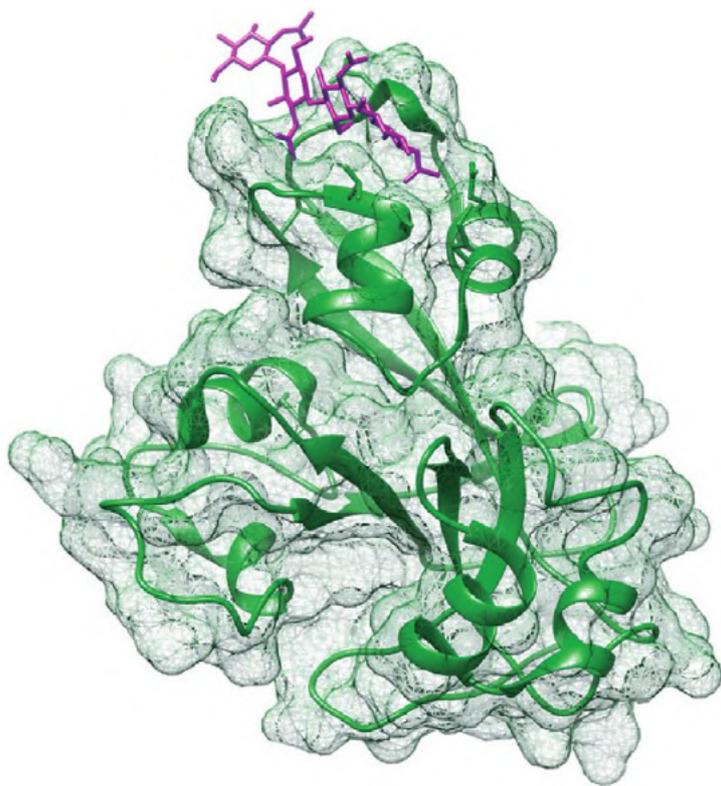
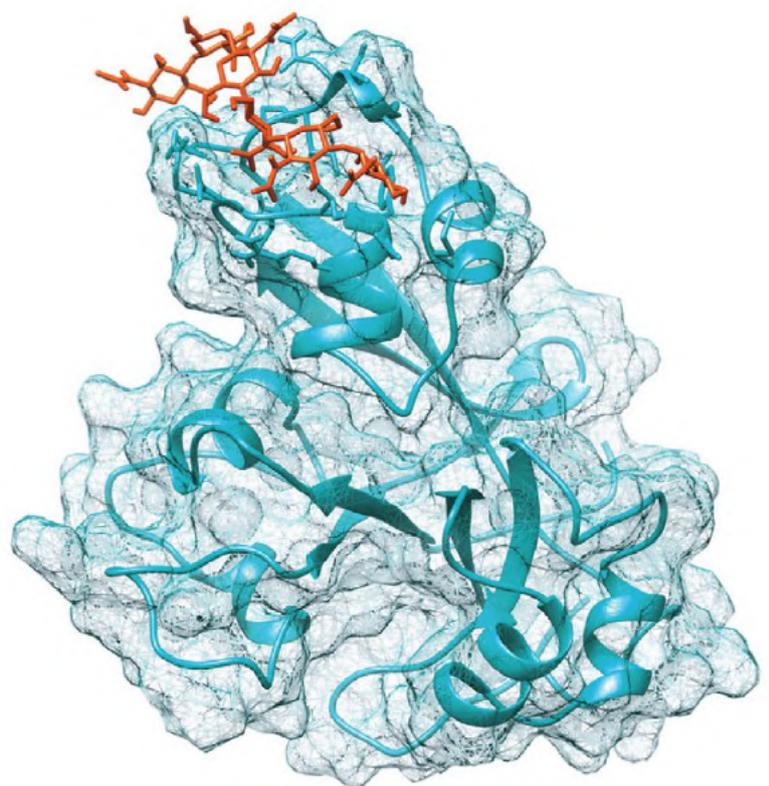


# *the plant journal*



CERK1 + chitohexahose

**Crystal structure**



CERK1 + chitohexahose

**Computational structure**

VOLUME 105 | NUMBER 6 | MARCH 2021  
<http://www.theplantjournal.com> | ISSN 1365-313X

**WILEY**  
Blackwell

## TECHNICAL ADVANCE

# Computational prediction method to decipher receptor–glycoligand interactions in plant immunity

Irene del Hierro<sup>1,2</sup>, Hugo Mélida<sup>1,†</sup> , Caroline Broyart<sup>3</sup>, Julia Santiago<sup>3</sup>  and Antonio Molina<sup>1,2,\*</sup> <sup>1</sup>Centro de Biotecnología y Genómica de Plantas (CBGP), Universidad Politécnica de Madrid (UPM), Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Campus de Montegancedo-UPM, 28223, Pozuelo de Alarcón, Madrid, Spain,<sup>2</sup>Departamento de Biotecnología-Biología Vegetal, Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas, Universidad Politécnica de Madrid (UPM), 28040, Madrid, Spain and<sup>3</sup>Département de Biologie Moléculaire Végétale (DBMV), University of Lausanne (UNIL), Biophore Building, UNIL Sorge, CH-1015, Lausanne, Switzerland

Received 3 July 2020; revised 30 November 2020; accepted 8 December 2020; published online 14 December 2020.

\*For correspondence (e-mail antonio.molina@upm.es).

<sup>†</sup>Present address: Área de Fisiología Vegetal, Departamento de Ingeniería y Ciencias Agrarias, Universidad de León, 24071, León, Spain

## SUMMARY

Microbial and plant cell walls have been selected by the plant immune system as a source of microbe- and plant damage-associated molecular patterns (MAMPs/DAMPs) that are perceived by extracellular ectodomains (ECDs) of plant pattern recognition receptors (PRRs) triggering immune responses. From the vast number of ligands that PRRs can bind, those composed of carbohydrate moieties are poorly studied, and only a handful of PRR/glycan pairs have been determined. Here we present a computational screening method, based on the first step of molecular dynamics simulation, that is able to predict putative ECD-PRR/glycan interactions. This method has been developed and optimized with Arabidopsis LysM-PRR members CERK1 and LYK4, which are involved in the perception of fungal MAMPs, chitohexaose (1,4- $\beta$ -D-(GlcNAc)<sub>6</sub>) and laminarihexaose (1,3- $\beta$ -D-(Glc)<sub>6</sub>). Our *in silico* results predicted CERK1 interactions with 1,4- $\beta$ -D-(GlcNAc)<sub>6</sub> whilst discarding its direct binding by LYK4. In contrast, no direct interaction between CERK1/laminarihexaose was predicted by the model despite CERK1 being required for laminarihexaose immune activation, suggesting that CERK1 may act as a co-receptor for its recognition. These *in silico* results were validated by isothermal titration calorimetry binding assays between these MAMPs and recombinant ECDs-LysM-PRRs. The robustness of the developed computational screening method was further validated by predicting that CERK1 does not bind the DAMP 1,4- $\beta$ -D-(Glc)<sub>6</sub> (cellohexaose), and then probing that immune responses triggered by this DAMP were not impaired in the Arabidopsis *cerk1* mutant. The computational predictive glycan/PRR binding method developed here might accelerate the discovery of protein–glycan interactions and provide information on immune responses activated by glycoligands.

**Keywords:** *Arabidopsis thaliana*, glycan, immunity, isothermal titration calorimetry, LysM domain, molecular dynamics, pattern recognition receptor, technical advance.

## INTRODUCTION

Plant terrestrial colonization and diversification was associated to the evolution of a set of plant protein receptors, called pattern recognition receptors (PRRs), that confer to plants the capacities to perceive environmental and developmental cues. These PRR-based surveillance systems have allowed plants to modulate their adaptive and physiological responses to environmental conditions and to

activate defense responses against the diversity of pathogens that can colonize them (Li *et al.*, 2016; Tang *et al.*, 2017). Plant PRRs comprise several groups of extracellular, membrane-anchored proteins that greatly exceed in number their PRR counterparts in animals (Zipfel, 2014). In addition, the plant immune system comprises a set of intracellular protein receptors, mainly resistance (R) proteins, encoded by *R* genes, that perceive pathogen

effectors and activate effector-triggered immunity (ETI), which dampens pathogen colonization (Baltrus *et al.*, 2011; Li *et al.*, 2016; Sonah *et al.*, 2016). Extracellular and membrane-anchored PRRs include three main classes of proteins: (i) receptor-like kinases (RLKs), which contain an extracellular ectodomain (ECD), a transmembrane region (TM) and an intracellular serine/threonine kinase domain (KD); (ii) receptor-like proteins (RLPs), with ECD and TM but lacking the KD; and (iii) receptor proteins (RPs), which either contain an ECD that can be attached to the plasma membrane by a glycosylphosphatidylinositol (GPI)-anchor (RPg) or are extracellular proteins (RPes) not attached to the membrane (Bellande *et al.*, 2017). In *Arabidopsis thaliana*, different subclasses of extracellular RLKs/RLPs/RPs can be considered based on their predicted ECD structures or sequence similarities: lectins (G, L and C-lectins), leucine-rich repeats (LRRs), CRinkly-Like (CR4L), pathogenic-related thaumatin-like (ThaumatinL/PR5), proline-rich extensin-like receptor kinase (PERK), wall-associated kinases (WAKs), malectins, *Catharanthus roseus* receptor-like kinase 1-like (*CrRLK1L*), lysin motif (LysM) and cysteine-rich kinases (CRK/DUF26, also known as stress-antifungal/salt-response receptors). These RLKs/RLPs/RPs comprise more than 600 members in *Arabidopsis*, representing approximately 2–3% of *Arabidopsis* genes (Shiu and Bleeker, 2003; Gish and Clark, 2011; Franck *et al.*, 2018).

RLKs/RLPs/RPs are involved in the perception of developmental cues (e.g., peptide ligands or hormones) (Santiago *et al.*, 2013; Santiago *et al.*, 2016; Stegmann *et al.*, 2017; Tang *et al.*, 2017) and in high-affinity recognition of different conserved structures from microbes called microbe-associated molecular patterns (MAMPs) (Boutrot and Zipfel, 2017). Additionally, these PRRs can recognize other types of patterns, like plant self-molecules that are released or synthesized upon pathogen infection or tissue damage, which are known as damage-associated molecular patterns (DAMPs) (Li *et al.*, 2020). Upon MAMP/DAMP recognition by specific PRRs, formation of protein complexes with other PRRs (co-receptors) takes place and pattern-triggered immunity (PTI) responses are activated (Greeff *et al.*, 2012; Dangl *et al.*, 2013).

Plant RLKs are very similar to animal receptor tyrosine kinases (RTKs), but RLKs have generally serine/threonine kinase specificity instead of the tyrosine specificity of animal RTKs (Shiu *et al.*, 2004; Greeff *et al.*, 2012). The RLK class has an ancient origin after the fungus–metazoan split and is not present in the fungal kingdom. Interestingly, domain fusion resulted in the creation of novel receptors, leading to a high diversity and the appearance of different protein subclasses, and gene duplications contributed to gene expansion, explaining the high number of RLKs in plants (Lehti-Shiu *et al.*, 2009; Li *et al.*, 2016). Most of these RLKs are found in large genomic clusters, supporting the

hypothesis of rapid evolution by duplication and gene shuffling under diversifying selection (Shiu and Bleeker, 2003; Fritz-Laylin *et al.*, 2005; Lehti-Shiu *et al.*, 2009; Gish and Clark, 2011; Li *et al.*, 2016). RLPs and RPs have been shown to be involved in the control of defensive and developmental processes, playing roles in RLK signaling as RLK counterparts in co-receptor mechanisms (Jeong *et al.*, 1999; Fritz-Laylin *et al.*, 2005; Gish and Clark, 2011; Bellande *et al.*, 2017; Tang *et al.*, 2017).

There is an extensive diversity of molecules that can be bound by PRRs to trigger plant immune responses, but most of the MAMPs and DAMPs described so far are peptides (Boutrot and Zipfel, 2017; Li *et al.*, 2020). In contrast, the number of glycan structures characterized as MAMPs/DAMPs is very low and accordingly the number of PRR/glycan pairs identified is very restricted (Bacete *et al.*, 2018; Li *et al.*, 2020). However, PRR/glycan interaction is a field in expansion as glycans are cell surface components of major plant pathogens like fungi, oomycete and bacteria (MAMPs) and they are also present in the plant cell walls and can be released as oligosaccharides (DAMPs) (Bacete *et al.*, 2018; Wanke *et al.*, 2020a). On the other hand, one of the reasons explaining the slow progress of this field is the diversity (thus complexity) of glycan ligands in terms of composition: (i) over 20 different monosaccharides can form the backbone and/or ramification building blocks of glycans through a high diversity of glycosidic linkages; (ii) glycans can differ in the degree of polymerization (DP); and (iii) monosaccharides can have different biochemical decorations (e.g., acetylation and methylation) and chemical modifications (e.g., reduction/oxidation) (Carpita and McCann, 2000; Latgé and Calderone, 2006; Mélida *et al.*, 2013; Srivastava *et al.*, 2017).

The interaction of PRRs with carbohydrate-based ligands is well studied in mammals, where several receptor/glycan complexes have been determined such as Dectin-1 (C-lectin ECD)/ $\beta$ -glucans or Galectin-3/ $\beta$ -galactosides (Brown *et al.*, 2003; Chen *et al.*, 2017; Díaz-Alvarez and Ortega, 2017). In contrast, the characterization of PRR–glycoligand interactions in plants has been mainly restricted to PRRs of the LysM subclass that harbor ECDs with lysin motifs, which are promiscuous motifs involved in the recognition of several ligands like chitin, peptidoglycans,  $\beta$ -1,3-glucans and lipopolysaccharides (Miya *et al.*, 2007; Willmann *et al.*, 2011; Desaki *et al.*, 2018; Mélida *et al.*, 2018). In the plasma membrane, CERK1 (LYK1) and LYK5 together with LYK4 are key components in plant immunity acting as co-receptors in the recognition of chitin, a polymer of 1,4- $\beta$ -D-GlcNAc (Liu *et al.*, 2012; Cao *et al.*, 2014). However, in the plasmodesmata region, LYM2 forms a complex with LYK4 upon chitin perception by a CERK1-independent mechanism (Faulkner *et al.*, 2013; Cheval *et al.*, 2020). In rice (*Oryza sativa*), the molecular mechanism of chitin recognition by the OsCEBiP receptor, a LysM-PRR, has also been

described (Liu *et al.*, 2016), and it has been shown that OsCERK1 cooperates with OsCEBiP module chitin-mediated signaling (Shimizu *et al.*, 2010). Interestingly, LysM domains have also been described in some fungal effector proteins, like CfAvr4F from *Cladosporium fulvum* and Mg1LysM from *Zymoseptoria tritici*, and their crystal structures have been obtained (Hurlburt *et al.*, 2018; Sánchez-Vallet *et al.*, 2020). These fungal proteins bind chitin as a mechanism of fungal virulence aiming to avoid fungal perception by plant LysM-PRRs.

Recent work has demonstrated that LysM-PRRs are involved in the perception by plants of 1,3- $\beta$ -glucans isolated from fungal cell walls (Mélida *et al.*, 2018; Wanke *et al.*, 2020b). Specifically, the hexasaccharide 1,3- $\beta$ -D-(Glc)<sub>6</sub> (laminarihexaose) is an immune-active structure whose recognition in Arabidopsis is CERK1-dependent. Moreover, molecular docking calculations suggested some interactions between 1,3- $\beta$ -D-(Glc)<sub>6</sub> and CERK1-ECD, though these interactions were not further validated by full molecular dynamics simulations or *in vitro* binding assays (Mélida *et al.*, 2018). CERK1 was also shown to be involved in the perception of bacterial peptidoglycans, with a contribution of LysM members LYM1 and LYM3 (Gust *et al.*, 2007; Willmann *et al.*, 2011). The bases of these additional roles of LysM-PRRs in the perception of these glycans structures are unknown since crystal structures of these ligand/LysM complexes have not been obtained.

Plant cell walls are mainly composed of cellulose, different types of hemicelluloses and pectins (Carpita and McCann, 2000). Upon pathogen infection, some plant cell wall-derived compounds can be released acting as DAMPs recognized by PRRs, activating DAMP-triggered immunity cascades. For example, cellulose-derived oligomers (1,4- $\beta$ -glucans) have been described as a novel group of plant DAMPs which trigger signaling cascades that share many similarities with the responses activated by the well-characterized plant DAMPs oligogalacturonides (OGs), derived from homogalacturonan pectins (Aziz *et al.*, 2007; de Azevedo Souza *et al.*, 2017; Johnson *et al.*, 2018; Locci *et al.*, 2019). No PRR receptor candidates have been proposed for 1,4- $\beta$ -glucans oligomers, whereas several PRRs (WAKs, THESEUS and FERONIA) have been suggested to bind pectins, though crystal structures of these ligand/PRR complexes have not been obtained (Hématy *et al.*, 2007; Boisson-Dernier *et al.*, 2011; Kohorn and Kohorn, 2012; Duan *et al.*, 2020). There are many open questions regarding glycan-triggered immunity that could be solved using biochemical and genetic approaches. However, these approaches are time consuming and face the redundant functions of PRR families, as illustrated by the lack of PTI-defective phenotypes of *lyk4* and *lyk5* Arabidopsis single mutants upon chitin and 1,3- $\beta$ -D-(Glc)<sub>6</sub> treatment (Cao *et al.*, 2014; Mélida *et al.*, 2018).

*In silico* prediction tools (computational modeling) could help in preliminary screening stages to select potential

PRRs for further characterization (Das *et al.*, 2018; Fratev *et al.*, 2018). However, structural conformations of glycans are not easy to determine due to their intrinsic mobility and the scarce capability of some structural techniques, like X-ray diffraction, to solve their conformations (Fadda and Woods, 2010; Gimeno *et al.*, 2020). Moreover, protein-glycan interactions are weak, with affinities ranging from the  $\mu$ M to the mM range, due to the formation of transient structural states that result in more dynamic interactions than protein-peptide ones (Otto *et al.*, 2011; Sapay *et al.*, 2013; Isaacson and Díaz-Moreno, 2019; Mende *et al.*, 2019; Gimeno *et al.*, 2020; Haab and Klamer, 2020). *In silico* approaches, particularly molecular dynamics, can aid to solve complex protein-glycan interactions, but simulation of carbohydrate-based structures can also be challenging since the initial protein-glycan conformation and the force field employed during the simulation might be critical steps to obtain reliable results. Moreover, molecular docking parameters in the simulation programs are not specifically designed for carbohydrates, but this limitation can be partially addressed by combining docking and molecular dynamics methods of structural analysis that proved to give excellent results when testing protein-ligand interactions (Woods and Tessier, 2010; Das *et al.*, 2018; Fratev *et al.*, 2018; Kumar *et al.*, 2019).

In this work we introduce an *in silico* pipeline designed to predict PRR-glycan interactions avoiding most of the heavy computing requirements and parameter configuration challenges of molecular dynamics. This method might help to screen the high number of putative interactions between plant PRRs and glycoligands, including MAMPs/DAMPs already identified and those to be discovered. Data obtained with this method would pave the way to design biochemical and genetic approaches to confirm the predicted interactions. Using this pipeline, we have validated the previously described structural interaction of the LysM-PRR member CERK1 with chitohexaose (1,4- $\beta$ -D-(GlcNAc)<sub>6</sub>), and we have discarded their direct interaction with laminarihexaose (1,3- $\beta$ -D-(Glc)<sub>6</sub>) and celohexaose (1,4- $\beta$ -D-(Glc)<sub>6</sub>), two glycoligands recently described to trigger immunity. These model predictions were confirmed by *in vitro* binding assays with purified ECDs of CERK1 and LYK4 PRR receptors, further validating the feasibility of the modeling method presented here to decipher PRR-glycoligand interactions.

## RESULTS

### PRRs with glycan-binding ECDs are highly represented in Arabidopsis

Previous studies have predicted different numbers of extracellular PRRs in the Arabidopsis genome (more than 600 members) and considered some of the following subclasses: lectins (C-, L- and G-lectins), LRRs, LysMs,

malectins, CrRLK1Ls, WAKs, CR4Ls, ThaumatinL/PR5, CRKs (also classified as stress-antifungal/salt-response receptors) and PERKs (Gish and Clark, 2011; Bellande *et al.*, 2017; Tang *et al.*, 2017). With this in mind, we decided to perform an additional comprehensive analysis of these proteins in order to establish an updated classification of Arabidopsis extracellular PRRs, comprising RLKs, RLPs and RPs (both RPs and RPe; Figure 1). The main aim was the identification of PRRs containing putative glycan-binding ECDs among these types of receptor proteins. Classification into RLKs, RLPs, RPs or RPe was performed depending on the presence of transmembrane regions, serine/threonine kinase domains or GPI-anchor motifs. This characterization allowed the identification of 617 putative PRRs with extracellular ECDs (Table S1), of which 329 (53.3%) have ECDs that putatively could interact with carbohydrates, indicating the high number of plant PRRs that might potentially bind glycoligands (Figure 1a and Table S2). The remaining 288 PRRs (46.7%) were described as putatively not related with carbohydrate binding based on their ECD topology, and comprise PRRs with different ECDs: LRRs (246 members), PERKs (23 members), CR4L (8 members) and non-structural classified ECDs (11 members) (Table S1).

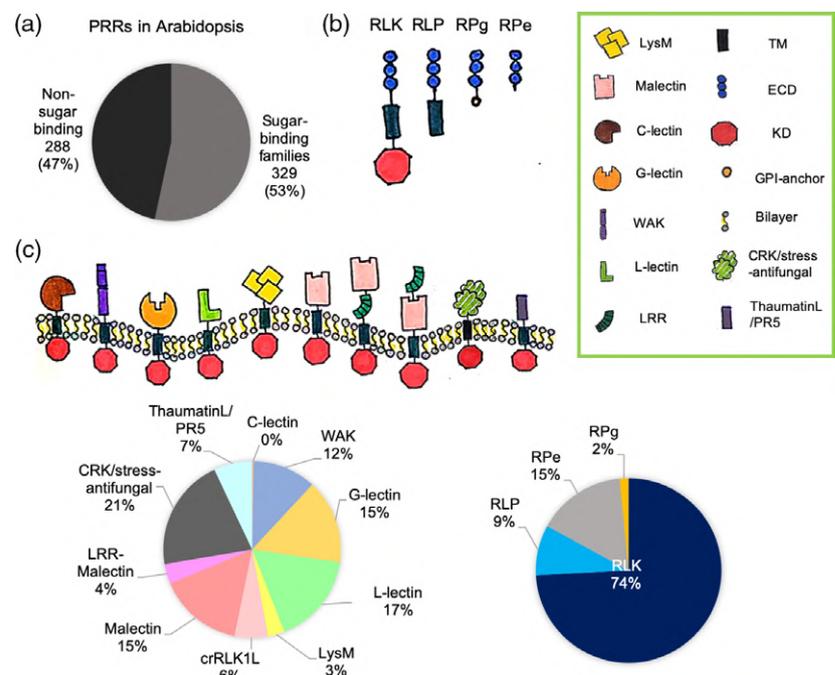
The potential Arabidopsis glycan-binding ECDs were classified based on their similarity with protein domains that have been reported to interact with carbohydrates. As shown in Figure 1(c), the putative glycan-binding PRR members were divided into 10 subclasses based on their different ECDs: C-lectin (1 member), G-lectins (51 members), L-lectins (55 members), WAKs (38 members), LysM (10 members), CrRLK1L (with malectin motif; 20 members),

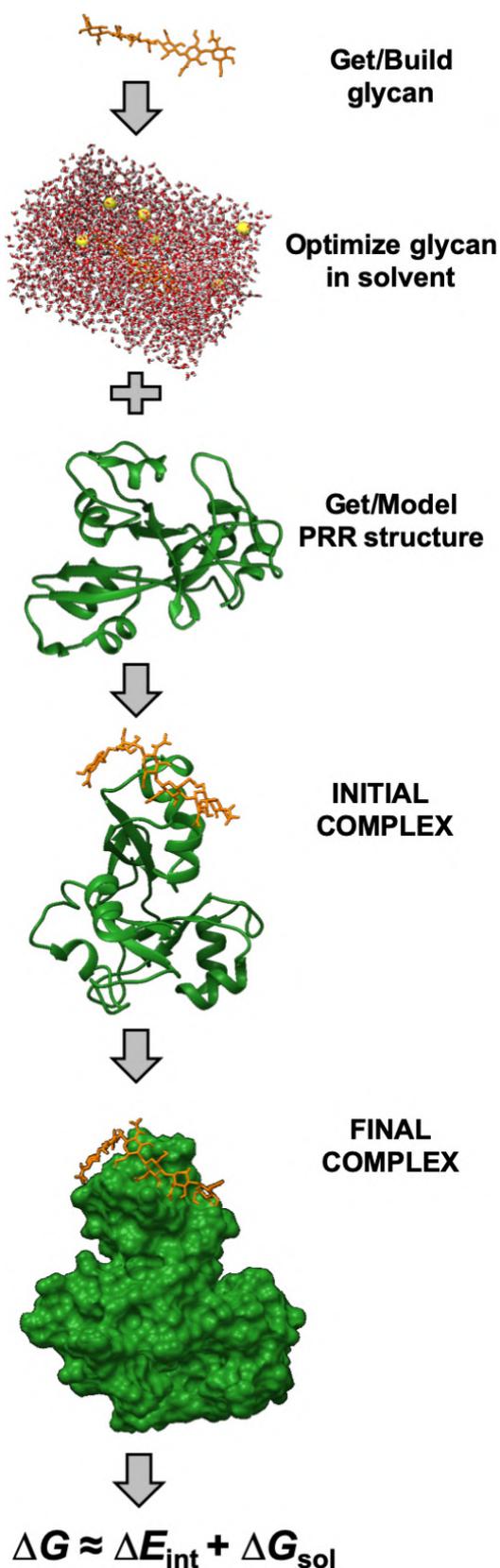
malectins (external malectin motif followed by an LRR; 51 members), LRR-malectins (external LRR domain followed by a malectin motif; 12 members), ThaumatinL/PR5 (23 members) and CRKs (68 members). All these subclasses include RLKs, RLPs and RPs, with the exception of C-lectin and LRR-malectins (which are all RLKs), WAKs (which are all RLKs except for one RPe) and CrRLK1Ls (which are all RLKs except for two RLPs) (Table S2). The LysM-PRR subclass accounts for 10 members in Arabidopsis, with 5 of them considered as RLKs (CERK1/LYK1, LYK2, LYK3, LYK4 and LYK5), 2 as RPe (AT5G62150 and AT4G25433) and 3 as RPs (LYM1, LYM2 and LYM3). Though LYM1 and LYM2 predictions suggest TM-anchor structures (Table S2), they have been previously classified as GPI-anchored proteins (Faulkner *et al.*, 2013; Bellande *et al.*, 2017).

### Computational prediction method for the identification of PRR/glycan complexes

In order to perform *in silico* predictions of PRR/glycan interactions, we developed and optimized a method that consisted on several steps (see Figure 2 and Experimental Procedures). First, the structure of the target glycan (ligand) was generated, either by retrieving it from crystallography or NMR experiments through the Protein Data Bank (PDB; <https://www.rcsb.org/>) or by generating it from scratch adding carbon rings and additional atoms and bonds since just a few crystallography or NMR data are available for glycans (Figure 2). The structures of 1,4- $\beta$ -D-(GlcNac)<sub>6</sub> and 1,3- $\beta$ -D-(Glc)<sub>6</sub> were obtained from the PDB database, whereas 1,4- $\beta$ -D-(Glc)<sub>6</sub> was both retrieved from crystallized data and built from scratch. In order to provide

**Figure 1.** Classification of *Arabidopsis thaliana* PRRs according to the glycan-binding domain of their ECDs. (a) Proportion of Arabidopsis PRRs putatively harboring glycan-binding domains in their ECDs (53.3% of the total 617 PRRs identified). (b) Scheme of the main domain structure of PRRs: receptor-like kinases (RLKs) that contain an extracellular ectodomain (ECD), a transmembrane region (TM) and an intracellular kinase domain (KD); receptor-like proteins (RLPs) with ECD and TM; and receptor proteins (RPs), which either contain an ECD that can be attached to the plasma membrane by a GPI-anchor (RPg) or are extracellular proteins (RPe) not attached to the membrane. (c) Graphical representation of putative glycan-binding PRR-RLK subclasses. Graph depicting the percentages and numbers of each subclass (left) and the proportions of RLKs, RLPs, RPs and RPe of the putative glycan-binding PRRs (right) are shown.





**Figure 2.** Protocol description of the computational method developed for glycan-PRR binding prediction. Glycan and PRR structures are retrieved through the PDB database, or alternatively glycoligands are built from scratch (with the Chimera *Build Structure* tool) and/or PRR-ECDs are modeled (see Experimental Procedures). Ligands are optimized in vacuum and in solvent using a 12-Å padding box (TIP3 water model, 0.15 M NaCl). Then five minimum different conformations (replicates) of the PRR/glycan complexes are calculated employing AutoDock Vina using a 27-Å padding spatial box. The final complex is then obtained for each replicate applying molecular dynamics with two minimization procedures (SolBox1 and SolBox2) and one full simulation procedure (Full) using the CHARMM force field and TIP3 water with 0.15 M NaCl and a 12-Å padding box. The  $\Delta G$  energies are obtained from each of these final complexes.

the ‘most biologically realistic’ glycan conformation, vacuum optimization of glycan structures was performed with Chimera followed by solvent minimization with Visual Molecular Dynamics (VMD)-AutoIMD. In parallel, the PRR-ECD structure of interest was obtained from X-ray or NMR data (e.g., 4EBZ for CERK1-ECD), or it was modeled *in silico* generally through homology modeling (Swiss-Model server; <https://swissmodel.expasy.org/>). For example, we modeled LYK4-ECD using the rice chitin elicitor-binding protein (OsCEBiP, 5JCE) structure (Liu *et al.*, 2016), while Arabidopsis ANXUR1 (6FIG) (Moussu *et al.*, 2018) was used as a template for modeling THESEUS1 (THE1), which was included as a negative control in our analyses. Next, the initial ECD-PRR/glycan complex was determined by docking the glycan in the putative binding site(s) of the PRR (those described or those that we obtained through sequence homology if they had not been previously identified). We considered the described putative binding sites for chitin in CERK1 (amino acids 109–115 and 137–143, encompassing Q109, E110, E114 and I141 directly involved in chitin binding) and LYK4 (amino acids 125–131 and 154–160) (Iizasa *et al.*, 2010; Wan *et al.*, 2012). THE1 binding sites were obtained by sequence alignment with *Xenopus laevis* malectin complexed with nigerose (2K46), as shown in previous studies (Moussu *et al.*, 2018). Docking was achieved with AutoDock Vina using a 27-Å padding box, retrieving a total of five docking complexes as possible different conformations of the glycan in the PRR-binding site (five initial complexes or replicates). Glycosylation positions and disulfide bonds in CERK1 were obtained from manual assertion inferred from a combination of experimental and computational evidence as shown in UniProt (<https://www.uniprot.org/>; Liu *et al.*, 2012). LYK4 and THE1 glycosylation patterns and disulfide bonds were identified by sequence analysis prediction and sequence similarity, respectively (UniProt). The identification of glycosylation sites in ECDs is crucial, since they can provoke false positive binding results and mislead predictions if they are close to the glycoligand.

Once the initial complexes were established they were parameterized in the CHARMM force field using CHARMM-GUI and solved using VMD software in order to prepare

them for molecular dynamics with scalable molecular dynamics (NAMD), following three standardized steps: minimization, equilibration and simulation (<http://www.iitg.ac.in/tamalb/karp/namd/>; [https://www.ch.embnet.org/MD\\_tutorial/](https://www.ch.embnet.org/MD_tutorial/)). The protocol to complete these three steps was optimized with the CHARMM force field at 298 K, pH 7 for 10 ns with canonical NVT ensemble (at temperature and volume constant allowing for pressure variation; see Table S3(c–e) for an example of configuration files). As full simulations with this standard parameterization tended to crash over time or to produce misleading binding results due to their high computing requirements, we decided to carry out only the minimization steps to test LysM-PRR/glycan interactions. These minimization procedures were SolBox1, with boundary conditions applied to NVT ensemble (constant volume and temperature) (Table S3a), and SolBox2, with boundary conditions applied to NPT ensemble (constant temperature and pressure) having a padding 2 Å bigger than SolBox1 to compensate for allowing volume variation (Table S3b). To properly compare, final energies were retrieved for SolBox1, SolBox2 and a full simulation of 10 ns (Full) as described above.

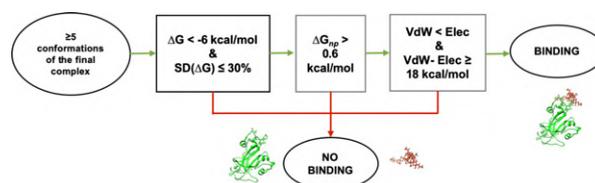
Once the final complexes were obtained, either from minimization or full simulation, we determined their free energies ( $\Delta G$ , kcal mol<sup>-1</sup>), which are used to predict binding events, and that depend on the contribution of internal and solvent energies ( $\Delta E_{\text{int}}$  and  $\Delta G_{\text{sol}}$ ; Figure S1). Entropy variations were not calculated since available calculations are still not reliable enough and they cause the majority of the computational cost in the energy calculations (Sadiq *et al.*, 2010; Hou *et al.*, 2011; Genheden and Ryde, 2015). Energy parameters of the final complexes were calculated following the MM/PBSA protocol with the steps summarized in Figure S1 (Hou *et al.*, 2011): (i) internal energies ( $\Delta E_{\text{int}}$ ), which are the sum of the electrostatic and Van der Waals interactions; (ii) solvent energies ( $\Delta G_{\text{sol}}$ ), which are the sum of polar and non-polar terms, with the polar terms being calculated applying the Poisson–Boltzmann potential to the complex, and then these terms were subtracted separately from the potential of the ECD-PRR alone in the complex and the potential of the glycan in the complex; (iii)  $\Delta G_{\text{sol}}$  non-polar terms, calculated applying a conversion factor into the SASA values ( $\text{gSASA}+b$ ) of the complex and the PRR alone (Gilson and Zhou, 2007; Genheden and Ryde, 2015; Das *et al.*, 2018); and then (iv) internal and solvent energies were summed up to obtain the final  $\Delta G$  energy.

The well-known complex CERK1/1,4- $\beta$ -D-(GlcNAc)<sub>6</sub> was used as positive control for optimizing the methodology (Liu *et al.*, 2012; Cao *et al.*, 2014; Cheval *et al.*, 2020). Three final complexes, two minimization procedures (SolBox1 and SolBox2) and one full simulation procedure (Full) were obtained for the CERK1/1,4- $\beta$ -D-(GlcNAc)<sub>6</sub> complex, which were in accordance with the described binding. Based on

the energies obtained with this well-characterized complex, we established some minimal conditions supporting putative glycoligand/PRR interaction assessment (Figure 3): (i) a minimum of five computational replicates were established to assess the binding with reliability; (ii)  $\Delta G$  energy average values below  $-6$  kcal mol<sup>-1</sup> (standard deviation of maximum 30%) were considered as evidence of interaction, while higher values may indicate non-specific binding, no binding or even repulsion; (iii) non-polar ( $\Delta G_{\text{np}}$ ) values above  $0.6$  kcal mol<sup>-1</sup> were considered as reliable and indicative of binding, while conformations with values below  $0.6$  kcal mol<sup>-1</sup> were discarded even though the final energy was negative since energy values below  $1$  kT (where  $k$  is the Boltzmann constant and  $T$  is the absolute temperature) are noisy and irrelevant because  $kT$  or thermic energy is around  $0.6$ – $0.7$  kcal mol<sup>-1</sup> at a thermodynamic  $T$  of 298 K; and (iv) Van der Waals energies must be more negative than the electrostatic ones with a sufficient difference amongst them of  $\geq 18$  kcal mol<sup>-1</sup> to consider binding conformation as reliable as noted in previous studies (Sadiq *et al.*, 2010; Hou *et al.*, 2011; Kumar *et al.*, 2019; Saravanan *et al.*, 2020; Peng *et al.*, 2020). As additional controls of this last parameter established in the model, we determined the energies of CERK1-, LYK4- and THE1-ECD against an amount of water molecules similar to the size of the ligand (box of 10.0 Å; absence of ligands) and we found that electrostatic terms were more negative than the Van der Waals ones (Table S4), indicating that when ECD-PRRs are bound to the glycan, electrostatic interactions lower their energy values below those of Van der Waals values that become more negative than electrostatic ones (Table 1).

### *In silico* binding prediction of selected LysM-PRR/glycan complexes

The above-described pipeline was applied to predict putative interactions between chitin oligosaccharides of different DP and proteins that have been described to bind chitin oligomers. We first applied the pipeline to determine the binding of chitin oligomers of DP2 and DP4 to



**Figure 3.** Decision tree for putative binding assessment for ECD-PRR/glycan pairs. Three conditions must be sequentially fulfilled to predict trustable energies and ECD-PRR/glycan putative binding (by using five or more conformations of the complex): (i) negative final energy values ( $< -6$  kcal mol<sup>-1</sup>) with  $<30\%$  of variation from average, (ii) non-polar term values higher than  $0.6$  kcal mol<sup>-1</sup> and (iii) electrostatic interaction energy values at least  $18$  kcal mol<sup>-1</sup> higher than the Van der Waals interaction counterparts.

**Table 1** Energy values of CERK1/glycan and LYK4/glycan trials obtained with the *in silico* computational minimization procedure

PRR/glycan	Trial <sup>a</sup>	$\Delta E_{elec}$	$\Delta E_{vdW}$	$\Delta G_{np}$	$\Delta G_{pol}$	$\Delta G$
CERK1/1,4- $\beta$ -D-(GlcNAc) <sub>6</sub>	SolBox1	-26.1 ± 2.4	-47.2 ± 3.5	1.8 ± 0.4	49.8 ± 8.9	<b>-21.6 ± 4.6</b>
	SolBox2	-26.2 ± 3.2	-46.1 ± 3.3	1.7 ± 0.3	51.3 ± 6.5	<b>-19.3 ± 4.2</b>
	Full	-12.6 ± 8.6	-35.1 ± 5.6	2.9 ± 0.4	41.4 ± 14.5	-3.3 ± 16.6
CERK1/1,3- $\beta$ -D-(Glc) <sub>6</sub>	SolBox1	-23.3 ± 7.7	-37.4 ± 5.1	1.0 ± 0.3	46.8 ± 10.8	-12.9 ± 10.8
	SolBox2	-24.4 ± 7.6	-37.4 ± 5.3	0.9 ± 0.2	49.8 ± 24.4	-11.1 ± 14.2
	Full	-4.9 ± 9.9	-11.6 ± 14.6	1.6 ± 1.8	16.1 ± 23.0	1.2 ± 4.1
CERK1/1,4- $\beta$ -D-(Glc) <sub>6</sub>	SolBox1	-21.2 ± 2.3	-34.8 ± 3.0	1.2 ± 0.2	46.2 ± 6.0	-8.7 ± 7.5
	SolBox2	-21.7 ± 2.3	-34.7 ± 3.5	1.2 ± 0.3	42.9 ± 9.8	-12.2 ± 10.7
	Full	-5.7 ± 6.5	-13.1 ± 12.8	1.5 ± 1.4	14.3 ± 26.2	-3.0 ± 22.4
LYK4/1,4- $\beta$ -D-(GlcNAc) <sub>6</sub>	SolBox1	-21.9 ± 4.1	-41.1 ± 11.7	2.3 ± 0.3	39.8 ± 9.7	-20.8 ± 15.5
	SolBox2	-21.8 ± 3.6	-42.6 ± 12.5	2.1 ± 0.3	37.8 ± 5.1	-24.4 ± 11.0
	Full	-6.9 ± 2.5	-22.5 ± 5.5	3.7 ± 0.5	26.2 ± 15.8	0.4 ± 14.5
LYK4/1,3- $\beta$ -D-(Glc) <sub>6</sub>	SolBox1	-19.7 ± 9.0	-26.1 ± 5.8	1.8 ± 0.2	37.2 ± 13.4	-8.2 ± 8.9
	SolBox2	-20.1 ± 8.9	-26.5 ± 3.2	1.6 ± 0.3	33.8 ± 13.3	-11.3 ± 11.6
	Full	-3.5 ± 4.5	-12.9 ± 12.4	2.5 ± 1.5	6.3 ± 8.8	-7.5 ± 9.5
LYK4/1,4- $\beta$ -D-(Glc) <sub>6</sub>	SolBox1	-25.1 ± 3.4	-29.1 ± 4.0	1.6 ± 0.3	46.7 ± 5.2	-6.0 ± 8.5
	SolBox2	-26.5 ± 2.1	-32.8 ± 3.8	1.2 ± 0.3	47.0 ± 7.1	-11.4 ± 7.8
	Full	-0.9 ± 2.0	-13.4 ± 12.1	1.6 ± 1.4	12.4 ± 10.9	-0.3 ± 3.3

<sup>a</sup>SolBox1, SolBox2 and full trials for every PRR–glycan pair are presented as the average ± SD of five replicates. Positive binding parameters fulfilling the criteria of Figure 3 are highlighted in bold. Only the CERK1/1,4- $\beta$ -D-(GlcNAc)<sub>6</sub> pair was assessed as binding in both SolBox1 and SolBox2 trials.

Arabidopsis CERK1 (CERK1) and we found that CERK1 was predicted to bind DP4, but not the DP2 oligomer (Figure S2), as described previously (Izasa *et al.*, 2010; Wan *et al.*, 2012). Then we applied the pipeline to predict binding of the DP4 oligomer to OsCEBiP, the rice chitin receptor, and its co-receptor OsCERK1, and we found that the pipeline predicted OsCEBiP binding to the DP4 oligomer, but not to OsCERK1 (Figure S2), as described previously (Liu *et al.*, 2016). Next, we applied the pipeline to the fungal effectors CfAvr4F and Mg1LysM, which bind to the DP6 chitin oligomer (Hurlburt *et al.*, 2018; Sánchez-Vallet *et al.*, 2020), and we found that the model positively predicted their binding to the DP6 oligomer (Figure S2), further confirming the robustness of the pipeline developed (Figure 3).

We next expanded the application of the pipeline to obtain predictions of the potential binding of ECDs from Arabidopsis CERK1 and LYK4 to several hexasaccharides that have been described to trigger PTI responses: 1,3- $\beta$ -D-(Glc)<sub>6</sub> and 1,4- $\beta$ -D-(Glc)<sub>6</sub>. ECD-THE1 from the CrRLK1L subclass was included in the analysis as a negative control (Table S5). THE1 has been proposed to function as a plant cell wall integrity hub, though it has been recently demonstrated to bind Rapid Alkalinization-Like Factor (RALF) peptides rather than glycans (Hématy *et al.*, 2007; Gonneau *et al.*, 2018). Three final complexes, two minimization procedures (SolBox1 and SolBox2) and one full simulation procedure (Full) were obtained for each PRR–glycan complex. In modeling calculations performed between CERK1 and 1,3- $\beta$ -D-(Glc)<sub>6</sub> we obtained final  $\Delta G$  energies and non-polar term values compatible with binding events, but

electrostatic and Van der Waals energy ratios and the standard deviation of the final binding  $\Delta G$  energy did not fulfill the established criteria based on CERK1/1,4- $\beta$ -D-(GlcNAc)<sub>6</sub> interaction determinations, which matched all the established criteria for binding (Table 1). In the case of LYK4, the standard deviation obtained for the calculated values was very high for both ligands, not allowing to predict direct binding of these ligands to LYK4 (Table 1). We next determined the potential binding of the cellulose-derived hexasaccharide DAMP 1,4- $\beta$ -D-(Glc)<sub>6</sub> to ECDs of these PRRs by obtaining the glycan structure from scratch, and the obtained data were incompatible with binding based on the established criteria (Table 1 and Figure 3). Similarly, modeling of ECD-THE1, included as a negative control, and of 1,3- $\beta$ -D-(Glc)<sub>6</sub>, 1,4- $\beta$ -D-(Glc)<sub>6</sub> and 1,4- $\beta$ -D-(GlcNAc)<sub>6</sub> yielded parameters that were not compatible with binding of these three glycans (Table S5). Moreover, full trials of the glycan/ECD-PRR analyses gave no clear attraction results, proving that most trajectories were not stable enough, further discarding the Full modeling trials as possible test settings for standard parameter binding prediction.

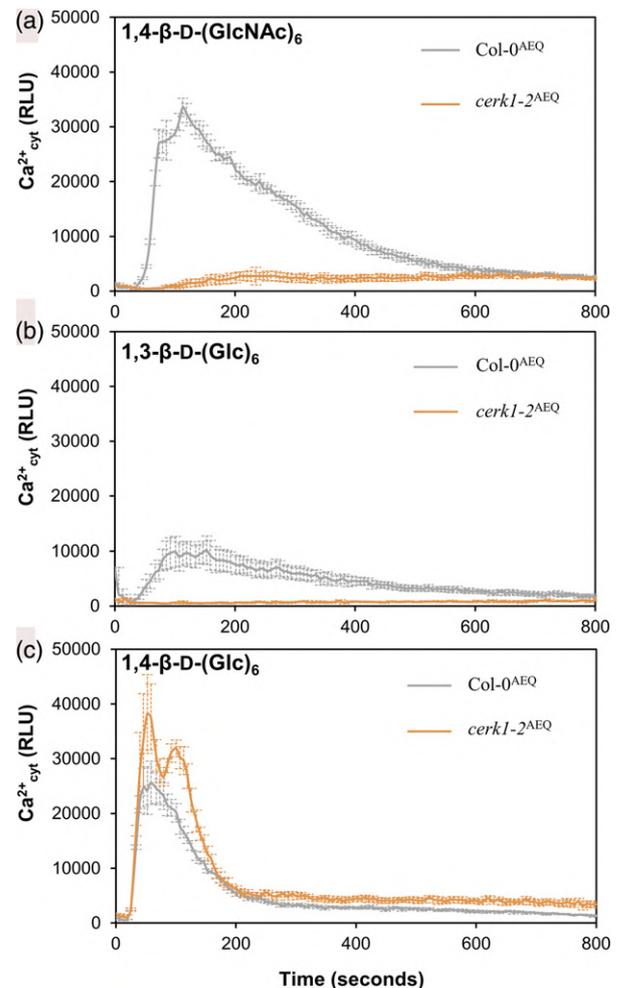
To further confirm that glycan structures generated from scratch did not affect the modeling calculations, we retrieved the glycan structure of cellohexaose (1,4- $\beta$ -D-(Glc)<sub>6</sub>) from the 3D crystal structure of cellobiohydrolase I from *Trichoderma reesei* (PDB: 7CEL; Divne *et al.*, 1998). Notably, we found that this structure is essentially identical to that of *in silico* modeled cellohexaose (Figure S3a), supporting the method developed to optimize glycan structures in solution from scratch (Figure 2). Then, we repeated the modeling calculation with the established

pipeline using this solved structure of cellohexaose and we confirmed the lack of binding of CERK1, LYK4 and THE1 to this glycan (Figure S3b), as predicted with the structure obtained from the *in silico* pipeline (Table 1).

#### *In vitro* binding assays confirmed *in silico* predictions

The *in silico* pipeline predicted a clear direct interaction of 1,4- $\beta$ -D-(GlcNAc)<sub>6</sub> with CERK1, but not with the remaining hexasaccharides tested (Table 1). These data were afterwards validated by assessing the capacity of these ligands to trigger an early plant immune event such as Ca<sup>2+</sup> influxes by using Arabidopsis wild-type and CERK1-defective sensor lines (Col-0<sup>AEQ</sup> and *cerk1-2*<sup>AEQ</sup>, respectively; Ranf *et al.*, 2012). Col-0<sup>AEQ</sup> and *cerk1-2*<sup>AEQ</sup> seedlings were incubated with the different ligands and Ca<sup>2+</sup> influxes were monitored through luminescence measurements (Figure 4). All three hexasaccharides triggered Ca<sup>2+</sup> influxes in Col-0<sup>AEQ</sup> lines, confirming that they were active glycan ligands in Arabidopsis, as described (Cao *et al.*, 2014; Mérida *et al.*, 2018; Locci *et al.*, 2019): 1,4- $\beta$ -D-(GlcNAc)<sub>6</sub> and 1,3- $\beta$ -D-(Glc)<sub>6</sub> produced a wide peak at about 90 sec after treatment followed by a maintained decrease in luminescence that lasted about 600 sec. Ca<sup>2+</sup> influx kinetics upon 1,4- $\beta$ -D-(Glc)<sub>6</sub> treatment were different, since cellohexaose triggered a faster response that was followed by a faster luminescence disappearance (Figure 4c). Notably, Ca<sup>2+</sup> influxes obtained with *cerk1-2*<sup>AEQ</sup> upon treatment with 1,4- $\beta$ -D-(Glc)<sub>6</sub> were similar to those observed in Col-0<sup>AEQ</sup> lines, indicating that CERK1 was not required for 1,4- $\beta$ -D-(Glc)<sub>6</sub> perception, whereas 1,4- $\beta$ -D-(GlcNAc)<sub>6</sub> did not active any Ca<sup>2+</sup> influxes in *cerk1-2*<sup>AEQ</sup>, as predicted by our *in silico* models and described previously (Table 1; Mérida *et al.*, 2018). In contrast *cerk1-2*<sup>AEQ</sup> results for 1,3- $\beta$ -D-(Glc)<sub>6</sub> demonstrated a full dependence on CERK1, as described previously (Mérida *et al.*, 2018), which was not consistent with the predictions of our molecular dynamics models (Table 1).

In view of these results and to further clarify this issue, we expressed CERK1 and LYK4 Arabidopsis ECDs in insect cells and purified them by affinity chromatography (Figure S4). Then, isothermal titration calorimetry (ITC) experiments (Sandoval and Santiago, 2020) were carried out to confirm the *in silico* binding predictions of both CERK1-dependent ligands 1,4- $\beta$ -D-(GlcNAc)<sub>6</sub> and 1,3- $\beta$ -D-(Glc)<sub>6</sub>. Our ITC results proved direct interactions of CERK1 with 1,4- $\beta$ -D-(GlcNAc)<sub>6</sub> ( $K_d$  values of  $37.5 \pm 10.0 \mu\text{M}$ ; Figure 5). The results obtained with 1,3- $\beta$ -D-(Glc)<sub>6</sub> clearly indicated that CERK1 did not bind, at least directly, to this glycan (Figure 5), confirming the *in silico* predictions (Table 1). Similar binding experiments were performed with LYK4 and these two glycans, but no direct binding was detected (Figure 5), which also supported our *in silico* prediction (Table 1). These data suggest that CERK1 takes part of the sensing complex for 1,3- $\beta$ -D-(Glc)<sub>6</sub> in an indirect manner.

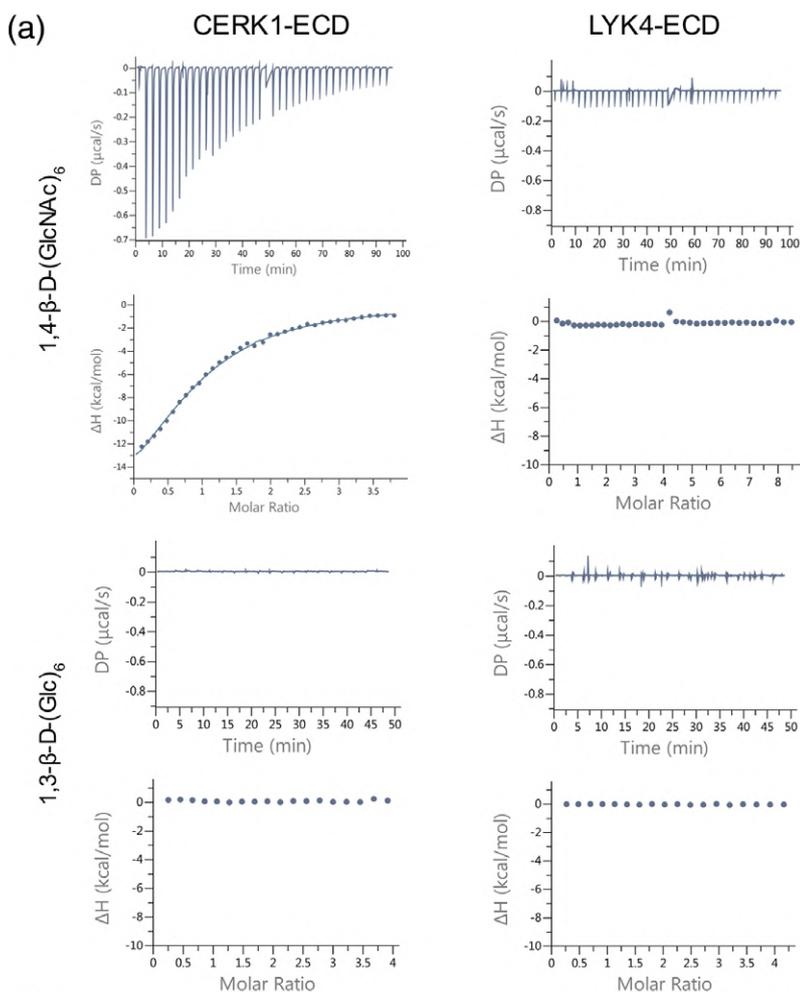


**Figure 4.** Elevations of cytoplasmic calcium concentrations over time in 8-day-old Arabidopsis Col-0<sup>AEQ</sup> and *cerk1-2*<sup>AEQ</sup> seedlings upon treatment with hexasaccharides. (a) Chitohexaose (250  $\mu\text{M}$  1,4- $\beta$ -D-(GlcNAc)<sub>6</sub>). (b) Laminarihexaose (250  $\mu\text{M}$  1,3- $\beta$ -D-(Glc)<sub>6</sub>). (c) Cellohexaose (250  $\mu\text{M}$  1,4- $\beta$ -D-(Glc)<sub>6</sub>). Data are presented as the mean  $\pm$  SD ( $n = 8$ ). Shown is one of three experiments that gave similar results.

Similarly, LYK4 may be involved as a co-receptor in the recognition of chitin glycans but may not physically bind chitin, as recently suggested (Cheval *et al.*, 2020). As predicted by our model, the THE1 ectodomain did not bind any of the CERK1-dependent ligands tested (Figure S5).

