Divergent receptor proteins confer responses to different karrikins in two ephemeral weeds

Yueming Kelly Sun[1,2], Jiaren Yao[1], Adrian Scaffidi[1], Kim T. Melville[1], Sabrina F Davies[1], Charles S Bond[1], Steven M Smith[3,4], Gavin R Flematti[1] and Mark T Waters[1,2,5]

1. School of Molecular Sciences, The University of Western Australia, 35 Stirling Hwy, Perth, WA 6009, Australia
2. Australian Research Council Centre of Excellence in Plant Energy Biology, The University of Western Australia, 35 Stirling Hwy, Perth, WA 6009, Australia
3. School of Natural Sciences, The University of Tasmania, Hobart, TAS 7000, Australia
4. Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, 1 West Beichen Road, Beijing 100101, PR China.
5. Author for correspondence: mark.waters@uwa.edu.au

Word count (Abstract, main text and methods): 7505 words

ABSTRACT

Wildfires can encourage the establishment of invasive plants by releasing potent germination stimulants, such as karrikins. Seed germination of *Brassica tournefortii*, a noxious weed of Mediterranean climates, is strongly stimulated by KAR_1, which is the archetypal karrikin produced from burning vegetation. In contrast, the closely-related yet non-fire-associated ephemeral *Arabidopsis thaliana* is unusual because it responds preferentially to KAR_2. The α/β-hydrolase KARRIKIN INSENSITIVE2 (KAI2) is the putative karrikin receptor identified in *Arabidopsis*. Here we show that *B. tournefortii* differentially expresses three *KAI2* homologues, and the most highly-expressed homologue is sufficient to confer enhanced responses to KAR_1 relative to KAR_2 when expressed in *Arabidopsis*. We further identify two variant amino acid residues near the KAI2 active site that explain the ligand selectivity, and show that this combination has arisen independently multiple times within dicots. Our results suggest that duplication and diversification of KAI2 proteins could confer upon weedy ephemerals and potentially other angiosperms differential responses to chemical cues produced by environmental disturbance, including fire. (162 words)

INTRODUCTION

Environmental disturbance promotes the establishment of invasive species, posing a potent threat to global biodiversity. Changing wildfire regimes, such as increasing frequency of fires, is one of the most relevant disturbance factors contributing to elevated invasion threat[1]. Wildfires create germination opportunities in part by releasing seed germination stimulants, such as karrikins, from burning vegetation[2, 3]. Karrikins comprise a family of butenolides with six known members. In samples of smoke water generated by burning grass straw, KAR_1 is the most abundant karrikin, while KAR_2 is six times less abundant[4], although these proportions may differ depending on source material, preparation method, age and storage conditions[5]. These two analogues differ only by the presence of a methyl group on the butenolide ring in KAR_1, which is absent in KAR_2 (Supplementary Fig. 1a). Invasive plant species that are responsive to karrikins could utilise natural and human-induced fires to facilitate their establishment[6, 7].

Karrikins can overcome seed dormancy and promote seed germination in a number of smoke-responsive species, as well as others that are not associated with fire regimes[8-10].

46 Furthermore, karrikins also influence light-dependent seedling development. Karrikins
47 enhance the sensitivity of *Arabidopsis thaliana* seedlings to light by inhibiting hypocotyl
48 elongation, stimulating cotyledon expansion, and promoting chlorophyll accumulation in a
49 dose-dependent manner[11]. Karrikins have been reported to enhance seedling survival and
50 biomass in species such as tomato, rice and maize[12, 13]. Such growth-promoting effects of
51 karrikins, especially at critical early stages of the life cycle, have the potential to further
52 encourage the establishment of invasive species after fire events.

53 *Brassica tournefortii* (Brassicaceae; Sahara mustard) is native to northern Africa and the
54 Middle East, but is an invasive weed that blights many ecosystems with a Mediterranean
55 climate and chaparral-type vegetation that are prone to wildfires in North America, Australia
56 and South Africa. *B. tournefortii* seeds can persist in the soil for many seasons, undergoing
57 wet-dry cycling that can influence dormancy and contribute to boom-bust cycles that
58 outcompete native species[9, 14]. *B. tournefortii* plants may radically alter fire frequency and
59 intensity by influencing fuel profiles[15, 16], further exacerbating the impact of fire on
60 susceptible native ecosystems. In addition, seeds of *B. tournefortii* are particularly responsive
61 to smoke-derived karrikins, and show a positive germination response to KAR_1 in the
62 nanomolar range[10]. Accordingly, *B. tournefortii* is particularly well positioned to invade areas
63 disturbed by fire events[17, 18].

64 The putative karrikin receptor KARRIKIN INSENSITIVE 2 (KAI2) was identified in
65 Arabidopsis, a weedy ephemeral that originated in Eurasia but is now widely distributed
66 throughout the northern hemisphere[19-21]. Arabidopsis is not known to colonise fire-prone
67 habitats, but nevertheless seeds germinate in response to karrikins in the micromolar range[22].
68 Unlike most smoke-responsive species that respond more readily to KAR_1[8, 23], Arabidopsis
69 responds preferentially to KAR_2[22]. KAI2 is an evolutionarily ancient α/β-hydrolase and a
70 paralogue of DWARF14 (D14), the receptor for strigolactones[24, 25]. Karrikins and
71 strigolactones are chemically similar by virtue of a butenolide moiety that is necessary for
72 bioactivity[26, 27]. KAI2 and D14 have dual functions as both enzyme and receptor, but the
73 functional significance of the enzymatic activity remains contested[28-32]. Furthermore, the
74 basis for ligand specificity by these two highly congruent proteins remains essentially
75 unknown.

76 Orthologues of *KAI2* are ubiquitous in land plants, and are normally present as a single gene
77 copy within an ancient and highly conserved "eu-KAI2" clade[33]. There is growing evidence

78 that, beyond its ability to mediate karrikin responses, KAI2 has a core ancestral role in
79 perceiving an endogenous KAI2 ligand ("KL") that regulates seed germination, seedling
80 development, leaf shape and cuticle development[34-36]. Since its divergence from the
81 Arabidopsis lineage, the tribe Brassiceae, which includes the genus *Brassica*, underwent a
82 whole genome triplication event 24–29 million years ago[37-39]. This process might have
83 allowed additional KAI2 copies to gain altered ligand specificity, potentially enhancing
84 perception of environmental signals such as karrikins from smoke. Here, we report that two
85 out of three KAI2 homologues expressed in *B. tournefortii* show distinct preferences for
86 different karrikins. We take advantage of the relatively recent genome triplication event in the
87 Brassiceae to identify two amino acids that are sufficient to explain these karrikin preferences
88 and confirm this by mutagenesis. Beyond demonstrating the potential ecological significance
89 of diversity among KAI2 homologues, our findings also reveal active site regions critical for
90 ligand selectivity among KAI2 receptor-enzymes that are found in all land plants.

91 # RESULTS

92 ### *B. tournefortii* is most sensitive to KAR_1

93 To characterise the karrikin response of *B. tournefortii*, we performed multiple physiological
94 and molecular assays comparing KAR_1 activity with that of KAR_2. First, germination of *B.*
95 *tournefortii* seeds was consistently more responsive to KAR_1 than KAR_2 at 10 nM, 100 nM and
96 1 μM (Fig. 1a and Supplementary Fig. 1b-d). Second, homologues of two karrikin-responsive
97 transcripts, *DWARF14-LIKE2* (*BtDLK2*) and *SALT TOLERANCE HOMOLOG7* (*BtSTH7*), identified
98 on the basis of close sequence homology to their Arabidopsis counterparts[11, 20], were
99 significantly more highly expressed when treated with 1 μM KAR_1 than with 1 μM KAR_2 (Fig.
100 1b). These observed differences in seed response are not due to differential karrikin uptake,
101 since both KAR_1 and KAR_2 were taken up from solution at similar rates by *B. tournefortii* seeds
102 during imbibition, as was also true for Arabidopsis seeds (Supplementary Fig. 1e-f). Besides
103 promoting germination of primary dormant seeds, karrikins also inhibited hypocotyl
104 elongation in *B. tournefortii* seedlings, as is the case in Arabidopsis; again, KAR_1 showed a
105 stronger effect than KAR_2 (Fig. 1c-d). Levels of *BtDLK2* transcripts in seedlings were also more
106 responsive to KAR_1 than KAR_2 at a given concentration (Fig. 1e). Therefore, we conclude that
107 *B. tournefortii* is more sensitive to KAR_1 than to KAR_2, a ligand preference that is a feature of
108 seed germination in many karrikin-responsive species from ecosystems prone to fires[8, 23].

109 ## Two out of three *B. tournefortii* KAI2 homologues are functional in Arabidopsis

110 To establish whether there are multiple *KAI2* homologues present in *B. tournefortii*, we
111 examined transcriptomes from seeds and seedlings. Three putative *KAI2* homologues were
112 identified (*BtKAI2a*, *BtKAI2b*, and *BtKAI2c*; Fig. 2a; Supplementary Fig. 2 and 3). *BtKAI2a*
113 grouped with *AtKAI2* and those of other Brassicaceae within a single clade, whereas *BtKAI2b*
114 and *BtKAI2c* grouped within a clade unique to *Brassica*. This taxonomic distribution implies
115 that *BtKAI2a* is most similar to *AtKAI2*, and that *BtKAI2b* and *BtKAI2c* are paralogues that
116 arose via genome triplication during the evolution of *Brassica*. All three *BtKAI2* transcripts
117 were expressed in *B. tournefortii* seeds, but only *BtKAI2a* and *BtKAI2b* could be detected in
118 seedlings (Fig. 2b). In seeds and seedlings, *BtKAI2b* transcripts were the most abundant of the
119 three, but there were no consistent effects on any of the transcripts upon treatment with 1 μM
120 KAR_1. We also identified two *BtD14* homologues, at least one of which is functionally
121 orthologous to *AtD14* (Supplementary Figs. 2 and 4).

122 We performed transgenic complementation of the Arabidopsis *kai2-2* null mutant by
123 expressing each of the three isoforms as a GFP fusion protein driven by the Arabidopsis *KAI2*
124 promoter and 5′UTR (*KAI2pro*:*GFP-BtKAI2*) to accurately reflect the native expression profile
125 of KAI2 in Arabidopsis. Such fusions of GFP with KAI2 proteins have been used previously to
126 analyse KAI2 activity[19, 40, 41]. Both *BtKAI2a* and *BtKAI2b* complemented the seedling and leaf
127 phenotypes of *kai2-2*, whereas *BtKAI2c* did not (Fig. 2c-d). GFP-BtKAI2a and GFP-BtKAI2b
128 accumulated at consistent levels when detected using both an anti-KAI2 antibody and an anti-
129 GFP antibody (which negated the effects of sequence differences in the anti-KAI2 epitope
130 region; Fig. 2e and Supplementary Fig. 3). However, we could not detect GFP-BtKAI2c protein
131 in three independent transgenic lines with either antibody. We verified the *GFP-BtKAI2c*
132 transcript sequences expressed in the transgenic plant material by RT-PCR and Sanger
133 sequencing (Supplementary Fig. 5). Since the *GFP-BtKAI2c* mRNA is processed faithfully in
134 Arabidopsis, and the apparent absence of protein is most likely a result of posttranslational
135 events. We then tested whether BtKAI2c was generally poorly expressed in plant cells by
136 transient overexpression in tobacco leaves using *Agrobacterium*-mediated infiltration and
137 tobacco mosaic virus-based plant expression vectors[42]. Even with the extremely high levels of
138 expression supported by this system, levels of myc-tagged BtKAI2c protein were substantially
139 lower than equivalently-tagged BtKAI2a and BtKAI2b (Supplementary Fig. 5). This result
140 implies that BtKAI2c protein may be inherently unstable and/or prone to degradation in plant
141 cells.

142 **BtKAI2c carries a destabilising mutation at position 98**

143 Numerous mutations identified by forward genetics destabilise AtKAI2 and render the
144 protein undetectable in plant extracts[43]. Ten residues in the BtKAI2c sequence are unique to
145 BtKAI2c within the genus *Brassica* (Supplementary Figure 3). We predicted the likely effect of
146 each of these residues upon protein function, relative to the *Brassica* consensus residues at
147 the same positions, using the PROVEAN tool[44]. Among the ten sites, R98 and E129 exceeded
148 the cutoff value of -2.5 and were thus predicted to be highly deleterious to protein function,
149 while D137 narrowly missed this threshold (Supplementary Table 2). Therefore based on this
150 predictive analysis, the native BtKAI2c protein carries at least two and possibly three
151 substitutions that have the potential to render the protein non-functional.

152 R98 is potentially highly deleterious because it is situated in the active site adjacent to the
153 catalytic serine, whereas E129 and D137 are located on the flexible hinge that links the lid and
154 core domains (Supplementary Fig. 6a). In AtKAI2, the G133E mutation is highly
155 destabilising[20, 41], suggesting that non-conserved substitutions in the hinge region have the
156 potential to be deleterious as well. To test the effect of these residues on BtKAI2c stability, we
157 constructed three variants that reverted the native BtKAI2c sequence to consensus amino
158 acids within *Brassica* at position 98 and/or in the hinge region: BtKAI2c^{R98V}, a quadruple
159 mutant BtKAI2c$^{E129D;D130V;E133Q;D137E}$ (BtKAI2c^{QUAD}), and a quintuple mutant that combined all
160 five substitutions (BtKAI2c^{QUINT}). Should these amino acids be crucial to protein stability, we
161 expected to restore BtKAI2c expression.

162 First, we tried expressing the native and mutated BtKAI2c variants in tobacco leaves using
163 *Agrobacterium*-mediated infiltration. The vectors were based on those used to generate the
164 *GFP-BtKAI2c* Arabidopsis transgenics, allowing us to compare expression of the same GFP
165 fusion proteins. As was the case for stably-transformed Arabidopsis, we found that native
166 GFP-BtKAI2c protein was undetectable, as was the BtKAI2c^{QUAD} variant (Supplementary Fig.
167 6b). However, BtKAI2c^{R98V} and BtKAI2c^{QUINT} were clearly detected. We also examined the
168 effect of the equivalent V96R mutation on GFP-AtKAI2, and found that this mutation
169 completely abolished expression compared to the native GFP-AtKAI2 control (Supplemental
170 Fig. 6b). These results suggest that the R98 residue causes the loss of protein expression in
171 BtKAI2c, and that arginine at this position is not tolerated by KAI2 proteins in general.

172 Early in our investigation, we had tried without success to express and purify BtKAI2c from *E.*
173 *coli* for biochemical characterisation. Having restored BtKAI2c expression in plants with the

174 R98V mutation, we next tried to express these proteins in bacteria. We found that SUMO-
175 BtKAI2c R98V was expressed at ~10-fold higher levels than native SUMO-BtKAI2c in crude
176 lysates, suggesting that the R98V mutation enhanced protein folding and/or stability in
177 bacterial cells. Although it was detectable in crude lysates, SUMO-BtKAI2c was consistently
178 intransigent to recovery following affinity chromatography; in contrast, the R98V and
179 quintuple mutant versions were stably recovered at high levels and purity (Supplemental
180 Figure 6c). From these results, we conclude that native BtKAI2c is non-functional due to
181 protein instability induced by a highly non-conservative substitution at the active site.

182 **BtKAI2a and BtKAI2b show differential ligand specificity**

183 We further characterised the ligand specificity of *BtKAI2* homologues by performing
184 physiological and molecular assays with the stable transgenic Arabidopsis lines. Primary-
185 dormant Arabidopsis seeds homozygous for the transgenes were tested for germination
186 response (Fig. 3a and Supplementary Fig. 7). None of the transgenic lines completely restored
187 germination levels to those of wild type, perhaps reflecting the need for additional, seed-
188 specific regulatory elements not included in our transgenes, or the importance of genomic
189 context for proper *KAI2* expression in seeds. Nevertheless, germination responses to karrikins
190 were evident in the transgenic seeds: germination of *BtKAI2b* seeds was more responsive to
191 KAR_1 than KAR_2 at 1 μM, whereas no significant difference was observed for *BtKAI2a* seeds. As
192 expected, *BtKAI2c* seeds were indistinguishable from the *kai2* control and insensitive to
193 karrikins. We also found that levels of *KUF1* transcripts, which are responsive to karrikins in
194 Arabidopsis seeds (Nelson 2010), were only slightly induced by karrikins in *BtKAI2a*
195 transgenic seeds, but were strongly enhanced by KAR_1 in *BtKAI2b* seeds (Fig. 3b).

196 In seedlings, both *BtKAI2a* and *BtKAI2b*, but not *BtKAI2c*, could complement the *kai2*
197 hypocotyl elongation phenotype (Fig 3c). Responses to karrikins in terms of hypocotyl
198 elongation (Fig. 3c) and levels of *DLK2* transcripts (Fig. 3d) broadly agreed with the
199 germination response with respect to karrikin preferences, with *BtKAI2b* transgenics showing
200 a clear preference for KAR_1 over KAR_2. However, *BtKAI2a* seedlings showed a preference for
201 KAR_2 that was not evident in seeds. From these experiments, we conclude that BtKAI2a is
202 similar to AtKAI2 in terms of ligand specificity, whereas BtKAI2b has a clear preference for
203 KAR_1 over KAR_2. As *BtKAI2b* is more highly expressed than *BtKAI2a* in *B. tournefortii* seeds
204 and seedlings (Fig. 2b), we conclude that the ligand specificity of BtKAI2b substantially

205 contributes to the enhanced KAR_1-responsiveness of this species at these stages of the life
206 cycle.

207 **Positions 98 and 191 account for ligand specificity between BtKAI2a and BtKAI2b**

208 To investigate interactions between BtKAI2 homologues and ligands, we performed
209 differential scanning fluorimetry (DSF) assays on purified recombinant proteins
210 (Supplementary Fig. 8). DSF has been used extensively for inferring the interaction of
211 strigolactone-like compounds with D14- and KAI2-related proteins[29, 41, 45-48]. Racemic GR24 is
212 a widely-used synthetic strigolactone analogue that consists of two enantiomers
213 (Supplementary Fig. 1a), of which GR24$^{ent\text{-}5DS}$ is bioactive via AtKAI2[40, 49]. Catalytically inactive
214 D14 and KAI2 variants do not respond to GR24 in DSF, suggesting that the shift in thermal
215 response results from ligand hydrolysis and a corresponding conformational change in the
216 receptor[29, 41, 46]. In DSF assays, AtKAI2 shows a specific response to >100 μM GR24$^{ent\text{-}5DS}$ but,
217 for unclear reasons, no response to karrikins[41]. Likewise, we found that both BtKAI2a and
218 BtKAI2b were also unresponsive to karrikins in DSF (Supplementary Fig. 9a). Therefore, we
219 used DSF assays with GR24$^{ent\text{-}5DS}$ as a surrogate substrate as a means to analyse differences in
220 ligand interactions between BtKAI2a and BtKAI2b. We used N-terminal 6xHIS-SUMO as a
221 solubility tag to aid expression in *E. coli* and enhance stability. We found that the presence of
222 the tag did not compromise the thermal destabilisation response of BtKAI2b to GR24$^{ent\text{-}5DS}$
223 (Supplementary Fig. 9b), and therefore we used intact SUMO fusion proteins for all other
224 experiments.

225 We found that AtKAI2 and BtKAI2a showed similar responses to >100 μM GR24$^{ent\text{-}5DS}$, but the
226 response of BtKAI2b was clear at >25 μM (Supplementary Fig. 9c). We then used a lower
227 range of GR24$^{ent\text{-}5DS}$ concentrations (0–50 μM) to determine the threshold for response, which
228 we defined as a statistically significant reduction in the maximal rate of change in
229 fluorescence at the melting temperature of the protein (T_m). Although BtKAI2a showed only a
230 weak and non-significant response at 40 and 50 μM GR24$^{ent\text{-}5DS}$, BtKAI2b responded
231 significantly at 10 μM and above (Fig. 4a). These results suggest that BtKAI2b is more
232 sensitive than BtKAI2a to GR24$^{ent\text{-}5DS}$, and that BtKAI2a is most like AtKAI2 in this respect.

233 BtKAI2a and BtKAI2b differ in primary amino acid sequence at just 14 positions
234 (Supplementary Fig. 3). We postulated that differences in ligand specificity might be
235 determined by amino acids in the vicinity of the ligand binding pocket. Protein structural
236 homology models revealed only two residues that differ in this region: V98 and V191 in

237 BtKAI2a, and L98 and L191 in BtKAI2b; the corresponding residues in AtKAI2 and AtD14 are
238 valines (Supplementary Fig. 8). Residue 98 is immediately adjacent to the catalytic serine at
239 the base of the pocket. Residue 191 is located internally on αT4 of the lid domain which, in
240 AtD14, is associated with a major rearrangement of protein structure upon ligand binding
241 that reduces the size of the pocket[31]. Homology modelling suggests only subtle differences in
242 the size and shape of the primary ligand binding pocket of BtKAI2a and BtKAI2b
243 (Supplementary Fig. 8). To determine if these residues are pertinent to ligand specificity, we
244 replaced the two valine residues of BtKAI2a with leucine residues, generating the variant
245 BtKAI2a$^{L98;L191}$, and *vice-versa* for BtKAI2b, generating the variant BtKAI2b$^{V98;V191}$
246 (Supplementary Fig. 8). In DSF assays, we found that exchanging the two residues was
247 sufficient to switch the original responses, such that BtKAI2a$^{L98;L191}$ responded sensitively to
248 GR24$^{ent\text{-}5DS}$, but BtKAI2b$^{V98;V191}$ did not (Fig. 4b). We also assessed the function of the stable
249 BtKAI2c^{R98V} and BtKAI2c^{QUINT} variants by DSF against GR24$^{ent\text{-}5DS}$ ligand, and found that they
250 were only weakly destabilised by this ligand, when compared to the highly sensitive BtKAI2b
251 (Supplemental Fig. 6e). Overall, these results indicate that BtKAI2a and BtKAI2b have distinct
252 response profiles, albeit to a synthetic ligand, and that residues 98 and 191 contribute to this
253 difference.

