Identification and molecular characterization of the alternative spliced variants of beta carbonic anhydrase 1 (βCA1) from *Arabidopsis thaliana*

Jinyu Shen[1,#], Zhiyong Li[2,3,#], Yajuan Fu[1], Jiansheng Liang[1,3]

[1] Department of Biology, Southern University of Science and Technology, Shenzhen, Guangdong, China
[2] Academy for Advanced Interdisciplinary Studies, Southern University of Science and Technology, Shenzhen, Guangdong, China
[3] Key Laboratory of Molecular Design for Plant Cell Factory of Guangdong Higher Education Institutes, Department of Biology, Southern University of Science and Technology, Shenzhen, Guangdong, China

[#] These authors contributed equally to this work

Corresponding Authors:
Zhiyong Li
No.1088, Xueyuan Road, Shenzhen, Guangdong, 510085, China
Email address: lizy6@sustech.edu.cn
Jiansheng Liang
No.1088, Xueyuan Road, Shenzhen, Guangdong, 510085, China
Email address: liangjs@sustech.edu.cn

Abstract

Carbonic anhydrases (CAs) are ubiquitous zinc metalloenzymes that catalyze the interconversion of carbon dioxide and bicarbonate. Higher plants mainly contain the three evolutionarily distinct CA families αCA, βCA, and γCA, with each represented by multiple isoforms. Alternative splicing (AS) of the CA transcripts is common. However, there is little information on the spliced variants of individual CA isoforms. In this study, we focused on the characterization of spliced variants of βCA1 from *Arabidopsis*. The expression patterns and subcellular localization of the individual spliced variants of βCA1 were examined. The results showed that the spliced variants of βCA1 possessed different subcellular and tissue distributions and responded differently to environmental stimuli. Additionally, we addressed the physiological role of βCA1 in heat stress response and its protein-protein interaction (PPI) network. Our results showed that *βCA1* was regulated by heat stresses, and *bca1* mutant was hypersensitive to heat stress, indicating a role for βCA1 in heat stress response. Furthermore, PPI network analysis revealed that βCA1 interacts with multiple proteins involved in several processes, including

40 photosynthesis, metabolism, and the stress response, and these will provide new avenues for
41 future investigations of βCA1.
42

Introduction

44   Carbonic anhydrases (CAs) are a group of Zn-containing enzymes that catalyze the
45 reversible hydration of carbon dioxide (CO_2), generating proton (H^+) and bicarbonate (HCO_3^-).
46 As CO_2 is the main source of carbon, these CA enzymes are involved in crucial physiological
47 processes, including almost all metabolic processes in higher plants and algae. Molecular,
48 biochemical, and genetic studies of CAs in a wide range of tissues across diverse plant species
49 have determined that CAs are involved in a wide range of diverse biological processes, including
50 the provision of bicarbonate for anaplerotic reactions (Werdan and Heldt, 1972; Espie and
51 Kimber, 2011), gas and ion exchange (Jacobson et al., 1975; Randall and Val, 1995), stomatal
52 movement and development (Hu et al., 2010; Engineer et al., 2014; Hu et al., 2015; Wang et al.,
53 2016), biotic and abiotic stress responses (Slaymaker et al., 2002; Restrepo et al., 2005; Yu et al.,
54 2007; Jung et al., 2008; Tianpei et al., 2015; Jing et al., 2019; Zhou et al., 2020), and lipid and
55 fatty acid biosynthesis (Price et al., 1994; Hoang et al., 2002a; Hoang et al., 2002b).
56   CAs can be classified into several evolutionarily independent families based on their
57 conserved nucleotide sequences, including alpha CAs (α), beta CAs (β), gamma CAs (γ) and
58 delta CAs (δ) (Hewett-Emmett and Tashian, 1996; Liljas and Laurberg, 2000). Higher plants
59 contain three evolutionarily distinct CA families, and each family is represented by multiple
60 isoforms (Moroney et al., 2001). A similar number of genes are present in the genomes of plant
61 species as diverse as mosses, monocots, and dicots. For example, *Arabidopsis* contains 19
62 carbonic anhydrase genes (8 αCA, 6 βCA, 5 γCA), and rice too has a similar number (9 αCA, 3
63 βCA, 4 γCA) (DiMario et al., 2017). However, there are 25 genes coding for CA in soybean due
64 to a past genome duplication (DiMario et al., 2017).
65   Alternative splicing (AS) is an important regulatory mechanism that substantially expands
66 eukaryotic transcriptome and proteome diversity and represents an additional level of cellular
67 regulation (Naftelberg et al., 2015). In many species, pre-mRNAs are alternatively spliced during
68 different developmental stages or under stress conditions, allowing organisms to reprogram their
69 regulatory networks (Mastrangelo, et al., 2012; Staiger and Brown, 2013; Laloum, et al., 2018).
70 AS has been found in some CA family genes. In human, CA IX, one of the 12 enzymatically
71 active carbonic anhydrase isoforms, includes one spliced variant lacking the catalytic domain at
72 the C-terminal, which functionally interferes with the full-length CA IX protein (Barathova et al.,
73 2008). AS was also occurred in another human CA gene, which led to produce two different
74 spliced protein forms with both linked to the aggressive behavior of cancer cells (Haapasalo et
75 al., 2008). Several isoforms of βCA from *Arabidopsis*, including *AtβCA1*, *AtβCA2*, and *AtβCA4*,
76 are predicted to encode at least two mRNA transcripts via AS (DiMario et al., 2017). Two
77 spliced variants of *βCA1* have been suggested to be present in *Arabidopsis* (Oh et al., 2014;
78 Rudenko et al., 2017). Further studies showed that the two spliced transcripts of *AtβCA4* are
79 expressed with different patterns: *AtβCA4.1* is only expressed in leaves, while *AtβCA4.2* is

80   expressed in both roots and leaves (Aubry et al., 2014; DiMario et al., 2016). AS of CA
81   transcripts was also detected in *Neurachne munroi*, in which four *βCA* transcripts derived from
82   two genes were expressed by AS (Clayton et al., 2017). Although AS of CA transcripts is
83   common, the knowledge of AS in CA family genes is still quite limited, especially the
84   identification and molecular characterization of spliced variants of individual isoform.
85   In this study, we investigated AS in *βCA1* homologs between plant species, focusing on
86   βCA1 spliced variants from *Arabidopsis*. We examined the expression patterns and subcellular
87   localization of individual spliced variants of βCA1. We also analyzed the role of βCA1 in
88   response to abiotic stress and the protein-protein interaction (PPI) network. This study will
89   contribute to our understanding of βCA1.
90
91   ## Materials & Methods
92   **Plant material and stress treatment**
93   All *Arabidopsis* plants generated in this study are of the Columbia-0 (*Col-0*) ecotype. The T-
94   DNA mutants lines used in this study have been described previously: *βca1* (Salk_106570) and
95   *βca2* (CS303346) (Huang et al., 2017). Seeds were surface sterilized and grown on plates with
96   Murashige and Skoog (MS) plus 1% sucrose and 0.8% phyto agar at 22°C under a long-day
97   (16/8 h light/dark) photoperiod with a photon flux density of 180 μmol photons m-2 s-1. 7-d-old
98   seedlings from the plates were transferred to the soil and grown in the growth chamber. Abiotic
99   stress and phytohormone treatments were subjected to two-week-old seedlings. For Abscisic acid
100 (ABA) 10 μM of ABA was supplied on the solid agar medium with seedling grown vertically.
101 For salt stress, seedlings were vertically placed in plates with 150 mM NaCl for the indicated
102 period of time. For heat treatment, seedlings were transferred to a new plate and placed within a
103 bake oven set at 40°C for the indicated period of time. For analysis of the seed germination after
104 heat treatment, seeds were sown in the dark for 2 d at 4°C, and then the seeds were heated at
105 50°C for 3 hours or 6 hours. After this heat treatment, the seeds were grown at 22°C for 7 days,
106 and then the germination rate and seedling with green cotyledon were measured.
107 **RNA extraction and reverse transcription-polymerase chain reaction (RT-PCR)**
108   Total RNA was extracted and purified from the one-week-old *Arabidopsis* seedlings using
109 ReliaPrep™ RNA Miniprep System (Promega, USA). For cDNA synthesis, the first-strand
110 cDNA was synthesized from 2 μg total RNA with NovoScript®Plus All-in-one 1st Strand cDNA
111 Synthesis SuperMix (Novoprotein, China) according to the instructions. For semi-quantitative
112 RT-PCR analysis primers were designed manually after identification of unique sites in the
113 cDNA of four spliced variants of *βCA1* (*βCA1.1*, *βCA1.2*, *βCA1.3* and *βCA1.4*). Specific primers
114 for amplification of *βCA1.2* and *βCA1.3* were applied, respectively. The primers used were
115 mentioned in Table S1. The *ACTIN2* was used as the endogenous control.
116 **Identification and analysis of βCA1 in plants**
117   To identify ortholog(s) of βCA1 from *Arabidopsis* in the representative plant genome
118 (including *Oryza sativa*, *Glycine max*, *Populus trichocarpa*, *Brachypodium distachyon*, *Sorghum*
119 *bicolor* and *Physcomitrella patens*), the amino acid sequence of AtβCA1 was used as the guide

