Genome-wide identification and characterization of the KCS gene family in sorghum [*Sorghum bicolor* (L.) Moench]

Aixia Zhang [1], **Jingjing Xu** [1], **Xin Xu** [1], **Junping Wu** [2], **Ping Li** [1], **Baohua Wang** [Corresp., 1], **Hui Fang** [Corresp. 1]

[1] Ministry of Agricultural Scientific Observing and Experimental Station of Maize in Plain Area of Southern Region, School of Life Sciences, Nantong University, Nantong, Jiang Su, China

[2] Nantong Changjiang Seed Co., Ltd, Nantong, Jiangsu, China

Corresponding Authors: Baohua Wang, Hui Fang
Email address: bhwang@ntu.edu.cn, fanghui8912@126.com

The aboveground parts of plants are covered with cuticle, a hydrophobic layer composed of cutin polyester and cuticular wax that can protect plants from various environmental stresses. β -Ketoacyl-CoA synthase (KCS) is the key rate-limiting enzyme in plant wax synthesis. Although the properties of *KCS* family genes have been investigated in many plant species, the understanding of this gene family in sorghum is still limited. Here, a total of 25 *SbKCS* genes were identified in sorghum genome, which were named from *SbKCS1* to *SbKCS25*. Evolutionary analysis among different species divided the *KCS* family into five subfamilies and the SbKCSs were more closely related to maize, implying a closer evolutionary relationship between sorghum and maize. All *SbKCS* genes were located on chromosomes 1, 2, 3, 4, 5, 6, 9 and 10, respectively, while Chr 1 and Chr 10 contained more *KCS* genes than other chromosomes. The prediction results of subcellular localization showed that SbKCSs were mainly expressed in plasma membrane and mitochondria. Gene structure analysis revealed that there was 0-1 intron in the sorghum *KCS* family and *SbKCSs* within the same subgroup were similar. Multiple *cis*-acting elements related to abiotic stress, light and hormones response were enriched in the promoters of *SbKCS* genes, which indicated the functional diversity among these genes. The three-dimensional structure analysis showed that a compact spherical space structure was formed by various secondary bonds to maintain the stability of SbKCS proteins, which was necessary for their biological activity. qRT-PCR results revealed that nine randomly selected *SbKCS* genes expressed differently under drought and salt treatments, among which *SbKCS8* showed the greatest fold of expression difference at 12 h after drought and salt stresses, which suggested that the *SbKCS* genes played a potential role in abiotic stress responses. Taken together, these results provided an insight into investigating the functions of *KCS* family in sorghum and in response to abiotic stress.

Genome-wide identification and characterization of the *KCS* gene family in sorghum [*Sorghum bicolor* (L.) Moench]

Aixia Zhang[1], Jingjing Xu[1], Xin Xu[1], Junping Wu[2], Ping Li[1], Baohua Wang[1*], Hui Fang[1*]

[1.]Ministry of Agricultural Scientific Observing and Experimental Station of Maize in Plain Area of Southern Region, School of Life Sciences, Nantong University, Nantong 226019, People's Republic of China.

[2] Nantong Changjiang Seed Co., Ltd, Nantong 226368, Jiangsu, People's Republic of China.

*Co-corresponding author: Hui Fang, Email: fanghui8912@hotmail.com

*Co-corresponding author: Baohua Wang, Email: bhwang@ntu.edu.cn

ABSTRACT

The aboveground parts of plants are covered with cuticle, a hydrophobic layer composed of cutin polyester and cuticular wax that can protect plants from various environmental stresses. β-Ketoacyl-CoA synthase (KCS) is the key rate-limiting enzyme in plant wax synthesis. Although the properties of *KCS* family genes have been investigated in many plant species, the understanding of this gene family in sorghum is still limited. Here, a total of 25 *SbKCS* genes were identified in sorghum genome, which were named from *SbKCS1* to *SbKCS25*. Evolutionary analysis among different species divided the *KCS* family into five subfamilies and the SbKCSs were more closely related to maize, implying a closer evolutionary relationship between sorghum and maize. All *SbKCS* genes were located on chromosomes 1, 2, 3, 4, 5, 6, 9 and 10, respectively, while Chr 1 and Chr 10 contained more *KCS* genes than other chromosomes. The prediction results of subcellular localization showed that SbKCSs were mainly expressed in plasma membrane and mitochondria. Gene structure analysis revealed that there was 0-1 intron in the sorghum *KCS* family and *SbKCSs* within the same subgroup were similar. Multiple *cis*-acting elements related to abiotic stress, light and hormones response were enriched in the promoters of *SbKCS* genes, which indicated the functional diversity among these genes. The three-dimensional structure analysis showed that a compact spherical space structure was formed by various secondary bonds to maintain the stability of SbKCS proteins, which was necessary for their biological activity. qRT-PCR results revealed that nine randomly selected *SbKCS* genes expressed differently under drought and salt treatments, among which *SbKCS8* showed the greatest fold of expression difference at 12 h after drought and salt stresses, which suggested that the *SbKCS* genes played a potential role in abiotic stress responses. Taken together, these results provided an insight into investigating the functions of *KCS* family in sorghum and in response to abiotic stress.

Subjects: Plant Science, Bioinformatics, Molecular Biology, Agricultural Science

Keywords: *Sorghum bicolor*, *KCS* gene family, Abiotic stress

INTRODUCTION

44

45 In higher plants, very long chain fatty acids (VLCFAs) are precursors for the synthesis of waxes

46 in pollen husks, fruits, stems and leaves (*Paul et al.*, *2006*). In addition, VLCFAs exist in plants

47 in the forms of several other substances, including triacylglycerol, suberin, phospholipids and

48 sphingolipids (*Haslam et al.*, *2013*). VLCFAs are synthesized by extending C16 or C18 fatty

49 acids through microsomal enzyme system. They participate in a variety of cellular processes

50 such as nuclear pore formation, and lipid and protein transport (*Paul et al.*, *2006*). Plant

51 cuticular wax has important physiological functions in resisting abiotic stresses, such as

52 preventing water loss, resisting ultraviolet radiation, drought and saline alkali stress (*Lee and*

53 *Suh*, *2015*). Therefore, VLCFAs play an important role in cell biology, whereas many aspects of

54 their synthesis and function remains to be further studied.

55 Plant cuticular waxes have a very complex chemical composition, mainly including VLCFAs

56 and alkanes, primary alcohols, secondary alcohols, aldehydes, ketones and esters (*Bernard and*

57 *Joubès*, *2013*). The synthesis and secretion of waxes need the participation of a variety of

58 organelles and enzymes. The synthesis of VLCFA is catalyzed by four enzymes, including β-

59 ketoyl-CoA synthase (KCS), β-Ketoyl-CoA reductase (KCR), β-hydroxyacyl-CoA dehydratase

60 (HCD) and enol-CoA reductase (ECR). Among them, KCS has substrate specificity and tissue

61 specificity, and is a key enzyme in the elongation process of VLCFA (*Fan et al.*, *2018*). It

62 participates in the synthesis of wax precursors and prolongs VLCFAs in higher plants, thereby

63 limiting non-stomatal water loss, protecting against UV radiation, resisting bacterial and fungal

64 pathogen invasion, and defending against drought stress (*Zhang et al.*, *2019*). Because of the

65 significance of *KCS* genes, their functions have been well studied and understood in some plant

66 species. Three *KCS* genes, including *KCS1* (*Todd et al.*, *1999*), *FDH* (*Pruitt et al.*, *2000*) and

67 *CER6* (*Hooker et al.*, *2002*), had been proved to be involved in the biosynthesis of cuticular

68 waxes in Arabidopsis. *FDH* was mainly expressed in the young leaves and flower organs of

69 Arabidopsis and related to the synthesis of VLCFAs in epidermal cells, regulating the formation

70 of plant morphological structure and improving the resistance to biotic or abiotic stresses (*Pruitt*

71 *et al.*, *2000*). *KCS2/DAISY* is a wax biosynthesis gene in Arabidopsis that is activated by direct

72 binding to *MYB94* (*Lee et al.*, *2015*) and *MYB96*. An increase in the accumulation of cuticular

73 wax decreased the rate of cuticular transpiration in the leaves of *MYB94*-OE lines, thus

74 improving the drought resistance of plants (*Seo et al.*, *2011*). Overexpression of *KCS* in wheat

75 significantly increased the resistance to adversity of wheat (*Hu et al.*, *2010*). In barley, the

76 mutation of *KCS1* gene affected the wax structure of stratum corneum, thereby reducing the

77 resistance to water and powdery mildew of cuticles (*Li et al.*, *2018*). Therefore, *KCS* genes

78 played an important role in the synthesis of wax components and in the response to stress

79 environments in plants. Besides the function in response to adverse stresses, the *KCS* genes

80 have potential to improve some quality traits, e.g., *KCS13* could promote the increase of fiber

81 length in cotton (*Qin et al.*, *2007*). Currently, *KCS* genes had been identified in several species.

82 For example, 21 *KCS* genes in Arabidopsis were identified and classified into four subgroups,

83 namely *FAE1*-like, *KCS1*-like, *FDH*-like and *CER6* (*Costaglioli et al.*, *2005*). In addition, 58

84 *KCS* genes in *Gossypium hirsutum* (*Xiao et al.*, *2016*), 28 *KCS* genes in *Malus domestica* (*Lian*

85 *et al.*, *2020*), 31 and 8 *KCS* genes in *Brassica rapa* and *Picea abies* (*Guo et al.*, *2016*),

86 respectively, have been reported.

87 Sorghum [*Sorghum bicolor* (L.) Moench], one of the earliest cultivated cereal crops in China,
88 has great edible value and is rich in nutrients with a lot of fat-soluble protein and unsaturated
89 fatty acids. The *KCS* gene family plays a crucial role in the synthesis of wax and VLCFAs in
90 many plants, but little is known about the *KCS* gene family in sorghum. Therefore, in this study,
91 we systematically identified and analyzed the characteristics of sorghum *KCS* gene family by
92 bioinformatics methods, and the expression levels of *SbKCS* genes under drought and salt
93 stresses were also investigated. This study provided useful information for further investigating
94 the molecular functions of *KCS* genes in response to abiotic stress in sorghum.