254 We reasoned that the most robust and informative means to assess the effect of residues 98
255 and 191 upon the response to karrikins was to test their function directly *in planta*. We
256 expressed BtKAI2a$^{L98;L191}$ and BtKAI2b$^{V98;V191}$ as GFP fusion proteins in the Arabidopsis *kai2-2*
257 null mutant background driven by the Arabidopsis *KAI2* promoter (*KAI2pro:GFP-*
258 *BtKAI2a$^{L98;L191}$* and *KAI2pro:GFP-BtKAI2b$^{V98;V191}$*), and selected two independent homozygous
259 transgenic lines for each, on the basis of protein expression level and segregation ratio
260 (Supplementary Fig. 10). Using two different assays, we found that substitutions between
261 BtKAI2a and BtKAI2b at positions 98 and 191 also reversed karrikin preference in
262 Arabidopsis seedlings both in terms of hypocotyl elongation and *DLK2* transcript levels (Fig.
263 4c, d; Supplementary Fig. 11). Most prominently, the clear preference of BtKAI2b for KAR_1
264 was unambiguously reversed to a preference for KAR_2 in *BtKAI2b$^{V98;V191}$* transgenics,
265 effectively recapitulating the response of the native BtKAI2a protein. We also examined the
266 responses of seeds in these transgenic lines, and found that *BtKAI2a$^{L98;L191}$* seeds germinated
267 with a clear preference for KAR_1, whereas *BtKAI2b$^{V98;V191}$* seeds showed no clear preference
268 for either karrikin (Supplementary Fig. 12). Notably, this germination response pattern is the
269 opposite to that observed for seeds expressing the native BtKAI2 proteins, with BtKAI2a

270 exhibiting no preference for either karrikin and BtKAI2b a preference for KAR_1 (Fig. 3a,
271 Supplementary Fig. 7). Taken together, these results demonstrate that the residues at
272 positions 98 and 191 largely determine differences in karrikin specificity between BtKAI2a
273 and BtKAI2b, at least in the context of Arabidopsis seedlings.

274 Taxonomic variability in KAI2 sequences suggests functional co-dependency between residues 275 96 and 189

276 We wished to gain further insight into the functional importance of diversity at positions 96
277 and 189 of KAI2 proteins (based on AtKAI2 notation, equivalent to positions 98 and 191 in
278 BtKAI2a and BtKAI2b). We analysed 476 distinct KAI2 homologues from 441 angiosperm
279 species (spanning eudicots, monocots and magnoliids, but excluding *B. tournefortii*) using
280 data collected from the 1000 Plant Transcriptomes project[50]. We found that at positions 96
281 and 189, valine was by far the most common amino acid (93.1% and 93.5% of sequences,
282 respectively; Figure 5a and Supplemental Table 3). L96 was observed in 6.1% of sequences,
283 while L189 was even rarer (2.1%). The L96;V189 combination was observed in 14 sequences
284 (2.9%), whereas V96;L189 was observed in only one homologue from *Ternstroemia*
285 *gymanthera* (Pentaphylacaceae; 0.2%). Notably, this species also expressed a second KAI2
286 homologue with the canonical V96;V189 combination. If the frequencies of valine and leucine
287 at positions 96 and 189 were independent, the L96;L189 combination would be expected to
288 occur less than once among 476 sequences (0.13%). However, the coincidence of L96 and
289 L189 was observed in nine sequences (1.9%), representing a significant 15-fold enrichment
290 for the L96;L189 pair over that expected by chance alone ($\chi^2 = 115.6$; $P<0.0001$).

291 Among the nine newly-identified $KAI2^{L96;L189}$ sequences, eight were restricted to eudicots,
292 distributed evenly between both the superrosids (Brassicales, Fabales and Saxifragales; four
293 species) and superasterids (Caryophyllales; four species). The ninth sequence, co-expressed
294 with another transcript encoding $KAI2^{V96;V189}$, was identified from the basal dicot *Hakea*
295 *drupacea* (syn. *Hakea suaveolens*; Proteaceae). Strikingly, this species is native to
296 southwestern Australia, but is recognised as an invasive weed in South Africa that relies upon
297 fire to liberate seeds from the canopy[51]. The L96;L189 combination therefore seems to have
298 arisen independently on multiple occasions. Furthermore, within the order Saxifragales,
299 *Daphniphyllum macropodum* ($KAI2^{L96;L189}$) belongs to a clade of species that all express
300 $KAI2^{L96;V189}$ homologues (Figure 5b). Although this clade is only represented by five
301 sequences and four species, this distribution pattern suggests that a V96L mutation occurred

302 first in the common ancestor of the Hamamelidaceae, Cercidiphyllaceae and
303 Daphyniphyllaceae, followed by V189L in the lineage leading to *D. macropodum*. Overall, this
304 analysis suggests that positions 96 and 189, while not in direct contact with each other in the
305 protein tertiary structure (Supplemental Fig. 8), likely exhibit functional interdependence.
306 Valine is most probably the ancestral state at both positions, but leucine at position 96 is also
307 tolerated. While rare, leucine at position 189 is generally found in combination with leucine at
308 position 96, and tentatively, a V96L mutation might potentiate a subsequent V189L mutation.

309 The above analysis suggested that a V96L substitution might be sufficient to confer altered
310 substrate specificity upon KAI2 proteins. To test this hypothesis, we generated transgenic
311 Arabidopsis seedlings expressing GFP-BtKAI2b proteins with just a single amino acid
312 substitution from leucine to valine at either position 98 or 191 (i.e. *KAI2pro:GFP-BtKAI2b*^{V98} or
313 *KAI2pro:GFP-BtKAI2b*^{V191}), and examined their response to KAR_1 and KAR_2 in hypocotyl
314 elongation assays. We found that the L98V substitution was sufficient to invert the karrikin
315 preference of GFP-BtKAI2b in three independent homozygous transgenic lines (Figure 5c;
316 Supplemental Figure 13). Although this result does not rule out an additional contribution
317 towards ligand specificity from L191, it is consistent with the higher frequency of L96
318 compared to L189 observed among the angiosperms, and suggests that substitution at
319 position 96 may have been selected relatively frequently in the evolution of KAI2 function.

320 DISCUSSION

321 The plant α/β-hydrolases KAI2 and D14 are characterised by responsiveness towards
322 butenolide compounds, including endogenous strigolactones and strigolactone-related
323 compounds, abiotic karrikins derived from burnt vegetation, and synthetic strigolactone
324 analogues with a wide array of functional groups. Direct evidence for a receptor-ligand
325 relationship between karrikins and KAI2 homologues stems from crystallography and *in vitro*
326 binding assays. Two crystal structures of KAR-responsive KAI2 proteins from Arabidopsis and
327 the parasitic plant *Striga hermonthica* reveal largely similar overall protein structure, but
328 surprisingly are non-congruent with respect to KAR_1 binding position and orientation[52, 53].
329 The affinity of AtKAI2 for KAR_1 is imprecisely defined, with estimates of dissociation
330 coefficients (Kd) ranging from 4.6 μM[54], to 26 μM[55], to 148 μM[56] using isothermal calorimetry,
331 and 9 μM using fluorescence microdialysis[52]. While variability in affinity estimates can be
332 explained in part by different experimental conditions and techniques, it should also be

333 considered that, depending on the homologue under examination, KAR_1 may not be the
334 optimal ligand. Furthermore, binding of a hydrophobic compound to a hydrophobic pocket
335 does not necessarily equate to ligand recognition or receptor activation. Our data, which are
336 derived from clear and distinct biological responses, provide strong evidence that KAI2 is
337 sufficient to determine ligand specificity, which in turn strengthens the case that KAI2 is the
338 receptor by which karrikins are perceived.

339 The predominance of valine and leucine at position 96 (98 in BtKAI2 proteins), observed in
340 99.2% of our sample of 476 KAI2 sequences, suggests that this site is critical to protein
341 function, which makes sense given its immediate proximity to the catalytic serine. This
342 position not only contributes to ligand preference, but also to protein stability, given the
343 negative effects of R98 in BtKAI2c. Compared with valine and leucine, arginine is relatively
344 bulky with a charged side chain; indeed, in the few KAI2 sequences in which position 96 is
345 neither leucine nor valine, the residue is nevertheless small and hydrophobic (either
346 methionine or alanine; Supplementary Table 3). Interestingly, substitutions in AtKAI2 at or
347 near the catalytic site, such as S95F, S119F and G245D, are not tolerated and render the
348 protein unstable in plants[43]. The likely signalling mechanism for KAI2 – by analogy of that for
349 D14 – involves ligand binding and/or hydrolysis, which triggers conformational changes in
350 the lid domain and formation of a protein complex with MAX2 and SMAX1/SMXL2 protein
351 partners[30, 31, 57]. As a result of signalling, KAI2 itself is degraded[40, 43]. It is possible that KAI2 is
352 an inherently metastable protein that is prone to conformational change and subsequent
353 degradation by necessity; perhaps this explains why it is so susceptible to instability through
354 mutation.

355 Ligand specificity is a key contributor to the functional distinction between karrikin and
356 strigolactone receptors, and elucidating the molecular mechanisms behind ligand specificity is
357 a significant research challenge. In certain root-parasitic weeds in the Orobanchaceae,
358 substitutions of bulky residues found in AtKAI2 with smaller hydrophobic amino acids that
359 increase the ligand-binding pocket size have likely improved the affinity for host-derived
360 strigolactone ligands, as opposed to the smaller-sized karrikin-type ligands[58-60]. Moreover, lid-
361 loop residues that affect the rigidity and size of the ligand entry tunnel determine the ligand
362 selectivity between KAR_1 and *ent*-5-deoxystrigol (a strigolactone analogue with non-natural
363 stereochemistry) among eleven KAI2 homologues in *Physcomitrella patens*[61]. Karrikin and
364 strigolactone compounds also show chemical diversity within themselves, yet ligand
365 discrimination among KAI2 and D14 receptors is not well characterised. Our data

366 demonstrate that, in the case of BtKAI2 proteins, the two KAI2 homologues respond
367 differently to subtly different KAR analogues, and that subtle changes in pocket residues likely
368 account for preferences between these ligands. Position 96 appears to play a primary role in
369 determining ligand specificity, while position 189 may improve ligand sensitivity, protein
370 stability or protein conformational dynamics following a prior mutation at position 96.
371 Although the identified residues L98 and L191 are not predicted to change dramatically the
372 pocket size of BtKAI2b in comparison to BtKAI2a, these residues would make the BtKAI2b
373 pocket more hydrophobic, which is consistent with KAR_1 being more hydrophobic than
374 KAR_2[4]. Therefore, ligand specificity between highly similar chemical analogues can be
375 achieved through fine-tuning of pocket hydrophobicity. Similarly subtle changes have been
376 reported for the rice gibberellin receptor GID1: changing Ile133 to Leu or Val increases the
377 affinity for GA_{34} relative to the less polar GA_4, which lacks just one hydroxyl group[62].

378 It is possible that what we learn about ligand specificity in KAI2 proteins may also be
379 informative about strigolactone perception by D14. Different D14 proteins may show ligand
380 specificity towards diverse natural strigolactones, resulting in part from multiplicity of
381 biosynthetic enzymes, such as the cytochrome P450 enzymes in the MAX1 family[45, 63, 64] and
382 supplementary enzymes such as LATERAL BRANCHING OXIDOREDUCTASE[65]. Although the
383 functional reasons underlying such strigolactone diversity are still unclear, it is possible that
384 variation among D14 homologues yields varying affinities for different strigolactone
385 analogues[66]. As an increasingly refined picture emerges of the features that determine ligand
386 specificity for KAI2 and D14 receptors, we envision the rational design of synthetic receptor
387 proteins with desirable ligand specificity in the future.

388 Gene duplication is a common feature in plant evolutionary histories as an initial step towards
389 neofunctionalisation[67]. In obligate parasitic weeds, duplication and diversification of *KAI2*
390 genes (also referred to as *HTL* genes) have shifted ligand specificity towards strigolactones[58,
391 59]. Karrikins, as abiotic molecules with limited natural occurrence, are unlikely to be the
392 natural ligand for most KAI2 proteins, which are found throughout land plants. Instead, the
393 evolutionary maintenance of a highly conserved receptor probably reflects the core KAI2
394 function of perceiving an endogenous ligand ("KL") that regulates plant development[34-36]. This
395 core function is presumably retained after gene duplication events through original, non-
396 divergent copies that are under purifying selection[58, 59]. KAI2 diversity may also reflect
397 differing affinity for KL variants in different species and at different life stages. Our results are
398 consistent with a scenario in which both BtKAI2a and BtKAI2b have retained the core KAI2

399 function of perceiving KL to support plant development, because both copies complement the
400 Arabidopsis *kai2* phenotype, especially at the seedling and later stages. However, BtKAI2b has
401 also acquired mutations that alter its ligand specificity, which in turn enhance sensitivity to
402 the archetypal karrikin released first discovered in smoke. The *BtKAI2b* gene is also more
403 highly expressed than *BtKAI2a* in seeds and seedlings, consistent with an adaptation to post-
404 fire seedling establishment. This diversity among KAI2 proteins may provide *B. tournefortii*
405 with a selective advantage in fire-prone environments, contributing to its invasive nature.

406 We have shown evidence for the independent evolution of valine-to-leucine substitutions in
407 KAI2 homologues throughout the dicots, albeit at low frequency, suggesting that these
408 changes may have adaptive significance in various ecological contexts. It will be of interest to
409 assess the function of such homologues in other species, to validate and extend the specific
410 conclusions we have made here. Recent findings indicate that KAI2 is an integrator that also
411 modulates germination in response to other abiotic environmental signals, including
412 temperature and salinity[68]. As germination is a critical life stage for seed plants, strategies for
413 environmental conservation, restoration and weed control will benefit from specific
414 knowledge of KAI2 sequence diversity and expression profiles.

415 # METHODS

416 ## Chemical synthesis

417 Karrikins (KAR_1 and KAR_2), $^{13}[C]_5$-labelled karrikins and GR24 enantiomers ($GR24^{5DS}$ and
418 $GR24^{ent\text{-}5DS}$) were prepared as previously described[69, 70].

419 ## Plant material

420 Arabidopsis *kai2-2* (L*er*) and *Atd14-1* (Col-0) mutants were previously described[20].
421 Arabidopsis plants were grown on a 6:1:1 mixture of peat-based compost (Seedling Substrate
422 Plus; Bord Na Mona, Newbridge, Ireland), vermiculite and perlite, respectively. Light was
423 provided by wide-spectrum white LED panels emitting 120-150 µmol photons.m^{-2}.s^{-1} with a
424 16 h light/8 h dark photoperiod, a 22 °C light/16 °C dark temperature cycle, and constant
425 60% relative humidity. The seeds of *Brassica tournefortii* used in this work were collected in
426 November and December 2009 from two sites in Western Australia (Kings Park in Perth, and
427 Merridin)[71]. *Brassica tournefortii* seeds were dried to 15% relative humidity for one month
428 prior to storage in air-tight bags at –20 °C.

429 ### Seed germination assays

430 Seed germination assays using Arabidopsis were performed on Phytagel as described
431 previously[49]. *Brassica tournefortii* seeds were sowed in triplicates (35-70 seeds each) on glass
432 microfibre filter paper (Grade 393; Filtech, NSW Australia) held in 9-cm petri dishes and
433 supplemented with mock or karrikin treatments (3 mL of aqueous treatment solution per
434 petri dish). Treatments were prepared by diluting acetone (mock) or karrikin stocks
435 (dissolved in acetone) 1:1000 with ultrapure water. Because germination of *B. tournefortii* is
436 inhibited by light[9], the seeds were imbibed in the dark at 22 °C. Numbers of germinated seeds
437 were counted each day until the germination percentages remained unchanged.

438 ### Hypocotyl elongation assays

439 Arabidopsis hypocotyl elongation assays were performed under red light as described
440 previously[20]. For *B. tournefortii*, assays were performed with the following modifications to
441 the Arabidopsis protocol: *B. tournefortii* seeds were sowed in triplicate on glass microfibre
442 filter paper (Filtech) held in petri-dishes supplemented with mock or karrikin treatments. The
443 seeds were imbibed in the dark for 22 h at 24 °C before exposing to continuous red light (5
444 μmol photons m^{-2} s^{-1}) for 4 days.

445 ### Transcript analysis

446 *Brassica tournefortii* seeds were imbibed on glass fibre filters in the dark at 22 °C and treated
447 with karrikins using the same procedure as described under seed germination assays, above.
448 For transcript quantification in *B. tournefortii* seeds, samples were taken at the indicated time
449 points, briefly blot-dried on paper towel, and frozen in liquid nitrogen. For *B. tournefortii*
450 seedlings, two methods were used. For data shown in Fig. 1e, seeds were sowed and seedlings
451 grown in triplicate under identical conditions to those described for hypocotyl elongation
452 assays. For data shown in Fig. 2b, seeds were first imbibed for 24 h in the dark, and then
453 transferred to the growth room (~120-150 μmol photons m^{-2} s^{-2} white light, 16 h light/8 h
454 dark, 22 °C light/16 °C dark) for three days. A sample of approximately 20 seedlings was then
455 transferred to a 250 mL conical flask containing 50 mL sterile ultrapure water supplemented
456 with 1 μM KAR_1, or an equivalent volume of acetone (0.1% v/v), with three replicate samples
457 per treatment. The flasks were then shaken for 24 h before seedlings were harvested.

458 Arabidopsis seedlings were grown on 0.5× MS agar under growth room conditions (wide-
459 spectrum white LED panels emitting 120-150 µmol photons.m^{-2}.s^{-1} with a 16 h light/8 h dark
460 photoperiod and a 22 °C light/16 °C dark temperature cycle) for seven days. On the seventh
461 day, seedlings were transferred to 3 mL liquid 0.5× MS medium in 12-well culture plates
462 (CORNING Costar 3513) and shaken at 70 rpm for a further 22 h under the same growth room
463 conditions. The medium was then removed by pipette and replaced with fresh medium
464 containing relevant compounds or an equivalent volume (0.1% v/v) of acetone. After a
465 further incubation period with shaking (8 h), the seedlings were harvested, blotted dry, and
466 frozen in liquid nitrogen.

467 RNA extraction, DNase treatment, cDNA synthesis and quantitative PCR were conducted as
468 previously described[72]. All oligonucleotides are listed in the Supplementary Table 4.

469 **Cloning and mutagenesis of *Brassica tournefortii KAI2* and *D14* homologues**

470 Full-length *BtKAI2* coding sequences (and unique 3′UTR sequences) were amplified from *B.*
471 *tournefortii* cDNA with Gateway-compatible *attB* sites using the universal forward primer
472 BtKAI2_universal_F and homologue-specific reverse primers RACE_R (BtKAI2a), Contig1_R
473 (BtKAI2b) and Contig5_R (BtKAI2c), before cloning into pDONR207 (Life Technologies). The
474 pDONR207 clones were confirmed by Sanger sequencing and recombined with pKAI2pro-
475 GFP-GW[41] to generate the binary plant expression plasmids pKAI2pro-GFP-BtKAI2a,
476 pKAI2pro-GFP-BtKAI2b and pKAI2pro-GFP-BtKAI2c. For transient overexpression in tobacco
477 (Supplementary Figure 5d), BtKAI2 coding sequences were transferred via Gateway-mediated
478 recombination into pSKI106, which drives very high expression of coding sequences placed
479 immediately downstream of a full-length cDNA clone of the tobacco mosaic virus U1 strain, all
480 under control of the CaMV *35S* promoter in a standard T-DNA binary vector[73]. pSKI106 also
481 encodes an N-terminal 3× *c*-myc tag. For transient expression in tobacco using standard
482 binary vectors (Supplementary Figure 6b), the *BtKAI2* coding sequences were transferred into
483 pMDC43, which is identical to pKAI2pro-GFP-GW except that the *AtKAI2* promoter & 5′UTR
484 sequences are replaced with a double CaMV *35S* promoter[74]. The *BtD14a* coding sequence
485 was amplified from cDNA using oligonucleotides BtD14_F and BtD14_R, cloned into
486 pDONR207 as above, and transferred into pD14pro-GW[41].

487 The full-length *BtKAI2* coding sequences (excluding 3′UTRs) were amplified from the
488 pDONR207 clones and reconstituted with the pE-SUMO vector by Gibson Assembly to
489 generate the heterologous expression plasmids pE-SUMO-BtKAI2a, -BtKAI2b and -BtKAI2c.

490 Site-directed mutagenesis generated pE-SUMO-BtKAI2a$^{L98;L191}$, pE-SUMO-BtKAI2b$^{V98;V191}$, and
491 the BtKAI2c variants. For expression of mutated versions of BtKAI2a, BtKAI2b, BtKAI2c and
492 AtKAI2 in tobacco and Arabidopsis, site-directed mutagenesis was performed on pDONR207
493 clones prior to recombination with pKAI2pro-GFP-GW or pMDC43. In the case of the double
494 substitutions in BtKAI2a and BtKAI2b, both targeted residues were mutated simultaneously
495 in one PCR product, while the remainder of the plasmid was amplified in a second PCR
496 product. These double mutated plasmids were reconstituted by Gibson assembly. For the
497 single mutants, the entire plasmid was amplified using a single pair of mutagenesis primers,
498 and the PCR product was self-ligated. The BtKAI2c^{QUINT} variant was generated by two
499 successive rounds of mutagenesis. Coding regions were confirmed by Sanger sequencing.