#### *In silico* determination of the binding interface residues of CERK1/1,4- $\beta$ -D-(GlcNAc)<sub>6</sub>

Since the *in silico* modeling pipeline was found to be robust, we next tested its capacity to predict the amino acids of the CERK1 binding pocket involved in chitin binding. It has been previously shown in the crystal structure of CERK1 (PDB: 4EBZ; Liu *et al.*, 2012) that residues E110, E114, Q109 and I141 seem to be essential for 1,4- $\beta$ -D-(GlcNAc)<sub>4</sub> binding. We carried out an initial comparison of



**Figure 5.** Direct binding assays between the PRR-ECDs and the corresponding glycans. (a) Isothermal titration calorimetry (ITC) experiments of CERK1 and LYK4 ECDs vs 1,4-β-D-(GlcNAc)<sub>6</sub> and 1,3-β-D-(Glc)<sub>6</sub>. (b) ITC table summaries of PRR-ECDs vs 1,4-β-D-(GlcNAc)<sub>6</sub> and 1,3-β-D-(Glc)<sub>6</sub>. The binding affinities between CERK1 and LYK4 and the corresponding glycans are reported as K<sub>d</sub> (dissociation constant, in micromoles); N indicates the reaction stoichiometry (N = 1 for a 1:1 interaction); ΔH indicates the enthalpy variation. Values indicated in the table are means ± SD of independent experiments (n = 2). N.d. indicates no binding detected.

the common residues involved in chitin binding of our *in silico* replicates of CERK1-1,4-β-D-(GlcNAc)<sub>6</sub> and the CERK1-1,4-β-D-(GlcNAc)<sub>4</sub> crystal and we discovered that in our *in silico* models E110, E114 and Q109 formed hydrogen bonds with the ligand in all the replicates and I141 in some models. Next, we performed *in silico* mutations of one or two of these residues by replacing them with alanine in CERK1 (Q109A, E110A, E114A, I141A, Q109A/E110A, E110A/E114A, E110A/I141A, Q109A/I141A and E114A/I141A). These *cerk1 in silico* mutants were subjected to modeling tests against 1,4-β-D-(GlcNAc)<sub>6</sub> using the SolBox2

procedure with five replicates. Notably, some individual mutations (e.g., Q109A and I141A) and double mutations (e.g., Q109A/E110A, E110A/E114A and Q109A/I141A) resulted in negative binding predictions of *cerk1* to 1,4-β-D-(GlcNAc)<sub>6</sub> (Figure S6), further supporting the relevance of binding of the residues identified in the 4EBZ PDB structure (Liu *et al.*, 2012). These results also indicate that the pipeline described here has the potential to predict the impact on binding of key residues mutations, thus enabling the generation of testable hypotheses for biochemical validation.

## DISCUSSION

The plant immune and mammal innate immune systems share a similar conceptual logic, but plants lack the adaptive immunity of mammals (Nürnberg *et al.*, 2004; Haney *et al.*, 2014). Many non-self- (e.g., from microorganisms) and self-macromolecules (e.g., from plants) harbor molecular patterns, MAMPs and DAMPs, respectively, that have been selected through evolution by plants to be perceived by their diverse set of PRRs (617 putative members in Arabidopsis: Figure 1). Modification of these macromolecules, precursors of MAMPs/DAMPs, by plant enzymes might mask, but also unmask, their immunogenic epitopes. In this context, PRRs are essential components for epitope monitoring, allowing plants to recognize and respond to the high diversity of signals that they are exposed to. Moreover, immune signaling upon pattern perception relies not only on the PRRs themselves, but also on immune signaling complexes consisting of membrane-bound and intracellular proteins that are involved in scaffolding or signal transduction (Albert *et al.*, 2020).

The molecular variety of MAMPs/DAMPs requires different ECD structures and properties for recognition (e.g., binding) and immune activation (Saijo *et al.*, 2018). In this work we present an updated classification of the 617 putative Arabidopsis PRRs with extracellular ECDs that comprises at least 14 different subclasses of receptors based on the ECD sequence and conformation (Table S1). LRR-PRRs is the largest group, and shows some similarities with some PRRs in animals, most notably Toll-like receptors (TLRs) that have an LRR-ECD, a transmembrane motif and a cytoplasmic Toll-like domain instead of the kinase domain of plant RLKs (Ronald and Beutler, 2010). Plant LRR-PRRs, which have received most attention from researchers, typically recognize proteinaceous ligands that trigger developmental and immunogenic processes (Smałowska-Luzan *et al.*, 2018). However, 53.3% of Arabidopsis PRRs harbor ECDs putatively able to interact with glycan-based ligands, but these PRRs have not received much attention. These putative glycan-binding PRRs were classified as C-, G- and L-lectins, CRK/stress-antifungal, CrRLK1L, LysM, LRR-malectins, malectins, ThaumatinL/PR5 and WAKs according to their ECD typology (Figure 1 and Table S1). CRKs constitute the most abundant subclass (68 members), followed by L- and G-lectins and malectins (55, 51 and 51 members, respectively; Figure 1). Some members of these groups have already been demonstrated to interact with carbohydrates and others have been associated with the binding of patterns of different nature (Schallus *et al.*, 2008; de Oliveira Figueiroa *et al.*, 2017; Bacete *et al.*, 2018; Moussu *et al.*, 2018; Cheung *et al.*, 2020).

WAK RLKs, with ECDs containing epidermal growth factor motifs, are the proposed receptors for OGs (He *et al.*, 1996; Anderson *et al.*, 2001; Kohorn and Kohorn, 2012).

Indeed, a recombinant peptide containing amino acids 67 to 254 of the extracellular domain of Arabidopsis WAK1 bound polygalacturonic acid, OGs, pectins and structurally related alginates (Decreux and Messiaen, 2005). However, a crystal structure of WAK-ECDs is not yet available and detailed structural work is still necessary in order to fully demonstrate OG–WAK interaction. On the other hand, considering that the WAK subclass includes 38 members in Arabidopsis, it would be expected that some of them could recognize other carbohydrate-based patterns, like recently described OG variants (Voxeur *et al.*, 2019), but this hypothesis requires further investigations.

Malectin-like domains in the ECDs of CrRLK1L members are known to bind di-glucose in their animal counterparts (Schallus *et al.*, 2008). Therefore, CrRLK1Ls were suggested to bind carbohydrates, but this hypothesis has not been experimentally demonstrated to date (Lindner *et al.*, 2012; Wolf, 2017). The crystal structures of the CrRLK1Ls ANX1 and ANX2 suggest that they are non-canonical malectins or carbohydrate-binding modules (CBMs) as they lack the conserved binding surfaces for carbohydrate ligands (Du *et al.*, 2018; Moussu *et al.*, 2018). Instead, some CrRLK1L members have been demonstrated to bind peptides. For example, FERONIA (FER), one of the best-characterized members of the CrRLK1L subclass, has been shown to bind the peptides RALF1 and RALF23 (Haruta *et al.*, 2014; Stegmann *et al.*, 2017). In spite of this RALF-binding ability, recent results also support FER's role as a sensor of cell wall damage through its interaction with pectins, which could postulate a dual interacting capacity with both peptides and carbohydrates (Feng *et al.*, 2018; Duan *et al.*, 2020). Similarly, lectins are proteins with domains that are well known to bind to carbohydrates and probably have been co-opted for immunity function from their original functions (e.g., catalytic) (Goldstein *et al.*, 1980). Plant lectin-PRRs were defined by homology to their mammalian relatives, and although several non-proteinaceous ligands are known to be perceived by lectin-type RLKs, no ligands of glycan nature have been described for these PRRs. The bulb-type G-lectin LIPOOLIGOSACCHARIDE-SPECIFIC REDUCED ELICITATION (LORE) was initially reported to sense bacterial lipopolysaccharides in Arabidopsis, but later on it was shown that ligands were indeed medium-chain 3-hydroxy fatty acids that copurify with lipopolysaccharides (Ranf *et al.*, 2015; Kutschera *et al.*, 2019). Extracellular ATP (eATP) is one of the best-studied DAMPs in animals. Identification in Arabidopsis of the eATP receptor, the L-lectin DOES NOT RESPOND TO NUCLEOTIDES1 (DORN1), was a major breakthrough in eATP biology and provided a key to address many questions about eATP in plants (Choi *et al.*, 2014). Also, LecRK-I.8 and LecRK-VI.2 L-lectin RLKs have been associated to extracellular NAD (eNAD) and NADP (eNADP) coenzyme perception in Arabidopsis (Singh *et al.*, 2012; Wang *et al.*, 2017), and LecRK-

I.8 has also been shown to be involved in early steps of egg recognition of the white butterfly (Gouhier-Darimont *et al.*, 2019).

LysMs PRRs are the only group of plant PRRs of which all members have been exclusively related to glycan perception, including chitin, peptidoglycans, 1,3- $\beta$ -glucans and lipopolysaccharides (Miya *et al.*, 2007; Willmann *et al.*, 2011; Mérida *et al.*, 2018). In contrast to WAKs, the availability of crystal structures of several LysM-ECDs (e.g., 4EBZ, 5JCE, 5BUM, 5K2L) made this subclass amenable for the optimization of the computational-based pipeline to study ECD–glycoligand interactions. Until very recently, the main ‘Achilles heel’ hampering the discovery of PRR–glycan pairs was considered to be the very limited number of identified MAMPs/DAMPs of this nature (Bacete *et al.*, 2018). However, recent discoveries have considerably increased the collection of active glycoligands in plants (de Azevedo Souza *et al.*, 2017; Claverie *et al.*, 2018; Johnson *et al.*, 2018; Mérida *et al.*, 2018, Mérida *et al.*, 2020; Locci *et al.*, 2019; Zang *et al.*, 2019; Wanke *et al.*, 2020b). Moreover, the development of tools such as synthetic glycan collections will soon facilitate growth of the list (Ruprecht *et al.*, 2020). However, none of these MAMPs/DAMPs, including cello-, laminari-, xyloglucan-, mannan- and arabinoxylo-oligosaccharides, have a single candidate PRR for their perception, whose elucidation will require the combination of heavy biochemical and genetic approaches. In particular, biomolecular interaction assays with pure ECDs of putative receptors have been shown to be the golden standard to allow quantification of potential MAMP/DAMP–ECD binding (Sandoval and Santiago, 2020). However, these approaches are time consuming and resource demanding, thus high-throughput screenings based on these technologies are of high risk and not always affordable.

Here we describe a computational minimization procedure that could favor preliminary screenings by narrowing down ECD candidates to be tested in further biomolecular assays. Using the proposed computational minimization procedures, either SolBox1 or SolBox2 determinations would be enough for a rough binding prediction between an ECD and a glycan. Both solvation boxes yield similar results, thus using any of them (or both) would be a user choice. However, ‘user-friendly’ access to full molecular dynamics simulations with standard parameters proved not to be reliable in the prediction results and is therefore not recommended for the pursued objective of the pipeline described here. The final purpose of the proposed protocol would then be to achieve minimization (the first step of molecular dynamics simulation), which at the end will allow users to obtain results in a computational time frame ranging from minutes to hours, while performing full simulations will increase the time frame from hours to days, depending on computing power and the overall difficulty

of the protocol. Additionally, we have developed an *in silico* method to optimize glycan structures in solvent (‘structures from scratch’), which has been proved here to generate structures that are similar to glycan crystal structures (Table 1 and Figure S3), thus expanding the possibility to test additional glycan structures without reported crystals, using the computational minimization procedure described here.

In Arabidopsis, chitin is perceived by a complex of LysM-RLKs comprising LYK4, LYK5 and CERK1 (Miya *et al.*, 2007; Cao *et al.*, 2014; Xue *et al.*, 2019). LYK4 and LYK5 interact constitutively (Xue *et al.*, 2019), whereas the LYK5–CERK1 interaction is ligand-dependent (Cao *et al.*, 2014). Our *in silico* computational minimization procedure as well as our *in vitro* analyses confirmed the direct interaction of CERK1 with chitin oligosaccharides of DP4 or DP6 (Figure S2), whereas we discarded the direct LYK4–chitin interaction without the support of other LysM partners (Table 1 and Figure 5). Therefore, it is suggested that LYK4 may help in the formation of a CERK1–LYK5 complex for the recognition of chitin at the plasma membrane. This agrees with recent data that suggested that LYK4 may be in complex with LYK5 at the plasma membrane, and when the chitin concentration rises, LYK4 would dissociate from LYK5 to associate with LYM2 in the plasmodesmata region (Cheval *et al.*, 2020). Our *in silico* computational minimization procedure further supports these recent data. Similarly, the computational calculations performed here with rice OsCEBiP and OsCERK1 and some fungal effectors (*Cf*Avr4 and *Mg*1LysM) that bind chitin (Figure S2) also confirmed published data (Liu *et al.*, 2016; Hurlburt *et al.*, 2018; Sánchez-Vallet *et al.*, 2020) and supported the robustness of glycoligand/PRR binding predictions by the *in silico* computational minimization procedure described here. Remarkably, this technical advance might also allow the identification of key PRR residues involved in binding and stabilization of glycan structures, as shown here with the *in silico* validation of the relevance of Q109 and I141 from CERK1 in 1,4- $\beta$ -D-(GlcNAc)<sub>6</sub> binding (Figure S6). Moreover, the methodology described might allow to simulate *in silico* point or multiple mutations in PRR binding pockets prior to their biochemical validation through expression of mutated ECD proteins.

$\beta$ -glucans with 1,3-glycosidic linkages have been shown to be perceived as molecular patterns by different plant species (Klarzynski *et al.*, 2000; Mérida *et al.*, 2013; Wanke *et al.*, 2020b). In mammals, 1,3- $\beta$ -glucans are recognized by Dectin-1, consisting of an extracellular C-type lectin domain connected to the plasma membrane by a stalk (Brown *et al.*, 2003). The extracellular lectin domain of Dectin-1 binds to 1,3- $\beta$ -glucans and mixed 1,3-/1,6- $\beta$ -glucans, with linear  $\beta$ -1,3-(Glc)<sub>10</sub> or a glucan heptasaccharide with one  $\beta$ -1,6-linked glucose side chain being the minimal structures required for Dectin-1 binding (Brown *et al.*,

2003; Palma *et al.*, 2006; Adams *et al.*, 2008). In *Arabidopsis*, the PRR required for 1,3- $\beta$ -glucan perception is yet unknown; however, recent works involved CERK1 in its perception, at least for those with a DP > 5 (Figure 4; Mérida *et al.*, 2018; Wanke *et al.*, 2020b). However, a direct CERK1–1,3- $\beta$ -D-(Glc)<sub>6</sub> interaction was fully discarded based on the molecular dynamics data obtained with the computational minimization method developed here (Figure 4). Notably, this *in silico* prediction of non-direct CERK1–1,3- $\beta$ -D-(Glc)<sub>6</sub> binding was fully confirmed by ITC experiments (Table 1 and Figure 5), further suggesting a co-receptor function of CERK1 in the perception of 1,3- $\beta$ -D-(Glc)<sub>6</sub>. This mechanism of 1,3- $\beta$ -D-(Glc)<sub>6</sub> perception by CERK1 has similarities to that described for the perception of chitin by OsCERK1 in rice through its interaction with the OsCEBiP receptor (Liu *et al.*, 2012), which has been also validated *in silico* here by demonstrating that OsCEBiP, but not OsCERK1, binds this ligand (Figure S2). Structural analyses have revealed a common activation mechanism for RLKs, in which the ECD domains of receptors (e.g., LRR and LysM subclasses) and shape complementary co-receptors heterodimerize in the presence of ligand (Albert *et al.*, 2020). Ligands promote dimerization either by binding both proteins directly as ‘molecular glue’ or allosterically through stabilization of an RLK island (Liu *et al.*, 2012; Sun *et al.*, 2013; Wang *et al.*, 2015; Hohmann *et al.*, 2017). It is known that such activation mechanism exists for LysM-RLKs upon chitin binding (Liu *et al.*, 2012), and our data and previous published data would lead to the hypothesis that CERK1 and maybe also its co-receptor partner LYK5 are involved in the recognition pathway as true receptors, or alternatively as co-receptors together with a *bona fide*, unknown receptor. Cellodextrins, oligosaccharides derived from cellulose, have recently emerged as a group of plant DAMPs showing a great potential for future investigations since they are in high abundance in all plant species and are active at low concentrations, at least in *Arabidopsis* (Aziz *et al.*, 2007; de Azevedo Souza *et al.*, 2017; Johnson *et al.*, 2018; Locci *et al.*, 2019). In spite of the interest generated by these DAMPs, little is known about their perception by plants and the PRRs involved. Our results clearly show that the main LysM-RLKs CERK1 and LYK4 are not involved in cello-oligosaccharide perception (Table 1).

The study of signal transduction in plants has expanded dramatically from the early efforts to define the basic components of signaling for most known hormones and environmental stresses to the current search for PRR/ligand pairs (Cheung *et al.*, 2020). Different types of ECD-PRRs mediate the perception of distinct MAMPs/DAMPs triggering immunity; however, the identification of receptor PRRs for a given ligand is only the tip of the iceberg. How ligand binding induces complex formation with co-receptors that are required for the activation of downstream immune signaling is really challenging (Wang and Chai, 2020). High-

throughput technologies are thus required to accelerate the identification of PRR–ligand pairs. The computational prediction method of glycan/PRR binding presented here might accelerate the discovery of protein–glycan interactions and provide information on immune and/or developmental responses activated by glycoligands. Future work will include (i) the optimization of the method to the specificities of different ECD subclasses and (ii) genetic and biochemical functional validations of wild-type and *in silico* predicted mutant versions of PRRs.

## EXPERIMENTAL PROCEDURES

### Classification of PRRs with putative glycan-binding ECDs

Bioinformatic classification of PRR-ECDs was performed using public databases at UniProt (<https://www.uniprot.org/>; Bateman, 2019), ScanProsite (<https://prosite.expasy.org/scanprosite/>; de Castro *et al.*, 2006) and NCBI (<https://www.ncbi.nlm.nih.gov/>; Agarwala *et al.*, 2017). Extracellular domains were screened and checked with PFAM (<https://pfam.xfam.org/>; El-Gebali *et al.*, 2019) and InterPro (<http://www.ebi.ac.uk/interpro/search/sequence/>; Mitchell *et al.*, 2019) online servers. Transmembrane prediction was done with TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>; Krogh *et al.*, 2001), TMPred ([https://embnet.vital-it.ch/software/TMPRED\\_form.html](https://embnet.vital-it.ch/software/TMPRED_form.html); Hofmann and Stoffel, 1993) and Das-TMFilter (<http://www.enzim.hu/DAS/DAS.html>; Cserzo *et al.*, 2004). GPI-anchored prediction was carried out with predGPI (<http://gpcr.biocomp.unibo.it/predgpi/>; Pierleoni *et al.*, 2008) and the presence of signal peptides was determined with SignalP (<http://www.cbs.dtu.dk/services/SignalP/>; Almagro Armenteros *et al.*, 2019) and PredSi (<http://www.predisi.de/>; Hiller *et al.*, 2004). Curated transmembrane and signal peptide prediction was made for all sugar-binding PRRs with Phobius (<https://www.ebi.ac.uk/Tools/pfa/phobius/>; Käll *et al.*, 2007).

### Theoretical protocol development

Glycan conformation was analyzed using the *Build Structure* tool of Chimera (Pettersen *et al.*, 2004), where carbohydrate structures were generated by adding atom by atom. A primary optimization in vacuum was performed with Chimera software and the *Minimize Structure* tool employing several rounds of a combined method of steepest descend (approximately 2000 steps per round) and conjugate gradient algorithms (approximately 300 steps per round). Hydrogen atoms were added considering hydrogen bonds and charges were computed with AMBER ff14SB and the Gasteiger algorithm. PSF and PDB glycan files were obtained from the *CHARMM-GUI PDB-Reader Tool* (Jo *et al.*, 2008; Lee *et al.*, 2016) and they were introduced into VMD 1.9.3 (Humphrey *et al.*, 1996) to create the solvent box with the *Add Solvation Box* tool (the maximum and minimum coordinates of the molecule are obtained with the ‘measure minmax [atomselect top all]’ command adding a 12-Å padding with TIP3 water). The models were neutralized and an NaCl concentration of 0.15 M was set with the *Add Ions* tool of VMD. Minimization was achieved with the *AutoMD* tool from VMD treating all atoms in the system as mobile.

PRRs were modeled through Swiss-Model (<https://swissmodel.expasy.org/>; Waterhouse *et al.*, 2018) and Phyre2 (<http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index>; Kelley *et al.*, 2015). All models and crystal structures were checked with MolProbity (<http://molprobity.biochem.duke.edu/>; Williams *et al.*,

2018). Binding sites that were not described were obtained with sequence alignment of different templates using Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>; Madeira *et al.*, 2019).

AutoDock Vina (Trot and Olsson, 2010) from Chimera was used to create the initial complexes, usually padding with 27 Å in the binding site. Docked PRR-glycan complexes were introduced into CHARMM-GUI *PDB-Reader* to retrieve the coordinate PSF file and its corresponding PDB, and were then introduced into VMD to get the solvation box as stated above. Molecular dynamics analysis was performed using NAMD2 2.13v software (Phillips *et al.*, 2005). For minimization, 5000 steps were used, and 5 000 000 steps were used for the full simulations. Full simulations were computed with a temperature of 298 K, neutral pH and NVT ensemble.

Parameter information and configuration files employed in the analyses are presented in Table S3.

### Energy calculations and binding assessment

Energies were calculated following a molecular mechanics approach with Poisson-Boltzmann and a surface area solvation MM/PBSA approach. PSF and corresponding DCD files were introduced into VMD to calculate electrostatic and Van der Waals interactions through the tool *NAMD energy*. Final coordinates were saved in a PDB file for further calculations. The PDB file was introduced in CHARMM-GUI *PBEQ Solver* to calculate the electrostatic potentials of the PRR-glycan complex and of the PRR and glycan alone. SASA calculations were retrieved for the PRR-glycan complex and for the PRR using SPDBV (<http://www.expasy.org/spdbv/>; Guex and Peitsch, 1997). Afterwards, a conversion was made to obtain  $gSASA + b$  (where  $g = 0.00526 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ ,  $b = 0.918 \text{ kcal mol}^{-1}$ ).

### In silico carbohydrates and PRRs

The 1,4- $\beta$ -D-(GlcNAc)<sub>6</sub> and 1,3- $\beta$ -D-(Glc)<sub>6</sub> structures were obtained from the 2PI8 and 1W9W PDB structures, respectively. The 1,4- $\beta$ -D-(Glc)<sub>6</sub> structure was built from scratch as described above and was additionally retrieved from the 7CEL crystal. The 1,4- $\beta$ -D-(GlcNAc)<sub>2</sub> and 1,4- $\beta$ -D-(GlcNAc)<sub>4</sub> structures were built *in silico* and retrieved from 5JCE, respectively. The CERK1-ECD structure corresponds to crystal code 4EBZ, while LYK4 was modeled by homology using Swiss-Model and 5JCE as template. THE1 was modeled with Swiss-Model using 6FIG as template. Mg1LysM, CfAvr4 and OsCEBiP correspond to 6Q40, 6BN0 and 5JCE, respectively. OsCERK1 was modeled using 4EBZ as template.

### Comparison of ligand structures and study of the binding interface of CERK1

Comparison of the *in silico* 1,4- $\beta$ -D-(Glc)<sub>6</sub> structure and the crystal 1,4- $\beta$ -D-(Glc)<sub>6</sub> structure was made possible with LigPlot+ (Laskowski and Swindells, 2011). LigPlot+ was also used to check CERK1 residues involved in chitin recognition. Each *in silico* mutation study was performed by changing the residues of interest to alanine in the amino acid sequence of CERK1. Then, those sequences were modeled with Swiss-Model against the CERK1 template (4EBZ) to check for binding alterations.

### Calcium influx assays

Eight-day-old, liquid-grown transgenic *Arabidopsis thaliana* seedlings of ecotype Col-0 carrying the calcium reporter aequorin (Col-0<sup>AEQ</sup>; Ranf *et al.*, 2012) were used for cytoplasmic calcium (Ca<sup>2+</sup><sub>cyt</sub>) measurements using a previously described method (Bacete *et al.*, 2017). The high-purity oligosaccharides used in

these assays (1,4- $\beta$ -D-(GlcNAc)<sub>6</sub>, 1,3- $\beta$ -D-(Glc)<sub>6</sub> and 1,4- $\beta$ -D-(Glc)<sub>6</sub>) were purchased from Megazyme Ltd., Bray, Ireland.

### Protein expression and purification

Codon-optimized synthetic genes coding for Arabidopsis CERK1 (residues 1–232), LYK4 (residues 1–275) and THE1 (residues 1–415) ectodomains were cloned into a modified pFastBac (Geneva Biotech) vector, providing a tobacco etch virus protease cleavable C-terminal StrepII-9xHis tag, for expression in *Spodoptera frugiperda* (Invitro GeneArt, Germany). For protein expression, *Trichoplusia ni* Tnao38 cells (Hashimoto *et al.*, 2012) were infected with a multiplicity of infection of 3 and incubated at 28°C for 1 day and at 22°C for another 2 days, at 110 rpm. The secreted ectodomains were purified from the supernatant by sequential nickel affinity chromatography (HisTrap excel; GE Healthcare, Boston, MA, USA); equilibrated in phosphate K buffer, which contains 50 mM K<sub>2</sub>HPO<sub>4</sub> and 500 mM NaCl, pH 7.8) and StrepII purification (Strep-Tactin Superflow high capacity; IBA; equilibrated in 25 mM Tris, pH 8.0, 250 mM NaCl, 1 mM EDTA). The proteins were further purified by size-exclusion chromatography on a Superdex 200 increase 10/300 GL column (GE Healthcare) or HiLoad 16/600 superdex 200 (GE Healthcare) equilibrated in 20 mM MES buffer pH 6.5, 150 mM NaCl. Peak fractions were analyzed by SDS-PAGE and then concentrated using Amicon Ultra concentrators (Millipore, MWCO 10 000) to reach a protein concentration of 50  $\mu$ M.

### Isothermal titration calorimetry

Experiments were performed at 25°C using a MicroCal PEAQ-ITC (Malvern Instruments, UK) with a 200- $\mu$ l standard cell and a 40- $\mu$ l titration syringe (Moussu *et al.*, 2020). Proteins were gel filtered into ITC buffer (20 mM MES buffer, pH 6.5, 150 mM NaCl) and carbohydrates (1,4- $\beta$ -D-(GlcNAc)<sub>6</sub> and 1,3- $\beta$ -D-(Glc)<sub>6</sub>) were dissolved in the same buffer. Each experiment was performed using an injection pattern of 2  $\mu$ l of the carbohydrate ligand, at 500  $\mu$ M or 1 mM, into 50  $\mu$ M protein in the cell at 150-sec intervals. ITC data were corrected for the heat of dilution by subtracting the mixing enthalpies for titrant solution injections into protein-free ITC buffer. Experiments were done at least in duplicate and data were analyzed using the MicroCal PEAQ-ITC Analysis Software provided by the manufacturer.

### Size-exclusion chromatography

Analytical gel filtration experiments were performed using a Superdex 200 increase 10/300 GL column (GE Healthcare) pre-equilibrated in 20 mM MES (pH 6.5) and 150 mM NaCl. Next, 100  $\mu$ l of the isolated CERK1, LYK4 and THE1 ectodomains was loaded sequentially onto the column and elution at 0.7 ml min<sup>-1</sup> was monitored by ultraviolet absorbance at 280 nm. The column was calibrated with a mixture of the high-molecular weight (HMW) and low-molecular weight (LMW) kits from GE Healthcare. Peak fractions were analyzed by SDS-PAGE.

### ACKNOWLEDGMENTS

This work was supported by grants BIO2015-64077-R of the Spanish Ministry of Economy and Competitiveness (MINECO) and RTI2018-096975-B-I00 of the Spanish Ministry of Science, Innovation and Universities to AM. This work was also financially supported by the 'Severo Ochoa Programme for Centers of Excellence in R&D(2017–2021) from the Agencia Estatal de Investigación of Spain (grant SEV-2016-0672 to CBGP). In the frame of this program HM was supported with a postdoctoral fellow supported by SEV-2016-0672. IdH was the recipient of a PhD FPU

fellow (FPU16/07118) from the Spanish Ministry of Education and from an EMBO Short-Term Fellowship (7985). Research in JS's lab was financially supported by the European Research Council (ERC) grant agreement no. 716358, the Swiss National Science Foundation grants no. 31003A\_173101 and the Programme Fondation Philanthropique Famille Sandoz. We thank Luis Fernández-Pacios (CBGP, Spain) for his advice on modeling calculations and the discussions of the results.

## AUTHOR CONTRIBUTIONS

AM and HM initiated, conceived and coordinated all the experiments except those related to isothermal titration calorimetry (ITC), which were conceived and initiated by JS. IdH performed the experiments described in Figures 1–3, S1–S3 and S6 and Tables 1 and S1–S5 with help from AM and HM. HM performed the experiments described in Figure 4. Experiments described in Figures 5, S4 and S5 were performed by IdH, CB and JS. IdH and HM prepared the tables and figures. IdH, HM and AM wrote the paper. JS edited the paper.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## DATA AVAILABILITY STATEMENT

All relevant data can be found within the manuscript and its supporting materials.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

**Figure S1.** Energy calculation procedure.

**Figure S2.** Energy values obtained in binding trials of 1,4- $\beta$ -D-(GlcNAc)<sub>2-6</sub> and proteins with LysM domains.

**Figure S3.** Comparison data of 1,4- $\beta$ -D-(Glc)<sub>6</sub> built *in silico* vs 1,4- $\beta$ -D-(Glc)<sub>6</sub> from the 7CEL crystal.

**Figure S4.** Scheme of the expression constructs and size-exclusion profiles of the expressed proteins.

**Figure S5.** ITC control experiment of THE1 vs 1,4- $\beta$ -D-(GlcNAc)<sub>6</sub> and 1,3- $\beta$ -D-(Glc)<sub>6</sub>.

**Figure S6.** *In silico* mutagenesis of key CERK1 residues for 1,4- $\beta$ -D-(GlcNAc)<sub>6</sub> stabilization in the binding pocket.

**Table S1.** Classification of Arabidopsis PRRs with different ECDs.

**Table S2.** Of 617 PRRs in Arabidopsis, 329 have ECDs that can putatively bind glycans.

**Table S3.** Example of configuration files for the *in silico* protocol.

**Table S4.** Energy results obtained in the control trials of CERK1, LYK4 and THE1 against a 10.0-Å box of water molecules.

**Table S5.** Energy results for the different binding trials performed for THE1.

## REFERENCES

- Adams, E.L., Rice, P.J., Graves, B. *et al.* (2008) Differential high-affinity interaction of Dectin-1 with natural or synthetic glucans is dependent upon primary structure and is influenced by polymer chain length and side-chain branching. *J. Pharmacol. Exp. Ther.* **325**, 115–123.
- Agarwala, R., Barrett, T., Beck, J. *et al.* (2017) Database resources of the national center for biotechnology information. *Nucleic Acids Res.* **45**, D12–D17.
- Albert, I., Hua, C., Nürnberger, T., Pruitt, R.N. and Zhang, L. (2020) Surface sensor systems in plant immunity. *Plant Physiol.* **182**, 1582.
- Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G. and Nielsen, H. (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**, 420–423.
- Anderson, C.M., Wagner, T.A., Perret, M., He, Z.H., He, D. and Kohorn, B.D. (2001) WAKs: cell wall-associated kinases linking the cytoplasm to the extracellular matrix. *Plant Mol. Biol.* **197**–206.
- de Azevedo Souza, C., Li, S., Lin, A.Z., Boutrot, F., Grossmann, G., Zipfel, C. and Somerville, S.C. (2017) Cellulose-derived oligomers act as damage-associated molecular patterns and trigger defense-like responses. *Plant Physiol.* **173**, 2383–2398.
- Aziz, A., Gauthier, A., Bézier, A., Poinssot, B., Joubert, J.M., Pugin, A., Heyraud, A. and Baillieu, F. (2007) Elicitor and resistance-inducing activities of  $\beta$ -1,4 cellobextrins in grapevine, comparison with  $\beta$ -1,3 glucans and  $\alpha$ -1,4 oligogalacturonides. *J. Exp. Bot.* **58**, 1463–1472.
- Bacete, L., Mérida, H., Pattathil, S., Hahn, M.G., Molina, A. and Miedes, E. (2017) Characterization of plant cell wall damage-associated molecular patterns regulating immune responses. In *Methods in Molecular Biology* (Shan, L., and He, P., eds). New York: Springer, pp. 13–23.
- Bacete, L., Mérida, H., Miedes, E. and Molina, A. (2018) Plant cell wall-mediated immunity: cell wall changes trigger disease resistance responses. *Plant J.* **93**, 614–636.
- Baltrus, D.A., Nishimura, M.T., Romanchuk, A. *et al.* (2011) Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 pseudomonas syringae isolates. *PLoS Pathog.* **7**, e1002132.
- Bateman, A. (2019) UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515.
- Bellande, K., Bono, J.J., Savelli, B., Jamet, E. and Canut, H. (2017) Plant lectins and lectin receptor-like kinases: How do they sense the outside? *Int. J. Mol. Sci.* **18**, 1164.
- Boisson-Dernier, A., Kessler, S.A. and Grossniklaus, U. (2011) The walls have ears: the role of plant CrRLK1Ls in sensing and transducing extracellular signals. *J. Exp. Bot.* **62**, 1581–1591.
- Boutrot, F. and Zipfel, C. (2017) Function, discovery, and exploitation of plant pattern recognition receptors for broad-spectrum disease resistance. *Annu. Rev. Phytopathol.* **55**, 257–286.
- Brown, G.D., Herre, J., Williams, D.L., Willment, J.A., Marshall, A.S.J. and Gordon, S. (2003) Dectin-1 mediates the biological effects of  $\beta$ -glucans. *J. Exp. Med.* **197**, 1119–1124.
- Cao, Y., Liang, Y., Tanaka, K., Nguyen, C.T., Jedrzejczak, R.P., Joachimiak, A. and Stacey, G. (2014) The kinase LYK5 is a major chitin receptor in Arabidopsis and forms a chitin-induced complex with related kinase CERK1. *Elife*, **3**, e03766.
- Carpita, N. and McCann, M. (2000) The cell wall. In *Biochemistry and Molecular Biology of Plants* (Buchanan, B.B., Gruissem, W. and Jones, R.L., eds). Rockville, MD: American Society of Plant Physiologists, pp. 52–108.
- de Castro, E., Sigrist, C.J.A., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P.S., Gasteiger, E., Bairoch, A. and Hulo, N. (2006) ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res.* **34**, W362–W365.
- Chen, S.M., Shen, H., Zhang, T. *et al.* (2017) Dectin-1 plays an important role in host defense against systemic *Candida glabrata* infection. *Virulence*, **8**, 1643–1656.
- Cheung, A.Y., Qu, L.-J., Russinova, E., Zhao, Y. and Zipfel, C. (2020) Update on receptors and signaling. *Plant Physiol.* **182**, 1527.
- Cheval, C., Samwald, S., Johnston, M.G. *et al.* (2020) Chitin perception in plasmodesmata characterizes submembrane immune-signaling specificity in plants. *Proc. Natl. Acad. Sci. USA*, **117**, 9621–9629.
- Choi, J., Tanaka, K., Cao, Y., Qi, Y., Qiu, J., Liang, Y., Lee, S.Y. and Stacey, G. (2014) Identification of a plant receptor for extracellular ATP. *Science*, **343**, 290–294.
- Claverie, J., Balacey, S., Lemaitre-Guillier, C. *et al.* (2018) The cell wall-derived xyloglucan is a new DAMP triggering plant immunity in *Vitis vinifera* and *Arabidopsis thaliana*. *Frontiers in Plant Science*, **9**, 1725.
- Cserzo, M., Eisenhaber, F., Eisenhaber, B. and Simon, I. (2004) TM or not TM: transmembrane protein prediction with low false positive rate using DAS-TMfilter. *Bioinformatics*, **20**, 136–137.
- Dangl, J.L., Horvath, D.M. and Staskawicz, B.J. (2013) Pivoting the plant immune system from dissection to deployment. *Science* **341**, 746–751.

- Das, S., Khatri, S., Siopsis, G. and Wilde, M.M. (2018) Fundamental limits on quantum dynamics based on entropy change. *J. Math. Phys.* **59**, 012205.
- Decreux, A. and Messiaen, J. (2005) Wall-associated kinase WAK1 interacts with cell wall pectins in a calcium-induced conformation. *Plant Cell Physiol.* **46**, 268–278.
- Desaki, Y., Miyata, K., Suzuki, M., Shibuya, N. and Kaku, H. (2018) Plant immunity and symbiosis signaling mediated by LysM receptors. *Innate Immun.* **24**, 92–100.
- Diaz-Alvarez, L. and Ortega, E. (2017) The many roles of galectin-3, a multifaceted molecule, in innate immune responses against pathogens. *Mediators Inflamm.* **2017**, 1–10. <https://doi.org/10.1155/2017/9247574>.
- Divne, C., Ståhlberg, J., Teeri, T.T. and Jones, T.A. (1998) High-resolution crystal structures reveal how a cellulose chain is bound in the 50 Å long tunnel of cellobiohydrolase I from *Trichoderma reesei*. *J. Mol. Biol.* **275**, 309–325.
- Du, S., Qu, L.J. and Xiao, J. (2018) Crystal structures of the extracellular domains of the CrRLK1L receptor-like kinases ANXUR1 and ANXUR2. *Protein Sci.* **27**, 886–892.
- Duan, Q., Liu, M.C.J., Kita, D. et al. (2020) FERONIA controls pectin- and nitric oxide-mediated male–female interaction. *Nature*, **579**, 561–566.
- El-Gebali, S., Mistry, J., Bateman, A. et al. (2019) The Pfam protein families database in 2019. *Nucleic Acids Res.* **47**, D427–D432.
- Faulkner, C., Petutschnig, E., Benitez-Alfonso, Y., Beck, M., Robatzek, S., Lipka, V. and Maule, A.J. (2013) LYM2-dependent chitin perception limits molecular flux via plasmodesmata. *Proc. Natl. Acad. Sci. USA*, **110**, 9166–9170.
- Feng, W., Kita, D., Peaucelle, A. et al. (2018) The FERONIA receptor kinase maintains cell-wall integrity during salt stress through Ca<sup>2+</sup> signaling. *Curr. Biol.* **28**, 666–675.
- Franck, C.M., Westermann, J. and Boisson-Dernier, A. (2018) Plant malectin-like receptor kinases: from cell wall integrity to immunity and beyond. *Annu. Rev. Plant Biol.* **69**, 301–328.
- Fratev, F., Steinbrecher, T. and Jónsdóttir, S.Ó. (2018) Prediction of accurate binding modes using combination of classical and accelerated molecular dynamics and free-energy perturbation calculations: an application to toxicity studies. *ACS Omega*, **3**, 4357–4371.
- Fritz-Laylin, L.K., Krishnamurthy, N., Tör, M., Sjölander, K.V. and Jones, J.D.G. (2005) Phylogenomic analysis of the receptor-like proteins of rice and Arabidopsis. *Plant Physiol.* **138**, 611–623.
- Genheden, S. and Ryde, U. (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **10**, 449–461.
- Gilson, M.K. and Zhou, H.X. (2007) Calculation of protein–ligand binding affinities. *Annu. Rev. Biophys. Biomol. Str.* **36**(1), 21–42.
- Gimeno, A., Valverde, P., Ardá, A. and Jiménez-Barbero, J. (2020) Glycan structures and their interactions with proteins. A NMR view. *Curr. Opin. Struct. Biol.* **62**, 22–30.
- Gish, L.A. and Clark, S.E. (2011) The RLK/Pelle family of kinases. *Plant J.* **66**, 117–127.
- Goldstein, I.J., Hughes, R.C., Monsigny, M., Osawa, T. and Sharon, N. (1980) What should be called a lectin? [2]. *Nature*, **285**, 66.
- Gonneau, M., Desprez, T., Martin, M. et al. (2018) Receptor kinase THESEUS1 is a rapid alkalization factor 34 receptor in Arabidopsis. *Curr. Biol.* **28**, 2452–2458.
- Gouhier-Darimont, C., Stahl, E., Glauser, G. and Reymond, P. (2019) The Arabidopsis lectin receptor kinase lecrk-i.8 is involved in insect egg perception. *Front. Plant Sci.* **10**, 623.
- Greff, C., Roux, M., Mundy, J. and Petersen, M. (2012) Receptor-like kinase complexes in plant innate immunity. *Front. Plant Sci.* **3**, 209.
- Guex, N. and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714–2723.
- Gust, A.A., Biswas, R., Lenz, H.D. et al. (2007) Bacteria-derived peptidoglycans constitute pathogen-associated molecular patterns triggering innate immunity in Arabidopsis. *J. Biol. Chem.* **282**, 32338–32348.
- Haab, B.B. and Klamer, Z. (2020) Advances in tools to determine the glycan-binding specificities of lectins and antibodies. *Mol. Cell. Proteomics*, **19**, 224–232.
- Haney, C.H., Urbach, J.M. and Ausubel, F.M. (2014) Differences and similarities: Innate immunity in plants and animals. *Biochemistry (Lond)*, **36**, 40–45.
- Haruta, M., Sabat, G., Stecker, K., Minkoff, B.B. and Sussman, M.R. (2014) A peptide hormone and its receptor protein kinase regulate plant cell expansion. *Science*, **343**, 408–411.
- Hashimoto, E., Yabuta, Y., Watanabe, F., Morimoto, M., Yamaguchi, Y. and Takenaka, H. (2012) Purification and characterization of phycobiliproteins from edible cyanobacterium *Nostochopsis* sp. *Food Sci. Technol. Res.* **18**, 485–490.
- He, Z.H., Fujiki, M. and Kohorn, B.D. (1996) A cell wall-associated, receptor-like protein kinase. *J. Biol. Chem.* **271**, 19789–19793.
- Hématy, K., Sado, P.E., Tuinen, A.V., Rochange, S., Desnos, T., Balzergue, S., Pelletier, S., Renou, J.P. and Höfte, H. (2007) A receptor-like kinase mediates the response of Arabidopsis cells to the inhibition of cellulose synthesis. *Curr. Biol.* **17**, 922–931.
- Hiller, K., Grote, A., Scheer, M., Münch, R. and Jahn, D. (2004) PrediSi: Prediction of signal peptides and their cleavage positions. *Nucleic Acids Res.* **32**, W375–W379.
- Hofmann, K. and Stoffel, W. (1993) TMbase: a database of membrane spanning protein segments. *Biol. Chem.* **374**, 166.
- Hohmann, U., Lau, K. and Hothorn, M. (2017) The structural basis of ligand perception and signal activation by receptor kinases. *Annu. Rev. Plant Biol.* **68**, 109–137.
- Hou, T., Wang, J., Li, Y. and Wang, W. (2011) Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Model.* **51**, 69–82.
- Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38.
- Hurlburt, N.K., Chen, L.H., Stergiopoulos, I. and Fisher, A.J. (2018) Structure of the *Cladosporium fulvum* Avr4 effector in complex with (GlcNAc)<sub>6</sub> reveals the ligand-binding mechanism and uncouples its intrinsic function from recognition by the Cf-4 resistance protein. *PLoS Pathog.* **14**, e1007263.
- Izasa, E., Mitsutomi, M. and Nagano, Y. (2010) Direct binding of a plant LysM receptor-like kinase, LysM RLK1/CERK1, to chitin in vitro. *J. Biol. Chem.* **285**, 2996–3004.
- Isaacson, R.L. and Diaz-Moreno, I. (2019) Editorial: Weak interactions in molecular machinery. *Front. Mol. Biosci.* **5**, 117.
- Jeong, S., Trotochaud, A.E. and Clark, S.E. (1999) The Arabidopsis CLAVATA2 gene encodes a receptor-like protein required for the stability of the CLAVATA1 receptor-like kinase. *Plant Cell*, **11**, 1925–1933.
- Jo, S., Kim, T., Iyer, V.G. and Im, W. (2008) CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865.
- Johnson, J.M., Thürich, J., Petutschnig, E.K. et al. (2018) A poly(A) ribonuclease controls the cellobiose-based interaction between *Piriformospora indica* and its host Arabidopsis. *Plant Physiol.* **176**, 2496–2514.
- Käll, L., Krogh, A. and Sonnhammer, E.L.L. (2007) Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. *Nucleic Acids Res.* **35**, W429–W432.
- Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. and Sternberg, M.J.E. (2015) The PyMol web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858.
- Klarzynski, O., Plesse, B., Joubert, J.M., Yvin, J.C., Kopp, M., Kloareg, B. and Fritig, B. (2000) Linear  $\beta$ -1,3 glucans are elicitors of defense responses in tobacco. *Plant Physiol.* **124**, 1027–1038.
- Kohorn, B.D. and Kohorn, S.L. (2012) The cell wall-associated kinases, WAKs, as pectin receptors. *Front. Plant Sci.* **3**, 88.
- Krogh, A., Larsson, B., Heijne, G.V. and Sonnhammer, E.L.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580.
- Kumar, A., Srivastava, G., Negi, A.S. and Sharma, A. (2019) Docking, molecular dynamics, binding energy-MM-PBSA studies of naphthofuran derivatives to identify potential dual inhibitors against BACE-1 and GSK-3 $\beta$ . *J. Biomol. Struct. Dyn.* **37**, 275–290.
- Kutschera, A., Dawid, C., Gisch, N. et al. (2019) Bacterial medium-chain 3-hydroxy fatty acid metabolites trigger immunity in Arabidopsis plants. *Science*, **364**, 178–181.
- Laskowski, R.A. and Swindells, M.B. (2011) LigPlot+: Multiple ligand–protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **51**, 2778–2786.
- Latgé, J.P. and Calderone, R. (2006) The fungal cell wall. In *Growth, Differentiation and Sexuality*. Berlin, Heidelberg: Springer, pp. 73–104.

- Lee, J., Cheng, X., Swails, J.M. *et al.* (2016) CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the charmm36 additive force field. *J. Chem. Theory Comput.* **12**, 405–413.
- Lehti-Shiu, M.D., Zou, C., Hanada, K. and Shiu, S.H. (2009) Evolutionary history and stress regulation of plant receptor-like kinase/pelle genes. *Plant Physiol.* **150**, 12–26.
- Li, L., Yu, Y., Zhou, Z. and Zhou, J.M. (2016) Plant pattern-recognition receptors controlling innate immunity. *Sci. China Life Sci.* **59**, 878–888.
- Li, Q., Wang, C. and Mou, Z. (2020) Perception of damaged self in plants. *Plant Physiol.* **182**, 1545–1565.
- Lindner, H., Müller, L.M., Boisson-Dernier, A. and Grossniklaus, U. (2012) CrRLK1L receptor-like kinases: not just another brick in the wall. *Curr. Opin. Plant Biol.* **15**, 659–669.
- Liu, S., Wang, J., Han, Z., Gong, X., Zhang, H. and Chai, J. (2016) Molecular mechanism for fungal cell wall recognition by rice chitin receptor OsCE-BiP. *Structure*, **24**, 1192–1200.
- Liu, T., Liu, Z., Song, C. *et al.* (2012) Chitin-induced dimerization activates a plant immune receptor. *Science*, **336**, 1160–1164.
- Locci, F., Benedetti, M., Pontiggia, D., Citterico, M., Caprari, C., Mattei, B., Cervone, F. and Lorenzo, G.D. (2019) An Arabidopsis berberine bridge enzyme-like protein specifically oxidizes cellulose oligomers and plays a role in immunity. *Plant J.* **98**, 540–554.
- Madeira, F., Park, Y.M., Lee, J. *et al.* (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641.
- Mélida, H., Bacete, L., Ruprecht, C., Rebaque, D., Del Hierro, I., López, G., Brunner, F., Pfrengle, F. and Molina, A. (2020) Arabinoxylan-Oligosaccharides Act as Damage Associated Molecular Patterns in Plants Regulating Disease Resistance. *Front Plant Sci.*, **11**, 1210. <https://doi.org/10.3389/fpls.2020.01210>
- Mélida, H., Sandoval-Sierra, J.V., Diéguez-Urbeondo, J. and Bulone, V. (2013) Analyses of extracellular carbohydrates in oomycetes unveil the existence of three different cell wall types. *Eukaryot. Cell*, **12**, 194–203.
- Mélida, H., Sopena-Torres, S., Bacete, L., Garrido-Arandia, M., Jordá, L., López, G., Muñoz-Barríos, A., Pacios, L.F. and Molina, A. (2018) Non-branched  $\beta$ -1,3-glucan oligosaccharides trigger immune responses in Arabidopsis. *Plant J.* **93**, 34–49.
- Mende, M., Bordoni, V., Tsouka, A., Loeffler, F.F., Delbianco, M. and Seeburger, P.H. (2019) Multivalent glycan arrays. *Faraday Discuss.* **219**, 9–32.
- Mitchell, A.L., Attwood, T.K., Babbitt, P.C. *et al.* (2019) InterPro in 2019: Improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* **47**, D351–D360.
- Miya, A., Albert, P., Shinya, T. *et al.* (2007) CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in Arabidopsis. *Proc. Natl. Acad. Sci. USA*, **104**, 19613–19618.
- Moussu, S., Augustin, S., Roman, A.O., Broyart, C. and Santiago, J. (2018) Crystal structures of two tandem malectin-like receptor kinases involved in plant reproduction. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 671–680.
- Moussu, S., Broyart, C., Santos-Fernandez, G., Augustin, S., Wehrle, S., Grossniklaus, U. and Santiago, J. (2020) Structural basis for recognition of RALF peptides by LRX proteins during pollen tube growth. *Proc. Natl. Acad. Sci. USA*, **117**, 7494–7503.
- Nürnberg, T., Brunner, F., Kemmerling, B. and Piater, L. (2004) Innate immunity in plants and animals: striking similarities and obvious differences. *Immunol. Rev.* **198**, 249–266.
- de Oliveira Figueiroa, E., Albuquerque da Cunha, C.R., Albuquerque, P.B.S. *et al.* (2017) Lectin-carbohydrate interactions: implications for the development of new anticancer agents. *Curr. Med. Chem.* **24**, 3667–3680.
- Otto, D.M.E., Campanero-Rhodes, M.A., Karamanska, R. *et al.* (2011) An expression system for screening of proteins for glycan and protein interactions. *Anal. Biochem.* **411**, 261–270.
- Palma, A.S., Feizi, T., Zhang, Y. *et al.* (2006) Ligands for the  $\beta$ -glucan receptor, dectin-1, assigned using “designer” microarrays of oligosaccharide probes (neoglycolipids) generated from glucan polysaccharides. *J. Biol. Chem.* **281**, 5771–5779.
- Peng, C., Zhu, Z., Shi, Y., Wang, X., Mu, K., Yang, Y., Zhang, X., Xu, Z. and Zhu, W. (2020) Exploring the binding mechanism and accessible angle of SARS-CoV-2 Spike and ACE2 by molecular dynamics simulation and free energy calculation. *ChemRxiv. Org.* **11877492**, v1.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.
- Phillips, J.C., Braun, R., Wang, W. *et al.* (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802.
- Pierleoni, A., Martelli, P. and Casadio, R. (2008) PredGPI: a GPI-anchor predictor. *BMC Bioinformatics*, **9**, 392.
- Ranf, S., Gisch, N., Schäffer, M. *et al.* (2015) A lectin S-domain receptor kinase mediates lipopolysaccharide sensing in Arabidopsis thaliana. *Nat. Immunol.* **16**, 426–433.
- Ranf, S., Grimmer, J., Pöschl, Y., Pecher, P., Chinchilla, D., Scheel, D. and Lee, J. (2012) Defense-related calcium signaling mutants uncovered via a quantitative high-throughput screen in Arabidopsis thaliana. *Mol. Plant*, **5**, 115–130.
- Ronald, P.C. and Beutler, B. (2010) Plant and animal sensors of conserved microbial signatures. *Science*, **330**, 1061–1064.
- Ruprecht, C., Bartzetzko, M.P., Senf, D. *et al.* (2020) A glycan array-based assay for the identification and characterization of plant glycosyltransferases. *Angew. Chemie - Int. Ed.* **59**, 12493–12498.
- Sadiq, S.K., Wright, D.W., Kenway, O.A. and Coveney, P.V. (2010) Accurate ensemble molecular dynamics binding free energy ranking of multidrug-resistant HIV-1 proteases. *J. Chem. Inf. Model.* **50**, 890–905.
- Saijo, Y., Loo, E.-P.-I. and Yasuda, S. (2018) Pattern recognition receptors and signaling in plant–microbe interactions. *Plant J.* **93**, 592–613.
- Sánchez-Vallet, A., Tian, H., Rodríguez-Moreno, L. *et al.* (2020) A secreted LysM effector protects fungal hyphae through chitin-dependent homodimer polymerization. *PLoS Pathog.* **16**, e1008652.
- Sandoval, P.J. and Santiago, J. (2020) Update on techniques to probe ligand-receptor interactions in vitro analytical approaches to study plant ligand-receptor interactions 1[OPEN]. *Plant Physiol. Ö.* **182**, 1697.
- Santiago, J., Brandt, B., Wildhagen, M., Hohmann, U., Hothorn, L.A., Butenko, M.A. and Hothorn, M. (2016) Mechanistic insight into a peptide hormone signaling complex mediating floral organ abscission. *Elife*, **5**, e15075.
- Santiago, J., Henzler, C. and Hothorn, M. (2013) Molecular mechanism for plant steroid receptor activation by somatic embryogenesis co-receptor kinases. *Science*, **341**, 889–892.
- Sapay, N., Nurisso, A. and Imberty, A. (2013) Simulation of carbohydrates, from molecular docking to dynamics in water. In *Biomolecular Simulations*. Totowa, NJ: Humana Press, pp. 469–483.
- Saravanan, K., Hunday, G. and Kumaradhas, P. (2020) Binding and stability of indirubin-3-monoxime in the GSK3 $\beta$  enzyme: a molecular dynamics simulation and binding free energy study. *J. Biomol. Struct. Dyn.* **38**, 957–974.
- Schallus, T., Jaechk, C., Fehér, K. *et al.* (2008) Malectin: a novel carbohydrate-binding protein of the endoplasmic reticulum and a candidate player in the early steps of protein N-glycosylation. *Mol. Biol. Cell*, **19**, 3404–3414.
- Shimizu, T., Nakano, T., Takamizawa, D. *et al.* (2010) Two LysM receptor molecules, CEBiP and OsCERK1, cooperatively regulate chitin elicitor signaling in rice. *Plant J.* **64**(2), 204–214. <https://doi.org/10.1111/j.1365-313X.2010.04324.x>
- Shiu, S.H. and Blecker, A.B. (2003) Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in Arabidopsis. *Plant Physiol.* **132**, 530–543.
- Shiu, S.H., Karlowski, W.M., Pan, R., Tzeng, Y.H., Mayer, K.F.X. and Li, W.H. (2004) Comparative analysis of the receptor-like kinase family in Arabidopsis and rice. *Plant Cell*, **16**, 1220–1234.
- Singh, P., Kuo, Y.C., Mishra, S. *et al.* (2012) The lectin receptor Kinase-VL2 is required for priming and positively regulates Arabidopsis pattern-triggered immunity. *Plant Cell*, **24**, 1256–1270.
- Smakowska-Luzan, E., Mott, G.A., Parys, K. *et al.* (2018) An extracellular network of Arabidopsis leucine-rich repeat receptor kinases. *Nature*, **553**, 342–346.
- Sonah, H., Deshmukh, R.K. and Bélanger, R.R. (2016) Computational prediction of effector proteins in fungi: opportunities and challenges. *Front. Plant Sci.* **7**, 126.
- Srivastava, V., McKee, L.S. and Bulone, V. (2017) *Plant Cell Walls*. in eLS. Chichester, UK: John Wiley & Sons Ltd. <https://doi.org/10.1002/9780470015902.a0001682.pub3>
- Stegmann, M., Monaghan, J., Smakowska-Luzan, E., Rovenich, H., Lehner, A., Holton, N., Belkhadir, Y. and Zipfel, C. (2017) The receptor kinase FER

- is a RALF-regulated scaffold controlling plant immune signaling. *Science*, **355**, 287–289.
- Sun, Y., Li, L., Macho, A.P., Han, Z., Hu, Z., Zipfel, C., Zhou, J.M. and Chai, J.** (2013) Structural basis for flg22-induced activation of the Arabidopsis FLS2-BAK1 immune complex. *Science*, **342**, 624–628.
- Tang, D., Wang, G. and Zhou, J.M.** (2017) Receptor kinases in plant-pathogen interactions: More than pattern recognition. *Plant Cell*, **29**, 618–637.
- Trott, O. and Olson, A.J.** (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**(2), 455–461.
- Voxeur, A., Habrylo, O., Guénin, S. et al.** (2019) Oligogalacturonide production upon *Arabidopsis thaliana*-*Botrytis cinerea* interaction. *Proc. Natl. Acad. Sci. USA*, **116**, 19743–19752.
- Wan, J., Tanaka, K., Zhang, X.C., Son, G.H., Brechenmacher, L., Nguyen, T.H.N. and Stacey, G.** (2012) LYK4, a lysin motif receptor-like kinase, is important for chitin signaling and plant innate immunity in Arabidopsis. *Plant Physiol.* **160**, 396–406.
- Wang, C., Zhou, M., Zhang, X., Yao, J., Zhang, Y. and Mou, Z.** (2017) A lectin receptor kinase as a potential sensor for extracellular nicotinamide adenine dinucleotide in *Arabidopsis thaliana*. *Elife*, **6**, e25474.
- Wang, J. and Chai, J.** (2020) Structural insights into the plant immune receptors PRRs and NLRs. *Plant Physiol.* **182**, 1566–1581.
- Wang, J., Li, H., Han, Z., Zhang, H., Wang, T., Lin, G., Chang, J., Yang, W. and Chai, J.** (2015) Allosteric receptor activation by the plant peptide hormone phytosulfokine. *Nature*, **525**, 265–268.
- Wanke, A., Malisic, M., Wawra, S. and Zuccaro, A.** (2020) Unraveling the sugar code: the role of microbial extracellular glycans in plant-microbe interactions. *J. Exp. Bot.*, <https://doi.org/10.1093/jxb/eraa414>.
- Wanke, A., Rovenich, H., Schwanke, F., Velte, S., Becker, S., Hehemann, J.H., Wawra, S. and Zuccaro, A.** (2020) Plant species-specific recognition of long and short  $\beta$ -1,3-linked glucans is mediated by different receptor systems. *Plant J.* **102**, 1142–1156.
- Waterhouse, A., Bertoni, M., Bienert, S. et al.** (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303.
- Williams, C.J., Headd, J.J., Moriarty, N.W. et al.** (2018) MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315.
- Willmann, R., Lajunen, H.M., Erbs, G. et al.** (2011) Arabidopsis lysin-motif proteins LYM1 LYM3 CERK1 mediate bacterial peptidoglycan sensing and immunity to bacterial infection. *Proc. Natl. Acad. Sci. USA*, **108**, 19824–19829.
- Wolf, S.** (2017) Plant cell wall signalling and receptor-like kinases. *Biochem. J.* **474**, 471–492.
- Woods, R.J. and Tessier, M.B.** (2010) Computational glycoscience: characterizing the spatial and temporal properties of glycans and glycan-protein complexes. *Curr. Opin. Struct. Biol.* **20**, 575–583.
- Xue, D.X., Li, C.L., Xie, Z.P., Staehelin, C. and Napier, R.** (2019) LYK4 is a component of a tripartite chitin receptor complex in *Arabidopsis thaliana*. *J. Exp. Bot.* **70**, 5507–5516.
- Zang, H., Xie, S., Zhu, B., Yang, X., Gu, C., Hu, B., Gao, T., Chen, Y. and Gao, X.** (2019) Mannan oligosaccharides trigger multiple defence responses in rice and tobacco as a novel danger-associated molecular pattern. *Mol Plant Pathol.* **20**(8), 1067–1079. <https://doi.org/10.1111/mpp.12811>
- Zipfel, C.** (2014) Plant pattern-recognition receptors. *Trends Immunol.* **35**, 345–351.

