Article

Reprogramming of flagellin receptor responses with surrogate ligands

Du-Hwa Lee [1,7] ✉, Ho-Seok Lee[1,2,8], Min-Soo Choi[1], Katarzyna Parys [1,9], Kaori Honda[3], Yasumitsu Kondoh[3], Jung-Min Lee [1], Natalie Edelbacher[1], Geon Heo [1], Balaji Enugutti[1], Hiroyuki Osada [3,4], Ken Shirasu [5,6] & Youssef Belkhadir [1,10] ✉

Receptor kinase (RK) families process information from small molecules, short peptides, or glycan ligands to regulate core cellular pathways in plants. To date, whether individual plant RKs are capable of processing signals from distinct types of ligands remains largely unexplored. Addressing this requires the discovery of structurally unrelated ligands that engage the same receptor. Here, we focus on FLAGELLIN-SENSING 2 (FLS2), an RK that senses a peptide of bacterial flagellin to activate antibacterial immunity in Arabidopsis. We interrogate >20,000 potential interactions between small molecules and the sensory domain of FLS2 using a large-scale reverse chemical screen. We discover two small molecules that interact with FLS2 in atypical ways. The surrogate ligands weakly activate the receptor to drive a functional antibacterial response channeled via unusual gene expression programs. Thus, chemical probes acting as biased ligands can be exploited to discover unexpected levels of output flexibility in RKs signal transduction.

Surface receptors transduce extracellular information inside cells via signal activation pathways[1]. In plants, a dramatically expanded family of sensors termed RKs presumably evolved to interact with a large number of diverse exogenous and endogenous ligands[1]. RKs in plants are classified into subfamilies based on structural features of the extracellular domains (ECDs) that enable detection of specific classes of ligands[1-8]. To date, only a handful of plant RKs have known functions, and even fewer have been assigned a corresponding ligand. RKs have been studied with a guiding principle in which a ligand or a ligand family activates a cognate receptor or receptor family, respectively. This simple 'one-to-one' receptor-ligand relationship, however, poses a repertoire conflict wherein the number of genetically encoded sensors in plant genomes would numerically be surpassed by a very large potential ligand space[3,9].To date, it is unknown if plants RKs evolved to detect structurally dissimilar ligands to solve this repertoire conflict. Whether the sensing of these ligands is coded into differential responses has not been reported and is therefore even less understood.

The most extensively studied plant RKs include members of the leucine-rich repeat receptor kinase (LRR-RKs) family, which are primarily characterized for their activation by short peptides and in few cases by small molecules[1]. In Arabidopsis, the prototypical LRR-RKs, FLS2 detects a short peptide derived from bacterial flagellin (flg22), to unleash a broad-spectrum antibacterial immune response[10,11]. flg22 interaction with the ECD of FLS2 (FLS2^ECD) is essential for recruitment of the co-receptor BRI1-ASSOCIATED KINASE 1 (BAK1)[12,13]. FLS2-BAK1 complex formation induces a canonical signal amplification cascade characterized by the phosphorylation of BOTRYTIS-INDUCED KINASE

[1]Gregor Mendel Institute (GMI), Austrian Academy of Sciences, Vienna BioCenter (VBC), Dr. Bohr-Gasse 3, Vienna, Austria. [2]Center for Genome Engineering, Institute for Basic Science, Daejeon, Korea. [3]RIKEN Center for Sustainable Resource Science, Wako, Saitama, Japan. [4]University of Shizuoka, Shizuoka, Japan. [5]RIKEN Center for Sustainable Resource Science, Yokohama, Japan. [6]Graduate School of Science, The University of Tokyo, Tokyo, Japan. [7]Present address: Department of Chemistry, University of Natural Resources and Life Sciences (BOKU), Vienna, Austria. [8]Present address: Department of Biology, Kyung Hee University, Seoul, Korea. [9]Present address: Faculty of Biology, Genetics, University of Munich (LMU), Martinsried, Germany. [10]Present address: Douar Ifraden, Residence Taghazout Ocean #13, Taghazout, Morocco. ✉e-mail: duhwa.lee@boku.ac.at; ybelkhad@gmail.com

1 (BIK1) and the Respiratory burst oxidase subunit RbohD (among other phosphorylation substrates), consequent production of reactive oxygen intermediates (ROIs), ion channel activation, defense-related mitogen activated protein kinases (MAPKs) activation and transcriptional reprogramming[12,14]. Prolonged exposure of Arabidopsis to flg22 results in seedling growth inhibition (SGI), a physiological response indicative of resources tradeoffs between development and defenses[12].

Here, we identified two synthetic small molecules that interact with and mildly activate FLS2 in Arabidopsis. We named these compounds after the Sanskrit word "Maya" (illusion) because they inform Arabidopsis of the presence of an illusory microbe via FLS2. By weakly activating FLS2, these flagellin surrogates compel atypical immune outputs that reduce bacterial colonization via gene expression programs distinct from those mediated by flg22.

Results and discussion

To identify surrogate ligands for FLS2, we used small-molecule-microarrays (SMMs) populated with a library of compounds from the RIKEN Natural Product Depository (NPDepo). The NPDepo library contains derivatives and analogs of natural products as well as drug-like synthetic compounds[15,16]. We screened pairwise interactions between FLS2^{ECD} and the NPDepo compounds at two different stringencies (Fig. 1a, Supplementary Data 1). Out of the 22,618 compounds tested (Supplementary Data 1), only 84 interacted with FLS2^{ECD} and were named FLS2^{ECD}-interacting compounds (FICs). Thus, 0.37% of the interrogated chemical space yielded a group of structurally diverse small molecules that are putative FLS2 ligands (Supplementary Fig. 1).

To assess the biological relevance of these physical interactions we devised parallel screening cascades (Supplementary Fig. 2). First, we assayed FIC-mediated induction of immune marker gene expression in Arabidopsis leaves. For this, we sourced the FICs from NPDepo, infiltrated the compounds and measured the resulting expression of *FLG22-INDUCED RECEPTOR KINASE 1 (FRK1)*, a gene rapidly induced at early stages of pathogen infection[17] (Fig. 1b). We then ranked the 84 hits by comparing *FRK1* induction levels and focused on the top10 inducers for subsequent assays (Supplementary Data 1). To eliminate false positives, we monitored the activities of the top candidates in *fls2* mutant leaves and excluded 4 compounds because they induced *FRK1* expression in the absence of receptor (Supplementary Data 1). Next, we performed additional rounds of assays with synthetic chemicals (SynChems) corresponding to the 6 NPDepo FICs that passed the *FRK1* induction triage in leaves. Only 4 SynChems induced *FRK1* expression. The two remaining compounds were set-aside because one was not commercially available and the other was unable to induce *FRK1* expression. To provide orthogonal confirmation that the interactions identified via SMMs are robust, we used microscale thermophoresis (MST) and established that one SynChem (FIC32; hereafter named Maya1), interacted robustly with FLS2^{ECD} in the micromolar range (Fig. 1c–e). Maya1 induced *FRK1* expression, albeit to lower levels than flg22 (Fig. 1f). Maya1 is a tetrapeptide-based peptidomimetic that shares high chemical similarity with FIC31, another peptidomimetic that only modestly induced *FRK1* in our screening cascade (Supplementary Data 1). Thus, our leaf triages yielded a weak, yet specific, peptidomimetic agonist for FLS2 (Fig. 1f).
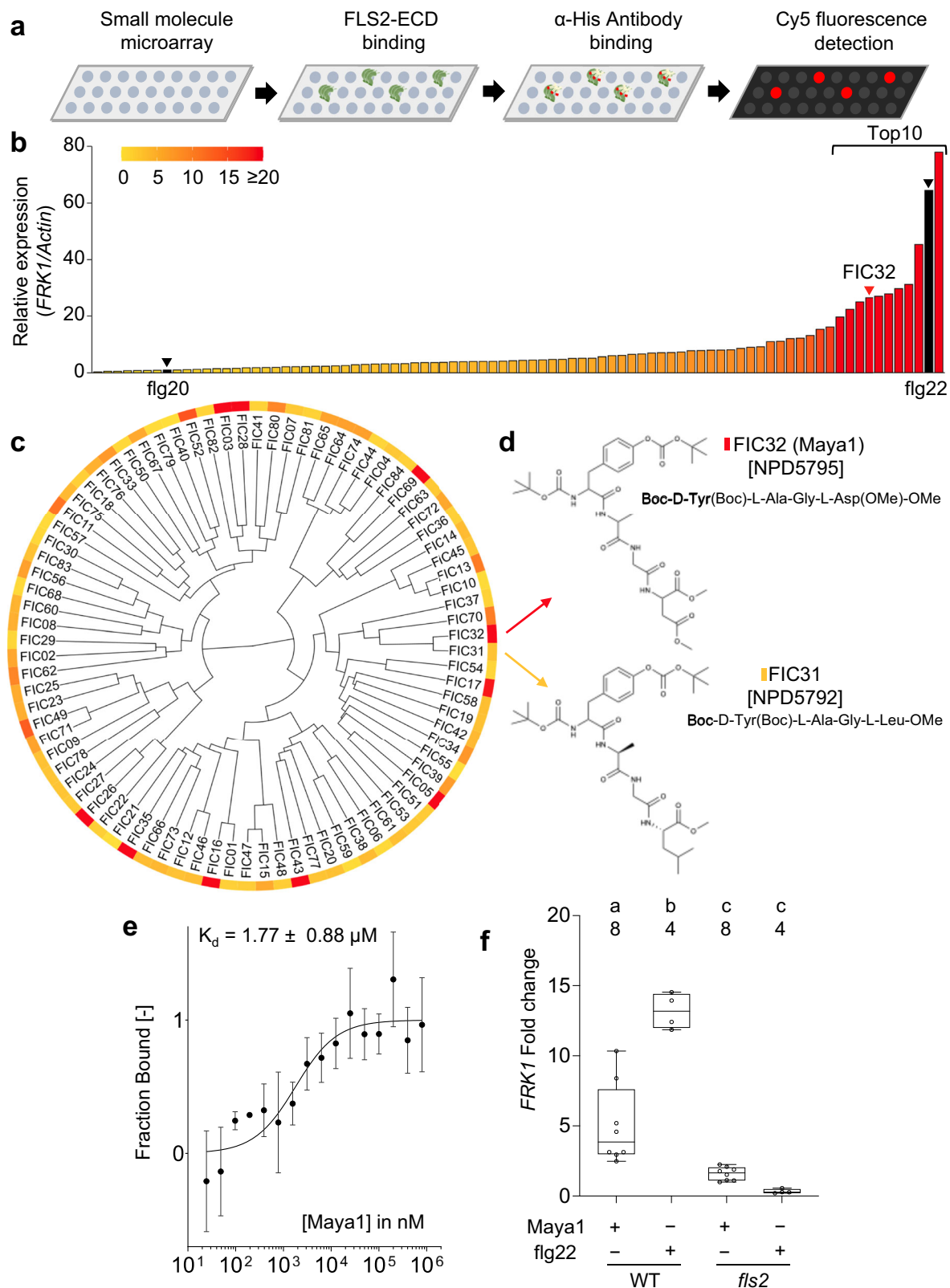
In parallel to the leaf triage, we used imaging to read-out *FRK1* expression in root cells with a fluorescent-based transcriptional reporter[17] (Supplementary Fig. 2). We quantified *FRK1* induction mediated by the 84 FICs sourced from NPDepo using fluorescence intensity and selected the top 10 inducers for further characterization (Fig. 2a, b). Interestingly, the groups containing the top inducers from leaves or roots were for the most part populated by non-overlapping sets of FICs (Supplementary Data 1). Yet, nearly half of the FICs from both groups shared a common scaffold displaying only differences in functional groups (Supplementary Fig. 3a, b). We attribute the bioactivity profiles of these structural analogs to their capacity to engage