95 ## MATERIALS AND METHODS

96 ### Identification of *SbKCS* genes in sorghum

97 The sorghum BTx623 genome dataset was downloaded from the Ensembl Genomes website

98 (http://ensemblgenomes.org/). The Hidden Markov Model (HMM) profiles of the 3-Oxoacyl-

99 [acyl-carrier-protein (ACP)] synthase III C terminal domain (ACP_syn_III_C, Pfam:PF08541)

100 and the FAE1/Type III polyketide synthase-like protein domain (FAE1_CUT1_RppA,

101 Pfam:PF08392) were obtained from the Pfam protein family database (http://pfam.xfam.org/)

102 (*Liu et al.*, *2018*). Firstly, after using the Hmmsearch program in the Linux system to search for

103 proteins containing the conserved domain (*Jeanmougin et al.*, *1998*), more than 50 protein

104 sequences were obtained with expected (E) values less than 1.2E-28. Then, to ensure the

105 accuracy   and   reliability   of   the   detected   proteins,   NCBI-CDD   website

106 (https://www.ncbi.nlm.nih.gov/cdd),   Pfam   and   SMART   database   (http://smart.embl-

107 heidel.berg.de/) were used to further detect the KCS-conserved domains in the non-redundant

108 sequences of each putative protein to confirm these sequences (*Mistry et al.*, *2021*). Finally, 25

109 *SbKCS* genes were identified and named according to their positions on chromosomes. Online

110 server PROTPARAM (http://web.expay.org/protparam/) was used to analyze the theoretical

111 isoelectric point (Pi), molecular weight (MW), and sequence length of all KCS proteins.

112 CELLO website (http://cello.life.nctu.edu.tw/) was used to predict the subcellular localization

113 of the KCS proteins by submitting the sequences of SbKCS proteins.

114

115 **Chromosomal location and gene structure analysis of *SbKCS* genes**

116 The physical locations of all *SbKCS* genes were obtained in the sorghum genome. The MapChart

117 (*Voorrips*, *2002*) and AI (Adobe Illustrator CS6) software were used to visualize the images of

118 chromosomal position. The online program of the gene structure display server (GSDS:

119 http://gsds.cbi.pku.edu.cn) was used to obtain the exon-intron structure of the *SbKCS* genes

120 according to their full-length CDS sequences.

121

122 **Multiple sequence alignment of SbKCS proteins**

123 To analyze the conserved domains of proteins, the multiple sequence alignment of 25 identified

124 SbKCS proteins and 21 Arabidopsis KCS proteins was performed using Clustal W in MEGA 7.0

125 (*Kumar et al.*, *2016*), and then the result of multiple sequence alignment was imported into

126 GeneDoc website (http://nrbsc.org/gfx/genedoc) to visualize the conserved sequences of the

127 SbKCS proteins.

128

129 **Phylogenetic and motif analysis of *SbKCS* genes**

130 To study the phylogenetic relationships of SbKCS proteins, the sequences of KCS proteins in

131 Arabidopsis, sorghum, maize, rice and *Brochypodium distachyon* were aligned with Clustal W

132 software under its default parameters. The phylogenetic tree was constructed using the

133 maximum-likelihood (ML) method of MEGA 7.0 software with 1000 bootstrap replicates.

134 Evolview online website (https://evolgenius.info//evolview-v2/#login) and AI software were

135 used to visualize the phylogenetic tree. An online MEME program (http://meme-suite.org/) was

136 used to analyze the conserved motifs in the identified sorghum KCS proteins and the parameter

137 settings were as follows: the optimum width of motif was between 6 and 50 residues, and

138 maximum number of motifs was 10. TBtools software was used to visualize the conserved

139 motifs based on output XML files (*Chen et al., 2020*).

140

141 **Secondary structure prediction and three-dimensional (3D) model construction of SbKCS**

142 **proteins**

143 Secondary structure of SbKCS proteins was predicted using an online website Prabi

144 (https://npsa-prabi.ibcp.fr/). To further analyze the protein structure of the SbKCS family, the

145 tertiary structure of SbKCS proteins was predicted according to the protein sequences. Then, the

146 online tools SWISS-MODEL (https://swissmodel.expasy.org/) was used to construct the 3D

147 models of SbKCS proteins by homologous protein modeling method (*Waterhouse et al.*, 2018).

148

149 **Prediction of *cis*-acting elements of *SbKCS* genes**

150 The 1.5 kb promoter sequences upstream of the initiation codon for all the 25 *SbKCS* genes were

151 extracted from sorghum genomic sequences, and then they were analyzed by PlantCARE

152 (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) (*Wang et al., 2020*) to survey the

153 potential *cis*-acting elements associated with stress response and hormones in the promoter

154 regions. The obtained *cis*-element information was then visualized through the GSDS online

155 website.

156

157 **Plant materials and stress treatments**

158 The seeds of sorghum variety BTx623, an elite inbred line whose genome have been sequenced

159 (*Paterson et al., 2009*), were planted in pots filled with Pindstrup substrate and vermiculite (3:1)

160 and cultivated in a controlled and adjusted green house at a temperature of 23 ℃/20 ℃ with a 16

161 h/8 h day and night regime. After growing to the three-leaf stage, sorghum seedlings with

162 consistent growth were selected and fixed on the planting basket, and cultured with 1/2 Hoagland

163 nutrient solution (PH 5.5). Two different abiotic stresses were applied to the seedlings at the

164 three-leaf stage. For salt stress, the sorghum seedlings were cultured in 150 mM NaCl solution

165 for 24 hours (*Wei et al., 2021*); for drought stress, the seedlings were hydroponically cultured

166 with 20% PEG6000 nutrient solution for 24 hours (*Wang et al., 2020*); the control group was

167 cultured with 1/2 Hoagland nutrient solution. The sorghum seedlings under stresses and control

168 were sampled at different time after treatment for 0, 6, 12 and 24h, respectively. Three biological

169 replicates for each sample were collected and quickly put into liquid nitrogen, and then stored in

170 an ultralow temperature freezer -80℃ for RNA extraction.

171

172 **RNA extraction, and quantitative real-time PCR (qRT-PCR) analysis**

173 Total RNA was extracted from leaves at different stages under stress treatment and control using

174 Ultrapure RNA Kit (CWbio, Jiangsu, China). About 1 μg of total RNA was used to synthesize

175 the first-strand cDNA using the HiScript III RT SuperMix for qPCR kit purchased from Nanjing

176 Vazyme Company. The qPCR primers were designed with Primer 5.0 software and the primer

177 sequences were shown in Table S1. The qRT-PCR was carried out for each sample using a

178 ChamQ SYBR qPCR Master Mix (Vazyme, Nanjing, China) with an ABI 7500 real-time PCR

179 system and the parameters were set as follows: stage1, 95 °C for 30 s; stage2, 40 cycles of 95 °C

180 for 10 s and 60 °C for 30 s; stage3, melting curve. The *actin1* (SORBI_3001G112600) was used

181 as an internal reference for standardization between samples (*Zheng et al., 2021*). For qRT-PCR

182 detection, three biological replicates and four technological replicates for a single sample were

183 performed to ensure the accuracy of qPCR analysis. Relative expression levels of each gene were

184 calculated using the $2^{-\Delta\Delta CT}$ method (*Livak et al., 2001*). The data were presented as mean ±

185 standard error (SE) and Student's *t* test was used for significance analysis, and *P* < 0.05 was

186 considered as significant. The histogram was drawn with GraphPad Prism software (*Swift, 1997*).

187

188 **RESULTS**

189 **Identification of *KCS* genes in sorghum**

190 More than 50 putative *KCS* genes were obtained by searching two conserved domains.

191 Subsequently, the protein sequences of the putative *SbKCS* genes were confirmed through

192 SMART, NCBI-CDD websites and Pfam databases, and some incomplete and uncertain

193 sequences were removed after screening. Ultimately, a total of 25 *SbKCS* genes were identified

194 as belonging to the *KCS* gene family and were named as *SbKCS1-SbKCS25* based on their

195 location on the sorghum chromosomes. Furthermore, the physicochemical properties of the

196 SbKCSs were characterized (Table S2). The SbKCS proteins with largest and smallest number of

197 amino acids were SbKCS5 and SbKCS21, containing 549 and 440 amino acids, respectively.

198 The molecular weight of these proteins was in the range of 48785.3 (SbKCS21) - 60617.4 Da

199 (SbKCS5), and the theoretical isoelectric point ranged from 7.1 (SbKCS17) to 9.91 (SbKCS9),

200 with a mean value of 8.92.

201

202 **Chromosomal location and subcellular localization prediction analysis of *SbKCS* genes**

203 To map the chromosomal location of the *SbKCS* genes, the physical location of the *SbKCS* genes

204 on the chromosomes was investigated. Each *SbKCS* gene was named in the light of its top-down

205 physical location on sorghum chromosome (Chr) 1 to 10 and the 25 *SbKCS* genes were unevenly

206 distributed on the chromosomes (Fig. 1). *SbKCS* genes were distributed separately or in clusters,

207 and mostly at both ends of chromosomes. Chr 1 had the largest number of *SbKCS* genes (8,

208 ~32.0%), followed by Chr 10 (6, ~24.0%); Chr 4 possessed the least *SbKCS* genes (1, ~4.0%).

209 Each of the other chromosomes (Chr 2, 3, 5, 6, 9) contained 2 (~8.0%) *SbKCS* genes. The

210 subcellular localization prediction analysis showed that SbKCS proteins were mainly localized in

211   the plasma membrane, followed by mitochondria and chloroplast, which suggested that SbKCS

212   proteins might be mostly expressed and function in these organelles (Fig. 2).

213

214   **Mutiple sequence alignment, phylogenetic and conserved motifs analysis**

215   To evaluate phylogenetic relationship among the members of the *KCS* family in sorghum,

216   Clustal W in MEGA 7.0 software was used to perform multiple sequence alignment for the 25

217   identified SbKCS and 21 AtKCS protein sequences. The results showed that it contained two

218   important conserved domains: the *FAE1*_CUT1_RppA domain and the ACP_syn_III_C domain

219   (Fig. S1). The sequences of these two domains were highly consistent and conserved in both

220   regions, indicating that they were essential to the function of *KCS* genes in sorghum. Meanwhile,

221   we also found that the C-terminal and intermediate regions of these proteins were highly

222   conserved, whereas the N-terminal was poorly conserved.

