The protein encoded by the TERMINAL FLOWER1 (TFL1) gene maintains indeterminacy in inflorescence meristem to repress flowering, and has undergone multiple duplications. However, basal angiosperms have one copy of a TFL1-like gene, which clusters with eudicot TFL1/CEN paralogs. Functional conservation has been reported in the paralogs CENTRORADIALIS (CEN) in eudicots, and ROOTS CURL IN NPA (RCNs) genes in monocots. In this study, long-term functional conservation and selective constraints were found between angiosperms, while the relaxation of selective constraints led to subfunctionalisation between paralogs. Long intron lengths of magnoliid TFL1-like gene contain more conserved motifs that potentially regulate TFL1/CEN/RCNs expression. These might be relevant to the functional flexibility of the non-duplicate TFL1-like gene in the basal angiosperms in comparison with the short, lower frequency intron lengths in eudicot and monocot TFL1/CEN/RCNs paralogs. The functionally conserved duplicates of eudicots and monocots evolved according to the duplication-degeneration-complementation model, avoiding redundancy by relaxation of selective constraints on exon 1 and exon 4. These data suggest that strong purifying selection has maintained the relevant functions of TFL1/CEN/RCNs paralogs on flowering regulation throughout the evolution of angiosperms, and the shorter introns with radical amino acid changes are important for the retention of paralogous duplicates.

TERMINAL FLOWER1 (TFL1) is a member of the phosphatidylethanolamine-binding protein (PEBP) family. It represses flowering by counteracting the action of another PEBP protein, FLOWERING LOCUS T (FT), which promotes flowering¹. The function of indeterminacy on shoot meristem of Antirrhinum majus suggests that the CENTRORADIALIS (CEN) gene is conserved and that its product is functionally identical to that of TFL1²;³. TFL1 and CEN are paralogous genes with conserved functions that involve the formation of inflorescences⁴;⁵ and the maintenance of indeterminacy in inflorescent meristems⁶. Most eudicot species possess low or one copy of TFL/CEN in their genomes². The monocot TFL1/CEN-like paralogous genes, named ROOTS CURL IN NPA (RCN1 and RCN2), also share the same function and are expressed in a similar pattern in rice, whereas another duplicated gene, RCN3, may be a non-functional chimera⁸. The TFL1-like gene was also found in the transition in inflorescent indeterminacy/determinacy in Phaseolus vulgaris⁸. The natural variation of TFL1-like gene may also be related to evolutionary transition of inflorescence architecture⁹;¹⁰. It has been suggested that gymnosperms lack orthologues of FT and TFL1/CEN. From functional analysis of the homologous FT/TFL1-like gene in gymnosperm, the repressive function of TFL1/CEN/RCNs in flowering is known to be plesiomorphic²⁵. The angiosperm TFL1/CEN/RCNs paralogs are the result of multiple gene duplications: (1) after the divergence between basal angiosperms (TFL1-like) and eudicots + monocots (TFL1/CEN/RCNs paralogs), (2) two-time duplications resulting in RCN1–3 in monocots, and (3) gene duplication causing the divergence of TFL1 and CEN in eudicots (phylogeny of angiosperms refers to Amborella Genome Project¹⁵, Fig. 1). The conserved function of TFL1/
CEN/RCNs prevents redundancy or silencing by functional divergence\textsuperscript{14,15}, which occurs by positive selection or through the relaxation of environmental constraints\textsuperscript{15}.

Different expression patterns of duplicated TFL1/CEN/RCNs genes in Arabidopsis\textsuperscript{16}, apple\textsuperscript{17}, tomato\textsuperscript{18,19}, and tobacco\textsuperscript{20} tissues have been reported. Such differential expression was suggested as complementary functions (subfunctionalisation)\textsuperscript{21,22}. Following functional divergence, genes normally experience a phase free from selective constraints\textsuperscript{23}. Because of the conserved properties of TFL1/CEN/RCNs, poor resolution of nucleotide phylogeny\textsuperscript{24} cannot explain their divergence. Nevertheless, a single reciprocally switched amino acid could cause functional interconversion between FT (flowering activator) and TFL (flowering repressor)\textsuperscript{25,26}. Therefore, a few changes in the amino acid sequence can alter protein function to escape the redundancy of duplicates\textsuperscript{27}. Therefore, determining radical amino acid changes between TFL1/CEN/RCNs paralogs (the type-II functional divergence of Gu)\textsuperscript{28,29} could be useful for predicting their functional divergence after duplication.

The functional conservation and divergence of paralogous genes is not only reflected in coding sequences, but also in exon-intron structure. Structural divergence is prevalent in duplicated genes and leads to functionally divergent paralogs\textsuperscript{30} and provide heterogeneous regulatory functions in duplicate\textsuperscript{31–33}. Highly expressed genes have longer introns than genes expressed at low levels\textsuperscript{33}. Exon length was also suggested to be associated with molecular functions in flowering development cf.\textsuperscript{34}. The Amborella trichopoda genome (http://www.amborella.org/)\textsuperscript{13} enables the comparison of gene structure and sequence variation in TFL1-like gene between basal angiosperms, and monocot and eudicot angiosperms. Comparisons of gene structure and intron lengths may enhance our understanding of evolution and its relevance among paralogs.

Genetic diversity among TFL1/CEN homologs played a key role in the diversification of flowering plants\textsuperscript{7,23}, which was probably driven by heterogeneous selective pressures on different gene regions. For example, strong selective sweeps in coding regions, and balancing selection of promoters were detected in Arabidopsis\textsuperscript{35}. Furthermore, epistatic selection was identified through a QTL closely linked to the Arabidopsis TFL1\textsuperscript{36}. In addition, latitudinal gradients adaptation was also inferred by nonsynonymous polymorphisms of TFL1\textsuperscript{37}. However, there have been limited studies focussed on the effects of selective pressures on TFL1/CEN/RCNs paralog duplication, as well as the TFL1-like gene in basal angiosperms. These functionally conserved paralogous gene duplicates may be subject to strong purifying selection pressures that constrain redundant functions, such as the floral-regulatory paralogs SEPALLATA 1 (SEP1) and SEPALLATA 2 (SEP2), and SHATTERPROOF 1 (SHP1) and SHATTERPROOF 2 (SHP2)\textsuperscript{38}. Selective constraints may be important in functionally redundant paralogous genes for buffering an organism’s phenotype against deleterious mutations\textsuperscript{39}.

