Perception of a conserved family of plant signalling peptides by the receptor kinase HSL3

Jack Rhodes[1], Andra-Octavia Roman[3], Marta Bjornson[2], Benjamin Brandt[2,5], Paul Derbyshire[1], Michele Wyler[4], Marc Schmid[4], Frank L.H. Menke[1], Julia Santiago[3] and Cyril Zipfel[1,2]

[1]The Sainsbury Laboratory, University of East Anglia, Norwich Research Park, NR4 7UH, Norwich, United Kingdom.

[2]Institute of Plant and Microbial Biology, Zurich-Basel Plant Science Center, University of Zurich, 8008 Zurich, Switzerland.

[3]The Plant Signaling Mechanisms Laboratory, Department of Plant Molecular Biology, University of Lausanne, 1015 Lausanne, Switzerland.

[4]MWSchmid GmbH, 8750 Glarus, Switzerland.

[5]Present address: Biozentrum der Ludwig-Maximilians-Universität München, Department Biologie I – Botanik, 82152 Planegg-Martinsried, Germany.

[*]Correspondence: Cyril Zipfel, cyril.zipfel@botinst.uzh.ch

Abstract

Plant genomes encode hundreds of secreted peptides; however, relatively few have been characterised. We report here an uncharacterised, stress-induced family of plant signalling peptides, which we call CTNIPs. Based on the role of the common co-receptor BRASSINOSTEROID INSENSITIVE 1-ASSOCIATED KINASE 1 (BAK1) in CTNIP-induced responses, we identified the orphan receptor kinase HAESA-LIKE 3 (HSL3) as the CTNIP receptor via a

proteomics approach. CTNIP binding, ligand-triggered complex formation with BAK1, and induced downstream responses all involve HSL3. Notably, the HSL3-CTNIP signalling module is evolutionarily ancient, predating the divergence of extant angiosperms. The identification of this signalling module will help establish its physiological role and provides a resource to understand further receptor-ligand co-evolution.

Introduction

Secreted plant peptides play major roles in growth, development and stress responses (Olsson et al., 2019). Whilst many hundreds of peptides are predicted to be encoded in plant genomes, relatively few have been characterised and their corresponding receptors are mostly unknown (Olsson et al., 2019).

Most signalling peptides are recognised by cell-surface localised receptors, especially by leucine-rich repeat receptor kinases (LRR-RKs). LRR-RKs generally function through the ligand-dependent recruitment of a shape complementary co-receptor to form an active signalling complex (Hohmann et al., 2017). The best characterised peptide receptors belong to LRR-RK subfamily XI, which recognize distinct families of plant peptides involved in growth, development or stress responses (Furumizu et al., 2021). Notably, the LRR-RK MIK2, which belongs to the closely related LRR-RK subfamily XIIb (an outgroup recently included within subfamily XI; (Liu et al., 2017; Man et al., 2020)) was recently shown to perceive SCOOP peptides (Hou et al., 2021; Rhodes et al., 2021). Despite intensive studies on the LRR-RK subfamily XI, the ligand for HAESA-like 3 (HSL3) has remained elusive, hindering our ability to investigate peptide-receptor coevolution across the family (Furumizu et al., 2021; Lee et al., 2020).

Results and discussion

Several peptides (PEPs, PIPs, SCOOPs, CLEs and IDLs) recognised by LRR-RKs from subfamily XI or XIIb are transcriptionally up-regulated by abiotic or biotic stresses (Bartels et al., 2013; Gully et al., 2019; Kim et al., 2021; Takahashi et al., 2018; Vie et al., 2015). In order to identify novel stress-induced signalling peptides, we searched for *Arabidopsis thaliana* (hereafter, Arabidopsis) transcripts encoding short proteins (<150 amino acids) with a predicted signal peptide, which were induced upon biotic elicitor treatment (Bjornson et al., 2021). Through this analysis, we identified an uncharacterised family of peptides with 5 predicted members, which we named CTNIP1 to 5 (pronounced catnip) based on relatively conserved residues within the peptides (Figure 1a-b; Figure 1-figure supplement 1a-b).

To determine whether CTNIPs function as signalling peptides, we synthetized peptides corresponding to the whole CTNIP proteins excluding the predicted signal peptide. CTNIP1-4 peptides were able to induce cytoplasmic Ca^{2+} influx and mitogen-activated protein kinase (MAPK) phosphorylation – hallmarks of peptide signalling (Figure 1c-e). However, a synthetic peptide derived from the divergent CTNIP5 peptide was inactive (Figure1-figure supplement 1a and 2c). Notably, the C-terminal 23 amino acids of CTNIP4, $CTNIP4^{48-70}$, were sufficient to induce responses (Figure1-figure supplement 2a-b), suggesting that the minimal bioactive peptide is contained within this region. Notably, this region contains two highly conserved cysteine residues (Fig 1b). Mutation of these cysteine residues revealed they are required for CTNIP4 activity (Figure1-figure supplement 2c). Going forward, we focused on CTNIP4 as a representative member of this peptide family, as its transcript was the most up-regulated upon elicitor treatment (Figure 1a).

We hypothesised that CTNIPs may be perceived by a cell-surface LRR-receptor. Typically LRR-receptors are dependent upon the SOMATIC EMBRYOGENESIS RECEPTOR KINASE (SERK) family of co-receptors (Hohmann et al., 2017). We therefore tested whether CTNIP-induced responses were affected in *bak1-5*, an allele of *BRASSINOSTEROID INSENSITIVE 1-ASSOCIATED KINASE 1* (*BAK1/SERK3*) that has a dominant-negative impact on SERK signalling (Perraki et al., 2018; Schwessinger et al., 2011). Concordant with perception by an LRR-receptor, we observed significantly impaired CTNIP4-induced reactive oxygen species production in *bak1-5* (Figure 1f-g).

Ligand-binding induces receptor-SERK heterodimerisation to activate signalling (Hohmann et al., 2017). To identify the CTNIP receptor, we therefore employed Arabidopsis lines expressing BAK1 tagged with green fluorescent protein (GFP) as a molecular bait to identify the CTNIP receptor. Using affinity-purification followed by mass spectrometry we looked for proteins specifically enriched into the BAK1 complex upon CTNIP4 treatment (Figure 2a) (Saur et al., 2016). In four independent biological replicates, the protein most enriched in the BAK1 complex upon CTNIP4 treatment was the LRR-RK HAESA-LIKE 3 (HSL3) (Figure 2b; Figure2-figure supplement 1; Supplementary file 1), making this a promising candidate for being the CTNIP receptor. We could independently confirm CTNIP-induced HSL3-BAK1 complex formation by co-immunoprecipitation (Figure 2c).

Consistent with a receptor function, the HSL3 ectodomain ($HSL3^{ECD}$, residues 22-627) heterologously-expressed in insect cells could directly bind CTNIP4 with a dissociation constant of ~4 µM in *in vitro* binding assays using isothermal titration calorimetry (Figure 2d-e; Figure2-figure supplement 2a-b). In the presence of CTNIP4, BAK1 strongly bound HSL3 with a dissociation constant in the mid-nanomolar range (~392 nM) (Figure 2d-f; Figure2-figure supplement 2b), consistent with its role as co-receptor. Furthermore, the two

96 conserved cysteine residues are required for receptor binding and co-receptor recruitment
97 explaining their loss of signalling activity (Figure 2d, g-h; Figure1-figure supplement 2c;
98 Figure2-figure supplement 2b).
99 Notably, we were unable to detect binding of INFLORESCENCE DEFICIENT IN ABSCISSION
100 (IDA), the ligand for the related receptors HAESA and HAESA-LIKE 2 (HSL2) (Meng et al., 2016;
101 Santiago et al., 2016), to $HSL3^{ECD}$ (Figure 2d; Figure2-figure supplement 2b), demonstrating
102 distinct ligand specificity. Accordingly, structural analysis of a $HSL3^{ECD}$ homology model
103 reveals that the HSL3 receptor lacks key conserved motifs required to recognise IDA peptides
104 (Figure2-figure supplement 3) (Santiago et al., 2016). Together, our data show that, while
105 HSL3 is phylogenetically related to HAE, HSL1 and HSL2, it perceives distinct peptides (*i.e.*
106 CTNIPs) most likely via different binding interfaces, which remain to be investigated in future
107 structural studies.
108 Having established biochemically that HSL3 is the CTNIP receptor, we tested its genetic
109 requirement for CNTIP-induced responses. As expected, we found that HSL3 is strictly
110 required for CTNIP-induced MAPK phosphorylation and whole genome transcriptional
111 reprogramming (Figure 3a-b; Figure3-figure supplement 1). Notably, whilst 30 min treatment
112 with 100 nM CTNIP4 led to differential expression of 1074 genes in wild-type Col-0, none were
113 differentially expressed in *hsl3-1* ($p<0.05$, $|Log2(FC)|>1$) (Fig 3b; Supplementary file 2).
114 We could additionally show that transient expression of HSL3 in *Nicotiana benthamiana* is
115 sufficient to confer responsiveness to CTNIPs (Figure 3c). Furthermore, whilst 500 nM CTNIP4
116 was unable to significantly inhibit growth in Col-0 seedlings, plants that over-express *HSL3*
117 became hypersensitive to active CTNIP4 (Figure 3 d-e; Figure3-figure supplement 2). Taken
118 together, our biochemical and genetic results demonstrate that HSL3 is the CTNIP receptor.
119

CTNIPs induce general early signalling outputs indicative of RK signalling, including cytoplasmic Ca^{2+} influx, MAPK phosphorylation and ROS production (Figure 1) (Olsson et al., 2019). In addition, CTNIP4 treatment induces significant HSL3-dependent transcriptional reprogramming (Figure 3b). Consistent with the up-regulation of *CTNIP* and *HSL3* expression by biotic elicitors (Figure 1a; Figure2-figure supplement 1), gene ontology analysis highlighted the enrichment of many defence- and stress-responsive pathways upon CTNIP4 treatment (Supplementary file 3). This is a pattern shared with other biotic elicitors (Figure3-figure supplement 3) indicative of a general stress response (Bjornson et al., 2021).

To investigate the biological consequence of HSL3 signalling, we fused the extracellular and transmembrane domains of BAK1-INTERACTING RECEPTOR-LIKE KINASE 3 (BIR3) to the cytoplasmic domain of HSL3 under the control of the *HSL3* promoter (Figure3-figure supplement 4a). This chimeric approach allows constitutive complex formation with SERKs, thus mimicking constitutive activation of an endogenous receptor kinase (Hohmann et al., 2020). Transgenic lines expressing this chimeric construct exhibited developmental defects, notably enhanced root curling (Figure 3f). Similarly, CTNIP4 treatment inhibited root growth and induced root skewing in a HSL3-dependent manner (Figure 3 g-i; Figure3-figure supplement 4c). In addition, *CTNIP4* overexpression, either with or without a C-terminal tag, was sufficient to induce a similar phenotype (Figure 3j-k; Figure3-figure supplement 4c). These data suggest that the HSL3-CTNIP signalling module modulates root development, similar to other LRR-RK subfamily XI signalling modules (Jeon et al., 2021; Jourquin et al., 2020).

Recent phylogenetic analyses indicate that HSL3 is conserved in angiosperms (Figure 4a; (Furumizu et al., 2021; Man et al., 2020). Having defined HLS3 as the CTNIP receptor, we

144 wondered whether CTNIPs were equally conserved. CTNIPs were identified in *Amborella*,
145 eudicots and early divergent monocots (Figure 4b-d).
146 Given the conservation of the HSL3-CTNIP signalling module, the lack of *At*CTNIP4 responses
147 in *N. benthamiana* suggests a co-evolution of ligand-receptor specificity, as for example
148 previously proposed for PLANT ELICITOR PEPTIDE (PEP)-PEP RECEPTOR (PEPR) pairs (Huffaker,
149 2015; Lori et al., 2015). Accordingly, *Medicago truncatula* HSL3 (*MtHSL3*) only induced a
150 cytoplasmic calcium influx upon treatment with a conspecific CTNIP (*Mt*CTNIP,
151 Medtr1g044470) (Figure 4e).
152 Our phylogenetic analysis however surprisingly revealed that no clear CTNIP could be found
153 in *Poaceae* genomes (Figure 4b-d). Interestingly, this absence is correlated with an expansion
154 of HSL3 paralogs within these genomes (Figure 4b). We can speculate that the HSL3-CTNIP
155 signalling module may have diverged considerably in this lineage. This is supported by the
156 divergence between eudicot and monocot CTNIPs (Figure 4c). Interestingly, over 40 % of the
157 CTNIPs identified were unannotated, including all monocot CTNIPs (Figure 4b), highlighting
158 how genome annotation still represents a significant challenge in the characterisation of
159 signalling peptides.
160

161 **Conclusion**

162 Here, we identified CTNIPs as a novel family of stress-induced signalling peptide. Using
163 affinity-purification and mass spectrometry based on ligand-induced association with the
164 BAK1 co-receptor, we identified the LRR-RK HSL3 as the CTNIP receptor. CTNIPs directly bind
165 the HSL3 ectodomain to promote BAK1 recruitment, and HSL3 is necessary and sufficient to
166 confer CTNIP perception. This ancient signalling module has been conserved for more than
167 180 million years (Furumizu et al., 2021; Kumar et al., 2017); however, its physiological role

remains elusive. HSL3 has recently been shown to play a role in regulating drought and disease resistance implicating HSL3 in multiple stress responses (Lee et al., 2020). Deorphanising HSL3 makes LRR-RK subfamily XI an exciting tool to understand receptor-ligand co-evolution and recognition specificity.

Acknowledgements

We thank the John Innes Centre Horticultural Services for plant care, especially T. Wells; M. Smoker, J. Taylor and A. Wawryk from the TSL Plant Transformation support group for plant transformation and all past and current members of the Zipfel and Santiago groups for technical help and fruitful discussions. N. Talbot is acknowledged for hosting J.R. for part of this study. This work was supported by the European Research Council under the Grant Agreements no. 773153 and no. 716358 (grant "IMMUNO-PEPTALK" to C.Z. and grant "WallWatchers" to J.S., respectively), The Gatsby Charitable Foundation (to C.Z.), the University of Zürich (to C.Z.), the Swiss National Science Foundation grants no. 31003A_182625 (to C.Z.) and no. 31003A_173101 (to J.S.), and the Fondation Philanthropique Famille Sandoz (to J.S.). M.B. was partially supported by the European Union's Horizon 2020 Research and Innovation Program under Marie Skłodowska-Curie Actions (grant agreement no. 703954).

References

Alexa A, Rahnenfuhrer J. 2021. topGO: Enrichment Analysis for Gene Ontology.

Andrews S, Krueger F, Seconds-Pichon A, Biggins F, Wingett S. 2015. FastQC.: A quality control tool for high throughput sequence data. *Babraham Inst*.

https://www.bioinformatics.babraham.ac.uk/projects/fastqc/%0Ahttp://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/

Bartels S, Lori M, Mbengue M, van Verk M, Klauser D, Hander T, Boni R, Robatzek S, Boller T. 2013. The family of Peps and their precursors in Arabidopsis: differential expression and localization but similar induction of pattern-triggered immune responses. *J Exp Bot* **64**:5309–5321. doi:10.1093/jxb/ert330

Bjornson M, Pimprikar P, Nürnberger T, Zipfel C. 2021. The transcriptional landscape of Arabidopsis thaliana pattern-triggered immunity. *Nat Plants 2021 75* **7**:579–586. doi:10.1038/s41477-021-00874-5

Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. 2008. Fast unfolding of communities in large networks. *J Stat Mech Theory Exp* **2008**:P10008. doi:10.1088/1742-5468/2008/10/P10008

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**:2114–2120. doi:10.1093/bioinformatics/btu170

Buchfink B, Xie C, Huson DH. 2014. Fast and sensitive protein alignment using DIAMOND. *Nat Methods 2014 121* **12**:59–60. doi:10.1038/nmeth.3176

Butenko MA, Patterson SE, Grini PE, Stenvik G-E, Amundsen SS, Mandal A, Aalen RB. 2003. Inflorescence deficient in abscission controls floral organ abscission in Arabidopsis and identifies a novel family of putative ligands in plants. *Plant Cell* **15**:2296–307. doi:10.1105/tpc.014365

Castel B, Tomlinson L, Locci F, Yang Y, Jones JDG. 2019. Optimization of T-DNA architecture for Cas9-mediated mutagenesis in Arabidopsis. *PLoS One* **14**:e0204778. doi:10.1371/JOURNAL.PONE.0204778

Cho SK, Larue CT, Chevalier D, Wang H, Jinn T-L, Zhang S, Walker JC. 2008. Regulation of

floral organ abscission in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **105**:15629–34. doi:10.1073/pnas.0805539105

Crook AD, Willoughby AC, Hazak O, Okuda S, VanDerMolen KR, Soyars CL, Cattaneo P, Clark NM, Sozzani R, Hothorn M, Hardtke CS, Nimchuk ZL. 2020. BAM1/2 receptor kinase signaling drives CLE peptide-mediated formative cell divisions in Arabidopsis roots. *Proc Natl Acad Sci* **117**:32750–32756. doi:10.1073/PNAS.2018565117

Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: A sequence logo generator. *Genome Res* **14**:1188–1190. doi:10.1101/gr.849004

Csardi G, Nepusz T. 2006. The igraph software package for complex network research. *InterJournal Complex Syst* **Complex Sy**:1695.

