	China	Purple
	Shang Chi Tsao Tao	China	Purple
	Yen No	China	Green
<i>O. barthii</i>	10412	Cameroon	Purple
<i>O. glaberrima</i>	96717	Senegal	Purple
<i>O. glumaepatula</i>	105668	Brazil	Purple
<i>O. meridionalis</i>	105300	Australia	Purple
<i>O. nivara</i>	103813	China	Green
<i>O. nivara</i>	104683	India	Purple
<i>O. rufipogon</i>	Taiwan type 1	Taiwan	Purple

rice and their wild progenitors, 2 accessions of *O. nivara* and 1 accession of *O. rufipogon*, formed 13 interconnected haplotypes. The African cultivated species *O. glaberrima* and its wild progenitor, *O. barthii*, had different haplotypes, H14 and H15; however, these two haplotypes were connected. *O. meridionalis* (from Australia) and *O. glumaepatula* (South America) each formed independent unique haplotypes, H16 and H17, respectively (Fig. 5).

*OsC1* in the 7 A-genome wild rices, as expected, exhibited rich nucleotide diversity, especially in *O. meridionalis*, as revealed by haplotype number (H), haplotype diversity ( $H_d$ ), polymorphic sites (S), pairwise nucleotide difference ( $\pi_T$ ,  $\pi_{sil}$ ), and Watterson's estimators of  $\theta_T$  and  $\theta_{sil}$  (Table 2). Nucleotide diversity of *OsC1* in cultivated *O. sativa* was apparently reduced. In general, more genetic variation was detected in *indica* than *japonica* rice. Greater sequence polymorphism was noted in PSH than GSH accessions in both *indica* and *japonica* rice despite characterizing more GSH accessions. Up to 90 % reduction of genetic diversity in GSH was revealed by the average number of pairwise nucleotide differences per site based on the total number of polymorphic sites ( $\pi_T$ ); however, less reduction in silent sites was indicated by  $\pi_{sil}$  and  $\theta_{sil}$  (Table 2).

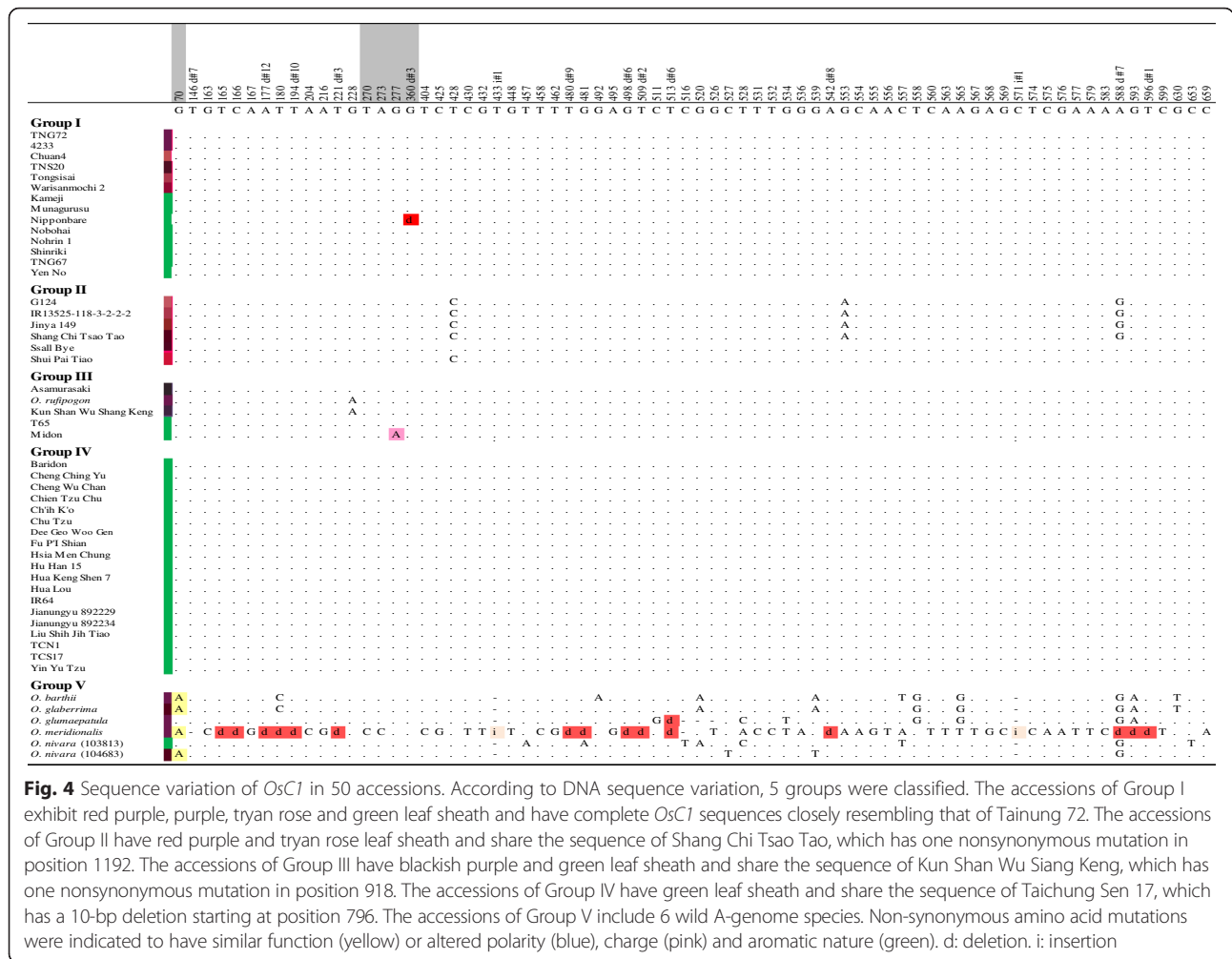
Tajima's D and Fu and Li's D & F were applied to test whether *OsC1* deviated from the neutral expectation of heterozygosity, but no values were statistically significant, indicating no evidence of strong selection (Table 3). The overall values were negative in *O. sativa* and the other A-genome species. Positive values, implying balancing selection, were noted in *indica* PSH accessions but not in *japonica* PSH accessions. Negative selection or purifying selection acting on *OsC1* could not be ruled out in *indica* GSH accessions ( $0.01 \leq p \leq 0.05$ ).



