

RESEARCH ARTICLE

Open Access



Comprehensive analysis of AGPase genes uncovers their potential roles in starch biosynthesis in lotus seed

Heng Sun^{1†}, Juanjuan Li^{2†}, Heyun Song^{1,3}, Dong Yang^{1,4}, Xianbao Deng^{1,4}, Juan Liu¹, Yunmeng Wang^{1,3}, Junyu Ma^{1,3}, Yaqian Xiong^{1,3}, Yanling Liu¹ and Mei Yang^{1,4*} 

Abstract

Background: Starch in the lotus seed contains a high proportion of amylose, which endows lotus seed a promising property in the development of hypoglycemic and low-glycemic index functional food. Currently, improving starch content is one of the major goals for seed-lotus breeding. ADP-glucose pyrophosphorylase (AGPase) plays an essential role in regulating starch biosynthesis in plants, but little is known about its characterization in lotus.

Results: We describe the nutritional compositions of lotus seed among 30 varieties with starch as a major component. Comparative transcriptome analysis showed that AGPase genes were differentially expressed in two varieties (CA and JX) with significant different starch content. Seven putative AGPase genes were identified in the lotus genome (*Nelumbo nucifera* Gaertn.), which could be grouped into two subfamilies. Selective pressure analysis indicated that purifying selection acted as a vital force in the evolution of AGPase genes. Expression analysis revealed that lotus AGPase genes have varying expression patterns, with *NnAGPL2a* and *NnAGPS1a* as the most predominantly expressed, especially in seed and rhizome. *NnAGPL2a* and *NnAGPS1a* were co-expressed with a number of starch and sucrose metabolism pathway related genes, and their expressions were accompanied by increased AGPase activity and starch content in lotus seed.

Conclusions: Seven AGPase genes were characterized in lotus, with *NnAGPL2a* and *NnAGPS1a*, as the key genes involved in starch biosynthesis in lotus seed. These results considerably extend our understanding on lotus AGPase genes and provide theoretical basis for breeding new lotus varieties with high-starch content.

Keywords: Lotus seed, Starch biosynthesis, Nutritional composition, AGPase gene, AGPase activity

Background

Nelumbo is a unique genus of Nelumbonaceae, which comprises two extant species: *N. nucifera* Gaertn. widely distributed in Asia and northern Australia, and *N. lutea* Pers. which is distributed in America [1]. As a long-

cultivated crop, lotus has versatile uses, parts of the plant have been used as a snack food, vegetable, medicine and ornamental horticulture. Based on distinct traits and agricultural uses, lotus cultivars can be classified into three groups: seed-lotus, rhizome-lotus and flower-lotus [2, 3]. Seed lotus can produce a higher yield of seed compared with the other two varieties. Currently, lotus seed products are very popular in daily diets in Asia, for example, the annual demand for lotus seed is about 50,000–75,000 tons in China. Mature lotus seed contains 40–60% of starch in the cotyledons, thus it is considered a rich source of starch, and is widely

* Correspondence: yangmei815815@wbpcas.cn

[†]Heng Sun and Juanjuan Li contributed equally to this work.

¹Key Laboratory of Plant Germplasm Enhancement and Specialty Agriculture, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China

⁴Center of Economic Botany, Core Botanical Gardens, Chinese Academy of Sciences, Wuhan 430074, China

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



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

employed in the production of low-glycemic index food. In addition, lotus seed is rich in protein, vitamins, essential amino acids, and a variety of bioactive components with important nutritional and medicinal value.

As a non-structural carbohydrate, starch represents the most significant form of carbohydrate storage in plants. It is composed of two polymers of glucose, amylose and amylopectin, which have different molecular structures, the former is composed of unbranched chains of glucose monomers, while the latter is a branched polysaccharide [4]. Starch can be accumulated in photosynthetic and non-photosynthetic tissues through a highly complex processes [5]. The biosynthesis of starch begins in leaves and then transported to other storage organs in the form of disaccharides for starch synthesis. Sucrose transported from leaves to storage organs is successively catalyzed by sucrose synthase (SuSy), UDPglucose pyrophosphorylase (UDPGase) and phosphoglucomutase (PGM) to produce glucose-6-phosphate, which is then transported into amyloplast for starch synthesis, which is catalyzed by PGM, ADP-glucose pyrophosphorylase (AGPase), starch synthase (SS) and starch branching enzyme (SBE) [6]. Furthermore, other important enzymes such as AGPase [7, 8], SS [9–11] and SBE [12, 13] have been identified and successfully applied in molecular breeding. Currently, the mechanism of starch biosynthesis process has been clarified plant species, such as *Arabidopsis*, potato, rice and corn [5, 14–16]. However, little is still known about the mechanism of starch biosynthesis in lotus seed.

Previous studies have shown that starch biosynthesis process plays an important role in the development of lotus seed [2, 17, 18]. Starch is rapidly accumulated in cotyledons from 9 to 20 days after pollination (DAP) in lotus [2, 17]. Proteomic analysis showed that AGPase was highly abundant at 25 DAP, and starch phosphorylase (StarchP) was significantly accumulated in mature lotus seed [17]. Comparative transcriptome between JX ('Jianxuan 17', a seed-lotus variety) with high starch content and CA ('China Antique') with low starch content showed that some enzyme encoding genes involved in starch biosynthesis, such as AGPase, soluble starch synthase (SSS) and SBE were up-regulated during lotus seed development [2].

AGPase plays an essential role in regulating glycogen and starch biosynthesis in bacteria and plants, respectively [19]. AGPase (EC:2.7.7.27) catalyzes the initial and major limiting step in the starch biosynthetic pathway, converting glucose-1-phosphate (Glc1P) and ATP to ADP-glucose and pyrophosphate (PPi) [20–22]. AGPase is a heterotetramer composed of two small/catalytic subunits and two large/modulatory subunits [8, 23]. There are two isoforms of AGPases, known as the cytosolic and plastidial isoforms, based on its cellular localization. In the grain endosperm of the grass family, ADP-glucose

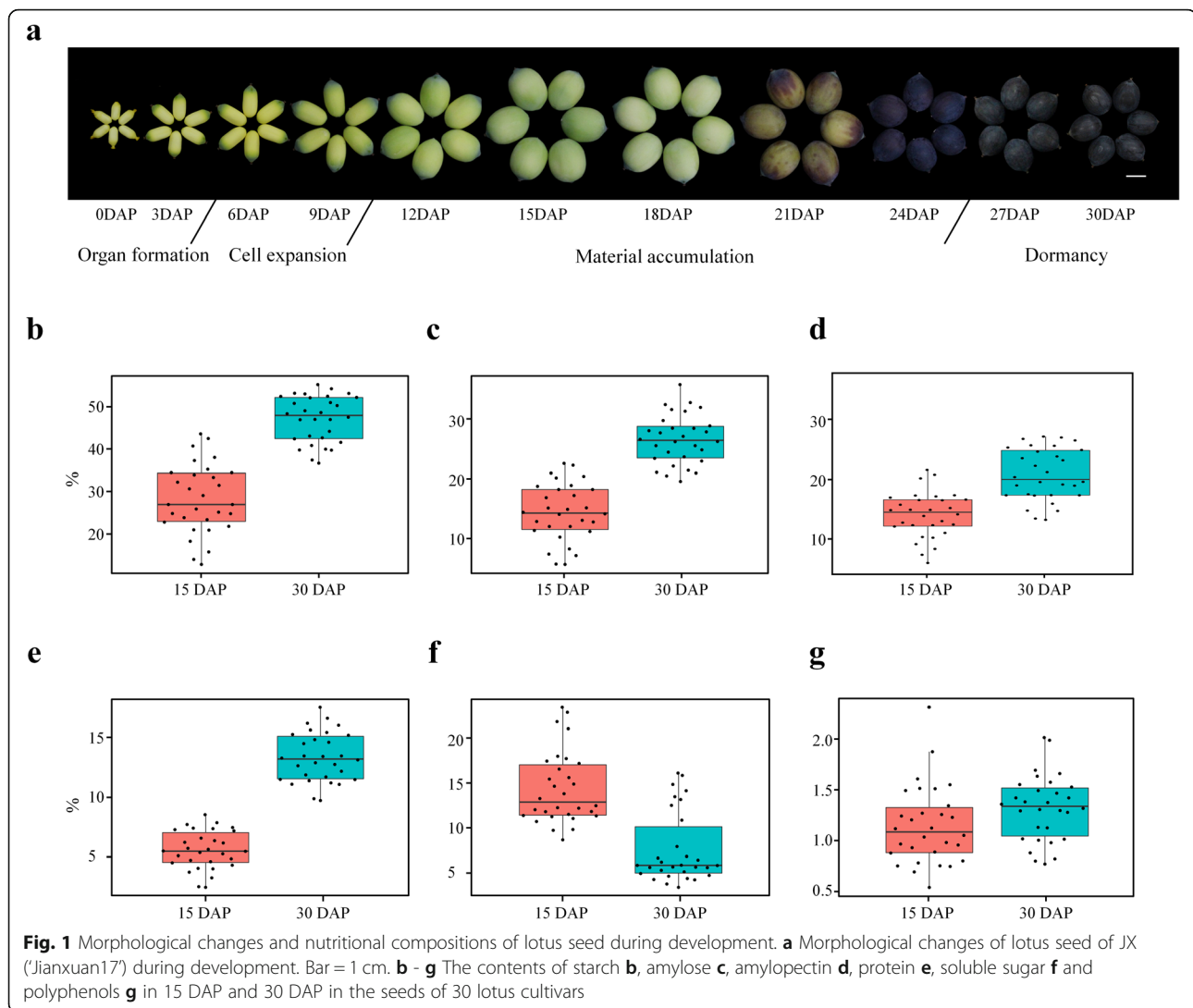
is formed in the cytosol by the cytosolic isoform, and imported into plastids by plastid envelope located ADP-glucose transporter for starch biosynthesis. In contrast, ADP-glucose in most dicotyledonous plants is formed exclusively in the plastids by plastidial AGPase isoform [24]. Currently, AGPase genes have been identified in many species, such as rice, maize, *Arabidopsis* and barley [14, 25, 26]. Previous studies have shown that AGPase mutants exhibited a reduction in starch content, for example, in a near-starchless *Arabidopsis* TL25 mutant with an *adg1* gene encoding small subunit structural gene of AGPase [22, 27]. This deficiency in starch biosynthesis was also observed in a barley mutant, Risø 16, which has a large deletion within the coding region of the small subunit of the cytosolic AGPase [27]. To date, little information is available on the role of AGPase genes in lotus starch biosynthesis.

