



HAL
open science

Three mutations repurpose a plant karrikin receptor to a strigolactone receptor

Amir Arellano-Saab, Michael Bunsick, Hasan Al Galib, Wenda Zhao, Stefan Schuetz, James Michael Bradley, Zhenhua Xu, Claresta Adityani, Asrinus Subha, Hayley Mckay, et al.

► **To cite this version:**

Amir Arellano-Saab, Michael Bunsick, Hasan Al Galib, Wenda Zhao, Stefan Schuetz, et al.. Three mutations repurpose a plant karrikin receptor to a strigolactone receptor. *Proceedings of the National Academy of Sciences of the United States of America*, 2021, 118 (30), pp.e2103175118. <10.1073/pnas.2103175118>. <hal-03299064>

HAL Id: hal-03299064

<https://hal.science/hal-03299064v1>

Submitted on 26 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

1 **Three mutations repurpose a plant karrikin receptor to a strigolactone receptor**

2

3 Amir Arellano-Saab^{1,7}, Shigeo Toh^{2,7}, Hasan Galib¹, Wenda Zhao¹, Stefan Schuetz¹, James
4 Michael Bradley¹, Asrinus Subha¹, Alexandre de Saint Germain³, Claresta Adityani¹, Michael
5 Bunsick¹, Hayley McKay¹, François-Didier Boyer⁴, Christopher S. P. McErlean⁵, Peter
6 McCourt^{1,8} Peter J. Stogios^{6,8} and Shelley Lumba^{1,8}

7

8 ¹ Department of Cell & Systems Biology, University of Toronto, 25 Willcocks Street, Toronto
9 M5S 3B2, Canada

10

11 ² Department of Environmental Bioscience, School of Agriculture, Meijo University
12 1-501 Shiogamaguchi, Tenpaku-ku, Nagoya, Japan, 468-8502

13

14 ³ Institut Jean-Pierre Bourgin, INRAE, AgroParisTech, Université Paris-Saclay, 78000,
15 Versailles, France.

16

17 ⁴ Université Paris-Saclay, CNRS, Institut de Chimie des Substances Naturelles, UPR 2301,
18 91198, Gif-sur-Yvette, France.

19

20 ⁵ School of Chemistry, The University of Sydney, NSW, 2006, Australia.

21

22 ⁶ Department of Chemical Engineering & Applied Chemistry, Banting & Best Department of
23 Medical Research, University of Toronto, 200 College Street, Toronto M5S 3E5,

24

25 ⁷ Both authors contributed equally to the project

26

27 ⁸ To whom correspondences should be addressed

28

29 **Uncovering the basis of small molecule hormone receptors evolution is paramount to a**
30 **complete understanding of how protein structure drives function. In plants, hormone**
31 **receptors for strigolactones are well suited to evolutionary inquiries because closely related**
32 **homologs have different ligand preferences. More importantly, because of facile plant**
33 **transgenic systems, receptors can be swapped and quickly assessed functionally *in vivo*.**
34 **Here, we show only three mutations are required to switch the non-strigolactone receptor,**
35 **KAI2, into a strigolactone receptor. This modified receptor still perceives KAI2 ligands and**
36 **does not require receptor hydrolysis for activity. Structural and molecular dynamic**
37 **modeling suggest receptor pocket flexibility is important for ligand specificity and**
38 **downstream signaling partner affinity. These findings indicate a few keystone mutations**
39 **link strigolactone signaling to germination, which explains how parasitic plants that**
40 **devastate African agriculture evolved SL receptors to sense the presence of a host plant.**

41

42 **Introduction**

43 Plants and animals use small molecule hormones to drive growth and development and for this
44 reason much effort is placed on understanding how hormone receptors perceive their ligands^{1,2}.
45 Most inquiries are tuned to identify receptor amino acids that contribute to ligand specificity
46 using “loss-of-function” approaches where an amino acid is replaced with a chemically inert
47 moiety, usually an alanine, and binding is assessed^{4,5}. The less explored “gain of function”
48 approach involves swapping amino acids between related receptors that respond to different
49 ligands^{6,7}. Gain-of-function approaches have the potential to not only answer questions about
50 receptor-ligand relationships but also can give insights into how receptors evolve. As selection
51 drives changes in the receptor sequence to recognize new ligands, important questions regarding
52 this molecular process can be posed: is this process gradual occurring through many amino acid
53 changes or are only a few keystone substitutions required? When receptors evolve the ability to
54 recognize new ligands, do they lose the ability to bind the ancestral ligands? Gleaning this type
55 of information through gain-of-function approaches, however, has imposing challenges⁸. First,
56 amino acids are not discrete entities acting in isolation but are a sea of interdependent residue. As
57 a result, swapping amino acids into a new context often can result in a non-functional protein due
58 to negative epistasis. Second, depending on the number of potential swaps, the number of

59 variants can quickly scale limiting analysis to simple *in vitro* assays rather than more complex *in*
60 *vivo* analysis.

61 With this said, the perception of the plant hormone strigolactone (SL) may be a useful
62 system to understanding the evolution of small molecule receptors. The root parasite *Striga*
63 *hermonthica* (*Striga*), for example, uses a group of α/β hydrolases designated
64 HYPOSENSITIVE TO LIGHT (ShHTL) to perceive host-derived SLs allowing the parasite to
65 coordinate its germination with the lifecycle of a host plant⁹. By contrast, the homologous
66 protein, HYPOSENSITIVE TO LIGHT/KARRIKIN INSENSITIVE 2 (KAI2), which is also
67 involved in germination in the model plant *Arabidopsis thaliana* (*Arabidopsis*), is not considered
68 an SL receptor^{10,11}. Plants only produce SLs with a 2'*R* enantiomeric configuration in D-ring but
69 this enantiomer is not recognized by KAI2 (**Supplementary Fig. 1**)^{12,13}. In *Arabidopsis*, (2'*R*)-
70 SLs are recognized by a more distantly related α/β hydrolase, DWARF14 (AtD14), a receptor
71 that is only involved in vegetative growth^{14,15}. Instead KAI2 responds to a collection of smoke-
72 derived butenolides collectively called karrikins (KARs)¹⁶, and strangely is activated by non-
73 natural (2'*S*)-SL enantiomers (**Supplementary Fig. 1**)¹³. This has led to suggestions that KAI2,
74 like ShHTL receptors, positively regulates germination not by using naturally occurring (2'*R*)-
75 SLs but by perceiving an unidentified butenolide-based ligand, designated KAI2-ligand (KL)¹⁷.
76 Because parasitic lifestyles are derived from nonparasites, this implies that ShHTL parasitic
77 receptors evolved the ability to perceive (2'*R*)-SLs from ancestor KAI2 receptors that bind KL¹⁸.

78 Structural analysis of KL and SL receptors generally conclude that ShHTL receptors have
79 larger binding sites with a broader set of active site amino acids allowing for greater
80 responsiveness to an array of (2'*R*)-SL molecules, a trait thought to increase *Striga* host range¹⁹.
81 There are also suggestions that receptor affinity to downstream signaling partners also
82 contributes to ligand responsiveness²⁰. The separation of ligand specificity versus partner affinity
83 helps to resolve the conundrum of how *Striga* SL receptors show decreased specificity but
84 increased sensitivity to the same ligand. These conclusions, however, are based solely on *in vitro*
85 experiments because *Striga* is an experimentally intractable genetic system²¹. Fortunately,
86 ShHTL receptors function in *Arabidopsis* mutants deficient in KAI2 function and under specific
87 conditions confers seed germination sensitivity to SLs^{11,22}. Furthermore, model plants like
88 *Arabidopsis*, unlike their animal counterparts, are easily transformed²³ making high throughput
89 analysis at the whole organism level feasible. This cross-species functionality means

90 homologous-substitution mutagenesis between KL-based KAI2 and SL-based ShHTL receptors
91 can be performed to probe evolutionary questions of ligand sensitivity, specificity and
92 promiscuity in an *in vivo* context.

93 With this in mind, here we systematically substituted key amino acids from the highly
94 responsive ShHTL7 SL receptor at homologous positions in KAI2 and assayed their ability to
95 sense SLs *in vivo*. We show using both yeast- and plant-based assays, that only three amino acid
96 changes were required to repurpose KAI2 into an SL receptor that recognizes (2'*R*)-SL
97 enantiomers. The emergent properties of this chimeric KAI2 receptor does not appear to require
98 loss of KL perception. Furthermore, ligand perception does not require receptor hydrolysis.
99 Structural and molecular dynamic modeling indicates substitution of three active site amino acids
100 from ShHTL7 into KAI2 produces a larger more flexible binding pocket that better
101 accommodates SLs including (2'*R*)-SL enantiomers and most likely allows better interactions
102 with downstream signaling partners. The relatively simple mutational path to convert KAI2 into
103 a SL receptor together with the central role of KAI2 in non-parasitic plant germination has
104 implications with respect to the molecular mechanism of how *Striga* co-opted this regulator to be
105 a germination sensor of host-derived SLs.

106

107 **Results**

108 **Identifying amino acids involved in SL binding.** KAI2 and ShHTL receptors share β -sheet
109 cores flanked by α -helices with their ligand binding pockets covered by a V-shaped cap²⁴⁻²⁷.
110 Within the binding pocket is a catalytic triad of amino acids that is involved in ligand
111 hydrolysis^{9,27-29}. To identify important amino acid differences between a KL and an SL receptor
112 we compared amino acid sequences within the ligand binding pocket of KAI2 to the *Striga* SL
113 receptor, ShHTL7, which both have experimental 3D structures available^{11,22}. Because the larger
114 ShHTL7 binding pocket can accommodate (2'*R*)-SLs^{19,20}, we focused on ShHTL7 residues that
115 were less bulky and/or more polar than their homologous counterparts in KAI2 (**Supplementary**
116 **Fig. 2**). This identified eight ShHTL7 amino acids (Leu26, Tyr124, Thr142, Leu153, Thr157,
117 Tyr174, Thr190, Cys194) that we substituted into the equivalent homologous position in KAI2 in
118 all single, double and triple combinations to generate 92 chimeric KAI2 gene variants.

119

120 **Three substitutions change KAI2 responsiveness to SL.** To systematically test the 92 chimeric
121 KAI2 receptors, we developed two *in vivo* high throughput SL receptor responsive assays. KAI2
122 interacts with its signaling partner MORE AXILLARY MERISTEMS 2 (MAX2)³⁰ making this
123 interaction a good candidate for a yeast two-hybrid assay (Y2H). KAI2, however, autoactivates
124 in Y2H assays in some conditions³¹ and MAX2 is also unstable in yeast¹⁹. Therefore, we
125 developed Y2H growth conditions where KAI2 did not autoactivate (**see Supplementary**
126 **Materials & Methods**) and fragmented MAX2 into nine constructs (N1-N9) to identify a
127 fragment that produced a positive Y2H signal in the presence of a SL (**Supplementary Fig. 3**).
128 Although KAI2 (Var1) and the chimeric receptor variants (Var2-Var93) showed no or poor
129 interactions with most MAX2 fragments, the N3 fragment interacted with some chimeric
130 receptors to varying degrees, in the presence of a racemic mixture (2'R, 2'S) of the artificial SL,
131 *rac*-GR24 (**Fig. 1a**). With a cut off of four standard deviations above the mean interaction
132 intensity this experiment identified seven variants (Var21, 37, 62, 64, 65, 66, 67) that showed
133 *rac*-GR24-dependent N3 interactions (**Fig. 1b**). Six of these variants contained a ShHTL7 amino
134 acid substitution at position 190 and four variants contained substitutions in positions 124 or 157
135 suggesting that these amino acids influence the *rac*-GR24-dependent interactions in yeast
136 (**Supplementary Fig. 3**). The identification of amino acids positioned at 124, 157 and 190 is
137 consistent with *in vitro* studies that show these amino acids are important for SL binding^{19,20}.