In this paper, a broad range of representative organisms from basal angiosperms, eudicots, and monocots were sampled to determine whether flowering plants exhibit divergent functions of TFL1/CEN/RCNs paralogs and how the selective pressure drove their evolution. General patterns of structural divergence in duplicated genes were analysed to represent the divergence/conservation of these paralogous genes. The aims of this research were to investigate (1) the evolution of intron variability in angiosperm TFL1/CEN/RCNs genes; (2) the effect of positive selection on angiosperm TFL1/CEN/RCNs coding sequences; and (3) the functional divergence between paralogs of angiosperm TFL1/CEN/RCNs, and thus infer the ancestral/derived type of TFL1/CEN/RCNs paralogs.

Figure 1. The hypothetical phylogenetic relationships of angiosperm TFL1/CEN/RCN paralogs. Values on the nodes indicate the ω of specific branches estimated under the free-ratio model, which suggest a pervasive purifying selection or selective constraints on the evolution of angiosperm TFL1/CEN/RCNs paralogs.
TFL1/CEN other paralogs. Furthermore, monocot was found to have relatively short but constant intron lengths compared with variable, and the monocot RCN1 respectively. Exon lengths were found to be constant and there were no significant differences between paralogs, -like genes are 496–2048 bps, 320–3273 bps, 312–384 bps, 509–1007 bps, 510–643 bps, and 1444–3539 bps, TFL1

Furthermore, magnoliid monocot paralogs were well grouped, and the basal lineages of eudicot and magnoliid paralogs, the monocot paralogs were well grouped, RCN

The misgrouping for monocot and magnoliid paralogs was also revealed in the bush-like topology of basal lineages of eudicot and magnoliid paralogs. Error bars represent one standard error. Different colors represent different TFL1/CEN/RCN lineages. Introns have greater length variation than exons, and the introns of monocot RCN1 are significantly shorter than other paralogs. Levels not connected by the same letter are significantly different based on Student’s t test.

Results

Sequence length variation. All sequences were confirmed as TFL1-like by the presence of histidine at the 92th amino acid position (corresponding position at the 88th site of Arabidopsis)25. Only one copy for each Magnoliid species was obtained after amplification, and this result is consistent with only one TFL1/CEN/RCNs member in EST-library of basal angiosperm database (accession number: gnl|Liriodendron|b4_c119764, Ancestral Angiosperm Genome Project, http://ancangio.uga.edu/index.php). The sequences amplified from Magnoliid shown the best hit to TFL1/CEN/RCNs family (Nelumbo nucifera CEN-like protein 2, E-value: 5e−90−2e−92). Exon lengths of eudicot TFL1 and CEN, monocot RCN1, RCN2, and RCN3, and basal angiosperm (magnolids + Amborella) TFL1-like gene range 519–609 bps, 447–531 bps, 522 bps, 522 bps and 516–522 bps, respectively. The length of introns from TFL1, CEN, RCN1, RCN2, RCN3, and basal angiosperm TFL1-like genes are 496–2048 bps, 320–3273 bps, 312–384 bps, 509–1007 bps, 510–643 bps, and 1444–3539 bps, respectively. Exon lengths were found to be constant and there were no significant differences between paralogs, with the exception of exon 4 between eudicots and monocots (Fig. 2). In contrast, the intron lengths were highly variable, and the monocot RCN1 was found to have relatively short but constant intron lengths compared with other paralogs. Furthermore, monocot RCNs had a higher number of intron length polymorphisms than eudicot TFL1/CEN (Fig. 2). Although only two TFL1-like full sequences were obtained from magnolids, the synapomorphy of the long intron lengths of TFL1-like genes in Lauraceae and Magnoliaceae were confirmed by PCR (Additional file 1: Fig. S1). Analysis revealed that the monocot Sorgium bicolor has lost intron 1, and that exon 2 has merged with exon 1, and this sequence is removed when estimating the exon/intron length variation. The exon/intron structures and lengths are shown in Fig. 3 and Additional file 1: Table S2.

Correlation between intron lengths and conserved motifs. Twelve conserved motifs, which are identical to the motifs of putative cis-acting elements, were identified in noncoding regions (Additional file 1: Table S3), and the four-base motifs CAAT box and WRKY were abundant and both presence frequencies (0.0054 and 0.0064, respectively) are higher than those predicted by random occurrence (1/256, Table S3), and the four-base motifs CAAT box and WRKY were abundant and both presence frequencies (0.0054 and 0.0064, respectively) are higher than those predicted by random occurrence (>1/256, p = 0.0245 and 0.0001, respectively) (Additional file 1: Table S3). The TFL1-like gene from magnolids was found to have longer introns and more abundant cis-element to motifs. A strong and significant positive correlation between the number of cis-element motifs and intron length were found (R² = 0.711, p < 0.0001, Fig. 4), suggesting that noncoding regions in TFL1/CEN/RCNs paralogs are relevant to intron length.

Phylogenetic tree of angiosperm TFL1/CEN/RCNs paralogs. The phylogenetic tree of angiosperm TFL1/CEN/RCNs paralogs was reconstructed using amino acid sequences (Fig. 3) and was inconsistent with the hypothetical tree (Fig. 1 and Additional file 1: Fig. S2). The magnoliid TFL1-like gene was misgrouped with monocot RCNs (Fig. 3). The misgrouping for monocot and magnoliod paralogs was also revealed in the bush-like tree topology for basal lineages by Bayesian inference (Additional file 1: Fig. S3). The misgrouping of Magnoliid with monocot or eudicot is common in phylogenetic analysis using certain genes, which is probably due to combination of the relatively old age of these taxa and long branches attraction40−42. In contrast to the unresolved topology of basal lineages of eudicot and magnoliod paralogs, the monocot RCN paralogs were well grouped, with relatively high bootstrap supports in the ML and Bayesian trees (Fig. 3 and Additional file 1: Fig. S3). Furthermore, RCN2 and RCN3 are grouped together in both ML and Bayesian analyses (Fig. 3), implying that the duplication sequence in monocots is RCN1 and RCN2/3 followed by RCN2 and RCN3.