FLS2 in these distinct organs. To assess the robustness of *FRK1* induction in roots with the top10 FICs, we performed multiple independent assays involving larger numbers of biological replicates than in the primary screen (Supplementary Fig. 2). We excluded two candidates because *FRK1* induction fluctuated across replicates (Supplementary Data 1). Next, we sourced SynChems corresponding to these FICs and obtained 7 compounds out of 8. Of these, only 4 induced *FRK1* in roots reliably, and only one (FIC44; hereafter referred to as Maya2) interacted with FLS2^{ECD} in MST assays (Kd = 8.56 ± 3.07 μM) (Fig. 2c, d). FIC04, a compound displaying high chemical similarity to Maya2 (Supplementary Fig. 4a, b), did not pass our cutoff a *FRK1* inducer in roots. Thus, we used FIC04 as an interaction control to assess the functional selectivity of this triage. We determined that FIC04 displayed a Kd at least an order of magnitude higher than that of Maya2 (Kd = 93.91 ± 19.51 μM) (Supplementary Fig. 4c). While Maya2 induced *FRK1* expression in an *FLS2*-dependent manner, FIC04 barely did even at higher doses (Fig. 2e–h, Supplementary Fig. 4d, e). The weaker affinity and activity of FIC04 compared to those of Maya2 highlights that differences in functional groups on the same scaffold can modulate interaction with and activation of FLS2. To extend these findings, we monitored the expression of two other flg22-responsive genes[17–19]. For this, we used transcriptional reporters to readout *SUGAR TRANS-PORT PROTEIN 13 (STP13)* and *PEROXIDASE 5 (PER5)* induction. Maya2 induced the expression of *PER5* and *STP13* in the root elongation zone and the root vascular system, respectively (Supplementary Figs. 5, 6). Maya2 mediated a weaker induction of *FRK1* and *PER5* when compared to flg22, whereas it showed an overall stronger induction of *STP13* compared to flg22 at the concentrations used. Thus, our screening cascade yielded another specific small molecule that interacts with FLS2 and weakly induces *FRK1* in roots.

The results of our cascading triages indicate that while FLS2 can interact with several dozen small molecules, only a subset of these act as receptor agonists and induce *FRK1* expression. It remains possible that some of the FICs did not pass our triages simply because they are receptor antagonists. Alternatively, the triage of FICs using other defense marker genes might have identified other functional ligands. Altogether, our results show that the extracellular LRR domains of FLS2 can be hijacked by small molecules or peptidomimetics to induce defense marker gene expression.

Next, we investigated the potential molecular mechanisms by which Maya1/2 drive FLS2 activation. In addition to BAK1, other members of the SOMATIC EMBRYOGENESIS RECEPTOR KINASES (SERKs) family of co-receptors (of which BAK1 is the third member (SERK3)) interact with and modulate the signaling competency of FLS2 in vivo[12]. Thus, we conducted a proteomic analysis using co-immunoprecipitations (Co-IPs) coupled to mass spectrometry to compare the FLS2 interactomes induced by the natural ligand or the surrogates (Supplementary Fig. 7a). Using a principal component and clustering analysis, we found that the interactomes resulting from surrogate treatments were distinct from that of flg22 (Supplementary Fig. 7b, c). In contrast to flg22, we found that none of the chemicals significantly induced FLS2-BAK1 or FLS2-SERKs heterocomplexes formation in our proteomic analysis (Fig. 3a, Supplementary Data 2). We further confirmed that the compounds were not able to induce a stable interaction between the FLS2^{ECD} and BAK1 in vitro or between the full-length proteins in Co-IPs assays in vivo (Fig. 3b, Supplementary Fig. 8b, c). Thus, we propose that due to their low affinity for FLS2, Maya1/2 induce very transient FLS2-BAK1 heterocomplex that falls below the threshold of detection of all the methods we employed to address this. Although unlikely, it is possible that Maya1/2 activate FLS2 via a mechanism that does not rely on BAK1.

To test this genetically, we measured if the loss of *BAK1* function altered the activities of the surrogates. While most of the activity in FLS2 signaling is due to BAK1, there is some level of functional overlap with other SERKs. These overlapping activities are reduced in the semi-

dominant *bak1-5* mutant when compared to the *bak1-4* null mutant. Thus, we treated the leaves of the semi-dominant *bak1-5* mutant, with the compounds and monitored *FRK1* induction. While flg22-induced expression of *FRK1* expression was significantly reduced in *bak1-5* and *fls2* when compared to wildtype, we found that Maya1 induced *FRK1* expression to wildtype levels in *bak1-5* but not in *fls2* (Fig. 3c). Consistent with the results of our primary leaf screen wherein

Maya2 showed no potential for *FRK1* induction, the compound failed again to induce *FRK1* expression in the leaves of all genotypes tested (Fig. 3c). Next, we used SGI to test if prolonged exposure to these compounds over days would reveal an involvement of BAK1. While Maya1 failed to induce SGI in all the genotypes assayed, Maya2 induced a moderate yet significant fresh weight reduction in wildtype and *bak1-5* mutant seedlings (Fig. 3d). Notably, the mild inhibitory effects of

Fig. 1 | Identification of a functional peptidomimetic-based ligand of FLS2 using large-scale reverse chemical genomics. a Chemical array interaction pipeline used in the study. **b** Bar plots distribution of *FRK1* expression in leaves treated for 2 h with FICs at 1 µM or with flg20 at 100 nM (black arrowhead). FICs are ordered from left to right by increasing *FRK1* expression. Bars represent the normalized mean of qPCR quantifications (*n* = 3). The color gradients indicate the level of *FRK1* expression levels. The top 10 hits are indicated in bracket. **c** Dendrogram showing chemical relatedness between FICs. The color gradients indicate the level of *FRK1* expression in (**b**). **d** Chemical structures of FIC32/Maya1 and FIC31. **e** Quantification

of interaction between Maya1 and FLS2^{ECD} using MST. Data points in the fitted dose-response binding curve indicate the fraction of labeled FLS2^{ECD} bound to Maya 1 (mean ± SE; *n* = 4). **f** qPCR analysis, showing relative expression levels of *FRK1* from wildtype (WT) and *fls2* leaves treated for 2 h with Maya1 (1 µM; *n* = 8) or flg22 (10 nM; *n* = 4). The box plot represents the first and third quartiles, centered by the median, and whiskers include the min-max of the data points. Letters above the boxes (a–c) indicate the results of one-way ANOVA followed by a Tukey's multiple comparison test. Groups with the same letter are indistinguishable at >95% confidence. *p*-values are provided in the source data file.

Maya2 on seedling development required *FLS2* and were as expected from the results of our triages more pronounced in roots than in shoots (Supplementary Fig. 9). However, Maya2 did not induce a stable interaction of FLS2-BAK1 complexes in Co-IP assays specifically performed on Arabidopsis root tissues. Thus, the distinct activities of Maya2 in leaves and roots are unlikely to stem from differences in BAK1-FLS2 heterocomplex formation (Supplementary Fig. 8d). Although our results indicate that *BAK1* does not seem to be required genetically for the activities of Maya1/2, the differences we observed are perhaps due to confounding changes in potency between the surrogates and flg22. Alternatively, other SERK family members functionally complement the loss of *BAK1* function. More potent Maya ligands will be essential to establish true mechanistic differences for BAK1/SERKs requirement in Maya1/2-induced FLS2 signaling.

In addition, we found that Maya1/2 treatments had no significant impact on BIK1 phosphorylation, ROI production, and MAPK activation even during an increase in treatment duration or ligand concentrations. Maya2 was only able to modestly activate MAPKs at the 15-min mark (Supplementary Fig. 10a–e). Quantitative differences in interaction with FLS2 which directly impinge on signaling strength and dynamics can possibly explain the differences in canonical outputs induction between flg22 and the Maya ligands. Overall, our results are reminiscent of the uncoupling of FLS2 response outputs mediated by synthetic and natural polymorphisms in flg22 epitopes[20,21].

Next, we investigated if Maya1/2 compete with flg22 for the same binding site on FLS2. First, we used flg22-induced FLS2-BAK1 complex formation as a readout to monitor competition of Maya1/2 on this molecular step. For this, we performed in vivo FLS2-BAK1 Co-IPs after subjecting seedlings to increasing doses of the surrogate ligands in the presence of flg22. While the effects were relatively mild, we observed that both surrogates consistently enhanced flg22-induced heterocomplex formation (Fig. 3e, f). While Maya2 readily enhanced BAK1-FLS2 complex formation at 1 µM, Maya1 required a 10-fold molar excess to barely reach the efficacy of Maya2 for this readout (Fig. 3e, f). To control that Maya1/2 activities on flg22-induced FLS2-BAK1 heterocomplex formation were specific, we performed a similar set of assays on PEP RECEPTOR 1 (PEPR1), another peptide binding immune receptor that requires BAK1 for signaling[22,23]. None of the surrogate ligands influenced pep ligand-induced PEPR1/BAK1 complex formation in vivo, indicating that the compounds are selective FLS2 agonists (Fig. 3g, h). Additionally, high concentration of Maya2, but not Maya1, mildly enhanced flg22-induced BIK1 phosphorylation (Supplementary Fig. 11a–d). Collectively, our results suggest that the surrogates, specifically Maya2 at high concentrations, can boost the activities of flg22 on BAK1-FLS2 complex formation.