**Fig. 3** Anthocyanin accumulation and gene expression of *OsC1* in the leaf sheath. **(a)** Phenotypes and anthocyanin contents of 1 green leaf sheath (GSH) accession and 14 purple leaf sheath (PSH) accessions with various degrees of purple. The 15 accessions are Tainung 67 (TNG67) (1), Chuan4 (2), G124 (3), Warisanmochi2 (4), Shui Pai Tiao (5), Tonsisai (6), IR13525-118-3-2-2-2 (7), 4233 8 (8), Jinya-149 (9), Tainung 72 (TNG72) (10), Ssall-Bye (11), Tainung Sen 20 (12), Shang Chi Tsao Tao (SCTT) (13), Asamurasaki (14), and Kun Shan Wu Siang Keng (KSWSK) (15). Leaf sheath phenotypes are characterized according to the RHS (Royal Horticultural Society, 5th edition) Color Charts. **(b)** Gene expression of *OsC1* in the leaf sheath of 4 cultivars at tillering. Cultivars TNG67, TNG72, SCTT and KSWSK, exhibiting diverse degrees of purple color, were selected to evaluate gene expression by real-time PCR. Letters indicate differences at significance level  $p < 0.05$  by using LSD analysis

Asian rice could be divided two major subgroups, with Group A including two haplotypes (H1, H2) and Group B including 11 haplotypes (H3-H13) (Fig. 5). All accessions of Group A exhibited GSH, and haplotype H1 containing 15 *indica* accessions and 3 *japonica* landraces all had the same 10-bp deletion. One *indica* landrace, Cheng Ching Yu, has an additional single nucleotide substitution, resulting in another haplotype, H2 (Figs. 4 and 5). Group B contained 27 accessions, 11 GSH and

16 PSH with various degrees of purple. According to the haplotype network, 3 major clusters, B1-B3, could be resolved within Group B. The 12 accessions of Cluster B1 shared the same sequence with PSH TNG72 but 7 were phenotypically GSH. All 6 accessions of Cluster B2 (4 *indica* and 2 *japonica* landraces), expressed pansy purple and red leaf sheath, similar to SCTT. Cluster B3 had 2 GSH and 3 PSH accessions sharing high sequence similarity with KSWSK. Interestingly, *O. rufipogon* (Taiwan



Type 1) and KSWSK were embodied in one haplotype, H7. The *indica* cultivar TNS20 and the *japonica* cultivar Nipponbare had their own additional mutations and were separated into two distinct haplotypes, H4 and H11, respectively. The two accessions of *O. nivara* formed different haplotypes, H12 and H13, and were split from Cluster B1 because of diverse nucleotide sequences in both upstream and genic regions (Figs. 4 and 5).