Positive selection analyses. To examine the effect of selective pressures on the angiosperm TFL1/CEN/RCNs paralogs, the ratio (ω) of missense (Ka) to silent mutation rates (Ks), an indicator of natural selection, was estimated. Likelihood ratio analysis revealed that the free ratio model was a better fit than the constant ratio model (2ΔL = 97.8117, df = 61, p = 0.0004), suggesting a strong and pervasive purifying selection on angiosperm TFL1/
CEN/RCNs paralogs (Fig. 1). To examine the grouping of magnoliid TFL1-like gene with eudicot TFL1/CEN and monocot RCNs (Table 1), the two and three ratio models were performed, and showed relaxation of selective constraints ($\omega_0 < \omega_1, \omega_2 < 1$) for eudicot TFL1, CEN, monocot RCN2 and RCN3, and magnoliid TFL1-like gene, but strong purifying selection ($\omega_0 > \omega_1$) for the monocot RCN1 (Table 2). The two ratio model was a better fit for the evolution of monocot, eudicot and magnoliids TFL1/TFL1-like paralogs than three ratio models. This suggests that the grouping of eudicot and magnoliid TFL1/TFL1-like paralogs is a consequence of functional constraint and that both paralogs suffered different selective pressures for magnoliids TFL1/TFL1-like paralogs (Table 2).

Evolutionary divergence between angiosperm TFL1/CEN/RCNs paralogs. The pairwise Ka/Ks ratio was calculated and plotted against Ks to reveal patterns of selection through time. No pairwise Ka/Ks > 1 were obtained suggesting that no positive divergent selection occurred between paralogs. Eudicot TFL1 and CEN were mostly distributed in the quadrant Ka/Ks < 1 and Ks > 1, indicating long-term purifying selection.
Radical functional divergence between angiosperm TFL1/CEN/RCNs paralogs. An *in silico* analysis of radical amino acid changes between duplicated genes was conducted for testing the functional divergence of TFL1/CEN/RCNs paralogs. Nonsignificant radical functional divergence, as estimated by the type-II functional divergence index (θ_{II})^{34}, was found between angiosperm TFL1/CEN/RCNs paralogs (Table 3). The proportion of fixed radical change between paralogs (F_{00,R}) was zero between eudicot CEN and paralogs of monocots and magnoliids, but more or less between paralogs between monocot RCNs and magnoliid TFL1-like genes (Table 3). This indicates that there is functional conservation of paralogs between eudicots and monocots, and functional specialization between paralogs within the eudicot and monocot species.
parameters have several protein binding sites. In addition, both AP1 and LFY can bind to the locus and directly sup-

eudicot sperm paralog with an average through intron evolution because, generally, there are longer introns in basal angiosperms than in both eudicots -like genes suggest absence of constraining variable intron lengths and sequences of angiosperm -ω three-ratio (ωII ) between paralogs supports the functional constraint hypothesis for angio-

TFL1 orthologs could Table 2. Summary of the ω estimation and likelihood ratio test (2ΔL) between two-ratio (ω1 = ω2 = ωRL) and three-ratio (ω1 = ω2 = ωRL) models. ω1, ω2, and ωRL are the Ka/Ks ratio of the branches of the eudicot TFL1 (or eudicot CEN, monocot RCNs), magnolid TFL1-like, and background lineages, respectively. np: number of parameters p: p-value obtained from fitted model using χ2 test.

| Hypothesis | np | lnL | 2ΔL | p   | ω  | Supporting hypothesis in Table 1 |
|------------|----|-----|-----|-----|-----|---------------------------------|
| TFL1 vs. magnoliids | 135 | -12594.3259 | 0.7256 | 0.3258 | ω1 = 0.1021, ω2 = 0.1361, ωRL = 0.1678 | 1. Functional constraint hypothesis |
| CEN vs. magnoliids | 136 | -12593.9631 | 0.0002 | 0.9887 | ω1 = 0.1054, ω2 = 0.1193, ωRL = 0.1674 | 1. Functional constraint hypothesis |
| RCN1 vs. magnoliids | 135 | -12597.2165 | 0.8682 | 0.2774 | ω1 = 0.1140, ω2 = 0.0841, ωRL = 0.1665 | 1. Functional constraint hypothesis |
| RCN2 vs. magnoliids | 136 | -12595.6616 | 0.176 | 0.8708 | ω1 = 0.1137, ω2 = 0.1164 | 1. Functional constraint hypothesis |
| RCN3 vs. magnoliids | 135 | -12597.2484 | 1.1712 | 0.2052 | ω1 = 0.1126, ω2 = 0.1079, ωRL = 0.1535 | 1. Functional constraint hypothesis |
| Eudicots vs. magnoliids | 135 | -12592.2355 | 0.0915 | 0.1255 | 1. Functional constraint hypothesis |
| Monocots vs. magnoliids | 136 | -12592.0526 | 0.3658 | 0.5494 | ω1 = 0.0916, ω2 = 0.1085, ωRL = 0.1264 | 1. Functional constraint hypothesis |

Table 2. Summary of the ω estimation and likelihood ratio test (2ΔL) between two-ratio (ω1 = ω2 = ωRL) and three-ratio (ω1 = ω2 = ωRL) models. ω1, ω2, and ωRL are the Ka/Ks ratio of the branches of the eudicot TFL1 (or eudicot CEN, monocot RCNs), magnolid TFL1-like, and background lineages, respectively. np: number of parameters p: p-value obtained from fitted model using χ2 test.

Discussion

Exon length conservation and intron length variability. Exon length conservation implies constraints of gene functions among organisms. Eudicot and monocot TFL1/CEN/RCNs are functionally conserved and the inflorescence architecture was determined by comparison with the model organisms Arabidopsis and rice. Highly variable intron lengths and sequences of angiosperm CEN/RCNs/TFL1-like genes suggest absence of constraining reproductive function from noncoding function. It is not known whether intron fragments have been gained or lost through evolution, due to poor or failed alignment in introns. However, we suspect that there was a gradual deletion throughout intron evolution because, generally, there are longer introns in basal angiosperms than in both eudicots and monocots (Fig. 3 and Additional file 1: Fig. S1). A deletion of this type in introns could be the result of recombi-

Formation of gene loops is also relevant to the activation or maintenance of gene functions among organisms. Eudicot and monocot TFL1/CEN/RCNs are functionally conserved and the inflorescence architecture was determined by comparison with the model organisms Arabidopsis and rice. Highly variable intron lengths and sequences of angiosperm CEN/RCNs/TFL1-like genes suggest absence of constraining reproductive function from noncoding function. It is not known whether intron fragments have been gained or lost through evolution, due to poor or failed alignment in introns. However, we suspect that there was a gradual deletion throughout intron evolution because, generally, there are longer introns in basal angiosperms than in both eudicots and monocots (Fig. 3 and Additional file 1: Fig. S1). A deletion of this type in introns could be the result of recombi-