138 Parallel to our Y2H assays, we evaluated all 92 of the chimeric KAI2 receptors by
139 transforming them into *Arabidopsis* deficient for KAI2 function (*htl-3*)³⁰. Using an SL-dependent
140 germination assay involving GA-depleted seeds¹¹, we found the one line, Var64 (Trp153Leu,
141 Phe157Thr, Gly190Thr), germinated well upon *rac*-GR24 addition (**Fig. 1a**). Subsequent
142 analysis of three independent transgenic lines (64A, 64B, 64C) showed 50 percent germination
143 (EC₅₀) at high nanomolar concentrations of GR24_{rac} (**Fig. 1c**). Interestingly, Var19 lines
144 (Trp153Leu, Gly190Thr), which contains a subset of Var64 amino acids, showed the next
145 highest level of *rac*-GR24-dependent germination in our *in planta* assay (**Fig. 1a**,
146 **Supplementary Fig. 4**). Notably, the Var64 receptor variant was also identified in our Y2H
147 assay, indicating that two independent assays showed that substitution of Leu153, Thr157 and
148 Thr190 into KAI2 greatly improved its responsiveness to *rac*-GR24.

149 The identification of Leu153, Thr157 and Thr190 as having roles in SL perception
150 encouraged us to determine the prevalence of these amino acids in KAI2-related proteins across

151 a large collection of land plants. Focusing on clades of receptors that are not expected to bind
152 SLs (conserved) versus clades of receptors that are expected to bind SLs (divergent)¹⁸ we found
153 conserved clades were mostly devoid of ShHTL7-related amino acids whilst the divergent clades
154 were highly populated with ShHTL7-related residues (**Fig. 2**). Divergent KAI2 proteins were
155 particularly enriched for 153Leu or hydrophobic amino acids at this position, and proteins more
156 closely related to ShHTL7 possessed similar amino acids at positions 157Thr and 190Thr (**Fig.**
157 **2**). Interestingly, within the divergent group, the one clade of proteins depleted of ShHTL7
158 amino acid identities were most closely related to receptors ShHTL10 and ShHTL11 (e.g.
159 *SaKAI2d2, 10, 11, 16, ShKAI2d10, 11*) which do not have a role in germination as assayed by
160 *Arabidopsis*²². Together, the distribution of ShHTL7 amino acids across land plants suggests that
161 our *in vivo* screening identifies important amino acids involved in the evolutionary switch of
162 KAI2 to an SL receptor. Finally, the observation that both our assays led to the Var64 amino acid
163 combination suggests that the simpler and less time consuming yeast-based protein interaction
164 assays may be a good first screen before plant-based assays.

165

166 **KAI2 variant 64 recognizes natural occurring (2'R)-SL.** The SL responsiveness of Var64
167 seed may be due to increased sensitivity to either or both enantiomers of GR24_{rac}. To
168 differentiate between these possibilities, we measured the germination response of three
169 independent Var64 lines (64A, 64B, 64C) to either (2'S)- or (2'R)-GR24 in our GA-depletion
170 assay¹¹. Compared to misexpressed KAI2, the Var64 lines showed more sensitivity to 2'S-GR24
171 but importantly showed responsiveness to (2'R)-GR24, a property not present in the wild type
172 KAI2 receptor (**Fig. 3a**). We also noticed that compared to KAI2 seed, Var64 lines germinated at
173 low levels, even in the absence of *rac*-GR24_{rac}, suggesting a weakly constitutive germination
174 phenotype in the absence of SL addition (**Fig. 3a**). Suppressor analysis in rice identify GID1 α/β
175 hydrolase GA receptor mutants with weak GA constitutive phenotypes that are enhanced by GA
176 addition³². This is thought to be due to alterations in the conformation of the receptor's lid
177 subdomain that increases its affinity for interactions with its downstream signaling partners³². By
178 analogy, ShHTL7 amino acids may alter the active site of KAI2 to a conformation closer to the
179 "on state", which may improve downstream protein interactions. Consistent with this, the
180 positions 153 and 157 are important in MAX2 and SMAX1 interactions with ShHTL7 *in vitro*²⁰.

181 To biochemically characterize the molecular mechanism of Var64 signalling, we first
182 measured hydrolysis activity of Var64 variant receptors using the YLG assay⁹. Var64 receptors
183 showed weak YLG hydrolysis activity compared to WT KAI2 enzyme, but this activity was
184 inhibited by addition of *rac*-GR24_{rac}, indicating that SL indeed bind the variant receptor (**Fig.**
185 **3b**). To more directly measure SL hydrolysis we monitored D-ring production by HPLC from
186 hydrolysis of either (-)-(2'S)- or (+)-(2'R)-GR24 and again found neither GR24 enantiomer was
187 a substrate compared to KAI2, which has a preference for (2'S)-SLs, or ShHTL7, which has a
188 preference for (2'R)-SLs (**Fig. 3c**). In summary, although Var64 seed showed improved
189 germination on (2'S)-GR24 and an emergent ability to germinate on the (2'R)-GR24 the Var64
190 protein at best, has poor hydrolysis activity for these two enantiomers. This suggests this receptor
191 variant functions in response to SL *in vivo* but with minimal hydrolase activity.

192
193 **KAI2 variant 64 responds to karrikins.** Because the identity of KL is unknown, we cannot
194 directly assess its binding to our variant receptors. However, KL also has a role in *Arabidopsis*
195 leaf development³³, which means rescue of the *htl-3* leaf mutant phenotype by variant receptors
196 in transgenic plants can be used as a readout of endogenous KL binding to its receptor. Under
197 our growth conditions, *htl-3* leaf length to width ratios are smaller (≤ 1.85) versus wild type
198 leaves (≥ 1.92) and this difference can be used to quantify the degree to which chimeric variants
199 can complement *htl-3* leaf shape defects (**Supplementary Fig. 5**). Generally, higher order
200 mutant lines showed less rescue of *htl-3* leaf mutant phenotype than lower order mutants, which
201 could reflect a loss of endogenous KL recognition (**Fig. 1a**). This pattern, however, may also be
202 due to increasing negative epistasis as newly introduced amino acids pile up resulting in non-
203 functional proteins⁸. We noticed, however, that triple substitution variant lines that carried a
204 ShHTL7 amino acid at position 190 complemented *htl-3* leaf defects more often than other triple
205 mutant or even double mutants lacking this substitution (**Fig. 1a**). Deep sequencing analysis
206 suggest that amino acid substitutions can be neither positive nor negative in terms of protein
207 function, but simply stabilize proteins and allow for a larger number of substitutions to be
208 tolerated during evolution³⁴. Perhaps, Threonine at position 190 is a permissive substitution that
209 opens alternative mutational paths for ligand recognition. With this said, Var64 lines, which
210 contain a Thr190 substitution, complemented *htl-3* leaf defects suggesting that this chimeric
211 receptor retained an ability to recognize KL (**Fig. 1a**). Although this cannot be directly tested,

212 Var64 seed germinated on KAR₂ to a similar level seen for KAI2 seed, indicating that this
213 chimeric variant retains the capacity to recognize a native KAI2 ligand.

214

215 **KAI2 variant 64 protein-ligand dynamics.** Structural comparisons of the Var64 protein model
216 to a wild type KAI2 crystallographic structure (PDB 4IH1^{xx}) showed that replacing KAI2
217 aromatic residues Trp153 and Phe157 with Leu and Thr, disrupted several hydrogen bond
218 interactions that would occur between the α B, α C, and α D helices of the lid domain (**Fig. 4a, b**).
219 This change would increase the pocket mouth area from 45.5 Å² to 57 Å², which would allow for
220 larger substrates, such as GR24, to access the active site of the receptor more easily. These
221 disruptions would also result in the increase of the overall pocket solvent-accessible volume by
222 approximately 140 Å³ (**Fig. 4a, b**). Larger binding pockets should increase ligand accessibility
223 but to explore this in more detail we modeled (2'*R*)-GR24 ligand-receptor interactions using
224 molecular dynamics (MD) simulations. Unlike traditional structural docking analysis, which is a
225 static snapshot of the protein structure, MD simulations take into account the temporal evolution
226 and velocities of all of the atoms of a protein averaged over a specific period of time to give a
227 more realistic atomic-level detail of the protein. MD simulations of Var64 and KAI2 receptors in
228 the presence of (2'*R*)-GR24, as well as their apo states, were conducted in triplicate 10 ns
229 simulations to evaluate the dynamics of these ligand-receptor complexes in contrast with the
230 unbound receptor, calculating the number and distance of hydrogen bonds that form between
231 receptor and ligand, and measuring the fluctuation of the ligand positions throughout the
232 simulation. To quantitatively distinguish between ligand binding states, we defined productive
233 modes to include states in which (2'*R*)-GR24 was close to the receptor (≤ 5.0 Å) and its surface
234 area was mostly covered by the active site of the protein. The unproductive mode incorporates
235 instances where the ligand was further away from the pocket opening (5.1-7.0 Å) and not
236 interacting with the active site. Finally, a dissociated mode comprised states where (2'*R*)-GR24
237 was far from the protein (≥ 7.5 Å) thus any interactions are very unlikely. Using these criteria,
238 (2'*R*)-GR24 was in a productive binding mode with the Var64 receptor almost 50% of the
239 simulation time versus only 2.1% for KAI2 (**Fig. 4c**). Furthermore, when docked on KAI2,
240 (2'*R*)-GR24 spent approximately 65% of the simulation time in a dissociated mode versus 10%
241 for Var64. Finally, we calculated the magnitude of location change of the position of (2'*R*)-GR24
242 on the two receptors. For a productive ligand-receptor interaction, a minimal fluctuation is

243 expected, as the ligand is tightly bound until it is metabolized or released. The simulations
244 showed that the position of (2'R)-GR24 in KAI2 fluctuates between 0.02 and 0.18 nm once
245 initially bound to the protein, by contrast, this oscillation is greatly reduced in the Var64 protein,
246 confirming the ability of the new variant to form a stable interaction with an SL ligand (**Fig. 4d**).

247 Visualization of the trajectory of the MD simulations revealed that the motility of the α E
248 loop of unbound Var64 is increased, consistent with our model of increased flexibility of the
249 receptor's active site and the expansion of the cavity depth and volume (**Supplementary Movie**
250 **1**). The trajectories also show three main changes between bound and unbound receptors. First,
251 the hydrophobicity of the binding domain of Var64 was slightly reduced throughout the
252 simulation after binding of (2'R)-GR24, which is qualified by the change in coloration from
253 orange to blue in the binding domain region. Second, (2'R)-GR24 appears to be more attracted
254 by the polar ShHTL7 Thr190 in Var64 than the KAI2 Gly190. Possibly, this substitution is the
255 reason for the higher ligand stability provided by Var64. Lastly, we observe a conformational
256 change of the α D helix of Var64 upon binding to (2'R)-GR24 (**Supplementary Movie 2**). This
257 conformational change is expected in the SL perception pathway, as SL bound receptors have
258 been shown to adapt the conformation of the α D helix to accommodate interactions with its
259 downstream partners^{20,35,36}.

260

261 **Discussion**

262 Structurally guided mutational analysis is an effective approach to dissecting how a protein's
263 sequence contributes to biochemical activity, but how these biochemical properties evolve is
264 more experimentally challenging⁸. From an evolutionary perspective, loss-of-function
265 approaches are not necessarily informative and often experiments on protein evolution are not
266 performed in an *in vivo* context which allows proteins to interact with other cellular constituents
267 and processes. In this study, we used a gain-of-function approach with active site amino acid
268 targeted substitutions informed by structural and evolutionary considerations to determine if a
269 non-parasitic plant orphan receptor, KAI2, that is incapable of recognizing natural SL ligands
270 harnessed by parasitic plants, could be converted into a receptor that recognizes such ligands.
271 Importantly, we assayed these receptor variants using yeast- and *Arabidopsis*-based systems to
272 evaluate receptor function under *in vivo* conditions.