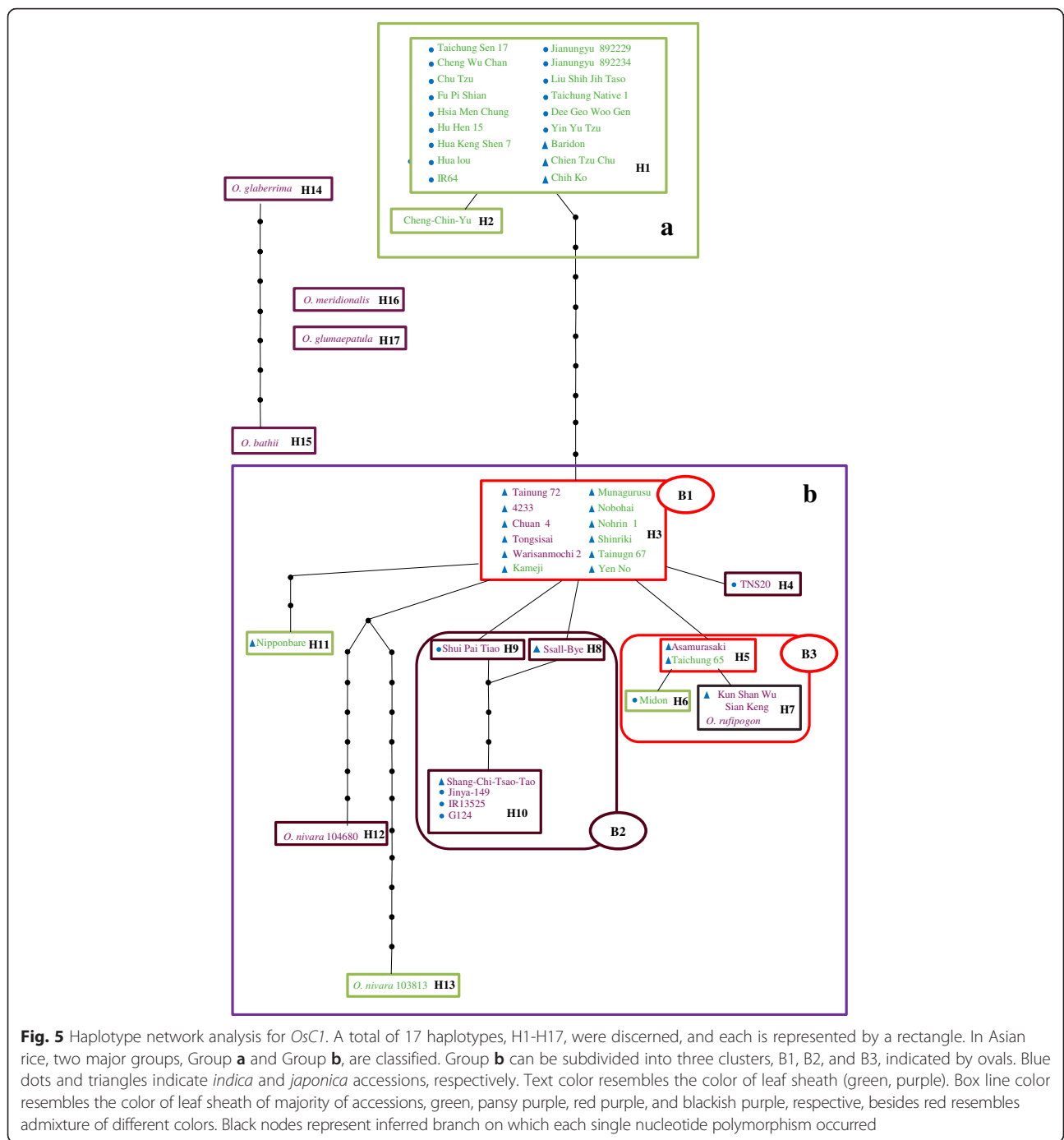
### Discussion

#### Allelic Diversity of *OsC1* Responsible for Variation in Leaf Sheath Colors

Plants accumulate anthocyanin in various tissues as an aid to survival and reproduction. In rice, anthocyanin is deposited in the root, leaf sheath, internode, leaf blade, lemma, palea, apiculus, stigma, and pericarp. Tissue-specific anthocyanin accumulation is common in numerous genotypes. Purple pigmented traits in different tissues do not always co-segregate (Sakamoto et al. 2001). However, purple leaf sheath, apiculus, and stigma cosegregated in the

F<sub>2</sub> population of PSH TNG72 × GSH TCS17 (Fig. 1) and other populations (Fan et al. 2008; Gao et al. 2011). Anthocyanin accumulation in the apiculus had previously been related to *OsC1* (Takahashi 1957; Saitoh et al. 2004) but in the leaf sheath was thought to be inherited by polygenes or a single gene (Fig. 2; Fan et al. 2008; Wang et al. 2009; Gao et al. 2011).

*OsC1*, conferring leaf sheath color, was isolated from the F<sub>2</sub> population of TNG72 × TCS17 by positional cloning (Fig. 2), differing from *PSH1* (t) but the same as a locus identified from the somaclonal line Z418 (Wang et al. 2009; Gao et al. 2011). The PSH allele was dominant to GSH in crosses between natural germplasm or mutant lines. PSH TNG72 had a full length *OsC1* allele encoding 272 amino acids while GSH TCS17 had a 10-bp deletion in the R3 Myb domain in exon 3, resulting in truncated translations of 207 amino acids. GSH cultivar Nipponbare had another mutated allele, a 3-bp deletion in the R3 Myb domain in exon 2 (Fig. 2). Among the 50 accessions analyzed herein, the 10-bp deletion was found in 17 *indica* landraces and improved cultivars



**Fig. 5** Haplotype network analysis for *OsC1*. A total of 17 haplotypes, H1-H17, were discerned, and each is represented by a rectangle. In Asian rice, two major groups, Group **a** and Group **b**, are classified. Group **b** can be subdivided into three clusters, B1, B2, and B3, indicated by ovals. Blue dots and triangles indicate *indica* and *japonica* accessions, respectively. Text color resembles the color of leaf sheath (green, purple). Box line color resembles the color of leaf sheath of majority of accessions, green, pansy purple, red purple, and blackish purple, respective, besides red resembles admixture of different colors. Black nodes represent inferred branch on which each single nucleotide polymorphism occurred

from Taiwan, China, and the Philippines, and 2 *japonica* landraces from Taiwan (Fig. 4). The 10-bp deletion was conserved in 17 indigenous varieties in Northeast India and only observed in *indica* varieties from Taiwan, China, India, and Indonesia (Saitoh et al. 2004; Choudhury et al. 2014).

Ten GSH accessions without the 10-bp deletion of *OsC1* had other mutated alleles or genes, leading to absence of anthocyanin pigmentation. Deletions of 3-bp

in exon 2 and 2-bp in exon 3, both in the R3 Myb domain, were found in *japonica* rice from Japan and China, respectively (Figs. 2 and 4; Saitoh et al. 2004). Two GSH accessions, Midon and T65, had an amino acid substitution at position 918. However, three other accessions with the same substitution, Asamurasaki, KSWSK, and *O. rufipogon*, had blackish purple leaf sheath. Thus, mutations in genes other than *OsC1* were suggested to also eliminate anthocyanin pigmentation, a hypothesis



**Table 2** Nucleotide divergence of *OsC1*

Taxon	N	H	$H_d$	S	$\pi_T$	$\theta_T$	$\pi_{sil}$	$\theta_{sil}$
<i>O. sativa</i>	43	10	0.481	11	0.00104	0.00196	0.00146	0.00206
<i>spp. indica</i>	23	7	0.522	10	0.00129	0.00208	0.00163	0.00201
purple leaf sheath	5	4	0.900	7	0.00275	0.00257	0.00325	0.00284
green leaf sheath	18	3	0.216	3	0.00026	0.00067	0.00016	0.00043
<i>spp. japonica</i>	20	5	0.442	7	0.00074	0.00152	0.00127	0.00251
purple leaf sheath	9	5	0.722	7	0.00144	0.00196	0.00246	0.00326
green leaf sheath	11	2	0.182	1	0.00014	0.00026	0.00027	0.00051
Total numbers of purple leaf sheath	14	8	0.857	9	0.00229	0.00216	0.00336	0.00279
Total numbers of green leaf sheath	29	4	0.200	3	0.00021	0.00059	0.00020	0.00038
A-genome wild rice	7	7	1.000	94	0.02584	0.03478	0.03568	N.A.
A-genome wild rice without <i>O. meridionalis</i>	6	6	1.000	37	0.01283	0.01325	0.01544	0.01612

$N$  total number of sequences,  $H$  total number of haplotypes at individual loci in each taxon,  $H_d$  haplotype diversity,  $S$  total number of polymorphic sites,  $\pi_T$  average number of pairwise nucleotide difference per site calculated based on the total number of polymorphic sites,  $\theta_T$  Watterson's estimator of  $\theta$  per base pair calculated based on the total number of polymorphic sites,  $\pi_{sil}$  average number of pairwise nucleotide difference per site calculated based on silent sites,  $\theta_{sil}$  Watterson's estimator of  $\theta$  per base pair calculated based on silent sites, N.A. Not applicable. When the proportion of differences is higher than 0.75, the Jukes and Cantor correction cannot be computed

that was supported by seven *japonica* accessions having the same *OsC1* coding sequence as PSH TNG72 (Fig. 4). *OsC1*, containing an R2R3 Myb domain, is thought to function as a transcription factor, regulating other genes involved in anthocyanin synthesis and enhancing another transcription factor, bHLH, regulating anthocyanin structural gene DFR (Dooner et al. 1991; Ithal and Reddy 2004). Failure of anthocyanin synthesis could result from malfunction of *OsC1* or any downstream proteins: for example, *PSH1* (t) conferring purple leaf sheath was mapped on chromosome 1 (Wang et al. 2009). One of two activator genes, *OsB1* and *OsB2* encoding basic helix-loop-helix (bHLH) transcription factors, incorporated with maize *C1* could induce anthocyanin synthesis in the aleurone layer; however, the lack of function of these two genes in T65 resulted in green leaf blade and leaf sheath (Sakamoto et al. 2001). One *japonica* GSH

accession, TNG67 which is a descendant of T65, had significantly reduced expression of *OsC1* despite having a full length amino acid encoding sequence (Fig. 3), suggesting feedback regulation of gene expression.

