PLOS ONE

A Large Insertion in bHLH Transcription Factor *BrTT8* Resulting in Yellow Seed Coat in *Brassica rapa*

Xia Li[1], Li Chen[2], Meiyan Hong[1], Yan Zhang[3], Feng Zu[4], Jing Wen[1], Bin Yi[1], Chaozhi Ma[1], Jinxiong Shen[1], Jinxing Tu[1]*, Tingdong Fu[1]

1 National Key Laboratory of Crop Genetic Improvement, National Center of Rapeseed Improvement in Wuhan, Huazhong Agricultural University, Wuhan, P.R. China, 2 Guizhou Rapeseed Institute, Guizhou Academy of Agricultural Sciences, Guiyang, P.R. China, 3 Vegetable Research Institute, Guangdong Academy of Agriculture Sciences, Guangdong, P.R. China, 4 Industrial Crop Research Institute, Yunnan Academy of Agricultural Science, Kunming, P.R. China

Abstract

Yellow seed is a desirable quality trait of the *Brassica* oilseed species. Previously, several seed coat color genes have been mapped in the *Brassica* species, but the molecular mechanism is still unknown. In the present investigation, map-based cloning method was used to identify a seed coat color gene, located on A9 in *B. rapa*. Blast analysis with the *Arabidopsis* genome showed that there were 22 *Arabidopsis* genes in this region including *at4g09820* to *at4g10620*. Functional complementation test exhibited a phenotype reversion in the *Arabidopsis thaliana tt8-1* mutant and yellow-seeded plant. These results suggested that the candidate gene was a homolog of *TRANSPARENT TESTA8 (TT8)* locus. *BrTT8* regulated the accumulation of proanthocyanidins (PAs) in the seed coat. Sequence analysis of two alleles revealed a large insertion of a new class of transposable elements, *Helitron* in *yellow sarson*. In addition, no mRNA expression of *BrTT8* was detected in the yellow-seeded line. It indicated that the natural transposon might have caused the loss in function of *BrTT8*. *BrTT8* encodes a basic/helix-loop-helix (bHLH) protein that shares a high degree of similarity with other bHLH proteins in the *Brassica*. Further expression analysis also revealed that *BrTT8* was involved in controlling the late biosynthetic genes (LBGs) of the flavonoid pathway. Our present findings provided with further studies could assist in understanding the molecular mechanism involved in seed coat color formation in *Brassica* species, which is an important oil yielding quality trait.

Introduction

B. rapa (AA) is an important crop and a model plant for studying *Brassica* genome evolution. Yellow seed is a desirable quality trait of the *Brassica* oilseed species. Compared with their dark-seeded counterparts, yellow seeds of *Brassica* have inherent advantages, such as higher oil content [1–3]. *Yellow sarson* (*Brassica rapa var.trilocularis*) is a valuable yellow-seeded variety of *B. rapa* in India. Moreover, *Yellow sarson* is a major yellow-seeded germplasm; used to create several artificial synthetic yellow-seeded varieties of *B. napus* [4–5]. Presently, it has been documented that there are two genes involved in controlling the seed color in *B. rapa* [6–8]. Seed color genes have been mapped and cloned in *B. rapa* during the past decade [9–11], but the molecular mechanism of the seed color formation has not been elaborated in *B. rapa*.

Map-based cloning has been performed on several genes in the *Brassica* species [12–13], but few reports describe the cloning of seed color genes. Additionally, the physiological functions of seed color gene in the *Brassica* species are still largely unknown. However, many seed color genes from other plant species such as maize, *Arabidopsis* and rice have been successfully cloned for genetic and molecular analyses. These seed color genes mainly correspond to enzymes and regulatory factors that participate in the flavonoid biosynthesis [14–18]. The main enzymes that are involved in the flavonoid synthesis pathway include chalcone synthase (CHS), chalcone isomerase (CHI), flavonoid 3-hydroxy-lase (F3H), flavonoid 3′-hydroxylase (F3′H), dihydroflavonol 4-reductase (DFR), flavonol synthase (FLS), leucoanthocyanidin reductase (LAR), leucocyanidin dioxygenase (LDOX), anthocya-nidin reductase (ANR), peroxidase (POD), and polyphenol oxidase (PPO).

To date, it has been established that the transcriptional regulation of the structural genes for flavonoid biosynthesis is controlled by MYB and basic helix–loop–helix (bHLH) transcription factors, together with WD40 proteins. Recently, the bHLH transcription factors are found to be involved in the transcriptional regulation of the flavonoid pathway [19–20]. The first bHLH transcription factor regulating the flavonoid pathway is identified in maize which regulates the anthocyanin biosynthesis pathway in seed [21]. The bHLH transcription factors *Rc/Rd* specifically control proanthocyanidins (PAs) synthesis in the grain pericarp in rice [22]. bHLH proteins in *Arabidopsis* such as TT8 control both anthocyanin and PA pathways [18], and is required for normal expression of the flavonoid "late" biosynthetic genes (LBGs) including *DIHYDROFLAVONOL 4-REDUCTASE (DFR)* and *BANYULS (BAN)*. DFR is the first enzyme leading to the production of anthocyanidin, and the *BAN* encoded ANR, the

core enzyme in PA biosynthesis [23–24]. In addition, through the yeast two- and three-hybrid experiments, TT8 (bHLH), TT2 (MYB), and TTG1 (WD40) can form a ternary complex which is involved in the regulation of *BAN* expression, and the expression of *TT8* itself is also controlled by different combinations of MYB and bHLH factors in the seed coat of *Arabidopsis thaliana* [25–26]. These data clearly indicate that bHLH transcription factors can regulate one or more branches of the flavonoid pathway, and the structural and functional similarities of transcription factors may slightly differ depending on its species.

Using the ethyl methanesulfonate (EMS) treatment or T-DNA insertion, a series of seed coat color mutants in *Arabidopsis thaliana* have been produced. However, in natural mutations, seed coat color change is due to a variety of factors: for example, changes in the length of the sequences that contain the insertion/deletion mutations in flavonoid biosynthetic genes, such as the *Rc* in rice and *TTG1* homolog in *B. rapa* [9,16]. Another factor is endogenous RNA interference (RNAi), which occurs in the *CHS* genes in soybean [27]. In addition, it is well known that the transposable elements affect the seed coat color in maize. The transposable elements participated in pigment metabolism have also been reported in other plant species [28–29].

