

RESEARCH ARTICLE

Open Access



# *De novo* transcriptome analysis in radish (*Raphanus sativus* L.) and identification of critical genes involved in bolting and flowering

Shanshan Nie<sup>1,2</sup>, Chao Li<sup>1,2</sup>, Liang Xu<sup>1,2</sup>, Yan Wang<sup>1,2</sup>, Danqiong Huang<sup>3</sup>, Everlyne M. Muleke<sup>1,2</sup>, Xiaochuan Sun<sup>1,2</sup>, Yang Xie<sup>1,2</sup> and Liwang Liu<sup>1,2\*</sup>

## Abstract

**Background:** The appropriate timing of bolting and flowering is pivotal for reproductive success in Brassicaceae crops including radish (*Raphanus sativus* L.). Although several flowering regulatory pathways had been described in some plant species, no study on genetic networks of bolting and flowering regulation was performed in radish. In this study, to generate dataset of radish unigenes for large-scale gene discovery and functional pathway identification, a cDNA library from mixed radish leaves at different developmental stages was subjected to high-throughput RNA sequencing (RNA-seq).

**Results:** A total of 54.64 million clean reads and 111,167 contigs representing 53,642 unigenes were obtained from the radish leaf transcriptome. Among these, 50,385 unigenes were successfully annotated by BLAST searching against the public protein databases. Functional classification and annotation indicated that 42,903 and 15,382 unique sequences were assigned to 55 GO terms and 25 COG categories, respectively. KEGG pathway analysis revealed that 25,973 unigenes were classified into 128 functional pathways, among which 24 candidate genes related to plant circadian rhythm were identified. Moreover, 142 potential bolting and flowering-related genes involved in various flowering pathways were identified. In addition, seven critical bolting and flowering-related genes were isolated and profiled by T-A cloning and RT-qPCR analysis. Finally, a schematic network model of bolting and flowering regulation and pathways was put forward in radish.

**Conclusions:** This study is the first report on systematic identification of bolting and flowering-related genes based on transcriptome sequencing and assembly in radish. These results could provide a foundation for further investigating bolting and flowering regulatory networks in radish, and facilitate dissecting molecular genetic mechanisms underlying bolting and flowering in Brassicaceae vegetable crops.

**Keywords:** Radish, *De novo* assembly, Bolting and flowering, Transcriptome, Flowering regulatory network

\* Correspondence: nauliulw@njau.edu.cn

<sup>1</sup>National Key Laboratory of Crop Genetics and Germplasm Enhancement, College of Horticulture, Nanjing Agricultural University, Nanjing 210095, People's Republic of China

<sup>2</sup>Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (East China) of the Ministry of Agriculture of P.R. China, Nanjing 210095, People's Republic of China

Full list of author information is available at the end of the article



## Background

The formation of bolting and flowering, marking the developmental transition from vegetative growth to flowering, is one of the most important developmental traits in plant life cycle. Appropriate bolting and flowering time is crucial to ensure reproductive success and high agricultural productivity for crop plants [1, 2]. In *Arabidopsis thaliana*, some molecular and genetic studies have revealed that a complex gene-network has evolved to orchestrate the initiation of bolting and flowering and is regulated by diverse environmental and endogenous factors such as temperature, light signals, day length, developmental stage and plant hormones [3–5].

Considerable studies in *Arabidopsis* revealed that a set of genes involved in flowering control have been discovered and characterized and integrated into the genetic circuitry of flowering and multiple flowering pathways, including pathways of photoperiod, vernalization, aging, autonomous and gibberellin (GA) [2, 3, 6]. The transition from vegetative growth to reproductive development is regulated by the interplays of central flowering genes involved in different flowering pathways [7, 8]. The pathways of autonomous and vernalization are integrated by *FLOWERING LOCUS C (FLC)*, which is a main flowering repressor and encodes a MADS-box transcription factor [9]. The expression level of *FLC* is correlated with plant flowering time and overexpression of *FLC* causes late flowering [2, 9]. In addition, *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1)*, *FLOWERING LOCUS T (FT)* and *LEAFY (LFY)* are showed to be convergence points for different flowering pathways and defined as flowering pathway integrators [7, 8, 10]. In the past decade, a number of functional genes and regulatory pathways related to flowering time had also been discovered in many crops including rice [11, 12], maize [13], strawberry [14] and soybean [15].

With the development of next-generation sequencing (NGS) technologies, several platforms such as Illumina Solexa, Roche 454 and ABI SOLiD, have been proven to be powerful and efficient tools for advanced genomic researches, including the whole genome re-sequencing and transcriptome sequencing [16, 17]. Recently, the advent of NGS-based RNA sequencing (RNA-seq) of transcriptome has provided an opportunity for large-scale discovery and profiling of functional genes involved in diverse biological processes [18, 19], such as floral scent biosynthesis in wintersweet [20], discovery of disease-resistance genes (NBS-LRR genes) in *Camelina sativa* [21], biosynthesis and accumulation of lignin in celery [22] and terpenoid biosynthesis in pitanga [23]. RNA-seq has also been used for the discovery of genes related to flowering time regulation and flower development in

some plant species including bamboo [24, 25], *Eichhornia paniculata* [26], *Lagerstroemia indica* [27], *Ipomoea nil* [28] and sweet potato [29]. However, genome-wide identification of flowering-related functional genes is still lacking in root vegetable crops.

Radish ( $2n = 2x = 18$ ) is an important annual or biennial root vegetable crop of Brassicaceae family and belongs to LD (long-day) plant. Bolting and flowering are integral phases in the complete life cycle of radish. Premature bolting is a destructive problem that limits vegetative growth and reduces yield and quality of economic products in Brassicaceae crops, especially the radish grown in spring. The timing of bolting and flowering is of great importance for high productivity. Recently, the draft genomic sequences of radish have been assembled [30], providing useful database for genomic research in radish. Moreover, some recent studies concerning the radish transcriptome assembly and analysis were reported, and a list of radish unique sequences were generated from various radish tissues [31, 32]. Previous studies reported that several flowering genes have been isolated and characterized by expression profiling and transgenic approaches [33–35]. However, the genetic networks of bolting and flowering regulation in radish are poorly understood. To characterize the radish leaf transcriptome and identify critical genes involved in bolting and flowering of radish, transcript sequences from radish leaves during different developmental phases were isolated and sequenced by Illumina sequencing method. The goal of this study was to obtain a complete set of assembled unigenes and transcripts in radish leaves, and to identify transition of bolting and flowering related genes involved in flowering-time regulatory networks. The full-length cDNA sequences of seven flowering related genes were isolated by T-A cloning and further compared with the assembled transcript sequences from radish leaf transcriptome. Moreover, expression patterns of seven isolated genes were validated by RT-qPCR analysis. Furthermore, a graphical network of bolting and flowering-time regulation was proposed. These findings could provide a solid foundation for better understanding of bolting and flowering-time regulatory networks in radish, and provide significant insights into molecular genetic mechanisms underlying bolting and flowering regulation in Brassicaceae vegetable crops.

## Results and discussion

**Illumina sequencing and *de novo* transcriptome assembly**  
RNA-seq is a useful approach for obtaining a complete set of transcripts from certain plant tissue at specific developmental stage or under certain physiological condition [18, 19]. To obtain a comprehensive overview of radish leaf transcriptome, a cDNA library from radish leaves of late-bolting 'NAU-LU127' at different

developmental stages was constructed and sequenced using Illumina RNA-seq. A total of 58,602,240 pair-end reads were generated from the radish leaf transcriptome, named 'NAU-LB'. After the removal of adapter sequences, ambiguous reads and low quality reads, 54,637,700 high-quality clean reads comprising of 4,917,393,000 nucleotides (nt) were obtained with an average length of 90 nt and average GC content of 47.12 % (Table 1). All clean reads were assembled into 111,167 contigs using the Trinity program [36], with a mean length of 429 nt and an N50 length of 476 nt (Table 2). Then, these contigs were assembled into 53,642 unigenes with a total length of 43,022,520 nt, an average length of 802 nt and an N50 length of 1169 nt (Table 2). Length distributions of assembled contigs and unigenes were analyzed and demonstrated in Fig. 1.

Using the RNA-seq platform and Trinity program, a number of assembled unigenes were also generated from radish leaves and roots [31, 32]. Zhang et al. [31] reported that 28,410 unigenes with an average length of 394 bp and an N50 of 422 bp was generated from radish leaves by transcriptome assembly. Wang et al. [32] showed that 73,084 unigenes with a mean length of 763 nt and an N50 of 1095 nt were obtained from radish root transcriptome. To evaluate the quality of assembly results and transcriptome data, comparison analysis showed that the average length and N50 length of unigenes obtained in this study were longer than that in previous studies on radish, indicating these sequencing data were of high quality and sufficient quantity for further analysis.