Cutler SR, Ehrhardt DW, Griffitts JS, Somerville CR. 2000. Random GFP::cDNA fusions enable visualization of subcellular structures in cells of Arabidopsis at a high frequency. *Proc Natl Acad Sci U S A* **97**:3718–3723. doi:10.1073/pnas.97.7.3718

Doblas VG, Smakowska-Luzan E, Fujita S, Alassimone J, Barberon M, Madalinski M, Belkhadir Y, Geldner N. 2017. Root diffusion barrier control by a vasculature-derived peptide binding to the SGN3 receptor. *Science (80-)* **355**:280–284. doi:10.1126/science.aaj1562

Doll NM, Royek S, Fujita S, Okuda S, Chamot S, Stintzi A, Widiez T, Hothorn M, Schaller A, Geldner N, Ingram G. 2020. A two-way molecular dialogue between embryo and endosperm is required for seed development. *Science (80-)* **367**:431–435. doi:10.1126/science.aaz4131

Eddy SR. 2011. Accelerated Profile HMM Searches. *PLOS Comput Biol* **7**:e1002195. doi:10.1371/JOURNAL.PCBI.1002195

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**:1792–7. doi:10.1093/nar/gkh340

240 Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative
241 genomics. *Genome Biol 2019 201* **20**:1–14. doi:10.1186/S13059-019-1832-Y
242 Furumizu C, Krabberød AK, Hammerstad M, Alling RM, Wildhagen M, Sawa S, Aalen RB.
243 2021. The sequenced genomes of nonflowering land plants reveal the innovative
244 evolutionary history of peptide signaling. *Plant Cell*. doi:10.1093/PLCELL/KOAB173
245 Futatsumori-Sugai M, Tsumoto K. 2010. Signal peptide design for improving recombinant
246 protein secretion in the baculovirus expression vector system. *Biochem Biophys Res*
247 *Commun* **391**:931–935. doi:10.1016/j.bbrc.2009.11.167
248 Gu Z, Eils R, Schlesner M. 2016. Complex heatmaps reveal patterns and correlations in
249 multidimensional genomic data. *Bioinformatics* **32**:2847–2849.
250 doi:10.1093/bioinformatics/btw313
251 Gully K, Pelletier S, Guillou M-C, Ferrand M, Aligon S, Pokotylo I, Perrin A, Vergne E, Fagard
252 M, Ruelland E, Grappin P, Bucher E, Renou J-P, Aubourg S. 2019. The SCOOP12 peptide
253 regulates defense response and root elongation in *Arabidopsis thaliana*. *J Exp Bot*
254 **70**:1349–1365. doi:10.1093/jxb/ery454
255 Harrell Jr FE. 2021. Hmisc: Harrell Miscellaneous.
256 Hashimoto Y, Zhang Sheng, Zhang Shiying, Chen Y-R, Blissard GW. 2012. Erratum to: BTI-
257 Tnao38, a new cell line derived from Trichoplusia ni, is permissive for AcMNPV
258 infection and produces high levels of recombinant proteins. *BMC Biotechnol 2012 121*
259 **12**:1–4. doi:10.1186/1472-6750-12-12
260 Hohmann U, Lau K, Hothorn M. 2017. The Structural Basis of Ligand Perception and Signal
261 Activation by Receptor Kinases. *Annu Rev Plant Biol* **68**. doi:10.1146/annurev-arplant-
262 042916-040957
263 Hohmann U, Ramakrishna P, Wang K, Lorenzo-Orts L, Nicolet J, Henschen A, Barberon M,

264 Bayer M, Hothorn M. 2020. Constitutive Activation of Leucine-Rich Repeat Receptor
265 Kinase Signaling Pathways by BAK1-INTERACTING RECEPTOR-LIKE KINASE3 Chimera.
266 *Plant Cell* **32**:3311–3323. doi:10.1105/tpc.20.00138
267 Hou S, Liu D, Huang S, Luo D, Liu Z, Xiang Q, Wang P, Mu R, Han Z, Chen S, Chai J, Shan L, He
268 P. 2021. The Arabidopsis MIK2 receptor elicits immunity by sensing a conserved
269 signature from phytocytokines and microbes. *Nat Commun 2021 121* **12**:1–15.
270 doi:10.1038/s41467-021-25580-w
271 Hou S, Wang X, Chen D, Yang X, Wang M, Turrà D, Di Pietro A, Zhang W. 2014. The secreted
272 peptide PIP1 amplifies immunity through receptor-like kinase 7. *PLoS Pathog*
273 **10**:e1004331. doi:10.1371/journal.ppat.1004331
274 Huffaker A. 2015. Plant elicitor peptides in induced defense against insects. *Curr Opin Insect*
275 *Sci*. doi:10.1016/j.cois.2015.06.003
276 Jeon BW, Kim M-J, Pandey SK, Oh E, Seo PJ, Kim J. 2021. Recent advances in peptide
277 signaling during Arabidopsis root development. *J Exp Bot* **72**:2889–2902.
278 doi:10.1093/JXB/ERAB050
279 Jourquin J, Fukaki H, Beeckman T. 2020. Peptide-Receptor Signaling Controls Lateral Root
280 Development. *Plant Physiol* **182**:1645–1656. doi:10.1104/PP.19.01317
281 Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate
282 alignment of transcriptomes in the presence of insertions, deletions and gene fusions.
283 *Genome Biol* **14**:R36. doi:10.1186/gb-2013-14-4-r36
284 Kim JS, Jeon BW, Kim J. 2021. Signaling Peptides Regulating Abiotic Stress Responses in
285 Plants. *Front Plant Sci* **12**:704490. doi:10.3389/FPLS.2021.704490
286 Knight MR, Campbell AK, Smith SM, Trewavas AJ. 1991. Transgenic plant aequorin reports
287 the effects of touch and cold-shock and elicitors on cytoplasmic calcium. *Nature*

352:524–526. doi:10.1038/352524a0

Krol E, Mentzel T, Chinchilla D, Boller T, Felix G, Kemmerling B, Postel S, Arents M, Jeworutzki E, Al-Rasheid KAS, Becker D, Hedrich R. 2010. Perception of the *Arabidopsis* Danger Signal Peptide 1 Involves the Pattern Recognition Receptor *At* PEPR1 and Its Close Homologue *At* PEPR2. *J Biol Chem* **285**:13471–13479. doi:10.1074/jbc.M109.097394

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* **35**:1547–1549. doi:10.1093/molbev/msy096

Kumar S, Stecher G, Suleski M, Hedges SB. 2017. TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Mol Biol Evol* **34**:1812–1819. doi:10.1093/molbev/msx116

Lee Y, Lemtiri-Chlieh F, Wang M, Zhang W, Liu X, Liang C, Hou S, Wang X, Chen D, Shen J. 2020. The LRR-RLK Protein HSL3 Regulates Stomatal Closure and the Drought Stress Response by Modulating Hydrogen Peroxide Homeostasis. doi:10.3389/fpls.2020.548034

Lemoine F, Gascuel O. 2021. Gotree/Goalign: toolkit and Go API to facilitate the development of phylogenetic workflows. *NAR Genomics Bioinforma* **3**. doi:10.1093/NARGAB/LQAB075

Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* **47**:W256–W259. doi:10.1093/nar/gkz239

Liao Y, Smyth GK, Shi W. 2019. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res* **47**:e47–e47. doi:10.1093/NAR/GKZ114

312 Liu PL, Du L, Huang Y, Gao SM, Yu M. 2017. Origin and diversification of leucine-rich repeat
313 receptor-like protein kinase (LRR-RLK) genes in plants. *BMC Evol Biol* **17**:1–16.
314 doi:10.1186/s12862-017-0891-5
315 Lori M, Van Verk MC, Hander T, Schatowitz H, Klauser D, Flury P, Gehring CA, Boller T,
316 Bartels S. 2015. Evolutionary divergence of the plant elicitor peptides (Peps) and their
317 receptors: Interfamily incompatibility of perception but compatibility of downstream
318 signalling. *J Exp Bot* **66**:5315–5325. doi:10.1093/jxb/erv236
319 Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for
320 RNA-seq data with DESeq2. *Genome Biol* **15**. doi:10.1186/s13059-014-0550-8
321 Man J, Gallagher JP, Bartlett M. 2020. Structural evolution drives diversification of the large
322 LRR-RLK gene family. *New Phytol* nph.16455. doi:10.1111/nph.16455
323 Meng X, Zhou J, Tang J, Li B, de Oliveira MV V, Chai J, He P, Shan L. 2016. Ligand-Induced
324 Receptor-like Kinase Complex Regulates Floral Organ Abscission in Arabidopsis. *Cell Rep*
325 **14**:1330–1338. doi:10.1016/j.celrep.2016.01.023
326 Morita J, Kato K, Nakane T, Kondo Y, Fukuda H, Nishimasu H, Ishitani R, Nureki O. 2016.
327 Crystal structure of the plant receptor-like kinase TDR in complex with the TDIF
328 peptide. *Nat Commun* **7**:12383. doi:10.1038/ncomms12383
329 Mou S, Zhang X, Han Z, Wang J, Gong X, Chai J. 2017. CLE42 binding induces PXL2 interaction
330 with SERK2. *Protein Cell* **8**:612–617. doi:10.1007/s13238-017-0435-1
331 Nakayama T, Shinohara H, Tanaka M, Baba K, Ogawa-Ohnishi M, Matsubayashi Y. 2017. A
332 peptide hormone required for Casparian strip diffusion barrier formation in Arabidopsis
333 roots. *Science (80-)* **355**.
334 Ntoukakis V, Schwessinger B, Segonzac C, Zipfel C. 2011. Cautionary notes on the use of C-
335 terminal BAK1 fusion proteins for functional studies. *Plant Cell* **23**:3871–8.

doi:10.1105/tpc.111.090779

Ogawa M, Shinohara H, Sakagami Y, Matsubayash Y. 2008. Arabidopsis CLV3 peptide directly binds CLV1 ectodomain. *Science (80-)* **319**:294. doi:10.1126/science.1150083

Okuda S, Fujita S, Moretti A, Hohmann U, Doblas VG, Ma Y, Pfister A, Brandt B, Geldner N, Hothorn M. 2020. Molecular mechanism for the recognition of sequencedivergent CIF peptides by the plant receptor kinases GSO1/SGN3 and GSO2. *Proc Natl Acad Sci U S A* **117**:2693–2703. doi:10.1073/pnas.1911553117

Olsson V, Joos L, Zhu S, Gevaert K, Butenko MA, De Smet I. 2019. Look Closely, the Beautiful May Be Small: Precursor-Derived Peptides in Plants. *Annu Rev Plant Biol* **70**:153–186. doi:10.1146/annurev-arplant-042817-040413

Ou Y, Lu X, Zi Q, Xun Q, Zhang J, Wu Y, Shi H, Wei Z, Zhao B, Zhang X, He K, Gou X, Li C, Li J. 2016. RGF1 INSENSITIVE 1 to 5, a group of LRR receptor-like kinases, are essential for the perception of root meristem growth factor 1 in Arabidopsis thaliana. *Cell Res* **26**:686–698. doi:10.1038/cr.2016.63

Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M, Pérez E, Uszkoreit J, Pfeuffer J, Sachsenberg T, Yılmaz Ş, Tiwary S, Cox J, Audain E, Walzer M, Jarnuczak AF, Ternent T, Brazma A, Vizcaíno JA. 2019. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* **47**:D442–D450. doi:10.1093/NAR/GKY1106

Perraki A, DeFalco TA, Derbyshire P, Avila J, Séré D, Sklenar J, Qi X, Stransfeld L, Schwessinger B, Kadota Y, Macho AP, Jiang S, Couto D, Torii KU, Menke FLH, Zipfel C. 2018. Phosphocode-dependent functional dichotomy of a common co-receptor in plant signalling. *Nature*. doi:10.1038/s41586-018-0471-x

Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 – Approximately Maximum-Likelihood Trees

360 for Large Alignments. *PLoS One* **5**:e9490. doi:10.1371/JOURNAL.PONE.0009490

361 Qian P, Song W, Yokoo T, Minobe A, Wang G, Ishida T, Sawa S, Chai J, Kakimoto T. 2018. The

362 CLE9/10 secretory peptide regulates stomatal and vascular development through

363 distinct receptors. *Nat Plants* **4**:1071–1081. doi:10.1038/s41477-018-0317-4

364 Rhodes J, Yang H, Moussu S, Boutrot F, Santiago J, Zipfel C. 2021. Perception of a divergent

365 family of phytocytokines by the Arabidopsis receptor kinase MIK2. *Nat Commun*.

366 doi:10.1038/s41467-021-20932-y

367 Rojo E, Sharma VK, Kovaleva V, Raikhel N V., Fletcher JC. 2002. CLV3 is localized to the

368 extracellular space, where it activates the Arabidopsis CLAVATA stem cell signaling

369 pathway. *Plant Cell* **14**:969–977. doi:10.1105/tpc.002196

370 Roux M, Schwessinger B, Albrecht C, Chinchilla D, Jones A, Holton N, Malinovsky FG, Tör M,

371 de Vries S, Zipfel C. 2011. The Arabidopsis leucine-rich repeat receptor-like kinases

372 BAK1/SERK3 and BKK1/SERK4 are required for innate immunity to hemibiotrophic and

373 biotrophic pathogens. *Plant Cell* **23**:2440–2455. doi:10.1105/tpc.111.084301

374 Santiago J, Brandt B, Wildhagen M, Hohmann U, Hothorn LA, Butenko MA, Hothorn M.

375 2016. Mechanistic insight into a peptide hormone signaling complex mediating floral

376 organ abscission. *Elife* **5**. doi:10.7554/eLife.15075

377 Saur IML, Kadota Y, Sklenar J, Holton NJ, Smakowska E, Belkhadir Y, Zipfel C, Rathjen JP.

378 2016. NbCSPR underlies age-dependent immune responses to bacterial cold shock

379 protein in Nicotiana benthamiana. *Proc Natl Acad Sci U S A*.

