

Identification of QTLs and Putative Candidate Gene Involved in Rhizome Enlargement of Asian lotus (*Nelumbo Nucifera*)

Yanling Liu

Chinese Academy of Sciences Wuhan Botanical Garden

Heyun Song

Chinese Academy of Sciences Wuhan Botanical Garden

Minghua Zhang

Chinese Academy of Sciences Wuhan Botanical Garden

Dong Yang

Chinese Academy of Sciences Wuhan Botanical Garden

Xianbao Deng

Chinese Academy of Sciences Wuhan Botanical Garden

Heng Sun

Chinese Academy of Sciences Wuhan Botanical Garden

Juan Liu

Chinese Academy of Sciences Wuhan Botanical Garden

Mei Yang (✉ yangmei815815@wbgcas.cn)

Chinese Academy of Sciences <https://orcid.org/0000-0002-3629-6371>

Research Article

Keywords: Lotus (*Nelumbo nucifera*), rhizome enlargement, linkage map, quantitative trait loci, NnBEL6

Posted Date: January 6th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-937813/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Lotus (*Nelumbo*) is perennial aquatic plant with nutritional, pharmacological, and ornamental significance. Rhizome is an underground lotus stem that act as a storage organ and as a reproductive tissue for asexual production. The enlargement of lotus rhizome is an important adaptive strategy for surviving the cold winter. The aims of this study were to identify quantitative trait loci (QTLs) for rhizome enlargement traits including rhizome enlargement index (REI) and number of enlarged rhizome (NER), and to uncover candidate genes associated with these phenotypic traits. A high-density genetic linkage map was constructed, consisting of 2,935 markers binned from 236,840 SNPs. A total of 14 significant QTLs were detected for REI and NER, which explained 6.67–22.28% of trait variance. Three QTL regions were repeatedly identified in at least two years, and a major QTL, designated cqREI-LG2, with a rhizome-enlargement effect and about 20% of the phenotypic contribution was identified across the three climatic years. A candidate *NnBEL6* gene located within the confidence interval of cqREI-LG2 was considered to be putatively involved in lotus rhizome enlargement. The expression of *NnBEL6* was exclusively induced by rhizome swelling. Sequence comparison of *NnBEL6* among lotus varieties revealed a functional Indel site in its promoter that likely initiates the rhizome enlargement process. Transgenic potato assay was used to confirm the role of *NnBEL6* in inducing tuberization. The successful identification QTLs and functional validation of *NnBEL6* gene reported in this study will enrich our knowledge on the genetic basis of rhizome enlargement in lotus.

Key Message

QTL mapping studies identified the two reliable QTLs of rhizome enlargement in lotus. *NnBEL6* located within the confidence interval of the major QTL cqREI-LG2 is a key candidate gene enhancing rhizome enlargement.

Introduction

Lotus (*Nelumbo Adans.*) is a basal eudicot that belongs to the family Nelumbonaceae, that comprises two extant species including *N. nucifera* Gaertn. and *N. lutea* (Willd.) Pers. *N. nucifera* is distributed throughout Asia and Northern Australia, whereas *N. lutea* is found in North America and the northern areas of South America (Yang et al. 2012; Zhang et al. 2011). Asian lotus have been cultivated for over 7,000 years for its ornamental, edible and medicinal properties in Asian countries, especially in China, and are classified into three categories: flower, seed, and rhizomes lotus, according to their use (Ming et al. 2013; Shen-Miller 2002).

Lotus rhizome is a storage organ derived from modified subterraneous stem, and acts not only as an important nutritional vegetable, but also is reproductive tissue for lotus asexual production (Man et al. 2012). Therefore, the rhizome development has attracted increasing attention from researcher. Rhizome is generated through a complex developmental process that begins with the formation of a longitudinal underground stolon, which later swells to enlarge the girth. This process can be classified into four

stages, including stolon stage, initial swelling, middle swelling, and late stage of expansion. Important carbohydrates, such as starch is rapidly synthesized and accumulated in the rhizome during the last two stages of development (Cheng et al. 2013; Yang et al. 2015). Rhizome girth enlargement occurs under short day (SD), and cool temperature promotes dry weight (starch) accumulation at the late stage of development (Anderson 1970; Masuda et al. 2007a; Masuda et al. 2007b). Rhizome transition from longitudinal elongation to the girth enlargement is regulated via the biosynthesis of GA and/or ABA (Masuda et al. 2010; Masuda et al. 2012). The transcriptomics studies reveal that the genes related to phytohormone, photoperiod, and starch synthesis are involved in this transition (Cheng et al. 2013; Yang et al. 2015). The light and auxin signal can be transduced through Ca^{2+} , which serve as a secondary messenger (Cao et al. 2019). To date, there is still insufficient clarification on the complex process of rhizome enlargement in lotus.

The enlargement of lotus rhizome ensures winter survival and sprouting in the incoming spring season. Temperate and tropical lotus are the two ecotypes of *N. nucifera*, and each has different rhizome development process and adaptability to cold conditions (Zhang and Wang 2006). The underground stems of temperate ecotype show longitudinal growth in the early stage with increased girth at the late stages, and undergo all four developmental stages mentioned above. Tropical lotus ecotypes have no clear girth enlargement in their entire growth period, and are unable to survive the cold winter in subtropical and temperate regions like China (Yang et al. 2014; Yang et al. 2015). Thus, other than its role as a vital asexual production tissue, the enlargement of lotus rhizome acts as an adaptive strategy for surviving the cold winter. Development of storage organs have been extensively studied in potato (*Solanum tuberosum*) and onion (*Allium cepa*) (Abelenda et al. 2014; Lee et al. 2013; Navarro et al. 2015). Tuberization in potato is regulated by environmental factors and a series of genes. It normally occurs under SDs. *StSP6A* is an *Flowering lotus T(FT) like* protein that functions as tuberigen, and it is activated in the leaves only under SDs (Navarro et al. 2011). The coding protein is transported through the phloem to the stolon, where it interacts with St14-3-3s and StFDL1 to promote tuberization (Teo et al. 2017). *StBEL5* is a BEL1-like homeobox gene, and its RNA functions as a phloem-mobile RNA signal that activates the expression of *StSP6A* in the stolon (Sharma et al. 2016). In onion, bulb formation is also regulated by the antagonistic effect of two *FT-like* genes, *AcFT1* and *AcFT4*. Bulb formation is promoted by *AcFT1*, while *AcFT4* prevents the upregulation of *AcFT1*, thus inhibiting bulbing. Long-day photoperiod promotes bulbing by upregulating and downregulating the expression levels of *AcFT1* and *AcFT4*, respectively. It is predicted that *AcFT1* interact with 14-3-3, but *AcFT4* completely disrupt the interface with 14-3-3 (Lee et al. 2013; Lyngkhai et al. 2019; Manoharan et al. 2016). These studies provide important cues for studying rhizome formation in lotus.

Lotus germplasm show notable variation in the phenotypes of storage rhizome, including different branching, elongation and expansion (Lin et al. 2019), which provides vital materials for studying the genetic mechanisms for rhizome enlargement. Quantitative trait locus (QTL) analysis is a powerful genetic approach for dissecting the complex process/trait and investigating the genetic variation within a species. Recently, QTLs of rhizome yield-related traits were detected, and candidate genes were predicted

in lotus (Huang et al. 2021). In order to understand the genetic control of rhizome enlargement in lotus, the two lotus ecotypes were used in this study to generate an F₂ population and to construct an SNP-based genetic linkage map. QTLs mapping for rhizome enlargement were analyzed using a three-year phenotype data, and stable QTLs across the years were identified. Furthermore, a candidate *NnBEL6* gene that putatively control rhizome enlargement was predicted, and functionally validated in transgenic potato. Our results elucidate the genetic basis of rhizome formation and provide a foundation for future breeding of new cultivars by marker-assisted selection.