To investigate the potential binding sites of the surrogates on FLS2^{ECD}, we conducted molecular docking simulations (Supplementary Fig. 12). Arg^{440}, Arg^{533}, Asp^{605} and Tyr^{629} residues on FLS2^{ECD} were predicted as critical residues for interaction with the surrogates based on energetically favorable multiple docking poses (Fig. 4a, Supplementary Fig. 13, Supplementary Data 3). Notably, these residues are distinct from those required for interaction with flg22. Only one docking pose for Maya2 involved Arg^{294} (Supplementary Fig. 13b), a key residue involved in interaction with flg22[13]. To confirm these predictions, we introduced missense mutations targeting independently these
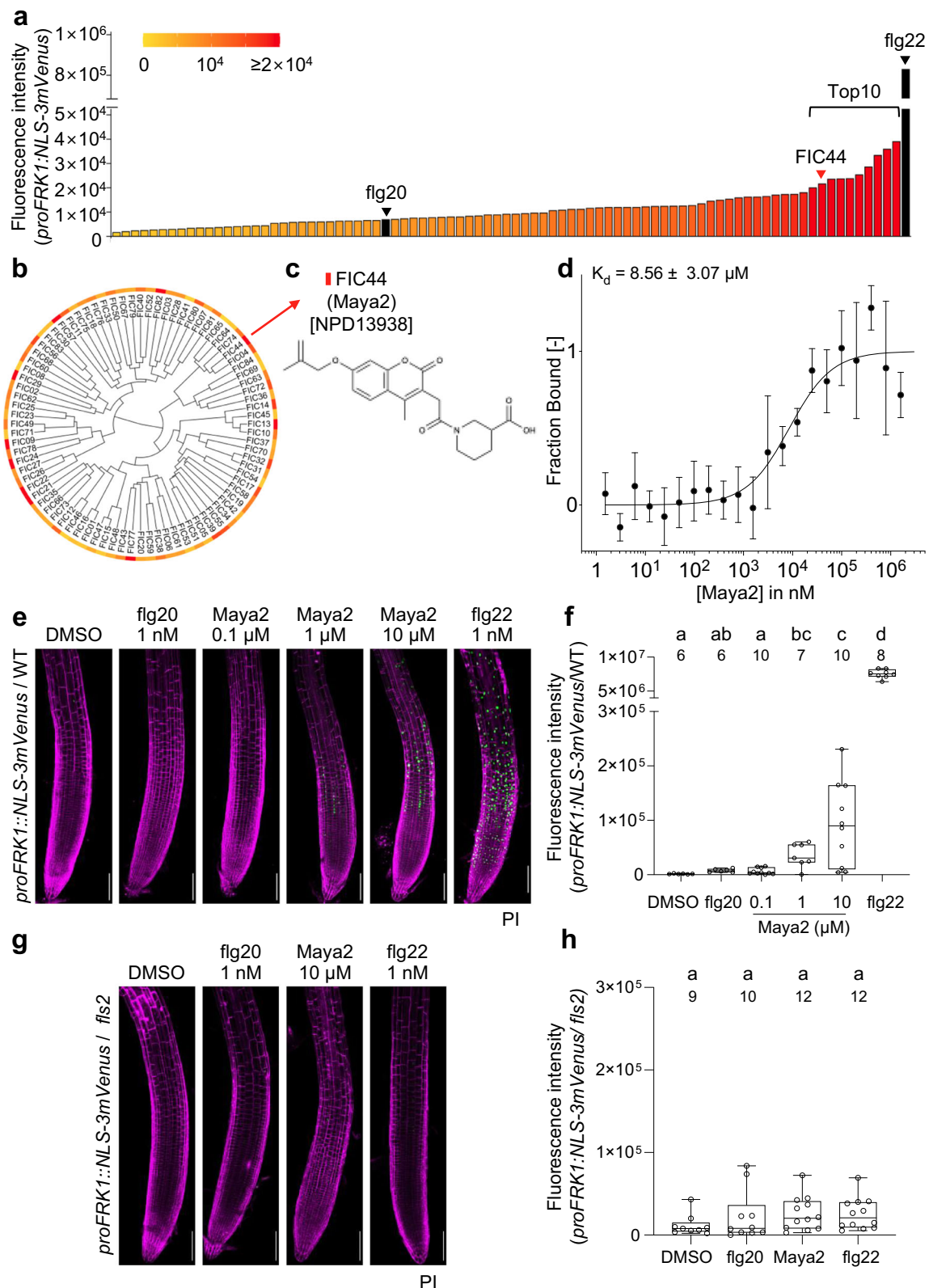
residues in *FLS2* and tested the functionality of the receptor variants by complementing *fls2* mutant lines (Fig. 4b–e, Supplementary Fig. 14). Alanine replacements at Arg^{440}, Arg^{533}, Asp^{605} and Tyr^{629}, but not at Arg^{294}, drastically decreased induction of *FRK1* by Maya1 (Fig. 4b). Conversely, flg22-mediated induction of *FRK1* was strongly reduced in FLS2^{R294A} but only mildly in the other receptor variants (Fig. 4c, e). Since Maya2 induction of *FRK1* was only reduced by mutations of Asp^{605} and Tyr^{629} (Fig. 4d), we conclude that Maya1/2 functionally interact with FLS2 in different ways. Our results also imply that the Maya ligands can interact with FLS2 by engaging residues that are not required for flg22 interaction and activity.

Next, we characterized the genome wide Maya1 and Maya2 transcriptional responses using RNA sequencing (RNAseq) in leaves and roots, respectively. First, we defined the repertoire of differentially expressed genes (DEGs) in each organ after treatment with flg22 or the surrogates in both wildtype and *fls2* mutant plants (Fig. 5a–h). Hierarchical clustering revealed a transcriptional sector composed of 156 *FLS2*-dependent Maya1-responsive genes (Fig. 5a, Supplementary Data 4). Gene Ontology (GO) analyses of the FLS2-dependent Maya1 regulon highlighted biological processes such as non-self-responses (Fig. 5b). Notably, 62.8% of the DEGs are co-regulated by flg22, whereas 37.2% are specific to Maya1 (Fig. 5c, d, Supplementary Data 5). Maya2 also induced a root transcriptional response involving 113 *FLS2*-dependent DEGs of which, only 49 are also regulated by flg22 (Fig. 5e, Supplementary Data 6). GO analyses of the *FLS2*-dependent Maya2 DEGs in roots revealed biological processes related to plant immune responses (Fig. 5f). 43.4% of these DEGs are also regulated by flg22, whereas 56.6% are specific to Maya2 (Fig. 5g, h, Supplementary Data 7). These Maya1/2-dependent gene expression changes were further validated by qPCR on a target set of Maya1 or Maya2 specific genes (Supplementary Fig. 15a, b). While the gene expression changes induced by the surrogates are less extensive than those induced by flg22, our RNAseq profiling of the Maya1/2 responses shows that the compounds preferentially activate target transcriptional programs, instead of a global immune program as does flg22[14].

Since our SGI and Co-IP results with Maya2 revealed that this molecule is also active in leaves (Fig. 3d, e), we also defined the genome-wide transcriptional changes mediated by the surrogate in leaves. In contrast with roots, Maya2 induced a larger transcriptional program involving 992 *FLS2*-dependent DEGs in leaves, of which only 3.25% are co-regulated by flg22 (Supplementary Fig. 15e, Supplementary Data 7, 8). Notably, Maya2 induced organ-specific transcriptional responses characterized by differential *FLS2*-dependent gene programs in leaves and roots (Supplementary Fig. 15f, Supplementary Data 9). Similarly, Maya1 and Maya2 transcriptional responses in leaves only marginally overlapped (Supplementary Fig. 15g, Supplementary Data 10). Altogether, our findings indicate that the modest activation of FLS2 by Maya ligands initiates distinct transcriptional responses, incorporating only a small portion of the flg22-mediated pathways, while also activating unique Maya-specific gene expression programs that are for the most part not directly involved in plant defense processes (Supplementary Fig. 16).

Finally, we tested if the reprogramming of FLS2 transcriptional outputs mediated by the surrogates was sufficient to enhance

a Fluorescence intensity (proFRK1:NLS-3mVenus)

b

c FIC44 (Maya2) [NPD13938]

d $K_d = 8.56 \pm 3.07\ \mu M$

Fraction Bound [-] vs [Maya2] in nM

e proFRK1::NLS-3mVenus / WT — DMSO, flg20 1 nM, Maya2 0.1 μM, Maya2 1 μM, Maya2 10 μM, flg22 1 nM, PI

f Fluorescence intensity (proFRK1:NLS-3mVenus/WT) — DMSO, flg20, 0.1, 1, 10, flg22; Maya2 (μM)

g proFRK1::NLS-3mVenus / fls2 — DMSO, flg20 1 nM, Maya2 10 μM, flg22 1 nM, PI

h Fluorescence intensity (proFRK1:NLS-3mVenus/fls2) — DMSO, flg20, Maya2, flg22

resistance against the foliar pathogen *Pseudomonas syringae* pv. *tomato* (*Pto*) DC3000 (hereafter *Ps*). Leaves pretreatments with flg22 prior to pathogen exposure prime resistance to subsequent *Ps* inoculation in Arabidopsis[10]. Thus, we assessed if Maya1/2 primed resistance against *Ps* after establishing that treatments with the compounds had no obvious effects on leaf development and *Ps* growth in vitro (Supplementary Fig. 17). Maya1/2 treatments resulted in a 10-fold reduction

of *Ps* growth in wildtype but not in *fls2* (Fig. 5i, j). Maya1/2 treatments also reduced the growth of *PsΔfliC*, a strain that lacks flagellin and no longer releases flg22 (Supplementary Fig. 18). Thus, our results show that Maya1/2, by regulating minimal *FLS2*-dependent gene sets, promote resistance against a bacterial pathogen even in the absence of flg22. Given the successful priming of defense by the Maya ligands, we speculate that the small fractions of overlapping genes induced by

Fig. 2 | Maya2 interacts with and partially activates FLS2 in roots. a Bar plots distribution of *FRK1* expression in roots after treatments of *proFRK1:NLS-3mVenus* seedlings for 48 hours with FICs at 1 μM or with flg20 at 100 nM (black bar). FICs are ordered from left to right by increasing mVenus fluorescence intensity. Bars represent the normalized mean fluorescence intensity from at least 5 independent roots. The color gradients indicate the level of fluorescence intensity. The top 10 hits are indicated in bracket. **b** Dendrogram showing chemical relatedness between FICs. The color gradients indicate the level of *FRK1* expression in (**a**). **c** Chemical structures of Maya2. **d** Quantification of interaction between FLS2^{ECD} and Maya2 using MST. Data points in the fitted dose-response binding curves indicate the fraction of labeled FLS2^{ECD} bound to Maya2 (mean ± SE; $n = 8$). **e** *FRK1*-induction patterns in *proFRK1:NLS-3mVenus* roots treated for 48 hours with flg20, flg22 or increasing concentrations of Maya2. Concentrations indicated on top. The nuclear-

localized mVENUS signals (green) are co-visualized with the plant cell wall marker propidium iodide (PI, magenta). Scale bar, 100 μm. **f** Quantitative analysis of mVENUS fluorescence intensity after treatment with the ligands indicated at the bottom. The box plot represents the first and third quartiles, centered by the median, and whiskers include the min-max of the data points. **g, h** same as (**e, f**) in the *fls2* mutant background. Statistical significance was assessed by linear mixed effect models; letters above the boxes (a–d) indicates the results of a one-way ANOVA followed by a Tukey's multiple comparison test (two-sided, adjusted for multiple comparisons, $p < 0.05$). Groups with the same letter are indistinguishable at >95% confidence. The numbers of biologically independent observations (n) are indicated on below the letters on the graph. *p*-values are provided in the source data file.

flg22 and any of the surrogates, are sufficient for a functional immune response (Supplementary Data 5, 7, 9).

Here, we demonstrate that FLS2, a prototypical member of the LRR-RK family—the most expanded subgroup of RKs in plants with over 220 evolutionarily related members in Arabidopsis—interacts with and is weakly activated by structurally distinct molecules. These molecules drive a biased signaling response, resulting in expression changes of genes that overlap with and are distinct from those normally regulated by the natural ligand. We propose that FLS2 can act as a 'biased' receptor by converging signals from structurally different ligands into functionally equivalent responses. While Maya1 and Maya2 act as weak agonists that do not induce the full spectrum of responses typically associated with more potent agonists like flg22, they define two separate gene programs of the flg22 output that are sufficient for driving antibacterial immunity. This partial activation can be insightful for understanding the threshold of activation necessary for effective immunity and may have implications for understanding signaling output flexibility for this large family of receptors.