273 Our approach unambiguously determined that only three amino acid substitutions were
274 necessary to convert KAI2 into a functional SL receptor. By introducing smaller, more polar
275 amino acids into KAI2, we produced a protein with a more flexible pocket and lid; this affects
276 receptor function in two ways. First, it increases the pocket volume to accommodate natural SL
277 conformations like that of (2*R*')-GR24. Second, it increases the overall elasticity of the protein,
278 which may allow better interactions with signaling partners^{20,35,36}. To this end, the identification
279 of positions 153, 157 and 190 in our *in vivo* experiments is consistent with the *in vitro*
280 conclusions about these amino acids^{19,20}. Our *in vivo* findings contribute to discussions
281 concerning the role of gradual versus keystone mutations in small molecule receptor evolution.
282 Animal studies using ancestral gene reconstructions involving steroid receptors show as few as
283 two mutations can switch ligand specificity³⁷. Our results now extend the importance of keystone
284 mutations to plant hormone receptor evolution.

285 General models of hormone evolution suggest promiscuous ancestral forms evolve
286 specificity so as to prevent deleterious ligand crosstalk³⁸. Three lines of evidence in this study,
287 however, suggest a path to SL perception in KAI2 does not necessarily require loss of
288 responsiveness to previous ligands. First, our variant receptor complements the KAI2 loss-of-
289 function phenotype suggesting it still recognizes the endogenous KAI2 ligand. Second, variant
290 receptor seeds are sensitive to (2'*S*)-GR24 enantiomers, a known substrate for KAI2. Finally,
291 variant lines respond to the KAI2-specific exogenous ligands, karrikins. The ability of variant
292 receptors to respond to both ancestral as well as derived ligands supports the notion that SL
293 perception can evolve through sufficient promiscuity rather than by a narrowing of specificity.
294 The sufficient promiscuity models receptor specificity as being adequate but not overly so to
295 distinguish endogenous ligands³⁹. This “just enough” specificity explains why receptors often
296 show sensitivity to drug structures unrelated to their natural ligands. The discovery of an
297 artificial femtomolar responsive ShHTL7 agonist with little similarity to SLs further supports a
298 sufficient promiscuity model for these receptors⁴⁰.

299 It also appears that the tight connection between SL hydrolysis and SL perception is not
300 necessary as Var64 variant seed responded to GR24 enantiomers as the Var64 protein showed
301 little hydrolase activity. The role of ligand hydrolysis in SL signaling is somewhat contentious
302 with structural studies producing conflicting models^{41,42}. Chemical biology suggests hydrolysis is
303 important to signaling sensitivity but again may not be essential^{30,40}. Finally, mutational analysis

304 of the D14 receptor in *Arabidopsis* indicates hydrolysis is both necessary²⁹ and not necessary⁴²
305 for SL signaling. Although our results support a less direct connection between of hydrolysis and
306 signaling, our variant receptor most likely does not reflect the natural progression by which
307 KAI2 genes evolved to become SL receptors in *Striga*. Our KAI2 variants were not under the
308 constant constraints seen in natural environments that are important for fitness. This may explain
309 why in contrast to the yeast-based interaction assay, we only found one variant with high SL
310 responsiveness *in planta*. We understand that monitoring KAI2 function using the output of
311 *Arabidopsis* germination is superior to *in vitro* based experiments but even this system almost
312 certainly overestimates the frequency of beneficial mutations relative to natural evolution. With
313 this said, the correlation of ShHTL7 153, 157 and 190 amino acids in divergent versus conserved
314 KAI2 is consistent with these residues playing key roles in SL perception. Finally, directed
315 protein engineering studies tell us that enzymes usually possess minor non-selective activities
316 that are often enhanced by laboratory evolution^{44,45}. If true, KAI2 may possess a to-date
317 undetected low level of responsiveness to 2'R enantiomers. *In vitro*, KAI2 only appears to bind
318 (2'S)-GR24²⁹ but *in vivo*, KAI2 can perceive (2'R)-GR24 but only when the SL receptor, D14, is
319 defective⁴⁶, suggesting that the absence of D14 uncovers the latent binding ability of KAI2 for
320 2'R enantiomers as it would be free from competition for (2'R)-SL substrates with endogenous
321 D14.

322 The role of KAI2 in germination together with its ability to perceive small molecules
323 makes it easy to understand why this gene was co-opted during the evolution of *Striga*
324 parasitism. *Striga* plants not only requires a host for growth but also requires a cue to signal the
325 proximity to a host, and a long-term survival strategy in the absence of a host. Since SLs are
326 needed for beneficial plant root-AM fungi interactions⁴⁷, secretion of these small molecules
327 makes it an excellent plant root positional cue. Seeds are an attractive solution to establishing
328 long-term viability in the absence of a host as dormant seeds can survive long periods under a
329 variety of environmental conditions. An unintended agricultural consequence of using dormant
330 seeds as the protective state in the absence of a host is that they easily contaminate soil and are
331 difficult to purge from farmers fields. For these reasons, *Striga* has contaminated over 20 African
332 nations and is the largest biological impediment to food security on the continent^{48,49}.
333 Understanding the steps by which KAI2 evolved to become a SL receptor will give insights not
334 only into the role of this key regulator of the *Striga* lifecycle but will also help define the

335 chemical space of this receptor, which should open leads to new chemical solutions to this
336 African scourge.

337 **Reference**

- 338 1. Evans, R. M. & Mangelsdorf, D. J. Nuclear Receptors, RXR, and the Big Bang. *Cell*,
339 **157**, 255–266 (2014).
- 340 2. Lumba, S., Cutler, S. & McCourt, P. Plant nuclear hormone receptors: a role for small
341 molecules in protein-protein interactions. *Annu Rev Cell Dev Biol.* **26**, 445-469 (2010).
- 342 3. Morrison, K. L. & Weiss, G. A. Combinatorial alanine-scanning. *Curr Opin Chem Biol.*
343 **5**, 302-307 (2001).
- 344 4. Howlader, M.T. et al. Alanine scanning analyses of the three major loops in domain II of
345 *Bacillus thuringiensis* mosquitocidal toxin Cry4Aa. *Appl Environ Microbiol.* **76**, 860-
346 865 (2010).
- 347 5. Zhao, L., et al. Destabilization of strigolactone receptor DWARF14 by binding of ligand
348 and E3-ligase signaling effector DWARF3. *Cell Res* **25**, 1219–1236 (2015).
- 349 6. Kobilka, B.K., et al. Chimeric alpha 2-,beta 2-adrenergic receptors: delineation of
350 domains involved in effector coupling and ligand binding specificity. *Science* **240**,
351 1310–16 (1988).
- 352 7. Cunningham, B. C., Jhurani, P., Ng, P. & Wells, J. A. Receptor and antibody epitopes in
353 human growth hormone identified by homolog-scanning mutagenesis. *Science* **243**,
354 1330–36 (1989).
- 355 8. Hochberg, G. K. A. & Thornton, J. W. Reconstructing Ancient Proteins to Understand
356 the Causes of Structure and Function. *Annu Rev Biophys.* **46**, 247-269 (2017).
- 357 9. Tsuchiya, Y., et al. PARASITIC PLANTS. Probing strigolactone receptors in *Striga*
358 *hermonthica* with fluorescence. *Science.* **349**, 864-868 (2015).
- 359 10. Scaffidi, A., et al. Strigolactone Hormones and Their Stereoisomers Signal through Two
360 Related Receptor Proteins to Induce Different Physiological Responses in Arabidopsis.
361 *Plant Physiol.* **165**, 1221-1232 (2014).
- 362 11. Bunsick M, et al. SMAX1-dependent seed germination bypasses GA signalling in
363 *Arabidopsis* and *Striga*. *Nat Plants.* **6**, 646-652 (2020).
- 364 12. Yoneyama, K., Xie, X., Yoneyama, K., et al. Which are the major players, canonical or
365 non-canonical strigolactones? *J Exp Bot.* **69**, 2231-2239 (2018).
- 366 13. Flematti, G. R., Scaffidi, A., Waters, M. T. & Smith, S. M. Stereospecificity in
367 strigolactone biosynthesis and perception. *Planta.* **243**,1361-1373 (2016).

- 368 14. Brewer, P. B., Koltai, H. & Beveridge C. A. Diverse roles of strigolactones in plant
369 development. *Mol Plant*. **6**, 18-28 (2013).
- 370 15. Waldie, T., McCulloch, H. & Leyser, O. Strigolactones and the control of plant
371 development: lessons from shoot branching. *Plant J*. **79**, 607-622 (2014).
- 372 16. Flematti, G. R., Ghisalberti, E. L., Dixon, K. W. & Trengove, R. D. A compound from
373 smoke that promotes seed germination. *Science* **305**, 977 (2004).
- 374 17. Conn, C. E. & Nelson, D. C. Evidence that KARRIKIN-INSENSITIVE2 (KAI2)
375 Receptors may Perceive an Unknown Signal that is not Karrikin or Strigolactone. *Front*
376 *Plant Sci*. **6**, 1219 (2016).
- 377 18. Conn, C.E., et al. PLANT EVOLUTION. Convergent evolution of strigolactone
378 perception enabled host detection in parasitic plants. *Science*. **349**, 540-543 (2015).
- 379 19. Xu, Y., et al. (2018). Structural analysis of HTL and D14 proteins reveals the basis for
380 ligand selectivity in *Striga*. *Nat Commun*. **9**, 3947 (2018).
- 381 20. Wang et al., *Plant Physiol*. 2020 in press
- 382 21. Spallek, T., Mutuku, M. & Shirasu K. The genus *Striga*: a witch profile. *Mol Plant*
383 *Pathol*. **14**, 861-869 (2013).
- 384 22. Toh, S., et al. Structure-function analysis identifies highly sensitive strigolactone
385 receptors in *Striga*. *Science*. **350**, 203-207 (2015).
- 386 23. Clough, S. J. & Bent, A. F. Floral dip: a simplified method for *Agrobacterium*-mediated
387 transformation of *Arabidopsis thaliana*. *Plant J*. **16**, 735-743 (1998).
- 388 24. Kagiya, M., et al. Structures of D14 and D14L in the strigolactone and karrikin
389 signaling pathways. *Genes Cells*. **18**, 147-160 (2013).
- 390 25. Guo Y., et al Smoke-derived karrikin perception by the α/β -hydrolase KAI2 from
391 *Arabidopsis*. *Proc Natl Acad Sci U S A*. **110**, 8284-8289. (2013).
- 392 26. Shahul Hameed, U., et al. Structural basis for specific inhibition of the highly sensitive
393 ShHTL7 receptor. *EMBO Rep*. **19**, e45619 (2018).
- 394 27. Hamiaux, C., et al. DAD2 is an α/β hydrolase likely to be involved in the perception of
395 the plant branching hormone, strigolactone. *Curr Biol*. **22**, 2032–2036. (2012).
- 396 28. de Saint Germain, A., et al. An histidine covalent receptor and butenolide complex
397 mediates strigolactone perception. *Nat Chem Biol*. **12**, 787–794 (2016).