Long introns with more conserved motifs could have a complicated folding structure, as well as alternative splicing sites that affect transcription, particularly for the basal angiosperms (such as Lauraceae, Magnoliaceae, and Amborellales). Alternative splicing in TFL1/CEN paralogs was reported to influence terminal flowering and flowering time. Formation of gene loops is also relevant to the activation or maintenance of Arabidopsis TFL1 expression. Therefore, gene lengths are hypothesized to be a key factor affecting the expression efficiency of TFL1 orthologs. TFL1 is targeted by several MADS-box genes, which have different functions during floral transition, and they coordinate the timing of flowering and floral development with TFL1, indicating that the TFL1 orthologs could have several protein binding sites. In addition, both AP1 and LFY can bind to the TFL1 locus and directly suppress TFL1 expression. Suppression of TFL1 in inflorescence branching regulation by MADS-box genes also affects LFY and API expression. No AP1 binding sites (CArG box) or LFY binding sites were found in either eudicot TFL1 and monocot RCN1, but were present in basal angiosperms (Additional file 1: Table S3). This might suggest the functional relevance of long introns in the TFL1-like gene in basal angiosperms. In contrast, the short introns of eudicot and monocot paralogs could reflect their low expression, which may facilitate the retention of duplicated genes and the conservation of their ancestral functions.

Pervasive purifying selection and relaxation of selective constraints on eudicot and monocot TFL1/CEN/RCNs paralogs. It was suggested that the functions of angiosperm flowering development genes have been conserved under selective constraint in eudicots and monocots. Strong purifying selection of the TFL1 paralog with an average ω = 0.097 was inferred based on site-model analysis, which is similar to the average ω of other floral-regulatory paralogs (SEP1 vs. SEP2 and SHP1 vs. SHP2, both ω ≈ 0.16). Nonsignificant radical functional divergence (h9) between paralogs supports the functional constraint hypothesis for angiosperm TFL1/CEN/RCNs paralogs (Table 3). However, in the phylogenetic analysis (Fig. 3 and Additional file 1: Table S3). This might suggest the functional relevance of long introns in the TFL1-like gene in basal angiosperms. In contrast, the short introns of eudicot and monocot paralogs could reflect their low expression, which may facilitate the retention of duplicated genes and the conservation of their ancestral functions.
TFL1/CEN (Fig. 5), particularly on exon 2 and exon 3 of all angiosperm paralogs. Relaxation of selective constraints on eudicot TFL1/CEN/RCNs paralogs suggests that constraints were relaxed, particularly in exon 1 and exon 4 (Figs. 5 and 6), which conserved with no amino acid changes (Fig. 7). However, relatively higher pairwise Ka/Ks within monocot and magnolid paralogs suggests that constraints were relaxed, particularly in exon 1 and exon 4 (Figs. 5 and 6), which is also supported by the analysis of site-specific radical functional change between paralogs of both monocots and magnolids (Fig. 7). Residues 133–151 in exon 4 form an external loop, and this conformation determines the functional specificities of floral regulators. Loss of the hydrogen bond between the external loop (exon 4) and the activator regions (exon 3) may be responsible for the functional conversion of activators of FT to floral repressors of TFL1. One radical change in RCN1/RCN2, RCN2/RCN3, RCN2/magnloid TFL1-like, and three radical changes in RCN1/RCN3 within the external loop were estimated (Fig. 7), suggesting that the paralogous divergence occurred by relaxation of selective constraints, particularly between the monocot RCNs.

Limited fixed radical differences (F_{00,R}) could suggest the maintenance of ancestral function between paralogs of different taxa and imply that the eudicot TFL1/CEN and monocot RCNs do not fit the neo-functionalisation hypothesis of duplicate genes. Instead, the duplication could be a case of sub-functionalisation due to the relaxation of functional constraints, because one-third to a half frequency of radical change (GR) was detected (Table 3). The duplication-degeneration-complementation (DDC) and escape from adaptive conflict (EAC) models are commonly used to explain subfunctional divergence of duplicates and retention of duplicates. The main difference between the DDC and EAC models is that the EAC puts more emphasis on positive selection for gene specialisation during or after duplication, while positive selection is not required for DDC. In the case of eudicot TFL1/CEN and monocot RCNs, all duplicates retained plesiomorphic functionality with slight differences, by relaxation of selective constraints. However, no specific paralogs suffered positive selective pressures, suggesting a more likely evolutionary fit to DDC. The EAC hypothesis therefore, was rejected.

**Relaxation of selective constraints and phylogenetic convergence of magnolid TFL1-like genes.** No duplication of TFL1-like genes was found in basal angiosperms. The constructed phylogenetic tree showed that the magnolid TFL1-like genes are grouped with eudicot TFL1/CEN paralogs (Fig. 3 and Additional file 1: Fig. S3). This may suggest: (1) constraining ancestral functions of the eudicot TFL1 with the basal-angiosperm TFL1-like gene (functional constraints hypothesis), (2) identical selection pressures acted on both eudicot TFL1 and magnolid TFL1-like genes synchronously (synchronous selection hypothesis), or (3) eudicot TFL1 and magnolid TFL1-like genes evolved in parallel independently, resulting in phylogenetic convergence (phylogenetic convergence hypothesis). The LRTs for the two ratio and three ratio models showed the functional constraints between magnolid TFL1-like and other paralogs (ω_1 = ω_2 ≤ 1) except the monocot RCN1 (ω_1 ≠ ω_2 ≠ ω_3) (Table 2). This indicates that (1) eudicot and monocot TFL1/CEN/RCNs could share ancestral polymorphisms and functions with TFL1-like gene of basal angiosperms, and (2) magnolid TFL1-like and monocot RCN1 could functionally converge under heterogeneous evolutionary rates. The basic functions of TFL1/CEN/RCNs paralogs in magnolids, eudicots, and monocots do not alter, but there is division of labour by small fractions of neutral or nearly neutral amino acid replacements, which is consistent with the functional divergence analysis (Table 3).
Figure 6. Ka/Ks sliding windows of 50 nucleotides with a 10-bp step size between angiosperm TFL1/CEN/RCN paralogs. Comparisons (A) between eudicot TFL1/CEN paralogs, (B) between eudicot TFL1 and monocot RCNs, (C) between eudicot CEN and monocot RCNs, (D) between monocot RCN paralogs, and (E) between magnoliid TFL1-like, and eudicot and monocot TFL1/CEN paralogs. (F) The corresponding alignment positions of exons, revealing selective constraints on exon 2 and exon 3. The midposition of windows were listed in base pair (bp).