120 sequence to perform a BLASTp search of the database whole genome sequences in the
121 Phytozome database (https://phytozome.jgi.doe.gov). Coding sequences (CDS) and
122 corresponding genomic DNA (gDNA) sequences of the βCA1s from the selected plant species
123 were retrieved from the databases. The GSDS tool (GSDS v2.0, http://gsds.cbi.pku.edu.cn/) was
124 employed to analyze the exon-intron structures for plant *βCA1* genes.

125 Plasmids construction and plant transformation

126   All DNAs and cDNAs were amplified using Tks Gflex™ DNA Polymerase (Takara, Japan).
127 The CDS of four *βCA1* spliced variants (*βCA1.1*, *βCA1.2*, *βCA1.3* and *βCA1.4*) were amplified
128 and cloned into pDONOR221 using Gateway™ BP Clonase™ Enzyme Mix (Thermo Scientific,
129 USA). Each βCA1 spliced variant was further cloned into the pGWB405 plant expression vector
130 designed for the production of C-terminal GFP-tagged fusion proteins under the control of the
131 35S Cauliflower Mosaic Virus (35S CaMV) promoter using Gateway™ LR Clonase™ II
132 Enzyme mix (Thermo Scientific, USA). To generate the plasmid for the native expression of
133 βCA1, the genomic region including 1.5 kb upstream of ATG plus introns and exons were
134 cloned. For insertion of specific tag including HA or Myc, overlapping PCR was applied by
135 designed primers. These DNA fragments were cloned into pDONOR223 and pGWB4 vectors
136 with the same method used for βCA1 spliced variants cloning. The constructs were confirmed by
137 sequencing. All the plant expression constructs were introduced into the *Agrobacterium*
138 *tumefaciens* GV3101 strain. Plant transformation was performed with floral dipping method
139 (Clough and Bent, 1998). All the primers used for cloning were listed in Supplementary Table
140 S1.

141 Total protein extraction and western blotting

142   To analyze the GFP-tagged βCA1 spliced variants, seedlings of stable transgenic plants
143 expressing the corresponding GFP fusions were freeze grounded into powder and homogenized
144 in total protein buffer (20 mmo1/L Tris-HCl (pH 7.5), 150 mmol/L NaCl, 2 mmol/L EDTA) with
145 protease inhibitor cocktail in DMSO (Yeasen, China). Lysates were incubated on ice for 20 min
146 and clarified by centrifugation at 18407 g for 15 min at 4°C. For immunoblotting, samples were
147 separated on 12% SDS polyacrylamide gel and transferred to PVDF membranes. The
148 membranes were blocked with 5% (g/v) defatted milk in TBST buffer (10 mM Tris-HCl (pH7.4),
149 150 mM NaCl, 0.05% Tween 20) and probed with using appropriate antibodies including 1:6000
150 dilution α-GFP conjugated with HRP (MBL, USA), 1:2000 α-HA (MBL, USA) and 1:2000 α-
151 Myc (MBL, USA) overnight at 4°C. Then the samples were washed with TBST buffer for three
152 times and visualized by using the ECL (Amersham™, USA). Actin protein was used as the
153 internal control.

154 Subcellular localization analysis

155   The protoplasts from the four-week-old stable transgenic plants leaves with GFP-tagged
156 βCA1 were isolated following the previous used methods (Yoo et al., 2007). The subcellular
157 localization of the GFP fusion proteins from both the protoplast and seedlings was determined by
158 confocal laser scanning microscopy Leica SP8. The excitation wavelength was 488 nm for GFP
159 and the emission window was set at 500-520 nm.

160 **Co-immunoprecipitation (Co-IP) and MS analysis**
161   The samples of 0.5 g 7-d-old seedlings tissue were freeze grounded to powder and
162 homogenized in 2 ml IP buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 20%
163 glycerol, 0.2% CA-360, 1× protease inhibitor cocktail in DMSO). Lysates were clarified by
164 centrifugation at 14,000 rpm for 15 min at 4°C and were incubated with Magnetic GFP beads
165 (MBL, USA) for overnight at 4°C in a top to end rotator. After incubation, the beads were
166 washed five times with ice-cold washing buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM
167 $MgCl_2$, 20% glycerol and 0.2% CA-360) and then eluted by boiling in reducing SDS sample
168 buffer. Samples were separated by SDS-PAGE. For mass spectrometry analysis, the gel was
169 stained with Coomassie blue after SDS-PAGE separation, and the visible stained protein bands
170 were cut out for in-gel trypsin digestion, followed by tandem liquid chromatograph-mass
171 spectrometry (LC-MS/MS) at Shanghai Applied Protein Technology Co. Ltd.
172 **Proteins enrichment and protein-protein interaction analysis**
173   Annotation and GO enrichment analysis for proteins identified by MS was performed using
174 the Metascape tools, a free online platform for data analysis (Zhou et al., 2019). The protein
175 interaction network analysis was applied in STRING Version 11.0 (minimum required
176 interaction score with high confidence of 0.700) (Szklarczyk et al., 2019).
177 **Gene IDs**
178   Sequence data from this article can be found in The Arabidopsis In-formation Resource
179 (http://www.arabidopsis.org/) under the following gene IDs: βCA1 (At3g01500), βCA1.1
180 (At3g01500.1), βCA1.2 (At3g01500.2), βCA1.3 (At3g01500.3), βCA1.4 (At3g01500.4).
181

182 # Results
183 **AS analysis in *Arabidopsis* βCA1**
184   A genome-wide analysis of the βCA1 gene family was performed based on the complete
185 genome sequences. Using the Phytozome database and the βCA1 from *Arabidopsis* as the guide
186 sequence for a BLASTp search, we first retrieved available βCA1 sequences from representative
187 sequenced genomes, including three monocots (*Oryza sativa*, *Brachypodium distachyon*,
188 *Sorghum bicolor*), two dicots (*Glycine max*, *Populus trichocarpa*), and *Physcomitrella patens*.
189 Further analysis of genome annotation obtained from the Phytozome database revealed that all
190 *βCA1* genes from selected plant species possessed several transcripts resulting from AS (Fig.
191 S1). At least four spliced variants were found for the *βCA1* gene. These spliced variants seemed
192 to result from pre-mRNA AS, mostly via the use of either alternative 5′ or 3′ splice site. By
193 analyzing the latest *Arabidopsis thaliana* genome database, four spliced variants of *βCA1* were
194 found: *βCA1.1*, *βCA1.2*, *βCA1.3*, and *βCA1.4* (Fig. 1A). A previous study showed that the full-
195 length βCA1 (same as βCA1.2) contains the chloroplast transit peptide at the N-terminus (Hu et
196 al., 2015), whereas βCA1.1 and βCA1.3 coded for truncated proteins, lacking the chloroplast
197 transit peptide at the N-terminus and 11 amino acids at the C-terminus, respectively (Fig. 1B).
198 Moreover, the shortest βCA1.4 was truncated at both ends of the protein.

