Dirigent proteins in plants – modulating cell wall metabolism during abiotic and biotic stress exposure

Candelas Paniagua[1], Anna Bilkova[1,2], Phil Jackson[1], Siarhei Dabravolski[1], Willi Riber[1], Vojtech Didi[1], Josef Houser[3], Nora Gigli Bisceglia[4], Michaela Wimmerova[3], Eva Budínská[2], Thorsten Hamann[4] and Jan Hejatko[1]

[1]Laboratory of Molecular Plant Physiology and Functional Genomics and Proteomics of Plants, CEITEC - Central European Institute of Technology and National Centre for Biomolecular Research, Masaryk University Kamenice 5, CZ-625 00 Brno, Czech Republic.

[2]Research Centre for Toxic Compounds in the Environment (RECETOX), Masaryk University Kamenice 5, CZ-625 00 Brno, Czech Republic.

[3]Glycobiochemistry, CEITEC - Central European Institute of Technology and National Centre for Biomolecular Research, Masaryk University Kamenice 5, CZ-625 00 Brno, Czech Republic.

[4]Department of Biology, Norwegian University of Science and Technology 5, Hogskoleringen, N-7491 Trondheim, Norway.

Corresponding author:

Jan Hejátko, hejatko@sci.muni.cz, tel. +420 5 4949 4165

Running title: Dirigent proteins – insights into the control of plant stress response

Author's emails:

Candelas Paniagua: candelas.paniagua@ceitec.muni.cz ; Anna Bílková: ana.bilkova@gmail.com; Siarhei Dabravolski: sergedobrowolski@gmail.com; Phil Jackson: phil@itqb.unl.pt; Willi Riber: willi.riber@web.de; Vojtěch Didi: didi@sci.muni.cz; Eva Budínská: budinska@iba.muni.cz; Josef Houser: houser@mail.muni.cz; Michaela

Wimmerová: michaw@chemi.muni.cz; Nora Gigli Bisceglia: nora.g.bisceglia@ntnu.no; Thorsten Hamann: thorsten.hamann@ntnu.no

Date of submission: 30th November 2016
Number of figures: 6
Total Word count: 8249
Supplementary Data table: 2
Supplementary Data figures: 7

Highlight

DIRIGENT proteins mediate lignan and lignin biosynthesis and could be important regulators of plant development, connecting hormonal-mediated stress responses with (not only) cell wall metabolism.

Abstract

Dirigent (DIR) proteins were found to mediate regio- and stereo-selectivity of bimolecular phenoxy radical coupling during lignan biosynthesis. Here we summarize the current knowledge of the importance of DIR proteins in lignan and lignin biosynthesis and highlight their possible importance in plant development. We focus on the still rather enigmatic Arabidopsis DIR gene family, discussing the few members with known functional importance. We comment on recent discoveries describing the detailed structure of two DIR proteins with implications in the mechanism of DIR-mediated catalysis. Further, we summarize the ample evidence for stress-induced dirigent gene expression, suggesting the role of DIRs in adaptive responses. In the second part of our work, we present a preliminary bioinformatics-based characterization of the AtDIR family. The phylogenetic analysis of AtDIRs complemented by comparison with DIR proteins of mostly known function from other species allowed us to suggest possible roles for several members of this family and identify interesting AtDIR targets for further study. Finally, based on the available metadata and our *in silico* analysis of *AtDIR*

promoters, we hypothesize about the existence of specific transcriptional controls for individual *AtDIRs* and implicate them in various stress responses, hormonal regulations and developmental processes.

Key words:

Dirigent protein, cell walls, lignin, lignan, stereoselectivity, regioselectivity, biotic and abiotic stress response.

Introduction

Dirigent proteins – when the stereoselectivity matters

Pasteur (1860) demonstrated that many organic molecules form enantiomeric pairs of non-superimposable mirror-image molecular structures, characterized by their oppositely signed optical rotation. However, while in vitro synthesis produces racemic mixtures consisting of equal numbers of left- and right-handed molecules, biosynthesis often provides just one of the two enantiomers. Based on this, Pasteur concluded that biosynthesis involves a chiral force. That was further exemplified in the beginning of the 20th century by Fischer (1890-1919), who demonstrated that functional biomolecules are composed specifically of the D-sugars and the L-amino acids (Mason, 1991). Since then, homochirality, i.e. using preferentially one of the possible enantiomers, was proposed as an intrinsic feature to all living systems and stereospecificity of biosynthetic reactions is considered as a necessary prerequisite for life's origin on the Earth (Weber and Pizzarello, 2006).

The regio- and stereo-selectivity of bimolecular phenoxy radical coupling reactions takes place in various biosynthetic pathways, including the production of suberin, flavonolignans and alkaloids in plants, fruiting body development in fungi, cuticle melanization and sclerotization in insects, the formation of pigments in aphids and cell wall polymers in algae (reviewed by (Davin *et al.*, 1997)). In vascular plant development, the bimolecular phenoxy radical coupling has especial importance in the biosynthesis of

lignin (Nose *et al.*, 1995) and lignan (Davin and Lewis, 2000; Paré *et al.*, 1994).

It has been established that the oxidative coupling of coniferyl alcohol by peroxidases or laccases *in vitro* lacks regio- and stereo-selectivity (Fig. 1), resulting in a mixture of (+/-) 8,8', (+/-) 8,5' and 8-O-4' linkages (Fig. 1). However, vascular plants are known to be capable of forming high proportions of regio- and stereo-selectivity products of this reaction, including 8,8' coupled (+)-pinoresinol, indicating the presence of a coordinating factor, which was initially named (+)-pinoresinol synthase (Paré *et al.*, 1994). The factors determining the stereoselective coupling of coniferyl alcohol (CA) radicals in the biosynthesis of (+)-pinoresinol were further investigated using crude cell wall extracts of *Forsythia suspensa* (Davin *et al.*, 1997). Nonspecific (lacking both regio- and stereo-selectivity) bimolecular radical coupling was obtained when a protein fraction containing a laccase was added to CA *in vitro,* resulting in the production of a racemic mixture of (+/-)-dehydrodiconiferyl alcohols, (+/-)-pinoresinols and (+/-)-eryrhro/threo guaiacylglycerol 8-0-4' coniferyl alcohol ethers (Fig. 1). However, when the reaction mixture was combined with fraction containing a ~27 kDa protein, both regio- and stereo-selectivity was restored, resulting in the production of essentially only (+)-pinoresinol. The ~27 kD protein of (at that time) unknown nature was designated as a DIRIGENT (DIR) protein (from Latin *dirigere,* to align, guide) and suggested to function by capturing CA radicals produced by an oxidizing compound to mediate stereoselective coupling of (+)-pinoresinol (Davin *et al.*, 1997).

In order to further characterize DIR proteins and to establish their involvement in lignan and lignin formation, Gang *et* al. (1999) cloned several genes encoding DIR proteins from different species. The recombinant proteins produced were able to confer strict regio- and stereo- selectivity of the monolignol free-radical coupling. The *DIR* expression in developing xylem and in other lignified tissues indicated the roles of DIRs in lignification (Burlat *et al.*, 2001).

Further studies identified DIR proteins in at least 104 species throughout the plant kingdom, suggesting that DIR families are probably present in all vascular plants (Davin *et al.*, 2008). In addition to DIR proteins involved in (+)-pinoresinol formation, the presence of (-)-pinoresinol forming DIRs has been suggested in species like Daphne

(Daphne tangutica) and flax (*Linum usitatissimum)* (reviewed by (Kim *et al.*, 2012a). Direct evidence for the involvement of DIR proteins from *Arabidopsis thaliana* in the formation of (-)-pinoresinol has been provided in case of AtDIR6, one of the few functionally characterized DIRs from *Arabidopsis* (Pickel *et al.*, 2010; Vassão *et al.*, 2010).

In the following text we provide an overview of the known and possible contributions of DIR proteins to lignan and lignin biosynthesis. We discuss recent structural insights into the hypothetical mechanism of the action of DIR proteins and recapitulate current evidence and implications for the role of DIR proteins in CW signaling as well as in biotic and abiotic stress responses. We provide a basic bioinformatics-based characterization of the *Arabidopsis thaliana* DIR family (AtDIRs) and suggest possible approaches useful in our and others' efforts to decipher the role of the fascinating DIR proteins in plant development and stress responses.

Function of DIR proteins

Stereoselective radical-radical coupling in lignan biosynthesis

Lignins and lignans are derived from phenylpropanoid metabolism. This pathway leads to the production of monolignols (coniferyl, sinapyl and p-coumaryl alcohols), which are precursors in both lignan and lignin biosynthesis. (Buchanan *et al.*, 2000) (Supplementary Fig. S1). The term lignan specifies a class of dimeric phenylpropanoid (C_6C_3) metabolites linked by 8-8´bond, while alternatively linked dimers are known as neolignans (Buchanan *et al.*, 2000). Twenty-three lignan and neolignan types across the plant kingdom have been described (Teponno *et al.*, 2016), indicating their wide range of usage. The lignan (and neolignan) biosynthetic pathway starts with the synthesis of phenylalanine, a precursor of coniferyl alcohol (Barros *et al.*, 2015; Hao and Mohnen, 2014). The dimerization/radical coupling of two coniferyl alcohol molecules leading to (+/-)-pinoresinol formation is mediated by oxidases, such as peroxidases or laccases with the assistance of DIRs, which ensure the stereoselectivity of coniferyl alcohol

dimerization (Davin *et al.*, 1997; Halls *et al.*, 2004; Halls and Lewis, 2002). This step is crucial because the optical activity is a property-determining feature in most of the lignans (Akiyama *et al.*, 2007a; Akiyama *et al.*, 2009). Pinoresinol can be converted to other lignan types including piperitol, laciresinol, sesamin, secoisolaresinol and their glucosides (Dinkova-Kostova *et al.*, 1996; Satake *et al.*, 2015).

Lignin initiation

Lignins are three-dimensional, amorphous heteropolymers with species-specific compositions. In angiosperms, lignin consists mainly of coniferyl and sinapyl alcohols, with *p*-coumaryl alcohol contributing only approx. 2% (Bonawitz *et al.*, 2014). Lignin polymerization can involve at least five different linkage sites to form p-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) units, which differ in their degree of methoxylation (Supplementary Fig. S1). The monolignol polymerization is initiated by the formation of lignin oligomers through an 'end-wise' reaction, which involves the 8 carbon-site of one monomeric phenylpropanoid to a different acceptor site in the growing lignin oligomer and occurs via enzyme-mediated formation of phenylpropanoid radicals (reviewed by (Behr *et al.*, 2015)). This reaction differs from phenylpropanoid dimerization in lignan biosynthesis, where a 8-8' bond is preferentially formed resulting in lignan precursors. In addition, lignin, in contrast to lignan, is an optically inactive complex polymer. However, even if the stereoselectivity is unnecessary, the regioselectivity of the coupling reaction is still needed to confer the observed predominance of 8-O-4' bonds in plant-isolated lignin (Davin and Lewis (2005b) and references therein). Moreover, monolignols can be targeted to precise sites, called lignin initiation sites. For example, during tracheid formation, *p*-coumaryl alcohol is targeted toward the middle lamella, whereas coniferyl alcohol is initially deposited in the S1 sublayer and cell corners, where the lignification is initiated and extends back toward the plasma membrane (Davin and Lewis, 2000). Although the detailed molecular mechanism is still unclear, specific gene expression seems to be involved in the control over the lignin composition of each cell (Liu, 2012).

The evidence supporting the involvement of protein(s) harboring an array of dirigent

(monolignol radical) binding sites in lignin biosynthesis was provided by a co-localizing signal generated by αDIR polyclonal antibodies with lignin initiation sites, i.e. in the S1 layer of the secondary cell walls of lignifying tracheary elements. It was suggested that a protein with similarity to DIR proteins, i.e. able to bind monolignol radicals, could be responsible for both directing the targeting of coniferyl alcohol to the lignification initiation sites and the regioselectivity resulting in the predominance of 8-O-4' bonds (Davin and Lewis, 2000). Davin *et* al (2008) highlighted that such a protein and its dirigent sites have to be distinctly different from those involved in stereoselective radical–radical coupling leading to lignan biosynthesis, such as to (+)-pinoresinol or (-)-pinoresinol. This prediction seems to be in line with recent findings implicating ENHANCED SUBERIN1 (ESB1/AtDIR10) in Casparian strips (CS) formation in Arabidopsis (Hosmani *et al.*, 2013), *vide infra*. Further, one of the major quantitative trait loci in soya encodes a dirigent-like protein PDH1, involved in the regulation of soybean pod dehiscence. Based on the gene-expression data, PDH1 seems to be also involved in the control of lignin deposition (Funatsuki *et al*., 2014).