In this study, we reported the isolation and functional characterization of the *BrTT8*. *BrTT8* regulated the accumulation of pigment in the seed coat. The sequence analysis of two alleles showed that a transposable element could affect the seed coat color. *BrTT8* encodes a protein exhibiting strong similarity with other bHLH proteins in *Brassica*. Additional experiments also demonstrated that *BrTT8* modulated the expression of flavonoid "LBGs". These results have provided useful information which could assist in the current understanding of the molecular mechanism of seed coat color formation in yellow-seeded *Brassica* crops.

Results

Phenotypic and genetic characterization of seed color in *B. rapa* BC$_5$ population

The population BC$_5$ was developed by backcrossing 3H219 (black-seeded parent) as a donor to the *Yellow sarson* (yellow-seeded parent). The black-seeded plant exhibited brown color seeds (Figure 1A), but the yellow-seeded individuals showed bright yellow color seeds (Figure 1D). Histological analysis of the immature seeds showed that there were three inner integument cell layers (ii1, ii2, ii3) in the black seed coat (Figure 1C), but the PA-accumulating cell layer (ii1) was completely absent in yellow seed coats, and there were some unknown fragments present (Figure 1F) when stained with Safranine O and Fast Green treatment. This shown that the absence of PAs was responsible for yellow color of the seeds.

The genetic analysis showed that the yellow seed trait was recessive, monogenic, and maternal. A total of 1183 individuals were obtained from the BC$_4$ black-seeded individuals, of which 565 plants were yellow-seeded and 618 plants were black-seeded. This segregation of yellow and black was consistent with the expected Mendelian ratio of 1:1 ($\chi^2 = 2.28$, $\chi^2_{0.05} = 3.84$), confirming that only one seed color gene is present in the BC$_5$ population.

Identification of candidate gene for seed coat color trait

AFLP markers were screened in the BC$_1$ population, and the markers that tightly linked to the seed color gene were selected for sequencing. Blast analysis suggested that the molecular marker EA11MG10 shared a very high sequence similarity with the BAC KBrB072E02, which is located in the linkage Group A9 (e-171). Therefore, BAC sequences within the target region were collected from the publicly available *B. rapa* genomic sequences (http://www.brassica-rapa.org/BRGP/index.jsp) and BRAD (http://brassicadb.org/brad/). Based on the BAC sequences, SSR primers were designed using the SSR finder tool. Using 22 SSR markers from the A9 BAC sequence of *B. rapa* (Table S1), the seed coat color gene was located on linkage group 9 corresponding with the two molecular markers lsr126 and bsr1 that are located at 1.3 cM and 2.7 cM, respectively (Figure 2A). In addition, we developed co-separation SSR molecular markers that linked to the seed coat color gene from the Scaffold000135 sequence on A9. Blast analysis using the *Arabidopsis* genome showed that the sequence was similar to a region of chromosome 4 (Figure 2B). There are 22 *Arabidopsis* genes in the region including *at4g09820* to *at4g10620*. It has been well established that *at4g09820* (*TT8*) is involved in the regulation of flavonoid biosynthesis in *Arabidopsis*, and *tt8* mutations possess transparent testa [18]. Moreover, the *TT8 Brassica* homolog showed no recombination with the seed coat color gene. These results suggested that the candidate gene was likely a *TT8* homolog.

Large insertion in the yellow-seeded line

The molecular data allowed us to design specific primers for the cloning of the full-length *TT8* ortholog by PCR amplification. For this purpose, we used the primers Tu-L and Tu-R (Table S2), which were designed from the predicted homologous sequence of *at4g09820* in the Scaffold000135. The amplification of the corresponding genomic sequence from the black-seeded line resulted in a 5420-bp fragment (Figure S1) that spanned the entire 3555-bp (Figure S2) putative *TT8* homolog open reading frame (Figure 3A). However, using the same primers, no amplification was observed in the yellow-seeded line. The primers TL1 and TR1 (Table S2), based on the *B. rapa* sequences showed high homology with the first and fifth exon sequences of *Arabidopsis*, respectively were used to amplify a fragment of that gene in yellow-seeded. Surprisingly, a fragment was amplified that was much larger in length than anticipated (Figure 3B). It was speculated that there was a large insertion in this region.

Due to its complex secondary structure, the insertion sequence was acquired through restriction enzyme digestion and sequencing. The sequence (Figure S3) analysis showed that the inserted fragment contained the structural characteristics of a recently discovered class of transposable elements in eukaryotes, termed *Helitron*. The insertion was 4320 bp, starting with 5′ TC, ending with 3′ CTAG, and containing two short palindromic sequences that were possibly formed by the 17-bp hairpin that was located near its 3′ terminus (Figure 3A).

For further verification, we again designed the primer TL2 in the second exon and the primer YCR1 (Table S2) in the insertion sequence. For the three genotypes, the same fragment was amplified in the homozygous yellow-seeded line and heterozygous black-seeded line (Figure 3C1, 2), but no fragment was amplified in the homozygous black-seeded line (Figure 3C3). We determined the level of mRNA expression in the seeds of the yellow-seeded line in comparison with that of the black-seeded line, to detect the expression changes of the *Brassica TT8* orthologue due to the insertion. There was no mRNA detected in the immature seeds of yellow-seeded line (Figure 3D). This indicated that the inserted fragment disturbed the normal transcription of *BrTT8* in the yellow-seeded line.
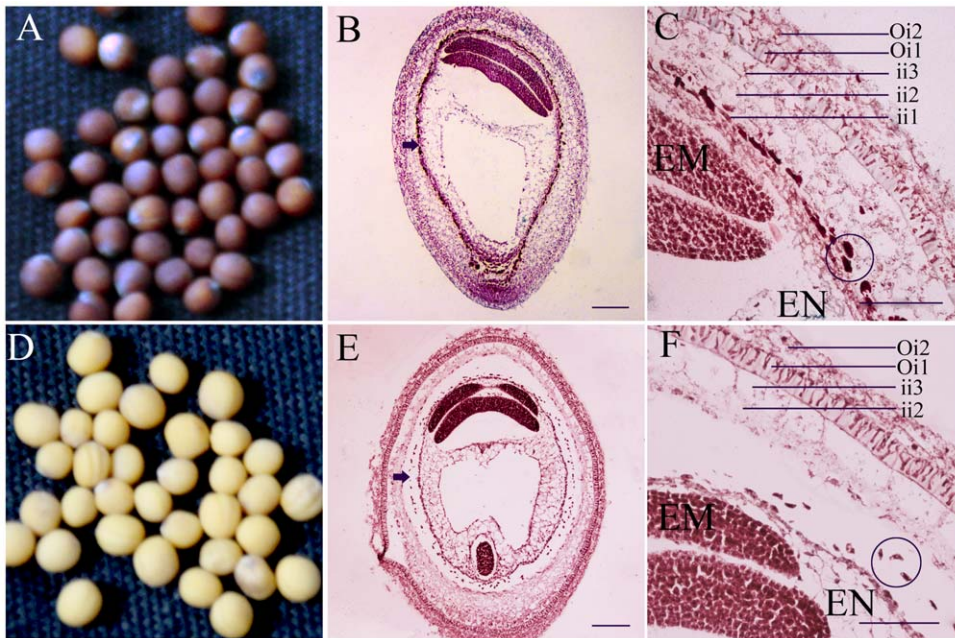
Figure 1. Seed coat structure of the *B. rapa*. (A) The black seeds of the *B. rapa*. (B) The whole black seed treatment with Safranine O and Fast Green. Arrowheads show the Phenolic compounds are localised in the seed coat of immature seed (20 days after flower). (C) Magnified image of (B). Phenolic compounds stain red, formed granules are localised in the endothelium cell of the ii1 layer (circle in C). (D) The yellow seeds of the *B. rapa*. (E) The whole yellow seed treatment with Safranine O and Fast Green (20 days after flower). Phenolic compounds are absent in the endothelium of the yellow seed (arrowhead in E). (F) Magnified image of (E). There are only unknown fragmented structures in the region of the endothelium cell layer (circle in F). EM, embryo; EN, endosperm; ii, inner integument; Oi, outer integument. Bar(N) = 100 μm for (B) and (E); 50 μm for (C) and (F).
doi:10.1371/journal.pone.0044145.g001