- 398 29. Waters, M. T., Scaffidi, A., Flematti, G., & Smith, S. M. Substrate- induced degradation
399 of the α/β -fold hydrolase KARRIKIN INSENSITIVE2 requires a functional catalytic
400 triad but is independent of MAX2. *Mol Plant*. **8**, 814–817 (2015).
- 401 30. Toh S, et al. (2014). Detection of Parasitic Plant Suicide Germination Compounds Using
402 a High-Throughput *Arabidopsis* HTL/KAI2 Strigolactone Perception System. *Chem*
403 *Biol*. **2**, 988-998 (2014).
- 404 31. Yao, J., et al. An allelic series at the KARRIKIN INSENSITIVE 2 locus of *Arabidopsis*
405 *thaliana* decouples ligand hydrolysis and receptor degradation from downstream
406 signalling. *Plant J*. **96**, 75-89 (2018).
- 407 32. Yamamoto, Y. et al. A rice *gid1* suppressor mutant reveals that gibberellin is not always
408 required for interaction between its receptor, GID1, and DELLA proteins. *Plant Cell*. **22**,
409 3589-3602 (2010).
- 410 33. Waters, M. T., et al. Specialisation within the DWARF14 protein family confers distinct
411 responses to karrikins and strigolactones in *Arabidopsis*. *Development*. **139**, 1285-1295
412 (2012).
- 413 34. Starr, T. N. & Thornton, J. W. Epistasis in protein evolution. *Protein Sci*. **25**, 1204-1218
414 (2016).
- 415 35. Bürger, M., et al. Structural Basis of Karrikin and Non-natural Strigolactone Perception
416 in *Physcomitrella patens*. *Cell Rep*. **26**, 855-865 (2019).
- 417 36. Lee, H. W., et al. Flexibility of the petunia strigolactone receptor DAD2 promotes its
418 interaction with signaling partners. *J Biol Chem*. **295**, 4181-4193 (2020).
- 419 37. Harms, M. J., et al. Biophysical mechanisms for large-effect mutations in the evolution
420 of steroid hormone receptors. *Proc Natl Acad Sci U S A*. **110**, 11475-11480 (2013).
- 421 38. Tawfik, D. S. Messy biology and the origins of evolutionary innovations. *Nat Chem Biol*
422 **6**, 692–696 (2010).
- 423 39. Eick, G. N., et al. Evolution of minimal specificity and promiscuity in steroid hormone
424 receptors. *PLoS Genet*. **8**, e1003072 (2012).
- 425 40. Uraguchi, D., et al. (2018). A femtomolar-range suicide germination stimulant for the
426 parasitic plant *Striga hermonthica*. *Science*. **362**, 1301-1305 (2018).
- 427 41. Yao, R., et al. (2016). DWARF14 is a non-canonical hormone receptor for strigolactone.
428 *Nature* **536**, 469–473 (2016).

- 429 42. Shabek, N., et al. Structural plasticity of D3-D14 ubiquitin ligase in strigolactone
430 signalling. *Nature*. **563**, 652-656 (2018).
- 431 43. Seto, Y., et al. Strigolactone perception and deactivation by a hydrolase receptor
432 DWARF14. *Nat Commun*. **10**, 191 (2019).
- 433 44. Bloom JD, Arnold FH (2009) In the light of directed evolution: pathways of adaptive
434 protein evolution. *Proc Natl Acad Sci U S A* **106** Suppl 1, 9995–10000.
- 435 45. Khersonsky O, Tawfik DS (2010) Enzyme promiscuity: a mechanistic and evolutionary
436 perspective. *Annu Rev Biochem* **79**, 471–505.
- 437 46. Villaecija-Aguilar, J. A., et al. (2019) SMAX1/SMXL2 regulate root and root hair
438 development downstream of KAI2-mediated signalling in *Arabidopsis*. *PLoS Genetics*,
439 15 (8) (2019).
- 440 47. Akiyama, K., Matsuzaki, K. & Hayashi, H. Plant sesquiterpenes induce hyphal
441 branching in arbuscular mycorrhizal fungi. *Nature*. **435**, 824-827 (2005).
- 442 48. Ejeta, G. (2007) The *Striga* scourge in Africa: A growing pandemic. Integrating New
443 Technologies for *Striga* Control, pp. 3-16
- 444 49. Parker, C. Observations on the current status of *Orobanchae* and *Striga* problems
445 worldwide. *Pest Manag. Sci.* **65**, 453–459 (2009).

446 **Figure 1. Identification of a variant KAI2 that perceives SL.** **a.** Heatmap representation of
447 SL-dependence of the 93 KAI2 chimeric variants. Var1 is the KAI2 wild type control. Lanes 1-2;
448 Yeast two hybrid assays where increased purple shading represent an increased interaction
449 intensity on *rac*-GR24 relative to the DMSO control for KAI2 chimeric variants queried against
450 the N3 MAX2 fragments (**Supplementary Fig. 3**). Interaction strengths were quantified using
451 colony color intensity after 72 hours. Representative colonies for Var27 and Var64 are shown on
452 the left inserts. Lanes 3-4; the mean germination percentage of three independent KAI2 chimeric
453 variants lines on PAC (20 μ M), DMSO (0.1%) or *rac*-GR24 (1 μ M). Lane 5-12; Graphic
454 representation of 93 insertion lines. Each blue box represents the amino acid position in KAI2
455 where a ShHTL7 amino acid is substituted. Lane 13; degree of rescue of the *htl-3* leaf shape
456 phenotype by each KAI2 chimeric variant (**Supplementary Fig 5**). Colors are from 0% rescued
457 (yellow) to 100% rescued (dark green). Grey indicates that the variant was not available.
458 Representative pictures of *htl-3*, *35S::KAI2*, Var19 and Var64 plants. Are shown on the right
459 inserts. **b.** Scatterplot showing the log-fold change scores calculated for each KAI2 chimeric
460 variant with full-length protein, empty vector control and MAX2 fragments using a Y2H protein
461 interaction assay (see **Supplementary Fig. 3**). 22 scores were calculated for each variant. Red
462 dotted line represents four standard deviations from the mean. **c.** SL dose response germination
463 curves. Percent germination of mis-expressed *ShHTL7*, *KAI2* and three independent Var64 seed
464 (64A, 64B, 64C) on PAC (20 μ M) and increasing *rac*-GR24 concentrations. Four-parameter
465 logistic curves were fitted to the data, and EC₅₀ values, \pm standard error, of *rac*-GR24
466 concentrations have been included for those lines showing response to SL. Bases on these curves
467 the effective concentration that germinates 50% of seed EC₅₀ was calculated and is shown in the
468 box.

469

470 **Fig. 2. Conservation of Var64 amino acids in land plants.** **a.** Phylogenetic tree built using
471 *KAI2/DLK/DDK* sequences from across the land plant phylogeny, rooted using *KAI2* sequences
472 from hornworts. Sequences within each of the three coloured clades were investigated in more
473 detail for their similarity with variant 64 at positions 153, 157 and 190. **b.** *KAI2* conserved
474 sequences from angiosperms. **c.** *KAI2* divergent sequences from parasitic Orobanchaceae.
475 Coloured boxes beside the clades indicate amino acid similarity with variant 64 at positions 153

476 (blue), 157 (green) or 190 (purple). **James, we need a list of what the similar amino acids are**
477 **for each category!**

478

479 **Figure 3. A. Var64 responds to both GR24 enantiomers a.** Percent germination of *htl-3*, mis-
480 expressed *KAI2* and three independent Var64 seed lines (64A, 64B, 64C) on PAC (20 μ M) with
481 DMSO, 1 μ M (-)-(2'*S*)-GR24 or (+)-(2'*R*)-GR24. **b.** SL inhibits YLG hydrolysis. Competitive
482 binding assay using YLG and *rac*-GR24 to evaluate the changes in YLG hydrolysis by *KAI2*
483 (black) and Var64 (blue) after incubation with 0.5 μ M *rac*-GR24. Error bars represent SD values
484 from three independent experiments. **c.** Hydrolysis activity towards GR24 enantiomers. UPLC-
485 UV (260 nm) analysis was used to detect the remaining amount of GR24 isomers. Box plots of
486 $n=3$ replicates represent the hydrolysis rate calculated from the remaining GR24 isomers and
487 analogs taking into account the hydrolysis in the buffer alone (non protein sample), quantified
488 using indanol as internal standard. Letters indicate different statistical groups (ANOVA, post-hoc
489 Tukey test). **d.** KAR_2 dose response germination curves. Mis-expressed *ShHTL7*, *KAI2* and three
490 independent Var64 seed (64A, 64B, 64C) on PAC (20 μ M) and increasing KAR_2 concentrations.
491 Four-parameter logistic curves were fitted to the data, and EC_{50} values, \pm standard error, of
492 KAR_2 concentrations have been included for those lines showing response to karrikin. Based on
493 these curves the effective concentration that germinates 50% of seed (EC_{50}) was calculated and is
494 shown in the box. **Alex I need your data to be converted to box plots and I only need the**
495 **data for GR24. Also I need the stats included in the figure.**

496

497 **Figure 4. Var64 derived amino acids create a SL binding domain. a.** The substitution of
498 *KAI2* aromatic residues for *ShHTL7* equivalents results in the disruption of several interactions
499 (purple dashed lines) along the X-axis of the protein. This disruption translates into increased
500 motility of the lid domain (Supplemental Movie 1), consistent with the receptor's new ability to
501 accommodate SL molecules. **b.** Pocket alignment of *KAI2* and Var64 receptors reveals an 11.5
502 \AA larger pocket mouth (orange shading) and a 29.7% increased pocket volume in the Var64
503 receptor. The pocket mouths are shown in yellow shading with width shown in \AA^2 . **c.** A
504 summary of the three most common binding states between Var64, *KAI2*, and 2'*R*-GR24.
505 Molecular dynamic simulations of Var64 and *KAI2* receptors to determine binding modes of
506 ligand-protein complexes over a 10 ns simulation. Var64-(2'*R*)-GR24 complex was found in a

507 productive binding mode 47.8% of the time whilst the same state was only found 2.1% of the
508 time for the KAI2 complex. **d.** Calculation of the magnitude change in the pose of the ligand
509 throughout the MD simulation. Both proteins stabilized the ligand after 1 ns of simulation, but its
510 position on the KAI2 pocket is compromised from the 4 ns mark as the interaction is not strong
511 nor stable.