380 doi:10.1073/pnas.1511847113

381 Schwessinger B, Roux M, Kadota Y, Ntoukakis V, Sklenar J, Jones A, Zipfel C. 2011.

382 Phosphorylation-Dependent Differential Regulation of Plant Growth, Cell Death, and

383 Innate Immunity by the Regulatory Receptor-Like Kinase BAK1. *PLoS Genet* **7**:e1002046.

doi:10.1371/journal.pgen.1002046

Segonzac C, Feike D, Gimenez-Ibanez S, Hann DR, Zipfel C, Rathjen JP. 2011. Hierarchy and roles of pathogen-associated molecular pattern-induced responses in nicotiana benthamiana. *Plant Physiol* **156**:687–699. doi:10.1104/pp.110.171249

Shinohara H, Mori A, Yasue N, Sumida K, Matsubayashi Y. 2016. Identification of three LRR-RKs involved in perception of root meristem growth factor in Arabidopsis. *Proc Natl Acad Sci U S A* **113**:3897–3902. doi:10.1073/pnas.1522639113

Smakowska-Luzan E, Mott GA, Parys K, Stegmann M, Howton TC, Layeghifard M, Neuhold J, Lehner A, Kong J, Grünwald K, Weinberger N, Satbhai SB, Mayer D, Busch W, Madalinski M, Stolt-Bergner P, Provart NJ, Mukhtar MS, Zipfel C, Desveaux D, Guttman DS, Belkhadir Y. 2018. An extracellular network of Arabidopsis leucine-rich repeat receptor kinases. *Nature* **553**:342–346. doi:10.1038/nature25184

Song W, Liu L, Wang J, Wu Z, Zhang H, Tang J, Lin G, Wang Y, Wen X, Li W, Han Z, Guo H, Chai J. 2016. Signature motif-guided identification of receptors for peptide hormones essential for root meristem growth. *Cell Res* **26**:674–85. doi:10.1038/cr.2016.62

Stanke M, Diekhans M, Baertsch R, Haussler D. 2008. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* **24**:637–644. doi:10.1093/BIOINFORMATICS/BTN013

Stephens M. 2017. False discovery rates: A new deal. *Biostatistics* **18**:275–294. doi:10.1093/biostatistics/kxw041

Tabata R, Sumida K, Yoshii T, Ohyama K, Shinohara H, Matsubayashi Y. 2014. Perception of root-derived peptides by shoot LRR-RKs mediates systemic N-demand signaling. *Science (80-)* **346**:343–346. doi:10.1126/science.1257800

Takahashi F, Suzuki T, Osakabe Y, Betsuyaku S, Kondo Y, Dohmae N, Fukuda H, Yamaguchi-

Shinozaki K, Shinozaki K. 2018. A small peptide modulates stomatal control via abscisic acid in long-distance signalling. *Nature* **556**:1. doi:10.1038/s41586-018-0009-2

Van der Does D, Boutrot F, Engelsdorf T, Rhodes J, McKenna JF, Vernhettes S, Koevoets I, Tintor N, Veerabagu M, Miedes E, Segonzac C, Roux M, Breda AS, Hardtke CS, Molina A, Rep M, Testerink C, Mouille G, Höfte H, Hamann T, Zipfel C. 2017. The Arabidopsis leucine-rich repeat receptor kinase MIK2/LRR-KISS connects cell wall integrity sensing, root growth and response to abiotic and biotic stresses. *PLOS Genet* **13**:e1006832. doi:10.1371/journal.pgen.1006832

Vie AK, Najafi J, Liu B, Winge P, Butenko MA, Hornslien KS, Kumpf R, Aalen RB, Bones AM, Brembu T. 2015. The IDA/IDA-LIKE and PIP/PIP-LIKE gene families in Arabidopsis: phylogenetic relationship, expression patterns, and transcriptional effect of the PIPL3 peptide. *J Exp Bot* **66**:5351–65. doi:10.1093/jxb/erv285

Wagih O. 2017. Ggseqlogo: A versatile R package for drawing sequence logos. *Bioinformatics* **33**:3645–3647. doi:10.1093/bioinformatics/btx469

Wei T, Simko V. 2021. R package “corrplot”: Visualization of a Correlation Matrix.

Wheeler TJ, Eddy SR. 2013. nhmmer: DNA homology search with profile HMMs. *Bioinformatics* **29**:2487–2489. doi:10.1093/BIOINFORMATICS/BTT403

Yamaguchi Y, Huffaker A, Bryan AC, Tax FE, Ryan CA. 2010. PEPR2 Is a Second Receptor for the Pep1 and Pep2 Peptides and Contributes to Defense Responses in *Arabidopsis*. *Plant Cell* **22**:508–522. doi:10.1105/tpc.109.068874

Yamaguchi Y, Pearce G, Ryan CA. 2006. The cell surface leucine-rich repeat receptor for AtPep1, an endogenous peptide elicitor in Arabidopsis, is functional in transgenic tobacco cells. *Proc Natl Acad Sci U S A* **103**:10104–9. doi:10.1073/pnas.0603729103

Zhang H, Lin X, Han Z, Qu L-J, Chai J. 2016. Crystal structure of PXY-TDIF complex reveals a

conserved recognition mechanism among CLE peptide-receptor pairs. *Cell Res* **26**:543–555. doi:10.1038/cr.2016.45

Figure and Table Legends

Figure 1. CTNIPs are a novel family of plant signalling peptide

(**a**) Heat map showing log_2(FC) expression levels of CTNIP1-4 in response to a range of elicitors (data obtained from Bjornson *et al.* (2021)).

(**b**) Sequence probability logo from *Arabidopsis* CTNIP1-4 generated using WebLogo3. Signal peptide (as predicted by SignalP5.0) and CTNIP motif are indicated, and conserved cysteine residues are highlighted in yellow. Amino acids are coloured based on their biochemical properties: red = acidic; blue= basic; black = hydrophobic and green = polar.

(**c-d**) cytoplasmic calcium influx measured in *p35S::AEQUORIN* seedlings after treatment with 1 µM CTNIP relative to pre-treated levels (n = 8 seedlings). (c) Points represent mean; error bars represent S.E.M. (d) represents mean relative Ca^{2+} influx between 5 and 15 min. A line represents mean; error bars represent S.D. *P*-values indicate significance relative to the WT control in a Dunnett's multiple comparison test following one-way ANOVA.

(**e**) Western blot using α-p42/p44-ERK recognizing phosphorylated MAP kinases in seedlings treated with 100 nM CTNIPs or mock for 15 min. The membrane was stained with CBB, as a loading control.

(**f-g**) ROS production in leaf disks collected from 4-week-old Arabidopsis plants induced by 1 µM CTNIP4 application ($n \geq 8$). (f) Points represent mean; error bars represent S.E.M. (g) Integrated ROS production over 40 min. Line represents mean; error bars

represent S.D. *P*-values indicate significance relative to the WT control in a two-tailed T-test.

All experiments were repeated and analysed three times with similar results. ROS, reactive oxygen species; CBB, Coomassie brilliant blue.

Figure 2. HLS3 forms a CTNIP-induced receptor complex with BAK1

(**a**) Schematic representation of BAK1-GFP immunoprecipitation in the (1) absence or (2) presence of CTNIP4 treatment to identify protein associations induced by CTNIP. Figure generated using Biorender.

(**b**) HSL3-specific spectral counts identified in four independent biological replicates where BAK1-GFP was pulled-down in the presence or absence of 1 µM CTNIP4 treatment. Circle diameter is proportional to the number of replicates. Red lines indicate the mean spectral counts for each treatment. *P*-values indicate significance relative to the untreated control in a two-tailed T-test.

(**c**) Co-immunoprecipitation of BAK1 with HSL3–GFP from HSL3-GFP seedlings treated with 1 µM CTNIP4$^{48-70}$ or water for 10 min. Western blots were probed with antibodies α-GFP and α-BAK1. This experiment was repeated 3 three times with similar results.

(**d**) ITC summary table of HSL3 vs CTNP4$^{48-70}$, CTNP4$^{C58S/C68S}$ and IDA peptides, and contribution of the BAK1 co-receptor to the ternary complex formation. K_d, (dissociation constant) indicates the binding affinity between the two molecules considered (nM). The N indicates the reaction stoichiometry (N=1 for a 1:1 interaction). The values indicated in the table are the mean ± S.E.M. of two independent experiments.

(**e**) Isothermal titration calorimetry (ITC) experiments of HSL3 vs CTNIP4 and CTNIP4$^{C58S/C68S}$, in the absence and presence of the co-receptor BAK1.

Figure 3. HSL3 is strictly required for CTNIP perception and growth regulation

(**a**) Western blot using α-p42/p44-ERK recognizing phosphorylated MAP kinases in seedlings treated with 100 nM CTNIPs or mock for 15 min. The membrane was stained with CBB, as a loading control.

(**b**) Heat map showing all significantly differentially expressed genes ($p<0.05$, $|Log_2(FC)|>1$) in *Arabidopsis* WT or *hsl3-1* seedlings treated with or without 100 nM CTNIP4$^{48-70}$ for 30 min relative to a mock control.

(**c**) Mean relative cytoplasmic Ca^{2+} influx in leaf disks of *N. benthamiana* transiently expressing the defined constructs induced by 1 µM CTNIP or mock application ($n = 8$ leaf disks). A line represents mean; error bars represent S.D. *P*-values indicate significance relative to the GUS-transformed control in a Dunnett's multiple comparison test following one-way ANOVA.

(**d**) Fresh weight of 14-day-old seedlings grown in the presence of 500 nM CTNIP4 for 10 days relative to mock ($n = 8$ seedlings). A line represents mean; error bars represent S.D. *P*-values indicate significance relative to the WT control in a two-tailed T-test.

(**e**) Representative images of (d).

(**f-g**) Nine-day-old vertically grown *Arabidopsis* seedlings on 1/2 MS agar medium with 1 % sucrose. Pictures were taken from the front of the plate.

(**h-k**) Root parameters were quantified from the base of the hypocotyl to the root tip using ImageJ (h) Root angle is shown relative to mock. Negative values indicate leftward root skewing. (j) Absolute root angle with 90° representing the gravity vector. Angles <90°

represent skewing to the left. A line represents mean; error bars represent S.D. *P*-values indicate significance relative to the WT control in a Dunnett's multiple comparison test following one-way ANOVA. All experiments were repeated and analysed three times with similar results.

Figure 4. The HSL3-CTNIP signalling module is ancient and conserved

(**a**) Phylogeny of the full-length amino acid sequences of HAE/HSL/CEPR/RLK7/IKU2 clade of receptor kinases. Eudicot sequences are indicated in blue, monocot sequences in green and *Amborella* sequences in red. Clades are named based upon the *Arabidopsis* genes. Alignment shown in Supplementary file 11. Further details of species, sequence identification, alignment and phylogeny generation are described in the material and methods.

(**b**) Species tree with number of CTNIP and HSL3 orthologs identified. Annotated CTNIPs are shown in grey whilst unannotated CTNIPs are shown in black. Sequences are shown in Supplementary file 9 and Supplementary file 11.

(**c**) Phylogeny of the full-length amino acid sequences of CTNIPs. Eudicot sequences are indicated in blue, monocot sequences in green and *Amborella* sequences in red. Sequences shown in Supplementary file 9. Further details of species, sequence identification, alignment and phylogeny generation are described in the material and methods.

(**d**) Sequence Logo generated from CTNIP alignment from (c) using the R-package ggseqlogo. Amino acids are coloured based on their biochemical properties: red = acidic; blue= basic; black = hydrophobic; purple = neutral and green = polar.

(**e**) Cytoplasmic calcium influx measured after treatment with 1 µM CTNIP in *p35S::AEQUORIN N. benthamiana* leaf disks transiently expressing the defined construct, relative to pre-treatment (n = 8 leaf disks). Points represent mean; error bars represent S.E.M. Experiments were repeated and analysed three times with similar results.

Figure 1 – figure supplement 1. Alignment and phylogeny of *Arabidopsis* CTNIPs

(**a**) Phylogeny of *Arabidopsis* CTNIPs. Full length protein sequences were aligned using MUSCLE and a phylogeny was inferred using the Maximum-likelihood method and JTT matrix-based model conducted in MEGAX. 1000 bootstraps were performed and values shown in blue. Branch lengths are shown in black.

(**b**) Alignment used to generate (a). CTNIP motif is highlighted in red.

Figure 1 -figure supplement 2. Characterisation of CTNIP fragments

(**a**) Alignment of CTNIP4 fragments used in this manuscript.

(**b**) Western blot using α-p42/p44-ERK recognizing phosphorylated MAP kinases in seedlings treated with 100 nM CTNIP4 fragments or mock for 15 min. The membrane was stained with CBB, as a loading control.

(**c**) Mean relative Ca^{2+} influx induced by 1 µM CTNIP in *p35S::AEQUORIN* seedlings between 5 and 15 min, relative to pre-treatment (n = 8 seedlings). A line represents mean; error bars represent S.D. *P*-values indicate significance relative to the WT control in a Dunnett's multiple comparison test following one-way ANOVA.

Figure 2 – figure supplement 1. *Arabidopsis* LRR-RK subfamily XI

Phylogeny of full-length protein sequences of the *Arabidopsis* LRR-RK subfamily XI. Sequences were aligned using MUSCLE and a phylogeny was inferred using the Maximum-likelihood method and JTT matrix-based model conducted in MEGAX. 1000 bootstraps were performed and are indicated based on the size of the blue circles. Expression of these receptors in response to 1 µM flg22 treatment was extracted from Bjornson *et al*. (2021) and is represented in a heat map. Known ligands for LRR-RK subfamily XI are highlighted to the right (Butenko et al., 2003; Cho et al., 2008; Crook et al., 2020; Doblas et al., 2017; Doll et al., 2020; Hou et al., 2014; Krol et al., 2010; Morita et al., 2016; Mou et al., 2017; Nakayama et al., 2017; Ogawa et al., 2008; Okuda et al., 2020; Ou et al., 2016; Qian et al., 2018; Rojo et al., 2002; Santiago et al., 2016; Shinohara et al., 2016; Song et al., 2016; Tabata et al., 2014; Yamaguchi et al., 2010, 2006; Zhang et al., 2016).

Figure 2 – figure supplement 2. ITC independent experiments and purification of HSL3 and BAK1 used in the binding experiments.

(**a**) Analytical size-exclusion chromatography (SEC) of the ectodomains of HSL3 and BAK1. An SDS PAGE of the two proteins in shown alongside.

(**b**) ITC raw thermograms of experiments shown in the ITC table summary in Figure 2d.

Figure 2 – figure supplement 3. Structural comparison of the binding pockets between the receptors HAESA and HSL3

(**a**) The hydroxyproline pocket required for anchoring the IDA peptide to the HAESA receptor is missing in HSL3. Close view of the binding pocket of the structural superimposition of the HAESA-IDA complex (PDB:5IXQ) and a homology model of HSL3 (AlphaFold: https://alphafold.ebi.ac.uk/). The HAESA receptor is depicted in surface

574 representation in teal blue, IDA in yellow sticks and HSL3 in magenta cartoon. In HSL3, the
575 hydroxyproline pocket is replaced by the bulky residue Phe286, colliding with the
576 potential anchoring of the IDA peptide to the receptor.
577 (**b**) The conserved RxR motif necessary for the coordination of the COO^- group the last Asn
578 in IDA is not present in the HSL3 receptor. Zoom in of the C-terminal region of the peptide
579 binding surface of HAESA (teal blue) (left panel) and HSL3 (magenta) (right panel). In
580 HAESA the motif RxR closes the binding pocket allowing for the coordination of the C-
581 terminal of IDA. In HSL3 this structural motif is substituted by the residues Thr406 and
582 Gln408, leaving the binding surface open to potentially accommodate a longer peptide
583 ligand. Figures were done using the PyMOL Molecular Graphics System, Version 2.0
584 Schrödinger, LLC.
585
586 **Figure 3 – figure supplement 1. Genetic characterization of *hsl3* mutants**
587 (**a**) Gene model showing the location of T-DNA inserts.
588 (**b**) PCR confirming T-DNA insertion and mutant homozygosity.
589
590 **Figure 3 – figure supplement 2. CTNIP-induced seedling growth inhibition**
591 (**a-b**) Fresh weight of 14-day-old seedlings grown in the presence of 500 nM CTNIPs for 10
592 days relative to mock (*n* = 8 seedlings). A line represents mean; error bars represent S.D.;
593 *P*-values indicate significance relative to the WT control in a Dunnett's multiple
594 comparison test following one-way ANOVA.
595
596 **Figure 3 – figure supplement 3. Correlation of CTNIP4-induced transcriptomic response with**
597 **that of elicitors at 30 min**

598 CTINP4-induced gene expression is well correlated with elicitor-induced gene expression
599 from Bjornson *et al.* (2021). Circle colour and size are proportional to the Spearman
600 correlation coefficient (R-squared value) of each pairwise comparison of log_2(fold
601 changes).
602
603 **Figure 3 – figure 4. Characterisation of CTNIP and chimeric receptor lines**
604 (**a**) Western blot using α-FLAG recognizing $BIR3_{ecto}$-$HSL3_{cyto}$-FLAG in seedlings to confirm
605 expression. The membrane was stained with CBB, as a loading control.
606 (**b**) Documentation of the Cas9-induced mutations observed within the *ctnip1-4*
607 polymutant and their predicted effects on protein products.
608 (**c**) qRT-PCR documenting the overexpression of CTNIP lines. Expression of *CTNIP4* is
609 shown relative to *U-Box* (*At5g15400*). Points represent independent biological replicates
610 each comprising two technical replicates. Lines represent the mean of biological
611 replicates.
612
613 **Supplementary file 1. Spectral counts of peptides identified through affinity-purification of**
614 **the BAK1 complex**
615
616 **Supplementary file 2. Differential gene expression induced by 30 min $CTNIP4^{48-70}$ treatment**
617
618 **Supplementary file 3. Gene ontology enrichment following 30 min $CTNIP4^{48-70}$ treatment**
619 **Supplementary file 4. Primers**
620
621 **Supplementary file 5. Synthetic peptides**

Supplementary file 6. Species induced in CTNIP and RK search

Supplementary file 7. Initial CTNIP candidates used to search

Supplementary file 8. Identified CTNIPs relaxed

Supplementary file 9. Identified CTNIPs confident

Supplementary file 10. Initial RK candidates used to search

Supplementary file 11. Alignment of RKs identified

Material and Methods

Plant material and growth conditions

Arabidopsis plants for ROS burst assays were grown in individual pots at 21 °C with a 10-h photoperiod. Seeds grown on plates were surface sterilized using chlorine gas for 5–6 h and sown on 1/2 Murashige and Skoog (MS) media, 1 % sucrose, and 0.8 % agar and stratified at 4 °C for 2–3 days. Plates were then transferred to 22 °C under a 16-h photoperiod. For root growth assays plates were placed in an upright position under a 10° angle relative to the direction of gravity and images were taken 9 days later (Van der Does et al., 2017). *Nicotiana benthamiana* plants were grown on peat-based media at 24 °C, with 16-h photoperiod.