Materials And Methods

Plant materials and phenotyping

An F₂ population consisting of 138 individuals was used to construct genetic linkage map and to detect QTL. Two parents, 'Baige' ('BG', female parent, temperate lotus) and 'Winter Red 1' ('WR1', male parent, tropical lotus) with different rhizome girth enlargements were used (Fig. 1). The mapping population was separately planted in a trial plot at Wuhan Botanical Garden of the Chinese Academy of Sciences (N30°32'44.02", E114°24'52.18"), Hubei Province, China, in April, between 2014 and 2016. Each individual was planted in a separate 60-cm diameter and 45-cm depth pots in order to avoid mixing of the rhizomes. Paddy soil was added in each pot to a 20-cm-depth, and 200 g organic fertilizer (mature soybean seedcake) was applied before planting. While 20 g chemical fertilizer was top dressed at the early developing stage of standing leaf, the onset of flowering, full-blooming, and the late stage of flowering. The depth of water in the pot was kept at a 20–25 cm throughout the whole growth season.

The three phenotypic traits including number of enlarged rhizome (NER), maximum diameter of enlarged internode (MDE) and length of enlarged internode (LER) were investigated in each accession in three biological replicates. Rhizome enlargement Index (REI) was calculated as maximum diameter / length of enlarged internode according to Masuda et al. (2007b). Statistical analyses of the traits were conducted using ANOVA procedure in the SAS 8.1 system (SAS Institute, Cary, NC, USA).

Resequencing analysis and genotyping

Young leaves of the mapping population were collected for DNA extraction using the cetyltrimethylammonium bromide (CTAB) method described by Doyle and Doyle (1990), and slightly modified with the addition of RNase A and proteinase K treatment to prevent RNA and protein contamination. DNA concentration was checked using a NanoDrop 2000 (Thermo) spectrophotometry and quality was analyzed with electrophoresis on an 0.8% agarose gels with a lambda DNA standard, then stored at – 20°C until use.

Genomic DNA was fragmented to 300-500bp by sonication for library construct and for Illumina resequencing following the manufacturer's standard instructions. Quantification and quality assessment of library were carried out on an Agilent DNA 1000 LabChip analyzer (Agilent Technology 2100 Bioanalyzer). Paired-end sequencing was conducted on an Illumina HiSeq 2000 platform. Low quality (Q

score < 20), contaminant sequences were filtered using the NGS QC Toolkit (<http://www.nipgr.res.in/ngsqctoolkit.html>). The clean data was then mapped to 'China Antique' masked genome using the BWA program. Only sequences with less than two mismatches were aligned to the reference genome. The SAM output file was converted to BAM format and sorted using the SAM tools suite. These high-quality reads were used for SNP calling using Perl scripts (<http://www.genomics.ceh.ac.uk/msatfinder>). SNPs of low quality or with clustered SNP sites were removed from the genotype data, and a missing value was used to replace genotypes occupying lower reliability (< 20).

All SNPs markers in each individual were scored as aaxbb, nnxnp, lmxll, and hkxhk. For aaxbb marker, the genotype of homozygous individual with similar genotype to the female or male parent was labeled as "a" or "b", the genotype of the heterozygous individual was labeled as "h". For nnxnp marker, the genotype of homozygous individual with similar genotype to female parent was labeled as "nn", while the genotype of heterozygous individual was labeled as "np". For lmxll marker, the genotype of homozygous individual with similar genotype to the male parent was labeled as "ll", while the genotype of heterozygous individual was labeled as "lm". For hkxhk marker, the genotype of individual 1, 2 and 3 were labeled as "hh", "kk" and "hk" respectively.

Construction of genetic linkage map

Firstly, all SNPs with > 70% missing data (more than 95 F₂ individuals) and with allelic distortion (the ratio between minor allele and major allele frequency > 0.2) were discarded. Secondly, missing data was imputed in the filtered SNPs to create bins using an in-house Perl script. Four steps were included in these scripts: (1) Based on the position of SNP markers on reference genome, the SNPs in every 30 kb were assigned into one bin using the majority rule; (2) the missing genotypes were impute using the majority rule; (3) the mis-scored genotypes were corrected with the following criteria: when the marker was different from the adjacent ten markers (five markers above or below) in the same genotype, it was considered as a mis-scored genotype, and then corrected to the same genotype; (4) the markers with 100% identical genotype across all individuals were merged to one bin. Thirdly, the bins of aaxbb, lmxll and nnxnp were merged to construct genetic linkage map using JoinMap 4.0 software (VanOoijen 2006). The regression mapping algorithm was used for map construction. The goodness of fit threshold was set to ≤ 5.0 with logarithm of the odds ratio (LOD) scores > 1.0 and a combination frequency < 0.4. Recombination frequency was converted to linkage distances in centimorgan (cM) using the Kosambi mapping function (VanOoijen 2006). Bins on the linkage map were aligned to the reference genome of lotus 'China Antique' variety, based on their position (Ming et al. 2013).

QTL analysis and candidate gene prediction

QTLs were detected by the composite interval method (CIM) using WinQTL Cartographer 2.5 software (Wang et al. 2012). CIM was used to scan the genetic map and estimate the likelihood of a QTL and its corresponding effect at every 2 cM interval. For each trait, the threshold for the detection of a significant

QTL ($P < 0.05$) was estimated by 1,000 permutations. The coincidental significant QTL for the same traits from the three climatic years were integrated into consensus QTL using the BioMercator2.1 software (Arcade et al. 2004). For nomenclature of significant QTL, “q” was used to denote QTL, followed by an abbreviation of the trait (e.g., NER), then the three climatic years (2014, 2015 and 2016), the linkage group number (LG1–LG8), and finally arbitrary letter codes (a, b, c...). For nomenclature of consensus QTL, “cq” was used to denote consensus QTL, followed by an abbreviation of the trait, and then the linkage group number, finally arbitrary lower-case letter code.

Based on the alignment between the linkage map and the reference genome of lotus ‘China Antique’ variety, the genes in the interval of consensus QTLs were identified by referring to the lotus genome annotation (Ming et al. 2013). The expression patterns of identified genes were then analyzed from transcriptomic profile data (Yang et al. 2014).

Molecular cloning and sequence analysis of *NnBEL6*

To isolate full-length complementary DNA (cDNA) of *NnBEL6*, the youngest rhizome at the initial swelling stage was sampled from ‘BG’, ‘WR1’ and a ‘ZO’ lotus variety with REI of 0.46 (Fig. 1). Total RNA was prepared using TRIzol® reagent (Cat. no. 15596-026, Invitrogen) according to the manufacturer’s instructions. First-strand cDNA was synthesized from 1 µg total RNA and used as a template for reverse transcriptase (RT)-PCR. The coding regions of the candidate genes were amplified from the cDNA using Pfu polymerase (TakaRa). The PCR products were purified and sequenced directly. The pair of primers used to amplify cDNA are listed in Table S1.

The genomic regions of up to 8,485bp from 1.5 kb upstream to the stop codon of *NnBEL6* were sequenced in the two parental genotypes and 23 other lotus varieties which were phenotypically divided into three classes based on REI as, class I, with $REI < 0.1$; class II, with $0.1 < REI < 0.4$; and class III, with $REI > 0.4$. Of the tested lotus accessions, seven had class I rhizome, while class II rhizome and class III rhizome each had nine lotus members. The sequence comparison from these lotus accessions was performed in Bioedit software (<http://www.mbio.ncsu.edu/bioedit>).

Expression analysis of *NnBEL6*

The two varieties, ‘WR1’ and ‘ZO’, were used to analyze the spatial and temporal expression of *NnBEL6*. The leaf, petiole, flower-bud, pedicel, rhizome and root samples were collected from three different, randomly selected plants at the stolon and initial swelling stages, and immediately frozen in liquid nitrogen and stored at -80°C prior to RNA extraction. Total RNA from each tissue sample was extracted for quantitative RT-PCR using TB Green Premix (www.takara-bio.com) on a Step One Plus Real-time PCR System (Applied Biosystems), following the manufacturers’ instructions. Specific primer pairs sequences for *NnBEL6*, and the control actin gene used are listed in Table S1. All reactions were carried out in three independent biological replicates. The comparative critical threshold ($2^{-\Delta\Delta\text{CT}}$) method was used to analyze the relative *NnBEL6* gene expression level (Livak and Schmittgen 2001).

