Genetic diversity, functional properties and expression analysis of *NnSBE* genes involved in starch synthesis of lotus (*Nelumbo nucifera* Gaertn.)

Fenglin Zhu [1, 2, 3], **Han Sun** [1, 2, 3], **Ying Diao** [1, 3], **Xingwen Zheng** [1, 2, 3, 4], **Keqiang Xie** [4], **Zhongli Hu** [Corresp. 1, 2, 3]

[1] College of Life Sciences, Wuhan University, Wuhan, China

[2] State Key Laboratory of Hybrid Rice, Wuhan, China

[3] Hubei Lotus Engineering Center, Wuhan, China

[4] Guangchang Bailian Institute of Jiangxi Province, Guangchang, China

Corresponding Author: Zhongli Hu
Email address: huzhongli@whu.edu.cn

Background Starch branching enzyme (SBE) is one of the key enzymes in starch biosynthetic metabolism, determining amylopectin structure.

Methods Full length coding sequences (CDS) of *SBE* genes were cloned using reverse transcription PCR (RT-PCR) technology, and neighbor-joining (NJ) tree was used for phylogenetic analysis. Single nucleotide polymorphisms (SNPs) were determined to assess the genetic polymorphisms and variation indexes between individuals and clusters. Quantitative real time PCR (qRT-PCR) was performed to analyze the spatial and temporal expression of *NnSBE* genes. The effect of *NnSBE* genes on amylopectin's fine structures was explored using affinity and the enzyme activity analysis of two isoforms in amylopectin and amylose.

Results In this study, two SBE family genes, *NnSBEI* and *NnSBEIII*, were identified in lotus (*Nelumbo nucifera* Gaertn.). Phylogenetic analysis sorted NnSBEI into SBE family B and NnSBEIII into SBE family A. UPGMA phylogenetic tree divided 45 individuals of lotus into three classes. The homozygous haplotype (A G G A G) of *NnSBEIII* was observed in seed lotus. During the seed embryo development stage, *NnSBEIII* reached the peak in the middle of the development stage, while *NnSBEI* increased in the mid-late developmental stage. The different affinity activity of the two isozymes binding amylopectin and amylose assay indicated NnSBEI has higher activity and wider affinity.

Discussion Genetic diversity showed that *NnSBE* genes received artificial selection during the process of cultivation and domestication in lotus seeds. Furthermore, the expression pattern and affinity activity analysis indicated that *NnSBE* genes were related to the chain length of amylopectin.

1 **Genetic diversity, functional properties and expression analysis of *NnSBE* genes involved**

2 **in starch synthesis of lotus (*Nelumbo nucifera* Gaertn.)**

3 Fenglin Zhu [1, 2, 3], Han Sun [1, 2, 3], Ying Diao [1, 3], Xingwen Zheng [1,2,3,4], Keqiang Xie [4], Zhongli

4 Hu[Corresp. 1, 2, 3]

5 1 College of Life Sciences, Wuhan University, Wuhan, China

6 2 State Key Laboratory of Hybrid Rice, Wuhan, China

7 3 Hubei Lotus Engineering Center, Wuhan, China

8 4 Guangchang Bailian Institute of Jiangxi Province, Guangchang, China

9 *Correspondence: Zhongli Hu. Fax: +86-27-68753606 E-mail huzhongli@whu.edu.cn

10 **Abstract**

11 **Background** Starch branching enzyme (SBE) is one of the key enzymes in starch biosynthetic

12 metabolism, determining amylopectin structure.

13 **Methods** Full length coding sequences (CDS) of *SBE* genes were cloned using reverse

14 transcription PCR (RT-PCR) technology, and neighbor-joining (NJ) tree was used for

15 phylogenetic analysis. Single nucleotide polymorphisms (SNPs) were determined to assess the

16 genetic polymorphisms and variation indexes between individuals and clusters. Quantitative real

17 time PCR (qRT-PCR) was performed to analyze the spatial and temporal expression of *NnSBE*

18 genes. The effect of *NnSBE* genes on amylopectin's fine structures was explored using affinity

19 and the enzyme activity analysis of two isoforms in amylopectin and amylose.

20 **Results** In this study, two SBE family genes, *NnSBEI* and *NnSBEIII*, were identified in lotus

21 (*Nelumbo nucifera* Gaertn.). Phylogenetic analysis sorted NnSBEI into SBE family B and

22   NnSBEIII into SBE family A. UPGMA phylogenetic tree divided 45 individuals of lotus into

23   three classes. The homozygous haplotype (A G G A G) of *NnSBEIII* was observed in seed lotus.

24   During the seed embryo development stage, *NnSBEIII* reached the peak in the middle of the

25   development stage, while *NnSBEI* increased in the mid-late developmental stage. The different

26   affinity activity of the two isozymes binding amylopectin and amylose assay indicated NnSBEI

27   has higher activity and wider affinity.

28   **Discussion** Genetic diversity showed that *NnSBE* genes received artificial selection during the

29   process of cultivation and domestication in lotus seeds. Furthermore, the expression pattern and

30   affinity activity analysis indicated that *NnSBE* genes were related to the chain length of

31   amylopectin.

32   **Introduction**

33   Lotus (*Nelumbo nucifera* Gaertn) is an ancient perennial aquatic plant and important crop in

34   Asia. Archaeological research has estimated that the history of lotus is more than 7,000 years

35   old, and it has been cultivated for more than 2000 years (Guo 2009; Shen-Miller 2002). In

36   general, lotus are grouped into three clusters corresponding with the three different important

37   organs, namely ornament lotus, seed lotus and rhizome lotus, respectively (Guo 2009; Wu et al.

38   2007). As a type of aquatic plant with high photosynthetic efficiency and high carbon

39   conversion, lotus has high starch content. Because of their high amount of starch, the edible

40   rhizomes and seeds play a key role in a daily diet and cultural activities (Chen et al. 2007).

41   Studies have shown that fresh rhizomes, on average, are comprised of 10–20% starch in their

42   total fresh weight, while the amount is 40–60% in mature seeds (Shen-Miller et al. 1995).

43   Starch is an important polysaccharide and the major form of carbohydrate storage in plants

44   (Slattery, Kavakli, Okita 2000). It is a necessary part of the human diet in terms of nutrition and

45   calories. Starch is comprised of two glucan polymers, amylose and amylopectin, which have

46   different characteristics because of their starch molecular structure. Amylose and amylopectin

47   synthesis are regulated by the coordinated action of a series of enzymes. AGPase produces

48   substrate that plays a role in the synthesis of amylose and amylopectin. Amylose synthesis is

49   mainly controlled by granule-bound starch synthase (GBSS), while amylopectin is generated by

50   the successive work of starch synthase (SS), starch branching enzyme (SBE), and debranching

51   enzyme (DBE) (Fujita et al. 2008; Jeon et al. 2010; Subasinghe et al. 2014; Tetlow Ian, Emes

52   Michael 2014).

53   Starch-branching enzymes, formerly known as Q-enzymes, have dual catalytic functions and

54   determine the structure of amylopectin. SBE cleaves the internal a-1,4 linkage from polyglucans

55   and then transfers the reducing ends to C-6 hydroxyls to generate a-1,6-branch linkage in the

56   liner chain (Martin, Smith 1995). Their catalytic function not only catalyzes the formation of a

57   new branch, but they also add new nonreducing ends in the starch molecule. Thus, SBE

58   determines the branching pattern in amylopectin, which is believed to affect the fine structure of

59   plants and influence the amount of starch (Satoh et al. 2003).

60   SBEs have been researched in various plants. In particular, two cDNA coding *PsSBE*s were

61   identified from the embryo of *Pisum sativum* by Burton et al (Burton Rachel et al. 1995). Later,

62   two or three SBE members were identified in various plants and were classified into SBE A and

63   SBE B based on phylogenetic analysis. SBE A and SBE B play different roles in the influence of

64   amylopectin synthesis. In the process of starch synthesis, the SBE A family tends to

65   amylopectin, while SBE B family show higher affinity for amylose. SBE A prefers amylose as a

66   substrate and predominantly transfers relatively longer chains (>14DP), while SBE B tends to

67   amylopectin and transfers shorter chains (<14DP) (Guan et al. 1997). In addition, different gene

68   expression patterns of *SBE A* and *SBE B* were detected with a range of species, indicating that

69   *SBE B* is expressed earlier than *SBE A* in the development stage (Gao et al. 1996; Larsson et al.

70   1998; Morell et al. 1997; Mutisya et al. 2003).

71   To fit the rapid development of the lotus processing industry, the work of breeding lotus

72   with high starch content is extremely urgent. Therefore, many studies have focused on the starch

73   of lotus. The *ADPase* and *GBSS* genes of lotus were isolated and characterized, but little

74   information is currently available about the *NnSBE* gene (Cheng et al. 2014; Liang-Jun et al.