Other forms of stereoselective and regioselective coupling mediated by DIRs: in *Atropselective synthesis of gossipol*

Besides the established roles of DIR proteins in lignan and lignin formation, an additional one has been demonstrated in the formation of the phenolic terpenoid (+)-gossypol in cotton (*Gossypium sp.*) (Effenberger *et al.*, 2015; Liu *et al.*, 2008). (+)-gossypol is a phenolic aldehyde found in flowers, seeds, roots and the foliage of cotton plants and it has an important role in the response to pathogens (Gao *et al.*, 2013). Dirigent proteins from *Gossypium hirsutum* (GhDIR14) (Liu *et al.*, 2008)*, G. barbadense* (GbDIR2) *and G. hirsutum (*GhDIR3*)* (Effenberger et al., 2016) confer atropselectivity to this terpenoid, Gossipol, formed from C-C coupling of two hemigossypol radicals. The possible rotation around the binaphthyl C-C bond allows the existence of two atropisomers with different properties and optical activities. When hemigossypol is incubated with peroxidase/H_2O_2, laccase/O_2 or ammonium persulfate, only a racemic mixture is obtained (Benedict *et al.*, 2006). However, when hemigossypol is incubated

with crude extracts containing DIR activity, the (+)-gossypol form was preferentially formed (Liu *et al.*, 2008). Both + and - isomers are involved in plant defense. Nevertheless, only (-)-gossypol has antispermatogenic as well as antiviral activities and only the (+) isomer is toxic to non-ruminant animals (Effenberger *et al.*, 2016). In this case, the control of atropselectivity could be of biotechnological relevance. For instance, the generation of cotton transgenic plants lacking the toxic (-)-gossypol but retaining (+)-gossypol (with role in plant defense) might enable use of cotton seeds as a source of food protein (Effenberger *et al.*, 2015).

The importance of lignans and lignin

Lignans have an important role in plant defense against pathogens (Davin *et al.*, 2008; Davin and Lewis, 2005b), by inhibiting microbe-derived degradative enzymes such as cellulases, polygalacturonases, glucosidases and laccases (MacRae and Towers, 1984). Additionally, it has been suggested that lignans can function as insect antifeedants by disrupting the insect endocrine system (Harmatha and Dinan, 2003). Importantly, lignans could be also used as drugs and chemopreventive agents in conventional medicine. Podophyllotoxin (from *Podophyllum peltatum*) has antiviral properties and one of its derivatives (Etopophos®) has applications in cancer chemotherapy (reviewed by (Davin *et al.*, 2008)).

Moreover, lignans could serve as storage pool of monolignols for lignification. Expression of lignan synthesis genes was upregulated during xylogenesis in lignified tissues of maritime pine and flax (Huis *et al.*, 2012; Villalobos *et al.*, 2012). Immunolabelling experiments revealed the presence of lignans in the secondary cell walls of flax (Attoumbre *et al.*, 2010). Interestingly, functional disruption of PINORESINOL REDUCTASE 1 (PrR1), which catalyzes the conversion of pinoresinol to lacinoresinol, leads to decreased lignin content and altered distribution in the inflorescence stem of Arabidopsis (Ruprecht *et al.*, 2011; Zhao *et al.*, 2015). The promoter of *PrR1* was also shown to be under the transcriptional control of secondary

cell wall-specific (and thus lignification inducing) TFs, NAC secondary wall thickening promoting factor (NST3) and MYB46 (Zhao *et al.*, 2015).

Lignin confers stability and hydrophobicity to the plant vascular system, forms a barrier against microbial pathogens to limit the spread of pathogen-derived toxins and enzymes into the host by altering compressibility and porosity of the cell wall (Bonello *et al.*, 2003; Miedes *et al.*, 2014). Moreover, lignin plays an important role in the spatial delimitation of silique shattering for seed release, as well as in seed protection (reviewed in Barros *et al.* 2015). The ability of lignin to control the permeability of cell wall is of especial importance in the case of Casparian strip (CS) formation, where lignification occurs before secondary cell wall formation and is thought to form a diffusion barrier regulating vascular solute transport (*vide infra*).

Dirigent proteins in Arabidopsis

AtDIRs in lignan biosynthesis

Arabidopsis DIRIGENT and DIRIGENT-LIKE (AtDIR) proteins constitute a protein family with 26 members whose specific functions are still barely understood (Kim *et al.*, 2012; Ralph *et al.*, 2006). Kim *et* al (2012) investigated 16 out of 26 AtDIRs. In this study, several *AtDIR* genes were cloned and their recombinant proteins biochemically analyzed. In particular, AtDIR5 and AtDIR6 (but not AtDIR10/ESB1 or AtDIR13) were found to be able to generate (-)-pinoresinol *in vitro* when incubated together with a laccase. Accordingly, increasing *AtDIR6* expression levels enhanced the abundance of (-)-pinoresinol and (-)-lariciresinol (a derivative of (-)-pinoresiol), indicating that *AtDIR6* was involved in preferential coupling leading to (-)-pinoresinol *in vivo*.

The analysis of promoter::GUS reporter constructs suggested that *AtDIR6*, *AtDIR10/ESB1* and *AtDIR13* are strongly but not exclusively expressed in the root (Kim *et al.*, 2012). *AtDIR10/ESB1* was also expressed in the lignifying leaf vasculature and hydathodes and activity of *AtDIR13* was detectable in cotyledons (Vassão *et al.*, 2010, Kim *et al.*, 2012). In contrast, the expression of A*tDIR12/DP1* seems to be specific for

the outer-seed coat at 7 days post anthesis (Esfandiari *et al.*, 2013). Accordingly, knocking-out of *AtDIR12/DP1* resulted in a lack of seed-specific neolignans such as 3-(4-[2-hydroxy-2-(4-hexosyloxy-3-methoxyphenyl)-1-hydroxymethylethoxy]-3,5 dimethoxyphenyl} acryloylcholine (Bottcher *et al.*, 2008; Matsuda *et al.*, 2010) suggesting a role of *AtDIR12/DP1* in neolignan biosynthesis.

AtDIRs in Casparian strip formation

The Casparian strip (CS) is mainly composed of lignin and forms a longitudinally-oriented belt passing through both transversal and anticlinal cell walls in the root endodermis (Naseer *et al.*, 2012). The CS extends across cell junctions and integrates middle lamella of adjacent epidermal cells to form a continuous ring of lignin. Consequently, the CS constitutes a physical and chemical barrier (analogous to tight junctions in animal), tightly controlling water and nutrient transport while providing protection against soil-borne pathogens (Geldner, 2013).

The lignification of the CS occurs in a tightly controlled manner indicative of a precise targeting mechanism with subcellular resolution. Recent studies demonstrated that the localization of CS domain proteins (CASPs) to specific cell membrane domains is indicative of the site of CS formation. Here, CASPs are thought to form a protein scaffold directing transport across the plasma membrane and recruiting proteins required for CS formation and lignification (Roppolo *et al.*, 2011). The DIR domain-containing protein, AtDIR10/ESB1 was shown to be targeted to the CS in a CASP-dependent manner. The *esb1-1* mutant showed that ESB1 is required both in the early deposition of lignin patches and their fusion in generating the mature CS. Loss of *AtDIR10/ESB1* resulted in the disruption of CASP1 localization, suggesting a reciprocal requirement for both AtDIR10/ESB1 and CASPs in the spatial control CS lignification. The lack of AtDIR10/ESB1 also provoked an ectopic deposition of suberin suggesting a possible cross-talk between CS and suberin biosynthesis (Hosmani *et al.*, 2013). Alternatively, the phenotype could be caused by the activation of the receptor-like cytoplasmic kinase (SCHENGENES (SGN)1-SGN3) pathway by the CASPARIAN STRIP INTEGRITY (CIF)

peptides lost from the stele due to the defect in the CS barrier in the *esb1-1* mutant (Doblas *et al.*, 2017).

Recently, MYB36 has been suggested as the transcriptional regulator controlling the expression of genes required for CS formation. Analysis of microarray data of *myb36* knock out mutants predicted 23 endodermal-expressed genes to be regulated by MYB36. Among these are genes encoding the DIR proteins *ESB1-ESB5 and ESB-like DIRS* (*AtDIR9, 10, 16, 18, 19, 24*), CASPs, PEROXIDASE 64 (PER64), and a leucine-rich repeat receptor-like kinases (LRR-RLK). Additionally, chromatin immunoprecipitation (ChIP)-qPCR using the MYB36 genome-GFP/myb36-1 showed a direct association with CASP1, PER64 and ESB1 promoters (Kamiya *et al.*, 2015). These findings suggest that other members of the AtDIR family could be important for CS and/or secondary CW formation. That, however, remains to be shown.

Structure and mode of action of pinoresinol-forming AtDIRs

One of the first published attempts to decipher the structural features of DIR proteins was the homology-based model of AtDIR6 (Pickel *et al.*, 2012). The model predicted monomer of AtDIR6 as an eight-stranded antiparallel β-barrel with a central hydrophobic cavity for substrate binding, structurally resembling allene oxide cyclase (Hofmann *et al.*, 2006; Pickel *et al.*, 2012). Based on earlier biochemical and kinetic data (Davin and Lewis, 2005a; Halls *et al.*, 2004; Halls and Lewis, 2002), it was suggested that DIR proteins form dimers, where each monomer binds a single CA radical (produced by laccase or oxidase) in a way that favors 8-8'-coupling (Halls *et al.*, 2004). However, the crystal structure obtained for the (+)-pinoresinol-forming DISEASE RESISTANCE RESPONSE 206 (PsDRR206) from pea, indicated a tightly packed homotrimer with the hydrophobic binding pockets placed on the outer surface of each monomer and too distant from each other to allow inter-pocket CA side-chain interactions. Based on this, it was proposed that each binding site enables stereoselective coupling (using either two CA radicals or a radical and a monolignol) (Kim *et al.*, 2015). The trimeric nature of DIRs and the docking of two CA molecules per DIR monomer was recently supported from

311 structural studies of AtDIR6, a (-)-pinoresinol-forming DIR (Gasper *et al.*, 2016; Kim *et al*
312 ., 2012).

313 In AtDIR6 (Fig. 2), the binding cavity of each monomer was seen to consist of two lobes
314 (pockets A and B), with each lined with a set of hydrophilic and potentially catalytic
315 residues (Figs. 3). These residues are conserved between (+)- and (-)-pinoresinol-
316 forming DIRs and are required for DIR activity (Gasper *et al.*, 2016; Kim *et al* ., 2012).

317 Comparing the structures of PsDRR206 and AtDIR6 (Supplementary Fig. S2) highlights
318 several important differences. First, there is important difference in the architecture of
319 the β1-β2 loop of both proteins. In AtDIR6, there is one more short β sheet (β1') in the
320 region corresponding to the β1-β2 loop of PsDRR206. Further, the β1-β2 loop of
321 PsDRR206 traverses the interaction interface between the neighboring monomers. This
322 places the N-terminal part of the β2 sheet of one monomer of PsDRR206 next to the β2
323 of the interacting monomer (Fig. S2). These two β2 strands contribute to the interaction
324 interface by seven hydrogen bonds between their backbone chains (Kim *et al.*, 2015).
325 No such domain-swapping could be seen in AtDIR6 (Gasper *et al.*, 2016). Second, there
326 are striking differences in the architecture of the active sites of both proteins. In
327 PsDRR206, the loops surrounding the active center are more flexible and therefore not
328 well resolved in the structure. Notably, in contrast to AtDIR6, where the loops are bent
329 inward, the loops of AtDRR206 bent outwards of the active center, making it more open
330 (Supplementary Fig. S2, B-D) (Gasper *et al.*, 2016). This has an important consequence
331 for the positioning of the key residues of pocket A, which are well (ideally for catalysis)
332 positioned in case of AtDIR6 (Asp137, Arg144). However, this is not the case for
333 corresponding residues of PsDRR206 (Asp134, Arg141), which are protruding out to the
334 solvent and forming the hydrogen bond with β5, respectively (Fig. S2B). Based on that,
335 Apo-DRR206 was proposed to exist in a noncatalytic state, requiring the binding of the
336 two CA radicals for the rearrangement of the active residues and DIR activation. In
337 comparison, free AtDIR6 is probably in a precatalytic conformation with the active site
338 residues lined up for catalysis (Gasper *et al.*, 2016).

339 Both AtDIR6 and PsDRR206 contain omega loop (part of the β2-β3 loop,
340 Supplementary Fig. S2) that comprises a cluster of highly conserved aa residues (His39,

Thr84, and Ser91 in PsDRR206). The omega loop is tightly adjacent to the active center and might have a role in proper substrate positioning (Kim *et al.*, 2015).

Based on the AtDIR6 structure and extensive mutational analysis, an updated model describing how AtDIR6 facilitates lignan synthesis was proposed (Gasper *et al.*, 2016). Each subunit of the DIR homotrimer binds two CA radicals. The binding to pockets A and B accommodates both radicals in their extended, planar *all-trans* conformation, in which they are precisely positioned to enable 8-8' coupling at the *re-re* face. This determines the regioselectivity of 8-8' pinoresinol coupling over other coupling options. Also the enantioselectivity of DIRs seems to be a direct consequence of the precise substrate positioning allowing regio-selective 8-8' coupling. Gasper *et* al. (2016) suggested that if coupling occurs at *re-re* face, the *S,S*-configured *bis*-quinone methide (bisQM) intermediate (see also below) is formed as the precursor of (-)-pinoresinol. However, *si-si face* coupling yields the *R, R*-bisQM and, consequently, (+)-pinoresinol. Interestingly, larger portions of the DIR protein must be changed (not simply individual amino acids) to change its stereoselectivity (Gasper *et al.*, 2016; Kim *et al* ., 2012).