With the development and improved potential commercial application of lotus seed starch, the work of breeding new lotus varieties with high-starch content is crucial. In this study, we measured the nutritional composition of lotus seed among 30 varieties, and quantified starch, soluble sugar, protein and polyphenol content. As a key rate-limiting enzyme in starch biosynthesis, AGPase encoding genes were systematically identified and characterized by phylogenetic, expression pattern and co-expression network analyses based on the completed lotus genome (*Nelumbo nucifera* Gaertn.) sequencing data. The dynamic change of starch content and AGPase activity in lotus seed was detected in our study. Furthermore, two promising genes for starch biosynthesis, *NnAGPL2a* and *NnAGPS1a*, were identified in lotus seed. This study establishes a foundation for the understanding of starch biosynthesis pathway in lotus, and offer theoretical basis for molecular breeding of new lotus varieties with high-starch content.

Results

The nutritional compositions of lotus seed

Lotus seed reaches mass maturity in about 30 days after pollination across four developmental stages, with organ formation at 1–3 DAP, cell expansion at 4–9 DAP, material accumulation at 10–25 DAP, and dormancy at 26–30 DAP. Every stage is accompanied by morphological changes, such as seed size and color (Fig. 1a). Seeds of 30 seed-lotus varieties were collected at 15 DAP and 30 DAP to determine the nutritional components, including total starch, amylose, amylopectin, protein, soluble sugar and polyphenol (Additional file 1: Table S1). For seeds collected at 15 DAP, total starch ranged from 12.79–43.60% with an average of 28.20%, amylose ranged from 5.67–22.58% with an average of 14.41%, amylopectin ranged from 5.68–21.33% with an average of 13.80%, protein ranged from 2.43–8.52% with an average of



5.55%, soluble sugar ranged from 8.66–23.43% with an average of 14.39%, and polyphenols ranged from 0.54–2.31% with an average of 1.15% (Fig. 1b - g). For seeds collected at 30 DAP, total starch ranged from 36.67–55.28% with an average of 47.12%, amylose ranged from 19.57–35.80% with an average of 26.58%, amylopectin ranged from 12.96–26.96% with an average of 20.54%, protein ranged from 9.7–17.53% with an average of 13.32%, soluble sugar ranged from 3.39–16.11% with an average of 7.56%, and polyphenols ranged from 0.77–2.01% with an average of 1.32% (Fig. 1b - g).

The contents of total starch, amylose, amylopectin and protein in 30 DAP lotus seeds were significantly higher than those in 15 DAP seeds (ANOVA, $P \leq 0.01$), while the soluble sugar content in 15 DAP lotus seeds was significantly higher than that in 30 DAP lotus seeds ($P \leq 0.01$), in contrast, no obvious change in the content of polyphenols was observed. In addition, there was no significant difference between amylose and amylopectin

content in 15 DAP lotus seeds, while amylose content was significantly higher than the amylopectin content in 30 DAP lotus seeds ($P \leq 0.01$). We found great differences in the nutritional composition among the 30 lotus varieties. For example, the coefficient of variation (CV%) for starch of seeds at 15 DAP, soluble sugar of seeds at 30 DAP and protein of seeds at 15 DAP was 29.10, 52.45 and 28.90, respectively (Additional file 1: Table S1).

The AGPase genes are differentially expressed during lotus seed development

The comparative transcriptome analysis was performed to explore the differences in the molecular mechanism for seed development between CA and JX [2]. A total of 4416 and 6916 differentially expressed genes (DEGs) were identified in CA and JX from 9 DAP to 15 DAP, respectively, with 2895 common DEGs (Additional file 2: Table S2). KEGG analysis showed that these common DEGs were mainly involved in 17 pathways (corrected P

value ≤ 0.05), including metabolic pathways, starch and sucrose metabolism and flavonoid biosynthesis (Fig. 2a). We found 44 genes involved in starch and sucrose metabolism pathway, and their expression patterns showed that 23 genes, including AGPase (NNU_05331, NNU_20629, NNU_06174), granule-bound starch synthase (NNU_04661) and 1,4- α -glucan-branching enzyme (NNU_25320, NNU_23975) were simultaneously up-regulated in CA and JX (Fig. 2b). Gene annotation showed that NNU_05331 and NNU_06174 encode the large subunit of AGPase, and NNU_20629 encode the small subunit. Interestingly, NNU_05331 and NNU_20629 were differentially expressed in CA and JX. The expression of NNU_05331 in JX was 16.90 and 27.13 fold higher than in CA, at 12 DAP and 15 DAP, respectively. Whilst, the expression of NNU_20629 in JX was 4.49 and 4.45 fold higher than in CA, at 12 DAP and 15 DAP, respectively. Previous study showed that JX could biosynthesize more starch than CA at 12 DAP and 15 DAP [2]. Therefore, it is tempting to speculate that the differences in AGPase expression could help explain the differences in starch accumulation in these two varieties.

Identification of AGPase subunit genes in lotus genome

After a Blast search in the sacred lotus genome (*Nelumbo nucifera* Gaertn.), a total of seven AGPase genes were identified, including three *NnAGPL1*, *NnAGPL2a*, *NnAGPL2b* genes encoding the large subunit, and four *NnAGPS1a*, *NnAGPS1b*, *NnAGPS2a*, *NnAGPS2b* genes encoding the

small subunit (Table 1; Additional file 3: Table S3). Four genes *NnAGPL1*, *NnAGPL2b*, *NnAGPS1a* and *NnAGPS2a* were distributed on Megasc scaffold_1, while *NnAGPL2a*, *NnAGPS2b* and *NnAGPS1b* were distributed on Megasc scaffold_6, Megasc scaffold_11 and Megasc scaffold_96, respectively. The full length of the AGPase subunit proteins ranged from 311 to 614 amino acids, while the molecular weights (Mw) of the large and small subunits ranged from 58.6 to 67.75 and from 35.21 to 65.2 kDa, respectively. The isoelectric points (pI) of the AGPase subunit proteins ranged from 6.18 to 9.01. Five conserved motifs were identified in six lotus AGPase genes, except for *NnAGPS2b* which only has two conserved motifs (Fig. 3a). Gene structure analysis showed that the large subunit contains 14 to 17 exons, while the small subunit contains 5 to 9 exons (Fig. 3b). In addition, sequence alignment of AGPase proteins against the PSS (the potato small subunit gene) revealed that the critical amino acids for catalysis (Arginine, R and Lysine, K) are in both AGPase subunit proteins. Five proteins including *NnAGPL1*, *NnAGPL2a*, *NnAGPL2b*, *NnAGPS1a* and *NnAGPS1b* had the same conserved binding site (Lysine, K) for Glc-1-P (Fig. 3c, d).

Evolutionary analysis of AGPase subunit genes

Protein sequence analysis showed that the amino acid homology of the large subunit genes was 50.46–75.36%, while for the small subunit genes was 17.69–78.75% (Fig. 4a), with *NnAGPL2a*-*NnAGPL2b* and *NnAGPS1a*-*NnAGPS1b* pairs showing the highest homology amino