199   To investigate these spliced variants of βCA1, the expression pattern of each variant was
200   analyzed. It was not possible to obtain RT-PCR primers suitable for *βCA1.1*, so the
201   corresponding transcripts of *βCA1.1* and *βCA1.2* were analyzed together. The same strategy was
202   applied to *βCA1.3* and *βCA1.4*. For instance, primers F1 and R1 were used for the total
203   transcripts of *βCA1.1* and *βCA1.2*, whereas F2 and R1 were used for the analysis of *βCA1.2* (Fig.
204   1A). We investigated the transcript expression in four tissues from flowering plants: leaves,
205   inflorescences, flowers, and fresh fruits (Fig. 1C). *βCA1.1* showed the highest global expression
206   levels, with more transcripts from leaves and inflorescences. On the contrary, the expression of
207   *βCA1.2* was higher in flowers and fruits than in leaves and inflorescences. The expression levels
208   of *βCA1.3* were weaker than the expression levels of *βCA1.4*, and *βCA1.3* transcripts were rarely
209   detected in flowers.
210   To further study the accumulation of each βCA1 spliced variant *in vivo*, two plasmids were
211   generated. The first (*βCA1:βCA1-HA*) had an HA tag (YPYDVPDYA) directly attached to the
212   last exon before the stop codon. The second (*βCA1:βCA1-Myc*) had a Myc tag (QILFRDEFLL)
213   linked to the end of the second-to-last exon, which was designed for expression analysis of four
214   individual spliced variants of βCA1 *in vivo* (Fig. 1D). The spliced variant of βCA1 was detected
215   in the protein extracts from stable transgenic seedlings. The total translation accumulation of
216   βCA1.1 and βCA1.2 were comparable, and protein levels of βCA1.4 were much higher than
217   βCA1.3. However, it is not possible to compare protein levels of all four spliced variants of
218   βCA1 due to the different genetic backgrounds. Taken together, these results suggest that βCA1
219   is regulated by AS.
220   **Subcellular localization of spliced variants for βCA1**
221   As the chloroplast targeting peptide was absent in βCA1.1 and βCA1.4, we speculated that
222   the different spliced variants might have different subcellular locations. We then investigated the
223   subcellular localization of each βCA1 spliced variant tagged with GFP in stable transgenic
224   plants. In one-week-old seedlings, βCA1.1 and βCA1.2 showed a similar distribution with a
225   visible GFP fluorescence signal in leaf chloroplasts localized on stomata (Fig. 2A and B).
226   βCA1.3 was also localized in leaf chloroplasts, however, it seemed to accumulate in the
227   chloroplast envelope and was also observed within the protoplast (Fig. 2B). Surprising, βCA1.4
228   presented a very different pattern. Strong GFP fluorescence signals in the cytoplasm were
229   observed in transgenic plants with βCA1.4-GFP, whereas no visible signal was detected in
230   chloroplasts either from stomata or leaf protoplasts.
231   In the root tips, a strong fluorescent signal was detected only for βCA1.4, whereas βCA1.1,
232   βCA1.2, and βCA1.3 were almost undetectable (Fig. 2C). Instead, the latter three variants
233   seemed to accumulate only in small, punctuated structures in the cytoplasm of the root mature
234   zone (Fig. 2D). Subsequent analysis showed that the signals from these small punctuated
235   structures were dynamic (Fig. S2). Immunoblotting analysis confirmed the accumulation of the
236   four spliced variants in the corresponding tissues and organs (Fig. S3).
237   **Stress influences the transcript abundance of βCA1 spliced variants**

238   The CA1 family is induced under multiple abiotic stresses and phytohormones; therefore, we
239   investigated the effect of some abiotic stresses (salt, heat) and ABA on the expression pattern of
240   *βCA1* spliced variants. The results showed that none of the stresses affected the two βCA1
241   spliced variant pools (*βCA1.1+βCA1.2*, *βCA1.3+βCA1.4*), whereas salt and heat stresses
242   significantly repressed the transcriptional expression of *βCA1.2* and *βCA1.3* (Fig. 3A). This
243   indicated that salt and heat stresses significantly enhanced the induction of *βCA1.1* and *βCA1.4*.
244   Previous studies had reported that a mature rice CA protein could confer the heat stress
245   tolerance in *E. coli* recombinants and that CA1 proteins from poplar were induced under heat
246   stress (Tianpei et al., 2015; Shi et al., 2017). Thus, we characterized the response of a T-DNA
247   insertion mutant *βca1* to heat stress (Fig. S4). For the heat stress germination assay, the
248   vernalized Col-0 and *βca1* mutant seeds were pretreated at 50°C for 3 or 6 hours before moving
249   them to normal growth conditions. The results showed that the germination rate of Col-0 was
250   almost unaffected after heat treatment for 3 hours, while it decreased to 85% in the *βca1* mutant
251   (Fig. 3B). Under 6 hours heat stress treatment, a more significant reduction in the germination
252   rate of *βca1* was observed, whereas 80% of Col-0 seeds still germinated. This suggests that the
253   *βca1* mutant is heat-sensitive during seed germination.
254   **Protein-protein inertaction (PPI) network analysis for *Arabidopsis* βCA1**
255   Knowledge of all direct and indirect interactions between proteins will provide new insights
256   into the complex molecular mechanisms inside a cell. To better understand the roles of βCA1 in
257   Arabidopsis, we used co-IP to identify proteins that interact with the most redundant spliced
258   variant βCA1.4. The results showed with high confidence that 109 proteins interact with βCA1.4
259   directly or indirectly (Table S2).
260   Among the interacting proteins of βCA1.4, two other βCA family proteins, βCA2 and βCA4,
261   were identified. This interaction was confirmed *in vivo* by bimolecular fluorescence
262   complementation (BiFC) (Huang et al., 2017). To gain further insight into their protein
263   functions, the 109 proteins interacting with βCA1.4 were analyzed using Gene Ontology (GO)
264   classifications using Metascape. The analysis revealed that 19 clusters were significantly
265   enriched (P value cutoff of 0.01). These proteins played various roles, most of which were
266   related to the functions of βCA1, including photosynthesis and stress response (Fig. 4A).
267   Surprisingly, 29 of these proteins were predicted in response to cadmium ions, which indicated
268   that *Arabidopsis* βCA1 may also function as a cadmium enzyme. In addition, GO analysis
269   showed that 17 proteins, including βCA1, were involved in the response to temperature, which
270   suggests a potential molecular basis for βCA1 in the heat stress response (Fig. 4B).
271   Additional PPI networks for βCA1.4 associated proteins were constructed using Search Tool
272   for the Retrieval of Interacting Genes/Proteins (STRING 11.0). The full set of 109 proteins
273   produced an interactome map composed of 108 nodes (proteins) and 516 edges, with an average
274   node degree of 9.56 (avg. local clustering coefficient = 0.7) (Fig. 5). The various PPI sub-
275   networks are associated with photosystem (green), cellular response to stress (red), and ribosome
276   pathway (cyan), which was consistent with the results from Metascape analysis. Further
277   subcellular localization analysis by STRING showed that most of these proteins are localized in