BrTT8 encodes a bHLH-Domain protein

The *BrTT8* gene encodes a putative bHLH protein consisting of 520 amino acids (Figure S4) with a predicted molecular weight of 59.5 kD, and a pI of 5.45 (http://www.expasy.org/tools/protparam/). The BrTT8 protein sequence contains a typical bHLH signature near the C terminus that corresponds with a putative binding domain (http://www.expasy.ch/tools/

scanprosite/). The bHLH structure and most of its invariant amino acid residues are conserved, consisting of a basic region (14 amino acids) and two α-helices separated by a loop of variable length (Figure 4A).

Distance analysis of bHLH sequences was performed (Figure 4B), which suggested the existence of 4 groups. *BrTT8* encodes the bHLH protein belonging to group 3, which occurs in five species of Brassicaceae: AEA03281, containing the bHLH
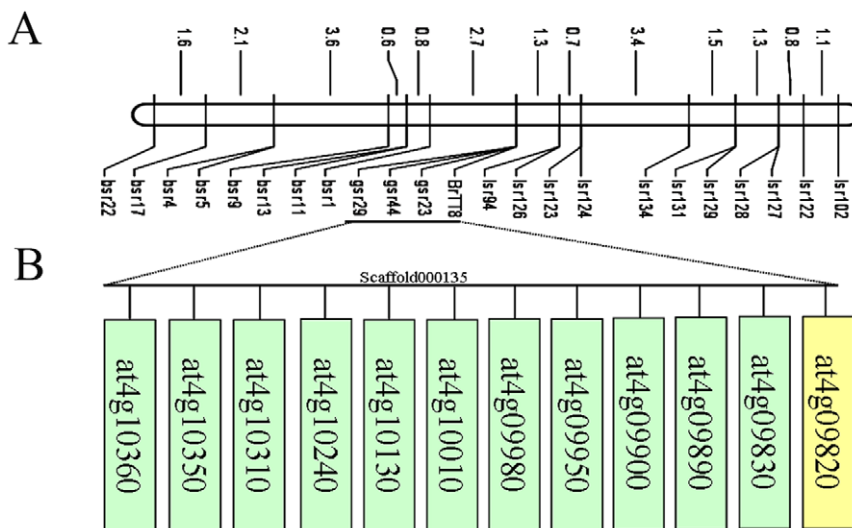
Figure 2. Mapping of the *BrTT8* gene. (A) The genetic linkage map of the *BrTT8* gene. The markers gsr23, gsr44, and gsr29 are derived from Scaffold000135. (B) BLAST analysis with the *Arabidopsis* genome showed that Scaffold000135 shared similarity with a region on chromosome 4. The rectangles containing *Arabidopsis* genes present several representative genes (E<10^{-30}) in this region.
doi:10.1371/journal.pone.0044145.g002
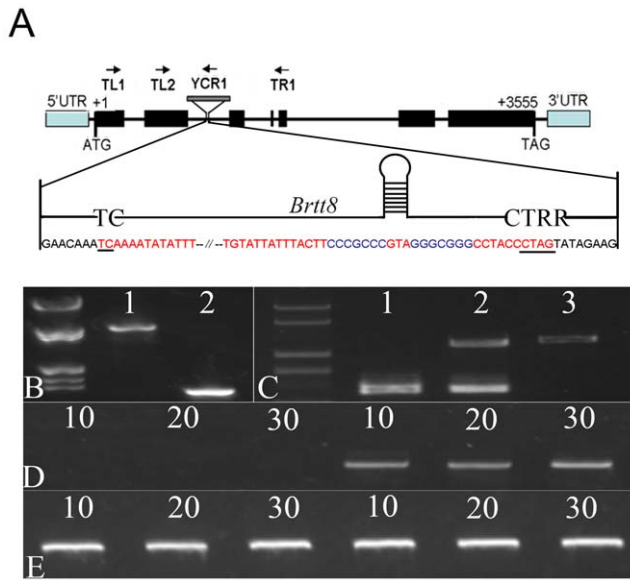
Figure 3. Differences in DNA and mRNA expression. (A) The insertion location is shown in gray rectangles in the ORF of the *BrTT8*. The black rectangles represent the exons, and TL1, TL2, TR1 and YCR1 are the primers that were developed from the corresponding exon sequences and insert sequence. The arrows are used to indicate the directions of the primers. The insertion sequences and flanking *BrTT8* intron 2 are shown in red and black, respectively. The conserved sequences at the termini of the element are underlined. Palindromic sequences that are capable of forming a hairpin are shown in blue. (B) The amplification products of the genomic DNA of the yellow-seeded (1) and black-seeded line (2) using the primers TL1 and TR1. (C) The primers TL2, TR1 and YCR1 amplified the genomic DNA from the three genotypes: (1) the homozygous yellow-seeded, (2) heterozygous black-seeded, (3) homozygous black-seeded plants. (D) mRNA levels in the immature seeds of the yellow-seeded line (three on the left) and black-seeded line (three on the right). The numbers 10, 20, and 30 signify the number of days after pollination. (E) 18S control.
doi:10.1371/journal.pone.0044145.g003

protein found in *Brassica rapa*, ABY59772.1, with the bHLH protein from *Brassica napus*, ADP76654.1, with the bHLH protein in *Brassica oleracea*, AEO53065.1, with the bHLH protein from *Raphanus sativus*, and NP 192720.2(AtTT8), containing the TT8 protein that is found in *Arabidopsis thaliana*. In addition, within the large family of plant bHLH-domain proteins, the four protein sequences of the *Brassica* shared the highest similarity with the BrTT8.