Potato transformation

The cDNA sequence of *NnBEL6* gene from 'ZO' was first cloned into the pGEM-T Easy vector (Promega). KpnI/XbaI-digested fragments were then inserted in the intron-GUS of pCambia-2300-35S-EGFP vector, to ensure the expression of transgene is driven by the cauliflower mosaic virus CaMV35S promoter. These constructs were introduced into *Agrobacterium tumefaciens* strain Gv3101, which was then transformed into the 'E3' line, as previously described (Zhou et al. 2019). PCR with primer specific to the 35S promoter was carried out to confirm the presence of the transgene (Table S1).

For *in vitro* tuberization test, transgenic plants overexpressing *NnBEL6* were cultured using single stem nodes on MS medium (Murashige and Skoog 1962) supplemented with 8% sucrose at 22°C, under a photoperiod of 8 h of light/16 h of dark (SD), with the light intensity ranging from 800 to 1000 $\mu\text{mol m}^{-2} \text{sec}^{-1}$. The experiment was carried out in five plants, each with three replicates. Microtubers were harvested approximately 4 months after transplantation. To detect the endogenous expression levels of *StSP6A* and *StSP5G* in WT (wild type) 'E3' controls and transgenic lines, leaves were harvested from 3-week-old plantlets for qRT-PCR. The potato actin gene (GenBank accession number: XM006350963) was used as internal control.

Results

Sequencing and mining of potential SNPs in lotus

A total of 138 F_2 individuals were developed from a cross between 'BG' and 'WR1'. Two parents were sequenced at a relatively high coverage using WGS. In total, 10.8 Gb (11.6-fold genome coverage) and 9.9 Gb (10.7-fold genome coverage) high-quality short-read sequences were generated for 'BG' and 'WR1', respectively (Table S2). All high-quality reads from 'BG' and 'WR1' were aligned to the repeat masked reference genome of 'China Antique' to detect putative SNPs. As results, 70.5 % and 85.4 % of the reads ($Q > 20$) from 'BG' and 'WR1' were aligned to the reference genome. A total of 949,840 and 900,010 putative SNPs were identified in 'BG' and 'WR1', respectively (Table 1). Among them, 296,706 SNPs were polymorphic between the two parents. For the polymorphic SNPs, 244,926 and 44,802 were heterozygous for 'BG' and 'WR1', respectively, indicating a high rate of heterozygosity in the maternal variety. Polymorphic SNPs markers were grouped as, aaxbb, lmxll, nnxnp, and hxxhk, based on parental genotypes, each with 27,738, 224,166, 24,042, and 20,760 markers, respectively (Table 1).

Table 1
Number of SNP detected in the parents and F₂ individuals

| Marker type | Number of SNP in 'BG' | Number of SNP in 'WR1' | Number of SNP in F ₂ | Number of bin in F ₂ |
|--|-----------------------|------------------------|---------------------------------|---------------------------------|
| Total | 949840 | 900010 | | |
| Missing marker | 155924 | 205754 | | |
| Non polymorphic marker between the parents | 447380 | | | |
| Polymorphic marker between the parents | 296706 | | 285895 | 4031 |
| aaxbb | 27738 | | 27527 | 1069 |
| lmxll | 224166 | | 214643 | 1933 |
| nnxnp | 24042 | | 23811 | 638 |
| hkxhk | 20760 | | 19914 | 391 |
| Homozygous SNP | 51780 | 251904 | 27527 | |
| Heterozygous SNP | 244926 | 44802 | 258368 | |

An overall genome sequence of 482.6 Gb was generated in the 138 F₂ individuals, with an average of 5.3 Gb for each individual, which was equivalent to approximately 5.7× coverage of the reference genome (Table S2). As a result, 64.2% of high-quality reads (Q > 20) of the F₂ individuals was mapped on the reference genome. Based on the parental polymorphic SNPs, a total of 285,895 SNPs were identified, with 27,527, 214,643, 23,811, and 19,914 in aaxbb, lmxll, nnxnp, and hkxhk, respectively. Moreover, these SNPs were binned as 1,069, 1,933, 638, and 391 markers in aaxbb, lmxll, nnxnp, and hkxhk, respectively (Table 1).

Construction of linkage map and comparison with the draft genome

Due to high heterozygosity of both parental genomes, most loci in the F₂ progenies were heterozygous. The hkxhk marker could not be distinguished from the parents, thus it was not used for genetic linkage map construction. As a result, the bins from aaxbb, lmxll, and nnxnp SNP were merged to construct genetic linkage map (Tables 2 and S3). This genetic map contained 2,935 bins generated from 236,840 SNPs. These bins were distributed on eight genetic groups spanning 896.11 cM with an average marker interval of 0.27 cM. Marker distribution and linkage length varied among linkage groups. LG1 was the largest with 170.99 cM, and contained 617 markers. The shortest group LG8 was 33.30 cM, and contained 114 markers. The highest marker density was on LG6, with an average marker interval of 0.21 cM. The lowest marker density was on LG4, with an average marker interval of 0.47 cM. The largest gap

among all linkage groups was 9.63 cM on LG7. The information of all markers including marker name, linkage group, genetic distance, and marker type are listed in Table S3.

Table 2
Summary of genetic linkage map

| Linkage group | No. of SNP | No. of bin | Length (cM) | Average distance (cM) | Max. gap (cM) |
|---------------|------------|------------|-------------|-----------------------|---------------|
| LG1 | 55633 | 617 | 170.99 | 0.28 | 7.24 |
| LG2 | 44251 | 567 | 161.53 | 0.28 | 9.55 |
| LG3 | 28368 | 292 | 120.38 | 0.41 | 5.97 |
| LG4 | 12997 | 247 | 117.07 | 0.47 | 5.89 |
| LG5 | 35287 | 356 | 112.08 | 0.31 | 8.63 |
| LG6 | 27924 | 520 | 106.64 | 0.21 | 7.10 |
| LG7 | 14917 | 222 | 74.12 | 0.33 | 9.63 |
| LG8 | 17463 | 114 | 33.30 | 0.30 | 4.61 |
| Total | 236840 | 2935 | 896.11 | | |

The genetic map was compared with the lotus draft genome (Fig. 2). The eight linkage groups were integrated on megascaffolds 1 to 15. All linkage groups except LG8 were anchored to at least two megascaffolds. LG1 covered the 1–35 Mb, 49–60 Mb, 73–77 Mb, 114–255 Mb regions of megascaffolds 1 and 8, respectively. LG2 was anchored to the 70–110 Mb region of megascaffolds 2 and 6, and to 1–20 Mb or 26–50 Mb region on megascaffolds 10 and 15, respectively. LG3 covered the 34–48 Mb region of megascaffolds 2 and 5, 20–25 Mb region of megascaffolds 10, and 13. LG4 covered the 60–73 Mb region of megascaffolds 1, 4 and 9. LG5 covered the 50–69 Mb region of megascaffolds 2 and 3, and 11–28 Mb region of megascaffold 11. LG6 was anchored to the 1–35 Mb and 107–133 Mb regions of megascaffold 2 and to 1–10 Mb region of megascaffolds 11, and 14. LG7 was anchored to megascaffolds 7 and 12. Therefore, megascaffolds 1, 2, 7, 9, 10, and 11 were divided into several blocks, and resettled with megascaffolds 3, 4, 5, 6, 8, 12, 13 14 and 15. Taken together, the original 15 megascaffolds in the draft genome were integrated into novel 8 megascaffolds (Figure S1).

QTL analysis for rhizome enlargement

The variation in phenotypic traits among segregating population was used as a key parameter for QTL mapping. Four phenotypic traits, NER, MDE, LER and REI, were distributed continuously in the F₂ population. Significant transgressive segregation was observed for each trait (Figure S2). ANOVA results showed that genotype, environment and their interactions had highly significant effects ($P < 0.001$) on these traits, indicating that the three traits were easily influenced by environmental factors (Table S4).