500 **Plant transformation**

501 Homozygous *kai2-2* plants were transformed by floral dip. Primary transgenic seedlings were
502 selected on sterile 0.5× MS medium supplemented with 20 µg/mL hygromycin B. T_2 lines
503 exhibiting a 3:1 ratio of hygromycin resistant-to-sensitive seedlings were propagated further
504 to identify homozygous lines in the T_3 generation. Experiments were performed from the T_3
505 generation onwards.

506 For transient expression in tobacco, *Agrobacterium* (GV3101) carrying pSKI106 or pMDC43
507 variants was grown in LB medium (25 mL) supplemented with antibiotics and 20 µM
508 acetosyringone until OD_{600} reached 1.0. The bacteria were then harvested by centrifugation
509 (15 min, 5000 × *g*) and resuspended in 10 mM $MgCl_2$, 10 mM MES (pH 5.6) and 100 µM
510 acetosyringone. The optical density was adjusted to 0.4, and the suspension was left standing
511 at 22 °C overnight (approximately 14 h). Leaves of three-week-old *Nicotiana benthamiana*
512 were then infiltrated with a 5-mL syringe, through the abaxial leaf surface. After four days, the
513 leaves were collected and frozen in liquid nitrogen.

514 **Karrikin uptake measurements**

515 Fifteen samples of seeds were sowed for each karrikin treatment (five time points, each in
516 triplicate). In each sample, approximately 40 mg of *Brassica tournefortii* seeds were imbibed
517 in 3 mL ultrapure water for 24 h in 5-mL tubes. After centrifugation (2 min at 3220 × *g*),
518 excess water was removed by pipette and the volume of residual water (mostly absorbed into
519 the seeds) was calculated by weighing the seeds before and after imbibition. Fresh ultrapure
520 water was added to the fully-imbibed seeds to reach a total volume of 980 µL. Then 20 µL of

521 100 μM KAR_1 or KAR_2 was added to a final concentration of 2 μM. The seeds were imbibed at
522 22 °C in darkness. At the indicated time point (0, 2, 4, 8 or 24 h post-treatment), 500 μL of the
523 imbibition solution was removed and combined with 100 ng of either $^{13}[C]_5$-KAR_1 or $^{13}[C]_5$-
524 KAR_2 (100 μL at 1 μg/mL) as an internal standard for quantification purposes. The sample
525 was then extracted once with ethyl acetate (500 μL), and 1 μL of this organic layer was
526 analysed using GC-MS in selective ion monitoring (SIM) mode as previously described[75].

527 The amount of KAR_1 in each sample was calculated by the formula $\frac{A(Ion150)}{A(Ion155)}$×100 ng and
528 converted to moles, where *A(Ion150)* indicates the peak area of the ion 150 (KAR_1 to be
529 measured), *A(Ion151)* indicates the peak area of the ion 151 ($^{13}[C]_5$-KAR_1), and 100 ng is the
530 amount of $^{13}[C]_5$-KAR_1 spiked in before the ethyl acetate extraction. Similarly, the amount of
531 KAR_2 in each sample was calculated by the formula $\frac{A(Ion136)}{A(Ion141)}$×100 ng and converted to moles.
532 The uptake percentage adjusted to 40 mg of *B. tournefortii* seeds was calculated by the
533 formula: $(1 - \frac{N(x)}{N(0)}) \times \frac{40\ mg}{m\ (seed)}$, where N(0) indicates moles of karrikins at time 0, $N(x)$ indicates
534 moles of karrikins at time point x, and $m(seed)$ indicates the dry weight (mg) of seeds tested in
535 each replicate. For Arabidopsis seeds, the procedure was scaled down for a smaller mass of
536 seeds (20 mg).

537 **Transcriptome assembly and analysis**

538 Twenty milligrams of dry *B. tournefortii* seeds were imbibed in water for 24 h and incubated
539 at 22 °C in the dark. The seeds were collected by centrifugation, blotted dry, and frozen in
540 liquid nitrogen. A separate sample of seeds was sown on glass filter paper, imbibed for 24 h as
541 above, and then incubated for 96 h under continuous red light (20 μmol m^{-2} s^{-1}) with a 22 °C
542 (16 h)/16 °C (8h) temperature cycle. A single sample of seedlings (50 mg fresh weight) was
543 harvested and frozen in liquid nitrogen. Total RNA was extracted from both seed and seedling
544 samples using the Spectrum Plant RNA kit (Sigma-Aldrich), including an on-column DNase
545 step. PolyA$^+$ mRNA was purified using oligo(dT) magnetic beads (Illumina), and cDNA
546 libraries for sequencing were generated as described[76]. Sequencing was performed on the
547 Illumina HiSeq 2000 platform at the Beijing Genomic Institute, Shenzhen, China. Raw reads
548 were filtered to remove adapters and low-quality reads (those with >5% unknown
549 nucleotides, or those in which greater than 20% of base calls had quality scores ≤10). After
550 filtering, both libraries generated reads with >99.9% of nucleotides attaining Q20 quality
551 score. Transcriptome *de novo* assembly was performed with Trinity[77]. For each library,

552 contigs were assembled into Unigenes; Unigenes from both libraries were then combined,
553 yielding a total of 45,553 predicted coding region sequences with a mean length of 1011 nt.
554 The combined Unigenes were then interrogated for homology to *AtKAI2*, *AtD14*, *AtDLK2*,
555 *AtSTH7* and *AtCACS* using BLASTn searches.

556 **Phylogenetic analysis and protein sequence analysis**

557 *KAI2* and *D14* homologues in *Brassica* species were identified from BLAST searches using
558 Arabidopsis coding sequences as a query. Additional sequences were sampled from an
559 existing phylogenetic analysis[33]. Multiple sequence alignments were performed using MAFFT
560 plugin implemented in Geneious R10 (Biomatters). The alignment was trimmed slightly at the
561 5′ end to remove non-aligned regions of monocot D14 sequences. Maximum likelihood
562 phylogenies were generated using PHYML (GTR +G +I substitution model, combination of NNI
563 and SPR search, and 100 bootstraps). The choice of substitution model was guided by Smart
564 Model Selection in PhyML[78] (http://www.atgc-montpellier.fr/sms). A list of all sequences, and
565 their sources, is provided in Supplementary Table 1.

566 For analysis of BtKAI2c using PROVEAN, all ten positions that were unique to BtKAI2c when
567 compared to AtKAI2 and other *Brassica* sequences (Supplementary Fig. 3) were manually
568 edited to the consensus identity at each position. This was the "reverted" sequence against
569 which each of the native BtKAI2c amino acid substitutions were tested using the PROVEAN
570 server (http://provean.jcvi.org) using the default cut-off score of -2.5. The PROVEAN output is
571 presented in Supplementary Table 2.

572 KAI2 homologues from The 1000 Plant Transcriptomes Project[50] were identified by TBLASTN
573 searches using *Arabidopsis thaliana* KAI2 protein sequence as a query against "onekp
574 database v5", which comprises 1328 individual samples (https://db.cngbdb.org). The search
575 was configured to return a maximum of 500 target sequences with Expect values <0.01, using
576 a word size of 6 and the BLOSUM62 scoring matrix. Nucleotide sequences were collated, and
577 open reading frames (ORFs) >700 nucleotides were predicted in Geneious software. A single
578 ORF was selected for each sequence; those with multiple ORFs were screened manually for
579 the one with homology to AtKAI2. The ORFs were translated and aligned using the MAFFT
580 multiple alignment tool. A few misaligned or mis-translated sequences were removed from
581 the alignment manually, resulting in a dataset of 476 unique sequences. Extracted parts of this
582 alignment centred on positions 96 and 189 were used to generate WebLogos

583 (http://weblogo.threeplusone.com[79]). A list of identified KAI2 homologues and their
584 positional analysis are presented in Supplementary Table 3.

585 ### Protein homology modelling

586 KAI2 structures were modelled using the SWISS-MODEL server
587 (https://swissmodel.expasy.org) using the alignment mode[80] and the *Arabidopsis thaliana*
588 KAI2 structure 3w06 as a template[54]. Figures of protein structure and homology models were
589 generated using PyMOL v1.3 (Schrödinger LLC). Cavity surfaces were visualised using the
590 "cavities & pockets (culled)" setting in PyMOL and a cavity detection cut-off value of four
591 solvent radii. Cavity volumes were calculated using the CASTp server v3.0[81]
592 (http://sts.bioe.uic.edu/castp) with a probe radius of 1.4Å. Values indicate Connolly's solvent-
593 excluded volumes. Cavities were inspected visually using Chimera v1.12
594 (https://www.cgl.ucsf.edu/chimera/). For both BtKAI2a and BtKAI2b models, CASTp
595 erroneously included surface residues near the primary pocket entrance in the calculation of
596 the pocket volumes. This issue was resolved by the artificial placement of a free alanine
597 residue adjacent to the cavity entrance, as described previously[82].

598 ### Protein expression and purification

599 BtKAI2 proteins were generated as N-terminal 6×HIS-SUMO fusion proteins. All proteins were
600 expressed in BL21 Rosetta DE3 pLysS cells (Novagen) and purified using IMAC as described in
601 detail previously[41].

602 ### Differential scanning fluorimetry

603 DSF was performed in 384-well format and thermal shifts were quantified as described
604 previously[41]. Reactions (10 µL) contained 20 µM protein, 20 mM HEPES pH 7.5, 150 mM NaCl,
605 1.25% (v/v) glycerol, 5× SYPRO Tangerine dye (Molecular Probes) and varying
606 concentrations of ligand that resulted in a final concentration of 5% (v/v) acetone.

607 ### Statistical analysis

608 Data were analysed using one- or two-way ANOVA (α = 0.05, with Tukey's multiple
609 comparisons test). For Figure 5a, in which data from three experimental replicates were
610 combined, data were analysed using a mixed effects model with experimental replicate as a
611 random effect, and genotype and treatment as fixed effects. Prior to ANOVA, germination data

612 were arcsine-transformed, and gene expression data were log-transformed. Tests were
613 implemented in GraphPad Prism version 7.0 or 8.0 (GraphPad Software, graphpad.com).

614 DATA AVAILABILITY

615 All relevant data and materials are available from the authors upon reasonable request.
616 Sequence data are available at NCBI Genbank under the following accessions: *BtKAI2a*,
617 MG783328; *BtKAI2b*, MG783329; *BtKAI2c*, MG783330; *BtD14a*, MG783331; *BtD14b*,
618 MG783332; *BtDLK2*, MG783333; *BtSTH7*, MK756121; *BtCACS*, MK756122. Raw RNA sequence
619 data from *Brassica tournefortii* seed and seedlings are available in the NCBI SRA database
620 under accession SRP128835. The original data underlying the following figures are provided
621 as a Source Data file: Figs 1a,b,d,e; 2b,e; 3a–d; 4a–d; 5a–c; Supplementary Figs 1b–f; 2; 3; 4c;
622 5b–d; 6b,c,e; 7; 8a–e; 9a–c; 10a; 11; 12; 13. A reporting summary for this Article is available as
623 a Supplementary Information File.

624 REFERENCES

625 1. Early R, *et al.* Global threats from invasive alien species in the twenty-first century and
626 national response capacities. *Nat Commun* **7**, 12485-12485 (2016).

627
628 2. Flematti GR, *et al.* Karrikin and cyanohydrin smoke signals provide clues to new
629 endogenous plant signaling compounds. *Molecular Plant* **6**, 29-37 (2013).

630
631 3. Keeley JE, Pausas JG. Evolution of 'smoke' induced seed germination in pyroendemic
632 plants. *South African Journal of Botany* **115**, 251-255 (2018).

633
634 4. Flematti GR, Ghisalberti EL, Dixon KW, Trengove RD. Identification of Alkyl Substituted
635 2 H-Furo[2,3- c]pyran-2-ones as Germination Stimulants Present in Smoke. *Journal of*
636 *Agricultural and Food Chemistry* **57**, 9475-9480 (2009).

637
638 5. Hrdlička J, *et al.* Quantification of karrikins in smoke water using ultra-high
639 performance liquid chromatography-tandem mass spectrometry. *Plant Methods* **15**, 81
640 (2019).

641
642 6. Long RL, Stevens JC, Griffiths EM, Adamek M, Powles SB, Merritt DJ. Detecting
643 karrikinolide responses in seeds of the Poaceae. *Australian Journal of Botany* **59**, 610-
644 620 (2011).

7. Milberg P, Lamont BB. Fire enhances weed invasion of roadside vegetation in southwestern Australia. *Biological Conservation* **73**, 45-49 (1995).

8. Flematti GR, Goddard-Borger ED, Merritt DJ, Ghisalberti EL, Dixon KW, Trengove RD. Preparation of 2H-furo[2,3-c]pyran-2-one derivatives and evaluation of their germination-promoting activity. *Journal of Agricultural and Food Chemistry* **55**, 2189-2194 (2007).

9. Long RL*, et al.* Seeds of Brassicaceae weeds have an inherent or inducible response to the germination stimulant karrikinolide. *Annals of Botany* **108**, 933-944 (2011).

10. Stevens J, Merritt D, Flematti G, Ghisalberti E, Dixon K. Seed germination of agricultural weeds is promoted by the butenolide 3-methyl-2H-furo[2,3-c]pyran-2-one under laboratory and field conditions. *Plant Soil* **298**, 113-124 (2007).

11. Nelson DC, Flematti GR, Riseborough J-A, Ghisalberti EL, Dixon KW, Smith SM. Karrikins enhance light responses during germination and seedling development in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences* **107**, 7095-7100 (2010).

12. Kulkarni MG, Sparg SG, Light ME, Staden JV. Stimulation of Rice (Oryza sativa L.) Seedling Vigour by Smoke‐water and Butenolide. *J Agron Crop Sci* **192**, 395-398 (2006).

13. Staden Jv, Sparg SG, Kulkarni MG, Light ME. Post-germination effects of the smoke-derived compound 3-methyl-2H-furo[2,3-c]pyran-2-one, and its potential as a preconditioning agent. *Field Crop Res* **98**, 98-105 (2006).

14. Bangle DN, Walker LR, Powell EA. Seed germination of the invasive plant Brassica tournefortii (Sahara mustard) in the Mojave Desert. *Western North American Naturalist* **68**, 334-342 (2008).

15. Brooks ML*, et al.* Effects of Invasive Alien Plants on Fire Regimes. *BioScience* **54**, 677 (2004).

16. Sanders R, Minnich R. Brassica tournefortii. In: *Invasive Plants of California's Wildlands* (eds Bossard CC, Randall JM, Hoshovsky MC). University of California Press (2000).

17. Schiermeier Q. Pall hangs over desert's future as alien weeds fuel wildfires. *Nature* **435**, 724 (2005).

686
687 18. Steers RJ. Invasive Plants, Fire Succession, and Restoration of Creosote Bush Scrub in
688 Southern California.). UNIVERSITY OF CALIFORNIA RIVERSIDE (2008).

689
690 19. Sun X-D, Ni M. HYPOSENSITIVE TO LIGHT, an Alpha/Beta Fold Protein, Acts
691 Downstream of ELONGATED HYPOCOTYL 5 to Regulate Seedling De-Etiolation.
692 *Molecular Plant* **4**, 116-126 (2011).

693
694 20. Waters MT*, et al.* Specialisation within the DWARF14 protein family confers distinct
695 responses to karrikins and strigolactones in Arabidopsis. *Development* **139**, 1285-1295
696 (2012).

697
698 21. Durvasula A*, et al.* African genomes illuminate the early history and transition to
699 selfing in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences* **114**,
700 5213-5218 (2017).

701
702 22. Nelson DC*, et al.* Karrikins discovered in smoke trigger Arabidopsis seed germination
703 by a mechanism requiring gibberellic acid synthesis and light. *Plant Physiology* **149**,
704 863-873 (2009).

705
706 23. Flematti GR, Ghisalberti EL, Dixon KW, Trengove RD. A compound from smoke that
707 promotes seed germination. *Science* **305**, 977 (2004).

708
709 24. Waters MT, Gutjahr C, Bennett T, Nelson DC. Strigolactone Signaling and Evolution.
710 *Annual Review of Plant Biology* **68**, 291-322 (2017).

711
712 25. Yao R, Chen L, Xie D. Irreversible strigolactone recognition: a non-canonical
713 mechanism for hormone perception. *Current Opinion in Plant Biology* **45**, 155-161
714 (2018).

715
716 26. Scaffidi A*, et al.* Exploring the molecular mechanism of karrikins and strigolactones.
717 *Bioorganic & Medicinal Chemistry* **22**, 3743-3746 (2012).

718
719 27. Zwanenburg B, Pospíšil T. Structure and Activity of Strigolactones: New Plant
720 Hormones with a Rich Future. *Molecular Plant* **6**, 38-62 (2013).

721
722 28. de Saint Germain A*, et al.* An histidine covalent receptor and butenolide complex
723 mediates strigolactone perception. *Nature Chemical Biology* **12**, 787-794 (2016).

724
725 29. Seto Y*, et al.* Strigolactone perception and deactivation by a hydrolase receptor
726 DWARF14. *Nat Commun* **10**, 191 (2019).

727
728 30. Shabek N, Ticchiarelli F, Mao H, Hinds TR, Leyser O, Zheng N. Structural plasticity of
729 D3–D14 ubiquitin ligase in strigolactone signalling. *Nature* **563**, 652-656 (2018).

730
731 31. Yao R, *et al.* DWARF14 is a non-canonical hormone receptor for strigolactone. *Nature*
732 **536**, 469-473 (2016).

733
734 32. Marzec M, Brewer P. Binding or Hydrolysis? How Does the Strigolactone Receptor
735 Work? *Trends in Plant Science* **24**, 571-574 (2019).

736
737 33. Bythell-Douglas R, *et al.* Evolution of strigolactone receptors by gradual neo-
738 functionalization of KAI2 paralogues. *BMC Biology* **15**, 52 (2017).

739
740 34. Conn CE, Nelson DC. Evidence that KARRIKIN-INSENSITIVE2 (KAI2) Receptors may
741 Perceive an Unknown Signal that is not Karrikin or Strigolactone. *Frontiers in Plant*
742 *Science* **6**, 1219 (2015).

743
744 35. Li W, *et al.* The karrikin receptor KAI2 promotes drought resistance in Arabidopsis
745 thaliana. *PLoS Genetics* **13**, e1007076 (2017).

746
747 36. Sun YK, Flematti GR, Smith SM, Waters MT. Reporter Gene-Facilitated Detection of
748 Compounds in Arabidopsis Leaf Extracts that Activate the Karrikin Signaling Pathway.
749 *Frontiers in Plant Science* **7**, 1799 (2016).

750
751 37. Liu S, *et al.* The Brassica oleracea genome reveals the asymmetrical evolution of
752 polyploid genomes. *Nat Commun* **5**, 3930 (2014).

753
754 38. Lysak MA, Koch MA, Pecinka A, Schubert I. Chromosome triplication found across the
755 tribe Brassiceae. *Genome Res* **15**, 516-525 (2005).

756
757 39. Wang X, *et al.* The genome of the mesopolyploid crop species Brassica rapa. *Nature*
758 *Genetics* **43**, 1035-1039 (2011).

759
760 40. Waters MT, Scaffidi A, Flematti G, Smith SM. Substrate-Induced Degradation of the α/β-
761 Fold Hydrolase KARRIKIN INSENSITIVE2 Requires a Functional Catalytic Triad but Is
762 Independent of MAX2. *Molecular Plant* **8**, 814-817 (2015).

763
764 41. Waters MT, Scaffidi A, Moulin SLY, Sun YK, Flematti GR, Smith SM. A Selaginella
765 moellendorffii Ortholog of KARRIKIN INSENSITIVE2 Functions in Arabidopsis
766 Development but Cannot Mediate Responses to Karrikins or Strigolactones. *Plant Cell*
767 **27**, 1925-1944 (2015).

768
769 42. Kagale S, *et al.* TMV-Gate vectors: Gateway compatible tobacco mosaic virus based
770 expression vectors for functional analysis of proteins. *Scientific Reports* **2**, 874 (2012).

771
772 43. Yao J, *et al.* An allelic series at the KARRIKIN INSENSITIVE 2 locus of Arabidopsis
773 thaliana decouples ligand hydrolysis and receptor degradation from downstream
774 signalling. *Plant J* **96**, 75-89 (2018).

775
776 44. Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino
777 acid substitutions and indels. *Bioinformatics* **31**, 2745-2747 (2015).

778
779 45. Abe S, *et al.* Carlactone is converted to carlactonoic acid by MAX1 in Arabidopsis and its
780 methyl ester can directly interact with AtD14 in vitro. *Proceedings of the National*
781 *Academy of Sciences of the United States of America* **111**, 18084-18089 (2014).

782
783 46. Hamiaux C, *et al.* DAD2 Is an α/β Hydrolase Likely to Be Involved in the Perception of
784 the Plant Branching Hormone, Strigolactone. *Current Biology* **22**, 2032-2036 (2012).

785
786 47. Hamiaux C, *et al.* Inhibition of strigolactone receptors by N-phenylanthranilic acid
787 derivatives: structural and functional insights. *Journal of Biological Chemistry* **293**,
788 6530-6543 (2018).

789
790 48. Végh A, *et al.* Comprehensive Analysis of DWARF14-LIKE2 (DLK2) Reveals Its
791 Functional Divergence from Strigolactone-Related Paralogs. *Frontiers in Plant Science*
792 **8**, 1641 (2017).

793
794 49. Scaffidi A, *et al.* Strigolactone hormones and their stereoisomers signal through two
795 related receptor proteins to induce different physiological responses in Arabidopsis.
796 *Plant Physiology* **165**, 1221-1232 (2014).