Functional complementation of the mutation

A modified pC2300 vector was used to conduct a functional complementation experiment to confirm that the phenotype was caused by a mutation in the *Brassica TT8* ortholog gene. 5420-bp genomic clone containing the *BrTT8* gene, along with a 1.2-kb promoter and a 0.6-kb 3′ flanking region from the black-seeded line, was introduced into an *Arabidopsis tt8-1* mutant (SALK_030966). The T_2 seed progeny originating from 25 independent T_1 kanamycin-resistant transformants exhibited phenotypic reversions in seed color compared with the wild type (Figure 5A). Furthermore, analysis of PAs deposition during seed development was performed by staining PAs and flavan-3-ol precursors with vanillin in the wild type and T_2. It showed that PAs accumulated in the seed body of the T_2 (Figure 5D) similar to the wild type (Figure 5B) indicating that the PAs production defect of *tt8-1* mutant was indeed restored by *BrTT8*. Clone contained

the *BrTT8* gene was introduced into the yellow-seed line, and a phenotypic reversion was exhibited in one positive transformant. These data demonstrate that the yellow-seeded coat in *B. rapa* was conferred by the mutant of *BrTT8*.

BrTT8 regulated flavonoid gene expression in *B. rapa*

Expression analysis was conducted to compare the effects of *BrTT8* on flavonoid metabolism in *B. rapa* with those of *TT8* in *Arabidopsis thaliana*, both of which modify seed pigmentation pattern. The expression level of five flavonoid genes, from two group EBGs and LBGs, were analyzed during the development of seeds in yellow-seed line and black-seed line, respectively (Figure 6). Quantitative Real-time PCR (QRT-PCR) analysis revealed that the yellow-seeded line contained similar amounts of mRNA as the wild-type for *BrTT6* and *BrTT7* which encoded the flavonoid 3-hydroxylase (F3H) and flavonoid 3′-hydroxylase (F3′H), respectively in the flavonoid metabolism pathway, and have been classified as flavonoid EBGs. Conversely, transcripts of three flavonoid LBGs, *BrDFR*, *BrBAN* and *BrLDOX* (encoded leucocyanidin dioxygenase) were statistically insignificant in the yellow-seeded line. Therefore, our results indicated that *BrTT8* is also involved in the genetic control of flavonoid late metabolism in the developing seed.

Discussion

BrTT8 is essential for formation of normal seed coat color in *B. rapa*

Seed color formation is due to the accumulation of flavonoids, mainly consists of flavonols, anthocyanins, phlobaphenes, isoflavones, and proanthocyanidins [30]. The brown to dark color of the mature seeds were due to the PAs oxidations during seed desiccation that accumulate as colorless compounds in the seed coat [31–32]. At present, 20 genes involved in the PA biosynthesis pathway have been identified in the *Arabidopsis thaliana* [30]. However, only a few genes such as *Bna.BAN* and *BnTT2* involved in proanthocyanidin biosynthesis have been identified through genetic and molecular studies in *Brassica* [33–34]. *Bna.BAN* genes were located on oilseed rape genetic maps and co-localised with a potential seed color quantitative trait loci. *ProBna.BAN* was activated in proanthocyanidin-accumulating cells, namely the innermost layer of the inner integument [33].

In our study, histological analysis showed that the black color of seeds was also due to the PAs accumulation in the plant endothelium layers of immature seeds (Figure 1B). However, there was no pigment present in the seed coats of the yellow-seeded plants (Figure 1E). Compared with the black-seeded plants, there was no accumulation of red granules in the endothelium layers of yellow seed coat after Safranine O and Fast Green treatments (Figure 1E). In addition, *BrTT8* was necessary for the expression of flavonoid LBGs, such as *LDOX* and *BAN* in the developing seeds (Figure 6), similar to *TT8* in *Arabidopsis* [18,24]. These results indicated that *BrTT8* was involved in the PA biosynthesis pathway, and the mutation of the seed color in *B. rapa* was due to impairment of the PAs accumulation in the seed coat.

Furthermore, in the yellow seeds of the *B. rapa*, because of the insertion mutation of *BrTT8*, histological analysis showed that the ii1 layer was completely absent (Figure 1F). To our present knowledge, the mechanism of the structural change was still unknown in *B. rapa*. Recently, research has shown that epidermal cell fate, seed-coat development and flavonoid biosynthesis are linked in *Arabidopsis* [35]. Moreover, the bHLH transcription factors have multiple functions in different biosynthetic pathway in plant species [36], so we speculated that the change of cell layer
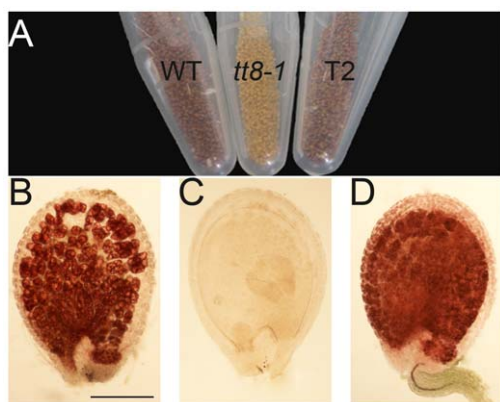
A