# Cell wall-derived mixed-linked $\beta$ -1,3/1,4-glucans trigger immune responses and disease resistance in plants

Diego Rebaque<sup>1,2,3</sup>, Irene del Hierro<sup>1,2</sup>, Gemma López<sup>1</sup>, Laura Bacete<sup>1,2,†</sup> , Francisco Vilaplana<sup>4</sup>, Pietro Dallabernardina<sup>5</sup>, Fabian Pfrengle<sup>5,8</sup>, Lucía Jordá<sup>1,2</sup>, Andrea Sánchez-Vallet<sup>1</sup>, Rosa Pérez<sup>3</sup>, Frédéric Brunner<sup>3</sup>, Antonio Molina<sup>1,2,†</sup>  and Hugo Mérida<sup>1,\*,†,†</sup> 

<sup>1</sup>Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid (UPM) - Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Campus de Montegancedo UPM, Pozuelo de Alarcón (Madrid), Spain,

<sup>2</sup>Departamento de Biotecnología-Biología Vegetal, Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas, UPM, Madrid, Spain,

<sup>3</sup>Plant Response Biotech, Centro de Empresas, Campus de Montegancedo UPM, Pozuelo de Alarcón (Madrid), Spain,

<sup>4</sup>Division of Glycoscience, School of Biotechnology, Royal Institute of Technology (KTH), Stockholm, Sweden, and

<sup>5</sup>Department of Biomolecular Systems, Max Planck Institute of Colloids and Interfaces, Potsdam, Germany

Received 27 August 2020; revised 26 January 2021; accepted 28 January 2021; published online 5 February 2021.

\*For correspondence (e-mail h.merida@upm.es and antonio.molina@upm.es).

†These authors should be considered joint senior.

<sup>‡</sup>Present address: Institute for Biology, Faculty of Natural Sciences, Norwegian University of Science and Technology, Trondheim, Norway

<sup>§</sup>Present address: Department of Chemistry, University of Natural Resources and Life Sciences, Vienna, Austria

<sup>¶</sup>Present address: Área de Fisiología Vegetal, Departamento de Ingeniería y Ciencias Agrarias, Universidad de León, León, Spain

## SUMMARY

Pattern-triggered immunity (PTI) is activated in plants upon recognition by pattern recognition receptors (PRRs) of damage- and microbe-associated molecular patterns (DAMPs and MAMPs) derived from plants or microorganisms, respectively. To understand better the plant mechanisms involved in the perception of carbohydrate-based structures recognized as DAMPs/MAMPs, we have studied the ability of mixed-linked  $\beta$ -1,3/1,4-glucans (MLGs), present in some plant and microbial cell walls, to trigger immune responses and disease resistance in plants. A range of MLG structures were tested for their capacity to induce PTI hallmarks, such as cytoplasmic  $\text{Ca}^{2+}$  elevations, reactive oxygen species production, phosphorylation of mitogen-activated protein kinases and gene transcriptional reprogramming. These analyses revealed that MLG oligosaccharides are perceived by *Arabidopsis thaliana* and identified a trisaccharide,  $\beta$ -D-cellobiosyl-(1,3)- $\beta$ -D-glucose (MLG43), as the smallest MLG structure triggering strong PTI responses. These MLG43-mediated PTI responses are partially dependent on LysM PRRs CERK1, LYK4 and LYK5, as they were weaker in *cerk1* and *lyk4 lyk5* mutants than in wild-type plants. Cross-elicitation experiments between MLG43 and the carbohydrate MAMP chitohexaose [ $\beta$ -1,4-D-(GlcNAc)<sub>6</sub>], which is also perceived by these LysM PRRs, indicated that the mechanism of MLG43 recognition could differ from that of chitohexaose, which is fully impaired in *cerk1* and *lyk4 lyk5* plants. MLG43 treatment confers enhanced disease resistance in *A. thaliana* to the oomycete *Hyaloperonospora arabidopsidis* and in tomato and pepper to different bacterial and fungal pathogens. Our data support the classification of MLGs as a group of carbohydrate-based molecular patterns that are perceived by plants and trigger immune responses and disease resistance.

**Keywords:** *Arabidopsis thaliana*, *Capsicum annuum*, cell wall, disease resistance, mixed-linked glucan, *Hyaloperonospora arabidopsidis*, pattern triggered immunity, plant immunity, *Solanum lycopersicum*.

## INTRODUCTION

As sessile organisms, plants have evolved a complex immune system comprised by several defence layers. One of them is known as pattern-triggered immunity (PTI), which is based on the recognition of damage- and

microbe-associated molecular patterns (DAMPs and MAMPs) derived from plants or microorganisms, respectively, by plasma membrane-resident pattern recognition receptors (PRRs). Upon DAMP/MAMP recognition, PRRs activate protein kinase signalling cascades that trigger

gene reprogramming processes, which ultimately result in plant surveillance to pathogen/pest attack (Bigeard *et al.*, 2015; Boutrot and Zipfel, 2017). PTI relevance is well-illustrated by the fact that immune responses and disease resistance to pathogens are compromised in plants defective in PRRs perceiving DAMPs or MAMPs of a peptidic nature, such as PEPR1 and FLS2 Receptor Like Kinases that recognize Arabidopsis AtPep1 DAMP and bacterial flg22 MAMP peptides, respectively (Gómez-Gómez and Boller, 2000; Yamaguchi *et al.*, 2006). Many PRR/peptidic DAMP or MAMP pairs triggering PTI have been elucidated (Boutrot and Zipfel, 2017; Tang *et al.*, 2017). However, the specific mechanisms of plant defence activation by carbohydrate-based DAMPs and MAMPs, which are highly abundant in plant and microbial extracellular layers, clearly lags behind our knowledge of peptide ligand recognition (Bacete *et al.*, 2018). The first bottleneck in the identification of carbohydrate-based ligand-PRR pairs is the relatively low number of carbohydrates known to trigger plant immune responses in contrast to the mammal counterpart (Bacete *et al.*, 2018). Among the carbohydrates recognized by the plant immune system are chitin and  $\beta$ -1,3-glucan from fungal/oomycete cell walls, peptidoglycan from bacterial walls, and cellulose ( $\beta$ -1,4-glucan), xyloglucan, mannan, xylan and homogalacturonan from plant cell walls (Aziz *et al.*, 2007; Claverie *et al.*, 2018; Galletti *et al.*, 2008; Gust *et al.*, 2007; Kaku *et al.*, 2006; Klarzynski *et al.*, 2000; Mélida *et al.*, 2018, 2020; Wanke *et al.*, 2020; Zang *et al.*, 2019).

In plants, PRR/carbohydrate-based ligand characterization at the structural level has been mainly limited to PRRs of the LysM family, which are involved in the recognition of several glycoligands such as chitin, peptidoglycans and lipopolysaccharides (Cao *et al.*, 2014; Desaki *et al.*, 2018; Liu *et al.*, 2012; Miya *et al.*, 2007; Willmann *et al.*, 2011). Recent work demonstrated that LysM-PRRs are also implicated in the perception by plants of  $\beta$ -1,3-glucans (Mélida *et al.*, 2018; Wanke *et al.*, 2020). Specifically, a  $\beta$ -1,3-glucan hexasaccharide (laminarihexaose or Lam6) is an immune-active structure whose recognition is dependent in Arabidopsis on the LysM-PRR CERK1 (Chitin Elicitor Receptor Kinase 1) (Mélida *et al.*, 2018). However, a direct binding of laminarihexaose to CERK1 extracellular ectodomain (ECD) has not been either observed in isothermal titration calorimetry binding assays performed with purified ECD-CERK1 nor predicted using recently developed *in silico* structural molecular dynamics simulations (del Hierro *et al.*, 2020). These data suggest that CERK1 function as co-receptor rather than a receptor in  $\beta$ -1,3-glucan hexasaccharide perception. Other plant species, such as rice and tobacco, have been shown to also recognize  $\beta$ -1,3-glucans with a higher degree of polymerization (DP) than 6 and this recognition has been demonstrated to be CERK1-independent, suggesting that  $\beta$ -glucan recognition may be mediated by multiple receptor/co-receptor proteins (Wanke

*et al.*, 2020). CERK1 does not seem to be involved in the perception of  $\beta$ -1,4-glucans (e.g. cellohexaose) in Arabidopsis as immune responses triggered by this DAMP were not impaired in *cerk1* mutant, as predicted by *in silico* structural molecular dynamics simulations (del Hierro *et al.*, 2020).

Glucans represent a group of widely distributed polysaccharides, mainly found in the extracellular layers of numerous phylogenetic groups across the tree of life (Latgé and Calderone, 2006; McIntosh *et al.*, 2005; Mélida *et al.*, 2013; Srivastava *et al.*, 2017). These include a wide variety of structures, mainly with  $\beta$ -linkages, although  $\alpha$ -linked glucans also occur in many species. Mixed-linked glucans [MLGs;  $\beta$ -1,3/1,4-glucans; (1,3;1,4)- $\beta$ -D-glucans] consist of unbranched and unsubstituted chains of  $\beta$ -1,4-glucosyl residues interspersed by  $\beta$ -1,3 linkages (Burton and Fincher, 2009). MLGs are widely distributed as matrix polysaccharides in cell walls of the plants from the Poaceae group, but have also been reported in other species such as in the walls of *Equisetum* spp. and other vascular plants outside the Poaceae (Fincher and Stone, 2004; Fry *et al.*, 2008; Smith and Harris, 1999; Sørensen *et al.*, 2008; Trethewey *et al.*, 2005), bryophytes and algae (Popper and Fry, 2003; Salmeán *et al.*, 2017). MLGs have been also described in lichen-forming ascomycete symbionts (Perlin and Suzuki, 1962; Gorin *et al.*, 1988; Stone and Clarke, 1992), in fungi and oomycetes (Fontaine *et al.*, 2000; Pettolino *et al.*, 2009; Samar *et al.*, 2015) and bacteria (Lee and Hollingsworth, 1997; Pérez-Mendoza *et al.*, 2015).  $\beta$ -Glucans are well-known modulators of the immune system in mammals, but less is known about their roles in the plant counterpart (Fesel and Zuccaro, 2016; Mélida *et al.*, 2018; Rovenich *et al.*, 2016; Sharp *et al.*, 1984; Wanke *et al.*, 2020). Interestingly, this field of  $\beta$ -glucan perception by plants has regained momentum due to recent discoveries demonstrating that glucans containing either  $\beta$ -1,3 or  $\beta$ -1,4 glycosidic linkages in their main backbones trigger PTI responses in plants (Claverie *et al.*, 2018; del Hierro *et al.*, 2020; Johnson *et al.*, 2018; Mélida *et al.*, 2018; Souza *et al.*, 2017; Wanke *et al.*, 2020; Wawra *et al.*, 2016). However, it remained elusive whether glucans containing both types of linkages (MLGs) can be perceived by plant cells as well (Aziz *et al.*, 2007; Johnson *et al.*, 2018; Klarzynski *et al.*, 2000; Locci *et al.*, 2019; Mélida *et al.*, 2018; Ménard *et al.*, 2004; Souza *et al.*, 2017).

In an effort to characterize additional carbohydrate-based structures (MAMPs or DAMPs) that are able to activate the plant immune system, we have determined the ability of different MLG structures to trigger early immune responses in Arabidopsis. We tested the capacity to induce PTI hallmarks (Boudsocq *et al.*, 2010; Ranf *et al.*, 2011) of several MLG-enriched fractions and purified MLGs. Our results demonstrate that MLGs are a group of glycoligands that activate plant immunity and we characterized  $\beta$ -

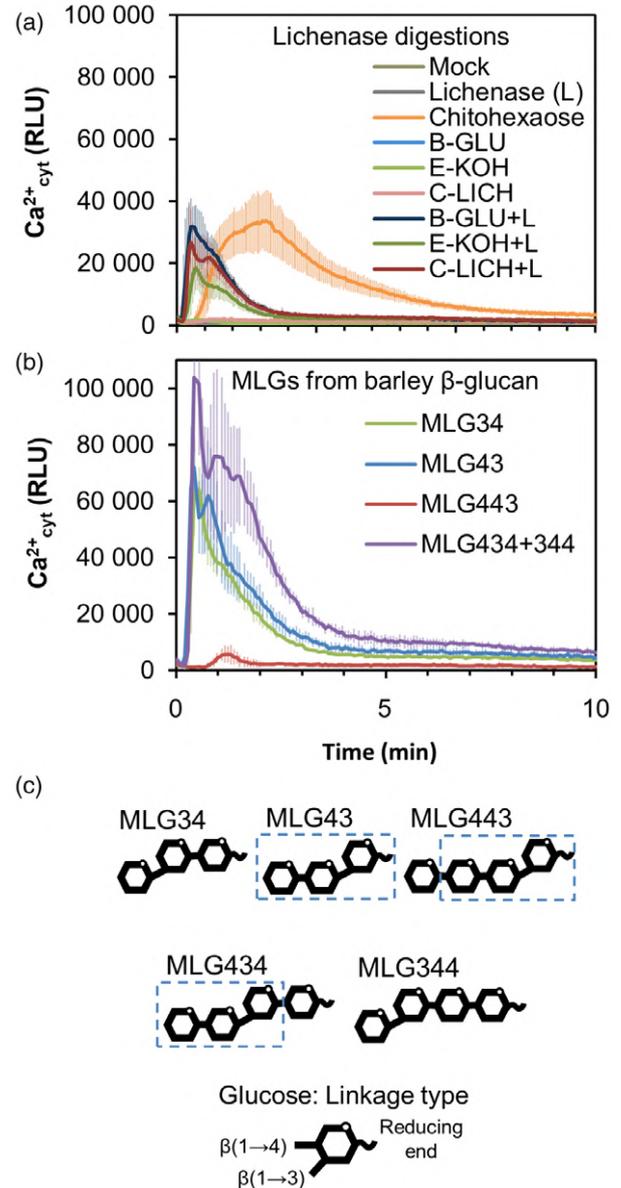
D-cellobiosyl-(1,3)- $\beta$ -D-glucose (MLG43) as the smallest active structure triggering immune responses in Arabidopsis and boosting Arabidopsis and crops (e.g. tomato and pepper) disease resistance to different pathogens.

## RESULTS

### MLG oligosaccharides trigger $\text{Ca}^{2+}$ influxes in Arabidopsis

Aequorin-based  $\text{Ca}^{2+}$  Columbia-0 (Col-0<sup>AEQ</sup>) sensor Arabidopsis system (Ranf *et al.*, 2011) was used to monitor whether early cytoplasmic  $\text{Ca}^{2+}$  influxes (burst) occurred upon treatment of seedlings with cell wall fractions enriched in MLGs, such as  $\beta$ -1,3/1,4-glucans from barley (B-GLU), an alkali-extracted fraction isolated from *Equisetum arvense* cell wall (E-KOH) and lichenan (a mixture of  $\beta$ -1,3/1,4-glucans) from *Cetraria islandica* (C-LICH). None of these polymeric fractions induced  $\text{Ca}^{2+}$  influxes in comparison with pure MAMPs, such as the chitin hexasaccharide chitohexaose [ $\beta$ -1,4-D-(GlcNAc)<sub>6</sub>] (Figure 1a). To generate MLG oligosaccharides with lower DP than these polymeric fractions, MLGs were treated with lichenase from *Bacillus subtilis* (EC 3.2.1.73), a glucan endohydrolase that catalyses the hydrolysis of  $\beta$ -1,4 bonds immediately following  $\beta$ -1,3 bonds, but does not catalyse the hydrolysis of purely  $\beta$ -1,3- or  $\beta$ -1,4-linked glucans (Henrissat and Bairoch, 1993; Planas, 2000). MLG polysaccharides digested with lichenase released MLGs with a single  $\beta$ -1,3 linkage placed next to the reducing end. Interestingly, products from lichenase digestions of these MLG polysaccharides activated plant  $\text{Ca}^{2+}$  influxes (B-GLU + L, E-KOH + L or C-LICH + L; Figure 1a), indicating that lichenase-released MLG oligosaccharides can trigger early immune responses in Arabidopsis.

To confirm, that lichenase-released oligosaccharides were plant immune-active structures, we screened the activity of MLG oligosaccharides from commercial sources obtained by the purification of B-GLU after enzymatic digestions with lichenase and cellulase (Figure 1b). Results from the commercial source clearly demonstrated that, at least, the shortest MLG oligosaccharides with DP 3 (e.g. MLG43 and MLG34) and some with DP 4 (MLG434 + MLG344 mixture and MLG443 in a lesser extent) were able to trigger  $\text{Ca}^{2+}$  influxes (Figure 1b). As MLGs consist of unbranched and unsubstituted chains of  $\beta$ -1,4-glucosyl residues interspersed by  $\beta$ -1,3 linkages, the minimal MLGs structures containing both types of linkages would be MLG43 and MLG34 (DP 3; Figure 1c). We decided to compare  $\text{Ca}^{2+}$  influxes triggered by MLG43 and MLG34 trisaccharides and by their constituent disaccharides in the Col-0<sup>AEQ</sup> sensor lines, and we observed that while cellobiose ( $\beta$ -1,4-linked disaccharide) triggered a slight  $\text{Ca}^{2+}$  influx, the  $\beta$ -1,3-linked glucan disaccharide did not (Lam2; Figure S1). In contrast, cellotriase (Cello3), a well characterized DAMP (Johnson *et al.*, 2018; Locci *et al.*, 2019),



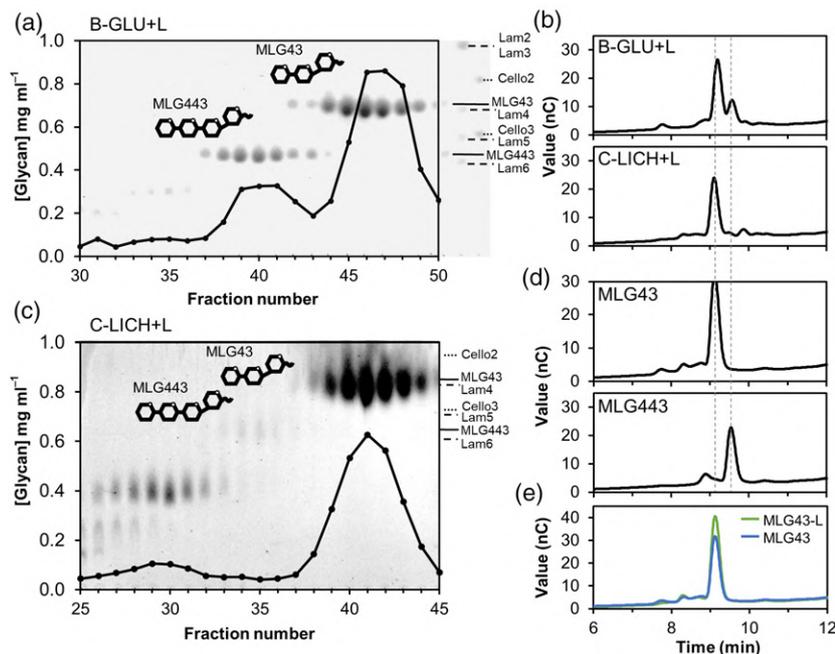
**Figure 1.** Mixed-linked glucans (MLG) oligosaccharides trigger cytoplasmic  $\text{Ca}^{2+}$  elevations in Arabidopsis.

$\text{Ca}^{2+}$  influxes were measured as relative luminescence units (RLU) over time in 8-day-old Arabidopsis Col-0<sup>AEQ</sup> seedlings after treatment with (a)  $\beta$ -1,3/1,4-glucans of barley ( $\beta$ -glucan; B-GLU), 4% KOH fraction of *Equisetum arvense* cell wall (E-KOH) and lichenan from *Cetraria islandica* (C-LICH) (0.25 mg ml<sup>-1</sup> final concentration) untreated or treated with lichenase enzyme (+L). Chitohexaose 50  $\mu$ M was used as positive control. Undigested materials (0.25 mg ml<sup>-1</sup>), lichenase suspension (L) and distilled water (mock) were used as negative controls.

(b)  $\text{Ca}^{2+}$  influx after treatment with commercial MLGs (50  $\mu$ M) purified from barley (Megazyme). Data represent mean  $\pm$   $\sigma$  ( $n = 8$ ) in all panels.

(c) Structural scheme of the different MLG oligosaccharides used in the experiments. A  $\beta$ -D-cellobiosyl-(1,3)- $\beta$ -D-glucose (MLG43) structure is highlighted over the different structures. These data are from one representative experiment of at least three performed that gave similar results.

triggered  $\text{Ca}^{2+}$  influxes that were slightly higher than those activated by MLG43 and MLG34, whereas the  $\beta$ -1,3-linked trisaccharide (Lam3) was not active at the concentration



**Figure 2.** Purification of mixed-linked glucans (MLG) oligosaccharides from lichenase (L)-digested polysaccharides.

(a,c) Size exclusion chromatography elution profiles of 5 mg of lichenase digestion products from different  $\beta$ -1,3/1,4-glucans sources. (a) Barley ( $\beta$ -glucan; B-GLU) and (c) lichenan from *Cetraria islandica* (C-LICH). Thin-layer chromatography profiles of each of the fractions are shown overlaid. Markers indicate the migration of laminarin-oligosaccharides (Lam), cello-oligosaccharides (Cello) and MLG oligosaccharides (MLG).

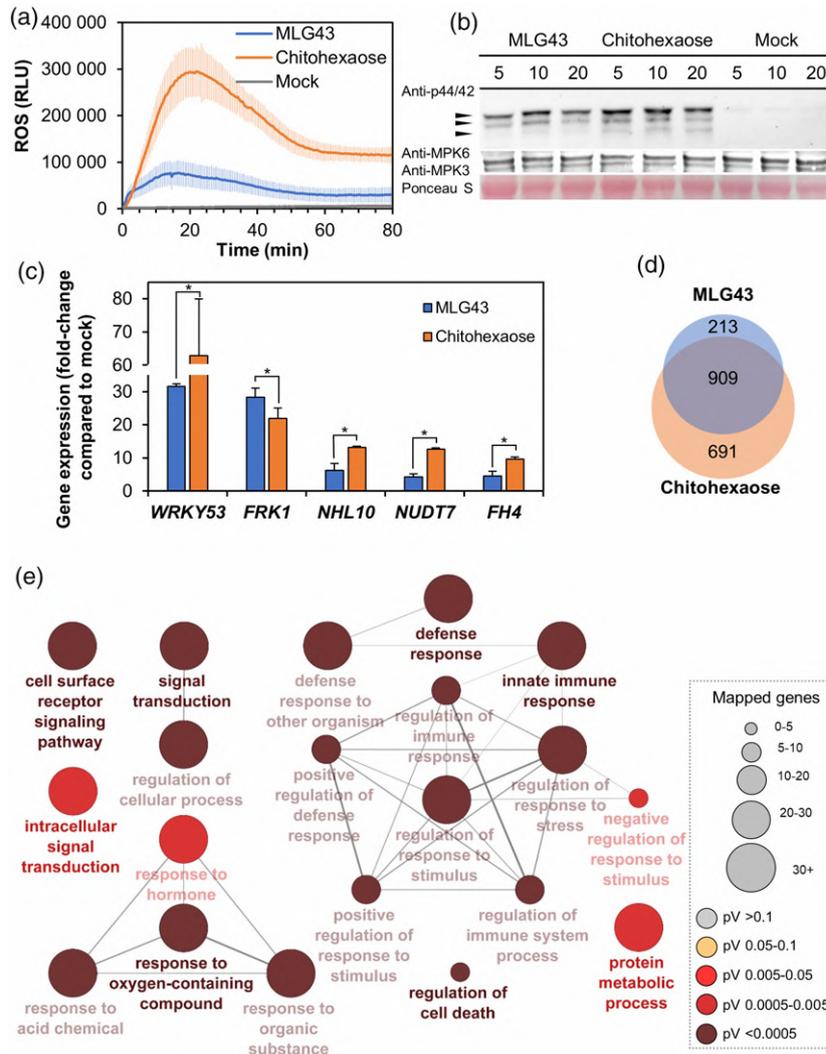
(b,d,e) High pressure anion exchange chromatography profiles of 5 mg of lichenase (L)-digested polysaccharides and commercial MLG oligosaccharides. (b) B-GLU + L and C-LICH + L; (d) MLG43 and MLG443. (e) Overlaid chromatograms comparing commercial MLG43 and purified MLG43 from lichenan (MLG43-L; pooled fractions 38–44 of the elution profile shown in c). Data shown in all panels are representative chromatograms from one experiment of at least 10 performed that gave similar results.

tested, as reported previously (Mélida et al. 2018; Figure S1). MLG43 used in these studies can be obtained after MLG (e.g. B-GLU) digestion with different enzymes, such as lichenases, whereas MLG34 must be obtained through cellulase digestion, which cleave  $\beta$ -1,4 bonds within the inner cello-oligosaccharide backbone of the MLG. As illustrated in the thin-layer chromatography (TLC) at Figure S2, while lichenase digestion products follow a pattern (MLG43, MLG443, MLG4443, etc.), more randomized oligosaccharide structures are released by cellulases, including  $\beta$ -1,4-linked (Cello2–Cello6) glucans, which have been described to trigger immune responses (Johnson et al., 2018). Based on these results, we selected the trisaccharide MLG43 as the minimal active MLG oligosaccharide (Figure 1b,c) and lichenase as the best enzyme to release this structure from MLG polymers.

Aiming to support the results obtained with oligosaccharides from commercial sources, we also purified MLG43 from C-LICH + L, B-GLU + L and E-KOH + L by size exclusion chromatography (SEC) (Figure 2; Figure S3). While B-GLU digestion yielded MLG43 and MLG443, the highest yield of MLG43 was obtained from *C. islandica* lichenan digestion as revealed by TLC analyses (MLG43-L in Figure 2a,c). MLG43-L clearly co-eluted with the respective commercial compound in a high-pressure anion exchange

(HPAE) chromatogram (Figure 2b,d,e). HPAE chromatography (HPAEC) data also corroborated that, while lichenase hydrolysis of B-GLU released a clearly noticeable amount of MLG443 tetrasaccharide in addition to the trisaccharide, lichenan structure clearly favoured the release of the trisaccharide (Figure 2c). In *Equisetum* fractions, according to SEC profiles, oligosaccharides of higher DP than the tetrasaccharide were released (Figure S3). These high DP oligosaccharides are compatible with hexasaccharides and nonasaccharides, previously described for this species (Simmons et al., 2013). Cross-elicitation experiments in Col-0<sup>AEQ</sup> seedlings, by subsequent application of two ligands in 600 sec interval, evidenced that MLG43-L and commercial MLG43 had a refractory period of Ca<sup>2+</sup> influx, indicating that both MLGs have equivalent activity (Figure S4). Dose-dependence and estimated effective dose (EED; 50% of total signal) of 265  $\mu$ M were determined for MLG43 on Arabidopsis seedlings using a concentration range between 200 nM and 5 mM (Figure S5). All these concentrations activated a Ca<sup>2+</sup> burst to a different extent, and based on Ca<sup>2+</sup> kinetics and EED value, 50  $\mu$ M was selected as an adequate final concentration for further experiments.

In addition, we decided to synthesize chemically the pure MLG structures to characterize the PTI activity of additional structures further. The analysis of Ca<sup>2+</sup> influxes



**Figure 3.** Pattern-triggered immunity hallmarks activation by mixed-linked glucans,  $\beta$ -D-cellobiosyl-(1,3)- $\beta$ -D-glucose (MLG43 50  $\mu$ M) in Arabidopsis. Chitohexaose (5  $\mu$ M (a,b) and 50  $\mu$ M for gene expression (c,d) and distilled water (mock) were used as controls in all the experiments. (a) Reactive oxygen species (ROS) production in Arabidopsis leaf-discs of 5-week-old Col-0 plants by Luminol reaction measured as relative luminescence units (RLU) over time. Data represent mean  $\pm$   $\sigma$  ( $n = 8$ ) from one experiment of three performed that produced similar results. (b) Mitogen-activated protein kinases (MAPK) phosphorylation in 12-day-old seedlings determined by western blot using anti-pTEpY antibody for phosphorylated MAPK moieties at different time points (5, 10 and 20 min). Arrows indicate the position of phosphorylated MPK6 (top), MPK3 (middle) and MPK4/11 (bottom). Ponceau red-stained membranes show equal loading. Data shown are from one experiment of the five performed that gave similar results. (c) Quantitative reverse transcription–polymerase chain reaction analysis in 12-day-old Arabidopsis seedlings. Relative expression levels to *UBC21* (*At5g25769*) gene at 30 min normalized to their expression levels in mock-treated seedlings were shown. Data represent mean  $\pm$   $\sigma$  of three technical replicates from two biological replicates ( $n = 6$ ) of three independent biological replicates analysed that gave similar results. Statistically significant differences between MLG43 and chitohexaose according to Student’s *t*-test ( $*P < 0.05$ ). (d) Venn diagram of shared overexpression between MLG43 and chitohexaose (both at 50  $\mu$ M). RNA-sequencing data were obtained from the combination of three biological replicates of 12-day-old Arabidopsis Col-0 plants at 30 min after treatment with MLG43 or chitohexaose. (e) Biological process Gene Ontology (GO) term enrichment map of the overexpressed genes in 50  $\mu$ M MLG43. GO term enrichment is expressed by node size. Enrichment *P*-value determined by enrichment/depletion (two-sided hypergeometric) test and corrected by the Bonferroni step down method is represented by colour scale. Only GO terms at  $P < 0.01$  are shown. Links between groups indicate shared genes ( $\kappa$  score level  $\geq 0.4$ ).

triggered by these pure, synthetic structures indicated that MLG oligosaccharides with DP3 and DP4 were not active (Figure S6a), probably due to the presence of an aminoalkyl linker at their reducing end, which is required for the synthesis process (Figure S6c; Bartetzko and Pfengle, 2019; Dallabernardina *et al.*, 2017) and that would affect the tri-dimensional structures of these synthetic

oligosaccharides and may interfere with their perception by PRRs. However, MLG oligosaccharides of higher DP (6 and 8) were active in triggering  $\text{Ca}^{2+}$  influxes (Figure S6b; MLG44434', MLG44434', MLG43344', MLG3434443'). Some of the immune active synthetic MLG structures tested cannot be generated through lichenase digestions, that generate MLGs with a single  $\beta$ -1,3 linkage placed next to the

reducing end, further indicating that 'non-canonical' MLG oligosaccharides containing the MLG43/MLG34 'signature' can be also recognized by plant cells.

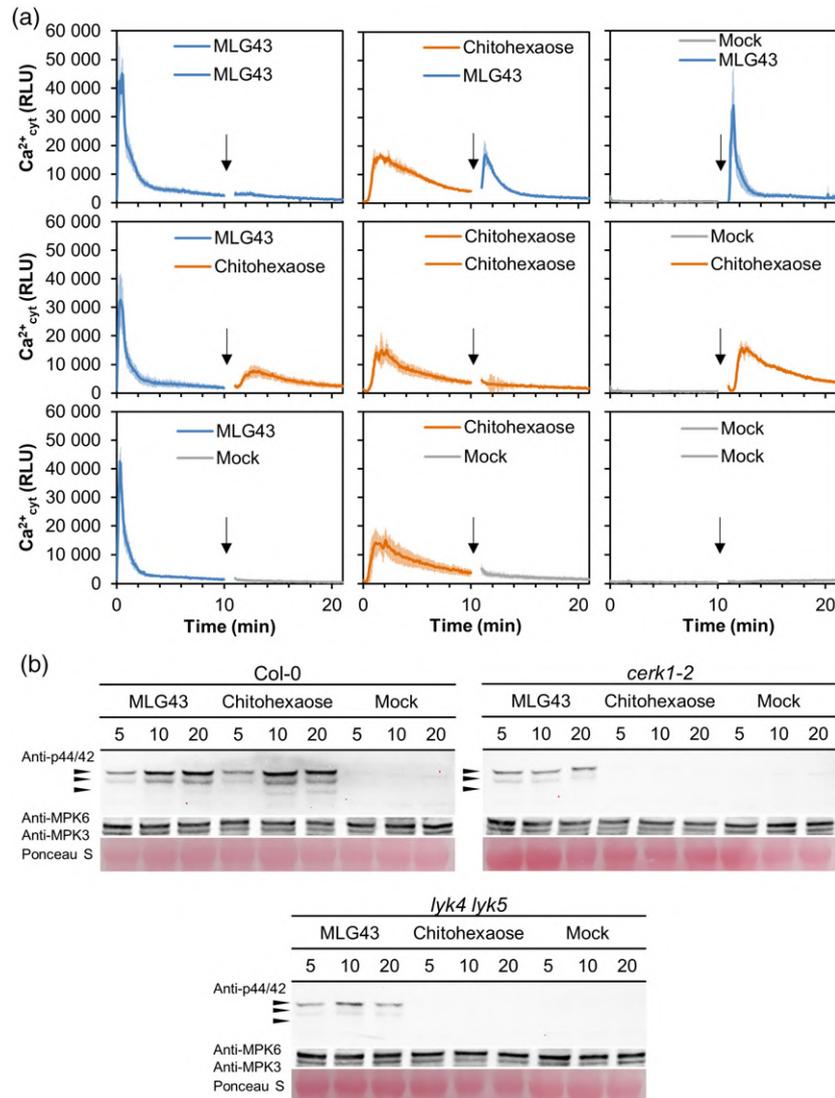
### MLG43 trisaccharide activate PTI hallmarks

Previous reports using luminol-based assays to quantify reactive oxygen species (ROS) production in leaf discs treated with a  $\beta$ -1,3-glucan hexasaccharide (laminarihexaose) or a  $\beta$ -1,4-glucan disaccharide (cellobiose) revealed an absence of ROS production, even after applying high concentrations (0.5–1 mM) of these glucans (Mélida *et al.*, 2018; Souza *et al.*, 2017). In contrast, we found that MLG43 triggered an ample ROS burst when applied at 50  $\mu$ M on Arabidopsis leaf discs, although this ROS burst was weaker than that triggered by chitohexaose (Figure 3a). To confirm PTI activity of MLG43, we next dissected the phosphorylation of protein kinases (MPK3/MPK6/MPK4/MPK11) and the upregulation of PTI-reporter genes upon treatment of Arabidopsis seedlings with MLG43 (Figure 3b,c). Western blotting showed MPK3 and MPK6 phosphorylation after MLG43 application (50  $\mu$ M) to Arabidopsis seedlings, with a phosphorylation peak at 10 min post-treatment (Figure 3b). MPK4/11 phosphorylation was almost undetected in MLG43-elicited plants, which contrasted with the observed phosphorylation upon chitohexaose treatment (Figure 3b). Expression of five PTI-marker genes upregulated by chitohexaose (*WRKY53*, *FRK1*, *NHL10*, *NUD7* and *FH4*; Mélida *et al.*, 2018) was assessed by quantitative reverse transcription–polymerase chain reaction (qRT-PCR), and the results indicated that expression of all these genes was similarly induced after MLG43 or chitohexaose elicitation in comparison with mock-treated seedlings (Figure 3c), suggesting that these compounds trigger similar transcriptional responses.

To characterize the global gene reprogramming triggered by MLG43 further, we performed transcriptomic analyses [RNA-Sequencing (RNA-seq)] of Arabidopsis seedlings treated for 30 min with MLG43 or chitohexaose (Figure 3d,e; Tables S1–S4). Incubation with MLG43 induced changes in the expression of 1229 genes, most of which (1122) were upregulated (Figure 3d; Table S1). On the other hand, treatments with chitohexaose resulted in 1988 genes whose expression levels were significantly altered, with 691 genes being upregulated exclusively by chitohexaose and 909 after treatment with both structures (Figure 3d; Figure S7, Tables S2 and S3). Gene Ontology (GO) classification of MLG43-induced genes showed that these mainly grouped in terms related to immune system processes, response to different stimuli, including biotic and abiotic stresses, signal transduction and cell surface receptors signalling pathways, among other GOs (Figure 3e), which are quite similar GOs to those found upon treatment with chitohexaose (Figure S7). These analyses indicate that MLG43-triggered responses are highly similar

to those induced by the well characterized MAMP chito-hexaose (Figure 3; Figure S7).

LysM-PRRs have been described in various species as co-receptors for glycan-based molecular patterns such as chitin, peptidoglycan,  $\beta$ -1,3-glucans (laminarins) and lipopolysaccharides (Desaki *et al.*, 2018; Mélida *et al.*, 2018; Miya *et al.*, 2007; Willmann *et al.*, 2011). In particular, the LysM-PRR CERK1 has a crucial role in glycan-based-MAMP perception. Given the high similarity in global gene reprogramming triggered by MLG43 and the CERK1-dependent ligand chitohexaose, we wondered whether refractory stages would exist between the application of these two glyco-ligands to Arabidopsis seedlings. Notably, cross-elicitation experiments in Col-0<sup>AEQ</sup> seedlings demonstrated the absence of such a refractory period, further suggesting that the mechanisms of perception of these glycans and the PRRs involved in their perception are not identical (Figure 4a). In addition to CERK1, the LysM-PRRs LYK5 and LYK4 have been involved in chitohexaose perception as receptor and co-receptor, respectively (Cao *et al.*, 2014; del Hierro *et al.*, 2020; Liu *et al.*, 2012). To characterize the molecular mechanisms of MLG43-mediated immunity further, phosphorylation of mitogen-activated protein kinase (MAPK) was tested by western blots in wild-type plants and *cerk1* single and *lyk4 lyk5* double mutants upon MLG43 treatment and we found that MAPK phosphorylation levels were weaker in these mutants than in wild-type plants, whereas they were fully impaired upon chito-hexaose treatment, as reported previously (Cao *et al.*, 2014; Figure 4b). These data indicate that MLG43 perception partially depends on CERK1, LYK5 and LYK4 PRRs, that might function as co-receptors, as previously described for the  $\beta$ -1,3-glucan elicitor laminarihexaose (Mélida *et al.*, 2018). In addition to LysM-PRRs, BAK1 and SOBIR1 are frequently involved as PRR co-receptors in the activation of signal transduction following perception of different DAMP/MAMPs (van der Burgh *et al.*, 2019; Perraki *et al.*, 2018). We tested phosphorylation of MAPK in wild-type plants and mutant lines impaired in BAK1 and SOBIR1, and we found that MAPK phosphorylation levels were similar in these mutants and wild-type plants upon MLG43 treatment (Figure S8), suggesting that MLG43 perception does not depend on BAK1 and SOBIR1. These data were in line with those obtained by determining protein–glycan interaction energies applying a molecular dynamics simulation methodology recently described (del Hierro *et al.*, 2020) that confronted optimized CERK1, BAK1 and SOBIR1 ECD structures and MLG43 in solvent boxes (Table S7). ECD-glycan  $\Delta G$  determinations resulted in positive or close to zero energy values, which indicate that no direct binding events between the ECDs of these PRRs and MLG43 took place during the molecular dynamics simulations (Table S7). Together, these analyses with the ECDs of main co-receptors indicated that MLG43 perception and



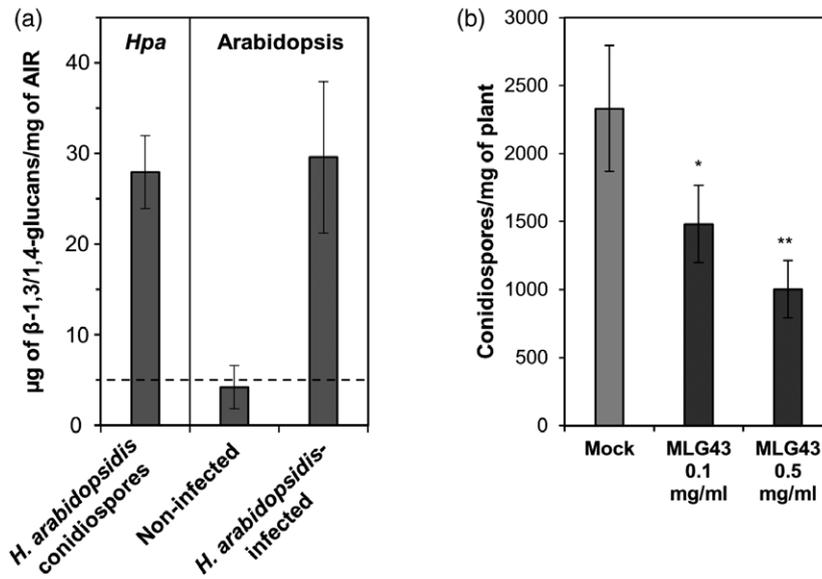
**Figure 4.** MLG43-triggered immunity in Arabidopsis is partially dependent on LysM-PRRs CERK1, LYK4 and LYK5. (a) Cross-elicitation during the refractory period of  $Ca^{2+}$  signalling upon application of 50  $\mu$ M MLG43, 50  $\mu$ M chitohexaose and distilled water (mock). Data show the elevation of cytoplasmic  $Ca^{2+}$  concentration, measured as relative luminescence units (RLU), over time in 8-day-old Arabidopsis Col-0<sup>AEQ</sup> seedlings after treatments. Arrow indicates the application time of the second treatment within the refractory period of the first elicitation. Data represent mean  $\pm$   $\sigma$  ( $n = 8$ ) in all panels. (b) Mitogen-activated protein kinases (MAPK) phosphorylation in 12-day-old Arabidopsis seedlings of Col-0 plants and *cerk1-2*, and *lyk4 lyk5* mutants impaired in LysM-PRR co-receptors. Western blot using anti-pTEpY antibody for phosphorylated MAPK moieties at different time points (5, 10 and 20 min). Black arrows indicate the position of phosphorylated MPK6 (top), MPK3 (middle) and MPK4/11 (bottom). Anti-MPK6 and anti-MPK3 were used as total protein control. Ponceau red-stained membranes show equal loading. Chitohexaose (5  $\mu$ M) and distilled water (mock) were used as controls. These results are from one representative experiment of the three performed that gave similar results.

signal transduction is mediated by a not yet characterized immune complex that does not involve BAK1 or SOBIR1 and that probably involves LysM-PRRs as redundant co-receptors.

#### MLG43 pre-treatments diminish plant disease symptoms caused by inoculations of pathogens

MLGs are undoubtedly present in the wall of several plants, algae, lichen-forming ascomycete symbionts, fungi

and bacteria, but most likely not in dicot plant species such as Arabidopsis (Burton *et al.*, 2006; Zablackis *et al.*, 1995). Several lines of evidence also point to the presence of MLGs in the cell wall of several plant pathogens harbouring glucan-rich extracellular envelopes such as oomycetes, fungi and some bacteria, but the presence of MLGs in these organisms has been reported only in a few cases (Fontaine *et al.*, 2000; Lee and Hollingsworth, 1997; Mélida *et al.*, 2013; Pérez-Mendoza *et al.*, 2015; Pettolino *et al.*,



**Figure 5.** Mixed-linked glucans are components of the cell walls of *Hyaloperonospora arabidopsidis* and trigger Arabidopsis disease resistance to this oomycete pathogen.

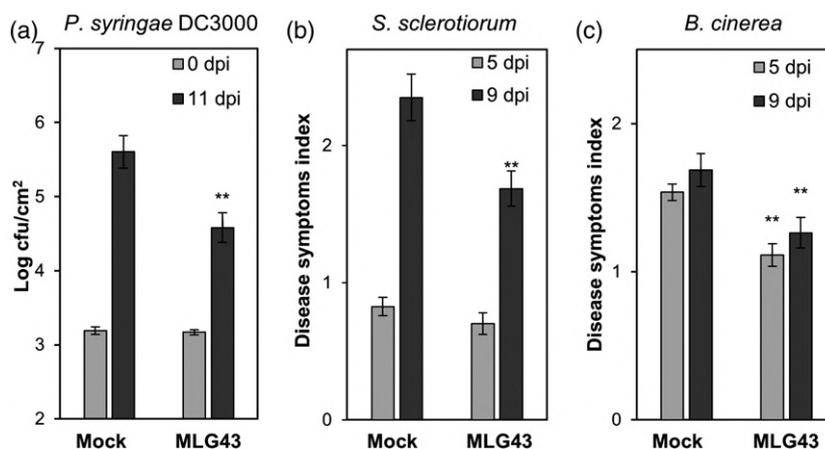
(a)  $\beta$ -1,3/1,4-glucan quantification in alcohol insoluble residues (AIRs) of *H. arabidopsidis* (*Hpa* isolate Noco2) conidiospores and of Arabidopsis Col-0 plants non-infected (mock-treated) or infected with *H. arabidopsidis* (Noco2) 7 days post-inoculation (dpi) with conidiospores (24-day-old Arabidopsis plants). Inoculated plants were extensively washed before AIR preparations in order to release *H. arabidopsidis* conidiospores from plant tissues. Data represent average  $\pm$   $\sigma$  ( $n = 3$ ). Dashed line indicates the detection limit of the method used.

(b) Arabidopsis plants were foliar pre-treated with MLG43 at two different concentrations 2 days before inoculation with *H. arabidopsidis* (Noco2). Presence of *H. arabidopsidis* in plants was quantified at 7 dpi (24 days old) as the abundance of conidiospores in inoculated per mg plant fresh weight. Data represent mean  $\pm$  SE ( $n = 30$ ). Statistically significant differences according to the Student's *t*-test (\* $P < 0.05$ ; \*\* $P < 0.01$ ).

2009; Samar *et al.*, 2015). We searched for the presence of MLGs in the cell wall of the Arabidopsis oomycete pathogen *Hyaloperonospora arabidopsidis*, belonging to the Peronosporales order, such as the *Phytophthora* species, that contain up to 85% of glucans in their cell walls (Mélida *et al.*, 2013). We obtained alcohol insoluble residues (AIR; equivalent to partially purified cell walls) from *H. arabidopsidis* Noco2 conidiospores, and found that these AIRs contained on average 27.9  $\mu$ g of MLGs per mg of dry weight (Figure 5a). Next, we purified AIR fractions from non-inoculated and *H. arabidopsidis*-inoculated Arabidopsis plants, which were extensively washed, before mechanical disruption for AIR preparation, to remove *H. arabidopsidis* sporangiophores and those conidiospores from spray inoculation of the plants that had not germinated from leaves surface. Interestingly, we found that only AIR from *H. arabidopsidis*-inoculated plants contained MLGs (29.6  $\mu$ g per mg of AIR dry weight; Figure 5a), which would necessarily derive from intracellularly grown *H. arabidopsidis* hyphae inside Arabidopsis leaves, as the conidiospores had been washed out. These data showed that, at least in the Arabidopsis–*H. arabidopsidis* pathosystem, plant cells are exposed to this oomycete's MLGs. In sight of these data, we tested whether pre-treatment with MLG43 of Arabidopsis Col-0 wild-type plants before infection with the oomycete would improve Arabidopsis resistance to the

compatible and virulent Noco2 isolate. Of note, we observed a reduction of up to 60% in conidiospore  $\text{mg}^{-1}$  of plant fresh weight at 7 days post-inoculation (dpi) in plants pre-treated with MLG43 in comparison with untreated plants (Figure 5b). These data would favour the classification of MLGs as MAMPs triggering PTI in Arabidopsis.

Next, we asked whether part of the knowledge gained using the model species Arabidopsis could be translated to crops such as tomato and pepper. In a first instance, we evaluated the MLG43-protection capacity by pre-treating tomato plants with the trisaccharide 2 days before challenging them with the bacterium *Pseudomonas syringae* pv. *tomato* DC3000. Notably, bacterial growth was significantly reduced in the MLG43-pre-treated tomato plants at 11 dpi compared with mock-treated plants (Figure 6a). Moreover, MLG43 treatment led to the upregulation of two tomato PTI-marker genes, *SIWRKY53* and *SIPTI5* (Liu *et al.*, 2019) (Figure S9). A similar approach was followed in pepper plants that were pre-treated with MLG43 2 days before inoculation with the necrotrophic fungi *Sclerotinia sclerotiorum* or *Botrytis cinerea*. MLG43-treated pepper plants showed, in comparison with mock-treated plants, a reduction in their disease symptoms index (Figure S10) at 9 dpi with *S. sclerotiorum* (Figure 6b), and at 5 and 9 dpi with *B. cinerea* (Figure 6c). The activation of MLG43-triggered immune responses in Arabidopsis, tomato and



**Figure 6.** MLG43 pre-treatment confers enhanced disease resistance against bacterial and fungal pathogens to tomato and pepper plants. Plants were sprayed with MLG43 ( $0.25 \text{ mg plant}^{-1}$ ) 2 days before pathogen challenge.

(a) Colony forming units (cfu) of *Pseudomonas syringae* pv. *tomato* DC3000 per leaf area at 0 and 11 days post-inoculation (dpi) in tomato plants. Data represent mean  $\pm$  SE ( $n = 8$ ).

(b) Disease symptoms index produced by *Sclerotinia sclerotiorum* at 5 and 9 dpi in leaves of pepper plants. Data represent mean  $\pm$  SE ( $n = 24$ ).

(c) Disease symptoms index produced by *Botrytis cinerea* at 5 and 9 dpi in leaves of pepper plants. Data represent mean  $\pm$  SE ( $n = 12$ ). Disease indexes (0–5) examples for (b) and (c) are shown in Figure S10. Statistically significant differences according to the Student's *t*-test (\* $P < 0.05$ ; \*\* $P < 0.01$ ). All the disease experiments were performed at least four times and one representative experiment is shown.

pepper suggests that these species have the PRRs and co-PRRs required for MLG perception and PTI activation.