278   chloroplasts and plastids, where βCA1 proteins are also common. Moreover, several proteins,
279   including βCA1, were predicted to be present in stromules, implying that they may function in
280   morphological maintenance and stromule regulation of non-mesophyll plastids (Fig. S5).
281

Discussion

283   The CA family plays various roles in many biological processes. Both mammalian and plant
284   cells mainly possess two (α and β) or three (α, β, and γ) families with multiple isoforms. In each
285   family, AS within individual isoforms further generates diversity, which significantly contributes
286   to the structural and functional diversification of CA proteins. However, there is still a lack of
287   information regarding AS of CA in plants.
288   The β family CAs are the most abundant with high expression intensity in plant leaves.
289   Here, βCA1 from the model plant *Arabidopsis* was selected for a comprehensive study. We
290   found that the *βCA1* gene contained at least four spliced variants from all the representative
291   plants, indicating a conserved evolutionary pattern for AS in *βCA* genes (Fig. S1). Interestingly,
292   different expression patterns of spliced variants in *CA* have been observed. In *Arabidopsis*, most
293   of the spliced variants of *βCA2* were detected in leaves using qRT-PCR and the spliced forms of
294   *βCA4* were found to be differentially expressed in tissues (DiMario et al., 2016; Wang et al.,
295   2014). Alternatively spliced *βCA* transcripts were also detected in the leaves of *N. munroi*
296   (Clayton et al., 2017). Here, we found different patterns of transcriptional expression in the *βCA1*
297   spliced variants of *Arabidopsis* (Fig. 1C). *βCA1.1* was found to be the dominant expressed
298   spliced variant. In contrast, *βCA1.3* was weakly expressed and was almost undetectable in
299   flowers. We further investigated whether these spliced variants would undergo accurate
300   translation into proteins. Immunoblotting assays showed the presence of translated proteins for
301   four spliced variants, indicating they may function *in vivo* (Fig. 1D). The previous study had
302   suggested cleavage sites of the chloroplast targeting sequence in *Arabidopsis* βCA1 (Hu et al.,
303   2015), indicating that the mature βCA1.1 and βCA1.2 proteins and the mature βCA1.3 and
304   βCA1.4 proteins are the same size (due to identical peptide content). However, they ran at two
305   different protein sizes on their respective gels (Fig. 1D). We suspected that two bands detected
306   from the *βCA1:βCA1-HA* (or *βCA1:βCA1-Myc*) seedlings might be due to posttranslational
307   modification of βCA1 variants, as βCA1 has been reported to be phosphorylated by the leucine-
308   rich repeat receptor-like kinase EXCESS MICROSPOROCYTES1 (EMS1) (Huang et al., 2017).
309   However, the results from immunoblotting analysis were the same with the extracts from
310   *βCA1:βCA1-HA* (or *βCA1:βCA1-Myc*) seedlings in the absence or presence of alkaline
311   phosphatase, suggesting two different protein sizes not result from phosphorylation modification
312   (data not shown). On the other hand, different expression pattern of βCA1 variants were
313   observed. The subcellular localization of them could be same if the mature βCA1.3 and βCA1.4
314   proteins are the same size. However, our results showed that localization of βCA1.3 is quite
315   different from that of βCA1.4 (Fig. 3). In addition, βCA1.1 was almost undetectable in root,
316   while strong signals of βCA1.1 were observed in leaf. On the contrary, βCA1.2 was weakly
317   expressed in leaf but much strong in root compared with that of βCA1.1 or βCA1.3. Taking

318 together, these results indicated that AS could result in different variants of functional proteins of
319 βCA1 in plant.
320   The correct inter- and intracellular placement of CAs is essential for efficient physiological
321 function. A previous study showed that the long form of βCA4, βCA4.1, localizes to the plasma
322 membrane while the short form, βCA4.2, is cytosolic (DiMario et al., 2016; Fabre et al., 2007).
323 This suggests that different spliced variants may show remarkable location diversity, which may
324 enable them to fulfill specific roles. Thus, the subcellular localization of individual spliced
325 variants of βCA1 was investigated. Amino acid sequence alignment analysis showed that
326 βCA1.1 and βCA1.4 were short of the chloroplast transit peptide (Fig. 1B). Consistent with
327 previous report (Hu et al., 2015), in our stable transgenic plants βCA1.1 was observed specially
328 localized in chloroplast (Fig. 2A, 2B) and the shortest spliced variant βCA1.4, which is 11 amino
329 acids shorter at the C-terminal in compared to βCA1.1, was mainly targeted in the cytoplasm.
330 This implies the contribution of this short peptide at C-terminus for chloroplast targeting. Eleven
331 amino acids present in the N-terminal region of the β-carbonic anhydrase *NmuCA1a* from
332 *Neurachne* are important for chloroplast targeting and have been suggested to be functional
333 chloroplast transit peptides (Clayton et al., 2017). However, the short peptide from *Arabidopsis*
334 (KDVATILHWKL) was different from that in *Neurachne* (ASLGTPAPSSS) and was possibly
335 involved in chloroplast transition. In addition, the short peptide from *Arabidopsis* is only found
336 in CA genes from dicots, while the other one from *Neurachne* is known only from monocots,
337 indicating evolutionary differences between these plants (Fig. S6). This remains to be
338 investigated further. At the organ/tissue levels, only strong expression of βCA1.4 was only
339 observed in mature leaves and root tips (Fig. 2A, 2B). In contrast, localization of the other three
340 spliced variants in the root tips was quite different from that of βCA1.4. As well as being present
341 in chloroplasts, plasma membranes, and cytoplasm, CA proteins such as βCA6 from *Arabidopsis*
342 and βCA2b from *Neurachne* have also been observed being imported into mitochondria (Clayton
343 et al., 2017; Fabre et al., 2007; Jiang et al., 2014). Moreover, in the red alga *Gracilariopsis*
344 *chorda*, some CAs showed multiple subcellular locations such as the ER, mitochondria, vacuole,
345 and cytosol (Razzak et al., 2019). As the subcellular location of homologous or orthologous
346 genes may be conserved, it would be interesting to further investigate the exact subcellular
347 localization of these βCA1 spliced variants in the root tip. To finding alternatively spliced
348 variants are resulting in different βCA1 proteins that have different subcellular locations is
349 important because: 1) CA proteins may be more redundant within subcellular locations (i.e. there
350 may be more CAs in a subcellular location that can compensate for loss of CA activity due to
351 mutation) and 2) this may bolster the potential physiological roles of CA.
352   The CA gene family has been linked to the stress response of plants. A previous study
353 showed that the transcription level and enzyme activity of the CA family could be induced by
354 various biotic and abiotic stresses, including pathogens, insect herbivores, salinity, and severe
355 temperature (Yu et al., 2007; Wang et al., 2009; Collins et al., 2010; Kravchik et al., 2013; Chen
356 et al., 2020). *Arabidopsis* βCA1 has been reported to function together with βCA4 in disease
357 resistance, and was induced under cold stress at the translational level (Jing et al., 2019; Zhou et