Table 3. Summary of type-II functional divergence analysis for angiosperm TFL1/CEN/RCNs/TFL1-like paralogs. $\theta_{II}$, coefficient of type-II functional divergence (SE: standard error); $p$-value, significance test based on Z-score test to test the hypothesis of deviation of $\theta_{II}$ from zero; $a_R/\pi_R$: the ratio of radical change under functional divergence versus nonfunctional divergence; $G_R$ and $G_C$, proportion of radical change and conserved change, respectively; $F_{00,N}$, $F_{00,R}$, and $F_{00,C}$, proportion of none change, radical change, and conserved change of amino acids between clusters but no change within clusters, respectively.
The long-term constrained evolution of floral development genes across divergent species was inferred by comparative analyses of 18 angiosperm species. However, the evolutionary pattern of these genes in basal angiosperms, such as Amborella, Lauraceae, and Magnoliaceae, has not yet been investigated. Although the functional constraint hypothesis was supported between magnoliid TFL1-like genes and most other paralogs, the ω of foreground branches are larger than background lineages, supporting the hypothesis of relaxation of constraints for flexing the non-duplicated magnoliid TFL1-like genes in shaping floral diversity inferred by both Figure 7. Site-specific profile of type II functional divergence between angiosperm TFL1/CEN/RCN paralogs. Only the comparisons between the magnoliid TFL1-like and other paralogs, between eudicot TFL1 and CEN, and between monocot RCN1, RCN2, and RCN3 are shown. The full bars indicate the critical posterior ratio with a posterior probability > 0.7.
pairwise Ka/Ks (Fig. 5) and functional divergence analysis (Table 3). The relaxation of selective constraints was common for duplicated genes at the phase of early duplication that accelerated evolution of duplicated genes to escape from redundancy, while most gene duplicates were stochastically silenced with few survivors subsequently experiencing strong (10-fold efficiency) purifying selection. Here, we provide at least two novel discoveries regarding the evolution of TFL1-like genes in basal angiosperms: (1) Lauraceae and Magnoliaceae TFL1-like genes are divergent from those of Amborella and are phylogenetically similar to the eudicot TFL1/CEN; (2) purifying selection prevailed over the magnoloid TFL1-like genes as well as the eudicot and monocot paralogs, but the unfixed paralogous radical replacement enabled their differentiation through the relaxation of selective constraints.

Conclusions
In this work, we inferred evolution and functional divergence of TFL1/CEN/RCN among 18 angiosperm species, including basal angiosperm species to elucidate the duplication history of TFL1/CEN/RCN genes. We found long-term retention of functionally redundant duplicates TFL1/CEN/RCNs in the angiosperm genomes. Based on the results of purifying selection on exon, radical amino acid changes and various intron lengths with cis-acting element analysis, the maintenance and conservation of their ancestral function could be explained by duplication-degeneration-complementation model. The ancestral function of TFL1/CEN/RCNs might be preserved and divided into each duplicates. Therefore, the strong selection pressure against removing any duplicates may cause the permanent establishment of duplicates during evolution of flowering plants. Consequently, these two duplicates together maintain the conservative mechanism in inflorescence architectures, and expansion of the PEPP gene members may be important factor for driving morphological divergence among angiosperms.

Intron length of TFL1 paralogs was various. TFL1 introns of basal angiosperm tend to have longer intron and more predicted cis-acting than monocot and eudicot. On the other hand, exon length was conserved with low amino acid substitution rate. These data suggest that strong purifying selection has maintained the relevant functions of TFL1/CEN/RCNs paralogs on flowering regulation throughout the evolution of angiosperms, and the shorter introns with radical amino acid changes are important for the retention of paralogous duplicates.

Methods
Data collection and phylogenetic tree reconstruction. The full lengths of angiosperm TFL1/CEN/RCNs genes were obtained from NCBI GenBank. Organisms without complete paralogs (e.g. only TFL1 of eudicot and RCN1 of monocot organisms) were excluded. Due to high similarity among PEPP family, many sequences named with TFL1 or CEN are belonged to FT/BFT/MFT. For missing-inferring of phylogenetics of TFL1/CEN, we only included sequences which were previously identified as TFL1/CEN in our subsequent analysis e.g. ref.7. The TFL1 and CEN gene sequences from five eudicot species (Arabidopsis thaliana, A. lyrata [Brassicaceae], Citrus clementina [Rutaceae], Fragaria vesca [Rosaceae], Glycine max [Fabaceae], Medicago truncatula [Fabaceae], Populus trichocarpa [Salicaceae], Solanum lycopersicum [Solanaceae], Solanum tuberosum [Solanaceae], Vitis vinifera [Vitaceae], Linum usitatissimum [Linaceae], Kalanche-fedtschenkoi [Crassulaceae], Mimulus guttatus [Phrymynaceae], Salix purpurea [Salicaceae], Trifolium pratense [Fabaceae], Vigna unguiculata [Fabaceae], Lactuca sativa [Asteraceae], Coffea arabica [Rubiceae],), and RCN1–3 from four monocot species (Musa acuminata [Musaceae], Ananas comosus [Bromeliaceae], Zostera marina [Zosteraceae], Oryza sativa, Sorghum bicolor, Setaria italic, and Brachypodium distachyon, Panicum hallii [Poaceae]), and the TFL1-like gene from Amborella trichopoda (Amborellaceae) were obtained from GenBank. We also amplified complete TFL1-like sequences from two basal angiosperm species Lindear megaphylla (Lauraceae) and Liriodendron sp. (Magnoliaceae) using primers (MaLaTFL1-F1: 5′-ATGGCAAGAATGTTAGACGC-3′; MaLaTFL1-R1: 5′-CAATCTCTCCTCATCAGCTG-3′). Intron positions were rechecked based on the GT-AG rule. Exon-intron structures were drawn by Exon-Intron Graphic Maker (http://wormweb.org/exonintron). Exons of Litsea cubeba, Neolitsea phanerophlebia, Persea sp., and Michelia compresa (GenBank accession number: KY933631-KY933636) were also sequenced for coding region analyses. The identification of the exon sequences were conducted using BLAST. Sequences without best hit to TFL1/CEN/RCNs were discarded (e.g. FT/BFT/MFT). The phylogenetic tree of TFL1/CEN/RCNs was reconstructed by exons using the Maximum likelihood method with the GTR+I+G model, gamma distribution (α = 0.46) for substitution rate among sites using PhyML 3.0.64. The tree bisection and reconnection (TBR) was adopted for tree rearrangement and fast bootstrap method aLRT was adopted for branch supports.