Figure 1

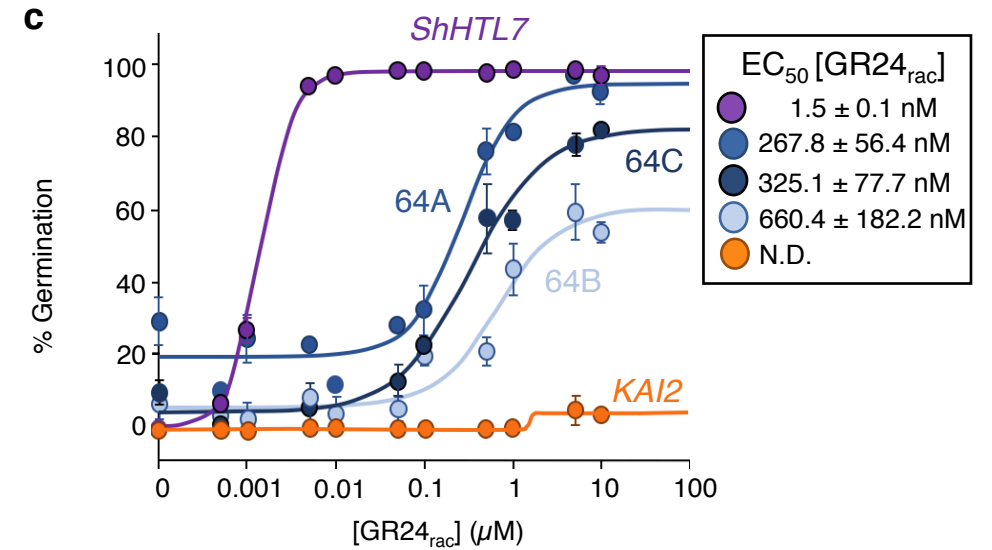
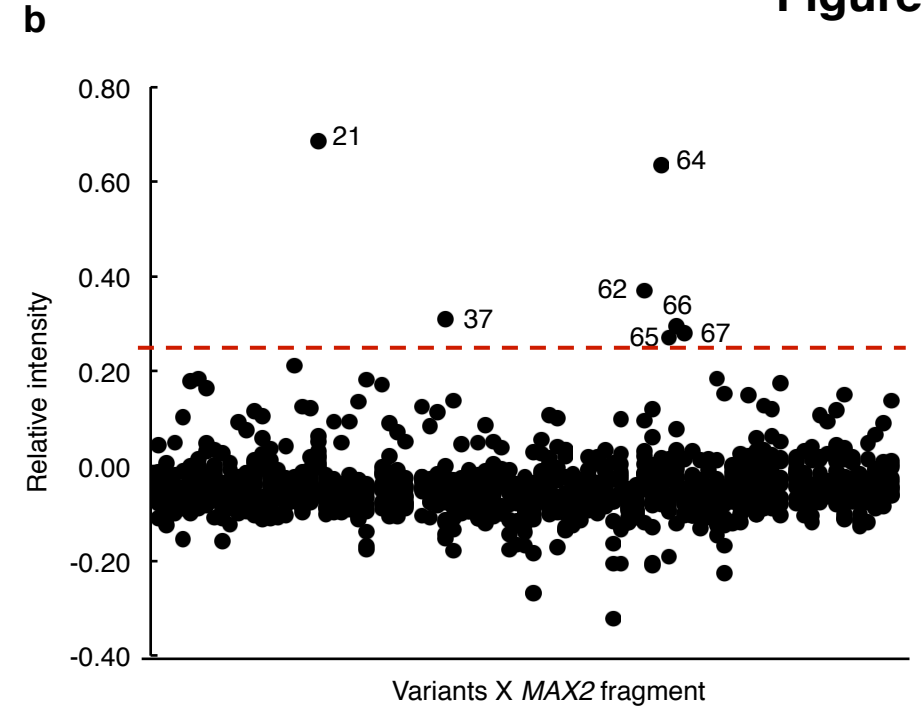
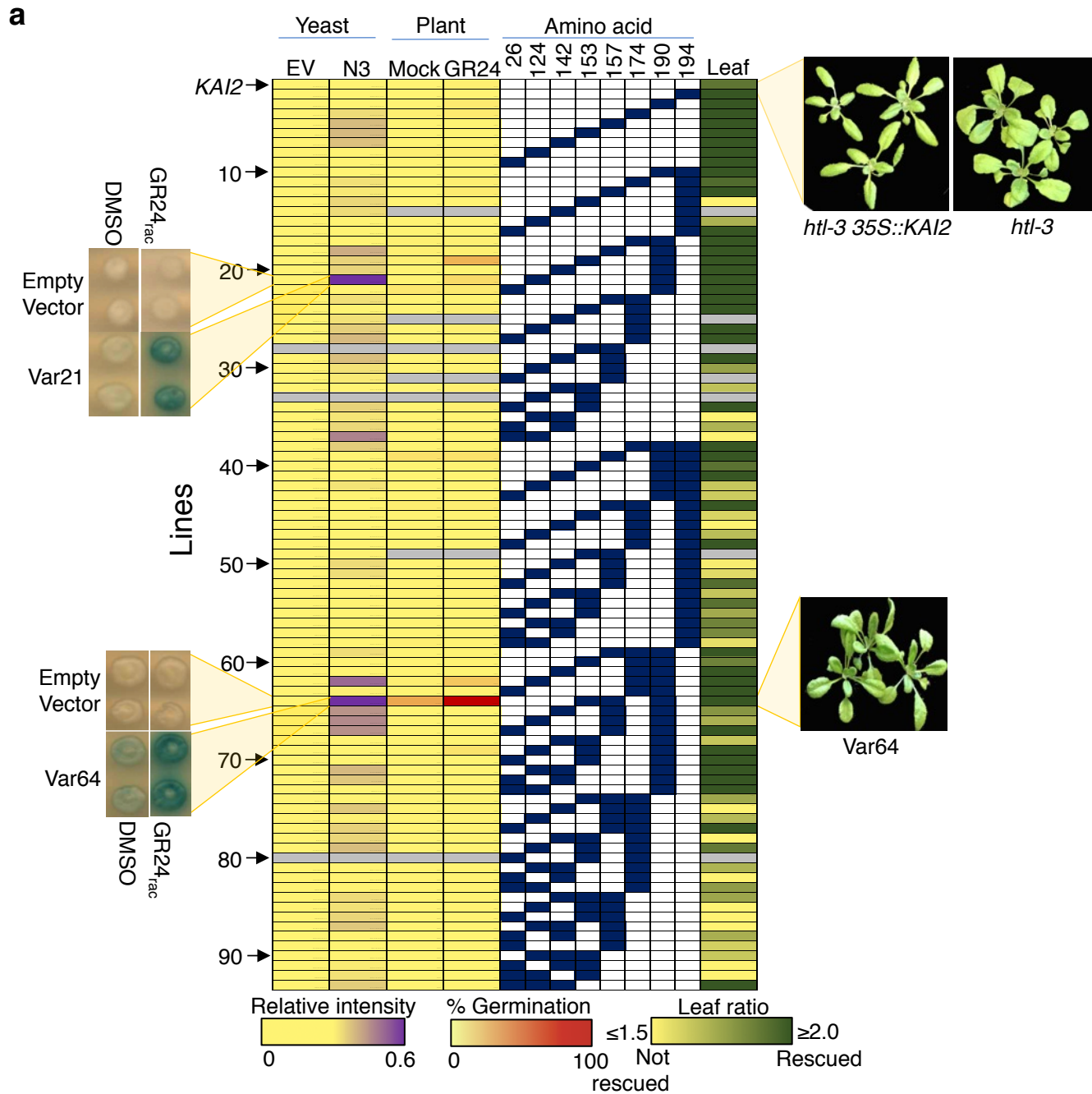


Figure 2

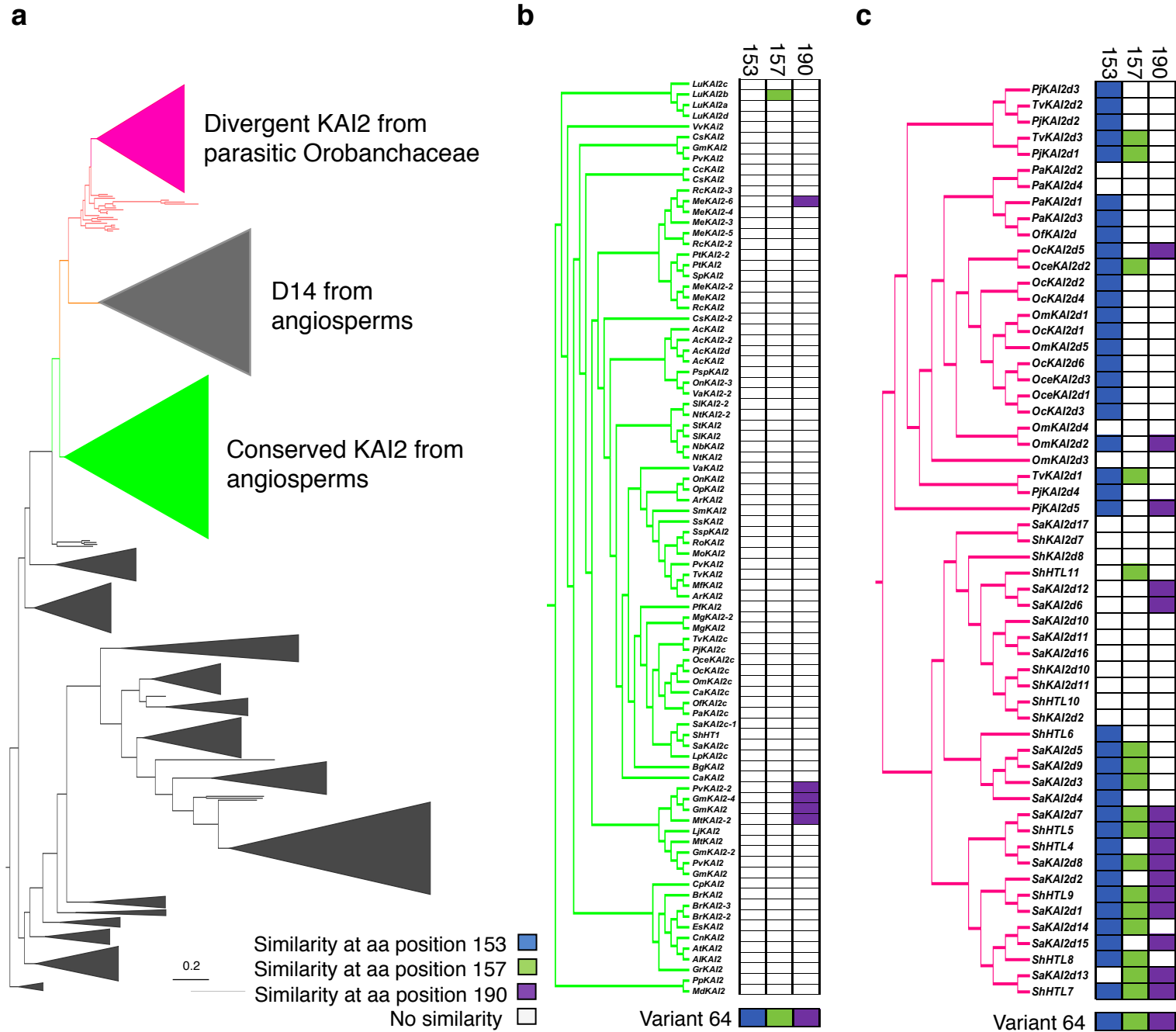
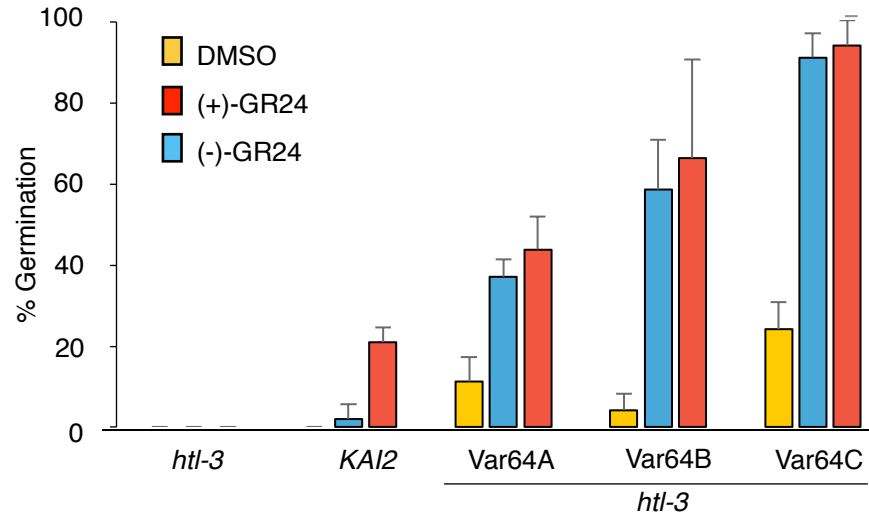
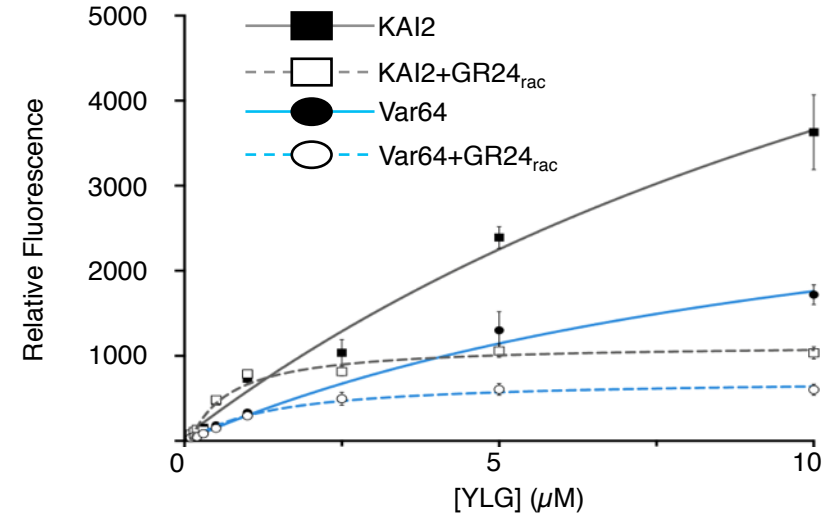