Originally it was hypothesized that DIRs actually lack enzymatic activity and are in reality responsible mostly for the correct positioning of linked CA radicals. Surprisingly, the solved structure suggests a direct involvement of AtDIR6 in catalyzing the cyclization of bisQM intermediate that is electron-deficient at C7 and thus susceptible to nucleophilic attack (Freudenberg, 1959; Gasper *et al.*, 2016; Ralph *et al.*, 2009). In the process of pinoresinol formation, the terminal OH groups of the propionyl side chains can act as nucleophiles. 1,6- addition to the *p*-quinone methides allows the cyclization of the furan rings and re-aromatization of the cyclohexadienones. In the new mode of action proposed by Gasper et al (2016), AtDIR6 contributes to the catalysis of pinoresinol formation by donating a proton to the hexadienone ring carbonyl of the bisQM (Fig. 4). The authors suggested that this reaction could be mediated via hydrogen bond formation or acid catalysis. Conservative amino acid Asp137, Arg144 and Thr166 in pocket A, and Asp49, Tyr104 and Tyr106 in pocket B are critical for this DIR function. Based on the physico-chemical properties and positioning of those conserved amino acid residues, most probable candidates are Asp137, Arg144 and Asp49.

The role of DIR proteins in stress responses

DIRs in a response to biotic stress

The involvement of lignin, and phenylpropanoid biosynthesis in general, in plant defense against pathogens is well known (Caño-Delgado *et al.*, 2003; Miedes *et al.*, 2014). Several publications have described increased concentrations of lignin after pathogen infection (Mohr and Cahill, 2007; Quentin *et al.*, 2009), or highlighted the importance of soluble phenylpropanoids in Arabidopsis pathogen defense (König *et al.*, 2014). An important contribution of lignin to defense against pathogens was demonstrated in *Medicago sativa,* where selective down-regulation of lignin biosynthesis resulted in the constitutive expression of defense response genes. It was suggested that lignin deficiency triggers the defense response via the enhanced release of bioactive cell wall fragments as a result of impaired cell wall integrity (Gallego-Giraldo *et al.*, 2011) (see also below). Conversely, changes in flax resistance to pathogens putatively correlate better with the antioxidant potential of seed lignans, whereas the structural barriers provided by lignin and cellulose appear to be not as important (Zeitoun *et al.*, 2014).

Direct evidence connecting AtDIRs with pathogen defense in Arabidopsis is still missing. Nevertheless, numerous examples clearly demonstrate the involvement of DIRs in the pathogen response in various species. For example, the infection of the moss *Physcomitrella patens* with *Colletotrichum gleosporioides* results in cell wall modifications including increased incorporation of phenolic compounds, which was correlated with the induction of a DIR-like protein encoding gene (Reboledo *et al.*, 2015). Recently, it has been reported that *P. patens* also exhibits cell wall reinforcements upon treatment with *Pectobacterium carotovorum* derived elicitors. Expression analysis showed that four genes encoding putative DIR-like proteins (*DIR-1, DIR-2, DIR-3* and *DIR-4*) were induced during the treatment with different dynamics (Alvarez *et al.*, 2016). *DIR-4* seems to be induced also by *B. cinerea* and *C. aloeosporioides* infection (Ponce De Leon *et al.*, 2012; Reboledo *et al.*, 2015). In Sitka spruce (*Picea sitchensis*), transcript profiling and micro-array based gene expression analysis suggested that several DIR genes were induced upon attack by stem-boring weevils or after mechanical

400 wounding. Similarly, *BrDIR12* and several *DIR-like* genes were differentially expressed
401 in cabbage (*Brassica oleracea*) infected by *Fusarium oxysporum* (Arasan *et al.*, 2013).

402 There is also evidence that DIRs and DIR-like proteins are spatially targeted during the
403 response to pathogen infection (Ma, 2014). The rice mannose-binding jacalin-related
404 lectin gene (OsJAC1) was shown to consist of a DIR domain and an N-terminal jacalin-
405 related lectin domain (JRL), which is predicted to bind mannose-containing
406 oligosaccharides prevalent in fungi. The overexpression of *OsJAC1* in rice resulted in
407 broad-spectrum resistance against major rice pathogens and seems to require the re-
408 localization of OsJAC1 to infection sites. Similarly, overexpression of *OsJAC1* or the
409 wheat orthologue (*TaJA1*) in barley results in enhanced resistance to powdery mildew
410 fungus. Interestingly, both the DIR and JRL domains are required to generate
411 resistance, as indicated by transient expression experiments (Weidenbach *et al.*, 2016).

412 The large number of different *DIRs* expressed during wounding and pathogen infection
413 supports the idea that DIR proteins participate in pathogen defense. This can be by
414 generating defense compounds and/or participate in dynamic reorganization of the cell
415 wall since both require radical coupling during the formation of pinoresinol and other
416 lignans as well as their stereoisomers. This hypothesis is further supported by the large
417 number of antibacterial activities observed for various lignan stereoisomers (Akiyama *et*
418 *al.*, 2007a; Akiyama *et al.*, 2009). However, plant pathogen defense via lignan
419 production is not delimited to antibacterial effects. Lignans isolated from *Piperaceae*
420 have also toxic effects on phytophagous insects (Bernard *et al.*, 1995). Other lignans
421 were shown to act as antifungal agents (Akiyama *et al.*, 2007b; Carpinella *et al.*, 2005)
422 and contribute to wound responses (Harju *et al.*, 2009). Interestingly the highest
423 transcriptional change observed upon fungal infection in *Vitis vinifera* was the
424 upregulation of a putative *DIR* gene (180 fold) (Borges *et al.*, 2013). The authors
425 hypothesized that low *DIR* expression during the early stages of fungal infection favors
426 lignin biosynthesis via peroxidase-mediated production of all three possible isomers
427 from coniferyl alcohol: pinoresinol, dehydrodiconiferyl alcohol and guaiacylglycerol 8-*O*-
428 4'-coniferyl ether, since the latter two are precursors of lignin biosynthesis. In

comparison, the up-regulation of *DIR* genes during later stages of infection increases antifungal activities through activation of lignan biosynthesis (Borges et al 2013).

DIRs during the response to abiotic stress

DIRs and peroxidases have frequently been implicated in modulation of lignification levels upon exposure to abiotic stress. The expression of several of the *DIR-like* genes was responsive to water, ABA and cold stress. More importantly, in the case of water stress, the expression of the most responsive *DIR* genes could be correlated with increased lignification (Arasan *et al.*, 2013).

In soybean roots, Mn toxicity is known to enhance peroxidase activity and wall lignification (Morita *et al.*, 2006). A recent proteomic study demonstrated that Mn toxicity also induced elevated accumulation of H_2O_2 in roots, which coincided with up-regulation of PEROXIDASE5- and DIR2-like protein levels, whereas levels of another DIR protein were reduced (Chen *et al.*, 2016). Similarly, in *Medicago sativa* application of cold stress resulted in the transcriptional down-regulation of two peroxidases and a *DIR* gene, while another *DIR* gene was upregulated by heat stress (Behr *et al.*, 2015). A *DIR* gene from *Boea hygrometrica* (*BhDIR1)* was implicated in the response to water and temperature stresses (Wu *et al.*, 2009). A *Saccharum* spp. dirigent gene ***(****ScDIR)* exhibiting stem-specific expression has been reported to respond to drought, salt and oxidative stresses (Jin-Long *et al.*, 2012). Additionally, CaiQiu *et al.*, (2010) reported enhanced expression levels of *TaDIR* in *Tamarix androssowii* after exposure to salinity–alkalinity stress.

The above examples imply that the mobilization of specific peroxidases and DIRs in response to particular stress conditions is possibly ensuring specificity of the plant response. The DIR-related molecular mechanisms underpinning the ability of plants to cope with the given abiotic stress type, however remain to be identified.

DIRs and plant cell wall integrity maintenance

As we discussed above, the spatial control of lignin deposition is important in defense responses to biotic or abiotic stress. Several groups have implicated DIR proteins in the response to pathogens causing physical damage to the cell wall during infection and abiotic stress (osmotic and drought) in different plant species (Ralph *et al.*, 2006; Jin-long *et al.*, 2012; Arasan *et al.*, 2013). Bearing in mind the current research on CS formation, it is conceivable that DIR proteins could mediate the spatial control of lignin deposition during development and stress responses. This would make them an essential element controlling cell wall modification/reinforcement during cell wall integrity maintenance.

During exposure to both biotic and abiotic stress, the conformation/structure of the differing plant cell walls needs to be adaptively altered to maintain stress-compromised functional characteristics essential for development and defense. The maintenance of wall integrity can often also neutralize the effects of targeted genetic manipulation through adaptive (compensatory) changes in other processes impacting cell wall composition and structure (Doblin *et al.*, 2014). It also implies that a mechanism exists to monitor cell wall status and to initiate specific compensatory responses. It is reasonable to assume that such responses probably require a tight spatial control at the cellular and subcellular level to target compromised wall matrix sites of specific architecture to ensure that novel cell wall components required for the restoration of function are integrated correctly.

A dedicated cell wall integrity maintenance mechanism has been well documented in *Sacharomyces cerevisiae* (Levin, 2011). It involves osmo, mechano and cell wall damage perception, signal translation via Rho GEFs and MAP kinase cascades as well as adaptive changes in cytoskeleton organization and cell wall metabolic processes. Interestingly it has been shown that certain *A. thaliana* genes can rescue yeast strains deficient in components of the yeast maintenance mechanism (Nakagawa *et al.*, 2007; Reiser *et al.*, 2003). Several groups have shown in parallel that in Arabidopsis a similar mechanism exists (Caño-Delgado *et al.*, 2003; Ellis *et al.*, 2002; Hematy *et al.*, 2007; Wolf and Hofte, 2014). It involves at least a plasma-membrane-localized receptor-like kinase (THESEUS1) as well as a putative stretch activated calcium channel (MID1-

COMPLEMENTING ACTIVITY1), which seem to perceive plant cell wall integrity impairment. Homologs of these genes have been found in a large number of both mono- and dicotyledonous plant species including crops like rice and strawberry as well as *P. patens* (N. Gigli-Bisceglia, J. Antsiferova, pers. communication), suggesting that the mechanism they are involved in exists throughout the plant kingdom (Kurusu *et al.*, 2012; Nguyen *et al.*, 2014; Zhang *et al.*, 2016).

Jasmonic acid and reactive oxygen species are required for signal translation in cell wall integrity maintenance while cell wall modifications can involve the production of pectic polysaccharides as well as targeted deposition of lignin in cell types that normally do not exhibit lignin deposition (Caño-Delgado *et al.*, 2003; Denness *et al.*, 2011)). To meet the specific functional requirements of the compensatory process, both the individual biosynthetic processes and the locations where they take place have to be tightly controlled. While the need for this tight control is obvious, our knowledge of the molecular machinery responsible for it is still very limited. However, as mentioned in the above text, DIR proteins might represent one of the potent effectors acting downstream of the cell wall integrity signaling cascade. That, however, remain to be demonstrated.

Bioinformatic analysis of Arabidopsis DIR family

Phylogenetic relationships of AtDIR proteins

In order to identify potential functional relationships in the AtDIR family, a phylogenetic tree was created based on the amino acid alignment of DIR domains from individual AtDIR sequences. As AtDIR17 contains just a partial DIR domain, it was excluded from the phylogenetic analysis.

As shown in the Fig. 5, AtDIRs might be divided in several distinct subclades. Some of them partially correspond to DIR (DIR-a) and DIR-like (DIRb - DIR-e) subfamilies previously distinguished by Ralph et al. (2006) by comparing 72 DIRs and DIR-like proteins from several species including AtDIRs. In the tree based exclusively on DIR domains of all 26 members of the AtDIR family, we were able to reliably identify three

514 main subclades. The subclade consisting of AtDIR5, 6, 12, 13, 14 corresponds to the
515 Ralphs DIR-a subfamily and contains both of the known (-)-pinoresinol forming
516 Arabidopsis AtDIRs, AtDIR5 and AtDIR6. Further reliably distinguishable subclades
517 include AtDIR1, 2, 11, 21, which seems to correspond to DIR-like subfamily DIR-d, in
518 Ralph et al (2006) represented by AtDIR11 and AtDIR2. Finally, the last well-specified
519 subclade consisting of AtDIR9, 10, 16, 18, 24, and 25 largely overlaps with DIR-like
520 subfamily DIR-e (Ralph *et al.*, 2006). Further classification of AtDIR proteins would be
521 rather approximate due to low bootstrap support of deep nodes (Fig. 5).