797
798 50. Leebens-Mack JH, *et al.* One thousand plant transcriptomes and the phylogenomics of
799 green plants. *Nature* **574**, 679-685 (2019).

800
801 51. Richardson DM, Wilgen BWV, Mitchell DT. Aspects of the reproductive ecology of four
802 australian Hakea species (Proteaceae) in South Africa. *Oecologia* **71**, 345-354 (1987).

803
804 52. Guo Y, Zheng Z, La Clair JJ, Chory J, Noel JP. Smoke-derived karrikin perception by the
805 α/β-hydrolase KAI2 from Arabidopsis. *Proceedings of the National Academy of Sciences*
806 *of the United States of America* **110**, 8284-8289 (2013).

807

808 53. Xu Y, *et al.* Structural basis of unique ligand specificity of KAI2-like protein from
809 parasitic weed Striga hermonthica. *Scientific Reports* **6**, 31386 (2016).

810
811 54. Kagiyama M, *et al.* Structures of D14 and D14L in the strigolactone and karrikin
812 signaling pathways. *Genes to Cells* **18**, 147-160 (2013).

813
814 55. Toh S, Holbrook-Smith D, Stokes ME, Tsuchiya Y, McCourt P. Detection of parasitic
815 plant suicide germination compounds using a high-throughput Arabidopsis HTL/KAI2
816 strigolactone perception system. *Chem Commun* **21**, 988-998 (2014).

817
818 56. Lee I, *et al.* A missense allele of KARRIKIN-INSENSITIVE2 impairs ligand-binding and
819 downstream signaling in Arabidopsis thaliana. *Journal of Experimental Botany* **69**,
820 3609-3623 (2018).

821
822 57. Yao J, Waters MT. Perception of karrikins by plants: a continuing enigma. *Journal of*
823 *Experimental Botany*, 10.1093/jxb/erz1548 (2019).

824
825 58. Conn CE, *et al.* Convergent evolution of strigolactone perception enabled host detection
826 in parasitic plants. *Science* **349**, 540-543 (2015).

827
828 59. Toh S, *et al.* Structure-function analysis identifies highly sensitive strigolactone
829 receptors in Striga. *Science* **350**, 203-207 (2015).

830
831 60. Xu Y, *et al.* Structural analysis of HTL and D14 proteins reveals the basis for ligand
832 selectivity in Striga. *Nat Commun* **9**, 3947 (2018).

833
834 61. Bürger M, *et al.* Structural Basis of Karrikin and Non-natural Strigolactone Perception
835 in *Physcomitrella patens*. *Cell Reports* **26**, 855-865 (2019).

836
837 62. Shimada A, *et al.* Structural basis for gibberellin recognition by its receptor GID1.
838 *Nature* **456**, 520-523 (2008).

839
840 63. Zhang Y, *et al.* Rice cytochrome P450 MAX1 homologs catalyze distinct steps in
841 strigolactone biosynthesis. *Nature Chemical Biology* **10**, 1028-1033 (2014).

842
843 64. Challis RJ, Hepworth J, Mouchel C, Waites R, Leyser O. A Role for MORE AXILLARY
844 GROWTH1 (MAX1) in Evolutionary Diversity in Strigolactone Signaling Upstream of
845 MAX2. *Plant Physiology* **161**, 1885-1902 (2013).

846

847 65. Brewer PB, *et al. LATERAL BRANCHING OXIDOREDUCTASE* acts in the final stages of
848 strigolactone biosynthesis in *Arabidopsis*. *Proceedings of the National Academy of*
849 *Sciences* **113**, 6301-6306 (2016).

850
851 66. Yoneyama K, *et al.* Which are the major players, canonical or non-canonical
852 strigolactones? *Journal of Experimental Botany* **69**, 2231-2239 (2018).

853
854 67. Panchy N, Lehti-Shiu MD, Shiu S-H. Evolution of gene duplication in plants. *Plant*
855 *Physiology* **171**, 2294-2316 (2016).

856
857 68. Wang L, Waters MT, Smith SM. Karrikin-KAI2 signalling provides Arabidopsis seeds
858 with tolerance to abiotic stress and inhibits germination under conditions
859 unfavourable for seedling establishment. *New Phytologist* **219**, 605-618 (2018).

860
861 69. Goddard-Borger ED, Ghisalberti EL, Stick RV. Synthesis of the germination stimulant 3-
862 methyl-2H-furo[2,3-c]pyran-2-one and analogous compounds from carbohydrates.
863 *European Journal of Organic Chemistry*, 3925-3934 (2007).

864
865 70. Mangnus E, Vanvliet LA, Vadenput D, Zwanenburg B. Structural modifications of strigol
866 analogs - influence of the B and C rings on the bioactivity of the germination stimulant
867 GR24. *Journal of Agricultural and Food Chemistry* **40**, 1222-1229 (1992).

868
869 71. Gorecki MJ, Long RL, Flematti GR, Stevens JC. Parental environment changes the
870 dormancy state and karrikinolide response of Brassica tournefortii seeds. *Annals of*
871 *Botany* **109**, 1369-1378 (2012).

872
873 72. Waters MT, Smith SM. KAI2- and MAX2-mediated responses to karrikins and
874 strigolactones are largely independent of HY5 in Arabidopsis seedlings. *Molecular*
875 *Plant* **6**, 63-75 (2013).

876
877 73. Kagale S, *et al.* TMV-Gate vectors: gateway compatible tobacco mosaic virus based
878 expression vectors for functional analysis of proteins. *Scientific Reports* **2**, 874 (2012).

879
880 74. Curtis MD, Grossniklaus U. A Gateway Cloning Vector Set for High-Throughput
881 Functional Analysis of Genes in Planta. *Plant Physiology* **133**, 462-469 (2003).

882
883 75. Flematti GR, Scaffidi A, Dixon KW, Smith SM, Ghisalberti EL. Production of the seed
884 germination stimulant karrikinolide from combustion of simple carbohydrates. *Journal*
885 *of Agricultural and Food Chemistry* **59**, 1195-1198 (2011).

886

887 76. Liang C, Liu X, Yiu S-M, Lim BL. De novo assembly and characterization of Camelina
888 sativa transcriptome by paired-end sequencing. *BMC Genomics* **14**, 146 (2013).

889
890 77. Grabherr MG*, et al.* Full-length transcriptome assembly from RNA-Seq data without a
891 reference genome. *Nature Biotechnology* **29**, 644-652 (2011).

892
893 78. Lefort V, Longueville J-E, Gascuel O. SMS: Smart Model Selection in PhyML. *Molecular*
894 *Biology and Evolution* **34**, 2422-2424 (2017).

895
896 79. Crooks GE, Hon G, Chandonia J-M, Brenner SE. WebLogo: A Sequence Logo Generator.
897 *Genome Res* **14**, 1188-1190 (2004).

898
899 80. Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T. Protein structure
900 homology modeling using SWISS-MODEL workspace. *Nature Protocols* **4**, 1-13 (2009).

901
902 81. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J. CASTp: computed atlas of
903 surface topography of proteins with structural and topographical mapping of
904 functionally annotated residues. *Nucleic Acids Research* **34**, W116-W118 (2006).

905
906 82. Lopez-Obando M*, et al.* Structural modelling and transcriptional responses highlight a
907 clade of PpKAI2-LIKE genes as candidate receptors for strigolactones in Physcomitrella
908 patens. *Planta* **243**, 1441-1453 (2016).

909

910

911 ## ACKNOWLEDGEMENTS

912 This work was supported by funding from the Australian Research Council (DP130103646
913 to SMS and GRF; DP140104567 to SMS; FT150100162 to MTW, DP160102888 to GRF and
914 MTW). Y.K.S was recipient of a Research Training Program Scholarship from the Australian
915 Government. We thank Dr Rowena Long for the original provision of *B. tournefortii* seeds, Dr
916 Rohan Bythell-Douglas for advice on homology modelling, and Mr Yongjie Meng with
917 assistance with RNA extractions. We are grateful to Dr Kevin Rozwadowski (Agriculture and
918 Agri-Food Canada) for the provision of pSKI106.

919 ## CONTRIBUTIONS

920 Y.K.S., S.M.S., C.S.B., G.R.F. and M.T.W. conceived and designed the research. Y.K.S. and
921 M.T.W. performed the majority of experiments with assistance from J.Y. (hypocotyl
922 elongation assays and screening of transgenic lines), A.S. (chemical synthesis and seed
923 germination assays), K.T.M. and S.F.D. (protein expression and purification). Y.K.S. and
924 M.T.W. analysed the data. Y.K.S., S.M.S. and M.T.W. wrote the manuscript.

925 ## ADDITIONAL INFORMATION

926 **Supplementary Figure 1** Germination of *Brassica tournefortii* seed treated with karrikins

927 **Supplementary Figure 2** Extended phylogeny of eu-KAI2 and D14 proteins in angiosperms

928 **Supplementary Figure 3** Alignment of Brassica KAI2 sequences

929 **Supplementary Figure 4** BtD14a is functionally homologous to AtD14

930 **Supplementary Figure 5** The *GFP-BtKAI2c* transgene is faithfully transcribed in Arabidopsis

931 **Supplementary Figure 6** The R98V mutation restores stability of BtKAI2c

932 **Supplementary Figure 7** Germination profiles of transgenic Arabidopsis seeds expressing
933 native BtKAI2 homologues

934 **Supplementary Figure 8** BtKAI2 homology models and SDS-PAGE of SUMO-BtKAI2 fusion
935 proteins used for DSF

936 **Supplementary Figure 9** BtKAI2a and BtKAI2b do not respond to karrikins in DSF assays

937 **Supplementary Figure 10** Stable transgenic expression of BtKAI2 valine–leucine double
938 exchange proteins in Arabidopsis

939 **Supplementary Figure 11** Three experimental replicates of hypocotyl elongation assays
940 with BtKAI2 Arabidopsis transgenics (double exchange of residues 98 and 191)

941 **Supplementary Figure 12** Germination profiles of transgenic Arabidopsis seed expressing
942 BtKAI2 homologues

943 **Supplementary Figure 13** Three experimental replicates of hypocotyl elongation assays
944 with BtKAI2b transgenics (individual exchange of residues 98 and 191)

945 **Supplementary Table 1** List of sequences identified in this study

946 **Supplementary Table 2** Analysis of BtKAI2c sequence using PROVEAN

947 **Supplementary Table 3** Analysis of KAI2 homologues identified from The 1000 Plant
948 Transcriptomes project

949 **Supplementary Table 4** List of oligonucleotides

950 **Supplementary References**

951 **Reporting Summary**

952 **Source Data file**

953 ## FIGURE LEGENDS

954 **Figure 1. *Brassica tournefortii* is highly sensitive to KAR_1, the major karrikin analogue**
955 **isolated from plant-derived smoke**
956 **a,** Germination responses of *B. tournefortii* seed to KAR_1 and KAR_2. Data are cumulative
957 germination after 11 days (mean ± SE; n = 3 biological replicates per treatment, ≥35 seeds per
958 replicate). **b**, Levels of karrikin-responsive transcripts *BtDLK2* and *BtSTH7* in *B. tournefortii*
959 seed. Seed were imbibed for 24 hours in the dark supplemented with KAR_1 and KAR_2.
960 Transcripts were normalised to *BtCACS* reference transcripts. Data are means ± SE, n = 3
961 biological replicates, ≥50 seeds per replicate. **c**, **d**, Hypocotyl elongation responses of *B.*
962 *tournefortii* seedlings grown for four days under continuous red light on water-saturated
963 glass filter paper containing 0.1% acetone (mock), KAR_1 or KAR_2. Data are means ± 95% CI of
964 n = 18 to 24 seedlings. Scale bar: 10 mm. **e**, Levels of *BtDLK2* transcripts in *B. tournefortii*
965 seedlings grown under the same conditions as for **d**. Data are means ± SE, n = 3 biological
966 replicates, ≥20 seedlings per replicate. In all panels, asterisks denote significance levels
967 (ANOVA) between indicated conditions: * $P<0.05$; ** $P<0.01$; *** $P<0.001$; **** $P<0.0001$.
968 Source data are provided as a Source Data file.

969 **Figure 2. Two differentially expressed *B. tournefortii KAI2* homologues are functional in**
970 ***Arabidopsis***
971 **a**, Maximum likelihood phylogeny of KAI2 homologues in the Brassicaceae, based on
972 nucleotide data. Node values represent bootstrap support from 100 replicates. A eu-KAI2
973 sequence from *Gossypium raimondii* (Malvaceae) serves as an outgroup. Tree shown is a
974 subset of a larger phylogeny in Supplementary Fig. 2. **b**, Transcript levels of the three *BtKAI2*
975 homologues in *B. tournefortii* seeds imbibed for 24 h (left) and four-day-old seedlings (right)
976 treated with 0.1% acetone or with 1 μM KAR_1 for 24 h. **c**, **d**, Seedling and rosette phenotypes
977 of two independent transgenic lines of Arabidopsis homozygous for *KAI2pro:GFP-BtKAI2*
978 transgenes. Scale bars: 5 mm (**c**); 50 mm (**d**). **e**, Immunoblots of soluble proteins challenged
979 with antibodies against KAI2 (upper panel), GFP (middle panel) or actin as a loading control
980 (lower panel). Anti-KAI2 detects both the native AtKAI2 protein (30 kDa) and the GFP-BtKAI2
981 fusion proteins (58 kDa) but detects BtKAI2b relatively poorly. Non-specific bands are
982 marked with asterisks. Protein was isolated from pools of approximately fifty 7-day-old
983 seedlings. Source data are provided as a Source Data file.

984 **Figure 3. Functional divergence between BtKAI2 homologues**
985 **a**, Germination responses of primary dormant *Arabidopsis* seed homozygous for *KAI2pro:GFP-*
986 *BtKAI2* transgenes in the *kai2-2* background. Germination values were determined 120 h after
987 sowing. Extended germination curves are shown in Supplementary Fig. 7. Data are means ±
988 SE, n = 3 independent seed batches, 75 seed per batch. **b**, Levels of *KUF1* transcripts in
989 *KAI2pro:GFP-BtKAI2* seeds treated with 1 μM KAR_1 or KAR_2. Expression was normalised to
990 *CACS* reference transcripts and scaled to the value for mock-treated seed within each
991 genotype. Data are means ± SE of n = 3 biological replicates. **c**, Hypocotyl elongation
992 responses of *KAI2pro:GFP-BtKAI2* seedlings treated with KAR_1 or KAR_2. Data are means ± SE
993 of n = 3 biological replicates, 12–18 seedlings per replicate. **d**, Levels of *DLK2* transcripts in 8-
994 day-old *KAI2pro:GFP-BtKAI2* seedlings treated with KAR_1 or KAR_2 for eight hours. Expression
995 was normalised to *CACS* reference transcripts and scaled to the value for mock-treated
996 seedlings within each genotype. Data are means ± SE of n = 3 biological replicates. Pairwise
997 significant differences: * $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$; ns, $P > 0.05$ (ANOVA). Source data
998 are provided as a Source Data file.

999 **Figure 4. Two residues account for ligand specificity between BtKAI2a and BtKAI2b**
1000 **a**, DSF curves of SUMO-BtKAI2a and SUMO-BtKAI2b fusion proteins treated with 0-50 μM
1001 $GR24^{ent\text{-}5DS}$, a KAI2-bioactive ligand. Each curve is the average of three sets of reactions, each

1002 comprising four technical replicates. Insets plot the minimum value of –(dF/dT) at the melting
1003 point of the protein as determined in the absence of ligand (means ± SE, n = 3). Significant
1004 differences from untreated control: * $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$ (ANOVA). **b**, DSF curves
1005 of SUMO-BtKAI2a$^{L98;L191}$ and SUMO-BtKAI2b$^{V98;V191}$ fusion proteins treated with 0-50 μM
1006 GR24$^{ent\text{-}5DS}$. **c**, Hypocotyl elongation responses of Arabidopsis expressing GFP-BtKAI2a$^{L98;L191}$
1007 and GFP-BtKAI2b$^{V98;V191}$ fusion proteins and treated with KAR_1 or KAR_2. Data are a summary
1008 of three experimental replicates performed on separate occasions, each comprising 25-40
1009 seedlings per genotype/treatment combination. Data for each replicate are shown in
1010 Supplementary Figure 11. Error bars are SE, n = 3 experimental replicates; each dot
1011 corresponds to the mean value derived from each replicate. Asterisks denote significant
1012 differences: * P<0.05, ** P<0.01, *** P<0.001, **** P<0.0001 (linear mixed model with
1013 experimental replicate as a random effect; specific pairwise comparisons using Tukey's HSD
1014 correction). **d**, Levels of *DLK2* transcripts in the same transgenic lines as above treated with 1
1015 μM KAR_1, 1 μM KAR_2, or 0.1% acetone and harvested 8 hours later. Expression was
1016 normalised to *CACS* reference transcripts. Data are means ± SE of n = 3 pools of ~50 seedlings
1017 treated in parallel. Asterisks denote significant differences as above (two-way ANOVA;
1018 specific pairwise comparisons using Tukey's HSD correction). Source data are provided as a
1019 Source Data file.

1020 **Figure 5. Diversity at positions 96 and 189 in KAI2 homologues across angiosperms**
1021 a, Sequence logos generated from 476 aligned KAI2 homologues sampled from eudicots,
1022 monocots and magnoliids and centred on position 96 (left) and 189 (right). b, Species tree of
1023 the Saxifragales based upon 410 single-copy nuclear genes estimated using the ASTRAL gene
1024 tree method[50]. Branch support values are shown only where less than 1. The four-letter taxon
1025 prefix corresponds to the 1000 Plant Transcriptomes project Sample ID. Taxa are highlighted
1026 by colour according to identity of amino acids of KAI2 orthologues at positions 96 and 189
1027 respectively: black, VV; blue, LV; red, LL; magenta, other. Postulated ancestral identities at
1028 specific nodes are indicated with filled circles. c, Hypocotyl elongation responses to karrikins
1029 in three independent transgenic lines homozygous for *GFP-BtKAI2a$^{L98;V191}$* and three lines for
1030 *GFP-BtKAI2b$^{V98;L191}$*. Data shown are a summary of three experimental replicates performed
1031 on separate occasions, each comprising ~20 seedlings per genotype/treatment combination.
1032 Data from each replicate are presented in Supplementary Figure 13. Error bars are SE, n = 3
1033 experimental replicates; each dot corresponds to the mean value derived from each replicate.
1034 Asterisks denote significant differences: * P<0.05, ** P<0.01, *** P<0.001, **** P<0.0001

1035 (linear mixed model with experimental replicate as a random effect; specific pairwise
1036 comparisons using Tukey's HSD correction). Source data are provided as a Source Data file.

1037 SUPPLEMENTARY FIGURE LEGENDS

1038 **Supplementary Figure 1. Germination of *Brassica tournefortii* seed treated with**
1039 **karrikins**
1040 **a**, Structures of KAR_1, KAR_2 and two enantiomers of GR24. **b**, Germination response of the
1041 "Merridin" batch of *B. tournefortii* seed treated with KAR_1 and KAR_2 after three days. **c-d**,
1042 Germination response of the "Perth" batch to KAR_1 (**c**) and KAR_2 (**d**). Data in Figure 1a are
1043 derived from the data shown in **c**. **e-f**, Uptake of KAR_1 and KAR_2 by imbibed seed, as
1044 determined by GC-MS. All error bars are mean ± SE of n = 3 batches of 75 seed (**b-d**) or 3
1045 samples of 40 mg (**e**) or 20 mg (**f**) seed as described in Methods. Source data are provided as a
1046 Source Data file.

1047 **Supplementary Figure 2. Extended phylogeny of eu-KAI2 and D14 proteins in**
1048 **angiosperms**
1049 Maximum likelihood phylogeny of KAI2 and D14 homologues in the Brassicaceae and
1050 monocots, based on nucleotide data. Node values represent bootstrap support from 100
1051 replicates. A KAI2 sequence from *Selaginella moellendorffii* (SmKAI2a) serves as an outgroup
1052 for the eu-KAI2 clade. Source data are provided as a Source Data file.

1053 **Supplementary Figure 3. Alignment of Brassica KAI2 sequences**
1054 Full length protein coding regions of *KAI2* homologues from four *Brassica* species (*Brassica*
1055 *tournefortii*, *B. rapa*, *B. nigra* and *B. oleracea*) were translated from database nucleic acid
1056 sequences and aligned to *Arabidopsis thaliana* KAI2 using MAFFT[1] implemented in Geneious
1057 R10 software (Biomatters Ltd). Amino acid residues are coloured according to polarity:
1058 yellow, non-polar (G, A, V, L, I, F, W, M, P); green, polar & uncharged (S, T, C, Y, N, Q); red, polar
1059 & acidic (D, E); blue, polar & basic (K, R, H). Residues that are unique to BtKAI2c but
1060 otherwise invariant are highlighted with asterisks (*); residues 98 and 191 are highlighted
1061 with red boxes. Source data are provided as a Source Data file.

1062 **Supplementary Figure 4. BtD14a is functionally homologous to AtD14**
1063 Three independent transgenic lines of Arabidopsis *Atd14-1* were analysed for functional
1064 complementation of the mutant phenotype by an *AtD14pro:BtD14a* transgene. **a**, Rosette and

1065 leaf morphology at 31 days post-germination. **b**, Plant height and number of primary rosette
1066 branches at 45 days post-germination. **c**, Quantification of height and branching parameters, *n*
1067 = 10 plants per genotype. Scale bars: 50 mm. Source data are provided as a Source Data file.