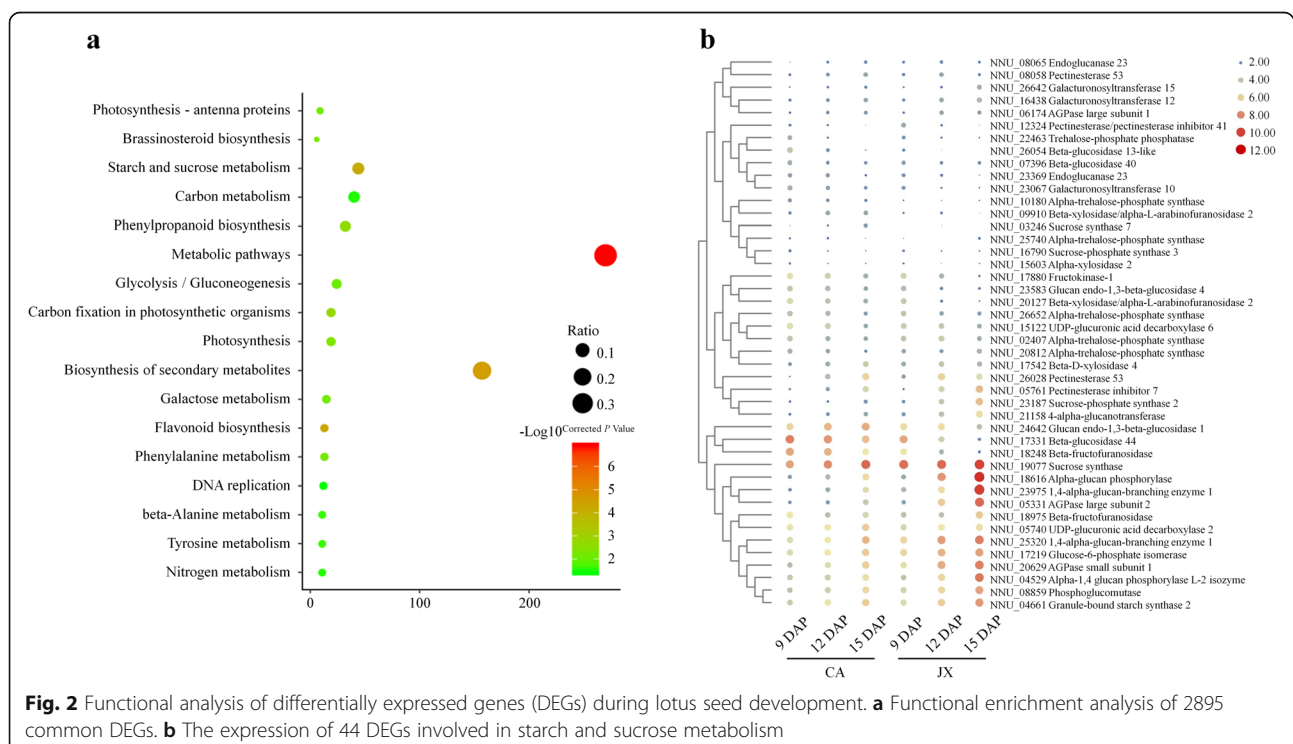
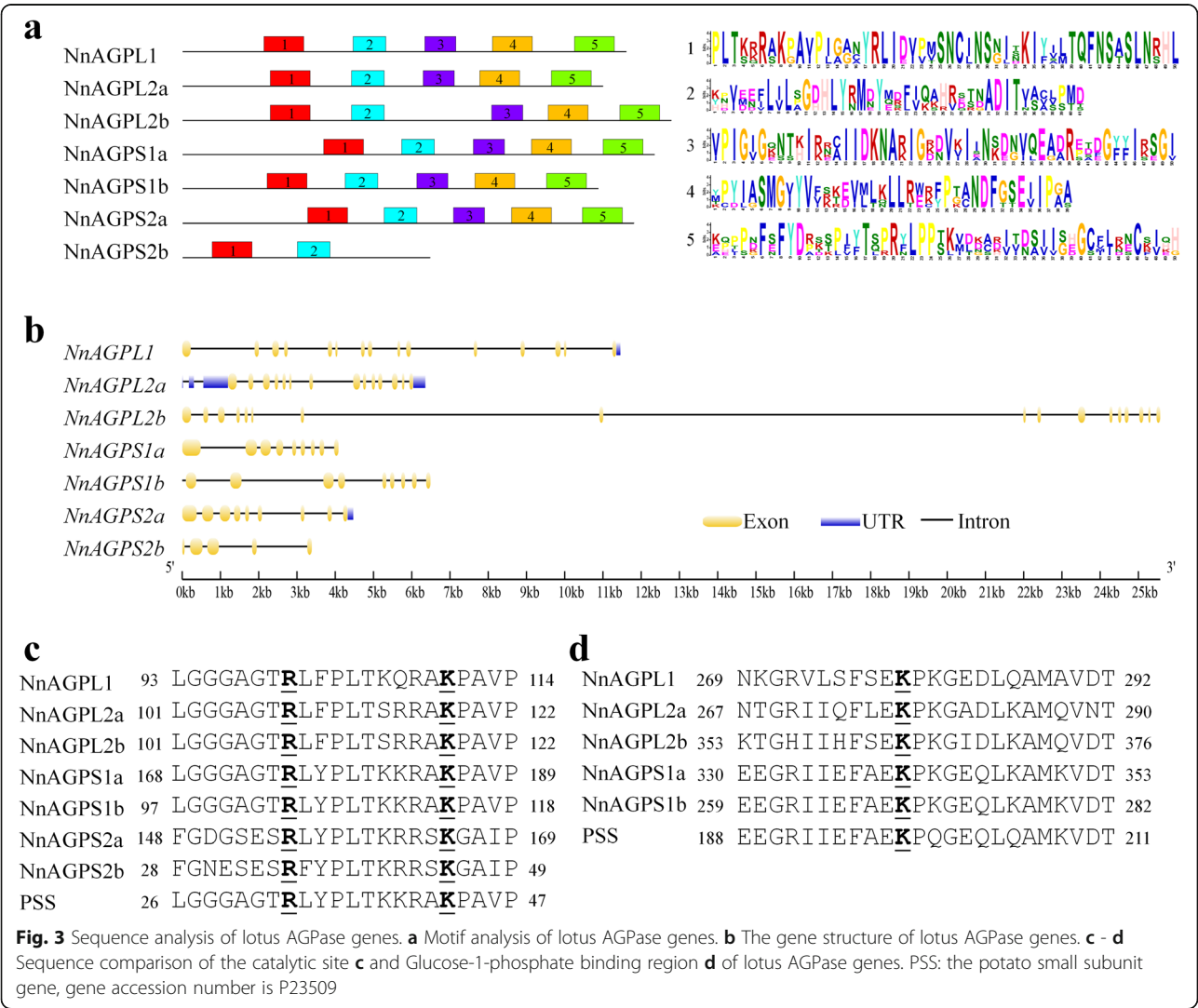


Table 1 Characteristics of lotus AGPase genes

Gene name	Gene ID	Loci	Length	Mw (kDa)	pI
NnAGPL1	NNU_06174	Megascaffold_1, 241,881,173–241,892,629	557	61.67	8.83
NnAGPL2a	NNU_05331	Megascaffold_6, 27,427,206–27,433,563	528	58.6	8.46
NnAGPL2b	NNU_17149	Megascaffold_1, 155,540,325–155,565,896	614	67.75	6.18
NnAGPS1a	NNU_20629	Megascaffold_1, 92,634,342–92,639,171	593	65.2	6.73
NnAGPS1b	NNU_21708	Megascaffold_96, 40,333–46,894	522	57.01	7.62
NnAGPS2a	NNU_10080	Megascaffold_1, 3,549,994–3,554,266	567	63.33	6.34
NnAGPS2b	NNU_11450	Megascaffold_11, 10,160,382–10,163,772	311	35.21	9.01

Mw molecular weight, pI Isoelectric points

acid sequence for the large and small subunits, respectively. In contrast, lowest homology was observed between *NnAGPS2b* and any of the other six identified AGPase gene members. The non-synonymous (dN) and synonymous (dS) substitution rates of homologous gene pairs were calculated to explore evolutionary dynamics and selection pressures between lotus and other plants, which including *Arabidopsis*, rice and maize. All dN/dS < 1 was observed, indicating that purifying selection is acting on AGPase genes (Fig. 4b). A phylogenetic tree was constructed to gain insights into the evolutionary relationships of AGPase genes. All AGPase genes could be grouped into two subfamilies, and genes showed obvious differentiation between the large subunit and the



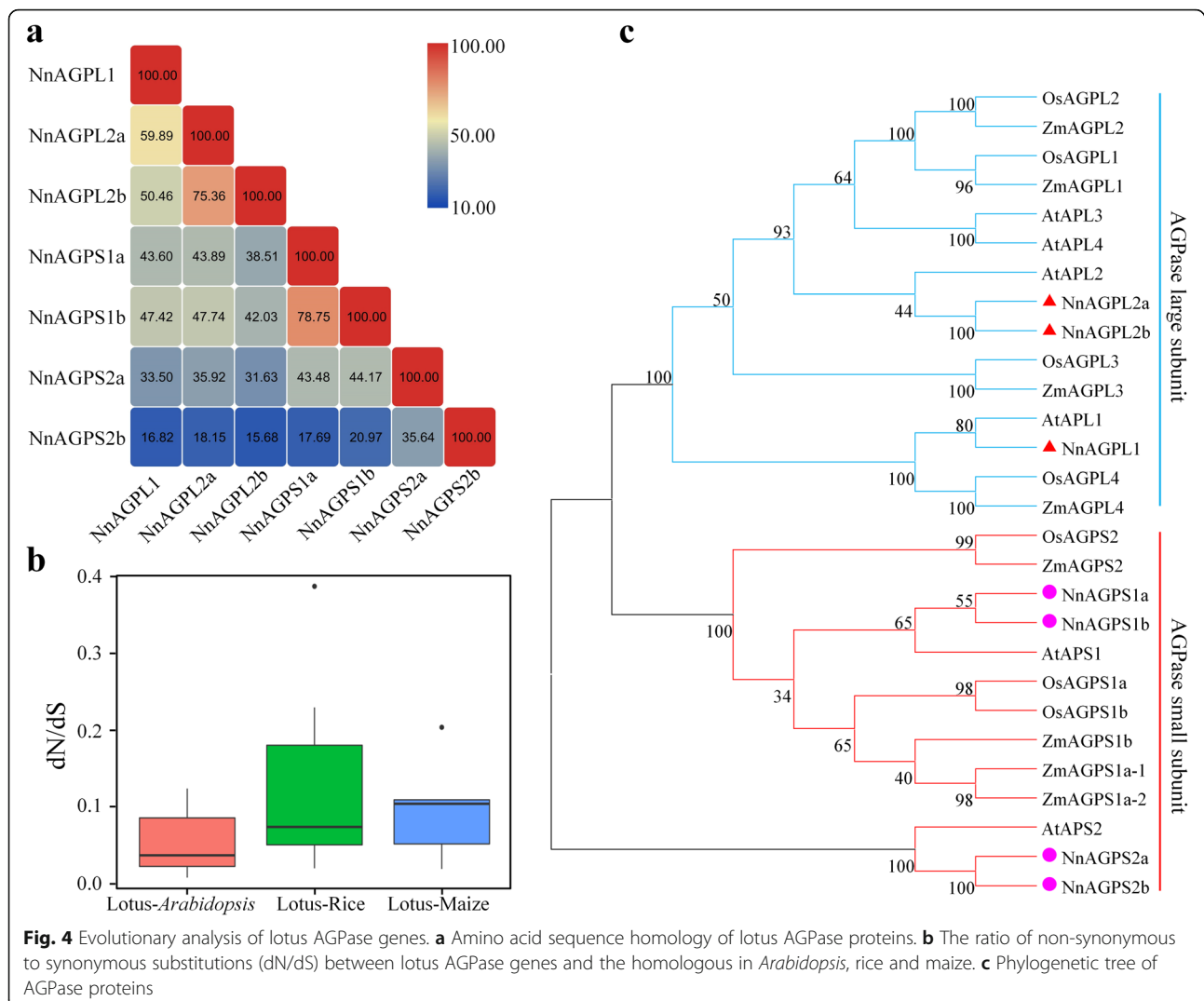


Fig. 4 Evolutionary analysis of lotus AGPase genes. **a** Amino acid sequence homology of lotus AGPase proteins. **b** The ratio of non-synonymous to synonymous substitutions (dN/dS) between lotus AGPase genes and the homologous in *Arabidopsis*, rice and maize. **c** Phylogenetic tree of AGPase proteins

small subunit (Fig. 4c). Compared to rice and maize, a closer evolutionary relationship of AGPase genes was detected between lotus and *Arabidopsis*.