B

Figure 4. BrTT8 shows features of a bHLH DNA-Binding domain protein. (A) Amino acid comparison of the bHLH domain of *BrTT8*, *B. napus* (ABY59772.1), *B. rapa* (AEA03281), *B. oleracea* (ADP76654.1), *Arabidopsis* (NP_192720.2) and *Populus* (XP_0023067). (B) Dendrogram of the relationships among the bHLH domains from several bHLH-related proteins. For the construction of the tree, the BrTT8 protein sequence and other selected bHLH-related proteins were used. The sequence similarity was calculated using the MEGA programme to generate a branching pattern. The numbers below the branches indicate the percentages of bootstrap support after 1000 replicates. The sequences used are *Brassica rapa* AEA03281, *Arabidopsis* NP_192720.2(AtTT8), *Brassica napus* ABY59772.1, *Brassica oleracea* ADP76654.1, *Populus* XP_002306769.1, *Vitis vinifera* CBI32369.3, *Lotus* BAH28881.1(LjTT8), *Raphanus* AEO53065.1, *Pisum sativum* ADO13282.1, *Perilla* BAC56998.1(F3G1), *Ricinus* XP_002520758.1, *Malus* AEI84807.1, *Petunia* AAG25927.1(AN1), *Nicotiana* AEE99260.1, *Nicotiana* AEE99258.1(NtAN1b), *Hordeum vulgare* BAJ92594.1, *Lilium* BAE20058.1, *Sorghum* XP_002448313.1, *Dahlia* BAJ33515.1, *Oryza* NP_001053530.2, *Zea mays* NP_001105706.1, *Cornus* AAR21675.1, *Cornus* AAS86268.1, *Oryza* EEC77782.1, *Mimulus* ACA04013.1, and *Gynura bicolor* BAJ17663(GbMYC1).
doi:10.1371/journal.pone.0044145.g004

structure in yellow seed coat might be due to *BrTT8* and it also influence the development of PAs accumulation cell layer in the seed coat.

A natural insertion leads to the mutation of *BrTT8* in *B. rapa*

Yellow sarson is not only a valuable yellow-seeded variety of *B. rapa* but also a major yellow-seeded germplasm, used to produce several artificial yellow-seeded varieties in *B. napus* [4–5]. Several seed coat color genes have been mapped, but few have been cloned by map-based cloning. Using BC$_5$ population, a seed coat color gene was isolated and found to be located in Scaffold000135 on A9 of *B. rapa* which contains a homologous of *TT8* gene (*at4g09820*). *TT8* gene (*at4g09820*) is involved in the regulation of flavonoid biosynthesis in *A. thaliana*.

Interestingly, the sequence analysis of alleles revealed that the seed coat color mutation in *B. rapa*, similar to the *tt8-3* mutation in *A. thaliana*, was caused through insertion in the second intron (Figure 3B and 3C). In *A. thaliana*, no *TT8* mRNA expression was detected in siliques of *tt8-3* [18]. Accordingly, we investigated the expression of *BrTT8* in *B. rapa* and found that no mRNA expression was detected in the immature seeds of the yellow-seeded line (Figure 3D). These data demonstrated that the disturbance in the transcription was caused by the insertion of the second intron.

The sequence analysis showed that the inserted fragment was a recently discovered class of transposable elements in eukaryotes, termed *Helitron*. *Helitrons* are quite large, ranging from approximately 5 to 15 kb. They lack terminal repeats, do not duplicate host insertion sites, and insert consistently between the nucleotides A and T. The *Helitron* structure consists of 5′ TC and 3′ CTRR termini, and a 16- to 20-bp predicted small hairpin structure located 10 to 12 bp upstream of the 3′ end [37]. The insertion was 4320 bp in our study, starting with 5′ TC, ending with 3′ CTAG, and containing two short palindromic sequences that were possibly formed by the 17-bp hairpin located near 3′ terminus (Figure 3A). This shown that the new transposable element is, in all probability, responsible for the changes in seed coat colour in *B. rapa*.

Although, there is no report of *Helitron* transposon in *Brassica*, and *Helitrons* have been found in seven additional flowering plant species [37]. Moreover, *Helitron* transposable elements participated in pigment metabolism has been reported in morning glory [29]. The extensive research on its presence has been done in maize. Researchers have found that the *Helitron* transposon often results in exon shuffling or the duplication of gene sequences [38–39]. Moreover, *Helitrons* also participate in the rearrangement and duplication of genomic regions, contributing to the evolution of novel eukaryotic genomic functions [40–41].
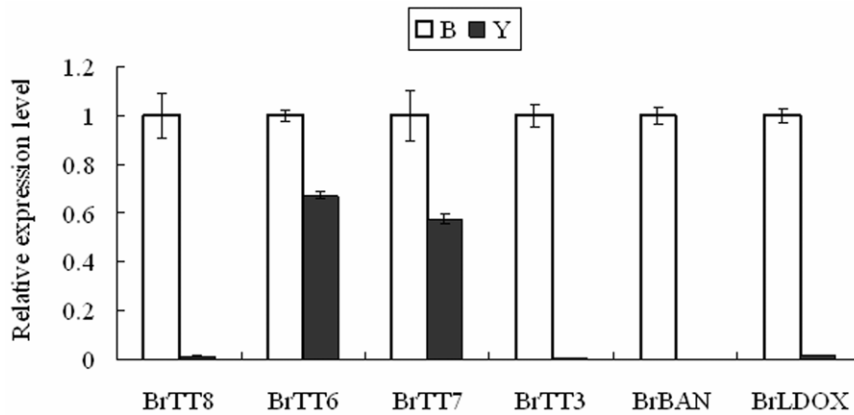
BrTT8 is a highly conserved gene

BrTT8 in *B. rapa* displayed the typical features of a transcription factor with a bHLH signature at its C terminus. The bHLH domain consists of 50–60 amino acids that form two distinct segments: the basic region which contains 10–15 predominantly basic amino acids, and the helix-loop-helix region which form two predicted amphipathic α-helices that are separated by a loop of variable length [42]. The retrieved amino acid sequences were aligned with those of five other plant speceis in the conserved bHLH region and found that the key amino acid residues were highly conserved in the bHLH domain (Figure 4A). Moreover, the sequences of BrTT8 and other bHLH proteins in *Brassica* species were almost completely identical.

The coding sequence of *BrTT8* from the black-seeded line of *B. rapa* was used to complement the *Arabidopsis tt8* mutant. *BrTT8* completely restored the wild-type phenotype of the *tt8* mutant, suggesting that *BrTT8* has a conserved function (Figure 5A). In addition, BLAST analysis showed that *BrTT8* is a single-copy gene in the *B. rapa* genome. Thus, the analytical results demonstrated that *BrTT8* was highly conserved gene in *B. rapa*.