358   al., 2020). These data suggest that the *Arabidopsis* βCA1 plays an important role in the response
359   of plants to environmental stress conditions. However, it is still unknown whether or how these
360   spliced variants of *Arabidopsis* βCA1 respond and contribute to the protection of organisms
361   under different stress conditions. In the present study, *βCA1.1* and *βCA1.4* were induced by salt
362   and heat stresses, whereas *βCA1.2* and *βCA1.3* were significantly repressed (Fig. 3A). This
363   implies that βCA1 is possibly in response to salt and heat stresses. However, no significant
364   difference was observed between wild-type and *βca1* mutants under salt stress (data not shown).
365   On the other hand, we found that loss-of-function *βca1* mutants were much more sensitive to
366   heat stress during seed germination (Fig. 3B). We also analyzed the stress response of *βca2*, the
367   closed homolog of βCA1. However, no obvious phenotype was found in *βca2* under same stress
368   conditions (data not shown). This indicates functional diversity in the CA family. Additionally,
369   the contribution to heat stress tolerance of CA has been found in other plant species, such as rice
370   and *Pyropia haitanensis* (Yu et al., 2007; Tianpei et al., 2015; Shi et al., 2017).
371   CA associates other proteins by transmitting signals to activate downstream events. Here, we
372   found that the most redundant spliced variant βCA1.4 immunoprecipitated at least 100 proteins
373   in seedlings. Several proteins respond to heat stress, including heat shock proteins, the
374   antioxidant enzyme catalase 3, and peroxidase 45. Heat stress has been found to increase
375   oxidative damage in cells, whereas antioxidant enzymes can lead to protection. This implies that
376   the βCA1 protein could protect cells from damage by reducing peroxides to some extent under
377   high-temperature stress. On the other hand, several studies have shown that coordination
378   between CA and aquaporins could facilitate CO_2 transport (Terashima and Ono, 2002; Uehlein et
379   al., 2003; Uehlein et al., 2008; Yaneff et al., 2014; Zhao et al., 2017). For instance, aquaporin
380   PIP2;1 in *Arabidopsis* interacts with βCA4 to facilitate CO_2 permeability across the plasma
381   membrane (Wang et al., 2016). Aquaporin NtAQP1 from tobacco plants has also been identified
382   as a CO_2 pore not only in the plasma membrane but also in the inner chloroplast envelope
383   membranes (Uehlein et al., 2003). In this study aquaporin PIP1;2 was identified as the iterator of
384   βCA1.4. Comprehensive interaction between CAs and aquaporins may be present in plants and
385   will be interesting to investigate in the future
386
Conclusions
388   The βCA gene family has been the most intensely studied in plants. Here, we found that
389   *Arabidopsis* βCA1 was involved in the heat stress response and AS of *βCA1* genes from different
390   plant species. Further investigation showed that four spliced variants of *Arabidopsis* βCA1 were
391   differentially expressed in tissues and in response to stresses. In particular, βCA1.4 showed
392   different localization compared to the other three variants, probably due to loss of a chloroplast
393   transit peptide at the N-terminal and a short peptide at the C-terminal. Interestingly, a short
394   peptide at the C-terminal was specifically appeared in dicots. A more detailed understanding of
395   βCA1 spliced variants will enable us to build a basic working model of their function.
396   Meanwhile, several proteins that interact with βCA1 were identified, which established a basis

397 for further research on the CA family and PPI network in order to understand their physiological
398 roles.
399

Acknowledgements

401   We thank Professor Dazhong Zhao (University of Wisconsin-Milwaukee, Wisconsin,
402 USA) for providing the mutants *βca1* (SALK_106570) and *βca2* (CS303346).

403

References

405 1. Aubry, S.; Smith-Unna, R.D.; Boursnell, C.M.; Kopriva, S.; Hibberd, J.M. Transcript
406   residency on ribosomes reveals a key role for the *Arabidopsis thaliana* bundle sheath in
407   sulfur and glucosinolate metabolism. *Plant J* **2014**, 78, 659–673.
408 2. Barathova, M.; Takacova, M.; Holotnakova, T.; Gibadulinova, A.; Ohradanova, A.;
409   Zatovicova, M.; Hulikova, A.; Kopacek, J.; Parkkila, S.; Supuran, C.T.; Pastorekova, S.;
410   Pastorek, J. Alternative splicing variant of the hypoxia marker carbonic anhydrase IX
411   expressed independently of hypoxia and tumour phenotype. *Br J Cancer* **2008**, 98(1), 129-
412   136.
413 3. Chen, Z.; Wang, W.; Dong, X.; Pu, X.; Gao, B.; Liu, L. Functional redundancy and
414   divergence of β-carbonic anhydrases in *Physcomitrella patens*. *Planta* **2020**, 252(2):20.
415 4. Clayton, H.; Saladié, M.; Rolland, V.; Sharwood, R.; Macfarlane, T.; Ludwig, M. Loss of
416   the Chloroplast Transit Peptide from an Ancestral C3 Carbonic Anhydrase Is Associated
417   with C4 Evolution in the Grass Genus *Neurachne*. *Plant Physiol* **2017**, 173, 1648-1658.
418 5. Clough, S.J.; Bent, A.F. Floral dip: a simplified method for Agrobacterium-mediated
419   transformation of *Arabidopsis thaliana*. *Plant J* **1998**, 16(6), 735-743.
420 6. Collins, R.M.; Afzal, M.; Ward, D.A.; Prescott, M.C.; Sait, S.M.; Rees, H.H.; Tomsett, A.B.
421   Differential proteomic analysis of *Arabidopsis thaliana* genotypes exhibiting resistance or
422   susceptibility to the insect herbivore, *Plutella xylostella*. *PLoS One* **2010**, 5, e10103.
423 7. DiMario, R.J.; Clayton, H.; Mukherjee, A.; Ludwig, M.; Moroney, J.V. Plant Carbonic
424   Anhydrases: Structures, Locations, Evolution, and Physiological Roles. *Mol Plant* **2017**,
425   10(1), 30-46.
426 8. DiMario, R.J.; Quebedeaux, J.C.; Longstreth, D.J.; Dassanayake, M.; Hartman, M.M.;
427   Moroney, J.V. The cytoplasmic carbonic anhydrases bCA2 and bCA4 are required for
428   optimal plant growth at low CO_2. *Plant Physiol* **2016**, 171, 280-293.
429 9. Engineer, C.B.; Ghassemian, M.; Anderson, J.C.; Peck, S.C.; Hu, H.; Schroeder, J.I.
430   Carbonic anhydrases, EPF2 and a novel protease mediate CO_2 control of stomatal
431   development. *Nature* **2014**, 513, 246-250.
432 10. Espie, G.S.; Kimber, M.S. Carboxysomes: cyanobacterial Rubisco comes in small packages.
433   *Photosynth Res* **2011**, 109, 7-20.
434 11. Fabre, N.; Reiter, I.M.; Becuwe-Linka, N.; Genty, B.; Rumeau, D. Characterization and
435   expression analysis of genes encoding alpha and beta carbonic anhydrases in *Arabidopsis*.
436   *Plant Cell Environ* **2007**, 30, 617-629.
437 12. Haapasalo, J.; Hilvo, M.; Nordfors, K.; Haapasalo, H.; Parkkila, S.; Hyrskyluoto, A.;
438   Rantala, I.; Waheed, A.; Sly, W.S.; Pastorekova, S.; Pastorek, J.; Parkkila, A.K.
439   Identification of an alternatively spliced isoform of carbonic anhydrase XII in diffusely
440   infiltrating astrocytic gliomas. *Neuro Oncol* **2008**, 10(2), 131-138.
441 13. Hanson, M.R.; Hines, K.M. Stromules: Probing Formation and Function. *Plant Physiol*
442   **2018**, 176(1), 128-137.

