Promoter Variation and Transcript Divergence in Brassicaceae Lineages of FLOWERING LOCUS T

Jing Wang¹, Clare J. Hopkins², Jinna Hou¹, Xiaoxiao Zou¹, Chongnan Wang¹, Yan Long¹, Smita Kurup³, Graham J. King⁴, Jinling Meng¹*

¹ National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, People’s Republic of China, ² Department of Pathology, The University of Melbourne, Victoria, Australia, ³ Rothamsted Research, Harpenden, Herts, United Kingdom, ⁴ Southern Cross Plant Science, Southern Cross University, Lismore, New South Wales, Australia

Abstract

Brassica napus (AACC, 2n = 38), an oil crop of world-wide importance, originated from interspecific hybridization of B. rapa (AA, 2n = 20) and B. oleracea (CC, 2n = 18), and has six FLOWERING LOCUS T (FT) paralogues. Two located on the homeologous chromosomes A2 and C2 arose from a lineage distinct from four located on A7 and C6. A set of three conserved blocks A, B and C, which were found to be essential for FT activation by CONSTANS (CO) in Arabidopsis, was identified within the FT upstream region in B. napus and its progenitor diploids. However, on chromosome C2, insertion of a DNA transposable element (TE) and a retro-element in FT upstream blocks A and B contributed to significant structural divergence between the A and C genome orthologues. Phylogenetic analysis of upstream block A indicated the conserved evolutionary relationships of distinct FT genes within Brassicaceae. We conclude that the ancient At2-like whole genome duplication contributed to distinct ancestral lineages for this key adaptive gene, which co-exist within the same genus. FT-A2 was found to be transcribed in all leaf samples from different developmental stages in both B. rapa and B. napus, whereas FT-C2 was not transcribed in either B. napus or B. oleracea. Silencing of FT-C2 appeared to result from TE insertion and consequent high levels of cytosine methylation in TE sequences within upstream block A. Interestingly, FT-A7/C6 paralogues were specifically silenced in winter type B. napus but abundantly expressed in spring type cultivars under vernalization-free conditions. Motif prediction indicated the presence of two CO protein binding sites within all Brassica block A and additional sites for FT activation in block C. We propose that the ancestral whole genome duplications have contributed to more complex mechanisms of floral regulation and niche adaptation in Brassica compared to Arabidopsis.

Citation: Wang J, Hopkins CJ, Hou J, Zou X, Wang C, et al. (2012) Promoter Variation and Transcript Divergence in Brassicaceae Lineages of FLOWERING LOCUS T. PLoS ONE 7(10): e47127. doi:10.1371/journal.pone.0047127

Introduction

Initiation of flowering is a key to the reproduction, evolution and survival of plants, and is determined by complex genetic pathways interacting with environmental and developmental signals. These are mediated through primary photoperiod, vernalization, gibberellin and autonomous pathways [1–3]. In Arabidopsis, FLOWERING LOCUS T (FT) expression causes early flowering independent of environmental and endogenous signals, whereas loss of FT function results in late flowering under long day (LD) conditions [4,5]. FT plays a central and indispensable role for floral induction, as an integrator of different pathways [6]. CO binds a unique sequence element containing the TGTG (N2-3) ATG motif within the FT promoter to induce expression in Arabidopsis under long day conditions [7–9], whereas the FLOWERING LOCUS C (FLC) protein binds directly to the CArg box of FT intron 1 to repress transcription [10]. Three highly conserved blocks, A, B and C were found within an approximated 5 kb upstream promoter region, and the proximal block A and distant block C were found to be essential for Arabidopsis FT activation by CO binding [11]. Although the FT protein is expressed in leaves it is mobile within the plant, combining with a bZIP transcription factor FD to form a complex of FT/FD heterodimer in the shoot apical meristem. This then activates the floral meristem identity genes APETALA 1 (AP1) and FRUITFUL (FUL) to promote flowering [7,8,12,13].

The cultivated Brassica species belong to the monophyletic Brassicaceae tribe within the Brassicaceae family [14]. Among the six Brassica crops in the U-triangle [15], B. carinata (2n = 34, BBCC), B. juncea (2n = 36, AABB) and B. rapa (2n = 38, AACC) are allohexaploids, which originated from pairwise hybridization of the three diploid species, B. nigra (2n = 16, BB), B. oleracea (2n = 18, CC) and B. rapa (2n = 20, AA). Comparative mapping have demonstrated the presence of numerous chromosomal regions homologous to the Arabidopsis genome, which are effectively triplicated within the diploid Brassica species, consistent with a common hexaploid ancestor in the evolution of Brassica and the tribe Brassiceae [16–19]. In contrast, partial and whole genome comparative sequence analysis indicates that gene content is variable within paralogous regions, with interspersed gene losses and insertions compared with Arabidopsis [19–22]. Interestingly, less than 10% of coding sequences of predicted gene models from
Arabidopsis were recently found to be retained as systemic orthologues in each of the triplicated A subgenomes (Brassica rapa Genome Sequencing Project Consortium [23]).