522 AtDIR 10, 16, 18, 25 together with isoleucyl-tRNA synthetase (OVA2), a rather atypical
523 member of the AtDIR family, comprise a group of multidomain DIRs. *OVA2* genes
524 represent a special case among *AtDIRs*, which is also reflected by its clustering outside
525 the DIR-e subclade. One of three possible *OVA2* isoforms extends to the upstream
526 transcription start site, thereby including the coding sequence of the *AtDIR3* gene. The
527 stop codon of *AtDIR3* is spliced out, resulting in a predicted OVA2 (At5g49030.3) protein
528 with an N-terminal DIR3 domain. Interestingly, AtDIR10, 16, 18, 25 contain both a
529 complete DIR domain and a partial N-terminal DIR domain.

530 AtDIR25 also contains a part of a JACALIN protein (represented in green in the
531 diagram), reminiscent of the recently described OsJAC1 protein of rice (Weidenbach *et*
532 *al.*, 2016). However, it is questionable if AtDIR25 and OsJAC1 have similar functions
533 since the JACALIN domain of AtDIR25 is not complete. Based on the clustering of all
534 but one multidomain AtDIRs into the DIR-e subclade suggests that the presence of
535 additional domains in the AtDIR family is “encoded” in the sequence of the DIR domain.
536 The conservancy of this mechanism is supported by the inclusion of DIR domain
537 sequences from DIR proteins with extended N-terminals from species other than
538 Arabidopsis (*Gossypium hirsutum* (Gh), *Theobroma cacao* (Tc), *Corchorus capsularis*
539 (Cc) and *Nicotiana sylvestris* (Ns)). All these sequences clustered together in the DIR-e
540 (multidomain) subclade (Supplementary Fig. S5).

541 One defining characteristic of this sub-clade in Arabidopsis could be the lack of the N-
542 terminal portion of the aligned DIR domains (Supplementary Fig. S3). Whether there is
543 any functional link between an absence of the N-terminal amino acid stretches

(corresponding to a part of the β1 sheet) and functional properties of the members of the DIR-e subfamily remains to be investigated. Interestingly, another common feature of the subfamily is the absence of one of the key functional residues, corresponding to Arg144 in AtDIR6. The residue is located at the N-terminus of β6 and lines the pocket A, where it is supposed to act as one of the potential proton donors facilitating cyclization of bisQM by AtDIR6 (Figs. 2, 4). In all AtDIRs, the residue corresponding to Arg144 in AtDIR10/ESB1 is conserved, but it is replaced by Ser in all members of the multidomain DIR-e subfamily (Supplementary Fig. S3).

Interestingly, AtDIR22 is the only AtDIR protein predicted not to contain a signal peptide for extra-cellular secretion, suggesting that this protein could be involved in the intracellular monolignol coupling and/or other role. Remarkably, a metabolic pathway involving intracellular coupling of monolignol radical has been described recently (Dima *et al.*, 2015).

A comparison of the chromosome map (Supplementary Fig. S4) with the phylogenetic tree shows that some *AtDIRs* are adjacent to each other on their respective chromosomes and also very similar at the sequence level (e.g. *AtDIR12, 13, 14; AtDIR2*, *AtDIR3, OVA2*). However, some of the *AtDIRs* locating to the same region on the chromosome are rather distantly related based on the phylogenetic tree (e.g. *AtDIR4* and *AtDIR23; AtDIR7* and *AtDIR22*). Vice versa, some of the almost identical *AtDIRs* are positioned far from each other on the chromosome map (e.g. *AtDIR16* and *AtDIR18*; *AtDIR9* and *AtDIR24*; *AtDIR10* and *AtDIR25*). This is implying the existence of several independent gene duplication events during the *AtDIR* family evolution and should be studied in more detail.

Identification of (+/-)-pinoresinol forming AtDIRs

In order to identify AtDIRs potentially involved in stereoselective lignan biosynthesis, the DIR domains from DIR proteins known to mediate stereoselective pinoresinol formation from other species were added to the phylogenetic analysis. These include (+)-

pinoresinol forming FiDIR1 from *Forsythia intermedia* (Davin *et al.*, 1997), TpDIR2 and TpDIR7 from *Thuja plicata* (Kim *et al.*, 2002), ScDIR1 from *Schizandra chinensis* (Kim *et al.*, 2012), PsDRr206 from *Pisum sativum* (Seneviratne *et al.*, 2015) and LuDIR1 from *Linum usitatissimum* (Dalisay *et al.*, 2015). The (-)-pinoresinol forming DIR proteins LuDIR5 and LuDIR6 from *Linum usitatissimum* (Dalisay *et al.*, 2015) were also included.

The resulting phylogenetic tree (Supplementary Fig. S5) clearly splits into four major subclades. Subclade I comprises the (+)-pinoresinol forming DIR proteins FiDIR1, TpDIR2 and TpDIR7, ScDIR1 and PsDRr206. The subclade II comprises the (-)-pinoresinol forming DIR proteins LuDIR5, LuDIR6, AtDIR5 and AtDIR6, and additionally AtDIRs 12, 13 and 14, whose role in stereoselective pinoresinol formation is uncertain (see further in the text). Subclades III consists of the multidomian DIRs, whereas the remaining sequences belong to subclade IV.

Several amino acids are differentially conserved in (+) and (-)-pinoresinol forming DIR proteins from different species, indicating a functional importance of those amino acid residues for stereoselective pinoresinol formation (Gasper *et al.*, 2016; Kim *et al.*, 2015). Using this information, we found that only AtDIR5 and AtDIR6 contain all residues conserved in (-)-pinoresinol forming DIRs (Supplementary Fig. S3). Remarkably, none of the conserved amino acid residues associated with (+)-pinoresinol formation activity is present in any AtDIR protein.

In conclusion, these data indicate that AtDIR5 and AtDIR6 might be the only AtDIRs involved in (-)-pinoresinol formation, while the (+)-pinoresinol-forming DIRs seem to be absent in Arabidopsis. These findings are in agreement with published experimental evidence, showing that both AtDIR5 and AtDIR6 were able to produce (-)-pinoresinol, both in vitro and in vivo. In contrast, AtDIR13, in spite being a member of the same subclade (DIR-a subfamily, Fig. 5), wasn't able to catalyze (-)-pinoresinol formation (Kim *et al.*, 2012), further highlighting the functional importance of conserved residues (Gasper *et al.*, 2016; Kim *et al.*, 2012).

Transcriptional regulation of AtDIRs

In order to analyze the transcriptional regulation of *AtDIR* genes in different tissues and in response to hormones and stress conditions, Genevestigator (Hruz *et al.*, 2008) and eFP browser (Winter *et al.*, 2007) were used to analyze the available transcriptome microarray and RNAseq datasets (summarized in Supplementary Table 1 and Fig. 6). Most of *AtDIR* genes reveal highest expression levels in the roots (*AtDIR1, 2, 5, 6, 7, 9, 10, 13, 14, 16, 18, 19, 20, 23, 24, 25*), whereas only a few show highest expression levels in other organs (*AtDIR3* and *AtDIR12* in the seed, *AtDIR7* in the hypocotyl, *AtDIR20* in the inflorescence and flower, *AtDIR8* and *AtDIR20* in flower or pollen and *AtDIR22* in anther and stamen). *AtDIR4, AtDIR11* and *AtDIR15* exhibited rather low expression levels in all organs. A more detailed comparison of the *AtDIR* expression levels in different root cell types is shown in Supplementary Fig. S6.

To investigate the transcriptional responses of *AtDIRs* to hormone treatments and different types of stress, microarray-based expression profiling datasets from AtGenExpress were analyzed using again Genevestigator (fold-change >2, p-value <0.05). Certain *AtDIRs* responded to abscisic acid (ABA; *AtDIR19, 23*), methyl jasmonate (*AtDIR5, 13*) and t-zeatin (*AtDIR13*). Different *AtDIRs* responded to biotic stress-related treatments (*AtDIR6, 7, 11, 20, 21*), i.e. following treatments with the organisms *Pseudomonas syringae* (*AtDIR6, 7, 11, 20, 21*), *Pseudomonas infestans* (*AtDIR20, 21*), *Golovinomyces orontii* treatment (*AtDIR20*) or the elicitors GST-necrosis-inducing phytophthora protein 1 (*AtDIR 7, 11, 20*), flagellin fragment 22 (*AtDIR 7*) and hairpin (hrpZ) *(AtDIR 11, 20)*. In parallel, the data showed that *AtDIRs* also change their expression levels in a response to osmotic stress (*AtDIR1, 2, 5, 21*), salt stress (*AtDIR5, 7, 9*), drought (*AtDIR5*), wounding (*AtDIR5, 7*), oxidative (*AtDIR5*) and heat stress (*AtDIRs7, 19*).

To expand our knowledge regarding the molecular mechanisms responsible for transcriptional regulation of individual *AtDIR* family members, we performed an *in silico* TF binding assay. 345 DNA-binding motif matrices of TFs from the TRANSFAC database were used for a calculation of binding affinities to each of the 25 *AtDIR* gene promoters (see Materials and Methods section and (Manke *et al.*, 2008; Roider *et al.*,

2007)). Out of 345 motifs analyzed, 281 (81%) were found to have a significant affinity (p-value < 0.05) to at least one of the *AtDIR* promoters. The results of this analysis (clustered according to the TF binding profiles of individual *AtDIRs*) are summarized in the heat map provided in Supplementary Fig. S7. All detected significant affinities are provided in Supplementary Table 2.

Since the composition of the TRANSFAC database used for this analysis might be biased due to overrepresentation of TFs associated with certain biological functions, or TF families comprising TFs with very similar DNA-binding motif matrices, the results have to be considered critically. This is exemplified by ETHYLENE RESPONSE FACTORS (ERFs), which show a high degree of conservation with respect to their recognized binding sites. This leads to a possibly artificially enhanced number of TFs recognizing being implicated with certain *AtDIR* promoters carrying particular binding site (data not shown). Bearing in mind these considerations, a comparative analysis of transcriptome data (Supplementary table 1) with predicted TF/promoter interactions (Supplementary Fig. S7, Supplementary Table 2) yields the following observations.

AtDIR3 is expressed in the seed

The expression of the *AtDIR3* gene seems to be seed specific and the motif matrices of the TFs ABI3 (At3g24650) and FUS3 (At3g26790) have significant affinities to the *AtDIR3* promoter. Interestingly these TFs have been shown independently to regulate seed specific gene expression (Mönke *et al.*, 2004).

AtDIR5 responds to stress and methyl jasmonate

AtDIR5 gene expression might respond to methyl jasmonate and a wide range of stresses such as drought, salt stress and wounding. Motif matrices of the TFs MYC2, MYC3 and MYC4 exhibit significant affinities to the *AtDIR5* promoter. These TFs have been demonstrated previously to mediate methyl jasmonate responses and are induced

by drought, salt stress and wounding (reviewed in (Kazan and Manners, 2013)).

AtDIR6 associates with secondary cell wall formation

AtDIR6 has been predicted to be co-expressed with *PrR1,* the pinoresinol reductase encoding gene (Kim *et al.*, 2012). Previous work showed that *PrR1* expression is under transcriptional control of NST3 and MYB46, the master regulators of secondary cell wall formation (Zhao *et al.*, 2015). We found that the motif matrices of MYB46 and additionally MYB52 exhibit significant affinities to the *AtDIR6* promoter and both TFs are involved in regulation of secondary cell wall formation (Zhong *et al.*, 2008).

AtDIR13 responds to cytokinins

Expression of the *AtDIR13* gene changes in response to zeatin treatment (Taniguchi *et al.*, 2007). Interestingly, the motif matrices of ARR10 (our analysis) and ARR1 (Taniguchi *et al.*, 2007), type-B ARRs and members of the cytokinin signaling pathway (Hwang *et al.*, 2012) exhibit significant affinities to the *AtDIR13* promoter.

Concluding remarks and future outlines

Originally identified as responsible for the regio- and stereo-selectivity of phenoxy radical coupling reactions, DIR proteins could be important regulators of plant development and plant stress response. Nowadays, AtDIR proteins family is still uncharacterized. Clearly, AtDIR5 and AtDIR6 are involved in the stereoselective radical-radical coupling leading to (-) pinoresinol. Unfortunately, the role of these genes in the Arabidopsis development is not known. In contrary to that, the role of AtDIR10/ESB1 in CS formation is indisputable. The inability of AtDIR10/ESB1 to mediate regio- and stereo-selectivity of coniferyl alcohol coupling mediated by its close homologues AtDIR5 and AtDIR6 suggests that very small changes in amino acid sequence may cause a loss

683 of stereoselectivity. Alternatively, AtDIR10/ESB1 could have ability to bind phenoxy
684 monolignol radicals, which might be necessary for the role of AtDIR10/EBS1 in the
685 initiation of lignification. The role of AtDIR10/ESB1 in targeting CASPs for precise
686 positioning of the entire CS forming machinery implies its ability to recognize specific
687 plasma membrane domains predetermined to guide lignin impregnations of adjacent cell
688 wall. Whether the additional N-terminal sequence of AtDIR10/ESB1 contains sequence
689 motifs responsible for this and what the nature of the molecular targets of the interaction
690 remains to be determined. Our preliminary bioinformatics analysis implies that the other
691 members of the Arabidopsis DIR-e subclade reveal i) similarity at the level of DIR
692 domain (missing the β1 sheet), ii) replacement of the conserved Arg by Ser at the N-
693 terminus of β6 and iii) most of them contain an additional N-terminally located,
694 incomplete DIR domain. Whether these protein structural characteristics are essential
695 for the role in the lignification is still to be determined.