223   To clarify the phylogenetic relationship among sorghum KCS proteins and obtain more detailed

224   classification of the KCS proteins, we constructed a phylogenetic tree using the MEGA 7.0 based

225   on the full-length protein sequences of the 25 obtained SbKCS and KCS proteins from the

226   representative species: Arabidopsis (21), rice (19), maize (29) and *Brochypodium distachyon* (32)

227   (Fig. 3). Subsequently, the 126 KCS proteins were divided into 5 subfamilies named from Group

228   1 to Group 5. SbKCS proteins were distributed in all the five groups. Group 1 had the largest

229   number (10) of SbKCS members, followed by Group 4 with the number of 9. Group 5 contained

230   4 members of the KCS family. Besides, Group 2 and Group 3 each contained only 1 SbKCS

231   protein. In previous studies, 21 AtKCS proteins were classified into 4 subfamilies, namely

232   FAE1-like, CER6, KCS1-like, and FDH-like (*Costaglioli et al.*, *2005*). Similarly, we also found

233   that 9 *SbKCSs* in Group 1 appeared in the KCS1-like subfamily and the remaining one gene in

234   Group 1, *SbKCS18*, was clustered in FAE1-like subfamily; 9 *SbKCSs* in Group 4 appeared in the

235   CER6 subfamily; 4 *SbKCSs* in Group 5 appeared in the FDH-like subfamily. The remaining two

236   *SbKCSs*, *SbKCS6* and *SbKCS7*, were branched separately into two new branches, Group 2 and

237   Group 3. These results indicated that the evolutionary relationship of *KCS* family genes in

238   sorghum was more complex than that in Arabidopsis. Furthermore, KCS proteins in sorghum

239 and maize were always in close branches in all subgroups, which implied that the SbKCSs were

240 more closely related to maize in evolutionary relationship. In addition, genes from the same

241 subgroup could be thought to have similar functions; consequently, the five distinct clades might

242 indicate that the functions of *KCS* gene were diverse.

243 In order to obtain the characteristic regions of SbKCS proteins, the analysis of conserved motifs

244 was performed using MEME software. Ten motifs in the SbKCS proteins were predicted and

245 finally, 5 distinct conserved motifs were found (Fig. 4A). Each of the 25 SbKCS proteins

246 contained motifs 1, 2, 4, 6 and 8, with the same order that they were arranged on the protein

247 sequences, indicating that the motifs of almost all the identified *KCS* family genes were

248 conserved.

249

250 **Gene structure, secondary structure and 3D modelling analysis of the SbKCS proteins**

251 To analyze the structural composition of each *SbKCS* gene, the exon-intron organizations were

252 obtained by comparing the cDNA and DNA sequences of a single *SbKCS* gene (Fig. 4B). From

253 the results, we discovered that most of the *SbKCS* genes were relatively simple in structure, e.g.,

254 the number of introns in the 25 *SbKCS* genes ranged from 0 to 1. Most of the *SbKCS* genes (17,

255 ~68.0%) contained no introns in their open reading frame regions, and 8 (~32.0%) *SbKCS* genes

256 contained only 1 intron. The 17 intronless genes were distributed in Group 1, 4 and 5, while the

257 *SbKCS* genes containing only one intron were distributed in all the five subgroups. As shown in

258 Fig 4B, the exon-intron structure of members within the same subgroup tended to be similar. The

259 number of *SbKCS* genes with and without introns in Group 1 was about the same. In Group 1, 4

260 genes including *SbKCS2*, *SbKCS8*, *SbKCS14* and *SbKCS15* each had one intron, whereas the

261 remaining 6 *SbKCS* genes had no intron. All the *SbKCS* genes in Group 2 and 3 have no introns.

262 In Groups 4 and 5, only one gene in each group contained one intron, and the others had no

263 introns. Generally speaking, genes in the same subfamily contained similar genetic structures and

264 the position of introns tended to be conserved in each subgroup. These results further illustrated

265 the authenticity of the evolutionary tree, and it was speculated that the intron-exon structure of

266 these *SbKCS* genes was also related to the evolution of the *KCS* gene family.

267   In view of the fact that the structure and function of proteins were closely related, to have a

268   clearer understanding of the function of *SbKCS* genes, we predicted their secondary structure and

269   constructed a tertiary structure model of all SbKCS proteins. We found that alpha helices,

270   extended strands, beta turns and random coils constituted the secondary structure of SbKCS

271   proteins (Table S3), of which the alpha helix (40.84%-50.00%) accounted for the largest

272   proportion, followed by random coils (32.12%-37.84%) and extended strands (11.99%-16.49%),

273   and the beta turns (4.46%-7.51%) accounted for the smallest proportion. In addition, from the

274   tertiary structure of SbKCS proteins constructed by the homology modeling (Fig. 5), we could

275   see that the SbKCS proteins had similar 3D structures, and the composition and position of

276   secondary structure of the SbKCS proteins could be clearly observed. The α-helix, random coils,

277   β-turn and extended strands that made up the secondary structure of SbKCS proteins were further

278   coiled and folded through the interaction of side chain groups, and a compact spherical space

279   structure was formed by the maintenance of various secondary bonds. The stability of tertiary

280   structure was necessary for proteins to exert biological activity.

281

282   ***Cis*-acting elements analysis of the *SbKCS* gene promoters**

283   To further study the underlying molecular function of *SbKCS* genes, the 1.5 kb promoter

284   sequences of *SbKCS* genes were extracted, and the PlantCARE online site was used to analyze

285   the possible *cis*-acting regulation elements (CAREs). A total of 10 *cis*-acting elements related to

286   abiotic stress response and hormones were found, including two stress-related elements (W-box

287   and MBS) and eight hormone-related elements (STRE, ABRE, I-box, MYB, G-Box, Gap-box,

288   TCT-motif and GA-motif) (Fig. 6). The ABRE related to the abscisic acid response was detected

289   in 19 *SbKCS* genes and the STRE, involved in the salicylic acid response, was revealed in 20

290   *SbKCS* genes. Furthermore, 11 and 14 *SbKCS* genes contained W-box and MBS elements, which

291   were involved in pathogen trauma response and drought-induced processes, respectively. The

292   *cis*-acting elements of G-box, TCT-motif, GA-motif, Gap-box and I-box related to light

293   responsiveness were identified in 20, 9, 4, 2 and 5 *SbKCS* genes, respectively. Besides, the *cis*-

294   acting element MYB was identified in 24 *SbKCS* genes, which played a role in controlling plant

295   secondary metabolism (*Uimari et al.*, *1997*), regulating cell morphogenesis (*Yang et al.*, *2007*)

296   and environmental factor responses (*Slimestad et al.*, *2007*). These results showed that the

297   *SbKCS* genes might play a role in hormone signaling pathways and biotic or abiotic stress

298   responses.

299

300   **Expression analysis of *SbKCS* genes under drought and salt stresses**

301   Previous researches have shown that *KCS* genes are involved in the response to plant abiotic

302   stresses (drought and salt) (*Lee and Suh*, *2015*) and plant hormones (ABA) (*Macková et al.*,

303   *2013*). The identification of *cis*-acting elements related to abiotic stress response in the promoter

304   region of *SbKCS* genes also indicated their potential functions in response to different abiotic

305   stresses. Consequently, to further determine the potential molecular function of *SbKCS* genes in

306   response to abiotic stresses in sorghum, a total of 9 *SbKCS* genes were randomly selected from

307   the five subgroups, and their expression levels at 0, 6, 12 and 24 h after drought and salt

308   treatments were analyzed (Table S4). The results showed that nine randomly selected *SbKCS*

309   genes were up-regulated or down-regulated within 24 h under drought and salt stresses. Under

310   drought stress, the expression of 7 genes except *SbKCS6* and *SbKCS8* were up-regulated at

311   different times after treatment (Fig. 7). The expression level of *SbKCS3* was significantly down-

312   regulated at 24h after treatment (Fig. 7B). In addition, the expression levels of *SbKCS1*, *SbKCS3*,

313   *SbKCS14* and *SbKCS16* significantly increased at 6 h after drought treatment, suggesting that

314   these *SbKCS* genes could rapidly respond to drought stress. Furthermore, *SbKCS14* and

315   *SbKCS18* showed a continuous response at 6-12h and 12-24 h after drought treatment, indicating

316   the key role in drought resistance (Figs. 7F; 7H).

317   Under salt treatment, the expression levels of the 9 selected *SbKCS* genes except *SbKCS7*

318   showed significant differences at different time after treatment compared to the control (Fig. 8).

319   *SbKCS3* and *SbKCS6* showed similar expression pattern under salt treatment, in which the

320   expression levels of the two genes increased firstly at 6 h after salt treatment and then declined

321   (Figs. 8B-8C), suggesting the complex of their function in response to salt stress. In addition,

322   *SbKCS14*, *SbKCS16* and *SbKCS18* were significantly up-regulated at 6 h or 12 h after salt

323 treatment (Figs. 8F; 8H), and *SbKCS1*, *SbKCS8* and *SbKCS21* were down-regulated after

324 treatment (Figs. 8A; 8E; 8I), of which the expression of *SbKCS8* declined significantly at both 6

325 h and 12 h under salt treatment. Interestingly, among all the differentially-expressed genes, no

326 matter under drought or salt treatment, *SbKCS8* showed the largest differentially expressed fold

327 change (at 12h of drought and salt treatments), which implied that *SbKCS8* might be an

328 important gene to resist drought and salt stresses in sorghum. Taken together, we found that

329 different genes showed different expression patterns under stress conditions and some of the

330 *SbKCS* genes might be involved in response to abiotic stresses with different mechanisms.

331

332 **DISCUSSION**

333 Cuticular waxes play a significant role in coping with abiotic stresses in plants and form an

334 important protective barrier in the long-term ecological adaptation process of resisting harsh

335 ecological environment, and biotic and abiotic stresses (*Bernard and Joubès*, *2013*). In recent

336 years, the functions of many genes related to wax synthesis have been widely validated. For

337 example, the *CER1* gene was related to the alkanes biosynthesis process and it was highly linked

338 to responses to biotic and abiotic stresses (*Bourdenx et al.*, *2011*). The overexpression of *CER1*

339 in Arabidopsis could increase the amounts of n-alkanes of chain lengths between 29 and 33

340 carbon atoms and the *cer1* mutant exhibited an altered n-alkanes biosynthesis (*Bourdenx et al.*,

341 *2011*). CER1-LIKE1 in Arabidopsis could control the alkane biosynthesis and wax

342 crystallization (*Pascal et al.*, *2019*). KCS was considered to be a key enzyme that determines the

343 tissue and substrate specificities in the biochemical process of fatty acid elongation in higher

344 plants (*Guo et al.*, *2020*). The biological functions of several *KCS* genes have been clarified by

345 using some Arabidopsis mutants in previous studies. For example, *FAE1* cloned from

346 Arabidopsis was shown to catalyze the biosynthesis of C20 and C22 VLCFA storing lipids in

347 seeds and it was only expressed in seeds (*James et al.*, *1995*). *CUT1/CER6* played a key role in

348 the biosynthesis of cuticular wax C26 and longer VLCFAs during epidermis and root hair

349 development (*Pang et al.*, *2010*). Loss of *FDH1* function led to ectopic organ fusions, suggesting

350   that *KCS* played a critical role in organ development (*Voisin et al.*, *2009*). Furthermore, complete

351   loss of *DAISY/KCS2* function resulted in the accumulation of C20 VLCFA, which suggested that

352   *KCSs* might be necessary for the biosynthesis of cuticular wax (*Franke et al.*, *2009*). So far,

353   limited wax-related genes have been reported in sorghum, and the molecular and biological

354   mechanism of wax synthesis and its response to abiotic stress of sorghum remain unclear.

