




Fern cell walls and the evolution of arabinogalactan proteins in streptophytes

Kim-Kristine Mueller^{1,†}, Lukas Pfeifer^{1,†} , Lina Schuldt¹, Péter Szövényi^{2,3}, Sophie de Vries⁴, Jan de Vries^{4,5,6} , Kim L. Johnson⁷ and Birgit Classen^{1,*} 

¹Pharmaceutical Institute, Department of Pharmaceutical Biology, Christian-Albrechts-University of Kiel, Gutenbergstr. 76, 24118, Kiel, Germany,

²Department of Systematic and Evolutionary Botany, University of Zurich, Zollikerstr. 107, 8008, Zurich, Switzerland,

³Zurich-Basel Plant Science Center (PSC), ETH Zürich, Tannenstrasse 1, 8092, Zürich, Switzerland,

⁴Department of Applied Bioinformatics, Institute of Microbiology and Genetics, University of Goettingen, Goldschmidtstr. 1, 37077, Goettingen, Germany,

⁵Department of Applied Bioinformatics, University of Goettingen, Goettingen Center for Molecular Biosciences (GZMB), Goldschmidtstr. 1, 37077, Goettingen, Germany,

⁶Campus Institute Data Science (CIDAS), University of Goettingen, Goldschmidtstr. 1, 37077, Goettingen, Germany, and

⁷Department of Animal, Plant and Soil Science, La Trobe Institute for Agriculture & Food, La Trobe University, AgriBio Building, Bundoora, Victoria 3086, Australia

Received 15 December 2022; revised 22 February 2023; accepted 6 March 2023; published online 9 March 2023.

*For correspondence (e-mail bclassen@pharmazie.uni-kiel.de)

[†]These authors contributed equally to this work.

SUMMARY

Significant changes have occurred in plant cell wall composition during evolution and diversification of tracheophytes. As the sister lineage to seed plants, knowledge on the cell wall of ferns is key to track evolutionary changes across tracheophytes and to understand seed plant-specific evolutionary innovations. Fern cell wall composition is not fully understood, including limited knowledge of glycoproteins such as the fern arabinogalactan proteins (AGPs). Here, we characterize the AGPs from the leptosporangiate fern genera *Azolla*, *Salvinia*, and *Ceratopteris*. The carbohydrate moiety of seed plant AGPs consists of a galactan backbone including mainly 1,3- and 1,3,6-linked pyranosidic galactose, which is conserved across the investigated fern AGPs. Yet, unlike AGPs of angiosperms, those of ferns contained the unusual sugar 3-O-methylrhannose. Besides terminal furanosidic arabinose, Ara (Araf), the main linkage type of Araf in the ferns was 1,2-linked Araf, whereas in seed plants 1,5-linked Araf is often dominating. Antibodies directed against carbohydrate epitopes of AGPs supported the structural differences between AGPs of ferns and seed plants. Comparison of AGP linkage types across the streptophyte lineage showed that angiosperms have rather conserved monosaccharide linkage types; by contrast bryophytes, ferns, and gymnosperms showed more variability. Phylogenetic analyses of glycosyltransferases involved in AGP biosynthesis and bioinformatic search for AGP protein backbones revealed a versatile genetic toolkit for AGP complexity in ferns. Our data reveal important differences across AGP diversity of which the functional significance is unknown. This diversity sheds light on the evolution of the hallmark feature of tracheophytes: their elaborate cell walls.

Keywords: arabinogalactan protein, *Azolla filiculoides*, cell wall, *Ceratopteris richardii*, glycosyltransferases, plant evolution, ferns, phylogenetic analysis, polysaccharides, *Salvinia molesta*.