696 Based on the published evidence and our *in silico* TF binding assay it seems that the
697 (GO) term best characterizing AtDIRs is the stress response. However, TF binding
698 based on TRANSFAC database (with overrepresentation of TF related with the
699 development) is only a basic bioinformatics approach and further analyses should be
700 done. Nonetheless, the ability of lignans to mediate pathogen resistance and stress-
701 induced expression of numerous *DIRs* in various plant species seems to be in line with
702 that hypothesis. The fact that several of the *AtDIRs* are proposed to be responsive to
703 plant hormones, is in agreement with an important role for hormonal regulations in the
704 mediation of adaptation to stress. The evidence that *DIRs* could be a target of CW
705 integrity signaling (largely under hormonal control) also fits the concept. However,
706 experimental approaches have to be put into action to confirm these speculations and
707 deepen our knowledge on the possible role of the role hormonal regulations of DIRs
708 expression.

709 The high homology in protein sequence of individual AtDIRs and the similar expression
710 patterns implies possible functional redundancy in frame of the family, potentially
711 hampering the functional characterization. Thus, taking into advantage of using the up-
712 to-date molecular biology approaches like amiRNA-based knock-down of multiple

AtDIRs might represent a suitable strategy.

We hope we have convinced the readers that *AtDIRs* constitute an essential and fascinating protein family, enabling plants to adapt to dynamically changing environmental conditions. Based on the available evidence it seems that the primary role of DIRs occurs at the level of control over the CW metabolism and/or production of antibacterial compounds. That determines DIRs as a potent target in a number of biotechnological applications as powerful tools in improving plant stress resistance or production of pharmaceutically interesting compounds.

Supplementary data

Supplementary Figure S1. Monolignol production via phenylpropanoid pathway.

Supplementary Figure S2. Comparison of AtDIR6 and PsDRR206.

Supplementary Figure S3. Multiple alignment of amino acid residues of AtDIR proteins.

Supplementary Figure S4. Chromosome map showing the positions of *AtDIR* genes.

Supplementary Figure S5. **Identifying stereospecific pinoresinol-forming *AtDIR*s.**

Supplementary Figure S6. Root cell type specific expression of *AtDIRs*.

Supplementary Figure S7. Heatmap of p-values of transcription factor binding affinities.

Supplementary Table S1: Transcriptional regulation and proposed roles of *AtDIR* genes from *Arabidopsis thaliana*.

Supplementary Table S2: TRANSFAC analysis of DNA binding motif affinities.

Acknowledgements

The work was supported by CEITEC 2020 (LQ1601), the Czech Science Foundation (13-25280S) and the Czech-Norwegian Research Programme (7F14155).

Methods

TRANSFAC (Transcription factor database)

To score TF binding, a physical affinity-based model was used (Manke *et al.*, 2008; Roider *et al.*, 2007) and position weight matrices describing 345 *Arabidopsis thaliana* TF binding preferences from TRANSFAC release 2015.1 (Matys *et al.*, 2003). 1500 bp upstream sequences of the 25 *DIR* genes were downloaded from TRANSFAC.

For each binding matrix, the affinity of the matrix for each sequence was calculated, and then transformed into *P*-values as described before (Manke *et al.*, 2008) These *P*-values represent the probability that the observed binding affinity is greater than would be expected from a random sequence from an *Arabidopsis*-promoter-based background model.

DIR domain alignment and phylogenetic tree analysis

DIR domains were extracted from the cdd database (https://www.ncbi.nlm.nih.gov/structure/cdd/cdd.shtml). The alignment of amino acid sequences was conducted in MEGA5 with the MUSCLE algorithm

The evolutionary relationships were inferred using the Maximum Parsimony method. The percentage of replicate trees in which the associated sequences clustered together in the bootstrap test (% of 500 replicates) are shown next to the branches (Felsenstein, 1985). The MP tree was created using the Tree-Bisection-Regrafting (TBR) algorithm (Nei and Kumar, 2000). Phylogenetic analyses were conducted in MEGA5 (Tamura *et*

761 *al.*, 2011).

762 Alignment of amino acid sequences was conducted in MEGA5 with the MUSCLE
763 algorithm and visualized with UGENE toolkit (Okonechnikov *et al.*, 2012).

764 **Protein structure**

765 All the structural figures shown were generated with PyMOL 4, using the PDB codes
766 4REV for PsDRR206 and 5LAL for AtDIR6.

767

References

Akiyama K, Maruyama M, Yamauchi S, Nakashima Y, Nakato T, Tago R, Sugahara T, Kishida T, Koba Y. 2007a. Antimicrobiological activity of lignan: effect of benzylic oxygen and stereochemistry of 2,3-dibenzyl-4-butanolide and 3,4-dibenzyltetrahydrofuran lignans on activity. *Biosci Biotechnol Biochem* **71**, 1745-1751.

Akiyama K, Yamauchi S, Maruyama M, Sugahara T, Kishida T, Koba Y. 2009. Antimicrobial activity of stereoisomers of morinols a and B, tetrahydropyran sesquineolignans. *Biosci Biotechnol Biochem* **73**, 129-133.

Akiyama K, Yamauchi S, Nakato T, Maruyama M, Sugahara T, Kishida T. 2007b. Antifungal activity of tetra-substituted tetrahydrofuran lignan, (-)-virgatusin, and its structure-activity relationship. *Biosci Biotechnol Biochem* **71**, 1028-1035.

Alvarez A, Montesano M, Schmelz E, Ponce de Leon I. 2016. Activation of Shikimate, Phenylpropanoid, Oxylipins, and Auxin Pathways in Pectobacterium carotovorum Elicitors-Treated Moss. *Frontiers in plant science* **7**, 328.

Arasan SKT, Park JI, Ahmed NU, Jung HJ, Hur Y, Kang KK, Lim YP, Nou IS. 2013. Characterization and expression analysis of dirigent family genes related to stresses in Brassica. *Plant Physiology and Biochemistry* **67**, 144-153.

Attoumbre J, Bienaime C, Dubois F, Fliniaux MA, Chabbert B, Baltora-Rosset S. 2010. Development of antibodies against secoisolariciresinol--application to the immunolocalization of lignans in Linum usitatissimum seeds. *Phytochemistry* **71**, 1979-1987.

Barros J, Serk H, Granlund I, Pesquet E. 2015. The cell biology of lignification in higher plants. *Ann Bot* **115**, 1053-1074.

Behr M, Legay S, Hausman J-F, Guerriero G. 2015. Analysis of Cell Wall-Related Genes in Organs of Medicago sativa L. under Different Abiotic Stresses. *International journal of molecular sciences* **16**, 16104-16124.

Benedict CR, Liu J, Stipanovic RD. 2006. The peroxidative coupling of hemigossypol to (+)- and (-)-gossypol in cottonseed extracts. *Phytochemistry* **67**, 356-361.

Bernard CB, Krishanmurty HG, Chauret D, Durst T, Philogene BJ, Sanchez-Vindas P, Hasbun C, Poveda L, San Roman L, Arnason JT. 1995. Insecticidal defenses of Piperaceae from the neotropics. *Journal of chemical ecology* **21**, 801-814.

Bonawitz ND, Kim JI, Tobimatsu Y, Ciesielski PN, Anderson NA, Ximenes E, Maeda J, Ralph J, Donohoe BS, Ladisch M, Chapple C. 2014. Disruption of Mediator rescues the stunted growth of a lignin-deficient Arabidopsis mutant. *Nature* **509**, 376-380.

Bonello P, Storer AJ, Gordon TR, Wood DL, Heller W. 2003. Systemic effects of Heterobasidion annosum on ferulic acid glucoside and lignin of presymptomatic ponderosa pine phloem, and potential effects on bark-beetle-associated fungi. *Journal of chemical ecology* **29**, 1167-1182.

Borges AF, Ferreira RB, Monteiro S. 2013. Transcriptomic changes following the compatible interaction Vitis vinifera-Erysiphe necator. Paving the way towards an enantioselective role in plant defence modulation. *Plant Physiol Biochem* **68**, 71-80.

Bottcher C, von Roepenack-Lahaye E, Schmidt J, Schmotz C, Neumann S, Scheel D, Clemens S. 2008. Metabolome analysis of biosynthetic mutants reveals a diversity of metabolic changes and allows identification of a large number of new compounds in Arabidopsis. *Plant Physiol* **147**, 2107-2120.

Buchanan BB, Gruissem W, Jones RL. 2000. *Biochemistry & molecular biology of plants*. Rockville, Md.: American Society of Plant Physiologists.

Burlat V, Kwon M, Davin LB, Lewis NG. 2001. Dirigent proteins and dirigent sites in lignifying tissues. *Phytochemistry* **57**, 883-897.

CaiQiu G, GuiFeng L, YuCheng W, Jing J, ChuanPing Y. 2010. Cloning and analysis of dirigent-like protein in gene from Tamarix androssowii. *Bulletin of Botanical Research* **30**, 81-86.

Cano-Delgado A, Penfield S, Smith C, Catley M, Bevan M. 2003. Reduced cellulose synthesis invokes lignification and defense responses in Arabidopsis thaliana. *The Plant journal : for cell and molecular biology* **34**, 351-362.

Carpinella MC, Ferrayoli CG, Palacios SM. 2005. Antifungal synergistic effect of scopoletin, a hydroxycoumarin isolated from Melia azedarach L. fruits. *Journal of agricultural and food chemistry* **53**, 2922-2927.

Dalisay DS, Kim KW, Lee C, Yang H, Rübel O, Bowen BP, Davin LB, Lewis NG. 2015. Dirigent Protein-Mediated Lignan and Cyanogenic Glucoside Formation in Flax Seed: Integrated Omics and MALDI Mass Spectrometry Imaging. *Journal of natural products* **78**, 1231-1242.

Davin LB, Jourdes M, Patten AM, Kim KW, Vassao DG, Lewis NG. 2008. Dissection of lignin macromolecular configuration and assembly: comparison to related biochemical processes in allyl/propenyl phenol and lignan biosynthesis. *Natural product reports* **25**, 1015-1090.

Davin LB, Lewis NG. 2000. Dirigent proteins and dirigent sites explain the mystery of specificity of radical precursor coupling in lignan and lignin biosynthesis. *Plant Physiology* **123**, 453-462.

Davin LB, Lewis NG. 2005a. Dirigent phenoxy radical coupling: advances and challenges. *Current Opinion in Biotechnology* **16**, 398-406.

Davin LB, Lewis NG. 2005b. Lignin primary structures and dirigent sites. *Current Opinion in Biotechnology* **16**, 407-415.

Davin LB, Wang HB, Crowell AL, Bedgar DL, Martin DM, Sarkanen S, Lewis NG. 1997. Stereoselective bimolecular phenoxy radical coupling by an auxiliary (dirigent) protein without an active center. *Science* **275**, 362-366.

Dima O, Morreel K, Vanholme B, Kim H, Ralph J, Boerjan W. 2015. Small glycosylated lignin oligomers are stored in Arabidopsis leaf vacuoles. *Plant Cell* **27**, 695-710.

Dinkova-Kostova AT, Gang DR, Davin LB, Bedgar DL, Chu A, Lewis NG. 1996. (+)-Pinoresinol/(+)-lariciresinol reductase from Forsythia intermedia. Protein purification, cDNA cloning, heterologous expression and comparison to isoflavone reductase. *J Biol Chem* **271**, 29473-29482.