Transposable elements (TEs) are ubiquitous amongst eukaryotic genomes. Of the most prevalent classes, retroelements and DNA elements [24] are believed to be major drivers of genome and gene evolution [25–27]. In plants, most genome size variability is associated with differences in repetitive DNA content, primarily ascribed to differential amplification of TEs [28]. The CACTA element of the En/Spm transposon system was first isolated and characterized at the molecular level in maize [29], and is estimated to be the most abundant DNA transposon family in B. oleracea [30]. B. oleracea transposon 1 (Bow), a C genome-specific CACTA element, is believed to have a major role in the recent divergence of the B. rapa and B. oleracea genomes [31]. In Arabidopsis, the major effect of TE DNA methylation is to silence transposition. However, methylated sequences may also affect the transcription of the flanking genes, typically reducing their expression [32–34].

We previously identified six BnFT paralogues in B. napus [35], and showed that four of these were located within inverted duplicated blocks of homoeologous sections of chromosomes A7 and C6 (BnC6.FT.a and BnC6.FT.b, BnA7.FT.a and BnA7.FT.b). The remaining two paralogues were located on homoeologous chromosomes A2 and C2 (BnA2.FT and BnC2.FT) with deletion or variation of a putative CArG box for the FLC protein binding site that had been identified in Arabidopsis within intron 1 [10]. The phylogenetic tree based on the coding sequence clearly demonstrates that BnFT paralogues share a common ancestor with Arabidopsis FT, but had undergone triplication events consistent with the hypothesis of a common hexaploid ancestor for the Brassica species.

Here, we report the isolation and comparison of the FT promoter sequences from A2 and C2 that have a common lineage, and explore their evolutionary history both within the genus and the Brassicaceae family. FT transcripts from the distinct A2/C2 and A7/C6 lineages were then analyzed from a sample of eleven Brassica diploid and tetraploid cultivars at different developmental stages and environments, revealing divergence of expression patterns. Finally, the potential regulatory mechanisms, including DNA methylation, are further explored and discussed.

**Results**

**Characterization of homoeologous FT genes on Brassica A2/C2 chromosomes**

Our previous results indicated that Brassica FT orthologues on chromosome A2 and C2 arose from a common lineage [35]. We isolated the complete promoter region of FT-A2 and partial promoter of FT-C2 from two BACs of B. rapa Tapidor DH, JBnB0435N00 and JBnB006F10. Based on the established up-stream blocks A, B and C required for FT activation in Arabidopsis [11], we aligned the corresponding upstream regions of FT-A2 and FT-C2 from B. napus, FT-A2 from B. rapa BAC KBrB070G11 and FT-C2 from B. oleracea Scafol000001 (unpublished data, available upon request from lissy@oilcrops.cn). This demonstrated the conservation of these blocks in the FT upstream regions of the Brassica homoeologous chromosomes A2 and C2 (Figure 1A).

Interestingly, a CACTA element of approximately 6 kb was found inserted within FT-C2 block A, and a 5.2 kb retrotransposon within block B of the same upstream region (Figure 1A). The sequences of four promoters are available in NCBI Genbank (accessions JX193765 (FT-A2), JX193766 (FT-C2), JX193767 (FT-C6a) and JX193768 (FT-C6b)).

We then analyzed micro-synteny in the genome segments containing FT-A2 and FT-C2 from B. rapa, B. oleracea and B. napus, as well as the orthologous region from Arabidopsis. This indicated that TEs were more abundant in the neighboring region of FT-C2 than the homologous region of FT-A2. Unfortunately, we had limited access to data from the corresponding regions of FT-C2 in B. napus compared with B. oleracea (Figure 1B). However, we did identify potential hot spots for the insertion of TEs within the intergenic region between orthologues of Atlg65290 and Atlg62925 on Brassica chromosome C2 (Figure 1B).

We designed specific primer sets to distinguish the upstream regions containing the transposon and retrotransposon within block A and B from the A and C genome. These were used for PCR amplification of samples from a range of accessions, including B. rapa (17), B. oleracea (17), B. napus (51), B. nigra (1), B. juncea (6), B. carinata (6) and 9 species representing a range of taxa from the Brassicaceae. Based on the evidence from these marker assays, TEs were only detected within the Brassica C genome, and no insertion was found in the homologous regions of Brassica A and B genomes, nor from other Brassicaceae (Table 1). We concluded that these transposon insertions most likely occurred after the divergence of the Brassica A and C genomes.

**Evolutionary analysis of FT upstream block A in Brassicaceae**

The sequences of FT upstream block A from a sample of diploid and tetraploid Brassica cultivars were determined and aligned, based on use of specific primer pairs for FT-A2 and FT-C2. The derived phylogenetic tree indicates that most B. napus and B. juncea accessions share high sequence similarity for block A with B. rapa, apart from two B. juncea accessions, Bj-6 and Bj-7 (Figure 2A). Accessions of B. juncea and B. carinata containing sequences of B genome origin were clearly allocated to two distinct groups, with only one variety of B. nigra grouped with B. juncea (Figure 2B). Three distinct groups were identified in the C genome phylogenetic tree, one including only B. oleracea and B. napus, and the other two including B. oleracea, B. napus and B. carinata (Figure 2C). This is consistent with the FT-A2/C2 lineage being conserved during the allopolyploidisation of Brassica species.