Conserved motifs in introns. Conserved motifs in introns were found by searching the database of plant cis-acting regulatory DNA elements, NEW PLACE65. From 212 types of predicted motifs like cis-acting elements, 12 putative functional cis-acting elements that have been reported to regulate the expression of TFL1/CEN/RCNs paralogs were identified (Additional file 1: Table S1)63,64 and the number of these putative cis-acting elements were calculated. Correlation between the number of cis-acting elements and the total intron length (i.e. intron1 + intron2 + intron3) was estimated.

Detection of positive selection on angiosperm TFL1/CEN/RCNs paralogs. To examine the effect of selective pressures, the ω ratio, which can be used for testing the gene neutrality hypothesis Ka/Ks (ω) = 1, was estimated by maximum likelihood approaches implemented in PAML 4.766. First, the ω under the free-ratio model was estimated, which allows varied ω of selective pressures, the ratio, which can be used for testing the gene neutrality hypothesis. Consequently, these two duplicates together maintain the conservative mechanism in inflorescence architectures, and expansion of the PEPP gene members may be important factor for driving morphological divergence among angiosperms.

Intron length of TFL1 paralogs was various. TFL1 introns of basal angiosperm tend to have longer intron and more predicted cis-acting than monocot and eudicot. On the other hand, exon length was conserved with low amino acid substitution rate. These data suggest that strong purifying selection has maintained the relevant functions of TFL1/CEN/RCNs paralogs on flowering regulation throughout the evolution of angiosperms, and the shorter introns with radical amino acid changes are important for the retention of paralogous duplicates.
we estimated the $\omega$ and evaluated the goodness-of-fit of two-ratio model ($\omega = \frac{\omega_{H}}{\omega_{B}}$) and three-ratio model ($\omega = \frac{\omega_{H}}{\omega_{B}}$) by LRT. $\omega_{H}$ are the Ka/Ks ratio of background branches; $\omega_{B}$ and $\omega_{H}$ are Ka/Ks of foreground branches that allowed $\omega > 1$, where $\omega_{H}$ are the $\omega$ of eudicot TFL1/CEN or monocot RCNs paralogs and $\omega_{H}$ are the $\omega$ of magnoliid TFL1-like genes. Detailed hypotheses and selection scenarios are listed in Table 1.

In addition, pairwise Ka/Ks comparisons between angiosperm TFL1/CEN/RCNs paralogs were calculated to examine the evolutionary divergence and represented by (1) the Ka/Ks against Ks plot and (2) sliding window analysis by DnaSP 5.067. The Ka/Ks against Ks plot helps to determine the degrees and relative times of paralogous divergence and the sliding windows provide details for clarifying the divergent regions from the regions under selective constraints.

**Functional divergence between angiosperm TFL1/CEN/RCNs paralogs.** Functional divergence between paralogs caused by radical amino acid changes was assessed by the type-II divergence function implemented in DIVERGE 3.068 with 500 bootstrap replications. Substitutions between amino acids with different radical biochemical properties (charge positive/negative, hydrophilic/hydrophobic) are classified as a radical change, and all others are classified as a conserved change. The Z-test was conducted to test the deviation of coefficients of type II functional divergence ($\theta_{H}$) from zero, and value of $\theta_{H}$ greater than zero implied radical shifts in amino acid physicochemical properties after duplication. The fold of radical change related to non-functional change was calculated using the ratio of radical change under functional divergence versus nonfunctional divergence (ar/πR). The proportion of such as fixed radical change, conserved change and none change sites were also calculated. Site-specific estimation of posterior probability of radical changes was performed to assess the probable regions and shifts of biochemical properties between paralogous groups.

**Availability of data and materials.** The sequences have been submitted to GenBank with the accession number KY933631-KY933636.