In a natural context, direct modulation of FLS2 signaling by endogenous or exogenous metabolites produced either by plants or microbes could conceivably produce similar immune outputs. While both Maya1 and Maya2 are synthetic compounds from NPDepo, Maya2 has a 2H-chromen-2-one structural core, characteristic of coumarins—a group of plant-produced secondary metabolites[24]. Coumarins are secreted by roots in large quantities to assist in iron uptake and to establish a niche for balanced microbiome colonization in the rhizosphere[24–27]. Other natural coumarins have also been reported to display protective effects against certain types of phytopathogens[28]. While the natural coumarins we have tested in this work showed either lower or no *FRK1* induction compared to Maya2 in roots (Supplementary Fig. 19), it is still tempting to speculate that other naturally occurring coumarins could boost immunity strongly via FLS2 activities. Based on our current results, we cannot decisively conclude that an flg22-independent FLS2-mediated defense pathway occurs in nature. Further studies will be required for exploring this possibility, and the discovery of Maya2 serves a promising starting point for these investigations.

We previously demonstrated that extensive changeover in LRR-RKs interaction networks provides additional determinants of response specificity to ligands[29]. Here, we expand on these findings to propose that combinatorial interactions among LRR-RKs, coupled with the ability of these sensors to perceive structurally distinct ligands, occur to accommodate myriad ligands and signaling responses. Our results not only challenge the 'one-to-one' architectures of LRR-RK signaling pathways, but also raise evolutionary questions that include how multiligand sensory capacities have initially emerged. While HYDROGEN-PEROXIDE-INDUCED Ca²⁺ INCREASES 1/ CANNOT RESPONSE TO DMBQ1 (HPCA1/CARD1) provides a precedent for an LRR-RK that can act as a receptor for hydrogen peroxide[9], as well as a sensor for 2,6-dimethoxy-p-benzoquinone[3], our findings expand on this by demonstrating that a single LRR-RK can serve as a receptor for both, small molecules and short peptides.

Our work lays a foundation for the discovery and engineering of more potent synthetic Maya ligands displaying higher affinities for FLS2 to elucidate the precise mode of activation mediated by these compounds. Chemical biology approaches have been instrumental to study hormone receptor pathways and endomembrane trafficking in plants[30–33]. The chemical engineering of abscisic acid receptor agonists has also demonstrated that chemical biology offers great possibilities to improve plant water use for translational aspects[34,35]. Here, we provide proof-of-concept demonstration for a research pipeline that could be used to discover biased ligands that target the signal sensing modules of RKs regulating plant responses to fungi, parasitic plants, and oomycetes for translational research[1].

Methods

Plant materials and growth conditions
Arabidopsis thaliana accession Columbia (Col-0) was used as wildtype (WT). The mutant line *fls2* (SAIL_691_C4)[10], and *bak1-5*[36] were previously described. The transgenic lines *proFRK1:NLS-3mVenus/WT*, *proFRK1:NLS-3mVenus/fls2*, *proPER5:NLS-3mVenus/WT* [17], *proFLS2:FLS2-GFP* [37], *proBIK1:BIK1-HA*[38] and *proPEPR1:PEPR1-GFP/pepr1/pepr2*[22] were described previously. *proSTP13:HTA6-GFP* was generated in this study. Plant seeds were surface-sterilized and grown on soil in a growth chamber at 22 °C under short-day light conditions (12 h light/12 h dark) or on Petri dishes containing 1/2 Muraschige and Skoog (MS) medium including Gamborg B₅ vitamins (M0231, Duchefa), 1% sucrose (AppliChem), and 0.8% plant agar (Duchefa) in a growth room at 22 °C under long-day light conditions (16 h light/8 h dark).

Generation of transgenic lines
The combined promoter (2631 bp) and the 5′ UTR (227 bp) sequence of *FLS2* (*proFLS2$_{long}$*) was cloned into the 5′ entry clone (pENTR-L4R1-MCS) using the In-Fusion HD Cloning kit. The full-length genomic sequence of *FLS2* lacking a stop codon was cloned into an entry clone using pDONR221. The *FLS2* entry clone was used for site-directed mutagenesis to generate R294A, R440A, R553A, D605A, and Y629A variants. See Supplementary Table 1 for primer information. The wildtype and mutated versions of the *proFLS2$_{long}$:gFLS2-2mcherry-2Myc* expression clone were generated by multisite gateway cloning using the following constructs: destination vector (pDEST-R4R3-pAlli:YFP), the 5′ entry clone (pENTR-L4-*proFLS2$_{long}$*-R1), respective CDS entry clone (pENTR-L1-*gFLS2*-L2), and 3′ entry clone (pENTR-R2-2mcherry-2Myc-L3). All expression clones were transformed into *Agrobacterium tumefaciens* GV3101 pSoup strain using the freeze-and-thaw method, and then transformed into the *fls2* and *proFRK1:NLS-3mVENUS/fls2* backgrounds by floral dipping method[39]. Transgenic T1 seeds were selected using fluorescence. For each construct, a mix of independent T1 plants were then tested for responses to Maya1/2 and flg22. For leaf tests, a total of 8 independent T1 plants per constructs were used across 2 separate experiments. For root tests, at least 5 independent T1 plants
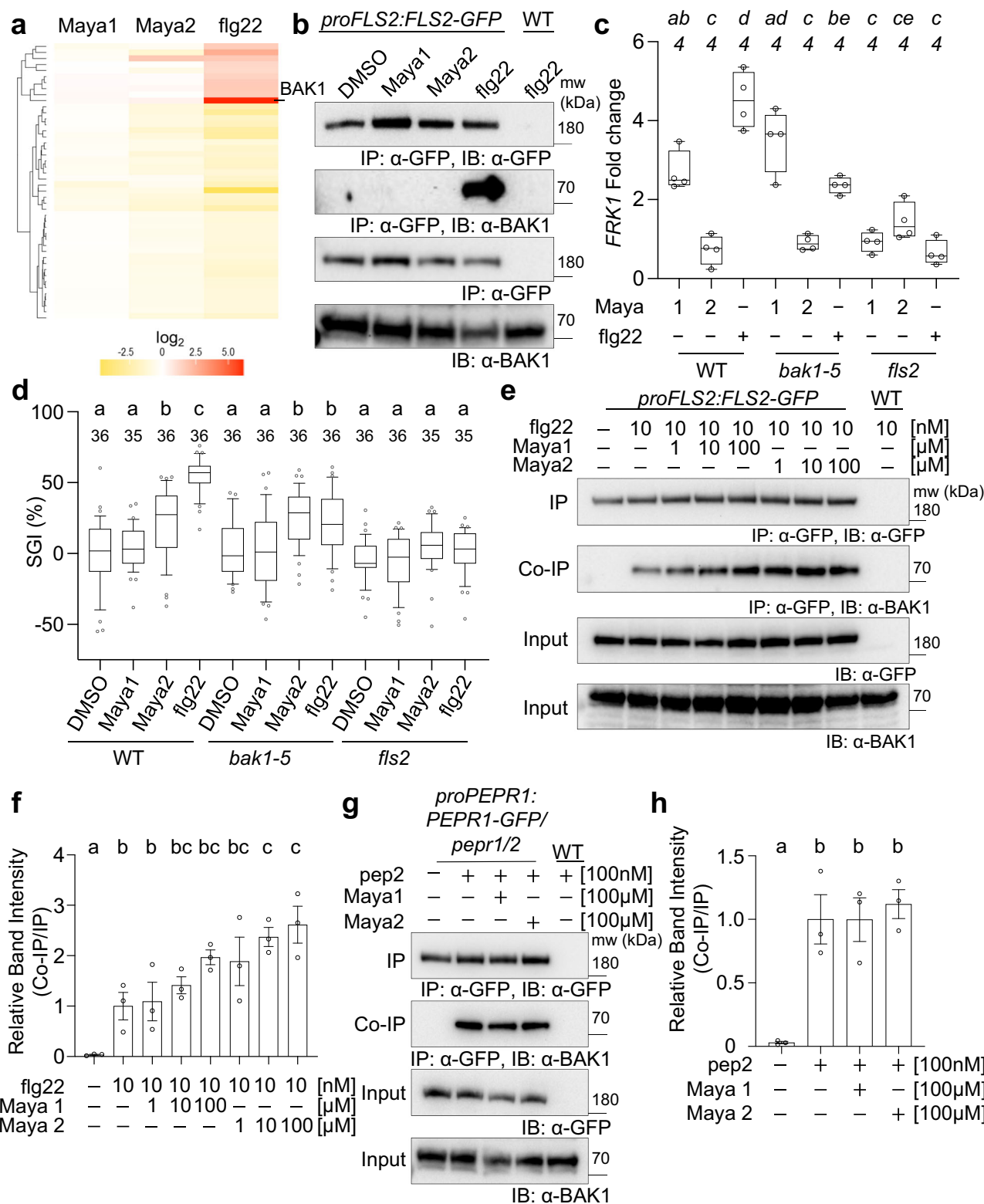
per genotypes were used across 2 separate experiments. The *STP13* transcriptional combined promoter (1181 bp) and the 5′ UTR (546 bp) sequence of *STP13* (*proSTP13*) was cloned into the entry clone using the pENTR™/D-TOPO cloning kit, and then the entry clone was subcloned into the pCGTAG vector to express the nuclear reporter HTA6(Histone H2A 6)-GFP under the *proSTP13* promoter. The expression clone was transformed into wildtype Col-0 plants by Agrobacteria-mediated floral dipping method and T3 homozygotes were used for experiments.