443   14. Hewett-Emmett, D.; Tashian, R. Functional diversity, conservation, and convergence in the
444   evolution of the α, β, and γ-carbonic anhydrase gene families. *Mol Phylogenet Evol* **1996**,
445   65, 50-77.
446   15. Hoang, C.V.; Chapman, K.D. Biochemical and molecular inhibition of plastidial carbonic
447   anhydrase reduces the incorporation of acetate into lipids in cotton embryos and tobacco cell
448   suspensions and leaves. *Plant Physiol* **2002**, 128, 1417-1427.
449   16. Hoang, C.V.; Chapman, K.D. Regulation of carbonic anhydrase gene expression in
450   cotyledons of cotton (*Gossypium hirsutum L.*) seedlings during post-germinative growth.
451   *Plant Mol Biol* **2002**, 49, 449-458.
452   17. Hu, H.; Boisson-Dernier, A.; Israelsson-Nordstro¨ m, M.; Bo¨ hmer, M.; Xue, S.; Ries, A.;
453   Godoski, J.; Kuhn, J.M.; Schroeder, J.I. Carbonic anhydrases are upstream regulators of
454   CO_2-controlled stomatal movements in guard cells. *Nat Cell Biol* **2010**, 12(1), 87-93.
455   18. Hu, H.H.; Rappel, W.J.; Occhipinti, R.; Ries, A.; Bohmer, M.; You, L.; Xiao, C.L.;
456   Engineer, C.B.; Boron, W.F.; Schroeder, J.I. Distinct cellular locations of carbonic
457   anhydrases mediate carbon dioxide control of stomatal movements. Plant Physiol **2015**, 169,
458   1168-1178.
459   19. Huang, J.; Li, Z.Y.; Biener, G.; Xiong, E.H.; Malik, S.; Eaton, N.; Zhao, C.Z.; Raicu, V.;
460   Kong, H.Z.; Zhao, D.Z. Carbonic Anhydrases Function in Anther Cell Differentiation
461   Downstream of the Receptor-Like Kinase EMS1. *Plant Cell* **2017**, 29(6), 1335-1356.
462   20. Jacobson, B.S.; Fong, F.; Heath, R.L. Carbonic anhydrase of spinach: studies on its location,
463   inhibition, and physiological function. *Plant Physiol* **1975**, 55(3), 468-474.
464   21. Jiang, C.Y.; Tholen, D.; Xu, J.M.; Xin, C.P.; Zhang, H.; Zhu, X.G.; Zhao, Y.X. Increased
465   expression of mitochondria-localized carbonic anhydrase activity resulted in an increased
466   biomass accumulation in *Arabidopsis thaliana*. *J Plant Biol* **2014**, 57, 366-374.
467   22. Jing Y.; Shi L.; Li X.; Zheng H.; Gao J.; Wang M.; He L.; Zhang W. OXS2 is Required for
468   Salt Tolerance Mainly through Associating with Salt Inducible Genes, CA1 and Araport11,
469   in *Arabidopsis*. *Sci Rep* **2019**, 9(1), 20341.
470   23. Jung, H.W.; Lim, C.W.; Lee, S.C.; Choi, H.W.; Hwang, C.H.; Hwang, B.K. Distinct roles of
471   the pepper hypersensitive induced reaction protein gene CaHIR1 in disease and osmotic
472   stress, as determined by comparative transcriptome and proteome analyses. *Planta* **2008**,
473   227, 409-425.
474   24. Kravchik, M.; Bernstein, N. Effects of salinity on the transcriptome of growing maize leaf
475   cells point at cell-age specificity in the involvement of the antioxidative response in cell
476   growth restriction. *BMC Genomics* **2013**, 14, 24.
477   25. Laloum, T.; Martín, G.; Duque, P. Alternative Splicing Control of Abiotic Stress Responses.
478   *Trends Plant Sci* **2018**, 23, 140–150.
479   26. Liljas, A.; Laurberg, M. A wheel invented three times. The molecular structures of the three
480   carbonic anhydrases. *EMBO Reports* **2000**, 1(1), 16-17.
481   27. Mastrangelo, A.M.; Marone, D.; Laido, G.; De Leonardis, A.M.; De Vita, P. Alternative
482   splicing: Enhancing ability to cope with stress via transcriptome plasticity. *Plant Sci* **2012**,
483   185, 40-49.
484   28. Moroney, J.V.; Bartlett, S.G.; Samuelsson, G. Carbonic anhydrases in plants and algae.
485   *Plant Cell Environ* **2001**, 24, 141-153.
486   29. Naftelberg, S.; Schor, I.E.; Ast, G.; Kornblihtt, A.R. Regulation of alternative splicing
487   through coupling with transcription and chromatin structure. *Annu Rev Biochem* **2015**, 84,
488   165-198.
489   30. Oh, D.H.; Hong, H.; Lee, S.Y.; Yun, D.J.; Bohnert, H.J.; Dassanayake, M. Genome
490   structures and transcriptomes signify niche adaptation for the multiple-ion-tolerant
491   extremophyte Schrenkiella parvula. *Plant Physiol* **2014**, 164, 2123-2138.

31. Price, G.D.; von Caemmerer, S.; Evans, J.R.; Yu, J.W.; Lloyd, J.; Oja, V.; Kell, P.; Harrison, K.; Gallagher, A.; Badger, M.R. Specific reduction of chloroplast carbonic anhydrase activity by antisense RNA in transgenic tobacco plants has a minor effect on photosynthesis. *Planta* **1994**, 193, 331-340.

32. Randall, D.J., and Val A.L. The Role of Carbonic Anhydrase in Aquatic Gas Exchange. In: Heisler N. (eds) Mechanisms of Systemic Regulation. *Advan Compara Environ Physiol.* **1995**, 21. Springer, Berlin, Heidelberg.

33. Razzak, M.A.; Lee, J.; Lee, D.W.; Kim, J.H.; Yoon, H.S.; Hwang, I. Expression of seven carbonic anhydrases in red alga *Gracilariopsis chorda* and their subcellular localization in a heterologous system, *Arabidopsis thaliana*. *Plant Cell Rep* **2019**, 38(2), 147-159.

34. Restrepo, S.; Myers, K.L.; del Pozo, O.; Martin, G.B.; Hart, A.L.; Buell, C.R.; Fry, W.E.; Smart, C.D. Gene profiling of a compatible interaction between *Phytophthora infestans* and *Solanum tuberosum* suggests a role for carbonic anhydrase. *Mol Plant Microbe Interact* **2005**, 18, 913-922.

35. Rudenko, N.N.; Vetoshkina, D.V.; Fedorchuk, T.P.; Ivanov, B.N. Effect of Light Intensity under Different Photoperiods on Expression Level of Carbonic Anhydrase Genes of the α- and β-Families in Arabidopsis thaliana Leaves. *Biochemistry (Mosc)* **2017**, 82(9):1025-1035.

36. Shi, J.; Chen, Y.; Xu, Y.; Ji, D.; Chen, C.; Xie, C. Differential Proteomic Analysis by iTRAQ Reveals the Mechanism of *Pyropia haitanensis* Responding to High Temperature Stress. *Sci Rep* **2017**, 7, 44734.

37. Slaymaker, D.H.; Navarre, D.A.; Clark, D.; del Pozo, O.; Martin, G.B.; Klessig, D.F. The tobacco salicylic acid-binding protein 3 (SABP3) is the chloroplast carbonic anhydrase, which exhibits antioxidant activity and plays a role in the hypersensitive defense response. *Proc Natl Acad Sci.USA* **2002**, 99, 11640-11645.

38. Staiger, D.; Brown, J.W. Alternative splicing at the intersection of biological timing, development, and stress responses. *Plant Cell* **2013**, 25, 3640-3656.

39. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; Jensen, L.J.; Mering, C.V. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* **2019**, 47, D607-D613.

40. Terashima, I.; Ono, K. Effects of $HgCl_2$ on CO_2 dependence of leaf photosynthesis: Evidence indicating involvement of aquaporins in CO_2 diffusion across the plasma membrane. *Plant Cell Physiol* **2002**, 43, 70-78.

41. Tianpei, X.; Mao, Z.; Zhu, Y.; Li, S. Expression of Rice Mature Carbonic Anhydrase Gene Increase E. coli Tolerance to Heat Stress. *Appl Biochem Biotechnol* **2015**, 176(2), 625-635.

42. Uehlein, N.; Lovisolo, C.; Siefritz, F.; Kaldenhoff, R. The tobacco aquaporin NtAQP1 is a membrane CO_2 pore with physiological functions. *Nature* **2003**, 425(6959), 734-737.