## DISCUSSION

Cell walls are dynamic and highly controlled structures that are the first point of contact during a plant–microbe interaction (Bacete *et al.*, 2018, 2020; Geoghegan *et al.*, 2017; Lampugnani *et al.*, 2018; Rui and Dinneny, 2020). The evolutionary arms race has provided plants and their microbial interactors with a large collection of cell wall-degrading enzymes to shoot down the opponent's wall (Rovenich *et al.*, 2016). Therefore, cell walls are rich sources of carbohydrate-based defence signalling molecules (DAMPs and MAMPs), which remain poorly characterized. Here, we demonstrate that some structures contained in MLGs are perceived as molecular patterns by different plant species triggering immune responses. In particular, we identified a minimal structure, the trisaccharide MLG43, which triggers several PTI-hallmarks in *Arabidopsis* at low concentrations. Indeed, the transcriptomic study showed that MLG43 and the well-known MAMP chitohexaose activate transcriptional responses that share almost 81% of the overexpressed genes, further demonstrating a very high gene reprogramming overlap upon treatment with these glycoligands. These data are in line with previous transcriptomic analyses comparing the *Arabidopsis*-responsive genes to single-linked  $\beta$ -1,3-glucans (DAMP/MAMP) and  $\beta$ -1,4-glucans (DAMPs) and chitin (MAMP) that revealed that a significant percentage of mis-expressed genes were shared between glucans and chitin treatments (Johnson *et al.*, 2018; Mélida *et al.*, 2018; Souza *et al.*, 2017).

Structurally, the most similar oligosaccharides to MLG43 are glucans with solely  $\beta$ -1,3 or  $\beta$ -1,4 linkages (Stone and Clarke, 1992). Cellulose-derived oligomers ( $\beta$ -1,4-glucans) are plant cell wall-derived DAMPs that trigger signalling cascades sharing some similarities with the responses triggered by chitin and the well-characterized DAMP oligogalacturonides derived from plant pectic polysaccharides (Aziz *et al.*, 2007; Benedetti *et al.*, 2015; Ferrari *et al.*, 2013; Johnson *et al.*, 2018; Souza *et al.*, 2017). Cello-oligomers (DP3 or higher) and MLG43, as well as chitin, are active on plants at nanomolar concentrations (Johnson *et al.*, 2018; Kaku *et al.*, 2006), which are significantly lower than those required to trigger PTI responses by  $\beta$ -1,3-glucans and xyloglucan (containing a  $\beta$ -1,4-linked glucose backbone), which are in the high micromolar range (Aziz *et al.*, 2003; Claverie *et al.*, 2018; Klarzynski *et al.*, 2000; Mélida *et al.*, 2018). Interestingly, in spite of the high  $\beta$ -1,3-glucans and xyloglucans doses required to be perceived by plants, they were able to improve protection in several pathosystems. For instance, xyloglucan increased grapevine and *Arabidopsis* resistance against the fungus *B. cinerea* or the oomycete *H. arabidopsidis*, respectively, while  $\beta$ -1,3-glucans improved, among others, tobacco and grapevine protection against bacterial (*Erwinia carotovora*), fungal (*B. cinerea*) and oomycete (*Plasmopara viticola*) pathogens (Aziz *et al.*, 2003; Claverie *et al.*, 2018; Klarzynski *et al.*, 2000). Given the high abundance of  $\beta$ -1,3-glucans in the cell walls of brown seaweed, laminarin-based products have been successfully developed to be used in agriculture as activators of plant natural defence against pathogens. Similarly, pre-treatments with the single-linked  $\beta$ -1,4-glucan cellobiose reduced

*P. syringae* growth on Arabidopsis seedlings, although high doses were required to observe such an effect (Souza *et al.*, 2017).

Notably, we show here that a glucan trisaccharide combining both  $\beta$ -1,3 and  $\beta$ -1,4 linkages (MLG43) is also perceived by Arabidopsis and some crops (tomato and pepper) at lower concentrations than the single-linked counterparts  $\beta$ -1,3- and  $\beta$ -1,4-glucans. Indeed, pre-treatments of Arabidopsis, tomato and pepper with MLG43 before pathogen inoculation confers enhanced disease resistance and significant protection against oomycete, bacterial and fungal pathogens, supporting that different plant species in addition to Arabidopsis harbour the PRRs required for recognition of the MLGs. This perception, at least in Arabidopsis, seems to be independent of the co-receptors BAK1 and SOBIR1, and only partially dependent on CERK1, LYK4 and LYK5 LysM-PRRs, contrary to chitin and  $\beta$ -1,3-glucans, which were fully impaired in the *cerk1-2* mutant (Cao *et al.*, 2014; Liu *et al.*, 2012; Mérida *et al.*, 2018). Despite these differences in PRR complexes involved in the perception of these carbohydrate-based MAMPs, they transcriptionally activate a very high similar downstream gene reprogramming, as reported in other MAMPs comparisons (Bjornson *et al.*, 2020).

Our data support that MLGs represent a group of molecular patterns perceived by plants, but it raises now several important biological questions about the different source of MLGs that plants are exposed to during plant/pathogen interactions, and the mechanisms of MLG recognition by plants. Notably, some plant phylogenetic groups, such as Poaceae species, harbour MLGs in their cell walls that can release immune-active MLGs (MLG43) upon the catalytic breakdown during infection by fungal/oomycete cell wall degrading enzymes, such as cellulases, and accordingly MLG43 could be classified as a self-molecular pattern or DAMP. However, other plant species, including Arabidopsis, tomato and pepper used in this work do not seem to contain MLGs in their cell walls, but we show here that they are able to perceive MLG43 and to trigger immune responses. In these cases, MLG43 will be perceived as non-self-molecular pattern or MAMP. We show here that the oomycete *H. arabidopsidis* harbours MLGs in the cell walls of conidiospores or intracellularly growing hyphae in Arabidopsis-infected plants. This finding is in line with previous reports indicating the presence of MLGs in the cell wall of several microorganisms, although these glucan structures remain poorly characterized (Fontaine *et al.*, 2000; Lee and Hollingsworth, 1997; Mérida *et al.*, 2013; Pérez-Mendoza *et al.*, 2015; Pettolino *et al.*, 2009; Samar *et al.*, 2015). There are many groups of microbes whose cell wall contains high proportions of glucans, but as MLG constituent units (1,4- and 1,3-linked glucosyl residues) can be the building blocks of other better microbial characterized polymers (e.g. laminarin, glycogen, cellulose), MLG

presence has probably been underestimated to date. For instance, plant pathogens such as oomycetes *Phytophthora infestans* and *Phytophthora parasitica* contain over 85% of glucans in their cell walls, including a high proportion of both 1,4- and 1,3-linked glucosyl units (Mérida *et al.*, 2013). However, the presence of polymers combining both types of linkages (MLGs) has not been investigated in detail yet. A good example illustrating this is that the cell wall composition of the closely related oomycete *H. arabidopsidis* is unknown, but we have demonstrated here that it contains at least a 3% of MLGs in extracted AIR wall fraction. The enhanced resistance to *H. arabidopsidis* of Arabidopsis plants pre-treated with MLG43 clearly demonstrate that MLGs are perceived as MAMPs. In this sense, further work unveiling perception mechanisms and the specific immune pathways triggered by MLGs in different species will help to decipher their biological functions further.

In conclusion, our data expand the current knowledge of the diversity of glycan-based molecular patterns recognized by plant immune systems, support the use of them as products for the modulation of crop immunity and anticipate that their application in agriculture could help towards the transition to a more sustainable agriculture.

## EXPERIMENTAL PROCEDURES

### Plant growth conditions

Col-0 background lines were used for all the Arabidopsis experiments in the present work. Arabidopsis seedlings used for  $\text{Ca}^{2+}_{\text{cyt}}$  (Col-0<sup>AEQ</sup>), MAPK phosphorylation and gene expression analyses were grown in liquid MS medium and plants in soil-vermiculite (3:1) under 10 h light/14 h dark conditions at 21–20°C (Mérida *et al.*, 2018). Tomato plants (*Solanum lycopersicum*, MoneyMaker) were grown in a greenhouse in soil-vermiculite (3:1) under 14 h of light/10 h of dark at 24–22°C. Pepper plants (*Capsicum annuum*, Murano) were grown in a greenhouse in soil-vermiculite (3:1) under 14 h of light/10 h of dark at 24–19°C.

### Carbohydrates used in the experiments

B-GLU (#P-BGBL), MLG43, MLG34, MLG443, MLG434 + 344 and hexaacetyl-chitohexaose (chitohexaose;  $\beta$ -1,4-D-(GlcNAc)<sub>6</sub>; #O-CHI6) were acquired from Megazyme (Wicklow, Ireland). Lichenan from *Cetraria islandica* (#GLU602) was purchased from Elicityl. *Equisetum arvense* raw materials were kindly provided by Bio-search Life (Granada, Spain; #COPMCOLO001). MLG34', MLG434', MLG44343', MLG44434', MLG43344', MLG34443' and MLG3434443' were chemically synthesized using automated glycan assembly as previously described (Dallabernardina *et al.*, 2017). More details can be found in Table S5.

### Preparation and digestion of $\beta$ -1,3/1,4-glucan polysaccharides and oligosaccharides purification

*Equisetum arvense* raw materials were fine-powdered using a kitchen blender and extracted with MeOH/CHCl<sub>3</sub> (1:1) four times during 4 h at 4°C. A vacuum pump filtration was used to separate the soluble fraction after each step. Soluble fractions were

discarded, and the insoluble residues were extracted with 70% (v/v) ethanol twice (1 h at 90°C and overnight at room temperature). The residue after filtration was then treated with distilled water twice (1 h at 90°C and overnight at room temperature). The residue after filtration was considered AIR. AIR polysaccharides were chemically extracted using 4% (w/v) KOH (E-KOH fraction) as previously described (Mélida *et al.*, 2009). B-GLU, E-KOH fraction from *E. arvense* and lichenan (C-LICH) were suspended in distilled water (5 mg ml<sup>-1</sup>) containing 1.4 U ml<sup>-1</sup> of lichenase (+L; EC 3.2.1.73; Megazyme) and stirred 72 h at 60°C. Digestion products were freeze-dried and fractionated by SEC (140 cm<sup>3</sup> bed volume in a 1.6 cm diameter column; Biogel P2 Extrafine; Bio-Rad, Hercules, CA, USA) (Mélida *et al.*, 2020). Total carbohydrates in each fraction were quantified by phenol-sulphuric acid method (Dubois *et al.*, 1956).

### $\beta$ -1,3/1,4-glucans content determination

For the determination of MLGs in *H. arabidopsidis* conidiospores, these were recovered from 24-day-old Arabidopsis plants inoculated ( $4 \times 10^4$  conidiospores ml<sup>-1</sup>) 7 days before tissue harvesting. Plant tissues were extensively washed with distilled water and a conidiospore suspension was collected. The suspension was centrifuged for 10 min at 5000 *g* to obtain a conidiospore pellet, upon discarding both supernatant and green upper layer of the pellet. Conidiospore pellets were homogenized and then extracted three times with 80% (v/v) ethanol for 1 h, overnight and 1 h. Air-dried pellets after acetone washings were considered as AIRs (Pettolino *et al.*, 2009). To determine MLGs in Arabidopsis plants inoculated ( $4 \times 10^4$  conidiospores ml<sup>-1</sup>) or mock-inoculated with *H. arabidopsidis*, plants were extensively washed with water to release conidiospores from inoculated plants, and then plant tissues were immediately frozen with liquid nitrogen, and AIRs were prepared as described by Bacete *et al.* (2017). MLGs were quantified from the different AIR materials using a Mixed Linkage  $\beta$ -Glucan Assay Kit (Megazyme; #K-BGLU).

### Carbohydrate analyses

Oligosaccharides were analysed by TLC and HPAEC. MLGs (5  $\mu$ g) were spotted onto TLC plates (Silicagel 60; Merck, Darmstadt, Germany) and run twice using 1-propanol/ethyl-acetate/water (9:7:4 by volume). TLC plates were developed by using the thymol-H<sub>2</sub>SO<sub>4</sub> method (Mélida *et al.*, 2020). HPAEC separations were performed using a CarboPac PA-200 anion exchange column (4.6  $\times$  250 mm; Dionex, Oakville, ON, Canada) mounted on a Dionex ICS 3000 HPAEC-PAD system and a pulsed amperometric detector. Oligosaccharides were eluted at 0.5 mL min<sup>-1</sup> using a linear saline gradient of 30 mM NaOH to 30 mM NaOH/300 mM sodium acetate over 16 min and equilibration at the initial conditions for 5 min.

### Aequorin luminescence measurements

Eight-day-old Arabidopsis 'aequorin-plants' (Col-0<sup>AEQ</sup>; Ranf *et al.*, 2012) were used for Ca<sup>2+</sup> influxes measurements (Bacete *et al.*, 2017). Dose–response curves and EED were calculated using total relative luminescence unit values (areas under kinetic curves) (Mélida *et al.*, 2018).

### ROS

H<sub>2</sub>O<sub>2</sub> production upon elicitation was monitored on 4-mm diameter leaf discs carefully obtained from 5-week-old Arabidopsis plants by using the luminol-peroxidase method (Escudero *et al.*, 2017).

### Immunoblot analysis of MAPK activation

Arabidopsis seedlings (12-day-old) were treated with different oligosaccharides and distilled water (mock) for 0, 5, 10 and 20 min, and fast-frozen with liquid nitrogen. Seedlings were homogenized using a FastPrep Bead Beating System (MP Biomedicals, Santa Ana, CA, USA) in extraction buffer (50 mM Tris-HCl pH 7.5, 200 mM NaCl, 1 mM EDTA, 10 mM NaF, 2 mM sodium orthovanadate, 1 mM sodium molybdate, 10% (v/v) glycerol, 0.1% (v/v) Tween-20, 1 mM 1,4-dithiothreitol, 1 mM phenylmethylsulfonyl fluoride and phosphatase inhibitor cocktail #P9599; Sigma-Aldrich, St. Louis, MO, USA). Total protein extracts were quantified by the Bradford assay (Bio-Rad). Proteins (40  $\mu$ g) were separated using 10% Mini-PROTEAN TGX Pre-cast protein gels and transferred to nitrocellulose membranes using the Invitrogen iBlot Gel Transfer Device. Membranes were blocked with Protein-Free Blocking Buffer [Tris-buffered saline (TBS); Thermo Fisher Scientific, Waltham, MA, USA] for 2 h at room temperature. Membranes were incubated overnight at 4°C in TBS containing phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204) antibody (Cell Signaling Technology, Danvers, MA, USA) (1:1000) or anti-AtMPK3 (1:2500) and anti-AtMPK6 (1:10 000) antibodies (Sigma-Aldrich). Membranes were washed with TBS containing 0.1% Tween-20 and incubated with horseradish peroxidase-conjugated anti-rabbit antibody (GE Healthcare, Chicago, IL, USA) (1:5000) in TBS. Blots were finally developed using the ECL western blotting substrate (Thermo Fisher Scientific) and imaged using an iBright FL1000 Image System (Thermo Fisher Scientific). Membranes were also stained with Ponceau-S Red (Sigma-Aldrich).

### Gene expression analyses

Twelve-day-old Arabidopsis seedlings were treated with different oligosaccharides and distilled water (mock) for 0 and 30 min and used for qRT-PCR and RNA-seq gene expression analysis. qRT-PCR analyses were performed as described by Mélida *et al.* (2020). Oligonucleotides used in the analysis are shown in Table S6.

RNA-seq analyses were performed by sequencing and analysing three biological replicates for each treatment as previously described (Mélida *et al.*, 2020). RNA-seq read data can be retrieved from the NCBI Sequence Read Archive (SRA) under BioProject accession ID PRJNA625401 (BioSample accession SAMN15682114). Significant ( $P < 0.05$ ) enrichments were determined using the hypergeometric test with Bonferroni step down correction. To determine differentially expressed genes, *t*-tests were performed for the treatments against mock values. N-fold  $\geq 2$  was used to prove upregulation and an n-fold  $\leq 0.5$  was applied to select downregulated genes. ClueGO 2.5.6 app for Cytoscape was used to determine which GO categories were statistically overrepresented in the differentially expressed set of genes.

### Molecular dynamics simulations

Geometries for MLG43-PRRs complexes were firstly obtained with docking calculations using the crystal structures of AtCERK1-ECD (PDB code: 4ebz), AtBAK1-ECD (PDB code: 4mn8) and AtSOBIR1-ECD (PDB code: 6rih). MLG43 ligand was built and optimized in vacuum as described (del Hierro *et al.*, 2020). Energy terms contributing to protein–ligand interactions were computed using molecular mechanics following a previously established pipeline (del Hierro *et al.*, 2020).

### Crop protection assays

For *H. arabidopsidis* experiments, Arabidopsis plants were grown in soil as indicated above but at a higher humidity (75%). Two-

week-old *Arabidopsis* plants were treated by foliar spray using 0.1 ml of MLG43 solution in water (0.1 mg ml<sup>-1</sup> or 0.5 mg ml<sup>-1</sup>). Two days after treatment, plants were spray-inoculated with 0.1 ml conidiospore suspension (4 × 10<sup>4</sup> conidia ml<sup>-1</sup>) of a *H. arabidopsidis* isolate Noco2. *H. arabidopsidis* level of infection in inoculated plants was quantified at 7 dpi (24-day-old *Arabidopsis* plants) as the abundance of conidiospores per mg of plant fresh weight. Conidiospores were recovered from inoculated plants by extensively washing them with distilled water, and released conidiospores in water suspensions were counted using a Neubauer chamber, and relativized to mg of plant fresh weight.

Tomato plants (*S. lycopersicum*, Moneymaker) were grown in a greenhouse in soil-vermiculite (3:1) under 14 h of light/10 h of dark at 24–22°C. Three-week-old plants were sprayed with 2 ml of a MLG43 solution (0.125 mg ml<sup>-1</sup>) containing 2.5% UEP-100 (Croda, Snaith, UK) and 2.5% Tween 24 MBAL (Croda) as adjuvants. Adjuvant solutions were used as mocks. Two days after treatments plants were challenged with *Pseudomonas syringae* pv. *tomato* DC3000 as described by Santamaría-Hernando *et al.*, (2019). Tomato leaf discs were collected from four different plants at 0 and 11 dpi and colony forming units (cfu) per foliar area were determined as described (Mélida *et al.*, 2020).

For *S. sclerotiorum* and *B. cinerea* experiments, pepper plants (*C. annuum*, Murano) were grown in a greenhouse in soil-vermiculite (3:1) under 14 h of light/10 h of dark at 24–22°C. Five-week-old plants were treated using 5 ml of a MLG43 solution (0.05 mg ml<sup>-1</sup>; for *S. sclerotiorum* experiments) or 2 ml of a MLG43 solution (0.125 mg ml<sup>-1</sup>; for *B. cinerea* experiments) containing 0.5% UEP-100 and 0.05% Tween 24 MBAL as adjuvants. Adjuvant solutions were used as mocks. Two days after treatment, plants were spray-inoculated with 5 ml of 250 colony forming units ml<sup>-1</sup> suspension of *S. sclerotiorum* homogenized mycelia according to Chen and Wang (2005) or with 3 ml of Gomborg's B5 medium containing 10<sup>6</sup> *B. cinerea* conidia (Benito *et al.*, 1998) and moved to a greenhouse high-humidity (75%) chamber. Disease symptoms were determined at 5 and 9 dpi in the first eight leaves of each plant (*n* = 24 plants for *S. sclerotiorum* and *n* = 12 for *B. cinerea*) using a scale from 0 to 5 where 0 = no symptoms; 1 = little necrotic spots (<10% of leaf area); 2 = two or more notable necrotic spots (10–25% of leaf area); 3 = big necrotic area (25–50% of leaf area); 4 = >50% of leaf area affected and 5 = leaf senescence. Representative images of the disease index scales used to evaluate both pathosystems are shown in Figure S10. All the disease experiments were performed at least four times and one representative experiment is shown.

## ACKNOWLEDGEMENTS

This work was supported by grants IND2017/BIO-7800 of the Comunidad de Madrid Regional Government, BIO2015-64077-R of the Spanish Ministry of Economy and Competitiveness (MINECO), RTI2018-096975-B-I00 of Spanish Ministry of Science, Innovation and Universities, to AM. This work has been also financially supported by the 'Severo Ochoa Programme for Centres of Excellence in R&D' from the Agencia Estatal de Investigación of Spain (grant SEV-2016-0672 (2017–2021) to the CBGP). Within the framework of this program HM was supported with a postdoctoral fellow. DR was the recipient of an Industrial PhD Fellow (IND2017/BIO-7800), IdH was the recipient of a PhD FPU fellow from the Spanish Ministry of Education (FPU16/07118) and ASV was the recipient of the RYC2018-025530-I grant of Spanish Ministry of Science, Innovation and Universities. FP thanks the Max Planck Society and the German Research Foundation (DFG, Emmy Noether program PF850/1-1 to FP) for financial support.

## CONFLICT OF INTEREST

The authors declare that they have no competing interests.

## AUTHOR CONTRIBUTIONS

HM and AM initiated, conceived and coordinated all the experiments except those related to *S. sclerotiorum* and *B. cinerea* which were conceived and initiated by DR, RP and FB. DR performed the experiments described in Figures 1–6 and Figures S1–S8 and S10 with help of GL, LB and HM. IdH performed the RNA-seq data analysis and the experiments described in Table S7. FV collaborated in HPAEC experiments. PD and FP performed the MLG oligosaccharides synthesis. LJ collaborated in *P. syringae* experiments and performed those shown in Figure S9. AS-V collaborated in experiments shown in Figures 5 and 6. DR and HM prepared the tables and figures. HM and AM wrote the paper. FV, FP, LJ and FB edited the paper.

## DATA AVAILABILITY STATEMENT

RNA-seq raw data can be obtained from BioProject PRJNA625401 (BioSample accession SAMN15682114) at the NCBI Sequence Read Archive (SRA) site. Materials and data are available upon request to corresponding authors.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

**Figure S1.** Cytoplasmic calcium elevation triggered by dimer and trimer mixed-linked glucans. Calcium influx measured as relative luminescence units (RLU) over time in 8-day-old *Arabidopsis* Col-0<sup>AEO</sup> seedlings after treatment with 50 μM of β-D-cellobiosyl-(1,3)-β-D-glucose (MLG43), β-D-glucosyl-(1,3)-β-D-cellobiose (MLG34), cellobiose (Cello2), cellotriose (Cello3) and laminaribiose (Lam2) and laminaritriose (Lam3). Data represent mean ± σ (*n* = 8). These results are from one representative experiment out of the three performed that gave similar results.

**Figure S2.** Thin-layer chromatography of enzymatic hydrolysis of different mixed-linked glucans sources with either lichenase (EC 3.2.1.73) or cellulase (EC 3.2.1.4): barley β-glucan (B-GLU), lichenan from *Cetraria islandica* (C-LICH) and 4% KOH fraction of *Equisetum arvense* cell wall (E-KOH). Enzyme dilution used and undigested substrates were loaded as controls. Markers at the right side indicate the migration of glucose, laminarin-oligosaccharides (Lam), cello-oligosaccharides (Cell) and MLG oligosaccharides (MLG).

**Figure S3.** Size exclusion chromatography elution profiles of lichenase digestion products (+L) from different β-1,3/1,4-glucans sources. (a) Barley β-glucan (B-GLU), (b) lichenan from *Cetraria islandica* (C-LICH) and (c) 4% KOH fraction of *Equisetum arvense* cell wall (E-KOH) were digested with lichenase (+L). Five mg of the digested fractions were loaded in the chromatography column. Data shown in all panels are representative chromatograms from one experiment of at least 10 performed that gave similar results.

**Figure S4.** Cross-elicitation during the refractory period of calcium signalling upon application of 50 μM of commercial MLG43 (Megazyme), purified MLG43 (MLG43-L) or distilled water (mock). Data show the elevation of cytoplasmic calcium concentration,

measured as relative luminescence units (RLU), over time in 8-day-old Arabidopsis Col-0<sup>AEQ</sup> seedlings after treatments. Arrow indicates the application time of the second elicitor within the refractory period of the first elicitation. Data represent mean  $\pm$   $\sigma$  ( $n = 4$ ) of a representative experiment of the three independent experiments performed that gave similar results.

**Figure S5.** MLG43 dose response analyses. (a) Increase of cytoplasmic calcium concentration in 8-day-old Arabidopsis Col-0<sup>AEQ</sup> seedlings measured as relative luminescence units (RLU) over time by increasing MLG43 concentrations (from 200 nM to 5  $\mu$ M). (b) Dose dependence of total cytoplasmic calcium influxes in Arabidopsis Col-0<sup>AEQ</sup> seedlings upon treatments with increasing MLG43 concentrations. Calcium saturation curves were adjusted by using Prism 6 Software, and the upper and lower lines represent the 95% confidence intervals. Arrow indicates the MLG43 estimated effective dose (EED = 265  $\mu$ M).

**Figure S6.** Calcium influx kinetics triggered by synthetic mixed-linked glucans. Calcium influx measured as relative luminescence units (RLU) over time in 8-day-old Arabidopsis Col-0<sup>AEQ</sup> seedlings after treatment with 50  $\mu$ M of different synthetic MLGs (a, b) or chitohexaose (a). (c) Structural scheme of the different synthetic MLG oligosaccharides used in the experiments. The oligosaccharide linker of the reducing end of the MLGs tested is shown. These data are from one representative experiment of at least three performed that gave similar results.

**Figure S7.** Biological process Gene Ontology (GO) term enrichment map of the overexpressed genes in 12-day-old Arabidopsis Col-0 plants treated MLG43 or chitohexaose. RNA-seq data were obtained from the combination of three biological replicates. (a) Common overexpressed genes 30 min after 50  $\mu$ M MLG43 or chitohexaose treatments. (b) Overexpressed genes 30 min after 50  $\mu$ M chitohexaose treatment. GO term enrichment is expressed by node size. Enrichment  $P$ -value determined by hypergeometric test and corrected by Benjamini-Hochberg false discovery rate is represented by colour scale. Links between groups indicates shared genes. Only GO terms with a  $P$ -value  $< 0.01$  are shown. Links between groups indicates shared genes ( $\kappa$  score level  $\geq 0.4$ ).

**Figure S8.** Mitogen-activated protein kinases (MAPK) phosphorylation in 12-day-old Arabidopsis seedlings of Col-0 plants and *cerk1-2*, *bak1-5* and *sobir1-12* mutants impaired in PRR co-receptors treated with 50  $\mu$ M MLG43. Western blot using anti-pTEpY antibody for phosphorylated MAPK moieties at different time points (5, 10 and 20 min). Black arrows indicate the position of MPK6 (top), MPK3 (middle) and MPK4/11 (bottom) proteins. Anti-MPK6 and anti-MPK3 were used as total protein control. Ponceau red-stained membranes show equal loading. Chitohexaose (5  $\mu$ M) and distilled water (mock) were used as controls. These results are from one representative experiment of the three performed that gave similar results.

**Figure S9.** qRT-PCR analyses of the expression of PTI-related genes upregulated in tomato MoneyMaker plants treated for 60 min with MLG43 (0.25 mg plant<sup>-1</sup>). Relative expression levels to the *LOC543683* (*SIUBC*) gene are shown. Values are means  $\pm$  SD,  $n = 3$  from two independent experiments. Asterisks indicate treatments with significant differences compared with non-treated control plants (Student's  $t$ -test analysis,  $*P < 0.05$ ). These results are from one of the two representative experiments performed that gave similar results.

**Figure S10.** Representative images of disease index scales used to evaluate pepper-*Sclerotinia sclerotiorum* (left) or pepper-*Botrytis cinerea* (right) pathosystems. Disease index ranges from 0 to 5, where 0 = no symptoms; 1 = little necrotic spots ( $< 10\%$  of leaf area); 2 = two or more notable necrotic spots (10–25% of leaf

area); 3 = big necrotic area (25–50% of leaf area); 4 = more than 50% of leaf area affected and 5 = leaf senescence.

**Table S1.** Differentially expressed genes under treatment with MLG43 in Arabidopsis.

**Table S2.** Differentially expressed genes under treatment with chitohexaose in Arabidopsis.

**Table S3.** Classification of Arabidopsis upregulated genes into common and specific after treatment with MLG43 or chitohexaose.

**Table S4.** Classification of Arabidopsis downregulated genes into common and specific after treatment with MLG43 or chitohexaose.

**Table S5.**  $\beta$ -1,3/1,4-glucan oligosaccharides (MLGs) used in this work.

**Table S6.** Oligonucleotides used in this work.

**Table S7.** MLG43-ectodomains interaction energies determinations.

## REFERENCES

- Aziz, A., Gauthier, A., Bézier, A., Poinssot, B., Joubert, J.M., Pugin, A. *et al.* (2007) Elicitor and resistance-inducing activities of beta-1,4 celloedextrins in grapevine, comparison with beta-1,3 glucans and alpha-1,4 oligogalacturonides. *Journal of Experimental Botany*, **58**, 1463–1472.
- Aziz, A., Poinssot, B., Daire, X., Adrian, M., Bézier, A., Lambert, B. *et al.* (2003) Laminarin elicits defense responses in grapevine and induces protection against *Botrytis cinerea* and *Plasmopara viticola*. *Molecular Plant-Microbe Interactions*, **16**, 1118–1128.
- Bacete, L., Mérida, H., López, G., Dabos, P., Tremousaygue, D., Denancé, N. *et al.* (2020) Arabidopsis Response Regulator 6 (ARR6) modulates plant cell wall composition and disease resistance. *Molecular Plant-Microbe Interactions*, **33**, 767–780.
- Bacete, L., Mérida, H., Miedes, E. & Molina, A. (2018) Plant cell wall-mediated immunity: cell wall changes trigger disease resistance responses. *The Plant Journal*, **93**, 614–636.
- Bacete, L., Mérida, H., Pattathil, S., Hahn, M.G., Molina, A. & Miedes, E. (2017) Characterization of plant cell wall damage-associated molecular patterns regulating immune responses. *Methods in Molecular Biology*, **1578**, 13–23.
- Bartetzko, M.P. & Pfrengle, F. (2019) Automated glycan assembly of plant oligosaccharides and their application in cell-wall biology. *ChemBioChem*, **20**, 877–885.
- Benedetti, M., Pontiggia, D., Raggi, S., Cheng, Z., Scalconi, F., Ferrari, S. *et al.* (2015) Plant immunity triggered by engineered in vivo release of oligogalacturonides, damage-associated molecular patterns. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, 5533–5538.
- Benito, E.P., ten Have, A., van't Klooster, J.W. and van Kan, J.A.L. (1998) Fungal and plant gene expression during synchronized infection of tomato leaves by *Botrytis cinerea*. *European Journal of Plant Pathology*, **104**, 207–220.
- Bigeard, J., Colcombet, J. & Hirt, H. (2015) Signaling mechanisms in pattern-triggered immunity (PTI). *Molecular Plant*, **8**, 521–539.
- Bjornson, M., Pimprikar, P., Nürnberger, T. & Zipfel, C. (2020). The transcriptional landscape of *Arabidopsis thaliana* pattern-triggered immunity. *bioRxiv*.
- Boudsocq, M., Willmann, M.R., McCormack, M., Lee, H., Shan, L., He, P. *et al.* (2010) Differential innate immune signalling via Ca(2+) sensor protein kinases. *Nature*, **464**, 418–422.
- Boutrot, F. & Zipfel, C. (2017) Function, discovery, and exploitation of plant pattern recognition receptors for broad-spectrum disease resistance. *Annual review of Phytopathology*, **55**, 257–286.
- Burton, R.A. & Fincher, G.B. (2009) (1,3;1,4)- $\beta$ -D-glucans in cell walls of the poaceae, lower plants, and fungi: a tale of two linkages. *Molecular Plant*, **2**, 873–882.
- Burton, R.A., Wilson, S.M., Hrmova, M., Harvey, A.J., Shirley, N.J., Medhurst, A. *et al.* (2006) Cellulose synthase-like CslF genes mediate the synthesis of cell wall (1,3;1,4)- $\beta$ -D-glucans. *Science*, **311**, 1940–1942.

- Cao, Y., Liang, Y., Tanaka, K., Nguyen, C.T., Jedrzejczak, R.P., Joachimiak, A. *et al.* (2014). The kinase LYK5 is a major chitin receptor in Arabidopsis and forms a chitin-induced complex with related kinase CERK1. *eLife*, **3**, e03766.
- Chen, Y. & Wang, D. (2005) Two convenient methods to evaluate soybean for resistance to *Sclerotinia sclerotiorum*. *Plant Disease*, **89**, 1268–1272.
- Clavierie, J., Balacey, S., Lemaître-Guillier, C., Brulé, D., Chiltz, A., Granet, L. *et al.* (2018) The cell wall-derived xyloglucan is a new DAMP triggering plant immunity in *Vitis vinifera* and *Arabidopsis thaliana*. *Frontiers in Plant Science*, **9**, 1725.
- Dallabernardina, P., Schuhmacher, F., Seeberger, P.H. & Pfrenge, F. (2017) Mixed-linkage glucan oligosaccharides produced by automated glycan assembly serve as tools to determine the substrate specificity of lichenase. *Chemistry*, **23**, 3191–3196.
- del Hierro, I., Mérida, H., Brodyart, C., Santiago, J. & Molina, A. (2020) Computational prediction method to decipher receptor-glycoligand interactions in plant immunity. *The Plant Journal*, **105**. <http://dx.doi.org/10.1111/tpl.15133>.
- Desaki, Y., Kouzai, Y., Ninomiya, Y., Iwase, R., Shimizu, Y., Seko, K. *et al.* (2018) OsCERK1 plays a crucial role in the lipopolysaccharide-induced immune response of rice. *New Phytologist*, **217**, 1042–1049.
- DuBois, M., Gilles, K.A., Hamilton, J.K., Rebers, P.A. & Smith, F. (1956) Colorimetric method for determination of sugars and related substances. *Analytical Chemistry*, **28**, 350–356.
- Escudero, V., Jordá, L., Sopena-Torres, S., Mérida, H., Miedes, E., Muñoz-Barrios, A. *et al.* (2017) Alteration of cell wall xylan acetylation triggers defense responses that counterbalance the immune deficiencies of plants impaired in the  $\beta$ -subunit of the heterotrimeric G-protein. *The Plant Journal*, **92**, 386–399.
- Ferrari, S., Savatin, D.V., Sicilia, F., Gramegna, G., Cervone, F. & Lorenzo, G.D. (2013) Oligogalacturonides: plant damage-associated molecular patterns and regulators of growth and development. *Frontiers in Plant Science*, **4**, 49.
- Fesel, P.H. & Zuccaro, A. (2016)  $\beta$ -glucan: crucial component of the fungal cell wall and elusive MAMP in plants. *Fungal Genetics and Biology*, **90**, 53–60.
- Fincher, G.B. & Stone, B.A. (2004) Chemistry of nonstarch polysaccharides. In: Wrigley, C., Corke, H. and Walker, C.E. (Eds.) *Encyclopedia of grain science*. Oxford, UK: Elsevier, pp. 206–223.
- Fontaine, T., Simenel, C., Dubreucq, G., Adam, O., Delepierre, M., Lemoine, J. *et al.* (2000) Molecular organization of the alkali-insoluble fraction of *Aspergillus fumigatus* cell wall. *Journal of Biological Chemistry*, **275**, 27594–27607.
- Fry, S.C., Nesselrode, B.H.W.A., Miller, J.G. & Mewburn, B.R. (2008) Mixed-linkage (1,3;1,4)- $\beta$ -D-glucan is a major hemicellulose of Equisetum (horsetail) cell walls. *New Phytologist*, **179**, 104–115.
- Galletti, R., Denoux, C., Gambetta, S., Dewdney, J., Ausubel, F.M., De Lorenzo, G. *et al.* (2008) The AtrbohD-mediated oxidative burst elicited by oligogalacturonides in Arabidopsis is dispensable for the activation of defense responses effective against *Botrytis cinerea*. *Plant Physiology*, **148**, 1695–1706.
- Geoghegan, I., Steinberg, G. & Gurr, S. (2017) The role of the fungal cell wall in the infection of plants. *Trends in Microbiology*, **25**, 957–967.
- Gómez-Gómez, L. & Boller, T. (2000) FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. *Molecular Cell*, **5**, 1003–1011.
- Gorin, P., Baron, M. & Iacomini, M. (1988) Storage products of lichens. In: Galun, M. (Ed.) *Handbook of Lichenology*. Boca Raton, FL, USA: CRC Press, pp. 9–23.
- Gust, A.A., Biswas, R., Lenz, H.D., Rauhut, T., Ranf, S., Kemmerling, B. *et al.* (2007) Bacteria-derived peptidoglycans constitute pathogen-associated molecular patterns triggering innate immunity in Arabidopsis. *Journal of Biological Chemistry*, **282**, 32338–32348.
- Henrissat, B. & Bairoch, A. (1993) New families in the classification of Glycosyl hydrolases based on amino acid sequence similarities. *The Biochemical Journal*, **293**, 781–788.
- Johnson, J.M., Thürich, J., Petutschnig, E.K., Altschmied, L., Meichsner, D., Sherameti, I. *et al.* (2018) Poly(A) ribonuclease controls the cellotriose-based interaction between *Piriformospora indica* and its host Arabidopsis. *Plant Physiology*, **176**, 2496–2514.
- Kaku, H., Nishizawa, Y., Ishii-Minami, N., Akimoto-Tomiyama, C., Dohmae, N., Takio, K. *et al.* (2006) Plant cells recognize chitin fragments for defense signalling through a plasma membrane receptor. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 11086–11091.
- Klarzynski, O., Plesse, B., Joubert, J.M., Yvin, J.C., Kopp, M., Kloreg, B. *et al.* (2000) Linear  $\beta$ -1,3 glucans are elicitors of defense responses in tobacco. *Plant Physiology*, **124**, 1027–1038.
- Lampugnani, E.R., Khan, G.A., Somssich, M. & Persson, S. (2018) Building a plant cell wall at a glance. *Journal of Cell Science*, **131**(2), jcs207373.
- Latgé, J.P. & Calderone, R. (2006). The fungal cell wall. In: Kües, U. & Fischer, R. (Eds). *The Mycota I. Growth, Differentiation and Sexuality*. Heidelberg, Germany: Springer Berlin, pp. 73–104.
- Lee, J. & Hollingsworth, R.I. (1997) Oligosaccharide beta-glucans with unusual linkages from *Sarcina ventriculi*. *Carbohydrate Research*, **304**, 133–141.
- Liu, D., Lu, J., Li, H., Wang, J. & Pei, Y. (2019) Characterization of the O-acetylserine(thiol)lyase gene family in *Solanum lycopersicum* L. *Plant Molecular Biology*, **99**, 123–134.
- Liu, T., Liu, Z., Song, C., Hu, Y., Han, Z., She, J. *et al.* (2012) Chitin-induced dimerization activates a plant immune receptor. *Science*, **336**, 1160–1164.
- Locci, F., Benedetti, M., Pontiggia, D., Citterico, M., Caprari, C., Mattei, B. *et al.* (2019) An Arabidopsis berberine bridge enzyme-like protein specifically oxidizes cellulose oligomers and plays a role in immunity. *The Plant Journal*, **98**, 540–554.
- McIntosh, M., Stone, B.A. & Stanisich, V.A. (2005) Curdlan and other bacterial (1 $\rightarrow$ 3)- $\beta$ -D-glucans. *Applied Microbiology and Biotechnology*, **68**, 163–173.
- Mérida, H., Bacete, L., Ruprecht, C., Rebaque, R., del Hierro, I., López, G. *et al.* (2020) Arabinoxylan-oligosaccharides act as Damage Associated Molecular Patterns in plants regulating disease resistance. *Frontiers in Plant Science*, **11**, 1210.
- Mérida, H., Garcia-Angulo, P., Alonso-Simón, A., Encina, A., Alvarez, J. & Acebes, J.L. (2009) Novel type II cell wall architecture in dichlobenil-habituated maize calluses. *Planta*, **229**, 617–631.
- Mérida, H., Sandoval-Sierra, J.V., Diéguez-Urbeondo, J. & Bulone, V. (2013) Analyses of extracellular carbohydrates in oomycetes unveil the existence of three different cell wall types. *Eukaryotic Cell*, **12**, 194–203.
- Mérida, H., Sopena-Torres, S., Bacete, L., Garrido-Arandia, M., Jordá, L., López, G. *et al.* (2018) Non-branched  $\beta$ -1,3-glucan oligosaccharides trigger immune responses in Arabidopsis. *The Plant Journal*, **93**, 34–49.
- Ménard, R., Alban, S., de Ruffray, P., Jamois, F., Franz, G., Fritig, B. *et al.* (2004)  $\beta$ -1,3 glucan sulfate, but not  $\beta$ -1,3 glucan, induces the salicylic acid signaling pathway in tobacco and Arabidopsis. *The Plant Cell*, **16**, 3020–3032.
- Miya, A., Albert, P., Shinya, T., Desaki, Y., Ichimura, K., Shirasu, K. *et al.* (2007) CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 19613–19618.
- Pérez-Mendoza, D., Rodríguez-Carvajal, M.A., Romero-Jiménez, L., Fariás Gde, A., Lloret, J., Gallegos, M.T. *et al.* (2015) Novel mixed-linkage  $\beta$ -glucan activated by c-di-GMP in *Sinorhizobium meliloti*. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, E757–E765.
- Perlin, A.S. & Suzuki, S. (1962) The structure of lichenin: selective enzymolysis studies. *Canadian Journal of Botany*, **40**, 50–56.
- Perraki, A., DeFalco, T.A., Derbyshire, P., Avila, J., Séré, D., Sklenar, J. *et al.* (2018) Phosphocode-dependent functional dichotomy of a common coreceptor in plant signalling. *Nature*, **561**, 248–252.
- Pettolino, F., Sasaki, I., Turbic, A., Wilson, S.M., Bacic, A., Hrmova, M. *et al.* (2009) Hyphal cell walls from the plant pathogen *Rhynchosporium secalis* contain (1,3;1,6)- $\beta$ -D-glucans, galacto- and rhamnmannans, (1,3;1,4)- $\beta$ -D-glucans and chitin. *FEBS Journal*, **276**, 4122–4133.
- Planas, A. (2000) Bacterial 1,3-1,4- $\beta$ -glucanases: structure, function and protein engineering. *Biochimica et Biophysica Acta*, **1543**, 361–382.
- Popper, Z.A. & Fry, S.C. (2003) Primary cell wall composition of bryophytes and charophytes. *Annals of Botany*, **91**, 1–12.
- Ranf, S., Eschen-Lippold, L., Pecher, P., Lee, J. & Scheel, D. (2011) Interplay between calcium signalling and early signalling elements during defence responses to microbe- or damage-associated molecular patterns. *The Plant Journal*, **68**, 100–113.
- Ranf, S., Grimmer, J., Pöschl, Y., Pecher, P., Chinchilla, D., Scheel, D. *et al.* (2012) Defense-related calcium signaling mutants uncovered via a

- quantitative high-throughput screen in *Arabidopsis thaliana*. *Molecular Plant*, **5**, 115–130.
- Rovenich, H., Zuccaro, A. & Thomma, B.P. (2016) Convergent evolution of filamentous microbes towards evasion of glycan-triggered immunity. *New Phytologist*, **212**, 896–901.
- Rui, Y. & Dinneny, J.R. (2020) A wall with integrity: surveillance and maintenance of the plant cell wall under stress. *New Phytologist*, **225**, 1428–1439.
- Salmeán, A.A., Duffieux, D., Harholt, J., Qin, F., Michel, G., Czjzek, M. *et al.* (2017) Insoluble (1 → 3), (1 → 4)- $\beta$ -D-glucan is a component of cell walls in brown algae (Phaeophyceae) and is masked by alginates in tissues. *Scientific Reports*, **7**, 2880.
- Samar, D., Kieler, J.B. & Klutts, J.S. (2015) Identification and deletion of Tft1, a predicted glycosyltransferase necessary for cell wall  $\beta$ -1,3;1,4-glucan synthesis in *Aspergillus fumigatus*. *PLoS ONE*, **10**, e0117336.
- Santamaría-Hernando, S., Senovilla, M., González-Mula, A., Martínez-García, P.M., Nebreda, S., Rodríguez-Palenzuela, P. *et al.* (2019) The *Pseudomonas syringae* pv. tomato DC3000 PSPTO\_0820 multidrug transporter is involved in resistance to plant antimicrobials and bacterial survival during tomato plant infection. *PLoS ONE*, **14**, e0218815.
- Sharp, J.K., McNeil, M. & Albersheim, P. (1984) The primary structures of one elicitor-active and seven elicitor-inactive hexa( $\beta$ -D-glucopyranosyl)-D-glucitols isolated from the mycelial walls of *Phytophthora megasperma* f. sp. glycinea. *Journal of Biological Chemistry*, **259**, 11321–11336.
- Simmons, T.J., Uhrin, D., Gregson, T., Murray, L., Sadler, I.H. & Fry, S.C. (2013) An unexpectedly lichenase-stable hexasaccharide from cereal, horsetail and lichen mixed-linkage  $\beta$ -glucans (MLGs): implications for MLG subunit distribution. *Phytochemistry*, **95**, 322–332.
- Smith, B.G. & Harris, P.J. (1999) The polysaccharide composition of Poales cell walls: Poaceae cell walls are not unique. *Biochemical Systematics and Ecology*, **27**, 33–53.
- Sørensen, I., Pettolino, F.A., Wilson, S.M., Doblin, M.S., Johansen, B., Bacic, A. *et al.* (2008) Mixed-linkage (1,3;1,4)- $\beta$ -D-glucan is not unique to the poales and is an abundant component of *Equisetum arvense* cell walls. *The Plant Journal*, **54**, 510–521.
- Souza, C., Li, S., Lin, A.Z., Boutrot, F., Grossmann, G., Zipfel, C. *et al.* (2017) Cellulose-derived oligomers act as damage-associated molecular patterns and trigger defense-like responses. *Plant Physiology*, **173**, 2383–2398.
- Srivastava, V., McKee, L.S. & Bulone, V. (2017) *Plant Cell Walls*. eLS. Chichester, UK: John Wiley & Sons Ltd.
- Stone, B.A. & Clarke, A.E. (1992) *Chemistry and biology of (1,3)- $\beta$ -D-glucans*. Victoria, Australia: La Trobe University Press.
- Tang, D., Wang, G. & Zhou, J.M. (2017) Receptor kinases in plant-pathogen interactions: more than pattern recognition. *The Plant Cell*, **29**, 618–637.
- Trethewey, J.A.K., Campbell, L.M. & Harris, P.J. (2005) (1,3;1,4)- $\beta$ -D-Glucans in the cell walls of the poales (sensu lato): an immunogold labeling study using a monoclonal antibody. *American Journal of Botany*, **92**, 1660–1674.
- van der Burgh, A.M., Postma, J., Robatzek, S. & Joosten, M.H.A.J. (2019) Kinase activity of SOBIR1 and BAK1 is required for immune signalling. *Molecular Plant Pathology*, **20**, 410–422.
- Wanke, A., Rovenich, H., Schwanke, F., Velte, S., Becker, S., Hehemann, J.H. *et al.* (2020) Plant species-specific recognition of long and short  $\beta$ -1,3-linked glucans is mediated by different receptor systems. *The Plant Journal*, **102**, 1142–1156.
- Wawra, S., Fesel, P., Widmer, H., Timm, M., Seibel, J., Leson, L. *et al.* (2016) The fungal-specific  $\beta$ -glucan-binding lectin FGB1 alters cell-wall composition and suppresses glucan-triggered immunity in plants. *Nature Communications*, **7**, 13188.
- Willmann, R., Lajunen, H.M., Erbs, G., Newman, M.-A., Kolb, D., Tsuda, K. *et al.* (2011) Arabidopsis lysin-motif proteins LYM1 LYM3 CERK1 mediate bacterial peptidoglycan sensing and immunity to bacterial infection. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 19824–19829.
- Yamaguchi, Y., Pearce, G. & Ryan, C.A. (2006) The cell surface leucine-rich repeat receptor for AtPep1, an endogenous peptide elicitor in Arabidopsis, is functional in transgenic tobacco cells. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 10104–10109.
- Zablackis, E., Huang, J., Müller, B., Darvill, A.G. & Albersheim, P. (1995) Characterization of the cell-wall polysaccharides of *Arabidopsis thaliana* leaves. *Plant Physiology*, **107**, 1129–1138.
- Zang, H., Xie, S., Zhu, B., Yang, X., Gu, C., Hu, B. *et al.* (2019) Mannan oligosaccharides trigger multiple defence responses in rice and tobacco as a novel danger-associated molecular pattern. *Molecular Plant Pathology*, **20**, 1067–1079.



OPEN ACCESS

**Edited by:**

Martin Heil,  
Instituto Politécnico Nacional de  
México (CINVESTAV), Mexico

**Reviewed by:**

Johannes Werner Stratmann,  
University of South Carolina,  
United States  
Maxime Versluys,  
KU Leuven, Belgium

**\*Correspondence:**

Antonio Molina  
antonio.molina@upm.es

**<sup>†</sup>Present address:**

Laura Bacete,  
Faculty of Natural Sciences, Institute  
for Biology, Norwegian University of  
Science and Technology,  
Trondheim, Norway  
Colin Ruprecht,  
Department of Chemistry, University of  
Natural Resources and Life Sciences,  
Vienna, Austria  
Fabian Pfrengle,  
Department of Chemistry, University of  
Natural Resources and Life Sciences,  
Vienna, Austria

**Specialty section:**

This article was submitted to  
Plant Pathogen Interactions,  
a section of the journal  
Frontiers in Plant Science

**Received:** 30 April 2020

**Accepted:** 24 July 2020

**Published:** 07 August 2020

**Citation:**

Mélida H, Bacete L, Ruprecht C,  
Rebaque D, del Hierro I, López G,  
Brunner F, Pfrengle F and Molina A  
(2020) Arabinoxylan-Oligosaccharides  
Act as Damage Associated  
Molecular Patterns in Plants  
Regulating Disease Resistance.  
Front. Plant Sci. 11:1210.  
doi: 10.3389/fpls.2020.01210

# Arabinoxylan-Oligosaccharides Act as Damage Associated Molecular Patterns in Plants Regulating Disease Resistance

Hugo Mélida<sup>1</sup>, Laura Bacete<sup>1,2†</sup>, Colin Ruprecht<sup>3†</sup>, Diego Rebaque<sup>1,2,4</sup>, Irene del Hierro<sup>1,2</sup>, Gemma López<sup>1</sup>, Frédéric Brunner<sup>4</sup>, Fabian Pfrengle<sup>3†</sup> and Antonio Molina<sup>1,2\*</sup>

<sup>1</sup> Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid (UPM)—Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Pozuelo de Alarcón (Madrid), Spain, <sup>2</sup> Departamento de Biotecnología-Biología Vegetal, Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas, UPM, Madrid, Spain, <sup>3</sup> Department of Biomolecular Systems, Max Planck Institute of Colloids and Interfaces, Potsdam, Germany, <sup>4</sup> PlantResponse Biotech S.L., Campus de Montegancedo UPM, Pozuelo de Alarcón (Madrid), Spain

Immune responses in plants can be triggered by damage/microbe-associated molecular patterns (DAMPs/MAMPs) upon recognition by plant pattern recognition receptors (PRRs). DAMPs are signaling molecules synthesized by plants or released from host cellular structures (e.g., plant cell walls) upon pathogen infection or wounding. Despite the hypothesized important role of plant cell wall-derived DAMPs in plant-pathogen interactions, a very limited number of these DAMPs are well characterized. Recent work demonstrated that pectin-enriched cell wall fractions extracted from the cell wall mutant impaired in *Arabidopsis Response Regulator 6 (arr6)*, that showed altered disease resistance to several pathogens, triggered more intense immune responses than those activated by similar cell wall fractions from wild-type plants. It was hypothesized that *arr6* cell wall fractions could be differentially enriched in DAMPs. In this work, we describe the characterization of the previous immune-active fractions of *arr6* showing the highest triggering capacities upon further fractionation by chromatographic means. These analyses pointed to a role of pentose-based oligosaccharides triggering plant immune responses. The characterization of several pentose-based oligosaccharide structures revealed that  $\beta$ -1,4-xylooligosaccharides of specific degrees of polymerization and carrying arabinose decorations are sensed as DAMPs by plants. Moreover, the pentasaccharide 3<sup>3</sup>- $\alpha$ -L-arabinofuranosyl-xylotetraose (XA3XX) was found as a highly active DAMP structure triggering strong immune responses in *Arabidopsis thaliana* and enhancing crop disease resistance.

**Keywords:** arabinoxylan, cell wall, damage-associated molecular pattern (DAMP), plant immunity, pattern triggered immunity

## INTRODUCTION

Plants are sessile organisms that need to develop robust disease resistance mechanisms to efficiently defend from pathogens and pests. Activation of plant defense responses requires the perception of molecules from the pathogen (“non-self” signals) and from the plant (“damaged-self” signals) that trigger specific resistance responses through diverse molecular monitoring systems (Atkinson and Urwin, 2012). Among these monitoring mechanisms are pattern- and effector-triggered immunity (PTI and ETI) (Dodds and Rathjen, 2010). PTI is based in the recognition by pattern recognition receptors (PRRs) of microbe/pathogen-associated molecular patterns (MAMPs/PAMPs) from microorganisms or of plant-derived damage-associated molecular patterns (DAMPs) (Boller and Felix, 2009). MAMPs and DAMPs structures, with different biochemical composition (e.g., proteins, carbohydrates, lipids, and nucleic acids) have been identified, thus reflecting the diversity of immunogenic structures recognized by plants (Boutrot and Zipfel, 2017). In comparison with the high number of MAMPs characterized so far, much less DAMPs derived from plants have been identified to date (Choi and Klessig, 2016; Duran-Flores and Heil, 2016; Bacete et al., 2018; De Lorenzo et al., 2018; Li et al., 2020).

