Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology) *2021 22(6):476-491*
www.jzus.zju.edu.cn; www.springer.com/journal/11585
E-mail: jzus_b@zju.edu.cn

Research Article

https://doi.org/10.1631/jzus.B2000545

Molecular cloning, characterization and expression analysis of three key starch synthesis-related genes from the bulb of a rare lily germplasm, *Lilium brownii* var. *giganteum*

Yun WU[1,2*], Minyi SUN[2*], Shiqi LI[2], Ruihan MIN[2], Cong GAO[2], Qundan LYU[3], Ziming REN[2✉], Yiping XIA[2✉]

[1]*Department of Landscape Architecture, School of Civil Engineering and Architecture, Zhejiang Sci-Tech University, Hangzhou 310018, China*
[2]*Genomics and Genetic Engineering Laboratory of Ornamental Plants, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310058, China*
[3]*Chemical Biology Center, Lishui Institute of Agriculture and Forestry Sciences, Lishui 323000, China*

Abstract: Starch is the predominant compound in bulb scales, and previous studies have shown that bulblet development is closely associated with starch enrichment. However, how starch synthesis affects bulbification at the molecular level is unclear. In this study, we demonstrate that *Lilium brownii* var. *giganteum*, a wild lily with a giant bulb in nature, and *L. brownii*, the native species, have different starch levels and characteristics according to cytological and ultra-structural observations. We cloned the complete sequence of three key gene-encoding enzymes (LbgAGPS, LbgGBSS, and LbgSSIII) during starch synthesis by rapid amplification of 5' and 3' complementary DNA (cDNA) ends (RACE) technology. Bioinformatics analysis revealed that the proteins deduced by these genes contain the canonical conserved domains. Constructed phylogenetic trees confirmed the evolutionary relationships with proteins from other species, including monocotyledons and dicotyledons. The transcript levels of various tissues and time course samples obtained during bulblet development uncovered relatively high expression levels in bulblets and gradual increase expression accompanying bulblet growth. Moreover, a set of single nucleotide polymorphisms (SNPs) was discovered in the *AGPS* genes of four lily genotypes, and a purifying selection fashion was predicted according to the non-synonymous/synonymous (Ka/Ks) values. Taken together, our results suggested that key starch-synthesizing genes might play important roles in bulblet development and lead to distinctive phenotypes in bulblet size.

Key words: *Lilium*; Starch synthesis; Bulblet development; Single nucleotide polymorphism (SNP); Expression pattern

1 Introduction

Starch is a polyglucan that is stored as granules within plastids (e.g., chloroplasts in leaves and amyloplasts in storage organs) and is a compound that plays a vital role in determining plant yields (Lloyd and Kossmann, 2019). Normally, there are transitory starch reserves and storage starch reserves. The highly complex, hierarchical structure of starch granules arises from the activity of a large suite of enzymes (Goren et al., 2018). First, the glucose donor adenosine-5'-diphosphate-glucose (ADPG) is generated by the enzyme adenosine-5'-diphosphate-glucose pyrophosphorylase (AGPase), which consists of two subunits, the large subunit (AGPL) and the small subunit (AGPS) (Goren et al., 2018). The glucose-1-phosphate (G1P) substrate for AGPase is derived either from the Calvin-Benson cycle in chloroplasts or from hexose phosphates imported from the cytosol. Second, ADPG is used for the initiation and elongation of two types of starch (amylopectin and amylose) by the coordinated reactions of soluble starch synthase (SS) and granule-bound starch synthase (GBSS) (Fig. 1). The starch synthesis pathway also involves the participation of starch-branching enzymes (SBEs) and debranching enzymes that eventually form the dedicated starch granule matrix (Goren et al., 2018; Smith and Zeeman, 2020).

✉ Yiping XIA, ypxia@zju.edu.cn
Ziming REN, zimingren@zju.edu.cn
* The two authors contributed equally to this work
Yun WU, https://orcid.org/0000-0002-1612-5620
Ziming REN, https://orcid.org/0000-0001-9130-6272
Yiping XIA, https://orcid.org/0000-0003-1335-2593

Received Sept. 10, 2020; Revision accepted Nov. 10, 2020;
Crosschecked May 12, 2021

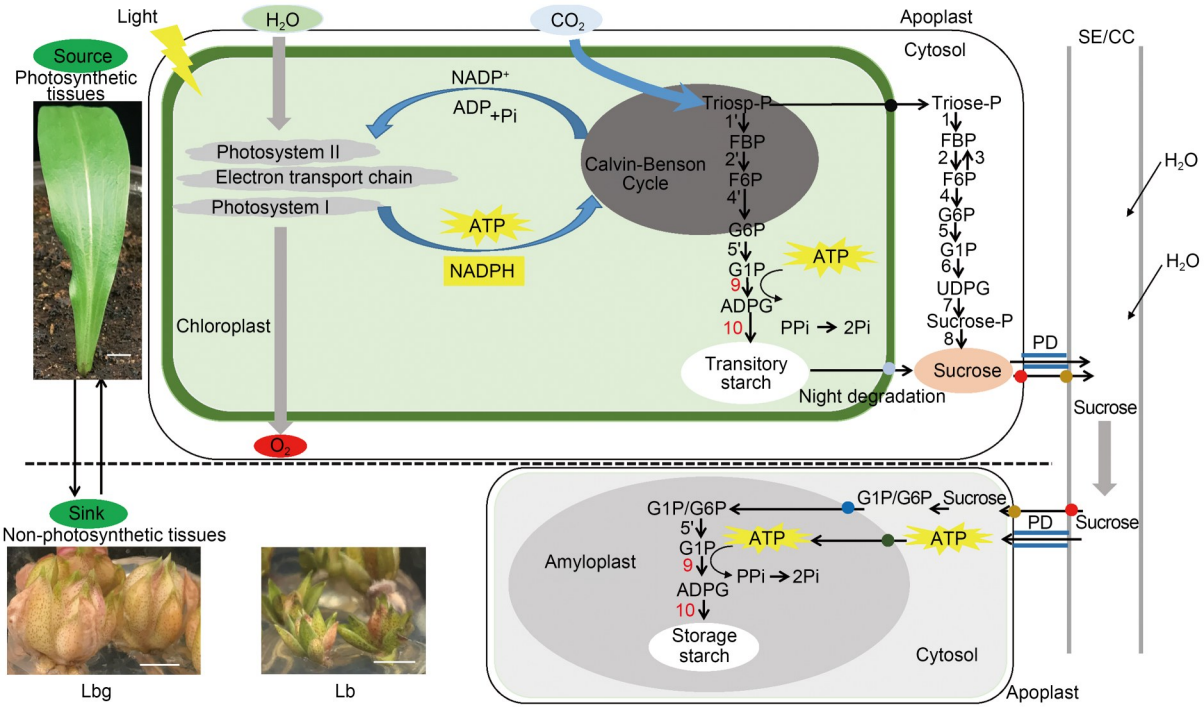

Fig. 1 Classic interpretation of possible source-sink-conversion pathway involving starch synthesis in bulblet (adapted from Bahaji et al. (2014) with permission from Elsevier). The enzymes are numbered as follows: 1 and 1', fructose-1, 6-bisphosphate aldolase (EC 4.1.2.13); 2 and 2', fructose-1,6-bisphosphatase (FBP; EC 3.1.3.11); 3, 6-phosphofructokinase (EC 2.7.1.11); 4 and 4', plastidial phosphoglucose isomerase (EC 5.3.1.9); 5 and 5', plastidial phosphoglucomutase (EC 5.4.2.2); 6, uridine diphosphate glucose pyrophosphorylase (EC 2.7.7.9); 7, sucrose-phosphate synthase (EC 2.4.1.14); 8, sucrose-phosphate phosphatase (EC 3.1.3.24); 9, adenosin-5'-diphosphate-glucose pyrophosphorylase (EC 2.7.7.27); 10, starch synthase (EC 2.4.1.21). The different colors of solid cylinder mean different membrane transporters. Lbg, *Lilium brownii* var. *giganteum*; Lb, *L. brownii*; PD, plasmodesma; ADP, adenosine diphosphate; ATP, adenosine triphosphate; NADPH, nicotinamide adenine dinucleotide phosphate; Triose-P, triose phosphate; F6P, fructose-6-phosphate; G6P, glucose-6-phosphate; G1P, glucose-1-phosphate; ADPG, adenosin-5'-diphosphate-glucose; UDPG, uridine-5'-diphosphoglucose; PPi, pyrophosphoric acid; Pi, phosophoric acid; SE/CC, sieve element/companion cell. White bar=1 cm (Note: for interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

The metabolic pathways of starch have been intensively studied. van Harsselaar et al. (2017) identified 77 genomic loci coding enzymes involved in starch metabolism, while nine genes of SSs were confirmed in *Saccharum spontaneum* (Ma et al., 2019). Analysis of the evolution of core starch synthesis genes also revealed the conservation and divergence of these genes (Mishra et al., 2017; Qu et al., 2018).

Many key genes have been cloned and functionally characterized in higher plants, especially in the model plant *Arabidopsis thaliana* and species with starchy organs such as rice, corn, potato, and cassava (Lloyd and Kossmann, 2019; Smith and Zeeman, 2020). Zhou et al. (2016) reported the comprehensive cloning of AGPase isoforms in *Ipomoea batatas*, including two for AGPS and four for AGPL. Similar work has been performed in *Nelumbo nucifera* (Cheng et al., 2015). Several studies presented the cloning of genes encoding starch synthase enzymes, such as granule-bound starch synthase I (*GBSSI*) in banana (Miao et al., 2014) and starch synthase III (*SSIII*) in lotus (Zhu et al., 2020). Some starch-related genes have been functionally verified. The overexpression of *IbAGPaseS* in sweet potato could affect carbohydrate gene regulation (Seo et al., 2015). The silencing of the *SSIII* gene by RNA interference (RNAi) resulted in altered characteristics of starch in potato tubers (Du et al., 2012). However, the current studies of starch turnover in non-photosynthetic organs (e.g., tuber) are still limited compared to the studies in *Arabidopsis* (Smith and Zeeman, 2020).