Aequorin lines of *Arabidopsis* and *N. benthamiana* were described previously (Knight et al., 1991; Segonzac et al., 2011). *hsl3* mutants have been previously described and were obtained from the Eurasian Arabidopsis Stock Centre (uNASC) (Hou et al., 2014; Lee et al., 2020). *bak1-4/*pBAK1::BAK1-GFP, *bak1-5* and p35S::GFP-Lti6B lines have also been described previously (Cutler et al., 2000; Ntoukakis et al., 2011; Schwessinger et al., 2011).

Synthetic peptides

All synthetic peptides were ordered at >80 % purity from either EZbiolabs or Genscript. Sequences of all peptides can be found in Supplementary file 5.

Alignment and phylogeny of Arabidopsis CTNIPs and LRR-RK subfamily XI

Full length protein sequences were aligned using MUSCLE (Edgar, 2004) and a phylogeny was inferred using the Maximum-likelihood method and JTT matrix-based model conducted in MEGAX (Kumar et al., 2018). 1000 bootstraps were performed. Trees were visualised in iTOL (Letunic and Bork, 2019). The sequence logo was generated using WebLogo3 (Crooks et al., 2004).

Molecular cloning

In-planta expression

For overexpression of *At*HSL3-GFP and *MtHSL3* in *N. benthamiana* and Arabidopsis the genomic DNA sequence was amplified from Arabidopsis ecotype Columbia and *M. truncatula* ecotype A11, domesticated and directly ligated into pICSL86977 downstream of a 35S promoter and with an in-frame C-terminal GFP tag.

669 Fragments for the pHSL3::BIR3ecto-HSL3cyto-FLAG construct were amplified from genomic
670 DNA using the indicated primers and ligated into pICSL86955 (Supplementary file 4).
671 Fragments were designed according to Hohmann *et al.* (2020). Clones were verified by Sanger
672 sequencing.
673
674 ***CRISPR-Cas9 mutagenesis***
675 CRISPR-Cas9 induced mutagenesis was performed as described by Castel *et al.* (2019). The
676 *RPS5a* promoter drove Cas9 expression and FASTred selection was used for positive and
677 negative selection. Primers used to generate the vector can be found in Supplementary file 4.
678 Mutants were screened by Sanger sequencing.
679
680 ***ROS measurements***
681 Leaf disks were harvested from 4-week-old *Arabidopsis* plants into white 96-well-plates
682 (655075, Greiner Bio-One) containing 100 µL water using a 4-mm diameter biopsy punch
683 (Integra™ Miltex™). Leaf disks were rested overnight. Prior to ROS measurement, the water
684 was removed and replaced with ROS assay solution (100 µM Luminol (123072, Merck),
685 20 µg mL^{-1} horseradish peroxidase (P6782, Merck)) with or without elicitors. Immediately
686 after light emission was measured from the plate using a HIGH-RESOLUTION PHOTON
687 COUNTING SYSTEM (HRPCS218, Photek) equipped with a 20 mm F1.8 EX DG ASPHERICAL RF
688 WIDE LENS (Sigma Corp).
689
690 ***Cytoplasmic calcium measurements***
691 Seedlings were initially grown on 1/2 MS plates for 3 days before being transferred to 96-well
692 plates (655075, Greiner Bio-One) in 100 µL liquid MS for 5 days. The evening before calcium

693 measurements the liquid MS was replaced with 100 µL 20 µM coelenterazine (EC14031,
694 Carbosynth) and the seedlings incubated in the dark overnight. The following morning the
695 coelenterazine solution was replaced with 100 µL water and rested for a minimum of 30 min
696 in the dark. Readings were taken in a VARIOSKAN™ MUTIPLATE READER (ThermoFisher)
697 before and after adding 50 µL of 3× concentrated elicitor solution or mock. For each well
698 readings were normalised to the average RLU value before elicitor addition (L_0).
699
700 ***Seedling growth inhibition***
701 Four-day-old seedlings growing on 1/2 MS plates were transferred into individual wells of a
702 transparent 48-well tissue culture plate (Greiner Bio-One) containing 500 µL of liquid MS
703 media with/without elicitor addition. The plates were returned to the growth conditions for
704 an additional 10 days before seedlings were blot-dried and weighed.
705
706 ***Protein extraction and western blot***
707 Two-week-old seedlings grown in liquid MS media (MAPK phosphorylation) or leaf disks from
708 4-week-old plants were flash-frozen in liquid nitrogen. Frozen plant tissue was ground in a
709 Genogrinder® with 2mm glass beads (1500 strokes/min, 1.5 min) prior to boiling in 2×
710 Laemmli sample buffer (4 % SDS, 20 % glycerol, 10 % 2-mercaptoethanol, 0.004 %
711 bromophenol blue, and 0.125 M Tris-HCl; (10 µL.mg^{-1} tissue)) for 10 min at 95 °C. The samples
712 were then spun at 13,000 rcf for 5 min prior to loading and running on SDS-PAGE gels.
713 Proteins were transferred using semi-dry transfer onto PVDF membrane (ThermoFisher),
714 blocked in 5 % (w/v) Bovine serum albumin prior to incubation with appropriate antibodies
715 (α-pMAPK ((p44/42 MAPK (Erk1/2) antibody #9102; 1:4000); α-FLAG-HRP (A8592, Merck;
716 1:5,000) and α-rabbit-HRP (A-0545, Merck; 1:10000). Western blots were imaged with a LAS

4000 IMAGEQUANT SYSTEM (GE Healthcare) or on X-ray film before being developed. Staining of the blotted membrane with Coomassie brilliant blue was used to confirm loading.

Co-immunoprecipitation

All steps involving the protein extract and subsequent protein isolation were carried out on ice or at 4 °C and all buffers and tubes were pre-cooled.

Seeds were sown on 1/2 MS agar and stratified for 3 days as described above. When seedlings had germinated, they were transferred 6 seedlings per well into 6 well plates containing 5 mL of liquid MS media and grown for a further 12 days. Seedlings were then transferred into MS media either with or without CTNIP4 addition and vaccum infiltrated for 2 min and left in the solution for a further 10 min. Seedlings were rapidly dried and flash frozen in liquid nitrogen and ground. Proteins were extracted using by addition of ~2:1extraction buffer (50 mM Tris pH 7.5, 150 mM NaCl, 2.5 mM EDTA, 10 % Glycerol, 1 % IGEPAL, 5 mM DTT, 1 % plant protease inhibitor cocktail (P9599, Sigma)):ground tissue (v/v). Proteins were solubilised at 4 °C with gentle agitation for 30 min before filtering through miracloth. The filtrate was centrifuged at 30000 rcf for 30 min at 4 °C. Protein concentrations were normalised using Bradford assay. An input sample was taken. To each 15 ml of protein extract 40 µl of GFP-TRAP AGAROSE BEADS (50 % slurry, ChromoTek) washed in extraction buffer were added and incubated with gentle agitation for 4 h at 4 °C. Bead were harvested by centrifugation at 1500 x g for 2 min and washed 3 times in extraction buffer. Beads were then resuspended in fifty microliters of 1.5 x elution (NuPage) buffer and incubated at 80 °C for 8 min. Samples were subsequently used from MS analysis or Western blotting.

739 Western blotting was performed as described previously (α-BAK1 (Roux et al., 2011);
740 1:2000), α-GFP-HRP (sc-9996, Santa Cruz; 1:5000) and α-rabbit-HRP (A-0545, Merck;
741 1:10000).
742
743 ***Sample Preparation for Mass Spectrometry***
744 Co-immunoprecipitated protein samples were ran approximately 1 cm into an SDS-PAGE
745 gel. This portion of the gel was then excised, cut into smaller pieces and washed three times
746 with acetonitrile (LC-MS-Grade):ammonium bicarbonate (50 mM), pH 8.0 (1:1, v/v), 30 min
747 each, followed by dehydration in acetonitrile, 10 min. Gel pieces were then reduced with 10
748 mm DTT for 30 min at 45 °C followed by alkylation with 55 mm iodoacetamide for 20 min at
749 room temperature, and a further three washes with acetonitrile:ammonium, 30 min each.
750 Gel pieces were dehydrated again with acetonitrile before rehydration with 40 µL trypsin
751 (Pierce Trypsin Protease, MS-Grade, catalog no. 90058) working solution (100 ng trypsin in
752 50 mM ammonium bicarbonate, 5% (v/v) acetonitrile). Where required, gel pieces were
753 covered with 50 mM ammonium bicarbonate to a final volume before incubation at 37 °C
754 overnight. Tryptic peptides were extracted from the gel pieces three times in an equal
755 volume of 50% acetonitrile, 5% formic acid (Pierce LC-MS-Grade, catalog no. 85178), 30 min
756 each. Extracted peptides were dried in a speed-vac and resuspended in (v/v) 2%
757 acetonitrile/0.2% trifluoroacetic acid (Merck, catalog no. 302031). A total of four biological
758 replicates for each sample type was submitted.
759
760 ***LC-MS/MS Analysis***
761 Approximately 35% of each sample was analysed using an Orbitrap Fusion™ Tribrid™ Mass
762 Spectrometer (Thermo Fisher Scientific) coupled to a U3000 nano-UPLC (Thermo Fisher

763 Scientific). The dissolved peptides were injected onto a reverse phase trap column
764 NanoEase m/z Symmetry C18, beads diameter 5 µm, inner diameter 180 µm x 20 mm length
765 (Waters). The column was operated at the flowrate 20µl/min in 2% acetonitrile, 0.05% TFA,
766 after 2.5min the trap column was connected to the analytical column NanoEase m/z HSS
767 C18 T3 Column, beads diameter 1.8 µm, inner diameter 75 µm x 250 mm length (Waters).
768 The column was equilibrated with 3% B (B: 80% acetonitrile in 0.05% formic acid (FA), A:
769 0.1% FA) before subsequent elution with the following steps of a linear gradient: 2.5min 3%
770 B, 5min 6.3% B, 13min 12.5% B, 50min 42.5% B, 58min 50% B, 61min 65% B, 63min 99% B,
771 66min 99% B, 67min 3% B, 90min 3% B. The flow rate was set to 200 nL/min. The mass
772 spectrometer was operated in positive ion mode with nano-electrospray ion source.
773 Molecular ions were generated by applying voltage +2.2kV to a conductive union coupling
774 the column outlet with fused silica PicoTip emitter, ID 10 µm (New Objective, Inc.) and the
775 ion transfer capillary temperature was set to 275°C. The mass spectrometer was operated in
776 data-dependent mode using a full scan, m/z range 300–1,800, nominal resolution of
777 120,000, target value 1×10^6, followed by MS/MS scans of the 40 most abundant ions.
778 MS/MS spectra were acquired using normalized collision energy of 30%, isolation width of
779 1.6 m/z, resolution of 120,000, and a target value set to 1×10^5. Precursor ions with charge
780 states 2-7 were selected for fragmentation and put on a dynamic exclusion list for 30
781 seconds. The minimum automatic gain control target was set to 5×10^3 and intensity
782 threshold was calculated to be 4.8×10^4. The peptide match feature was set to the
783 preferred mode and the feature to exclude isotopes was enabled.
784
785 ***Data Processing and Peptide Identification***

Peak lists in the form of Mascot generic files were prepared from raw data files using MS Convert (Proteowizard) and sent to a peptide search on Mascot server v2.7 using Mascot Daemon (Matrix Science, Ltd.) against an in-house contaminants database and the Araport 11 protein database. Tryptic peptides with up to 1 possible mis-cleavage and charge states +2, +3 were allowed in the search. The following peptide modifications were included in the search: carbamidomethylated Cysteine (fixed) and oxidized Methionine (variable). Data were searched with a monoisotopic precursor and fragment ion mass tolerance 10 ppm and 0.8 Da respectively. Decoy database was used to validate peptide sequence matches. Mascot results were combined in Scaffold v4.4.0 (Proteome Software Inc.) and peptide and protein identifications accepted if peptide probability and protein threshold was ≥ 80.0% and 99% respectively. Under these conditions the False Discovery Rate was 0.47%. Data was then exported to Excel (Microsoft) for further processing. Proteins were accepted if identified by at least 2 peptides and present in 2 or more biological replicates. Spectral counts from the 4 biological replicates were summed and used to derive a ratio of CTNIP4 treatment:mock treatment. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2016) partner repository with the dataset identifier PXD029264 and 10.6019/PXD029264

Transient expression in *Nicotiana benthamiana*

Agrobacterium tumefaciens strain GV3101 transformed with the appropriate construct were grown overnight in L-media and spun-down. The bacteria were resuspended in 10 mM $MgCl_2$ and adjusted to $O.D._{600} = 0.2$ prior to infiltration into the youngest fully expanded leaves of 3-week-old plants. Leaf disks were collected 24 h later, and calcium assays were

809 performed as described for seedlings with leaf disks being floated overnight in the dark in
810 20 µM coelenterazine (EC14031, Carbosynth).
811
812 ***Protein expression and purification***
813 The ectodomains expressed and purified were coded from Arabidopsis genes *HSL3* (22-627,
814 AT5G25930) and *BAK1* (residues 20 – 637, AT4G33430). Codon-optimised synthetic genes
815 were cloned into a modified pFastBac vector (Geneva Biotech) vector, providing a TEV
816 (tobacco etch virus protease) cleavable C-terminal StrepII-9xHis tag. Expression of HSL3 and
817 BAK1 was driven by the signal peptides 30K (Futatsumori-Sugai and Tsumoto, 2010) or
818 *Drosophila BiP* (Smakowska-Luzan et al., 2018), respectively. The baculovirus were generated
819 in DH10 cells and *Spodoptera frugiperda* Sf9 cells were used for viral amplification. For protein
820 expression *Trichoplusia ni* Tnao38 cells (Hashimoto et al., 2012), were infected with HSL3 and
821 BAK1 viruses with a multiplicity of infection (MOI) of 3. The cells were grown 1 day at 28 °C
822 and two days at 22 °C at 110 rpm The secreted proteins were purified separately by sequential
823 Ni^{2+} (HisTrap excel, GE Healthcare, equilibrated in 25 mM KP_i pH 7.8 and 500 mM NaCl) and
824 StrepII (Strep-Tactin Superflow high-capacity, (IBA, Germany) equilibrated in 25 mM Tris pH
825 8.0, 250 mM NaCl, 1 mM EDTA) affinity chromatography. Recombinant Strep-tagged TEV
826 protease was used in 1:50 ratio to remove the affinity tags. The cleaved tag and the protease
827 were separated from the protein ectodomains by Ni^{2+} affinity chromatography. Proteins were
828 further purified by size exclusion chromatography on a Superdex 200 Increase 10/300 GL
829 column (GE Healthcare) equilibrated in 20 mM citric acid pH 5.0, 150 mM NaCl. Peak fractions
830 containing the complex were concentrated using Amicon Ultra concentrators (Millipore,
831 MWCO 10,000 for BAK1 and 30,000 for HSL3). Proteins were analysed for purity and structural
832 integrity by SDS-PAGE.