Materials and Methods

Plant materials

Two backcross populations, BC$_1$ and BC$_5$, were developed by backcrossing *3H219* (black-seeded parent) as a donor to *Yellow sarson* (yellow-seeded parent). Markers linked with the target gene were used to select black-seeded individuals in each generation of backcrossing. The BC$_1$ population, containing 202 individuals,

Figure 5. Complementary test and PAs localization in T$_2$ seeds. (A) Seeds of the TT8 wild-type genotype (left), mutant (middle) and T$_2$ progeny of a tt8 homozygous-transformed *Arabidopsis* with the *BrTT8* genomic region (right). (B–D) Detection of PAs and their precursors in immature seed (heart stage) treated with vanillin HCl. The vanillin test stains the PAs and their precursors (leucoanthocyanidins and catechins) red in the endothelium of the wild type (B); the complete absence of these compounds in *tt8-1* immature seed (C); the recovery of these compounds in T$_2$ immature seeds (D). Bar (N) = 100 μm for the (B), (C) and (D).
doi:10.1371/journal.pone.0044145.g005

Figure 6. Expression of flavonoid biosynthetic genes in developing seed. Seeds were obtained from the yellow-seeded and black-seeded plants 10 days after pollination. The expression of the different genes was detected by quantitative Real-time PCR. Transcripts for two flavonoid EBGs *BrTT6* (encode F3H, flavanone 3-hydroxylase) and *BrTT7* (encode F3′H, flavanone 3′-hydroxylase), three flavonoid LBGs *BrDFR* (encode dihydroflavonol reductase), *BrBAN* (encode anthocyanidin reductase) and *BrLDOX* (encode LDOX, leucoanthocyanidin dioxygenase) in *B. rapa*. The comparative Ct method was used to calculate the levels of transcripts relative to black-seeded plant. ("B" in the legend is the black seed and "Y" is the yellow seed).
doi:10.1371/journal.pone.0044145.g006

was used to identify the amplified fragment length polymorphisms (AFLP) and simple sequence repeats (SSR) markers that are linked to the seed coat color gene. The BC_5 population, containing a total of 1183 individuals, was used for molecular mapping of the seed coat color gene.

Markers development

DNA was extracted using a CTAB modified method according to Lei *et al.* (2007) [43]. AFLP analysis was conducted as described by Liu *et al.* (2005) with minor modifications [44]. The two bulks along with the two parental lines were subjected to AFLP analysis. The AFLP fragments cloning and sequence analyses were performed as described by Lei *et al* (2007) [43]. The AFLP sequences were analyzed using BLAST of the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). Bacterial artificial chromosome (BAC) sequences within the target region were collected from *B. rapa* genomic sequences (http://www.brassica-rapa.org/BRGP/index.jsp) and the *Brassica* database (BRAD) (http://brassicadb.org/brad/). All of the SSR primers were designed using the web-based SSR finder tool (http://www.geboc.org/index). A local linkage map of the region on the chromosome surrounding the gene was drawn using Mapdraw V2.5 [45].

Comparison analysis and candidate gene cloning

The sequences that contained the co-segregate SSR markers were aligned with the *A. thaliana* genome sequences from the TAIR database (http://www.arabidopsis.org/). The sequences with a threshold value of $E<10^{-30}$ were regarded as homologous loci of the *A. thaliana* genome, and located on a physical map. According to the conserved co-linearity with *A. thaliana*, we developed other co-segregate markers, and finally identified the candidate gene. Based on the BAC sequence that contained the seed coat color gene, we designed the primers and cloned the gene.

Expression analysis

Black and yellow developing seeds were harvested from 5 plants each for studying the expression analysis. We tagged individual flowers on the primary inflorescence on the first day of flower pollination and again every 10 days. Seeds were removed from the siliques at each stage, including 10_DAP, 20_DAP, and 30_DAP

(DAP: days after pollination). The tissue samples were stored in liquid nitrogen, and total RNA was extracted with the highly efficient hot CTAB-LiCl synthesis [46]. 0.4 g of grated seeds was mixed with 4 ml 65°C CTAB in a 10 ml centrifuge tube, and incubated in a 65°C water bath heat-treatment for 30 min. The supernatant was then collected by centrifugation after mixing with an equal volume of chloroform/isoamylalcohol (24:1) at 4°C. After the addition of 1/4 volume 4 M LiCl, the samples were incubated at −20°C for 2 hours. They were eluted with 1.35 mL diethyl pyrocarbonate-treated water, and then 150 μl NaAc (pH 5.2, 3 M) and 3.75 ml 70% ethanol (freezing-treated) were added. Finally, the RNA was dried and dissolved in the DEPC water. The RNA extracts were converted to first-strand cDNA by using MMLV reverse transcriptase (MBI Fermentas, USA) according to the manufacturer's instruction and the products were diluted 100-fold with sterilized ddH$_2$O for subsequent PCR reactions. The RT-PCR was performed in 20_μl reactions using 2_μl of diluted first-strand cDNA as the template. The PCR was conducted using the following settings: 35 cycles with 94°C for 30 sec (3 min for the first cycle), 94°C for 30 s, 58°C for 30 s, and 72°C for 40 s, followed by an extension at 72°C for 5 min, after which samples were held at 4°C. The primers for the RT and Q-PCR tests were in the Table S3.

GenBank accession numbers

The GenBank accession numbers were as follows (Figure 4): *Brassica rapa* (AEA03281), *Arabidopsis* (NP_192720.2), *Brassica napus* (ABY59772.1), *Brassica oleracea* (ADP76654.1), *Populus* (XP_002306769.1), *Vitis vinifera* (CBI32369.3), *Lotus* (BAH28881.1), *Raphanus* (AEO53065.1), *Pisum sativum* (ADO13282.1), *Perilla* (BAC56998.1), *Ricinus* (XP_002520758.1), *Malus* (AEI84807.1), *Petunia* (AAG25927.1), *Nicotiana* (AEE99260.1), *Nicotiana* (AEE99258.1), *Hordeum vulgare* (BAJ92594.1), *Lilium* (BAE20058.1), *Sorghum* (XP_002448313.1), *Dahlia* (BAJ33515.1), *Oryza* (NP_001053530.2), *Zea mays* (NP_001105706.1), *Cornus* (AAR21675.1), *Cornus* (AAS86268.1), *Oryza* (EEC77782.1), *Mimulus* (ACA04013.1), and *Gynura bicolor* (BAJ17663).