Lily (*Lilium* spp.) has great economic significance due to its versatile uses (van Tuyl et al., 2018).

The various species (cultivars) can be used as cut flowers, potted and garden plants worldwide, and also as food and medicine in some countries (Kamenetsky and Okubo, 2012; Huang et al., 2020). Although more than 300 new cultivars are registered each year (van Tuyl et al., 2011), releasing new cultivar is a lengthy process (ca. 12−15 years) in large part because bulb production is exceedingly slow (de Klerk, 2012).

Researchers have recently been addressing the important problems in bulb development (Podwyszyńska, 2012; Moreno-Pachón, 2017; Askari et al., 2018), finding that starch metabolism might act as a major and indispensable player during bulb initiation and growth (Li et al., 2014; Wu et al., 2016; Yang et al., 2017; Wu et al., 2019a). However, the details of gene function and therefore of the molecular mechanism driving bulb development remain poorly understood.

Lilium brownii var. *giganteum* (Lbg) is native to Zhejiang Province, China, with a giant bulb (10–12 cm in diameter) compared to the species of its origin, *L. brownii* (Lb) and *L. brownii* var. *viridulum* (Lbv), each with a bulb diameter of only approximately 3 cm (Li et al., 2007; Wu et al., 2019b) (Figs. 2a–2f). All three of these wild lilies are diploid with relatively similar genetic backgrounds, and their classification at the molecular level was confirmed in our previous studies (Du et al., 2015, 2017). Therefore, we considered Lbg as a natural mutant due to the difference in bulb size trait to explore bulb development issues. However, whether starch metabolism has a special role in Lbg is

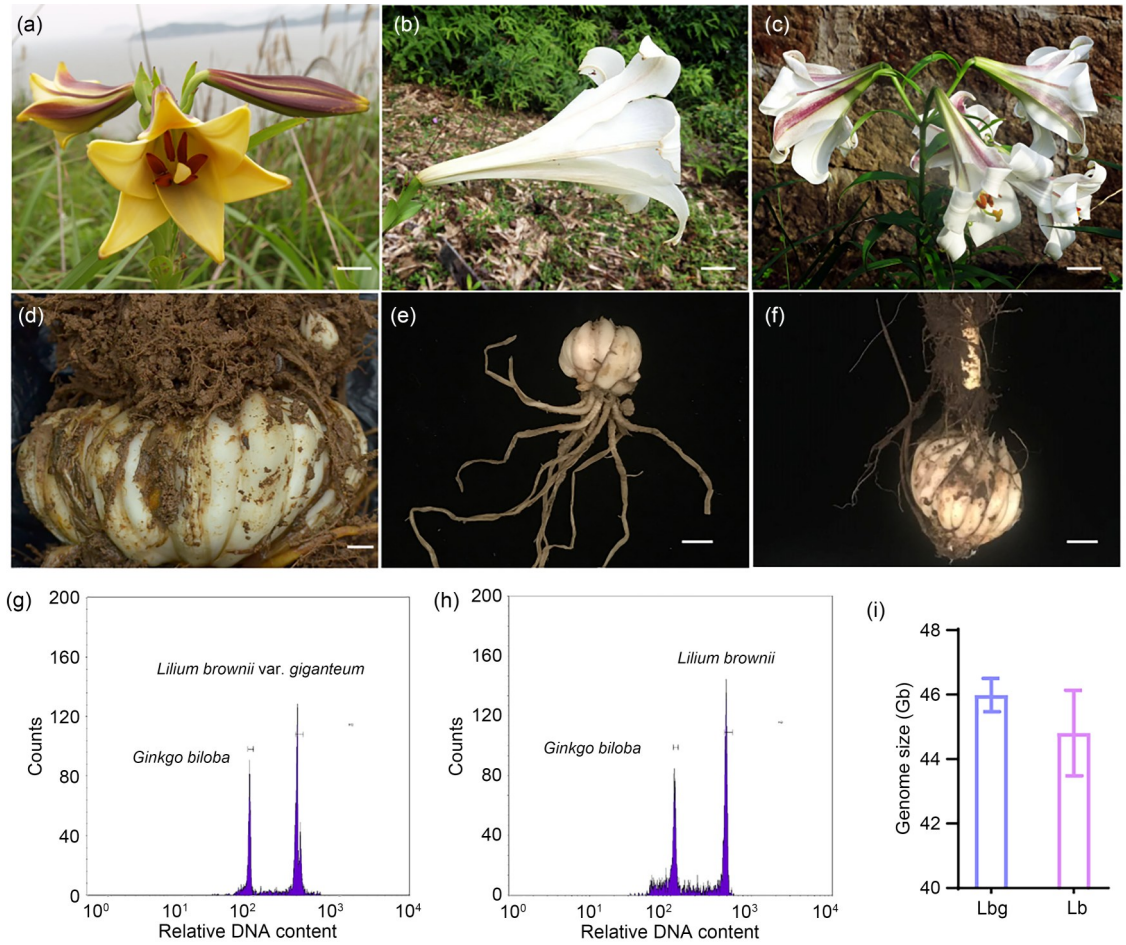

Fig. 2 Characteristics of *Lilium brownii* var. *giganteum* (Lbg) and its close wild relatives. Different flower performance under natural condition for Lbg (a), *L. brownii* (Lb) (b), *L. brownii* var. *viridulum* (Lbv) (c). Different bulb sizes of flowering plants for Lbg (d), Lb (e), and Lbv (f). (g, h) Histograms of the relative flow-cytometric genome size measurement. (i) Comparison of genome size between Lbg and Lb. Data are represented as mean±standard deviation (SD) (n=3 biological replicates). No significance was observed at P<0.05 according to Duncan's multiple range test. White bar=1 cm.

unclear. As the rapid resource degradation in the natural habitat for Lbg, we have established tissue culture procedures (Wu et al., 2019b). In this study, based on our data on starch features in different lilies and our unpublished RNA-seq data (Illumina and PacBio), we used rapid amplification of 5' and 3' complementary DNA (cDNA) ends (RACE) technology to isolate three major genes in the starch synthesis pathway in Lbg using in vitro tissues. Then, bioinformatics analysis and spatiotemporal expression profiling were performed and the results were compared. During this process, some single nucleotide polymorphism (SNP) sites were discovered.

2 Materials and methods

2.1 Preparation of different tissues and bulblet developmental stages

The initial plantlet of Lbg was obtained using the protocol described by Wu et al. (2019b) and subcultured in bulbing medium (BM) with Murashige and Skoog (MS) medium (Murashige and Skoog, 1962) as the basal medium containing 70 g/L sucrose and 3 g/L Phytagel (Sigma-Aldrich Chemical Company, St. Louis, MO, USA) to develop 10-mm diameter sterile mature bulblets. These mature plantlets were divided into four parts (leaf, petiole, bulblet, and root) to test the tissue-specific gene expression profile and were used to quantify the non-structure carbohydrate (NSC) content. The outer scales of the mature bulblets were detached and inoculated in induction medium (IM; MS+1.0 mg/L benzylaminopurine+0.1 mg/L naphthaleneacetic acid+30 g/L sucrose+3 g/L phytagel) to obtain primary shoots as mentioned before (Wu et al., 2019b). The primary shoots underwent a bulbification process in BM and three distinct developmental stages were derived: Stage 1 (S1, the initial shoot status), Stage 2 (S2, the early bulblet formation stage), and Stage 3 (S3, the later bulblet developmental stage). Bulblets of Lb were acquired in BM as well. Samples were pooled together from at least 10 plantlets as one biological replicate and three independent biological replications were included in current study. All the cultures were incubated at (24±2) ℃ under a 12-h light/dark photoperiod with a photosynthetic photon flux density (PPFD) of 110 μmol/(m^2·s) from light-emitting diode (LED) tubes.

2.2 Genome size estimation

Nuclei were prepared by sampling 50 mg fresh young leaves (plantlets with 10-mm diameter of bulblet) of Lbg and Lb in extraction buffer as described previously (Doležel et al., 2007). The leaf was cut with a razor blade and incubated for 10 min on ice. The homogenate was filtered through a 30-μm mesh filter and mixed with 50 μg/mL propidium. Flow cytometry measurement was taken by BD FACSCalibur™ platform (Becton Dickinson, San Jose, CA, USA) equipped with a 488-nm laser canon. To estimate the genome size (GS) of the investigated taxa, three individuals were analyzed. A DNA content of 10 000 stained nuclei was determined for each sample. The relative fluorescence histograms were analyzed on ModFit 2.3 software (Verity Software House, Topsham, ME, USA). As the GS of lily is very large (Shahin et al., 2012), the GS of *Ginkgo biloba* (1 C=10.61 Gb) was used as an internal standard (Guan et al., 2016). Based on the peak of sample and internal standard, experimental GS was calculated following equation: sample GS= (sample peak mean/standard peak mean)×standard GS.

2.3 Non-structure carbohydrate content assays

The NSC in this study included sucrose, total soluble sugar (TSS), and starch. A total of 0.5 g (fresh weight) samples of bulblets (ca. 10 mm in diameter) were used for extraction by the modified anthrone method as described by Wu et al. (2019a). Samples were detected by a spectrophotometric analysis with an Enspire™ 2300 Multilabel Reader (Perkin Elmer Corporation, MA, USA) at 620 nm.

2.4 Cytological observation and transmission electron microscopy

Histological observations were determined by the modified periodic acid-Schiff (PAS) method as described by Wu et al. (2017). The outer scales of in vitro bulblets (ca. 10 mm in diameter) for two lilies were fixed, dehydrated, and embedded. Semithin 12-μm sections were cut by a microtome (Leica RM2016, Leica Microsystems, Heidelberg, Germany) and mounted onto glass slides. The PAS Kit (G1008) was purchased from Wuhan Goodbio Technology Co., Ltd. (Wuhan, China), and the sections were examined under a Nikon ECLIPSE Ci microscope (Japan) and photographed.