INTRODUCTION

One of the most important steps in the evolution of life on Earth happened at about 500 million years ago (mya), when descendants of streptophyte algae colonized the terrestrial habitat (e.g., Becker & Marin, 2009; Bowman

et al., 2017; de Vries & Archibald, 2018; Delwiche & Cooper, 2015; Harrison, 2017; Kenrick & Crane, 1997). By the end of the Devonian (360 mya), the extant lineages of land plants comprising bryophytes, lycophytes, monilophytes (=ferns), and spermatophytes have evolved and the major

organ and tissue types known from extant plants (e.g., vasculature, roots, leaves, seeds, wood, secondary growth) were already present (Bowman, 2013). Vascular plants (tracheophytes) comprise lycophytes, ferns, and seed plants, and are characterized by features such as tracheids and sieve elements for water and nutrient transport and a highly structured and dominant sporophyte. Lycophytes separated from all other living vascular plant lineages already in the nearly mid Devonian (approximately 400 mya; Banks et al., 2011; Pryer et al., 2004), ferns first appeared in the early Carboniferous (Galtier & Scott, 1985). They comprise approximately 12 000 extant species belonging to the five lineages Equisetales, Psilotales, Ophioglossales, Marattiales, and the dominant group of leptosporangiate ferns (Nitta et al., 2022; PPG, 2016; Pryer et al., 2004). Ferns form the sister lineage to seed plants; to understand the evolutionary history of innovations specific to seed plants, knowledge of this group is necessary to infer (i) the characteristics of the last common ancestor (LCA) shared by ferns and seed plants, and (ii) define the evolutionary trajectory of seed plant-specific characteristics (Rensing, 2017).

Ferns are known for possessing large genomes with numerous chromosomes. To date, the genome sequence of the heterosporous ferns *Azolla filiculoides* and *Salvinia cucullata*, the homosporous ferns *Ceratopteris richardii*, *Adiantum capillus-veneris*, as well as the tree fern *Alsophila spinulosa* are publicly available (Fang et al., 2022; Huang et al., 2022; Li et al., 2018; Marchant et al., 2022). *Azolla* and *Salvinia* are floating aquatic ferns with high growth rates and the potential to be significant carbon sinks. Data from the Arctic Ocean revealed that approximately 50 mya, an abundance of *Azolla* characterized an 800 000-year interval called the “*Azolla* event,” which possibly had a role in global cooling by sequestering atmospheric carbon dioxide (Brinkhuis et al., 2006; Speelman et al., 2009). *Azolla* is further remarkable due to its obligate symbiosis with the N-fixing cyanobacterium *Nostoc azollae* in specialized leaf cavities (de Vries et al., 2018; de Vries & de Vries, 2022; Li et al., 2018). *Ceratopteris richardii* is also adapted to water and is known as the genetically tractable fern model organism (“C-Fern”; Plackett et al., 2015) and as a teaching tool in biology (Renzaglia & Warne, 1995). *Azolla*, *Salvinia*, and *Ceratopteris* all belong to the leptosporangiate ferns, which account for approximately 80% of non-flowering vascular plant species (Plackett et al., 2015). Among them *Ceratopteris* was suggested as the model for research on adaptive cell wall modification in vascular plants (Leroux, Eeckhout, et al., 2013).

Plant cell walls are important at various levels of plant morphology and as such are expected to have changed during evolution (Sørensen et al., 2011). According to current knowledge, fern cell walls share basic features with those of seed plants, e.g., the occurrence of cellulose,

hemicelluloses, and pectic polysaccharides (Leroux, Eeckhout, et al., 2013; Matsunaga et al., 2004; Popper, 2008; Popper & Fry, 2004). However, differences exist, e.g., in lignin structures; whereas secondary cell walls of lycophytes and angiosperms are reinforced with lignin containing mainly syringyl monomers, lignin of ferns, and gymnosperms is derived primarily from guaiacyl monomers (Weng et al., 2008). Besides polysaccharides and lignin, hydroxyproline-rich-glycoproteins (HRGPs) are also important components of plant cell walls. They are broadly classified into the three groups arabinogalactan proteins (AGPs), extensins, and proline-rich proteins (PRPs); however, in reality a continuum exists (Johnson et al., 2018). PRPs are only minimally glycosylated, extensins possess a glycan part of about 50% and AGPs consist of up to 90% polysaccharides. In AGPs, several arabinogalactan (AG) moieties are covalently linked to the protein via hydroxyproline. The characteristic structure of the AG in angiosperms is a backbone of 1,3-linked β -D-Galp, branched at position 6 to 1,6-linked β -D-Galp side chains, which are substituted with α -L-Araf and often terminal β -D-GlcPA (for review see Ma et al., 2018; Strasser et al., 2021). AGPs are key constituents of the extracellular matrix of seed plants and are involved in several processes such as cell growth, cell proliferation, pattern formation, sexual reproduction, and plant-microbe interactions (for review see Ma et al., 2018; Seifert & Roberts, 2007).