FLS2 extracellular domain expression and purification

The expression and purification of the FLS2^{ECD} was performed as described in Parys et al.[20]. Briefly, FLS2^{ECD} (amino acid 24–804) was cloned into the baculovirus pMel Bac B1 (Invitrogen) with C-terminal StrepII-9×His tag. The FLS2^{ECD}-StrepII-9His was produced by secreted expression in baculovirus-infected *Trichoplusia ni* High Five insect cells (94-002 F, Expression Systems) and harvested 72 h post-infection. The secreted protein was purified using Ni-NTA affinity chromatography (HisTrap™ Excel, Cytiva) with Ni buffer (20 mM HEPES pH 7.4,

Fig. 3 | Maya1/2 act as biased ligands to partially activate FLS2 signaling.
a Heatmap of differential interactomes induced by 10 mins treatments of *proFLS2:FLS2-GFP* seedlings with Maya1 (10 μM), Maya2 (10 μM), or flg22 (10 nM). The pixels corresponding to BAK1 are indicated on the right. The color scale shows log₂ fold changes in protein abundances compared to the DMSO control.
b Immunoblotting (IB) analyses of FLS2–BAK1 co-immunoprecipitations in Arabidopsis transgenic seedlings (*proFLS2:FLS2-GFP*) treated as described in (a). Antibodies were used for blotting as indicated. An independent biological replicate is in Supplementary Fig. 8a. **c** FRK1-induction measured by qPCR analysis of WT, *bak1-5*, and *fls2* leaves infiltrated with DMSO, Maya1 (1 μM), Maya2 (1 μM), or flg22 (10 nM) for 2 hours. The box plot represents the first and third quartiles, centered by the median, and whiskers include the min-max of the data points. **d** Seedling Growth inhibition (SGI) percentage values after 6 days of incubation with DMSO, Maya1 (5 μM), Maya2 (5 μM), and flg22 (1 nM). Genotypes are indicated at the bottom. SGI percentage indicates the relative reduction in growth compared between the average fresh weight of the mock treatment controls with that of the treated

samples, where 100% indicates no inhibition. The Box plot displays the first and third quartiles, split by the median. Whiskers extend to include the 10th and 90th percentile. **e** IB analysis of FLS2-BAK1 co-immunoprecipitation in *proFLS2:FLS2-GFP* seedlings untreated, treated with flg22, or co-treated with Maya1 or Maya2. Concentrations are indicated on top. This experiment was repeated twice with similar results. **f** Quantification of relative band intensity (Co-IP/IP) from (e) with two independent experiments using Fiji (mean ± SE; $n = 3$). **g** Similar to (e), but for PEPR1-BAK1 in proPEPR1:PEPR1-GFP/*pepr1/pepr2* seedlings untreated, treated with pep2 alone, or co-treated with Maya1 or Maya2. This experiment was repeated twice with similar results. **h** Same as (f), but from (g) with two independent experiments (mean ± SE; $n = 3$). For (**c**, **d**, **f**, **h**), statistical significance was assessed by linear mixed effect models; letters above the boxes (**a**–**e**) indicate the results of a one-way ANOVA followed by a Tukey's multiple comparison test (two-sided, adjusted for multiple comparisons, $p < 0.05$). Groups with the same letter are indistinguishable at >95% confidence. Numbers of biologically independent observations (n) are indicated below the letters. p-values are provided in the source data file.
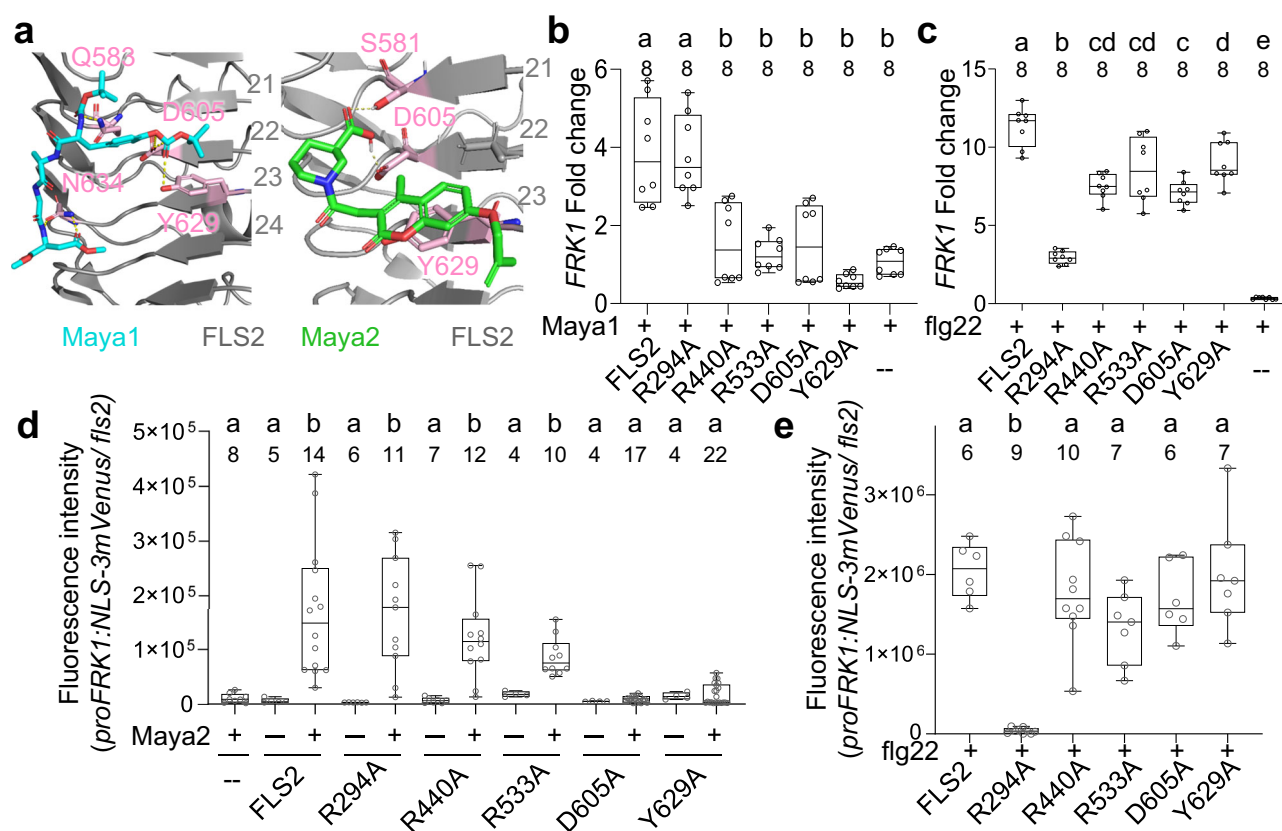
Fig. 4 | Analysis of potential Maya1/2 ligand binding sites from molecular docking simulation. **a** Representative snapshot showing the most probable binding pose of Maya1 (left) and Maya2 (right) on the FLS2^ECD structure (PDB ID: 4MNA). The residues predicted to be involved in interaction with Maya1 (turquoise) and Maya2 (green) are shown in pink. FLS2 LRR numbers are shown in gray on the right of the snapshots. **b** Maya1-mediated induction of *FRK1* in leaves of *fls2* plants left untransformed (--) or transformed with either wildtype or mutant *FLS2* variants (residue changes indicated at the bottom). The box plot represents the first and third quartiles, centered by the median, and whiskers include the min-max of the data points. **c** Same as (**b**) but with flg22. **d** Maya2-mediated induction of *FRK1* in

roots of *proFRK1:NLS-3mVenus/fls2* seedlings. The box plot represents the first and third quartiles, centered by the median, and whiskers include the min-max of the data points. **e** Same as (**d**) but with flg22. For (**b**–**e**), statistical significance was assessed by linear mixed effect models; letters above the boxes (**a**–**e**) indicate the results of a one-way ANOVA followed by a Tukey's multiple comparison test (two-sided, adjusted for multiple comparisons, $p < 0.05$). Groups with the same letter are indistinguishable at >95% confidence. Numbers of biologically independent observations (n) are indicated below the letters. p-values are provided in the source data file.

500 mM NaCl, and 20 mM Imidazole) and size-exclusion chromatography (SEC) (HiLoad® 16/600 Superdex® 200 pg, GE Healthcare) with either the SEC buffer A (10 mM Bis-Tris pH 6.0 and 100 mM NaCl) for chemical array binding assays or SEC buffer B (20 mM NaH₂PO₄/Na₂HPO₄ pH 7.5, 200 mM NaCl, and 5% glycerol) for microscale thermophoresis (MST) binding assays.

Chemical array interaction assay
The chemical arrays were prepared as reported in ref. 16. Solutions of the 22,618 compounds (2.5 mg/ml or 5 mM in DMSO) from the RIKEN Natural Products Depository (NPDepo) were arrayed onto 8 separate photoaffinity linker-coated glass slides using a chemical arrayer developed at RIKEN[15]. The slides were blocked with TBS-T buffer
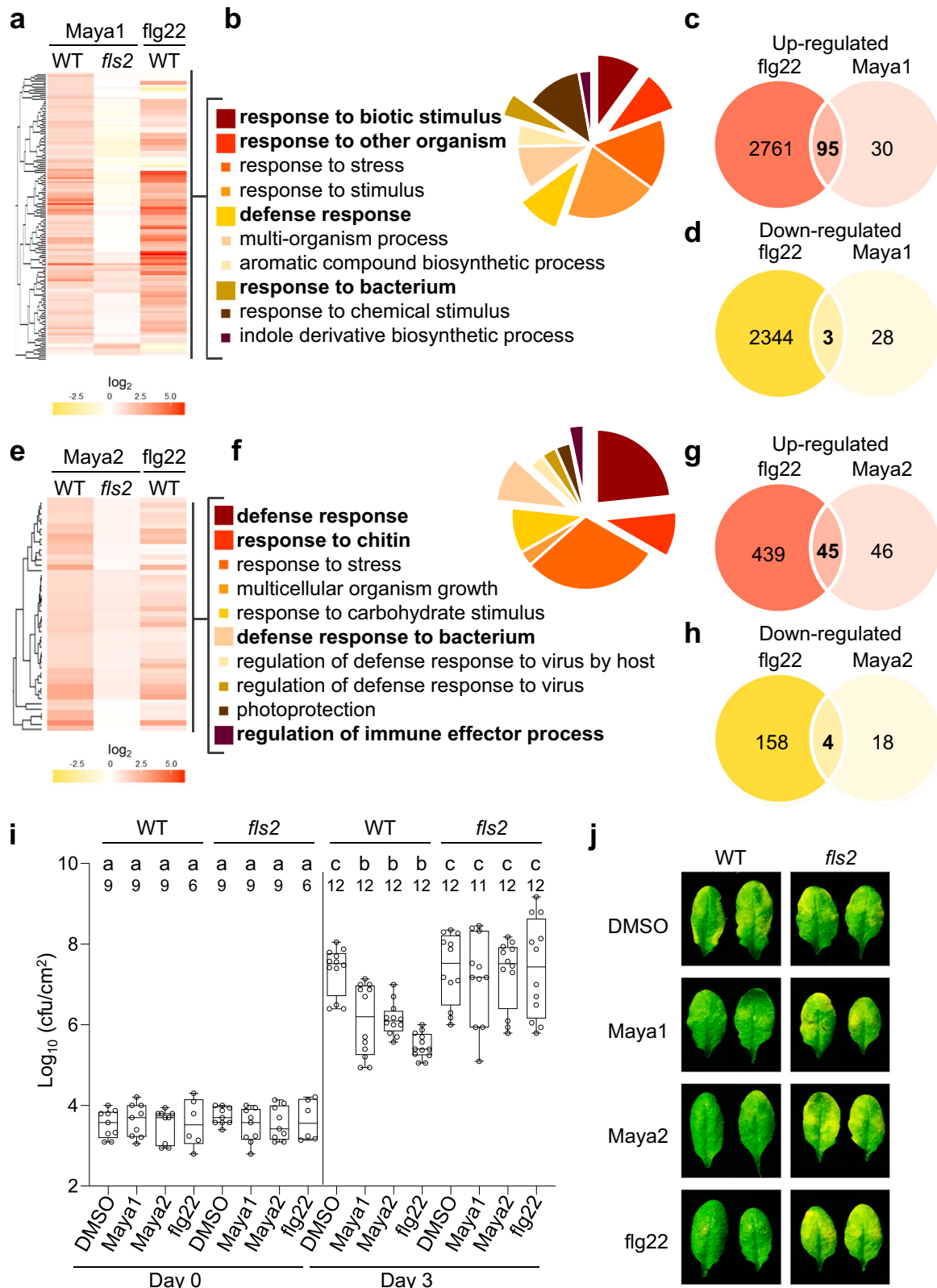
(10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.05% Tween-20) containing 1% skim milk for 1 h at room temperature. Next, the slides were probed with a purified FLS2^{ECD} in TBS-T buffer containing 1% skim milk at 30 °C for 1 h. After a TBS-T wash, the slides were incubated with an anti-His antibody (1:100, Cytiva) in TBS-T containing 1% skim milk at 30 °C for 1 h. This step was followed by another washing step prior to incubation with a secondary antibody for fluorescence

detection (A10523, goat anti-rabbit IgG, Cy5 conjugate, 20 µg/ml, ThermoFisher) at 30 C for 1 h. After a final wash, the slides were scanned with a GenePix microarray scanner (Molecular Devices) using the Cy5 channel (an excitation wavelength of 635 nm and an emission wavelength of 675 nm). The fluorescence signals were quantified using the GenePixPro 7.3 software (Molecular Devices) with local background subtraction. Data obtained from slides treated