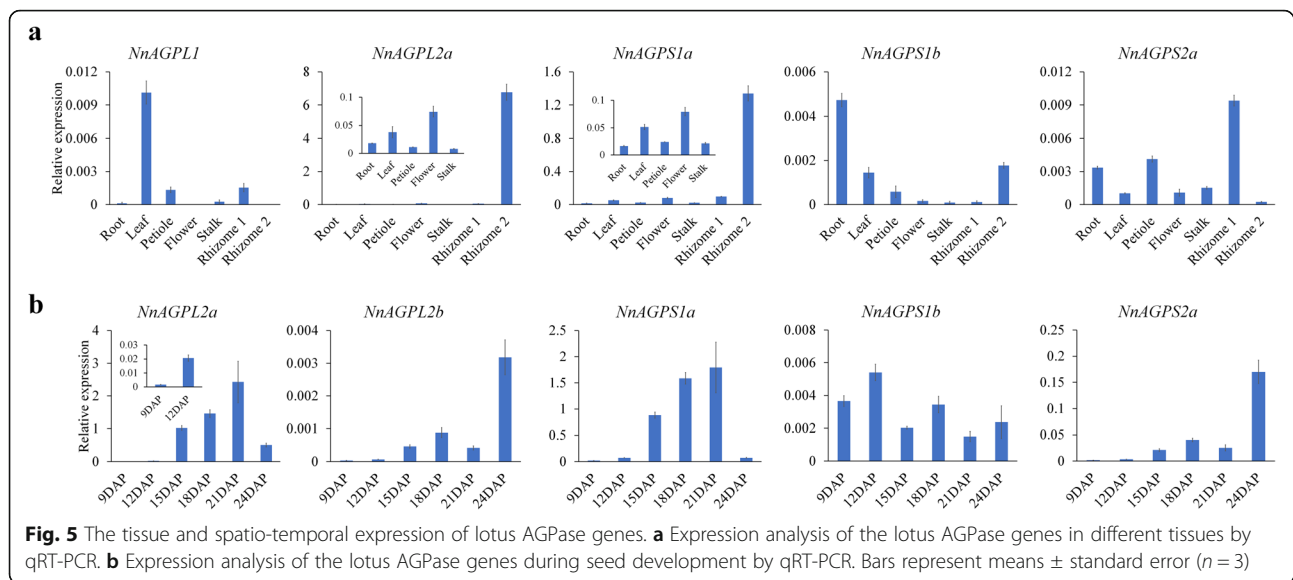
***NnAGPL2a* and *NnAGPS1a* are predominantly expressed in lotus**

The expression patterns of lotus AGPase genes were investigated in different tissues of the cultivar JX, including root, leaf, petiole, flower, stalk, the stolon stage rhizome (rhizome 1) and the swelling stage rhizome (rhizome 2) by Real-time PCR. The expression of five genes were detected in at least one tissue except for *NnAGPL2b* and *NnAGPS2b* (Fig. 5a). *NnAGPL1* was highly expressed in leaf, and also found in petiole and rhizome 1, but not expressed in flower and rhizome 2. *NnAGPL2a*, *NnAGPS1a*, *NnAGPS1b* and *NnAGPS2a* were expressed in all tissues, with up-regulation of *NnAGPL2a* and *NnAGPS1a* in rhizome 2 than in rhizome 1. In addition,

NnAGPL2a and *NnAGPS1a* showed overall higher expression abundances than other genes.

The expression patterns of lotus AGPase genes were also investigated during lotus seed development. For the large subunit genes, *NnAGPL2a* was continuously up-regulated from 9 DAP to 21 DAP, and *NnAGPL2b* showed an obvious expression change at 24 DAP (Fig. 5b). However, the expression abundances of *NnAGPL1* in developing seed was hardly detected. For the small subunit genes, three genes showed expression abundance except for *NnAGPS2b* (Fig. 5b). *NnAGPS1a* and *NnAGPS2a* were induced during seed development, however, *NnAGPS1b* showed no obvious change in trend. In addition, *NnAGPS1a* and *NnAGPS2a* showed similar expression patterns with *NnAGPL2a* and *NnAGPL2b*, respectively.

Based on the above results, we speculate that *NnAGPL2a* and *NnAGPS1a* are the predominantly expressed genes in AGPase large subunit and small subunit,



respectively, especially in response to the development process in lotus seed and rhizome.

Co-expression network analysis of *NnAGPL2a* and *NnAGPS1a*

Co-expression network analysis is a powerful method for predicting gene function [28]. The predominantly expressed genes during lotus development, *NnAGPL2a* and *NnAGPS1a* were selected for this study. As a result, a total of 408 and 444 genes were co-expressed with *NnAGPL2a* and *NnAGPS1a*, respectively ($|PCC| \geq 0.9$) (Fig. 6a; Additional file 4: Table S4), with 359 commonly co-expressed genes identified. KEGG analysis showed that eight pathways were enriched among these co-expressed genes (corrected P value ≤ 0.05), including the biosynthesis of secondary metabolites, starch and sucrose metabolism and carbon metabolism (Fig. 6b). Fourteen genes were involved in starch and sucrose metabolism, including granule-bound starch synthase (NNU_04661), alpha-1,4 glucan phosphorylase (NNU_04529) and 1,4-alpha-glucan-branching enzyme (NNU_23975, NNU_25320) (Additional file 5: Figure S1). Five genes were selected to verify the expression patterns by qRT-PCR, and all were significantly up-regulated from 9 DAP to 18 DAP and with similar expression pattern to *NnAGPL2a* and *NnAGPS1a* (Figs. 5b; 6c). Thus, it is likely that the role of *NnAGPL2a* and *NnAGPS1a* could be linked with these co-expressed genes to contribute to starch biosynthesis in lotus seed.

Sequence variation of *NnAGPL2a* and *NnAGPS1a* between CA and JX

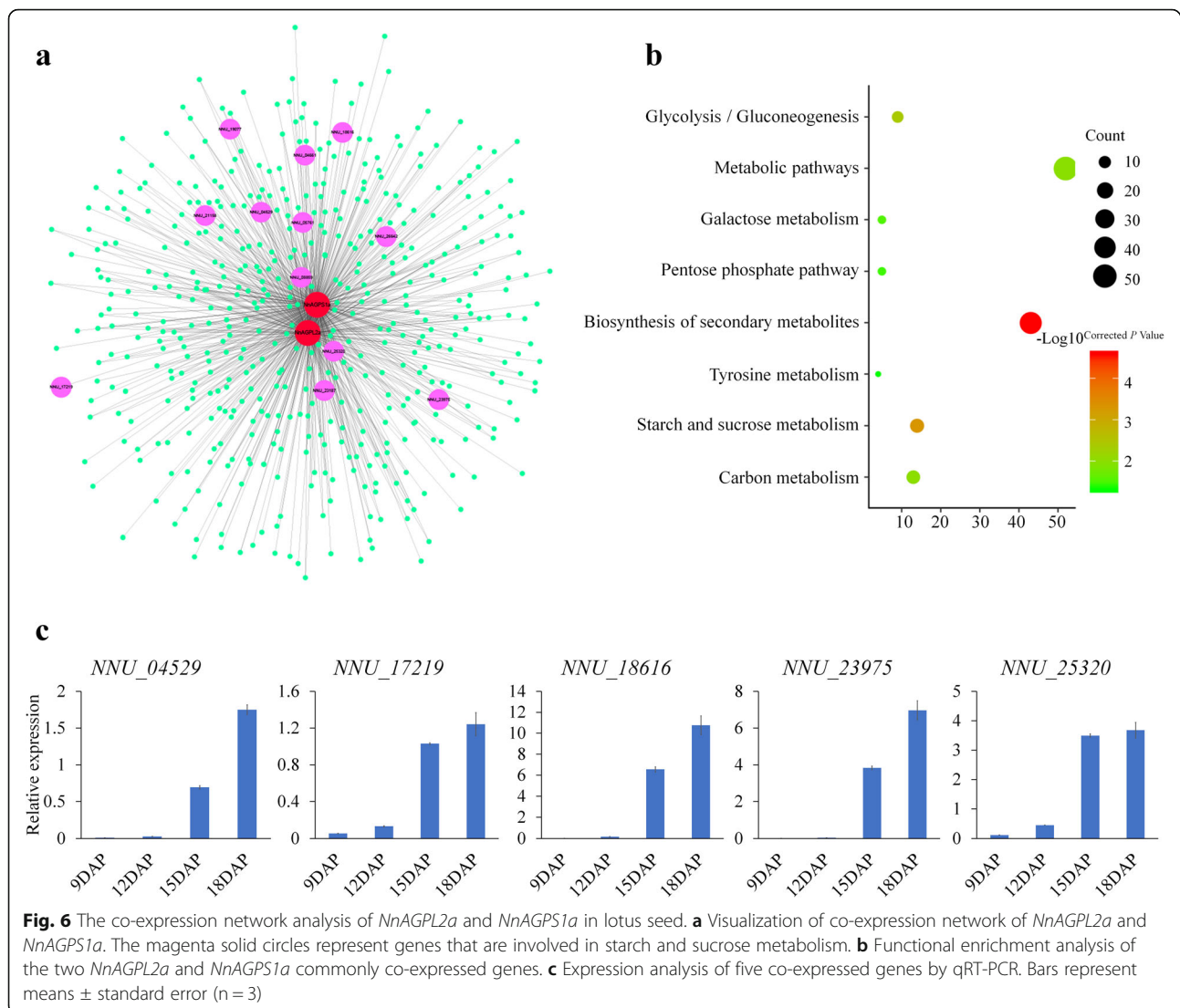
Significant differences in seed size and starch content were detected between CA and JX varieties [2]. In order to investigate whether *NnAGPL2a* and *NnAGPS1a* have

sequence variation in different lotus varieties, we cloned the coding sequences (CDS) and promoter regions of these two genes from CA and JX, respectively (Additional file 6: Figure S2; Additional file 7: Table S5). For *NnAGPL2a*, three nonsynonymous mutation and four synonymous mutation were identified in the CDS between CA and JX, the nucleotide polymorphism of C¹¹⁷/G¹¹⁷, T²⁶⁹/G²⁶⁹ and G⁹⁰³/A⁹⁰³ resulted in Asn³⁹/Lys³⁹, Ile⁹⁰/Arg⁹⁰ and Met³⁰¹/Ile³⁰¹, respectively (Fig. 7a; Additional file 6: Figure S2C). A total of 14 SNP and two Indel were identified in the promoter region of *NnAGPL2a*, and 12 variations were found to occur in the -1000 to -2000 bp region (Fig. 7b; Additional file 6: Figure S2d). Three variations caused changes in cis-elements, the SNP at position -833 bp (C/T) in JX caused a change in G-box (CACGTT→CATGTT), the deletion at position -1762 bp (G/-) in JX caused a change in I-box (GGATAAGCTG→GATAAGCTG), the deletion at position -1862 bp (TC/-) in CA caused a change in LTR (TTTCGG→TTGG) (Fig. 7b).