In order to explore the wider evolutionary relationships of FT genes within the Brassicaceae, we isolated the sequences of upstream block A from 9 representative genera, using the specific primer pairs for the FT-A2 orthologues. The block A sequences of the FT-A2/C2 lineage from the diploid and tetraploid Brassica species and representative Brassicaceae species were aligned with those of Arabidopsis FT, BnC6.FT.a and BnC6.FT.b. There is a clear differentiation of block A within the inverted duplication blocks of the A7/C6 and FT-A2/C2 lineages. The Brassica FT-A7/C6 genes had high identity with that of Lactuca indigotica and Arabidopsis (Figure 2D), whilst the Brassica FT-A2 parologue had higher identity with that of Oxytropis violacea, Rorippa indica, Capsella bursa-pastoris, Diplotaxis teffuolia, Crambe abyssinica and Sinapis arruca, than the corresponding Brassica B lineage (Figure 2D). These results strongly suggest that the distinct lineages observed within the Brassica genomes arose prior to the ancestral whole genome triplication events.

**Differential expression and DNA methylation analysis of FT genes in Brassica diploids and tetraploids**

Since we had observed strong evidence for lineage-specific variation in upstream sequences within the paralogous FT genes, we then explored potential consequences on the pattern of gene expression. We designed specific primer sets to detect transcription
from each of the FT-N2/FT-C2 and the FT-A7/C6 paralogues, using semi-quantitative RT-PCR. Within TapidorDH (B. napus, winter type DH line) and Ningyou7DH (B. napus, semi-winter type DH line) FT-A2 was transcribed in all leaf samples from different developmental stages and different photoperiod treatments, whereas FT-C2 was not transcribed (Figure 3A). Moreover, FT-A2 was found to be transcribed in leaf 2 from three other winter type cultivars, and three spring type cultivars under vernalization-free conditions, whilst FT-C2 was not (Figure 3B). Interestingly, under the same conditions FT-A7/C6 was not transcribed in leaf 2 of TapidorDH and three other winter type cultivars, but was transcribed in Ningyou7DH and three other spring type cultivars (Figure 3A-B).

We then checked FT transcription in leaves from B. rapa and B. oleracea at three different developmental stages without vernalization. FT-A2 was transcribed in all samples from two crop forms of B. rapa, yellow sarson (R-o-18) and Chinese cabbage (Chiifu-401), whereas FT-C2 was not transcribed in B. oleracea var. albovagina, where TEs were inserted in block A and B (Figure 3C). The FT-A7/C6 was transcribed in all samples from all diploid accessions, and the transcript level from the C genome appeared to be higher than that from the A genome (Figure 3C).

We were interested to determine whether the TE insertion in FT-C2 had a significant effect on the neighbouring DNA methylation status. Cytosine methylation analysis of bisulphite treated DNA from leaves of diploid and tetraploid accessions provides strong circumstantial evidence for their role as potential functional motifs within block C. Of seven sites (designated C-1 to C-7) identified, six were universal within Brassicaceae, and another (C-2) specific to the Brassica FT-A2/C2 lineage (Figure 6A). This comparative taxonomic analysis demonstrates that block C is highly conserved across the Brassicaceae, although consistently located between 4.5 and 7.5 kb upstream of the start codon. We were able to designate a series of 10–30 bp regions, highly conserved within the Brassicaceae, as potential functional motifs within block C. Of seven sites (designated C-1 to C-7) identified, six were universal within Brassicaceae, and another (C-2) specific to the Brassica FT-A2/C2 lineage (Figure 6A). This comparative taxonomic analysis provides strong circumstantial evidence for their role as potential motifs for FT activation.

Discussion

Structure and diversification of the Brassica FT-A2/C2 promoter

We previously established that all extant Brassica FT genes share a common ancestry with Arabidopsis FT [35]. Here, we isolated and aligned promoter sequences from the diploid and amphidiploid representatives of the FT-A2/C2 lineage and demonstrated conservation of upstream blocks A and C previously identified in Arabidopsis [11] as being of key functional significance for activation of FT transcription, and hence flowering. However, since the divergence of the B. rapa A and B. oleracea C genomes around 4 MYA [20,38], two class TEs have inserted specifically within FT-C2 upstream blocks A and B. Our micro-synteny analysis indicated that TE insertion appears generally to be more abundant in the vicinity of B. oleracea FT-C2 of the homoeologous population.

| Table 1. Detection of TE insertion in block A and block B of Brassicaceae species. |

| Gene location | Number of accession with TE insertion in corresponding FT paralogous |
|---------------|-------------------------------------------------------------------|
| B. rapa (AA)  | B. oleracea (CC) | B. napus (AACC) | B. nigra (BB) | B. juncea (AABB) | B. carinata (BBCC) | Wild species |
| Block A, A2   | (17 accessions) | (15 accessions) | (51 accessions) | (1 accession) | (8 accessions) | (6 accessions) | (9 accessions) |
| Block A, C2   | /                  | 15              | 51              | 0             | /               | 6              | 0             |
| Block B, A2   | /                  | /               | 0               | 0             | /               | 6              | 0             |
| Block B, C2   | /                  | 15              | 51              | 0             | /               | 6              | 2             |

doi:10.1371/journal.pone.0047127.t001
regions of *B. rapa* FT-A2. This may have resulted from genome-wide TE insertion events (Figure 1B) which are more abundant in the C genome, as revealed by homoeologous BAC sequencing [38,39]. The presence of a CACTA element within FT-C2 block A appears to have contributed to non-functionalization of this key floral regulator. CACTA element insertion leading to gene function diversification is common in plants [40–43]. For example, in soybean, a 20.5 kb insertion in the flavonoid 3'-hydroxylase (*F3'H*) gene led to a new allele conferring a novel phenotype of stable gray trichome color. The 20.5 kb insertion has the molecular structure of a putative autonomous transposon of the CACTA family [42].