75   2006; Lu et al. 2012). For researching the *NnSBE*s further, this study undertook the isolation of

76   cDNA and genomic clones by encoding two related *SBE* genes from lotus. Then, genomic

77   variation and evolution were analyzed using DNA and protein sequences of the two NnSBEs.

78   Expression patterns and the affinity of two isozymes were described, as well as the enzyme

79   activity of both NnSBEs in development and different tissues. These data provide valuable

80   information for understanding the processes involved in starch synthesis, and offer some

81   fundamental information for further study about improving the edible quality of lotus.

82   **Material and methods**

83   **Plant material and treatments**

84   *Nelumbo nucifera* cv. Taikong lotus 36, the highest strain of selective lotus breed after space

85   mutagenesis, was selected in this study (Wu et al. 2007). Seeds were sprouted by soaking in

86   water for germination. Five days after, plants were provided with 50 cm depth pots in a

87   greenhouse for the entire growing season. Functional leaves, petiole, rhizomes and roots were

88   separately collected in the 8th week, 10th week, 12th week. Seeds from plants at 12, 16, 20, 24 and

89   28 DAF (days after fertilization) were also collected from the Taikong lotus 36 in the genetic

90   experimental base of Wuhan University. Various materials from lotus were quick-frozen in

91   liquid nitrogen, then stored at -80 °C for next manipulation. The fresh leaves of 45 lotus

92   individuals (Supplement 1) were collected from the genetic experimental base of Wuhan

93   University and stored in silica gel (Sinopharm, China).

94   **RNA isolation and cDNA synthesis**

95   Total RNA and genomic DNA were isolated from plant tissue samples using a plant RNA

96   extraction kit and Plant Genomic DNA Kit (TIANGEN China), according to their respective

97   operation manuals. The quantities were subjected to 1% agarose gel electrophoresis and the

98   results were analyzed using a UV Transilluminator (Eppendorf, Germany). The first-strand

99   cDNA of all materials was constructed using the FastKing RT Kit (with gDNase) (TIANGEN,

100   China). The products were stored at -20 °C for later use.

101   **Characterization of a genetic polymorphism**

102   Genetic polymorphism of *NnSBE* cDNA from 45 individuals of lotus were also analyzed using

103   the gene-specific primers. PCR reactions were conducted in 50ul volumes containing 2 ul of total

104   cDNA, 5 ul of 10×PCR buffer, 5 ul of 2 mM dNTPs mixture, 3ul of 25 mM $MgSO_4$, 1.6 ul of 10

105   pM of each primer, 1 ul of 1 U/μl KOD DNA polymerase (TOYOBO, Japan) and 30.8 μl ddH_2O,

106   for a total volume of 50 μl. Amplification conditions followed the two-step amplification

107   procedure: 94 ℃ for 2 min, 36 cycles of 98 °C for 10 s, (Tm) ℃ for 30 s, and 68 °C for 1 min.

108   The sizes of the PCR products were assessed using 1.0% agarose gel electrophoresis. PCR

109   products were sequenced by Sanger sequencing (Augct, China). Sequences were aligned by

110   DNAman to find molecular markers.

111   **Bioinformatics analysis**

112   All sequence data were obtained from the National Center for Biotechnology Information

113   (NCBI) GenBank database (http://www.ncbi.nlm.nih.gov/). All primers were designed by Primer

114   Premier 5.0v. Genomic structures were performed with the program Gene Structure Display

115   Server 2.0 (http://gsds.cbi.pku.edu.cn/) (Hu et al. 2015). The GO (Gene Ontology,

116   http://www.geneontology.org) was used for annotation of gene products. The conserved domains

117   were predicted by Pfam sever (http://pfam.xfam.org/search/sequence;) (Finn et al. 2016). All

118   sequences were aligned using Cluxal-X(Thompson et al. 1997) and DNAman. Subsequently the

119   phylogenetic neighbor-joining tree was generated by MEGA6.0v (Tamura et al. 2013). The

120   bootstrap consensus tree inferred from 1,000 replicates was taken to depict the evolutionary

121   history of the analyzed taxa. The H_O, H_E, Shannon Index was calculated by popGen32 (Nei

122   1978; Nei, Li 1973; Nei, Li 1979). A dendrogram of the cluster analysis was based on Nei's

123   genetic distance using the UPGMA method.

124   **Real-time PCR analysis**

125   The cDNA of reverse transcription was diluted for 1:10 and used for RT-PCR and qRT-PCR

126   assays. A house-keeping gene, *CYP* (cyclophilin, GenBank accession no.EU131153), was

127   selected as the reference gene in this experiment. The primers of *CYP* were based on the

128   sequence of CDS and designed by Primer Primer 5.0 (Table 1). The total PCR reaction mixture

129   contained 2 μl cDNA (1:10 diluted), 0.4 μl forward primer (10 M), 0.4 μl reverse primer (10 M),

130   10 μl 2×SYBR qPCR Master Mix (with ROX Premixed) (Vazyme, China) and 7.2 μl ddH$_2$O, for

131   a total volume of 20 μl. For the qRT-PCR experiment, the two-step amplification procedure was

132   used: 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. The relative

133   gene expression data was calculated using the $2^{-\Delta\Delta Ct}$ method with the guidance of the StepOne

134   software v2.1 (ABI, US). The experimental design followed MIQE (minimum information for

135   publication of quantitative real-time PCR experiments) (Bustin et al. 2010; Bustin et al. 2009).

136   All measurements were processed three times for biological and technical paralleled repetition.

137   **Construction of plasmids**

138   The full length CDS of *NnSBEI* and *NnSBEIII*, which were cloned from lotus, were subcloned

139   into the plasmids pET-28a and pET-32a respectively. The pET-28a-*NnSBEI* was generated using

140   the following primers:5'- CCCAAGCTTATGTACAGTTTTTCTGGGT-3' (HindIII site

141   underlined), and 5'-CATGCCATGGTCAGTCATCCAATCCCA-3' (Nco I site underlined).

142   The pET-32a-*NnSBEIII* was obtained by primers: 5'-

143   CGGGATCCATGGCTACTACAGTTGCGCT-3' (BamHI site underlined) and 5'-

144   GCGATATCTCATATTCGCAAAATCCGAG-3' (EcoR V site underlined). The full length and

145   sequence of the nucleotide acid of the inserted gene in the extracted recombinant plasmid were

146   completely identical to the genes. Then, the different recombinant plasmids were transformed

147   into Escherichia coli BL21 (DE3) (TransGen Biotech, China) for expressing and into DH5α for

148   storing.

149   **Enzyme activity assay**

150   Cheng et al.'s protocol of measuring SBE activity was used in this study(Cheng et al. 2001).

151   First, total enzyme isolation: about 0.4 g powder of plant material was added into a 4 ml

152   HEPES–NaOH (PH7.5) buffer. After proper shaking, the tube was centrifuged at 10000 g for 15

153   min, and supernatant was extracted to check the enzyme activity of SBE. We took 100 μl

154   dilution, added 1280 ul of HEPES–NaOH (PH7.5) buffer, and 120 ul 0.75% of soluble starch.

155   Then we incubated the total reaction mixture at 37 ℃, and gave it a water bath for 20 min. Next,

156   we gave it a boiling water bath for 1 min to stop the reaction, diluted the mixture with 2 ml water

157   (containing 0.2% HCl), added 150 ul of an iodine solution (0.1%I_2-1%KI), blending it at room

158   temperature for 10 min. The boiled crude enzyme was used as the control. Finally, we measured

159   absorbance at 660 nm. The enzyme activity of SBE was expressed as a percentage of decrease in

160   absorbance at the wavelength of 660 nm to be compared with the control. All measurements

161   were processed three times for biological and technical paralleled repetition.

162   **The affinity of two isozymes that bind amylopectin and amylose assay**

163   Recombinant proteins were expressed in the Escherichia coli BL21(DE3) cells. The transformed

164   cells were grown in LB at 37 ℃ until the OD600 reached 0.6-1.0, using IPTG to induce

165   expression with a final concentration of 0.5 mM, and were further incubated at 16 ℃ for 12 h.

166   Cells were harvested by centrifugation at 12000 rpm for 5 min at 4 ℃. The products were

167   suspended in a buffer containing 10 mM PBS in the radio of 200:1 (ml/g), and cells were

168   disrupted (Diagenode, Belgium) by sonication for 25 min and centrifuged for 5 min at 4 ℃. We

169   took 30 µl supernatant, added 0, 12.5, 16, 20, 25, 35 ul 0.75% amylopectin (Sigma, America) or

170   amylose (Sigma, America), and added HEPES–NaOH (PH7.5) buffer into 90 ml. Then the total

171   reaction mixtures were incubated at 37 ℃ for 30 min. We diluted the mixture with 100 ul water

172   (containing 0.2% HCl), and added 10 ul iodine solution ($0.1\%I_2$-$1\%KI$), blending at room

173   temperature for 10 min. Finally, we measured absorbance of amylopectin at 560 nm and amylose

174   at 620 nm. The blank pET-28a and pET-32a which were transformed into BL21 and cultured in

175   the same condition were used as the control. The affinity of the two isozymes that bind

176   amylopectin and amylose were measured by enzyme activity. All measurements were processed

177   three times for technical repetition and 2~3 paralleled repetition.