1068 **Supplementary Figure 5. The GFP-BtKAI2c transgene is faithfully transcribed in**
1069 **Arabidopsis**

1070 **a**, Structure of the *AtKAI2pro:mGFP6-BtKAI2c* transgene. Primers used for RT-PCR are shown
1071 with arrows. The promoter and 5′UTR are derived from *At4g37470* (*AtKAI2*). attB1, Gateway
1072 recombination site that links *mGFP6* and *BtKAI2c* regions; *nos ter*, nopaline synthase
1073 terminator. Not drawn to scale.

1074 **b**, RT-PCR analysis of *GFP-BtKAI2c* transcripts after 35 cycles of amplification. Primers
1075 MW406 + BK010 target the transgene, while a second primer pair (MW275 + MW278) serves
1076 as a control and spans five introns of *At1g03055*. *kai2-2* serves as a non-transgenic control
1077 genotype. Templates: R, total RNA; D, genomic DNA; W, water only. RT, reverse transcriptase
1078 (Supercript III). DNA size standards (in base pairs) are indicated on the left, with anticipated
1079 PCR product sizes shown on the right and defined in the table.

1080 **c**, The RT-PCR product generated with proof-reading polymerase (Q5, New England Biolabs)
1081 and primers MW406 and BK010 was cloned into pCR4-TOPO (Life Technologies). Five
1082 dideoxy sequence traces were aligned against the *GFP-BtKAI2c* transgene reference. No
1083 disagreements with the reference sequence were observed. Red bars indicate trimmed
1084 regions of sequence traces to remove low quality data.

1085 d, Transient expression of BtKAI2a, BtKAI2b and BtKAI2c proteins in tobacco. Plasmids
1086 encoding N-terminal, c-myc-tagged proteins were transferred to Agrobacterium, and the
1087 resulting strains used to infiltrate tobacco leaves. After 96 h, samples were harvested in
1088 triplicate (two to three leaves per sample). Mock-treated leaves were transformed with a
1089 plasmid encoding a non-tagged protein. Sixty micrograms of total protein were separated by
1090 SDS-PAGE, blotted and challenged with anti-c-myc antibody (Genscript A00704). Band
1091 intensity was measured using ImageJ, and expression was normalised to intensity of the
1092 Rubisco large subunit (RbcL) band on the "stain-free" gel imaged under UV light. Error bars
1093 indicate SE, n = 3 replicates. Source data are provided as a Source Data file.

1094 **Supplementary Figure 6. The R98V mutation restores stability of BtKAI2c**

1095 **a**, Homology model of the native BtKAI2c protein. Highlighted in yellow are five residues and
1096 their corresponding mutations that were assessed for effects on BtKAI2c stability. The
1097 combination of E129D, D130V, E133Q and D173E in the hinge region between the lid domain
1098 (orange) and core domain (green) is the quadruple ("quad") mutant; combining this with
1099 R98V yields the quintuple ("quint") mutant. Red residues are the Ser-His-Asp catalytic triad. **b**,
1100 Transient expression of AtKAI2 and BtKAI2c variants in tobacco leaves following infiltration
1101 by *Agrobacterium* strains encoding GFP-KAI2 variants driven by the CaMV *35S* promoter.
1102 Three leaves were infiltrated on each of two different plants. After five days, 20 µg soluble
1103 protein was electrophoresed, blotted and challenged with antibodies against either anti-GFP
1104 (top) or anti-actin (bottom). **c**, Immunoblot of crude bacterial lysates expressing SUMO fusion
1105 proteins and challenged with anti-SUMO antibody. Negative values above each lane indicate
1106 $\log_{10}$ dilution factors of lysate; U, lysate from uninduced bacterial culture expressing SUMO-
1107 AtKAI2. **d**, SDS-PAGE analysis of protein purification runs of SUMO-BtKAI2c, -BtKAI2c^{R98V} and
1108 -BtKAI2c^{QUINT} harvested from 900 mL bacterial culture. Lys, clarified lysate; FT, column flow
1109 through after immobilisation on Co^{2+} affinity column; W, wash with 10 mM imidazole.
1110 Proteins were eluted with 200 mM imidazole in eight successive fractions. **e**, DSF curves of
1111 SUMO fusion proteins treated with 0-200 µM GR24$^{ent\text{-}5DS}$. Insets plot the minimum value of –
1112 (dF/dT) at the melting point of the protein as determined in the absence of ligand (means ±
1113 SE, n = 3). Significant differences from untreated control: * $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$
1114 (ANOVA). Source data are provided as a Source Data file.

1115 **Supplementary Figure 7. Germination profiles of transgenic Arabidopsis seeds**
1116 **expressing BtKAI2 homologues**

1117 Freshly harvested seed (three batches per genotype, each batch harvested from four plants)
1118 were removed from freezer storage, surface-sterilised and sown on 1% Phytagel
1119 supplemented with 0.1% acetone (mock), 1 µM KAR_1 or 1 µM KAR_2. Seed were incubated
1120 under constant light at 25 °C. Seed were examined for germination (radicle protrusion) 72 h
1121 after sowing and every 24 h thereafter. Data are means ± SE of three independent seed
1122 batches and 75 seed per batch. For each transgene, two independent, homozygous transgenic
1123 lines were analysed. Data presented in Figure 3 of the main manuscript are derived from
1124 these data. Source data are provided as a Source Data file.

1125 **Supplementary Figure 8. BtKAI2 homology models and SDS-PAGE of SUMO-BtKAI2**
1126 **fusion proteins used for DSF**

1127 **a-d**, Solved structure of AtKAI2 (PDB: 3w06, ref. 2), AtD14 (PDB: 4IH4, ref. 3) and predicted
1128 homology models of BtKAI2a and BtKAI2b. Coloured surfaces depict internal cavities; values
1129 indicate the volumes of the primary ligand-binding cavities, adjacent to the catalytic Ser-His-
1130 Asp residues. Also shown is a variable secondary pocket, to the left of the primary pocket in
1131 these images.
1132 **e**, To assess purity after affinity chromatography, five micrograms of each purified protein
1133 was electrophoresed on a 12% acrylamide gel containing 2,2,2-trichloroethanol and
1134 visualised under UV light. Protein size standards at 75 and 25 kDa (Bio-Rad Precision Plus
1135 Dual Colour) fluoresce strongly under UV light. Source data are provided as a Source Data file.

1136 **Supplementary Figure 9. BtKAI2a and BtKAI2b do not respond to karrikins in DSF**
1137 **assays**
1138 **a**, Differential scanning fluorimetry curves of SUMO-BtKAI2a and SUMO-BtKAI2b in presence
1139 of 0–200 μM KAR_1 or KAR_2. **b**, DSF responses of SUMO-tagged BtKAI2b (left) or untagged
1140 BtKAI2b (right) to 0–200 μM $GR24^{ent\text{-}5DS}$. **c**, DSF responses to 0–200 μM of the two
1141 enantiomers of GR24. Data are means of eight (**a** and **c**) or four (**b**) technical replicates at each
1142 concentration of ligand. Source data are provided as a Source Data file.

1143 **Supplementary Figure 10. Stable transgenic expression of BtKAI2 valine–leucine**
1144 **exchange proteins in Arabidopsis**
1145 **a**, Immunoblots of total soluble protein extracted from 7-day-old seedlings of independent
1146 transgenic lines segregating in a 3:1 ratio for hygromycin resistance. Transgene expression in
1147 six lines expressing *GFP-BtKAI2a*$^{L98;L191}$ (upper panels) and five expressing *GFP-BtKAI2b*$^{V98;V191}$
1148 (lower panels) were compared to a representative unmodified control (*GFP-BtKAI2a* 1F and
1149 *GFP-BtKAI2b* 9A respectively). Based on expression level two lines of each construct (3-5 and
1150 3-14; 6-4 and 6-13) were selected and brought to homozygosity for further experiments.
1151 Protein blots were challenged with anti-GFP antibody. Equal gel loading was assessed by
1152 imaging total protein prior to blotting; the RbcL band is shown. **b**, Rosette phenotypes of
1153 homozygous individuals expressing native and modified *GFP-BtKAI2* transgenes. Plants were
1154 25 days old and grown under long day conditions as described in Methods. Scale bar: 50 mm.
1155 Source data are provided as a Source Data file.

1156 **Supplementary Figure 11. Three experimental replicates of hypocotyl elongation**
1157 **assays with BtKAI2a and BtKAI2b transgenics (double exchange of residues 98 & 191)**

1158 Each panel depicts data from an independent experiment performed on a separate date,
1159 which are shown in summarised format in Figure 4. Data are means ± SE, n=24 to 40
1160 seedlings. Each dot corresponds to an individual seedling. Source data are provided as a
1161 Source Data file.

1162 **Supplementary Figure 12. Germination profiles of transgenic Arabidopsis seeds**
1163 **expressing BtKAI2 homologues (double exchange of residues 98 and 191)**
1164 Freshly harvested seeds (three batches per genotype, each batch harvested from three plants)
1165 were removed from freezer storage, surface-sterilised and sown on 0.7% agar supplemented
1166 with 0.02% acetone (mock), 1 μM KAR_1 or 1 μM KAR_2. Seeds were incubated under constant
1167 light at 25 °C. Seeds were examined for germination (radicle protrusion) 72 h after sowing
1168 and every 24 h thereafter. Data are means ± SE of three independent seed batches and 100
1169 seed per batch. For the $BtKAI2a^{L98;L191}$ and $BtKAI2b^{V98;V191}$ transgenes, two independent,
1170 homozygous transgenic lines were analysed. Source data are provided as a Source Data file.

1171 **Supplementary Figure 13. Three experimental replicates of hypocotyl elongation**
1172 **assays with BtKAI2b transgenics (individual exchange of residues 98 and 191)**
1173 Each panel depicts data from a separate experiment performed on the indicated date, which
1174 are shown in summarised format in Figure 5. Data are means ± SE, n = 19-21 seedlings. Each
1175 dot corresponds to an individual seedling. Source data are provided as a Source Data file.

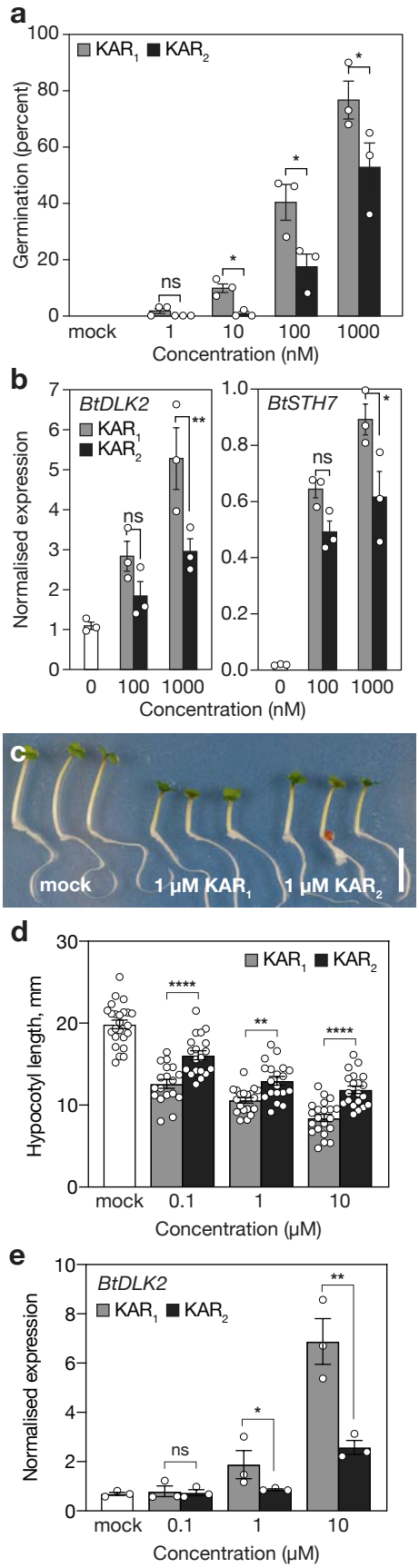


Figure 1. *Brassica tournefortii* is most sensitive to KAR_1, the major karrikin analogue isolated from plant-derived smoke
a, Germination responses of *B. tournefortii* seed to KAR_1 and KAR_2. Data are cumulative germination after 11 days (means ± SE; n = 3 biological replicates per treatment, ≥35 seeds per replicate). **b**, Levels of karrikin-responsive transcripts *BtDLK2* and *BtSTH7* in *B. tournefortii* seed. Seed were imbibed for 24 hours in the dark supplemented with KAR_1 and KAR_2. Transcripts were normalised to *BtCACS* reference transcripts. Data are means ± SE, n = 3 biological replicates, ≥50 seeds per replicate. **c**, **d**, Hypocotyl elongation responses of *B. tournefortii* seedlings grown for four days under continuous red light on water-saturated glass filter paper containing 0.1% acetone (mock), KAR_1 or KAR_2. Data are means ± 95% CI of n = 18 to 24 seedlings. Scale bar: 10 mm. **e**, Levels of *BtDLK2* transcripts in *B. tournefortii* seedlings grown under the same conditions as for d. Data are means ± SE, n = 3 biological replicates, ≥20 seedlings per replicate. In all panels, asterisks denote significance levels (ANOVA) between indicated conditions: * P<0.05; ** P<0.01; *** P<0.001; **** P<0.0001. Source data are provided as a Source Data file.

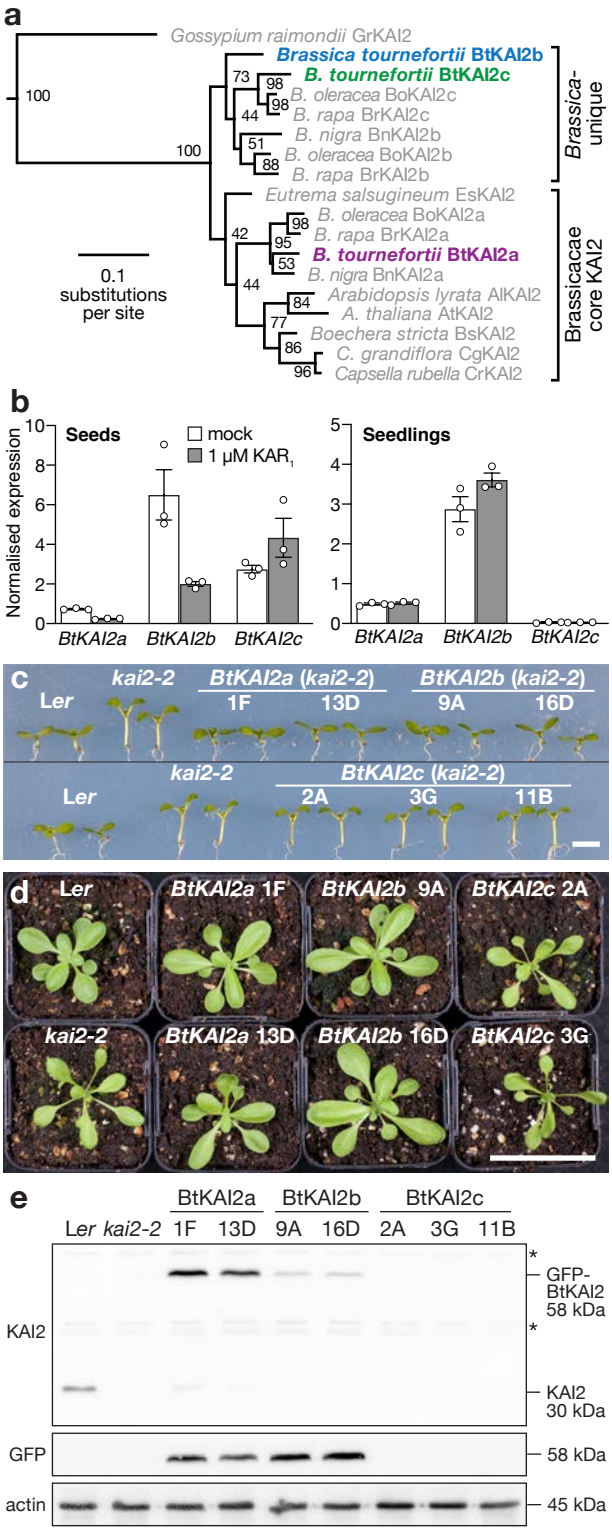


Figure 2. Two differentially expressed *B. tournefortii* KAI2 homologues are functional in Arabidopsis
a, Maximum likelihood phylogeny of KAI2 homologues in the Brassicaceae, based on nucleotide data. Node values represent bootstrap support from 100 replicates. A *KAI2* sequence from *Gossypium raimondii* (Malvaceae) serves as an outgroup. Tree shown is a subset of a larger phylogeny in Supplementary Figure 2. **b**, Transcript levels of the three *BtKAI2* homologues in *B. tournefortii* seed imbibed for 24 h (left) and four-day-old *B. tournefortii* seedlings (right) treated with 0.1% acetone or 1 μM KAR_1 for 24 h. Data are means ± SE, n = 3 biological replicates. **c, d,** Seedling and rosette phenotypes of two independent transgenic lines of Arabidopsis homozygous for *KAI2pro:GFP-BtKAI2* transgenes. Scale bars: 5 mm (**c**); 50 mm (**d**).**e**, Immunoblots of soluble proteins challenged with antibodies against KAI2 (upper panel), GFP (middle panel) or actin as a loading control (lower panel). Anti-KAI2 detects both the native AtKAI2 protein (30 kDa) and the GFP-BtKAI2 fusion proteins (58 kDa) but detects BtKAI2b relatively poorly. Non-specific bands are marked with asterisks. Protein was isolated from pools of approximately fifty 7-day-old seedlings. Source data are provided as a Source Data file.

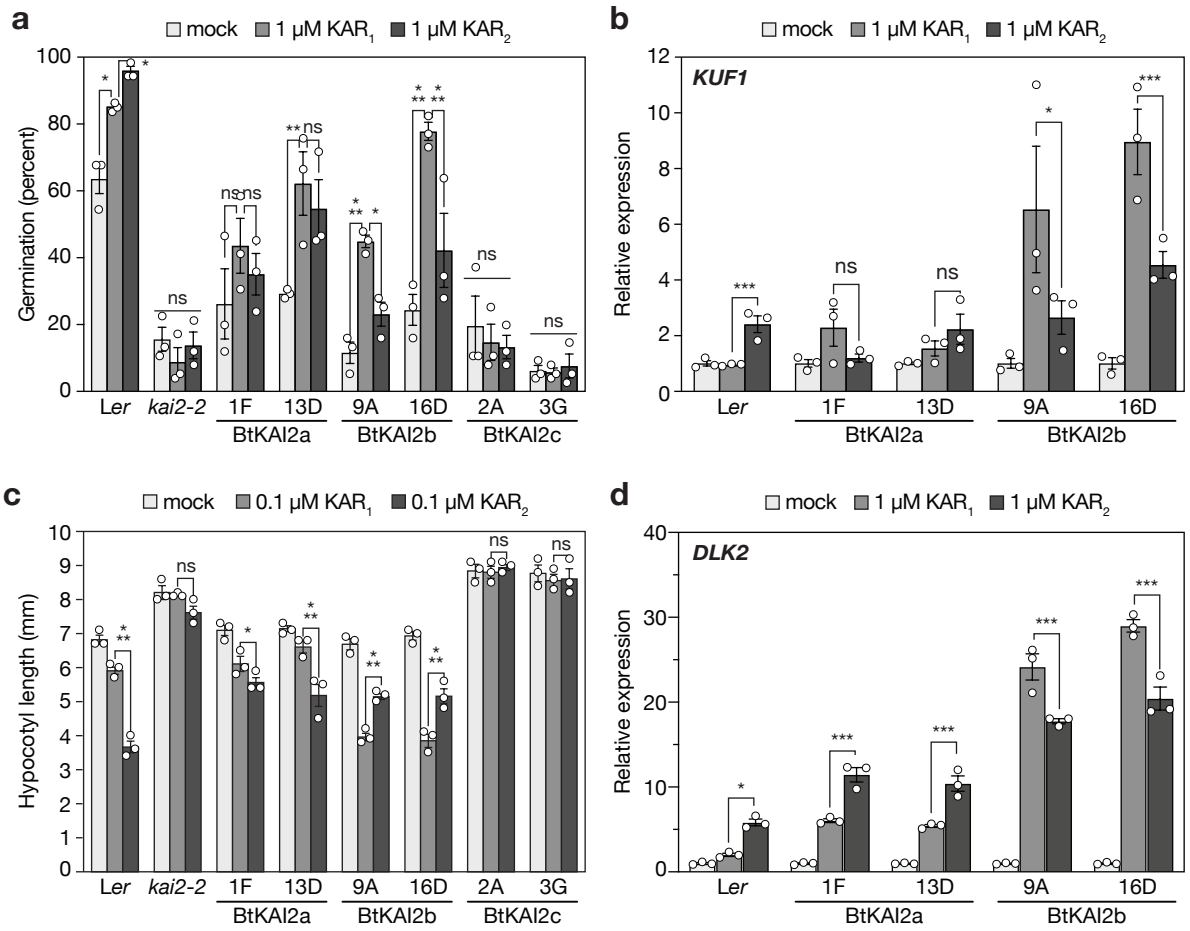


Figure 3. Functional divergence between BtKAI2 homologues in seeds and seedlings of transgenic Arabidopsis
a, Germination responses of primary dormant Arabidopsis seed homozygous for *KAI2pro:GFP-BtKAI2* transgenes in the *kai2-2* background. Germination values were determined 120 h after sowing. Extended germination curves are shown in Supplementary Fig. 7. Data are means ± SE of n = 3 independent seed batches, 75 seed per batch. **b**, Levels of *KUF1* transcripts in *KAI2pro:GFP-BtKAI2* seeds treated with 1 μM KAR_1 or KAR_2. Expression was normalised to *CACS* reference transcripts and scaled to the value for mock-treated seed within each genotype. Data are means ± SE of n = 3 biological replicates. **c**, Hypocotyl elongation responses of *KAI2pro:GFP-BtKAI2* seedlings treated with KAR_1 or KAR_2. Data are means ± SE of n = 3 biological replicates, 12–18 seedlings per replicate. **d**, Levels of *DLK2* transcripts in 8-day-old *KAI2pro:GFP-BtKAI2* seedlings treated with KAR_1 or KAR_2 for eight hours. Expression was normalised to *CACS* reference transcripts and scaled to the value for mock-treated seedlings within each genotype. Data are means ± SE of n = 3 biological replicates. Pairwise significant differences: * $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$; ns, $P > 0.05$ (two-way ANOVA). Source data are provided as a Source Data file.