Mapping analysis for NER and REI identified 14 significant QTLs in the F₂ population in three climatic years (Table 3, Fig. 3A). For NER, four QTLs distributed on LG5 and LG6 were detected, with their phenotypic variation ranging from 7.5 to 14.8%. The region which covered 9.4–11.8 cM on LG6 was repeatedly found across the three years. Ten QTLs on LG1, LG2, LG5 and LG6 were identified for REI in the three years, with phenotypic variation of these QTLs ranging from 6.6 to 22.8%. A region covering 28.4–29.9 cM on LG1 was consistently detected in two years, and another region covering 12.3–13.5 cM on LG1 was detected in all the three climatic years (Table 3).

Table 3
Significant QTLs identified for the traits in three climatic years

| Trait | Name of QTL | Year | LG | Peak ^a | C.I. ^b | LOD ^c | R ² ^d | Additive ^e |
|---|--------------|------|----|-------------------|-------------------|------------------|-----------------------------|-----------------------|
| Number of enlarged rhizome | qNER-LG6-14 | 2014 | 6 | 10 | 9.4–11.8 | 2.73 | 7.45 | 0.35 |
| | qNER-LG6-15 | 2015 | 6 | 10 | 9.4–11.8 | 2.82 | 11.53 | 0.31 |
| | qNER-LG8-15 | | 5 | 29 | 28.0–29.7 | 3.94 | 13.59 | -0.38 |
| | qNER-LG6-16 | 2016 | 6 | 10 | 8.6–11.8 | 3.27 | 14.85 | 0.39 |
| Rhizome enlargement Index | qREI-LG1-14 | 2014 | 1 | 29 | 28.4–29.9 | 4.09 | 11.48 | 0.05 |
| | qREI-LG2-14a | | 2 | 52 | 51.3–53.5 | 3.45 | 8.75 | 0.03 |
| | qREI-LG2-14b | | 2 | 12 | 11.6–13.5 | 3.61 | 20.00 | 0.05 |
| | qREI-LG6-14 | | 6 | 10 | 9.5–10.5 | 3.44 | 9.55 | 0.05 |
| | qREI-LG5-14 | | 5 | 29 | 28.0–29.7 | 4.98 | 14.52 | -0.05 |
| | qREI-LG1-15 | 2015 | 1 | 29 | 28.4–29.9 | 2.65 | 10.91 | 0.03 |
| | qREI-LG2-15 | | 2 | 13 | 12.3–14.0 | 2.72 | 22.80 | 0.04 |
| | qREI-LG1-16a | 2016 | 1 | 162 | 160.9–165.8 | 2.53 | 9.14 | 0.03 |
| | qREI-LG1-16b | | 1 | 46 | 45.9–47.1 | 2.84 | 6.67 | 0.03 |
| | qREI-LG2-16 | | 2 | 12 | 11.6–13.5 | 3.65 | 19.96 | 0.06 |
| ^a The peak position of QTL; ^b Confidence interval at $P=0.05$; ^c LOD score as calculated by WinQTLCart 2.5. ^d Represents the phenotypic variation accounted for; | | | | | | | | |
| ^e Additive effect, positive value (+) means the allele is derived from the female parent 'BG', negative value (-) means the allele is derived from the male parent 'WR1'. | | | | | | | | |

These significant QTLs with overlapping confidence intervals were integrated into the consensus QTLs (Fig. 3A). As a result, one consensus NER QTL designated cqNER-LG1, with a positive additive effect and

covering 2.4 cM, was integrated. Two consensus QTLs, cqREI-LG1 and cqREI-LG2, were identified for REI covering a 1.5 cM and 1.2 cM regions on LG1 and LG2, respectively. cqREI-LG2 explained about 20% of the phenotypic variation, and was the major QTL for REI. The two consensus REI QTLs had a positive additive effect, suggesting that alleles from the female parent of 'BG' contributes to rhizome enlargement.

Candidate genes underlying the major cqREI-LG2 QTL interval

A major cqREI-LG2 QTL showing 20% of the phenotypic variation and with a rhizome-enlargement effect were identified across the three-year QTL mapping period. The QTL interval of cqREI-LG2 covered a 4.5 Mb region on megascaffold 6, which harbored 59 annotated genes (Table S5). Of these genes, 11 had low or no expression in all tested tissues including leaves, petiole, petals, pistils, stamens, roots and rhizomes, while the remaining 48 genes were highly expressed in most tissues (Figure S3). A previous transcriptome data showed an increase in the number of candidate genes controlling the conversion of stolon to rhizome between the ZT1 and ZT2 or RT2 and RT3 developmental stages in two lotus cultivars of 'RL' (tropical lotus with stolon) and 'ZO' (temperate lotus with enlarged rhizome), and these genes were considered candidate for rhizome girth enlargement (Yang et al. 2015). Of the 48 highly expressed genes in this study, nine were predicted as candidate genes for rhizome girth enlargement (Fig. 3B). Among them, four including NNU_05272, NNU_05296, NNU_05305 and NNU_05314 involved in cell and lipid metabolism were observed. A soluble starch synthase-like gene, NNU_05273, was also identified. In addition, two genes, NNU_05287 and NNU_05299, encoding ubiquitin E2 enzyme and peroxidase, respectively, were detected. Moreover, a BEL1-like homeobox protein 6 homolog gene, NNU_05289, and a bHLH transcription factor, NNU_05313, were identified. Expression levels of NNU_05273, NNU_05289 and NNU_05305 increased at the RT3 and ZT2, while the remaining six genes decreased at these two stages (Fig. 3B).

StBEL5 has been shown to be a transcriptional activator of *StSP6A* which acts as a tuberization signal in potato (Abelenda et al. 2014; Sharma et al. 2016). The BEL1-like homeobox protein 6 homolog gene, NNU_05289, designated as *NnBEL6*, is most likely candidate gene that potentially control rhizome enlargement. Thus we sequenced 8,485bp of the genomic region of *NnBEL6* from 1.5 kb upstream to the stop codon in the two parental genotypes and other 23 lotus accessions phenotypically classified into class I, II and III with varying REI as mentioned above (Fig. 3C). A total of 17 polymorphic sites were identified, with four in the promoter region, one in the first exon (+ 124), ten in the first intron, one in the second intron (5,923) and one in the fourth exon (6,695). Only two polymorphic sites at -1,347 and 3,741 were detected between class I and II lotus. The remaining 15 polymorphic sites were observed between class III lotus and those of class I or II, and were highly conserved in most varieties, except for two lotus accessions (Fig. 3C). These results suggested that two polymorphic sites at -1347 and 3741 in *NnBEL6* could be crucial for determining rhizome phenotype in lotus. Thus, *NnBEL6* is a key candidate gene regulating rhizome enlargement in lotus.

Functional validation of *NnBEL6* gene in promoting rhizome enlargement

NnBEL6 encodes a putative protein of 751 amino acids and contains four conserved regions: the conserved BELL domain (Ger340-Ile424), the homeodomain (Trp393-Glu540), the N-terminal SKY, and C-terminal VSLTLGL boxes (Figure S4A). Phylogenetic analysis constructed with the alignments of *NnBEL6* protein with the BEL-like genes from *Arabidopsis thaliana* and *Solanum tuberosum* revealed a close relationship between *NnBEL6* and *AtBEL6* or *AtBEL7* (Figure S4B). The tissue distribution and expression pattern of *NnBEL6* were determined by qRT-PCR analyses (Fig. 4). *NnBEL6* was expressed in the leaf, petiole, flowerbud, pedicel, rhizome and root at the stolon and initial swelling stage of 'ZO' (Fig. 4B). At the stolon stage, *NnBEL6* transcript was highly accumulated in the rhizome and roots, but with decreased levels in the leaf, petiole, flower, and pedicel. At the initial swelling stage, *NnBEL6* was primarily expressed in the rhizome with a four-fold higher transcript levels than that at the stolon stage. In contrast, the expression of *NnBEL6* in other tested tissues showed no significant variation. *NnBEL6* expression in first rhizome internode was higher in 'ZO' than that in 'WR1' at the initial swelling stage (Fig. 4C). Comparison of rhizome change and *NnBEL6* mRNA levels over the developmental time-course indicated its strict correlation with the developmental process, suggesting that *NnBEL6* might play a positive role in controlling the initial rhizome enlargement. Subcellular localization of *NnBEL6* performed in tobacco epidermal cells indicated that *NnBEL6* was a nuclear protein (Figure S5).