To estimate the expression levels of assembled unigenes in radish leaf transcriptome, the FPKM (Fragments Per kb per Million fragments) method [37] was used for the calculation of unigene expression level. The global detailed information of all the assembled unigenes was showed in Additional file 1. It is well known that the *de novo* transcriptome sequencing had been successfully applied for many vegetable crops including *B. oleracea* [38], *B. napus* [39] and celery [22]. Consistent with these studies, the transcript sequences from radish leaf transcriptome could

**Table 1** Summary of Illumina transcriptome sequencing from radish leaves

Samples	NAU-LB
Total raw reads	58,602,240
Total clean reads	54,637,700
Total clean nucleotides (nt)	4,917,393,000
Average length of clean reads (nt)	90
Q20 percentage	98.47 %
N percentage	0.01 %
GC percentage	47.12 %

**Table 2** Summary of radish leaf transcriptome assembly

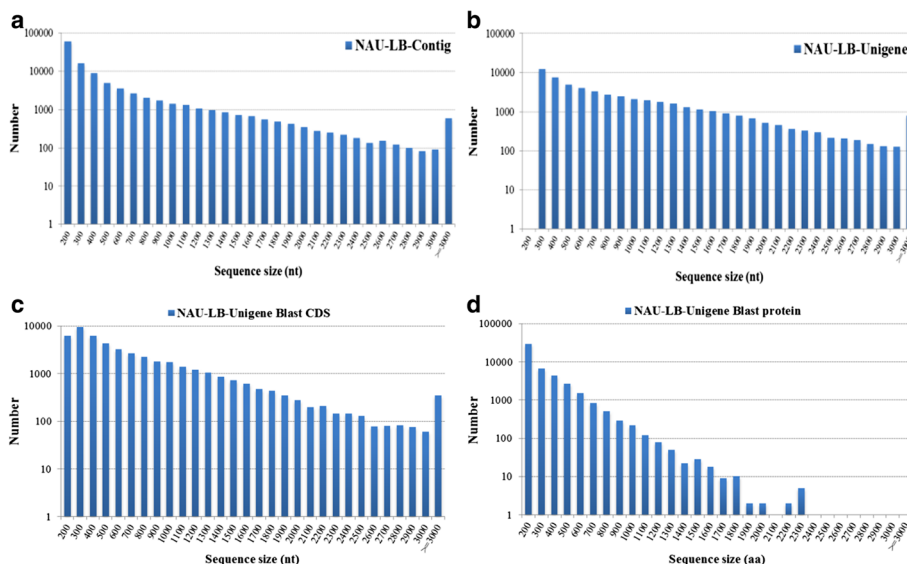
	Contig	Unigene
Total number	111,167	53,642
Total Length(nt)	47,730,600	43,022,520
Mean Length(nt)	429	802
N50	476	1169
Total consensus sequences	–	53,642
Distinct clusters	–	22,665
Distinct singletons	–	30,977

provide a valuable resource for comprehensively identifying specific developmental processes, pathways and functional genes in radish.

### Functional annotation of radish leaf transcriptome

To assign putative gene functions and annotations of all the assembled unigenes, BLASTx alignment was performed with an *E*-value threshold of 1.0E-05 against the public protein databases including NCBI non-redundant protein (NR), Swiss-Prot, Clusters of Orthologous Groups (COG) and Kyoto Encyclopedia of Genes and Genomes (KEGG). A total of 50,385 unigenes (93.93 % of all unigenes) were annotated and matched to one or more of the public protein databases (Table 3; Additional file 1). The best aligning results were used to decide the sequence direction of each assembled unigene. The size distributions for the coding sequence (CDS) and predicted proteins were analyzed and shown in Fig. 1. Nevertheless, a small proportion of sequences (3,257 unigenes, 6.07 %) was unannotated and showed no matches to any above protein databases, which might represent radish-specific transcriptomic sequences in radish leaves.

By sequence similarity searching against the NCBI NR database, 46,660 annotated sequences (86.98 % of all unigenes) were detected and showed significant similarities to known proteins (Table 3). According to the *E*-value distribution of significant hits against the NR database, 59.38 % of the matched sequences showed high homologies with the *E*-value < 1.0E-45, while 40.62 % of the top hits had *E*-value in the range of 1.0E-05 to 1.0E-45 (Fig. 2a). The similarity distribution of the top BLAST hits revealed that 57.55 % of the matched sequences were in similarity higher than 80 % (Fig. 2b). Further analysis of homologies among different plant species showed that 42.69 % of the annotated sequences matched to sequences from *Arabidopsis lyrata*, followed by *A. thaliana* (42.47 %), *Thellungiella halophila* (3.27 %) and *B. napus* (2.12 %) (Fig. 2c). The species distribution revealed that most of annotated unigenes had high hits with sequences from the Brassicaceae species, suggesting that the assembly and annotation of radish leaf transcriptome are correct and reliable.



**Fig. 1** Overview of the radish leaf transcriptome assembly. The length distribution of assembled contigs (a) and unigenes (b), and the CDS (c) and predicted proteins (d) by BLASTx alignment in radish

**GO and COG classification**

Based on the BLASTx results against the NR database, Blast2GO program was performed to obtain Gene Ontology (GO) functional annotations and categorizations of these assembled unigenes. As a result, 42,903 unique sequences were annotated and classified into 55 GO classes including 22 biological processes, 17 cellular components and 16 molecular functions (Fig. 3). Under the biological process category, ‘cellular process’ (30,031 sequences), ‘metabolic process’ (28,182) and ‘single-organism process’ (23,665) were prominently represented. Within these cellular components, ‘cell’ (39,426), ‘cell part’ (39,426) and ‘organelle’ (33,017) were the most highly represented categories. For the molecular function category, the largest proportion of genes were clustered into ‘binding’ (22,010), ‘catalytic activity’ (18,298) and ‘nucleic acid binding transcription factor activity’ (3,242). Additionally, only a few sequences were assigned to ‘virion’, ‘virion part’, ‘channel regulator activity’ and

‘translation regulator activity’ terms with less than ten sequences.

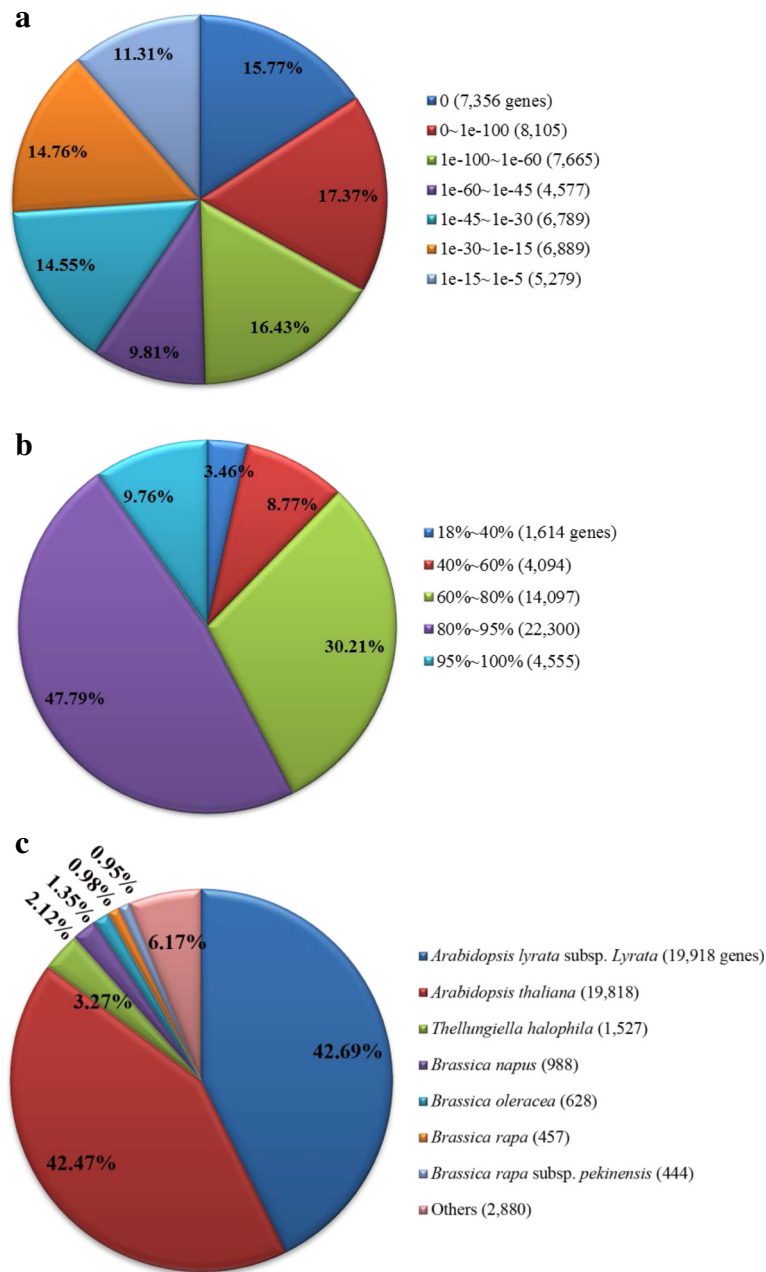
To further predict and classify the possible functions of radish unigenes, these unique sequences from radish leaf transcriptome were aligned to the COG database. The results showed that 15,382 unigenes (28.68 % of all unigenes) were assigned to 25 COG categories (Table 3; Fig. 4). Of these, the largest category was ‘General function prediction only’ (5,088 sequences), followed by ‘Transcription’ (3,019), ‘Replication, recombination and repair’ (2,568), ‘Posttranslational modification, protein turnover, chaperones’ (2,362) and ‘Signal transduction mechanisms’ (2,091); whereas the clusters of ‘Extracellular structures’ (3) and ‘Nuclear structure’ (7) represented the smallest classifications. These assigned functions of unigenes covered a wide range of GO and COG classifications, indicating that the transcriptome data from radish leaves represented a broad variety of transcripts in radish.