833

834 ***Isothermal titration calorimetry (ITC)***

835 A MicroCal PEAQ-ITC (Malvern Instruments) was used to performing the ITC binding assays.
836 Experiments were performed at 25 °C with a 200 µL standard cell and a 40 µL titration syringe.
837 HSL3 and BAK1 proteins were gel-filtrated into pH 5 ITC buffer (20 mM sodium citrate pH 5.0,
838 150 mM NaCl). Protein concentrations for HSL3 and BAK1 were calculated using their molar
839 extinction coefficient and a calculated molecular weight of ~75,000 for HSL3 and ~ 25,000 Da
840 for BAK1. Experiments were performed with 20µM of HSL3 protein in the cell and between
841 200-450 µM of indicated peptide ligand in the syringe, following an injection pattern of 2 µL
842 at 150 s intervals and 500 r.p.m. stirring speed. The BAK1 vs HSL3- peptide experiments were
843 performed by titrating 100 µM of BAK1 in the cell, using the same injection pattern. ITC data
844 were corrected for the heat of dilution by subtracting the mixing enthalpies for titrant solution
845 injections into protein free ITC buffer. Experiments were done in replicates and data were
846 analyzed using the MicroCal PEAQ-ITC Analysis Software provided by the manufacturer. All
847 ITC runs used for data analysis had an N ranging from 0.7 to 1.3. The N values were fitted to
848 1 in the analysis.

849

850 ***RNA sequencing and qRT-PCR***

851 Two 3-day-old seedlings per well were transferred into transparent 24-well plates (Grenier
852 Bio-One) containing 1 mL liquid MS media, sealed with porous tape and grown for a further 9
853 days. For qRT-PCR seedlings were harvested at this point. For RNA-seq experiments media
854 was then exchanged for 500 µL fresh MS media and left overnight. In the morning a further
855 480 µL of fresh media was added. 9.5 h later 20 µL treatment/mock was added and seedlings
856 were harvested after 30 min. All seedlings were ground in liquid nitrogen.

Total RNA was extracted using Trizol reagent (Merck) according to the manufacturer's instructions and DNAase/RNA cleanup treatment was performed using the RNeasy kit (Qiagen). RNA sequencing was performed by Novogene. The RNA-seq datasets generated and analysed in the current study have been deposited in the ArrayExpress database at EMBL-EBI (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-11093. qRT-PCR was performed on cDNA synthesised using The RevertAid first strand cDNA synthesis kit (Thermofisher) according to the manufacturer's instructions. cDNA was amplified by quantitative PCR using SYBR Green JumpStart Taq ReadyMix (Roche) and the CFX96 Real-Time PCR Detection System (Bio-Rad Laboratories, Hercules, CA, USA).

The read data were analysed using FastQC, trimmed using trimmomatic (Bolger et al., 2014) and mapped to the *Arabidopsis* TAIR10 genome via TopHat2 (Andrews et al., 2015; Kim et al., 2013). The mapped reads were assigned to genes by featureCounts from package Rsubread in R (Liao et al., 2019), and differential expression analysis was performed using DESeq2 with ashr L2FC shrinkage (Love et al., 2014; Stephens, 2017). Changes in gene expression were visualised using the R package ComplexHeatmap (Gu et al., 2016).

GO enrichment

GO term enrichment was calculated using the R package topGO (Alexa and Rahnenfuhrer, 2021), with arguments method= weight.01 and statistic=Fisher.

Correlation of expression

Pairwise comparisons of gene expression differences (log_2(FC)) was performed in R using the rcorr function from package Hmisc (Harrell Jr, 2021), type=Spearman, and correlations were plotted using corrplot (Wei and Simko, 2021).

881

882

883 ***Genome data retrieval***

884 Whole genome sequences and protein sequences were retrieved from Ensembl (release 50),
885 Phytozome (version 13), NCBI, and marchantia.info. Species and individual assembly versions
886 are listed in Supplementary file 6 (SI_table_species_data.csv).

887

888 ***CTNIP identification***

889 *Peptide search*

890 Protein sequences from all species were first filtered for a maximum length of 300 amino
891 acids and merged into a single file. The initial set of CTNIP peptide sequences is given in
892 Supplementary file 7 (SI_data_initial_CTNIP_candidates.fasta). Additional candidates were
893 searched with 1) jackhmmer (version 3.1b2, (Eddy, 2011)), 2) diamond (version 0.9.26,
894 options -e 1e-8 -k 100, (Buchfink et al., 2014)), and 3) hmm profile search (3.1b2, (Wheeler
895 and Eddy, 2013)). For the hmm profile search, the initial set of candidates and the candidates
896 from the diamond search were aligned with muscle (v3.8.31, (Edgar, 2004)) to generate an
897 hmm profile (hmmbuild) that was then used to search more candidates (hmmsearch).
898 Candidates from all approaches were merged and grouped with a sequence similarity
899 network. For this, sequences were matched to each other with diamond (options -e 0.01 -k
900 100). The pairwise percent similarity scores above 20 % were used to construct a network.
901 The community structure of the network was resolved with a modularity optimization
902 algorithm (Blondel et al., 2008) implemented by the function cluster_louvain in the R package
903 igraph (version 1.0.1, (Csardi and Nepusz, 2006)). Candidates within the same communities
904 as the original candidate sequences were used as protein candidates.

905 *DNA search*

906 To search novel peptides that were previously not annotated, we extracted all transcript
907 sequences of the protein candidates and aligned them with muscle to generate an HMM
908 profile (hmmbuild) that was used to search all genomes with nhmmer (3.1b2,(Wheeler and
909 Eddy, 2013)). Candidate regions were filtered for already annotated genes and used as input
910 to restrict *de novo* gene prediction with Augustus (version 3.3.3, (Stanke et al., 2008)). Finally,
911 candidates from both, protein and DNA search, were merged to generate the final set of
912 CTNIP candidates (Supplementary file 8 (CTNIP_relaxed.align)). This "relaxed" set of
913 candidates was further filtered for having two Cysteins with a 9-11 amino acid spacing. Few
914 candidates were also removed by a visual inspection of the alignment, resulting in the
915 "confident" CTNIP candidates (Supplementary file 9 (CTNIP_confident.align)). Phylogeny and
916 clade identification was done with the "relaxed" set of candidates using muscle and FastTree
917 (version 2.1.11 SSE3, option -lg, (Price et al., 2010)) using an age cutoff of 9. The resulting
918 phylogenetic tree was rooted using a similar sequence from *M. polymorpha* (chr5:16052258-
919 16053618) as outgroup with gotree (v0.4.2, (Lemoine and Gascuel, 2021)) and graphically
920 represented using FigTree (v1.4.3, http://tree.bio.ed.ac.uk/software/figtree). The sequence
921 logo was generated with the alignment of the "confident" CTNIP candidates and the R-
922 package ggseqlogo (v.1, (Wagih, 2017)). Amino acids with a low occurrence (i.e., seen in less
923 than 5 % of the peptides) were trimmed from the alignment to generate a gap-free logo.

924

925

926 ***RK identification***

927 Protein sequences from all species were first filtered for a minimum length of 500 amino acids
928 and merged into a single file. The initial set of RK sequences was taken from the alignment

929 provided by Furumizu et al. (2021), but with the outgroups removed (*P. margaritaceaum*, *S.*
930 *muscicola*, and *M. endlicherianum*). The sequences are given in Supplementary file 10
931 (SI_data_initial_RK_candidates.fasta). Sequences were aligned with muscle to build and
932 search an HMM profile (hmmsearch options -E 1e-10 --incE 1e-10). Candidates were matched
933 to each other with diamond (options -e 1e-11 --id 20 --query-cover 80). The pairwise percent
934 similarity scores above 50 % were used to construct a network and communities were defined
935 as described as above. Likewise, only candidates within the same communities as the original
936 candidates were kept. Candidates were further filtered for the presence of an LRR and a
937 kinase domain with hmmsearch (options -E 1e-5) and PFAMv33 (Supplementary file 11
938 (RECEPTOR.align)). Phylogeny and clade identification was done with muscle and FastTree
939 (option -lg) using an age cutoff of 5.5. The resulting phylogenetic tree was rooted using the
940 sequences from *P. margaritaceaum* as outgroup (Furumizu et al., 2021).
941
942 ***Phylogenetic tree of all species***
943 The species tree was calculated using OrthoFinder (v2.5.4, (Emms and Kelly, 2019)) with all
944 protein sequences of all plants species.

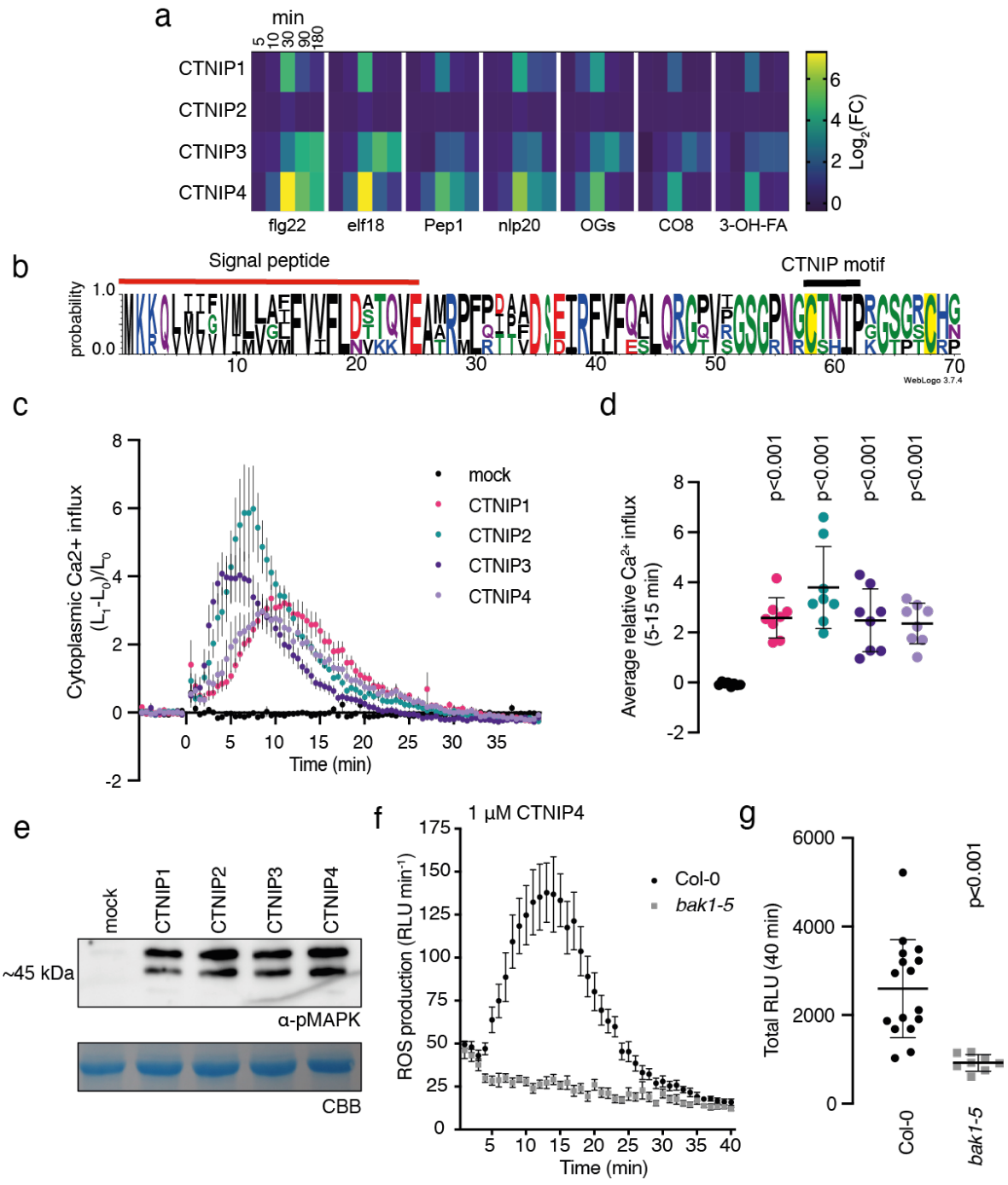


Figure 1. CTNIPs are a novel family of plant signalling peptide

(**a**) Heat map showing log_2(FC) expression levels of CTNIP1-4 in response to a range of elicitors (data obtained from Bjornson *et al.* (2021)).

(**b**) Sequence probability logo from *Arabidopsis* CTNIP1-4 generated using WebLogo3. Signal peptide (as predicted by SignalP5.0) and CTNIP motif are indicated, and conserved

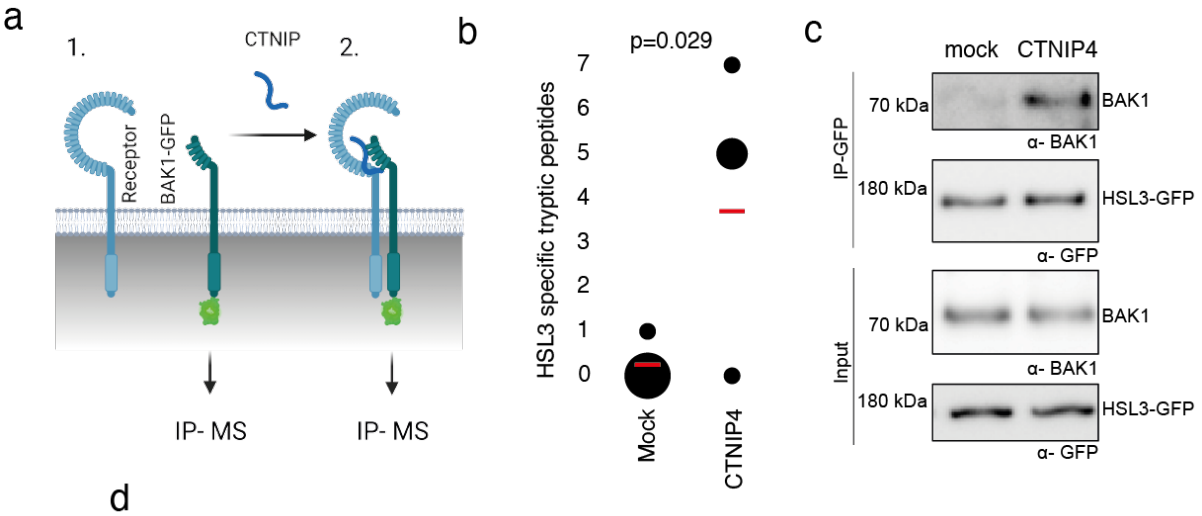
cysteine residues are highlighted in yellow. Amino acids are coloured based on their biochemical properties: red = acidic; blue= basic; black = hydrophobic and green = polar.

(**c-d**) cytoplasmic calcium influx measured in *p35S::AEQUORIN* seedlings after treatment with 1 µM CTNIP relative to pre-treated levels (n = 8 seedlings). (c) Points represent mean; error bars represent S.E.M. (d) represents mean relative Ca^{2+} influx between 5 and 15 min. A line represents mean; error bars represent S.D. *P*-values indicate significance relative to the WT control in a Dunnett's multiple comparison test following one-way ANOVA.

(**e**) Western blot using α-p42/p44-ERK recognizing phosphorylated MAP kinases in seedlings treated with 100 nM CTNIPs or mock for 15 min. The membrane was stained with CBB, as a loading control.

(**f-g**) ROS production in leaf disks collected from 4-week-old Arabidopsis plants induced by 1 µM CTNIP4 application ($n \geq 8$). (f) Points represent mean; error bars represent S.E.M. (g) Integrated ROS production over 40 min. Line represents mean; error bars represent S.D. *P*-values indicate significance relative to the WT control in a two-tailed T-test.