Fig. 5 | FLS2 activation by Maya1/2 induces minimal gene expression programs that are sufficient to promote antibacterial immunity. a Heatmap of differentially expressed genes (DEGs) in wildtype (WT) and *fls2* leaves treated for 2 hours with Maya1 (1 μM) or flg22 (10 nM). The color scale bar at the bottom shows the $\log_2$ fold change values of Maya1- or flg22-responsive genes compared to DMSO controls. **b** Gene Ontology analysis showing enrichment in biological process categories upon treatments of wildtype leaves with Maya1. Pie charts show the fractions of Maya1-responsive gene functions enriched in each category. **c, d** Venn diagrams showing the overlap of flg22- (left) and *FLS2*-dependent Maya1-responsive (right) DEGs. Upregulated and downregulated DEGs are shown at the top and bottom, respectively. DEGs numbers are indicated in the circles. **e–h** same as (**a–d**) but for 1-h treatments of wildtype and *fls2* roots with Maya2 (10 μM) or flg22 (10 nM). **i** Growth of *Ps* DC3000 in wildtype and *fls2* leaves primed with DMSO, Maya1 (1 μM),

Maya2 (1 μM), or flg22 (10 nM) a day prior bacterial inoculation. The number of bacteria per area of leaf (cfu/cm²) is plotted on a $\log_{10}$ scale for day 0 and day 3. The box plot represents the first and third quartiles, centered by the median, and whiskers include the min-max of the data points. Numbers of biologically independent observations (*n*) are indicated below the letters. Statistical significance was assessed using linear mixed effect modeling; letters above the boxes (**a–c**) indicate the results of one-way ANOVA followed by a Tukey's multiple comparison test (two-sided, adjusted for multiple comparisons, $p < 0.05$). Groups with the same letter are indistinguishable at >95% confidence. Numbers of biologically independent observations (*n*) are indicated below the letters. p-values are provided in the source data file. **j** Representative pictures of leaves taken 5 days after inoculation with *Ps* DC3000. Pretreatments and genotypes are indicated at the bottom and the top, respectively.

with antibodies only were used as a reference. The data was analyzed as described previously[16].

Microscale thermophoresis (MST) binding assays

The purified FLS2^{ECD} was labeled with a fluorescent dye, Monolith His-Tag Labeling Kit RED-tris-NTA 2nd Generation (MO-L018, NanoTemper). The labeled FLS2^{ECD} was kept at constant 50 nM concentration, in MST buffer (50 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$ pH 7.5, 200 mM NaCl, 5% Glycerol, and 0.005% Tween-20), and varying SynChem concentrations were serially added. Approximately 4-6 μL of each sample was loaded in Monolith Premium Capillaries (MO-K025, NanoTemper). Measurements were performed at 25 °C in a Monolith NT.115 instrument setting with LED power of 80% and MST power of 20%, 40%, and 60%. To ensure binding reproducibility, measurements were repeated with independent protein preparations. The data was analyzed by plotting $\log_{10}$ values of SynChem concentrations against changes of normalized thermophoresis (ΔFnorm [‰]) using MO. Affinity Analysis software (NanoTemper). The data was subsequently analyzed with GraphPad Prism 10 software to obtain the K_d fit and K_d values.

Cheminformatics analysis

Structural analysis of the 84 FICs was performed using R package RDKit (ver 2021_09_3). The chemical structures were imported as SMILES and transformed into RDKit mol object. Fingerprinting was performed using FingerprintMol for similarity calculation. The dendrogram describing the relative similarity of chemicals was produced with scipy package and visualized with iTOL (ver 6).

Source of chemicals

The chemicals used in this study were sourced directly from RIKEN for the primary leaf and root screenings. For confirmatory analysis and to rule out that the phenotypes observed during our primary screenings were due to contaminations or impurities, the chemicals were acquired form different commercial suppliers. Maya1 (Vitas-M Lab ID STL525579) and Maya 2 (Vitas-M Lab ID STL535086) were sourced from Vitas-M Laboratory (https://vitasmlab.biz). The top10 compounds selected in the leaf and root triages were sourced from Vitas-M Laboratory or InterBioScreen ltd (https://www.ibscreen.com). The vendor information and product codes are listed in the Supplementary Data 1. Natural coumarins were sourced from Sigma (Coumarin, C4261; 4-Methylumbelliferone, M1381; Esculin hydrate, E8250).

Chemical treatments for qPCR analysis

The leaves of 4-weeks-old Arabidopsis plants grown on soil were infiltrated with DMSO, 1 μM of the chemicals sourced from RIKEN or from the commercial suppliers (SynChems), or 10 nM of flg22 peptide dissolved in DMSO. 2 h after treatment, total RNA was isolated from pools of leaves obtained from independent plants. For flg22 in leaves, total RNA from 4–5 leaves was isolated using the GeneMATRIX Universal RNA Purification Kit (EURx). For other

treatment in leaves and roots, total RNA was isolated from pools of 15 leaves using RNeasy Plant Mini Kit (QIAGEN). For both, one microgram of total RNA was used for cDNA synthesis using the High-Capacity cDNA Reverse Transcription kit (Applied Biosystems™). qPCR analyses were performed with FastStart Essential DNA Green Master (Roche Life Science) using a Roche LightCycler 96 instrument (Roche Life Science), following the manufacturer's instructions. *Actin* (*Actin 2/8*) mRNA level was used as an endogenous control for normalization. *FRK1* fold induction was determined by comparing it to the normalized mean values of DMSO treatments in each genotype. See Supplementary Table 1 for primer information for qPCR analyses.

Treatments for RNA-Seq analysis

For leaves, 4-week-old wildtype and *fls2* Arabidopsis plants grown on soil were treated for 2 h with DMSO, 1 μM Maya1 (SynChem) or 1 μM Maya2 (SynChem). As a positive control, DMSO and 10 nM of flg22 dissolved in DMSO were also applied to 4-week-old wildtype Arabidopsis for 2 hours. Total RNA was isolated from pools of 15 leaves isolated from independent plants using RNeasy Plant Mini Kit (QIAGEN). For roots, 5-day-old seedlings grown vertically on plates were isolated and treated by gently pouring 1/2 MS liquid medium containing 0.01% tween-20 with or without 10 μM Maya2 (SynChem), or 10 nM flg22. After 1 hour of incubation, the root sectors corresponding to the meristematic zone, elongation zone, and a small portion of the maturation zone were isolated. 30 independent roots were pooled to extract total RNA with the RNeasy Plant Mini Kit (QIAGEN). All samples were prepared with triple replication from independent experiments, except for flg22 in roots which had two replicates. Five hundred nanograms of total RNA from each sample was used for cDNA and NGS library preparations with QuantSeq 3′ mRNA-Seq Library Prep kit (FWD) (015, Lexogen) or NEBNext Poly(A) mRNA Magnetic Isolation Module with NEBNext Ultra RNA Library Prep kit for Illumina (New England Biolabs). The individual NGS libraries were pooled in equimolar ratios and sequenced on a Novaseq S4 in single end- or paired end-150 base mode. The quality check of raw reads was conducted by FastQC (ver 0.11.8) for subsequent read trimming by trimmomatic (ver 0.38). Trimmed reads were aligned to TAIR10 *A.thaliana* transcripts using HISAT2 (ver 2.1.0). The aligned reads were counted via feature-Counts (ver 2.0.1) and differential expression analysis was performed using the R package DESeq2 (ver 1.44.0). Expression levels were compared within the same genotypes. Cut-offs for differential expression were set as adjusted p-value < 0.1 and |$\log_2$(Fold Change)| > 1 to investigate the widest range of transcriptional changes. Differentially expressed genes were clustered using R package stats (ver 4.4.1) and displayed as heatmaps created by R package ggplot2 (ver 3.5.1). Gene Ontology (GO) enrichment analysis was done via BinGO (ver 3.0.3) in Cytoscape (ver 3.8.0.). Enrichment was determined by the Hypergeometric test corrected with the Benjamini & Hochberg False Discovery Rate correction. FDR smaller than 0.1 was considered significant.

Chemical treatment for fluorescent imaging using transcriptional reporter lines

5-day-old Arabidopsis seedlings grown vertically on plates were transferred to liquid medium (1/2 MS, 1% sucrose, and 0.01% Tween-20) supplemented with 1 μM of Maya1 or Maya 2, 1 nM flg20, or 1 nM flg22 (all dissolved in DMSO). After 48 h, the roots were observed after staining with or without Propidium iodide (PI) using a confocal laser scanning microscope (Zeiss LSM880). Pictures were taken with a 10 × OR 20 × objective. For a large area of interest, imaging was performed combining Z-stack with tile-scan (overlap 10%). 488 nm excitation and 500-530 nm detection windows were used for mVENUS, and 520 nm excitation and 590 nm detection windows were used for PI. Images were merged from maximal projections of all z-stacks for mVENUS and middle 4 z-stacks for PI. Identical root positions, laser intensity, scanning area, and interval/number of Z-stack slices were used for each measurement. Confocal images were processed and analyzed using Fiji software[40].

Protein extraction and co-immunoprecipitation

12-day-old *proFLS2:FLS2-GFP* and wild-type Arabidopsis seedlings grown vertically on plates were transferred to a 6-well plate and rinsed with sterile water. The water was exchanged with a treatment solution (0.01% tween-20 in sterile water) containing DMSO, 10 μM Maya1, 10 μM Maya2, or 10 nM flg22. For each treatment, 0.6 grams of seedlings were then vacuum infiltrated for 10 min in the treatment solution and quickly frozen in liquid N_2. For the competition Co-IP assays, the rinsed seedlings were first vacuum infiltrated in the treatment solution containing DMSO or increasing concentration of chemicals for 10 min and then subjected to 10 nM flg22. Samples were ground in liquid N_2 using mortar/pestle and then incubated in 2 mL of extraction buffer (50 mM HEPES pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, 2 mM EDTA, 10% Glycerol, 1 mM Na_2MoO_4, 20 mM NaF, 1 mM DTT, EDTA-free protease inhibitor cocktail (Roche), 1% IGEPAL) for 30 min at 4 °C. After a brief centrifugation to sediment the cell debris, the supernatant was collected, filtered with Miracloth (475855-1, Merck Millipore), and then diluted 1:1 in extraction buffer without IGEPAL. 15 μL of GFP-Trap agarose beads (gta-100, Chromotek, Lot 111102-01-02) equilibrated in the same buffer were added to the solution. The samples were incubated for 3 hours at 4 °C on a rotor and then washed 4 times with 1 mL of washing buffer (50 mM HEPES pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, 0.1% IGEPAL). For immunoblot analyses, the beads were resuspended in 2X Laemmli sample buffer, boiled for 2 min at 95 °C, and then centrifuged. The supernatant was subjected to 8% SDS-PAGE. Immunoblotting was performed using the anti-GFP-HRP (130-091-833, Miltenyi Biotec; 1:5000, Lot 5201006689) or primary anti-BAK1 antibodies (AS12 1858, Agrisera; 1:5000, Lot 1904) with secondary goat anti-rabbit-HRP antibodies (A6154, Sigma-Aldrich; 1:5000). Signals were detected with the ChemiDoc instrument (Bio-Rad). FLS2-myc protein levels were quantified using the anti-Myc-HRP antibody (130-092-113, Miltenyi Biotec; 1:5000) after functional complementation of *fls2* or *proFRK1:NLS-3mVenus/ fls2* mutant plants. For both lines, total proteins were extracted from a pool of 8 T2 seedlings derived from 2 independent T1 plants. Roots Co-IP experiment was performed similarly but used 0.25 grams of roots treated without vacuum infiltration.