**Table 3** The annotations of radish leaf unigenes against the public databases

Sequence file	NAU-LB-Unigene	Percent (%)
NR	46,660	86.98
NT	49,051	91.44
Swiss-Prot	30,955	57.71
KEGG	25,973	48.42
COG	15,382	28.68
GO	42,903	79.98
ALL	50,385	93.93

**Identification of functional genes involved in circadian rhythm**

To systematically study the complex biological functions of genes and identify active biological pathways in radish leaves, the assembled unigenes were mapped against the KEGG database using BLASTx analysis with a cut-off *E*-value of 1.0E-05. In total, 25,973 unigenes (48.42 % of all unigenes) had significant matches to the database and were assigned to 128 KEGG pathways (Table 3; Additional file 2). The most dominant pathways were ‘metabolic pathways’ (ko01100; 5,522 unigenes), ‘biosynthesis of secondary metabolites’ (ko01110; 2,573),



**Fig. 2** Characteristics of sequence homology of radish unigenes against NR database. **a** E-value distribution of BLAST hits for each unigene with an E-value cutoff of 1.0E-05. **b** Similarity distribution of the top BLAST hits for each unigene. **c** Species distribution of the top BLAST hits

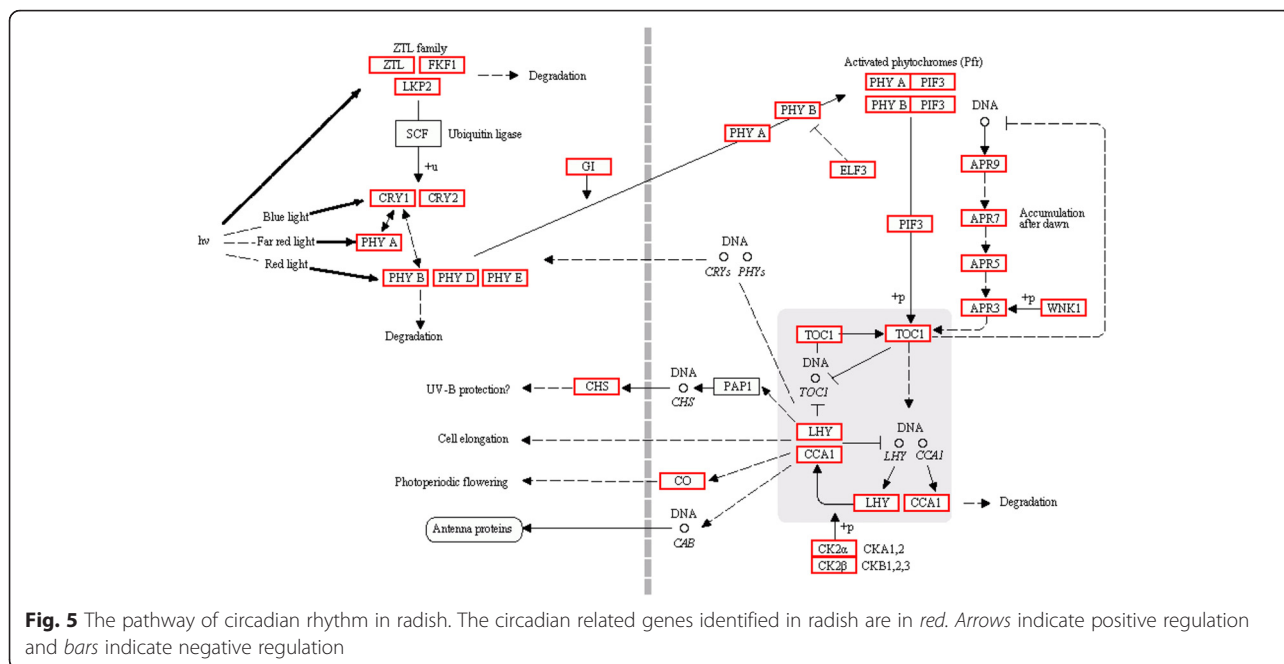
'plant hormone signal transduction' (ko04075; 1,495), 'plant-pathogen interaction' (ko04626; 1,436) and 'RNA transport' (ko03013; 1,088) (Additional file 2). The results were comparable with the previous transcriptome profiles in radish root, in which 33,567 unigenes were clustered into 128 KEGG pathways [32].

Significantly, pathway-based analysis revealed that a total of 249 transcript sequences representing 24 rhythm related candidate genes were identified for the 'circadian rhythm-plant' pathway (ko04712) (Fig. 5; Additional file 3).

These rhythm related genes including *CONSTANS* (*CO*, K12135), *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*, K12134), *EARLY FLOWERING 3* (*ELF3*, K12125), *LATE ELONGATED HYPOCOTYL* (*LHY*, K12133) and *GIGANTEA* (*GI*, K12124), were involved in many rhythmic processes including photoperiodic flowering, UV-B protection and cell elongation (Fig. 5). In addition, several identified potential genes including *LOV KELCH PROTEIN 2* (*LKP2*, K12117), *CRYPTOCHROME 1* (*CRY1*, K12118), *CRYPTOCHROME 2* (*CRY2*, K12119), *PHYTOCHROME A* (*PHYA*,







**Fig. 5** The pathway of circadian rhythm in radish. The circadian related genes identified in radish are in red. Arrows indicate positive regulation and bars indicate negative regulation

[2, 3]. In this study, to comprehensively identify the candidate genes putatively implicated in bolting and flowering regulation in radish, a local BLASTx similarity search was performed against *A. thaliana* flowering genes which were downloaded from NCBI database. A total of 142 potential bolting and flowering-related genes including 373 transcript sequences were identified and shared higher similarity to the homologous genes from *Arabidopsis* (Additional files 4 and 5). According to the reported flowering pathways and regulatory networks, these identified bolting and flowering-related genes were characterized and involved in various flowering pathways (Table 4), including photoperiod/circadian clock, vernalization, autonomous, GA and aging pathways [2, 3]. The BLAST searching showed that 71 potential functional genes including *CONSTANS-LIKE 1 (COL1)*, *TEMPRANILLO 1 (TEM1)*, *TEM2*, *TERMINAL FLOWER 2 (TFL2)* and *FLOWERING LOCUS D (FD)*, were identified and implicated in photoperiod pathway (Table 4; Additional file 4). In addition, the pathways of vernalization, autonomous, GA and aging contained 28, 8, 16 and 4 candidate genes, respectively. These results revealed that the known genetic flowering pathways and many critical flowering genes shared a high degree of conservation in radish and *Arabidopsis*.

Moreover, to summarize and characterize the identified functional genes and different flowering pathways, a putative schematic network of bolting and flowering regulation in radish was proposed (Fig. 6). More importantly, many key players of flowering regulation including *SOC1*, *CO*, *FT* and *FLC* and several functional genes

associated with floral meristem identity and flower development including *APETALA1 (API)*, *SEPALLATA1-3 (SEP1-3)*, *AGAMOUS-LIKE 24 (AGL24)* and *LFY*, were identified in this study (Table 4; Fig. 6). Similar results were also found in previous genetic analysis studies on rice [12], barley [47], soybean [15], maize [13] and strawberry [14]. These results provide essential information for functional analysis of bolting and flowering genes and regulatory pathways, and facilitate further genetically investigating in radish.