178   **Results**

179   **Cloning of *NnSBE* genes**

180   Based on the reference sequence of whole-genome sequencing, the CDS of two *SBE* genes were

181   obtained using RT-PCR technique from the embryo of the lotus seed. The CDS of *NnSBEI*

182   (FJ592190.1) was 2577bp, and the CDS of *NnSBEIII* (XM_010254179) was 2691bp (Fig. 1).

183   Complete genomic structures of *NnSBEI* (comprising of 14 exons and 13 introns) and *NnSBEIII*

184   (comprising of 21 exons and 20 introns) were separately distributed over 14.3 and 32.1kb. The

185   gene structures are shown in figure 2 and the gene information is shown in Table 2. The cDNA

186   of *NnSBEI* showed the highest (85%) identity with *Castanea mollissima*, and 80–84% identify

187   with *SBEI* genes of other plant species, while *NnSBEIII* showed the highest (85%) identity with

188   *Juglans regia* and 80–84% identity with the others.

189   The deduced protein of NnSBEI comprised 858 amino acids with a predicted molecular

190   mass of 97.144 kDa, and NnSBEIII comprised 896 amino acids with a predicted molecular mass

191   of 103.135 kDa. GO analysis indicated NnSBE participated in biological processes

192   (GO:0005978) and 1 molecular function (GO:0043169; GO:0003844; GO:0004553). Pfam

193   analysis showed that NnSBEs have three domains of secondary structure: the central catalytic A

194   domain, an N-terminal domain and the C-terminal domain. The central catalytic A-domain of

195   SBEs is α-amylase which covered four conserved amino acid regions and six active sites spread

196   in those conserved regions respectively. The characterization of catalytic A-domain showed

197   significant homology (47.83%) between NnSBEI and NnSBEIII, but highly dissimilar sequences

198   in N-terminus (21.84%) and C-terminus (26.47%).

199   **Phylogenetic analysis**

200   SBE protein sequences from 20 other species were used for phylogenetic analysis, shown in

201   figure 3. The NJ phylogenetic tree indicates that SBEs in these plants could be classified into two

202   families: SBE A and SBE B. SBE A was further divided into four classes: algae SBEII, dicot

203   SBEII, monocot SBEII and SBEIII, and SBE B was subdivided into algae, dicot SBEI, monocot

204   SBEI. NnSBEI belonged to the dicot SBEI of family B; NnSBEIII belonged to the SBEIII class

205   of family A. Evolutionary divergence reveal that SBEIII came before SBEI, and differentiation

206   of NnSBE occurred before that of the other dicots, but later than the monocots.

207   **Characterization of genetic diversity**

208   Genetic diversity was identified based on the ORF of two *NnSBE* genes from 45 lotus

209   individuals belonging to four clusters (Supplement 1). This revealed six polymorphic sites from

210   the coding region: five SNPs in *NnSBEIII* and one in *NnSBEI*. The SNPs of *NnSBEIII* was

211   analyzed and the detailed parameters of genetic diversity are listed in Table 3. Three SNPs

212   resulted in missense mutations which changed the amino acid sequence, while the other two

213   SNPs were synonymously mutated. The observed heterozygosity ranged from 0.2 to 0.4444, and

214   the expected heterozygosity ranged from 0.4012 to 0.4583. The Shannon-Wiener Index ranged

215   from 0.5908 to 0.6508. In further association analysis of starch content (amylopectin, amylose,

216   total starch) in lotus seeds, no SNP showed highly significant associations. Total of 12 genotypes

217   were detected in those 45 individuals, each cluster had different several kinds of genotypes

218   (Supplement 1). A homozygous genotype (AA GG GG AA GG) of *NnSBEIII* was observed in

219   most individuals in seed lotusand a haplotype (A G G A G) was identified from seed lotus.

220   Cluster analysis according to the UPGMA method is shown in figure 4. The 45 individuals of

221   lotus were copolymerized into three classes. The first class can be divided into two subclasses

222   with a similarity coefficient of 9.8, and most seed lotus showed together in a same subgroup.

223   These results paved a way to apply the useful allelic variations or gene haplotypes in cultivar

224   lotus and quality breed programs.

225   **Expression pattern**

226   The temporal and spatial expression of both genes were analyzed to investigate their expression

227   patterns. The results of qRT-PCR demonstrated that the transcriptional expression level of

228   *NnSBEI* and *NnSBEIII* was detectable in all tissues. Tissue-specific expression analysis showed

229   that the highest transcript levels of *NnSBEI* and *NnSBEIII* were observed in leaves. The

230   transcript level of *NnSBEI* was higher in rhizomes while *NnSBEIII* expressed strongly in

231   petioles. The expression profile of *NnSBEI* in the rhizome showed strong temporal differences,

232   and the relative expression level increased gradually from initial development to late stage.

233   *NnSBEIII* showed temporal differences in the petiole, which enhanced in the middle and

234   decreased in the late stage (Fig.5). During the seed embryo developing stage, *NnSBEI* and

235   *NnSBEIII* expressed significant differences. Transcripts of *NnSBEIII* reached the peak at 20 DAF

236   in the middle of development stage, and then decreased gradually, while *NnSBEI* increased in 16

237   DAF to 28 DAF and expressed strongly in the mid-late developmental stage (Fig.6). In the peak

238   stage, *NnSBEI* expressed ten times higher than *NnSBEIII*.

239   **Dynamic changes of enzyme activities**

240   The dynamic changes of enzyme activity of the SBEs were analyzed in different temperatures,

241   tissues and developing stages. Incubation in different temperatures revealed that the highest

242   enzyme activity was generated under 37 ℃. The enzyme activity of inter-organizational

243   measurement and assessment showed the highest catalytic activity in leaves, followed by

244   rhizomes and petioles which increased steadily, and activity in the roots was the weakest. During

245   seed development, enzyme activity increased rapidly, and the activity peak appeared at 20 DAF,

246   then decreased slightly (Fig. 7).

247   **The affinity of two isozymes that bind amylopectin and amylose**

248   To explore the affinity activity, recombinant DNA techniques were used to generate the pET-

249   28a-*NnSBEI* and pET-32a-*NnSBEIII* plasmids. Testing of the inducible expression vector

250   showed that both proteins were soluble, and the apparent molecular weights of pET-28a-NnSBEI

251   and pET-32a-NnSBEIII were about 105 kDa and 120 kDa (with His tag) respectively. The

252   affinity activity of NnSBEI and NnSBEIII to amylopectin and amylose were assayed at various

253   starch concentrations. As we can see in figure 8, the activity of both isozymes correlated with the

254   types and concentration of substrate. Rate of reaction changed with the increase of the substrate

255   concentration until the enzyme was saturated. NnSBEI showed higher affinity activity when the

256   substrates were amylose and amylopectin. However, another isozyme, NnSBEIII, only worked

257   on amylopectin, and the branching efficiency was only half of NnSBEI.

258   **Discussion**

259   Starch is an important edible component in lotus, and the starch branch enzyme is involved in

260   the synthesis of amylopectin. In this study, we focused on the SBE family genes in lotus, and

261   revealed two isoforms of *NnSBE*s: *NnSBEI* and *NnSBEIII*. In this study, SBE phylogenetic tree

262   included the SBE gene families from algae, monocots and dicots according to the protein

263   sequences of the sequenced species. Two or three SBE isoforms were isolated from different

264   plants, and divided into two families: family A and family B (Fig. 4). NnSBEI belonged to

265   family B, which has been reported in a range of species. NnSBEIII belonged to family A, which

266   has only been identified in maize, wheat, rice and lotus, and little research has been done about

267   this isoform (Kang et al. 2013; Tian et al. 2009).

268   Over the last decades, many investigations have been devoted to exploring genomic variation

269   and evolution among different germplasms. Whole genome re-sequencing reveals the

270   evolutionary patterns of sacred lotus, rhizome lotus had the lowest genomic diversity and a

271   closer relationship to wild lotus, whereas the genomes of seed lotus and ornament lotus were

272   admixed (Huang et al. 2018). In this study, genetic diversity and genetic variation of *NnSBEIII*

273   were investigated among 45 individuals from four lotus subgroups. Wild lotus had the higher

274   genomic diversity, rhizome lotus was admixed with wild and flower lotus respectively, and lotus

275   seed formed a homozygous single haplotype, which might be the result of continuous selection

276   in the cultivation process of lotus in the long-term evolution process (Doebley, Gaut, Smith

277   2006; Yue, Melamud, Moult 2006). This different performance speculates that the *NnSBE* gene

278   has more effect on starch accumulation in seed lotus, but less in rhizome lotus.