The basal scale tissues (2−3 mm in length) were selected for transmission electron microscopy (TEM)

analysis as in the previous study in *Lycoris* (Wu et al., 2017) and were fixed in 4% (volume fraction) glutaraldehyde in 0.1 mol/L phosphate-buffered saline (PBS; pH 7.0) for 6 h, and post-fixed in 1% OsO_4 for 1 h. Subsequently, dehydration was conducted by a degressive ethanol series, and embedded in Spurr's resin. Ultrathin sections (80 nm) were prepared and mounted on copper grids for viewing with a transmission electron microscope (TEM-1200EX, JEOL, Tokyo, Japan).

2.5 Isolation of three key genes encoding starch synthetic pathways

Total RNA was extracted by the EASYspin Plus Plant RNA kit (RN38, Aidlab Bio, Beijing, China) and treated with DNase I (4716728001, Roche, Basel, Switzerland) according to the manufacturer's protocol. Integrity and purity of RNA were determined by agarose gel electrophoresis and ultraviolet spectroscopy (NanoDrop 2000, Thermo Scientific, Madison, WI, USA). Twelve primers were designed based on selected unigenes encoding AGPase, GBSS, and SS proteins from our unpublished transcriptome data (Table S1). cDNA synthesis was performed according to SMARTer™ RACE cDNA Amplification Kit (634858, Clontech, Mountain View, CA, USA), and nested polymerase chain reaction (PCR) was used in obtaining the 5' and 3' ends of the cDNAs. At least, three positive clones from ampicillin-containing media were selected and sequenced by Sangon Biotech (Shanghai, China). Based on the merged 5' RACE and 3' RACE sequencing results, the forward and reverse primers were designed to amply the entire open reading frame (ORF) from the cDNA.

2.6 Sequence bioinformatics and phylogenetic analysis

The ORFs were confirmed by BLASTp and ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). The molecular weight and theoretical isoelectric point were predicted by ExPASy analysis system (https://web.expasy.org/protparam). The membrane protein was predicted by TMHMM 2.0 (http://www.cbs.dtu.dk/services/TMHMM) (Krogh et al., 2001). The three-dimensional (3D)-structural model of protein is predicted by homology-modeling with SWISS-MODE (https://swissmodel.expasy.org) (Arnold et al., 2006) and secondary structures by SOPMA (Deléage, 2017). The subcellular localization of the deduced polypeptides was predicted by Cell-PLoc 2.0 (http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2) (Chou and Shen, 2010). The conserved domain of three proteins was anticipated by SMART (http://smart.embl.de) (Letunic and Bork, 2018).

The homologous genes in Lb were searched from our unpublished PacBio transcriptome data. Protein sequences of *AGPase* and *SS* gene families in *A. thaliana*, *Solanum tuberosum*, and *Oryza sativa* were downloaded from the PLAZA database (Version 4.5; https://bioinformatics.psb.ugent.be/plaza) to verify the classification of cloned genes. Moreover, to elucidate the phylogenetic relationship within each protein, sequences from at least 15 species were downloaded from the National Center for Biotechnology Information Database (NCBI; http://www.ncbi.nlm.nih.gov). All the sequences were aligned by Clustal-X 1.8 (Thompson et al., 1997) and Neighbor-Joining (NJ) trees were subsequently constructed in MEGA (Version 10.1.8) by bootstrap analysis with 1000 replications (Kumar et al., 2016). The tree was collapsed and formatted by iTOL v4 (Letunic and Bork, 2019). Sequence alignment and calculation of identity were conducted by DNAMAN software (Version 9.0.1.116).

2.7 Expression pattern of the key starch synthesis genes in various tissues and developmental stages

Quantitative PCR primers were designed according to the full-length cDNA sequences of three genes (Table S2). The cDNA synthesis from extracted RNA was transcribed by Prime Script kit (RR047A, TaKaRa, Dalian, China) and the reactions were performed by TB Green® Premix Ex Taq™ (RR420A, TaKaRa). The signal was monitored by the CFX Connect™ Real-Time PCR Detection System (Bio-Rad, CA, USA). The average quantification cycle (C_q) value was calculated from three biological and triplicate technical runs. Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) was selected as the reference gene for normalization. The relative gene expression levels were analyzed by the $2^{-\Delta\Delta C_T}$ method (Livak and Schmittgen, 2001) and plotted with GraphPad Prism 8.0 software (CA, USA). Statistical analyses were conducted by one-way analysis of variance (ANOVA) in IBM SPSS Statistics 20.0 software (NY, USA) with Duncan's multiple range test (DMRT).

2.8 Characterization of genetic polymorphism of *AGPS* cDNA in Lb and its varieties

Based on our data and released sequence for Lbv in NCBI, we carried out the sequence alignment by Clustal-X 1.8 to find the potential SNPs. The coding sequences (CDSs) of *AGPS* gene of four genotypes were used for estimating non-synonymous/synonymous (Ka/Ks) values. The sequences were firstly aligned by MEGA and then calculated by the KaKs_Calculator (Version 2.0) (Wang et al., 2010).

3 Results

3.1 Genome size prediction and starch feature characterization

Since the extremely large bulb size of Lbg compared to its close relatives Lb and Lbv, we examined the GS by flow cytometry to determine whether this specific trait has any association with GS expansion. The results showed that the GS of Lbg ((45.98±0.89) Gb) was slightly larger than that of Lb ((44.81±2.30) Gb), and the difference was not significant (Figs. 2g–2i). The NSC content assays revealed that more starch accumulated in Lb than in Lbg. However, the sugars, including sucrose and TSS, exhibited an opposite trend in Lbg and Lb (Fig. 3a). Further histological and ultrastructure observation of the outer scales of same size bulblets showed that most of the parenchyma cells and starch grains (SGs) were smaller but the number of SG was greater in Lbg than in Lb, indicating a potential difference in starch formation between the two species (Figs. 3b–3g). However, whether the difference is related to the bulb size remains unclear.

3.2 Cloning and sequence analysis of key genes in the starch synthesis pathway

The full-length cDNA sequence was amplified by RACE and the intermediate products are shown in Fig. S1. All three genes were deposited in GenBank with the accession numbers MT740689, MT740690, and MT740691. The complete ORFs were 1569-, 1818-, and 3576-nt long, encoding proteins with 522, 605, and 1191 amino acids (Figs. 4a–4c, and S2–S4), with molecular weights of 56.98, 66.56, and 135.07 kDa and isoelectric points of 6.53, 6.26, and 5.54, respectively (Table 1). Additionally, all three proteins are located in chloroplasts, as predicted by Cell-PLoc 2.0, which was consistent with previous reports indicating plastid locations (Smith and Zeeman, 2020). In addition, all three predicted proteins did not belong to transmembrane ones. The deduced proteins were classified by comparing whole gene families from model and non-model

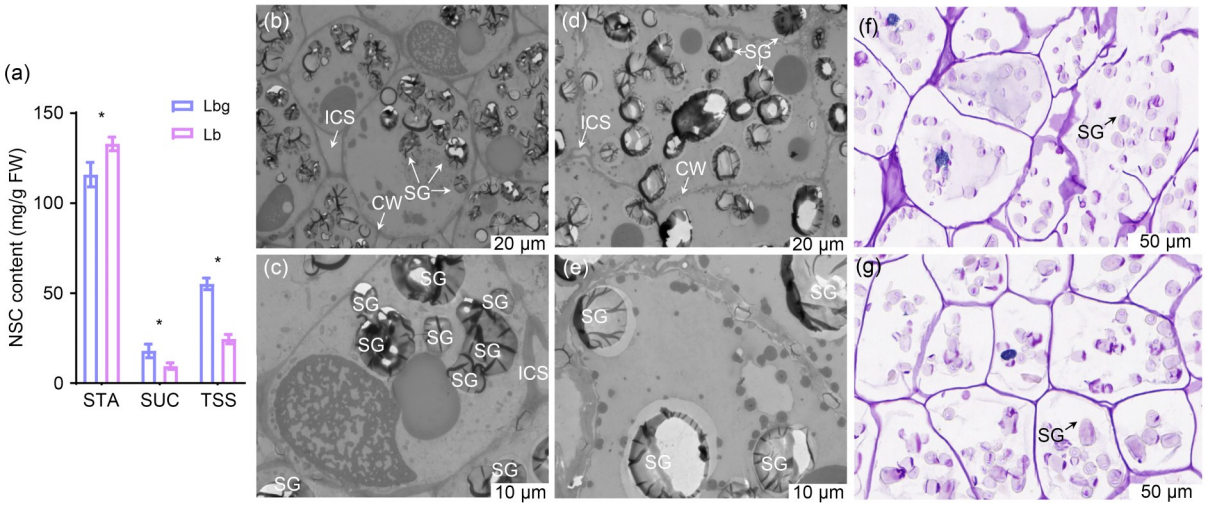

Fig. 3 Comparison of starch content and characteristics between *Lilium brownii* var. *giganteum* (Lbg) and *L. brownii* (Lb) in bulblet (around 10 mm). (a) Data of non-structure carbohydrate (NSC) content are represented as mean± standard deviation (SD) (n=3 biological replicates). * P<0.05 indicates significant difference between Lbg and Lb within each physiological trait (Duncan's multiple range test). (b–e) Transmission electron micrographs of the starch granules in the outer scale tissues of Lbg (b, c) and Lb (d, e). (f, g) Histological analysis of the outer scale tissues of Lbg (f) and Lb (g). FW, fresh weight; STA, starch; SUC, sucrose; TSS, total soluble sugar; ICS, intercellular space; SG, starch grain; CW, cell wall.