**Characterization of bolting and flowering genes and regulatory pathways in radish**

Studies on developmental switch of *Arabidopsis* flowering have demonstrated that multiple different flowering pathways converge at several flowering pathway integrators including *FT*, *SOC1* and *LFY* to determine the flowering time [2, 7, 10]. These integrators are regulated by two key upstream genes *CO* and *FLC*, which antagonistically control flowering [1, 7]. *FLC* as a main flowering repressor integrates autonomous and vernalization pathways and negatively regulates the expressions of *FT* and *SOC1* [48, 49]. Previous studies had identified the homologous genes of *FLC* in *B. rapa* [50], sweet potato [29] and orange [51]. In this study, three unique sequences of *FLC* gene, CL1584.Contig1, CL1584.Contig2 and CL1584.Contig3, were identified and matched to the homologs of *AtFLC* (Additional file 4). After undergoing vernalization process, a long exposure to low temperature, the plants permit to initiate the phase transition of flowering. While the process could be repressed by the high level of *FLC* [9, 52], which

**Table 4** The identified candidate genes involved in various flowering pathways of radish

Gene name	Number of unigenes	Flowering pathway
<i>ABH1/ENS</i>	1	Photoperiod
<i>AGL12</i>	1	Others
<i>AGL14</i>	1	Others
<i>AGL15</i>	2	Photoperiod
<i>AGL16</i>	1	Others
<i>AGL18</i>	3	Photoperiod
<i>AGL19</i>	2	Vernalization
<i>AGL24</i>	2	Photoperiod
<i>AGL6</i>	7	Photoperiod
<i>AGL8/FUL</i>	4	Photoperiod
<i>AP1</i>	5	Photoperiod
<i>AP2</i>	6	Photoperiod
<i>AP3</i>	3	Photoperiod
<i>ARP4</i>	3	Photoperiod
<i>ARP6</i>	1	Vernalization
<i>ARR3</i>	2	Photoperiod
<i>ATH1</i>	3	flower development
<i>ATX1</i>	6	flower development
<i>CBF1</i>	1	Vernalization
<i>CBF2</i>	1	Vernalization
<i>CCA1</i>	3	Photoperiod/Circadian clock
<i>CCR2</i>	3	Autonomous
<i>CDF1</i>	5	Photoperiod
<i>CDF2</i>	1	Photoperiod
<i>CDF3</i>	1	Photoperiod
<i>CO</i>	3	Photoperiod/Circadian clock
<i>COL1</i>	2	Photoperiod/Circadian clock
<i>COL2</i>	1	Photoperiod/Circadian clock
<i>COL3</i>	1	Photoperiod
<i>COL4</i>	3	Photoperiod
<i>COL5</i>	4	Photoperiod
<i>COL9</i>	2	Photoperiod/Circadian clock
<i>COP1</i>	1	Photoperiod
<i>CRY1</i>	4	Photoperiod/Circadian clock
<i>CRY2</i>	1	Photoperiod/Circadian clock
<i>CSTF64</i>	2	Vernalization
<i>CSTF77</i>	4	Vernalization
<i>DDF1</i>	7	GA
<i>DDF2</i>	2	GA
<i>DDL</i>	1	flower development
<i>EBS</i>	3	Photoperiod
<i>EFS</i>	2	flower development
<i>ELF3</i>	6	Photoperiod/Circadian clock

**Table 4** The identified candidate genes involved in various flowering pathways of radish (Continued)

Gene name	Number of unigenes	Flowering pathway
<i>ELF4</i>	1	Photoperiod/Circadian clock
<i>ELF5</i>	1	Photoperiod
<i>ELF6</i>	1	Photoperiod
<i>ELF7</i>	2	Photoperiod
<i>ELF8</i>	1	Photoperiod
<i>ELF9</i>	1	Photoperiod
<i>EMF1</i>	1	Vernalization
<i>ESD4</i>	1	Vernalization
<i>FCA</i>	2	Autonomous
<i>FD</i>	4	Photoperiod
<i>FES1</i>	2	Vernalization
<i>FIE</i>	3	others
<i>FIP1</i>	2	Vernalization
<i>FIP2</i>	1	Vernalization
<i>FLC</i>	3	integrator
<i>FLD</i>	1	Autonomous
<i>FLK</i>	4	Autonomous
<i>FLM/MAF1</i>	1	Photoperiod
<i>FLP1</i>	1	flower development
<i>FPA</i>	3	Autonomous
<i>FPF1</i>	1	GA
<i>FRI</i>	1	Vernalization
<i>FT</i>	1	integrator
<i>FVE/MSI4</i>	7	Autonomous
<i>FY</i>	1	Autonomous
<i>GA3</i>	1	GA
<i>GAI</i>	2	GA
<i>GI</i>	5	Photoperiod/Circadian clock
<i>GID1A</i>	2	GA
<i>GID1B</i>	4	GA
<i>GID1C</i>	2	GA
<i>GNC</i>	2	GA
<i>GNL</i>	2	GA
<i>HUA2</i>	5	Vernalization
<i>HYL1/DRB1</i>	2	others
<i>LCL1</i>	3	Photoperiod/Circadian clock
<i>LD</i>	4	Autonomous
<i>LDL2</i>	1	flower development
<i>LFY</i>	3	integrator
<i>LHP1</i>	3	Vernalization
<i>LHY</i>	2	Photoperiod/Circadian clock
<i>LKP2</i>	3	Photoperiod/Circadian clock
<i>LWD1/ATAN11</i>	1	Photoperiod/Circadian clock
<i>LWD2</i>	1	Photoperiod/Circadian clock



**Table 4** The identified candidate genes involved in various flowering pathways of radish (*Continued*)

<i>MBD9</i>	6	Photoperiod
<i>MYB33</i>	2	GA
<i>MYB65</i>	3	GA
<i>NF-YA1/HAP2A</i>	1	Photoperiod
<i>NF-YA4</i>	3	Photoperiod
<i>NF-YB1</i>	2	Photoperiod
<i>NF-YB2</i>	3	Photoperiod
<i>pEARLI</i>	2	Vernalization
<i>PHYA</i>	2	Photoperiod/Circadian clock
<i>PHYB</i>	2	Photoperiod/Circadian clock
<i>PHYC</i>	3	Photoperiod/Circadian clock
<i>PHYD</i>	1	Photoperiod/Circadian clock
<i>PHYE</i>	9	Photoperiod/Circadian clock
<i>PI</i>	2	flower development
<i>PIE1</i>	4	Photoperiod
<i>PRC2/CLF/EMF2</i>	5	Vernalization
<i>PRR3</i>	3	Photoperiod/Circadian clock
<i>PRR7</i>	7	Photoperiod/Circadian clock
<i>PRR9</i>	1	Photoperiod/Circadian clock
<i>REF6</i>	6	Vernalization
<i>RGA</i>	2	GA
<i>RGL1</i>	2	GA
<i>RGL2</i>	1	GA
<i>SDG10/SWN</i>	2	Vernalization
<i>SEP1/AGL2</i>	4	Others
<i>SEP2/AGL4</i>	4	Others
<i>SEP3/AGL9</i>	6	flower development
<i>SHP1/AGL1</i>	3	flower development
<i>SHP2/AGL5</i>	3	flower development
<i>SKB1</i>	2	Vernalization
<i>SOC1/AGL20</i>	2	integrator
<i>SPA</i>	9	Photoperiod
<i>SPL</i>	2	Age
<i>SPL3</i>	4	Age
<i>SPL5</i>	2	Age
<i>SPL9</i>	4	Age
<i>SPY</i>	3	GA
<i>SUF4</i>	8	Vernalization
<i>SVP/AGL22</i>	3	Vernalization
<i>TEM1</i>	6	Photoperiod
<i>TEM2</i>	7	Photoperiod
<i>TFL2</i>	3	Photoperiod
<i>TIC</i>	10	Photoperiod/Circadian clock
<i>TOE1</i>	3	Photoperiod

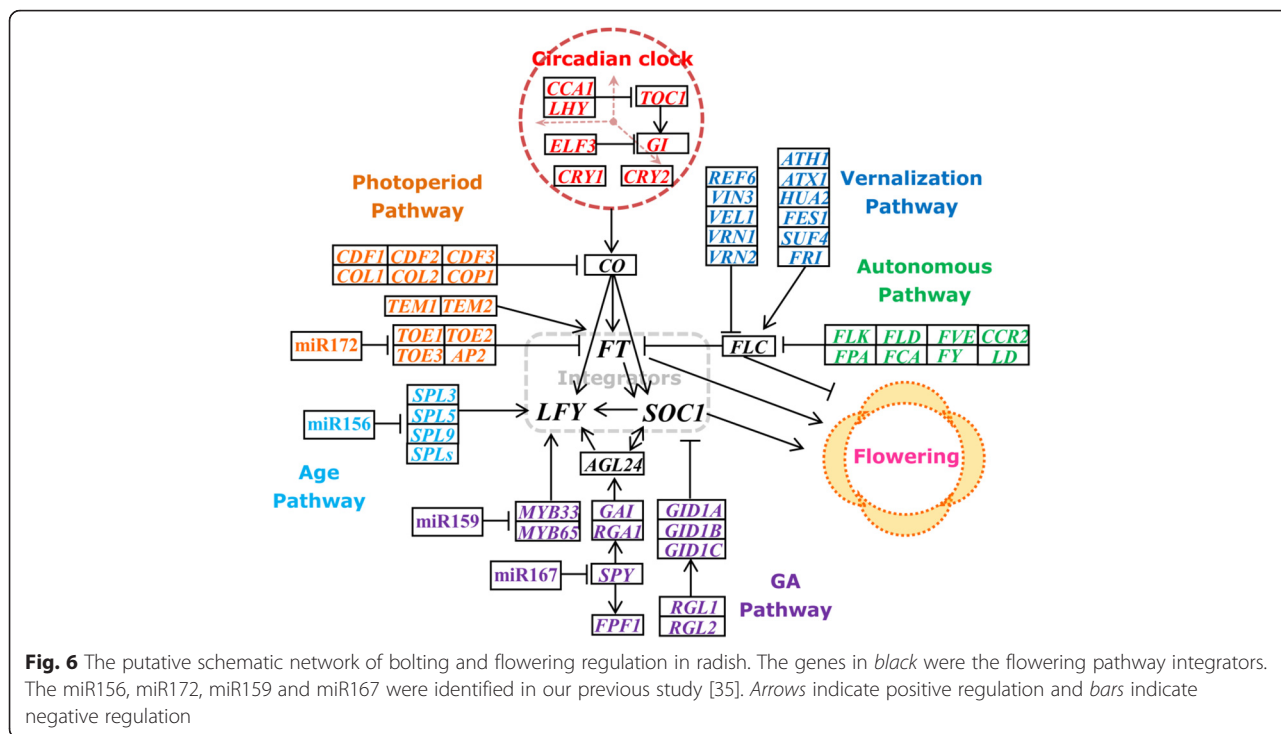
**Table 4** The identified candidate genes involved in various flowering pathways of radish (*Continued*)