Fourteen PSH accessions exhibited variation in intensity of purple coloring of the leaf sheath, classified as tyran rose, pansy purple, red purple, and blackish purple. The anthocyanin contents of these 14 PSH accessions differed significantly (Fig. 3). However, *OsC1* gene expression among three accessions with diverse anthocyanin contents, TNG72, SCTT, and KSWSK, showed only slight and non-significant differences. These three accessions represented three different haplotypes, H3, H7, and H10, respectively (Fig. 5). Allelic variation of *OsC1* accounted for anthocyanin accumulation and pigmentation in the leaf sheath, for which nonsynonymous mutations at the C-terminal domain were found herein.

**Table 3** Neutrality tests of *OsC1*

Taxon	N	Tajima's D	Fu and Li's D	Fu and Li's F
<i>O. sativa</i>	43	-1.41027	-1.32676	-1.59411
<i>spp. indica</i>	23	-1.27639	-1.07742	-1.32111
purple leaf sheath	5	0.49788	0.49788	0.51896
green leaf sheath	18	-1.71304*	-2.30153*	-2.45793*
<i>spp. japonica</i>	20	-1.69169*	-1.82926	-2.06967
purple leaf sheath	9	-1.18930	-1.02031	-1.18277
green leaf sheath	11	-1.12850	-1.28946	-1.39919
Total number of purple leaf sheath	14	0.23043	0.02189	0.08865
Total number of green leaf sheath	29	-1.53343	-1.45773	-1.71236
A-genome wild rice	7	-1.33215	-1.32405	-1.46956
A-genome wild rice without <i>O. meridionalis</i>	6	-0.22043	-0.17585	-0.20295

\*The values of Tajima's D test and Fu and Li's D&F test are at significance levels of  $0.10 < P < 0.05$

The C terminus of R2R3 Myb transcription factors is an activation and repression domain (Dubos et al. 2010). Allelic variation of *OsC1* and *Purple leaf* (*Pl*) coding regions also caused diverse intensities of purple apiculus pigmentation (Sakamoto et al. 2001; Saitoh et al. 2004). Allelic variation resulting in diversified phenotypes is very common, as exemplified by the great impact of different alleles of *Wx* and starch synthesis related genes on rice grain appearance, cooking and eating quality (Tian et al. 2009; Zhang et al. 2012; Wu et al. 2015). However, other genes participating in anthocyanin synthesis and environmental effects cannot be neglected (Sakamoto et al. 2001). Anthocyanin accumulation is frequently responsive to abiotic stresses, such as UVB radiation, temperature, and soil acidity (Reddy et al. 2004). A *cis*-element related to light regulation (-10PEHVPSBD, Thum et al. 2001) was found in the promoter region of *OsC1* in TNG72 by using the PLACE database (<https://sogo.dna.affrc.go.jp/>, Higo et al. 1999). We noted that the leaf sheath accumulated more anthocyanin when rice plants were grown in soil under natural light than in hypotonic solution in the growth chamber.

#### The Selection and Genealogy of *OsC1*

Traits considered part of domestication syndromes are favored by artificial selection during domestication, and consequently distinguish cultivated plants from their wild progenitors. During domestication, genetic diversity of whole genomes can be dramatically reduced because of the 'genetic bottleneck' effect of selecting a few individuals as a founder population. The genetic diversities of genes contributing to domestication syndromes and improvement traits are often accompanied by artificial selection for production and culture, and the geographic distribution of alleles might be altered by human migration (Kovach et al., 2007; Kovach and McCouch 2008; Olsen and Wendel 2013). In rice, *OsC1* conferring purple leaf sheath, apiculus, and stigma has been suggested as a domestication gene (Choudhury et al. 2014). Wild types are generally purple while cultivars tend to be green, although exceptions exist as noted above. The nucleotide diversity of *OsC1* was higher in *O. rufipogon* than in *O. sativa*, and *OsC1* of the haplotype of Asian cultivated accessions showed evidence of selection (Saitoh et al. 2004). In addition to perennial *O. rufipogon* and annual *O. rufipogon* (also called *O. nivara*), *OsC1* sequences of 4 other wild species, *O. glaberrima*, *O. barthii*, *O. glumaepatula*, and *O. meridionalis* which had not been previously reported were subjected to allelic variation analysis. Sequence variation existed both in noncoding and coding regions. While these species had PSH, a few non-synonymous amino acid substitutions were detected (Fig. 4). *O. meridionalis* (from Australia) had extremely divergent *OsC1* DNA sequence, forming a distinct haplotype by itself,

and all parameters regarding nucleotide divergence of *OsC1* decreased when it was not included in analysis (Table 2). African cultivated species, *O. glaberrima*, its progenitor *O. barthii*, and South American species, *O. glumaepatula*, also had their own specific sequences and formed 2 different haplotypes (Figs. 4 and 5). On the other hand, *OsC1* in Asian rice shared sequence similarity and formed 13 interconnected haplotypes. Allelic variation of *OsC1* in Asian, African, South American, and Australian rice reflected independent mutation without gene flow because of their geographic distribution. Two accessions of *O. nivara* and one accession of *O. rufipogon* together with 24 accessions of *O. sativa* were in Group B. The genealogy of *OsC1* in the collected Asian rice accessions herein and other studies showed no distinct correlation with geographic distribution. The 10-bp deletion conferring GSH was prevalent in *indica* landraces and improved cultivars and rare in *japonica* accessions from many countries (Saitoh et al. 2004; Choudhury et al. 2014). The accessions in haplotypes H3 and H10 were collected from several countries (Fig. 5). Gene flow due to human activities might be an important factor in the geographic distribution of *OsC1* alleles.