**References**

1. Ahn, J. H. et al. A divergent external loop confers antagonistic activity on floral regulators FT and TFL1. *EMBO J.* 25, 605–614 (2006).
2. Bradley, D. et al. Control of inflorescence architecture in *Antirrhinum* Nature 379, 791–797, https://doi.org/10.1038/379791a0 (1996).
3. Bradley, D., Ratcliffe, O., Vincent, C., Carpenter, R. & Coen, E. Inflorescence commitment and architecture in *Arabidopsis*. *Science* 275, 80–83, https://doi.org/10.1126/science.275.5296.80 (1997).
4. Zhang, D. & Yuan, Z. Molecular control of grass inflorescence development. *Annu. Rev. Plant Biol.* 65, 553–578, https://doi.org/10.1146/annurev-arplant-052013-040108 (2014).
5. Benlloch, R., Berbel, A., Serrano-Mislata, A. & Madueno, F. Floral initiation and inflorescence architecture: A comparative view. *Ann. Bot.* 100, 659–676 (2007).
6. Prusinkiewicz, P., Erasmus, Y., Lane, B., Harder, L. D. & Coen, E. Evolution and development of inflorescence architectures. *Science* 316, 1452–1456, https://doi.org/10.1126/science.1140429 (2007).
7. Karrigan, A. et al. Evolution of the PEBP gene family in plants: functional diversification in seed plant evolution. *Plant Physiol.* 156, 1967–1977, https://doi.org/10.1104/pp.111.176206 (2011).
8. Nakagawa, M., Shimamoto, K. & Kyozuka, J. Overexpression of RCNI and RCN2, rice TERMINAL FLOWER 1/CENTRORADIALIS homologs, confers delay of phase transition and altered panicle morphology in rice. *Plant J.* 29, 743–750, https://doi.org/10.1046/j.1365-310x.2002.01255.x (2002).
9. Repinski, S. L., Kowak, M. & Gepts, P. The common bean growth habit gene PvTFL1 is a functional homolog of *Arabidopsis*. *TFL1, Theor. Appl. Genet.* 124, 1539–1547, https://doi.org/10.1007/s00122-012-1808-8 (2012).
10. Chen, M.-S. et al. Comparative transcriptome analysis between gynoecious and monoecious plants identifies regulatory networks controlling sex determination in *Jatropha curcas*. *Front Plant Sci.* 7, 1953, https://doi.org/10.3389/fpls.2016.01953 (2016).
11. Gare, R. & Wilde, H. D. Natural allelic variation in blueberry *FLOWER1* homologue. *Theor. Appl. Genet.* 129, 1967–1977, https://doi.org/10.1007/s00122-012-1808-8 (2012).
12. Prusinkiewicz, P., Erasmus, Y., Lane, B., Harder, L. D. & Coen, E. Evolution and development of inflorescence architectures. *Science* 316, 1452–1456, https://doi.org/10.1126/science.1140429 (2007).
13. Karrigan, A. et al. Evolution of the PEBP gene family in plants: functional diversification in seed plant evolution. *Plant Physiol.* 156, 1967–1977, https://doi.org/10.1104/pp.111.176206 (2011).
14. Nakagawa, M., Shimamoto, K. & Kyozuka, J. Overexpression of RCNI and RCN2, rice TERMINAL FLOWER 1/CENTRORADIALIS homologs, confers delay of phase transition and altered panicle morphology in rice. *Plant J.* 29, 743–750, https://doi.org/10.1046/j.1365-310x.2002.01255.x (2002).
15. Repinski, S. L., Kowak, M. & Gepts, P. The common bean growth habit gene PvTFL1 is a functional homolog of *Arabidopsis*. *TFL1, Theor. Appl. Genet.* 124, 1539–1547, https://doi.org/10.1007/s00122-012-1808-8 (2012).
16. Chen, M.-S. et al. Comparative transcriptome analysis between gynoecious and monoecious plants identifies regulatory networks controlling sex determination in *Jatropha curcas*. *Front Plant Sci.* 7, 1953, https://doi.org/10.3389/fpls.2016.01953 (2016).
17. Gare, R. & Wilde, H. D. Natural allelic variation in blueberry *FLOWER1* homologue. *Theor. Appl. Genet.* 129, 1967–1977, https://doi.org/10.1007/s00122-012-1808-8 (2012).
18. Prusinkiewicz, P., Erasmus, Y., Lane, B., Harder, L. D. & Coen, E. Evolution and development of inflorescence architectures. *Science* 316, 1452–1456, https://doi.org/10.1126/science.1140429 (2007).
19. Karrigan, A. et al. Evolution of the PEBP gene family in plants: functional diversification in seed plant evolution. *Plant Physiol.* 156, 1967–1977, https://doi.org/10.1104/pp.111.176206 (2011).
20. Nakagawa, M., Shimamoto, K. & Kyozuka, J. Overexpression of RCNI and RCN2, rice TERMINAL FLOWER 1/CENTRORADIALIS homologs, confers delay of phase transition and altered panicle morphology in rice. *Plant J.* 29, 743–750, https://doi.org/10.1046/j.1365-310x.2002.01255.x (2002).
21. Repinski, S. L., Kowak, M. & Gepts, P. The common bean growth habit gene PvTFL1 is a functional homolog of *Arabidopsis*. *TFL1, Theor. Appl. Genet.* 124, 1539–1547, https://doi.org/10.1007/s00122-012-1808-8 (2012).
22. Chen, M.-S. et al. Comparative transcriptome analysis between gynoecious and monoecious plants identifies regulatory networks controlling sex determination in *Jatropha curcas*. *Front Plant Sci.* 7, 1953, https://doi.org/10.3389/fpls.2016.01953 (2016).
23. Gare, R. & Wilde, H. D. Natural allelic variation in blueberry *FLOWER1* homologue. *Theor. Appl. Genet.* 129, 1967–1977, https://doi.org/10.1007/s00122-012-1808-8 (2012).
24. Ballesteri, E. S. & Kramer, E. M. In the light of evolution: a reevaluation of conservation in the CO-FT regulon and its role in photoperiodic regulation of flowering time. *Plant Cell* 2, 81, Art1 81 https://doi.org/10.1103/Fpl.2011.00081 (2011).
25. Hanzawa, Y., Money, T. & Bradley, D. A single amino acid converts a repressor to an activator of flowering. *P Natl. Acad. Sci. USA* 102, 7748–7753, https://doi.org/10.1073/pnas.0500932102 (2005).
26. Guo, Y., Hans, H., Christian, J. & Molina, C. Mutations in single FF- and TFL1-paralogs of rapeseed (*Brassica napus*) L. and their impact on flowering time and yield components. *Plant Sci. 5*, 282, https://doi.org/10.3389/fpls.2014.00282 (2014).
27. Song et al., Johnston, C. R. & Shields, D. C. Significantly different patterns of amino acid replacement after gene duplication as compared to after speciation. *Mol. Biol. Evol.* 20, 484–490, https://doi.org/10.1093/molbeevomo.95 (2003).
28. Gu, X. Maximum-likelihood approach for gene family evolution under functional divergence. *Mol. Biol. Evol.* 18, 453–464 (2001).
29. Gu, X. A simple statistical method for estimating Type-II (Cluster-Specific) functional divergence of protein sequences. *Mol. Biol. Evol.* 23, 1937–1945 (2006).
30. Xu, G. X., Guo, C. C., Shao, H. Y. & Kong, H. Z. Divergence of duplicate genes in exon-intron structure. *P Natl. Acad. Sci. USA* 109, 1187–1192, https://doi.org/10.1073/pnas.0904110109 (2012).
31. Duester, G., Jornvall, H. & Hatfield, G. W. Intron-dependent evolution of the nucleotide-binding domains within alcohol dehydrogenase and related enzymes. *Nucl. Acids Res.* 14, 1931–1941, https://doi.org/10.1093/nar/14.5.1931 (1986).