<i>TOE2</i>	2	Photoperiod
<i>TOE3</i>	3	Photoperiod
<i>TPS1</i>	4	Others
<i>VEL1</i>	2	Photoperiod
<i>VIN3</i>	1	Vernalization
<i>VIP1</i>	2	Vernalization
<i>VIP2</i>	2	Vernalization
<i>VRN1</i>	9	Vernalization
<i>VRN2</i>	2	Vernalization
<i>WNK1</i>	3	Photoperiod/Circadian clock

requires its activator *FRIGIDA (FRI)*, a plant-specific gene conferring vernalization response in *Arabidopsis* [53]. In addition, genomic analysis suggested that many genes were involved in vernalization pathway and played roles in flowering control [2, 15]. As expected, the vernalization-response genes including *FRI*, *FRIGIDA ESSENTIAL 1 (FES1)*, *VERNALIZATION INSENSITIVE 3 (VIN3)*, *VERNALIZATION 1 (VRN1)* and *VRN2*, were also identified in this study (Fig. 6; Table 4), which was consistent with previous studies on flowering gene discovery [15, 24, 25, 51].

Additionally, we also found transcript sequences matching to the candidate genes of *SOC1* (CL4258.Contig1 and CL4258.Contig2), *AGL24* (CL2892.Contig1 and CL2892.Contig2) and *LFY* (Unigene29702, Unigene30559 and Unigene28363) (Additional file 4). Both *SOC1* and *AGL24* belonged to MADS-box genes, and their interaction directly regulated the expression of *LFY* and determined the flowering time [10, 49, 54, 55]. The family of MADS-box transcription factor is a major group of regulators controlling floral transition in *Arabidopsis* and act in floral organ specification and floral development in flowering plants [56–58]. Our studies also identified some members of MADS family such as *AGLs*, *API*, *AP2*, *SHORT VEGETATIVE PHASE (SVP)* and *FLOWERING LOCUS M (FLM)*, which might participate in the development and flowering regulation of radish (Fig. 6).

Moreover, radish transcripts which showed to be homologous genes in GA pathway included *FLOWERING PROMOTING FACTOR 1 (FPF1)*, *GIBBERELLIC ACID INSENSITIVE (GAI)*, *MYB33*, *MYB65* and *SPINDLY (SPY)* (Table 4). It was reported that *MYB33* and *MYB65* were the target genes of miR159 which could control the flowering time *via* the GA pathway [59, 60], while *SPY* targeted by miR167 was a negative regulator of gibberellin signaling and functioned in mediating flowering time [61]. Furthermore, some bolting and flowering-related critical miRNAs including miR156, miR172, miR159 and



miR167 were also identified in our previous study [35] (Fig. 6). These findings revealed that the transcriptome dataset of radish leaves could be used for the discovery and profiling of functional genes involved in various developmental pathways and processes in root vegetable crops.

**Isolation and validation of bolting and flowering-related genes**

Based on the annotation analysis and BLAST search of radish leaf transcriptomic sequences, the full-length sequences of functional genes related to flowering regulation were obtained by transcript sequence assembly. To evaluate the quality of assembled sequences and annotation results from the NGS sequencing, the sequences of several flowering genes from T-A cloning and transcript assembly were subjected to comparative analysis. The full-length cDNA sequences of seven randomly selected

bolting and flowering-related genes including *AGL24*, *AP2*, *PPF1*, *LUMINIDEPENDENS (LD)*, *VRN2*, *SOC1* and *SVP*, were isolated from radish late-bolting line ‘NAU-LU127’ using specific primers and sequenced by the Sanger methods (Table 5). Comparative analysis showed that the full length of these isolated gene sequences ranged from 365 bp to 2,504 bp, and all of them contained the complete open reading frame (ORF) (Table 5). These assembled transcripts had more than 92 % coverage to the corresponding full-length genes, with more than 91 % similarity of ORF. These results revealed that the transcript sequences from radish leaf transcriptome could be used for the identification and isolation of more functional genes related to bolting and flowering in radish.

Remarkably, the assembled transcript sequence of *SOC1* had 100 % coverage and 100 % ORF similarity to that from T-A cloning (Table 5). The role

**Table 5** Sequence analysis of the isolated genes related to bolting and flowering in radish

Gene name	Full length cDNA (bp)	Coverage (%)	ORF length (bp)	ORF similarity (%)
<i>RsAGL24</i>	828	99.28	666	99.85
<i>RsAP2</i>	1334	94.15	1302	94.01
<i>RsPPF1</i>	365	98.18	330	98.18
<i>RsLD</i>	2504	99.40	2448	99.58
<i>RsSOC1</i>	724	100	642	100
<i>RsSVP</i>	801	92.52	726	91.74
<i>RsVRN2</i>	1342	99.93	1281	99.92

and function of *SOC1* is mainly promoting flowering and floral meristem identity, and the photoperiod, vernalization and GA pathways converge on *SOC1* to control the flowering time [10]. Phylogenetic relationship between *RsSOC1* and the homologous genes in other plants was analyzed using MEGA 6.0 software (Fig. 7). The phylogenetic analysis indicated that the *SOC1* homologs were clustered into two subgroups, and *RsSOC1* shared the closest relationship with the *BrAGL20* from *B. rapa* and was clearly separated from *ZmMADS1*, a *ZmSOC1*-like gene from *Zea mays*.

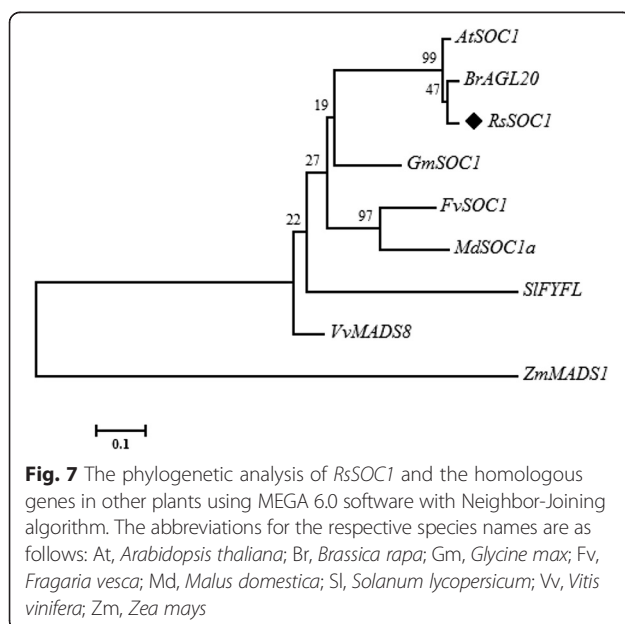
### Expression profiles by RT-qPCR

To investigate the spatial-temporal expression patterns of bolting and flowering-related genes, the relative expression levels of seven isolated candidate genes in various tissues and at different developmental stages were validated by RT-qPCR (Reverse transcription quantitative real-time PCR) analysis (Fig. 8). The expression patterns showed that seven selected genes related to bolting and flowering were differentially expressed in root, stem, leaf and flower of radish (Fig. 8a). Most of them showed a low expression in radish root, but the lowest expression levels of *RsAGL24*, *RsSOC1* and *RsSVP* were in the flower. The relative expression levels of *RsAGL24*, *RsSOC1*, *RsSVP* and *RsVRN2* were significantly higher in the leaf than in the other three tissues, whereas *RsAP2* was mainly expressed in the flower. The transcript levels of *RsFPF1* and *RsLD* were higher in leaf and flower, and no significant difference was detected in expression of these two tissues. The results indicate that these genes may play different roles in various tissues and thus affect the growth and development of radish. The expression

pattern differences of these genes were also observed in previous studies [2, 14].