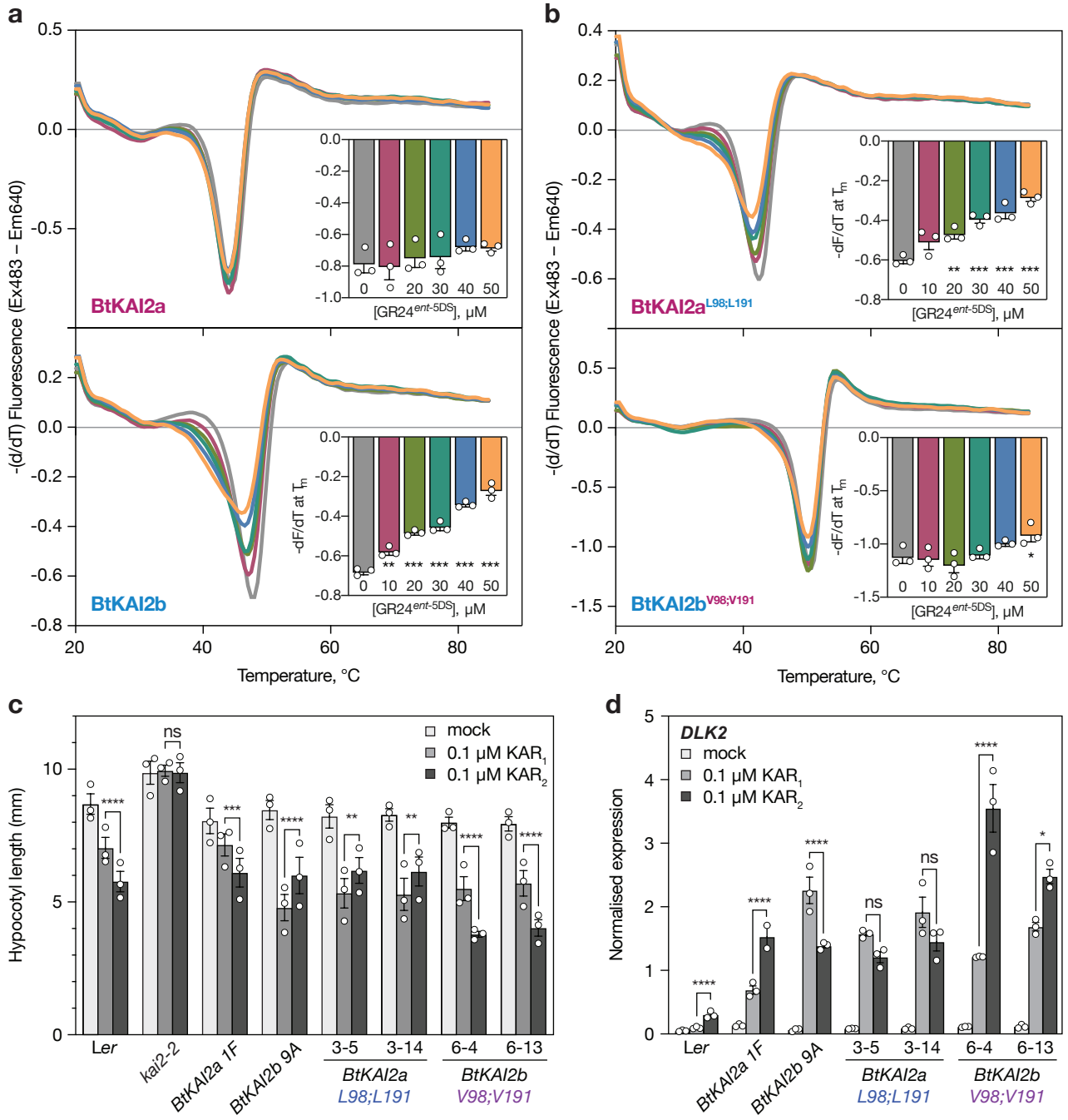


Figure 4. Two residues contribute to ligand specificity between BtKAI2a and BtKAI2b
a, DSF curves of SUMO-BtKAI2a and SUMO-BtKAI2b fusion proteins treated with 0-50 µM GR24$^{ent\text{-}5DS}$, a KAI2-bioactive ligand. Each curve is the average of three sets of reactions, each comprising four technical replicates. Insets plot the minimum value of –(dF/dT) at the melting point of the protein as determined in the absence of ligand (means ± SE, n = 3). Significant differences from untreated control: * $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$ (ANOVA). **b**, DSF curves of SUMO-BtKAI2a$^{L98;L191}$ and SUMO-BtKAI2b$^{V98;V191}$ fusion proteins treated with 0-50 µM GR24$^{ent\text{-}5DS}$. **c**, Hypocotyl elongation responses of Arabidopsis expressing GFP-BtKAI2a$^{L98;L191}$ and GFP-BtKAI2b$^{V98;V191}$ fusion proteins and treated with KAR_1 or KAR_2. Data are a summary of three experimental replicates performed on separate occasions, each comprising 25-40 seedlings per genotype/treatment combination. Data for each replicate are shown in Supplementary Figure 11. Error bars are SE, n = 3 experimental replicates; each dot corresponds to the mean value derived from each replicate. Asterisks denote significant differences: * P<0.05, ** P<0.01, *** P<0.001, **** P<0.0001 (linear mixed model with experimental replicate as a random effect; specific pairwise comparisons using Tukey's HSD correction). **d**, Levels of *DLK2* transcripts in the same transgenic lines as above treated with 1 µM KAR_1, 1 µM KAR_2, or 0.1% acetone and harvested 8 hours later. Expression was normalised to *CACS* reference transcripts. Data are means ± SE of n = 3 pools of ~50 seedlings treated in parallel. Asterisks denote significant differences as above (two-way ANOVA; specific pairwise comparisons using Tukey's HSD correction). Source data are provided as a Source Data file.

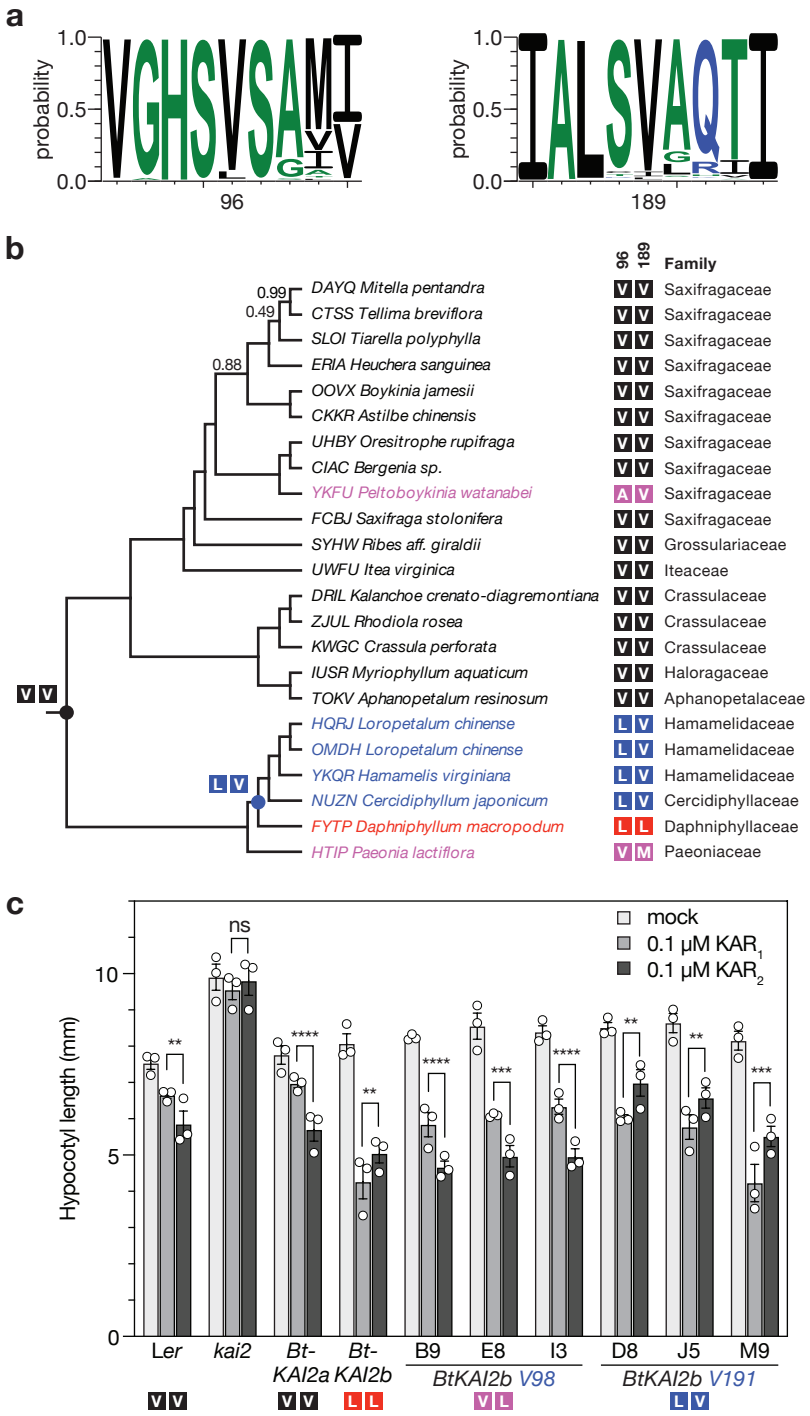
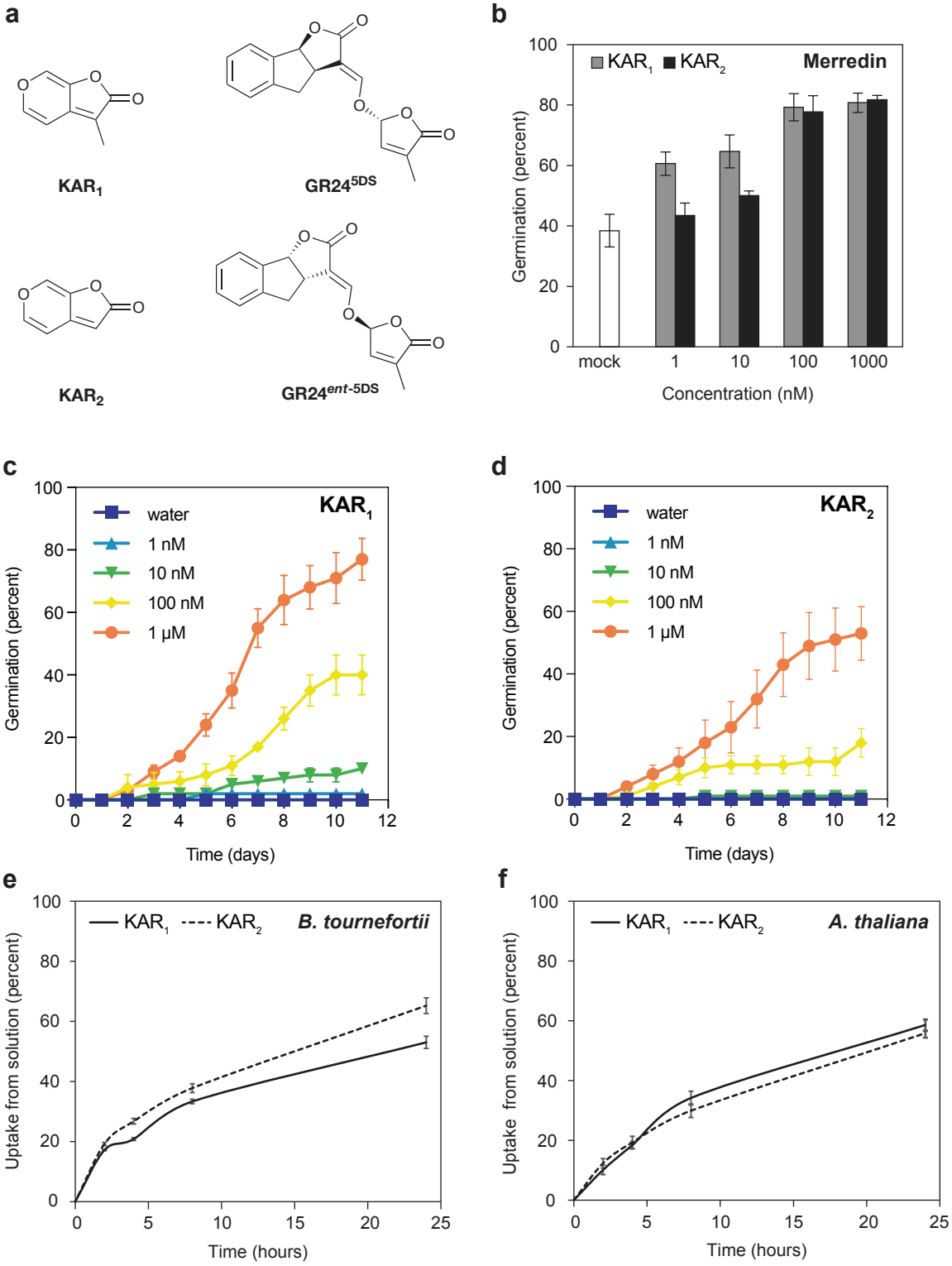
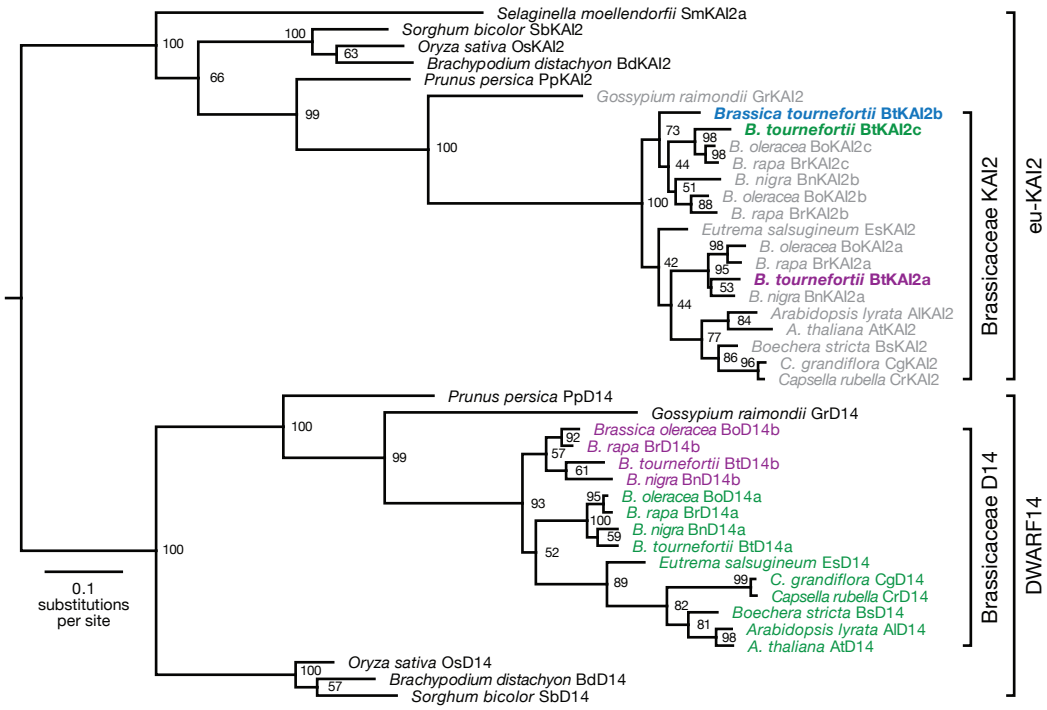


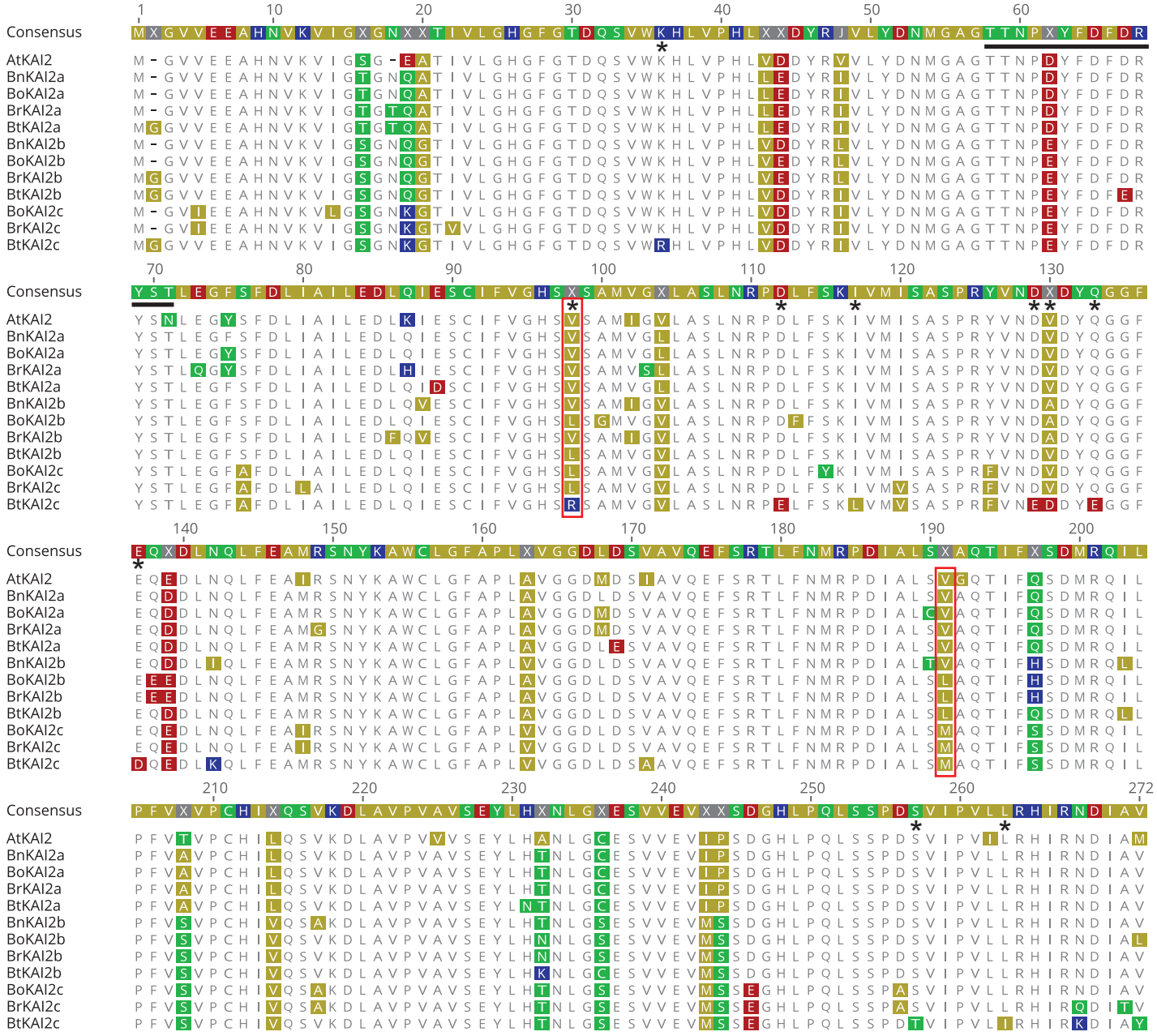
Figure 5. Diversity at positions 96 and 189 in KAI2 homologues across angiosperms
a, Sequence logos generated from 476 aligned KAI2 homologues sampled from eudicots, monocots and magnoliids and centred on position 96 (left) and 189 (right). **b**, Species tree of the Saxifragales based upon 410 single-copy nuclear genes estimated using the ASTRAL gene tree method (Leebens-Mack et al., 2019). Branch support values are shown only where less than 1. The four-letter taxon prefix corresponds to the 1000 Plant Transcriptomes project Sample ID. Taxa are highlighted by colour according to identity of amino acids of KAI2 orthologues at positions 96 and 189 respectively: black, VV; blue, LV; red, LL; magenta, other. Postulated ancestral identities at specific nodes are indicated with filled circles. **c**, Hypocotyl elongation responses to karrikins in three independent transgenic lines homozygous for *GFP-BtKAI2a*$^{L98;V191}$ and three lines for *GFP-BtKAI2b*$^{V98;L191}$. Data shown are a summary of three experimental replicates performed on separate occasions, each comprising ~20 seedlings per genotype/treatment combination. Data from each replicate are presented in Supplementary Figure 13. Error bars are SE, n = 3 experimental replicates; each dot corresponds to the mean value derived from each replicate. Asterisks denote significant differences: * P<0.05, ** P<0.01, *** P<0.001, **** P<0.0001 (linear mixed model with experimental replicate as a random effect; specific pairwise comparisons using Tukey's HSD correction). Source data are provided as a Source Data file.