The plant cell wall is a dynamic and highly regulated structure mainly consisting of carbohydrate-based polymers, essential for growth, and development (Srivastava et al., 2017). Cellulose is the main load-bearing component in all plant cell walls, whereas different types of hemicelluloses and pectins are found in different plant phylogenetic groups (Carpita and Gibeau, 1993; Carpita and McCann, 2000). Xylans are a diverse group of hemicelluloses with the common feature of a backbone of  $\beta$ -1,4-linked xylose residues (Scheller and Ulvskov, 2010). Monocot xylans usually contain many arabinose residues attached to the backbone and are known as arabinoxylans (AXs). Arabinofuranose substitutions are, in principle, less frequent in dicot xylans, but exceptions are found (Darvill et al., 1980; Fischer et al., 2004; Naran et al., 2008). Instead, dicot xylans are more commonly substituted with  $\alpha$ -1,2-linked glucuronosyl and 4-O-methyl glucuronosyl residues known as glucuronoxylans which are the dominating non-cellulosic polysaccharides in the secondary walls of dicots. This variability in the fine structure of wall polymers exists not only among phylogenetic groups of plants, but also even between different tissues of a given plant. Cell wall heterogeneity may have had an evolutionary impact in the diversity of mechanisms that pathogens have evolved to breach plant cell walls, including the secretion of numerous cell wall-degrading enzymes (CWDE), such as cellulases, polygalacturonases, or xylanases (Annis and Goodwin, 1997). The functional integrity of cell walls is controlled by cell wall integrity monitoring systems (Bacete et al., 2018; Vaahter et al., 2019). These systems trigger countervailing responses to cell wall restructuring which occurs upon pathogen infection, abiotic stress, and cell expansion during growth and development. The plant cell wall integrity pathway is strongly involved in the regulation of

growth, immune responses and resource allocation between development and immunity (Hamann et al., 2009; Wolf et al., 2012; Engelsdorf et al., 2018). Alterations in cell wall composition or integrity by genetic or chemical means have a significant impact on plant resistance to different pathogens and/or abiotic stresses, since they typically lead to the activation of defensive signaling pathways, some of which are regulated by hormones (Miedes et al., 2014; Mélida et al., 2015; Nafisi et al., 2015; Houston et al., 2016; Bacete et al., 2020). For example, enhanced resistance to pathogens has been observed in *Arabidopsis thaliana* (*Arabidopsis*) mutants defective in specific cellulose synthases, enzymes involved in xylan decoration and in lignin biosynthesis (Ellis et al., 2002; Hernández-Blanco et al., 2007; Delgado-Cerezo et al., 2012; Xu et al., 2014; Escudero et al., 2017; Gallego-Giraldo et al., 2020; Molina et al., 2020).

Given both the complexity of the plant cell wall and the fact that many pathogens secrete a wide range of CWDE, it would be expected that the breakdown products of cell wall polymers could act as DAMPs that regulate immune responses. Confirming this hypothesis, pectic oligogalacturonides (OGs) were first cell wall DAMPs to be characterized (Nothnagel et al., 1983). OGs are derived from homogalacturonan, the main component of pectins, as a result of the activity of CWDE released by the pathogens during the colonization process (Ridley et al., 2001; Benedetti et al., 2015; Voxeur et al., 2019). Also, the overexpression or inactivation of genes encoding enzymes involved in the control of pectin structure [e.g., pectin methyl esterases (PME) and PME inhibitors], results in the modification of the degree of OGs release upon infection, and alterations of disease resistance phenotypes (Ferrari et al., 2008; Raiola et al., 2011; Lionetti et al., 2017; Benedetti et al., 2018; De Lorenzo et al., 2019). Another group of cell wall-derived carbohydrates recently characterized as DAMPs in *Arabidopsis* are cellulose-derived oligomers ( $\beta$ -1,4-glucans), which trigger signaling cascades sharing many similarities with the responses activated by OGs (Aziz et al., 2007; Souza et al., 2017; Johnson et al., 2018; Locci et al., 2019). With around 20 different monosaccharide moieties building the polysaccharides of the plant cell wall, other carbohydrate-based cell wall molecules in addition to OGs and cello-oligosaccharides should have been selected by plants as DAMPs. In line with this hypothesis, recent works have also nominated xyloglucan and mannan cell wall-derived oligosaccharides as plant DAMPs (Claverie et al., 2018; Zang et al., 2019), and  $\beta$ -1,3-glucan oligosaccharides present in plant callose but also in fungal cell walls, as dual DAMPs/MAMPs (Mélida et al., 2018).

Thus, growing evidences have awarded the cell wall with prominent novel roles in plant immunity. In this line, we have recently proposed a novel link between the cytokinin signaling pathway, cell wall composition control, and disease resistance responses through *Arabidopsis* Response Regulator 6 (ARR6) protein (Bacete et al., 2020). Cytokinins have emerged as an important hub integrating defense responses mediated by other hormones, and have been shown to regulate the activation of immune responses (Choi et al., 2011). In *Arabidopsis*, cytokinins

are perceived by *Arabidopsis* Histidine Kinase receptors, that are two-component system proteins which initiate a downstream phosphotransfer cascade that leads to the phosphorylation of *Arabidopsis* Response Regulator (ARR) proteins (To et al., 2007). We showed a novel function for ARR6, as impairment of *ARR6* gene affect plant cell wall composition, which impact plant-pathogen interactions, and might lead to the accumulation of differential or increased levels of DAMPs in *arr6* in comparison to wild-type plants that would favor a “defense-ready” state instead of a resting one. Remarkably, pectin-enriched cell wall fractions extracted from *arr6* cell walls triggered, when applied to wild-type *Arabidopsis* plants, more intense immune responses than those activated by similar wall fractions from wild-type plants, suggesting that *arr6* pectin fraction is enriched in wall-derived DAMPs. In an effort toward a better understanding of plant mechanisms involved in cell wall-mediated immunity, we have further purified *arr6* pectin fraction. Results from such purifications suggested that pentose-based oligosaccharides co-extracted with pectins (using calcium chelators) could play a role as plant DAMPs. Afterwards, we purified several pentose-based oligosaccharides, generated by enzymatic digestion from a natural material source rich in that type of hemicelluloses, that were biochemically analyzed and tested for their capacity to induce PTI hallmarks (Boller and Felix, 2009; Boudsocq et al., 2010; Ranf et al., 2011). Using this strategy, we identified AX-oligosaccharides as a novel group of DAMPs active on plants and characterized 3<sup>3</sup>- $\alpha$ -L-arabinofuranosyl-xylotetraose (XA3XX) as a highly active structure triggering strong immune responses in *Arabidopsis* and enhancing crop disease resistance.

## RESULTS

### Low Molecular Weight Pectic Fractions From *Arabidopsis* Enriched in Oligopentoses Contain Active Plant DAMPs

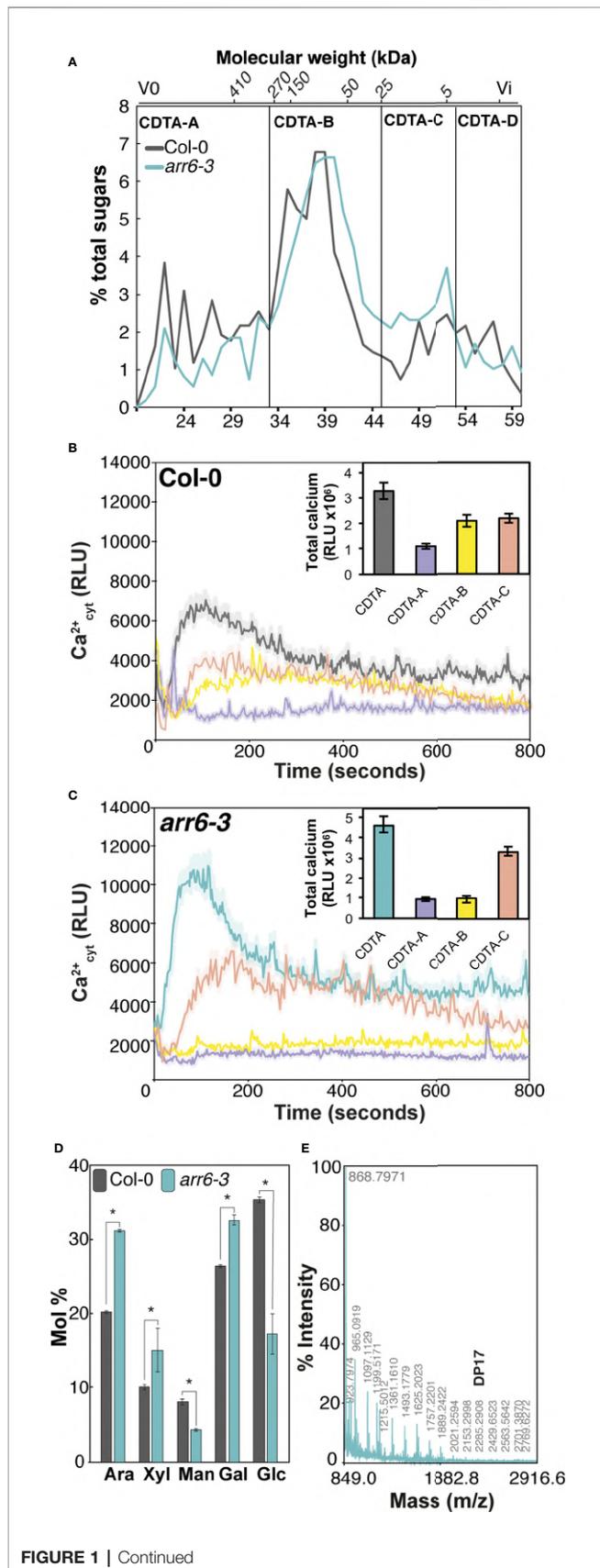
In a previous work, we hypothesized that the molecular basis of the differential disease resistance responses in the *Arabidopsis arr6-3* mutant allele could be associated with the enhanced and differential presence of carbohydrate-based DAMPs in the pectin-enriched fractions derived from their cell walls (Bacete et al., 2020). These DAMPs, when released, would activate immune responses, thus triggering disease resistance. Pectic fractions isolated from *arr6-3* cell walls triggered more intense Ca<sup>2+</sup> influxes and MAPK phosphorylation than the fractions from wild-type plants (Bacete et al., 2020), thus they were selected for further analyses in order to characterize the putative DAMPs responsible for the observed differential immune responses in *arr6-3* plants. Pectin-enriched fractions from *arr6-3* and Col-0 plants, extracted with 1,2-cyclohexylenedinitrotetraacetic acid (CDTA) from purified cell walls, were further fractionated by size exclusion chromatography to obtain samples containing carbohydrates with distinct molecular weights. Four sub-fractions (CDTA-A to CDTA-D) were obtained, containing molecules with different theoretical sizes: *i*) CDTA-A: >270 kDa;

*ii*) CDTA-B: 270-25 kDa; *iii*) CDTA-C: 25-5 kDa; *iv*) CDTA-D: <5 kDa (Figure 1A). These masses are estimated as the Sepharose column was calibrated with commercial dextrans of known weight-average relative molecular mass, which may not display similar conformations as pectic polymers. Total sugar quantifications showed that even after long dialysis procedures, the CDTA-D fractions contained very low amounts of carbohydrates and were most likely composed of the solvent used to obtain this fraction (CDTA; Mort et al., 1991), and therefore, they were excluded for further analyses. CDTA-A, -B, and -C were tested for their capacity to trigger intracellular Ca<sup>2+</sup> entry, an early immune response, in Col-0<sup>AEQ</sup> sensor lines (Ranf et al., 2012). CDTA-C sub-fractions from both Col-0 and *arr6-3* retained most of the activity of the complete CDTA-pectin fractions (Figures 1B, C), whereas CDTA-A or -B from *arr6-3* did not present any activity and CDTA-B from Col-0 still presented some activity (Figures 1B, C). Thus, we concluded that potential active DAMPs were most abundant in the *arr6-3* CDTA-C sub-fractions.

Neutral sugar analyses by GC/MS revealed that CDTA-C sub-fractions were still very complex in terms of monosaccharide composition, challenging further predictions about the identity of novel DAMPs that they would contain (Figure 1D). However, the enrichment of *arr6-3* CDTA-C sub-fractions in arabinose and xylose (Figure 1D) was in line with MALDI-TOF/TOF mass spectrometry analyses, which showed only the presence of pentose oligosaccharides (m/z shifts of 132) with degree of polymerization (DP) up to 17 (Figure 1E). Other oligosaccharide signatures were not found in the MALDI-TOF/TOF, clearly indicating that pentose-containing carbohydrates could be novel DAMPs present in the PTI-active, CDTA-extractable pectin-enriched fractions of *Arabidopsis*. However, further fractionation of CDTA-subfractions resulted non-viable due to their complexity in terms of composition and polydispersity combined with the low yields obtained. Thus, we decided to investigate the capacity of different oligosaccharides containing arabinose and xylose which could be obtained from commercial AXs using specific glycosyl hydrolases (GH).

### Arabinoxylan Oligosaccharides With Different DP Trigger Calcium Influxes in *Arabidopsis*

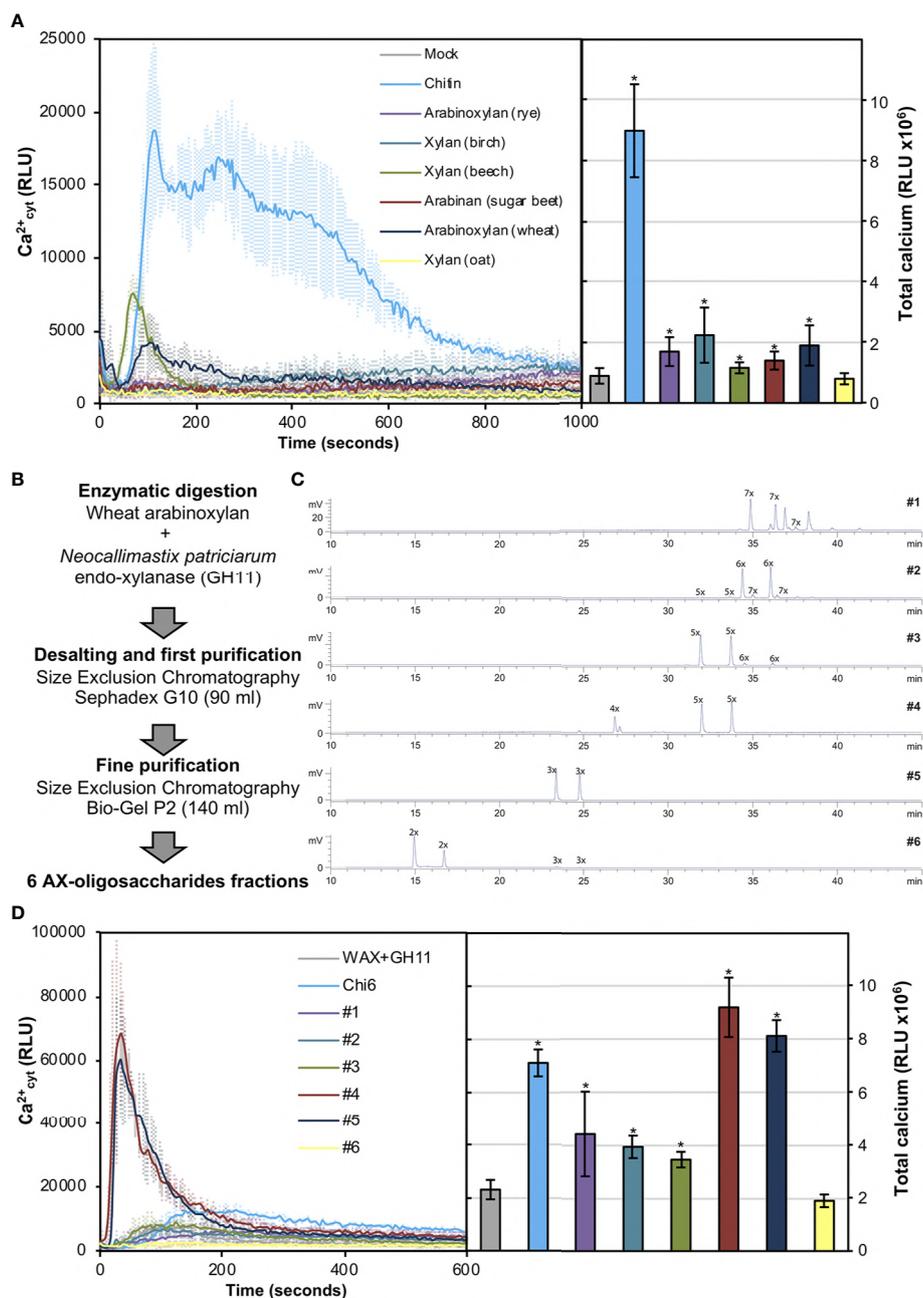
In order to investigate whether pentose-based structures could be sensed by *Arabidopsis*, we decided to analyze in a first instance the capacity to trigger Ca<sup>2+</sup> influxes of different commercial polymeric structures (Figure 2A). As previously described, polysaccharides often need to be solubilized to smaller entities in order to trigger early immune responses in plants (Mélida et al., 2018). In this regard, partial solubilization by heating of water-dissolved polysaccharides can help to expose ligands which may not be accessible in their insoluble counterparts. Heat-solubilized xylan from beech and AX from wheat triggered subtle calcium influxes compared to chitin, but still represented good candidates as pentose-DAMP sources (Figure 2A). Based on these results, we selected wheat AX as the polymeric source to



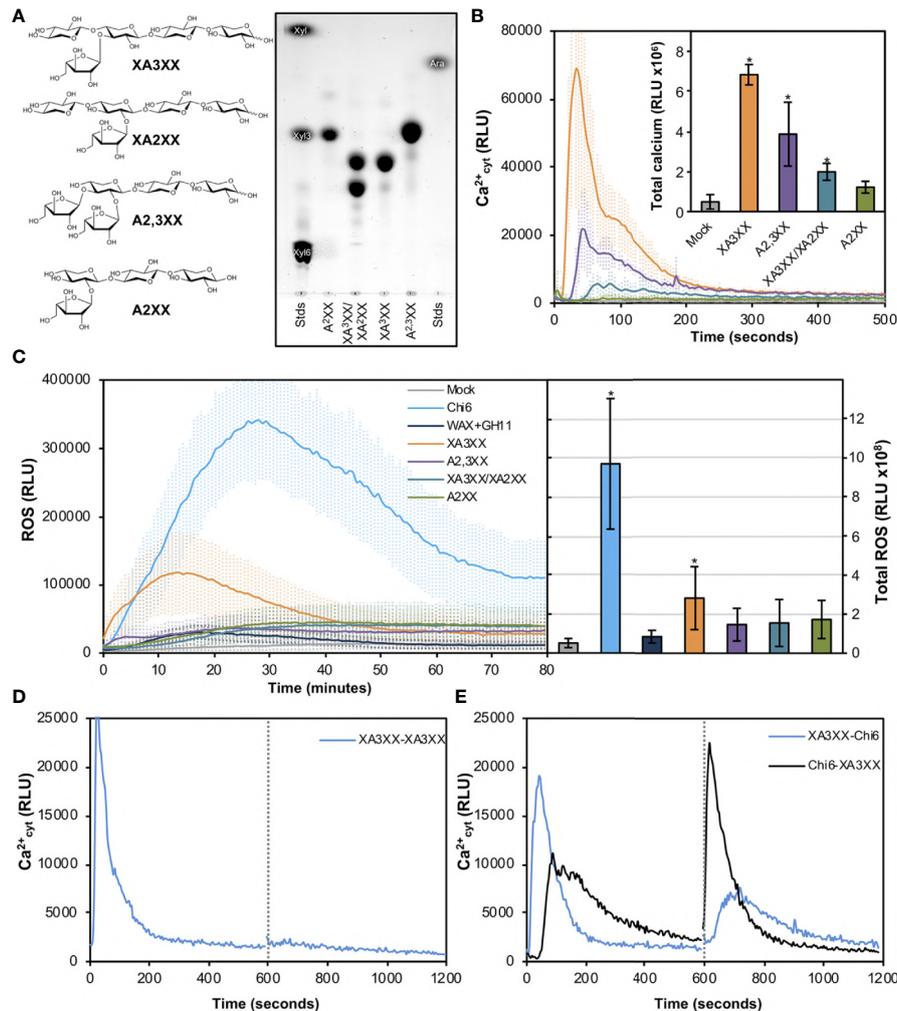
**FIGURE 1 |** Pectin-CDTA sub-fractions between 25 and 5 kDa retain most of the activity triggering  $\text{Ca}^{2+}$  influxes and are enriched in pentose oligosaccharides. **(A)** Size exclusion chromatography (SEC) elution profile (Sepharose CL-6B) of pectin fractions (CDTA extract) from wild type (Col-0) and *arr6-3* plants. The Sepharose column was calibrated with commercial dextrans of known weight-average relative molecular mass and the elution fraction number (bottom) of some of them is indicated on the top of the chromatogram. The elution profiles were monitored by total sugar quantification (phenol-sulfuric method). Sub-fractions were defined as: [A] > 270 kDa, [B] 270–25 kDa, [C] 25–5 kDa, [D] < 5 kDa. Profiles are representative of ten independent preparations. **(B, C)**  $\text{Ca}^{2+}$  influx kinetics triggered by CDTA sub-fractions A-C from Col-0 **(B)** and *arr6-3* **(C)** plants in Col-0<sup>AEC</sup> seedlings. Elevations of cytoplasmic calcium concentrations over 800 s were measured as relative luminescence units (RLU). Data are means (n=8) from one experiment representative of three independent ones with similar results. The total areas-under-the-curves were integrated and their average values  $\pm$  SD (n=8) are represented at the right side of each panel. **(D)** Monosaccharide composition (Mol %  $\pm$  SD, n=3) of Col-0, and *arr6-3* CDTA-C sub-fraction. Ara: arabinose; Xyl: xylose; Man: mannose; Gal: galactose; Glc: glucose. Statistically significant differences between genotypes according to Student's t-test (\* $p < 0.05$ ). **(E)** MALDI-TOF mass spectrum of CDTA-C sub-fraction. M/z shifts are coherent with the presence of pentose oligosaccharides of different degree of polymerization (DP). The spectrum shown (*arr6-3*) is representative of all analyzed pectin I-C sub-fractions (n $\geq$ 10).

be hydrolyzed to oligosaccharides given that arabinose decorations could mean an advantage compared to non-decorated xylans when enzymatic hydrolysis is used to generate different oligosaccharides of desired DP (McCleary et al., 2015). Wheat AX was hydrolyzed with an endo-xylanase (GH11) from *Neocallimastix patriciarum* and 6 AX-oligosaccharide fractions (#1 to #6) were purified through two rounds of size exclusion chromatography (**Figure 2B**). Purified fractions contained pentose-oligosaccharides ranging from DP 2 to 9 as demonstrated by HPLC-ELSD and MS/MS (**Figure 2C** and **Supplementary Figure S1**). Interestingly, we found that  $\text{Ca}^{2+}$  burst in treated plants was activated by all fractions except #6, which contained mainly a disaccharide and minor amounts of a trisaccharide (**Figure 2D** and **Supplementary Figure S1**). Since fraction #5 also contained a trisaccharide but triggered intense  $\text{Ca}^{2+}$  influxes, it seems that a DP above 2 is required by Arabidopsis perception machinery in case of pentose-based oligosaccharides. Together with fraction #5, fraction #4 resulted the most active and according to HPLC-ELSD and MS/MS these contained pentose oligosaccharides ranging from DP 3 to 5 (**Figure 2D** and **Supplementary Figure S1**).

In view of the results obtained with our purified oligosaccharides, we decided to investigate well-defined and highly-pure commercial structures which most likely resemble those from our purifications (McCleary et al., 2015). These included the pentasaccharides 3<sup>3</sup>- $\alpha$ -L-arabinofuranosyl-xylo-tetraose (XA3XX), 2<sup>3</sup>- $\alpha$ -L-arabinofuranosyl-xylo-tetraose (XA2XX), 2<sup>3</sup>,3<sup>3</sup>-di- $\alpha$ -L-arabinofuranosyl-xylo-triose (A2,3XX), and the tetrasaccharide 2<sup>3</sup>- $\alpha$ -L-arabinofuranosyl-xylo-triose (A2XX) (**Figure 3A**). Readouts from two early PTI events, such as  $\text{Ca}^{2+}$  influxes and production of reactive oxygen species (ROS), upon plant treatment with these oligosaccharides indicated that the different pentasaccharides tested were able to trigger immune responses on Arabidopsis seedlings and plants, the responses induced by XA3XX being the most intense ones



**FIGURE 2 |** Pentose-based oligosaccharides trigger cytoplasmic calcium elevations. **(A)** Calcium influx measured as relative luminescence units (RLU) over the time in 8-d old *Arabidopsis Col-0<sup>AEC</sup>* seedlings after treatment with 0.5 mg/ml of chitin and arabinoxylan polysaccharide preparations. The total areas-under-the-curves were integrated and their values are represented at the right side of the panel. Data represent mean  $\pm$  SD (n=8) from one experiment representative of three independent ones with similar results. Statistically significant differences according to Student's t-test ( $p < 0.05$ ) compared to negative control (mock) are shown. **(B)** Preparation and fractionation pipeline of pentose-based oligosaccharides from wheat arabinoxylan (AX). **(C)** HPLC-ELSD chromatograms of purified oligosaccharide preparations. Peaks are labeled as nX, where "n" correspond to the number of pentoses contained. Double peaks correspond to alpha-/beta-anomeric isomers at the reducing end of each detected oligosaccharide. **(D)** Calcium influxes after treatment with 0.5 mg/ml of the GH11-digested wheat AX before chromatographic purifications (WAX+GH11), chitohexaose (Chi6) and the purified pentose-based oligosaccharides (#1–6). The total areas-under-the-curves were integrated and their values are represented at the right side of the panel. Data represent mean  $\pm$  SD (n=8) from one experiment representative of three independent ones with similar results. Statistically significant differences according to Student's t-test ( $p < 0.05$ ) compared to negative control (WAX+GH11) are shown.



**FIGURE 3 |** Pure arabinoxylan (AX) oligosaccharides trigger early immune responses in Arabidopsis. **(A)** Molecular structures of the different AX oligosaccharides used in the experiments. 3<sup>3</sup>- $\alpha$ -L-arabinofuranosyl-xylotetraose (XA3XX), 2<sup>3</sup>- $\alpha$ -L-arabinofuranosyl-xylotetraose (XA2XX), 2<sup>3,3</sup>-di- $\alpha$ -L-arabinofuranosyl-xylotriose (A2,3XX), and 2<sup>3</sup>- $\alpha$ -L-arabinofuranosyl-xylotriose (A2XX). Thin layer chromatography profiles of the pure commercial AXs used in the experiments using 1-propanol/ethyl-acetate/water (9:7:4 by volume) as mobile phase. Left and right lanes show xylose (Xyl), xylotriose (Xyl3), xylohexaose (Xyl6) and arabinose (Ara) which were used as markers. TLC shown is from one run representative of three independent ones with similar results. **(B)** Calcium influx measured as relative luminescence units (RLU) over the time in 8-d-old Arabidopsis Col-0<sup>AEQ</sup> seedlings after treatment with 500  $\mu$ M of pure AX oligosaccharides. Water (mock) was used as negative control. The total areas-under-the-curves were integrated and their values are represented at the right side of the panel. Data represent mean  $\pm$  SD ( $n=8$ ) from one experiment representative of three independent ones with similar results. Statistically significant differences according to Student's  $t$ -test ( $p < 0.05$ ) compared to negative control are shown. **(C)** Reactive oxygen species (ROS) production (by Luminal reaction) after treatment with 500  $\mu$ M of pure AX oligosaccharides in Arabidopsis leaf-discs measured as RLU over the time. Water (mock), GH11-digested wheat AX before chromatographic purifications (WAX+GH11) and chitohexaose (Chi6; 100  $\mu$ M) were used as negative (mock and GH11+WAX) and positive (Chi6) controls. The total areas-under-the-curves were integrated and their values are represented at the right side of the panel. Data represent mean  $\pm$  SD ( $n=8$ ) from one experiment representative of three independent ones with similar results. Statistically significant differences according to Student's  $t$ -test ( $p < 0.05$ ) compared to water control are shown. **(D, E)** Cross elicitation during the refractory period of calcium signaling between XA3XX (250  $\mu$ M) and chitohexaose (Chi6; 100  $\mu$ M). Data show the elevation of cytoplasmic calcium concentration, measured as relative luminescence units (RLU), over the time in 8-d-old Arabidopsis Col-0<sup>AEQ</sup> seedlings after treatments. Dashed line (600 s) indicates the application time of the second elicitor within the refractory period of the first elicitation. In **(E)**, blue line represents a first treatment of XA3XX followed by Chi6 after 600 s, while black line represents a first treatment with Chi6 followed by a second of XA3XX. Data represent mean ( $n=8$ ) from one experiment representative of three independent ones with similar results.

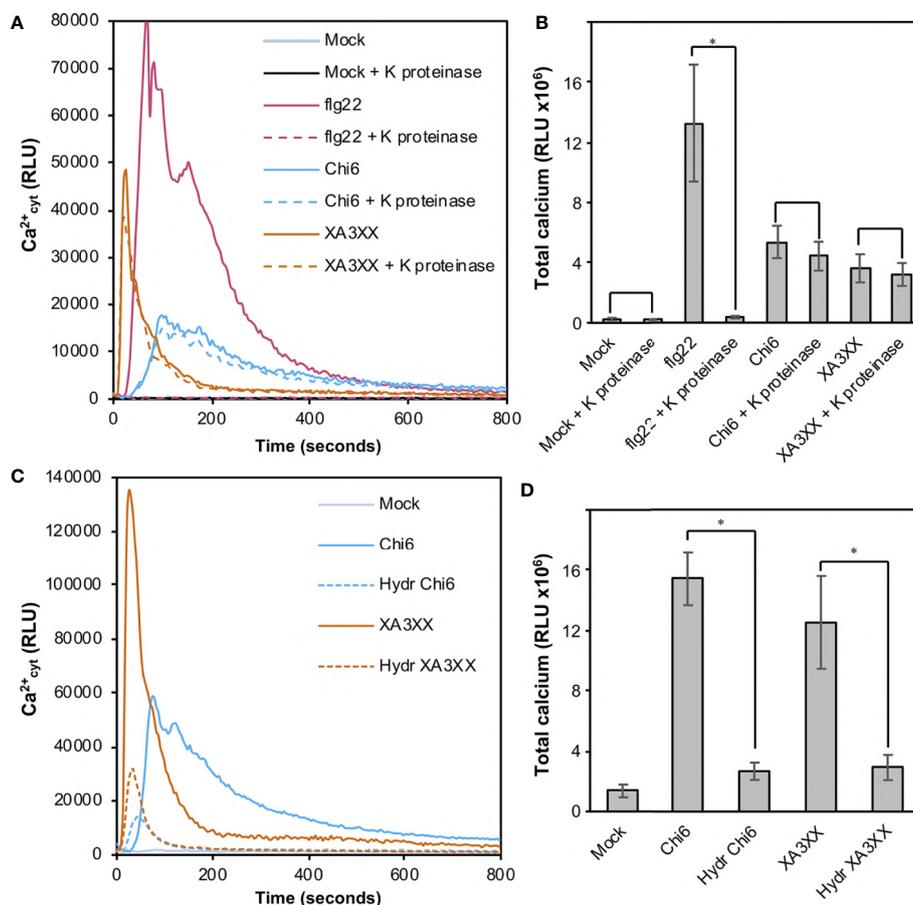
(Figures 3B, C). Interestingly, cross-elicitation experiments, by sequential application of two compounds in 600 s interval, demonstrated that fractions #4 and #5 and commercial XA3XX had a refractory period of Ca<sup>2+</sup> influx. Notably, this effect was not observed when the well characterized MAMP chitohexaose was

used in the experiments (Figures 3D, E and Supplementary Figure S2). Although a refractory state does not necessarily indicate the same perception mechanism or receptor, these results indicated that these pentose oligosaccharides ranging from DP 3 to 5 have equivalent activities and, at least in

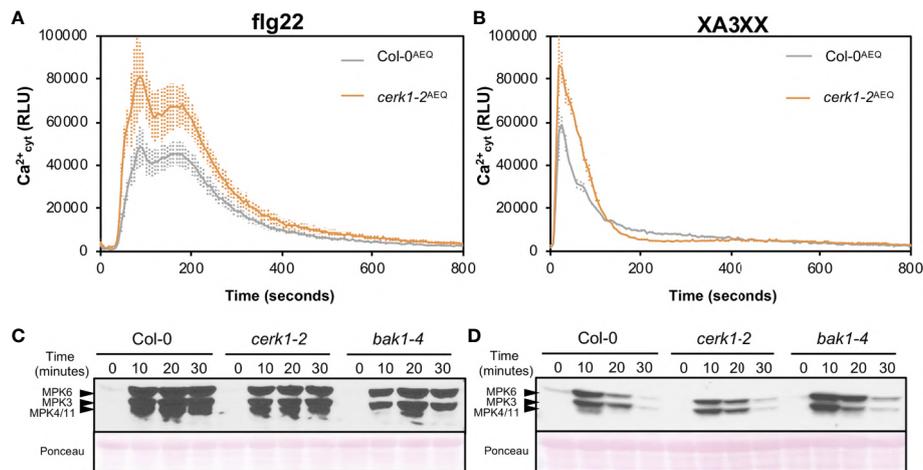
Arabidopsis, differ from chitin-based signaling. According to manufacturer XA3XX is a pure carbohydrate, but given the low required doses of peptide MAMP/DAMPs to be perceived by plants (Stegmann et al., 2017), we performed proteinase K proteolytic digestions of XA3XX solutions, and of two MAMPs solutions (chitohexaose and flg22), used as controls (Figures 4A, B). Of note, both carbohydrate-based elicitors remained fully active in the  $\text{Ca}^{2+}$  system after proteolytic treatment, which contrasted with the immune activity abolishment observed in case of flg22 after proteinase K treatment (Figures 4A, B). On the other hand, in order to confirm that XA3XX elicitor activity was linked to its oligomer structure, acid hydrolysis of the oligosaccharide was performed. Our results indicated that the hydrolyzed XA3XX lost its capacity to trigger  $\text{Ca}^{2+}$  influxes as it was the case for chitohexaose (Figures 4C, D). All these data, together with the observation that XA3XX was also able to trigger a ROS burst in soybean plants (Supplementary Figure S3), clearly indicated that this pentasaccharide deserved a more detailed investigation.

## Arabinofuranosyl-Xylotetraose (XA3XX) Activates Several PTI Hallmarks Through a Novel Plant Sensing Mechanism

In order to further characterize the early immune responses triggered by XA3XX, we performed a more detailed analysis of the  $\text{Ca}^{2+}$  kinetics following the application of the pentasaccharide and the peptide MAMP flg22 (Figures 5A, B). Flg22 induced a double  $\text{Ca}^{2+}$  burst peak at about 90 and 180 s followed by a maintained decrease in luminescence that lasted about 600 s (Figure 5A). However, XA3XX kinetics was very different, with a very fast single peak at 20 s post-application and a rapid signal lost at about 200 s (Figure 5B). In Arabidopsis, lysin motif-(LysM)-PRR CERK1 (Chitin Elicitor Receptor Kinase 1) plays a central role as a co-receptor for several glycan MAMPs such as chitin, peptidoglycan a  $\beta$ -1,3-glucans (Willmann et al., 2011; Liu et al., 2012; Mélida et al., 2018). The use of *cerk1-2* mutant aequorin lines demonstrated that the perception of XA3XX, like that of flg22, was CERK1-independent, which was in line with the refractory experiments with XA3XX/chitin (Figures 5A, B).



**FIGURE 4 |** Proteolysis and acid hydrolysis effects on XA3XX calcium signaling. Variation of intracellular  $\text{Ca}^{2+}$  concentration after 800 s of treatment of Col-0<sup>AEG</sup> seedlings with untreated flg22 (1  $\mu\text{M}$ ), chitohexaose (Chi6; 100  $\mu\text{M}$ ), or XA3XX (250  $\mu\text{M}$ ) (A, B) with or without proteinase K previous digestion and (C, D) with or without acid hydrolysis previous digestion. Data represent mean  $\pm$  SD from one experiment representative of two independent ones with similar results. Statistically significant differences according to Student's t-test ( $p < 0.05$ ).



**FIGURE 5 |** Pattern-triggered immunity hallmarks comparison between flg22 and arabinofuranosyl-xylotetraose (XA3XX). Flg22 (A, B) and XA3XX (C, D) final concentrations were 1 and 500  $\mu\text{M}$  respectively. (A, B) Elevations of cytoplasmic calcium concentrations over time in 8-d-old Arabidopsis Col-0<sup>AEG</sup> and *cerk1-2*<sup>AEG</sup> seedlings upon treatments. Data represent mean  $\pm$  SD (n=8) from one experiment representative of three independent ones with similar results. (C, D) MAPK activation in 12-d-old Arabidopsis seedlings of the indicated genotypes. The phosphorylation of MPK6, MPK3, and MPK4/MPK11 was determined by Western blot, using the anti-pTepY antibody, at the indicated time points (minutes) after treatment. Ponceau red-stained membranes show equal loading. Western-blot shown is from one experiment representative of three independent ones with similar results.

Next, we monitored phosphorylation of downstream protein kinases (MPK3/MPK6/MPK4/MPK11), a PTI hallmark, in Col-0 wild-type plants and *cerk1-2* and *bak1-4* mutants (in Col-0 background) impaired in PRR co-receptors required for the perception of chitin and flg22, respectively (Figures 5C, D). These analyses confirmed that XA3XX recognition by Arabidopsis plants was CERK1-independent and demonstrated that it was also BAK1-independent, while a partial BAK1-dependence for flg22 was observed, as described (Chinchilla et al., 2007). Western-blot assays confirmed MPK3- and MPK6-phosphorylation after application of a XA3XX solution to Arabidopsis seedlings, reaching the highest level of phosphorylation at 10 min post-treatment (Figure 5D). MPK4/11-phosphorylation was almost not-detectable in XA3XX-treated plants. MPKs phosphorylation levels of plants treated with XA3XX was weaker than phosphorylation of MPK3/MPK6/MPK4/MPK11 at all time points tested after elicitation with flg22 (Figure 5C).

Global gene reprogramming is the expected output of earlier PTI events such as  $\text{Ca}^{2+}$  influxes, ROS production and MAPK phosphorylation. Such alteration in the expression patterns of specific genes would determine the adaptation ability of a given plant to respond to a potential infection by pathogens. To further characterize the basis of XA3XX-mediated immunity, we performed RNA-seq analyses of Col-0 seedlings treated for 30 min with XA3XX (Figure 6 and Supplementary Tables S1 and S2). Elicitation with XA3XX changed the expression of 511 genes, most of which (460) were up-regulated (Supplementary Table S1). XA3XX up-regulated genes mainly grouped into gene ontology (GO) terms related to innate immune and defense response to different stimuli, kinase and signal transduction activities, and indole-containing compound metabolic processes

(Figure 6A), further corroborating the function of XA3XX in modulating PTI. We validated RNA-seq data of five PTI-marker genes (*CYP81F2*, *WRKY53*, *PHI1*, *FRK1*, and *NHL10*) by qRT-PCR in seedlings 30 min after treatment, and we found that all these genes were up-regulated after XA3XX elicitation compared to mock-treated seedlings (Figure 6B) confirming a full PTI response of Arabidopsis plants treated with XA3XX. Together, these analyses suggest that XA3XX-induced responses are addressed to a global immune response.

### XA3XX Crop Pre-Treatment Diminishes Pathogen Disease Symptoms Through a Non-Yet Characterized PRR Complex

Exposure of plants to active MAMP/DAMPs prior to subsequent pathogen attack may allow a more efficient plant defense activation through PTI activation (Héloir et al., 2019; Schellenberger et al., 2019). We showed that XA3XX was perceived by Arabidopsis and soybean (Figure 3 and Supplementary Figure S3). Therefore, we tested the elicitor activity of XA3XX in three-week-old tomato plants (MoneyMaker) treated by foliar spray with XA3XX 2 d before inoculation with the biotroph *Pseudomonas syringae* pv *tomato* DC3000 ( $10^8$  cfu/ml). Notably, bacterial population, determined as colony forming units (cfu) per leaf area, was significantly reduced in tomato XA3XX-pretreated plants compared to mock-treated plants (Figure 7A). Bacterial growth reduction found at 5 d post-inoculation (dpi) were in the order of 0.8–0.9 log of cfu/cm<sup>2</sup> when 0.25 and 0.5 mg of XA3XX per tomato plant were applied as pre-treatment, respectively (Figure 7A). Previous studies using similar approaches have also shown protection results of carbohydrate-based DAMPs against fungal necrotrophs (Claverie et al., 2018).

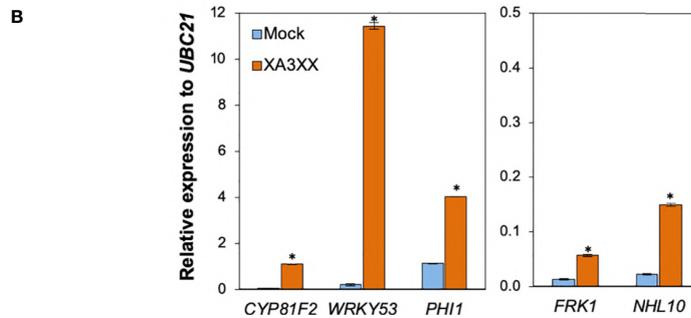
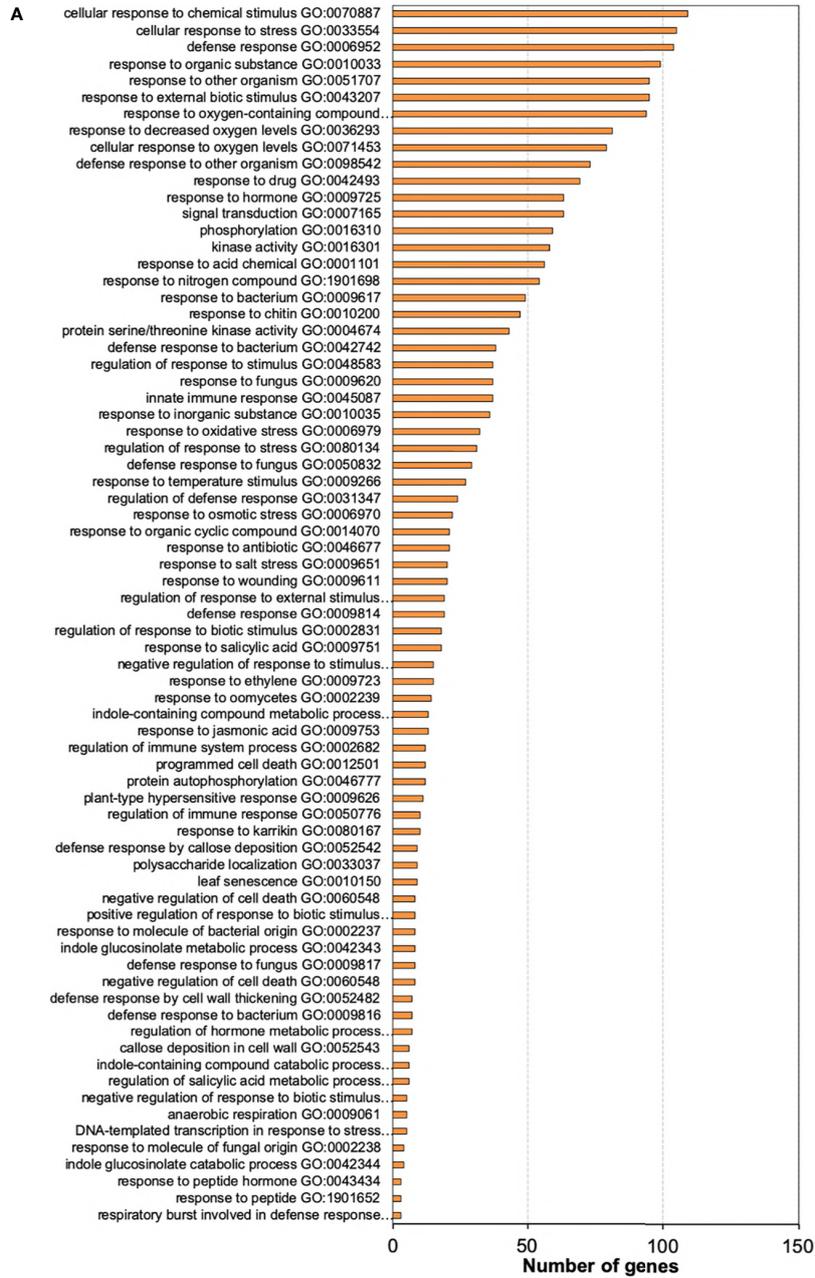
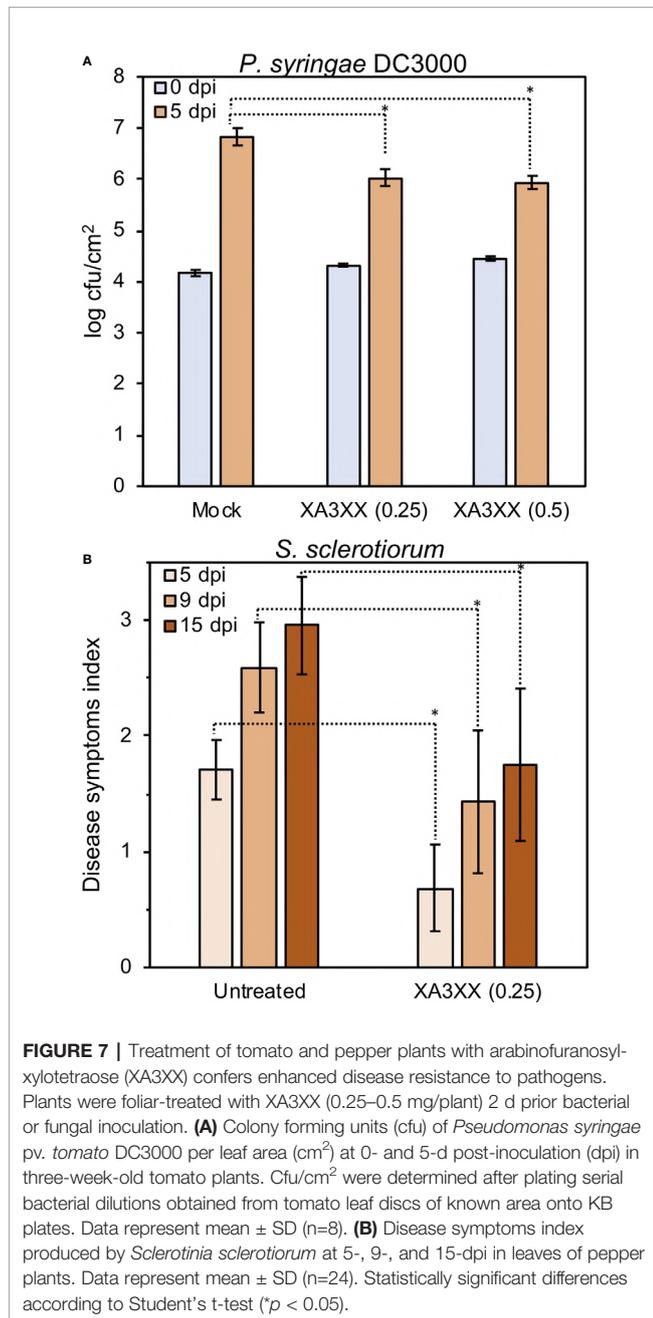


FIGURE 6 | Continued

**FIGURE 6** | Functional classification of arabinofuranosyl-xylotetraose (XA3XX)-differentially expressed genes in Arabidopsis. **(A)** Biological process Gene Ontology (GO) term enrichment map of the 460 overexpressed genes in 12-d-old Arabidopsis Col-0 plants at 30 min after 250  $\mu$ M XA3XX treatment. GO term enrichment is expressed as number of mapped genes. Data used to build the histogram can be retrieved from **Supplementary Table S2**. **(B)** RNA-seq data validation by quantitative RT-PCR analysis in 12-d-old Arabidopsis seedlings. Relative expression levels to *UBC21* (*At5g25769*) gene at 30 min are shown. Data represent mean  $\pm$  SD (n=3) from one experiment representative of two independent ones with similar results. Statistically significant differences between XA3XX and mock according to Student's t-test ( $p < 0.005$ ).



We next tested XA3XX-pre-treated pepper plants against the necrotroph fungi *Sclerotinia sclerotiorum*. Treated pepper plants showed a reduction in the disease symptoms index at 5, 9, and 15 dpi in comparison to control plants (**Figure 7B**). These data

indicate that XA3XX is able to trigger immune responses in some dicot crops conferring disease resistance.

## DISCUSSION

Plant pathogens and their hosts have co-evolved an arsenal of CWDE to break-down the opponent's cell wall during the interactions (Rovenich et al., 2016). Thus, plant and microbial cell walls are rich sources of carbohydrate-based defense signaling molecules (DAMPs/MAMPs), that are under-characterized. Recently, we have showed that impairment of *ARR6* gene in Arabidopsis affects cell wall composition, which may lead to the accumulation of DAMPs that would favor a "defense-ready" state, thus affecting plant-pathogen interactions (Bacete et al., 2020). Remarkably, pectin-enriched cell wall fractions extracted from *arr6* cell walls resulted to be enriched in carbohydrate-based DAMPs compared to wild-type fractions. However, the composition of such DAMPs could not be deciphered. In this work we have attempted to unveil the nature of these DAMPs. Unexpectedly, analytical data obtained from size-exclusion chromatography purifications suggested, for the first time, that pentose-based oligosaccharides co-extracted with pectins (using calcium chelators) could play a role as plant DAMPs. Next, we asked whether pentose-based oligosaccharides could be a novel group of plant DAMPs, and indeed we have demonstrated here that AXs can be perceived as molecular patterns by plants. In particular, we identified several active oligosaccharides structures, with pentasaccharide XA3XX being the most active one (**Figure 3**). In particular, XA3XX triggers  $\text{Ca}^{2+}$  influxes, ROS production, MAPK phosphorylation, and a global gene reprogramming in Arabidopsis at micromolar concentrations. Bearing in mind that the presence of glucuronoAXs has been suggested to be a component of Arabidopsis cell walls (Zablackis et al., 1995), XA3XX and related structures characterized in this work could be considered as plant DAMPs even in Arabidopsis model species.

Structurally, the most similar plant DAMPs characterized so far would be other  $\beta$ -1,4-linked hemicelluloses such as xyloglucan and mannan (Claverie et al., 2018; Zang et al., 2019). Xyloglucan recently proposed as DAMP consists of a  $\beta$ -1,4-glucan backbone associated with xylosyl, galactosyl, and fucosyl-type branching, mainly of DP 7 (Claverie et al., 2018). The purified xyloglucan triggered MAPK phosphorylation and immune-associated gene expression in Arabidopsis and *Vitis vinifera*, but no ROS production was found (Claverie et al., 2018). In another recent work, Zang et al., produced mannan oligosaccharides (DP 2–6) by enzymatic hydrolysis of locust bean gum and demonstrated their DAMP potential on *Nicotiana*

*benthamiana* and *Oryza sativa* (Zang et al., 2019). Mannan oligosaccharides triggered  $\text{Ca}^{2+}$  influxes, ROS production, stomata closure, and over-expression of defense-related genes such as *PR-1a* and *LOX*. Interestingly, these novel groups of hemicellulosic DAMPs, including xyloglucans, mannans, and AXs display the same type of glycosidic linkage in their backbone ( $\beta$ -1,4-linked) as the previously characterized plant DAMPs pectic OGs and cello-oligosaccharides (Aziz et al., 2007; Ferrari et al., 2013; Benedetti et al., 2015; Souza et al., 2017; Johnson et al., 2018). In contrast to cello-oligomers (DAMP) and chitin oligosaccharides (MAMP), which are actively triggering broad immune responses in the low micromolar range, the rest of glycan cell wall derived DAMPs characterized (including XA3XX) are only active at high micromolar concentrations, although the activities of the DAMPs/MAMPs are difficult to compare since experiments were performed in different labs, using different experimental setups and species (Aziz et al., 2007; Claverie et al., 2018; Johnson et al., 2018; Mélida et al., 2018; Zang et al., 2019; and this work). It should also be noted that cell wall polysaccharides can be very abundant in plant cells and they could yield very high concentrations of their main components, such as cellulose and xylans (over 5% of *Arabidopsis* fresh weight; Sakamoto and Mitsuda, 2015), and of the DAMPs derived from these polymers. Anyway, in spite of the high doses required of these novel groups of hemicellulosic DAMPs to be perceived by plants as such, they were able to enhance plant protection against different plant pathogens. Xyloglucan effectively protected grapevine and *Arabidopsis* against the necrotrophic fungus *Botrytis cinerea* or the oomycete *Hyaloperonospora arabidopsidis* pathogens while mannans improved rice protection against the bacteria *Xanthomonas oryzae* and the oomycete *Phytophthora nicotianae*, respectively (Claverie et al., 2018; Zang et al., 2019). In this work we have shown the protection capacity of XA3XX on tomato and pepper plants against bacterial plant pathogen *P. syringae* and the fungus *S. sclerotiorum* (Figure 7).

Xylans are main hemicelluloses of dicot secondary cell walls whose presence is essential for plant development, as exemplified in *Arabidopsis* plants with reduced xylan quantity, which show weakened cell walls and are unable to develop a vascular system (Brown et al., 2007; Wu et al., 2009). The importance of xylans in plant resistance to pathogens has been suggested previously, though the molecular bases of xylan-associated resistance phenotypes were largely unknown. For example, *Arabidopsis* plants with enhanced levels of xylose in their cell walls, as it occurs in *Arabidopsis de-etiolated3 (det3)* and *irregular xylem6 (irx6)* mutants (Brown et al., 2005; Rogers et al., 2005) or with modifications in their xyloglucan structure, as it is the case of the *Arabidopsis xyl1-2* mutant (Sampedro et al., 2010), show an enhanced resistance to the necrotrophic fungus *Plectosphaerella cucumerina* (Delgado-Cerezo et al., 2012). In contrast, *Arabidopsis er* plants, impaired in ERECTA Receptor-Like Kinase, and *agb1* and *agg1 agg2* mutants, impaired in the G $\beta$  and G $\gamma$  subunits of heterotrimeric G proteins, that are hypersusceptible to the same necrotrophic fungus, show a reduced xylose content (Llorente et al., 2005; Sánchez-

Rodríguez et al., 2009; Delgado-Cerezo et al., 2012). Also, alteration of cell wall xylan acetylation caused by *Arabidopsis* ESKIMO1 impairment was shown to enhance plant disease resistance to several pathogens, including *P. cucumerina* (Escudero et al., 2017). Whether these modifications in cell wall xylans are linked to an enhanced pentose-based DAMPs release from weakened walls (increased resistance) or to the alteration of pathogen capability to penetrate host tissues upon secretion of their CWDE repertoires are two interesting questions to address in future works. CWDE able to hydrolyze xylan polysaccharides to DAMPs such as those described in this work, are endo-1,4- $\beta$ -xylanases belonging to GH families 10 and 11 (McCleary et al., 2015). In particular, studies on GH11  $\beta$ -xylanases crystal structures, such as that from *N. patriciarum* used in this work, showed that the  $\alpha$ -L-Araf can be accommodated on O-2 and O-3, thus being able to release structures such as XA3XX and XA2XX (Vardakou et al., 2008). However, *Arabidopsis* only displays a handful of GH10 endo-xylanases in its genome, but any GH11 (see CAZy database at [www.cazy.org/e1.html](http://www.cazy.org/e1.html)). GH10 endo-xylanases would cleave arabinose-decorated non-reducing ends instead of xylose-free ones as is the case of XA3XX (Suzuki et al., 2002; McCleary et al., 2015). Therefore, our hypothesis is that the activity of GH10 and GH11 xylanases from pathogens might be associated to the release of xylan-derived DAMPs such as XA3XX and that such release might be differential in cell wall mutants displaying a modified architecture which could make some structures more or less accessible to GHs from pathogens, that could be the case of *arr6* mutants (Bacete et al., 2020). Indeed, bacterial and fungal endo-1,4- $\beta$ -xylanases have been shown to be required for full virulence of plant pathogens such as *B. cinerea* and *Xanthomonas* (Brito et al., 2006; Santos et al., 2014).