Greater sequence conservation was detected in *NnAGPS1a* than in *NnAGPL2a* between CA and JX. A single SNP in the CDS and two in the promoter region were identified (Fig. 7; Additional file 6: Figure S2). The nucleotide polymorphism of T³⁷/C³⁷ in CDS resulted in Cys¹³/Arg¹³, while the SNP at position -409 bp (A/T) in CA caused a change in ARE cis-element (AAAC CA→AAACCT) (Fig. 7b).

AGPase activity is increased during starch biosynthesis in lotus seed

Due to its potential role in regulating starch synthesis in plants, the dynamic change in starch content and AGPase activity in lotus seed of JX was detected in our study. The starch content and enzyme activity consistently increased



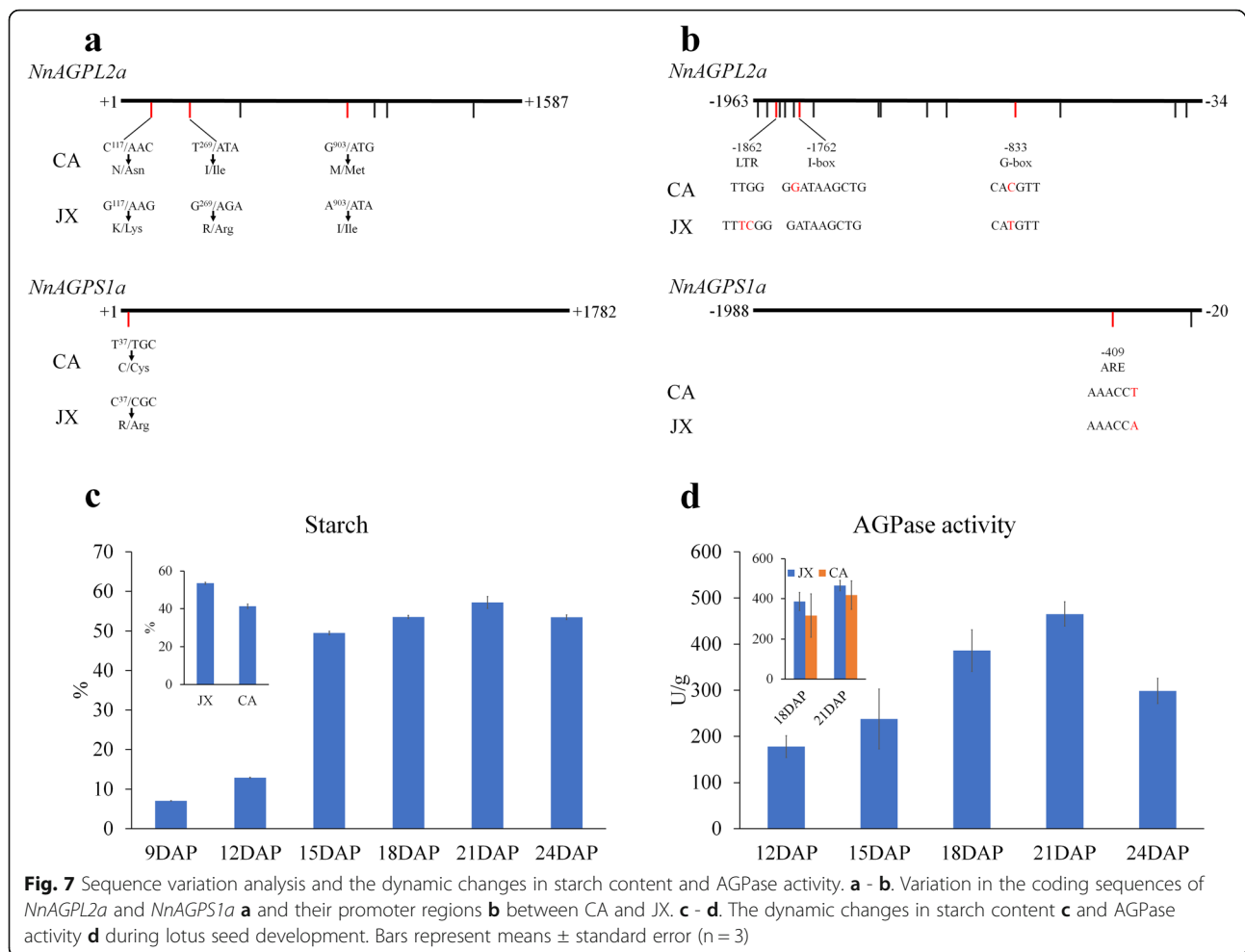
from 9 DAP to 21 DAP (Fig. 7c, d). Starch synthesis peaked during 12DAP to 15DAP, then accumulation slowed until 21 DAP. JX had a higher starch content than CA, which is consistent with previous studies (Fig. 7c). Compared with starch biosynthesis, the enzyme activity showed a steady increasing trend until 21DAP, and then rapid decreased at 24 DAP (Fig. 7d). In addition, JX had a higher enzyme activity than CA at 18 DAP and 21 DAP in lotus seed (Fig. 7d).

Discussion

Lotus seed is a product of sexual reproduction, and is largely consumed across Asia due to its rich in nutritional constituents. Previous studies on lotus seed has mainly focused on the processing technology, the structure and physicochemical properties of starch and the functional analysis of nutritional components, while the research progress on molecular mechanism of lotus seed

development and quality traits has been slow [29–32]. In this study, we determined the content of key components in 30 seed-lotus varieties, these data will provide an important information on the scientific evaluation of lotus seed quality and the screening of germplasm resources for seed-lotus breeding.

Starch can be divided into three categories as low (< 20% amylose), medium (21–25%) and high (> 26%) based on amylose content [33]. Unlike most cereal crops, the starch in lotus seed belongs to natural high amylose, and some varieties contain more than 32% amylose, such as 'Jianxuan 35' (35.80%), 'WBG_S1' (32.78%) and 'Honghua Jian Lian' (32.44%). The starch in lotus seed can be easily retrograded to produce resistant starch, which has the potential to develop hypoglycemic functional food and low glycemic index food [34–36]. Our results indicate that the starch content in lotus seed still needs to be improved, with some main cultivars having their



content lower than the average level of 47.12%, such as 'Taikong 3' (40.00%) and 'Jingguang 1' (42.66%). Previous studies have shown that understanding the mechanism of starch biosynthesis can successfully provide the molecular basis for breeding new varieties with high-starch content in crops, such as rice, corn and wheat [37–40]. Although the research on the mechanism of starch biosynthesis in lotus seed has attracted the attention of plant breeders, little is still known on the topic [2, 17, 18]. Previous studies have shown that starch accumulation is accompanied with a high mRNA and protein abundance of AGPase in lotus seed, indicating that AGPase plays an important role in this process [2, 17], thus warranting our study to analyze the mechanism of starch biosynthesis in lotus seed.

Here, we exploited the available genome data to systematically identify seven AGPase subunit genes in lotus. The number of AGPase genes in lotus was similar to those found in other plants, for example, seven in rice, eight in maize and six in *Arabidopsis* [14]. High variation was observed in small AGPase subunit genes than in large subunit genes, which is not consistent with previous reports

that showed high amino acid sequence conservation of small subunit genes than large subunit genes [41]. The AGPase genes could be distinctly divided into two sub-families in plants, indicating that large and small AGPase subunit genes have different evolutionary patterns (Fig. 4c). The lotus AGPase genes showed closer evolutionary relationship with genes in *Arabidopsis* than with rice and maize, thus supporting the conclusion that sacred lotus is a basal eudicot [1]. Selective pressure analysis showed purifying selection is acting as a primary force in the evolution of AGPase genes, which could suggest that the functions of AGPase homologous genes in different plant species are conserved.

The spatio-gene expression patterns can partly be used to predict gene function. In lotus, AGPase genes showed varied expression patterns in different tissues except for *NnAGPS2b*, suggesting that AGPase might play an important role in regulating the growth of lotus, especially in the starch synthesis tissues, such as seed and rhizome (Fig. 5). Two homologous gene pairs, *NnAGPL2a* - *NnAGPL2b* and *NnAGPS1a* - *NnAGPS1b*, showed an inconsistent expression patterns, which could suggest that

functional differentiation of AGPase homologous genes has occurred in lotus. As a key enzyme in starch biosynthesis pathway, the regulatory properties of AGPase synergistically interact between the two subunits [7]. It is noteworthy that *NnAGPL2a* and *NnAGPS1a* are the predominantly expressed genes in lotus seed, suggesting these two are the key genes that might be involved in the regulation of starch biosynthesis and seed development. Differential expression of these two genes has been linked with difference in starch accumulation between CA and JX [2]. Phylogenetic analysis revealed that *NnAGPS1a* and *NnAGPL2a* are homologous to *AtAPS1* and *AtAPL2*, respectively (Fig. 4c). *AtAPS1* is the main small subunit in *Arabidopsis* predominantly regulating starch biosynthesis, with its mutant showing low starch phenotype, while *AtAPL2* is the minor regulatory subunit in leaf with catalytic activity that could contribute to ADP-glucose synthesis in planta [7, 22]. We speculate that *NnAGPL2a* and *NnAGPS1a* are likely involved role in the regulation of starch biosynthesis in lotus seed, however, further studies to define the biological function and molecular mechanisms of these two genes are needed.

Domestication and breeding lead significant genetic changes in most crops, and identification of favored haplotypes can be used in molecular breeding [42]. In wheat, the haplotypes of two AGPase genes, *TaAGP-S1-7A* and *TaAGP-L-1B*, which are associated with thousand kernel weight (TKW), and the favored haplotypes underwent strong positive selection [37]. Here, we identified sequence variations in *NnAGPL2a* and *NnAGPS1a* between CA and JX. No variation was detected in the catalysis and Glc-1-P binding site, thus it would be worthy to determine whether variations in the CDS could affect the function of AGPase (Fig. 7a). The observed differential expression of *NnAGPL2a* and *NnAGPS1a* between CA and JX during lotus seed development, could be associated with the variation in the promoter region of these two genes (Fig. 7b). For example, the G-box mutation in *NnAGPL2a* promoter may affect the binding and regulatory activity of some environmental stress responsive transcription factors in JX, such as MYC and bZIP genes [43]. In addition, G-box is also closely associated with regulation of starch biosynthesis process, for example, a previous study showed that G-box plays a role in regulation of SBE and AGPase gene expression [44, 45]. Overall, these results partly reveal the genetic differentiation of *NnAGPL2a* and *NnAGPS1a*, and provide an important reference for identifying favored haplotype at the population level, which could be used in molecular breeding.