43. Uehlein, N.; Otto, B.; Hanson, D.T.; Fischer, M.; McDowell, N.; Kaldenhoff R. Function of Nicotiana tabacum aquaporins as chloroplast gas pores challenges the concept of membrane CO_2 permeability. *Plant Cell* **2008**, 20(3), 648–657.

44. Wang, C.; Hu, H.; Qin, X.; Zeise, B.; Xu, D.; Rappel, W.J.; Boron, W.F.; Schroeder, J.I. Reconstitution of CO_2 regulation of SLAC1 anion channel and function of CO_2-permeable PIP2;1 aquaporin as CARBONIC ANHYDRASE4 interactor. *Plant Cell* **2016**, 28, 568-582.

45. Wang, M.; Zhang, Q.; Liu, F.C.; Xie, W.F.; Wang, G.D.; Wang, J.; Gao, Q.H.; Duan, K. Family-wide expression characterization of *Arabidopsis* beta-carbonic anhydrase genes using qRT-PCR and Promoter::GUS fusions. *Biochimie* **2014**, 97, 219-227.

46. Wang, Y.Q.; Feechan, A.; Yun, B.W.; Shafiei, R.; Hofmann, A.; Taylor, P.; Xue, P.; Yang, F.Q.; Xie, Z.S.; Pallas, J.A.; Chu, C.C.; Loake, G.J. S-nitrosylation of AtSABP3 antagonizes the expression of plant immunity. *J Biol Chem* **2009**, 284, 2131-2137.

541 47. Werdan, K.; Heldt, H.W. Accumulation of bicarbonate in intact chloroplasts following a pH
542   gradient. *Biochim Biophys Acta* **1972**, 283, 430-441.
543 48. Yaneff, A.; Sigaut, L.; Marquez, M.; Alleva, K.; Pietrasanta, L.I.; Amodeo, G.
544   Heteromerization of PIP aquaporins affects their intrinsic permeability. *Proc Natl Acad Sci*
545   *U S A* **2014**, 111, 231-236.
546 49. Yoo S.D.; Cho Y.H.; Sheen J. *Arabidopsis* mesophyll protoplasts: a versatile cell system for
547   transient gene expression analysis. *Nat Protoc* **2007**, 2(7), 1565-1572.
548 50. Yu, S.; Zhang, X.; Guan, Q.; Takano, T.; Liu, S. Expression of a carbonic anhydrase gene is
549   induced by environmental stresses in rice (*Oryza sativa L.*). *Biotechnol Lett* **2007**, 29, 89-94.
550 51. Zhao, M.; Tan, H.T.; Scharwies, J.; Levin, K.; Evans, J.R.; Tyerman, S.D. Association
551   between water and carbon dioxide transport in leaf plasma membranes: assessing the role of
552   aquaporins. *Plant Cell Environ* **2017**, 40(6), 789-801.
553 52. Zhou, Y.; Vroegop-Vos, I.A.; Van Dijken, A.J.H.; Van der Does, D.; Zipfel, C.; Pieterse,
554   C.M.J.; Van Wees S.C.M. Carbonic anhydrases CA1 and CA4 function in atmospheric CO_2-
555   modulated disease resistance. Planta **2020**, 251(4), 75.
556 53. Zhou, Y.; Zhou, B.; Pache, L.; Chang, M.; Khodabakhshi, A.H.; Tanaseichuk, O.; Benner,
557   C.; Chanda, S.K. Metascape provides a biologist-oriented resource for the analysis of
558   systems-level datasets. *Nat Commun* **2019**, 10(1), 1523.

Figure 1

Sequence structure features of the *Arabidopsis* βCA1.

(**A**) The exon-intron structures of βCA1 spliced variants were generated by comparing the coding sequences and the corresponding genomic sequences using the GSDS website (http://gsds.gao-lab.org/). The black boxes represent exons, solid lines represent introns, and bright blue boxes represent untranslated regions (UTRs). Primers used for RT-PCR were indicated by arrows. (**B**) Amino acids alignment analysis for βCA1 spliced variants. The chloroplast transit peptide was in green and the 11 amino acids at the C-terminal were in red. (**C**) Reverse transcription-polymerase chain reaction (RT-PCR) validation of expression profiling of representative βCA1 spliced variants in different tissues. The RT-PCR products were obtained as: F1+R1 primers for βCA1.1 + βCA1.2, F2+R1 primers for βCA1.2, F1+R2 primers for βCA1.3 + βCA1.4, and F2+R2 primers for βCA1.3. (**D**) Expression profiling of βCA1 spliced variants. The protein accumulation of βCA1.1 and βCA1.2 was analyzed by western blotting with an anti-HA antibody from the βCA1:βCA1-HA seedlings. And the proteins of βCA1.3 and βCA1.4 were detected with an anti-Myc antibody from the βCA1:βCA1-Myc seedlings. The wild type plants (CK) was used as control.
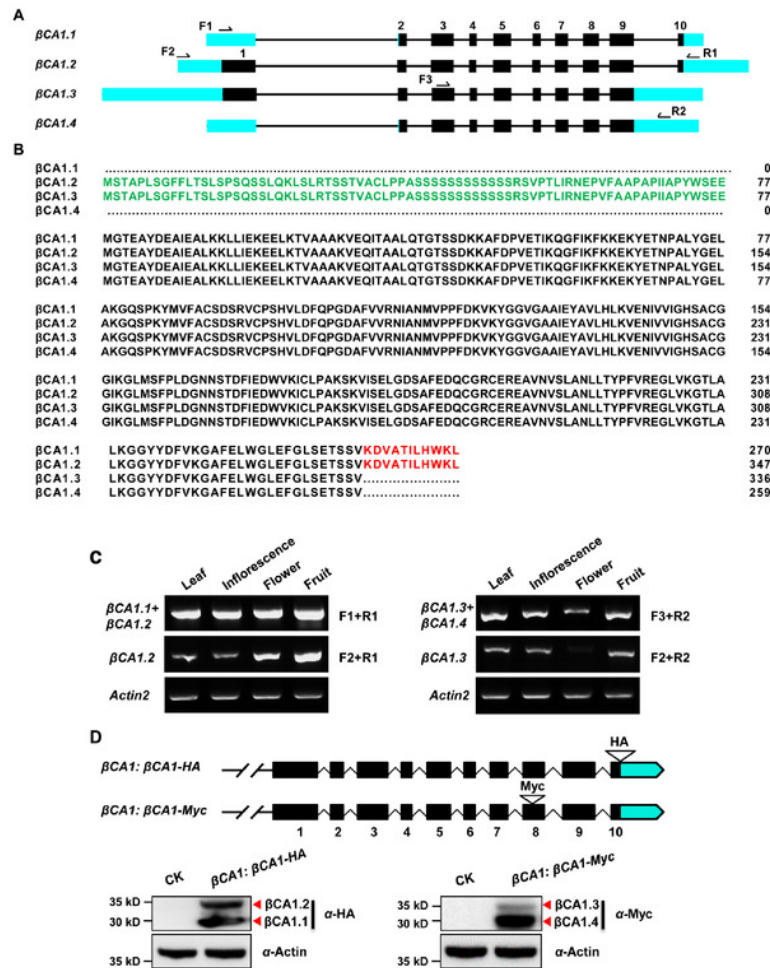
Figure 2

Subcellular localization of βCA1 spliced variants in transgenic plants.