Notably, we show here that XA3XX is perceived by dicot crops, like soybean, tomato, and pepper, supporting that other plant species than *Arabidopsis* harbor the mechanisms required to perceive xylan-derived DAMPs. This perception, at least in the case of *Arabidopsis*, is independent of the co-receptors CERK1 and BAK1, further indicating that the mechanism of AX perception differs from that of chitin and  $\beta$ -1,3-glucans (Liu et al., 2012; Cao et al., 2014; Mélida et al., 2018). Taken together that arabinofuranose substitutions are less frequent in dicot than in monocot xylans and that these hemicelluloses are quite more abundant in the monocot branch, it will be interesting to test whether these and related pentose-based DAMPs trigger stronger or lower (if any) responses in plant species at different phylogenetic positions than those included in this study (all dicots). Regarding MAMPs, only a minor fraction of them (flg22, peptidoglycan, and chitin) are recognized by PRRs that are widespread among plants and can be found in both monocot and dicot plant species (Albert et al., 2020). Sensor systems for such patterns are considered an ancient set of PRRs, however the majority of PRRs known to date exhibit genus-specific distribution patterns. On the other hand, PRR-independent perception mechanisms for carbohydrate-based DAMPs could have been evolutionary selected. Therefore, future work in the characterization of the perception mechanisms and the specific

immune pathways triggered by AXs in different species will be necessary to unveil their functions and to determine if this is part of an additional mechanism of inter plant species recognition. Our findings support the use of carbohydrate-based DAMPs/MAMPs as biological products for the regulation of crops immunity and disease resistance responses. The use of these biological products in agriculture production can contribute to reach the social demand of a more sustainable agriculture.

## MATERIALS AND METHODS

### Biologic Material and Growth Conditions

All *Arabidopsis* lines used in this study were in the Columbia-0 (Col-0) background. *Arabidopsis* plants used for  $[Ca^{2+}]_{cyt}$ , MAPKs and gene expression analyses were grown in 24-well plates (10 seedlings per well) under long day conditions (16 h of light) at 20–22°C in liquid MS medium [0.5x Murashige & Skoog basal salt medium (Duchefa), 0.25% sucrose, 1 mM MES, pH 5.7]. Soil-grown *Arabidopsis* plants used for cell wall isolation and ROS assays were maintained under short day conditions (10 h of light). Tomato (*Solanum lycopersicum*, Moneymaker), pepper (*Capsicum annuum*, Murano), and soybean (*Glycine max*, Annushka) plants were grown in soil under greenhouse conditions.

### Statistical Methods

As a general rule, data shown are means  $\pm$  standard deviation (SD) from a given number of replicates ( $n \geq 3$ ). Data was normally retrieved from one representative independent out of three, however, given the particularity of each assay specific details are indicated in figure footnotes and in specific method subsections below. Asterisks indicate significant differences according to Student's t-test analysis, \*  $p \leq 0.05$  (R software).

### Carbohydrates

Details about carbohydrates used in this work can be found in **Supplementary Table S3**. AX polysaccharides (from wheat and rye), oligosaccharides (XA3XX, XA2XX, XUXX, A23XX, A2XX), and chitohexaose ( $\beta$ -1,4-D-(GlcNAc)<sub>6</sub>; Chi6) were purchased from Megazyme. Xylan (from birch, beet, and oat) and arabinan (from sugar beet) were purchased from Sigma-Aldrich.

### Arabidopsis Cell Wall Fractionation

Three-week-old *Arabidopsis* plants ( $n > 50$ ) were collected and immediately frozen in liquid nitrogen. Cell walls and their fractions were prepared as previously described (Bacete et al., 2017). The pectin-I fractions were size-fractionated by size-exclusion chromatography on Sepharose CL-6B (GE Healthcare, 140 ml bed-volume in a 1.6 cm diameter column) in 0.33 M sodium acetate buffer (pH 5.0). The column was connected to a Biologic-LP instrument (Bio-Rad) and the flow rate was 1.8 ml/min. The Sepharose column was calibrated with commercial dextrans (Sigma) of known weight-average relative molecular mass. Resulting sub-fractions were dialysed (Spectra/Por MWCO 1000 Daltons, Repligen) against deionized water to remove solutes of a small molecular mass (dialysis tubings were

thoroughly washed before use to eliminate any contaminants potentially associated to the membranes). The entire process was repeated three times.

### Xylan Enzymatic Digestion and Oligosaccharides Purification

Five hundred mg of low viscosity wheat flour AX (P-WAXYL, Ara : Xyl 38:62) were added to 24.5 ml of deionized water at 60°C and dissolved by stirring on a magnetic stirrer until complete dissolution. Then, the solution was equilibrated to 40°C and 0.5 ml of 0.5 M sodium phosphate buffer, pH 6, were added. This solution was placed in a water bath at 40°C and 97.5 U of endo-1,4- $\beta$ -D-xylanase from *Neocallimastix patriciarum* (Megazyme #E-XYLNP) were added and incubated at 40°C for 16 h. Reactions were terminated by incubating the solutions at 95°C for 5 min. Solutions were centrifuged at 9,400 g for 10 min to remove insoluble materials. Digestion products were freeze-dried, desalted and pre-purified using a Sephadex G-10 column (90 cm<sup>3</sup> bed-volume in a 1.5 cm diameter column; Merck) and size-fractionated using a Biogel P2 Extrafine column (140 cm<sup>3</sup> bed-volume in a 1.6 cm diameter column; BioRad). Columns were connected to a Biologic-LP instrument, distilled water was used as mobile phase and the flow rates were 0.24 ml/min. The entire process was repeated three times.

### Carbohydrate Analysis

The dried purified cell wall fractions (0.5 mg) were hydrolyzed in the presence of 2 M trifluoroacetic acid (TFA) at 121°C for 3 h. *Myo*-inositol was used as an internal standard. The resulting monosaccharides were converted to alditol acetates (Albersheim et al., 1967). Derivatized monosaccharides were separated and analyzed by gas chromatography (GC) on a SP-2380 capillary column (30 m x 0.25 mm i.d.; Supelco) using a Scion 450-GC system equipped with EVOQ triple quadrupole (Bruker). The temperature programme increased from 165°C to 270°C at a rate of 2°C min<sup>-1</sup>. MALDI-TOF MS analyses were performed using a 4800 Plus Proteomics Analyzer MALDI-TOF/TOF mass spectrometer (Applied Biosystems, MDS Sciex) as described (Mélida et al., 2018). Technical replicates were considered from the TFA hydrolysis step of a given cell wall fraction from the same extraction procedure.

Purified oligosaccharides were monitored by thin layer chromatography (TLC) and high-performance liquid chromatography (HPLC). Spotted-AXs were run twice on TLC Silicagel 60 plates (Merck) using 1-propanol/ethyl-acetate/water (9:7:4 by volume) as mobile phase. TLC plates were developed by dipping in a solution of 0.5% (w/v) thymol and 5% (v/v) H<sub>2</sub>SO<sub>4</sub> in 96% (v/v) ethanol and heated at 80 °C for 5–8 min. The HPLC-ELSD analysis was performed as previously described (Senf et al., 2017). The oligosaccharides were injected into an Agilent 1200 Series HPLC equipped with an Agilent 6130 quadrupole mass spectrometer (MS) and an Agilent 1200 Evaporative Light Scattering Detector (ELSD). The purified oligosaccharides were separated on a graphitized carbon Hypercarb column (150 x 4.6 mm, Thermo Scientific) using a water (including 0.1% formic acid)-acetonitrile (ACN) gradient.

The peaks in the ELSD traces were assigned based on their retention time and the corresponding masses in the MS. For additional MS analyses, a fraction of each oligosaccharide sample was injected directly into an Agilent 1260 Infinity II Series, LC/MSD XT (Single Quadrupole mit ESI-Jetstream-source).

### Aequorin Luminescence Measurements

*Arabidopsis* 8-d-old liquid-grown transgenic seedlings of ecotype Col-0 carrying the calcium reporter aequorin (Col-0<sup>AEQ</sup>; Ranf et al., 2012) were used for cytoplasmic calcium ([Ca<sup>2+</sup>]<sub>cyt</sub>) measurements using the method previously described (Bacete et al., 2017). Negative controls (water) were included in all the experiments. Aequorin luminescence was recorded with a Varioskan Lux Reader (Thermo Scientific). Data shown represent mean ± SD (n=8 seedlings) from one experiment representative of at least three independent ones with similar results.

### Reactive Oxygen Species

Five-week-old *Arabidopsis* or 6-week-old soybean plants were used to determine ROS production after treatments using the luminol assay (Escudero et al., 2017) and a Varioskan Lux luminescence reader (Thermo Scientific). Data shown represent mean ± SD (n=8 leaf discs from at least 4 different plants) from one experiment representative of three independent ones with similar results.

### Immunoblot Analysis of MAPK Activation

Twelve-day-old seedlings (n=10) grown on liquid MS medium in 24-well plates were treated with water (mock) and oligosaccharides for 0, 10, 20, and 30 min, and then harvested in liquid nitrogen. Protein extraction and detection of activated MAPKs using the Phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204) antibody (Cell Signaling Technology) were performed as described (Ranf et al., 2011). Western-blot shown is from one experiment representative of three independent ones with similar results.

### Gene Expression Analyses

For gene expression analysis (qRT-PCR and RNA sequencing), 12-d-old seedlings grown on liquid MS medium were treated with the oligosaccharide or water (mock) solutions for 0 and 30 min. Total RNA was purified with the RNeasy Plant Mini Kit (Qiagen) according to the manufacturer's protocol. qRT-PCR analyses were performed as previously reported (Delgado-Cerezo et al., 2012). *UBC21* (*At5g25760*) expression was used to normalize the transcript level in each reaction. Oligonucleotides used for detection of gene expression are detailed on **Supplementary Table S4**. Analysis of mock-treated seedlings showed no alterations in the expression levels of the marker genes used in this study. Data shown represent mean ± SD (n=3) from one experiment representative of two independent ones with similar results.

For RNA-seq analyses, samples from three biological replicates for each treatment were sequenced using 50bp Illumina HiSeq 2500. RNA-seq read raw data can be retrieved from the NCBI Sequence Read Archive (SRA) under BioProject accession ID PRJNA639010 (<http://www.ncbi.nlm.nih.gov/bioproject/639010>). Transcripts obtained were aligned against *Arabidopsis* annotation Araport11 (Cheng et al., 2017) using Hisat2 2.10.0 release (Kim et al., 2019). Afterwards, they were processed using Stringtie v1.3.6 (Pertea et al., 2015) and Ballgown R packages (Frazee et al., 2015) as previously described (Pertea et al., 2016). Differential expression analysis was performed with FPKM (Fragments Per Kilobase of transcript per Million mapped reads) values from the treatment against FPKM mock values in order to obtain the n-fold. For the up-regulated genes, a coverage cutoff of 50% of the dataset was applied to the treatment genes while for the down-regulated genes it was applied to the mock genes. N-fold of above or equal than 2 was used to prove up-regulation and an n-fold below or equal than 0.5 was applied to look for down-regulated genes. ClueGO 2.5.6 app for Cytoscape (Bindea et al., 2009) was used to determine which Gene Ontology (GO) categories were statistically overrepresented in the differentially expressed set of genes. Significant enrichments were determined using the Enrichment/Depletion (Two-sided hypergeometric) test and Bonferroni step down corrected *p* values are represented. Additional parameters are detailed in **Supplementary Table S2**.

### Crop Protection Assays

Three-week-old tomato plants (*Solanum lycopersicum*, MoneyMaker) were sprayed with 2 ml of a XA3XX solution (0.125 or 0.25 mg/ml) containing 2.5% UEP-100 (Croda) and 2.5% Tween 24 MBAL (Croda) as adjuvants. Adjuvant solution was used as mock. *Pseudomonas syringae* pv. *tomato* DC3000 infections were performed 48 h after pre-treatments according mainly to Santamaría-Hernando et al., 2019. Briefly, plants were sprayed with a suspension of the bacterium (10<sup>8</sup> cfu/ml) and two tomato leaf discs were collected from four different plants at 0- and 5-d post-infection (dpi). Colony forming units (cfu) per foliar area (cm<sup>2</sup>) were determined after plating serial bacterial dilutions obtained from tomato leaf discs of known area onto KB plates with rifampicin (25 µg/ml). Data shown represent mean ± SD (n=8) from one experiment representative of three independent ones with similar results. For *Sclerotinia sclerotiorum* experiments, 5-weeks-old pepper plants (*Capsicum annuum*, Murano) were treated using 5 ml of a XA3XX solution (0.05 mg/ml) containing 0.5% UEP-100 and 0.05% Tween 24 MBAL as adjuvants. Two-days after treatment, plants were moved to a 75% humidity greenhouse chamber and spray-inoculated with 5 ml of a 250 cfu/ml suspension of *S. sclerotiorum* homogenized mycelia according to Chen and Wang (2005). Disease symptoms were determined at 5 and 9 dpi in all the leaves of each plant (n=24) using a scale from 0 to 4 where 0 = no symptoms; 1 = little necrotic spots (< 20% of leaf area); 2 = two or more notable necrotic spots (20–50% of leaf area); 3 = more than 50% of leaf area affected, 4 = leaf senescence. Data shown represent mean ± SD (n=24) from three experiments.

## DATA AVAILABILITY STATEMENT

RNA-seq read raw data can be retrieved from the NCBI Sequence Read Archive (SRA) under BioProject accession ID PRJNA639010.

## AUTHOR CONTRIBUTIONS

HM and AM initiated, conceived and coordinated all the experiments. HM performed most of the experiments with support of GL, LB, and DR. IH performed the RNA-seq data analysis. CR and FP performed the HPLC-ELSD and MS experiments. HM prepared figures and tables. HM and AM wrote the paper with contributions of all the authors. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by grants IND2017/BIO-7800 of the Comunidad de Madrid Regional Government, BIO2015-64077-

## REFERENCES

- Albersheim, P., Nevins, D. J., English, P. D., and Karr, A. (1967). A method for the analysis of sugars in plant cell-wall polysaccharides by gas-liquid chromatography. *Carbohydr. Res.* 5, 340–345. doi: 10.1016/S0008-6215(00)80510-8
- Albert, I., Hua, C., Nürnberger, T., Pruitt, R. N., and Zhang, L. (2020). Surface Sensor Systems in Plant Immunity. *Plant Physiol.* 182, 1582–1596. doi: 10.1104/pp.19.01299
- Annis, S. L., and Goodwin, P. H. (1997). Production and regulation of polygalacturonase isozymes in Canadian isolates of *Leptosphaeria maculans* differing in virulence. *Can. J. Plant Pathol.* 19, 358–365. doi: 10.1080/07060669709501060
- Atkinson, N. J., and Urwin, P. E. (2012). The interaction of plant biotic and abiotic stresses: from genes to the field. *J. Exp. Bot.* 63, 3523–3543. doi: 10.1093/jxb/ers100
- Aziz, A., Gauthier, A., Bézier, A., Poinssot, B., Joubert, J. M., Pugin, A., et al. (2007). Elicitor and resistance-inducing activities of beta-1,4 cellobextrins in grapevine, comparison with beta-1,3 glucans and alpha-1,4 oligogalacturonides. *J. Exp.* 58, 1463–1472. doi: 10.1093/jxb/erm008
- Bacete, L., Mélida, H., Pattathil, S., Hahn, M. G., Molina, A., and Miedes, E. (2017). “Characterization of Plant Cell Wall Damage-Associated Molecular Patterns Regulating Immune Responses,” in *Plant Pattern Recognition Receptors, Methods in Molecular Biology*, vol. 1578. Eds. L. Shan and P. He (New York, NY: Humana Press), 13–23.
- Bacete, L., Mélida, H., Miedes, E., and Molina, A. (2018). Plant cell wall-mediated immunity: cell wall changes trigger disease resistance responses. *Plant J.* 93, 614–636. doi: 10.1111/tpj.13807
- Bacete, L., Mélida, H., López, G., Dabos, P., Tremousaygue, D., Denancé, N., et al. (2020). Arabidopsis Response Regulator 6 (ARR6) modulates plant cell wall composition and disease resistance. *Mol. Plant Microbe Interact.* 33, 767–780. doi: 10.1094/MPMI-12-19-0341-R
- Benedetti, M., Pontiggia, D., Raggi, S., Cheng, Z., Scaloni, F., Ferrari, S., et al. (2015). Plant immunity triggered by engineered *in vivo* release of oligogalacturonides, damage-associated molecular patterns. *Proc. Natl. Acad. Sci. U.S.A.* 112, 5533–5538. doi: 10.1073/pnas.1504154112
- Benedetti, M., Verrascina, I., Pontiggia, D., Locci, F., Mattei, B., De Lorenzo, G., et al. (2018). Four Arabidopsis berberine bridge enzyme-like proteins are specific oxidases that inactivate the elicitor-active oligogalacturonides. *Plant J.* 94, 260–273. doi: 10.1111/tpj.13852

R of the Spanish Ministry of Economy and Competitiveness (MINECO), RTI2018-096975-B-I00 of Spanish Ministry of Science, Innovation and Universities, to AM. This work has been also financially supported by the “Severo Ochoa Programme for Centres of Excellence in R&D” from the Agencia Estatal de Investigación of Spain (grant SEV-2016-0672 (2017–2021) to the CBGP). In the frame of this program HM was supported with a postdoctoral fellow. DR was the recipient of an Industrial PhD Fellow (IND2017/BIO-7800) and IH was the recipient of an PhD FPU fellow from the Spanish Ministry of Education (FPU16/07118). FP thanks the Max Planck Society and the German Research Foundation (DFG, Emmy Noether program PF850/1-1 to FP) for financial support.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2020.01210/full#supplementary-material>

- Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., et al. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 25, 1091–1093. doi: 10.1093/bioinformatics/btp101
- Boller, T., and Felix, G. (2009). A Renaissance of Elicitors: Perception of Microbe-Associated Molecular Patterns and Danger Signals by Pattern-Recognition Receptors. *Annu. Rev. Plant Biol.* 60, 379–406. doi: 10.1146/annurev.arplant.57.032905.105346
- Boudsocq, M., Willmann, M. R., McCormack, M., Lee, H., Shan, L., He, P., et al. (2010). Differential innate immune signalling via Ca(2+) sensor protein kinases. *Nature* 464, 418–422. doi: 10.1038/nature08794
- Boutrot, F., and Zipfel, C. (2017). Function, Discovery, and Exploitation of Plant Pattern Recognition Receptors for Broad-Spectrum Disease Resistance. *Annu. Rev. Phytopathol.* 55, 257–286. doi: 10.1146/annurev-phyto-080614-120106
- Brito, N., Espino, J. J., and González, C. (2006). The endo- $\beta$ -1,4-xylanase Xyn11A is required for virulence in *Botrytis cinerea*. *Mol. Plant Microbe Interact.* 19, 25–32. doi: 10.1094/MPMI-19-0025
- Brown, D. M., Zeef, L. A. H., Ellis, J., Goodacre, R., and Turner, S. R. (2005). Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell* 17, 2281–2295. doi: 10.1105/tpc.105.031542
- Brown, D. M., Goubet, F., Wong, V. W., Goodacre, R., Stephens, E., Dupree, P., et al. (2007). Comparison of five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis. *Plant J.* 52, 1154–1168. doi: 10.1111/j.1365-313X.2007.03307.x
- Cao, Y., Liang, Y., Tanaka, K., Nguyen, C. T., Jedrzejczak, R. P., Joachimiak, A., et al. (2014). The kinase LYK5 is a major chitin receptor in Arabidopsis and forms a chitin-induced complex with related kinase CERK1. *Elife* 3, e03766. doi: 10.7554/eLife.03766
- Carpita, N. C., and Gibeau, D. M. (1993). Structural models of primary-cell walls in flowering plants - consistency of molecular-structure with the physical-properties of the walls during growth. *Plant J.* 3, 1–30. doi: 10.1111/j.1365-313X.1993.tb00007.x
- Carpita, N., and McCann, M. (2000). *The cell wall* (Rockville, MD: American Society of Plant Physiologists).
- Chen, Y., and Wang, D. (2005). Two Convenient Methods to Evaluate Soybean for Resistance to Sclerotinia sclerotiorum. *Plant Dis.* 89, 1268–1272. doi: 10.1094/PD-89-1268
- Cheng, C. Y., Krishnakumar, V., Chan, A. P., Thibaud-Nissen, F., Schobel, S., and Town, C. D. (2017). Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. *Plant J.* 89, 789–804. doi: 10.1111/tpj.13415

- Chinchilla, D., Zipfel, C., Robatzek, S., Kemmerling, B., Nürnberger, T., Jones, J. D., et al. (2007). A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defense. *Nature* 448, 497–500. doi: 10.1038/nature05999
- Choi, H. W., and Klessig, D. F. (2016). DAMPs, MAMPs, and NAMPs in plant innate immunity. *BMC Plant Biol.* 16, 232. doi: 10.1186/s12870-016-0921-2
- Choi, J., Choi, D., Lee, S., Ryu, C. M., and Hwang, I. (2011). Cytokinins and plant immunity: old foes or new friends? *Trends Plant Sci.* 16, 388–394. doi: 10.1016/j.tplants.2011.03.003
- Claverie, J., Balacey, S., Lemaître-Guillier, C., Brulé, D., Chiltz, A., Granet, L., et al. (2018). The Cell Wall-Derived Xyloglucan Is a New DAMP Triggering Plant Immunity in *Vitis vinifera* and *Arabidopsis thaliana*. *Front. Plant Sci.* 9, 1725. doi: 10.3389/fpls.2018.01725
- Darvill, J. E., McNeil, M., Darvill, A. G., and Albersheim, P. (1980). Structure of plant cell walls. XI. Glucuronarabinoxylan, a second hemicellulose in the primary cell walls of suspension-cultured sycamore cells. *Plant Physiol.* 66, 1135–1139. doi: 10.1104/pp.66.6.1135
- De Lorenzo, G., Ferrari, S., Cervone, F., and Okun, E. (2018). Extracellular DAMPs in Plants and Mammals: Immunity, Tissue Damage and Repair. *Trends Immunol.* 39, 937–950. doi: 10.1016/j.it.2018.09.006
- De Lorenzo, G., Ferrari, S., Giovannoni, M., Mattei, B., and Cervone, F. (2019). Cell wall traits that influence plant development, immunity, and bioconversion. *Plant J.* 97, 134–147. doi: 10.1111/tj.14196
- Delgado-Cerezo, M., Sánchez-Rodríguez, C., Escudero, V., Miedes, E., Fernández, V. P., Jorda, L., et al. (2012). Arabidopsis Heterotrimeric G-protein Regulates Cell Wall Defense and Resistance to Necrotrophic Fungi. *Mol. Plant* 5, 98–114. doi: 10.1093/mp/ssp082
- Dodds, P. N., and Rathjen, J. P. (2010). Plant immunity: towards an integrated view of plant–pathogen interactions. *Nat. Rev. Gen.* 11, 539–548. doi: 10.1038/nrg2812
- Duran-Flores, D., and Heil, M. (2016). Sources of specificity in plant damaged-self recognition. *Curr. Opin. Plant Biol.* 32, 77–87. doi: 10.1016/j.pbi.2016.06.019
- Ellis, C., Karafyllidis, I., Wasternack, C., and Turner, J. G. (2002). The Arabidopsis mutant *cev1* links cell wall signaling to jasmonate and ethylene responses. *Plant Cell* 14, 1557–1566. doi: 10.1105/tpc.002022
- Engelsdorf, T., Gigli-Bisceglia, N., Veerabagu, M., McKenna, J. F., Vaahtera, L., Augstein, F., et al. (2018). The plant cell wall integrity maintenance and immune signaling systems cooperate to control stress responses in *Arabidopsis thaliana*. *Sci. Signal.* 11 (536), pii: eao3070. doi: 10.1126/scisignal.aao3070
- Escudero, V., Jordá, L., Sopena-Torres, S., Mélida, H., Miedes, E., Muñoz-Barrios, A., et al. (2017). Alteration of cell wall xylan acetylation triggers defense responses that counterbalance the immune deficiencies of plants impaired in the  $\beta$ -subunit of the heterotrimeric G-protein. *Plant J.* 92, 386–399. doi: 10.1111/tj.13660
- Ferrari, S., Galletti, R., Pontiggia, D., Manfredini, C., Lionetti, V., Bellincampi, D., et al. (2008). Transgenic expression of a fungal endo-polygalacturonase increases plant resistance to pathogens and reduces auxin sensitivity. *Plant Physiol.* 146, 669–681. doi: 10.1104/pp.107.109686
- Ferrari, S., Savatin, D. V., Sicilia, F., Gramagna, G., Cervone, F., and Lorenzo, G. (2013). Oligogalacturonides: plant damage-associated molecular patterns and regulators of growth and development. *Front. Plant Sci.* 4, 49. doi: 10.3389/fpls.2013.00049
- Fischer, M. H., Yu, N., Gray, G. R., Ralph, J., Anderson, L., and Marlett, J. A. (2004). The gel-forming polysaccharide of psyllium husk (*Plantago ovata* Forsk.). *Carbohydr. Res.* 339, 2009–2017. doi: 10.1016/j.carres.2004.05.023
- Fraze, A. C., Perrea, G., Jaffe, A. E., Langmead, B., Salzberg, S. L., and Leek, J. T. (2015). Ballgown bridges the gap between transcriptome assembly and expression analysis. *Nat. Biotechnol.* 33, 243–246. doi: 10.1038/nbt.3172
- Gallego-Giraldo, L., Liu, C., Pose-Albacete, S., Pattathil, S., Peralta, A. G., Young, J., et al. (2020). ARABIDOPSIS DEHISCENCE ZONE POLYGALACTURONASE 1 (ADPG1) releases latent defense signals in stems with reduced lignin content. *Proc. Natl. Acad. Sci. U.S.A.* 117, 3281–3290. doi: 10.1073/pnas.1914422117
- Hamann, T., Bennett, M., Mansfield, J., and Somerville, C. (2009). Identification of cell-wall stress as a hexose-dependent and osmosensitive regulator of plant responses. *Plant J.* 57, 1015–1026. doi: 10.1111/j.1365-313X.2008.03744.x
- Héloir, M. C., Adrian, M., Brulé, D., Claverie, J., Cordelier, S., Daire, X., et al. (2019). Recognition of elicitors in grapevine: from MAMP and DAMP perception to induced resistance. *Front. Plant Sci.* 10, 1117. doi: 10.3389/fpls.2019.01117
- Hernández-Blanco, C., Feng, D. X., Hu, J., Sanchez-Vallet, A., Deslandes, L., Llorente, F., et al. (2007). Impairment of Cellulose Synthases Required for Arabidopsis Secondary Cell Wall Formation Enhances Disease Resistance. *Plant Cell* 19, 890–903. doi: 10.1105/tpc.106.048058
- Houston, K., Tucker, M. R., Chowdhury, J., Shirley, N., and Little, A. (2016). The plant cell wall: a complex and dynamic structure as revealed by the responses of genes under stress conditions. *Front. Plant Sci.* 7, 984. doi: 10.3389/fpls.2016.00984
- Johnson, J. M., Thürich, J., Petutschnig, E. K., Altschmied, L., Meichsner, D., Sherameti, I., et al. (2018). Poly(A) ribonuclease controls the cellobiose-based interaction between *Piriformospora indica* and its host Arabidopsis. *Plant Physiol.* 176, 2496–2514. doi: 10.1104/pp.17.01423
- Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotech.* 37, 907–915. doi: 10.1038/s41587-019-0201-4
- Li, Q., Wang, C., and Mou, Z. (2020). Perception of Damaged Self in Plants. *Plant Physiol.* 182, 1545–1565. doi: 10.1104/pp.19.01242
- Lionetti, V., Fabri, E., De Caroli, M., Hansen, A. R., Willats, W. G. T., Piro, G., et al. (2017). Three pectin methyltransferase inhibitors protect cell wall integrity for Arabidopsis immunity to Botrytis. *Plant Physiol.* 173, 1844–1863. doi: 10.1104/pp.16.01185
- Liu, T., Liu, Z., Song, C., Hu, Y., Han, Z., She, J., et al. (2012). Chitin-Induced Dimerization Activates a Plant Immune Receptor. *Science* 336, 1160–1164. doi: 10.1126/science.1218867
- Llorente, F., Alonso-Blanco, C., Sánchez-Rodríguez, C., Jorda, L., and Molina, A. (2005). ERECTA receptor-like kinase and heterotrimeric G protein from Arabidopsis are required for resistance to the necrotrophic fungus *Plectosphaerella cucumerina*. *Plant J.* 43, 165–180. doi: 10.1111/j.1365-313X.2005.02440.x
- Locci, F., Benedetti, M., Pontiggia, D., Citterico, M., Caprari, C., Mattei, B., et al. (2019). An Arabidopsis berberine bridge enzyme-like protein specifically oxidizes cellulose oligomers and plays a role in immunity. *Plant J.* 98, 540–554. doi: 10.1111/tj.14237
- McCleary, B. V., McKie, V. A., Draga, A., Rooney, E., Mangan, D., and Larkin, J. (2015). Hydrolysis of wheat flour arabinoxylan, acid-debranched wheat flour arabinoxylan and arabino-xylo-oligosaccharides by  $\beta$ -xylanase,  $\alpha$ -L-arabinofuranosidase and  $\beta$ -xylosidase. *Carbohydr. Res.* 407, 79–96. doi: 10.1016/j.carres.2015.01.017
- Mélida, H., Largo-Gosens, A., Novo-Uzal, E., Santiago, R., Pomar, F., García, P., et al. (2015). Ectopic lignification in primary cellulose-deficient cell walls of maize cell suspension cultures. *J. Integr. Plant Biol.* 57, 357–372. doi: 10.1111/jipb.12346
- Mélida, H., Sopena-Torres, S., Bacete, L., Garrido-Arandia, M., Jordá, L., López, G., et al. (2018). Non-branched  $\beta$ -1,3-glucan oligosaccharides trigger immune responses in Arabidopsis. *Plant J.* 93, 34–49. doi: 10.1111/tj.13755
- Miedes, E., Vanholme, R., Boerjan, W., and Molina, A. (2014). The role of the secondary cell wall in plant resistance to pathogens. *Front. Plant Sci.* 5, 358x. doi: 10.3389/fpls.2014.00358
- Molina, A., Miedes, E., Bacete, L., Rodríguez, T., Mélida, H., Denancé, N., et al. (2020). Arabidopsis cell wall composition determines disease resistance specificity and fitness. *bioRxiv*. doi: 10.1101/2020.05.21.105650
- Mort, A. J., Moerschbacher, B. M., Pierce, M. L., and Maness, N. O. (1991). Problems encountered during the extraction, purification, and chromatography of pectic fragments, and some solutions to them. *Carbohydr. Res.* 215, 219–227. doi: 10.1016/0008-6215(91)84022-7
- Nafisi, M., Fimognari, L., and Sakuragi, Y. (2015). Interplays between the cell wall and phytohormones in interaction between plants and necrotrophic pathogens. *Phytochemistry* 112, 63–71. doi: 10.1016/j.phytochem.2014.11.008
- Naran, R., Chen, G., and Carpita, N. C. (2008). Novel rhamnogalacturonan I and arabinoxylan polysaccharides of flax seed mucilage. *Plant Physiol.* 148, 132–141. doi: 10.1104/pp.108.123513
- Nothnagel, E. A., McNeil, M., Albersheim, P., and Dell, A. (1983). Host-Pathogen Interactions: XXII. A galacturonic acid oligosaccharide from plant cell walls elicits phytoalexins. *Plant Physiol.* 71, 916–926. doi: 10.1104/pp.71.4.916
- Perrea, M., Perrea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122
- Perrea, M., Kim, D., Perrea, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* 11, 1650–1667. doi: 10.1038/nprot.2016.095

- Raiola, A., Lionetti, V., Elmaghraby, I., Immerzeel, P., Mellerowicz, E. J., Salvi, G., et al. (2011). Pectin methyltransferase is induced in Arabidopsis upon infection and is necessary for a successful colonization by necrotrophic pathogens. *Mol. Plant Microbe Interact.* 24, 432–440. doi: 10.1094/MPMI-07-10-0157
- Ranf, S., Eschen-Lippold, L., Pecher, P., Lee, J., and Scheel, D. (2011). Interplay between calcium signalling and early signalling elements during defence responses to microbe- or damage-associated molecular patterns. *Plant J.* 68, 100–113. doi: 10.1111/j.1365-313X.2011.04671.x
- Ranf, S., Grimmer, J., Pöschl, Y., Pecher, P., Chinchilla, D., Scheel, D., et al. (2012). Defense-related calcium signaling mutants uncovered via a quantitative high-throughput screen in *Arabidopsis thaliana*. *Mol. Plant* 5, 115–130. doi: 10.1093/mp/ssr064
- Ridley, B. L., O'Neill, M. A., and Mohnen, D. (2001). Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. *Phytochemistry* 57, 929–967. doi: 10.1016/S0031-9422(01)00113-3
- Rogers, L. A., Dubos, C., Surman, C., Willment, J., Cullis, I. F., Mansfield, S. D., et al. (2005). Comparison of lignin deposition in three ectopic lignification mutants. *New Phytol.* 168, 123–140. doi: 10.1111/j.1469-8137.2005.01496.x
- Rovenich, H., Zuccaro, A., and Thomma, B. P. (2016). Convergent evolution of filamentous microbes towards evasion of glycan-triggered immunity. *New Phytol.* 212, 896–901. doi: 10.1111/nph.14064
- Sakamoto, S., and Mitsuda, N. (2015). Reconstitution of a Secondary Cell Wall in a Secondary Cell Wall-Deficient Arabidopsis Mutant. *Plant Cell Physiol.* 56, 299–310. doi: 10.1093/pcp/pcu208
- Sampedro, J., Pardo, B., Gianzo, C., Guitian, E., Revilla, G., and Zarra, I. (2010). Lack of alpha-Xylosidase Activity in Arabidopsis Alters Xyloglucan Composition and Results in Growth Defects. *Plant Physiol.* 154, 1105–1115. doi: 10.1104/pp.110.163212
- Sánchez-Rodríguez, C., Estevez, J. M., Llorente, F., Hernández-Blanco, C., Jordá, L., Pagan, I., et al. (2009). The ERECTA Receptor-Like Kinase Regulates Cell Wall Mediated Resistance to Pathogens in Arabidopsis thaliana. *Mol. Plant Microbe Interact.* 22, 953–963. doi: 10.1094/MPMI-22-8-0953
- Santamaría-Hernando, S., Senovilla, M., González-Mula, A., Martínez-García, P. M., Nebreda, S., Rodríguez-Palenzuela, P., et al. (2019). The *Pseudomonas syringae* pv. tomato DC3000 PSPTO\_0820 multidrug transporter is involved in resistance to plant antimicrobials and bacterial survival during tomato plant infection. *PLoS One* 14, e0218815. doi: 10.1371/journal.pone.0218815
- Santos, C. R., Hoffmann, Z. B., de Matos Martins, V. P., Zanphorlin, L. M., de Paula Assis, L. H., Honorato, R. V., et al. (2014). Molecular mechanisms associated with xylan degradation by Xanthomonas plant pathogens. *J. Biol. Chem.* 289, 32186–32200. doi: 10.1074/jbc.M114.605105
- Schellenberger, R., Touchard, M., Clément, C., Baillieul, F., Cordelier, S., Crouzet, J., et al. (2019). Apoplastic invasion patterns triggering plant immunity: plasma membrane sensing at the frontline. *Mol. Plant Pathol.* 20, 1602–1616. doi: 10.1111/mpp.12857
- Scheller, H. V., and Ulvskov, P. (2010). Hemicelluloses. *Annu. Rev. Plant Biol.* 61, 263–289. doi: 10.1146/annurev-arplant-042809-112315
- Senf, D., Ruprecht, C., de Kruijff, G. H., Simonetti, S. O., Schuhmacher, F., Seeberger, P. H., et al. (2017). Active Site Mapping of Xylan-Deconstructing Enzymes with Arabinoxylan Oligosaccharides Produced by Automated Glycan Assembly. *Chemistry* 23, 3197–3205. doi: 10.1002/chem.201605902
- Souza, C., Li, S., Lin, A. Z., Boutrot, F., Grossmann, G., Zipfel, C., et al. (2017). Cellulose-derived oligomers act as damage-associated molecular patterns and trigger defense-like responses. *Plant Physiol.* 173, 2383–2398. doi: 10.1104/pp.16.01680
- Srivastava, V., McKee, L. S., and Bulone, V. (2017). “Plant Cell Walls,” in *eLS* (Chichester, UK: John Wiley & Sons Ltd).
- Stegmann, M., Monaghan, J., Smakowska-Luzan, E., Rovenich, H., Lehner, A., Holton, N., et al. (2017). The receptor kinase FER is a RALF-regulated scaffold controlling plant immune signaling. *Science* 355, 287–289. doi: 10.1126/science.aal2541
- Suzuki, M., Kato, A., Nagata, N., and Komeda, Y. (2002). A xylanase, AtXyn1, is predominantly expressed in vascular bundles, and four putative xylanase genes were identified in the Arabidopsis thaliana genome. *Plant Cell Physiol.* 43, 759–767. doi: 10.1093/pcp/pcf088
- To, J. P., Deruere, J., Maxwell, B. B., Morris, V. F., Hutchison, C. E., Ferreira, F. J., et al. (2007). Cytokinin Regulates Type-A Arabidopsis Response Regulator Activity and Protein Stability via Two-Component Phosphorelay. *Plant Cell* 19, 3901–3914. doi: 10.1105/tpc.107.052662
- Vaahtera, L., Schulz, J., and Hamann, T. (2019). Cell wall integrity maintenance during plant development and interaction with the environment. *Nat. Plants* 5, 924–932. doi: 10.1038/s41477-019-0502-0
- Vardakou, M., Dumon, C., Murray, J. W., Christakopoulos, P., Weiner, D. P., Juge, N., et al. (2008). Understanding the structural basis for substrate and inhibitor recognition in eukaryotic GH11 xylanases. *J. Mol. Biol.* 375, 1293–1305. doi: 10.1016/j.jmb.2007.11.007
- Voxeur, A., Habrylo, O., Guenin, S., Miart, F., Soulie, M. C., Rihouey, C., et al. (2019). Oligogalacturonide production upon Arabidopsis thaliana-Botrytis cinerea interaction. *Proc. Natl. Acad. Sci. U. S. A.* 116, 19743–19752. doi: 10.1073/pnas.1900317116
- Willmann, R., Lajunen, H. M., Erbs, G., Newman, M. A., Kolb, D., Tsuda, K., et al. (2011). Arabidopsis lysin-motif proteins LYM1 LYM3 CERK1 mediate bacterial peptidoglycan sensing and immunity to bacterial infection. *Proc. Natl. Acad. Sci. U.S.A.* 108, 19824–19829. doi: 10.1073/pnas.1112862108
- Wolf, S., Hématy, K., and Höfte, H. (2012). Growth Control and Cell Wall Signaling in Plants. *Annu. Rev. Plant Biol.* 63, 381–407. doi: 10.1146/annurev-arplant-042811-105449
- Wu, A. M., Rihouey, C., Seveno, M., Hörnblad, E., Singh, S. K., Matsunaga, T., et al. (2009). The Arabidopsis IRX10 and IRX10-LIKE glycosyltransferases are critical for glucuronoxylan biosynthesis during secondary cell wall formation. *Plant J.* 57, 718–731. doi: 10.1111/j.1365-313X.2008.03724.x
- Xu, F., Liu, Z., Xie, H., Zhu, J., Zhang, J., Kraus, J., et al. (2014). Increased drought tolerance through the suppression of eskmo1 gene and overexpression of CBF-related genes in Arabidopsis. *PLoS One* 9, e106509. doi: 10.1371/journal.pone.0106509
- Zablackis, E., Huang, J., Müller, B., Darvill, A. G., and Albersheim, P. (1995). Structure of plant-cell walls: characterization of the cell-wall polysaccharides of *Arabidopsis thaliana* leaves. *Plant Physiol.* 107, 1129–1138. doi: 10.1104/pp.107.4.1129
- Zang, H., Xie, S., Zhu, B., Yang, X., Gu, C., Hu, B., et al. (2019). Mannan oligosaccharides trigger multiple defence responses in rice and tobacco as a novel danger-associated molecular pattern. *Mol. Plant Pathol.* 20, 1067–1079. doi: 10.1111/mpp.12811

**Conflict of Interest:** Authors DR and FB were employed by the company PlantResponse Biotech S.L.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Mélida, Bacete, Ruprecht, Rebaque, del Hierro, López, Brunner, Pfrengle and Molina. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Differential Expression of Fungal Genes Determines the Lifestyle of *Plectosphaerella* Strains During *Arabidopsis thaliana* Colonization

Antonio Muñoz-Barrios,<sup>1,2</sup> Sara Sopeña-Torres,<sup>1</sup> Brisa Ramos,<sup>1</sup> Gemma López,<sup>1</sup> Irene del Hierro,<sup>1,2</sup> Sandra Díaz-González,<sup>1,2</sup> Pablo González-Melendi,<sup>1,2</sup> Hugo Mélida,<sup>1</sup> Vanessa Fernández-Calleja,<sup>1</sup> Verónica Mixão,<sup>3,4</sup> Marina Martín-Dacal,<sup>1,2</sup> Marina Marcet-Houben,<sup>3,4</sup> Toni Gabaldón,<sup>3,4,5</sup> Soledad Sacristán,<sup>1,2,†</sup> and Antonio Molina<sup>1,2,†</sup>

<sup>1</sup> Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid (UPM)-Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Campus Montegancedo-UPM, 28223-Pozuelo de Alarcón (Madrid), Spain

<sup>2</sup> Departamento de Biotecnología-Biología Vegetal, Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas, Universidad Politécnica de Madrid, 28040-Madrid, Spain

<sup>3</sup> Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology, Dr. Aiguader 88, 08003 Barcelona, Spain

<sup>4</sup> Universitat Pompeu Fabra (UPF), 08003 Barcelona, Spain

<sup>5</sup> ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain

Accepted 25 July 2020.

The Whole-Genome Shotgun projects have been deposited in the DDBJ/ENA/GenBank database under accessions JACAFV000000000 (*Plectosphaerella* sp. strain P0831 Biosample SAMN14235337 TaxID: 40657), JACAFW000000000 (*P. cucumerina* Pc2127 Biosample SAMN14233823 TaxID: 40658), and JACAFX000000000 *P. cucumerina* (PcBMM Biosample SAMN14233534) TaxID: 40658, and BioProject number PRJNA609142. The versions described by this paper are versions JACAFV010000000, JACAFW010000000 and JACAFX010000000. The raw RNA-seq data in this study are deposited in the National Center for Biotechnology Information Sequence Read Archive (SRA), BioProject number PRJNA614936 and BioSamples SAMN14444085 to SAMN14444096, SAMN14444149, SAMN14444167, and SAMN14444168, with SRA accessions SRR11668188 to SRR11668202.

Current address for Toni Gabaldón: Barcelona Supercomputing Centre and Institute for Research in Biomedicine, Jordi Girona 29, 08034 Barcelona, Spain

†Corresponding authors: S. Sacristán; soledad.sacristan@upm.es and A. Molina; antonio.molina@upm.es

**Funding:** This work was supported by the Spanish Ministry of Economy and Competitiveness (MINECO) grant BIO2015-64077-R and the Spanish Research Agency (AEI) grant RTI2018-096975-B-I00 to A. Molina and by the “Severo Ochoa Programme for Centers of Excellence in R&D” grant SEV-2016-0672 (2017-2021) to the CBGP (UPM-INIA). In the frame of SEV-2016-0672 program, H. Mélida was supported with a postdoctoral contract. A. Muñoz-Barrios was financially supported by the Universidad Politécnica de Madrid (UPM) Ph.D. students PIF program, I. del Hierro was a FPU fellow (Spanish Ministry of Education, Culture and Sports grant FPU16/07118), V. Fernández-Calleja was supported by the Consejería de Educación e Investigación de Comunidad de Madrid YEI program for postdoctoral researchers (PEJD-2016/BIO-3327), and the work was further supported through a Comunidad de Madrid YEI program for laboratory technicians grant (PEJ16/BIO/TL-1570).

\*The e-Xtra logo stands for “electronic extra” and indicates there are supplementary figures and tables published online.

The author(s) declare no conflict of interest.

© 2020 The American Phytopathological Society

The fungal genus *Plectosphaerella* comprises species and strains with different lifestyles on plants, such as *P. cucumerina*, which has served as model for the characterization of *Arabidopsis thaliana* basal and nonhost resistance to necrotrophic fungi. We have sequenced, annotated, and compared the genomes and transcriptomes of three *Plectosphaerella* strains with different lifestyles on *A. thaliana*, namely, PcBMM, a natural pathogen of wild-type plants (Col-0), Pc2127, a nonpathogenic strain on Col-0 but pathogenic on the immunocompromised *cyp79B2 cyp79B3* mutant, and P0831, which was isolated from a natural population of *A. thaliana* and is shown here to be nonpathogenic and to grow epiphytically on Col-0 and *cyp79B2 cyp79B3* plants. The genomes of these *Plectosphaerella* strains are very similar and do not differ in the number of genes with pathogenesis-related functions, with the exception of secreted carbohydrate-active enzymes (CAZymes), which are up to five times more abundant in the pathogenic strain PcBMM. Analysis of the fungal transcriptomes in inoculated Col-0 and *cyp79B2 cyp79B3* plants at initial colonization stages confirm the key role of secreted CAZymes in the necrotrophic interaction, since PcBMM expresses more genes encoding secreted CAZymes than Pc2127 and P0831. We also show that P0831 epiphytic growth on *A. thaliana* involves the transcription of specific repertoires of fungal genes, which might be necessary for epiphytic growth adaptation. Overall, these results suggest that in-plant expression of specific sets of fungal genes at early stages of colonization determine the diverse lifestyles and pathogenicity of *Plectosphaerella* strains.

**Keywords:** *Arabidopsis*, CAZyme, epiphytic fungus, genome, immunity, pathogenicity, necrotroph, *Plectosphaerella*

Plants are continuously encountering a diverse array of microorganisms that differ in their ability to infect and cause disease (Brader et al. 2017; Newton et al. 2010; Zeilinger et al. 2016). Microbial pathogens have genetic repertoires required for the colonization of plant tissues and molecular tools to overcome plant immune responses (Rai and Agarkar 2016).

Nonpathogenic microorganisms may be able to survive on host surfaces, growing epiphytically and establishing specific types of interactions with their hosts (Leveau 2015; Lindow and Brandl 2003; Vorholt 2012; Whipps et al. 2008). Little is known about the genetic mechanisms underlying the epiphytic interactions of fungi with plants and their differences with pathogenic fungi, and similarly, our knowledge about the function of plant immunity in these epiphytic interactions is scarce (Rastogi et al. 2013; Vorholt 2012; Yao et al. 2019). However, fungal epiphytes represent a considerable fraction of the phyllosphere microbiome that can greatly affect plant fitness, either positively or negatively (Hongsanan et al. 2016; Vorholt 2012; Whipps et al. 2008). For example, it has been shown that fungal epiphytes can protect plants against foliar pathogens, either as microbial antagonists (Zhou et al. 2017) or by priming plant resistance (Buxdorf et al. 2013). On the other hand, epiphytic growth could be an important phase in the interaction with the plant prior to infection or could be used as a survival mechanism in the absence of a suitable host (Vorholt 2012). Comparing pathogenic and nonpathogenic plant-fungus interactions is an excellent strategy to find molecular determinants of pathogenicity and virulence, as shown by several recent studies (Baetsen-Young et al. 2020; Hacquard et al. 2016; Plett and Martin 2018; Zeilinger et al. 2016). Since the nature of nonpathogenic interactions is diverse, additional comparative studies are needed in other fungus-plant interactions to better understand the genetic bases of nonpathogenic lifestyles of fungi.

Fungal plant pathogens are classified as biotrophs, hemibiotrophs, and necrotrophs, each having different modes of interaction with their host plants (Horbach et al. 2011). Biotrophic pathogens are usually obligate parasites that do not kill host cells and establish sophisticated interactions with the host that include the secretion of effectors that manipulate the plant metabolism, suppress the immune responses, and promote the trophic interaction (Spanu and Panstruga 2017). Necrotrophic pathogens rapidly cause substantial tissue damage, killing host cells by a combination of mechanisms involving, among others, the expression of cell wall-degrading enzymes (CWDEs) that hydrolyze plant cell-wall polymers, the production of reactive oxygen species (ROS) and the secretion of toxins (Wang et al. 2014). These activities lead to cell wall and membrane disruption in the host cells and to the release of nutrients, which favor extensive colonization of a plant host by the fungus and tissue decomposition (Zeilinger et al. 2016). Despite being less-studied than biotrophic interactions, increasing evidence shows that the molecular mechanisms of plant and necrotrophic fungi interactions are complex and involve fungal effectors (Wang et al. 2014).

The Ascomycete genus *Plectosphaerella* includes several species commonly found in the rhizosphere and strains isolated from very different hosts, mainly plants, but also from insects, crustaceans, or nematodes (Giraldo and Crous 2019; Yu and Coosemans 1998). Pathogenic *Plectosphaerella* spp. have been widely reported to cause fruit, root, and collar rot on several crops, becoming an emergent pathogen in recent decades (Carlucci et al. 2012; Dillard et al. 2005; Jimenez and Zitter 2005; Su et al. 2017; Usami and Katagiri 2017; Vitale et al. 2007). However, *Plectosphaerella* species can also exhibit other lifestyles in plants, as endophytes colonizing plant tissues without causing visible symptoms (D'Amico et al. 2008; Götz et al. 2006) or as epiphytes showing antagonistic effects against bacterial pathogens (Zhou et al. 2017). *Plectosphaerella* strains have also been found in natural populations of *Arabidopsis thaliana*, either in pathogenic associations (Durán et al. 2018; Ton and Mauch-Mani 2004) or as endophytes in asymptomatic plants (García et al. 2013; Junker et al. 2012; Thiergart et al. 2020).

The interaction between *Arabidopsis* and the species *P. cucumerina* is a well-established pathosystem for the study of

plant basal and nonhost resistance to necrotrophic fungi (Ramos et al. 2013; Sánchez-Vallet et al. 2010). The analysis of this pathosystem has contributed to the identification of novel components of plant defense mechanisms. For example, it has been found that the biosynthesis of tryptophan (Trp)-derived metabolites (depleted in *cyp79B2 cyp79B3*, *pen2*, *cyp81f2*, and *cyp71A12* mutants) and their targeted delivery at pathogen contact sites (impaired in *pen3* mutant) are required for *Arabidopsis* basal resistance to both pathogenic (e.g., PcBMM) and nonpathogenic strains (e.g., Pc2127) of *P. cucumerina* (Bednarek et al. 2009; Hernández-Blanco et al. 2007; Lipka et al. 2005; Pastorczyk et al. 2020; Sánchez-Vallet et al. 2010; Stein et al. 2006). Additional signaling pathways are also involved in *Arabidopsis* resistance to the pathogenic strain PcBMM, like those mediated by the defense hormones salicylic acid (SA), jasmonate (JA), ethylene (ET), abscisic acid (ABA), and cytokinin, by heterotrimeric G protein and ERECTA receptor like kinase, or by signaling mechanisms triggered upon alteration of plant cell-wall integrity (Bacete et al. 2020; Berrocal-Lobo et al. 2002; Delgado-Cerezo et al. 2012; Hernández-Blanco et al. 2007; Llorente et al. 2005, 2008; Sánchez-Vallet et al. 2012). Also, immune responses triggered by microbe-associated molecular patterns (MAMPs) are required for *Arabidopsis* resistance to *P. cucumerina*. The cell walls of spores and mycelium of PcBMM contain several glycans, like chitin and 1,3- $\beta$ -glucans, that are perceived as MAMPs by plant pattern recognition receptors (PRRs) like CERK1, triggering transcriptional regulation of immune-related genes (Bacete et al. 2018; Mérida et al. 2018). Accordingly, *Arabidopsis cerk1* mutant is immunocompromised and shows enhanced susceptibility to *P. cucumerina* and to other fungi and oomycetes (Mérida et al. 2018; Wan et al. 2008).

To characterize the genetic determinants of the interactions between *Plectosphaerella* spp. and *Arabidopsis*, we have sequenced, annotated, and compared the genomes and transcriptomes of three fungal strains with different lifestyles (PcBMM, Pc2127, and P0831). PcBMM was isolated from *Arabidopsis* and is an adapted necrotrophic pathogen in all *Arabidopsis* genotypes tested, whereas the nonadapted Pc2127, isolated from a different host, is unable to colonize *Arabidopsis* wild-type plants (Col-0) but is pathogenic on immunocompromised double mutant *cyp79B2 cyp79B3* plants (Bednarek et al. 2009; Sánchez-Vallet et al. 2010). Notably, we show here that strain P0831, isolated from a leaf of an asymptomatic *Arabidopsis* plant from a natural population in central Spain (García et al. 2013), is nonpathogenic on Col-0 and *cyp79B2 cyp79B3* plants but is able to grow epiphytically on leaves of these genotypes. Here, we describe the whole sequencing and characterization of the genomes of these strains, showing that they have minor differences in the number of genes encoding pathogenesis-related functions. Our data point to secreted carbohydrate-active enzymes (CAZymes) as one of the main pathogenicity determinants of strain PcBMM. Our analyses also show that there are significant differences in the number and predicted function of fungal genes expressed in the interaction of the epiphytic P0831 strain with *Arabidopsis* plants compared with those established by the adapted and nonadapted strains.