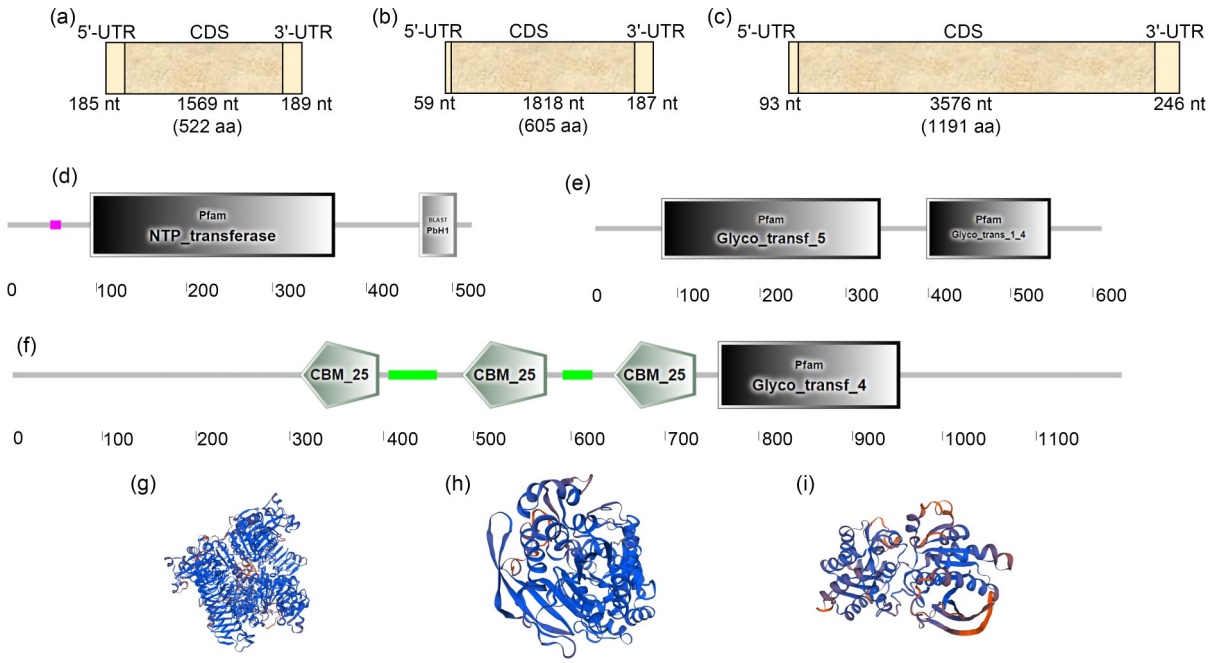

Fig. 4 Structures of LbgAGPS, LbgGBSS, and LbgSSIII. (a–c) Putative complementary DNA (cDNA) sequence structure of LbgAGPS (a), LbgGBSS (b), and LbgSSIII (c); (d–f) Conserved domain prediction for the protein encoded by LbgAGPS (d), LbgGBSS (e), and LbgSSIII (f); (g–i) Protein tertiary structure of LbgAGPS (g), LbgGBSS (h), and LbgSSIII (i). UTR, untranslated region; CDS, coding sequence.

Table 1 Analysis of protein physical and chemical properties

Protein	MW (kDa)	pI	TNNCR/ TNPCR	GRAVY	II	Predicted location	Secondary structure (%)			
							Random coil	Extended strand	α-Helix	β-Turn
LbgAGPS	56.98	6.53	60/58	−0.164	37.65	Chloroplast	49.23	19.92	24.90	5.94
LbgGBSS	66.56	6.26	68/64	−0.093	27.55	Chloroplast	42.31	13.88	37.02	6.78
LbgSSIII	135.07	5.54	177/143	−0.536	45.66	Chloroplast	42.65	17.30	34.17	5.88

MW, molecular weight; pI, isoelectric point; TNNCR, total number of negatively charged residues (Asp+Glu); TNPCR, total number of positively charged residues (Arg+Lys); GRAVY, grand average of hydropathicity; II, instability index.

plants, including monocotyledons and dicotyledons, and they were named LbgAGPS, LbgGBSS, and LbgSSIII (Fig. 5). The deduced proteins were analyzed and canonical conserved domains were identified, including NTP_transferase (PF00483) in LbgAGPS, Glyco_transf_5 (PF08323) and Glyco_trans_1_4 (PF13692) in LbgGBSS, and CBM_25 (PF03423) and Glyco_transf_4 (PF13439) in LbgSSIII (Figs. 4d−4f). The secondary structure compositions of the three proteins were similar and dominated by random coils, extended strands, and α-helixes (Table 1). The 3D structures of the cloned genes were predicted by the SWISS-MODEL platform. The resolution was determined to be 2.1 Å (1 Å=10^{-10} m), and the protein with the closest sequence identity (93.11%) was potato tuber AGPase, with an oligo state determined to be a homo-tetramer. The sequence identity of LbgGBSS with rice was only 70.40% and that of LbgSSIII with *Arabidopsis* was 51.38%. All the predicted 3D structures are shown in Figs. 4g–4i.

To illuminate the phylogenetic relationship among cloned proteins, NJ trees were constructed using complete protein sequences from various monocots and dicots. The analysis revealed that four AGPSs from *Lilium* were closely clustered and similar to that in *Gladiolus* hybrids, also a flowering bulb. The similarity was as high as 99.57% according to a DNAMAN sequence alignment determination (Figs. 6a and S5a). For GBSS and SSIII, sequences from *Lilium* joined

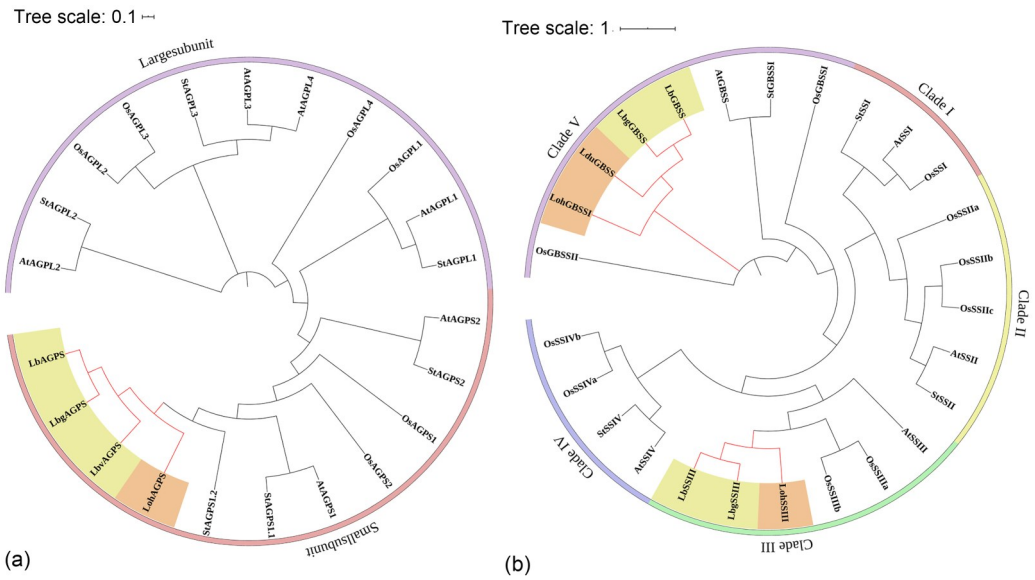

Fig. 5 Classification of deduced amino acids of adenosine-5'-diphosphate-glucose pyrophosphorylase (AGPase) and soluble starch synthase (SS) by comparing with gene families from *Arabidopsis thaliana*, *Oryza sativa*, and *Solanum tuberosum*. (a) AGPase; (b) SS. AGPS, small subunit of AGPase; AGPL, large subunit of AGPase; GBSS, granule-bound starch synthase. The clade line in red indicates the proteins from *Lilium* (Note: for interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

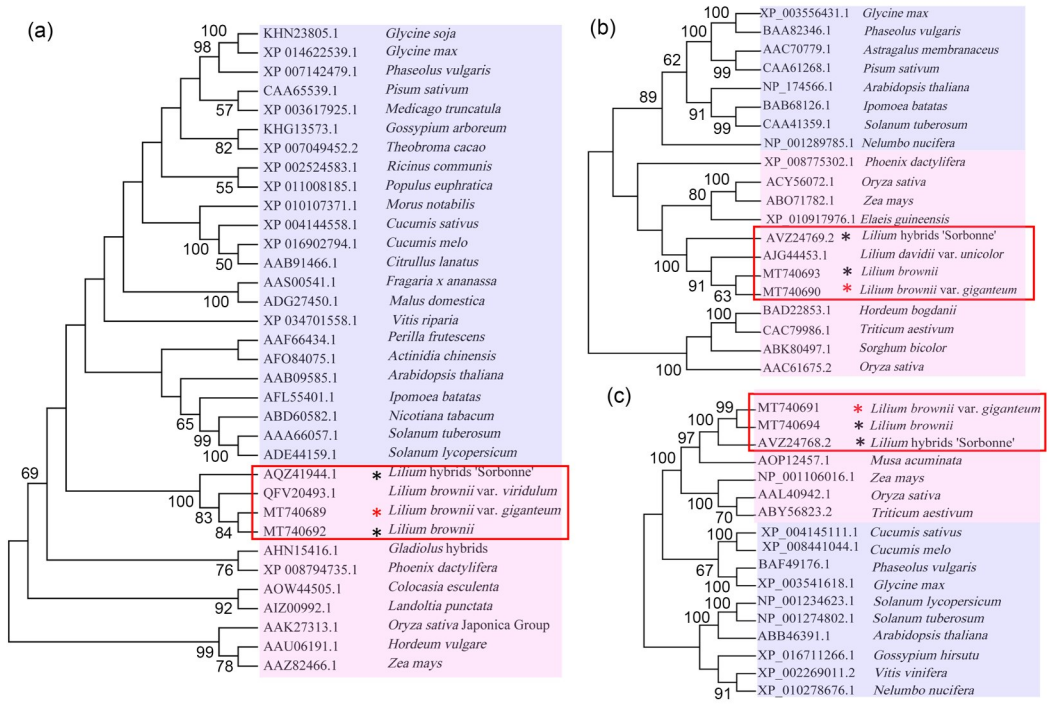

Fig. 6 Phylogenetic trees of the predicted amino acids: (a) AGPS; (b) GBSS; (c) SS. Proteins shaded with light blue and light purple are sequences from dicotyledons and monocotyledons, respectively. The red box indicates the sequences from *Lilium*. Asterisk indicates the sequence from our group. The red asterisk indicates the sequence cloned by the RACE in the present study. AGPS, small subunit of adenosine-5'-diphosphate-glucose pyrophosphorylase (AGPase); GBSS, granule-bound starch synthase; SS, soluble starch synthases; RACE, rapid amplification of 5' and 3' complementary DNA (cDNA) ends (Note: for interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

together, and both bootstrap values were 100, indicating high reliability. Furthermore, the dicots and monocots were divided into two clear subclades (Figs. 6b and 6c). Whereas, the divergence in SSIII was greater than that in GBSS in lily (Figs. S5b and S5c). Taken together, the isolated genes were predicted to encode an *AGPS* and two starch synthase genes which were highly conserved among different lily species (cultivar).