All experiments were repeated and analysed three times with similar results. ROS, reactive oxygen species; CBB, Coomassie brilliant blue.



d

Ligand (Syringe)	Protein (cell)	Kd(nM)	ΔH(kcal/mol)	N
CTNIP4$^{48-70}$	HSL3	4050 ± 1260	-0.89 ± 0.06	1
CTNIP4$^{C58S/C68S}$	HSL3	n.d.	n.d.	
IDA	HSL3	n.d.	n.d.	
BAK1	HSL3+CTNIP4$^{48-70}$	392 ± 7	-6.74 ± 0.7	1
BAK1	HSL3+CTNIP4$^{C58S/C68S}$	n.d.	n.d.	
BAK1	HSL3	n.d.	n.d.	

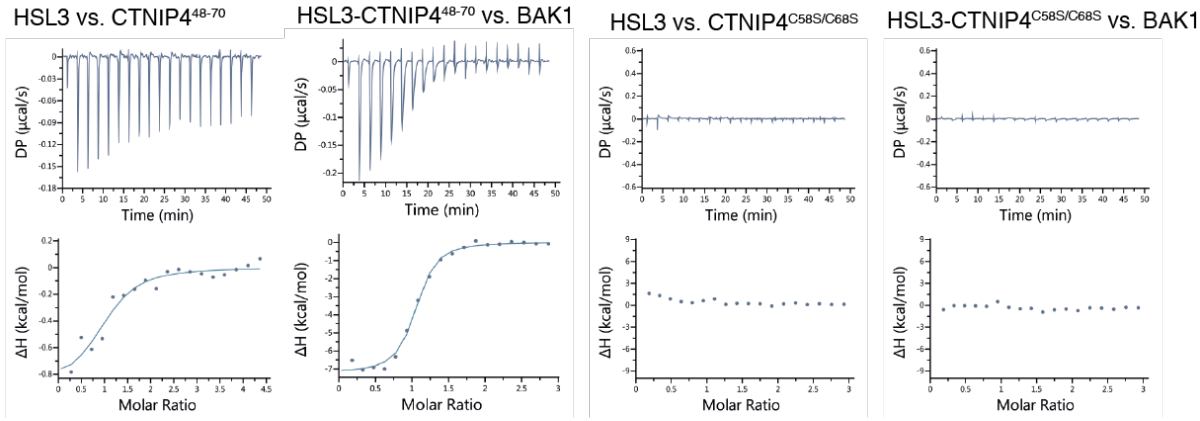


Figure 2. HLS3 forms a CTNIP-induced receptor complex with BAK1

(**a**) Schematic representation of BAK1-GFP immunoprecipitation in the (1) absence or (2) presence of CTNIP4 treatment to identify protein associations induced by CTNIP. Figure generated using Biorender.

(**b**) HSL3-specific spectral counts identified in four independent biological replicates where BAK1-GFP was pulled-down in the presence or absence of 1 µM CTNIP4 treatment. Circle diameter is proportional to the number of replicates. Red lines indicate the mean

spectral counts for each treatment. *P*-values indicate significance relative to the untreated control in a two-tailed T-test.

(**c**) Co-immunoprecipitation of BAK1 with HSL3–GFP from HSL3-GFP seedlings treated with 1 µM CTNIP4$^{48-70}$ or water for 10 min. Western blots were probed with antibodies α-GFP and α-BAK1. This experiment was repeated 3 three times with similar results.

(**d**) ITC summary table of HSL3 vs CTNP4$^{48-70}$, CTNP4$^{C58S/C68S}$ and IDA peptides, and contribution of the BAK1 co-receptor to the ternary complex formation. K_d, (dissociation constant) indicates the binding affinity between the two molecules considered (nM). The N indicates the reaction stoichiometry (N=1 for a 1:1 interaction). The values indicated in the table are the mean ± S.E.M. of two independent experiments.

(**e**) Isothermal titration calorimetry (ITC) experiments of HSL3 vs CTNIP4 and CTNIP4$^{C58S/C68S}$, in the absence and presence of the co-receptor BAK1.

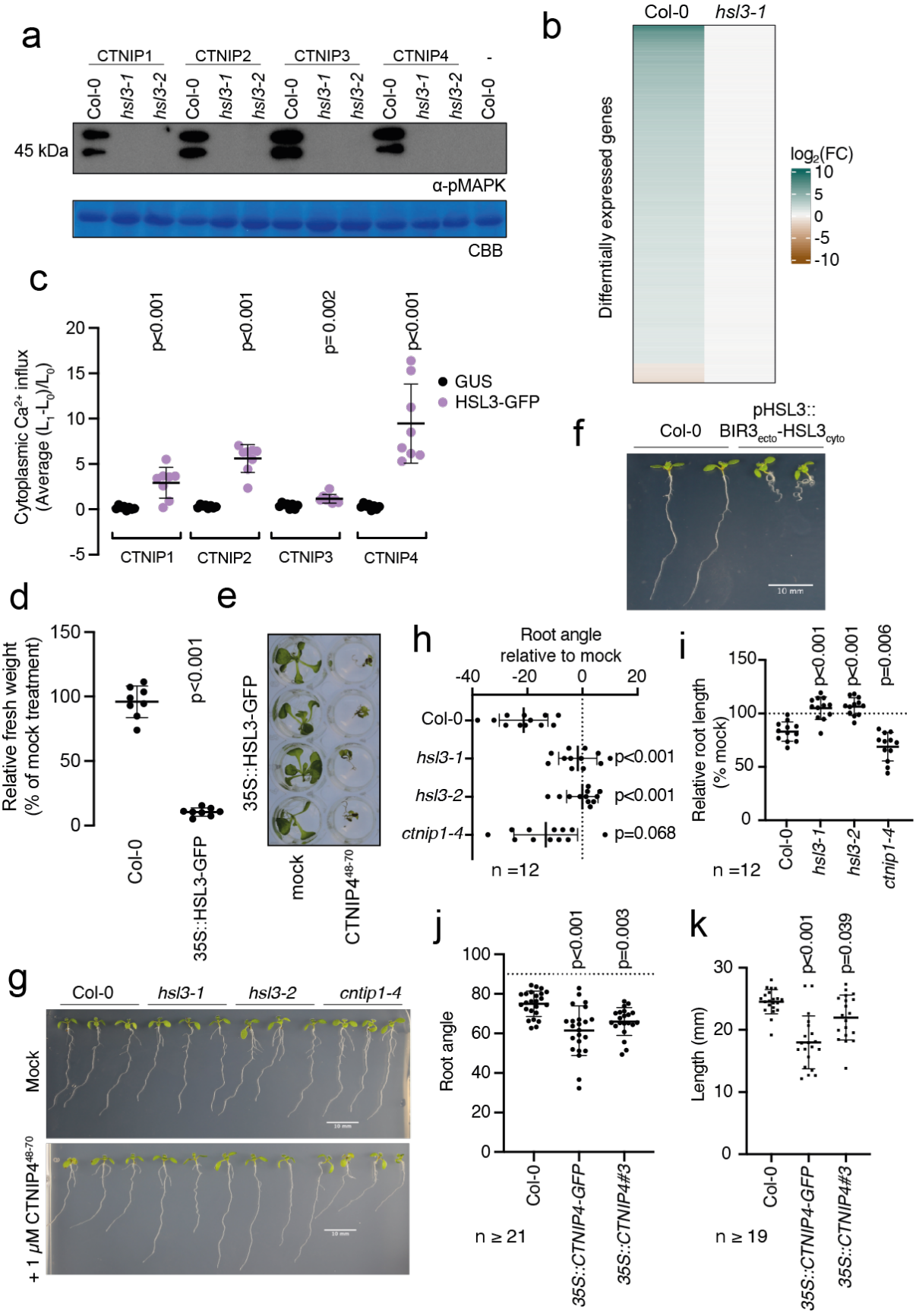


Figure 3. HSL3 is strictly required for CTNIP perception and growth regulation

(**a**) Western blot using α-p42/p44-ERK recognizing phosphorylated MAP kinases in seedlings treated with 100 nM CTNIPs or mock for 15 min. The membrane was stained with CBB, as a loading control.

(**b**) Heat map showing all significantly differentially expressed genes ($p<0.05$, $|Log_2(FC)|>1$) in *Arabidopsis* WT or *hsl3-1* seedlings treated with or without 100 nM $CTNIP4^{48-70}$ for 30 min relative to a mock control.

(**c**) Mean relative cytoplasmic Ca^{2+} influx in leaf disks of *N. benthamiana* transiently expressing the defined constructs induced by 1 µM CTNIP or mock application (*n* = 8 leaf disks). A line represents mean; error bars represent S.D. *P*-values indicate significance relative to the GUS-transformed control in a Dunnett's multiple comparison test following one-way ANOVA.

(**d**) Fresh weight of 14-day-old seedlings grown in the presence of 500 nM CTNIP4 for 10 days relative to mock (*n* = 8 seedlings). A line represents mean; error bars represent S.D. *P*-values indicate significance relative to the WT control in a two-tailed T-test.

(**e**) Representative images of (d).

(**f-g**) Nine-day-old vertically grown *Arabidopsis* seedlings on 1/2 MS agar medium with 1 % sucrose. Pictures were taken from the front of the plate.

(**h-k**) Root parameters were quantified from the base of the hypocotyl to the root tip using ImageJ (h) Root angle is shown relative to mock. Negative values indicate leftward root skewing. (j) Absolute root angle with 90° representing the gravity vector. Angles <90° represent skewing to the left. A line represents mean; error bars represent S.D. *P*-values indicate significance relative to the WT control in a Dunnett's multiple comparison test

following one-way ANOVA. All experiments were repeated and analysed three times with similar results.

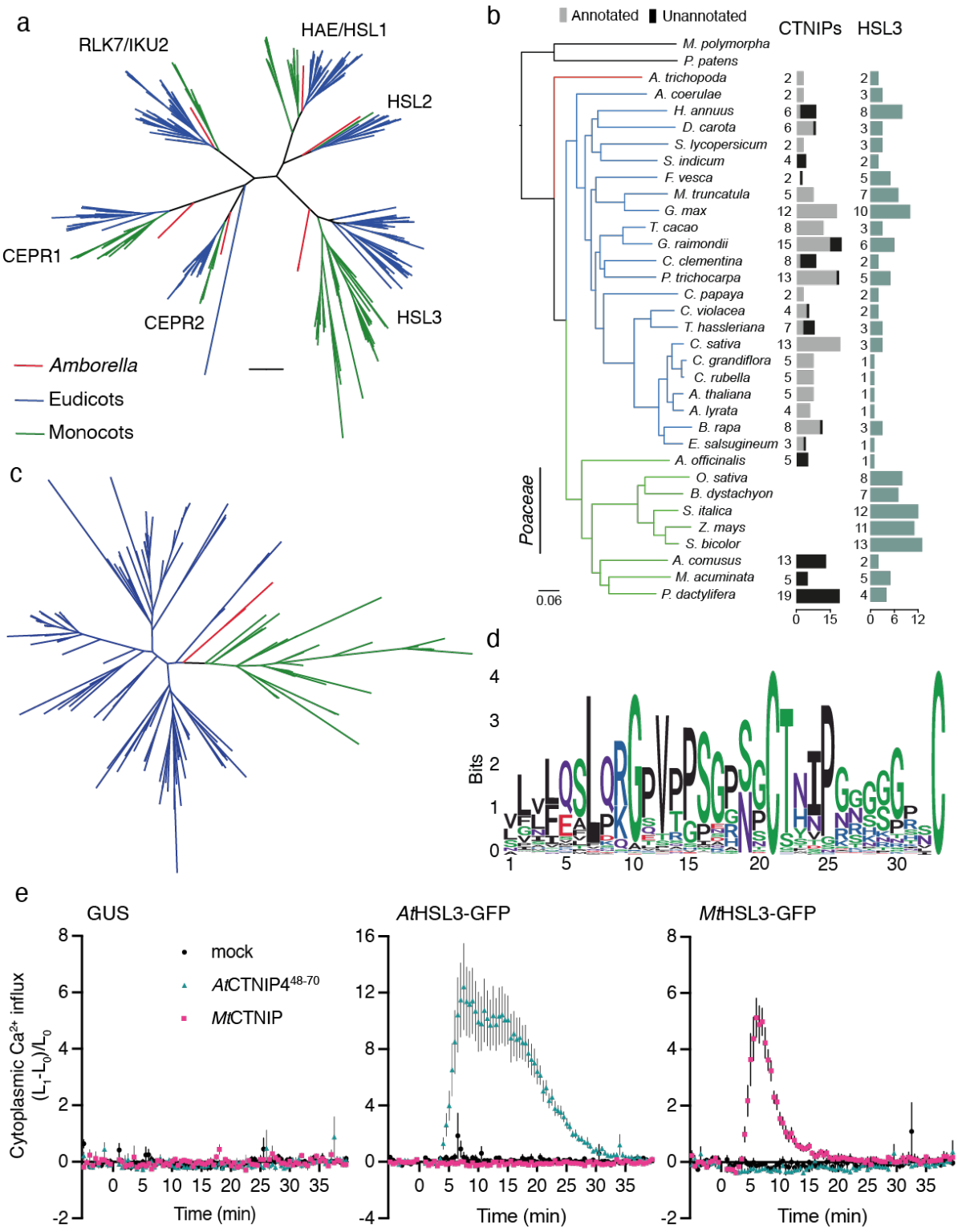


Figure 4. The HSL3-CTNIP signalling module is ancient and conserved

(**a**) Phylogeny of the full-length amino acid sequences of HAE/HSL/CEPR/RLK7/IKU2 clade of receptor kinases. Eudicot sequences are indicated in blue, monocot sequences in green and *Amborella* sequences in red. Clades are named based upon the *Arabidopsis* genes.

Alignment shown in Supplementary file 11. Further details of species, sequence identification, alignment and phylogeny generation are described in the material and methods.

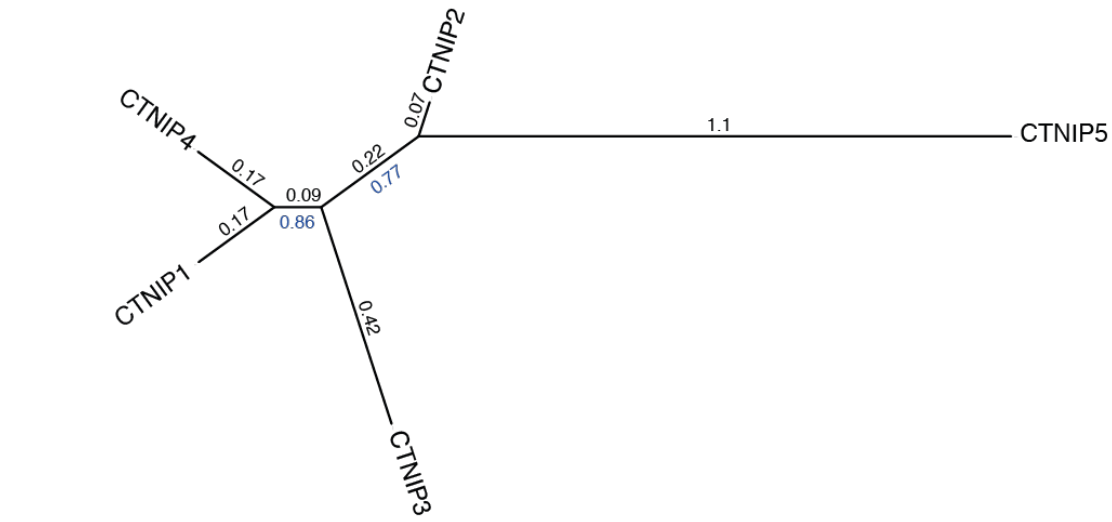
(**b**) Species tree with number of CTNIP and HSL3 orthologs identified. Annotated CTNIPs are shown in grey whilst unannotated CTNIPs are shown in black. Sequences are shown in Supplementary file 9 and Supplementary file 11.

(**c**) Phylogeny of the full-length amino acid sequences of CTNIPs. Eudicot sequences are indicated in blue, monocot sequences in green and *Amborella* sequences in red. Sequences shown in Supplementary file 9. Further details of species, sequence identification, alignment and phylogeny generation are described in the material and methods.