Conservation of FT upstream block A within the Brassicaceae

*FT* genes are located in genomic regions corresponding to Arabidopsis chromosome 1 block E [35]. In Brassica species, extensive comparative genomic studies have shown that copies of block E are found on chromosomes A2, B2, C2, A7, B7 and C6 [19,44–46]. Since block E is inversely duplicated on A7 and C6 [35,47] and there is extensive homoeology among the two sets of linkage groups A2/B2/C2 and A7/B7/C6 [45], we were confident that the large block E on B7 is also inversely duplicated. The presence of two *FT* lineages could be inferred within the B genome, with FT-A2/C2/B2 representing homeologues from one lineage and FT-A7/B7/C6 from the other.

The sequences of FT-A2/C2 upstream block A were grouped according to genomic content of Brassica diploid and tetraploid species within the U triangle (Figure 2A–C), which suggests that the FT-A2/C2 lineage was conserved during the process of the amphidiploid formation and domestication. A wider phylogenetic analysis of FT-A2/C2 and FT-A7/C6 upstream block A, that included a representative range of Brassicaceae species, demonstrated the clear distinction between the different groups (Figure 2D), further confirming the independent FT lineages. It is particularly surprising that block A of *I. indigotica* clearly grouped with the FT-A7/C6 lineage even though the sequence was isolated with A2 specific primers and the FT-A2/C2 lineage cluster was
events leading to paleopolyploidy [50]. The most recent *At*-α (~40 MYA) duplication was considered to explain the drastically increased species number in the core *Brassicaceae* [51]. The paleopolyploid *At*-α duplication was followed by later lineage-specific mesopolyploid WGD events. In *Brassicaceae*, at least four independent lineage-specific WGD events have been revealed and the whole-genome triplication (*Br*-τ) in *Brassica* was first proposed and then shown to have occurred before the radiation of the whole tribe *Brassicaceae* [16,18,19,52]. This triplication event further complicates analysis of the evolutionary relationships between different genera and species within the tribe *Brassicaceae*, as shown by our results above.

**Transcript differentiation amongst *Brassica* FT genes**

Duplicated genes may undergo three primary evolutionary fates, including silencing, sub-functionalization and neo-functionalization [53–55]. Within polyploids, and in particular allopolyploids, a wider repertoire may be available in terms of increased variation in dosage-regulated gene expression, altered regulatory interactions, and rapid genetic and epigenetic changes [56–58]. The ability to detect relative expression for all six *Brassica* FT copies is challenging due to high sequence similarity of transcripts, especially for the four FT paralogues within the duplication blocks of *A7/C6* [35]. The lack of transcripts for FT-C2 in all accessions of *B. napus* and *B. oleracea* contrasted with the constitutive expression of FT-A2 in *B. rapa* and *B. napus* (Figure 3A–C). Moreover, whereas the four FT genes within the duplication block of *A7/C6* were all silenced prior to vernalization in winter type *B. napus*, they were expressed in at least some spring type varieties (Figure 3A–B). These different expression patterns of FT-A2 might be attributed to the deletion or variation of FLC protein binding site within intron 1 of A2/C2 origin [10,35], which would result in constitutive expression of FT-A2. However, silencing of FT-C2 may have resulted from TE insertion and consequent conserved high levels of DNA methylation in the TE and flanking sequences which are most found at CG and CHG sites. So, the TE itself and DNA methylation appears to have generated a functional pseudogene FT-C2, given the absence of expression in all materials analyzed. TEs are recognized as a primary target of cytosine methylation in eukaryotes, where it serves primarily to silence these repeat sequences [59]. DNA methylation is able to silence genes by blocking transcription initiation either by preventing protein binding or as a consequence of DNA methylation-induced chromatin remodeling [60]. In plants, TE methylation may occur adjacent to upstream regions where TEs have inserted, with ‘spreading’ of methylation then affecting cis-regulation, such as occurs in tomato, with the *CNR* gene affecting fruit ripening [61]. Although expression of FT-A7/C6 in *B. oleracea* appeared more abundant than in *B. rapa* (Figure 3C), and may compensate for the silencing of FT-C2, this requires further investigation. The difference in transcript abundance for FT-A7/C6 between winter and spring type *B. napus* under vernalization-free conditions may be caused by different control mechanisms for flowering within these two crop types, and in particular the interactions between FLC and FT. However, to clarify this, it is essential to sample leaf at multiple developmental time points from plants subject to vernalising and non-vernalising conditions.