Construct preparation

A pCAMBIA2300 vector [47] was digested with the EcoRI and HindIII restriction enzymes (NEB, USA). The seed coat color gene was amplified from the black-seeded lines using high fidelity polymerase (NEB, USA) with the primers Tu-L and Tu-R, and suitable restriction enzyme cleavage sites were added for use with the pC2300 vector. The PCR products from the coding sequences were treated with T4_ligase and mixed with the digested vector. Chemical transformation was used to introduce the mixture of PCR fragments and vector DNA into chemically competent *E. coli* DH5a (Invitrogen, USA). Positive clones were selected through PCR and the insert was confirmed by sequencing.

Microscopy

Immature seed samples were harvested after 20 days of flowering. Seeds were directly fixed in FAA (Formalin 10 ml, Acetic acid 3 ml, 50% Ethanol 87 ml) for 24 hours. After fixation, the material was dehydrated through a series of graded ethanol solutions (50, 70, 90, 100%). Then the material transferred to the graded chloroform-ethanol solutions (25, 50, 75, and 100%) for transparent processing. Finally, the material was paraffin-embedded after infiltration of graded paraffin solutions at 42°C, 56°C, 60°C each one hour later. The embedded samples were sectioned to a thickness of 8 μm using an automatic microtome (Microm HM 360, Thermo). Selected sections were stained for fast green and counterstain with safranine. The vanillin test [48] was performed by direct incubation of immature siliques samples in a freshly prepared solution of 1% (w/w) vanillin (methanol) in 6 N HCl for 30 min at room temperature.

References

1. Shirzadegan M, Robbelen G (1985) Infiuence of seed colour and hull proportions on quality properties of seeds in *Brassica napus* L. Fette Seifen Anstrichm 87: 235–237.
2. Tang ZL, Li JN, Zhang XK, Chen L, Wang R (1997) Genetic variation of yellow-seeded rapeseed lines (*Brassica napus* L.) from different genetic sources. Plant Breeding 116: 471–474.
3. Meng JL, Shi SW, Gan L, Li ZY, Qu XS (1998) The production of yellow-seeded *Brassica napus* (AACC) through crossing interspecific hybrids of *B. campestris* (AA) and *B. carinata* (BBCC) with *B. napus*. Euphytica 103: 329–333.
4. Chen BY, Heneen WK, Jonsson R (1988) Resynthesis of *Brassica napus* L. through interspecific hybridization between *B. alboglabra* bailey and *B. campestris* L. with special emphasis on seed color. Plant Breeding 101: 52–59.
5. Rahman MH (2001) Production of yellow-seeded through interspecific crosses. Plant Breeding 120: 463–472.
6. Schwetka A (1982) Inheritance of seed color in turnip rape (*Brassica campestris* L.). Theoretical and Applied Genetics 62: 161–169.
7. Zaman MW (1989) Inheritance of seed colour in *Brassica campestris*. Sveriges Utsadesförenings Tidskrift 99: 205–207.
8. Rahman MH (2001) Inheritance of petal colour and its independent segregation from seed color in *Brassica rapa*. Plant breeding 120: 197–200.
9. Chen BY, Jorgensen RB, Cheng BF, Heneen WK (1997) Identification and chromosomal assignment of RAPD marker linked with a gene for seed coat color in a *Brassica campestris-alboglabra* addition line. Hereditas 126: 133–138.
10. Rahman MH, Peter B.E . McVetty, Li GY (2007) Development of SRAP, SNP and Multiplexed SCAR molecular markers for the major seed coat color gene in *Brassica rapa* L. Theoretical and Applied Genetics 115: 1101–1107.
11. Zhang JF, Ying L, Yuan YX, Zhang XW, Geng JF, et al. (2009) Map-based cloning and characterization of a gene controlling hairiness and seed coat color traits in *Brassica rapa*. Plant Molecular Biology 69(5): 53–63.
12. Yi B, Zeng FQ, Lei SL, Chen YN, Yao XQ, et al. (2010) Two duplicate CYP704B1 homologous genes *BnMs1* and *BnMs2* are required for pollen exine formation and tapetal development in *Brassica napus*. The Plant Journal 63: 925–938.
13. Dun XL, Zhou ZF, Xia SQ, Wen J, Yi B, et al. (2011) BnaC.Tic40, a plastid inner membrane translocon originating from *Brassica oleracea*, is essential for tapetal function and microspore development in *Brassica napus*. The Plant Journal 68: 532–545.
14. Carey CC, Strahle JT, Selinger DA, Chandler VL (2004) Mutations in the pale aleurone color1 regulatory gene of the *Zea mays* anthocyanin pathway have distinct phenotypes relative to the functionally similar TRANSPARENT TESTA GLABRA1 gene in *Arabidopsis thaliana*. The Plant Cell 16: 450–464.
15. Kitamura S, Shikazono N, Tanaka A (2004) TRANSPARENT TESTA 19 is involved in the accumulation of both anthocyanins and proanthocyanidins in *Arabidopsis*. The Plant Journal 37: 104–114.
16. Furukawa T, Maekawa M, Oki T, Suda I, Iida S, et al. (2007) The *Rc* and *Rd* genes are involved in proanthocyanidin synthesis in rice pericarp. The Plant Journal 49: 91–102.
17. Nesi N, Debeaujon I, Jond C, Stewart AJ, Jenkins GI, et al. (2002) The TRANSPARENT TESTA16 locus encodes the ARABIDOPSIS BSISTER MADS domain protein and is required for proper development and pigmentation of the seed coat. The Plant Cell 14: 2463–2479.
18. Nesi N, Debeaujon I, Jond C, Pelletier G, Caboche M, et al. (2000) The TT8 gene encodes a basic helix-loop-helix domain protein required for expression of *DFR* and *BAN* genes in *Arabidopsis* siliques. The Plant Cell 12: 1863–1878.
19. Hichri I, Barrieu F, Bogs J, Kappel C, Delrot S, et al. (2011) Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. Journal of Experimental Botany 62(8): 2465–2483.
20. Zhang F, Gonzalez A, Zhao M, Payne CT, Lloyd A (2003) A network of redundant bHLH proteins functions in all TTG1-dependent pathways of *Arabidopsis*. Development 130: 4859–4869.
21. 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. The Plant Cell 1: 1175–1183.
22. 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. The Plant Cell 18: 283–294.
23. Pelletier MK, Murrell JR, Shirley BW (1997) Characterization of flavonol synthase and leucoanthocyanidin dioxygenase genes in *Arabidopsis*: Further evidence for differential regulation of "early" and "late" genes. Plant Physiology 113: 1437–1445.
24. Abrahams S, Lee E, Walker AR, Tanner GJ, Larkin PJ, et al. (2003) The *Arabidopsis* TDS4 gene encodes leucoanthocyanidin dioxygenase (LDOX) and is essential for proanthocyanidin synthesis and vacuole development. The Plant Journal 35: 624–636.
25. Baudry A, Caboche M, Lepiniec L (2006) TT8 controls its own expression in a feedback regulation involving TTG1 and homologous MYB and bHLH factors, allowing a strong and cell-specific accumulation of flavonoids in *Arabidopsis thaliana*. The Plant Journal 46: 768–779.
26. Baudry A, Heim MA, Dubreucq B, Caboche M, Weisshaar B, et al. (2004) TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. The Plant Journal 39: 366–38.