3.3 Spatiotemporal-specific expression pattern analysis

As shown in Fig. 3, the starch content and properties differed between Lbg and Lb. We therefore cloned three starch synthesis genes by RACE. To further exploit the expression pattern of those genes at the transcriptional level, we used quantitative real-time PCR (qRT-PCR) to determine the relative expression levels. For various tissues, differential expression was observed, with bulblet possessing the highest abundance of transcripts for all the tested genes (Fig. 7a). The patterns in *LbgAGPS* and *LbgGBSS* were very similar, and the difference between the highest and lowest transcript expression levels was remarkable; for example, the fold change for *LbgGBSS* in bulblet was 39 times that that in leaf (Figs. 7b and 7c). In contrast, *LbgSSIII* expression difference was only 1.5-fold between bulblet and leaf (Fig. 7d). We speculate that the bulblet is a major starch enrichment pool, while relatively low starch accumulates in the root, according to the tissue expression profile.

To investigate the transcript expression levels during the bulblet developmental process, three time points were chosen to represent the three stages: 0, 15, and 35 DAT (days after transplanting) (Figs. 8a and

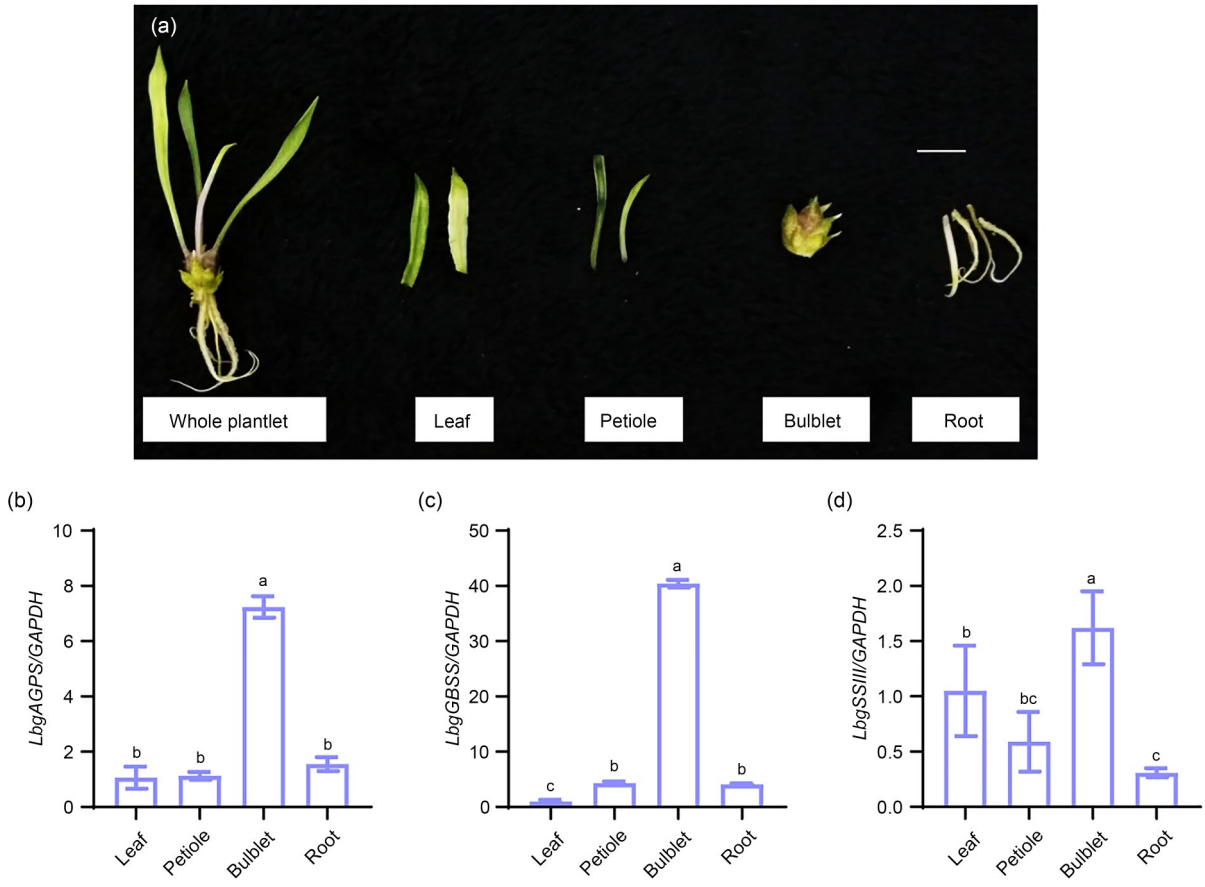

Fig. 7 Relative expression levels of cloned genes in various tissues in vitro of *Lilium brownii* var. *giganteum* (Lbg). (a) Different tissues, white bar=1 cm; (b–d) Expression of *LbgAGPS* (b), *LbgGBSS* (c), and *LbgSSIII* (d). The relative expression level represents the fold changes as compared with the expression level in the leaf tissue. Data are presented as mean±standard deviation (SD) (*n*=3 biological replicates). The different lowercase letters indicate significant difference at *P*<0.05 according to Duncan's multiple range test. *GAPDH*, glyceraldehyde-3-phosphate dehydrogenase.

8b). Overall, the results of the qRT-PCR analysis demonstrated upward trends for the three major starch synthesis genes in both species as the bulblet enlarged. However, the extent of the increase differed between Lbg and Lb. For instance, the abundance of *LbgAGPS* was elevated significantly in S1, S2, and S3. Notably, the relative expression levels of the *AGPS* and *GBSS* genes in Lb were lower than those in Lbg. Additionally, the fold changes of these two genes in both lilies were higher compared to those of *SSIII*, which is consistent with the results of tissue-specific expression, demonstrating that these genes might be more active in vitro during bulbification (Figs. 8c–8e).

3.4 Characterization of genetic diversity in the *AGPS* gene

SNPs are the most common forms of DNA sequence variation between alleles, and some are closely associated with economically important traits (Mammadov et al., 2012). Therefore, we identified SNPs using the *AGPS* gene as the referent sequence, as four complete CDSs were available at the time of the current study, including those for the three species of interest. In total, 41 SNPs were found within these four different lilies, most of which were synonymous SNPs (78%); that is, they did not affect the protein sequence. Notably, all the SNPs located in the NTP_transferase domain were synonymous SNPs (shade), indicating the high conservation of this typical feature. Moreover, the SNPs found only in *Lilium* Oriental hybrids 'Sorbonne' (Loh) accounted for 68% of all the SNPs identified, which aligns with previous identification in this pedigree (Du et al., 2015). The ratio of non-synonymous/synonymous SNPs was 0.28 (Table 2). The average Ka/Ks was 0.1587 for the four lilies, suggesting a purifying selection of this functional gene (Fig. 9).