(**d**) Sequence Logo generated from CTNIP alignment from (c) using the R-package ggseqlogo. Amino acids are coloured based on their biochemical properties: red = acidic; blue= basic; black = hydrophobic; purple = neutral and green = polar.

(**e**) Cytoplasmic calcium influx measured after treatment with 1 µM CTNIP in *p35S::AEQUORIN N. benthamiana* leaf disks transiently expressing the defined construct, relative to pre-treatment ($n = 8$ leaf disks). Points represent mean; error bars represent S.E.M. Experiments were repeated and analysed three times with similar results.

a



b

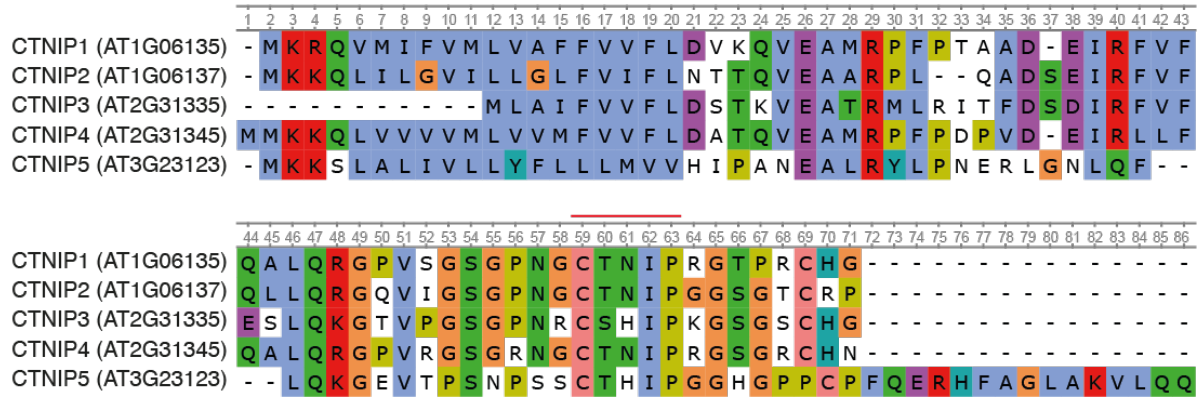


Figure 1 – figure supplement 1. Alignment and phylogeny of *Arabidopsis* CTNIPs

(**a**) Phylogeny of *Arabidopsis* CTNIPs. Full length protein sequences were aligned using MUSCLE and a phylogeny was inferred using the Maximum-likelihood method and JTT matrix-based model conducted in MEGAX. 1000 bootstraps were performed and values shown in blue. Branch lengths are shown in black.

(**b**) Alignment used to generate (a). CTNIP motif is highlighted in red.

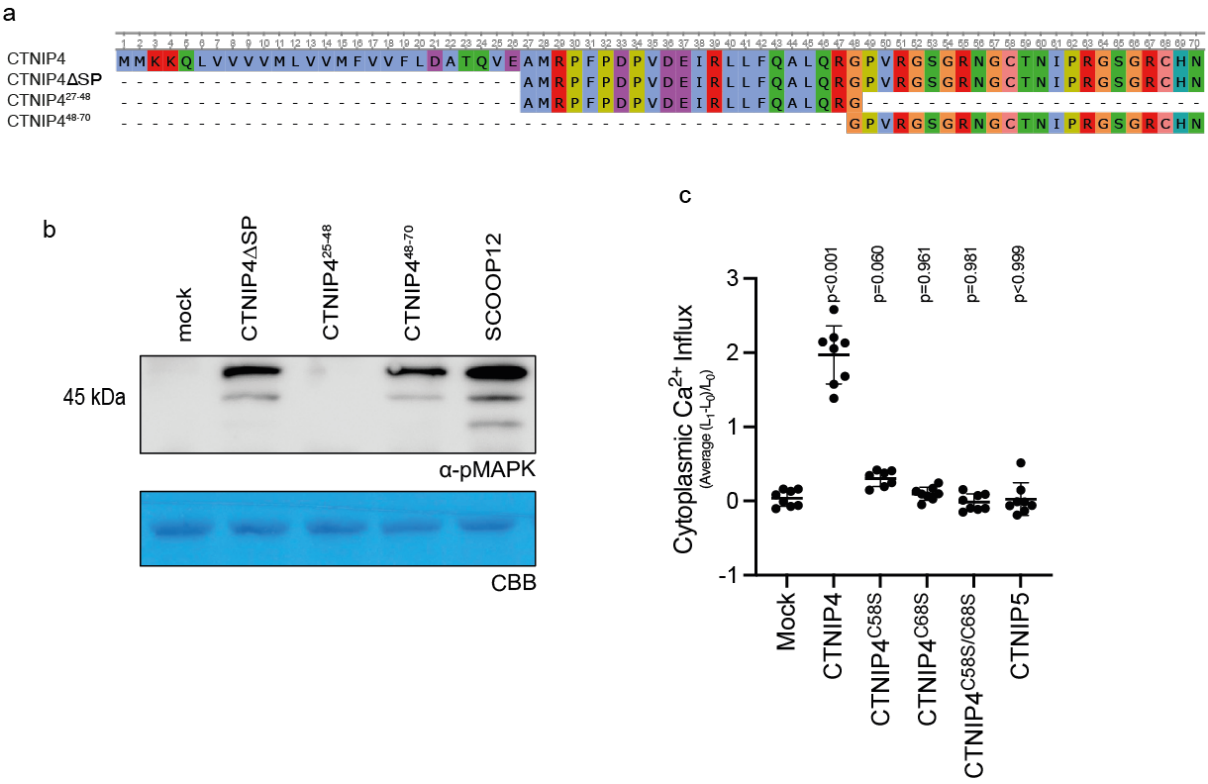


Figure 1 -figure supplement 2. Characterisation of CTNIP fragments

(**a**) Alignment of CTNIP4 fragments used in this manuscript.

(**b**) Western blot using α-p42/p44-ERK recognizing phosphorylated MAP kinases in seedlings treated with 100 nM CTNIP4 fragments or mock for 15 min. The membrane was stained with CBB, as a loading control.

(**c**) Mean relative Ca^{2+} influx induced by 1 µM CTNIP in *p35S::AEQUORIN* seedlings between 5 and 15 min, relative to pre-treatment (*n* = 8 seedlings). A line represents mean; error bars represent S.D. *P*-values indicate significance relative to the WT control in a Dunnett's multiple comparison test following one-way ANOVA.

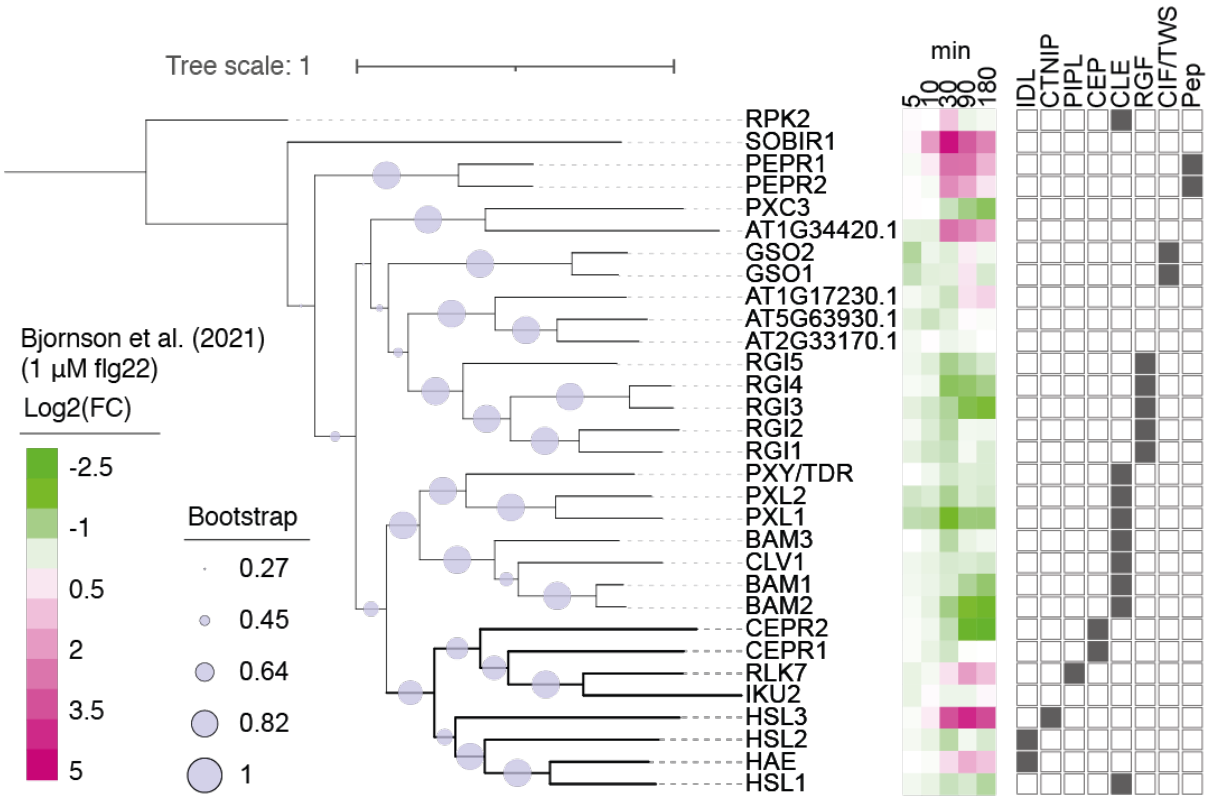


Figure 2 – figure supplement 1. *Arabidopsis* LRR-RK subfamily XI

Phylogeny of full-length protein sequences of the *Arabidopsis* LRR-RK subfamily XI. Sequences were aligned using MUSCLE and a phylogeny was inferred using the Maximum-likelihood method and JTT matrix-based model conducted in MEGAX. 1000 bootstraps were performed and are indicated based on the size of the blue circles. Expression of these receptors in response to 1 µM flg22 treatment was extracted from Bjornson *et al*. (2021) and is represented in a heat map. Known ligands for LRR-RK subfamily XI are highlighted to the right (Butenko et al., 2003; Cho et al., 2008; Crook et al., 2020; Doblas et al., 2017; Doll et al., 2020; Hou et al., 2014; Krol et al., 2010; Morita et al., 2016; Mou et al., 2017; Nakayama et al., 2017; Ogawa et al., 2008; Okuda et al., 2020; Ou et al., 2016; Qian et al., 2018; Rojo et al., 2002; Santiago et al., 2016; Shinohara et al., 2016; Song et al., 2016; Tabata et al., 2014; Yamaguchi et al., 2010, 2006; Zhang et al., 2016).

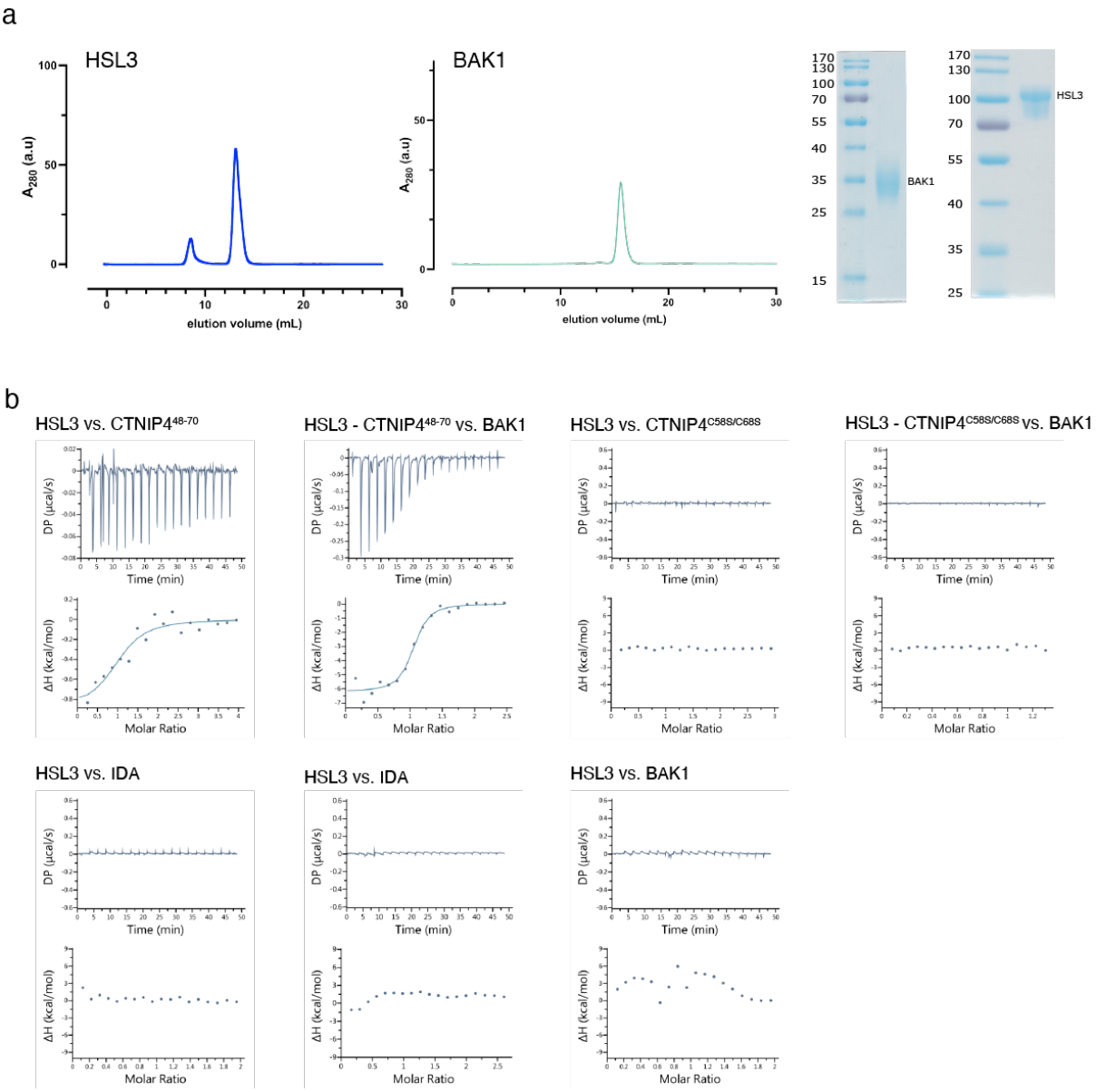


Figure 2 – figure supplement 2. ITC independent experiments and purification of HSL3 and BAK1 used in the binding experiments.

(**a**) Analytical size-exclusion chromatography (SEC) of the ectodomains of HSL3 and BAK1. An SDS PAGE of the two proteins in shown alongside.

(**b**) ITC raw thermograms of experiments shown in the ITC table summary in Figure 2d.