Supplementary Figure 1. Germination of *Brassica tournefortii* seeds treated with karrikins
a, Structures of KAR_1, KAR_2 and two enantiomers of GR24. **b,** Germination response of the "Merridin" batch of *B. tournefortii* seed treated with KAR_1 and KAR_2 after three days. **c-d,** Germination response of the "Perth" batch to KAR_1 (**c**) and KAR_2 (**d**). Data in Figure 1a are derived from the data shown in **c** and **d.** **e-f**, Rates of KAR_1 and KAR_2 uptake by imbibed seed, as determined by GC-MS. All error bars are mean ± SE of n = 3 batches of 75 seed (b-d) or 3 samples of 40 mg (e) or 20 mg (f) seed as described in Methods. Source data are provided as a Source Data file.



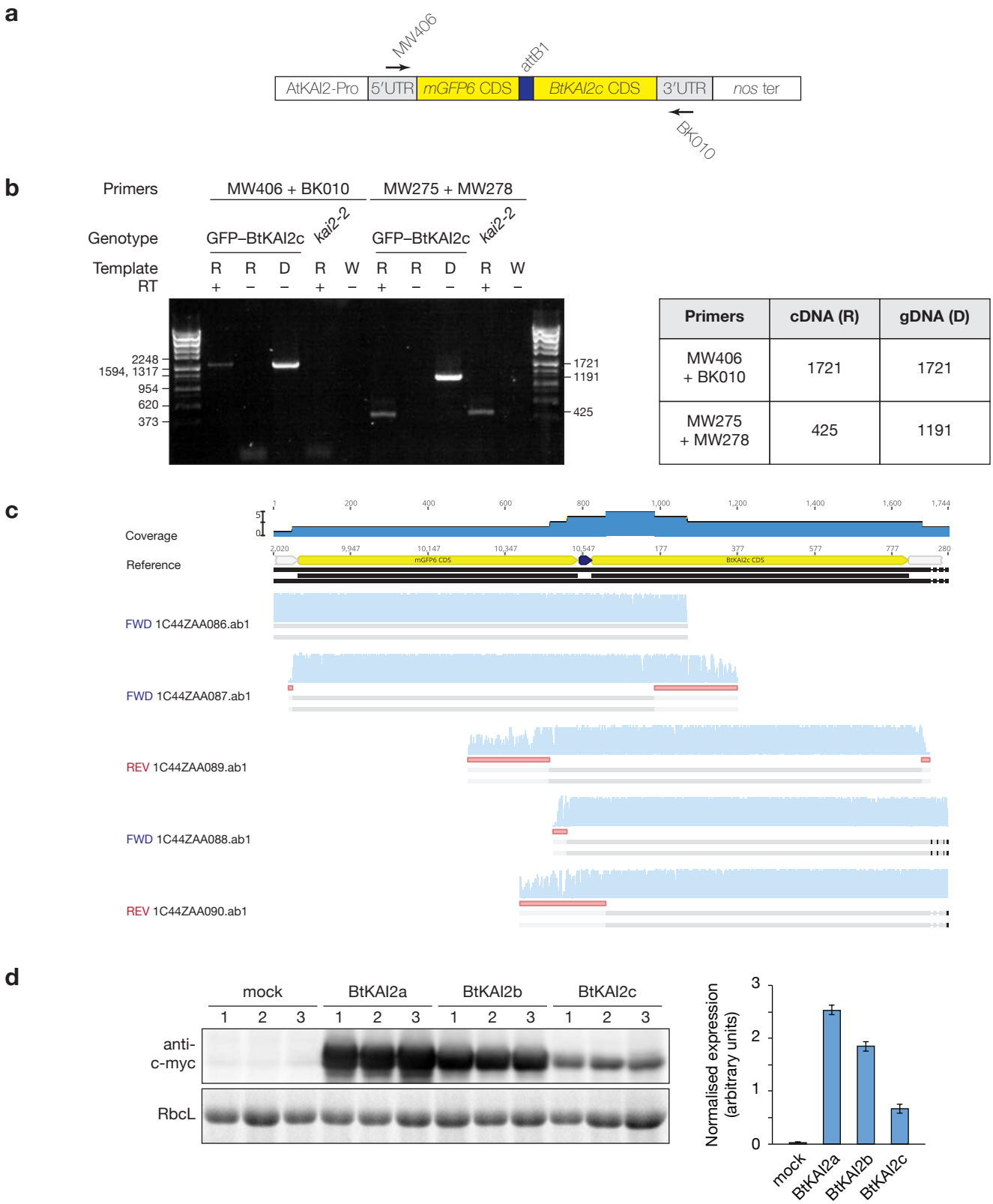
Supplementary Figure 2. Extended phylogeny of eu-KAI2 and D14 proteins in angiosperms
Maximum likelihood phylogeny of KAI2 and D14 homologues in the Brassicaceae and monocots, based on nucleotide data. Node values represent bootstrap support from 100 replicates. A KAI2 sequence from *Selaginella moellendorffii* (SmKAI2a) serves as an outgroup for the eu-KAI2 clade. Source data are provided as a Source Data file.



Supplementary Figure 3. Alignment of *Brassica* KAI2 sequences

Full length protein coding regions of KAI2 homologues from four *Brassica* species (*Brassica tournefortii*, *B. rapa*, *B. nigra* and *B. oleracea*) were translated from database nucleic acid sequences and aligned to Arabidopsis KAI2 using MAFFT (ref. 1) implemented in Geneious R10 software (Biomatters Ltd). Amino acid residues are coloured according to polarity: yellow, non-polar (G, A, V, L, I, F, W, M, P); green, polar & uncharged (S, T, C, Y, N, Q); red, polar & acidic (D, E); blue, polar & basic (K, R, H). Residues that are unique to BtKAI2c but otherwise invariant are highlighted with asterisks (*); residues 98 and 191 are highlighted with red boxes. The underlined region indicates the epitope used to raise the antibody against AtKAI2. Source data are provided as a Source Data file.

Supplementary Figure 4. BtD14a is functionally homologous to AtD14
Three independent transgenic lines of Arabidopsis *Atd14-1* were analysed for functional complementation of the mutant phenotype by an *AtD14pro:BtD14a* transgene. **a**, Rosette and leaf morphology at 31 days post-germination. **b**, Plant height and number of primary rosette branches at 45 days post-germination. **c**, Quantification of height and branching parameters, means ± SE, *n* = 10 plants per genotype. Scale bars: 50 mm. Source data are provided as a Source Data file.



Primers	cDNA (R)	gDNA (D)
MW406 + BK010	1721	1721
MW275 + MW278	425	1191

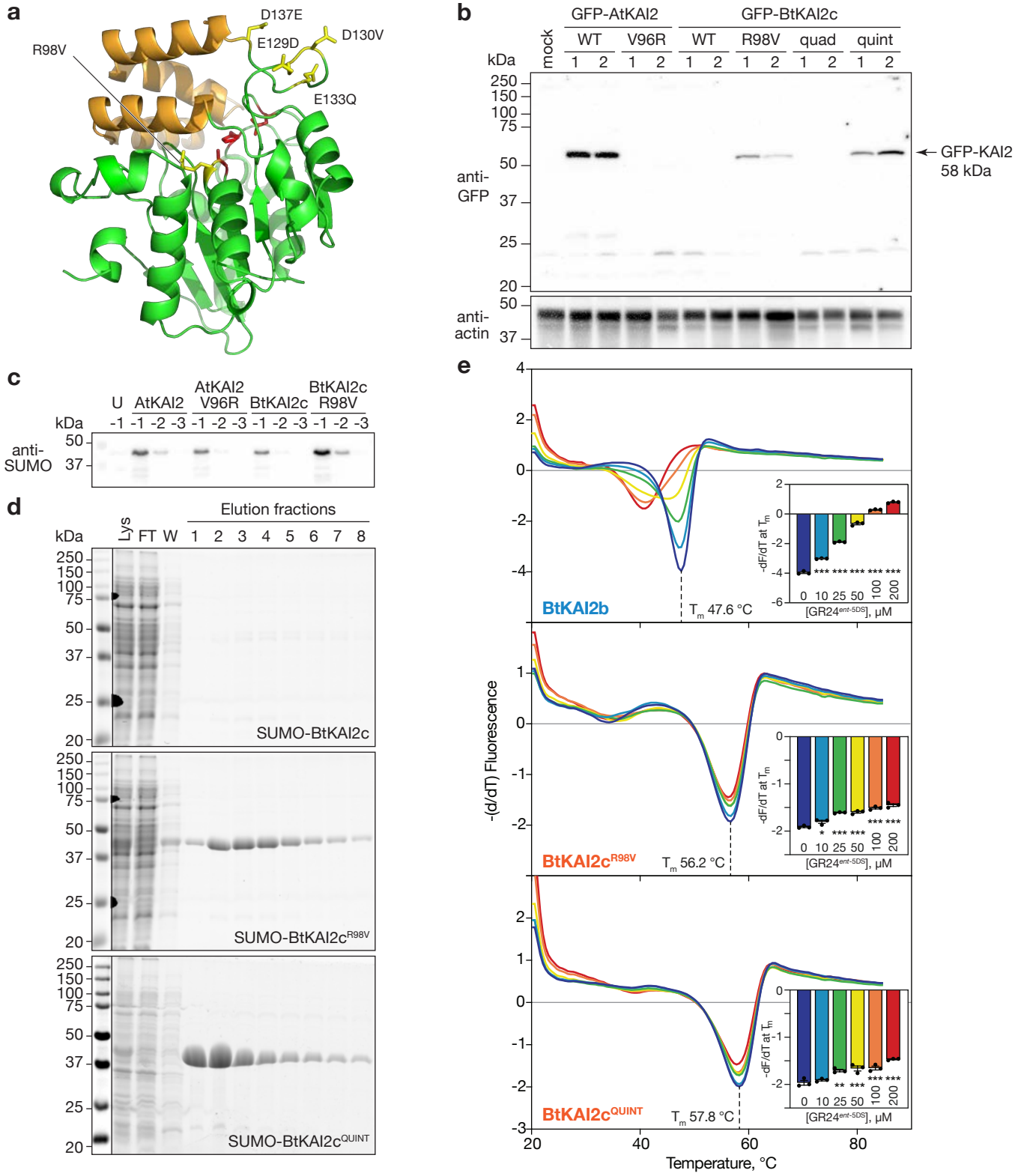
Supplementary Figure 5. The *GFP-BtKAI2c* transgene is faithfully transcribed in Arabidopsis

a, Structure of the *AtKAI2pro:mGFP6-BtKAI2c* transgene. Primers used for RT-PCR are shown with arrows. The promoter and 5′UTR are derived from *At4g37470* (*AtKAI2*). attB1, Gateway recombination site that links mGFP6 and BtKAI2c regions; *nos* ter, nopaline synthase terminator. Not drawn to scale.

b, RT-PCR analysis of *GFP-BtKAI2c* transcripts after 35 cycles of amplification. Primers MW406 + BK010 target the transgene, while a second primer pair (MW275 + MW278) serves as a control and spans five introns of *At1g03055*. *kai2-2* serves as a non-transgenic control genotype. Templates: R, total RNA; D, genomic DNA; W, water only. RT, reverse transcriptase (Supercript III). DNA size standards (in base pairs) are indicated on the left, with anticipated PCR product sizes shown on the right and defined in the table.

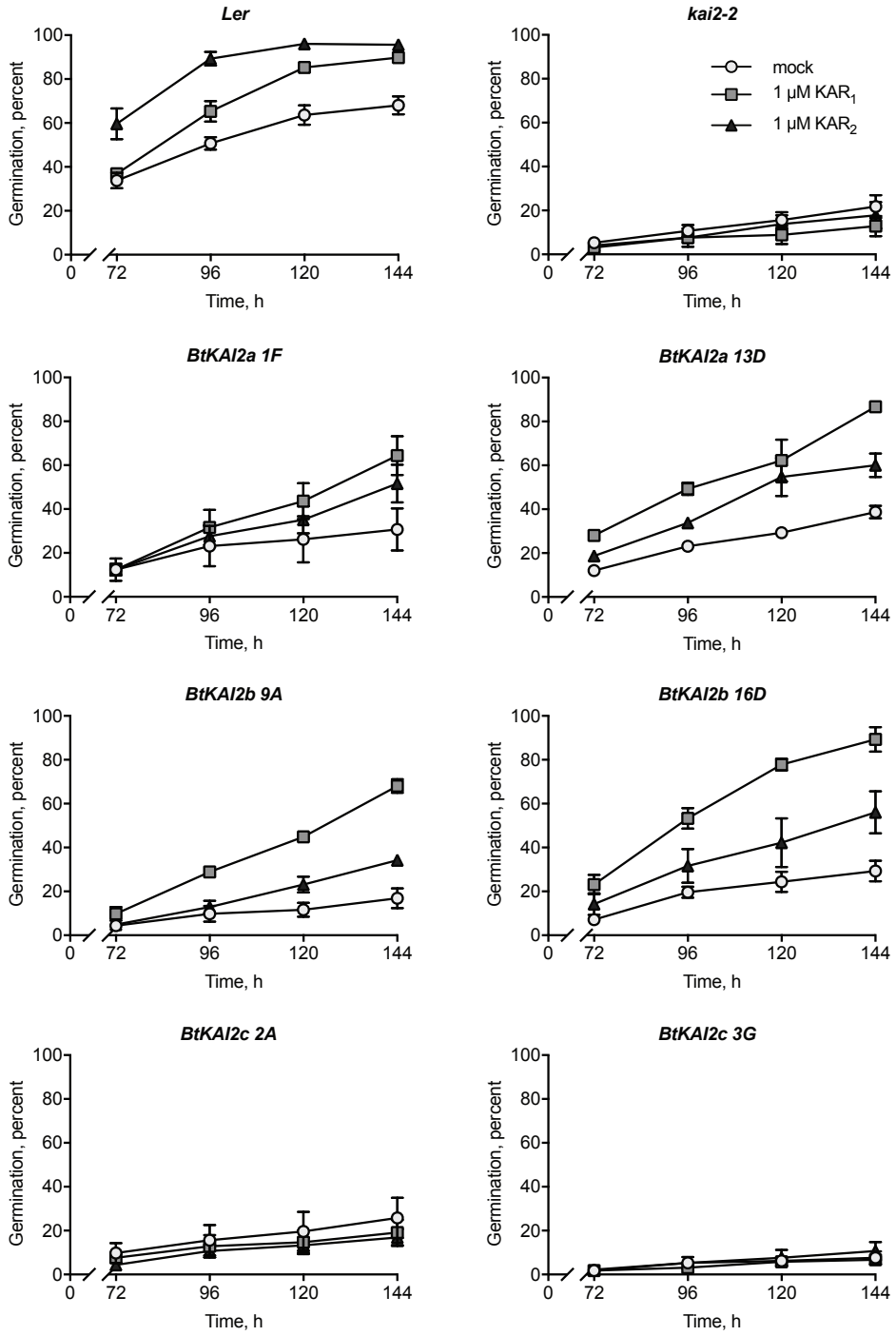
c, The RT-PCR product generated with proof-reading polymerase (Q5, New England Biolabs) and primers MW406 and BK010 was cloned into pCR4-TOPO (Life Technologies). Five dideoxy sequence traces were aligned against the *GFP-BtKAI2c* transgene reference. No disagreements with the reference sequence were observed. Red bars indicate trimmed regions of sequence traces to remove low quality data.

d, Transient expression of BtKAI2a, BtKAI2b and BtKAI2c proteins in tobacco. Plasmids encoding N-terminal, c-myc-tagged proteins were transferred to Agrobacterium, and the resulting strains used to infiltrate tobacco leaves. After 96 h, samples were harvested in triplicate (two to three leaves per sample). Mock-treated leaves were transformed with a plasmid encoding a non-tagged protein. Sixty micrograms of total protein were separated by SDS-PAGE, blotted and challenged with anti-c-myc antibody (Genscript A00704). Band intensity was measured using ImageJ, and expression was normalised to intensity of the Rubisco large subunit (RbcL) band on the “stain-free” gel imaged under UV light. Error bars indicate SE, $n = 3$ replicates. Source data are provided as a Source Data file.



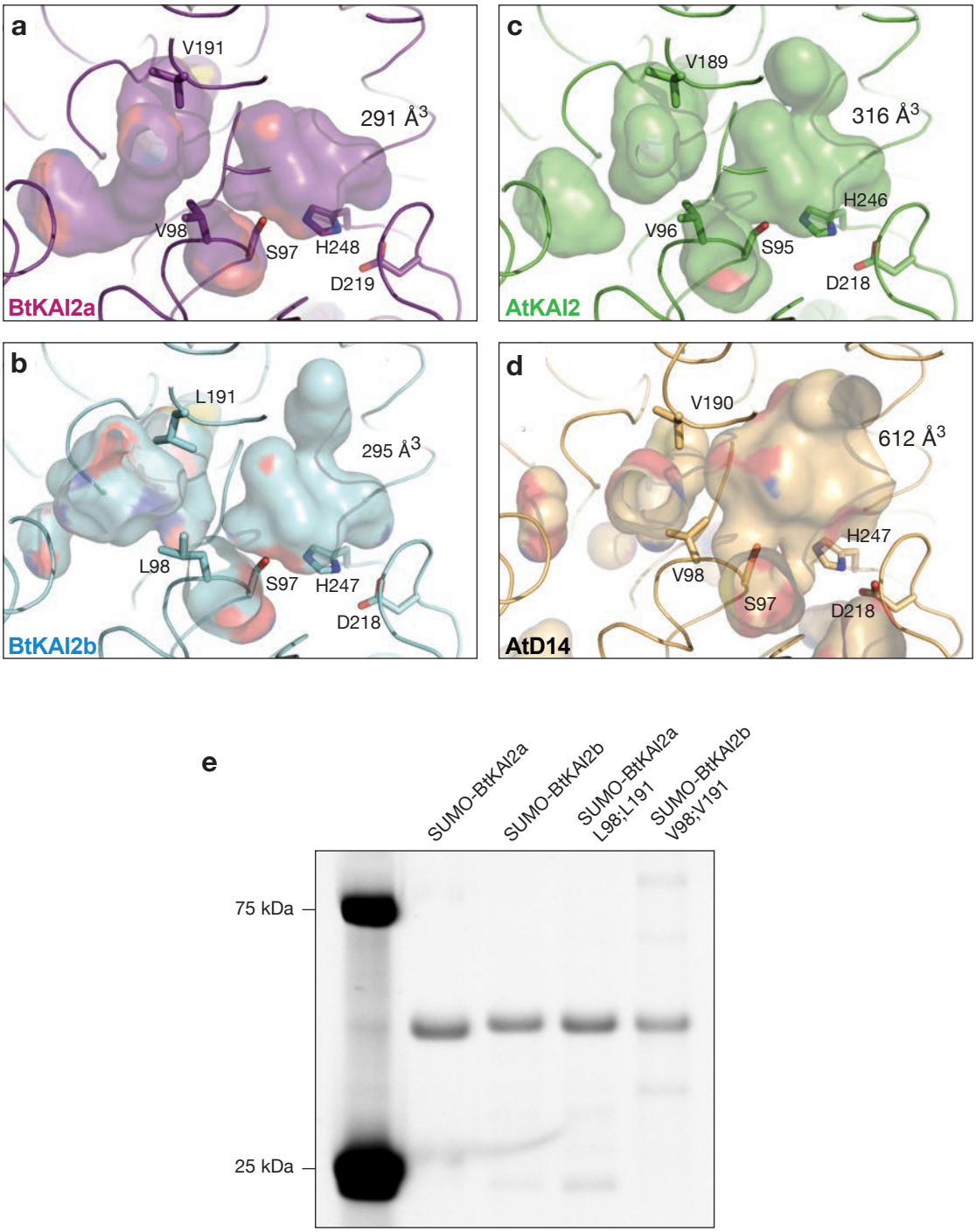
Supplementary Figure 6. The R98V mutation restores stability of BtKAI2c

a, Homology model of the native BtKAI2c protein. Highlighted in yellow are five residues and their corresponding mutations that were assessed for effects on BtKAI2c stability. The combination of E129D, D130V, E133Q and D173E in the hinge region between the lid domain (orange) and core domain (green) is the quadruple ("quad") mutant; combining this with R98V yields the quintuple ("quint") mutant. Red residues are the Ser-His-Asp catalytic triad. **b**, Transient expression of AtKAI2 and BtKAI2c variants in tobacco leaves following infiltration by Agrobacterium strains encoding GFP-KAI2 variants driven by the CaMV *35S* promoter. Three leaves were infiltrated on each of two different plants. After five days, 20 µg soluble protein was electrophoresed, blotted and challenged with antibodies against either anti-GFP (top) or anti-actin (bottom). **c**, Immunoblot of crude bacterial lysates expressing SUMO fusion proteins and challenged with anti-SUMO antibody. Negative values above each lane indicate $\log_{10}$ dilution factors of lysate; U, lysate from uninduced bacterial culture expressing SUMO-AtKAI2.**d**, SDS-PAGE analysis of protein purification runs of SUMO-BtKAI2c, -BtKAI2c^{R98V} and -BtKAI2c^{QUINT} harvested from 900 mL bacterial culture. Lys, clarified lysate; FT, column flow through after immobilisation on Co^{2+} affinity column; W, wash with 10 mM imidazole. Proteins were eluted with 200 mM imidazole in eight successive fractions. **e**, DSF curves of SUMO fusion proteins treated with 0-200 µM GR24$^{ent\text{-}5DS}$. Insets plot the minimum value of –(dF/dT) at the melting point of the protein as determined in the absence of ligand (means ± SE, n = 3). Significant differences from untreated control: * $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$ (ANOVA). Source data are provided as a Source Data file.