Doblas VG, Smakowska-Luzan E, Fujita S, Alassimone J, Barberon M, Madalinski M, Belkhadir Y, Geldner N. 2017. Root diffusion barrier control by a vasculature-derived peptide binding to the SGN3 receptor. *Science* **355**, 280-284.
Doblin MS, Johnson KL, Humphries J, Newbigin EJ, Bacic A. 2014. Are designer plant cell walls a realistic aspiration or will the plasticity of the plant's metabolism win out? *Curr Opin Biotechnol* **26**, 108-114.
Effenberger I, Harport M, Pfannstiel J, Klaiber I, Schaller A. 2016. Expression in Pichia pastoris and characterization of two novel dirigent proteins for atropselective formation of gossypol. *Appl Microbiol Biotechnol*.
Effenberger I, Zhang B, Li L, Wang Q, Liu Y, Klaiber I, Pfannstiel J, Wang Q, Schaller A. 2015. Dirigent Proteins from Cotton (Gossypium sp.) for the Atropselective Synthesis of Gossypol. *Angewandte Chemie International Edition* **54**, 14660-14663.
Ellis C, Karafyllidis I, Wasternack C, Turner JG. 2002. The Arabidopsis mutant cev1 links cell wall signaling to jasmonate and ethylene responses. *The Plant Cell* **14**, 1557-1566.
Esfandiari E, Jin Z, Abdeen A, Griffiths JS, Western TL, Haughn GW. 2013. Identification and analysis of an outer-seed-coat-specific promoter from Arabidopsis thaliana. *Plant Mol Biol* **81**, 93-104.
Felsenstein J. 1985. Confidence-Limits on Phylogenies - an Approach Using the Bootstrap. *Evolution* **39**, 783-791.
Freudenberg K. 1959. Biosynthesis and constitution of lignin. *Nature* **183**, 1152-1155.
Gallego-Giraldo L, Jikumaru Y, Kamiya Y, Tang Y, Dixon RA. 2011. Selective lignin downregulation leads to constitutive defense response expression in alfalfa (Medicago sativa L.). *The New phytologist* **190**, 627-639.
Gang DR, Costa MA, Fujita M, Dinkova-Kostova AT, Wang HB, Burlat V, Martin W, Sarkanen S, Davin LB, Lewis NG. 1999. Regiochemical control of monolignol radical coupling: a new paradigm for lignin and lignan biosynthesis. *Chemistry & Biology* **6**, 143-151.
Gao W, Long L, Zhu LF, Xu L, Gao WH, Sun LQ, Liu LL, Zhang XL. 2013. Proteomic and virus-induced gene silencing (VIGS) Analyses reveal that gossypol, brassinosteroids, and jasmonic acid contribute to the resistance of cotton to Verticillium dahliae. *Mol Cell Proteomics* **12**, 3690-3703.
Gasper R, Effenberger I, Kolesinski P, Terlecka B, Hofmann E, Schaller A. 2016. Dirigent protein mode of action revealed by the crystal structure of AtDIR6. *Plant Physiol*.
Geldner N. 2013. The endodermis. *Annual Review of Plant Biology* **64**, 531-558.
Halls SC, Davin LB, Kramer DM, Lewis NG. 2004. Kinetic study of coniferyl alcohol radical binding to the (+)-pinoresinol forming dirigent protein. *Biochemistry* **43**, 2587-2595.
Halls SC, Lewis NG. 2002. Secondary and quaternary structures of the (+)-pinoresinol-forming dirigent protein. *Biochemistry* **41**, 9455-9461.
Hao Z, Mohnen D. 2014. A review of xylan and lignin biosynthesis: foundation for studying Arabidopsis irregular xylem mutants with pleiotropic phenotypes. *Crit Rev Biochem Mol Biol* **49**, 212-241.
Harju AM, Venalainen M, Laakso T, Saranpaa P. 2009. Wounding response in xylem of Scots pine seedlings shows wide genetic variation and connection with the constitutive defence of heartwood. *Tree physiology* **29**, 19-25.
Harmatha J, Dinan L. 2003. Biological activities of lignans and stilbenoids associated with plant-insect chemical interactions. *Phytochemistry Reviews* **2**, 321–330.
Hematy K, Sado PE, Van Tuinen A, Rochange S, Desnos T, Balzergue S, Pelletier S, Renou JP, Hofte H. 2007. A receptor-like kinase mediates the response of Arabidopsis cells to the inhibition of cellulose synthesis. *Curr Biol* **17**, 922-931.
Hofmann E, Zerbe P, Schaller F. 2006. The crystal structure of Arabidopsis thaliana allene oxide cyclase: insights into the oxylipin cyclization reaction. *Plant Cell* **18**, 3201-3217.

Hosmani PS, Kamiya T, Danku J, Naseer S, Geldner N, Guerinot ML, Salt DE. 2013. Dirigent domain-containing protein is part of the machinery required for formation of the lignin-based Casparian strip in the root. *Proc Natl Acad Sci U S A* **110**, 14498-14503.
Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, Oertle L, Widmayer P, Gruissem W, Zimmermann P. 2008. Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv Bioinformatics* **2008**, 420747.
Huis R, Morreel K, Fliniaux O, Lucau-Danila A, Fenart S, Grec S, Neutelings G, Chabbert B, Mesnard F, Boerjan W, Hawkins S. 2012. Natural hypolignification is associated with extensive oligolignol accumulation in flax stems. *Plant Physiol* **158**, 1893-1915.
Hwang I, Sheen J, Muller B. 2012. Cytokinin signaling networks. *Annual Review of Plant Biology* **63**, 353-380.
Chen Z, Yan W, Sun L, Tian J, Liao H. 2016. Proteomic analysis reveals growth inhibition of soybean roots by manganese toxicity is associated with alteration of cell wall structure and lignification. *Journal of Proteomics* **143**, 151-160.
Jin-Long G, Li-Ping X, Jing-Ping F, Ya-Chun S, Hua-Ying F, You-Xiong Q, Jing-Sheng X. 2012. A novel dirigent protein gene with highly stem-specific expression from sugarcane, response to drought, salt and oxidative stresses. *Plant Cell Reports* **31**, 1801-1812.
Kamiya T, Borghi M, Wang P, Danku JM, Kalmbach L, Hosmani PS, Naseer S, Fujiwara T, Geldner N, Salt DE. 2015. The MYB36 transcription factor orchestrates Casparian strip formation. *Proc Natl Acad Sci U S A* **112**, 10533-10538.
Kazan K, Manners JM. 2013. MYC2: The Master in Action. *Molecular Plant* **6**, 686-703.
Kim K-W, Smith CA, Daily MD, Cort JR, Davin LB, Lewis NG. 2015a. Trimeric Structure of (+)-Pinoresinol-forming Dirigent Protein at 1.95 Ã… Resolution with Three Isolated Active Sites. *Journal of Biological Chemistry* **290**, 1308-1318.
Kim KW, Moinuddin SG, Atwell KM, Costa MA, Davin LB, Lewis NG. 2012. Opposite stereoselectivities of dirigent proteins in Arabidopsis and schizandra species. *The Journal of Biological Chemistry* **287**, 33957-33972.
Kim MK, Jeon JH, Davin LB, Lewis NG. 2002. Monolignol radical-radical coupling networks in western red cedar and Arabidopsis and their evolutionary implications. *Phytochemistry* **61**, 311-322.
König S, Feussner K, Kaever A, Landesfeind M, Thurow C, Karlovsky P, Gatz C, Polle A, Feussner IC. 2014. Soluble phenylpropanoids are involved in the defense response of Arabidopsis against Verticillium longisporum. *New Phytologist* **202**, 823-837.
Levin DE. 2011. Regulation of Cell Wall Biogenesis in Saccharomyces cerevisiae: The Cell Wall Integrity Signaling Pathway. *Genetics* **189**, 1145-1175.
Liu CJ. 2012. Deciphering the enigma of lignification: precursor transport, oxidation, and the topochemistry of lignin assembly. *Molecular Plant* **5**, 304-317.
Liu J, Stipanovic RD, Bell AA, Puckhaber LS, Magill CW. 2008. Stereoselective coupling of hemigossypol to form (+)-gossypol in moco cotton is mediated by a dirigent protein. *Phytochemistry* **69**, 3038-3042.
Ma Q-H. 2014. Monocot chimeric jacalins: a novel subfamily of plant lectins. *Critical reviews in biotechnology*, Vol. 34: Taylor & Francis, 300-306.
MacRae WD, Towers GHN. 1984. Biological activities of lignans. *Phytochemistry* **23**, 1207-1220.
Manke T, Roider HG, Vingron M. 2008a. Statistical modeling of transcription factor binding affinities predicts regulatory interactions. *PLoS computational biology* **4**, e1000039.
Mason SF. 1991. Origins of the handedness of biological molecules. *Ciba Foundation symposium* **162**, 3-10; discussion 10-15.
Matsuda H, Nakashima S, Abdel-Halim OB, Morikawa T, Yoshikawa M. 2010. Cucurbitane-type triterpenes with anti-proliferative effects on U937 cells from an egyptian natural medicine, Bryonia

cretica: structures of new triterpene glycosides, bryoniaosides A and B. *Chemical & pharmaceutical bulletin* **58**, 747-751.
Matys V, Fricke E, Geffers R, Gossling E, Haubrock M, Hehl R, Hornischer K, Karas D, Kel AE, Kel-Margoulis OV, Kloos DU, Land S, Lewicki-Potapov B, Michael H, Munch R, Reuter I, Rotert S, Saxel H, Scheer M, Thiele S, Wingender E. 2003. TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res* **31**, 374-378.
Miedes E, Vanholme R, Boerjan W, Molina A. 2014. The role of the secondary cell wall in plant resistance to pathogens. *Frontiers in plant science* **5**, 358.
Mohr PG, Cahill DM. 2007. Suppression by ABA of salicylic acid and lignin accumulation and the expression of multiple genes, in Arabidopsis infected with Pseudomonas syringae pv. tomato. *Functional & integrative genomics* **7**, 181-191.
Mönke G, Altschmied L, Tewes A, Reidt W, Mock H-P, Bäumlein H, Conrad U. 2004. Seed-specific transcription factors ABI3 and FUS3: molecular interaction with DNA. *Planta* **219**, 158-166.
Morita A, Yokota H, Ishka MR, Ghanati F. 2006. Changes in peroxidase activity and lignin content of cultured tea cells in response to excess manganese. *Soil Science and Plant Nutrition*, Vol. 52: Taylor & Francis, 26-31.
Nakagawa Y, Katagiri T, Shinozaki K, Qi Z, Tatsumi H, Furuichi T, Kishigami A, Sokabe M, Kojima I, Sato S, Kato T, Tabata S, Iida K, Terashima A, Nakano M, Ikeda M, Yamanaka T, Iida H. 2007. Arabidopsis plasma membrane protein crucial for Ca2+ influx and touch sensing in roots. *Proceedings of the National Academy of Sciences* **104**, 3639-3644.
Naseer S, Lee Y, Lapierre C, Franke R, Nawrath C, Geldner N. 2012. Casparian strip diffusion barrier in Arabidopsis is made of a lignin polymer without suberin. *Proc Natl Acad Sci U S A* **109**, 10101-10106.
Nei M, Kumar S. 2000. *Molecular evolution and phylogenetics*. New York ; Oxford: Oxford University Press.
Nose M, Bernards MA, Furlan M, Zajicek J, Eberhardt TL, Lewis NG. 1995. Towards the specification of consecutive steps in macromolecular lignin assembly. *Phytochemistry* **39**, 71-79.
Okonechnikov K, Golosova O, Fursov M, team U. 2012. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* **28**, 1166-1167.
Paré PW, Wang H-B, Davin LB, Lewis NG. 1994. (+)-Pinoresinol synthase: A stereoselective oxidase catalysing 8,8'-lignan formation in Forsythia intermedia. *Tetrahedron letters* **35**, 4731-4734.
Pickel B, Constantin MA, Pfannstiel J, Conrad J, Beifuss U, Schaller A. 2010. An enantiocomplementary dirigent protein for the enantioselective laccase-catalyzed oxidative coupling of phenols. *Angewandte Chemie* **49**, 202-204.
Pickel B, Pfannstiel J, Steudle A, Lehmann A, Gerken U, Pleiss J, Schaller A. 2012. A model of dirigent proteins derived from structural and functional similarities with allene oxide cyclase and lipocalins. *The FEBS journal* **279**, 1980-1993.
Ponce De Leon I, Schmelz EA, Gaggero C, Castro A, Alvarez A, Montesano M. 2012. Physcomitrella patens activates reinforcement of the cell wall, programmed cell death and accumulation of evolutionary conserved defence signals, such as salicylic acid and 12-oxo-phytodienoic acid, but not jasmonic acid, upon Botrytis cinerea infection. *Molecular plant pathology* **13**, 960-974.
Quentin M, Allasia V, Pegard A, Allais F, Ducrot PH, Favery B, Levis C, Martinet S, Masur C, Ponchet M, Roby D, Schlaich NL, Jouanin L, Keller H. 2009. Imbalanced lignin biosynthesis promotes the sexual reproduction of homothallic oomycete pathogens. *PLoS Pathog* **5**, e1000264.
Ralph J, Schatz PF, Lu F, Kim H, Akiyama T, Nelsen SF. 2009. Quinone methides in lignification. In: S.E. R, ed. *Quinone Methides*. Hoboken, NJ: John Wiley & Sons, 385-420.
Ralph S, Park J-Y, Bohlmann Jr, Mansfield SD. 2006. Dirigent Proteins in Conifer Defense: Gene Discovery, Phylogeny, and Differential Wound- and Insect-induced Expression of a Family of DIR and DIR-like Genes in Spruce (Picea spp.). *Plant Molecular Biology* **60**, 21-40.