## RESULTS

### *Plectosphaerella* strains display different lifestyles in *Arabidopsis* leaves.

We collected asymptomatic *Arabidopsis* plants from natural populations in central Spain and, by PCR, detected the genus *Plectosphaerella* in up to 50 and 27% of their leaves and roots, respectively (Supplementary Fig. S1). *Plectosphaerella* P0831 strain was isolated from the leaf of one of these asymptomatic

plants (García et al. 2013). The growth and pathogenicity patterns of strain P0831 on *Arabidopsis* Col-0 leaves were analyzed and compared with those of the well-characterized PcBMM and Pc2127 strains, which are pathogenic and non-pathogenic in Col-0 plants, respectively (Sánchez-Vallet et al. 2010; Ramos et al. 2013). P0831, like Pc2127, did not produce any disease symptoms on Col-0 plants, whereas PcBMM caused necrotic spots on leaves that spread and reached the vascular system, as previously described (Fig. 1A) (Ramos et al. 2013). By using fungal transformants of the three strains constitutively expressing the green fluorescence protein (GFP) (PcBMM-GFP and Pc2127-GFP [Ramos et al. 2013] and P0831-GFP [generated in this work]), we found that P0831-GFP spores germinated and formed a dense mycelium on Col-0 leaf surfaces that could be observed at 4 days postinoculation (dpi), whereas Pc2127 spores germinated on Col-0 leaf surfaces but did not form mycelia, as previously reported (Fig. 1B) (Ramos et al. 2013). Of note, besides the epiphytic growth of P0831-GFP, hyphae of P0831-GFP were observed in the intercellular space at the mesophyll at 4 dpi (Supplementary Fig. S2). Thus, we concluded that P0831 is not pathogenic on Col-0 leaves, being able to grow epiphytically and, occasionally, endophytically on this accession. Since the pathogenic PcBMM strain was also isolated from *Arabidopsis* plants and several *Plectosphaerella* sp. strains have recently been identified in the characterization of *Arabidopsis* fungal microbiomes (Durán et al. 2018; García et al. 2013; Junker et al. 2012; Thiergart et al. 2020; Ton and Mauch-Mani 2004), we can conclude that the genus *Plectosphaerella* can display different lifestyles on natural populations of *Arabidopsis*.

We next tested the interaction of P0831 with the *Arabidopsis* immunocompromised double mutant *cyp79B2 cyp79B3* that is depleted of all Trp-derived secondary metabolites required for basal resistance (Bednarek et al. 2009; Sánchez-Vallet et al. 2010). Notably, we found that P0831 also grew epiphytically, in this double mutant, without causing disease symptoms (Fig. 1A and B). This contrasted with the previously described lifestyle of the nonpathogenic strain Pc2127 that, like PcBMM, colonized and caused disease symptoms in *cyp79B2 cyp79B3* plants, as determined by quantitative PCR (qPCR) of the fungal  $\beta$ -tubulin gene (Fig. 1A and C) (Sánchez-Vallet et al. 2010). The slight growth of P0831 in *cyp79B2 cyp79B3* was further corroborated at different days postinoculation by trypan blue (TB) staining of inoculated leaves, which revealed that *cyp79B2 cyp79B3* leaves showed a very faint and localized TB staining, suggesting that necrosis was limited to a few plant cells. This contrasted with the intense TB stain, indicative of plant cell death, caused by Pc2127 and PcBMM strains (Supplementary Fig. S3). We also analyzed plant defense responses by determining the production of plant ROS upon fungal inoculation through diaminobenzidine (DAB) staining. We found a weak DAB stain in *cyp79B2 cyp79B3* leaves inoculated with Pc2127 and P0831, whereas no DAB staining was observed in Col-0 plants inoculated with P0831 (Supplementary Fig. S3). On the other hand, strong DAB stains were observed in Col-0 and *cyp79B2 cyp79B3* plants inoculated with PcBMM, which were indicative of an intense immune response triggered by the progression of fungal growth on leaves (Supplementary Fig. S3).

We next determined whether *Arabidopsis* plants were able to perceive the three *Plectosphaerella* strains and to activate immune responses such as the transcriptional upregulation of the *RbohD* gene, which encodes NADPH oxidase RBHOD involved in ROS production and that has been described to be upregulated by pathogen infection (e.g., PcBMM) or MAMP treatment (Mélida et al. 2018; Morales et al. 2016). *Arabidopsis pRbohD::LUC* lines (in Col-0 background) expressing the luciferase (*LUC*) gene under the control of RBOHD promoter (*pRbohD*) were inoculated with the fungal strains, and we

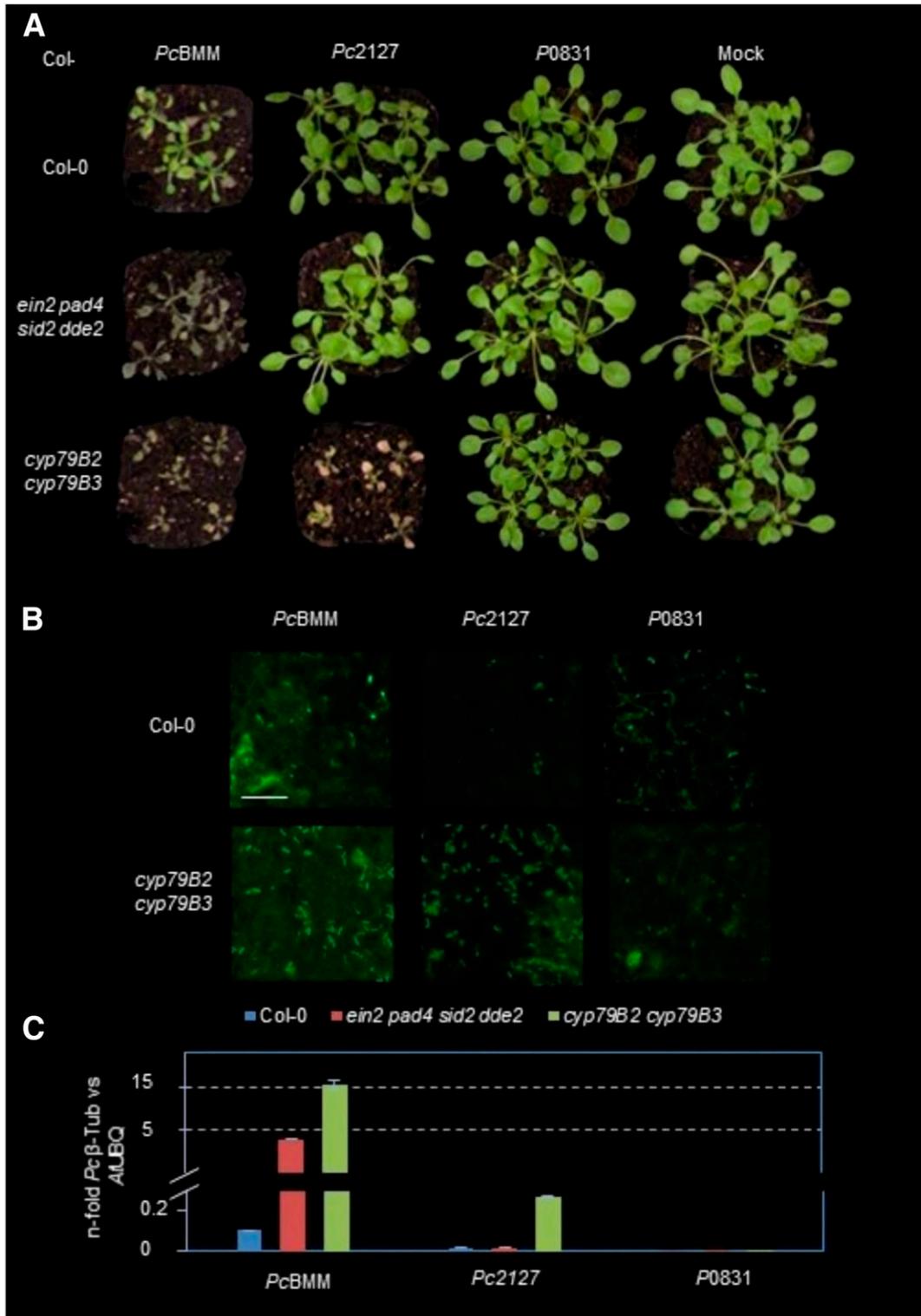
followed, in vivo, transcriptional regulation of *RbohD* by determining LUC bioluminescence. We found that bioluminescence at 3 dpi in P0831-inoculated plants was higher than in Pc2127-inoculated plants but lower than the bioluminescence observed in plants inoculated with PcBMM (Supplementary Fig. S4). Notably, LUC bioluminescence of P0831 in Col-0 was also stronger than that caused by the endophytic fungus *Colletotrichum tofieldiae* 0861 (Ct0861), which was also isolated from a natural *Arabidopsis* population in central Spain but is unable to grow on the leaf surface (García et al. 2013; Hiruma et al. 2016). These results indicated that Col-0 plants harbor immune mechanisms for the perception of the three *Plectosphaerella* strains (e.g., PRRs recognizing fungal MAMPs) and the transcriptional upregulation of genes associated with pathogen-associated molecular pattern-triggered immunity (PTI).

To further characterize the relevance of *Arabidopsis* defense mechanisms in shaping *Plectosphaerella* lifestyle, we tested the colonization of an *Arabidopsis ein2 pad4 sid2 dde2* quadruple mutant, defective in all three major phytohormone-dependent defense signaling pathways (ET, SA, and JA) and the effector triggered immunity (ETI) key regulator *PAD4* (Mine et al. 2018). Fungal biomass, determined by qPCR of the  $\beta$ -tubulin fungal gene, in the *ein2 pad4 sid2 dde2* mutant inoculated with either Pc2127 or P0831 was similar to that of Col-0 plants, which did not support fungal infection, whereas PcBMM growth in the mutant was higher than in Col-0 plants (Fig. 1C) (Berrocal-Lobo et al. 2002; Lipka et al. 2005; Sánchez-Vallet et al. 2010). Overall, these results confirm the central role of SA, ET, JA, and *PAD4*-mediated signaling pathways and Trp-derived metabolites in *Arabidopsis* basal resistance to the pathogenic PcBMM strain and corroborate that Trp-derived basal resistance is sufficient to limit the growth of the nonadapted Pc2127 strain (Sánchez-Vallet et al. 2010). Since impairing SA, ET, JA, and *PAD4* signaling pathways or Trp-derived basal resistance does not affect P0831 lifestyle and colonization of *Arabidopsis* plants, different immune mechanisms may be acting to limit the growth and potential virulence of P0831 epiphytic strain.

### Genomic features and phylogenetic relationships of *Plectosphaerella* strains.

High-molecular-weight genomic DNA was extracted from mycelia of the three *Plectosphaerella* strains grown in minimal media. This DNA was used for Illumina single-end (SE), paired-end (PE), and PacBio read-sequencing analyses, which produced sequences covering each fungal genome that were used for de novo genome assembly (Table 1). *Plectosphaerella* genomes sizes were estimated to be between 37.7 (PcBMM) and 35.9 (Pc2127) Mb (Table 1), which are comparable with the genome sizes of other ascomycete fungi reported (Supplementary Table S1). Gene models were predicted, using the Maker v2.31.10 (Cantarel et al. 2008) pipeline, that identified 11,323 (PcBMM), 11,007 (Pc2127) and 10,821 (P0831) genes in the fungal strains (Table 1). Next, we determined the gene space coverage, using BUSCO v3 (Simão et al. 2015), which indicated that 97.5 to 97.7% of the core conserved genes of the Sordariomycetes database were in the assembled genomes (Table 1).

We used a read mapping approach to compare Pc2127 and P0831 with PcBMM, and we detected the presence of 179,658 single nucleotide polymorphisms (SNPs) (163,473 exclusive) in Pc2127 and 812,707 SNPs (796,522 exclusive) in P0831. In the case of Pc2127, 94% of the proteins had more than 90% of the positions with reads in PcBMM and fewer than 0.2% of the proteins had no reads, which may correspond to putative gene deletions in Pc2127 with respect to PcBMM. P0831 had only 88% identity at the protein level with PcBMM, which contrasted with 98.6% identity found for Pc2127, suggesting that



**Fig. 1.** Differential interactions of PcBMM, Pc2127, and P0831 with *Arabidopsis* wild-type plants (Col-0) and immune-deficient mutants. Three-week-old Col-0 wild-type plants and *ein2 pad4 sid2 dde2* and *cyp79B2 cyp79B3* mutants were spray-inoculated with spore suspensions ( $4 \times 10^6$  spores per milliliter) of either PcBMM, Pc2127, and P0831 strains or fungal transformants (PcBMM-GFP, Pc2127-GFP, and P0831-GFP) or with water (mock). **A**, Macroscopic disease symptoms caused in the inoculated plants by PcBMM, Pc2127, and P0831 strains at 7 days postinoculation (dpi). **B**, Confocal microscopy maximum projections of PcBMM-GFP, Pc2127-GFP, and P0831-GFP spores and mycelium on leaves of wild-type plants (Col-0) and *cyp79B2 cyp79B3* mutants at 4 dpi. Scale bar = 50  $\mu$ m. **C**, Quantification of *Plectosphaerella cucumerina*  $\beta$ -tubulin DNA in inoculated plants at 3 dpi by quantitative PCR. Values are represented as the average ( $\pm$ standard deviation) of the *n*-fold fungal DNA levels relative to plant ubiquitin (At-UBQ) gene. These experiments were repeated at least three times with similar results.

P0831 can be considered a different species than PcBMM, while Pc2127 may represent a divergent strain of *P. cucumerina*. Multilocus phylogenetic analysis of the TUB, EF1, and ITS sequences grouped PcBMM and Pc2127 with other members of *P. cucumerina*, whereas P0831 was located outside the clade formed by *P. cucumerina* and *P. plurivora*, further confirming that the P0831 isolate belongs to a different species (Supplementary Fig. S5).

We used OrthoFinder (Emms and Kelly 2015) to cluster protein-coding sequences into sets of homologous genes, comparing *Plectosphaerella* genomes with those of 24 other fungal species, and these data were used to reconstruct the evolution of every single gene (Fig. 2A; Supplementary Table S1). The phylogenomic analysis confirmed the taxonomic position of the genus *Plectosphaerella* within the family *Plectosphaerellaceae*, together with the genus *Verticillium*, and its

proximity to other hemibiotrophic and endophytic species, such as those of the genus *Colletotrichum* (order *Glomerellales*), rather than to other necrotrophic fungi such as *Botrytis* spp. or *Sclerotinia* spp. (Fig. 2A). The OrthoFinder tool clustered the protein-coding sequences of the three *Plectosphaerella* isolates in a core of 9,001 gene families that contained orthologs in the three genomes (Fig. 2B), with an average number of 3.3 genes per family (e.g., one ortholog per genome in most of the families). Only a few gene families were exclusive of PcBMM (3), Pc2127 (3), and P0831 (2), accounting, respectively, for 27, 24, and 27 genes that encoded putative proteins of unknown functions (Fig. 2B). Strain P0831 contained more genes without family (487) assigned by OrthoFinder, followed by PcBMM (412) and Pc2127 (176) (Table 1).

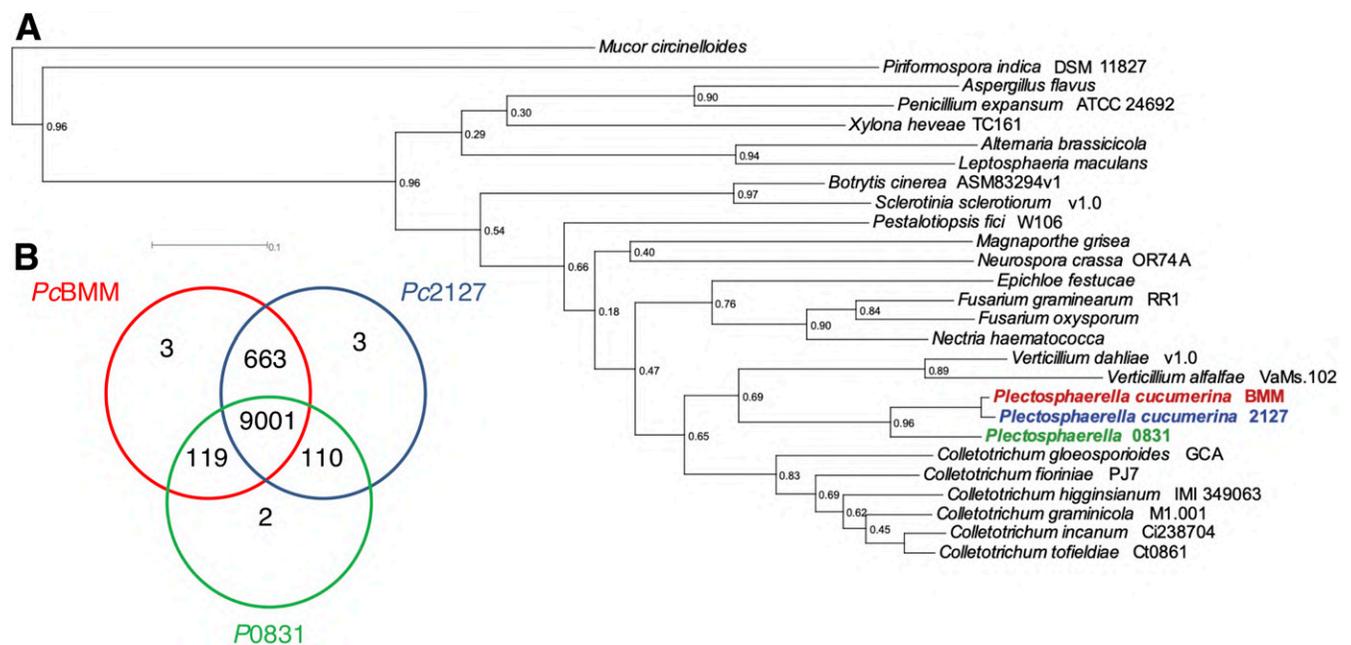
A phylome was reconstructed (stored in PhylomeDB [Huerta-Cepas et al. 2014]) using the proteomes of the three

**Table 1.** Genome sequencing and assembly statistics of *Plectosphaerella* strains

Parameter	PcBMM	Pc2127	P0831
Illumina single-end coverage	26.3	27.6	
Illumina paired-end coverage	27.4	28.6	27.6
PacBio coverage	14.6		
Assembly size	37.7 Mb	35.9 Mb	37.5 Mb
No. of contigs (>1 kb)	467	243	388
Largest contig	1.36 Mb	0.85 Mb	0.79 Mb
GC%	58.17	58.33	57.23
N50 length	338 kb	315 kb	191 kb
L50	36	41	56
N75 length	160 kb	158 kb	107 kb
L75	76	86	122
No. of 'N's in assembly per 100 kbp	1,054	113	1,910
No. of predicted gene models	11,323	11,007	10,821
Gene space coverage (BUSCO) <sup>a</sup>	94.0 to 97.7%	92.1 to 97.7%	93.2 to 97.5%
No. of genes in orthogroups	10,911	10,831	10,333
No. of orthogroups	9,786	9,777	9,232
No. of unassigned genes <sup>b</sup>	411 (3.6%)	176 (1.6%)	488 (4.5%)

<sup>a</sup> The two numbers indicate fully present and partially present genes.

<sup>b</sup> Percentage of unassigned genes to orthogroups was calculated over total predicted genes.



**Fig. 2.** Genomic relationships of PcBMM, Pc2127, and P0831 genomes with those of other fungi. **A**, Whole-genome phylogeny inferred by using the STAG method (Emms and Kelly 2018) with Orthofinder algorithm (Emms and Kelly 2015) within each orthogroup of genes present in all 27 compared genomes. A greedy consensus tree was achieved from all these individual estimates to get the final tree. **B**, Venn diagram distribution of the orthogroup genes generated by comparison of the three *Plectosphaerella* strains using Orthofinder algorithm.

*Plectosphaerella* strains and those of the 24 fungal species listed in Supplementary Table S1. Analysis to identify acquired genes in the genome of PcBMM allowed us to identify 159 orphan genes in PcBMM, 23 gained genes in the first node (shared by PcBMM and Pc2127), and 296 gained genes in the second one (shared by the three strains). Losses were also explored, and we identified 533 genes lost by Pc2127, and 1,112 genes lost by P0831 (Supplementary Fig. S6). For all the acquired genes, a BLASTp (Zhang et al. 2000) search against the UniProt database was performed, finding 10 putative cases of horizontal gene transfers corresponding to fungal genes that had homology with bacterial proteins (Supplementary Table S2). The percentage of identity between these pairs of homologous genes was very low (between 20 and 40%), suggesting they are not due to a contamination. Moreover, to find whether gained genes were enriched in any particular function, an enrichment analysis with FatiGO (Al-Shahrour et al. 2004) was performed for each of the referred nodes, but no enrichment was found in any of those analyses.

### Genes encoding putative secreted CAZymes are overrepresented among PcBMM pathogenicity and virulence-related genes.

We looked for predicted functions (proteins) related to pathogenesis in the three *Plectosphaerella* genomes by determining CAZymes, secreted proteins (secretome), membrane-associated transporters, and genes associated with secondary metabolism (Table 2). The proportions of these pathogenesis-related groups were practically identical among the three genomes, indicating that the genomic structure linked to known pathogenic mechanisms did not differ significantly among the three *Plectosphaerella* strains. However, we found within the secretome that the predicted secreted CAZymes were up to five times more abundant in the pathogenic PcBMM strain than in Pc2127 and P0831 (Table 2; Supplementary Table S3). Around 75% of the predicted secreted CAZymes of PcBMM had modules corresponding to CWDEs, such as carbohydrate esterases (CE), glycoside hydrolases (GH), and polysaccharide lyases (PL), that can potentially hydrolyze all the major plant cell-wall polysaccharides (e.g., pectin, hemicellulose, and cellulose [Supplementary Table S3]). Remarkably, secreted carbohydrate-binding modules (CBMs), which have functions related to adhesion to carbohydrates, and secreted auxiliary enzymes with redox activities (AA) were also highly enriched in PcBMM in comparison with Pc2127 and P0831. On the other hand, secreted glycosyl transferases (GT), which participate in the formation of glycosidic bonds, were similarly low-abundant in the three isolates (Supplementary Table S3).

In addition to CAZymes, we identified candidate secreted effector proteins (CSEPs) and secreted proteases in the secretomes of the three fungal strains (Table 2). CSEP was the most abundant

secreted category in the three genomes, representing between 44 and 47% of the predicted genes in the secretome (Table 2). No significant differences between the three isolates were found for CSEPs or proteases. We identified 87, 79, and 82 genes related to secondary metabolism pathways in the genomes of PcBMM, Pc2127, and P0831, respectively (Table 2; Supplementary Fig. S7). These pathways were represented by 15 core genes in each of the three genomes, which is half the average number of core genes found in other ascomycetes (Wang et al. 2015). P0831 contained more genes encoding nonribosomal peptide synthase (NRPS)-like genes and fewer genes encoding NRPSs and polyketide synthases than PcBMM and Pc2127 (Supplementary Fig. S6), but most of these genes had orthologous genes in the three genomes and only one gene was exclusive for each genome. These results suggest that the pathogenicity mechanisms of *Plectosphaerella* spp. in *Arabidopsis* (e.g., PcBMM strain) involve secreted CAZymes and probably proteinaceous toxins rather than toxins derived from secondary metabolites.

We also found 342, 264, and 387 genes, respectively, in PcBMM, Pc2127, and P0831 without homology to genes in other organisms. The proportion of these genes in the genome was significantly larger in P0831 ( $P < 0.00001$ ,  $\chi^2$ ). In PcBMM, one third of these genes (118) were isolate-exclusive, whereas only 35 and 17 genes were isolate-exclusive in Pc2127 and P0831, respectively.

### Differential repertoires of fungal genes are expressed during colonization of Col-0 and *cyp79B2 cyp79B3* plants by *Plectosphaerella* strains.

To identify *Plectosphaerella* genes determining the different types of interactions with *Arabidopsis* (Fig. 1), we performed genome-wide RNA-seq expression profiling. For that, we used total RNA extracted from Col-0 and *cyp79B2 cyp79B3* three-week-old plants at 10 and 16 h postinoculation (hpi) with spores ( $4 \times 10^6$  spores per milliliter) of PcBMM, Pc2127, or P0831 strains or after water treatment (noninoculated). In parallel, total RNA was also extracted from mycelia of the *Plectosphaerella* strains grown in vitro under minimal medium conditions. RNA-seq of in-vitro samples generated 59,794,266, 4,809,246, and 7,343,450 mapped reads, which corresponded to 10,586, 10,282, and 10,272 expressed genes of the PcBMM, Pc2127, and P0831 strains, respectively (Supplementary Table S4). RNA-seq of in-planta samples generated from 36,324 to 189,858 fungal mapped reads (1 to 5% of the total sequenced reads), which reflects the relatively low fungal RNA abundance and the differences in growth ability of the strains in planta (Supplementary Table S4). In Col-0 inoculated plants, we detected the expression of 7,687, 7,188 and 3,236 genes of PcBMM, Pc2127, and P0831, respectively, which means 73, 70, and 32% of the genes with detected expression (in vitro + in planta RNA-seq analyses [Fig. 3A and B]). These results showed dramatic differences for isolate P0831, which expressed in planta less than half of the number of genes expressed by PcBMM or Pc2127 (Fig. 3B).

We found specific clusters of genes from each fungal strain that were either up-regulated at different timepoints (10 or 16 hpi) or in one of the plant genotypes (Col-0 or *cyp79B2 cyp79B3*) (Fig. 3A). The number of PcBMM and Pc2127 genes expressed in *cyp79B2 cyp79B3* plants increased at 16 hpi in relation to 10 hpi, whereas it decreased in Col-0 inoculated with Pc2127, as expected from the observed failure of this strain to grow in Col-0 plants (Figs. 1 and 3B). In contrast, we observed an increase in the number of fungal expressed genes in P0831 in Col-0 at 16 hpi in comparison with 10 hpi, further supporting the idea that the nature of the interaction of P0831 with Col-0 plants differs from that of Pc2127 (Fig. 1B). Notably, about 50% of the P0831 upregulated genes were expressed exclusively

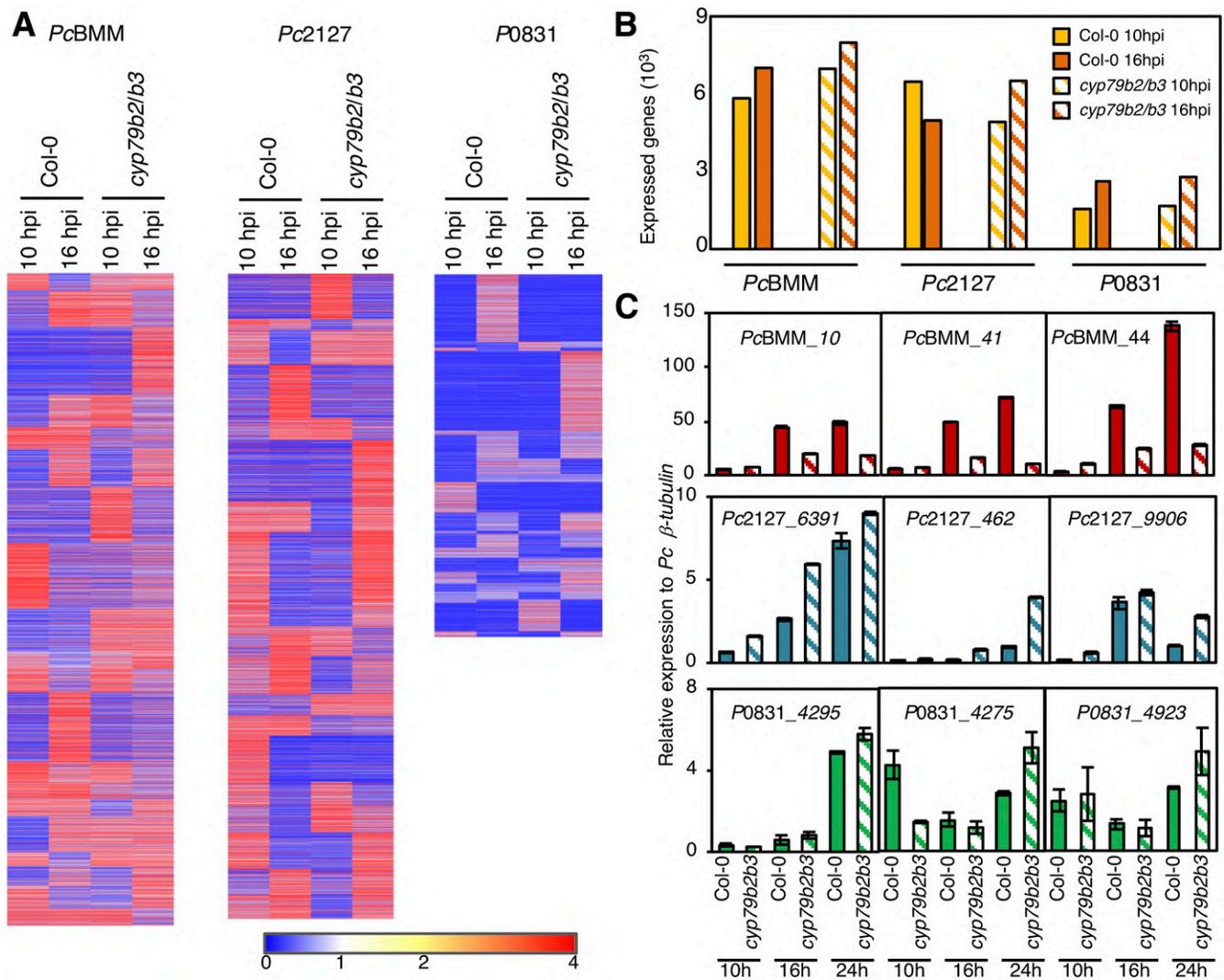
**Table 2.** Functional classification of pathogenesis related genes in the three *Plectosphaerella* genomes studied

Annotated function	Number of predicted genes		
	PcBMM	Pc2127	P0831
Carbohydrate-active enzymes (CAZymes)	782	782	768
Secretome	539	502	492
Secreted CAZymes	178	40	35
Candidate secreted effector proteins	253	221	217
Proteases	63	59	68
Secondary metabolism	87	79	82
Transport	1,032	1,015	1,114
Total number of predicted genes in each genome	11,323	11,007	10,821

at 16 hpi (Supplementary Fig. S8A), while most of the up-regulated genes of PcBMM and Pc2127 were expressed at both 10 and 16 hpi. Similarly, most PcBMM and Pc2127 upregulated genes were expressed in both Col-0 and *cyp79B2 cyp79B3*, whereas only 50% of the in planta-expressed P0831 genes were shared between both plant genotypes (Supplementary Fig. S8B). These data indicate that gene expression in P0831 was reprogrammed over time and depending on the immunity status of the host. To confirm RNA-seq data, in-planta expression of a representative subset of *Plectosphaerella* genes from some clusters were validated by quantitative reverse transcription-PCR (qRT-PCR) in independent experiments at 10, 16, and 24 hpi (Fig. 3C).

We next analyzed the in planta expressed fungal genes that could be related to pathogenicity, and we found that the most represented category in the three strains was “membrane-associated transporters,” followed by “secreted proteins,” and “secondary metabolism pathways,” which was in accordance with the relative abundance of these categories in the fungal genomes (Supplementary Fig. S9A; Table 2). The percentage of in

planta-expressed genes within each category was higher in PcBMM and Pc2127 than in P0831 (28 to 36% of PcBMM values [Supplementary Fig. S9A]), except for secondary metabolism pathways genes, whose expression was reduced about 60% in Pc2127 in comparison with PcBMM (Supplementary Fig. S9A). Main differences among strains were found within the number of genes of the secretome categories, particularly for the secreted CAZymes expressed in planta, which were significantly higher in PcBMM than in Pc2127 and P0831 (25 and 3% of PcBMM values, respectively) (Supplementary Fig. S9B). Among the top 10% of genes with the highest level of expression in planta, those encoding “secreted proteins,” particularly secreted CAZymes, and “membrane associated transporters” were the most abundant in PcBMM and Pc2127, while in P0831 the most represented category of highest expressed genes was “membrane associated transporters,” followed by “secreted proteins” (Supplementary Fig. S9C and D). These data support the relevant role of the secreted CAZymes in the pathogenicity of PcBMM and suggest that the survival strategy of P0831 on the plant surface could be based on the expression of membrane-associated transporters and on a weakened



**Fig. 3.** Differential expression of the transcriptomes of PcBMM, Pc2127, and P0831 in *Arabidopsis* inoculated plants. **A**, Heatmaps of gene expression showing the fungal genes significantly upregulated ( $\log_2$ -fold change [FC]  $\geq 2$ , false discovery rate  $< 0.05$ ) in Col-0 and *cyp79B2 cyp79B3* inoculated plants at 10 and 16 h postinoculation (hpi) in relation to in-vitro growth. Overrepresented (white to dark red) and underrepresented transcripts (white to dark blue) are shown as  $\log_2$ -FC relative to the expression of the genes in vitro. **B**, Number of in planta-expressed genes of PcBMM, Pc2127, and P0831 in the tested conditions (Col-0 and *cyp79B2 cyp79B3*) and timepoints (10 and 16 hpi). **C**, Validation by quantitative reverse transcription-PCR of the expression of some selected genes of PcBMM, Pc2127, and P0831 in Col-0 and *cyp79B2 cyp79B3* plants at different timepoints (10 to 24 hpi). Values are the relative expression of selected genes to fungal  $\beta$ -tubulin gene expression levels. Bars represent the average ( $\pm$ standard deviation) of two technical replicates.

transcriptional activation of CAZyme genes. In general, minor differences in the expression of genes of the most represented pathogenicity categories were observed between Col-0 and *cyp79b2 cyp79b3*, with the exception of strain Pc2127, which expressed more secretome-associated genes when it infected the mutant (Supplementary Fig. S9B and C). This probably contributes to the colonization of this mutant by Pc2127 and further supports the relevance of these genes in fungal pathogenicity.

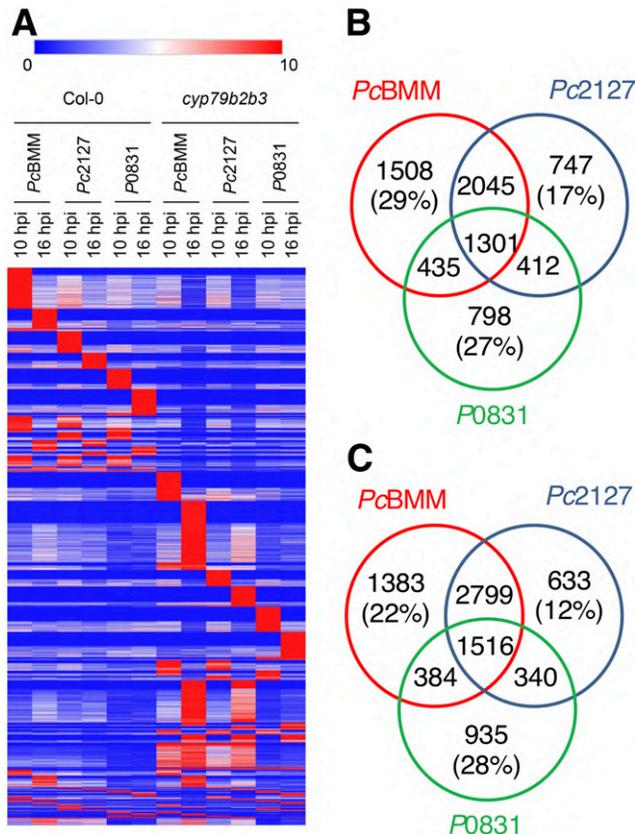
#### ***Arabidopsis* transcriptional responses during colonization by *Plectosphaerella* strains.**

Differential expression of *Arabidopsis* genes in Col-0 and *cyp79B2 cyp79B3* plants upon inoculation (10 and 16 hpi) with the three fungal strains was determined by comparing RNA-seq data of these samples with those of noninoculated (mock) plants (Supplementary Tables S5 and S6). The total number of mapped RNA-seq reads and *Arabidopsis*-expressed genes detected in Col-0 and *cyp79B2 cyp79B3* plants inoculated with each fungal strain was very similar, indicating that plant transcriptional responses were comparable (Fig. 4; Supplementary Table S5). We identified specific sets of differentially expressed genes (DEGs) (false discovery rate [FDR] < 0.05; log<sub>2</sub>-fold change [FC] ≥ 1) that were either upregulated or downregulated

in inoculated plants (Fig. 4A; Supplementary Table S6). Of note, the number of DEGs in P0831-inoculated plants was lower (from 35 to 48%) than in plants infected with PcBMM or Pc2127 (Supplementary Table S6; Fig. 4B). The number of DEGs also differed among plant genotypes, since it was larger in *cyp79B2 cyp79B3* than in Col-0 plants inoculated with any of the three strains (increases of 15, 17, and 7% for PcBMM, Pc2127, and P0831, respectively) (Supplementary Table S6; Fig. 4B). Moreover, the number of plant upregulated DEGs increased over time in *cyp79B2 cyp79B3* mutants inoculated with any of the strains and in Col-0 plants inoculated with PcBMM, whereas the number decreased in Col-0 plants inoculated with Pc2127 or P0831 (Supplementary Table S6; Supplementary Fig. S10). The number of *Arabidopsis* downregulated DEGs also increased over time in plants inoculated with PcBMM and Pc2127, especially in the *cyp79B2 cyp79B3* mutant, whereas it remained stable in plants inoculated with P0831 (Supplementary Table S6). Interestingly, profound changes in the sets of DEGs (up- and downregulated) at 10 and 16 hpi were found, especially in plants inoculated with P0831, in which only 15 and 10% of DEGs were shared between the two timepoints tested in Col-0 and *cyp79B2 cyp79B3* plants, respectively (Fig. 4A; Supplementary Fig. S10).

We identified two cores of 1,301 and 1,516 DEGs in Col-0 and *cyp79B2 cyp79B3* plants, respectively, which were shared upon inoculation with the three fungal strains (Fig. 4A, B, and C; Supplementary Fig. S10). However, we also found specific clusters of DEGs upon infection progression, depending on the plant genotype and fungal strain, that represented from 29% of total DEGs in Col-0 inoculated with PcBMM to 12% in *cyp79B2 cyp79B3* infected with Pc2127 (Fig. 4A, B, and C; Supplementary Fig. S10; Supplementary Table S6). The similarities between the transcriptional responses of plants inoculated with PcBMM and P0831 were lower (59% of shared DEGs) than those between PcBMM and Pc2127 (over 74% of shared genes [Fig. 4B and C; Supplementary Fig. S10]). Together, these data indicate that the transcriptional responses of *Arabidopsis* genotypes upon inoculation with each fungal strain differ significantly and that P0831 elicits in planta a weaker transcriptional response (lower number of DEGs) than the other two strains.

The putative functions of *Arabidopsis* DEGs in inoculated plants was analyzed by determining gene ontology (GO) term enrichment. DEGs in the Col-0-PcBMM interaction were enriched in 117 GO terms that were mainly related to “defense,” “signaling,” and “response to stress” functions (Supplementary Table S7; Supplementary Fig. S11A). A very similar pattern was found for DEGs in Col-0-Pc2127 and *cyp79B2 cyp79B3*-PcBMM or *cyp79B2 cyp79B3*-Pc2127 interactions, although with fewer GO terms enriched (about 70 in the three cases [Supplementary Table S7; Supplementary Fig. S11A and B]). Hence, a strong plant defense response is observed in these interactions independently of the final fungal outcome (e.g., disease progression or fungal growth inhibition). In *cyp79B2 cyp79B3* plants, a remarkable enrichment in GO terms related to aging was also observed, which could reflect the activation of molecular mechanisms related to cell death associated to fungal-induced necrosis. A very different GO enrichment pattern was observed among the upregulated DEGs in the Col-0-P0831 interaction, with only eight GO terms, mainly related to metabolic processes such as “biosynthesis of secondary metabolites and toxins” that could be related to defense functions (Supplementary Table S7; Supplementary Fig. S11A). Similar GO term enrichment analyses were performed with downregulated DEGs, but the number of significant GO terms identified was very low, 14 and nine GO terms in the *cyp79B2 cyp79B3* interactions with PcBMM and Pc2127, respectively, three GO terms in the Col-0-PcBMM interaction,



**Fig. 4.** Differential expression of the transcriptomes of Col-0 and *cyp79B2 cyp79B3* plants inoculated with *Plectosphaerella* strains. **A**, Heatmaps of gene expression showing the plant genes significantly upregulated (log<sub>2</sub>-fold change [FC] ≥ 2, false discovery rate [FDR] < 0.05) in the Col-0 and *cyp79B2 cyp79B3* inoculated plants at 10 and 16 h postinoculation (hpi). Overrepresented (white to dark red) and underrepresented transcripts (white to dark blue) are shown as log<sub>2</sub>-fold changes relative to the gene expression in mock-treated plants. **B** and **C**, Venn diagrams showing the number of genes upregulated (log<sub>2</sub>-FC ≥ 2, FDR < 0.05) in Col-0 (**B**) or *cyp79B2 cyp79B3* (**C**) plants at 10 and 16 h postinoculation with PcBMM, Pc2127, or P0831. The percentage of specific *Arabidopsis* genes upregulated by each fungal strain in inoculated Col-0 and *cyp79B2 cyp79B3* plants is indicated between parentheses in the Venn diagram.

and two in either Col-0 or *cyp79B2 cyp79B3* P0831-inoculated plants (Supplementary Table S7).

The analyses of the number of DEGs and GO term enrichment indicated that the response of Col-0 plants to P0831 was significantly different from those to PcBMM and Pc2127 and, also, suggest that genes related to canonical immune and disease resistance response are either not induced or weakly expressed in plants inoculated with P0831. To further validate this hypothesis, we performed independent experiments in inoculated Col-0 and *cyp79B2 cyp79B3* plants to determine, by qRT-PCR, the expression of marker genes of different defense pathways (e.g., *PR1* (SA), *PDF1.2* (ET + JA), *NCED3* (ABA), and *PAD3*, *FRK1*, and *CYP81F2*, which are PTI-associated genes). As shown in Supplementary Figure S12, the expression of the majority of these genes matched RNA-seq data and was strongly upregulated (up to 1,000-fold) in Col-0 and *cyp79B2 cyp79B3* plants upon inoculation with PcBMM and Pc2127. In contrast, these defense and immunity marker genes were, in accordance with GO term enrichment analyses (Supplementary Fig. S1A), either not induced or were slightly up-regulated in P0831-inoculated plants (up to 10-fold [Fig. S12]). These results further suggest that the immune responses of plants inoculated with the epiphytic strain are weakly induced or attenuated by the fungus.

## DISCUSSION

Plant-fungus interactions are diverse and largely dependent on the genetic determinants of fungus and host and on the environmental conditions (Brader et al. 2017; Rai and Agarkar 2016; Zeilinger et al. 2016). Comparing closely related fungal species with different abilities to cause disease in different hosts can provide novel cues about the specific genetic determinants of pathogenesis (Hacquard et al. 2016; Seidl et al. 2015; Wang et al. 2020; Xu et al. 2014). In this work, we have compared the initial colonization steps of three strains of the genus *Plectosphaerella* on wild-type and immunocompromised *Arabidopsis* genotypes, and we have sequenced and annotated the genomes and studied the transcriptional responses of these fungi upon plant colonization. The comparative analysis of the pathogenic PcBMM strain and the two nonpathogenic strains P0831 and Pc2127 revealed that, beyond the expected remarkable differences between pathogenic and nonpathogenic interactions, nonpathogenic interactions are also highly distinct between them (Malcom et al. 2013; Selosse et al. 2018).

Among nonpathogenic interactions, we can distinguish endophytic and epiphytic fungal lifestyles (Arnold 2007; Hardoim et al. 2015; Yao et al. 2019). Fungal endophytes are a highly diverse group that naturally colonize internal plant tissues without causing disease symptoms for at least part of their life cycle (Hardoim et al. 2015) and have received considerable attention in recent years due to their potential benefits to plants (Brader et al. 2017; Fesel and Zuccharo 2016; Lugtenberg et al. 2016; Plett and Martin 2018). Endophytes share infection mechanisms with plant pathogens and can be potentially pathogenic in immunocompromised plants, such as the *cyp79B2 cyp79B3* mutant (Hiruma et al. 2016). In contrast, the number of epiphytic fungal interactions characterized so far at the genomic and transcriptomic levels is scarce (Takashima et al. 2019; Wang et al. 2017; Xu et al. 2016). Fungal epiphytes are organisms growing on the leaf surface (Hongsanan et al. 2016), including pathogens and endophytes that can transiently colonize and be found on that habitat (Newton et al. 2010; Schulz and Boyle 2005). The leaf surface is a harsh environment, subject to abrupt temperature and humidity changes, scarcity of nutrients, and competitive forces in the diverse population of inhabiting microorganisms (Leveau 2015; Lindow and Brandl 2003). Host defenses are acting at the leaf surface, shaping the composition of epiphytic populations (Arnold 2007; Whipps

et al. 2008). Therefore, epiphytes must have adaptations that allow their growth on the leaf surface and must interact with the host (Vorholt 2012; Wang et al. 2017; Xu et al. 2016). Here, we have sequenced, annotated, and characterized the transcriptomic response of the epiphytic strain P0831 in its interaction with *Arabidopsis* wild type (Col-0) and a *cyp79B2 cyp79B3* mutant. Notably, we have found that strain P0831 is able to colonize Col-0 wild-type plants and also the immune-defective *dde2 ein2 pad4 sid2* and *cyp79B2 cyp79B3* mutants, without causing disease (Fig. 1). This contrasts with the observed phenotypes of PcBMM, which was more virulent in these immunocompromised mutants than in Col-0 plants, and of Pc2127, which, as previously described, was pathogenic on *cyp79B2 cyp79B3* plants but not on Col-0 (Sánchez-Vallet et al. 2010). Thus, our data reveal a novel lifestyle (epiphytic) of *Plectosphaerella* spp. on *Arabidopsis* plants that had not been previously studied. Since P0831 was isolated from a natural population of *Arabidopsis* (García et al. 2013) and fungi of genus *Plectosphaerella* are consistently associated with *Arabidopsis* in nature (Thiergart et al. 2020), we can conclude that the P0831-*Arabidopsis* interaction described here occurs in nature.

Several studies on the interaction of *P. cucumerina* with *Arabidopsis* have largely contributed to a precise understanding of plant basal immunity against fungal necrotrophs (Bednarek et al. 2009; Hernández-Blanco et al. 2007; Lipka et al. 2005; Pastorczyk et al. 2020; Sánchez-Vallet et al. 2010; Stein et al. 2006). These studies found dramatic differences between strains in their ability to overcome plant defenses and to cause disease, but the genetic determinants of pathogenesis of these fungi had not been characterized so far. The genomic and transcriptomic analyses of the different strains of genus *Plectosphaerella* described here place this genus very close to *Verticillium* and *Collectotrichum* species and not phylogenetically related to other fungal genera like *Botrytis* or *Sclerotinia* that contain strains with necrotrophic lifestyles in different plant species, including *Arabidopsis* (Dickman and Mitra 1992; Ge and Barbetti 2019; Windram et al. 2012). We have found that the pathogenicity determinants in genus *Plectosphaerella* include mainly CWDE (including secreted CAZymes) and proteinaceous toxins and effectors rather than secondary metabolites (Table 1; Supplementary Fig. S9), which contrasts with *Botrytis* or *Sclerotinia* species that used these metabolites as pathogenicity determinants (Amselem et al. 2011). These results suggest that the necrotrophic lifestyle of plant pathogenic fungi has been evolutionarily achieved through different genetic mechanisms. Our results also indicate that genetic differences among strains are more related to their taxonomic position than with their pathogenic ability, since we identified very few exclusive genes in each *Plectosphaerella* genome.

We have found a remarkable relationship between the number of secreted CAZymes in the *Plectosphaerella* strains and their pathogenicity on wild-type plants (Table 2; Supplementary Table S3; Supplementary Fig. S9), pointing to this group of enzymes as one of the main determinants of pathogenicity in *Plectosphaerella* spp. By contrast, the rest of the well-described pathogenicity determinants seem to be very well-conserved in the three fungal strains and similarly expressed in colonized plants (Table 2). CAZymes, which comprise several classes of proteins with different catalytic activities on carbohydrates (Lombard et al. 2014), are often classified as CWDEs, since they process or hydrolyze plant polysaccharides to either facilitate infection, gain access to nutrients, or both (Kameshwar et al. 2019; Zhao et al. 2013). For example, pathogens of dicots often contain more pectinases than fungi-infecting monocots, as dicots contain more pectins in their cell walls than monocots (Zhao et al. 2013). CAZymes have been previously identified as important pathogenicity or virulence factors of necrotrophic fungi (Chang et al. 2016; Douaiher et al. 2007; Kikot, et al.

2009; Lyu et al. 2015). Biotrophic fungi tend to have fewer CAZymes than necrotrophic and hemibiotrophic fungi, which is consistent with their lifestyles (Lyu et al. 2015; Zhao et al. 2013). While nutrient acquisition of necrotrophic and hemibiotrophic fungi is based on killing the host cell at the last stage of infection, biotrophic fungi absorb nutrients from living cells (Laluk and Mengiste 2010; Oliver and Ipcho 2004). The genomes of the three *Plectosphaerella* strains studied here harbor similar numbers of CAZymes (Table 2), also in line with those found in other necrotrophic and hemibiotrophic phytopathogenic fungi of dicots (Gazis et al. 2016; Knapp et al. 2018; Wang et al. 2020; Xu et al. 2014). However, the numbers of GT, CE, GH, and PL CAZymes in *Plectosphaerella* genomes are higher than in other phytopathogenic fungi, except for some *Colletotrichum* and *Fusarium* species (Hacquard et al. 2016). For example, CEs are up to three times more abundant in *Plectosphaerella* genomes than in other phytopathogenic fungi (Table 2). CEs catalyze the de-*O* or de-*N*-acylation of esters or amides and other substituted saccharides in which sugars play the role of alcohol and amine (Biely 2012). Acetylation of glycosyl residues of polysaccharides prevents hydrolysis of their glycosidic linkages by the corresponding GH, protecting plant cell walls against invading microorganisms. For instance, *Arabidopsis esk1* or *rwa/pmr* mutants show alterations in enzymes related to plant cell-wall polysaccharide acetylation and differential disease resistance responses to pathogens, including PcBMM (Escudero et al. 2017; Pawar et al. 2016; Yuan et al. 2013). Conversely, acetyl groups became targets of microbial CE that evolved to overcome the complexity of the plant cell walls and that cooperate with GH in plant polysaccharide degradation. Of note, plants overexpressing fungal esterases and acetylases show differential immune responses and disease-resistance phenotypes (Pawar et al. 2016).

In addition to the catalytic modules, around 7% of the CAZymes of genus *Plectosphaerella* also contain CBMs, which are the most common noncatalytic modules associated with enzymes active in cell-wall hydrolysis (Lombard et al. 2014). Among these CBMs are LysM-bearing proteins (CBM50 family) that have been described to function in scavenging chitin and other  $\beta$ -glucan fungal oligomers that are perceived as MAMPs by the plant immune system, thus preventing the recognition of the fungus by plant PRRs (Bolton et al. 2008; de Jonge and Thomma 2009). The number of CBM domains found in the three *Plectosphaerella* strains were not different, but the number of secreted proteins with these modules were seven and 18 times higher in PcBMM than in Pc2127 and P0831, respectively (Supplementary Table S3), probably explaining the ability of PcBMM to overcome plant immune responses.

Fungal secretomes play crucial roles during plant colonization by necrotrophic, hemibiotrophic, or biotrophic fungi and oomycetes (Kamoun 2009; Kim et al. 2016; Lyu et al. 2015; Stergiopoulos and de Wit 2009). The function of secreted proteins in necrotrophic fungi with a broad host range is poorly understood (Guyon et al. 2014). Secretome analysis of the necrotrophic fungus *Sclerotinia sclerotiorum* revealed large numbers of predicted effector candidates in the genome and the crucial role of non-CWDE secreted proteins during infection (Lyu et al. 2016). Here, we show that there are no significant differences between the number of small secreted proteins encoded by the genomes of the three *Plectosphaerella* strains. This contrasts with the differences observed in the number of genes encoding predicted secreted CAZymes (Table 2), which is much higher in PcBMM (Table 2). Moreover, these secreted CAZymes are expressed during the colonization process in higher numbers in pathogenic PcBMM than in the nonpathogenic strains (Supplementary Fig. S9). Therefore, we can conclude that secreted CAZymes are determinants of PcBMM pathogenicity.

RNA-seq analysis showed dramatic differences in fungal and plant gene expression at early stages of the interactions studied that not fully correlated with the ability of the different strains to grow and cause disease in the plant. The P0831 strain expressed in planta half of the genes expressed by PcBMM or Pc2127 (Fig. 3). The number of in planta-expressed genes of PcBMM or P0831 was in line with results obtained in inoculated plants with pathogenic and nonpathogenic strains of other fungi, such as *Colletotrichum* or *Fusarium* species (Baetsen-Young et al. 2020; Hacquard et al. 2016). However, in the case of Pc2127, the number of expressed genes was closer to the pathogenic interaction caused by PcBMM, suggesting an aggressive behavior of this strain at short timepoints. Pc2127, however, does not thrive at later timepoints (Fig. 3), and the number of its expressed genes decreased at 16 hpi in Col-0 inoculated plants, which contrasted with the stable or increased number of expressed genes of PcBMM or P0831 over time. These profound differences in gene expression dynamics between Pc2127 and P0831 strains in Col-0 *Arabidopsis* indicate the different nature of their interactions with *Arabidopsis* besides being both nonpathogenic.