Recent characterization of 12 *GPAT4* (sn-glycerol-3-phosphate acyltransferase 4) and seven *PSY* (*Phytoene synthase*) paralogues in *B. napus* and its two progenitors found that all were expressed, but with tissue specific expression exhibiting overlapping redundancy and signs of sub-functionalization [62,63]. In contrast, the expression patterns of FT genes differed with developmental stage and in different species/cultivars, consistent with selection for
different behaviors for this key regulator of the primary adaptive trait of flowering time, associated with the A and C genome divergence and more recent formation of *B. napus*. This behavior provides a wider range of options for adaption to different environments and domestication selection, such as the divergence between FT-A7/C6 which may correspond to the formation of

**Figure 4. DNA methylation profiles of block A.** (A) Scheme representing the location of TE insertion within blocks A of FT promoters on chromosome A2 and C2, together with primers used. (B) The methylation state of block A from C2 was analyzed using the DNA collected from leaves of *B. napus* and *B. oleracea* by bisulphite sequencing of ten random clones. T represents TapidorDH, N represents Ningyou7DH and Ba represents *B. oleracea var. alboaglabra*. One box represents one cytosine (C) and 0–100% represents the percentage of methylated cytosine in (5mC) detected in the ten random clones sequenced. Red arrows: TE insertion site in the C genome; Vertical line: boundary of TE. doi:10.1371/journal.pone.0047127.g004

**Figure 5. Diagram of conserved blocks within promoter of FT-A2/C2 and FT-A7/C6.** Block A and block C were detected in all FT promoters of *Brassica* whereas block B was not detected in *B. rapa* FT-A7.a from BAC (KBr092C03) of Chiflu-401, *B. napus* FT-C6.a and FT-C6.b from TapidorDH BAC J8nB104L19 and J8nB054L06, respectively.

doi:10.1371/journal.pone.0047127.g005
winter and spring *B. napus* crop types. Interestingly, *FT-A2* was constitutively expressed in all materials and developmental stages tested, and notably in winter-type *B. napus* before vernalization. *FT* genes have recently been found not only to regulate flowering but stomatal opening in *Arabidopsis*, vegetative growth in perennial poplar and storage organ formation in potato [64–66]. In sugar beet, two paralogues of *FT* have evolved antagonistic roles in the control of flowering, with the one functionally conserved with *Arabidopsis FT* being essential for flowering. However, the second parologue represses flowering, with its down-regulation crucial for regulating flowering but in a different manner.

### Potential FT regulatory motif e in Brassicaceae

Three conserved blocks in the promoter region of *FT* have been found to be essential for activation by CO in *Arabidopsis*, with block A subsequently found as a cis-element for promoting gene expression, block B probably enhancing *FT* expression in response to blue light [68] and block C as a potential enhancer for long-distance genomic regulation in conjunction with CO [11]. Interestingly, we did not detect block B in the *Brassica* *FT-A7/C6* lineage although these *FT* genes could be expressed normally (Figure 3A–B), indicating that block B is not required for *Brassica FT* activation. This may be explained by the shade avoidance regulation in *Arabidopsis* where the blue light is an important control of the shade avoidance syndrome responses [69]. However, for *Brassica* this may not be an issue, given the ecological niches it occupies compared with *Brassicas*. We found two potential motifs for CO protein binding in *Brassica* block A (Figure 6A), each associated with the distinct *FT-A2/C2* and *FT-A7/C6* lineages, which may help to distinguish multiple CO paralogues in the *Brassica* genomes. Recently, another transcription factor, *PHYTOCHROME INTERACTING FACTOR4* (*PIF4*), has been identified that binds directly to the *FT* promoter near the transcript start site to regulate *FT* expression in *Arabidopsis* [70]. Another level of regulation of *FT* by *PIF4* may also exist in *Brassica* species which requires further investigation. We found that block C was significantly more conserved across the *Brassica* species and other species of Brassicaceae, even when located about 4.5–7.5 kb upstream of the *FT* start codon (Figure 5), suggesting an important role in *FT* regulation. This finding suggests it will be important to determine whether the potential motifs within block C are combined with other regulatory proteins or small RNA, which then interact with CO to promote *FT* expression.

### Materials and Methods

#### Plant materials and growth conditions

Ninety-nine plant accessions representing diploid and tetraploid *Brassica* and a further nine non-domesticated Brassicaceae species were sampled, with DNA isolated from the leaves of plants grown in field or greenhouse (Table S1). No specific permits were required for the described field studies. Plants grown in the field at Huazhong Agricultural University were used solely for DNA extraction. *B. napus* plants of homozygous TapidorDH and Ningyou7DH were grown in soil in controlled environment (Sanyo Gallenkamp) at 18°C under LDs (16 h light/8 h dark) for 3 weeks, vernalized (8 h light/16 h dark) at 5°C for 7 weeks and 3 weeks, respectively, with half of the plants then transferred to LDs (16 h light/8 h dark, 18°C), and another plants to SDs (10 h light/14 h dark, 18°C) conditions.