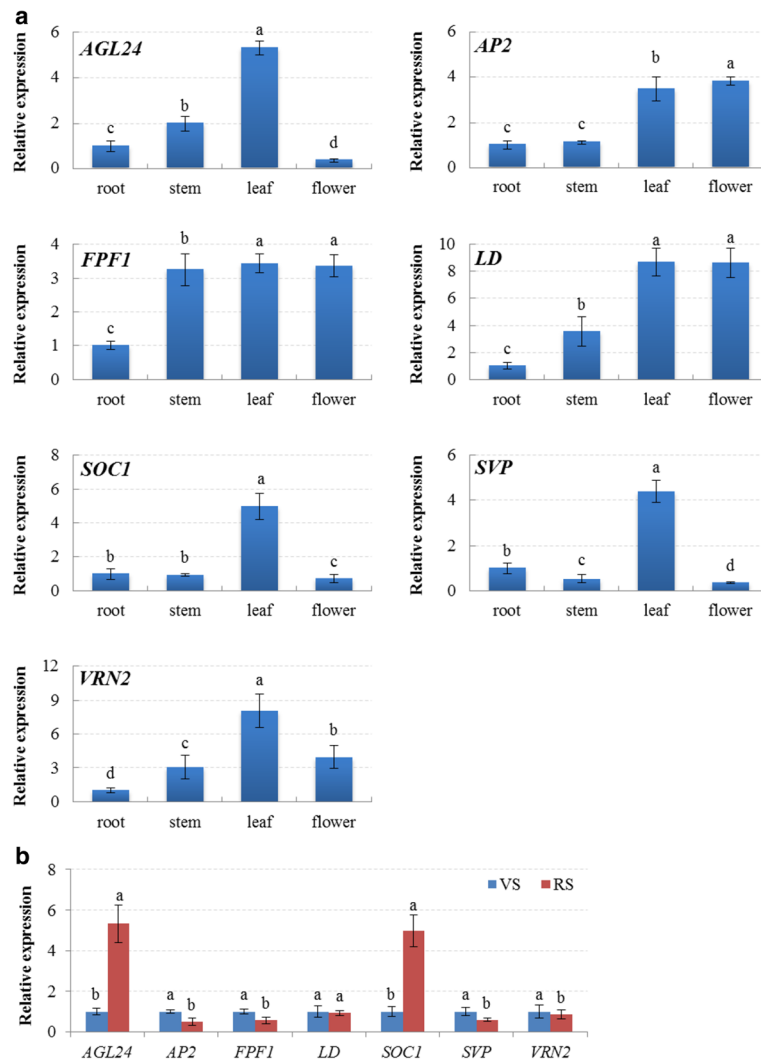
The expression profiling of these seven bolting and flowering-related genes also revealed that their expression patterns in radish leaves varied at vegetative stage and reproductive stage (Fig. 8b), which contributed to further explore their possible roles during radish bolting and flowering. The expression levels of *RsAGL24* and *RsSOC1* were lower at vegetative stage and increased at reproductive stage; whereas the expression levels of these four genes including *RsAP2*, *RsFPF1*, *RsSVP* and *RsVRN2*, were relative higher at vegetative stage and significantly decreased at reproductive stage. The up- or down-expression of these genes were in agreement with previous analysis in other species [24, 25, 51]. Flowering pathway analysis indicated that seven selected genes were implicated in different flowering pathways (Fig. 6). Among them, *AP2* belonging to photoperiod pathway is regulated by miR172 and represses the bolting and flowering [35]; while *SOC1* as a floral integrator negatively regulates *FLC* expression and plays vital roles in promoting flowering [10, 49]. These findings revealed that these differentially expressed genes might participate in the transition from vegetative stage to bolting and flowering in radish.

Additionally, four flower development-related genes including *RsAPI*, *RsSEP3*, *RsAGL24* and *RsLFY*, were selected for RT-qPCR analysis to explore their expression patterns during radish flower development and in different flower parts. The expression levels of *RsSEP3*, *RsAGL24* and *RsLFY* were relative higher at whole flower, while no significant difference of *RsAPI* expression was showed at three stages of flower development (Fig. 9a). The expression analysis revealed that four genes were differentially expressed in four flower parts (Fig. 9b), which was in consistent with previous studies [51, 56, 58], suggesting that these genes may play important roles in flower organ identity. Importantly, *API* belonging to A-class genes in the ABC model of flower organ identity, determines the development of sepal and petal [58, 62]. The results showed that *RsAPI* expression level was relative higher in sepal and petal than in stamen and pistil (Fig. 9b), suggesting that *API* gene is essential for sepal and petal development and specify the identity of the floral meristem [62].

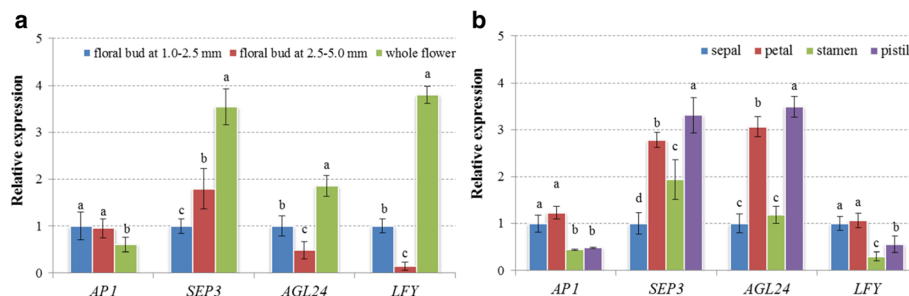


### Conclusions

To our knowledge, this is the first report on *de novo* transcriptome sequencing and assembly in late-bolting radish leaves using RNA-seq technology. A total of 53,642 unigenes were assembled in radish leaf transcriptome and 50,385 unique sequences were successfully annotated by GO, COG and KEGG databases. A total of



**Fig. 8** The spatial-temporal expression profiling of bolting and flowering-related genes by RT-qPCR analysis. **a** The expression patterns of seven genes in different radish tissues. **b** The expression patterns of seven genes at vegetative stage (VS) and reproductive stage (RS). Each bar shows the mean  $\pm$  SE of triplicate assays. The values with different letters indicate significant differences at  $P < 0.05$  according to Duncan's multiple range tests



**Fig. 9** The spatial-temporal expression profiling of flower development-related genes by RT-qPCR analysis. **a** The relative expression levels of four genes at three stages of flower development. **b** The relative expression levels of four genes at different flower parts (sepal, petal, stamen and pistil). Each bar shows the mean  $\pm$  SE of triplicate assays. The values with different letters indicate significant differences at  $P < 0.05$  according to Duncan's multiple range tests

24 candidate genes were involved in plant circadian rhythmic pathway. Moreover, 142 potential radish genes related to bolting and flowering-time regulation were identified by BLAST similarity search with *A. thaliana* flowering-related genes. RT-qPCR analysis revealed that several flowering-related genes showed spatial-temporal expression patterns in various radish tissues and at different development stages. The findings from this study provided insights into bolting and flowering-time regulatory networks in radish, and facilitated further dissecting molecular genetic mechanisms associated with bolting and flowering regulation in Brassicaceae vegetable crops.

## Methods

### Plant materials

The radish advanced inbred line 'NAU-LU127', late bolting and flowering, was used in this study. When cultivated during the autumn season, this late-bolting radish generally bolts until the middle and late of next April. The seeds were sown in plastic pots and cultured in plant growth chamber with 16 h light (25 °C) and 8 h dark (18 °C). The radish leaves were collected at vegetative stage and reproductive stage, respectively. For gene expression validation, the different radish tissues including root, stem, leaf, flower, floral buds at 1.0-2.5 mm and 2.5-5.0 mm, sepal, petal, stamen and pistil, were separately collected at reproductive stage and immediately frozen in liquid nitrogen and stored at -80 °C. All the tissues were harvested with three independent biological replicates.

### Library preparation and Illumina sequencing

Total RNA was isolated using Trizol reagents (Invitrogen, USA) according to the manufacturer's protocol. For cDNA library construction, a total of 20 µg RNA from radish leaf samples at two different stages were pooled in equimolar quantity. Briefly, the mRNA was isolated using magnetic beads with Oligo (dT), and fragmented to small pieces using fragmentation buffer (Ambion, USA). Then the mRNA fragments were used as templates to synthesize double-stranded cDNA with random hexamer primers using the SuperScript Double-Stranded cDNA Synthesis Kit (Invitrogen, USA). The synthesized cDNA were purified with QiaQuick PCR extraction kit and subjected to end reparation and single nucleotide A (adenine) addition. Thereafter, the short fragments were connected with adapters and the suitable fragments were screened as templates for PCR amplification. The transcriptome library was sequenced using Illumina HiSeq™ 2000 at Beijing Genomics Institute (Shenzhen, China).

### De novo assembly of radish leaf transcriptome

The raw reads were initially processed by Illumina sequencing and filtered to produce clean reads by removing adapter sequences, unknown or low quality reads.

Transcriptome *de novo* assembly was performed using the Trinity program [36] based on the de Bruijn graph algorithm, with k-mer of 25, minimum k-mer coverage of 1 and minimum contig length of 100. Default settings were used for all other parameters. Clean reads were firstly assembled to form longer fragments named contigs, and consequently the reads with overlapping regions were mapped back to the corresponding contigs. Based on the paired-end reads, the detection of different contigs from the same transcript sequences and the distance among them were detected and calculated, respectively. Then, these contigs were further assembled by the Trinity [36] and the obtained sequences were defined as unigenes with no extension on either end. According to the FPKM method (Fragments Per kb per Million fragments) previously described [37], the expression level of each unigene was calculated with the formula:  $FPKM = (10^6 \times C \times 10^3) / NL$ . Where C is the number of reads that uniquely aligned to certain unigene, N is the total number of reads that uniquely aligned to all unigenes, and L is the number of bases on this unigene.