355   In this study, 25 *SbKCS* genes were identified and characterized in sorghum, which could

356   provide important information about the roles of *SbKCS* in plant's growth and development and

357   abiotic stress resistance. The 25 *KCS* genes in sorghum contained the complete

358   FAE1_CUT1_RppA and ACP_SYN_III_C domain. It was speculated that the function of *KCS*

359   gene family was closely related to these two important conserved domains. The physicochemical

360   properties of 25 SbKCS proteins was further calculated and it was found that they were weakly

361   basic proteins with different isoelectric point. The discrepancies in structure and properties of the

362   SbKCS proteins were presumed to adapt to the changes of the environment in the process of

363   evolution. Lipid metabolism pathways are mainly located in plastid, endoplasmic reticulum (ER)

364   and peroxisome of plants (*Li et al.*, *2013*). In previous studies, it was confirmed that some genes

365   related to Arabidopsis wax synthesis were also expressed in the ER, such as *AtKCR1* and

366   *AtKCR2* (*Beaudoin et al.*, *2009*), *AtKCS9* (*Kim et al.*, *2013*), *CER10* (*Haslam et al.*, *2013*), and

367   *CER1* (*Pascal et al.*, *2019*). However, most of SbKCS proteins were predicted to be mainly

368   expressed in plasma membrane, mitochondria and chloroplast. The differences in subcellar

369   location of KCS proteins in sorghum might indicate that these genes function in different ways.

370   Genes with similar gene structure and protein structure often have similar functions. The

371   SbKCS proteins were revealed to be mainly classified into five distinct subgroups. Most of the

372   *SbKCS* genes have no or only one intron in their open reading frame regions and the typical

373   pattern of gene structure was in accord with those in Arabidopsis (*Joubès et al.*, *2008*) and Malus

374   (*Lian et al.*, *2020*). In addition to further confirming the reliability of the *KCS* superfamily

375   classification, these findings also suggested that the *KCS* superfamily genes had similar

376   characteristics in dicots and monocots, which might lead to their functional similarity of the

377   same families in different species. The conserved motifs of SbKCS proteins were also further

378   analyzed in our study, and all SbKCS proteins shared the same motif composition and similar

379   location distribution, suggesting that these proteins might have similar functions in plants.

380   However, the research on these motifs is still limited, and their regulatory functions needs to be

381   further explored.

382   The analysis of *SbKCS* gene chromosomal localization uncovered a gene cluster on Chr 10.

383   The most common amplification method of gene families is replication (*Cannon et al.*, *2004*). It

384   was speculated that the *KCS* gene family had undergone duplications during evolution.

385   Interestingly, *SbKCS* genes were not found on Chr 7 or Chr 8. We speculated that this might be

386   caused by chromosome shift or fragment loss in the process of evolution. Furthermore, the

387   regulatory mechanism of each gene could be better understood by identifying the upstream *cis*-

388   acting elements of *SbKCS* genes. *SbKCS* genes contained several *cis*-acting elements that could

389   regulate gene expression by combining with *trans*-acting factors. Genes containing different *cis*-

390   acting elements might have different regulatory mechanisms. For instance, the G-box, a light-

391   responsive *cis*-acting element, was involved in the regulation of plant flowering stage, which was

392   greatly affected by light (*Menkens et al.*, *1995*). ABRE, as a homeopathic element regulating

393   seed and bud dormancy, could regulate stomatal switch and control gene expression via abscisic

394   acid response, thus enhancing plant stress resistance (*Narusaka et al.*, *2003*). Genes containing

395   MYB *cis*-acting element were speculated to play a role in the response to cold and drought

396   resistance (*Shinozaki et al.*, *2003*).

397   *KCS* gene family plays a significant role in determining the quantity and composition of

398   VLCFAs, and the wax yield of aerial part of plants is affected by different environmental factors,

399   such as light, moisture and temperature (*Shepherd and Wynne Griffiths*, *2006*). Previous studies

400   had revealed that *AtKCS6* transcript level was affected by changes in different environmental

401   conditions, such as different light conditions or osmotic stress and these environmental signals

402   promoted wax synthesis by up-regulating the expression of key enzymes in wax biosynthesis,

403   thus improving plant stress resistance (*Hooker et al.*, *2002*). Therefore, many studies have

404   focused on the functional identification of *KCS* genes in plants, and the extensive response of

405   *KCS* family genes to abiotic stress has been confirmed in many plants. In Arabidopsis, the

406   expression levels of *KCS1*, *KCS3* and *KCS6* decreased under dark and low temperature

407   conditions, but increased under osmotic tolerance (*Joubès et al*., *2008*). The expression of

408   *AtKCS6* decreased in dark environment, indicating that light was necessary for the expression of

409   *AtKCS6* gene in stems and seedlings, while the expression of *AtKCS6* increased in stems and

410   seedlings after NaCl and PEG treatment or dehydration treatment, which suggested its critical

411   function in resistance to adverse environment (*Hooker et al*., *2002*). The expression of *MdKCS12*

412   and *MdKCS24* in Malus was up-regulated 5 and 10 times, respectively, after 20% PEG treatment

413   for 1 h and meanwhile, *MdKCS12* and *MdKCS24* could also be induced by salt stress (*Lian et al*.,

414   *2020*). These results suggested that plants could respond to abiotic stress through regulating the

415   expression levels of associated *KCS* genes. Therefore, to further study the potential functions of

416   *SbKCS* genes in response to abiotic stress, the expression levels of 9 randomly selected *SbKCS*

417   genes were studied under drought and salt stress conditions by qRT-PCR. Most of the 9 selected

418   *SbKCS* genes were obviously induced or repressed when subjected to drought or salt stresses,

419   clarifying that *SbKCSs* might play a role in response to drought and salt stresses in sorghum

420   seedling. Meanwhile, we found that some *SbKCS* genes showed similar expression patterns

421   under drought and salt treatments. For instance, the expression of *SbKCS8* was down-regulated

422   under drought and salt treatments, whereas *SbKCS18* was up-regulated under these two stress

423   treatments, which suggested that these two genes may respond to drought and salt stresses by a

424   similar mechanism. Remarkably, the expression of *SbKCS1* was up-regulated at 6 h and 12 h

425   after drought stress, but was down-regulated in the salt stress, indicating that *SbKCS1* might play

426   different roles under two different stresses. The results further indicated the facticity of these

427   genes and the possibility of their function in two stress responses. In brief, these results will

428   provide useful information for the subsequent functional studies of *SbKCS* genes, and contribute

429   to improve the stress resistance of sorghum germplasms.

430   **CONCLUSIONS**

431   In this study, we identified 25 *KCS* genes in sorghum and classified them into five subgroups

432   based on the phylogenetic tree, and the SbKCS and ZmKCS proteins showed a closer

433   evolutionary relationship. These *SbKCS* family genes were unevenly distributed on eight

434   chromosomes of Sorghum. The analysis of their conserved motif compositions and gene

435   structures showed high levels of consistency and similarity within the same subgroup. *Cis*-acting

436   element analysis of these identified *SbKCS* genes contributed to a further study of their

437   performance in plant growth and stress responses. Besides, gene expression levels of *SbKCS*

438   genes in leaves under drought and salt stresses were also analyzed by qRT-PCR, which hinted

439   their potential roles in response to abiotic stresses. The results found in this study provided a

440   theoretical basis for investigating the function of *KCSs* in sorghum and also provided valuable

441   gene resources for the application of wax-related genes to improve stress resistance and genetic

442   characteristics of sorghum.

443   **ADDITIONAL INFORMATION AND DECLARATIONS**

444   **Funding**

445   This work was supported by the Social Livelihood Science and Technology Project of Nantong

446   City, China (MS22020033), the Nantong University Scientific Research Start-up project for

447   Introducing Talents (135420609055), the Practice Innovation Training Program Projects for

448   College Students in 2021 (202110304081Y) and the Large Instruments Open Foundation of

449   Nantong University.

450

451   **Grant Disclosures**

452   The following grant information was disclosed by the authors:

453   Social Livelihood Science and Technology Project of Nantong City, China: MS22020033.

454   The Nantong University Scientific Research Start-up project for Introducing Talents:

455   135420609055.

456   The Practice Innovation Training Program Projects for College Students in 2021:

457   202110304081Y.

458

Competing Interests

The authors declare there are no competing interests. Junping Wu is employed by Nantong

Changjiang Seed Co., Ltd.

Author Contributions

- Aixia Zhang performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Jingjing Xu, Xin Xu and Junping Wu analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.
- Ping Li, Baohua Wang, and Hui Fang conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The raw data is available in the Supplemental Files.

REFERENCES

Bourdenx B, Bernard A, Domergue F, Pascal S, Léger A, Roby D, Pervent M, Vile D, Haslam RP, Napier JA, Lessire R, Joubès J. 2011. Overexpression of Arabidopsis *ECERIFERUM1* promotes wax very-long-chain alkane biosynthesis and influences plant response to biotic and abiotic stresses. *Plant Physiology* 156(1):29-45 DOI 10.1104/pp.111.172320.

Beaudoin F, Wu X, Li F, Haslam RP, Markham JE, Zheng H, Napier JA, Kunst L. 2009. Functional characterization of the Arabidopsis β-ketoacyl-coenzyme A reductase candidates of the fatty acid elongase. *Plant Physiology* 150(3):1174-91 DOI 10.1104/pp.109.137497.

Bernard A, Joubès J. 2013. Arabidopsis cuticular waxes: advances in synthesis, export and regulation. *Progress in Lipid Research* 52(1):110-29 DOI 10.1016/j.plipres.2012.10.002.

Cannon SB, Mitra A, Baumgarten A, Young ND, May G. 2004. The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biology* 4:10 DOI

487   10.1186/1471-2229-4-10.

488   Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. 2020. TBtools: an integrative toolkit

489   developed for interactive analyses of big biological data. *Molecular plant*, 13(8): 1194-1202 DOI

490   10.1016/j.molp.2020.06.009.