The plants of *B. rapa* var. *trilocularis* R-o-18 and var. *pekinesis* Chifū-401, and *B. oleracea* var. *alboglabra* Bailey were grown in cabinets (NK System Biotron, Tokyo, Japan) at 22°C under LDs (16 h light/8 h dark). Tissue was collected from leaves at different growth stages (leaf 2, 6, 9 and 14) for transcript and DNA methylation analysis. Leaves were excised from the plant, the midrib removed, and the left hand side lamina used for RNA, the right hand side for DNA.

### Sequence acquisition and analysis

GenScan software optimized for *Arabidopsis* was used to identify protein-coding genes within TapidorDH BAC sequences [71]. *B. rapa* gene models were obtained from the Brassica database (BRAD) ([http://brassicadb.org/brad/index.php](http://brassicadb.org/brad/index.php)) and coding sequences of *B. oleracea* and *B. napus* were compared to *B. rapa* and *Arabidopsis* genomes using Blastn and WU-BLAST against the BRAD and TIGR database ([www.tigr.org/tdb/e2k1/bog1/](http://www.tigr.org/tdb/e2k1/bog1/)). Transposable elements were predicted and located using "RepeatMasker" ([http://www.repeatmasker.org](http://www.repeatmasker.org)). Primers used for isolating the *FT* upstream blocks and estimating the insertion events are given in Table S2. For TE insertion detection, the primer sets specific to the A2 lineage are successful when there is no TE insertion in block A and block B, whereas the primers specific to the C2 lineage amplify products when TE insertions are present. ClustalW was used for multiple alignments of the block sequences ([http://www.ebi.ac.uk](http://www.ebi.ac.uk/)). Phylogenetic trees were constructed using MEGA version 4.1 [72]. Bootstraps with 500 replicates were performed to assess node support.

### RNA isolation, reverse transcription and RT-PCR assays

Total RNA was extracted from leaves using the RNaseasy Plant Mini Kit (Qiagen). cDNA was reverse-transcribed from total RNA (1 μg) with RevertAid™ First Strand cDNA Synthesis Kit (Fermentas) in a 20 μl reaction. Gene-specific primers were used for RT-PCR to detect Actin (30 cycles) and *Brassica FT* transcripts (30 cycles). The thermal cycling program was 94°C for 3 min followed by 30 cycles of 94°C for 30 s, annealing for 30 s, and extension at 72°C for 30 s, ending with a 10 min extension at 72°C. The primer sets used for RT-PCR detection are listed in Table S2.

### Bisulphite sequencing

DNA was extracted from leaves using the DNeasy Plant Mini Kit (Qiagen). 450 ng genomic DNA was subjected to two successive treatments of sodium bisulphite conversion using the EpiTect Bisulphite kit (Qiagen) according to the manufacturer’s instructions. The reaction was then purified once more using the PCR purification kit (Qiagen) and pooled together. Gene-specific primers were designed to amplify partial coding sequences of *FT-A2* and *FT-C2* and *FT-A7* using Primer3 [77]. PCR products were cloned into the pMD18-T vector (TaKaRa), and 10 individual clones were sequenced. Percentage methylation (% C) was calculated as 100 × C/C(TA).
Supporting Information

Table S1  Plant materials used in this study.

(XLS)

Table S2  Description of primer sets used in this study and the resultant PCR products.

(XLS)

Acknowledgments

The authors thank Graham Teakle (Warwick HR, University of Warwick, Wellesbourne, UK) for kindly providing the seeds of B. oleracea, Xianhong Ge for leaves of representative Brassicaceae species (Huazhong Agricultural University, Wuhan, China) and Shengyi Liu for the Scafin0000001 sequence of B. oleracea (Oil Crops Research Institute, Chinese Academy of Agricultural Sciences).

Author Contributions

Conceived and designed the experiments: JW GK JM. Performed the experiments: JW. Analyzed the data: JW. Contributed reagents/materials/analysis tools: CH JJ XZ CW YL SK. Wrote the paper: JW GK JM. Helped conceive the study and revised the manuscript: SK YL. Secured funds (20100480915): JW.