### Functional annotation and classification

All the assembly unigenes were searched and annotated against the publicly available protein databases including NR, Swiss-Prot, KEGG and COG, using BLASTx analysis with an *E*-value cut-off of 1.0E-05. Based on the BLAST results, the CDS of unigenes were predicted for proteins by using the best alignments, which conducted following a priority order of NR, Swiss-Prot, KEGG and COG. If a unigene could not be aligned to these databases, ESTScan software was used to decide the sequence direction [63], including the nucleotide (5'-3') and amino acid sequences of the coding regions. Based on NR annotations, the Blast2GO program was employed to gain GO annotations of unique transcripts at the second level, according to biological processes, cellular components and molecular functions [64]. GO functional classification and distribution of gene functions of each assembly unigene were performed using WEGO software at the macroscopic level [65].

### Identification and isolation of bolting and flowering-related genes

According to the reported flowering regulatory network in *A. thaliana* [2, 3], the complete cDNA sequences of flowering-related genes were downloaded from the public NCBI database. The assembly transcript sequences from 'NAU-LB' were subjected to a local BLASTx similarity search at *E*-value  $\leq 1.0E-05$  against the *A. thaliana* gene sequences. The phylogenetic analysis was performed using the MEGA 6.0 software (<http://www.megasoftware.net/>) with the Neighbor-Joining algorithm. The sequences of



*SOC1* homologs used for phylogenetic tree construction were retrieved from the NCBI database. For T-A cloning, first-strand cDNA was synthesized from total RNA in radish leaves with the M-MLV (RNase H<sup>-</sup>) reverse transcriptase (TaKaRa, Dalian, China). Specific PCR primers were designed using Primer Premier 5.0 and shown in Additional file 6. The PCR amplification was performed according to previously reported method [32]. Purified PCR products were ligated into the pMD18-T vector (TaKaRa, Beijing, China), and then transformed into *E. coli* DH5 $\alpha$ . Positive clones were sequenced on an ABI3730 sequencer (Applied Biosystem, USA).

### RT-qPCR analysis

RT-qPCR analysis was performed to validate the relative expression levels of several flowering genes in diverse radish tissues and at different stages. Total RNA from various tissue samples was isolated and reverse transcribed into cDNA as described above following the manufacturer's instructions, respectively. The specific primers for RT-qPCR were designed using Beacon Designer 7.0 (Premier Biosoft International, USA). Three biological replicates were performed for RT-qPCR assay to ensure the reliability of quantitative analysis. The PCR reaction was conducted using an iCycler Real-Time PCR Detection System (Bio-Rad, USA) and the *RsActin* gene was used as the reference gene [66]. The relative fold expression changes were calculated using the  $2^{-\Delta\Delta C_T}$  method [67]. The error bars representing the standard deviation were derived from each sample in triplicate. The statistical analysis with SAS Version 9.0 software (SAS Institute, Cary, North Carolina, USA) was performed using Duncan's multiple range test at the  $P < 0.05$  level of significance. All primer pairs used in this study are listed in Additional file 6.

### Additional files

**Additional file 1:** Summary of all the assembled unigenes in radish leaf transcriptome. (XLSX 14582 kb)

**Additional file 2:** KEGG pathways of the assembled unigenes in radish. (XLSX 130 kb)

**Additional file 3:** Candidate genes involved in circadian rhythm of radish. (XLSX 107 kb)

**Additional file 4:** Identification of flowering related genes in radish. (XLSX 57 kb)

**Additional file 5:** Annotation of transcript sequences related to flowering genes in radish. (XLSX 139 kb)

**Additional file 6:** Specific primer sequences of flowering related genes for T-A cloning and RT-qPCR. (XLSX 11 kb)

### Abbreviations

CO: *CONSTANS*; COG: Clusters of Orthologous Groups; FLC: *FLOWERING LOCUS C*; FPKM: Fragments Per kb per million fragments; FT: *FLOWERING LOCUS T*; GO: Gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; LFY: *LEAFY*; NGS: Next-generation sequencing technology; ORF: Open reading

frame; RNA-seq: RNA sequencing; RT-qPCR: Reverse transcription quantitative real-time PCR; *SOC1*: *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1*.

### Acknowledgements

This work was partially supported by grants from the National Natural Science Foundation of China (31171956, 31372064), the National Key Technologies R & D Program of China (2012BAD02B01), the Key Technologies R & D Program of Jiangsu Province (BE2013429) and the PAPD.

### Availability of data and materials

The datasets supporting the conclusions of this article are included within the article and its additional files. The sequencing data have been deposited in NCBI SRA (Sequence Read Archive, <http://www.ncbi.nlm.nih.gov/sra/>) under the accession number SRX1671013.

### Authors' contributions

NS and LL conceived and designed the study. NS, LC and LL designed the experiments and wrote the manuscript. NS, LC, XL and SX were responsible for sequencing, cloning and expression analysis. WY, HD, EM and XY participated in the design of the study and performed the data analysis. All authors read and approved the final manuscript.

### Competing interests

The authors declare that they have no competing interests.

### Consent for publication

Not applicable.

### Ethics approval and consent to participate

Not applicable.

### Author details

<sup>1</sup>National Key Laboratory of Crop Genetics and Germplasm Enhancement, College of Horticulture, Nanjing Agricultural University, Nanjing 210095, People's Republic of China. <sup>2</sup>Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (East China) of the Ministry of Agriculture of P.R. China, Nanjing 210095, People's Republic of China. <sup>3</sup>Department of Plant Sciences, North Dakota State University, Fargo, ND 58108, USA.

Received: 26 August 2015 Accepted: 21 April 2016

Published online: 23 May 2016

### References

- Amasino RM, Michaels SD. The timing of flowering. *Plant Physiol.* 2010;154:516–20.
- Srikanth A, Schmid M. Regulation of flowering time: all roads lead to Rome. *Cell Mol Life Sci.* 2011;68:2013–37.
- Fornara F, de Montaigu A, Coupland G. SnapShot: Control of flowering in *Arabidopsis*. *Cell.* 2010;141:550.
- Capovilla G, Schmid M, Posé D. Control of flowering by ambient temperature. *J Exp Bot.* 2014; 66: 59–69.
- Pajoro A, Biewers S, Dougali E, Valentim FL, Mendes MA, Porri A, et al. The (r)evolution of gene regulatory networks controlling *Arabidopsis* plant reproduction; a two decades history. *J Exp Bot.* 2014;65:4731–45.
- Wang JW. Regulation of flowering time by the miR156-mediated age pathway. *J Exp Bot.* 2014;65:4723–30.
- Moon J, Lee H, Kim M, Lee I. Analysis of flowering pathway integrators in *Arabidopsis*. *Plant Cell Physiol.* 2005;46:292–9.
- Parcy M. Flowering: a time for integration. *Int J Dev Biol.* 2005;49:585–93.
- Michaels S, Amasino R. *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell.* 1999;11:949–56.
- Lee J, Lee I. Regulation and function of *SOC1*, a flowering pathway integrator. *J Exp Bot.* 2010;61:2247–54.
- Tsuji H, Taoka KI, Shimamoto K. Regulation of flowering in rice: two florigen genes, a complex gene network, and natural variation. *Curr Opin Plant Biol.* 2011;14:45–52.
- Shrestha R, Gómez-Ariza J, Brambilla V, Fornara F. Molecular control of seasonal flowering in rice, *Arabidopsis* and temperate cereals. *Ann Bot.* 2014;114:1445–58.