491   Costaglioli P, Joubès J, Garcia C, Stef M, Arveiler B, Lessire R, Garbay B. 2005. Profiling candidate genes

492   involved in wax biosynthesis in *Arabidopsis thaliana* by microarray analysis. *Biochimica et biophysica*

493   *acta* 1734(3):247-58 DOI 10.1016/j.bbalip.2005.04.002.

494   Denic V, Weissman JS. 2007. A molecular caliper mechanism for determining very long-chain fatty acid

495   length. Cell 130(4):663-77 DOI 10.1016/j.cell.2007.06.031.

496   Fan Y, Yuan C, Jin Y, Hu GR, Li FL. 2018. Characterization of 3-ketoacyl-coA synthase in a nervonic acid

497   producing oleaginous microalgae Mychonastes afer. *Algal Research-biomass Biofuels and Bioproducts*

498   31:225-231 DOI 10.1016/j.algal.2018.02.017.

499   Franke R, Höfer R, Briesen I, Emsermann M, Efremova N, Yephremov A, Schreiber L. 2009. The *DAISY* gene

500   from Arabidopsis encodes a fatty acid elongase condensing enzyme involved in the biosynthesis of

501   aliphatic suberin in roots and the chalaza-micropyle region of seeds. *Plant Journal* 57(1):80-95 DOI

502   10.1111/j.1365-313X.2008.03674-x.

503   Guo HS, Zhang YM, Sun XQ, Li MM, Hang YY, Xue JY. 2016. Evolution of the KCS gene family in plants:

504   the history of gene duplication, sub/neofunctionalization and redundancy. *Molecular Genetics and*

505   *Genomics* 291(2):739-52 DOI 10.1007/s00438-015-1142-3.

506   Guo W, Wu Q, Yang L, Hu W, Liu D, Liu Y. 2020. Ectopic expression of *CsKCS6* from navel orange

507   promotes the production of very-long-chain fatty acids (VLCFAs) and increases the abiotic stress

508   tolerance of *Arabidopsis thaliana*. *Frontiers in Plant Science* 11:564656 DOI 10.3389/fpls.2020.564656.

509   Haslam TM, Kunst L. 2013. Extending the story of very-long-chain fatty acid elongation. *Plant Science*

510   210:93-107 DOI 10.1016/j.plantsci.2013.05.008.

511   Hooker TS, Millar AA, Kunst L. 2002. Significance of the expression of the *CER6* condensing enzyme for

512   cuticular wax production in Arabidopsis. *Plant Physiology* 129(4):1568-1580 DOI 10.1104/pp.003707.

513   Hu XJ, Zhang ZB, Fu ZY, Xu P, Hu SB, Li WQ. 2010. Significance of a β-ketoacyl-CoA synthase gene

514   expression for wheat tolerance to adverse environments. *Biologia Plantarum* 54:575-578 DOI

515   10.1007/s10535-010-0103-2.

516   James DW Jr, Lim E, Keller J, Plooy I, Ralston E, Dooner HK. 1995. Directed tagging of the Arabidopsis

517   *FATTY ACID ELONGATION1* (*FAE1*) gene with the maize transposon activator. *Plant Cell* 7(3):309-19

518   DOI: 10.1105/tpc.7.3.309.

519   Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ. 1998. Multiple sequence alignment with

520   Clustal X. *Trends in Biochemical Sciences* 23:403-5 DOI 10.1016/s0968-0004(98)01285-7.

521   Joubès J, Raffaele S, Bourdenx B, Garcia C, Laroche-Traineau J, Moreau P, Domergue F, Lessire R. 2008.

522   The VLCFA elongase gene family in *Arabidopsis thaliana*: phylogenetic analysis, 3D modelling and

523   expression profiling. *Plant Molecular Biology* 67(5):547-66 DOI 10.1007/s11103-008-9339-z.

524   Kim J, Jung JH, Lee SB, Go YS, Kim HJ, Cahoon R, Markham JE, Cahoon EB, Suh MC. 2013. Arabidopsis

525   3-ketoacyl-coenzyme a synthase9 is involved in the synthesis of tetracosanoic acids as precursors of

526   cuticular waxes, suberins, sphingolipids, and phospholipids. *Plant Physiology* 162(2):567-80 DOI

527   10.1104/pp.112.210450.

528   Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for

529   bigger datasets. *Molecular Biology and Evolution* 33(7):1870-74 DOI 10.1093/molbev/msw054.

530   Lea US, Slimestad R, Smedvig P, Lillo C. 2007. Nitrogen deficiency enhances expression of specific MYB

531   and bHLH transcription factors and accumulation of end products in the flavonoid pathway. *Planta*

532   225(5):1245-53 DOI 10.1007/s00425-006-0414-x.

533   Lee SB, Suh MC. 2015. Advances in the understanding of cuticular waxes in *Arabidopsis thaliana* and crop

534   species. *Plant Cell Reports* 34(4):557-72 DOI 10.1007/s00299-015-1772-2.

535   Lee SB, Suh MC. 2015. Cuticular wax biosynthesis is up-regulated by the *MYB94* transcription factor in

536   Arabidopsis. *Plant and Cell Physiology* 56(1):48-60 DOI 10.1093/pcp/pcu142.

537   Mistry J, Chuguransky S, Williams L, et al. 2021. Pfam: The protein families database in 2021. *Nucleic Acids*

538   *Research* 49(D1): D412-D419 DOI 10.1093/nar/gkaa913.

539   Li C, Haslam TM, Krüger A, Schneider LM, Mishina K, Samuels L, Yang H, Kunst L, Schaffrath U, Nawrath

540   C, Chen G, Komatsuda T, von Wettstein-Knowles P. 2018. The β-ketoacyl-CoA synthase Hv *KCS1*,

541   encoded by Cer-zh, plays a key role in synthesis of barley leaf wax and germination of barley powdery

542   mildew. *Plant and Cell Physiology* 59(4):806-22 DOI 10.1093/pcp/pcy020.

543 Li H, Peng Z, Yang X, Wang W, Fu J, Wang J, Han Y, Chai Y, Guo T, Yang N, Liu J, Warburton ML, Cheng

544   Y, Hao X, Zhang P, Zhao J, Liu Y, Wang G, Li J, Yan J. 2013. Genome-wide association study dissects

545   the genetic architecture of oil biosynthesis in maize kernels. *Nature Genetics* 45(1):43-50 DOI

546   10.1038/ng.2484.

547 Lian XY, Wang X, Gao HN, Jiang H, Mao K, You CX, Li YY, Hao YJ. 2020. Genome wide analysis and

548   functional identification of MdKCS genes in apple. *Plant Physiology and Biochemistry* 151:299-312 DOI

549   10.1016/j.plaphy.2020.03.034.

550 Liu M, Ma Z, Wang A, Zheng T, Huang L, Sun W, Zhang Y, Jin W, Zhan J, Cai Y, Tang Y, Wu Q, Tang Z,

551   Bu T, Li C, Chen H. 2018. Genome-wide investigation of the auxin response factor gene family in

552   Tartary buckwheat (*Fagopyrum tataricum*). *International Journal of Molecular Sciences* 19(11):3526

553   DOI 10.3390/ijms19113526.

554 Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR

555   and the $2^{-\Delta\Delta CT}$ method. *Methods* 25:402-8 DOI 10.1006/meth.2001.1262.

556 Macková J, Vašková M, Macek P, Hronková M, Schreiber L, Šantrůček J. 2013. Plant response to drought

557   stress simulated by ABA application: changes in chemical composition of cuticular waxes. *Environmental*

558   *and Experimental Botany* 86:70-75 DOI 10.1016/j.envexpbot.2010.06.005.

559 Menkens AE, Schindler U, Cashmore AR. 1995. The G-box: a ubiquitous regulatory DNA element in plants

560   bound by the GBF family of bZIP proteins. *Trends in Biochemical Sciences* 20(12):506-10 DOI

561   10.1016/s0968-0004(00)89118-5.

562 Narusaka Y, Nakashima K, Shinwari ZK, Sakuma Y, Furihata T, Abe H, Narusaka M, Shinozaki K,

563   Yamaguchi-Shinozaki K. 2003. Interaction between two *cis*-acting elements, ABRE and DRE, in

564   ABA-dependent expression of Arabidopsis *rd29A* gene in response to dehydration and high-salinity

565   stresses. *Plant Journal* 34(2):137-48 DOI 10.1046/j.1365-313x.2003.01708-x.

566 Pang CY, Wang H, Pang Y, Xu C, Jiao Y, Qin YM, Western TL, Yu SX, Zhu YX. 2010. Comparative

567   proteomics indicates that biosynthesis of pectic precursors is important for cotton fiber and Arabidopsis

568   root hair elongation. *Molecular & Cellular Proteomics* 9(9):2019-33 DOI 10.1074/mcp.M110.000349.

569 Pascal S, Bernard A, Deslous P, Gronnier J, Fournier-Goss A, Domergue F, Rowland O, Joubès J. 2019.

570   Arabidopsis CER1-LIKE1 functions in a cuticular very-long-chain alkane-forming complex. *Plant*

571   *Physiology* 179(2):415-432 DOI 10.1104/pp.18.01075.

572 Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros

573   T, Poliakov A, Schmutz J, Spannagl M, Tang H, Wang X, Wicker T, Bharti AK, Chapman J, Feltus FA,

574   Gowik U, Grigoriev IV, Lyons E, Maher CA, Martis M, Narechania A, Otillar RP, Penning BW, Salamov

575   AA, Wang Y, Zhang L, Carpita NC, Freeling M, Gingle AR, Hash CT, Keller B, Klein P, Kresovich S,

576   McCann MC, Ming R, Peterson DG, Mehboob-ur-Rahman, Ware D, Westhoff P, Mayer KF, Messing J,

577   Rokhsar DS. 2009. The Sorghum bicolor genome and the diversification of grasses. *Nature*

578   457(7229):551-556 DOI 10.1038/nature07723.

579 Paul S, Gable K, Beaudoin F, Cahoon E, Jaworski J, Napier JA, Dunn TM. 2006. Members of the Arabidopsis

580   FAE1-like 3-ketoacyl-CoA synthase gene family substitute for the Elop proteins of *Saccharomyces*

581   *cerevisiae*. *Journal of Biological Chemistry* 281:9018-9029 DOI 10.1074/jbc.M507723200.

582 Pruitt RE, Vielle-Calzada JP, Ploense SE, Grossniklaus U, Lolle SJ. 2000. FIDDLEHEAD, a gene required to

583   suppress epidermal cell interactions in Arabidopsis, encodes a putative lipid biosynthetic enzyme.