A unique feature of AGPs is their ability to precipitate with red-colored Yariv phenylglycosides, e.g., the β -glucosyl Yariv reagent (β GlcY). Already in the 1970s, the occurrence of AGPs in extracts of different bryophytes and ferns was shown by precipitation in gel-diffusion assays with Yariv's reagent (Clarke et al., 1978). In addition, AGPs in plant tissue or extracts can be detected using monoclonal antibodies directed against AG motifs. Using these antibodies in microscopy or enzyme-linked immunosorbent assay (ELISA) revealed AGP glycan epitopes in different bryophytes (Bartels et al., 2017; Berry et al., 2016; Happ & Classen, 2019; Kremer et al., 2004; Ligrone et al., 2002) and ferns (Bartels & Classen, 2017; Eeckhout et al., 2014; Lopez & Renzaglia, 2014).

Up to now, fine structures of AGPs from bryophytes (Bartels et al., 2017; Fu et al., 2007; Happ & Classen, 2019) and ferns (Akiyama et al., 1987; Bartels & Classen, 2017) were only sparsely investigated. To assess systematically the cell wall characteristics of the emerging model ferns, we isolated and characterized AGPs from the leptosporangiate ferns *Azolla filiculoides*, *Salvinia molesta*, and *Ceratopteris richardii*. To obtain broader insight into AGP biosynthesis in ferns, we searched for members of the glycosyltransferase family 31 (GT31) involved in AG glycan biosynthesis as well as identified sequences encoding AGP and other HRGP protein backbones. Knowledge of cell wall diversity contributes to our understanding of plant

evolution. Owing their crucial phylogenetic position as sister to land plants, ferns are key for understanding cell wall evolution of land plants.

RESULTS

We used *S. molesta*, *A. filiculoides*, and *C. richardii* to investigate the diversity and conservation of AGPs in cell walls of ferns. The three species were cultivated under greenhouse conditions, harvested, and dried (see Figure 1 for appearance of cultivated ferns and phylogeny of the green plant lineage). In the water-soluble fractions (aqueous extract [AE]) of all species (Tables S1–S3), galactose (Gal), and glucose (Glc) were present in high amounts. The content of arabinose (Ara) differed between the

investigated species with about 30% in *Salvinia*, 20% in *Ceratopteris*, and only 10% in *Azolla*. AE of *Azolla* further differed by higher amounts of mannans (Man), uronic acids, and fucose (Fuc) as well as lack of 3-O-methylrhhamnose (3-O-MeRha; trivial name acofriose). The content of xylose (Xyl) was another striking difference between the species with low amounts in *Salvinia* but over 20% in *Ceratopteris* and *Azolla*.

(Glyco)diversity in fern AGPs

In a gel diffusion assay with the water-soluble fractions (AE) of the three ferns and β GlcY, red precipitation lines occurred after 22 h of incubation in the dark at room temperature, indicating the presence of AGPs (Figure 2a). The