Immunoprecipitation coupled with mass spectrometry (IP-MS)

12-day-old *proFLS2:FLS2-GFP* and wild-type Arabidopsis seedlings grown vertically on plates were transferred to a 6-well plate and rinsed with sterile water. The water was exchanged with a treatment solution (0.01% tween-20 in sterile water) containing DMSO, 10 μM Maya1, 10 μM Maya2, or 10 nM flg22. For each treatment, 0.6 grams of seedlings were then vacuum infiltrated for 10 min in the treatment solution and quickly frozen in liquid N_2. A total of 15 samples were prepared for IP-MS analysis, including treatments with Maya1, Maya2, and flg22 in *proFLS2:FLS2-GFP*, DMSO in *proFLS2:FLS2-GFP* for treatment control,

and flg22 in WT for IP control, each with three biological replicates. Samples were ground in liquid N_2 using mortar/pestle and then incubated in 2 mL of extraction buffer (50 mM HEPES pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, 2 mM EDTA, 10% Glycerol, 1 mM Na_2MoO_4, 20 mM NaF, 1 mM DTT, EDTA-free protease inhibitor cocktail (Roche), 1% IGEPAL) for 30 min at 4 °C. After a brief centrifugation to sediment the cell debris, the supernatant was collected, filtered with Miracloth (475855-1, Merck Millipore), and then diluted 1:1 in extraction buffer without IGEPAL. 15 μL of GFP-Trap agarose beads (gta-100, Chromotek, Lot 111102-01-02) equilibrated in the same buffer were added to the solution. The samples were incubated for 3 hours at 4 °C on a rotor and then washed 4 times with 1 mL of washing buffer (50 mM HEPES pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, 0.1% IGEPAL), and then two rounds of additional washes with 1 mL of detergent-free washing buffer were added. 5% of the beads were analyzed by silver staining as control. Immunoprecipitated GFP-Trap agarose beads were resuspended in 60 μL of 100 mM ammonium bicarbonate (ABC), supplemented with 600 ng of lysyl endopeptidase (Lys-C, Fujifilm Wako Pure Chemical Corporation), and incubated for 4 hours at 37 °C with shaking. After centrifugation, the supernatant was transferred to a fresh tube and reduced with 0.5 mM Tris 2-carboxyethyl phosphine hydrochloride (TCEP, Sigma) for 30 min at 60 °C and then alkylated in 4 mM methyl methanethiosulfonate (MMTS, Fluka) for 30 min at room temperature. Subsequently, the samples were digested with 600 ng trypsin (Trypsin Gold, Promega) at 37 °C overnight. The digests were acidified by addition of 1% trifluoroacetic acid (TFA, Pierce). 40% of each sample was used for analysis by Liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS). The system used was an UltiMate 3000 RSLC nano system (Thermo Fisher Scientific) coupled to a Q Exactive HF-X mass spectrometer (Thermo Fisher Scientific), equipped with an Easyspray nanospray source (Thermo Fisher Scientific). Peptides were loaded onto a trap column (Thermo Fisher Scientific, Pep-Map C18, 5 mm × 300 μm ID, 5 μm particles, 100 Å pore size) at a flow rate of 25 μL/min using 0.1% TFA as mobile phase. After 10 min, the trap column was switched in line with the analytical column (Thermo Fisher Scientific, PepMap C18, 500 mm × 75 μm ID, 2 μm, 100 Å). Peptides were eluted using a flow rate of 230 nL/min, starting with the mobile phases 98% A (0.1% formic acid in water) and 2% B (80% acetonitrile, 0.1% formic acid) and linearly increasing to 35% B over the next 180 min, followed by a gradient to 90% B in 5 min, staying there for 5 min and decreasing in 2 min back to the gradient 98% A and 2% B for equilibration at 30 °C. The mass spectra were obtained using data-dependent MS^2 (dd-MS^2) in positive ion mode. The acquisition parameters were as follows: spray voltage was 2.3 kV, the capillary temperature was 275 °C, S-Lens RF level was 40.0, and probe heater temperature was 350 °C. For full MS, the scan range was 380-1500 m/z, the resolution was 60,000, AGC target was 1e6, and the maximum injection time was 60 ms. For dd-MS^2, the topN was 10, the scan range was 200–2000 m/z, the resolution was 30,000, the AGC target was 1e5, maximum injection time was 105 ms, and the isolation window was 1 m/z. The normalized collision energy was set to 28 %, the maximum AGC target was 5e3, and the intensity threshold was 4.8e4. Charge states of 1 and greater than 7 were excluded. The dynamic exclusion was set to 60 s.

Proteomic data analysis

Raw MS data from IP experiments were loaded into Proteome Discoverer (PD, version 2.5.0.400, Thermo Scientific). All MS/MS spectra were searched using MSAmanda v2.0.0.16129[41]. Peptide and protein identification was performed in two steps. An initial search was performed against the database TAIR10_pep_20101214.fasta (32,785 sequences; 14,482,855 residues) with common contaminants appended. Trypsin was specified as proteolytic enzyme without pro-line restriction and allowing for up to 2 missed cleavages. Mass tolerances were set to ±5 ppm at the precursor and ±10 ppm at the fragment

mass level. The minimum peptide length was set to 7. The result was filtered to 1 % FDR on protein level and a sub-database of proteins identified in this search was generated for further processing. Additionally, a minimum MS Amanda score of 150 was applied on peptide spectrum match (PSM) level and proteins had to be identified by a minimum of 2 PSMs. Peptide areas were quantified using the apQuant[42]. Proteins were quantified by summing unique and razor peptides and applying intensity-based absolute quantification, iBAQ[43]. Protein-abundances-normalization was done using sum normalization. Statistical significance of differentially interacting proteins was determined using limma[44]. Cut-offs for differential interaction were set as adjusted $p < 0.1$ and $|\log_2(\text{Fold Change})| > 1$. Differentially interacting proteins were clustered using R package stats (ver 4.4.1) and displayed as heatmaps created by R package ggplot2 (ver 3.5.1).

Seedling growth inhibition assays

5-day-old seedlings grown vertically on plates were transferred to 48-well-plates (Griner Bio one). The wells were filled with 1 mL of ½ MS liquid media containing 0.01% Tween-20 and incubated for 2 days. The same volume of sterile water (Mock), DMSO, 5 μM Maya1, 5 μM Maya2 or 10 nM flg22 was then added into the wells and the plates were further incubated for 6 days. The fresh weight of each seedling was measured using a high-precision scale. For each treatment and genotype, the SGI percentages were calculated by comparing the median fresh weight of the mock controls with that of the treated samples.

BIK1 phosphorylation assay

12-week-old *proBIK1:BIK1-HA* seedlings vertically grown on plates were transferred to 6-well plate containing sterile MonoQ water and rinsed. The water was replaced with 2 mL of 0.01% Tween-20 treatment buffer containing either 10 nM flg22 or 1–100 μM of Maya1 or Maya2. The seedlings were vacuum infiltrated for 10 min and total proteins extracted for immunoblotting assays. Total proteins were extracted in the protein extraction buffer (50 mM Tris-Cl pH 7.5, 150 mM NaCl, 2 mM Na$_2$MoO$_4$, 2 mM Na$_3$VO$_4$, 5 mM NaF, 2 mM phenylmethylsulfonyl fluoride, 1 mM DTT, 10% Glycerol, 1% IGEPAL, EDTA-free protease inhibitor cocktail (Roche)). Immunoblotting was performed using anti-HA-HRP antibodies (130-091-972, Miltenyi Biotec; 1:5000).

Reactive oxygen intermediate (ROI) burst assays

Leaf discs (diameter 4 mm) were punched from 6-week-old healthy wildtype plants using a cork borer. The discs were placed in a 96-well luminescence assay plate (Griner Bio one) containing 200 μL sterile MonoQ water, with the adaxial side up, and then vacuum infiltrated for 10 min. The plate was then incubated in the dark overnight at 21 °C. The next day the water was replaced with 100 μL of treatment solution containing either 100 nM flg22, 1 μM Maya1 or Maya2, L-012 (Wako Chemicals) and Horseradish peroxidase (Thermo Fisher Scientific). The relative luminescence was measured every minute for 50 min using a BiTec Synergy 4 microplate reader and then summed up to obtain total photon counts.

MAPK assays

5-day-old Arabidopsis seedlings vertically grown on plates were transferred to a 6-well plate containing 2 mL of 1/2 MS liquid media and incubated further for 1 week. For treatment, the previous media was replaced with a new 0.01% tween-20 buffer supplemented with either 10 nM flg22, or 1–100 μM of Maya1 or Maya2. The seedlings were incubated for 15 min and then total proteins were extracted for immunoblotting assays. Total proteins were extracted in the protein extraction buffer (50 mM Tris-Cl pH 7.5, 150 mM NaCl, 2 mM Na2MoO4, 2 mM Na$_3$VO$_4$, 5 mM NaF, 2 mM phenylmethylsulfonyl fluoride, 1 mM DTT, 10% Glycerol, 1% IGEPAL, EDTA-free protease inhibitor cocktail (Roche)). Immunodetection was performed using

the anti-p44/p42 MAPK antibodies (9102, Cell Signaling; 1:5000, Ref 07/2020) with anti-Rabbit-HRP antibodies (A6154, Sigma-Aldrich; 1:5000). Signals were detected with the ChemiDoc instrument (Bio-Rad).

Protein-ligand docking simulations

Schrödinger suites (ver 2021-4) was used as follows: the crystal structure of FLS2^{ECD} was obtained from the PDB database with accession code 4MNA and processed using the default settings on the Protein Preparation pipeline of the software. 5 different receptor grids (<20 Å) were generated to encompass every possible binding pocket on the FLS2 structure. SMILE strings of chemicals were processed into 3D structure for docking by LigPrep. The docking was conducted using Glide XP mode, with flexible ligand setting. 10 poses for each chemical isoforms were extracted based on binding energy prediction computed with MM-GBSA. Subsequently, the lateral chains of FLS2 amino acids participating in interaction with the chemicals within a 12 Å window were extracted. All steps were conducted under simulatory pH=5.3. Visualization of the estimated model was performed using PyMOL (ver 2.5.5).