To further confirm the function of *NnBEL6* in the formation of storage organ, an overexpressing *NnBEL6* construct driven by the CaMV 35S promoter was transformed into the potato cultivar 'E3', which can only form microtuber under SD in *vitro* (Fig. 5). The results demonstrated that under SD, overexpressing *NnBEL6* promoted tuber formation. The transgenic lines developed microtuber at approximately 95 days after transplantation (DAT), while microtuber formation started in the WT 'E3' controls at 109 DATs (Figs. 5A and B). The transgenic lines formed an average of 3.2 microtubers per plant at 120 DATs (Fig. 5C). To dissect the effect of *NnBEL6* on signal transduction in the photoperiodic tuberization pathway, we further assessed the endogenous expression of two reported key genes in the photoperiodic tuberization pathway, *StSP6A* and *StSP5G*, in the transgenic lines. The expression level of *StSP6A* was significantly up-regulated in 35S:*NnBEL6* transgenic plants (Fig. 5D), which was consistent with the observed early tuberization result. However, the expression level of *StSP5G* showed varied profiles, with a slight down-regulation in two transgenic plants, and an up-regulation in one transgenic plant (Fig. 5E).

Discussion

High-quality genetic map is important for identifying genes of complex traits. A series of genetic maps have been constructed in lotus using SSR and SNP molecular markers from several segregating populations (Liu et al. 2016; Yang et al. 2012; Zhang et al. 2014). The first genetic map of lotus was constructed based on 224 SSR markers using an F1 population derived from a cross between Asian lotus 'Chinese Antique' and American lotus 'AL1' (Yang et al. 2012). More recently, a much higher density genetic map for the same mapping population was constructed using restriction-site associated DNA sequencing (RADseq), and a total of 634 bins with 4,098 SNPs were distributed in the American lotus genetic map (Zhang et al. 2014). While a double digest restriction site-associated DNA sequencing (ddRADseq) was used to construct a linkage map of 581.3 cM, with 791 bin markers consisting of 8,971

SNPs sorted into eight LGs (Liu et al. 2016). Moreover, an F_2 population and genotyping-by-sequencing (GBS) technique was used to construct a genetic map with 1,475 bin markers containing 12,113 SNP markers (Huang et al. 2021).

In this study, we report for the first time a whole genome sequencing (WGS) approach for detecting SNPs between two parental lotus genotypes. A total of 2,935 bins consisting of 236,840 SNPs were distributed on eight LGs, spanning 896.11 cM (Table 2). The number of markers identified and length of the linkage map in this study are higher than those previously reported, suggesting the accuracy and uniformity of SNP detected by WGS. This could be due to the direct whole genome sequencing without restriction site digestion of associated DNA tags. Thus, WGS-based SNP discovery technique is a worthy tool for rapid construction of high-density genetic maps, which can be genotyped on a much larger scale (Huang et al. 2009; Scheben et al. 2017), and it has been applied in the construction of linkage maps and QTL detection in species such as, *Eucalyptus* (Bartholome et al. 2015), *Malus pumila* (Allard et al. 2016), *Cajanus cajan* (Yadav et al. 2019).

We observed a highly unbalanced distribution of SNP markers in the two lotus parents used in this study. The number of *lmxll* markers was extremely higher than for *nnxnp* markers (Table 1), this could be attributed to the unbalanced heterozygosity in parental genomes. 'BG' is a progeny of several cross between Chinese lotus, and exhibited a high ratio of heterozygosity, while 'WR1' is a wild Thai lotus, with lower level of heterozygosity (Huang et al. 2018). Thus, more SNPs with higher ratio of heterozygosity could be developed from their genomes. Normally, only homozygous *aaxbb* markers are used for the F_2 population linkage map construction (Huang et al. 2021; Liu et al. 2016). Due to the high heterozygosity of the two lotus parents used in this study, three heterozygous markers *lmxll*, *nnxnp* and *hxxhk* were detected in the F_2 population. The number of heterozygous marker were much higher than the homozygous *aaxbb* makers (Table 1), thus using *aaxbb* marker could risk omitting several chromosomal coverage and breakpoints, thereby reducing the accuracy and resolution of linkage map. This assumption was validated by our results that more gap and lower coverage was existed in a set of linkage maps individually constructed with *aaxbb*, *lmxll*, and *nnxnp* than in merged linkage map (Figs. 2 and S6). This suggested that, both heterozygous and homozygous markers should be used for construction of genetic maps of F_2 mapping population derived from highly heterozygous parents.

The consensus linkage map was compared with the draft genome. As a result, approximately 67.6% of 'China Antique' draft genome was anchored onto the ten largest megascaffolds, namely megascaffolds 1–10. In addition, putative false joins within the draft genome assembly were detected (Fig. 2). Megascaffold 1 was broken into eight blocks (A-H). Block A, C, E, H were linked with megascaffold 8, and assembled as the longest megascaffold. For megascaffolds 6, blocks 10A, 10C and 2E were reassembled as a new megascaffold (Figure S1). Similar results were also observed by fluorescence *in situ* hybridization (FISH) analysis in lotus (Meng et al. 2017). The FISH assay also revealed that megascaffolds 1–10 were assigned to eight chromosomes. In addition, the assay showed that megascaffolds 1 and 8 were located on the longest chromosome, while megascaffolds 6 and 10 were on the same chromosome. We concluded that some segement of the draft genome of 'China Antique' was

previously misplaced, which could be due to that the assembly of the draft genome was based on the genetic map of American lotus. While there are chromosomal rearrangements between the genomes of Asian lotus and American lotus. Moreover, the small population size of 51 individuals used in the American lotus could be responsible for the lack of identification of several recombination breakpoints in the chromosomes (Ming et al. 2013; Yang et al. 2012). Due to these factors, BioNano and PacBio were used to generate a more accurate draft genome of 'China Antique' variety (Gui et al. 2018; Shi et al. 2020). The high synteny between the two versions of genome further confirmed the accuracy of our generated linkage map (Figure S7).

Rhizome is an important lotus storage organ that is also used for asexual propagation. The enlargement of rhizome is a strategy that enables lotus plants to survive the cold winter. As mentioned above, the temperate lotus undergo four developmental stages, and produce enlarged internodes in late autumn, which can survive the winter in the temperate and sub-temperate regions (Yang et al. 2015). A previous study showed a notable phenotypic variation in rhizome branching, elongation and expansion of lotus germplasm (Lin et al. 2019). Higher REI ensures stronger survival ability against low winter temperature, thus studies on the mechanism of rhizome enlargement are vital for understanding the lotus survival and propagation. The increase in rhizome girth is associated with photoperiod and hormone, which are induced by SDs and ABA (Masuda et al. 2010; Masuda et al. 2007a, 2012; Masuda et al. 2007b). It has been reported that the genes in photoperiod pathway, starch metabolism and hormone signal transduction involved in rhizome girth enlargement (Cao et al. 2019; Cheng et al. 2013; Yang et al. 2015). Here, we used QTL mapping to uncover the genetic determinants of lotus rhizome enlargement. Fourteen significant QTLs were identified for NER and REI on LG1, LG2, and LG6 in the three climatic years. Three QTL regions, cqREI-LG1, cqREI-LG2, and cqNER-LG6, which anchored to megascaffolds 1, 6 and 2, were reliable for NER and REI (Table 3 and Fig. 3). Although Huang et al. (2021) also detected QTLs for REI on LG1 and LG2 in the 'WR1' and 'ZO' F₂ mapping population, these QTL were anchored to megascaffolds 3 and 4, and could only be found in one year. Thus, these QTLs identified in this study were no collinearity with those ones in our study that are more reliable.