279   Starch synthesis and accumulation are closely related to photosynthesis. It was found that the

280   *AGPase* and *GBSS* genes of lotus, related to starch synthesis, were expressed higher in leaves

281   (Cheng et al. 2014; Lu et al. 2012). As the tissue of photosynthesis, leaves support the first step

282   of synthesizing sugar and meet the carbon demand faster and it is the major tissue for

283   accumulating transient starch. The results of qRT-PCR showed that the two *NnSBE*s were

284   expressed throughout plant tissues. *NnSBEI* was strongly activated in the leaves and petiole;

285   *NnSBEIII* was highly expressed in the leaves and stems (Fig. 4). The transcriptional level kept

286   rising during the swelling stage in the petiole and rhizome, which represent the photosynthetic

287   tissue and storage tissues respectively. It is possible to adapt to the synthesis of transient starch in

288   photosynthetic tissue, and it is consistent with reserve starch synthesis in storage organs during

289   starch rapid development. This spatial and temporal variation from top to bottom may be related

290   to the process and the transport of starch synthesis to achieve higher efficiency of starch

291   accumulation. Such expression patterns of temporal and spatial controls have also been found in

292   many other species (Tetlow 2010). This spatial and temporal expression pattern of *NnSBE*s help

293   amylopectin to adapt requirements at different developmental stages.

294   During evolution, SBE A family is inclined to the amylopectin branch, while SBE B family

295   is more sensitive to amylose. Likewise, this experiment showed that NnSBEI has higher catalytic

296   activity for amylose and amylopectin, while NnSBEIII expressed catalytic activity only when the

297   substrate was amylopectin. Protein sequences of the SBE genes among a range of species

298   revealed that the SBEI and SBEIII subunits showed homology in their α-amylase catalytic

299   domain, but highly dissimilar sequences in N-terminus and C-terminus. Construction of chimeric

300   enzymes out of maize branching enzymes found that the N-terminal determined the specificity of

301   the transferred chain length, and the C-terminal domain participates in the specificity of the

302   substrate (Commuri, Keeling 2001; Hong, Mikkelsen, Preiss 2001; Kuriki, Stewart, Preiss 1997).

303   Therefore the N-terminal region of SBE is essential for maximum enzyme activity and

304   thermostability (Hamada et al. 2007), and C-terminus determines both substrate preference and

305   maximal catalytic activity (Hong, Preiss 2000). We speculated that the difference between

306   enzymatic characteristics of the two subunits are determined by the N-terminus and C-terminus.

307   These differences met the requirements of amylopectin to produce different grades of branches

308   and different sizes of side chains.

309   Expression and protein structure-activity analysis in the plant kingdom indicated that SBE

310   plays a critical role in affecting the fine structure of amylopectin (Satoh et al. 2003). In another

311   study of developing lotus seeds, a similar conclusion was drawn. In this study, we learned that

312   NnSBEI belong to the SBE B family and tend to transfer longer chains, while NnSBEIII

313   belonged to the SBE A family and prioritized short chain. Higher catalytic activity and

314   expression level of NnSBEI indicated that the long chain was preferred in the transfer during the

315   starch synthesis of lotus seed. Zhuo's research about the molecular structure of lotus seed found

316   that the branching degree of amylopectin was short, and the glucose residues of side chains were

317   longer, more than 30 (Zheng, Zheng, Zeng 2004). This is consistent with the level of genetic

318   research on *NnSBE*s. Therefore, transcription and activity of SBEs in lotus have a great influence

319   in the fine structure of amylopectin. Furthermore, the edible quality of lotus seed was affected,

320   making it easy for starch to retrograde.

321   **Conclusion**

322   This study undertook the preliminary study of *NnSBE* genes in lotus. Two isoforms which

323   encoding starch branching enzyme, were isolated and characterized from lotus. Genetic diversity

324   was analyzed by SNPs of two *NnSBE*s, revealed the genetic variation levels among varieties.

325   Difference of expression patterns and the affinity about two isozymes were described, as well as

326   the enzyme activity of SBE in development and different tissues, provided necessary information

327   for understanding of the processes involved in starch synthesis from the level of gene and

328   protein. This study revealed the selection of *NnSBE* genes during the cultivation process of lotus,

329   and the effect of *NnSBE* genes on the fine structure of starch in lotus seed. Although the

330   relationships between transcription level, enzyme activity and starch accumulation are complex,

331   our study provides us as much functional information as possible.

332   **References**

333   Burton Rachel A, Bewley JD, Smith Alison M, Bhattacharyya Madan K, Tatge H, Ring S, Bull V, Hamilton
334   William DO,Martin C. 1995. Starch branching enzymes belonging to distinct enzyme families are
335   differentially expressed during pea embryo development. *The Plant Journal* 7:3-15. 10.1046/j.1365-
336   313X.1995.07010003.x
337   Bustin SA, Beaulieu J-F, Huggett J, Jaggi R, Kibenge FSB, Olsvik PA, Penning LC,Toegel S. 2010. MIQE précis:
338   Practical implementation of minimum standard guidelines for fluorescence-based quantitative real-time
339   PCR experiments. *BMC Molecular Biology* 11:74. 10.1186/1471-2199-11-74

340   Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL,
341   Vandesompele J,Wittwer CT. 2009. The MIQE guidelines: minimum information for publication of
342   quantitative real-time PCR experiments. *Clin Chem* 55:611-622. 10.1373/clinchem.2008.112797
343   Chen Y, Fan G, Wu H, Wu Y,Mitchell A. 2007. Separation, identification and rapid determination of liensine,
344   isoliensinine and neferine from embryo of the seed of Nelumbo nucifera GAERTN. by liquid
345   chromatography coupled to diode array detector and tandem mass spectrometry. *Journal of Pharmaceutical*
346   *and Biomedical Analysis* 43:99-104. https://doi.org/10.1016/j.jpba.2006.06.016
347   Cheng F, Jiang D, Wu P,Shi C. 2001. The dynamic change of starch synthesis enzymes during the grain filling stage
348   and effects of temperature upon it. *Zuo wu xue bao* 27:201-206.
349   Cheng N, Zeng X-F, Zheng X-F, Diao Y, Wang Y-W, Xie K-Q, Zhou M-Q,Hu Z-L. 2014. Cloning and
350   characterization of the genes encoding the small and large subunit of the ADP-glucose pyrophosphorylase
351   in lotus (Nelumbo nucifera Gaertn). *Acta Physiologiae Plantarum* 37. 10.1007/s11738-014-1734-2
352   Commuri PD,Keeling PL. 2001. Chain-length specificities of maize starch synthase I enzyme: studies of glucan
353   affinity and catalytic properties. *The Plant Journal* 25:475-486. 10.1046/j.1365-313x.2001.00955.x
354   Doebley JF, Gaut BS,Smith BD. 2006. The Molecular Genetics of Crop Domestication. *Cell* 127:1309-1321.
355   https://doi.org/10.1016/j.cell.2006.12.006
356   Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-
357   Vegas A, Salazar GA, Tate J,Bateman A. 2016. The Pfam protein families database: towards a more
358   sustainable future. *Nucleic Acids Research* 44:D279-285. 10.1093/nar/gkv1344
359   Fujita N, Goto S, Yoshida M, Suzuki E,Nakamura Y. 2008. *The Function of Rice Starch Synthase I Expressed in*
360   *Escherichia coli*.
361   Gao M, Fisher DK, Kim K-N, Shannon JC,Guiltinan MJ. 1996. Evolutionary conservation and expression patterns
362   of maize starch branching enzyme I and IIb genes suggests isoform specialization. *Plant Molecular Biology*
363   30:1223-1232. 10.1007/BF00019554
364   Guan H, Li P, Imparl-Radosevich J, Preiss J,Keeling P. 1997. Comparing the Properties ofEscherichia
365   coliBranching Enzyme and Maize Branching Enzyme. *Archives of Biochemistry and Biophysics* 342:92-98.
366   https://doi.org/10.1006/abbi.1997.0115
367   Guo HB. 2009. Cultivation of lotus (Nelumbo nucifera Gaertn. ssp. nucifera) and its utilization in China. *Genetic*
368   *Resources and Crop Evolution* 56:323-330. 10.1007/s10722-008-9366-2
369   Hamada S, Ito H, Ueno H, Takeda Y,Matsui H. 2007. The N-terminal region of the starch-branching enzyme from
370   Phaseolus vulgaris L. is essential for optimal catalysis and structural stability. *Phytochemistry* 68:1367-
371   1375. https://doi.org/10.1016/j.phytochem.2007.02.024
372   Hong S, Mikkelsen R,Preiss J. 2001. Analysis of the Amino Terminus of Maize Branching Enzyme II by
373   Polymerase Chain Reaction Random Mutagenesis. *Archives of Biochemistry and Biophysics* 386:62-68.
374   https://doi.org/10.1006/abbi.2000.2179
375   Hong S,Preiss J. 2000. Localization of C-Terminal Domains Required for the Maximal Activity or for
376   Determination of Substrate Preference of Maize Branching Enzymes. *Archives of Biochemistry and*
377   *Biophysics* 378:349-355. https://doi.org/10.1006/abbi.2000.1845
378   Hu B, Jin J, Guo A-Y, Zhang H, Luo J,Gao G. 2015. GSDS 2.0: an upgraded gene feature visualization server.
379   *Bioinformatics* 31:1296-1297. 10.1093/bioinformatics/btu817
380   Huang L, Yang M, Li L, Li H, Yang D, Shi T,Yang P. 2018. Whole genome re-sequencing reveals evolutionary