Pc2127 was highly virulent on immunocompromised *cyp79B2 cyp79B3* plants but did not cause disease on *dde2 ein2 pad4 sid2* mutants (Fig. 1), indicating that Trp-derived metabolites rather than ET, SA, and JA and ETI pathways are required for basal resistance to nonadapted *Plectosphaerella* fungi, as previously described (Bednarek et al. 2009; Sánchez-Vallet et al. 2010). P0831 was able to grow on plant surfaces both of wild type and immunocompromised *Arabidopsis* mutants without causing disease symptoms (Fig. 1; Supplementary Figs. S2 and S3). These results suggest that alternative plant defense mechanisms might prevent plant infection by P0831. However, our results also showed a weaker upregulation of canonical plant immune responses upon P0831 epiphytic growth when compared with the other two strains (Supplementary Figs. S11 and S12). This weak immune response does not seem to be related to a defective perception of P0831, since in-vivo analyses at 3 dpi with *pRbohD::LUC* plants demonstrate that these plants are able to perceive this fungus and to activate transcriptional responses (Supplementary Fig. S4). These results suggest that the epiphytic P0831 strain might inhibit activation of the plant immune system at early stages of colonization through specific mechanisms that will require further characterization. The transcriptional response of P0831 upon interaction with *Arabidopsis* was distinct from those of PcBMM and Pc2127, and it changed considerably over time and depending on the immunity status of the host (Fig. 3; Supplementary Fig. S8). Strikingly, the transcriptional response of P0831 seemed more related to transport than with pathogenicity (Supplementary Fig. S9). In summary, we have shown here the diversity of interactions that *Plectosphaerella* fungal strains can establish with plants (*Arabidopsis*) and have characterized a novel type of epiphytic interaction. The characterization of *Plectosphaerella* genomes and transcriptome has also revealed that the expression of specific sets of fungal genes in planta determines the diverse lifestyles and pathogenicity of the *Plectosphaerella* strains studied.

## MATERIALS AND METHODS

### Detection of *Plectosphaerella* sp. from natural *A. thaliana* populations.

Asymptomatic *A. thaliana* leaves and roots were collected from natural populations in central Spain, Menasalbas (MEN) and Las Rozas, that have been described previously (García et al. 2013). Total DNA was extracted from the samples using the FastDNA SPIN Kit for Soil (MP Biomedicals) and 10 ng of

DNA template was used for real time PCR amplification using the specific primer pair that targets the *Plectosphaerella*  $\beta$ -tubulin gene, 5'-CAAGTATGTTCCCCGAGCCGT-3' and 5'-GGTCCCTTCGGTCAGCTCTTC-3' (Sánchez-Rodríguez et al. 2009), and the FS Universal SYBR Green Master Rox (Roche) as described (Delgado-Cerezo et al. 2012).

### Biological material and growth conditions.

*A. thaliana* genotypes used in this study were accession Columbia-0 (Col-0) and the mutants (in Col-0 background) *cyp79B2 cyp79B3* (Zhao et al. 2002), *dde2 ein2 pad4 sid2* (Tsuda et al. 2009), and the *pRBOHD::LUC* line (in Col-0 background [Morales et al. 2016]). Plants were grown on a 3:1 soil-vermiculite mixture under a 10-h day and 14-h night cycle, 24°C day and 22°C night temperatures, 65% relative humidity, and light intensity of 120 mE m<sup>-2</sup> s<sup>-1</sup>. PcBMM isolate was provided by B. Mauch-Mani (University of Neuchâtel, Switzerland [Tierens et al. 2001, Ton and Mauch-Mani 2004]), and Pc2127 was obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ) Collection (Braunschweig, Germany). P0831 strain was isolated from an asymptomatic *Arabidopsis* leaf from the wild MEN population in central Spain (García et al. 2013). Spores from these fungi were obtained as reported by Ramos et al. (2013). P0831-GFP transformant was obtained and selected following previously reported methodology used to obtain PcBMM-GFP and Pc2127-GFP transformants (Ramos et al. 2013).

### Arabidopsis inoculation and detection of *Plectosphaerella* isolates.

Inoculation of *Arabidopsis* leaves was performed by spraying 3-week-old soil-grown plants with a spore suspension (4 × 10<sup>6</sup> spores per milliliter) of the fungus as described (Llorente et al. 2005). At least 12 plants per genotype were inoculated and a minimum of three independent experiments were performed. For fungal DNA quantification, DNA from infected plants was extracted as described by Llorente et al. (2005) and fungal DNA was determined by qPCR using the FS Universal SYBR Green Master Rox (Roche), oligonucleotides for the *Plectosphaerella*  $\beta$ -tubulin gene, and the *Arabidopsis* UBIQUITIN21 (At5G25760, UBIQ21) gene, and by calculating the change in the cycle threshold ( $\Delta$ Ct) values, as reported previously (Delgado-Cerezo et al. 2012). The relative ratio was determined from the expression 2<sup>- $\Delta\Delta$ Ct</sup> (Rieu and Powers 2009). The qPCR results are the means ( $\pm$ standard deviation [SD]) of two technical replicates.

Confocal images were acquired on a TCS SP2 AOBS spectral confocal laser scanning microscope (Leica Microsystems) as described (Ramos et al. 2013). GFP fluorescence (green) was captured by excitation at 488 nm and emission in the range 490 to 520 nm, and chlorophyll autofluorescence (red) was captured with an excitation of 543 nm and the emission was collected in the range 600 to 720 nm. The dimensional xyz confocal stacks and orthogonal projections across selected xz and yz planes were obtained using Leica Confocal Software LCS Lite version 2.61 build 1538 (Leica Microsystems).

Lactophenol-TB staining of plant leaves was performed as described by Koch and Slusarenko (1990) and detection of H<sub>2</sub>O<sub>2</sub> by 3,3'-DAB (Sigma) staining was performed as described by Torres et al. (2002). For in vivo quantification of Rboh-promoter driven LUC activity, pRBOHD::LUC Col-0 plants were sprayed with a 0.5 mM solution of D-luciferin (Melford) and luminescence produced was measured using a NightOWL LB 983 in-vivo imaging system (Berthold) as described by Morales et al. (2016).

### Genome sequencing and assembly of *Plectosphaerella* strains.

Library construction, quality control, and DNA Illumina HiSeq sequencing (PE 125 bp and SE 50 bp) were performed at the

Centre for Genomic Regulation (CRG) (Barcelona, Spain) in an Illumina HiSeq2500, and PacBio sequencing was performed at Wageningen University and Research (Wageningen, The Netherlands), using 1  $\mu$ g of fungal genomic DNA. For fungal DNA extraction, cultures were grown at 28°C as described previously (Ramos et al. 2007). For PcBMM genome assembly, a hybrid strategy was used combining Illumina and PacBio data. SE and PacBio reads were first assembled using Velvet 1.2.10 (Zerbino and Birney 2008) and were further used for scaffolding of Illumina PE read assemblies from SPAdes assembler version 3.11.0 (Bankevich et al. 2012). The established SPAdes pipeline was used in careful mode providing Illumina and PacBio assemblies as untrusted contigs for scaffolding only and kmer scan using 21, 31, 41, 51, 61, 71, and 81. For Pc2127 and P0831, assemblies were constructed only from Illumina data using a combination of Velvet 1.2.10 (Zerbino and Birney 2008) and SPAdes version 3.11.0 (Bankevich et al. 2012). To identify and remove potential contaminating sequences, assemblies were aligned to the genomes of *A. thaliana* and *Homo sapiens* using MUMmer v. 3.0 (Kurtz et al. 2004) with default parameters settings. Contigs that aligned with more than 50% of their sequence or at least 85% sequence identity to any of the selected genomes were removed from the assemblies. Finally, assembly quality was assessed on the basis of L50/75 and N50/75 values, percentage of error-free bases estimated with quast 5.0.1 (Gurevich et al. 2013), and gene space coverage estimated with BUSCO v3 (Simão et al. 2015). The Whole-Genome Shotgun projects have been deposited in the DDBJ/ENA/GenBank database under accessions JACAFV000000000 (*Plectosphaerella* sp. strain P0831 Biosample SAMN14235337 TaxID: 40657), JACAFW000000000 (*P. cucumerina* Pc2127 Biosample SAMN14233823 and Pc2127 TaxID: 40658), SAMN14233823 and Pc2127 TaxID: 40658), and JACAFX000000000 (*P. cucumerina* PcBMM Biosample SAMN14233534 TaxID: 40658), and BioProject number PRJNA609142. The versions described by this paper are versions JACAFV010000000, JACAFW010000000, and JACAFX010000000.

### Gene annotation of *Plectosphaerella* strains.

The prediction of PcBMM, Pc2127, and P0831 gene models was performed using the Maker v2.31.10 pipeline (Cantarel et al. 2008) that integrates different *ab initio* gene prediction tools together with evidence from expressed sequence tag and protein alignments. In a first step for each genome, the pipeline was run using Augustus (Stanke et al. 2006) (with species model *Fusarium graminearum*) and GeneMark-ES (Hoff et al. 2016) for *ab initio* gene prediction together with transcript and protein alignment evidence. The resulting gene models from this first run were used as training set for a third *ab initio* prediction tool, SNAP (Korf 2004), and, subsequently, the annotation pipeline was rerun, this time including all three *ab initio* prediction tools together with the transcript and protein alignment evidence to yield the final gene models. The alignment evidence was created from BLAST and Exonerate (Zaharia et al. 2011) alignments of both protein and transcript sequences of each respective fungus obtained from corresponding RNA-seq data via a transcriptome de novo assembly. For this purpose, we extracted RNA-seq reads that did not align to the host plant genome from in-planta samples and combined these with the reads from in-vitro samples of the respective fungus. The combined RNA-seq reads were used as input for Trinity (Bryant et al. 2017), with default parameter, to assemble transcripts and extract peptide sequences of the best-scoring open reading frames. General functional annotations for the predicted gene models were obtained using Blast2GO (Conesa et al. 2005) with the local National Center for Biotechnology Information (NCBI) Non-redundant database downloaded (version: June 12, 2018).

### Read mapping and variant calling.

Read mapping and variant calling was performed using the PcBMM assembly as reference. Before aligning the reads

against the reference assembly, they were trimmed at the first base with a PHRED quality below 10 and pairs with one or both reads <31 bases after trimming were filtered out. Then, Pc2127 and P0831 reads were aligned against PcBMM assembly using BWA-MEM algorithm of BWA v0.7.13 (Li 2013) and were sorted with SAMtools v0.1.18 (Li et al. 2009). GATK v3.5 (McKenna et al. 2010) was used to call SNPs, and they were filtered according to GATK documentation (QD < 2.0; MQ < 40; FS > 60.0; haplotype score > 13.0; MQRankSum < -12.5; ReadPosRankSum < -8.0).

To detect putative deletions, a mpileup file of each strain (Pc2127 and P0831) was generated with SAMtools v0.1.18 (Li et al. 2009) and was compared with the reference gff file. A base was considered to be present in the strain if it was covered by at least one read. The percentage of total coverage for a region was calculated through the ratio between the number of positions with reads and the total size of the respective region. To obtain the identity between each strain and the reference genome, a BLASTp (Zhang et al. 2000) between the strain proteome and the reference and vice versa was performed. Only results in which both BLASTp best hits were the same were considered for further analysis. Then, the identity of each strain was determined with trimAl v1.4.rev15 (Capella-Gutierrez et al. 2009).

### Phylogeny of different species in the genus *Plectosphaerella*.

Phylogeny was obtained with the maximum likelihood method and Tamura-Nei model (Tamura and Nei 1993) on a multilocus alignment (TUB, EF1, and ITS). One or more initial trees for the heuristic search were obtained automatically by applying neighbor-joining and BioNJ algorithms to a matrix of pairwise distances estimated using the maximum composite likelihood (MCL) approach and then selecting the topology with superior log likelihood value. The tree was drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 28 nucleotide sequences and there were a total of 1,004 positions in the final dataset. Evolutionary analyses were conducted in MEGA X (Kumar et al. 2018).

### MCL analysis and whole-genome phylogeny.

Gene families and clusters of orthologous genes were inferred, using OrthoFinder (Emms and Kelly 2015) v2.2.7, with standard parameters, using Diamond (Buchfink et al. 2015) as a sequence aligner for protein and translated DNA searches. Whole-genome phylogeny was inferred by using the STAG method (Emms and Kelly 2018) with Orthofinder algorithm within each orthogroup of genes present in all the 27 compared genomes (Supplementary Table S1). A greedy consensus tree was achieved from all these individual estimates to get the final tree.

### Phylome reconstruction and gene gain and loss analysis.

The *Plectosphaerella* phylome was reconstructed with PhylomeDB (Huerta-Cepas et al. 2014). For each *Plectosphaerella* protein, a Smith-Waterman search was performed against the proteome database containing data on the 24 species used in this work (Supplementary Table S1), as well as of the three *Plectosphaerella* strains. To obtain a set of proteins with significant similarity, an e-value threshold of <1e-05 was established. Moreover, only sequences that aligned with a continuous region longer than 50% of the query sequence were chosen for the detection of homologs. At maximum, 150 sequences per gene were taken for further analysis. MUSCLE v3.8.31 (Edgar 2004), MAFFT v6.814b (Katoh et al. 2002), and DIALIGN-TX (Subramanian et al. 2008) were used to align, in forward and reverse directions, the homologous protein sequences. M-COFFEE (T-Coffee v8.80) (Wallace et al. 2006) was used to combine the six resulting alignments. The final

alignment was trimmed with trimAl v1.3 (consistency cutoff of 0.1667 and a gap score cutoff of 0.1) (Capella-Gutierrez et al. 2009). Phylogenetic trees were reconstructed through a neighbor-joining approach as implemented in BioNJ (Gascuel 1997). Eight different models (JTT, WAG, MtREV, VT, LG, Blosum62, CpREV, and DCMut) were used to compute the likelihood of this topology, allowing branch-length optimization, as in PhyML v3.0 (Guindon et al. 2010). Their likelihood was compared in accordance with the AIC criterion, and the two evolutionary models best fitting the data were selected to derive maximum likelihood trees. A discrete gamma-distribution model with four rate categories plus invariant positions was used in all cases, with parameters estimated from the data. Gene gain and loss analysis in the *Plectosphaerella* branch was performed based on the phylome results. For each gene from the starting genome, orthologs were mapped using a species overlap algorithm. Then the emergence of each gene was located at the common ancestor of all its orthologs. Losses were inferred by going from the node where the gene was gained to the tips and locating nodes where all species had lost the gene. GO term enrichment was calculated using an in-house adaptation of FatiGo (Al-Shahrour et al. 2004). Enrichment was searched between genes gained and lost in the nodes pertaining to the emergence and diversification of *Plectosphaerella* species.

### Annotation of specific gene categories.

To predict the repertoire of CAZymes encoded by the three strains, the corresponding genomes were scanned using the CAZymes analysis toolkit (Park et al. 2010). The secretomes were determined using the SECRETOOL pipeline (Cortázar et al. 2014) run on the proteome of each strain. CSEPs were defined as extracellular secreted proteins with no significant BLAST similarity (e-value < 1e-03). To identify genes encoding secreted proteases, sequences of predicted extracellular secreted proteins were subjected to MEROPS Batch BLAST analysis (Rawlings et al. 2016). Putative membrane transporter genes were identified and classified through BLAST searches against the Transporter Collection (TC) database (Saier et al. 2014). Genes with a sequence identity of at least 30% to their best hit in the database (e-value < 1e-03) were extracted. These genes were subsequently assigned to the TC family of their best TC database hit. Secondary metabolism was predicted using the antiSMASH fungal version web server (Medema et al. 2011).

### RNA-seq and gene expression analyses.

Library construction, quality control, and RNA Illumina HiSeq 50-bp SE sequencing were performed at the CRG in an Illumina HiSeq2500 using 1 µg of total RNA. To make sure the sequenced reads were of sufficiently high quality, an initial quality check was performed using the FastQC suite (Andrews 2010). All statistical analyses of plant and fungal gene expression were performed using Tophat and the Cufflinks pipeline (Trapnell et al. 2012), where the RNA-seq reads were mapped to the assembled and annotated genomes of either PcBMM, Pc2127, or P0831 and in parallel to the annotated genome of the host plant *A. thaliana* (ARAPORT). Differential gene expression in Col-0 and *cyp79B2 cyp79B3* plants inoculated with PcBMM, Pc2127, and P0831 was fitted for each analysis using FPKM (fragments per kilobase per million reads) values. Heatmaps of gene expression profiles were generated with the Morpheus analysis package. Gene expression data of fungal and plant genes are shown in Supplementary Tables S8 and S9. The raw RNA-seq data in this study are deposited in the NCBI Sequence Read Archive (SRA), BioProject number PRJNA614936 and BioSamples SAMN14444085 to SAMN14444096, SAMN14444149, SAMN14444167, and SAMN14444168, with SRA accessions SRR11668188 to SRR11668202.

The cytoscape plugin Cluego + Cluepedia (Bindea et al. 2009) was used to construct GO term enrichment networks and to visualize functionally grouped terms among significantly regulated genes (FDR <0.05; log<sub>2</sub>-FC ≥ 1) in *Arabidopsis* inoculated plants compared with mock-inoculated plants. Significant (*P* < 0.05) enrichments were determined using the hypergeometric test and were manually classified as defense, signaling, response to stress, metabolic process, transcription and translation, multiorganism process, or aging (Supplementary Table S7).

### Gene expression validation.

RNA extractions from *Plectosphaerella*- or mock-inoculated plants were done as described by Llorente et al. (2005). Real-time qRT-PCR analyses were done as described previously (Delgado-Cerezo et al. 2012), using the FS Universal SYBR Green Master Rox (Roche). The qRT-PCR results are the means (±SD) of two technical replicates. Oligonucleotides used for cDNA amplification were designed with Primer Express (version 2.0; Applied Biosystems) and are shown in Supplementary Table S10 (Jordá et al. 2016; Sánchez-Rodríguez et al. 2009). For fungal genes, relative expression was calculated related to the β-tubulin expression. For *Arabidopsis* genes, ubiquitin (AT5G25760) expression was used to normalize the transcript level in each sample of the genes analyzed. The relative ratio was then determined from the expression 2<sup>-ΔΔCt</sup> (Rieu and Powers 2009), using the mock-inoculated plants as calibrators.

### ACKNOWLEDGMENTS

We thank L. Faino and B. Thomma (Laboratory of Phytopathology, Wageningen University & Research, The Netherlands) for providing support with *Plectosphaerella* paired-ends sequencing, and A. Vlasova, F. Cámara and R. Guigó (CRG, Barcelona Institute of Science and Technology, Barcelona, Spain) for helping with the initial draft genome assemblies. We thank P. del Prado (CBGP, UPM-INIA) for technical assistance.

### AUTHOR-RECOMMENDED INTERNET RESOURCES

antiSMASH fungal version web server: <https://fungismash.secondarymetabolites.org/#!/start>  
 ARAPORT Arabidopsis Information Portal: <https://www.araport.org>  
 Augustus gene prediction program: <http://bioinf.uni-greifswald.de/augustus>  
 Blast2GO: <https://www.blast2go.com>  
 BUSCO v3: <https://busco.ezlab.org>  
 Cytoscape Cluego: <http://apps.cytoscape.org/apps/cluego>  
 Diamond: <https://github.com/bbuchfink/diamond>  
 Exonerate software: <https://www.ebi.ac.uk/about/vertebrate-genomics/software/exonerate>  
 FastQC suite: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>  
 GeneMark-ES: <http://exon.gatech.edu/GeneMark>  
 MEROPS database: [https://www.ebi.ac.uk/merops/submit\\_searches.shtml](https://www.ebi.ac.uk/merops/submit_searches.shtml)  
 Morpheus software: <https://software.broadinstitute.org/morpheus>  
 MUMmer v. 3.0: <http://mummer.sourceforge.net>  
 OrthoFinder: <http://www.stevkellylab.com/software/orthofinder>  
 SNAP prediction tool: <http://snap.cs.berkeley.edu>  
 SECRETOOL pipeline: <http://genomics.cicbiogune.es/SECRETOOL/Secretool.php>  
 Cufflinks pipeline: <http://cole-trapnell-lab.github.io/cufflinks/>  
 Transporter Collection database: <http://www.tcdb.org>  
 Trinity de novo assembly: <https://github.com/trinityrnaseq/trinityrnaseq/wiki>

### LITERATURE CITED

Al-Shahrour, F., Díaz-Urriarte, R., and Dopazo, J. 2004. FatiGO: A web tool for finding significant associations of gene ontology terms with groups of genes. *Bioinformatics* 20:578-580.  
 Amselem, J., Cuomo, C. A., van Kan, J. A., Viaud, M., Benito, E. P., Couloux, A., Coutinho, P. M., de Vries, R. P., Dyer, P. S., Fillinger, S., Fournier, E., Gout, L., Hahn, M., Kohn, L., Lapalu, N., Plummer, K. M., Pradier, J. M., Quévillon, E., Sharon, A., Simon, A., ten Have, A., Tudzynski, B., Tudzynski, P., Wincker, P., Andrew, M., Anhouard, V.,

Beever, R. E., Beffa, R., Benoit, I., Bouzid, O., Brault, B., Chen, Z., Choquer, M., Collémare, J., Cotton, P., Danchin, E. G., Da Silva, C., Gautier, A., Giraud, C., Giraud, T., Gonzalez, C., Grossetete, S., Güldener, U., Henrissat, B., Howlett, B. J., Kodira, C., Kretschmer, M., Lappartient, A., Leroch, M., Levis, C., Mauceli, E., Neuvéglise, C., Oeser, B., Pearson, M., Poulain, J., Poussereau, N., Quesneville, H., Rasclé, C., Schumacher, J., Ségurens, B., Sexton, A., Silva, E., Sirven, C., Soanes, D. M., Talbot, N. J., Templeton, M., Yandava, C., Yarden, O., Zeng, Q., Rollins, J. A., Lebrun, M. H., and Dickman, M. 2011. Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. *PLoS Genet.* 7:e1002230.  
 Andrews, S. 2010. FastQC: A quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.  
 Arnold, A. E. 2007. Understanding the diversity of foliar endophytic fungi: Progress, challenges, and frontiers. *Fungal Biol. Rev.* 21:51-66.  
 Bacete, L., Mérida, H., López, G., Dabos, P., Tremoussaygue, D., Denancé, N., Miedes, E., Bulone, V., Goffner, D., and Molina, A. 2020. *Arabidopsis* response regulator 6 (ARR6) modulates plant cell wall composition and disease resistance. *Mol. Plant-Microbe Interact.* 33:767-780.  
 Bacete, L., Mérida, H., Miedes, E., and Molina, A. 2018. Plant cell wall-mediated immunity: Cell wall changes trigger disease resistance responses. *Plant J.* 93:614-636.  
 Baetsen-Young, A., Man Wai, C., VanBuren, R., and Day, B. 2020. *Fusarium virguliforme* transcriptional plasticity is revealed by host colonization of maize vs. soybean. *Plant Cell* 32:336-351.  
 Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Pribelski, A. D., Pyskhin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., and Pevzner, P. A. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19:455-477.  
 Bednarek, P., Pislewski-Bednarek, M., Svatos, A., Schneider, B., Doubek, J., Mansurova, M., Humphry, M., Consonni, C., Panstruga, R., Sanchez-Vallet, A., Molina, A., and Schulze-Lefert, P. 2009. A glucosinolate metabolism pathway in living plant cells mediates broad-spectrum antifungal defense. *Science* 323:101-106.  
 Berrocal-Lobo, M., Molina, A., and Solano, R. 2002. Constitutive expression of ETHYLENE-RESPONSE-FACTOR1 in Arabidopsis confers resistance to several necrotrophic fungi. *Plant J.* 29:23-32.  
 Biely, P. 2012. Microbial carbohydrate esterases deacetylating plant polysaccharides. *Biotechnol. Adv.* 30:1575-1588.  
 Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W. H., Pagès, F., Trajanoski, Z., and Galon, J. 2009. ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 25:1091-1093.  
 Bolton, M. D., van Esse, H. P., Vossen, J. H., de Jonge, R., Stergiopoulos, I., Stulemeijer, I. J. E., van den Berg, G. C. M., Borrás-Hidalgo, O., Dekker, H. L., de Koster, C. G., de Wit, P. J. G. M., Joosten, M. H. A. J., and Thomma, B. P. H. J. 2008. The novel *Cladosporium fulvum* lysin motif effector Ecp6 is a virulence factor with orthologues in other fungal species. *Mol. Microbiol.* 69:119-136.  
 Brader, G., Compant, S., Vescio, K., Mitter, B., Trognitz, F., Ma, L.-J., and Sessitsch, A. 2017. Ecology and genomic insights into plant-pathogenic and plant-nonpathogenic endophytes. *Annu. Rev. Phytopathol.* 55:61-83.  
 Bryant, D. M., Johnson, K., DiTommaso, T., Tickle, T., Couger, M. B., Payzin-Dogru, D., Lee, T. J., Leigh, N. D., Kuo, T. H., Davis, F. G., Bateman, J., Bryant, S., Guzikowski, A. R., Tsai, S. L., Coyne, S., Ye, W. W., Freeman, R. M., Jr., Peshkin, L., Tabin, C. J., Regev, A., Haas, B. J., and Whited, J. L. 2017. A tissue-mapped axolotl de novo transcriptome enables identification of limb regeneration factors. *Cell Rep.* 18:762-776.  
 Buchfink, B., Xie, C., and Huson, D. H. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12:59-60.  
 Buxdorf, K., Rahat, I., Gafni, A., and Levy, M. 2013. The epiphytic fungus *Pseudozyma aphidis* induces jasmonic acid- and salicylic acid/nonexpressor of PR1-independent local and systemic resistance. *Plant Physiol.* 161:2014-2022.  
 Cantarel, B. L., Korf, I., Robb, S. M. C., Parra, G., Ross, E., Moore, B., Holt, C., Sánchez Alvarado, A., and Yandell, M. 2008. MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* 18:188-196.  
 Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. 2009. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972-1973.

- Carlucci, A., Raimondo, M. L., Santos, J., and Phillips, A. J. L. 2012. *Plectosphaerella* species associated with root and collar rots of horticultural crops in southern Italy. *Persoonia* 28:34-48.
- Chang, H. X., Yendrek, C. R., Caetano-Anolles, G., and Hartman, G. L. 2016. Genomic characterization of plant cell wall degrading enzymes and *in silico* analysis of xylanases and polygalacturonases of *Fusarium virguliforme*. *BMC Microbiol.* 16:147.
- Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674-3676.
- Cortázar, A. R., Aransay, A. M., Alfaro, M., Oguiza, J. A., and Lavín, J. L. 2014. SECRETOOL: Integrated secretome analysis tool for fungi. *Amino Acids* 46:471-473.
- D'Amico, M., Frisullo, S., and Cirulli, M. 2008. Endophytic fungi occurring in fennel, lettuce, chicory, and celery—Commercial crops in southern Italy. *Mycol. Res.* 112:100-107.
- de Jonge, R., and Thomma, B. P. 2009. Fungal LysM effectors: Extinguishers of host immunity? *Trends Microbiol.* 17:151-157.
- Delgado-Cerezo, M., Sánchez-Rodríguez, C., Escudero, V., Miedes, E., Fernández, P. V., Jordá, L., Hernández-Blanco, C., Sánchez-Vallet, A., Bednarek, P., Schulze-Lefert, P., Somerville, S., Estevez, J. M., Persson, S., and Molina, A. 2012. Arabidopsis heterotrimeric G-protein regulates cell wall defense and resistance to necrotrophic fungi. *Mol. Plant* 5: 98-114.
- Dickman, M. B., and Mitra, A. 1992. *Arabidopsis thaliana* as a model for studying *Sclerotinia sclerotiorum* pathogenesis. *Physiol. Mol. Plant Pathol.* 41:255-263.
- Dillard, H. R., Cobb, A. C., Shah, D. A., and Straight, K. E. 2005. Identification and characterization of russet on snap beans caused by *Plectosporium tabacinum*. *Plant Dis.* 89:700-704.
- Douaiher, M. N., Nowak, E., Durand, R., Halama, P., and Reignault, P. 2007. Correlative analysis of *Mycosphaerella graminicola* pathogenicity and cell wall-degrading enzymes produced *in vitro*: The importance of xylanase and polygalacturonase. *Plant Pathol.* 56:79-86.
- Durán, P., Thiergart, T., Garrido-Oter, R., Agler, M., Kemen, E., Schulze-Lefert, P., and Hacquard, S. 2018. Microbial interkingdom interactions in roots promote *Arabidopsis* survival. *Cell* 175:P973-983.e14.
- Edgar, R. C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792-1797.
- Emms, D. M., and Kelly, S. 2015. OrthoFinder: Solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 16:157.
- Emms, D. M., and Kelly, S. 2018. STAG: Species tree inference from all genes. *bioRxiv* 267914.
- Escudero, V., Jordá, L., Sopena-Torres, S., Mérida, H., Miedes, E., Muñoz-Barrios, A., Swami, S., Alexander, D., McKee, L. S., Sánchez-Vallet, A., Bulone, V., Jones, A. M., and Molina, A. 2017. Alteration of cell wall xylan acetylation triggers defense responses that counterbalance the immune deficiencies of plants impaired in the  $\beta$ -subunit of the heterotrimeric G-protein. *Plant J.* 92:386-399.
- Fesel, P. H., and Zuccaro, A. 2016. Dissecting endophytic lifestyle along the parasitism/mutualism continuum in *Arabidopsis*. *Curr. Opin. Microbiol.* 32:103-112.
- García, E., Alonso, A., Platas, G., and Sacristan, S. 2013. The endophytic mycobiota of *Arabidopsis thaliana*. *Fungal Divers.* 60:71-89.
- Gascuel, O. 1997. BIONJ: An improved version of the NJ algorithm based on a simple model of sequence data. *Mol. Biol. Evol.* 14: 685-695.
- Gazis, R., Kuo, A., Riley, R., LaButti, K., Lipzen, A., Lin, J., Amirebrahimi, M., Hesse, C. N., Spatafora, J. W., Henrissat, B., Hainaut, M., Grigoriev, I. V., and Hibbett, D. S. 2016. The genome of *Xylona heveae* provides a window into fungal endophytism. *Fungal Biol.* 120:26-42.
- Ge, X. T., and Barbetti, M. J. 2019. Host response of *Arabidopsis thaliana* ecotypes is determined by *Sclerotinia sclerotiorum* isolate type. *Eur. J. Plant Pathol.* 153:583-597.
- Giraldo, A., and Crous, P. W. 2019. Inside *Plectosphaerellaceae*. *Stud. Mycol.* 92:227-286.
- Götz, M., Nirenberg, H., Krause, S., Wolters, H., Draeger, S., Buchner, A., Lottmann, J., Berg, G., and Smalla, K. 2006. Fungal endophytes in potato roots studied by traditional isolation and cultivation-independent DNA-based methods. *FEMS Microbiol. Ecol.* 58:404-413.
- Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* 59:307-321.
- Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. 2013. QUASt: Quality assessment tool for genome assemblies. *Bioinformatics* 29: 1072-1075.
- Guyon, K., Balagué, C., Roby, D., and Raffaele, S. 2014. Secretome analysis reveals effector candidates associated with broad host range necrotrophy in the fungal plant pathogen *Sclerotinia sclerotiorum*. *BMC Genomics* 15:336.
- Hacquard, S., Kracher, B., Hiruma, K., Münch, P. C., Garrido-Oter, R., Thon, M. R., Weimann, A., Damm, U., Dallery, J.-F., Hainaut, M., Henrissat, B., Lespinet, O., Sacristán, S., Ver Loren van Themaat, E., Kemen, E., McHardy, A. C., Schulze-Lefert, P., and O'Connell, R. J. 2016. Survival trade-offs in plant roots during colonization by closely related beneficial and pathogenic fungi. *Nat. Commun.* 7:11362.
- Hardoim, P. R., van Overbeek, L. S., Berg, G., Pirttilä, A. M., Compant, S., Campisano, A., Döring, M., and Sessitsch, A. 2015. The hidden world within plants: Ecological and evolutionary considerations for defining functioning of microbial endophytes. *Microbiol. Mol. Biol. Rev.* 79: 293-320.
- Hernández-Blanco, C., Feng, D. X., Hu, J., Sánchez-Vallet, A., Deslandes, L., Llorente, F., Berrocal-Lobo, M., Keller, H., Barlet, X., Sánchez-Rodríguez, C., Anderson, L. K., Somerville, S., Marco, Y., and Molina, A. 2007. Impairment of cellulose synthases required for Arabidopsis secondary cell wall formation enhances disease resistance. *Plant Cell* 19:890-903.
- Hiruma, K., Gerlach, N., Sacristán, S., Nakano, R. T., Hacquard, S., Kracher, B., Neumann, U., Ramírez, D., Bucher, M., O'Connell, R. J., and Schulze-Lefert, P. 2016. Root endophyte *Colletotrichum tofieldiae* confers plant fitness benefits that are phosphate status dependent. *Cell* 165:464-474.
- Hoff, K. J., Lange, S., Lomsadze, A., Borodovsky, M., and Stanke, M. 2016. BRAKER1: Unsupervised RNA-Seq-based genome annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics* 32:767-769.
- Hongsanan, S., Sánchez-Ramírez, S., Crous, P. W., Ariyawansa, H. A., Zhao, R. L., and Hyde, K. D. 2016. The evolution of fungal epiphytes. *Mycosphere* 7:1690-1712.
- Horbach, R., Navarro-Quesada, A. R., Knogge, W., and Deising, H. B. 2011. When and how to kill a plant cell: Infection strategies of plant pathogenic fungi. *J. Plant Physiol.* 168:51-62.
- Huerta-Cepas, J., Capella-Gutiérrez, S., Pryszcz, L. P., Marcet-Houben, M., and Gabaldón, T. 2014. PhylomeDB v4: Zooming into the plurality of evolutionary histories of a genome. *Nucleic Acids Res.* 42:D897-D902.
- Jimenez, P., and Zitter, T. A. 2005. First report of *Plectosporium* blight on pumpkin and squash caused by *Plectosporium tabacinum* in New York. *Plant Dis.* 89:432.
- Jordá, L., Sopena-Torres, S., Escudero, V., Nuñez-Corcuera, B., Delgado-Cerezo, M., Torii, K. U., and Molina, A. 2016. ERECTA and BAK1 receptor like kinases interact to regulate immune responses in *Arabidopsis*. *Front. Plant Sci.* 7:897.
- Junker, C., Draeger, S., and Schulze, B. 2012. A fine line-endophytes or pathogens in *Arabidopsis thaliana*. *Fungal Ecol.* 5:657-662.
- Kameshwar, A. K. S., Pereira Ramos, L., and Qin, W. 2019. CAZymes-based ranking of fungi (CBRF): An interactive web database for identifying fungi with extrinsic plant biomass degrading abilities. *Bioresour. Bioprocess.* 6:51.
- Kamoun, S. 2009. The secretome of plant-associated fungi and oomycetes. *Plant Relationships. The Mycota (A Comprehensive Treatise on Fungi as Experimental Systems for Basic and Applied Research)*. H. B. Deising, ed. Vol. 5. Springer, Berlin.
- Katoh, K., Misawa, K., Kuma, K., and Miyata, T. 2002. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30:3059-3066.
- Kikot, G. E., Hours, R. A., and Alconada, T. M. 2009. Contribution of cell wall degrading enzymes to pathogenesis of *Fusarium graminearum*: A review. *J. Basic Microbiol.* 49:231-241.
- Kim, K.-T., Jeon, J., Choi, J., Cheong, K., Song, H., Choi, G., Kang, S., and Lee, Y.-H. 2016. Kingdom-wide analysis of fungal small secreted proteins (SSPs) reveals their potential role in host association. *Front. Plant Sci.* 7:186.
- Knapp, D. G., Németh, J. B., Barry, K., Hainaut, M., Henrissat, B., Johnson, J., Kuo, A., Lim, J. H. P., Lipzen, A., Nolan, M., Oh, R. A., Tamás, L., Grigoriev, I. V., Spatafora, J. W., Nagy, L. G., and Kovács, G. M. 2018. Comparative genomics provides insights into the lifestyle and reveals functional heterogeneity of dark septate endophytic fungi. *Sci. Rep.* 8:6321.
- Koch, E., and Slusarenko, A. 1990. Arabidopsis is susceptible to infection by a downy mildew fungus. *Plant Cell* 2:437-445.
- Korf, I. 2004. Gene finding in novel genomes. *BMC Bioinformatics* 5:59.
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol. Biol. Evol.* 35:1547-1549.
- Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., and Salzberg, S. L. 2004. Versatile and open software for comparing large genomes. *Genome Biol.* 5:R12.

- Laluk, K., and Mengiste, T. 2010. Necrotroph attacks on plants: Wanton destruction or covert extortion? *Arabidopsis Book* 8:e0136.
- Leveau, J. 2015. Life of microbes on aerial plant parts. *Principles of Plant-Microbe Interactions*. B. Lugtenberg, ed. Springer, Cham, Switzerland.
- Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv*:1303.3997.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. 2009. The sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* 25:2078-2079.
- Lindow, S. E., and Brandl, M. T. 2003. Microbiology of the phyllosphere. *Appl. Environ. Microbiol.* 69:1875-1883.
- Lipka, V., Dittgen, J., Bednarek, P., Bhat, R., Wiermer, M., Stein, M., Landtag, J., Brandt, W., Rosahl, S., Scheel, D., Llorente, F., Molina, A., Parker, J., Somerville, S., and Schulze-Lefert, P. 2005. Pre- and postinvasion defenses both contribute to nonhost resistance in *Arabidopsis*. *Science* 310:1180-1183.
- Llorente, F., Alonso-Blanco, C., Sánchez-Rodríguez, C., Jordá, L., and Molina, A. 2005. ERECTA receptor-like kinase and heterotrimeric G protein from *Arabidopsis* are required for resistance to the necrotrophic fungus *Plectosphaerella cucumerina*. *Plant J.* 43:165-180.
- Llorente, F., Muskett, P., Sánchez-Vallet, A., López, G., Ramos, B., Sánchez-Rodríguez, C., Jordá, L., Parker, J., and Molina, A. 2008. Repression of the auxin response pathway increases *Arabidopsis* susceptibility to necrotrophic fungi. *Mol. Plant* 1:496-509.
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42:D490-D495.
- Lugtenberg, B., Caradus, J., and Johnson, L. 2016. Fungal endophytes for sustainable crop production. *FEMS Microbiol. Ecol.* 92:fiw194.
- Lyu, X., Shen, C., Fu, Y., Xie, J., Jiang, D., Li, G., and Cheng, J. 2015. Comparative genomic and transcriptional analyses of the carbohydrate-active enzymes and secretomes of phytopathogenic fungi reveal their significant roles during infection and development. *Sci. Rep.* 5:15565.
- Lyu, X., Shen, C., Fu, Y., Xie, J., Jiang, D., Li, G., and Cheng, J. 2016. A small secreted virulence-related protein is essential for the necrotrophic interactions of *Sclerotinia sclerotiorum* with its host plants. *PLoS Pathog.* 12:e1005435.
- Malcolm, G. M., Kuldau, G. A., Gugino, B. K., and del Mar Jiménez-Gasco, M. 2013. Hidden host plant associations of soilborne fungal pathogens: An ecological perspective. *Phytopathology* 103:538-544.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M. A. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20:1297-1303.
- Medema, M. H., Blin, K., Cimermancic, P., de Jager, V., Zakrzewski, P., Fischbach, M. A., Weber, T., Takano, E., and Breitling, R. 2011. antiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* 39:W339-W346.
- Mélida, H., Sopena-Torres, S., Bacete, L., Garrido-Arandia, M., Jordá, L., López, G., Muñoz-Barrios, A., Pacios, L. F., and Molina, A. 2018. Non-branched  $\beta$ -1,3-glucan oligosaccharides trigger immune responses in *Arabidopsis*. *Plant J.* 93:34-49.
- Mine, A., Seyfferth, C., Kracher, B., Berens, M. L., Becker, D., and Tsuda, K. 2018. The defense phytohormone signaling network enables rapid, high-amplitude transcriptional reprogramming during effector-triggered immunity. *Plant Cell* 30:1199-1219.
- Morales, J., Kadota, Y., Zipfel, C., Molina, A., and Torres, M. A. 2016. The *Arabidopsis* NADPH oxidases RbohD and RbohF display differential expression patterns and contributions during plant immunity. *J. Exp. Bot.* 67:1663-1676.
- Newton, A. C., Fitt, B. D. L., Atkins, S. D., Walters, D. R., and Daniell, T. J. 2010. Pathogenesis, parasitism and mutualism in the trophic space of microbe-plant interactions. *Trends Microbiol.* 18:365-373.
- Oliver, R. P., and Ipcho, S. V. 2004. *Arabidopsis* pathology breathes new life into the necrotrophs-vs.-biotrophs classification of fungal pathogens. *Mol. Plant Pathol.* 5:347-352.
- Park, B. H., Karpinet, T. V., Syed, M. H., Leuze, M. R., and Uberbacher, E. C. 2010. CAZymes analysis toolkit (CAT): Web service for searching and analyzing carbohydrate-active enzymes in a newly sequenced organism using CAZy database. *Glycobiology* 20:1574-1584.
- Pastorczyk, M., Kosaka, A., Piślewska-Bednarek, M., López, G., Frerigmann, H., Kułak, K., Glawischnig, E., Molina, A., Takano, Y., and Bednarek, P. 2020. The role of CYP71A12 monooxygenase in pathogen-triggered tryptophan metabolism and *Arabidopsis* immunity. *New Phytol.* 225:400-412.
- Pawar, P. M.-A., Derba-Maceluch, M., Chong, S.-L., Gómez, L. D., Miedes, E., Banasiak, A., Ratke, C., Gaertner, C., Mouille, G., McQueen-Mason, S. J., Molina, A., Sellstedt, A., Tenkanen, M., and Mellerowicz, E. J. 2016. Expression of fungal acetyl xylan esterase in *Arabidopsis thaliana* improves saccharification of stem lignocellulose. *Plant Biotechnol. J.* 14:387-397.
- Plett, J. M., and Martin, F. M. 2018. Know your enemy, embrace your friend: Using omics to understand how plants respond differently to pathogenic and mutualistic microorganisms. *Plant J.* 93:729-746.
- Rai, M., and Agarkar, G. 2016. Plant-fungal interactions: What triggers the fungi to switch among lifestyles? *Crit. Rev. Microbiol.* 42:428-438.
- Ramos, B., Alves-Santos, F. M., García-Sánchez, M. A., Martín-Rodríguez, N., Eslava, A. P., and Díaz-Minguez, J. M. 2007. The gene coding for a new transcription factor (ftf1) of *Fusarium oxysporum* is only expressed during infection of common bean. *Fungal Genet. Biol.* 44:864-876.
- Ramos, B., González-Melendi, P., Sánchez-Vallet, A., Sánchez-Rodríguez, C., López, G., and Molina, A. 2013. Functional genomics tools to decipher the pathogenicity mechanisms of the necrotrophic fungus *Plectosphaerella cucumerina* in *Arabidopsis thaliana*. *Mol. Plant Pathol.* 14:44-57.
- Rastogi, G., Coaker, G. L., and Leveau, J. H. J. 2013. New insights into the structure and function of phyllosphere microbiota through high-throughput molecular approaches. *FEMS Microbiol. Lett.* 348:1-10.
- Rawlings, N. D., Barrett, A. J., and Finn, R. 2016. Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* 44 (D1):D343-D350.
- Rieu, I., and Powers, S. J. 2009. Real-time quantitative RT-PCR: Design, calculations, and statistics. *Plant Cell* 21:1031-1033.
- Saier, M. H., Jr., Reddy, V. S., Tamang, D. G., and Västermark, A. 2014. The transporter classification database. *Nucleic Acids Res.* 42:D251-D258.
- Sánchez-Rodríguez, C., Estévez, J. M., Llorente, F., Hernández-Blanco, C., Jordá, L., Pagán, I., Berrocal, M., Marco, Y., Somerville, S., and Molina, A. 2009. The ERECTA receptor-like kinase regulates cell wall-mediated resistance to pathogens in *Arabidopsis thaliana*. *Mol. Plant-Microbe Interact.* 22:953-963.
- Sánchez-Vallet, A., López, G., Ramos, B., Delgado-Cerezo, M., Riviere, M.-P., Llorente, F., Fernández, P. V., Miedes, E., Estevez, J. M., Grant, M., and Molina, A. 2012. Disruption of abscisic acid signaling constitutively activates *Arabidopsis* resistance to the necrotrophic fungus *Plectosphaerella cucumerina*. *Plant Physiol.* 160:2109-2124.
- Sánchez-Vallet, A., Ramos, B., Bednarek, P., López, G., Piślewska-Bednarek, M., Schulze-Lefert, P., and Molina, A. 2010. Tryptophan-derived secondary metabolites in *Arabidopsis thaliana* confer non-host resistance to necrotrophic *Plectosphaerella cucumerina* fungi. *Plant J.* 63:115-127.
- Schulz, B., and Boyle, C. 2005. The endophytic continuum. *Mycol. Res.* 109:661-686.
- Seidl, M. F., Faino, L., Shi-Kunne, X., van den Berg, G. C. M., Bolton, M. D., and Thomma, B. P. H. J. 2015. The genome of the saprophytic fungus *Verticillium tricorpus* reveals a complex effector repertoire resembling that of its pathogenic relatives. *Mol. Plant-Microbe Interact.* 28:362-373.
- Selosse, M.-A., Schneider-Maunoury, L., and Martos, F. 2018. Time to rethink fungal ecology? Fungal ecological niches are often prejudged. *New Phytol.* 217:968-972.
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210-3212.
- Spanu, P. D., and Panstruga, R. 2017. Editorial: Biotrophic plant-microbe interactions. *Front. Plant Sci.* 8:192.
- Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S., and Morgenstern, B. 2006. AUGUSTUS: Ab initio prediction of alternative transcripts. *Nucleic Acids Res.* 34:W435-W439.
- Stein, M., Dittgen, J., Sánchez-Rodríguez, C., Hou, B. H., Molina, A., Schulze-Lefert, P., Lipka, V., and Somerville, S. 2006. *Arabidopsis* PEN3/PDR8, an ATP binding cassette transporter, contributes to nonhost resistance to inappropriate pathogens that enter by direct penetration. *Plant Cell* 18:731-746.
- Stergiopoulos, I., and de Wit, P. J. 2009. Fungal effector proteins. *Annu. Rev. Phytopathol.* 47:233-263.
- Su, L., Deng, H., and Niu, Y.-C. 2017. Phylogenetic analysis of *Plectosphaerella* species based on multi-locus DNA sequences and description of *P. sinensis* sp. nov. *Mycol. Prog.* 16:823-829.
- Subramanian, A. R., Kaufmann, M., and Morgenstern, B. 2008. DIALIGN-TX: Greedy and progressive approaches for segment-based multiple sequence alignment. *Algorithms Mol. Biol.* 3:6.

- Takashima, M., Manabe, R.-I., and Ohkuma, M. 2019. Draft genome sequences of basidiomycetous epiphytic phylloplane yeast type strains *Dioszegia crocea* JCM 2961 and *Dioszegia aurantiaca* JCM 2956. *Microbiol. Resour. Announc.* 8:e01727-18.
- Tamura, K., and Nei, M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10:512-526.
- Thiergart, T., Durán, P., Ellis, T., Vannier, N., Garrido-Oter, R., Kemen, E., Roux, F., Alonso-Blanco, C., Ågren, J., Schulze-Lefert, P., and Hacquard, S. 2020. Root microbiota assembly and adaptive differentiation among European *Arabidopsis* populations. *Nat. Ecol. Evol.* 4: 122-131.
- Tierens, K. F., Thomma, B. P., Brouwer, M., Schmidt, J., Kistner, K., Porzel, A., Mauch-Mani, B., Cammue, B. P., and Broekaert, W. F. 2001. Study of the role of antimicrobial glucosinolate-derived isothiocyanates in resistance of *Arabidopsis* to microbial pathogens. *Plant Physiol.* 125: 1688-1699.
- Ton, J., and Mauch-Mani, B. 2004. Beta-amino-butyric acid-induced resistance against necrotrophic pathogens is based on ABA-dependent priming for callose. *Plant J.* 38:119-130.
- Torres, M. A., Dangl, J. L., and Jones, J. D. G. 2002. *Arabidopsis* gp91<sup>phox</sup> homologues *AtrbohD* and *AtrbohF* are required for accumulation of reactive oxygen intermediates in the plant defense response. *Proc. Natl. Acad. Sci. U.S.A.* 99:517-522.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H., Salzberg, S. L., Rinn, J. L., and Pachter, L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7:562-578.
- Tsuda, K., Sato, M., Stoddard, T., Glazebrook, J., and Katagiri, F. 2009. Network properties of robust immunity in plants. *PLoS Genet.* 5: e1000772.
- Usami, T., and Katagiri, H. 2017. Pathogenicity of *Plectosphaerella* species on lettuce and susceptibility of lettuce cultivars. *J. Gen. Plant Pathol.* 83: 366-372.
- Vitale, S., Maccaroni, M., and Belisario, A. 2007. First report of zucchini collapse by *Fusarium solani* f. sp. *cucurbitae* race 1 and *Plectosporium tabacinum* in Italy. *Plant Dis.* 91:325.
- Vorholt, J. A. 2012. Microbial life in the phyllosphere. *Nat. Rev. Microbiol.* 10:828-840.
- Wallace, I. M., O'Sullivan, O., Higgins, D. G., and Notredame, C. 2006. M-Coffee: Combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Res.* 34:1692-1699.
- Wan, J., Zhang, X. C., Neece, D., Ramonell, K. M., Clough, S., Kim, S. Y., Stacey, M. G., and Stacey, G. 2008. A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in *Arabidopsis*. *Plant Cell* 20:471-481.
- Wang, B., Liang, X., Gleason, M. L., Zhang, R., and Sun, G. 2017. Genome sequence of the ectophytic fungus *Ramichloridium luteum* reveals unique evolutionary adaptations to plant surface niche. *BMC Genomics* 18:729.
- Wang, M., Liu, B., Ruan, R., Zeng, Y., Luo, J., and Li, H. 2020. Genomic sequencing of *Phyllosticta citriasiana* provides insight into its conservation and diversification with two closely related *Phyllosticta* species associated with Citrus. *Front. Microbiol.* 10:2979.
- Wang, X., Jiang, N., Liu, J., Liu, W., and Wang, G. L. 2014. The role of effectors and host immunity in plant-necrotrophic fungal interactions. *Virulence* 5:722-732.
- Wang, X., Zhang, X., Liu, L., Xiang, M., Wang, W., Sun, X., Che, Y., Guo, L., Liu, G., Guo, L., Wang, C., Yin, W. B., Stadler, M., Zhang, X., and Liu, X. 2015. Genomic and transcriptomic analysis of the endophytic fungus *Pestalotiopsis fici* reveals its lifestyle and high potential for synthesis of natural products. *BMC Genomics* 16:28.
- Whipps, J. M., Hand, P., Pink, D., and Bending, G. D. 2008. Phyllosphere microbiology with special reference to diversity and plant genotype. *J. Appl. Microbiol.* 105:1744-1755.
- Windram, O., Madhou, P., McHattie, S., Hill, C., Hickman, R., Cooke, E., Jenkins, D. J., Penfold, C. A., Baxter, L., Breeze, E., Kiddle, S. J., Rhodes, J., Atwell, S., Kliebenstein, D. J., Kim, Y. S., Stegle, O., Borgwardt, K., Zhang, C., Tabrett, A., Legaie, R., Moore, J., Finkenstadt, B., Wild, D. L., Mead, A., Rand, D., Beynon, J., Ott, S., Buchanan-Wollaston, V., and Denby, K. J. 2012. *Arabidopsis* defense against *Botrytis cinerea*: Chronology and regulation deciphered by high-resolution temporal transcriptomic analysis. *Plant Cell* 24:3530-3557.
- Xu, C., Chen, H., Gleason, M. L., Xu, J. R., Liu, H., Zhang, R., and Sun, G. 2016. *Peltaster fructicola* genome reveals evolution from an invasive phytopathogen to an ectophytic parasite. *Sci. Rep.* 6:22926.
- Xu, X.-H., Su, Z.-Z., Wang, C., Kubicek, C. P., Feng, X.-X., Mao, L.-J., Wang, J.-Y., Chen, C., Lin, F.-C., and Zhang, C.-L. 2014. The rice endophyte *Harpophora oryzae* genome reveals evolution from a pathogen to a mutualistic endophyte. *Sci. Rep.* 4:5783.
- Yao, H., Sun, X., He, C., Maitra, P., Li, X.-C., and Guo, L.-D. 2019. Phyllosphere epiphytic and endophytic fungal community and network structures differ in a tropical mangrove ecosystem. *Microbiome* 7:57.
- Yu, Q., and Coosemans, J. 1998. Fungi associated with cysts of *Globodera rostochiensis*, *G. pallida*, and *Heterodera schachtii*; and egg masses and females of *Meloidogyne hapla* in Belgium. *Phytoprotection* 79:63-69.
- Yuan, Y., Teng, Q., Zhong, R., and Ye, Z. H. 2013. The *Arabidopsis* DUF231 domain-containing protein ESK1 mediates 2-O- and 3-O-acetylation of xylosyl residues in xylan. *Plant Cell Physiol.* 54:1186-1199.
- Zaharia, M., Bolosky, W. J., Curtis, K., Fox, A., Patterson, D., Shenker, S., Stoica, I., Karp, R. M., and Sittler, T. 2011. Faster and more accurate sequence alignment with SNAP. arXiv:1111.5572v1.
- Zeilinger, S., Gupta, V. K., Dahms, T. E. S., Silva, R. N., Singh, H. B., Upadhyay, R. S., Gomes, E. V., Tsui, C. K., and Nayak S, C. 2016. Friends or foes? Emerging insights from fungal interactions with plants. *FEMS Microbiol. Rev.* 40:182-207.
- Zerbino, D. R., and Birney, E. 2008. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18:821-829.
- Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. 2000. A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* 7:203-214.
- Zhao, Y., Hull, A. K., Gupta, N. R., Goss, K. A., Alonso, J., Ecker, J. R., Normanly, J., Chory, J., and Celenza, J. L. 2002. Trp-dependent auxin biosynthesis in *Arabidopsis*: Involvement of cytochrome P450s CYP79B2 and CYP79B3. *Genes Dev.* 16:3100-3112.
- Zhao, Z., Liu, H., Wang, C., and Xu, J.-R. 2013. Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. *BMC Genomics* 14:274.
- Zhou, J., Bi, S., Chen, H., Chen, T., Yang, R., Li, M., Fu, Y., and Jia, A.-Q. 2017. Anti-biofilm and antiviral activities of metabolites from *Plectosphaerella cucumerina* against *Pseudomonas aeruginosa*. *Front. Microbiol.* 8:769.