References

1. Jack T (2004) Molecular and genetic mechanisms of floral control. Plant Cell 16: S1–17.
2. Ross PK, Bastow RM, Mylne JS, Dean C (2004) Multiple pathways in the decision to flower: enabling, promoting, and resetting. Plant Cell 16: S18–31.
3. Philbrook I, Dean C (2006) The timing of developmental transitions in plants. Cell 125: 655–664.
4. Kardailsky I, Shulka VK, Ahn JH, Dagenais N, Christensen SK, et al. (1999) Activation tagging of the floral inducer FT. Science 286: 1962–1965.
5. Kiyosue Y, Kaya H, Goto K, Iwashuchi M, Araki T (1999) A pair of related genes with antagonistic roles in mediating flowering signals. Science 286: 1960–1962.
6. Moon J, Lee H, Kim M, Lee I (2005) Analysis of flowering pathway integrators in Arabidopsis. Plant Physiol 46: 293–299.
7. Wigge PA, Kim MC, Jaeger KE, Busch W, Schmid M, et al. (2005) Integration of spatial and temporal information during floral induction in Arabidopsis. Science 309: 1056–1059.
8. Turchi F, Fornara F, Coupland G (2008) Regulation and identity of orthologs: FLOWERING LOCUS T moves center stage. Annu Rev Plant Biol 59: 573–594.
9. Tiwari SB, Shen Y, Chang HC, Hou Y, Harris A, et al. (2010) The flowering time regulator CONSTANS is recruited to the FLOWERING LOCUS T promoter via a unique cis-element. New Phytologist 187: 57–66.
10. Scarpel I, He Y, Turch F, Vincent C, Fornara F, et al. (2006) The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signaling in Arabidopsis. Genes Dev 20: 1988–1998.
11. Adrian J, Farrona S, Reimer JJ, Albani MC, Coupland G, et al. (2010) cis-Regulatory elements and chromatin state coordinately control temporal and spatial expression of FLOWERING LOCUS T in Arabidopsis. Plant Cell 22: 1423–1440.
12. Abe M, Kobayashi Y, Yamamoto S, Daimon Y, Yamaguchi A, et al. (2005) FD, a bZIP protein mediating signals from the floral pathway integrator FLC, is recruited to the FLOWERING LOCUS T promoter at the shoot apex. Science 309: 1052–1056.
13. Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, et al. (2007) FD protein movement contributes to long-distance signaling in floral induction of Arabidopsis. Science 315: 1050–1053.
14. Al-Sherhaz IA, Belstein MA, Kellogg EA (2006) Systematics and phylogeny of the Brassicaceae (Cruciferae): An overview. Plant Syst Evol 259: 89–120.
15. U N (1953) Genome analysis in Brassica with special reference to the experimental formation of B. napus and peculiar mode of fertilization. Jpn J Biol 7: 389–452.
16. Lagercrantz U (1998) Comparative mapping between Arabidopsis thaliana and Brassica nigra indicates that Brassica genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics 150: 1217–1228.
17. Lukens L, Zou F, Lydiate D, Parkin I, Oshorn T (2003) Comparison of a Brassica oleracea genetic map with the genome of Arabidopsis thaliana. Genetics 164: 359–372.
18. Lysak MA, Koch MA, Pecinka A, Schubert I (2005) Chromosome triplication found across the tribe Brassicaceae. Genome Res 15: 516–525.
19. Parkin IA, Golden SM, Sharpe AG, Lukens L, Trick M, et al. (2005) Segmental structure of the Brassica napus genome based on comparative analysis with Arabidopsis thaliana. Genetics 171: 765–781.
20. Rana D, van den Boogaart T, O'Neill CM, Hynes L, Bent E, et al. (2004) Conservation of the microstructure of genome segments in Brassica napus and its diploid relatives. Plant J 40: 725–733.
21. Town CD, Cheung F, Trick M, Drou N, Lin YP, et al. (2009) Comparative analysis between homoeologous genome segments of Brassica napus and its progenitor species reveals extensive sequence-level divergence. Plant Cell 21: 1912–1928.
22. Town CD, Cheung F, Trick M, Drou N, Lin YP, Park JY, et al. (2009) Comparative analysis between homoeologous genome segments of Brassica napus and its progenitor species reveals extensive sequence-level divergence. Plant Cell 21: 1912–1928.
23. Qiin D, Gao MQ, Li YG, Quiros C (2009) Comparative sequence analysis for Brassica oleracea with similar sequences in B. rapa and Arabidopsis thaliana. Plant Cell Rep 28: 649–661.
24. Chopra S, Bredel V, Zhang J, Axell JD, Peterson T (1999) Molecular characterization of a mutant pigmentation phenotype and isolation of the first active transposable element from Sorgnum bicolor. Proc Natl Acad Sci USA 96: 1530–1535.
25. Wicker T, Yahiaoui N, Gaudillere J, Haas BJ, Liu ZD, et al. (2003) Rapid transposon insertion in an F3 generation plant. Plant Cell 15: 1530–1535.
26. Wicker T, Yahiaoui N, Gaudillere J, Haas BJ, Liu ZD, et al. (2003) Rapid transposon insertion in an F3 generation plant. Plant Cell 15: 1530–1535.
27. Zabala G and Vodkin L (2008) A putative autonomous 20.5 kb-CACTA transposon insertion in an F3’/H allele identifies a new CACTA transposon subfamily in Glycin max. BMC Plant Biol 8: 124.
28. Xu M, Brar HK, Groscie S, Palmer RG, Bhattacharyya MK (2010) Excision of an active CACTAlike transposable element from DFR2 causes variegated flowers in soybean [Glycin max (L.) Merr.]. Genetics 184: 53–63.
29. Schranz ME, Lysak MA, Mitchell-Olds T (2006) The ABCs of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. Trends Plant Sci 11: 535–542.
30. Panjabi P, Jagannath A, Bish NC, Padmaja K L, Sharma S, et al. (2008) Comparative mapping of Brassica juncea and Arabidopsis thaliana using Intron Polymorphism (IP) markers: Homoeologous relationships, diversification and evolution of the A, B and C Brassica genomes. BMC Genomics 9: 113.
31. Wang J, Lydiate DJ, Parkin IA, Falentin C, Delourme R, et al. (2011) Integration of linkage maps for the amphidiploid Brassica napus and comparative mapping with Arabidopsis and Brassica rapa. BMC Genomics 12: 101.
47. Howell EC, Armstrong SJ, Barker GC, Jones GH, King GJ, et al. (2005) Physical organization of the major duplication on *Brassica oleracea* chromosome O6 revealed through fluorescence in situ hybridization with Arabidopsis and *Brassica* BAC probes. Genome 48: 1093–1103.