Figure 3

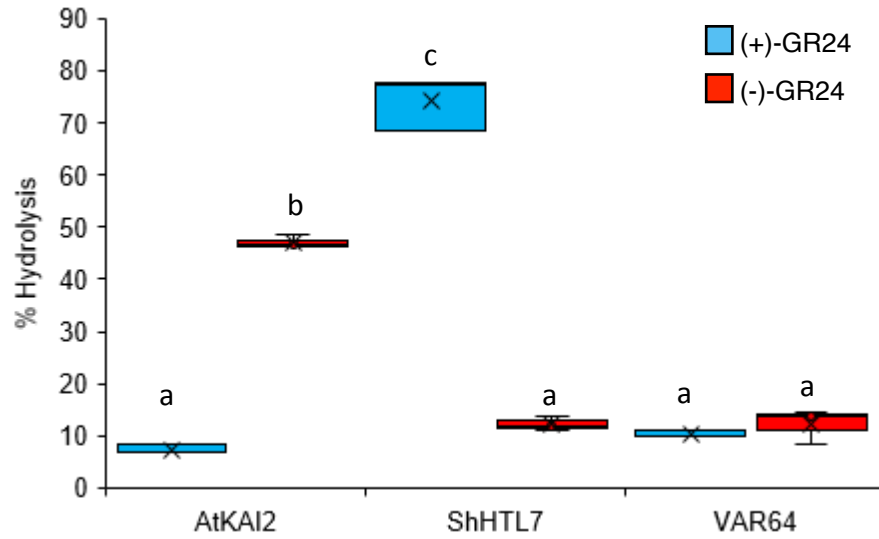
a



b



c



d

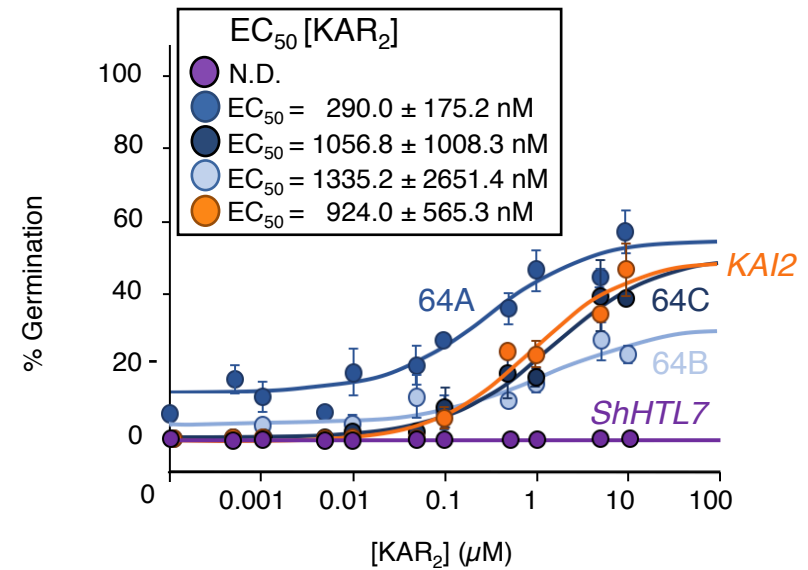
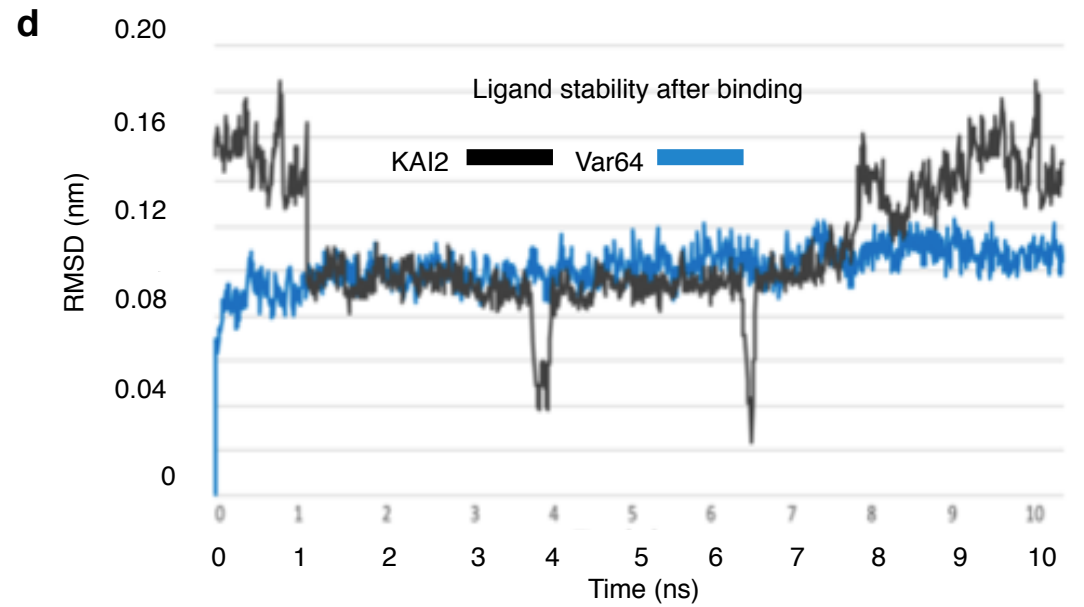
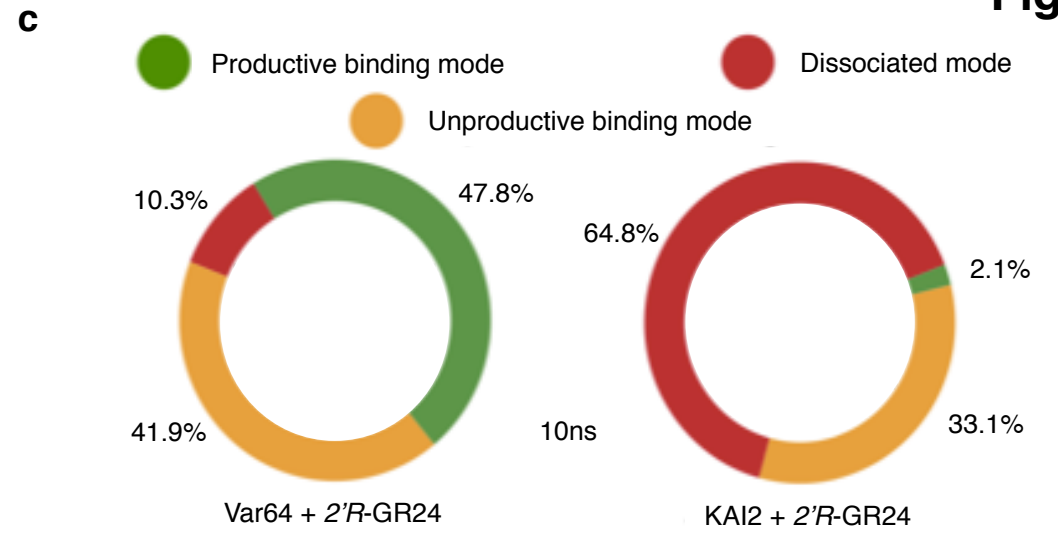
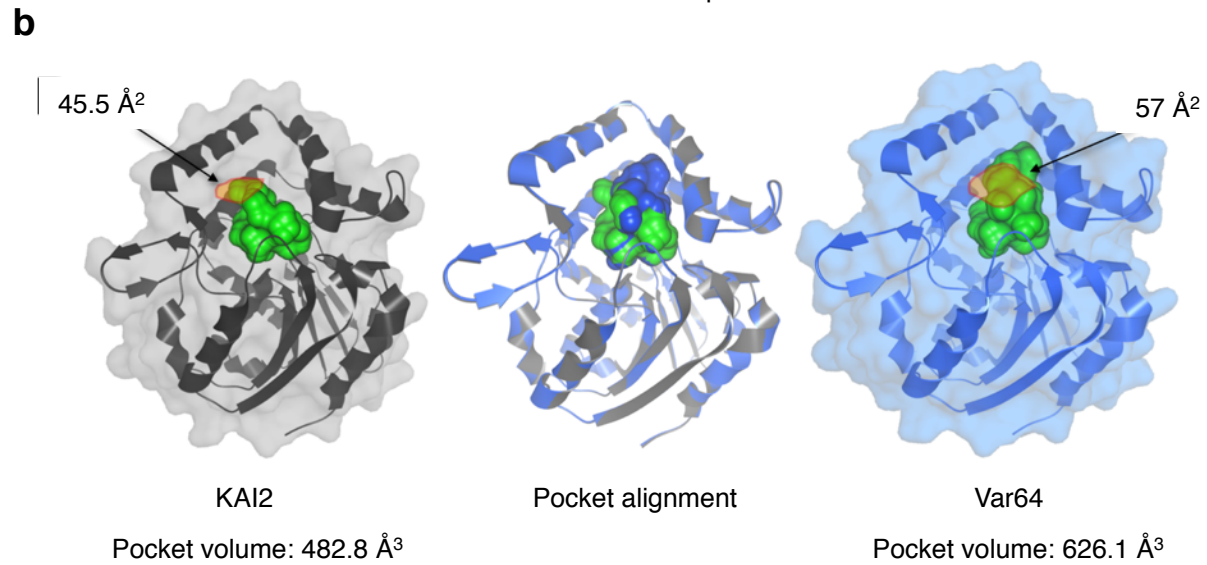
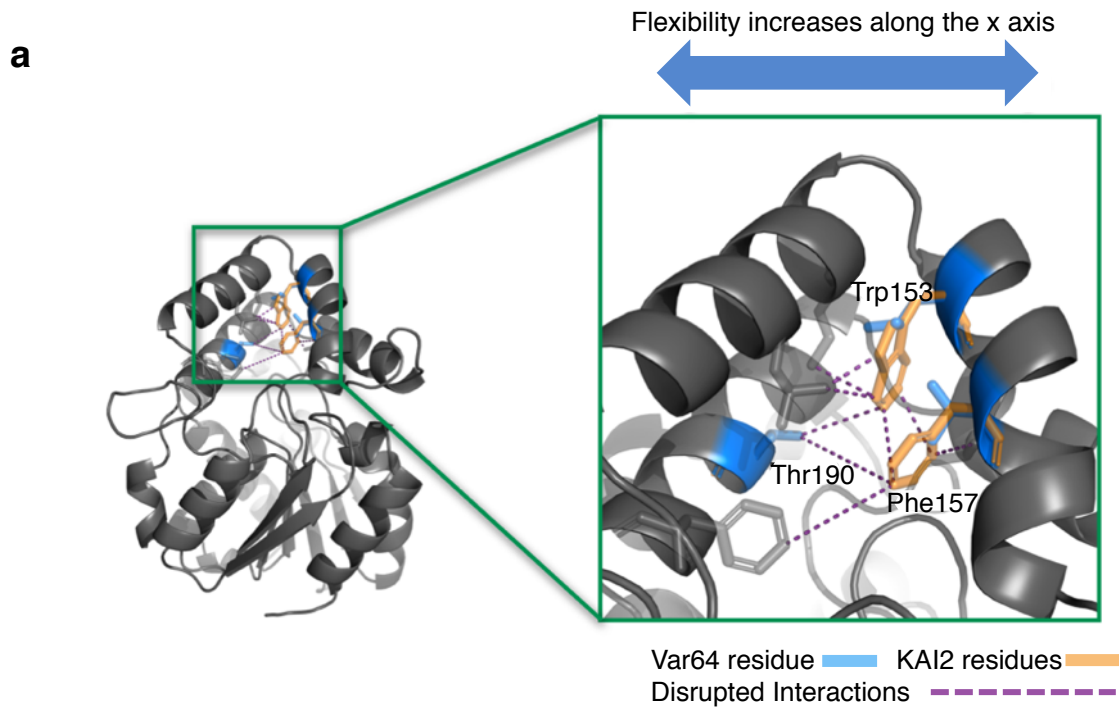
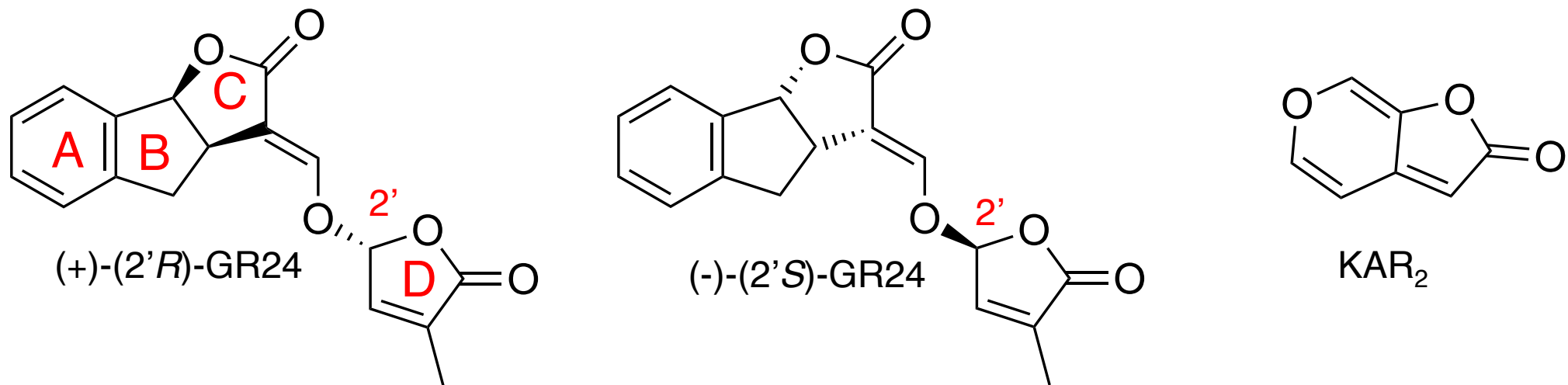