Purple leaf sheath, stigma, and apiculus are widespread in wild forms and often found in landraces, while green leaf sheath as well as colorless stigma and apiculus are common in modern cultivars. *OsC1* nucleotide divergence was up to 90 % lower in GSH than PSH although twice as many GSH accessions were studied (Table 2). However, no values from three neutral tests, Tajima's D, Fu and Li's D & F, were significantly differently from neutral expectations, indicating that *OsC1* had not been subjected to selection—a conclusion that is supported by findings regarding *OsC1* in indigenous rice varieties in Northeast India (Choudhury et al. 2014). Genetic diversity was higher in *indica* than *japonica* accessions both in 14 PSH and 20 GSH accessions, revealed by parameters reflecting nucleotide segregation at total polymorphic sites and silent sites,  $\pi_T$ ,  $\theta_T$ ,  $\pi_{sil}$ , and  $\theta_{sil}$  (Table 2, Fig. 4). This phenomenon is congruent with evidence that genetic diversity is larger in subspecies *indica* than *japonica* (Garris et al. 2005).

Selection might not be a driving force for reducing genetic diversity in GSH. There is no significant evidence that *OsC1* deviated from neutral expectations in *indica* PSH or *japonica* PSH and GSH accessions. Nevertheless, *indica* GSH accessions might be experiencing relaxed purifying selection, indicated by neutrality tests at significance levels of  $0.10 < P < 0.05$  (Table 3). Unlike other domestication syndromes directly related to productivity and other desirable traits that were selected for particular purposes over several thousand years, GSH might have been selected unintentionally.

The Asian cultivated species, *O. sativa*, evolved from Asian wild rice progenitors *O. nivara* (annual) and *O. rufipogon* (perennial). *O. sativa* was domesticated from divergent wild populations about 10,000 years ago and diversified into two major subspecies, *indica* and *japonica*, subsequently being subjected to a long period of natural and artificial diversifying selection (Gross and Zhao 2014). *Indica* and *japonica* subspecies are distinguishable in morphology and physiology, already recognized as Hsien (long grain) and Keng (short grain) in the Han dynasty, China, over 2,000 years ago (Oka 1988; Callaway 2014). Numerous genes related to differentiation between these two subspecies experienced mutation and diversifying selection, e.g., *Phr1* responsible for phenol reaction; and *GS3*, *qSW5* and *GS5* responsible for grain shape (Yu et al. 2008; Lu et al. 2013). The 10-bp deletion conferring GSH was prevalent in *indica* landraces and improved cultivars and rare in *japonica*, but not in its progenitor species. The *OsCI* allele with 10-bp deletion was suggested to have originated and been an early target of domestication in subspecies *indica* (Figs. 4 and 5; Saitoh et al. 2004; Choudhury et al. 2014). In addition, 3-bp and 2-bp deletions in exon 2 and exon 3 of the R3 Myb domain were found in *japonica* rice from Japan and China, respectively. These two alleles were independent from the gene lineage of *indica*, which suggested mutation after subspecies divergence (Figs. 2 and 4; Saitoh et al. 2004).

Although reproductive barriers such as hybrid sterility and hybrid breakdown impede gene flow between cultivated rice and its wild progenitors, numerous interspecific crosses and successful introgressive hybridizations have been performed to unravel useful alleles and genes of wild species. Gene flow confounded with selection has been revealed at the genome level (Zhao et al. 2010; He et al. 2011; Yang et al. 2011) and in domestication-related genes including *Wx*, *GS3*, *SD1*, and *qSH1* (Yamanaka et al. 2004; Konishi et al. 2006; Takano-Kai et al. 2009; Asano et al. 2011).

The genealogy of *OsCI* suggests some gene flow events. *O. rufipogon* (Taiwan type 1) had an *OsCI* sequence identical to that of Taiwan landrace KSWSK. One SNP aligned to *japonica* Asamursaki and Taichung 65 and *indica* Midon, accessions that were clustered in Group III and haplotype B3 (Fig. 4 and 5). Two *O. nivara* accessions also shared similar sequences to most *japonica* accessions and were classified as haplotype B1. Thus, *OsCI* in three Asian wild accessions closely resembled that of *japonica* but not *indica*, a finding which might support the hypothesis that *japonica* and *indica* were domesticated independently from *O. rufipogon* (Yang et al. 2011; Wei et al. 2014). Gene flow between subspecies was not rare, as revealed by genealogy of *OsCI* in 23 *indica* and 20 *japonica* accessions. In haplotype H1, three *japonica* landraces from Taiwan

also possessed the 10-bp deletion specific to *indica*; in haplotype H10, one *japonica* landrace (Shang Chi Tsao Tao) and three *indica* accessions (Taiwan landrace Jinya-149, India landrace G124 and improved line IR1535) had the same allele (Figs. 4 and 5).

The genealogy of *OsCI* might not be in agreement with rice phylogeography because of human behavior. In Taiwan, *O. rufipogon* and *O. nivara* were once found in several swamp sites (Chang 1976) but, unfortunately, all habitats were destroyed several decades ago. Archaeological evidence shows that tropical *japonica* or *indica* had been cultivated over 5,000 years by ancient indigenous peoples in Taiwan (Hsieh et al. 2011). In the early 17th century, numerous Chinese migrated and carried many landraces (mostly *indica*) from coastal regions of Fujian and Guangdong Provinces of China to Taiwan. By the early 20th century, 1,197 *indica* accessions were identified officially in Taiwan, and 1,256 *japonica* accessions were introduced from Japan (Iso 1964). More than 1,000 of these accessions were deposited in The T.T. Chang Genetic Resources Center at the International Rice Research Institute (IRRI), the Philippines. Taiwanese rice germplasm was thus an admixture of indigenous wild species, landraces, and introduced germplasm from China and Japan; and was subsequently spread over Southeast Asia via the germplasm deposited in IRRI. As a result, introgression of *OsCI* may have occurred by hybridization between subspecies and both artificial and natural selection, clouding the true genealogy of *OsCI*.