584   *Proceedings of the National Academy of Sciences of the United States of America* 97:1311-6 DOI

585   10.1073/pnas.97.3.1311.

586 Qin YM, Hu CY, Pang Y, Kastaniotis AJ, Hiltunen JK, Zhu YX. 2007. Saturated very-long-chain fatty acids

587   promote cotton fiber and Arabidopsis cell elongation by activating ethylene biosynthesis. *The Plant Cell*

588   19:3692-3704 DOI 10.1105/tpc.107.054437.

589 Richardson A, Boscari A, Schreiber L, Kerstiens G, Jarvis M, Herzyk P, Fricke W. 2007. Cloning and

590   expression analysis of candidate genes involved in wax deposition along the growing barley (*Hordeum*

591   *vulgare*) leaf. *Planta* 226:1459-73 DOI 10.1007/s00425-007-0585-0.

592 Seo PJ, Lee SB, Suh MC, Park MJ, Go YS, Park CM. 2011. The *MYB96* transcription factor regulates

593   cuticular wax biosynthesis under drought conditions in Arabidopsis. *The Plant Cell* 23:1138-52 DOI

594   10.1105/tpc.111.083485.

595 Shepherd T, Wynne Griffiths D. 2006. The effects of stress on plant cuticular waxes. *New Phytologist*
596   171:469-99 DOI 10.1111/j.1469-8137.2006.01826-x.

597 Shinozaki K, Yamaguchi-Shinozaki K, Seki M. 2003. Regulatory network of gene expression in the drought
598   and cold stress responses. *Current Opinion in Plant Biology* 6:410-17 DOI 10.1016/s1369-
599   5266(03)00092-x.

600 Swift ML. 1997. GraphPad prism, data analysis, and scientific graphing. *Journal of Chemical Information and*
601   *Computer Sciences* 37:411-12 DOI 10.1021/ci960402j.

602 Todd J, Post-Beittenmiller D, Jaworski JG. 1999. *KCS1* encodes a fatty acid elongase 3-ketoacyl-CoA
603   synthase affecting wax biosynthesis in *Arabidopsis thaliana*. *The Plant Journal* 17:119-30 DOI
604   10.1046/j.1365-313x.1999.00352-x.

605 Uimari A, Strommer J. 1997. Myb26: a MYB-like protein of pea flowers with affinity for promoters of
606   phenylpropanoid genes. *Plant Journal* 12:1273-84 DOI 10.1046/j.1365-313x.

607 Voisin D, Nawrath C, Kurdyukov S, Franke RB, Reina-Pinto JJ, Efremova N, Will I, Schreiber L, Yephremov
608   A. 2009. Dissection of the complex phenotype in cuticular mutants of Arabidopsis reveals a role of
609   SERRATE as a mediator. *Plos Genetics* 5: e1000703 DOI 10.1371/journal.pgen.1000703.

610 Voorrips RE. 2002. MapChart: software for the graphical presentation of linkage maps and QTLs. *Journal of*
611   *Heredity* 93:77-8 DOI 10.1093/jhered/93.1.77.

612 Wang J, Du Z, Huo X, Zhou J, Chen Y, Zhang J, Pan A, Wang X, Wang F, Zhang J. 2020. Genome-wide
613   analysis of PRR gene family uncovers their roles in circadian rhythmic changes and response to drought
614   stress in *Gossypium hirsutum* L. *Peer Journal* 8: e9936 DOI 10.7717/peerj.9936.

615 Wang JH, Gu KD, Han PL, Yu JQ, Wang CK, Zhang QY, You CX, Hu DG, Hao YJ. 2020. Apple ethylene
616   response factor MdERF11 confers resistance to fungal pathogen Botryosphaeria dothidea. *Plant Science*
617   291:110351 DOI 10.1016/j.plantsci.2019.110351.

618 Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C,
619   Bordoli L, Lepore R, Schwede T. 2018. SWISS-MODEL: homology modelling of protein structures and
620   complexes. *Nucleic Acids Research* 46(W1): W296-W303 DOI 10.1093/nar/gky427

621 Wei X, Liu L, Lu C, Yuan F, Han G, Wang B. 2021. *SbCASP4* improves salt exclusion by enhancing the root

622 apoplastic barrier. *Planta* 254(4):81 DOI 10.1007/s00425-021-03731-z.

623 Xiao GH, Wang K, Huang G, Zhu YX. 2016. Genome-scale analysis of the cotton KCS gene family revealed a

624 binary mode of action for gibberellin A regulated fiber growth. *Journal of Integrative Plant Biology*

625 58:577-89 DOI 10.1111/jipb.12429.

626 Yang XY, Li JG, Pei M, Gu H, Chen ZL, Qu LJ. 2007. Over-expression of a flower-specific transcription

627 factor gene *AtMYB24* causes aberrant anther development. *Plant Cell Reports* 26:219-228 DOI

628 10.1007/s00299-006-0229-z.

629 Zhang GY, Shan S, Wu YB, Huang SQ, Li DF, Deng JL, Qi JM. 2019. The KCS gene is involved in the

630 formation of chloroplast stromules and other physiological processes in jute (*Corchorus capsularis* L.).

631 *Industrial Crops and Products* 141:111781 DOI 10.1016/j.indcrop.2019.111781.

632 Zheng H, Sun X, Li J, Song Y, Song J, Wang F, Liu L, Zhang X, Sui N. 2021. Analysis of N[6]-

633 methyladenosine reveals a new important mechanism regulating the salt tolerance of sweet sorghum.

634 *Plant Science* 304:110801 DOI 10.1016/j.plantsci.2020.110801.

Figure 1

Chromosomal localization of the *SbKCS* genes.

25 *SbKCS* genes were mapped on 8 sorghum chromosomes. The chromosome numbers were indicated at the top of each vertical green bar. The gene names on each chromosome corresponded to the approximate locations of each *SbKCS* genes. The number shown to the left of each vertical green bar was the specific physical position of each *SbKCS* gene on the chromosome.
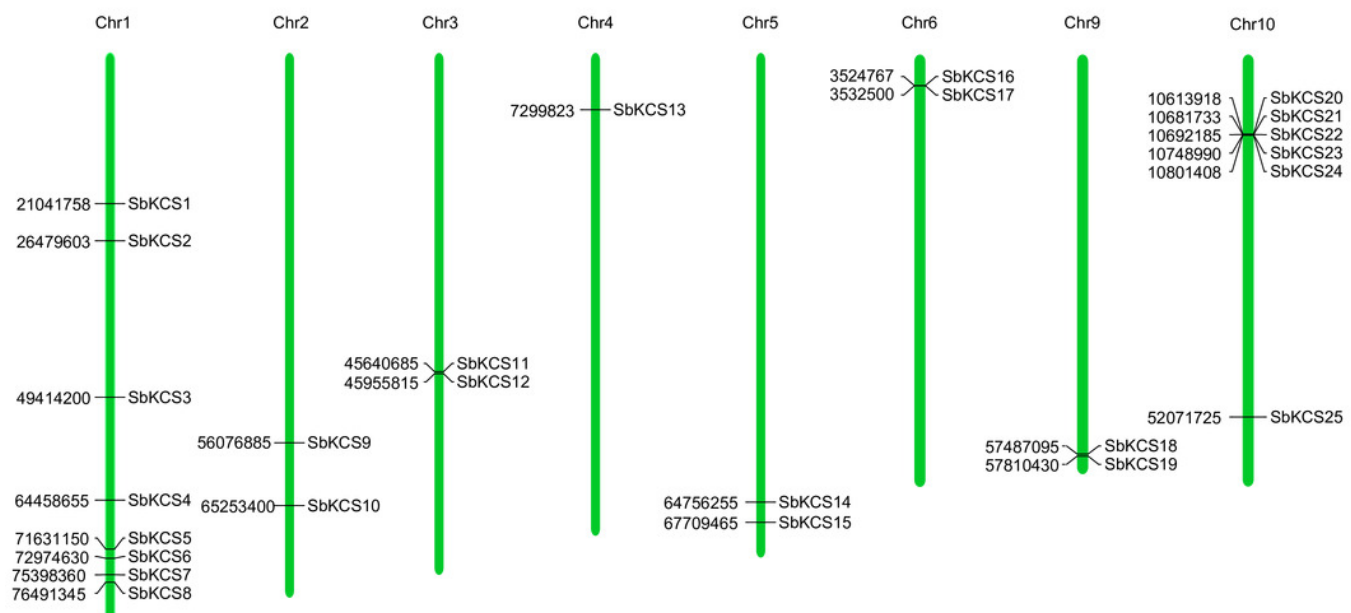
Figure 2

The prediction of subcellular localization for SbKCS proteins.