381   patterns of sacred lotus (Nelumbo nucifera). *Journal of Integrative Plant Biology* 60:2-15.
382   10.1111/jipb.12606

383   Jeon JS, Ryoo N, Hahn TR, Walia H,Nakamura Y. 2010. Starch biosynthesis in cereal endosperm. *Plant Physiology*
384   *and Biochemistry* 48:383-392. 10.1016/j.plaphy.2010.03.006

385   Kang G, Li S, Zhang M, Peng H, Wang C, Zhu Y,Guo T. 2013. Molecular cloning and expression analysis of the
386   starch-branching enzyme III gene from common wheat (Triticum aestivum). *Biochemical Genetics* 51:377-
387   386. 10.1007/s10528-013-9570-4

388   Kuriki T, Stewart DC,Preiss J. 1997. Construction of Chimeric Enzymes out of Maize Endosperm Branching
389   Enzymes I and II. *Journal of Biological Chemistry* 272:28999-29004. 10.1074/jbc.272.46.28999

390   Larsson C-T, Khoshnoodi J, Ek B, Rask L,Larsson H. 1998. Molecular cloning and characterization of starch-
391   branching enzyme II from potato. *Plant Molecular Biology* 37:505-511. 10.1023/A:1005908305456

392   Liang-Jun LI, Zhang XD, Pan EC, Sun L, Xie K, Li GU,Cao BS. 2006. Relationship of Starch Synthesis with It's
393   Related Enzymes'Activities During Rhizome Development of Lotus(Nelumbo nucifera Gaertn). *Scientia*
394   *Agricultura Sinica*.

395   Lu Y, Li L, Zhou Y, Gao Q, Liang G, Chen X,Qi X. 2012. Cloning and Characterization of the Wx Gene Encoding
396   a Granule-Bound Starch Synthase in Lotus (Nelumbo nucifera Gaertn). *Plant Molecular Biology Reporter*
397   30:1210-1217. 10.1007/s11105-012-0430-x

398   Martin C,Smith AM. 1995. Starch biosynthesis. *The Plant Cell* 7:971-985.

399   Morell MK, Blennow A, Kosar-Hashemi B,Samuel MS. 1997. Differential Expression and Properties of Starch
400   Branching Enzyme Isoforms in Developing Wheat Endosperm. *Plant Physiology* 113:201.

401   Mutisya J, Sathish P, Sun C, Andersson L, Ahlandsberg S, Baguma Y, Palmqvist S, Odhiambo B, Åman
402   Per,Jansson C. 2003. Starch branching enzymes in sorghum (Sorghum bicolor) and barley (Hordeum
403   vulgare): Comparative analyses of enzyme structure and gene expression. *Journal of Plant Physiology*
404   160:921-930. https://doi.org/10.1078/0176-1617-00960

405   Nei M. 1978. ESTIMATION OF AVERAGE HETEROZYGOSITY AND GENETIC DISTANCE FROM A
406   SMALL NUMBER OF INDIVIDUALS. *Genetics* 89:583.

407   Nei M,Li W-H. 1973. LINKAGE DISEQUILIBRIUM IN SUBDIVIDED POPULATIONS. *Genetics* 75:213.

408   Nei M,Li WH. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases.
409   *Proceedings of the National Academy of Sciences* 76:5269.

410   Satoh H, Nishi A, Yamashita K, Takemoto Y, Tanaka Y, Hosaka Y, Sakurai A, Fujita N,Nakamura Y. 2003. Starch-
411   Branching Enzyme I-Deficient Mutation Specifically Affects the Structure and Properties of Starch in Rice
412   Endosperm. *Plant Physiology* 133:1111.

413   Shen-Miller J. 2002. Sacred lotus, the long-living fruits of <I>China Antique</I>. *Seed Science Research* 12:131-
414   143. 10.1079/ssr2002112

415   Shen-Miller J, Mudgett MB, Schopf JW, Clarke S,Berger R. 1995. Exceptional Seed Longevity and Robust Growth:
416   Ancient Sacred Lotus from China. *American Journal of Botany* 82:1367-1380. 10.2307/2445863

417   Slattery CJ, Kavakli IH,Okita TW. 2000. Engineering starch for increased quantity and quality. *Trends in Plant*
418   *Science* 5:291-298. https://doi.org/10.1016/S1360-1385(00)01657-5

419   Subasinghe RM, Liu F, Polack UC, Lee EA, Emes MJ,Tetlow IJ. 2014. Multimeric states of starch phosphorylase
420   determine protein–protein interactions with starch biosynthetic enzymes in amyloplasts. *Plant Physiology*
421   *and Biochemistry* 83:168-179. https://doi.org/10.1016/j.plaphy.2014.07.016

422   Tamura K, Stecher G, Peterson D, Filipski A,Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis
423   version 6.0. *Molecular Biology and Evolution* 30:2725-2729. 10.1093/molbev/mst197
424   Tetlow Ian J,Emes Michael J. 2014. A review of starch-branching enzymes and their role in amylopectin
425   biosynthesis. *IUBMB Life* 66:546-558. 10.1002/iub.1297
426   Tetlow IJ. 2010. Starch biosynthesis in developing seeds. *Seed Science Research* 21:5-32.
427   10.1017/S0960258510000292
428   Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F,Higgins DG. 1997. The CLUSTAL_X Windows Interface:
429   Flexible Strategies for Multiple Sequence Alignment Aided by Quality Analysis Tools. *Nucleic Acids*
430   *Research* 25:4876-4882. 10.1093/nar/25.24.4876
431   Tian Z, Qian Q, Liu Q, Yan M, Liu X, Yan C, Liu G, Gao Z, Tang S, Zeng D, Wang Y, Yu J, Gu M,Li J. 2009.
432   Allelic diversities in rice starch biosynthesis lead to a diverse array of rice eating and cooking qualities.
433   *Proceedings of the National Academy of Sciences of the United States of America* 106:21760-21765.
434   10.1073/pnas.0912396106
435   Wu J-Z, Zheng Y-B, Chen T-Q, Yi J, Qin L-P, Rahman K,Lin W-X. 2007. Evaluation of the quality of lotus seed of
436   Nelumbo nucifera Gaertn from outer space mutation. *Food Chemistry* 105:540-547.
437   https://doi.org/10.1016/j.foodchem.2007.04.011
438   Yue P, Melamud E,Moult J. 2006. SNPs3D: Candidate gene and SNP selection for association studies. *BMC*
439   *Bioinformatics* 7:166. 10.1186/1471-2105-7-166
440   Zheng BD, Zheng JG,Zeng CX. 2004. Study on Three Functional Components of Chinese Main Lotus Seed
441   Varieties. *Acta Nutrimenta Sinica*.

442

443

Figure 1

The amplication of cDNA fragments of *NnSBE*s.

The amplication of cDNA fragments of *NnSBE*s *NnSBEI* and *NnSBEIII* using RT-PCR experiments. The products were analyzed by electrophoresis on 1.0% agarose gels, DL 5000 DNA marker.