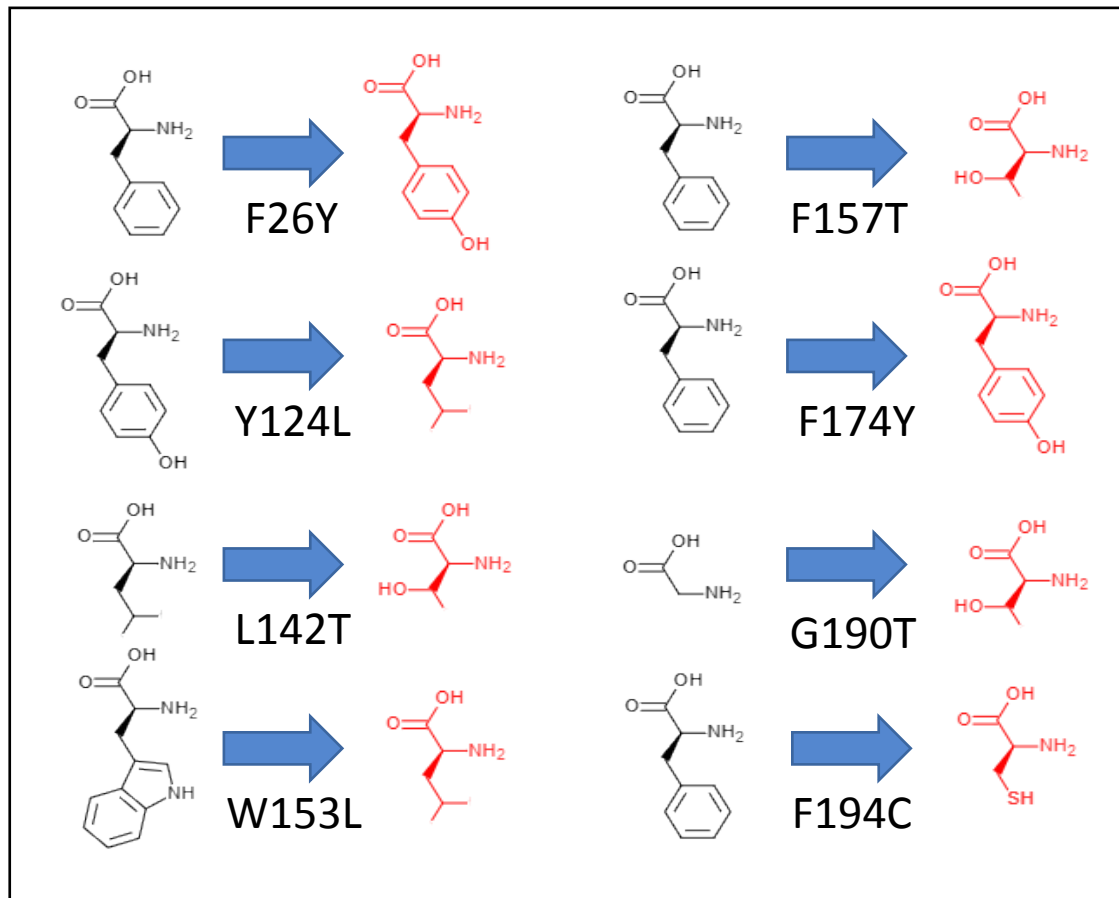
Figure 4



Supplementary Fig. 1. Structures of 2' *R*- and 2' *S*-enantiomeric forms of strigolactones (SLs) and karrikin (KAR₂). For SLs the synthetic GR24 is shown and the 2' chirality is at the enol-ether bridge between the C and D-ring. For karrikin KAR₂ is shown.

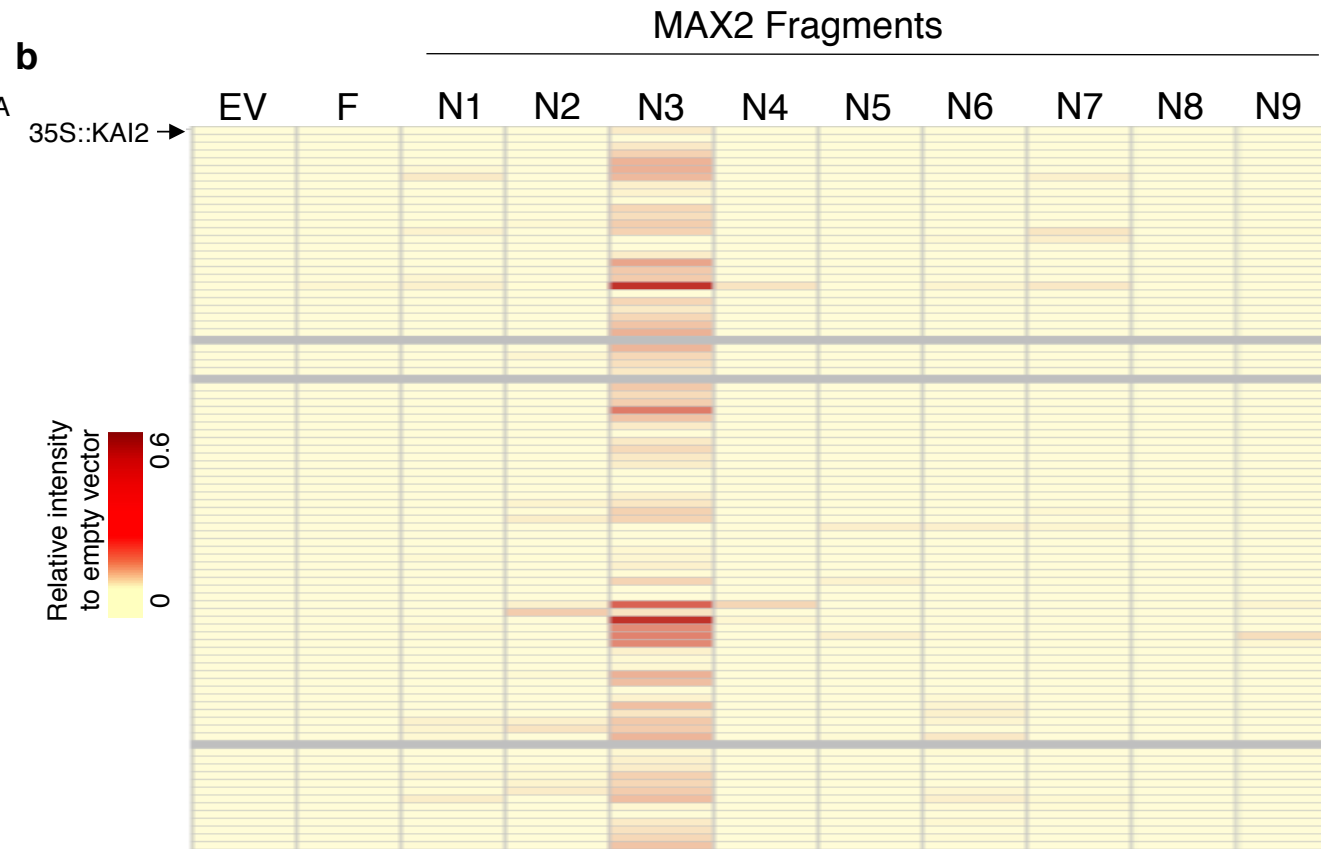
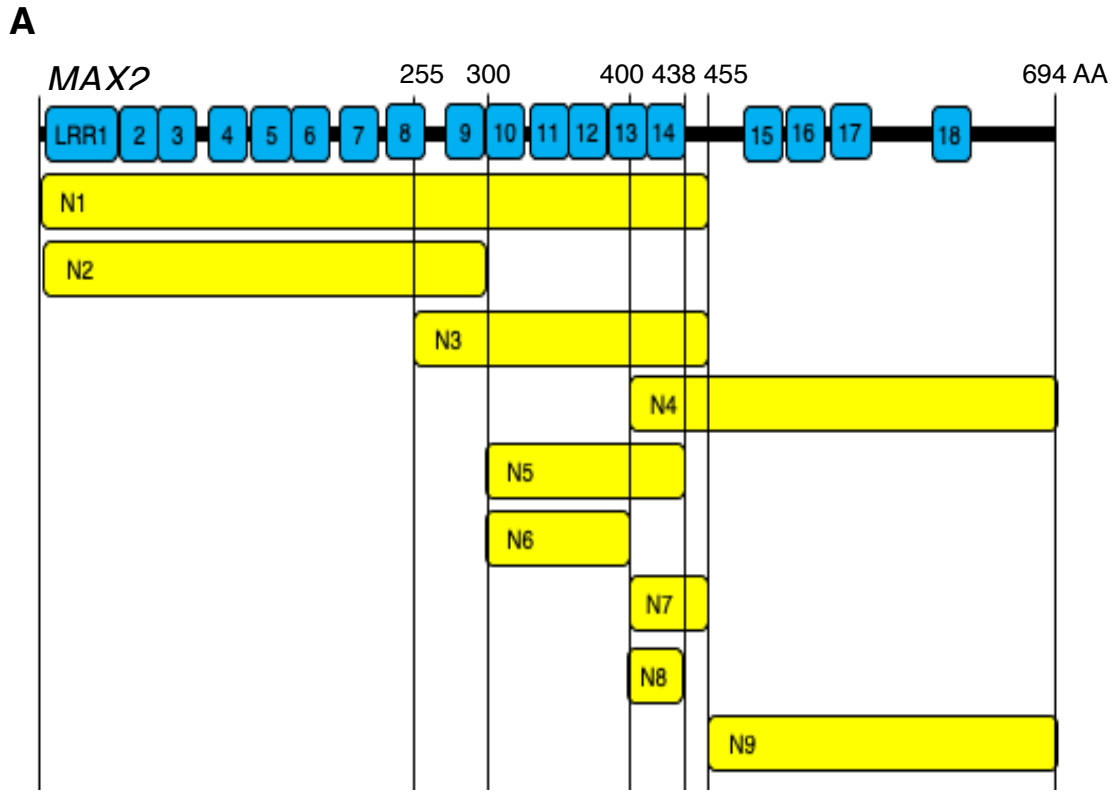
a