The color and the size of the circles indicated the reliability of the prediction results. The name of each protein was shown on the left. The site name for the predicted subcellular localization of each SbKCS protein was shown at the bottom.
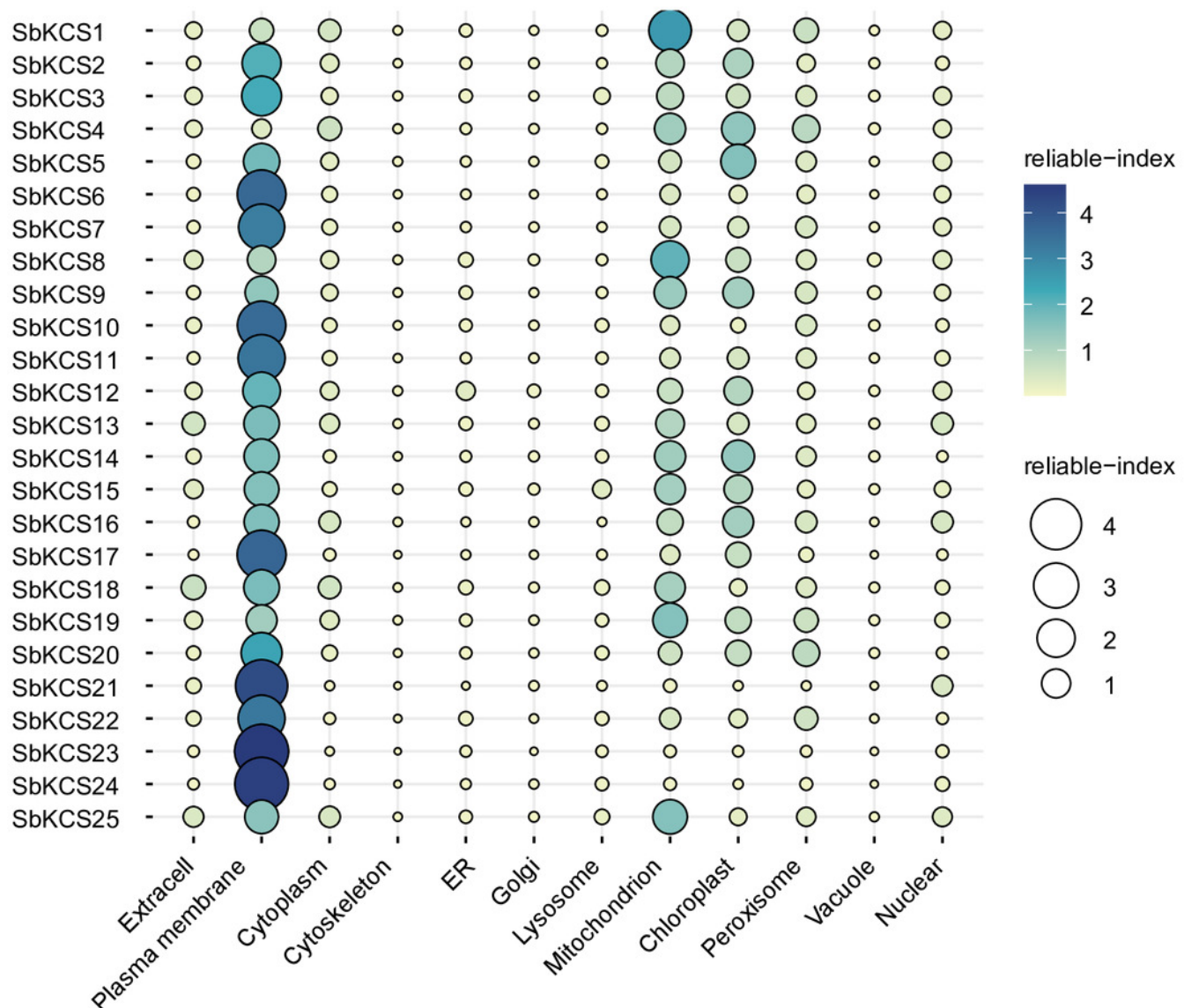
Figure 3

Phylogenetic tree of KCS gene family members in five species.

The full-length protein sequences from Arabidopsis, sorghum, maize, rice and *Brachypodium distachyon* were used to construct the phylogenetic tree by MEGA 7.0 software based on a Maximum Likelihood (ML) method with bootstrap replicates of 1000. Five groups were highlighted in different colors. The members of KCS family in sorghum were marked in red.
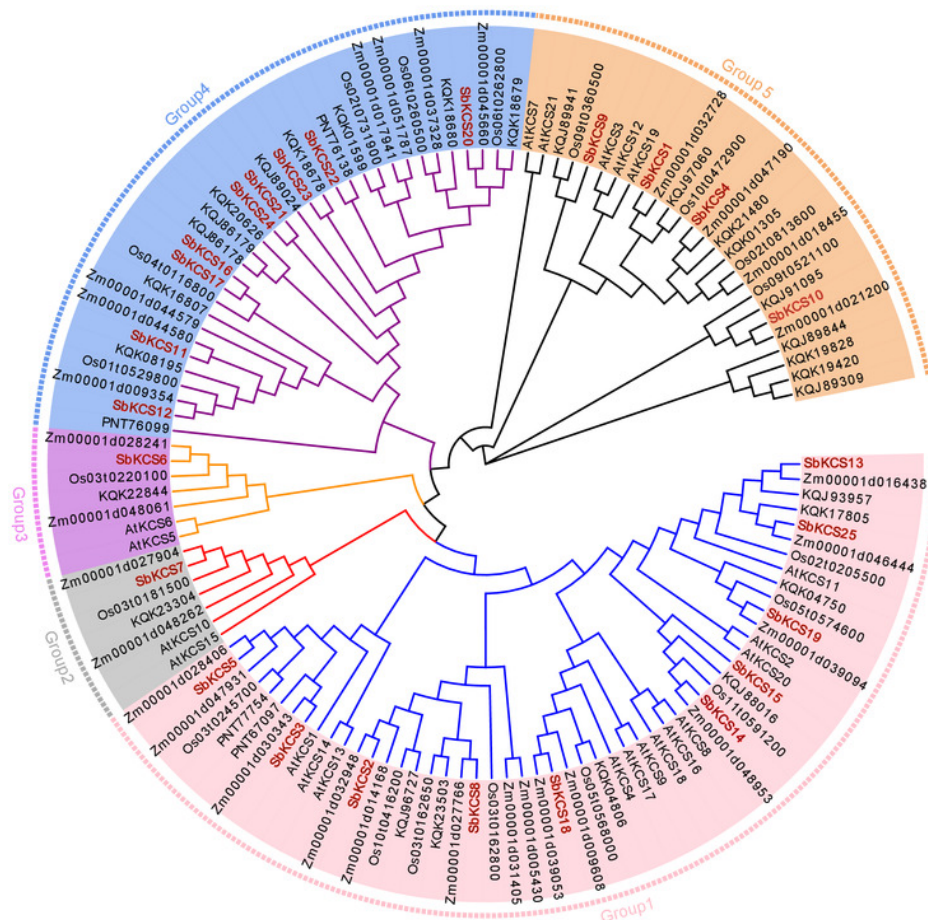
Figure 4

The motif and gene structure analysis of the KCS family members in sorghum.

(A) Motif analysis of the 25 SbKCS proteins. The conserved motifs in the SbKCS proteins were identified with MEME software. The grey lines denoted the non-conserved sequences, and each motif was showed by a different color. The length of motifs in each protein was presented in proportion. (B) The exon-intron structure analysis of the 25 *SbKCS* genes. Exons were represented by yellow rectangles, introns were represented by gray lines, and green rectangles represented UTR regions.
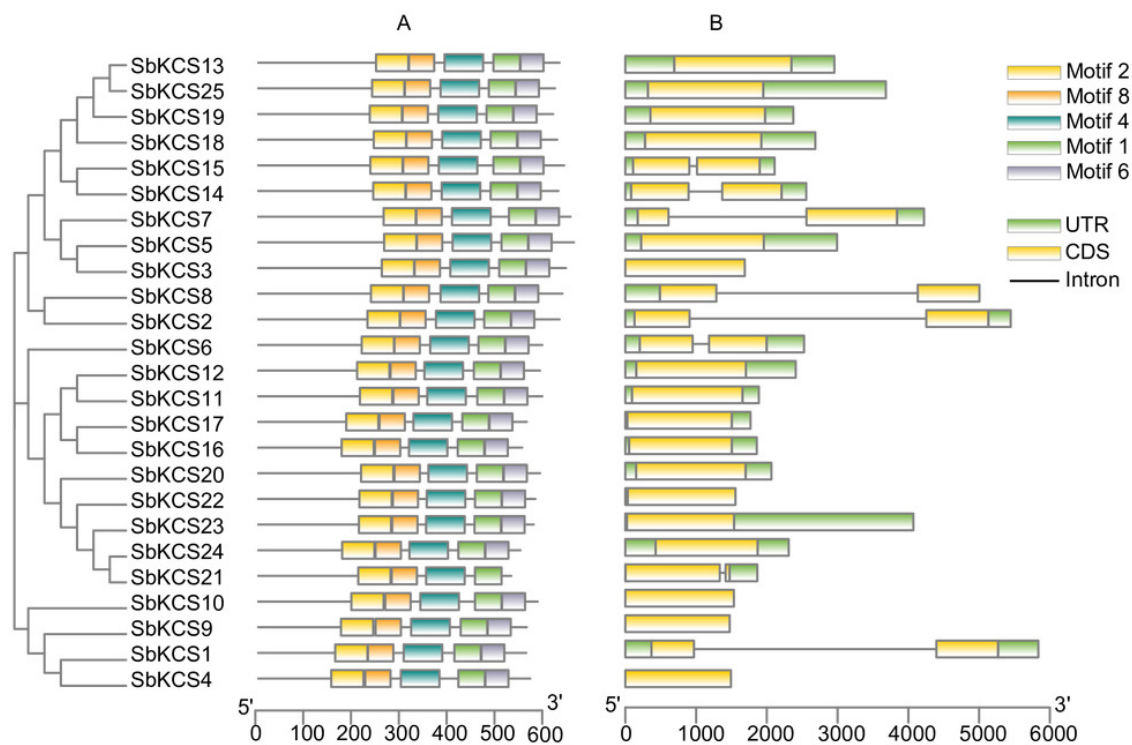
Figure 5

Three-dimensional (3D) models of the SbKCS proteins.

The 3D model of SbKCS protein was constructed by using the method of homologous protein modeling on SWISS-MODEL online website. Proteins of the different subfamily were shown in different colors at the bottom of each 3D model. Orange fonts represented the Group 1 subfamily; Blue fonts represented the Group 2 subfamily; Black fonts represent the Group 3 subfamily; Purple fonts represented the Group 4 subfamily and red fonts represented the Group 5 subfamily.