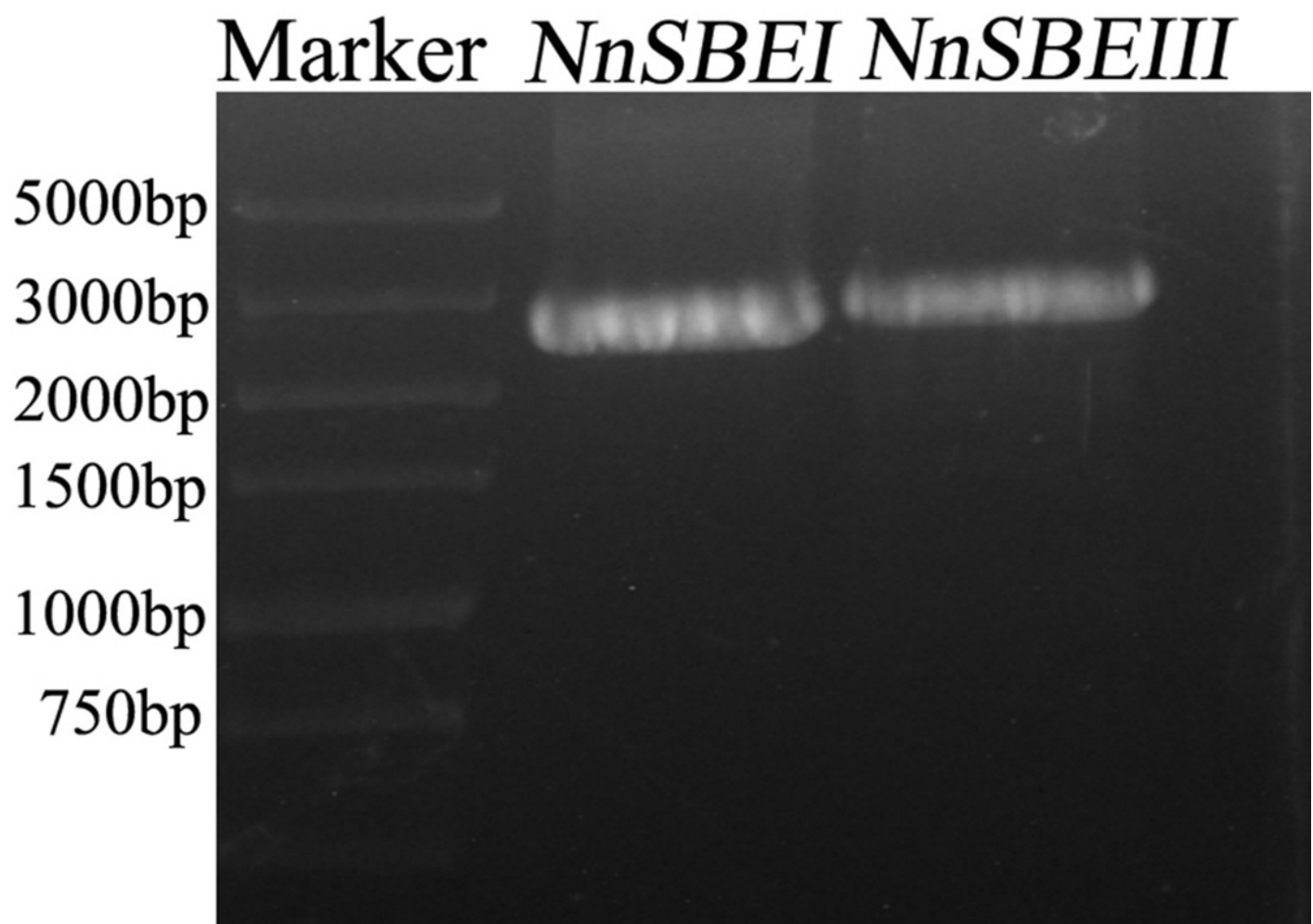



Figure 2

The gene structure of *NnSBEI* and *NnSBEIII*.

The blue boxes represent UTR, black boxes represent the exons, thick line represent intros.

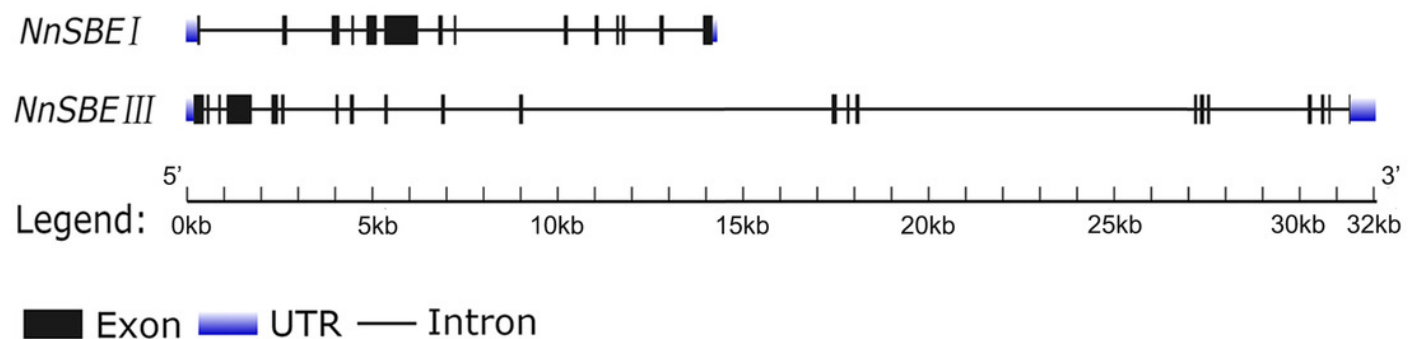



Figure 3

Phylogenetic analysis of the SBEs family.

The dendrogram was constructed using MEGA6.0 software with the neighbor-joining method. Sequences aligned included the SBE protein sequences of 20 other species, which identified or predicted from NCBI database were used for phylogenetic analysis. Different color regions represent two subclades of SBEs: the blue region cover SBE A family, and the gray cover SBE B family. The NnSBEs were labeled with "●".

NnSBEIII *Nelumbo nucifera*-XP 010252481.1●
100 ┌ TaSBEIII *Triticum aestivum*- AFH58741.1
100 └ ZmSBEIII *Zea mays*-ABY56822.1
99 ┌ SbSBEIIb *Sorghum bicolor*-AAP72267.1
92 └ ZmSBEII *Zea mays*-AAA18571.1
99 HvSBEIIb *Hordeum vulgare*-AAC69754.1
97 OsSBE *Oryza sativa*-ACE88265.1
EcSBE *Eleusine coracana*-ARA71544.1
56 99 ┌ HvSBEIIa *Hordeum vulgare*-AAC69753.1
69 100 └ TaSBEII *Triticum aestivum*-CAA72154.1
57 CeSBEA *Colocasia esculenta*-ABP96984.1
LdSBE *Lilium davidii*-AJG44456.1
100 ┌ LcSBEII *Lens culinaris*-AHW50661.1
51 └ PaSBEI *Pisum sativum*-CAA56319.1
99 ┌ PvSBE *Phaseolus vulgaris*-BAA82348.2
100 100 └ VrSBEII *Vigna radiata*-AAT76444.1
AtSBE2.1 *Arabidopsis thaliana*-NP 181180.1
88 AtSBE2.2 *Arabidopsis thaliana*-NP 195985.3
100 IbSBEII *Ipomoea batatas*-BAB64912.1
77 StSBE2-2 *Solanum tuberosum*-NP 001275467.1
CrSBE *Chlamydomonas reinhardtii*-XP 001696229.1
100 PkSBEII *Parachlorella kessleri*-BAF98234.1

SBE A

CrSBE *Chlamydomonas reinhardtii*-XP 001695391.1
100 ┌ HvSBE *Hordeum vulgare*-CAX51365.1
100 └ TaSBEI *Triticum aestivum*-AAG27621.1
71 OsSBEI *Oryza sativa*-ABL74558.1
86 ZmSBEI *Zea mays*-NP 001105370.1
100 NnSBEI *Nelumbo nucifera*-ACM07441.1●
39 ┌ IbSBEI *Ipomoea batatas*-BAE96956.1
99 └ StSBE *Solanum tuberosum*-NP 001275183.1
45 MeSBE *Manihot esculenta*-CAA54308.1
79 PvSBE *Phaseolus vulgaris*-BAA82349.1
87 MtSBEI *Medicago truncatula*-XP 013462275.1
98 ┌ LcSBEI *Lens culinaris*-AHW50662.1
100 └ PsSBEII *Pisum sativum*-CAA56320.1

SBE B

Figure 4

Dendrogram of 45 individuals lotus based on SNP data of *NnSBEIII*.

Dendrogram based Nei's genetic distance and the UPGMA method was using for Cluster analysis. The 45 individuals of lotus were copolymerized into three classes. The number of individuals were shown in the supplementary 1.