Bacterial infection assays

Arabidopsis infection assays with *Pseudomonas syringae pv. tomato* DC3000 strain and *ΔfliC* were performed using the direct infiltration method[10]. 4-weeks-old Arabidopsis plants grown on soil in short-day conditions were infiltrated with DMSO, 1 μM Maya1 (SynChem), 1 μM Maya2 (SynChem), or 10 nM flg22 using a 1 mL syringe, and incubated for 1 day. A bacterial inoculum of OD$_{600}$ = 0.0002 diluted in 10 mM MgCl$_2$ was then infiltrated in the pretreated leaves. 3 days after inoculation, 8 infiltrated leaf discs coming from 4 independent plants were bored, ground, and homogenized in 10 mM MgCl$_2$. The bacterial titer was assessed by serial dilution plating. Leaf pictures were acquired five days after inoculation.

Data visualization and statistical analysis

GraphPad PRISM 10 and R programming environment (version R & Bioconductor 4.4.1) were used for data analysis and heatmap visualizations. In all assays performed, data were combined from at least two independent biological experiments. The box plot represents the first and third quartiles, centered by the median, and whiskers include the min-max of the data points unless specified otherwise. Outliers were removed using the ROUT test and assessed using GraphPad PRISM (Q = 0.1%). DMSO treatment was used as a control sample and comparator in all statistical tests performed. For the linear mixed effect modeling, the lme4 package was used and DMSO treatment in WT was used as a fixed effect and each experiment was performed as a random effect using the formula Value ~ Treatment + (1|Trial). In case of fluorescence values were converted to log$_{10}$ to fit in a linear model. Subsequently, statistical significance was calculated using One-way ANOVA followed by Tukey's multiple comparisons. Statistical significance for data in fig. S8 was determined without linear mixed effect models. Small letters were obtained from all pair-wise comparisons using multicomp package ($p < 0.05$).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The RNA-seq datasets generated and analysed in the current study have been deposited in the Sequence Read Archive at NCBI under bioproject accession code PRJNA939365. The proteomics data generated in this study have been deposited in the jPOST under accession code PXD056313. The molecular docking structure generated in this study have been deposited in the Figshare [https://doi.org/10.6084/

m9.figshare.27157110]. The authors declare that the data supporting the findings of this study are available within the paper and its supplementary information files. Source data are provided with this paper.

Code availability

All source codes that we used in this study are also available in *Zenodo* [https://doi.org/10.5281/zenodo.13844042].

References

1. Lee, D. H., Lee, H. S. & Belkhadir, Y. Coding of plant immune signals by surface receptors. *Curr. Opin. Plant Biol.* **62**, 102044 (2021).
2. Kato, H. et al. Recognition of pathogen-derived sphingolipids in Arabidopsis. *Science* **376**, 857–860 (2022).
3. Laohavisit, A. et al. Quinone perception in plants via leucine-rich-repeat receptor-like kinases. *Nature* **587**, 92–97 (2020).
4. Kutschera, A. et al. Bacterial medium-chain 3-hydroxy fatty acid metabolites trigger immunity in Arabidopsis plants. *Science* **364**, 178–181 (2019).
5. Luu, D. D. et al. Biosynthesis and secretion of the microbial sulfated peptide RaxX and binding to the rice XA21 immune receptor. *Proc. Natl. Acad. Sci. USA* **116**, 8525–8534 (2019).
6. Martín-Dacal, M. et al. Arabidopsis immune responses triggered by cellulose and mixed-linked glucan oligosaccharides require a group of Leucine-Rich Repeat-Malectin Receptor Kinases. *Plant J. Cell Mol. Biol.* **113**, 833–850 (2023).
7. Roman, A. O. et al. HSL1 and BAM1/2 impact epidermal cell development by sensing distinct signaling peptides. *Nat. Commun.* **13**, 876 (2022).
8. Choi, J. et al. Identification of a plant receptor for extracellular ATP. *Science* **343**, 290–294 (2014).
9. Wu, F. et al. Hydrogen peroxide sensor HPCA1 is an LRR receptor kinase in Arabidopsis. *Nature* **578**, 577–581 (2020).
10. Zipfel, C. et al. Bacterial disease resistance in Arabidopsis through flagellin perception. *Nature* **428**, 764–767 (2004).
11. Buscaill, P. et al. Glycosidase and glycan polymorphism control hydrolytic release of immunogenic flagellin peptides. *Science* **364**, https://doi.org/10.1126/science.aav0748 (2019).
12. DeFalco, T. A. & Zipfel, C. Molecular mechanisms of early plant pattern-triggered immune signaling. *Mol. Cell* **81**, 4346 (2021).
13. Sun, Y. et al. Structural basis for flg22-induced activation of the Arabidopsis FLS2-BAK1 immune complex. *Science* **342**, 624–628 (2013).
14. Bjornson, M., Pimprikar, P., Nürnberger, T. & Zipfel, C. The transcriptional landscape of Arabidopsis thaliana pattern-triggered immunity. *Nat. Plants* **7**, 579–586 (2021).
15. Osada, H. Introduction of new tools for chemical biology research on microbial metabolites. *Biosci. Biotechnol. Biochem* **74**, 1135–1140 (2010).
16. Kondoh, Y., Honda, K. & Osada, H. Construction and application of a photo-cross-linked chemical array. *Methods Mol. Biol. (Clifton N. J.)* **1263**, 29–41 (2015).
17. Zhou, F. et al. Co-incidence of damage and microbial patterns controls localized immune responses in roots. *Cell* **180**, 440–453 e418 (2020).
18. Lee, H. G. & Seo, P. J. Transcriptional activation of SUGAR TRANSPORT PROTEIN 13 mediates biotic and abiotic stress signaling. *Plant Signal. Behav.* **16**, 1920759 (2021).
19. Yamada, K., Saijo, Y., Nakagami, H. & Takano, Y. Regulation of sugar transporter activity for antibacterial defense in Arabidopsis. *Science* **354**, 1427–1430 (2016).
20. Parys, K. et al. Signatures of antagonistic pleiotropy in a bacterial flagellin epitope. *Cell Host Microbe* **29**, 620–634.e629 (2021).
21. Colaianni, N. R. et al. A complex immune response to flagellin epitope variation in commensal communities. *Cell Host Microbe* **29**, 635–649.e639 (2021).
22. Ortiz-Morea, F. A. et al. Danger-associated peptide signaling in Arabidopsis requires clathrin. *Proc. Natl. Acad. Sci. USA* **113**, 11028–11033 (2016).
23. Tang, J. et al. Structural basis for recognition of an endogenous peptide by the plant receptor kinase PEPR1. *Cell Res.* **25**, 110–120 (2015).
24. Rajniak, J. et al. Biosynthesis of redox-active metabolites in response to iron deficiency in plants. *Nat. Chem. Biol.* **14**, 442–450 (2018).
25. Harbort, C. J. et al. Root-secreted coumarins and the microbiota interact to improve iron nutrition in Arabidopsis. *Cell Host Microbe* **28**, 825–837.e826 (2020).
26. Stringlis, I. A. et al. MYB72-dependent coumarin exudation shapes root microbiome assembly to promote plant health. *Proc. Natl. Acad. Sci. USA* **115**, E5213–E5222 (2018).
27. Voges, M. J. E. E. E., Bai, Y., Schulze-Lefert, P. & Sattely, E. S. Plant-derived coumarins shape the composition of an Arabidopsis synthetic root microbiome. *Proc. Natl. Acad. Sci. USA* **116**, 12558–12565 (2019).
28. Stringlis, I. A., De Jonge, R. & Pieterse, C. M. J. The age of coumarins in plant–microbe interactions. *Plant Cell Physiol.* **60**, 1405–1419 (2019).
29. Smakowska-Luzan, E. et al. An extracellular network of Arabidopsis leucine-rich repeat receptor kinases. *Nature* **553**, 342–346 (2018).
30. Uchida, N. et al. Chemical hijacking of auxin signaling with an engineered auxin-TIR1 pair. *Nat. Chem. Biol.* **14**, 299–305 (2018).
31. Dejonghe, W. et al. Disruption of endocytosis through chemical inhibition of clathrin heavy chain function. *Nat. Chem. Biol.* **15**, 641–649 (2019).
32. Ma, Q., Chang, M., Drakakaki, G. & Russinova, E. Selective chemical probes can untangle the complexity of the plant cell endomembrane system. *Curr. Opin. Plant Biol.* **68**, 102223 (2022).
33. Kinoshita, T., Toh, S. & Torii, K. U. Chemical control of stomatal function and development. *Curr. Opin. Plant Biol.* **60**, 102010 (2021).
34. Vaidya, A. S. et al. Dynamic control of plant water use using designed ABA receptor agonists. *Science* **366**, https://doi.org/10.1126/science.aaw8848 (2019).
35. Park, S. Y. et al. Agrochemical control of plant water use using engineered abscisic acid receptors. *Nature* **520**, 545–548 (2015).
36. Schwessinger, B. et al. Phosphorylation-dependent differential regulation of plant growth, cell death, and innate immunity by the regulatory receptor-like kinase BAK1. *PLoS Genet* **7**, e1002046 (2011).
37. Chinchilla, D. et al. A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. *Nature* **448**, 497–500 (2007).
38. Kadota, Y. et al. Direct regulation of the NADPH oxidase RBOHD by the PRR-associated kinase BIK1 during plant immunity. *Mol. Cell* **54**, 43–55 (2014).
39. Clough, S. J. & Bent, A. F. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J.* **16**, 735–743 (1998).
40. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
41. Dorfer, V. et al. MS Amanda, a universal identification algorithm optimized for high accuracy tandem mass spectra. *J. Proteome Res.* **13**, 3679–3684 (2014).
42. Doblmann, J. et al. apQuant: accurate label-free quantification by quality filtering. *J. Proteome Res.* **18**, 535–541 (2019).
43. Schwanhausser, B. et al. Global quantification of mammalian gene expression control. *Nature* **473**, 337–342 (2011).

44. Smyth, G. K. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**, https://doi.org/10.2202/1544-6115.1027 (2004).

Author contributions

Project Administration, Y.B.; Conceptualization, D.H.L., H.O., K.S. and Y.B.; Investigation, D.H.L, H.S.L., M.S.C., K.P., Y.K., K.H., J.M.L., N.E., G.H., and B.E.; Supervision, D.H.L., H.S.L., H.O., K.S., and Y.B.; Methodology, D.H.L., H.S.L., M.S.C., K.P., Y.K., K.H., B.E., H.O., K.S., and Y.B.; Resources, H.O., K.S. and Y.B.; Software, D.H.L., M.S.C., and Y.B.; Formal Analysis, D.H.L., and M.S.C.; Validation, D.H.L., and Y.B.; Writing-Original Draft, D.H.L., and Y.B.; Writing-Review & Editing, all other authors commented and agreed on the manuscript; Visualization, D.H.L., M.S.C., and Y.B.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-54271-5.

Correspondence and requests for materials should be addressed to Du-Hwa Lee or Youssef Belkhadir.

Peer review information *Nature Communications* thanks Zhi-Yong Wang, and the other, anonymous, reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.