13. Dong Z, Danilevskaya O, Abadie T, Messina C, Coles N, Cooper M. A gene regulatory network model for floral transition of the shoot apex in maize and its dynamic modeling. *PLoS One*. 2012;7:e43450.
14. Mouhu K, Hytönen T, Folta K, Rantanen M, Paulin L, Auvinen P, et al. Identification of flowering genes in strawberry, a perennial SD plant. *BMC Plant Biol*. 2009;9:122.
15. Jung CH, Wong CE, Singh MB, Bhalla PL. Comparative genomic analysis of soybean flowering genes. *PLoS One*. 2012;7:e38250.
16. Metzker ML. Sequencing technologies-the next generation. *Nat Rev Genet*. 2010;11:31–46.
17. Ward JA, Ponnala L, Weber CA. Strategies for transcriptome analysis in non-model plants. *Am J Bot*. 2012;99:267–76.
18. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*. 2009;10:57–63.
19. Strickler SR, Bombarely A, Mueller LA. Designing a transcriptome next-generation sequencing project for a nonmodel plant species. *Am J Bot*. 2012;99:257–66.
20. Liu D, Sui S, Ma J, Li Z, Guo Y, Luo D, et al. Transcriptomic analysis of flower development in wintersweet (*Chimonanthus praecox*). *PLoS One*. 2014;9:e86976.
21. Liang C, Liu X, Yiu SM, Lim BL. *De novo* assembly and characterization of *Camelina sativa* transcriptome by paired-end sequencing. *BMC Genomics*. 2013;14:146.
22. Jia X, Wang GL, Xiong F, Yu XR, Xu ZS, Wang F, et al. *De novo* assembly, transcriptome characterization, lignin accumulation, and anatomic characteristics: novel insights into lignin biosynthesis during celery leaf development. *Sci Rep*. 2015;5:8259.
23. Guzman F, Kulcheski FR, Turchetto-Zolet AC, Margis R. *De novo* assembly of *Eugenia uniflora* L. transcriptome and identification of genes from the terpenoid biosynthesis pathway. *Plant Sci*. 2014;229:238–46.
24. Zhang XM, Zhao L, Larson-Rabin Z, Li DZ, Guo ZH. *De novo* sequencing and characterization of the floral transcriptome of *Dendrocalamus latiflorus* (Poaceae: Bambusoideae). *PLoS One*. 2012;7:e42082.
25. Gao J, Zhang Y, Zhang CL, Qi FY, Li XP, Mu SH, et al. Characterization of the floral transcriptome of moso bamboo (*Phyllostachys edulis*) at different flowering developmental stages by transcriptome sequencing and RNA-seq analysis. *PLoS One*. 2014;9:e98910.
26. Ness RW, Siol M, Barrett SCH. *De novo* sequence assembly and characterization of the floral transcriptome in cross- and self-fertilizing plants. *BMC Genomics*. 2011;12:298.
27. Zhang ZY, Wang P, Li Y, Ma LL, Li LF, Yang RT, et al. Global transcriptome analysis and identification of the flowering regulatory genes expressed in leaves of *Lagerstroemia indica*. *DNA Cell Biol*. 2014;33:680–8.
28. Wei CH, Tao X, Li M, He B, Yan L, Tan XM, et al. *De novo* transcriptome assembly of *Ipomoea nil* using Illumina sequencing for gene discovery and SSR marker identification. *Mol Genet Genomics*. 2015;290:1873–84.
29. Tao X, Gu YH, Jiang YS, Zhang YZ, Wang HY. Transcriptome analysis to identify putative floral-specific genes and flowering regulatory-related genes of sweet potato. *Biosci Biotech Bioch*. 2013;77:2169–74.
30. Kitashiba H, Li F, Hirakawa H, Kawanabe T, Zou ZW, Hasegawa Y, et al. Draft sequences of the radish (*Raphanus sativus* L.) genome. *DNA Res*. 2014;21:481–90.
31. Zhang L, Jia H, Yin Y, Wu G, Xia H, Wang X, et al. Transcriptome analysis of leaf tissue of *Raphanus sativus* by RNA sequencing. *PLoS One*. 2013;8:e80350.
32. Wang Y, Pan Y, Liu Z, Zhu XW, Zhai LL, Xu L, et al. *De novo* transcriptome sequencing of radish (*Raphanus sativus* L.) and analysis of major genes involved in glucosinolate metabolism. *BMC Genomics*. 2013;14:836.
33. Curtis I, Nam H, Yun J, Seo K. Expression of an antisense *GIGANTEA* (*Gl*) gene fragment in transgenic radish causes delayed bolting and flowering. *Transgenic Res*. 2002;11:249–56.
34. Yi G, Park H, Kim J, Chae W, Park S, Huh J. Identification of three *FLOWERING LOCUS C* genes responsible for vernalization response in radish (*Raphanus sativus* L.). *Hort Environ Biotechnol*. 2014;55:548–56.
35. Nie S, Xu L, Wang Y, Huang D, Muleke E, Sun X, et al. Identification of bolting-related microRNAs and their targets reveals complex miRNA-mediated flowering-time regulatory networks in radish (*Raphanus sativus* L.). *Sci Rep*. 2015;5:14034.
36. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 2011;29:644–52.
37. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*. 2008;5:621–8.
38. Izzah NK, Lee J, Jayakodi M, Perumal S, Jin M, Park BS, et al. Transcriptome sequencing of two parental lines of cabbage (*Brassica oleracea* L. var. *capitata* L.) and construction of an EST-based genetic map. *BMC Genomics*. 2014;15:149.
39. Yan XH, Dong CH, Yu JY, Liu WH, Jiang CH, Liu JH, et al. Transcriptome profile analysis of young floral buds of fertile and sterile plants from the self-pollinated offspring of the hybrid between novel restorer line NR1 and *Nsa* CMS line in *Brassica napus*. *BMC Genomics*. 2013;14:26.
40. Fankhauser C, Staiger D. Photoreceptors in *Arabidopsis thaliana*: light perception, signal transduction and entrainment of the endogenous clock. *Planta*. 2002;216:1–16.
41. Johansson M, Staiger D. Time to flower: interplay between photoperiod and the circadian clock. *J Exp Bot*. 2015;66:719–30.
42. McClung CR. Plant circadian rhythms. *Plant Cell*. 2006;18:792–803.
43. Suarez-Lopez P, Wheatley K, Robson F, Onouchi H, Valverde F, Coupland G. CONSTANS mediates between the circadian clock and the control of flowering in *Arabidopsis*. *Nature*. 2001;410:1116–20.
44. Imaizumi T, Kay SA. Photoperiodic control of flowering: not only by coincidence. *Trends Plant Sci*. 2006;11:550–8.
45. Khan S, Rowe SC, Harmon FG. Coordination of the maize transcriptome by a conserved circadian clock. *BMC Plant Biol*. 2010;10:126.
46. Valverde F, Mouradov A, Soppe W, Ravenscroft D, Samach A, Coupland G. Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science*. 2004;303:1003–6.
47. Kikuchi R, Handa H. Photoperiodic control of flowering in barley. *Breeding Sci*. 2009;59:546–52.
48. Helliwell CA, Wood CC, Robertson M, James PW, Dennis ES. The *Arabidopsis* FLC protein interacts directly in vivo with SOC1 and FT chromatin and is part of a high-molecular-weight protein complex. *Plant J*. 2006;46:183–92.
49. Seo E, Lee H, Jeon J, Park H, Kim J, Noh YS, et al. Crosstalk between cold response and flowering in *Arabidopsis* is mediated through the flowering-time gene *SOC1* and its upstream negative regulator *FLC*. *Plant Cell*. 2009;21:3185–97.
50. Xiao D, Zhao JJ, Hou XL, Basnet RK, Carpio D, Zhang NW, et al. The *Brassica rapa* FLC homologue *FLC2* is a key regulator of flowering time, identified through transcriptional co-expression networks. *J Exp Bot*. 2013;64:4503–16.
51. Zhang JZ, Ai XY, Sun LM, Zhang DL, Guo WW, Deng XX, et al. Transcriptome profile analysis of flowering molecular processes of early flowering trifoliolate orange mutant and the wild-type [*Poncirus trifoliolate* (L.) Raf.] by massively parallel signature sequencing. *BMC Genomics*. 2011;12:63.
52. Kim DH, Doyle MR, Sung S, Amasino RM. Vernalization: winter and the timing of flowering in plants. *Annu Rev Cell Dev Biol*. 2009;25:277–99.
53. Choi K, Kim J, Hwang HJ, Kim S, Park C, Kim SY, et al. The FRIGIDA complex activates transcription of *FLC*, a strong flowering repressor in *Arabidopsis*, by recruiting chromatin modification factors. *Plant Cell*. 2011;23:289–303.
54. Lee J, Oh M, Park H, Lee I. SOC1 translocated to the nucleus by interaction with AGL24 directly regulates *LEAFY*. *Plant J*. 2008;55:832–43.
55. Liu C, Chen H, Er HL, Soo HM, Kumar PP, Han JH, et al. Direct interaction of AGL24 and SOC1 integrates flowering signals in *Arabidopsis*. *Development*. 2008;135:1481–91.
56. Theißen G. Development of floral organ identity, stories from the MADS house. *Curr Opin Plant Biol*. 2001;4:75–85.
57. Becker A, Theißen G. The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Mol Phylogenet Evol*. 2003;29:464–89.
58. Smaczniak C, Immink RG, Angenent GC, Kaufmann K. Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies. *Development*. 2012;139:3081–98.
59. Yamaguchi A, Abe M. Regulation of reproductive development by non-coding RNA in *Arabidopsis*: to flower or not to flower. *J Plant Res*. 2012;125:693–704.
60. Spanudakis E, Jackson S. The role of microRNAs in the control of flowering time. *J Exp Bot*. 2014;65:365–80.
61. Tseng TS, Salomé PA, McClung CR, Olszewski NE. SPINDLY and GIGANTEA interact and act in *Arabidopsis thaliana* pathways involved in light responses, flowering, and rhythms in cotyledon movements. *Plant Cell*. 2004;16:1550–63.

62. Mandel M, Gustafson-Brown C, Savidge B, Yanofsky M. Molecular characterization of the *Arabidopsis* floral homeotic gene *APETALA1*. *Nature*. 1992;360:273–7.
63. Iseli C, Jongeneel CV, Bucher P. ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *Proc Int Conf Intell Syst Mol Biol*. 1999;99:138–48.
64. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 2005;21:3674–6.
65. Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, et al. WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res*. 2006;34 Suppl 2:293–7.
66. Xu Y, Zhu X, Gong Y, Xu L, Wang Y, Liu LW. Evaluation of reference genes for gene expression studies in radish (*Raphanus sativus* L.) using quantitative real-time PCR. *Biochem Biophys Res Commun*. 2012;424:398–403.
67. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta C_T}$  method. *Methods*. 2001;25:402–8.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