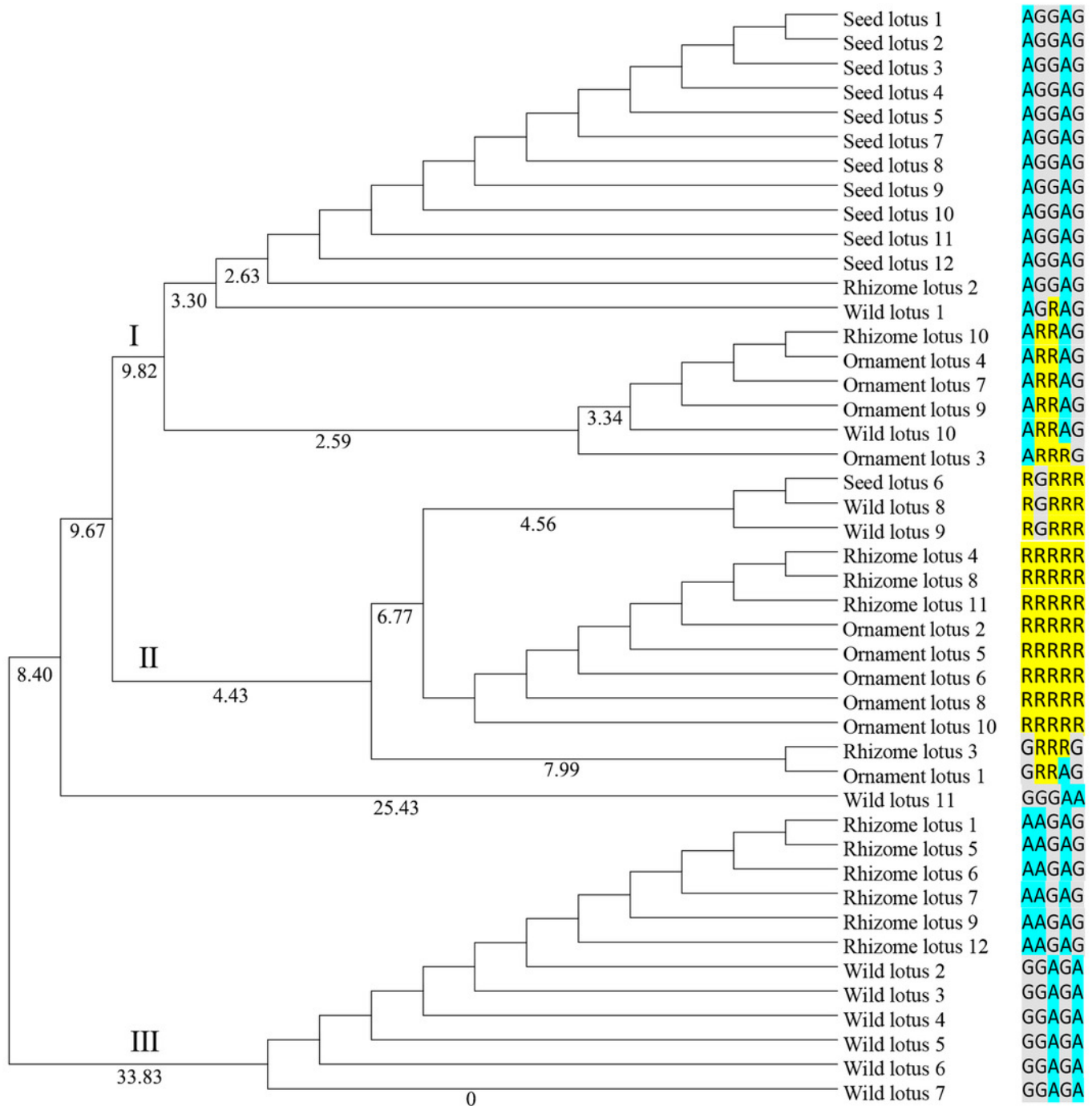



Figure 5

Tissue-differential expression of the *NnSBE*s.

The column chart confirmed expression pattern of leaves, petiole, rhizomes, roots which collected in 8th week after sowing at the early swelling stage (a, d), collected in 10th week after sowing at the middle swelling stage (b, e), collected in 12th week after sowing at the later swelling stage (c, f) in lotus. The gray boxes represent *NnSBEI* and the lightgray represent *NnSBEIII*. Error bars indicate standard error (n = 3) (p< 0.05).

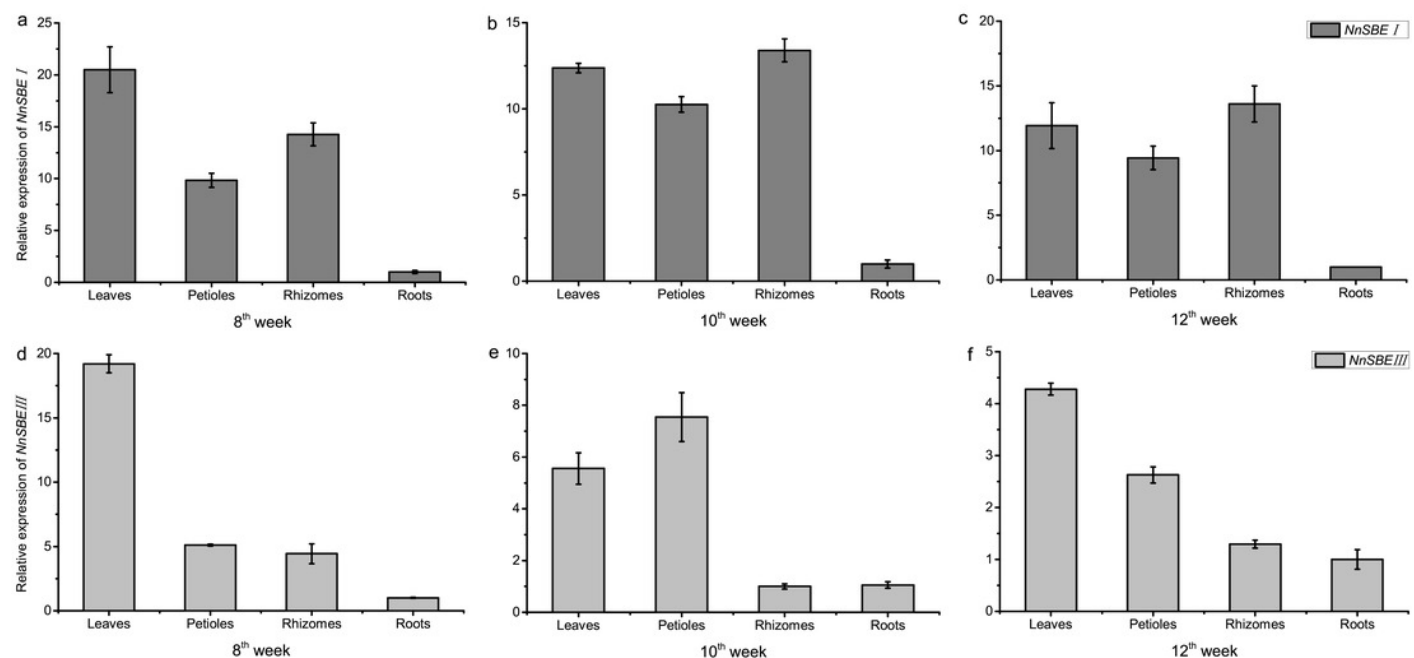



Figure 6

Expressed tendency of the *NnSBE*s during developing seed.

Relative expression of the *NnSBE*s were analyzed at 12, 16, 20, 24, 28 DAF. The gray boxes represent *NnSBEI* (a) and the lightgray represent *NnSBEIII* (b). Error bars indicate standard error (n = 3) (p< 0.05).

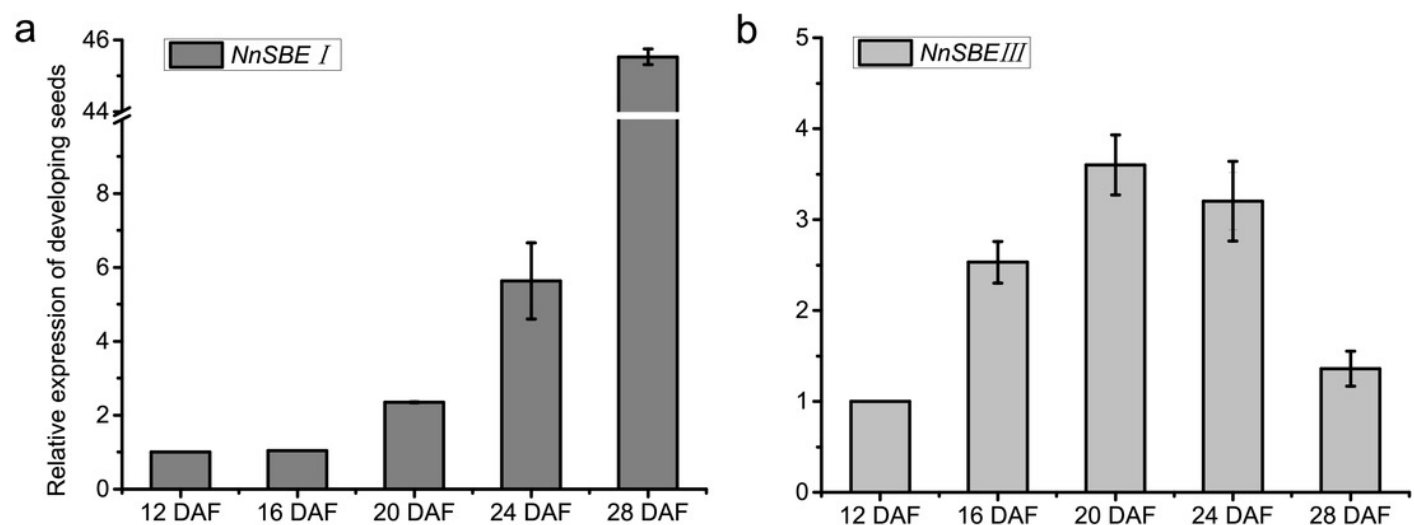



Figure 7

Enzyme activity of NnSBE.