Rhizome is an important lotus storage organ, and starch is one of its most abundant components accounting for 10–20% fresh weight [46]. Previous studies have shown that AGPase gene is involved in the starch biosynthesis process

in rhizome [3, 46]. Our data showed a higher AGPase enzyme activity at the initial swelling stage in comparison to stolon or late swelling stages in the rhizome of ZO ('Zhou Ou'), which is a rhizome-lotus cultivar. Since *NnAGPL2a* and *NnAGPS1a* were the predominantly expressed and up-regulated genes during the rhizome swelling stage in two cultivars, JX and ZO (Fig. 5a; Additional file 8: Figure S3). We speculate that these two AGPase genes also regulate the starch biosynthesis in rhizome.

Our results showed that the expression patterns of *NnAGPL2a* and *NnAGPS1a* are accompanied with the changes in AGPase activity in lotus seed and rhizome (Figs. 5b; 7d; Additional file 8: Figure S3b). Therefore, enhancing the activity of AGPase by genetic engineering these two genes are likely to be a feasible way to increase the starch content in lotus seed and rhizome. Similar research strategies have been successfully applied in crops, such as rice, corn and wheat [47–50]. Transgenic rice plants expressing the potato AGPase large subunit *UpReg1* gene exhibited elevated photosynthetic capacity and starch levels in leaves, and increased seed biomass [47]. Overexpression of AGPase large subunit *TaAGPL1* significantly enhanced AGPase activity and the rate of starch accumulation in wheat grains [48]. Here, we provide important gene resources for future genetic improvement of starch accumulation in lotus varieties.

Conclusions

Improving the starch content is currently one of the major goals for seed-lotus breeding. In this study, we have revealed the nutritional composition including starch, soluble sugar, protein and polyphenols in the seed of 30 lotus varieties. These data provide important information on the scientific evaluation of lotus seed quality and for screening germplasm resources for seed-lotus breeding. Comparative transcriptome analysis showed that AGPase genes were differentially expressed in two varieties with significantly different starch content. Seven AGPase genes were characterized in lotus (*Nelumbo nucifera* Gaertn.), and their sequence conservation, evolution, expression pattern and co-expression network were analyzed. The expression patterns of *NnAGPL2a* and *NnAGPS1a* AGPase genes were accompanied by the increase in starch content and enhanced AGPase activity in lotus seed, thus are to be the key genes regulating starch biosynthesis in lotus seed (Fig. 8). This study presents a starting point for functional evaluation of AGPase genes in starch biosynthesis in lotus, and provides theoretical basis for breeding new lotus varieties with high-starch content.

Methods

Transcriptome analysis of lotus seed

To investigate the regulatory mechanism of starch accumulation during lotus seed development, the public

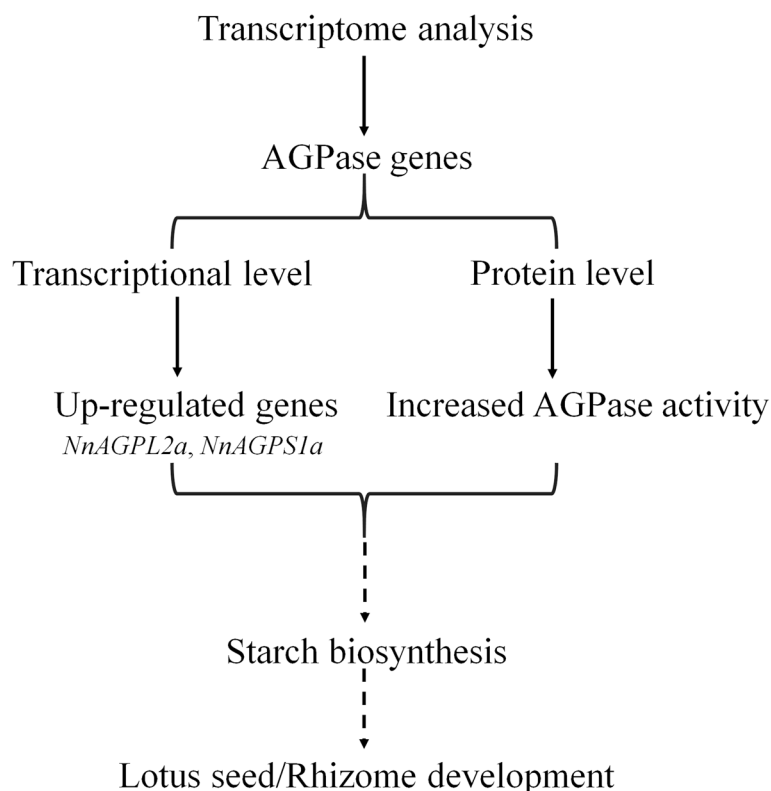


Fig. 8 Summary model of this study. We identified AGPase genes involved in lotus seed development by comparative transcriptome analysis in lotus (*Nelumbo nucifera* Gaertn.). The expression pattern of *NnAGPL2a* and *NnAGPS1a* were accompanied by the increased AGPase activity and starch content

transcriptome database corresponding to expression abundances in two lotus varieties CA and JX at various stages (9 DAP, 12 DAP, 15 DAP) were obtained from NCBI (<https://www.ncbi.nlm.nih.gov/sra?term=SRP127765>). The identification of differentially expressed genes (DEGs) was performed as previously described [2]. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was implemented by KOBAS 3.0 [51]. The heatmap analysis of gene expression was visualized using TBtools [52].

For gene co-expression analysis, pearson correlation coefficient (PCC) was calculated based on the public transcriptome database to measure the co-expression relationships between two AGPase genes (*NnAGPL2a* and *NnAGPS1a*) and other genes [28]. Genes with $|PCC| \geq 0.9$ were used for constructing co-expression network, and the network was visualized by Cytoscape software (3.4.0).

Identification of AGPase genes from the sacred lotus genome

To identify AGPase subunit genes from the lotus genome, we obtained all the predicted protein sequences from lotus genome (*Nelumbo nucifera* Gaertn.) using

gene annotation gff3 file [1]. The fully characterized AGPase subunit protein sequences of *Arabidopsis*, rice and maize were download from Tair (<https://www.arabidopsis.org/index.jsp>) and NCBI website (<https://www.ncbi.nlm.nih.gov/>). The protein BLAST program was performed to identify the lotus AGPase subunit genes. After removing the redundant sequences, the NTP transferase domains of each lotus AGPase subunit protein amino acid sequences were analyzed using SMART online program (http://smart.embl-heidelberg.de/smart/set_mode.cgi?NORMAL=1). The predicted molecular weight (Mw) and isoelectric points (pI) of AGPase subunit proteins were calculated using the ExPASy portal (<http://web.expasy.org/protparam/>). The complete amino acid sequences were analyzed using MEME software (<http://meme-suite.org/index.html>) to discover the conserved motifs of AGPase genes. Multiple sequence alignments of gene were carried out using ClustalX (ver.1.83) software with default settings.

Phylogenetic analysis of AGPase genes in lotus

AGPase subunit protein sequences from four plant species including lotus, *Arabidopsis*, rice and maize were used for the phylogenetic analysis. The tree was

constructed using MEGA7 software with Neighbour-Joining method as previously described [53, 54]. The homologous pairs of AGPase subunit genes between lotus and the three species were identified using BLASTP program. The homologous gene pairs were subsequently used to calculate non-synonymous (dN) and synonymous (dS) substitution rates to explore the evolutionary dynamics of AGPase genes. The dN/dS was calculated by Maximum Likelihood (PAML) yn00 program with the GMYN method [55].

RNA extraction and qRT-PCR analysis

The lotus cultivars were grown in the experimental field at Wuhan Botanical Garden (N30°30', E114°31'), Wuhan, China. Tissues and seeds were collected at different development stages, immediately frozen in liquid nitrogen, and stored at -80 °C for later use. Total RNA was extracted using the Plant Total RNA Isolation Kit (Beijing Zoman Biotechnology Co., Ltd., Beijing, China). High-quality RNAs were reverse transcribed to cDNA using TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix (Lot#M31212, Beijing TransGen Biotech Co., Ltd., Beijing, China). The qRT-PCR experiments were performed using StepOnePlus Real-time PCR System (Applied Biosystems, USA), and the relative gene expression level was calculated and normalized using *NnACTIN* (Gene ID NNU_24864) as the internal standard. Gene-specific primers for qRT-PCR were designed according to the gene coding sequences using Primer Premier 5.0 software and synthesized commercially (TIANYI HUIYUAN, Wuhan, China). The primers used for qRT-PCR are listed in Additional file 9: Table S6.

Cloning *NnAGPL2a* and *NnAGPS1a* genes and their promoter regions

Gene and promoter specific primers (Additional file 9: Table S6) of *NnAGPL2a* and *NnAGPS1a* were designed based on the published lotus genome (*Nelumbo nucifera* Gaertn.) sequence database [1]. The coding sequence (CDS) of *NnAGPL2a* and *NnAGPS1a* were amplified using cDNA from 18 DAP seed of CA and JX as the template, the promoter region of these two genes was amplified using genomic DNA from CA and JX. The corresponding PCR fragments were cloned into pDONR ZEO for sequencing (TIANYI HUIYUAN, Wuhan, China). Comparison of the CDS and promoter sequences from JX and CA clones was done by ClustalX (ver.1.83) software with default settings. The cis-elements were analyzed using PlantCARE program [56].

Assay of AGPase activity

The ADP-Glucose Pyrophosphorylase activity was detected using ADPG Pyrophosphorylase (AGP) Assay Kit according to the manufacturer's instructions (Cat#BC0430, Beijing

Solarbio Science & Technology Co., Ltd., Beijing, China). Samples from seed tissue at different stages (12, 15, 18, 21 and 24 DAP) and rhizome were collected and immediately frozen in liquid nitrogen, and stored at -80 °C until use. All AGPase activity assays were performed in triplicate with one of the replicates assayed using 0.1 g of fresh tissue for every stage. A specific enzyme unit was defined as the amount of the enzyme that catalyzes the conversion of one nmol of NADPH per minute, per gram of tissue under specified assay conditions. The absorption wavelength was set to 340 nm and detection performed with an Infinite M200 Luminometer (Tecan, Mannerdorf, Switzerland).