A-D: Representative confocal microscopy images were shown in leaves (A), protoplasts (B), root tips (C), and the mature zones of roots (D). The samples of leaves, root tips were from one week old seedlings, and the protoplasts were isolated from the leaves of four-week old seedlings.
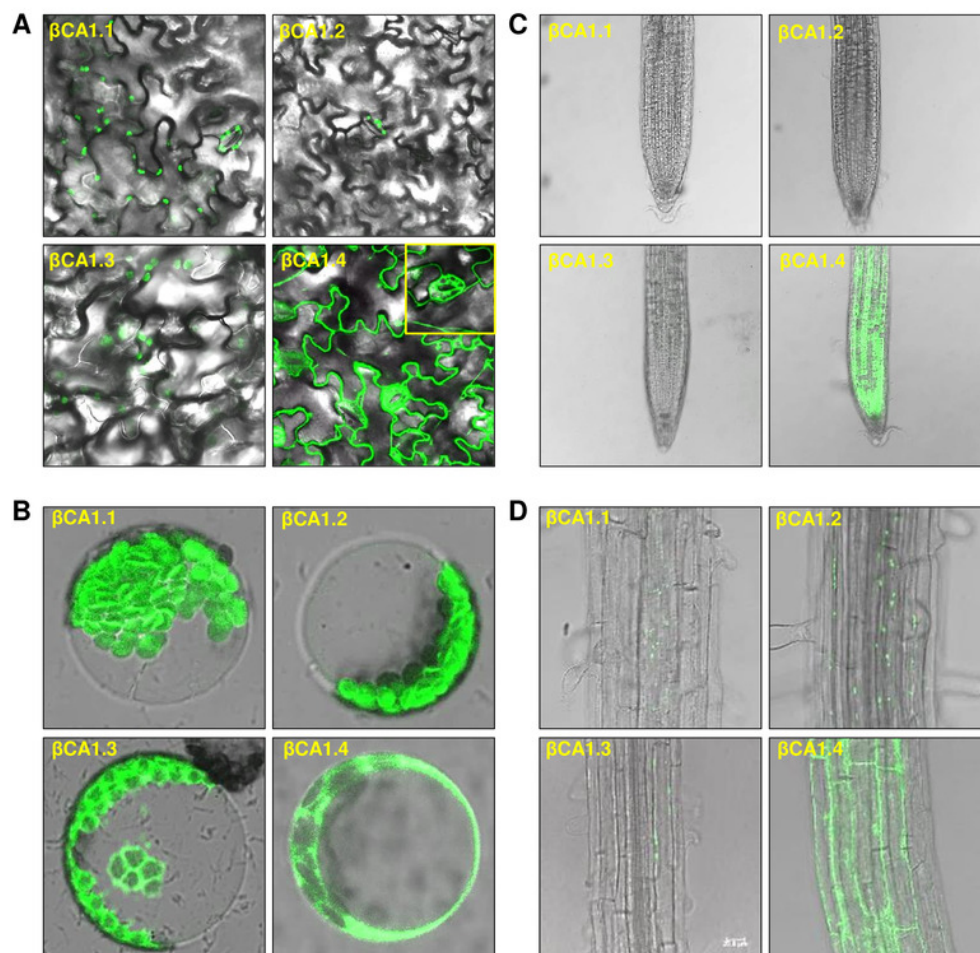
Figure 3

Functional analysis of *Arabidopsis* βCA1 in response to heat stress.

(**A**) RT-PCR svalidation of gene expression patterns of representative *βCA1* spliced variants in the absence (Mock) or presence of stresses. For various stresses treatment, one-week-old seedlings were subjected to 10 μM ABA, 150 mM NaCl (salt) and 40 ◦C (HT) on plates, respectively. (**B**) Representative images of wild-type (Col-0) and *βca1* mutant seedlings without or with heat treatment. Seedlings with green cotyledons were calculated.
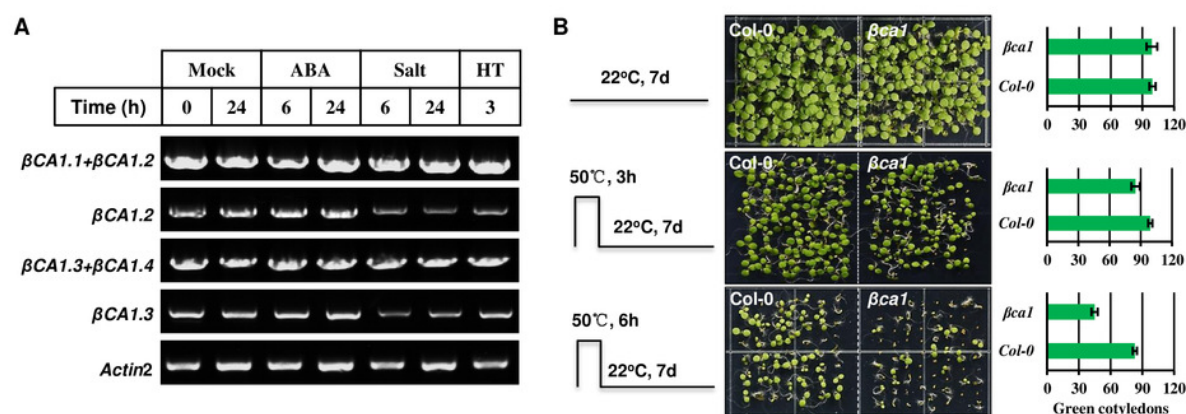
Figure 4

GO analysis of *Arabidopsis* βCA1.4 interacting proteins.

A-B: By using Metascape tool, top 19 clusters with their representative enriched terms (one per cluster) were presented (**A**). In which, 17 proteins were enriched in response to temperature including βCA1 (**B**). The Clusters were indicated by numbers, in which 20 and 21 represented "response to temperature stimulus" and "response to cold", respectively.
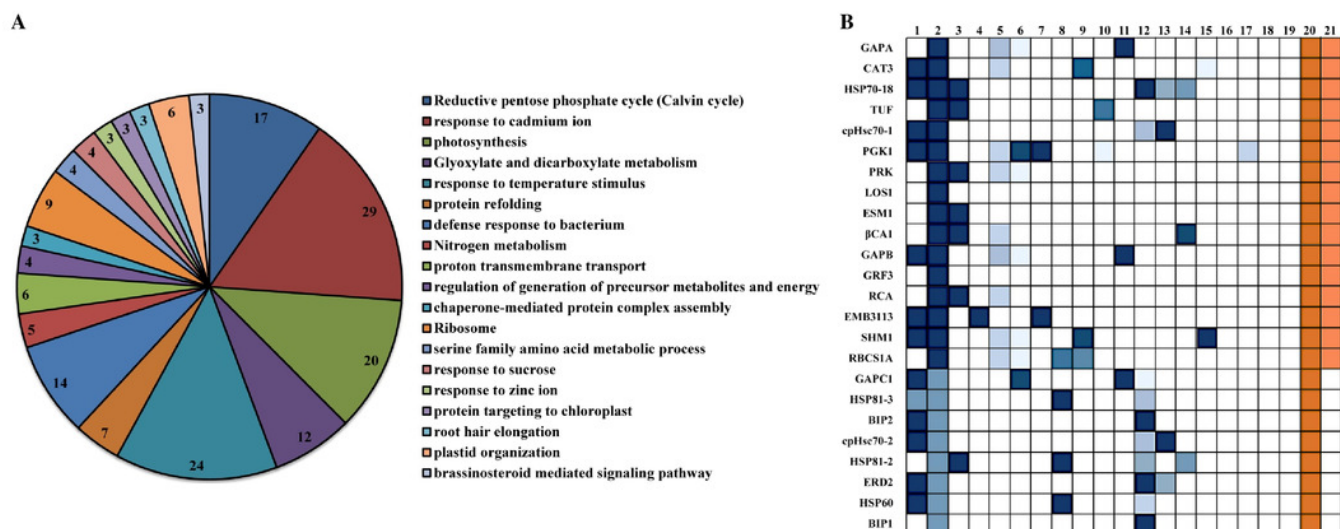
Figure 5

The PPI of *Arabidopsis* βCA1.4 interacting proteins.

The PPI sub-networks associated with photosystem (green), cellular response to stress (red) and ribosome pathway (cyan) were shown.