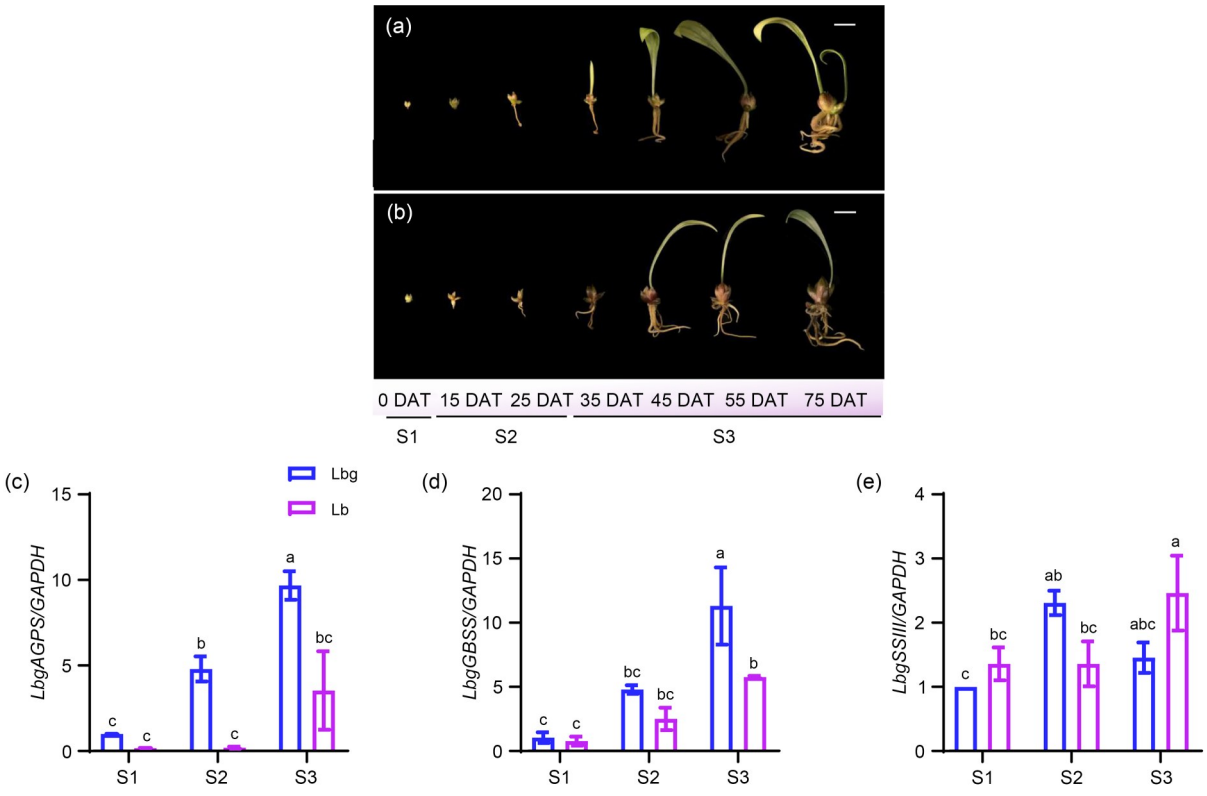

Fig. 8 Comparison of relative expression levels of cloned genes in various bulblet developmental stages between *Lilium brownii* var. *giganteum* (Lbg) and *L. brownii* (Lb) in vitro. (a, b) Whole bulblet developmental process of Lbg (a) and Lb (b), white bar=1 cm. S1, the initial shoot status; S2, the early bulblet formation stage; S3, the later bulblet developmental stage. (c–e) Expression of *LbgAGPS* (c), *LbgGBSS* (d), and *LbgSSIII* (e). The relative expression level represents the fold change as compared with the expression level at the S1 of Lbg. Data are presented as mean±standard deviation (SD) (n=3 biological replicates). The different lowercase letters indicate significant difference at P<0.05 according to Duncan's multiple range test. DAT, days after transplanting; *GAPDH*, glyceraldehyde-3-phosphate dehydrogenase.

Table 2 Sequence polymorphisms of four *AGPS* genes in *Lilium*

SNP	*AGPS* gene				Mutation of amino acid	SNP	*AGPS* gene				Mutation of amino acid
	Loh	Lbv	Lbg	Lb			Loh	Lbv	Lbg	Lb	
c.42 C>T	C	C	C	T	SM, R	c.927 T>C	T	T	T	C	SM, D
c.66 C>T	C	C	T	T	SM, P	c.1080 T>C	**C**	T	T	T	SM, N
c.69 G>A	G	A	A	A	SM, K	c.1152 G>A	**A**	G	G	G	SM, T
c.73 C>T	**T**	C	C	C	NSM, S/P	c.1200 G>A	**A**	G	G	G	SM, T
c.80 A>G	A	G	A	A	NSM, Q/R	c.1224 C>T	**T**	C	C	C	SM, C
c.89 A>G	**A**	G	G	G	NSM, K/R	c.1242 G>A	**G**	A	A	A	SM, K
c.121 G>T	G	T	G	G	NSM, A/S	c.1269 T>A	T	T	T	A	SM, R
c.210 G>T	G	G	G	T	SM, S	c.1299 G>A	**A**	G	G	G	SM, E
c.234 A>G	G	A	A	G	SM, K	c.1309 C>T	**T**	C	C	C	SM, L
c.423 A>G	**G**	A	A	A	SM, L	c.1353 C>T/A	A	C	C	T	NSM, L/F
c.348 G>A	**A**	G	G	G	SM, A	c.1358 A>C	**C**	A	A	A	NSM, S/Y
c.387 G>A	A	G	A	A	SM, P	c.1359 T>C	**C**	T	T	T	NSM, S/Y
c.447 C>T	**T**	C	C	C	SM, S	c.1388 C>T	**T**	C	C	C	SM, G
c.465 C>T	**T**	C	C	C	SM, H	c.1408 A>G	**A**	G	A	A	NSM, R/G
c.471 A>G	**G**	A	A	A	SM, S	c.1443 G>C	**G**	C	C	C	NSM, E/D
c.519 C>T	**T**	C	C	C	SM, V	c.1464 T>C	**T**	C	C	C	SM, T
c.528 C>T	**C**	T	T	T	SM, L	c.1485 G>A	**A**	G	G	G	SM, A
c.744 C>T	**C**	T	T	T	SM, R	c.1524 G>C	**C**	G	G	G	SM, V
c.820 C>A	**C**	A	A	A	SM, Q/K	c.1527 T>C	**C**	T	T	T	SM, T
c.894 T>C	T	T	T	C	SM, A	c.1554 T>C	**C**	T	T	T	SM, S
c.897 G>A	G	G	G	A	SM, S						

Loh, *Lilium* Oriental hybrids 'Sorbonne'; Lbv, *Lilium brownii* var. *viridulum*; Lbg, *L. brownii* var. *giganteum*; Lb, *L. brownii*; SM, synonymous mutation; NSM, non-synonymous mutation. Nucleotide in bold represents that the difference in single nucleotide occurred only in Loh. The row shaded with grey represents the SNPs within the NTP_transferase domain.

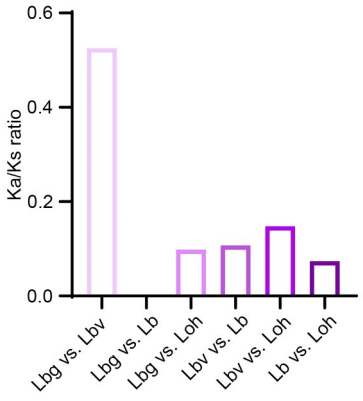

Fig. 9 Ratio of non-synonymous substitutions to synonymous substitutions (Ka/Ks) for *AGPS* gene in four different lilies. Lbg, *Lilium brownii* var. *giganteum*; Lbv, *L. brownii* var. *viridulum*; Lb, *L. brownii*; Loh, *Lilium* Oriental hybrids 'Sorbonne.' *AGPS*, small subunit of adenosine-5'-diphosphate-glucose pyrophosphorylase (AGPase).

4 Discussion

Lilies are among the top 10 commercial flowers worldwide, and the underground bulbs are crucial to the aboveground flower quality (Thakur et al., 2006; Xu et al., 2020). However, the bulbs need a rather long vegetative growth period (normally 2–3 years) to reach the size necessary for flowering, which is a main factor limiting bulb production and the introduction of newly bred cultivars (Askari et al., 2018). The lily bulb is mainly composed of fleshy scales with starch as the primary storage compound (approximately 85% in dry matter) (Matsuo and Mizuno, 1974). Starch is therefore shown to be highly associated with bulb development in different study models, i.e., in vitro bulbification in 'Sorbonne' (Wu et al., 2019a),

scaling in *Lilium davidii* var. *unicolor* (Li et al., 2014), and bulbil formation in *Lilium lancifolium* (Yang et al., 2017). Despite the great importance of starch, no experimental research is available to date regarding the properties of key starch synthesis genes in lily. In addition, in contrast to model plants, such as *Arabidopsis*, the lack of lily mutants is also a main restriction factor.

In the present study, a wild lily excellent for bulb development study, Lbg, was used as a natural mutant to compare with the native species, Lb, and close relative, *L. brownii* var. *viridulum*. We found that the starch content, as well as number and size of SGs differed between Lbg and Lb (Fig. 3), which raised the question, does starch synthesis played a critical role in the bulb development within this lily group, in which each species has a very similar genetic background? Three starch-synthesizing genes from Lbg, *LbgAGPS*, *LbgGBSS*, and *LbgSSIII*, were therefore cloned successfully by RACE (Figs. 4 and S2−S4). Bioinformatics analysis revealed that these genes contained sequences encoding the characteristic domains reported to date (Wang et al., 2014; Cheng et al., 2015; Zhu et al., 2020) (Figs. 4 and S2). For example, the catalytic domains for both LbgGBSS and LbgSSIII were found to belong to glycosyltransferase (GT) family 5 with two similar Rossmann-like subdomains (Lombard et al., 2014; Abt et al., 2020), while we found that only LbgSSIII has carbohydrate binding modules (CBMs) from family 53 (Abt et al., 2020). Additionally, we were able to show that the vast divergence in the deduced amino acids for all these genes was principally in the N-terminal region (not shown), which was consistent with the results obtained for lotus (Cheng et al., 2015). The N-terminus of AGPS was suggested to be important for its allosteric regulatory or catalytic properties (Bejar et al., 2006), and deletion or insertion in this region probably leads to alteration in the regulatory function of the AGPase (Ballicora et al., 1995; Linebarger et al., 2005). Further phylogenetic analysis demonstrated that *Lilium* clustered more closely to monocotyledons, which is in alignment with its taxa classification. Previously, Cheng et al. (2015) reported that AGPS sequences can be classified mainly into two subclades, dicots and monocots, with only one exception, from *Zea mays*; this result is slightly different from our clustering pattern. We speculated that this difference might be attributed to the difference in species selection, as all species in the Cheng et al. (2015) are from Poales. Notably, available sequence data for monocots are relatively limited, and a systematic illustration of the evolution of different starch synthesis-related genes is expected to be achieved through genome-wide analysis in the future (Ma et al., 2019). Previous identification and prediction results showed that these three proteins are plastid proteins (Beckles et al., 2001; Smith and Zeeman, 2020), and our prediction corroborated this finding (Table 1). For instance, the evidence shows the existence of a transit peptide that targets the plastid in AGPS (Fig. S2). Collectively, we are the first to report a comprehensive RACE cloning and sequence analysis of genes encoding key starch synthesis enzymes.