32. Kirby, D. A., Muse, S. V. & Stephan, W. Maintenance of pre-mRNA secondary structure by epistatic selection. *P Natl. Acad. Sci. USA* 92, 9047–9051 (1995).
33. Ren, X. Y., Vorst, O., Fiers, M. W. E. J., Stiekema, W. J. & Nap, J. P. In plants, highly expressed genes are the least compact. *BMC Genom.* 12, 557–564, https://doi.org/10.1186/1471-2164-13-S1-S10 (2012).
34. Lehnert-Mack, J. et al. Identifying the basal angiosperm node in chloroplast genome phylogenies: sampling one’s way out of the Felsenstein Zone. *Mol. Biol. Evol.* 22, 1948–1963, https://doi.org/10.1093/molbev/msi191 (2005).
35. Moore, M. J., Bell, C. D., Solis, P. S. & Solis, D. E. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proceedings of the National Academy of Sciences* 104, 19363–19368, https://doi.org/10.1073/pnas.0708072104 (2007).
36. Jansen, R. K. et al. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proceedings of the National Academy of Sciences* 104, 19369–19374, https://doi.org/10.1073/pnas.0709121104 (2007).
37. Gu, X. et al. An update of DIVERGE software for functional divergence analysis of protein family. *Mol. Biol. Evol.* 30, 1713–1719, https://doi.org/10.1093/molbev/msn069 (2013).
38. Fu, G. C. L. & Lin, W. C. Identification of gene-oriented exon orthology between human and mouse. *BMC Genomics* 13, 510, doi:10.1186/1471-2164-13-S1-S10 (2012).
39. Liu, C. et al. A conserved genetic pathway determines inflorescence architecture in Arabidopsis and rice. *Dev. Cell* 24, 612–622, https://doi.org/10.1016/j.devcel.2013.02.013 (2013).
40. Hu, K. J. Intron exclusion and the mystery of intron loss. *FEBS Lett.* 580, 6361–6365, https://doi.org/10.1016/j.fsl.2006.10.048 (2006).
41. Park, K. C., Kwon, S. J. & Kim, N. S. What genes are in pieces? A genomics perspective. *Genes Genom* 30, 439–437 (2008).
42. Rose, A. B., Carter, A., Korf, I. & Kojima, N. Intron sequences that stimulate gene expression in *Arabidopsis*. *Plant Mol. Biol. 92*, 337–346, https://doi.org/10.1007/s11103-016-0516-1 (2016).
43. Tsafiraris, A. et al. Isolation of a CENTRORADIALIS/TERMINAL FLOWER1 homolog in saffron (*Crocus sativus* L.): characterization and expression analysis. *Mol. Biol. Rep.* 39, 7899–7910, https://doi.org/10.1007/s11033-012-1634-8 (2012).
44. Kaufmann, K. et al. Orchestration of floral initiation by APETALA1. *Science* 328, 85–89, https://doi.org/10.1126/science.1185244 (2010).
45. Moyroud, E. et al. Prediction of regulatory interactions from genome sequences using a biophysical model for the Arabidopsis LEAFY transcription factor. *Plant Cell* 23, 1293–1306, https://doi.org/10.1105/tpc.110.083329 (2011).
46. Winter, C. M. et al. LEAFY target genes reveal floral regulatory logic, cis motifs, and a link to biotic stimulus response. *Dev. Cell* 20, 430–443, https://doi.org/10.1016/j.devcel.2011.03.019 (2011).
47. Qian, W. F., Liao, B. Y., Chang, A. Y. F. & Zhang, J. Z. Maintenance of duplicate genes and their functional redundancy by reduced mutational constraint. *Proc. Natl. Acad. Sci. USA* 102, 1641–1649, https://doi.org/10.1073/pnas.0406909102 (2005).
48. Kang, S., Lee, J., Lee, M. S. & Seok, C. Structural basis of functional conversion of a floral repressor to an activator: A molecular dynamics simulation study. *B Korean Chem Soc* 29, 402–412 (2008).
49. Force, A. et al. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151, 1531–1545 (1999).
50. Hittinger, C. T. & Carroll, S. B. Gene duplication and the adaptive evolution of a classic genetic switch. *Nature* 449, 677–681, https://doi.org/10.1038/nature06151 (2007).
51. Panzhi, N., Lehti-Shiu, M. & Shiu, S. H. Evolution of gene duplication in plants. *Plant Physiol.* 171, 2294–2316, https://doi.org/10.1104/pp.110.1600523 (2010).
52. Flagel, L. E. & Wendel, J. F. Gene duplication and evolutionary novelty in plants. *New Phytol.* 183, 557–564, https://doi.org/10.1111/j.1469-8137.2009.02923.x (2009).
53. Quignon, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321, https://doi.org/10.1093/sysbio/syp010 (2010).
54. Higo, K., Uwamoto, Y. & Korenaga, T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucl. Acids Res.* 27, 297–300, https://doi.org/10.1093/nar/27.1.297 (1999).
55. Benlloch, R. et al. Integrating long-day flowering signals: a LEAFY binding site is essential for proper photoperiodic activation of APETALA1. *Plant J.* 67, 1094–1102, https://doi.org/10.1111/j.1365-313X.2011.04660.x (2011).
56. West, A. G., Shore, P. & Sharrocks, A. D. DNA binding by MADS-box transcription factors: A molecular mechanism for differential DNA bending. *Mol. Cell. Biol.* 17, 2876–2887 (1997).
57. Roudoux, M. et al. Gibberellins regulate the transcription of the continuous flowering regulator, *RoKSN*, a rose *TFL1* homologue. *J. Exp. Bot.* 65, 6543–6554, https://doi.org/10.1093/jxb/ers310 (2012).
65. Fu, J. X. et al. Photoperiodic control of FT-like gene ClFT initiates flowering in Chrysanthemum lavandulifolium. Plant Physiol. Biochem. 74, 230–238, https://doi.org/10.1016/j.plaphy.2013.11.004 (2014).
66. Yang, Z. H. PAML 4: Phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24, 1586–1591, https://doi.org/10.1093/molbev/msm088 (2007).
67. Librado, P. & Rozas, J. DnaSPv5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452, https://doi.org/10.1093/bioinformatics/btp187 (2009).

Acknowledgements
We thank Li-Ping Ju for providing specimens, and the FuShan Botanical Garden, a long-term ecological research (LTER) site for providing samples. We also gratefully thank Chun-Neng Wang for his insightful comments on this paper. This research was financially supported by the Ministry of Science and Technology in Taiwan (MOST 105–2628-B-003-001-MY3 and MOST 105-2628-B-003-002-MY3) and was also subsidized by the National Taiwan Normal University (NTNU), Taiwan.

Author Contributions
P.C.L. conceived this study. B.H.H. and J.Y.C. collected materials. J.G. and Y.T.W. conducted experiments. B.H.H. and P.C.L. analysed the data. P.C.L. wrote the paper. J.G., B.H.H., and J.Q.L. critically reviewed the manuscript. All authors participated in the discussion and read and approved the final manuscript.

Additional Information
Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-017-13645-0.

Competing Interests: The authors declare that they have no competing interests.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2017