Reboledo G, del Campo R, Alvarez A, Montesano M, Mara H, Ponce de Leon I. 2015. Physcomitrella patens Activates Defense Responses against the Pathogen Colletotrichum gloeosporioides. *International journal of molecular sciences* **16**, 22280-22298.
Reiser Vr, Raitt DC, Saito H. 2003. Yeast osmosensor Sln1 and plant cytokinin receptor Cre1 respond to changes in turgor pressure. *The Journal of Cell Biology* **161**, 1035-1040.
Roider HG, Kanhere A, Manke T, Vingron M. 2007. Predicting transcription factor affinities to DNA from a biophysical model. *Bioinformatics* **23**, 134-141.
Roppolo D, De Rybel B, Tendon VD, Pfister A, Alassimone J, Vermeer JE, Yamazaki M, Stierhof YD, Beeckman T, Geldner N. 2011. A novel protein family mediates Casparian strip formation in the endodermis. *Nature* **473**, 380-383.
Ruprecht C, Mutwil M, Saxe F, Eder M, Nikoloski Z, Persson S. 2011. Large-scale co-expression approach to dissect secondary cell wall formation across plant species. *Front Plant Sci* **2**, 23.
Satake H, Koyama T, Bahabadi SE, Matsumoto E, Ono E, Murata J. 2015. Essences in metabolic engineering of lignan biosynthesis. *Metabolites* **5**, 270-290.
Seneviratne HK, Dalisay DS, Kim K-W, Moinuddin SGA, Yang H, Hartshorn CM, Davin LB, Lewis NG. 2015. Non-host disease resistance response in pea (Pisum sativum) pods: Biochemical function of DRR206 and phytoalexin pathway localization. *Phytochemistry* **113**, 140-148.
Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731-2739.
Taniguchi M, Sasaki N, Tsuge T, Aoyama T, Oka A. 2007. ARR1 directly activates cytokinin response genes that encode proteins with diverse regulatory functions. *Plant Cell Physiol* **48**, 263-277.
Teponno RB, Kusari S, Spiteller M. 2016. Recent advances in research on lignans and neolignans. *Natural product reports* **33**, 1044-1092.
Vassão DG, Kim K-W, Davin LB, Lewis NG. 2010. 1.23 - Lignans (Neolignans) and Allyl/Propenyl Phenols: Biogenesis, Structural Biology, and Biological/Human Health Considerations. *Comprehensive Natural Products II*. Oxford: Elsevier, 815-928.
Villalobos DP, Diaz-Moreno SM, Said el SS, Canas RA, Osuna D, Van Kerckhoven SH, Bautista R, Claros MG, Canovas FM, Canton FR. 2012. Reprogramming of gene expression during compression wood formation in pine: coordinated modulation of S-adenosylmethionine, lignin and lignan related genes. *BMC Plant Biol* **12**, 100.
Weber AL, Pizzarello S. 2006. The peptide-catalyzed stereospecific synthesis of tetroses: A possible model for prebiotic molecular evolution. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 12713-12717.
Weidenbach D, Esch L, MÃ¶ller C, Hensel G, Kumlehn J, HÃ¶fle C, HÃ¼ckelhoven R, Schaffrath U. 2016. Polarized Defense Against Fungal Pathogens Is Mediated by the Jacalin-Related Lectin Domain of Modular Poaceae-Specific Proteins. *Molecular Plant* **9**, 514-527.
Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ. 2007. An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS One* **2**, e718.
Wolf S, Hofte H. 2014. Growth Control: A Saga of Cell Walls, ROS, and Peptide Receptors. *Plant Cell* **26**, 1848-1856.
Wu RH, Wang L, Wang Z, Shang HH, Liu X, Zhu Y, Qi DD, Deng X. 2009. Cloning and expression analysis of a dirigent protein gene from the resurrection plant Boea hygrometrica. *Progress in Natural Science* **19**, 347-352.
Zeitoun AM, Preisner M, Kulma A, Dymińska L, Hanuza J, Starzycki M, Szopa J. 2014. Does biopolymers composition in seeds contribute to the flax resistance against the Fusarium infection? *Biotechnology Progress* **30**, 992-1004.

Zhao Q, Zeng Y, Yin Y, Pu Y, Jackson LA, Engle NL, Martin MZ, Tschaplinski TJ, Ding SY, Ragauskas AJ, Dixon RA. 2015. Pinoresinol reductase 1 impacts lignin distribution during secondary cell wall biosynthesis in Arabidopsis. *Phytochemistry* **112**, 170-178.

Zhong R, Lee C, Zhou J, McCarthy RL, Ye ZH. 2008. A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in Arabidopsis. *Plant Cell* **20**, 2763-2782.

Figure Legends

Figure 1. **Products of phenoxy radical coupling of E-coniferyl alcohol radicals.**

In the presence of peroxidases or laccases but absence of DIR proteins, the result is racemic mixture of approx. equal amounts of (+/-) 8,8', (+/-) 8,5' and 8-O-4' linked products. The presence of DIR proteins results into regio- and stereoselective coupling to give (-)-pinoresinol or (+)-pinoresinol depend on the specificity of DIR. LACs - Laccases; PRXs - Peroxidases; DIRs - dirigent proteins.

Figure 2. Structure of AtDIR6.

AtDIR6 trimer **(A)** top and **(B)** side view. Each monomer colored in different color. Glycosylation shown as sticks. Location of active site is highlighted by black circle in one of the monomers. **(C)** AtDIR6 active site. Residues forming pocket A are shown in green, pocket B shown in yellow and residues in between pockets are shown in orange.

Figure 3. Localization of functional and conserved residues of AtDIR6.
(A) AtDIR6 residues important for protein activity. Based on Gasper et al. (2016), mutation of green residues results in increase of activity, mutation of yellow residues in slight decrease of activity and of magenta residues in strong decrease of activity. **(B)** DIR6 monomer with residues conserved within DIR proteins highlighted. Color code: cyan – non conserved, grey – fully conserved across (+) and (-) DIRs, blue – (-) DIR conserved, magenta – (+) DIR conserved, red – differentially (+) and (-) conserved residues.

Figure 4. Schematic representation of the of bisQM intermediate cyclization catalyzed by AtDIR6.

The cartoon combines the side view on the active site as shown in Fig. 2 with the key functional residues highlighted (sticks). The mechanism of bisQM cyclization proposed by Gasper et al. (2016) via acid catalysis (left) or hydrogen bond formation (right) is schematically shown. Both mechanisms seem to result into partial or full positive charge on C7 that facilitates the nucleophilic attack during bisQM cyclization.

Figure 5. Phylogenetic analysis of AtDIRs.

The cluster analysis is based on the alignment of amino acid sequences of DIR domains of all 26 members of the AtDIR family. Three subfamilies identified with high reliability are highlighted (DIR-a, yellow; DIR-d, green; DIR-e, blue).

Figure 6. **Expression pattern of *AtDIR* genes**.

Schematic representation of highest expression values of individual AtDIRs as determined by the eFP browser tool (Winter *et al.*, 2007) and genevestigator (Hruz et al. 2008). Most of the *AtDIRs* show highest expression levels in the roots. However, few *AtDIR* genes show highest expression levels in flowers (floral organs) or seeds. The picture does not reflect the developmental stage in which the expression occurs (especially in case of the root), but is used only for visualization purposes. For the cell-specific expression in the root see Supplementary Fig. S6. For those *AtDIRs* which are not represented on Affymetrix microarray chips, RNAseq data from the Genevestigator database were analyzed. According to these data, *AtDIR3* shows highest expression levels in seeds, but the exact time point of expression during seed development is unclear, as indicated by the question mark.

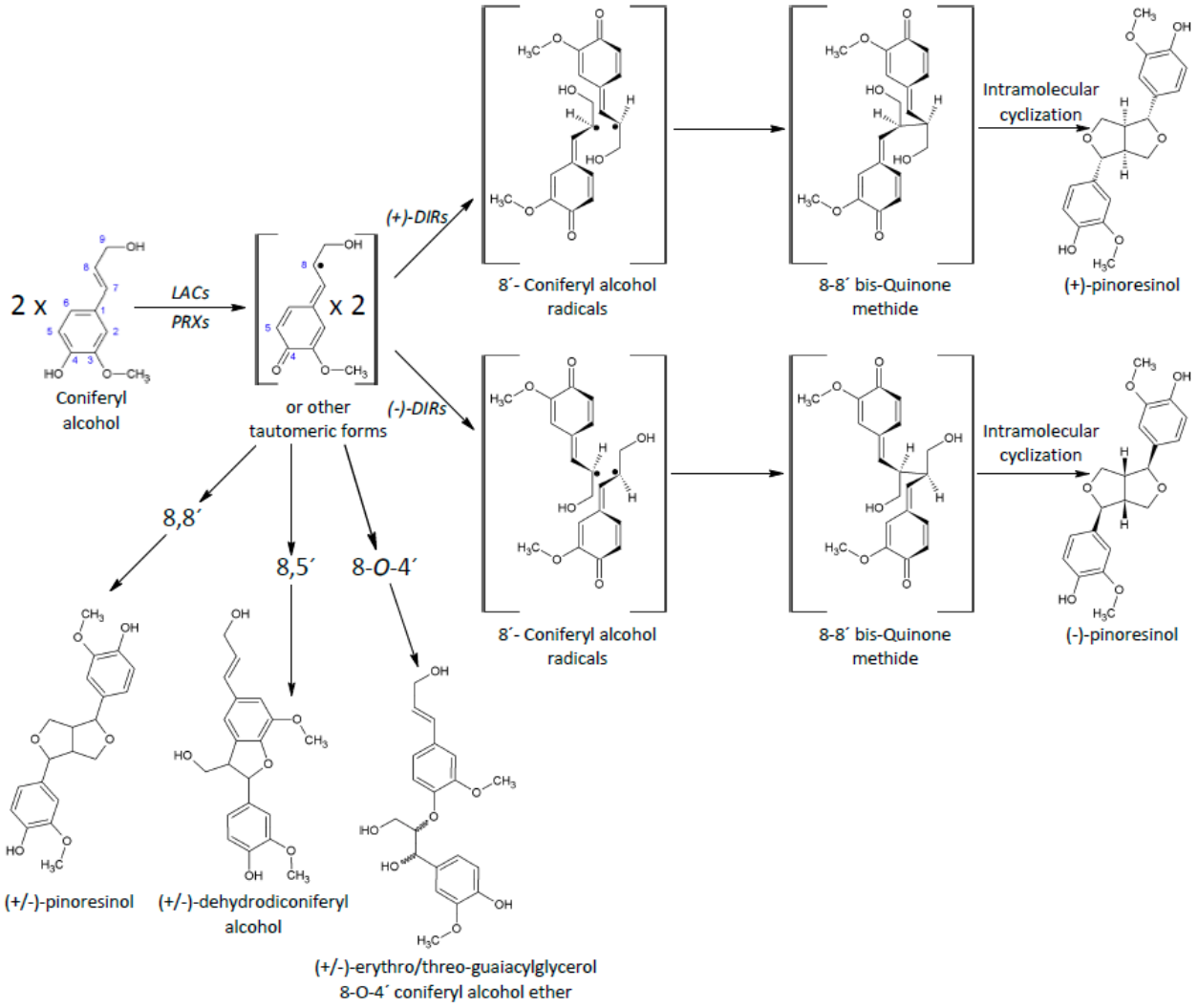

Figure 1. **Products of phenoxy radical coupling of E-coniferyl alcohol radicals.**

In the presence of peroxidases or laccases but absence of DIR proteins, the result is racemic mixture of approx. equal amounts of (+/-) 8,8', (+/-) 8,5' and 8-O-4' linked products. The presence of DIR proteins results into regio- and stereoselective coupling to give (-)-pinoresinol or (+)-pinoresinol depend on the specificity of DIR. LACs - Laccases;

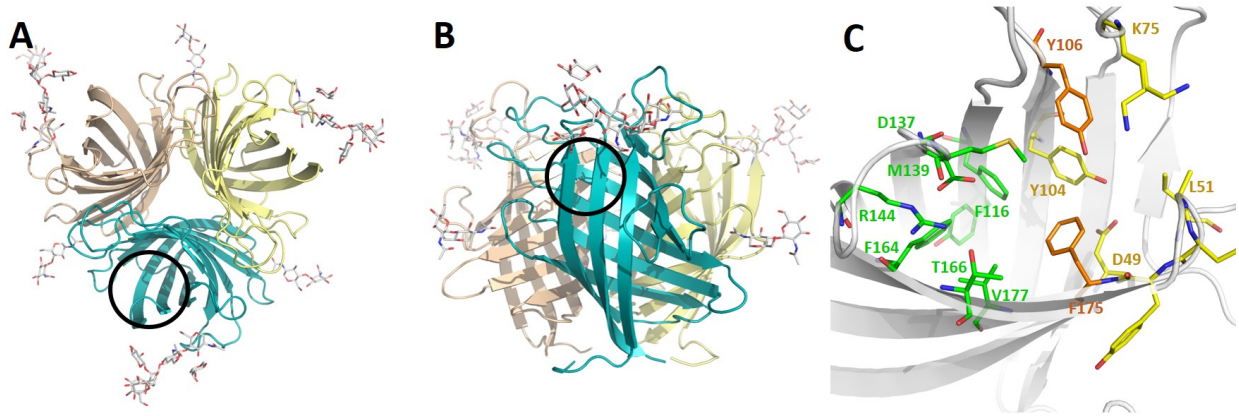

Figure 2. Structure of AtDIR6.