48. Bailey CD, Koch MA, Mayer M, Mummenhoff K, O’Kane SL Jr, et al. (2006) Toward a global phylogeny of the Brassicaceae. Mol Biol Evol 23: 2142–2160.

49. Lysak MA, Lexer C (2006) Towards the era of comparative evolutionary genomics in Brassicaceae. Plant Syst Evol 259: 175–196.

50. Franzke A, Lysak MA, Al-Shnbar IA, Koch MA, Mummenhoff K (2011) Cabbage family affairs: the evolutionary history of Brassicaceae. Trends Plant Sci 16: 108–116.

51. Henry Y, Bedhomme M, Blanc G (2006) History, protohistory and prehistory of the *Arabidopsis thaliana* chromosome complement. Trends Plant Sci 11: 267–275.

52. Lysak MA, Cheung K, Kitschke M, Bures P (2007) Ancestral chromosomal blocks are triplicated in Brassiceae species with varying chromosome number and genome size. Plant Physiol 145: 402–410.

53. Lynch M, Conery JS (2000) The evolutionary fate and consequences of duplicate genes. Science 290: 1151–1155.

54. Wendel JF (2000) Genome evolution in polyploids. Plant Mol Biol 42: 225–249.

55. Blanc G, Wolfe KH (2004) Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. Plant Cell 16: 1679–1691.

56. Osborn TC, Pires JC, Birchler JA, Auger DL, Chen ZJ, et al. (2003) Understanding mechanisms of novel gene expression in polyploids. Trends Genet 19: 141–147.

57. Chen ZJ (2007) Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. Annu Rev Plant Biol 58: 377–406.

58. Whittle CA, Krochko JE (2009) Transcript profiling provides evidence of functional divergence and expression networks among ribosomal protein gene paralogs in *Brassica napus*. Plant Cell 21: 2203–2219.

59. Suzuki MM, Bird A (2008) DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet 9: 465–476.

60. Curcurti M, Lezo A, Badaracco G, Landsberger N (2002) Molecular mechanisms of gene silencing by DNA methylation. Mol Cell Biol 22: 3157–3173.

61. Manning K, Torr M, Poole M, Hong Y, Thompson AJ, et al. (2006) A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. Nat Genet 38: 948–952.

62. Chen X, Truksa M, Snyder CL, El-Mezawy A, Shah S, et al. (2011) Three homologous genes encoding sn-glycerol-3-phosphate acyltransferase 4 exhibit different expression patterns and functional divergence in *Brassica napus*. Plant Physiol 155: 831–845.

63. Cárdenas PD, Gajardo HA, Haeber T, Parkin IA, Iniguez-Luy FL, et al. (2012) Retention of triplicated phytoene synthase (*PSY*) genes in *Brassica napus* L. and its diploid progenitors during the evolution of the Brassicaceae. Theor Appl Genet 124: 1215–1228.

64. Kinoshita T, Ono N, Hayashi Y, Morimoto S, Nakamura S, et al. (2011) *FLOWERING LOCUS T* regulates stomatal opening. Curr Biol 21: 1232–1238.

65. Hao CY, Adams JP, Kim H, No K, Ma C, et al. (2011) *FLOWERING LOCUS T* duplication coordinates reproductive and vegetative growth in perennial poplar. Proc Natl Acad Sci USA 108: 10756–10761.

66. Navarro C, Abendana JA, Cruz-Oro E, Guellar CA, Tamaki S, et al. (2011) Control of flowering and storage organ formation in potato by *FLOWERING LOCUS T*. Nature 478: 119–122.

67. Pin PA, Benfocch R, Bonnet D, Wremersch-Weich E, Kraft T, et al. (2010) An antagonistic pair of *FT* homologs mediates the control of flowering time in sugar beet. Science 330: 1397–1400.

68. Liu H, Yu X, Li K, Kleijnot J, Yang H, et al. (2008) Photoexcited CRY2 interacts with CIB1 to regulate transcription and floral initiation in *Arabidopsis*. Science 322: 1535–1539.

69. Keller MM, Jalsalais Y, Pedmale UV, Moreno JR, Chory J et al. (2011) Cryptochrome 1 and phytochrome B control shade-avoidance responses in *Arabidopsis* via partially independent hormonal cascades. Plant J 67: 195–207.

70. Kumar SV, Lucyshyn D, Jaeger KE, Alós E, Alvey E, et al. (2012) Transcription factor PIF4 controls the thermosensory activation of flowering. Nature 484: 242–245.

71. Burge C, Karlin S (1997) Prediction of complete gene structures in human genomic DNA. J Mol Biol 268: 78–94.

72. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596–1599.

73. Wang J, Wang C, Long Y, Hepkins C, Kurup S, et al. (2011) Universal endogenous gene controls for bisulphite conversion in analysis of plant DNA methylation. Plant Methods 7: 39.