Determination of nutritional components in lotus seed

Thirty seed-lotus cultivars were kept and grown in the experimental field at Wuhan Botanical Garden (Wuhan, China) to analyze the nutritional composition of the seeds. Seeds were collected at 15 DAP and 30 DAP, and the collection of the materials complied with local and national guidelines. Subsequently, the germ was removed and the seeds were dried in an oven (DHG-9146A, Shanghai) at 100 °C for 1 h, and then dried at 65 °C to a constant weight. The dried seeds were ground into a powder and filtered by a 100-mesh sieve. The determination of nutritional components was performed at Wuhan ProNets Biotechnology Co., Ltd. (Wuhan, China). Three replicates were used for this assay.

Soluble sugars and starch content were measured using anthrone-sulfuric acid colorimetry method. Approximately 0.1 g of seed powder was mixed thoroughly with 10 ml of 80% ethanol in a centrifuge tube, boiled in a water bath at 80 °C for 30 min and then centrifuged. The supernatant and pellet were used to determine the content of soluble sugar and starch content, respectively. A 2 ml volume of the supernatant was transferred to a new centrifuge tube, mixed with 0.5 ml anthrone-ethyl acetate (1 g anthrone dissolved in 50 ml of ethyl acetate) and 5 ml of concentrated sulfuric acid, then boiled in a water bath at 100 °C for 1 min, and cooled at room temperature. The absorbance was measured at 630 nm with a spectrophotometer (TU-1810D, Beijing, China). For starch content, the pellet was transferred into a 50 ml volumetric flask, mixed with 20 ml distilled water, and boiled in a water bath at 100 °C for 30 min. After boiling, 2 ml 9.2 mol/L perchloric acid were added. And extracted for 15 min, then cooled at room temperature. Finally, the tube was centrifuged and the supernatant was used to determine the content of starch by anthrone colorimetric assay.

The content of amylose was measured using iodine colorimetry [57]. Approximately 0.1 g of seed powder was mixed with 9 ml 1 mol/L NaOH solution in a centrifuge tube, boiled in a water bath at 100 °C for 10 min, after using distilled water for cooling capacity to 100 ml,

5 ml fluid was transferred to a new centrifuge tube, and 50 ml distilled water, 1 ml of 1 mol/L acetic acid and 1 ml iodine reagent (0.2 g iodine, 20 g potassium iodide dissolved in 100 ml distilled water) were added in turns. The absorbance was measured at 630 nm with a spectrophotometer.

Protein content was measured using coomassie brilliant blue method [58]. About 0.02–0.05 g seed powder was mixed with distilled water and ground to a homogenate, centrifuged at 4000 rpm/min for 10 min, then 1 ml of the supernatant was transferred to a new centrifuge tube, mixed with 5 ml of coomassie brilliant blue reagent (0.1 g coomassie bright blue dissolved in 50 ml 90% ethanol, added 100 ml 85% (W/V) phosphoric acid, and topped up to 1 L volume using distilled water). The absorbance was measured at 595 nm with a spectrophotometer.

Polyphenol content was measured using folin-ciocalteu colorimetry as previously described [59]. Approximately 0.02–0.05 g seed powder was mixed with 10 ml of 60% ethanol and hydrochloric acid (the final concentration is 0.024%) in a centrifuge tube, and boiled in a water bath at 75 °C for 50 min. The absorbance was measured at 765 nm with a spectrophotometer.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12870-020-02666-z>.

Additional file 1: Table S1. Summary statistics for nutritional components of lotus seed.

Additional file 2: Table S2. The expression of 2895 common DEGs during lotus seed development.

Additional file 3: Table S3. Protein sequences of lotus AGPase genes in lotus genome (*Nelumbo nucifera* Gaertn.).

Additional file 4: Table S4. The co-expressed genes of *NnAGPL2a* and *NnAGPS1a*.

Additional file 5: Figure S1. The expression of 14 co-expressed genes which were involved in starch and sucrose metabolism.

Additional file 6: Figure S2. Cloning of *NnAGPL2a* and *NnAGPS1a* genes and their promoter regions. a-b. The coding sequence (a) and promoter region (b) of *NnAGPL2a* and *NnAGPS1a* were amplified from CA and JX. c-d. Variation in the CDS of *NnAGPL2a* and *NnAGPS1a* (c) and promoter regions (d).

Additional file 7: Table S5. The CDS and promoter sequences of *NnAGPL2a* and *NnAGPS1a* in JX and CA.

Additional file 8: Figure S3. AGPase is involved in the development of rhizome. a. Illustrations of rhizome at different developmental stages. Bar = 1 cm. b. The activity of AGPase during rhizome development. Bars represent means ± standard error (n = 3). c. Expression analysis of *NnAGPL2a* and *NnAGPS1a* during rhizome development. Bars represent means ± standard error (n = 3).

Additional file 9: Table S6. Primer sequences used in this study.

Additional file 10. Original images for gels in this study.

Abbreviations

AGPase: ADP-glucose pyrophosphorylase; CDS: Coding sequences; DEGs: Differentially expressed genes; dN: Non-synonymous; dS: Synonymous; Glc1P: Glucose-1-phosphate; PGM: Phosphoglucosyltransferase; pI: Isoelectric points; PPi: Pyrophosphate; SBE: Starch branching enzyme; SS: Starch synthase;

SSS: Soluble starch synthase; SuSy: Sucrose synthase; UDPase: UDP-glucose pyrophosphorylase

Acknowledgements

We are grateful to Rodrigo Matus (University of Durham, Durham, UK) and Abid Ullah (Department of Botany, University of Malakand, Pakistan) for the language modification in article.

Authors' contributions

MY, HS and JYL conceived and designed the experiments. HS, JYL, DY, HYS, YMW, JYM and YQX performed the experiments. HS wrote the paper. MY, XBD, JL and YLL revised the manuscript. All authors have read and approved the manuscript.

Funding

This work was supported by the Key Research Program of Frontier Sciences CAS (QY2DB-SSW-SMC017), the Youth Innovation Promotion Association CAS (2017390) and the National Natural Science Foundation of China (31872136, 31772353). The funders provided the funds but were not involved in the study design, or data collection, analysis and interpretation, or manuscript writing.

Availability of data and materials

The main data supporting the results of this article are included within the article and the provided additional files. The genome data of *N. nucifera* 'China Antique' can be downloaded from GenBank (PID PRJNA168000, <http://www.ncbi.nlm.nih.gov/Traces/wgs/?val=AQOG01>). The public transcriptome data sets corresponding to expression abundances in two lotus varieties CA and JX are available in the Sequence Read Archive (SRA) of NCBI with accession number SRP127765.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Key Laboratory of Plant Germplasm Enhancement and Specialty Agriculture, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China. ²Hubei Province Research Center of Engineering Technology for Utilization of Botanical Functional Ingredients, Hubei Key Laboratory of Quality Control of Characteristic Fruits and Vegetables, College of Life Science and Technology, Hubei Engineering University, Xiaogan 432000, Hubei, China. ³University of Chinese Academy of Sciences, 19A Yuquanlu, Beijing 100049, China. ⁴Center of Economic Botany, Core Botanical Gardens, Chinese Academy of Sciences, Wuhan 430074, China.

Received: 9 June 2020 Accepted: 23 September 2020

Published online: 06 October 2020

References

- Ming R, Vanburen R, Liu YL, Yang M, Han YP, Li LT, Zhang Q, Kim MJ, Schatz MC, Campbell M, et al. Genome of the long-living sacred lotus (*Nelumbo nucifera* Gaertn.). *Genome Biol.* 2013;14:R41.
- Li JJ, Shi T, Huang LY, He DL, Nyong'A TM, Yang PF. Systematic transcriptomic analysis provides insights into lotus (*Nelumbo nucifera*) seed development. *Plant Growth Regul.* 2018;86:339–50.
- Yang M, Zhu LP, Pan C, Xu L, Liu YL, Ke WD, Yang PF. Transcriptomic analysis of the regulation of rhizome formation in temperate and tropical Lotus (*Nelumbo nucifera*). *Sci Rep.* 2015;5:13059.
- Tester RF, Karkalas J, Qi X. Starch - composition, fine structure and architecture. *J Cereal Sci.* 2004;39:151–65.
- Streb S, Zeeman SC. Starch metabolism in Arabidopsis. *Arabidopsis Book.* 2012;10:e0160.
- Stitt M, Zeeman SC. Starch turnover: pathways, regulation and role in growth. *Curr Opin Plant Biol.* 2012;15:282–92.