A *BEL 1-like* family transcription factor, *NnBEL6*, located within a confidence interval of a major QTL designated cqREI-LG2, with the highest phenotypic contribution, was identified as a key candidate gene (Fig. 3). The expression of *NnBEL6* was highly correlated with REI in the two cultivars of 'RL' and 'ZO' (Figs. 3 and 4). Members of the *BEL 1-like* family transcription factors are associated with aspects of plant growth such as meristem maintenance, tuberization, and floral and root development, and had been identified in all plant species (Brambilla et al. 2007; Hannapel et al. 2013; Kumar et al. 2007; Sharma et al. 2014). Our study revealed that *NnBEL6* is expressed in all tissues, with high transcript levels in rhizome. In addition, the level of RNA accumulation in enlarged lotus rhizome was enhanced from the stolon stage to the initial swelling stage (Fig. 4). Notably, one SNP at position 3741 and one Indel at position - 1374 showed dramatic differences between the sequences of temperate and tropical lotus varieties, which suggest they could be functional polymorphic sites for initiating lotus rhizome enlargement (Fig. 3C). Overall, our results revealed that *NnBEL6* is a key candidate gene potentially involved rhizome enlargement.

of temperate lotus varieties. *StSP6A* is a key tuberization signal in potato which is activated by *StBEL5* (Abelenda et al. 2014; Sharma et al. 2016). In this study, ectopic expression of *35S:NnBEL6* promoted tuberization in transgenic potato (Figs. 5A-C). The endogenous expression level of *StSP6A* in transgenic lines was highly induced, while the expression of *StSP5G* exhibited no variation in the transgenic lines (Fig. 5D and E). *StSP5G* is a homolog of *FT* that directly represses *StSP6A*, thus inhibiting tuberization (Abelenda et al. 2014). No difference in transcript abundance of *StSP5G* was observed in transgenic potato overexpressing *NnBEL6*, suggesting that *NnBEL6* did not regulate *StSP5G*, but directly activated the expression of *StSP6A* and induced tuberization in potato. Further studies to elucidate the relationship between *NnBEL6* and *StSP6A* are still needed. Overall, our functional analysis demonstrated the role of *NnBEL6* in promoting lotus rhizome enlargement.

Conclusions

We constructed a high-density SNP genetic map of *N. nucifera* with 2,935 markers binned from 236,840 SNPs. Fourteen significant QTLs for NER and REI were identified in three climatic years, which explained 6.67–22.28% of trait variance. Three QTL regions, cqNER-LG1, cqREI-LG1 and cqREI-LG2 were repeatedly identified in at least two years. cqREI-LG2 explained about 20% of the phenotypic variation, and was considered as the major QTL for REI. *NnBEL6* was identified as a putative candidate gene underlying cqREI-LG2. The role of *NnBEL6* in rhizome enlargement in lotus was validated through gene expression analysis, sequence comparison and genetic transformation in potato.

Declarations

Data availability

All data supporting the findings of this study are available within the paper and within its supplementary data published online.

Acknowledgements

We would like to thank Dr. Ray Ming for his help in resequencing analysis and linkage map construction.

Funding

This project was supported by Key Research Program of Frontier Science CAS (Grant No. QYZDB-SSW-SMC017), the National Natural Science Foundation of China (Grant No 31772353 and 31872136), and the Youth Innovation Promotion Association CAS (Grant no. 2017390).

Author's contributions

Y.L. and M.Y. conceived the project and planned the experiments, H.S. and M.Z. cloned and validated the function of the gene. D.Y. planted and managed all lotus materials in the field, X.D. investigated the

phenotypic traits. H.S. and J.L. analyzed SNP and constructed the linkage map. Y.L. and M.Y. analyzed the data and drafted the manuscript. All authors read and approved the final manuscript.

Conflict of interest

The authors declare no competing interest.

References

1. Abelenda JA, Navarro C, Prat S (2014) Flowering and tuberization: a tale of two nightshades. *Trends Plant Sci* 19:115-122
2. Allard A, Bink MCAM, Martinez S, Kelner JJ, Legave JM, di Guardo M, Di Pierro EA, Laurens F, van de Weg EW, Costes E (2016) Detecting QTLs and putative candidate genes involved in budbreak and flowering time in an apple multiparental population. *J Exp Bot* 67:2875-2888
3. Anderson RC (1970) Role of daylength and temperature in tuber formation and rhizome growth of *trientalis-borealis* raf. *Bot Gaz* 131:122-&
4. Arcade A, Labourdette A, Falque M, Mangin B, Chardon F, Charcosset A, Joets J (2004) BioMercator: integrating genetic maps and QTL towards discovery of candidate genes. *Bioinformatics* 20:2324-2326
5. Bartholome J, Mandrou E, Mabilia A, Jenkins J, Nabihoudine I, Klopp C, Schmutz J, Plomion C, Gion JM (2015) High-resolution genetic maps of *Eucalyptus* improve *Eucalyptus grandis* genome assembly. *New Phytol* 206:1283-1296
6. Brambilla V, Battaglia R, Colombo M, Masiero S, Bencivenga S, Kater MM, Colombo L (2007) Genetic and molecular interactions between BELL1 and MADS box factors support ovule development in *Arabidopsis*. *Plant Cell* 19:2544-2556
7. Cao DD, Damaris RN, Zhang Y, Liu MH, Li M, Yang PF (2019) Proteomic analysis showing the signaling pathways involved in the rhizome enlargement process in *Nelumbo nucifera*. *Bmc Genomics* 20
8. Cheng LB, Li SY, Yin JJ, Li LJ, Chen XH (2013) Genome-wide analysis of differentially expressed genes relevant to rhizome formation in lotus root (*Nelumbo nucifera* Gaertn). *Plos One* 8
9. Doyle J, Doyle J (1990) Isolation of plant DNA from fresh tissue. *Focus* 12:13-15
10. Gui ST, Peng J, Wang XL, Wu ZH, Cao R, Salse J, Zhang HY, Zhu ZX, Xia QJ, Quan ZW, Shu LP, Ke WD, Ding Y (2018) Improving *Nelumbo nucifera* genome assemblies using high-resolution genetic maps and BioNano genome mapping reveals ancient chromosome rearrangements. *Plant J* 94:721-734
11. Hannapel DJ, Sharma P, Lin T (2013) Phloem-mobile messenger RNAs and root development. *Front Plant Sci* 4
12. Huang LY, Li M, Cao DD, Yang PF (2021) Genetic dissection of rhizome yield-related traits in *Nelumbo nucifera* through genetic linkage map construction and QTL mapping. *Plant Physiol Bioch* 160:155-165