AtDIR6 trimer **(A)** top and **(B)** side view. Each monomer colored in different color. Glycosylation shown as sticks. Location of active site is highlighted by black circle in one of the monomers. **(C)** AtDIR6 active site. Residues forming pocket A are shown in green, pocket B shown in yellow and residues in between pockets are shown in orange.

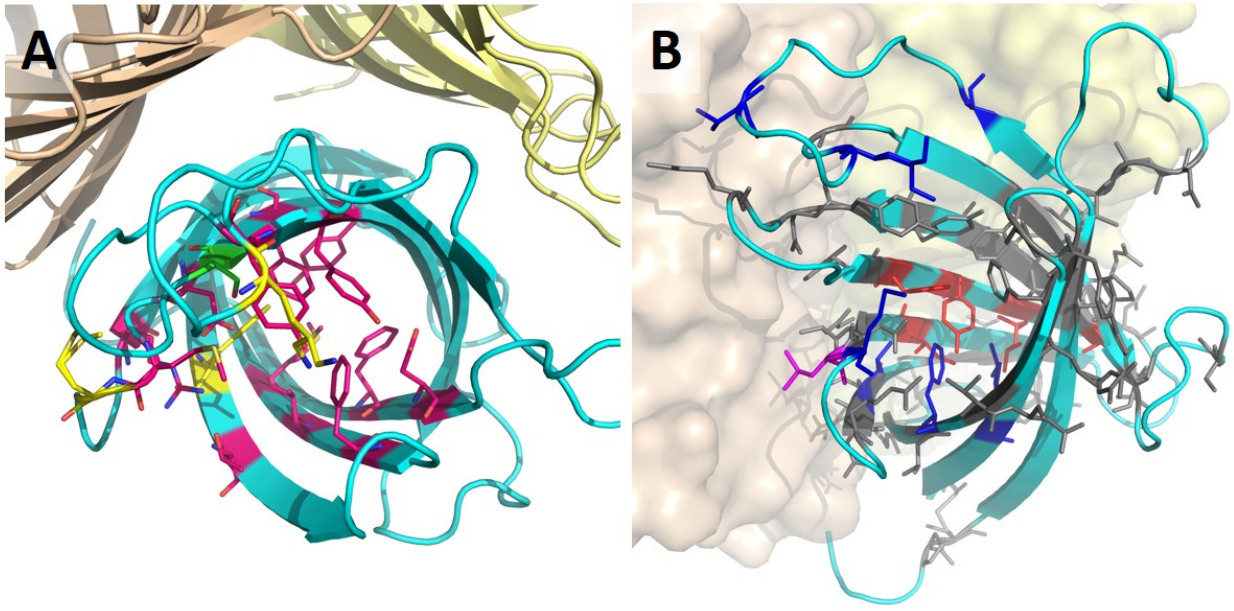

Figure 3. Localization of functional and conserved residues of AtDIR6.

(A) AtDIR6 residues important for protein activity. Based on Gasper et al. (2016), mutation of green residues results in increase of activity, mutation of yellow residues in slight decrease of activity and of magenta residues in strong decrease of activity. **(B)** DIR6 monomer with residues conserved within DIR proteins highlighted. Color code: cyan – non conserved, grey – fully conserved across (+) and (-) DIRs, blue – (-) DIR conserved, magenta – (+) DIR conserved, red – differentially (+) and (-) conserved residues.

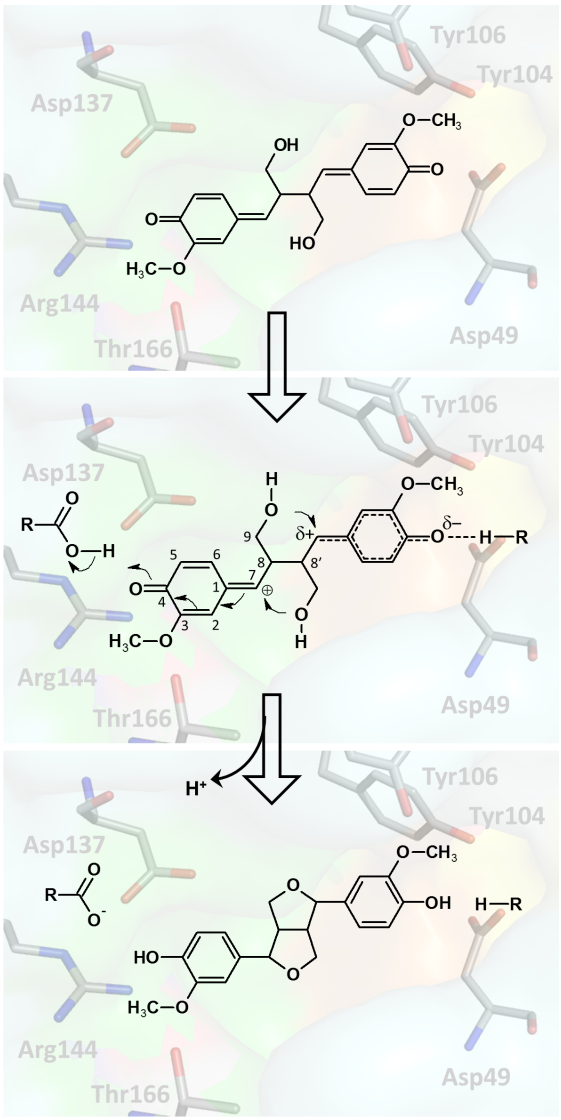

Figure 4. Schematic representation of the of bisQM intermediate cyclization catalyzed by AtDIR6.

The cartoon combines the side view on the active site as shown in Fig. 2 with the key functional residues highlighted (sticks). The mechanism of bisQM cyclization proposed by Gasper et al. (2016) via acid catalysis (left) or hydrogen bond formation (right) is schematically shown. Both mechanisms seem to result into partial or full positive charge on C7 that facilitates the nucleophilic attack during bisQM cyclization.

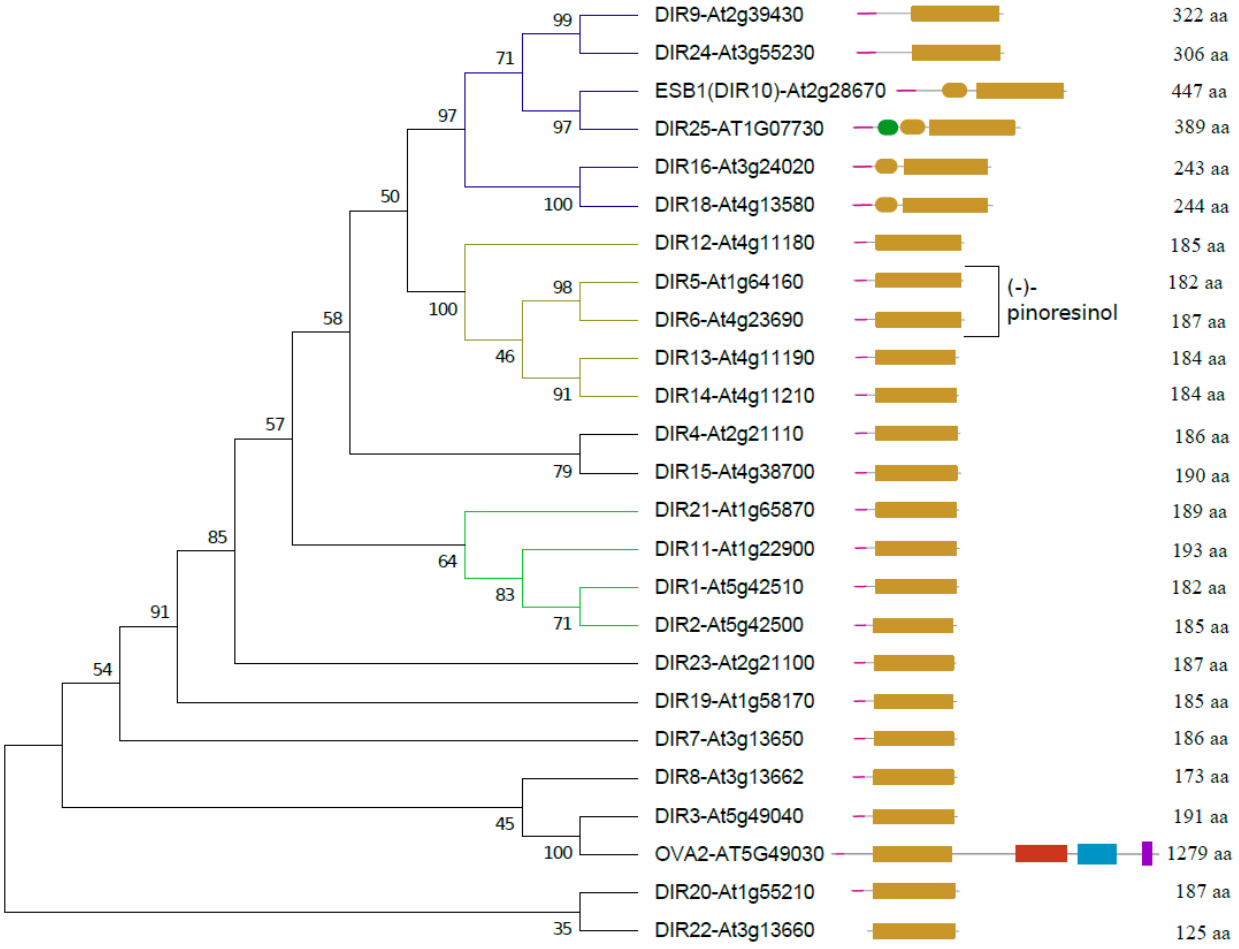

Figure 5. Phyllogenetic analysis of AtDIRs.

The cluster analysis is based on the alignment of amino acid sequences of DIR domains of all 26 members of the AtDIR family. Three subfamilies identified with high reliability are highlighted (DIR-a, yellow; DIR-d, green; DIR-e, blue).

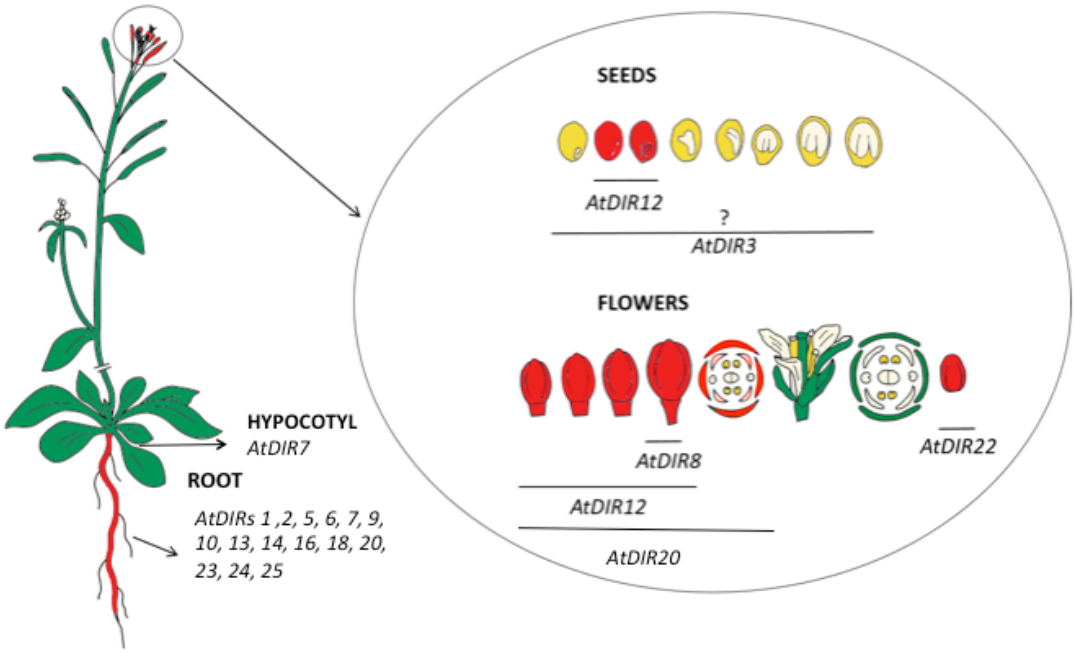

Figure 6. **Expression pattern of *AtDIR* genes**.

Schematic representation of highest expression values of individual AtDIRs as determined by the eFP browser tool (Winter *et al.*, 2007) and genevestigator (Hruz et al. 2008). Most of the *AtDIRs* show highest expression levels in the roots. However, few *AtDIR* genes show highest expression levels in flowers (floral organs) or seeds. The picture does not reflect the developmental stage in which the expression occurs (especially in case of the root), but is used only for visualization purposes. For the cell-specific expression in the root see Supplementary Fig. S6. For those *AtDIRs* which are not represented on Affymetrix microarray chips, RNAseq data from the Genevestigator database were analyzed. According to these data, *AtDIR3* shows highest expression levels in seeds, but the exact time point of expression during seed development is unclear, as indicated by the question mark.