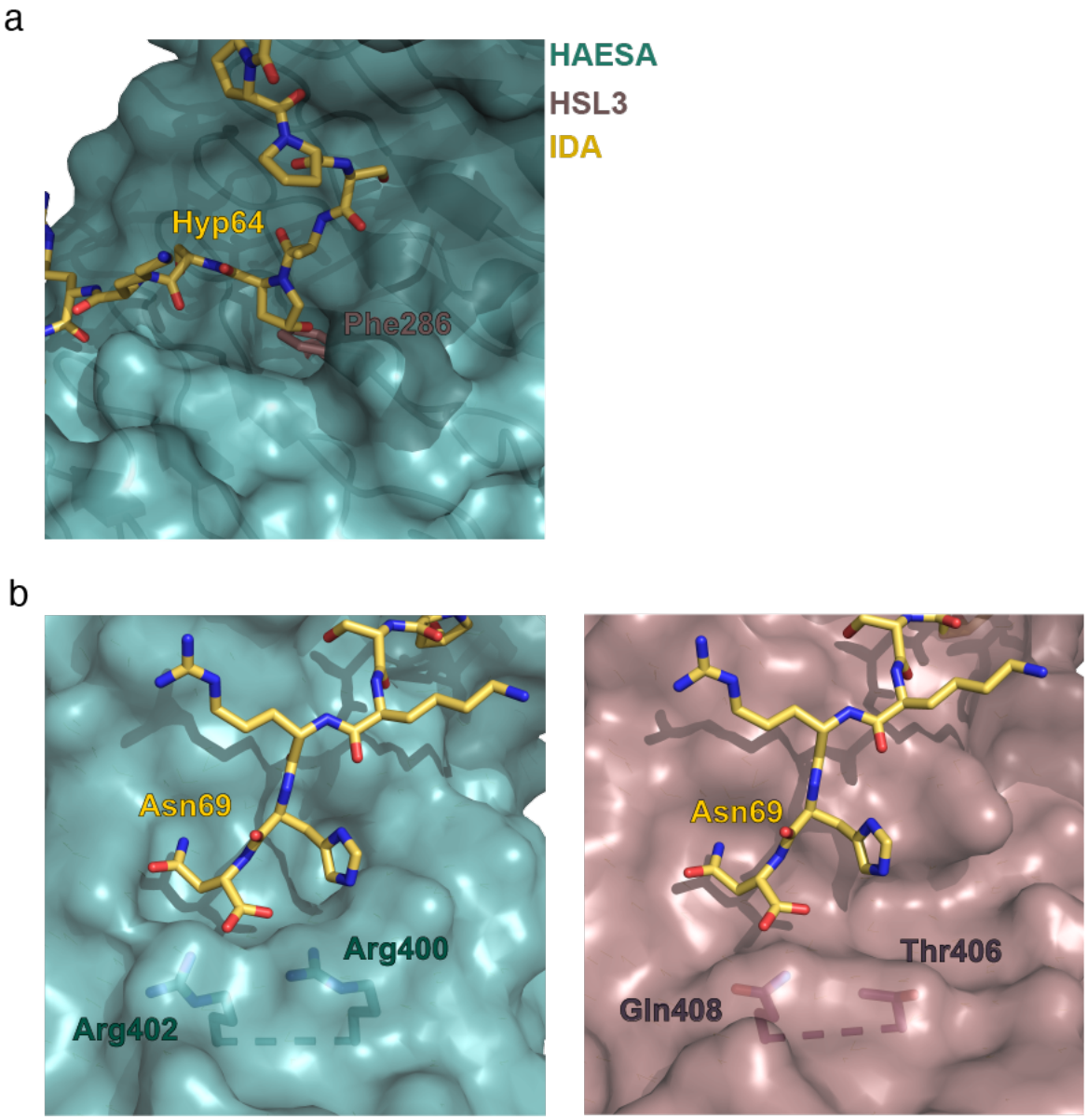


Figure 2 – figure supplement 3. Structural comparison of the binding pockets between the receptors HAESA and HSL3

(**a**) The hydroxyproline pocket required for anchoring the IDA peptide to the HAESA receptor is missing in HSL3. Close view of the binding pocket of the structural superimposition of the HAESA-IDA complex (PDB:5IXQ) and a homology model of HSL3 (AlphaFold: https://alphafold.ebi.ac.uk/). The HAESA receptor is depicted in surface representation in teal blue, IDA in yellow sticks and HSL3 in magenta cartoon. In HSL3, the

hydroxyproline pocket is replaced by the bulky residue Phe286, colliding with the potential anchoring of the IDA peptide to the receptor.

(b) The conserved RxR motif necessary for the coordination of the COO- group the last Asn in IDA is not present in the HSL3 receptor. Zoom in of the C-terminal region of the peptide binding surface of HAESA (teal blue) (left panel) and HSL3 (magenta) (right panel). In HAESA the motif RxR closes the binding pocket allowing for the coordination of the C-terminal of IDA. In HSL3 this structural motif is substituted by the residues Thr406 and Gln408, leaving the binding surface open to potentially accommodate a longer peptide ligand. Figures were done using the PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

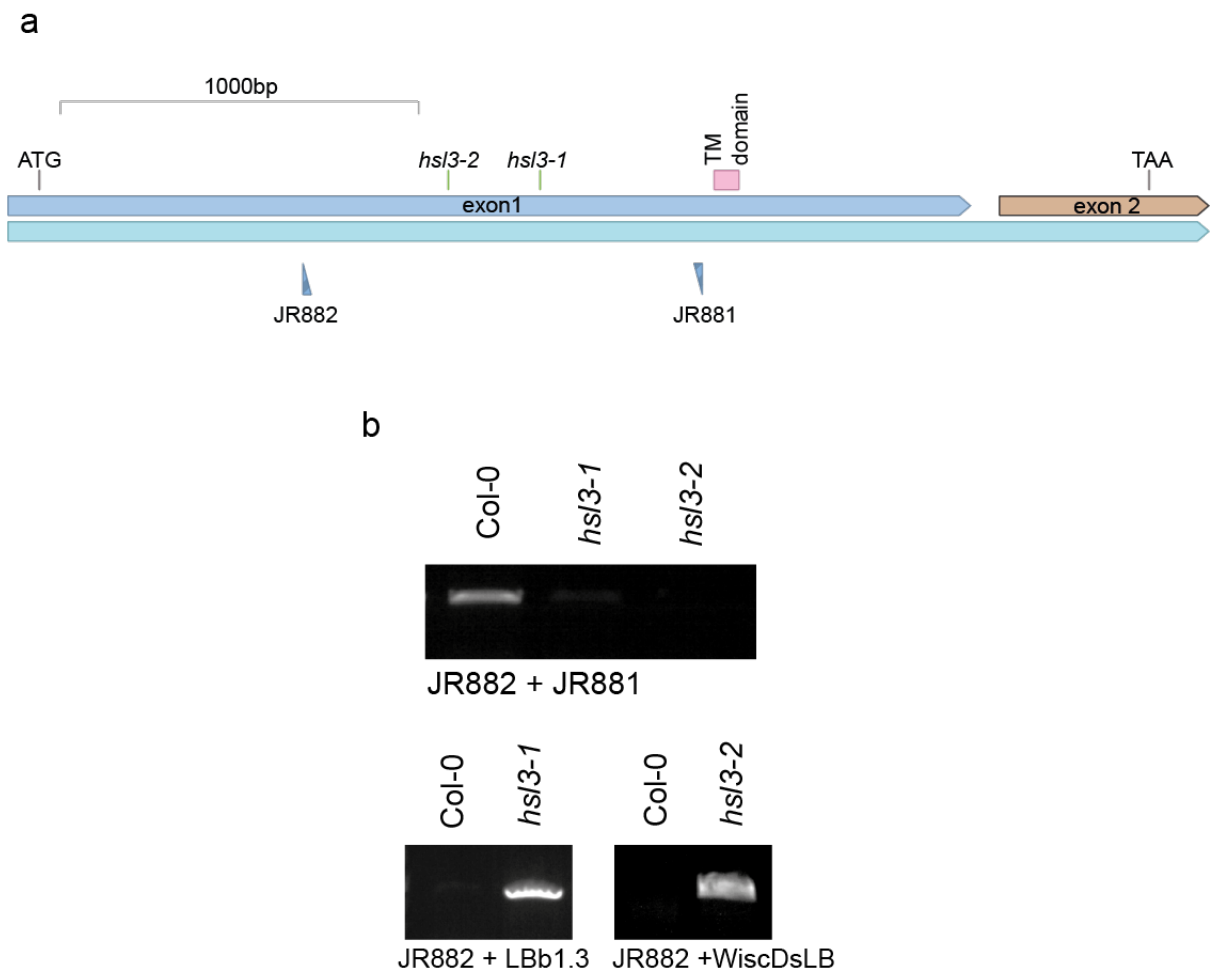


Figure 3 – figure supplement 1. Genetic characterization of *hsl3* mutants

(**a**) Gene model showing the location of T-DNA inserts.

(**b**) PCR confirming T-DNA insertion and mutant homozygosity.

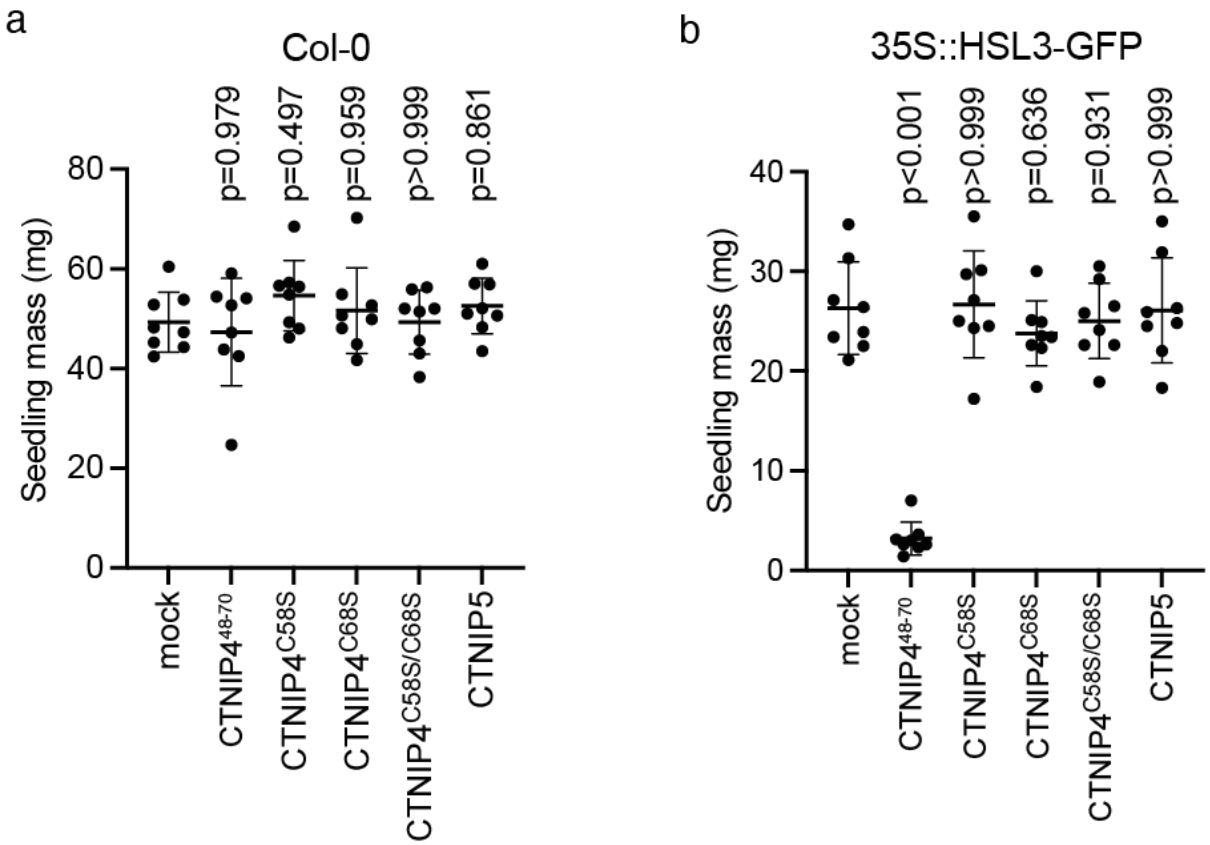


Figure 3 – figure supplement 2. CTNIP-induced seedling growth inhibition

(**a-b**) Fresh weight of 14-day-old seedlings grown in the presence of 500 nM CTNIPs for 10 days relative to mock (*n* = 8 seedlings). A line represents mean; error bars represent S.D.; *P*-values indicate significance relative to the WT control in a Dunnett's multiple comparison test following one-way ANOVA.

Figure 3 – figure supplement 3. Correlation of CTNIP4-induced transcriptomic response with that of elicitors at 30 min

CTINP4-induced gene expression is well correlated with elicitor-induced gene expression from Bjornson *et al.* (2021). Circle colour and size are proportional to the Spearman correlation coefficient (R-squared value) of each pairwise comparison of $\log_2$(fold changes).

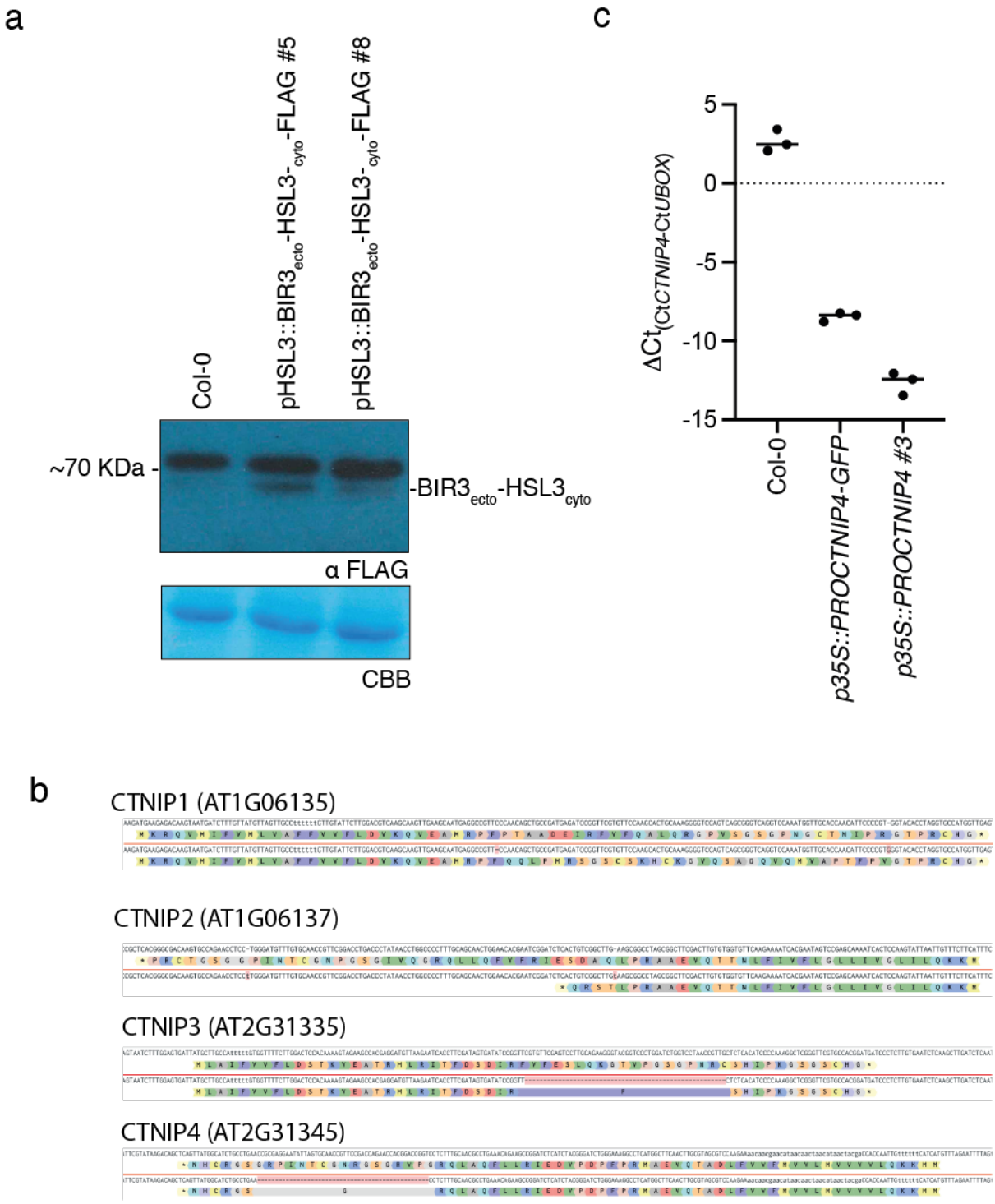


Figure 3 – figure 4. Characterisation of CTNIP and chimeric receptor lines

(**a**) Western blot using α-FLAG recognizing $BIR3_{ecto}$-$HSL3_{cyto}$-FLAG in seedlings to confirm expression. The membrane was stained with CBB, as a loading control.

(**b**) Documentation of the Cas9-induced mutations observed within the *ctnip1-4* polymutant and their predicted effects on protein products.

(**c**) qRT-PCR documenting the overexpression of CTNIP lines. Expression of *CTNIP4* is shown relative to *U-Box* (*At5g15400*). Points represent independent biological replicates each comprising two technical replicates. Lines represent the mean of biological replicates.