13. Huang LY, Yang M, Li L, Li H, Yang D, Shi T, Yang PF (2018) Whole genome re-sequencing reveals evolutionary patterns of sacred lotus (*Nelumbo nucifera*). *J Integr Plant Biol* 60:2-15
14. Huang XH, Feng Q, Qian Q, Zhao Q, Wang L, Wang AH, Guan JP, Fan DL, Weng QJ, Huang T, Dong GJ, Sang T, Han B (2009) High-throughput genotyping by whole-genome resequencing. *Genome Res* 19:1068-1076
15. Kumar R, Kushalappa K, Godt D, Pidkowich MS, Pastorelli S, Hepworth SR, Haughn GW (2007) The *Arabidopsis* *BEL 1-LIKE HOMEODOMAIN* proteins SAW1 and SAW2 act redundantly to regulate KNOX expression spatially in leaf margins. *Plant Cell* 19:2719-2735
16. Lee R, Baldwin S, Kenel F, McCallum J, Macknight R (2013) *FLOWERING LOCUS T* genes control onion bulb formation and flowering. *Nat Commun* 4
17. Lin Z, Zhang C, Cao D, Damaris RN, Yang P (2019) The latest studies on lotus (*Nelumbo nucifera*)-an emerging horticultural model plant. *Int J Mol Sci* 20:3680
18. Liu ZW, Zhu HL, Liu YP, Kuang J, Zhou K, Liang F, Liu ZH, Wang DP, Ke WD (2016) Construction of a high-density, high-quality genetic map of cultivated lotus (*Nelumbo nucifera*) using next-generation sequencing. *Bmc Genomics* 17
19. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* 25:402-408
20. Lyngkhai F, Khar A, Mangal M, Gaikwad AB, Thirunavukkarasu N (2019) Expression analysis and association of bulbing to *FLOWERING LOCUS T* (*FT*) gene in short day onion (*Allium cepa* L.). *Indian J Genet PI Br* 79:77-81
21. Man JM, Cai JW, Cai CH, Xu B, Huai HY, Wei CX (2012) Comparison of physicochemical properties of starches from seed and rhizome of lotus. *Carbohydr Polym* 88:676-683
22. Manoharan RK, Han JSH, Vijayakumar H, Subramani B, Thamilarasan SK, Park JI, Nou IS (2016) Molecular and functional characterization of *FLOWERING LOCUS T* homologs in *Allium cepa*. *Molecules* 21
23. Masuda J, Ozaki Y, Miyajima I, Okubo H (2010) Ethylene is not involved in rhizome transition to storage organ in lotus (*Nelumbo nucifera*). *J Fac Agr Kyushu U* 55:231-232
24. Masuda J, Ozaki Y, Okubo H (2007a) Rhizome transition to storage organ is under phytochrome control in lotus (*Nelumbo nucifera*). *Planta* 226:909-915
25. Masuda J, Ozaki Y, Okubo H (2012) Regulation in rhizome transition to storage organ in lotus (*Nelumbo nucifera* Gaertn.) with exogenous gibberellin, gibberellin biosynthesis inhibitors or abscisic acid. *J Jpn Soc Hortic Sci* 81:67-71
26. Masuda J, Yoshimizu S, Ozaki Y, Okubo H (2007b) Rhythmic response of rhizome growth to light-break in lotus (*Nelumbo nucifera*). *J Fac Agr Kyushu U* 52:35-38
27. Meng Z, Hu XX, Zhang ZL, Li ZJ, Lin QF, Yang M, Yang PF, Ming R, Yu QY, Wang K (2017) Chromosome nomenclature and cytological characterization of sacred Lotus. *Cytogenet Genome Res* 153:223-231

28. Ming R, VanBuren R, Liu YL, Yang M, Han YP, Li LT, Zhang Q, Kim MJ, Schatz MC, Campbell M, Li JP, Bowers JE, Tang HB, Lyons E, Ferguson AA, Narzisi G, Nelson DR, Blaby-Haas CE, Gschwend AR, Jiao YN, Der JP, Zeng FC, Han J, Min XJ, Hudson KA, Singh R, Grennan AK, Karpowicz SJ, Watling JR, Ito K, Robinson SA, Hudson ME, Yu QY, Mockler TC, Carroll A, Zheng Y, Sunkar R, Jia RZ, Chen N, Arro J, Wai CM, Wafula E, Spence A, Han YN, Xu LM, Zhang JS, Peery R, Haus MJ, Xiong WW, Walsh JA, Wu J, Wang ML, Zhu YJ, Paull RE, Britt AB, Du CG, Downie SR, Schuler MA, Michael TP, Long SP, Ort DR, Schopf JW, Gang DR, Jiang N, Yandell M, dePamphilis CW, Merchant SS, Paterson AH, Buchanan BB, Li SH, Shen-Miller J (2013) Genome of the long-living sacred lotus (*Nelumbo nucifera* Gaertn.). *Genome Biol* 14
29. Murashige T, Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol Plantarum* 15:473-497
30. Navarro C, Cruz-Oro E, Prat S (2015) Conserved function of *FLOWERING LOCUS T (FT)* homologues as signals for storage organ differentiation. *Curr Opin Plant Biol* 23:45-53
31. Scheben A, Batley J, Edwards D (2017) Genotyping-by-sequencing approaches to characterize crop genomes: choosing the right tool for the right application. *Plant Biotechnol J* 15:149-161
32. Sharma P, Lin T, Grandellis C, Yu M, Hannapel DJ (2014) The *BEL1-like* family of transcription factors in potato. *J Exp Bot* 65:709-723
33. Sharma P, Lin T, Hannapel DJ (2016) Targets of the *StBEL5* transcription factor include the *FT* ortholog *StSP6A*. *Plant Physiol* 170:310-324
34. Shen-Miller J (2002) Sacred lotus, the long-living fruits of China Antique. *Seed Sci Res* 12:131-143
35. Shi T, Rahmani RS, Gugger PF, Wang M, Li H, Zhang Y, Li Z, Wang Q, Van de Peer Y, Marchal K, Chen J (2020) Distinct expression and methylation patterns for genes with different fates following a single whole-genome duplication in flowering plants. *Mol Biol Evol* 37:2394-2413
36. Teo CJ, Takahashi K, Shimizu K, Shimamoto K, Taoka K (2017) Potato tuber induction is regulated by interactions between components of a tuberigen complex. *Plant Cell Physiol* 58:365-374
37. VanOoijen J (2006) JoinMap 4 software for the calculation of genetic linkage maps in experimental populations. Wageningen, Netherlands
38. Wang S, Basten CJ, Zeng ZB (2012) Windows QTL Cartographer 2.5. Department of Statistics. North Carolina State University, Raleigh, NC.
39. Yadav P, Saxena KB, Hingane A, Kumar CVS, Kandalkar VS, Varshney RK, Saxena RK (2019) An "Axiom Cajanus SNP Array" based high density genetic map and QTL mapping for high-selfing flower and seed quality traits in pigeonpea. *Bmc Genomics* 20
40. Yang M, Han YN, VanBuren R, Ming R, Xu LM, Han YP, Liu YL (2012) Genetic linkage maps for Asian and American lotus constructed using novel SSR markers derived from the genome of sequenced cultivar. *Bmc Genomics* 13
41. Yang M, Zhu L, Xu L, Pan C, Liu Y (2014) Comparative transcriptomic analysis of the regulation of flowering in temperate and tropical lotus (*Nelumbo nucifera*) by RNA-Seq. *Ann Appl Biol* 165:73-95

42. Yang M, Zhu LP, Pan C, Xu LM, Liu YL, Ke WD, Yang PF (2015) Transcriptomic analysis of the regulation of rhizome formation in temperate and tropical lotus (*Nelumbo nucifera*). *Sci Rep-Uk* 5
43. Zhang Q, Li LT, VanBuren R, Liu YL, Yang M, Xu LM, Bowers JE, Zhong CH, Han YP, Li SH, Ming R (2014) Optimization of linkage mapping strategy and construction of a high-density American lotus linkage map. *Bmc Genomics* 15
44. Zhang XY, Chen LQ, Wang QC (2011) New lotus flower cultivars in China. China forestry Publishing House, Beijing
45. Zhang XY, Wang QC (2006) Preliminary study of the eco-types of genetic resources of tropical lotus. *Landscape Plants* 22:82-85
46. Zhou TT, Song BT, Liu TF, Shen YL, Dong LP, Jing SL, Xie CH, Liu J (2019) Phytochrome F plays critical roles in potato photoperiodic tuberization. *Plant J* 98:42-54

Figures

Figure 1

Phenotypic variation of rhizomes in three lotus varieties, 'BG', 'WR1' and ZO. 'BG' and 'WR1' are the female and male parents, respectively, of F₂ population used in the study. (A) The rhizome of 'BG'. (B) The rhizome of 'WR1'. (C) The rhizome of 'ZO', which was also used to analyze the expression profiles of the candidate, *NnBEL6* gene. Bar =10 cm. (D) Number of enlarged rhizomes in the three lotus varieties in three climatic years (2014, 2015, and 2016). (E) Rhizome enlargement index (REI) of the three lotus varieties in climatic years (2014, 2015, and 2016). REI was estimated as the ratio of maximum diameter to length of enlarged internode, according to the method described by Masuda et al., (2007b).