7. Ventriglia T, Kuhn ML, Ruiz MT, Ribeiro-Pedro M, Valverde F, Ballicora MA, Preiss J, Romero JM. Two Arabidopsis ADP-glucose pyrophosphorylase large subunits (APL1 and APL2) are catalytic. *Plant Physiol.* 2008;148:65–76.
8. Ferrero DML, Asencio Diez MD, Kuhn ML, Falaschetti CA, Piattoni CV, Iglesias AA, Ballicora MA. On the roles of wheat endosperm ADP-glucose Pyrophosphorylase subunits. *Front Plant Sci.* 2018;9:1498.
9. Crofts N, Sugimoto K, Oitome NF, Nakamura Y, Fujita N. Differences in specificity and compensatory functions among three major starch synthases determine the structure of amylopectin in rice endosperm. *Plant Mol Biol.* 2017;94:1–19.
10. Gamez-Arjona FM, Li J, Raynaud S, Baroja-Fernandez E, Munoz FJ, Ovecka M, Ragel P, Bahaji A, Pozueta-Romero J, Merida A. Enhancing the expression of starch synthase class IV results in increased levels of both transitory and long-term storage starch. *Plant Biotechnol J.* 2011;9:1049–60.
11. Wang YN, Li Y, Zhang H, Zhai H, Liu QC, He SZ. A soluble starch synthase I gene, *lssSI*, alters the content, composition, granule size and structure of starch in transgenic sweet potato. *Sci Rep.* 2017;7:2315–29.
12. Brummell DA, Watson LM, Zhou J, McKenzie MJ, Hallett IC, Simmons L, Carpenter M, Timmerman-Vaughan GM. Overexpression of *STARCH BRANCHING ENZYME II* increases short-chain branching of amylopectin and alters the physicochemical properties of starch from potato tuber. *BMC Biotechnol.* 2015;15:28–41.
13. Sun M, Lee HJ, Abdula SE, Jee MG, Cho YG. Overexpression of starch branching enzyme 1 gene improves eating quality in japonica rice. *J Plant Biochem Biot.* 2013;40:88–101.
14. Yan HB, Pan XX, Jiang HW, Wu GJ. Comparison of the starch synthesis genes between maize and rice: copies, chromosome location and expression divergence. *Theor Appl Genet.* 2009;119:815–25.
15. Van Harselaar JK, Lorenz J, Senning M, Sonnewald U, Sonnewald S. Genome-wide analysis of starch metabolism genes in potato (*Solanum tuberosum* L.). *BMC Biotechnol.* 2017;18:37–54.
16. Nakamura Y. Rice starch biotechnology: Rice endosperm as a model of cereal endosperms. *Starch-Stärke.* 2018;70:1600375–94.
17. Wang L, Fu JL, Li M, Fragner L, Weckwerth W, Yang PF. Metabolomic and proteomic profiles reveal the dynamics of primary metabolism during seed development of Lotus (*Nelumbo nucifera*). *Front Plant Sci.* 2016;7:750.
18. Zhu FL, Sun H, Diao Y, Zheng XW, Xie KQ, Hu ZL. Genetic diversity, functional properties and expression analysis of *NnSBE* genes involved in starch synthesis of lotus (*Nelumbo nucifera* Gaertn.). *PeerJ.* 2019;7:e7750.
19. Ball SG, Morell MK. From bacterial glycogen to starch: understanding the biogenesis of the plant starch granule. *Annu Rev Plant Biol.* 2003;54:207–33.
20. Hwang SK, Singh S, Maharana J, Kalita S, Tuncel A, Rath T, Panda D, Modi MK, Okita TW. Mechanism underlying heat stability of the Rice endosperm cytosolic ADP-glucose Pyrophosphorylase. *Front Plant Sci.* 2019;10:70.
21. Ballicora MA, Iglesias AA, Preiss J. ADP-glucose Pyrophosphorylase: a regulatory enzyme for plant starch synthesis. *Photosynth Res.* 2004;79:1–24.
22. Lin TP, Caspar T, Somerville C, Preiss J. Isolation and characterization of a starchless mutant of Arabidopsis thaliana (L.) Heynh lacking ADPglucose pyrophosphorylase activity. *Plant Physiol.* 1988;86:1131–5.
23. Tetlow IJ, Davies EJ, Vardy KA, Bowsher CG, Burrell MM, Emes MJ. Subcellular localization of ADPglucose pyrophosphorylase in developing wheat endosperm and analysis of the properties of a plastidial isoform. *J Exp Bot.* 2003;54:715–25.
24. Comparot-Moss S, Denyer K. The evolution of the starch biosynthetic pathway in cereals and other grasses. *J Exp Bot.* 2009;60:2481–92.
25. Radchuk VV, Borisjuk L, Sreenivasulu N, Merx K, Mock HP, Rolletschek H, Wobus U, Weschke W. Spatiotemporal profiling of starch biosynthesis and degradation in the developing barley grain. *Plant Physiol.* 2009;150:190–204.
26. Batra R, Saripalli G, Mohan A, Gupta S, Gill KS, Varadwaj PK, Balyan HS, Gupta PK. Comparative analysis of AGPase genes and encoded proteins in eight monocots and three dicots with emphasis on wheat. *Front Plant Sci.* 2017;8:19.
27. Johnson PE, Patron NJ, Botttrill AR, Dinges JR, Fahy BF, Parker ML, Waite DN, Denyer K. A low-starch barley mutant, Riso 16, lacking the cytosolic small subunit of ADP-glucose pyrophosphorylase, reveals the importance of the cytosolic isoform and the identity of the plastidial small subunit. *Plant Physiol.* 2003;131:684–96.
28. Sun H, Hu ML, Li JY, Chen L, Li M, Zhang SQ, Zhang XL, Yang XY. Comprehensive analysis of NAC transcription factors uncovers their roles during fiber development and stress response in cotton. *BMC Plant Biol.* 2018;18:150.
29. Zheng YM, Li ZY, Zhang C, Zheng BD, Tian Y. Effects of microwave-vacuum pre-treatment with different power levels on the structural and emulsifying properties of lotus seed protein isolates. *Food Chem.* 2020;311:125932.
30. Ali NA, Dash KK, Routray W. Physicochemical characterization of modified lotus seed starch obtained through acid and heat moisture treatment. *Food Chem.* 2020;319:126513.
31. Yen GC, Duh PD, Su HJ. Antioxidant properties of lotus seed and its effect on DNA damage in human lymphocytes. *Food Chemistry Chem.* 2005;89:379–85.
32. Mukherjee D, Khatua TN, Venkatesh P, Saha BP, Mukherjee PK. Immunomodulatory potential of rhizome and seed extracts of *Nelumbo nucifera* Gaertn. *J Ethnopharmacol.* 2010;128:490–4.
33. Juliano BO. Structure, chemistry, and function of the rice grain and its fractions. *Cereal Food World.* 1992;37:772–9.
34. Wang Q, Zheng YF, Zhuang WJ, Lu X, Luo XL, Zheng BD. Genome-wide transcriptional changes in type 2 diabetic mice supplemented with lotus seed resistant starch. *Food Chem.* 2018;264:427–34.
35. Zeng HL, Zheng YX, Lin Y, Huang CC, Lin S, Zheng BD, Zhang Y. Effect of fractionated lotus seed resistant starch on proliferation of Bifidobacterium longum and Lactobacillus delbrueckii subsp bulgaricus and its structural changes following fermentation. *Food Chem.* 2018;268:134–42.
36. Zeng SX, Wu XT, Lin S, Zeng HL, Lu X, Zhang Y, Zheng BD. Structural characteristics and physicochemical properties of lotus seed resistant starch prepared by different methods. *Food Chem.* 2015;186:213–22.
37. Hou J, Li T, Wang YM, Hao CY, Liu HX, Zhang XY. ADP-glucose pyrophosphorylase genes, associated with kernel weight, underwent selection during wheat domestication and breeding. *Plant Biotechnol J.* 2017;15:1533–43.
38. Fan CF, Wang GY, Wang YM, Zhang R, Wang YT, Feng SQ, Luo KM, Peng LC. Sucrose synthase enhances Hull size and grain weight by regulating cell division and starch accumulation in transgenic Rice. *Int J Mol Sci.* 2019;20:4971.
39. Zhang X, von Mogel KJH, Lor VS, Hirsch CN, De Vries B, Kaeppler HF, Tracy WF, Kaeppler SM. Maize sugary enhancer1 (*se1*) is a gene affecting endosperm starch metabolism. *P Natl Acad Sci USA.* 2019;116:20776–85.
40. Zhu JH, Yu WW, Zhang CQ, Zhu Y, Xu JJ, Li EP, Gilbert RG, Liu QQ. New insights into amylose and amylopectin biosynthesis in rice endosperm. *Carbohydr Polym.* 2020;230:115656.
41. Saripalli G, Gupta PK. AGPase: its role in crop productivity with emphasis on heat tolerance in cereals. *Theor Appl Genet.* 2015;128:1893–916.
42. Shi JP, Lai JS. Patterns of genomic changes with crop domestication and breeding. *Curr Opin Plant Biol.* 2015;24:47–53.
43. Ezer D, Shepherd SJK, Brestovitsky A, Dickinson P, Cortijo S, Charoensawan V, Box MS, Biswas S, Jaeger KE, Wigge PA. The G-Box transcriptional regulatory code in Arabidopsis. *Plant Physiol.* 2017;175:628–40.
44. Kim KN, Fisher DK, Gao M, Gultinan MJ. Genomic organization and promoter activity of the maize starch branching enzyme I gene. *Gene.* 1998;216:233–43.
45. Song YH, Luo GB, Shen LS, Yu K, Yang WL, Li X, Sun JZ, Zhan KH, Cui DQ, Liu DC, Zhang AM. *TubZIP28*, a novel bZIP family transcription factor from Triticum urartu, and *TabZIP28*, its homologue from Triticum aestivum, enhance starch synthesis in wheat. *New Phytol.* 2020;226:1384–98.
46. Cheng LB, Liu X, Yin JJ, Yang JQ, Li Y, Hui LC, Li SY, Li LJ. Activity and expression of ADP-glucose pyrophosphorylase during rhizome formation in lotus (*Nelumbo nucifera* Gaertn.). *Bot Stud.* 2016;57:26.
47. Gibson K, Park JS, Nagai Y, Hwang SK, Cho YC, Roh KH, Lee SM, Kim DH, Choi SB, Ito H, Edwards GE, Okita TW. Exploiting leaf starch synthesis as a transient sink to elevate photosynthesis, plant productivity and yields. *Plant Sci.* 2011;181:275–81.
48. Kang GZ, Liu GQ, Peng XQ, Wei LT, Wang CY, Zhu YJ, Ma Y, Jiang YM, Guo TC. Increasing the starch content and grain weight of common wheat by overexpression of the cytosolic AGPase large subunit gene. *Plant Physiol Biochem.* 2013;73:93–8.
49. Oiestad AJ, Martin JM, Giroux MJ. Yield increases resulting from AGPase overexpression in rice are reliant on plant nutritional status. *Plant Growth Regul.* 2019;89:179–90.
50. Li N, Zhang SJ, Zhao YJ, Li B, Zhang JR. Over-expression of AGPase genes enhances seed weight and starch content in transgenic maize. *Planta.* 2011;233:241–50.
51. Xie C, Mao XZ, Huang JJ, Ding Y, Wu JM, Dong S, Kong L, Gao G, Li CY, Wei LP. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* 2011;39:W316–22.
52. Chen CJ, Chen H, Zhang Y, Thomas HR, Frank MH, He YH, Xia R. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant.* 2020;13:1194–202.

53. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 2016;33:1870–4.
54. Sun H, Chen L, Li JY, Hu ML, Ullah A, He X, Yang XY, Zhang XL. The JASMONATE ZIM-domain gene family mediates JA signaling and stress response in cotton. *Plant Cell Physiol.* 2017;58:2139–54.
55. Yang ZH. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 2007;24:1586–91.
56. Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouze P, Rombauts S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 2002;30:325–7.
57. Xu CJ, Chen WJ, Chen KS, Zhang SL. A simple method for determining the content of starch-iodine colorimetry. *Biotechnology.* 1998;8:41–3.
58. Bradford MM. Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding. *Anal Biochem.* 1976;72:248–54.
59. Lv QJ, Yi JR, Ding Y. Determination of polyphenol in Carambola by Folin-Ciocalteu Colorimetry. *Hunan Agric Sci.* 2009;7:99–101.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