Residue	KAI2	ShHTL7
26	Phe	Tyr
95	Ser	Ser
121	Ser	Thr
124	Tyr	Leu
134	Phe	Phe
139	Leu	Met
142	Leu	Thr
143	Phe	Leu
146	Ile	Leu
153	Trp	Leu
154	Cys	Ser
157	Phe	Thr
161	Ala	Leu
174	Phe	Tyr
190	Gly	Thr
193	Ile	Ile
194	Phe	Cys
196	Ser	Leu
217	Asp	Asp
219	Ala	Met
246	His	His

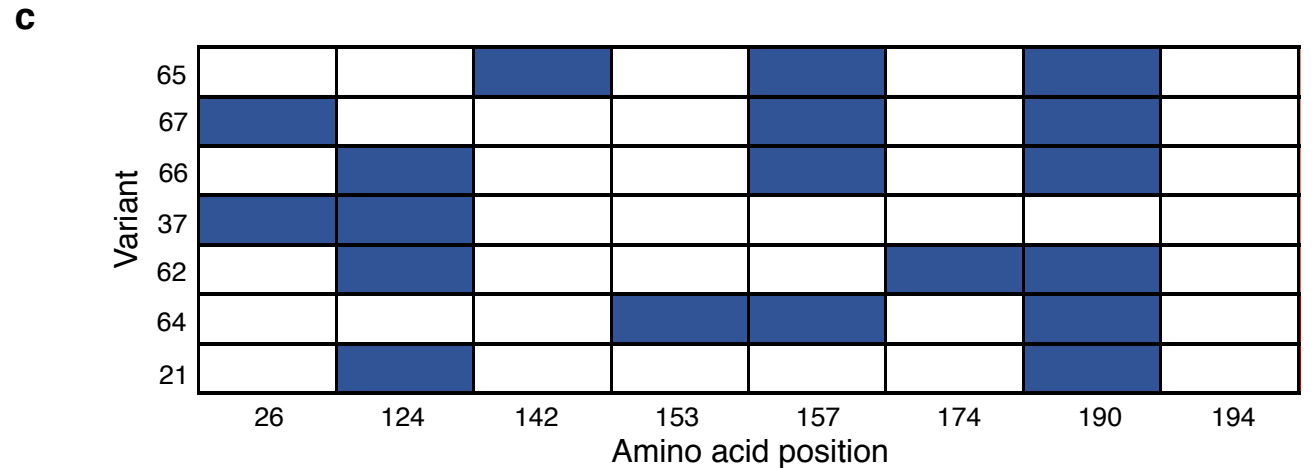
b**c**

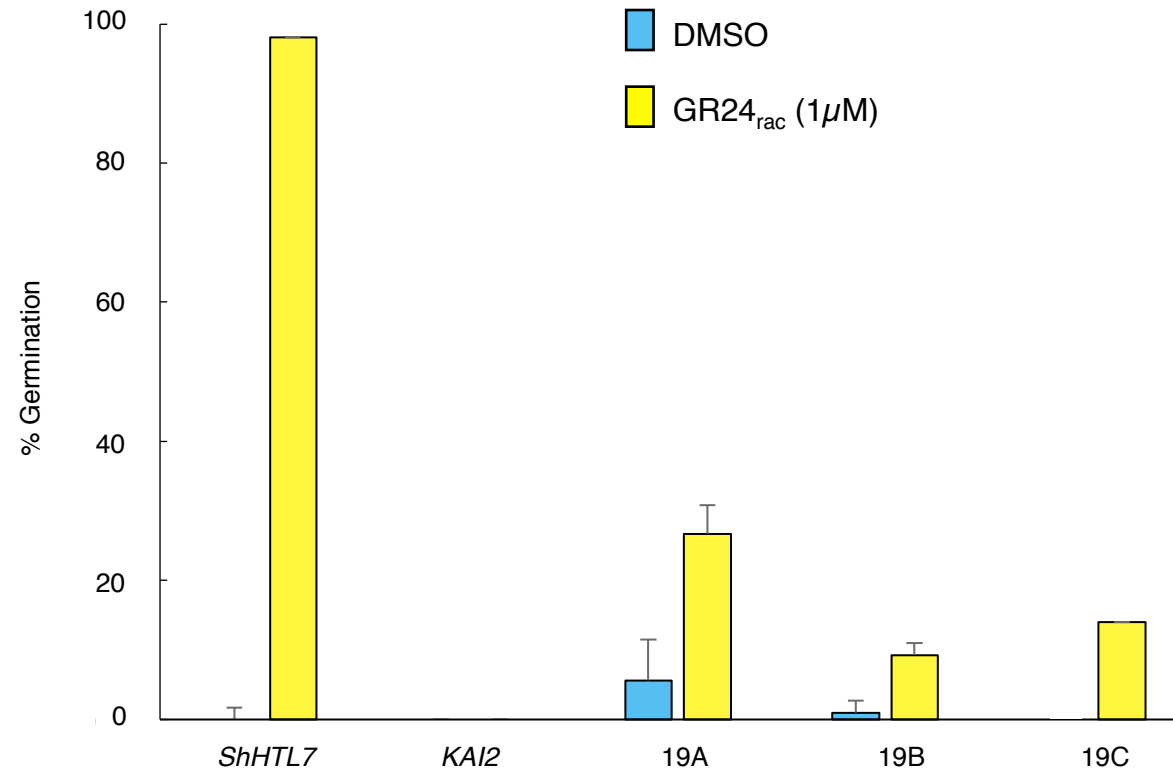
- Catalytic triad
- Polar side chain
- Non Polar side chain
- Electrically charged side chain

Supplementary Fig. 2 Selection of ShHTL7 residues for substitution in KAI2. **a.** The choice of the eight substitutions (pink boxes) were made based on increased polarity and decreased size (blue amino acids) in ShHTL7 versus the KAI2 equivalent. Red boxes represent the three catalytic residues. **b.** Chemical structures of the amino acids at the positions substituted in KAI2 to generate JGI variants. Structures in black represent the amino acids present at these positions in KAI2, structures in red represent their counterparts in ShHTL7. Letter codes and numbers below the arrows denote the substitutions that were made to produce the JGI variants **c.** Crystal models representing the eight residues highlighted in pink, The three catalytic residues highlighted in red, Structures are visualized using PYMOL version 2.3.2.

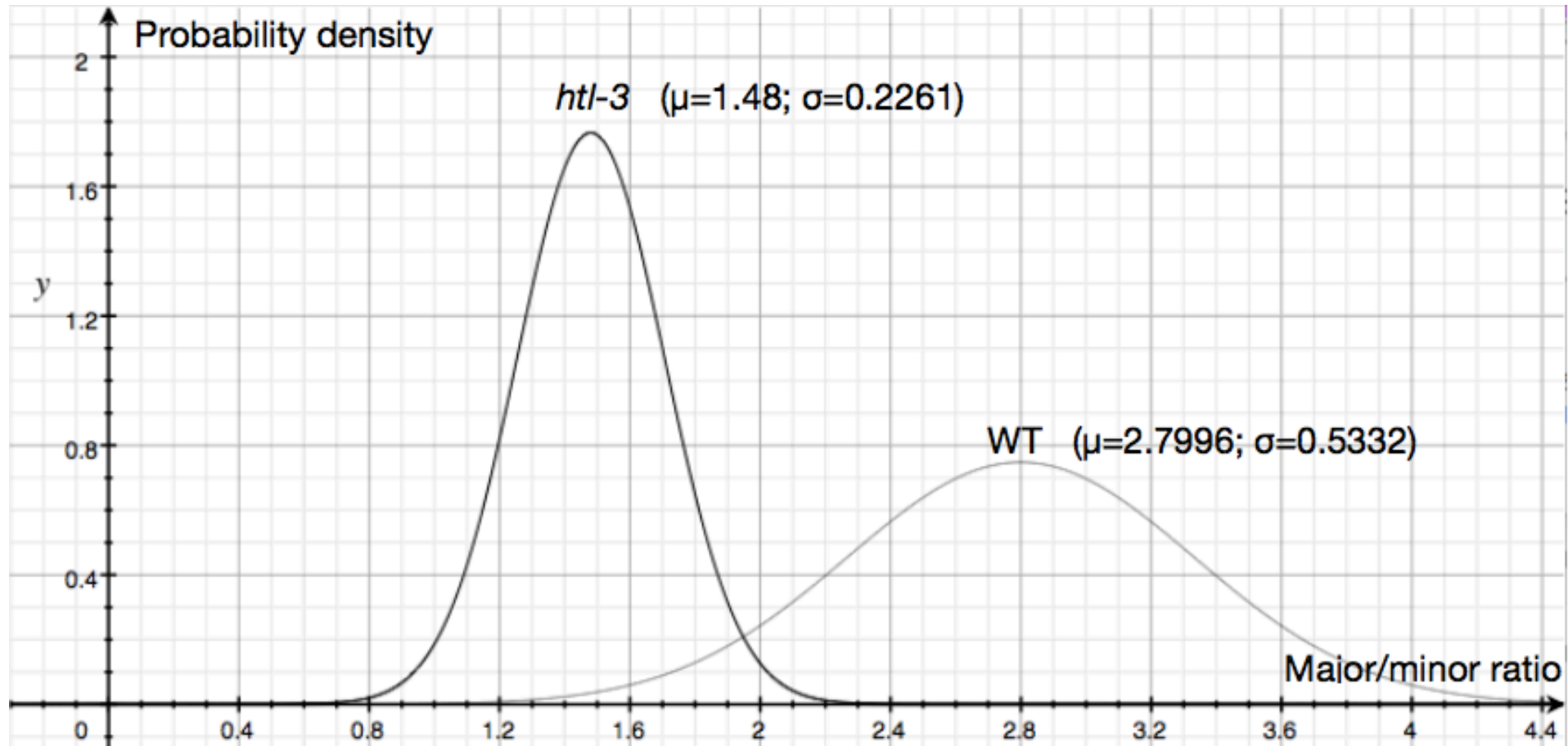


Supplementary Fig. 3 Interactions of KAI2 chimeric variants with MAX2 fragments using yeast-two-hybrid (Y2H) (a) Map of MAX2 fragments (N) used as Y2H preys. (b) Heatmap of log-fold change scores based on Y2H interactions relative to DMSO conditions for 1-, 2-, and 3-substitution variants queried against all MAX2 fragments. EV, empty vector, F, full-length protein. Redder shading indicates increased interaction. Grey lines represent construct that could not be cloned into Y2H vectors (c) Statistically significant hits in Y2H screen.





Supplementary Fig. 4 Var19 germination on SL. Mis-expressed *ShHTL7*, *KAI2* and three independent Var19 lines (19A, 19B, 19C) on PAC (20 µM) plus or minus Gr24_{rac} concentrations. Each line was tested three times Bar = S.D



Supplementary Fig. 5 Standard distribution of length/width (major/minor) ratio of *htl-3* and wild type Columbia (WT) leaves. Each distribution was a normal distribution based on the population mean (μ) and standard deviation (σ) approximated by that of the 26 control samples. WT: $y = (1 / (0.53332 \cdot \sqrt{2\pi})) \cdot e^{-(x - 2.7996)^2 / (2 \cdot (0.53332)^2)}$; *htl-3*: $y = (1 / (0.2261 \cdot \sqrt{2\pi})) \cdot e^{-(x - 1.48)^2 / (2 \cdot (0.2261)^2)}$. Three categories given by 5% cutoff: *htl-3*: $x < 1.85$; WT: $x > 1.92$; ambiguous: $1.85 < x < 1.92$. Based on the two distributions, in order to meet the 0.05 significance level, any sample data smaller than 1.85 would be considered as *htl-3*, while sample data larger than 1.92 would be considered as wild type. Note that there is an ambiguous region in between (1.82-1.92) where the phenotype can not be statistically determined in this case.