SbKCS1 SbKCS4 SbKCS9 SbKCS10 SbKCS7

SbKCS5 SbKCS8 SbKCS2 SbKCS3 SbKCS13

SbKCS14 SbKCS15 SbKCS19 SbKCS25 SbKCS18

SbKCS17 SbKCS12 SbKCS16 SbKCS20 SbKCS11

SbKCS22 SbKCS21 SbKCS23 SbKCS24 SbKCS6
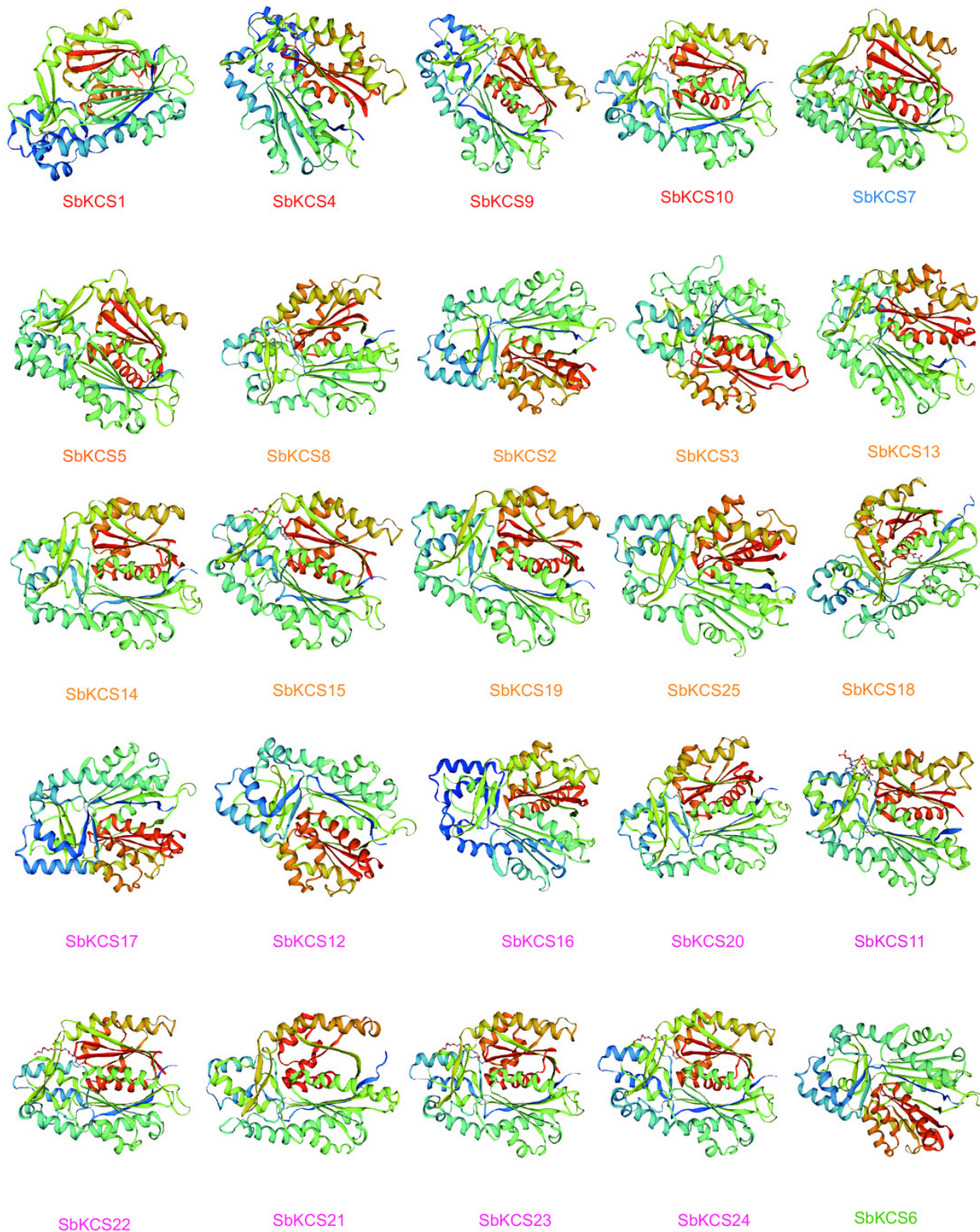
Figure 6

Cis-acting elements in the *SbKCS* gene promoters.

The different *cis*-acting elements were shown in different colors. The name of each gene was shown on the left. Promoter sequence length was displayed in proportion.
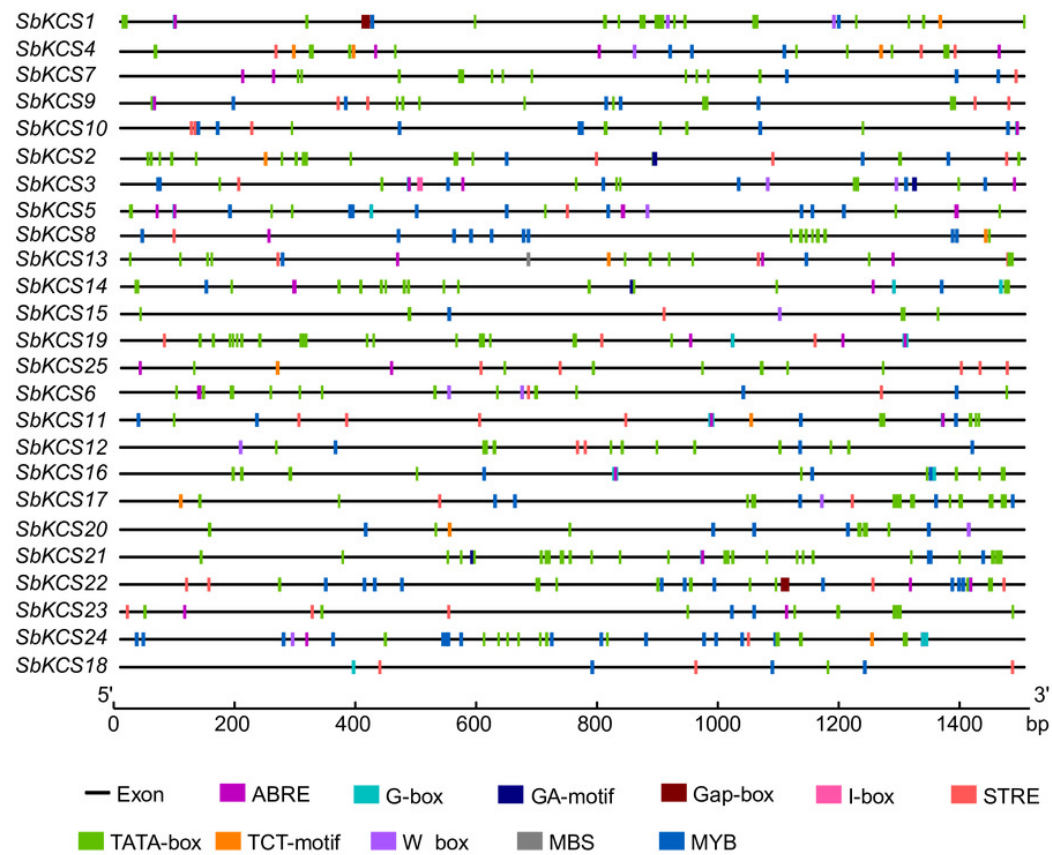
Figure 7

Expression profile of nine SbKCS genes under drought stress treatment.

The expression level of the 9 *SbKCS* genes (A-I) were validated by qRT-PCR. Seedings were treated with drought (20% PEG6000) and leaves were sampled at 0, 6, 12 and 24h. Date represent mean ± standard error (SE) of three biological replicates. Statistically significant differences between the control (CK) and treatment groups (PEG) are indicated using asterisks (* $p < 0.05$, ** $p < 0.01$, *** $P < 0.001$, **** $P < 0.0001$; independent Student's *t*-test).
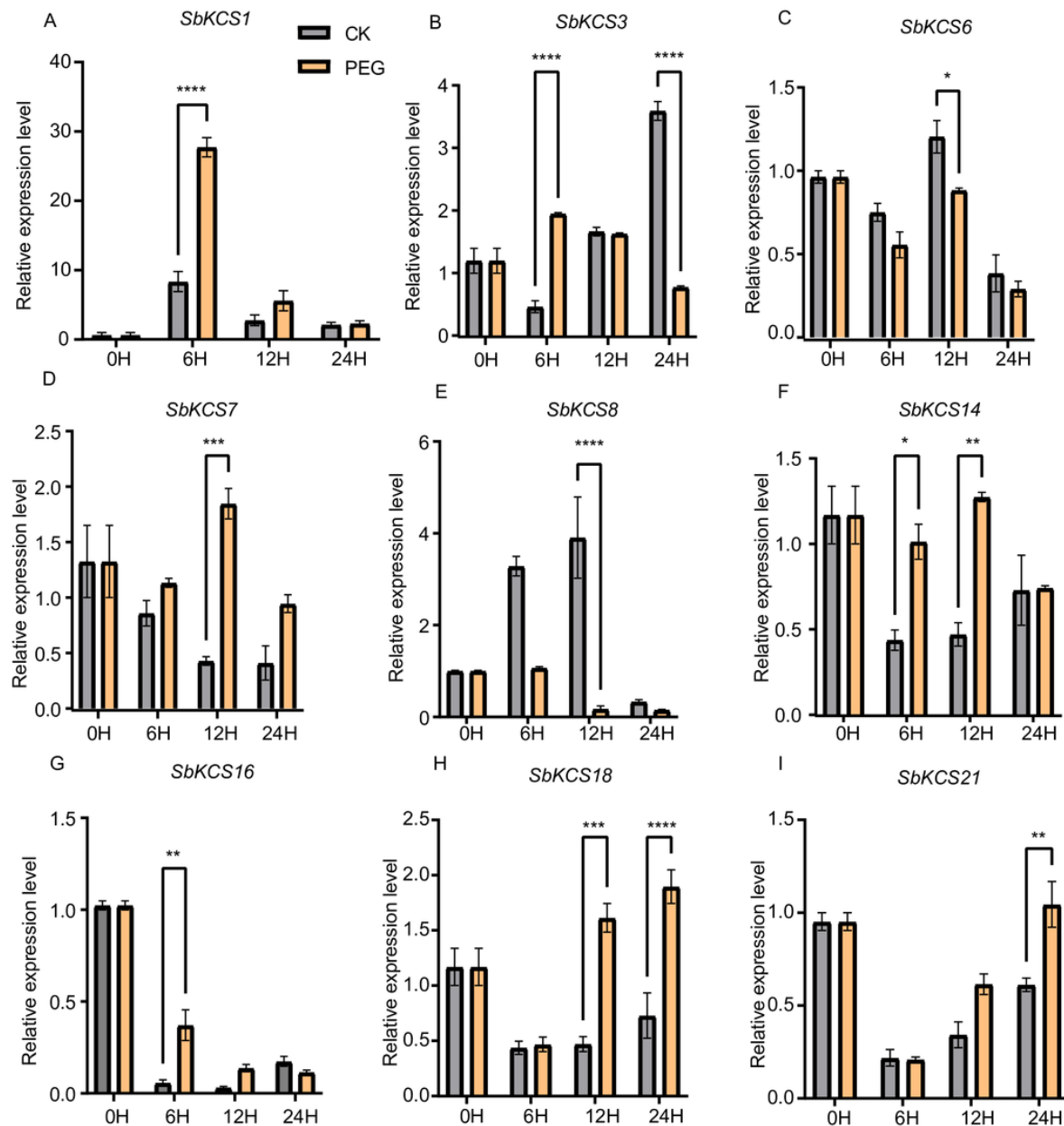
Figure 8

Expression profile of nine SbKCS genes under salt stress treatment.

The expression level of the 9 *SbKCS* genes (A-I) were validated by qRT-PCR. Seedings were treated with salt (150 mM NaCl) and leaves were sampled at 0, 6, 12 and 24h. Date represent mean ± standard error (SE) of three biological replicates. Statistically significant differences between the control (CK) and treatment groups (Salt) are indicated using asterisks (* $p < 0.05$, ** $p < 0.01$, *** $P < 0.001$, **** $P < 0.0001$; independent Student's *t*-test).