Supporting Information

Figure S1   Genomic sequence the *BrTT8*.
(TIF)

Figure S2   ORF sequence of *BrTT8*.
(TIF)

Figure S3   DNA sequence of the insertion.
(TIF)

Figure S4   The amino acid sequence of BrTT8.
(TIF)

Table S1   The SSR markers were developed in the study.
(DOC)

Table S2   The primers were used in the study.
(DOC)

Table S3   The primers were used for Q-PCR.
(DOC)

Acknowledgments

Author Contributions

Conceived and designed the experiments: JT. Performed the experiments: XL LC MH YZ FZ. Analyzed the data: XL. Contributed reagents/materials/analysis tools: JW BY CM JS TF. Wrote the paper: XL.

27. Kurauchia T, Kasaia A, Tougoub M, Sendaa M (2011) Endogenous RNA interference of chalcone synthase genes in soybean: Formation of double-stranded RNA of GmIRCHS transcripts and structure of the 5′ and 3′ends of short interfering RNAs. Journal of Plant Physiology 168: 1264–1270.

28. Ohno S, Hosokawa M, Hoshino A, Kitamura Y, Morita M, et al. (2011) bHLH transcription factor, DvIVS, is involved in regulation of anthocyanin synthesis in dahlia (*Dahlia variabilis*). Journal of Experimental Botany 62(14): 5105–5116.

29. Choi J, Hoshino A, Park K, Park I, Iida S (2007) Spontaneous mutations caused by a Helitron transposon, Hel-It1, in morning glory, *Ipomoea tricolor*. The Plant Journal 49: 924–934.

30. Lepiniec L, Debeaujon I, Routaboul J-M, Baudry A, Pourcel L, et al. (2006) Genetics and biochemistry of seed flavonoids. Annual Review of Plant Biology 57: 405–430.

31. Nesi N, Lucas MO, Auger B, Lécureuil A, Guerche P, et al. (2009) The promoter of the *Arabidopsis thaliana BAN* gene is active in proanthocyanidin-accumulating cells of the *Brassica napus* seed coat. Plant Cell Reports 28: 601–617.

32. Pourcel L, Routaboul JM, Kerhoas L, Caboche M, Lepiniec L, et al. (2005) TRANSPARENT TESTA10 encodes a Laccase-like enzyme involved in oxidative polymerization of flavonoids in *Arabidopsis* seed coat. The Plant Cell 17: 2966–2980.

33. Auger B, Baron C, Lucas MO, Vautrin S, Berge's H, et al. (2009) *Brassica* orthologs from BANYULS belong to a small multigene family, which is involved in procyanidin accumulation in the seed. Planta 230: 1167–1183.

34. Wei YL, Li JN, Lu J, Tang ZL, Pu DC, et al. (2006) Molecular cloning of *Brassica napus TRANSPARENT TESTA 2* gene family encoding potential MYB regulatory proteins of proanthocyanidin biosynthesis. Molecular Biology Reports 34(2): 105–120.

35. Pierre Broun (2005) Transcriptional control of flavonoid biosynthesis: a complex network of conserved regulators involved in multiple aspects of differentiation in *Arabidopsis*. Current Opinion in Plant Biology 8: 272–279.

36. Feller A, Machemer K, Braun EL, Grotewold E (2011) Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. The Plant Journal 66: 94–116.

37. Yang LX, Bennetzen J L (2009) Structure-based discovery and description of plant and animal *Helitrons*. Proceedings of the National Academy of Sciences, USA 106(31): 12832–12837.

38. Morgante M, Brunner S, Pea G, Fengler K, Zuccolo A, et al. (2005) Gene duplication and exon shuffling by Helitron-like transposons generate intraspecies diversity in maize. Nature Genetics 37: 997–1002.

39. Lal SK, Giroux MJ, Brendel V, Vallejos CE, Hannah LC (2003) The maize genome contains a *Helitron* insertion. The Plant Cell 15:381–391.

40. Yang L, Bennetzen JL (2009) Distribution, diversity, evolution, and survival of *Helitrons* in the maize genome. Proceedings of the National Academy of Sciences, USA 106: 19922–19927.

41. Bennetzen JL (2005) Transposable elements, gene creation and genome rearrangement in flowering plants. Current opinion in genetics & development 15: 621–627.

42. Pires N, Dolan L (2010) Origin and Diversification of Basic-Helix-Loop-Helix Proteins in Plants. Molecular Biology and Evolution 27(4): 862–874.

43. Lei SL, Yao XQ, Yi B, Chen W, Ma CZ, et al. (2007) Towards map-based cloning: fine mapping of a recessive genic male-sterile gene (*BnMs2*) in *Brassica napus* L. and syntenic region identification based on the *Arabidopsis thaliana* genome sequences. Theoretical and Applied Genetics 115: 643–651.

44. Liu ZW, Fu TD, Tu JX, Chen BY (2005) Inheritance of seed colour and identification of RAPD and AFLP markers linked to the seed colour gene in rapeseed (*Brassica napus* L.). Theoretical and Applied Genetics 110: 303–310.

45. Liu RH, Meng JL (2003) Mapdraw, a Microsoft Excel macro for draw genetic linkage maps based on given genetic linkage data. Hereditas (Beijing) 25: 317–321.

46. Ding Y (2006) Cloning, Characterisation and Expression of Caleosin Gene from *Brassica napus*. [Master thesis] Wuhan (China): Huazhong Agricultural University: 21–22.

47. Hajdukiewicz P, Svab Z, Maliga P (1994) The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. Plant Molecular Biology 25: 989–994.

48. Aastrup S, Outtrup H, Erdal K (1984) Location of the proanthocyanidins in the barley grain. Carlsberg Research Communication 49: 105–109.

9