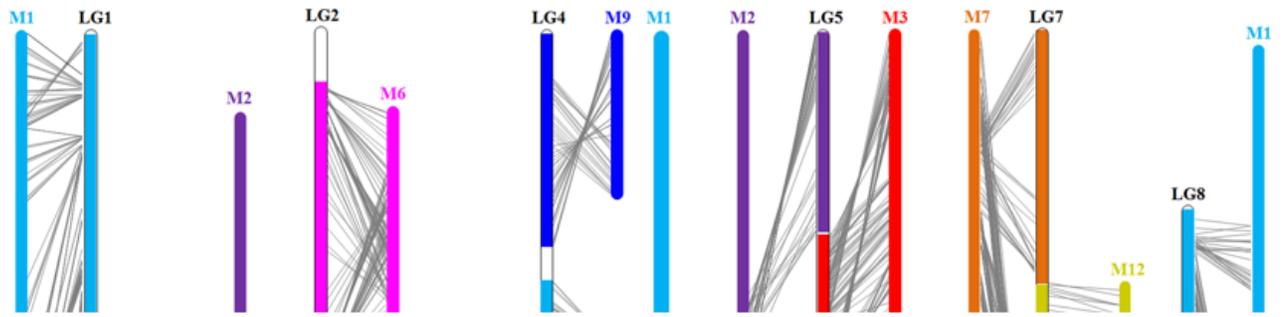


Figure 2

Integration of genetic linkage map and the draft genome of 'China Antique'. The blocks M1-M15 represent megascaffolds 1-15 of the draft genome. The blocks L1-LG8 represent linkage groups 1-8 of genetic linkage map.

Figure 3

Quantitative trait loci (QTLs) analysis and identification of genes within the confidence interval of cqREI-LG2. (A) Circos plot of QTL mapping results for the number of enlarged rhizome (NER) and rhizome enlargement index (REI) on eight linkage groups. The innermost circle represents the linkage group. The shaded black and grey areas represent molecular markers. The dotted circles from outside to inside represent QTL identified in 2014, 2015, and 2016 climatic years, and the red or blue colored areas within them indicate significant QTLs. The outermost circle indicate consensus QTLs. The confidence interval of each QTL is shown by the width of the bar. (B) The FPKM expression of nine genes located within the confidence interval of cqREI-LG2. RT1, RT2, and RT3 represent three developmental stages of 'RL' rhizome. ZT1, ZT2, and ZT3 represent three developmental stages of 'ZO' rhizome. (C) Sequence difference of NnBEL6 gene among various lotus germplasm. Numbers in the boxes represent the length in base pairs (bp) of the promoter and exons. Upper case letters indicate SNPs or Indels in three categories of lotus germplasm classified as, class I, II, and III based on REI values.

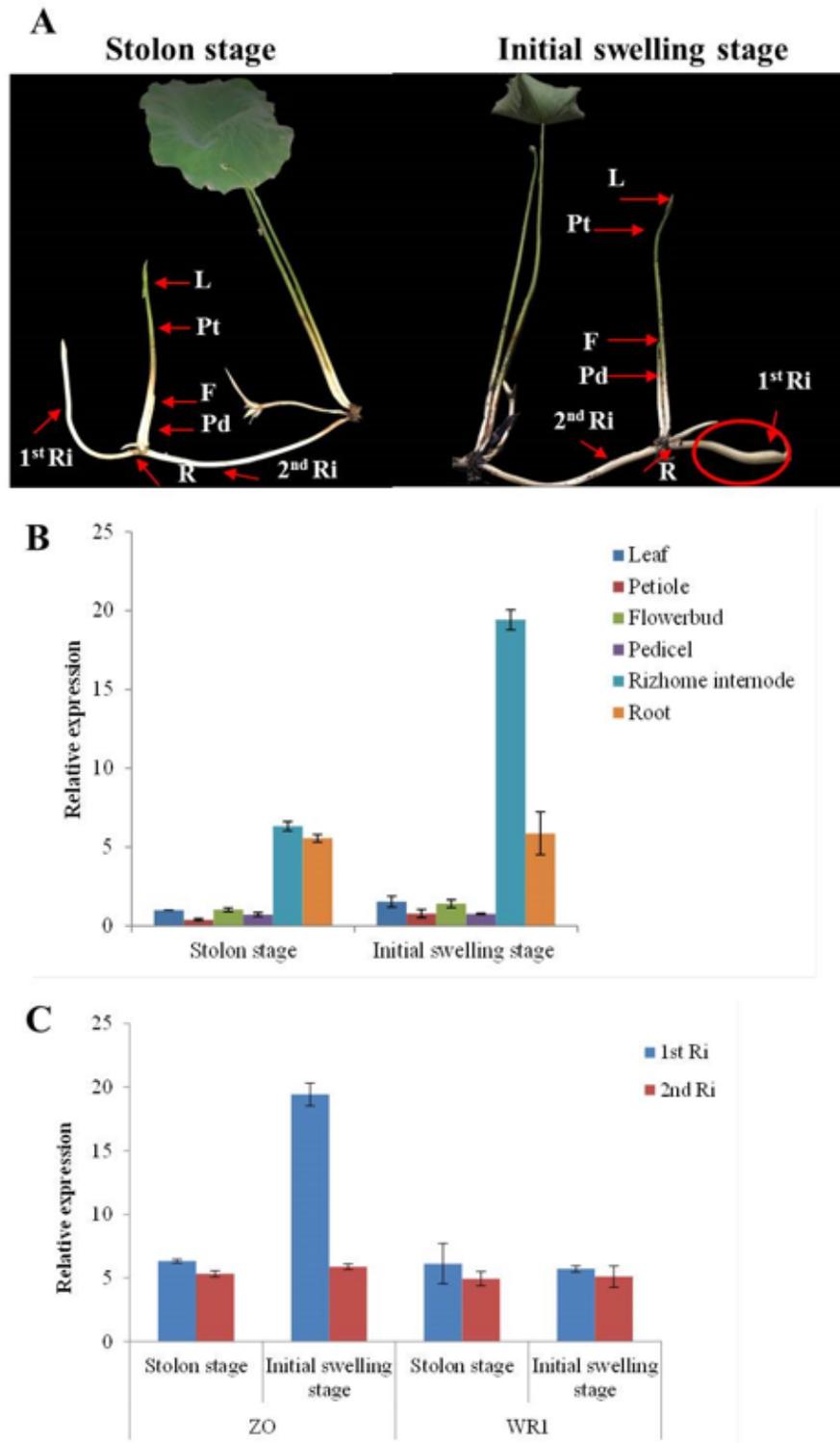


Figure 4

Expression pattern of *NnBEL6* revealed by qRT-PCR. (A) The stolon stage and initial swelling stage samples from 'ZO' were used to analyze the expression levels of *NnBEL6*. The abbreviations are as listed in Fig. 4B. (B) The expression of *NnBEL6* in leaves (L), petiole (Pt), flower-bud (F), pedicel (Pd), 1st rhizome internode (1st Ri) and root (R) from the stolon stage and initial swelling stage of 'ZO'. (C) The

expression of *NnBEL6* in the first (1st Ri) and second (2nd Ri) rhizome internode in the stolon stage and initial swelling stage from 'ZO' and 'WR1'.

Figure 5

Tuberization of 'E3' potato lines overexpressing *NnBEL6* and in WT 'E3' controls which can form tubers under short day conditions (SDs). (A) representative photographs of overexpressing *NnBEL6* lines and WT 'E3' control plants grown under SD for 120 days *in vitro* (tubers are indicated by red arrows). (B) The number of days until the start of microtubers development in overexpressing lines and WT 'E3' control plants grown under SDs. (C) Number of microtubers in 'E3' lines overexpressing *NnBEL6* and WT 'E3' control plants grown under SDs for 120 days. (D) Expression of *StSP6A* and *StSP5G* (E) in the overexpressing *NnBEL6* lines and WT 'E3' control plants grown under SD.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureS1.jpg](#)
- [FigureS2.jpg](#)
- [FigureS3.jpg](#)
- [FigureS4.jpg](#)
- [FigureS5.jpg](#)
- [FigureS6.jpg](#)
- [FigureS7.jpg](#)
- [TableS1.docx](#)
- [TableS2.docx](#)
- [TableS3.xlsx](#)
- [TableS4.docx](#)
- [TableS5.xlsx](#)