It has been demonstrated that compared with that in the non-photosynthetic rhizome tissue in *N. nucifera*, *NnAGPS1* showed significantly higher expression in photosynthetic tissues, including leaves and leafstalks (Cheng et al., 2015). However, discrepancies were found in the current study, with the highest expression of all these genes in the bulblet (Fig. 7), which is a major sink tissue under in vitro conditions. This result is probably due to the unigene sequences we used being derived from transcriptomic data of bulblet initiation and development biological processes. As more than two isoforms existed for these genes (Smith and Zeeman, 2020), we still do not know whether any genes are expressed at higher levels and/or expressed only in autotrophic tissues, such as leaves. These questions will likely be able to be addressed when the full and likely very large lily genome is released (Shahin et al., 2012) (Figs. 2g−2i). The starch content in bulblets increased gradually during the developmental process in lily, especially at the later stage (Li et al., 2014; Yang et al., 2017), as it does in other species with storage organs, e.g., *Sagittaria sagittifolia* (Gao et al., 2018) and *N. nucifera* (Yang et al., 2015). In this study, we found that the expression trends of *LbgAGPS* and *LbgGBSS* were consistent with starch accumulation during bulblet growth in both Lbg and Lb. By contrast, the transcript level of *LbgSSIII* peaked at the early bulblet formation stage (S2) in Lbg, while it increased to the highest level in Lb at the S3 stage (Fig. 8). Generally, the results indicated that major starch synthesis genes were positively correlated with starch enrichment during bulbification process. In addition, we suggest that transient starch may be predominant in Lbg because of the combined evidence from the qRT-PCR results and histological

observations (Fig. 3), which possibly leads to increased starch degradation during bulblet development to supply energy for scale initiation, which eventually results in a larger bulblet size in Lbg. However, the detailed mechanism needs further investigation in the future.

Currently, molecular markers are widely used in *Lilium* for a wide variety of scientific tasks such as evaluating genetic diversity, identifying cultivars, and constructing linkage map (Huang et al., 2009; Bakhshaie et al., 2016). Our group also reported 70 validated and transferable simple sequence repeats (SSRs) to establish parentage and taxon identity of landraces (Du et al., 2015). SNPs have been intensively studied in recent years due to their various advantages (Ganal et al., 2009). Nevertheless, few effective SNPs have been identified and validated, mainly due to the lack of genomic resources in *Lilium* (Shahin et al., 2012; Lee et al., 2016; van Tuyl et al., 2018). In this study, we found 41 SNPs according to the CDS of the *AGPS* homology among the four lily genotypes. The analysis of their interspecific relationships revealed higher genomic diversity in Loh, which agrees with our former classification generated by SSR markers (Du et al., 2015). Ka/Ks analysis confirmed that the AGPS was subjected to be a purifying selection, which theoretically eliminates harmful mutations in the population (Oteo-García and Oteo, 2019). Another starch synthesis gene, *NsSSSIII*, also showed evidence for a purifying selection (Zhu et al., 2020). However, to determine phenotypes, we will need to further analyze starch content and other bulb characteristics. Additionally, these potential SNPs need to undergo experimental genotyping to reveal those that are actually applicable ones.

5 Conclusions

In this study, we found differences in starch content and features according to physiological and histological observations of Lbg, a rare wild lily with a giant bulb, and the native species Lb. Then, we firstly identified three genes encoding major starch synthesis enzymes in Lbg based on our transcriptome data. Further gene expression profiling elaborated that the transcript levels of these genes showed obvious spatial- and temporal-preference patterns. Moreover, a set of SNPs in *AGPS* within the four lilies, namely Lbg, Lb, Lbv, and Loh, was creatively uncovered and awaits future verification and correlation analysis. In summary, our results will shed light on bulb initiation and development from the starch synthesis perspective using an elite wild germplasm.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 31772337), the Lishui Research Fund (No. 2020zdhz03), the China Postdoctoral Science Foundation (No. 2017M612005), and the National Key Research and Development Program of China (No. 2018YFD1000401). We gratefully thanks Tianyu HAN (Nanjing Forestry University, Nanjing, China) for his technical support.

Author contributions

Yiping XIA, Yun WU, and Ziming REN designed the experiments, wrote and edited the manuscript. Minyi SUN, Shiqi LI, Ruihan MIN, Cong GAO, and Qundan LYU performed the study and data analysis. All authors have read and approved the final version of the manuscript and, therefore, have full access to all the data in the study and take responsibility for the integrity and security of the data.

Compliance with ethics guidelines

Yun WU, Minyi SUN, Shiqi LI, Ruihan MIN, Cong GAO, Qundan LYU, Ziming REN, and Yiping XIA declare that they have no conflict of interest.

This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Abt MR, Pfister B, Sharma M, et al., 2020. STARCH SYNTHASE 5, a noncanonical starch synthase-like protein, promotes starch granule initiation in Arabidopsis. *Plant Cell*, 32(8):2543-2565. https://doi.org/10.1105/tpc.19.00946

Arnold K, Bordoli L, Kopp P, et al., 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, 22(2):195-201. https://doi.org/10.1093/bioinformatics/bti770

Askari N, Visser RGF, de Klerk GJ, 2018. Growth of lily bulblets in vitro, a review. *Int J Hortic Sci Technol*, 5(2):133-143. https://doi.org/10.22059/ijhst.2018.268870.263

Bahaji A, Li J, Sánchez-López ÁM, et al., 2014. Starch biosynthesis, its regulation and biotechnological approaches to improve crop yields. *Biotechnol Adv*, 32(1):87-106. https://doi.org/10.1016/j.biotechadv.2013.06.006

Bakhshaie M, Khosravi S, Azadi P, et al., 2016. Biotechnological advances in *Lilium*. *Plant Cell Rep*, 35(9):1799-1826. https://doi.org/10.1007/s00299-016-2017-8

Ballicora MA, Laughlin MJ, Fu Y, et al., 1995. Adenosine 5'-diphosphate-glucose pyrophosphorylase from potato tuber (significance of the N terminus of the small subunit for catalytic properties and heat stability). *Plant Physiol*,

109(1):245-251.
https://doi.org/10.1104/pp.109.1.245

Beckles DM, Craig J, Smith AM, 2001. ADP-glucose pyrophosphorylase is located in the plastid in developing tomato fruit. *Plant Physiol*, 126(1):261-266.
https://doi.org/10.1104/pp.126.1.261

Bejar CM, Ballicora MA, Iglesias AA, et al., 2006. ADPglucose pyrophosphorylase's N-terminus: structural role in allosteric regulation. *Biochem Biophys Res Commun*, 343(1): 216-221.
https://doi.org/10.1016/j.bbrc.2006.02.123

Cheng N, Zeng XF, Zheng XF, et al., 2015. Cloning and characterization of the genes encoding the small and large subunit of the ADP-glucose pyrophosphorylase in lotus (*Nelumbo nucifera* Gaertn). *Acta Physiolo Plant*, 37(1): 1734.
https://doi.org/10.1007/s11738-014-1734-2

Chou KC, Shen HB, 2010. Cell-PLoc 2.0: an improved package of web-servers for predicting subcellular localization of proteins in various organisms. *Nat Sci*, 2(10):1090-1103.
https://doi.org/10.4236/ns.2010.210136

de Klerk GJ, 2012. Micropropagation of bulbous crops: technology and present state. *Floricult Ornam Biotechnol*, 6: 1-8.

Deléage G, 2017. ALIGNSEC: viewing protein secondary structure predictions within large multiple sequence alignments. *Bioinformatics*, 33(24):3991-3992.
https://doi.org/10.1093/bioinformatics/btx521

Doležel J, Greilhuber J, Suda J, 2007. Estimation of nuclear DNA content in plants using flow cytometry. *Nat Protoc*, 2(9):2233-2244.
https://doi.org/10.1038/nprot.2007.310

Du F, Jiang J, Jia HM, et al., 2015. Selection of generally applicable SSR markers for evaluation of genetic diversity and identity in *Lilium*. *Biochem Syst Ecol*, 61:278-285.
https://doi.org/10.1016/j.bse.2015.05.002

Du HH, Yang T, Ma CY, et al., 2012. Effects of RNAi silencing of *SSIII* gene on phosphorus content and characteristics of starch in potato tubers. *J Integr Agric*, 11(12): 1985-1992.
https://doi.org/10.1016/S2095-3119(12)60455-9

Du YP, Bi Y, Zhang MF, et al., 2017. Genome size diversity in *Lilium* (Liliaceae) is correlated with karyotype and environmental traits. *Front Plant Sci*, 8:1303.
https://doi.org/10.3389/fpls.2017.01303

Ganal MW, Altmann T, Röder MS, 2009. SNP identification in crop plants. *Curr Opin Plant Bio*, 12(2):211-217.
https://doi.org/10.1016/j.pbi.2008.12.009

Gao MP, Zhang SW, Luo C, et al., 2018. Transcriptome analysis of starch and sucrose metabolism across bulb development in *Sagittaria sagittifolia*. *Gene*, 649:99-112.
https://doi.org/10.1016/j.gene.2018.01.075

Goren A, Ashlock D, Tetlow IJ, 2018. Starch formation inside plastids of higher plants. *Protoplasma*, 255(6):1855-1876.
https://doi.org/10.1007/s00709-018-1259-4

Guan R, Zhao YP, Zhang H, et al., 2016. Draft genome of the living fossil *Ginkgo biloba*. *Gigascience*, 5(1):49.
https://doi.org/10.1186/s13742-016-0154-1

Huang JH, Zhou RR, He D, et al., 2020. Rapid identification of *Lilium* species and polysaccharide contents based on near infrared spectroscopy and weighted partial least square method. *Int J Biol Macromol*, 154:182-187.
https://doi.org/10.1016/j.ijbiomac.2020.03.109

Huang YF, Yang MX, Zhang H, et al., 2009. Genetic diversity and genetic structure analysis of the natural populations of *Lilium brownii* from Guangdong, China. *Biochem Genet*, 47(7-8):503-510.
https://doi.org/10.1007/s10528-009-9258-y

Kamenetsky R, Okubo H, 2012. Ornamental Geophytes: from Basic Science to Sustainable Production. CRC Press, Boca Raton, Florida, USA, p.1-16.