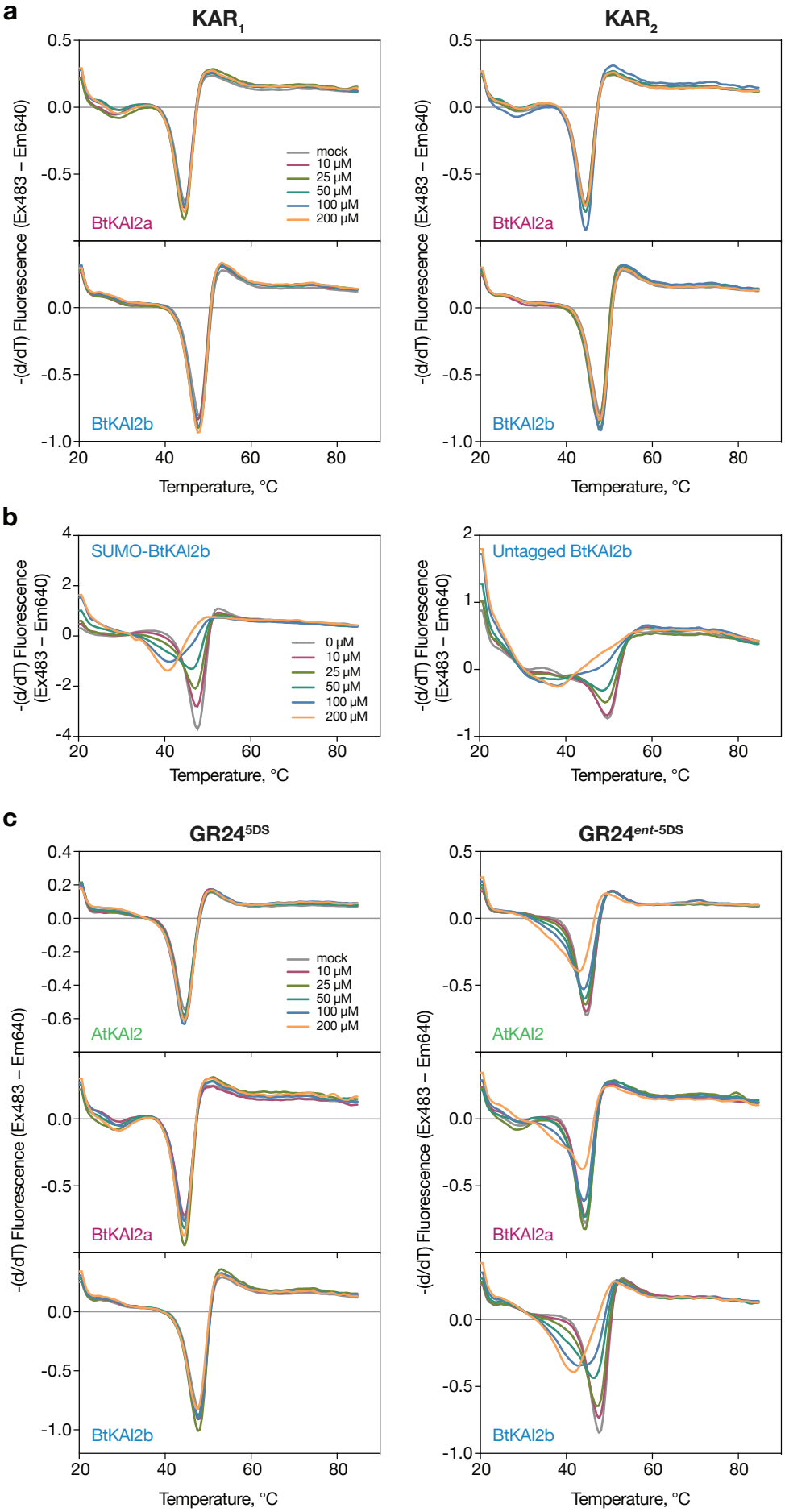


Supplementary Figure 7. Germination profiles of transgenic Arabidopsis seeds expressing native BtKAI2 homologues

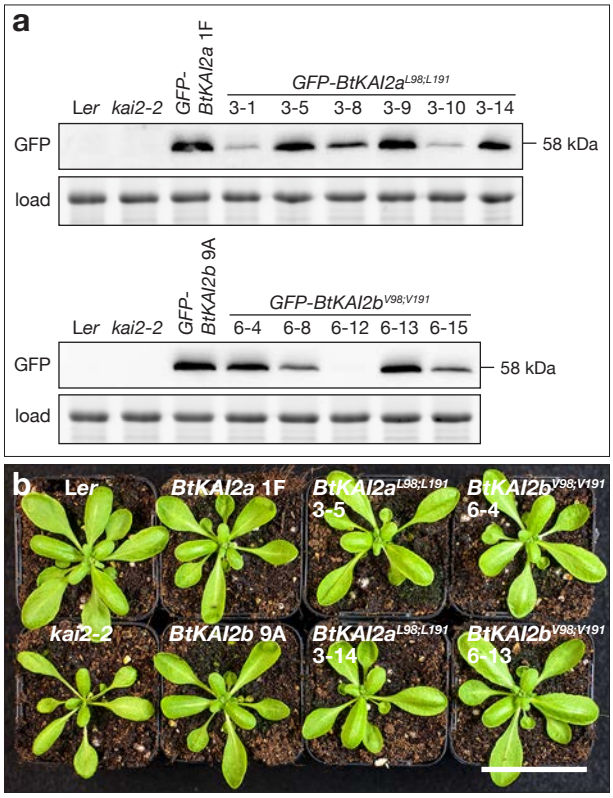
Freshly harvested seed (three batches per genotype, each batch harvested from four plants) were removed from freezer storage, surface-sterilised and sown on 1% Phytagel supplemented with 0.1% acetone (mock), 1 μM KAR_1 or 1 μM KAR_2. Seed were incubated under constant light at 25 °C. Seed were examined for germination (radicle protrusion) 72 h after sowing and every 24 h thereafter. Data are means ± SE of three independent seed batches and 75 seed per batch. For each transgene, two independent, homozygous transgenic lines were analysed. Data presented in Figure 3 of the main manuscript are derived from these data. Source data are provided as a Source Data file.



Supplementary Figure 8. BtKAI2 homology models and SDS-PAGE of SUMO-BtKAI2 fusion proteins used for DSF.
a-d, Solved structure of AtKAI2 (PDB: 3w06, ref. 2), AtD14 (PDB: 4IH4, ref. 3) and predicted homology models of BtKAI2a and BtKAI2b. Coloured surfaces depict internal cavities; values indicate the volumes of the primary ligand-binding cavities, adjacent to the catalytic Ser-His-Asp residues. Also shown is a variable secondary pocket, to the left of the primary pocket in these images.
e, To assess purity after affinity chromatography, five micrograms of each purified protein was electrophoresed on a 12% acrylamide gel containing 2,2,2-trichloroethanol and visualised under UV light. Protein size standards at 75 and 25 kDa (Bio-Rad Precision Plus Dual Colour) fluoresce strongly under UV light. Source data are provided as a Source Data file.

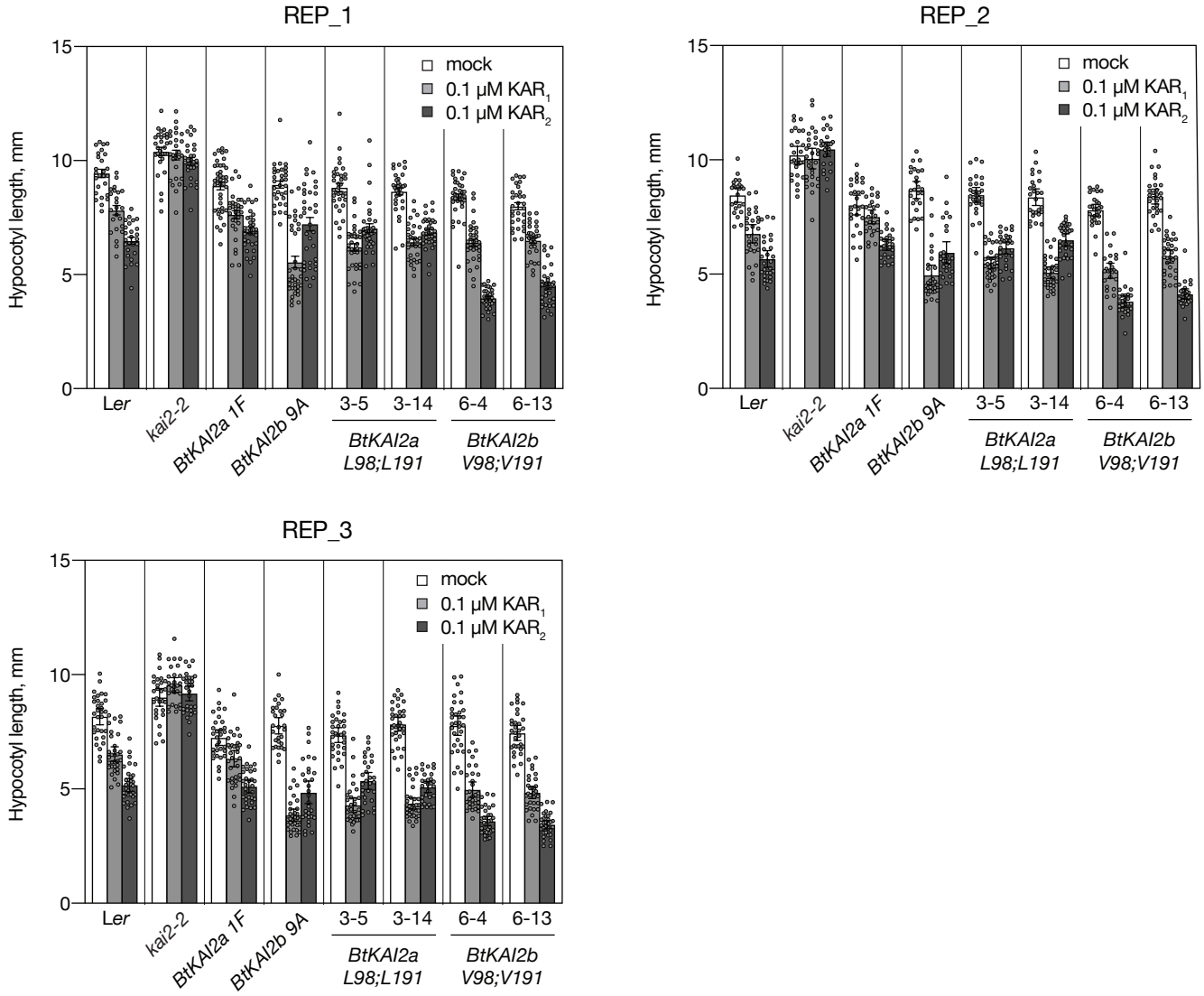


Supplementary Figure 9. BtKAI2a and BtKAI2b do not respond to karrikins in DSF assays
a, Differential scanning fluorimetry curves of SUMO-BtKAI2a and SUMO-BtKAI2b in presence of 0–200 μM KAR_1 or KAR_2. **b**, DSF responses of SUMO-tagged BtKAI2b (left) or untagged BtKAI2b (right) to 0–200 μM $GR24^{ent\text{-}5DS}$. c, DSF responses to 0–200 μM of the two enantiomers of GR24. Data are means of eight (a and c) or four (b) technical replicates at each concentration of ligand. Source data are provided as a Source Data file.



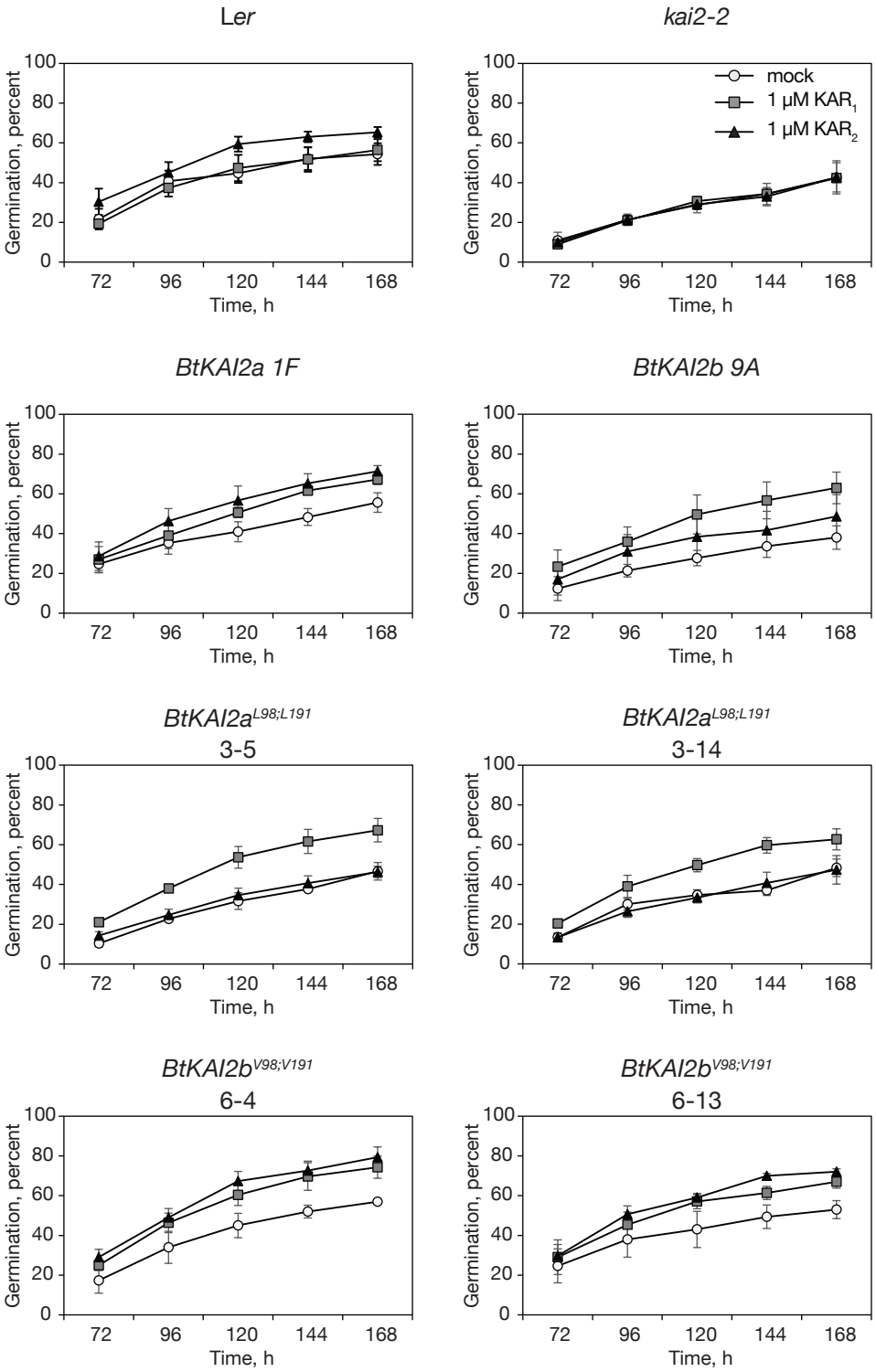
Supplementary Figure 10. Stable transgenic expression of BtKAI2 valine–leucine double exchange proteins in Arabidopsis

a, Immunoblots of total soluble protein extracted from 7-day-old seedlings of independent transgenic lines segregating in a 3:1 ratio for hygromycin resistance. Transgene expression in six lines expressing GFP-BtKAI2a$^{L98;L191}$ (upper panels) and five expressing GFP-BtKAI2b$^{V98;V191}$ (lower panels) were compared to a representative unmodified control (GFP-BtKAI2a 1F and GFP-BtKAI2b 9A respectively). Based on expression level two lines of each construct (3-5 and 3-14; 6-4 and 6-13) were selected and brought to homozygosity for further experiments. Protein blots were challenged with anti-GFP antibody. Equal gel loading was assessed by imaging total protein prior to blotting; the RbcL band is shown. **b**, Rosette phenotypes of homozygous individuals expressing native and modified *GFP-BtKAI2* transgenes. Plants were 25 days old and grown under long day conditions as described in Methods. Scale bar: 50 mm. Source data are provided as a Source Data file.



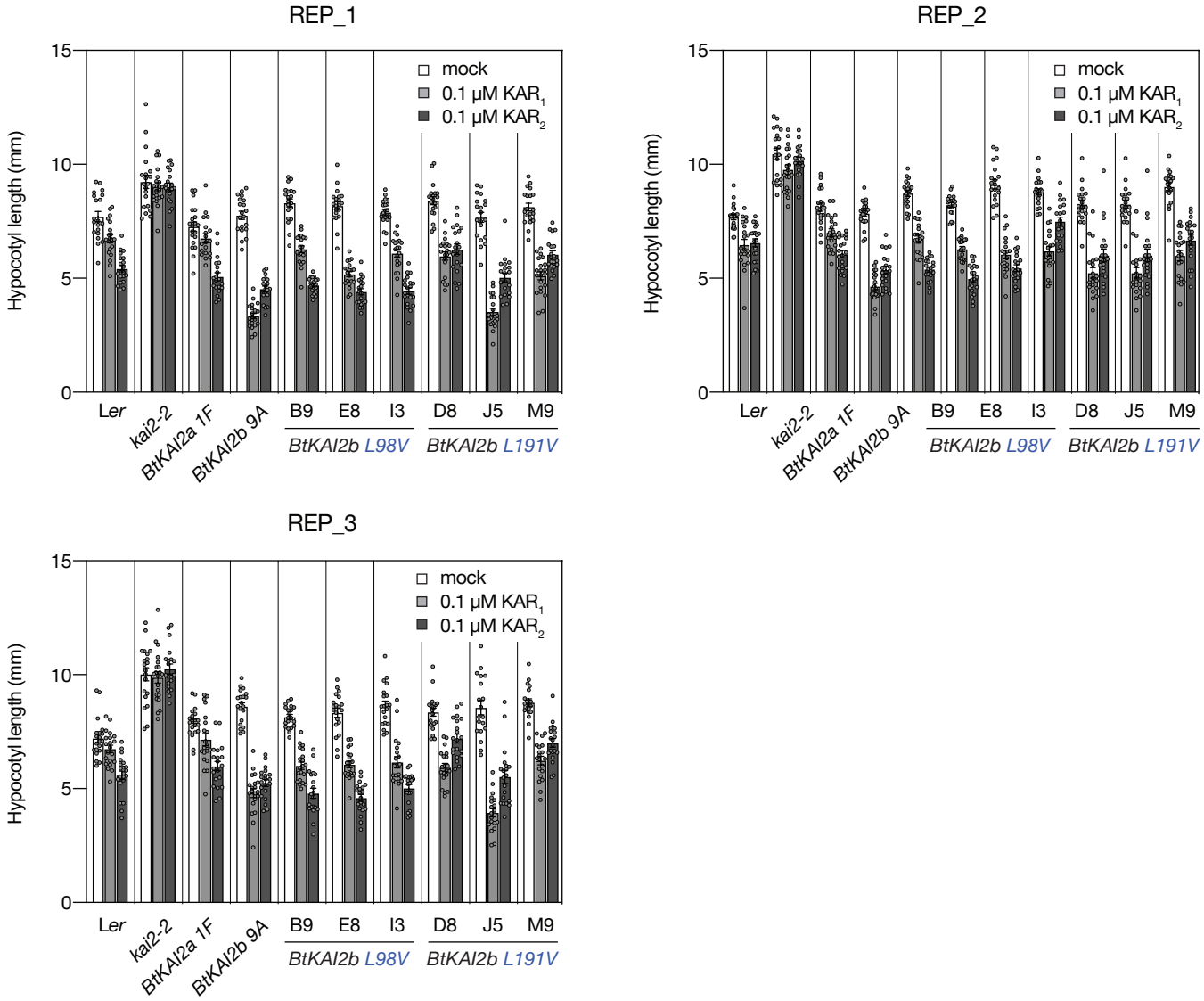
Supplemental Figure 11. Three experimental replicates of hypocotyl elongation assays with BtKAI2a and BtKAI2b transgenics (double exchange of residues 98 & 191)

Each panel depicts data from a separate experiment performed on the indicated date, which are shown in summarised format in Figure 4. Data are means ± SE, n=24 to 40 seedlings. Each dot corresponds to an individual seedling. Source data are provided as a Source Data file.



Supplementary Figure 12. Germination profiles of transgenic Arabidopsis seeds expressing BtKAI2 homologues (double exchange of residues 98 and 191)

Freshly harvested seeds (three batches per genotype, each batch harvested from three plants) were removed from freezer storage, surface-sterilised and sown on 0.7% agar supplemented with 0.02% acetone (mock), 1 μM KAR_1 or 1 μM KAR_2. Seeds were incubated under constant light at 25 °C. Seeds were examined for germination (radicle protrusion) 72 h after sowing and every 24 h thereafter. Data are means ± SE of three independent seed batches and 100 seed per batch. For the *BtKAI2a*$^{L98;L191}$ and *BtKAI2b*$^{V98;V191}$ transgenes, two independent, homozygous transgenic lines were analysed. Source data are provided as a Source Data file.



Supplemental Figure 13. Three experimental replicates of hypocotyl elongation assays with BtKAI2b transgenics (individual exchange of residues 98 & 191)

Each panel depicts data from a separate experiment performed on the indicated date, which are shown in summarised format in Figure 5. Data are means ± SE, n = 19-21 seedlings. Each dot corresponds to an individual seedling. Source data are provided as a Source Data file.