## Conclusion

We used positional cloning to isolate a PSH gene, revealing independent 10- and 3-bp deletions in the R2R3 Myb domain of *OsCI* that occurred in different lineages. Allelic diversity resulted in different *OsCI* protein function rather than gene expression levels, accounting for various degrees of anthocyanin accumulation and associated intensities of purple color. Allelic variation of *OsCI* among wild species, landraces, and cultivars revealed greatly reduced genetic variation in GSH phenotypes, but there was little evidence that *OsCI* had experienced recent selection, suggesting that it may have been enriched in cultivated forms by chance. Introgressions of *OsCI* between subspecies *indica* and *japonica* were frequent through natural and artificial hybridization. Because of admixture of rice germplasm by artificial and natural selection in Taiwan, the genealogy of *OsCI* might not be accurately reflected by the current distribution of alleles through lineage diversification.

## Materials and methods

### Plant Materials

Tainung 72 (TNG72), registered by Taiwan Agriculture Research Institute in 1987, is known as an aromatic elite

cultivar. Taichung Sen 17 (TCS17), registered by Taichung District Agricultural Research and Extension Station in 1984, is one of two leading *indica* varieties in Taiwan. An F<sub>2</sub> population of 624 individuals from a cross between these genotypes was planted in a paddy field at the Chiayi Agricultural Experiment Station (CAES), Taiwan, on January 21, 2008. The ratio of purple to green leaf sheath in the F<sub>2</sub> population was 2.92: 1 (465:159), indicating single gene inheritance with green sheath being recessive. The 159 F<sub>2</sub> individuals possessing GSH, which were predicted to be recessive homozygotes, were subjected to linkage analysis.

A panel of 50 accessions obtained from the National Plant Genetic Resources Center (NPGRC), Taiwan and The *Oryza* Map Alignment Project (OMAP) were used to reveal DNA sequence diversity of the gene conferring leaf sheath color. As detailed in Table 1, the panel included: (a) 43 accessions of Asian cultivated rice, comprised of 21 ssp. *japonica* accessions (8 cultivars, 13 landraces) and 22 ssp. *indica* accessions (9 cultivars, 13 landraces), and (b) 7 accessions of A-genome wild rice, including *O. rufipogon* (Taiwan type 1), *O. glaberrima* (96717), *O. barthii* (10412), *O. glumaepatula* (105668), *O. meridionalis* (105300), and *O. nivara* (2 accessions: 103813, 104680). Geographically, 26 accessions were from Taiwan (12 *indica*, 13 *japonica*, 1 *O. rufipogon*), 11 from China (7 *indica*, 3 *japonica*, 1 *O. nivara*), 5 from Japan (5 *japonica*), 2 from the Philippines (2 *indica*), 2 from India (1 *indica*, 1 *O. nivara*), and 1 each from Australia (*O. meridionalis*); Cameroon (*O. barthii*); Brazil (*O. glumaepatula*) and Senegal (*O. glaberrima*). Regarding leaf sheath color, 20 and 30 accessions exhibited purple and green, respectively.

#### Measurement of Anthocyanin Contents

Total anthocyanin contents of leaf sheath and leaf blade harvested at four-leaf seedling and active tillering stages were quantified as described by Padmavati et al. (1997). Briefly, 0.5 g of fresh tissue was homogenized with extraction buffer and 1 % HCl/Methanol (V/V), for 24 h at 4 °C with occasional shaking. The supernatant was saved after centrifugation at 10,000 rpm for 15 min at 4 °C and absorbance at 530 nm measured using a spectrophotometer (Metertech SP8001, Taiwan). The relative contents of anthocyanin were estimated from A<sub>530</sub> using an mM extinction coefficient of 31.6 (Padmavati et al. 1997).

#### Genetic and Physical Mapping of the Gene Conferring Purple Sheath

For linkage analysis, 159 GSH F<sub>2</sub> progeny, predicted to be recessive homozygotes, were genotyped with 121 polymorphic markers, including 67 SSR and 54 indel markers distributed over the 12 rice chromosomes. SSR primer sequences were retrieved from Gramene (<http://www.gramene.org>). The 54 indel markers were newly

designed based on sequence divergence between *Nipponbare* (*japonica*) and 93–11 (*indica*) (Additional file 1: Table S1). DNA extraction and genotype assays of PCR-based markers was as described (Hsu et al. 2014). The linkage map was constructed with a LOD threshold of 3.5 by using MapMaker Exp 3.0 (Lander et al. 1987). The physical map of the target interval containing candidate genes was retrieved from the Rice Genome Annotation Project (RGAP, <http://rice.plantbiology.msu.edu>).

#### Gene Expression of *OsC1*

Four accessions displaying various degrees of PSH were subjected to quantitative real-time PCR (qPCR) to investigate the relationship of *OsC1* gene expression with anthocyanin contents. These four accessions were *japonica* rice, Tainung 67 (TNG67, green), Tainung 72 (TNG72, light purple), Shang Chi Tsao Tao (SCTT, purple), and Kun Shan Wu Siang Keng (KSWSK, dark purple). Total RNA was extracted by using TRIZOL (Invitrogen, USA), and contaminating DNA was removed by using a TURBO DNA-free kit (Ambion, USA), according to manufacturer's protocols. Primers of *OsC1* cDNA for qPCR were designed based on the gene sequence of GenBank accession number Y15219, and were forward primer: 5'-CAACgAgCTggTTTgAggCggT-3' and reverse primer: 5'-TgAgAgACCACCTgTTgCCgAg-3'. Quantitation of gene expression and calculation of relative gene expression were as described (Hsu et al. 2014).

#### Genetic Diversity Assessment

A panel of 50 accessions (above) possessing a range of leaf sheath colors were used to assess genetic diversity of *OsC1*. The *OsC1* DNA sequences of 1,311 bp were amplified by use of KOD plus *Taq* polymerase (TOYOBO, Japan) and Kapa Hifi DNA polymerase with three sets of primers, set 1 (F: 5'-ACATCgTACggggCTACAAg-3', R: 5'-AgCgTTAgCCAgCTTCAAAT-3'), set 2 (F: 5'-ACTATCTCCggCCTAACATCAA-3', R: 5'-TAgTAGTCgCAGTCgACgTC-3'), and set 3 (F: 5'-ATgTTgTCAggTggTCTCTC-3', R: 5'-CACgTT CATgCAACCTTTg-3'). The amplicons were sequenced using an ABI3730 DNA Analyzer following manufacturer's protocol. The sequences were assembled and then subjected to multiple alignments using DNAMAN ver. 6.0 (Lynnon Biosoft) and consequently MAGA ver. 6.0 (Tamura et al. 2013).