Krogh A, Larsson B, von Heijne G, et al., 2001. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J Mol Biol*, 305(3): 567-580.
https://doi.org/10.1006/jmbi.2000.4315

Kumar S, Stecher G, Tamura K, 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol*, 33(7):1870-1874.
https://doi.org/10.1093/molbev/msw054

Lee GJ, Subburaj S, Kumari S, et al., 2016. Identification of single nucleotide polymorphism (SNP) markers associated with botrytis resistance in lily (*Lilium* spp.). *Hortscience*, 51:S320-S321.

Letunic I, Bork P, 2018. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res*, 46(D1): D493-D496.
https://doi.org/10.1093/nar/gkx922

Letunic I, Bork P, 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res*, 47(W1):W256-W259.
https://doi.org/10.1093/nar/gkz239

Li GY, Chen ZH, Yan FB, 2007. *Lilium brownin* var. giganteum: a new variety of *Lilium* from Wenling, Zhejiang. *J Zhejiang Forestry Coll*, 24(6):767-768 (in Chinese).
https://doi.org/10.3969/j.issn.2095-0756.2007.06.023

Li XY, Wang CX, Cheng JY, et al., 2014. Transcriptome analysis of carbohydrate metabolism during bulblet formation and development in *Lilium davidii* var. *unicolor*. *BMC Plant Biol*, 14:358.
https://doi.org/10.1186/s12870-014-0358-4

Linebarger CRL, Boehlein SK, Sewell AK, et al., 2005. Heat stability of maize endosperm ADP-glucose pyrophosphorylase is enhanced by insertion of a cysteine in the N terminus of the small subunit. *Plant Physiol*, 139(4):1625-1634.
https://doi.org/10.1104/pp.105.067637

Livak KJ, Schmittgen TD, 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods*, 25(4):402-408.
https://doi.org/10.1006/meth.2001.1262

Lloyd JR, Kossmann J, 2019. Starch trek: the search for yield. *Front Plant Sci*, 9:1930.
https://doi.org/10.3389/fpls.2018.01930

Lombard V, Ramulu HG, Drula E, et al., 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*, 42(D1):D490-D495. https://doi.org/10.1093/nar/gkt1178

Ma PP, Yuan Y, Shen QC, et al., 2019. Evolution and expression analysis of starch synthase gene families in *Saccharum spontaneum*. *Trop Plant Biol*, 12(3):158-173. https://doi.org/10.1007/s12042-019-09225-3

Mammadov J, Aggarwal R, Buyyarapu R, et al., 2012. SNP markers and their impact on plant breeding. *Int J Plant Genomics*, 2012:728398. https://doi.org/10.1155/2012/728398

Matsuo T, Mizuno T, 1974. Changes in the amounts of two kinds of reserve glucose-containing polysaccharides during germination of the Easter lity bulb. *Plant Cell Physiol*, 15(3):555-558. https://doi.org/10.1093/oxfordjournals.pcp.a075036

Miao HX, Sun PG, Liu WX, et al., 2014. Identification of genes encoding granule-bound starch synthase involved in amylose metabolism in banana fruit. *PLoS ONE*, 9(2): e88077. https://doi.org/10.1371/journal.pone.0088077

Mishra BP, Kumar R, Mohan A, et al., 2017. Conservation and divergence of *Starch Synthase III* genes of monocots and dicots. *PLoS ONE*, 12(12):e0189303. https://doi.org/10.1371/journal.pone.0189303

Moreno-Pachón N, 2017. Mechanisms of Vegetative Propagation in Bulbs: A Molecular Approach. PhD Thesis, Wageningen University, Wageningen, the Netherlands.

Murashige T, Skoog F, 1962. A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol Plant*, 15(3):473-497. https://doi.org/10.1111/j.1399-3054.1962.tb08052.x

Oteo-García G, Oteo JA, 2019. Evolutionary distances corrected for purifying selection and ancestral polymorphisms. *J Theor Biol*, 483:110004. https://doi.org/10.1016/j.jtbi.2019.110004

Podwyszyńska M, 2012. The mechanisms of in vitro storage organ formation in ornamental geophytes. *Floricult Ornam Biotechnol*, 6(1):9-23.

Qu JZ, Xu ST, Zhang ZQ, et al., 2018. Evolutionary, structural and expression analysis of core genes involved in starch synthesis. *Sci Rep*, 8:12736. https://doi.org/10.1038/s41598-018-30411-y

Seo SG, Bea SH, Jun BK, et al., 2015. Overexpression of *ADP-glucose pyrophosphorylase* (*IbAGPaseS*) affects expression of carbohydrate regulated genes in sweet potato [*Ipomoea batatas* (L.) Lam. cv. Yulmi]. *Genes Genom*, 37(7): 595-605. https://doi.org/10.1007/s13258-015-0289-y

Shahin A, van Kaauwen M, Esselink D, et al., 2012. Generation and analysis of expressed sequence tags in the extreme large genomes *Lilium* and *Tulipa*. *BMC Genomics*, 13:640. https://doi.org/10.1186/1471-2164-13-640

Smith AM, Zeeman SC, 2020. Starch: a flexible, adaptable carbon store coupled to plant growth. *Annu Rev Plant Biol*, 71:217-245. https://doi.org/10.1146/annurev-arplant-050718-100241

Thakur R, Sood A, Nagar PK, et al., 2006. Regulation of growth of *Lilium* plantlets in liquid medium by application of paclobutrazol or ancymidol, for its amenability in a bioreactor system: growth parameters. *Plant Cell Rep*, 25(5):382-391. https://doi.org/10.1007/s00299-005-0094-1

Thompson JD, Gibson TJ, Plewniak F, et al., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*, 25(24):4876-4882. https://doi.org/10.1093/nar/25.24.4876

van Harsselaar JK, Lorenz J, Senning M, et al., 2017. Genome-wide analysis of starch metabolism genes in potato (*Solanum tuberosum* L.). *BMC Genomics*, 18:37. https://doi.org/10.1186/s12864-016-3381-z

van Tuyl JM, Arens P, Ramanna MS, et al., 2011. *Lilium. In*: Kole C (Ed.), Wild Crop Relatives: Genomic and Breeding Resources. Springer, Berlin Heidelberg, Germany, p.161-183. https://doi.org/10.1007/978-3-642-21201-7_9

van Tuyl JM, Arens P, Shahin A, et al., 2018. *Lilium. In*: van Huylenbroeck J (Ed.), Ornamental Crops. Springer, Cham, Switzerland, p.481-512. https://doi.org/10.1007/978-3-319-90698-0_20

Wang DP, Zhang YB, Zhang Z, et al., 2010. KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genom Proteom Bioinf*, 8(1): 77-80. https://doi.org/10.1016/s1672-0229(10)60008-3

Wang X, Feng B, Xu ZB, et al., 2014. Identification and characterization of granule bound starch synthase I (*GBSSI*) gene of tartary buckwheat (*Fagopyrum tataricum* Gaertn.). *Gene*, 534(2):229-235. https://doi.org/10.1016/j.gene.2013.10.053

Wu Y, Xia YP, Zhang JP, et al., 2016. Low humic acids promote in vitro lily bulblet enlargement by enhancing roots growth and carbohydrate metabolism. *J Zhejiang Univ-Sci B (Biomed & Biotechnol)*, 17(11):892-904. https://doi.org/10.1631/jzus.B1600231

Wu Y, Li Y, Ma YD, et al., 2017. Hormone and antioxidant responses of *Lilium* Oriental hybrids 'Sorbonne' bulblets to humic acid treatments *in vitro*. *J Hortic Sci Biotech*, 92(2):155-167. https://doi.org/10.1080/14620316.2016.1240019

Wu Y, Sun MY, Zhang JP, et al., 2019a. Differential effects of paclobutrazol on the bulblet growth of oriental lily cultured in vitro: growth behavior, carbohydrate metabolism, and antioxidant capacity. *J Plant Growth Regul*, 38(2):359-372. https://doi.org/10.1007/s00344-018-9844-5

Wu Y, Ma YD, Li Y, et al., 2019b. Plantlet regeneration from primary callus cultures of *Lilium brownii* F.E.Br. ex Miellez var. *giganteum* G. Y. Li & Z. H. Chen, a rare bulbous germplasm. *In Vitro Cell Dev-Biol-Plant*, 55(1):44-59. https://doi.org/10.1007/s11627-018-09955-1

Xu H, Yang PP, Cao YW, et al., 2020. Cloning and functional

characterization of a flavonoid transport-related MATE gene in asiatic hybrid lilies (*Lilium* spp.). *Genes*, 11(4):418. https://doi.org/10.3390/genes11040418

Yang M, Zhu LP, Pan C, et al., 2015. Transcriptomic analysis of the regulation of rhizome formation in temperate and tropical lotus (*Nelumbo nucifera*). *Sci Rep*, 5:13059. https://doi.org/10.1038/srep13059

Yang PP, Xu LF, Xu H, et al., 2017. Histological and transcriptomic analysis during bulbil formation in *Lilium lancifolium*. *Front Plant Sci*, 8:1508. https://doi.org/10.3389/fpls.2017.01508

Zhou YX, Chen YX, Tao X, et al., 2016. Isolation and characterization of cDNAs and genomic DNAs encoding ADP-glucose pyrophosphorylase large and small subunits from sweet potato. *Mol Genet Genomics*, 291(2):609-620. https://doi.org/10.1007/s00438-015-1134-3

Zhu FL, Cheng N, Sun H, et al., 2020. Molecular cloning and characterization of a gene encoding soluble starch synthase III (*SSSIII*) in Lotus (*Nelumbo nucifera*). *Biologia*, 75(2):279-288. https://doi.org/10.2478/s11756-019-00341-9

Supplementary information

Figs. S1–S5; Tables S1 and S2