Tissue-differential enzyme activity of NnSBE in leaves, petiole, rhizomes, roots which collected in 8th week after sowing at the early swelling stage, 10th week after sowing at the middle swelling stage, 12th week after sowing at the later swelling stage, and the different colors (black, red, blue, purple) represent various organs(leaf, petiole, rhizome and root) in an orderly way from top to bottom (a). Enzyme activity of developing seed of NnSBE at 12, 16, 20, 24, 28 DAF, the green lines represent seed(b). Error bars indicate standard error (n = 3) (p< 0.05).

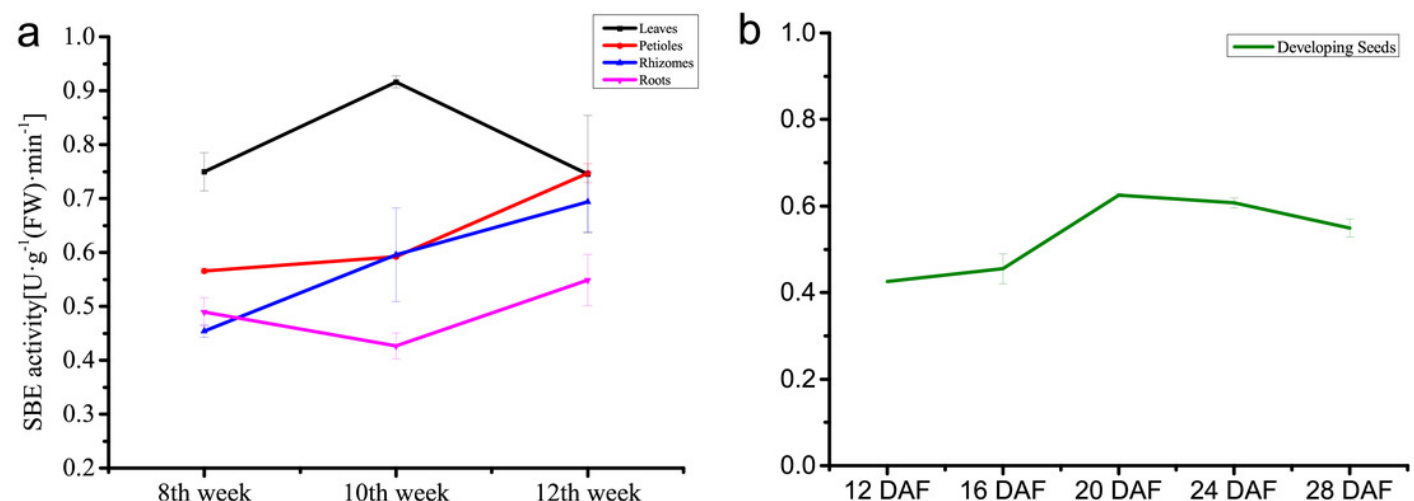
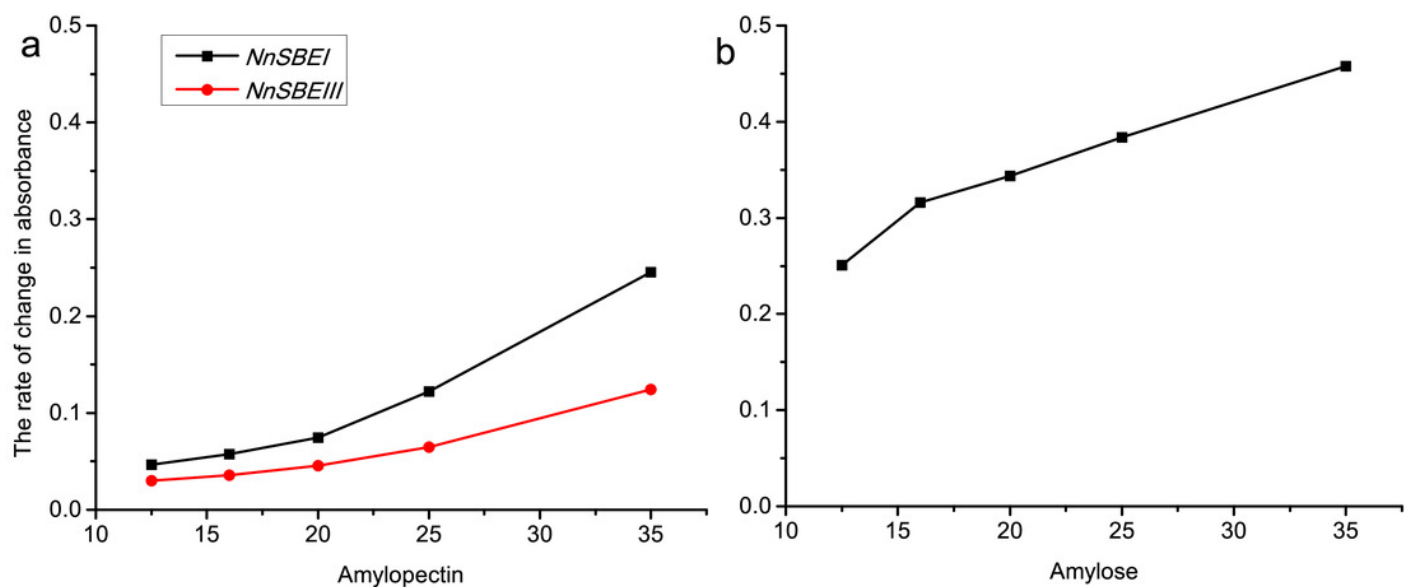



Figure 8

The rate of change in absorbance about NnSBEI and NnSBEIII reacted among different substrate concentration.

The substrate is amylopectin (a); the substrate is amylose(b). The black lines represent NnSBEI; the red line represents NnSBEIII.

Primers used in qRT-PCR

Gene	Sequence forward & reverse primers (5'-3')	Amplicon length (bp)	R²	Primer efficiency
q*CYP*-F	GTACCCAGAAGAATGCCCTA	102	0.998	96.222
q*CYP*-R	ATGAAGCCCTTGATGACTCG			
q*NnSBEI*-F	GTAGACCATTTCACATCGC	114	0.999	93.132
q*NnSBEI*-R	TAATAAGCCACACATGTACGAG			
q*NnSBEIII*-F	TATGCATGGCTAGTTCCAC	110	0.999	86.246
q*NnSBEIII*-R	TTATGCCAAAATGCCTCGT			

1   **Table 1** Primers used in qRT-PCR

Gene information of *NnSBEI* and *NnSBEIII*

Gene name	Gene ID	Genomic length(kb)	ORF length (bp)	Number of exons	Number of intros	Number of amino acid (Aa)
NnSBEI	104603742	14.3	2352	14	13	783
NnSBEIII	104594060	32.1	2691	21	20	896

1   **Table 2** Gene information of *NnSBEI* and *NnSBEIII*

Genetic diversity based on SNPs of *NnSBEIII* in the test population.

H$_O$: the observed heterozygosity; H$_E$: the expected heterozygosity.

SNPs of *NnSBEIII*	Aa	Allelic frequency		Genotypic frequency		H$_O$	H$_E$	Shannon Index
c.282 G>A	M	A	0.6778	AG	0.2444	0.2444	0.4368	0.6285
		G	0.3222	AA	0.5556			
				GG	0.2			
c.535 G>A	E/K	A	0.3111	AG	0.3556	0.3556	0.4286	0.6200
		G	0.6889	AA	0.1333			
				GG	0.5111			
c.880 G>A	V/L	A	0.3556	AG	0.4444	0.4444	0.4583	0.6508
		G	0.6444	AA	0.1333			
				GG	0.4222			
c.2308 G>A	O/N	A	0.7000	AG	0.2889	0.2889	0.4200	0.6109
		G	0.3000	AA	0.5556			
				GG	0.1555			
c.2559 G>A	G	A	0.2778	AG	0.2000	0.2000	0.4012	0.5908
		G	0.7222	AA	0.1778			
				GG	0.6222			

1 Table 3 Genetic diversity based on SNPs of *NnSBE III* in the test population. H$_O$: the observed heterozygosity; H$_E$: the
2 expected heterozygosity