Genetic divergence of the *OsC1* gene region was estimated with  $\pi$  (average nucleotide diversity per site, Nei and Li 1979) and  $\theta_w$  (number of segregating sites, Watterson 1975) according to DnaSP version 5.10.1 (Librado and Rozas 2009). Two neutrality tests, 1) Tajima's D and 2) Fu and Li's D\* & F\*, were conducted to investigate whether there was any deviation from neutrality by using DnaSP version 5.10.1. Tajima's D estimates the difference between the mean pairwise differences ( $\pi$ ) and Watterson's estimator

( $\theta_w$ ) (Tajima 1989). Fu and Li's  $D^*$  &  $F^*$  test reveals the discrepancy between the number of polymorphic sites in external and internal groups (Fu and Li 1993). A haplotype network, based on the probability of parsimony and calculated for pairwise differences until it exceeds 0.95, was constructed using TCS1.21 (Templeton et al. 1992).

## Additional file

**Additional file 1: Table S1.** Newly designed markers used in coarse mapping. (DOCX 17 kb)

## Competing Interests

The authors declare that they have no competing interests.

## Authors' Contributions

YW and YL designed and supervised the research; YW conducted the genetic materials and field trials; HC, and CH performed the experiments; HC, AH and YL analyzed the data. HC, AH, and YL wrote the manuscript. All authors read and approved the final manuscript.

## Acknowledgments

This work was supported by the National Science Council (NSC 100-2313-B-002-055, NSC 101-2313-B-002-005) of Taiwan. The authors are thankful to Drs. Yue-ie C. Hsing and Yuan-chin Tsai for providing leaf tissue of wild rice species and Su-chen Kuo for field management. We also appreciate the services of Professor Hawkeye, LLC for critical review and English editing.

## Author details

<sup>1</sup>Department of Agronomy, National Taiwan University, Taipei, Taiwan. <sup>2</sup>Department of Agronomy, Chiayi Agricultural Experiment Station, Taiwan Agricultural Research Institute, Chiayi, Taiwan. <sup>3</sup>Department of Life Science, Fu-Jen Catholic University, Xinbei, Taiwan. <sup>4</sup>Department of Agricultural Chemistry, National Taiwan University, Taipei, Taiwan.

Received: 13 November 2015 Accepted: 19 February 2016

Published online: 27 February 2016

## References

- Asano K, Yamasaki M, Takuno S, Miura K, Katagiri S, Ito T, et al (2011) Artificial selection for a green revolution gene during japonica rice domestication. *Proc Natl Acad Sci USA* 108:11034–11039
- Asano K, Takashi T, Miura K, Qian Q, Kitano H, Matsuoka M et al (2007) Genetic and molecular analysis of utility of *sd1* alleles in rice breeding. *Breeding Sci* 57:53–58
- Brenda WS (2001) Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol* 126:485–493
- Callaway E (2014) Domestication: the birth of rice. *Nature* 514:558–559
- 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* 1:175–183
- Chang TT (1976) The origin, evolution, cultivation, dissemination, and diversification of Asian and African rices. *Euphytica* 25:425–441
- Chalker-Scott L (1999) Environmental significance of anthocyanins in plant stress responses. *Photochem Photobiol* 70:1–9
- Choudhury BI, Khan ML, Dayanandan S (2014) Patterns of nucleotide diversity and phenotypes of two domestication related gene (*OsC1* and *Wx*) in indigenous rice varieties in northeast India. *BMC Genet* 15:71–82
- Dooner HK, Robbins TP, Jorgensen RA (1991) Genetics and development control of anthocyanin biosynthesis. *Annu Rev Genet* 25:173–199
- Dubos C, Stracke R, Grotewold E, Weissshaar B, Martin C, Lepiniec L (2010) MYB transcription factors in Arabidopsis. *Trends Plant Sci* 15:573–581
- Fan FJ, Fan YY, Du JH, Zhuang JY (2008) Fine mapping of *C* (Chromogen of anthocyanin) gene in rice. *Rice Sci* 15:1–6
- Fu YX, Li WH (1993) Statistical tests of neutrality mutations. *Genetics* 133:693–709
- Gao D, He B, Zhou Y, Sun L (2011) Genetic and molecular analysis of a purple sheath somaclonal mutant in *japonica* rice. *Mol Breeding* 7:73–83
- Gross BL, Zhao Z (2014) Archaeological and genetic insights into the origins of domesticated rice. *Proc Natl Acad Sci U S A* 111:6190–6197
- Garris AJ, Tai TH, Coburn J, Kresovich S, McCouch S (2005) Genetic structure and diversity in *Oryza sativa* L. *Genetics* 169:1631–1638
- Hadagal BN, Manjunath A, Goud JV (1980) Linkage of genes for anthocyanin pigmentation in rice (*Oryza sativa* L.). *Euphytica* 30:747–754
- He Z, Zhai W, Wen H, Tang T, Wang Y, Lu X et al (2011) Two evolutionary histories in the genome of rice: the roles of domestication genes. *PLoS Genet* 7:e1002100
- Higo K, Ugawa Y, Iwamoto M, Korenaga T. (1999) Plant cis-acting regulatory DNA elements (PLACE) database. *Nucleic Acids Res.* 27:297–300
- Hsieh JS, Hsing YI, Hsu TF, Li JK, Li KT, Tsang CH (2011) Studies on ancient rice-where botanists, agronomists, archeologists, linguists, and ethnologists meet. *Rice* 4:178–183
- Hsu YC, Tseng MC, Wu YP, Lin MY, Wei FJ, Hwu KK et al (2014) Genetic factors responsible for eating and cooking qualities of rice grains in a recombinant inbred population of an inter-subspecific cross. *Mol Breeding* 34:655–673
- Iso E (1964) Talks on Horai rice. Amayomikai, Yamakuchi, Japan, p 89, in Japanese
- 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* 166:1503–1513
- Koes R, Verweij W, Quattrocchio F (2005) Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends Plant Sci* 10:1360–1385
- Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T et al (2006) An SNP caused loss of seed shattering during rice domestication. *Science* 312:1392–1396
- Kovach MJ, Calingacion MN, Fitzgerald MA, McCouch SR (2009) The origin and evolution of fragrance in rice. *Proc Natl Acad Sci U S A* 106:14444–14449
- Kovach MJ, McCouch SR (2008) Leveraging natural diversity: back through the bottleneck. *Curr Opin Plant Biol* 11:193–200
- Kovach MJ, Sweeney MT, McCouch SR (2007) New insights into the history of rice domestication. *Trends Genet* 23:578–587
- Lander ES, Green P, Abrahamson J, Barlow A, Day MJ, Lincoln SE et al (1987) MapMarker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1:174–181
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452
- Larson G, Piperno DR, Allaby RG, Purugganan MD, Andersson L, Arroyo-Kalin M et al (2014) Current perspectives and the future of domestication studies. *Proc Natl Acad Sci U S A* 111:6139–6146
- Lu L, Shao D, Qiu X, Sun L, Yan W, Zhou X et al (2013) Natural variation and artificial selection in four genes determine grain shape in rice. *New Phytol* 200:1269–1280
- Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci U S A* 76:5269–5273
- Oka HI (1988) Origin of cultivated rice. Japan Sci Soc Elsevier, Tokyo
- Olsen KM, Purugganan MD (2002) Molecular evidence on the origin and evolution of glutinous rice. *Genetics* 162:941–950
- Olsen KM, Wendel JF (2013) A bountiful harvest: genomic insights into crop domestication phenotypes. *Annu Rev Plant Biol* 64:47–70
- Padmavati M, Sakthivel N, Tahara KV, Reddy AR (1997) Differential sensitivity of rice pathogens to growth inhibition by flavonoids. *Phytochem* 46:499–502
- Reddy KR, Kakani VG, Zhao D, Koti S, Gao W (2004) Interactive effects of ultraviolet-B radiation and temperature on cotton physiology, growth, development and hyperspectral reflectance. *Photochem Photobiol* 79:416–427
- 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
- Reddy VS, Goud KV, Sharma R, Reddy AR (1994) Ultraviolet-B-responsive anthocyanin production in a rice cultivar is associated with a specific phase of phenylalanine ammonia lyase biosynthesis. *Plant Physiol* 105:1059–1066
- Reddy VS, Scheffler BE, Wienand U, Wessler SR, Reddy AR (1998) Cloning and characterization of the rice homologue of the maize *C1* anthocyanin regulatory gene. *Plant Mol Biol* 36:497–498
- Saitoh K, Onishi K, Mikami I, Thidar K, Sano Y (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<sup>W</sup>* 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

- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30:2725–2729
- Takahashi ME (1957) Analysis on apiculus color genes essential to anthocyanin coloration rice. *J Facu Agr Hokkaido Univ* 50:266–362
- Takano-Kai N, Jiang H, Kubo T, Sweeney M, Matsumoto T, Kanamori H et al (2009) Evolutionary history of *GS3*, a gene conferring grain length in rice. *Genetics* 182:1323–1334
- Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III Cladogram estimation. *Genetics* 132:619–633
- Thum KE, Kim M, Morishige DT, Eibl C, Koop HU, Mullet JE (2001) Analysis of barley chloroplast *psbD* light-responsive promoter elements in transplastomic tobacco. *Plant Mol Biol* 47:353–366
- Tian Z, Qian Q, Liu Q, Yan M, Liu X, Yan C et al (2009) Allelic diversities in rice starch biosynthesis lead to a diverse array of rice eating and cooking qualities. *Proc Natl Acad Sci U S A* 106:21760–21765
- Wang WY, Ding HF, Li GX, Jiang MS, Li RF, Liu X et al (2009) Delimitation of the *PSH1* (t) gene for rice purple leaf sheath to 23.5 kb DNA fragment. *Genome* 52:268–274
- Watterson GA (1975) On the number of segregating sites in genetic models without recombination. *Theoret Pop Biol* 7:256–276
- Wei X, Qian W, Yuan N, Chen Y, Wang R, Cao L et al (2014) Domestication and association analysis of *Hd1* in Chinese mini-core collections of rice. *Genet Resour Crop Evol* 61:121–142
- Wu YP, Pu CH, Lin HY, Huang HY, Huang YC, Hong CY et al (2015) Three novel alleles of *FLOURY ENDOSPERM2* (*FLO2*) conferring dull rice grains with low amylose content. *Plant Sci* 233:44–52
- Yamanaka S, Nakamura I, Watanabe N, Sato YI (2004) Identification of SNPs in the waxy gene among glutinous rice cultivars and their evolutionary significance during the domestication process of rice. *Theor Appl Genet* 108:1200–1204
- Yang CC, Kawahara Y, Mizuno H, Wu J, Matsumoto T, Itoh T (2011) Independent domestication of Asian rice followed by gene flow from *japonica* to *indica*. *Mol Biol Evol* 29:1471–1479
- Yu Y, Tang T, Qian Q, Wang Y, Yan M, Zeng D et al (2008) Independent losses of function in a polyphenol oxidase in rice: differentiation in grain discoloration between subspecies and the role of positive selection under domestication. *Plant Cell* 20:2946–2959
- Yue B, Cui KH, Yu SB, Xue WY, Luo LJ, Xing YZ (2006) Molecular marker-assisted dissection of quantitative trait loci for seven morphological traits in rice (*Oryza sativa* L.). *Euphytica* 150:131–139
- Zhang LB, Zhu Q, Wu ZQ, Ross-Ibarra J, Gaut BS, Ge S et al (2009) Selection on grain shattering genes and rates of rice domestication. *New Phytol* 184:708–720
- Zhang Z, Li M, Fang Y, Liu F, Lu Y, Meng Q et al (2012) Diversification of the Waxy gene is closely related to variations in rice eating and cooking quality. *Plant Mol Biol Rep* 30:462–469
- Zhao K, Wright M, Kimball J, Eizenga G, McClung A, Kovach M et al (2010) Genomic diversity and introgression in *O. sativa* reveal the impact of domestication and breeding on the rice genome. *PLoS One* 5, e10780

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)

---