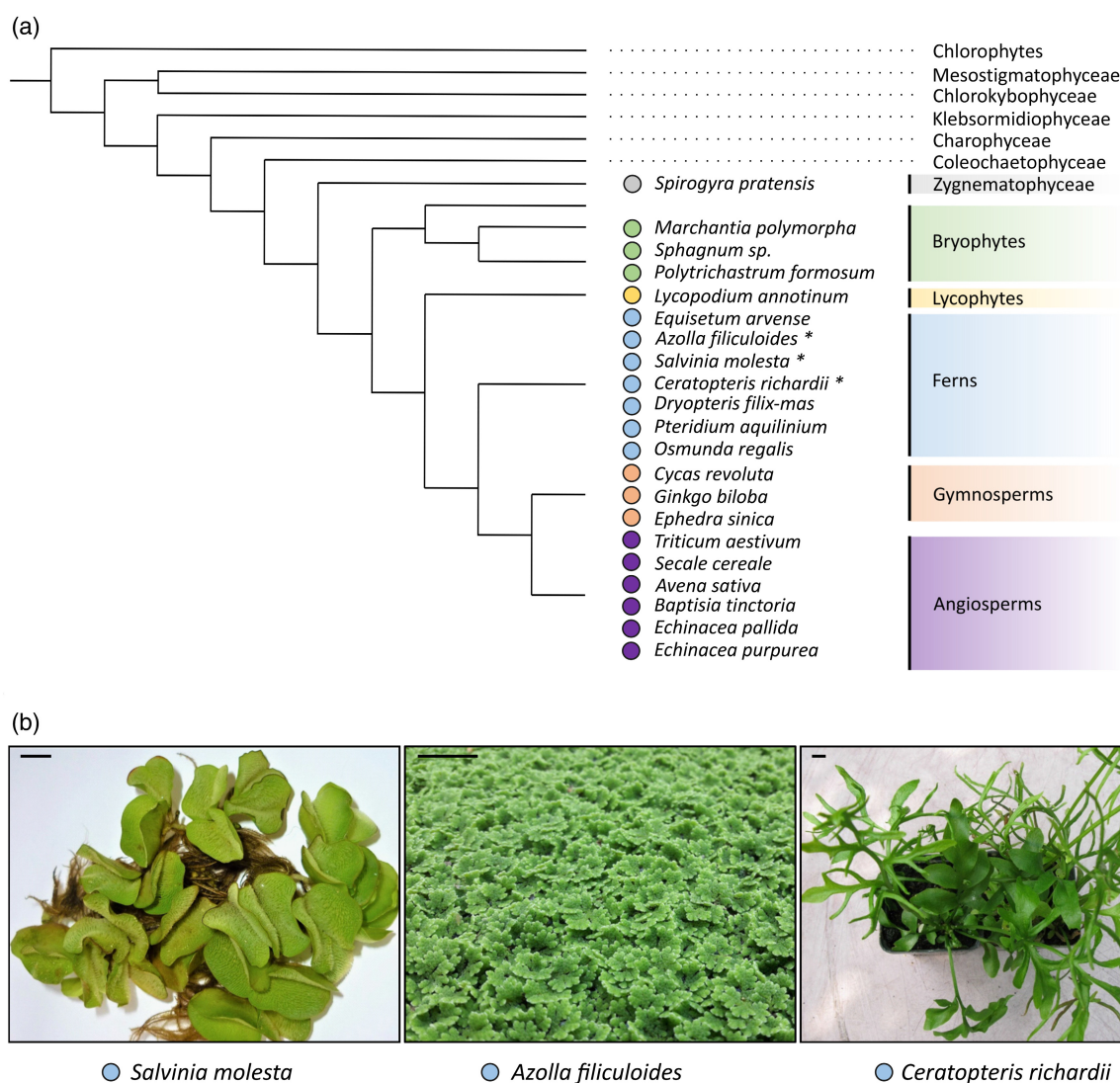


Figure 1. Overview of the investigated plant species.

(a) Phylogenetic position of the analyzed plants according to de Vries and Archibald (2018) and Puttick et al. (2018). Colored dots indicate inclusion in the principal components analysis (see Figure 3). Fern species that were cultured for this study are marked with asterisks.

(b) Cultivation of the ferns *Salvinia molesta*, *Azolla filiculoides*, and *Ceratopteris richardii* in the greenhouse of the Pharmaceutical Institute of Kiel University. Scale bars = 1 cm.

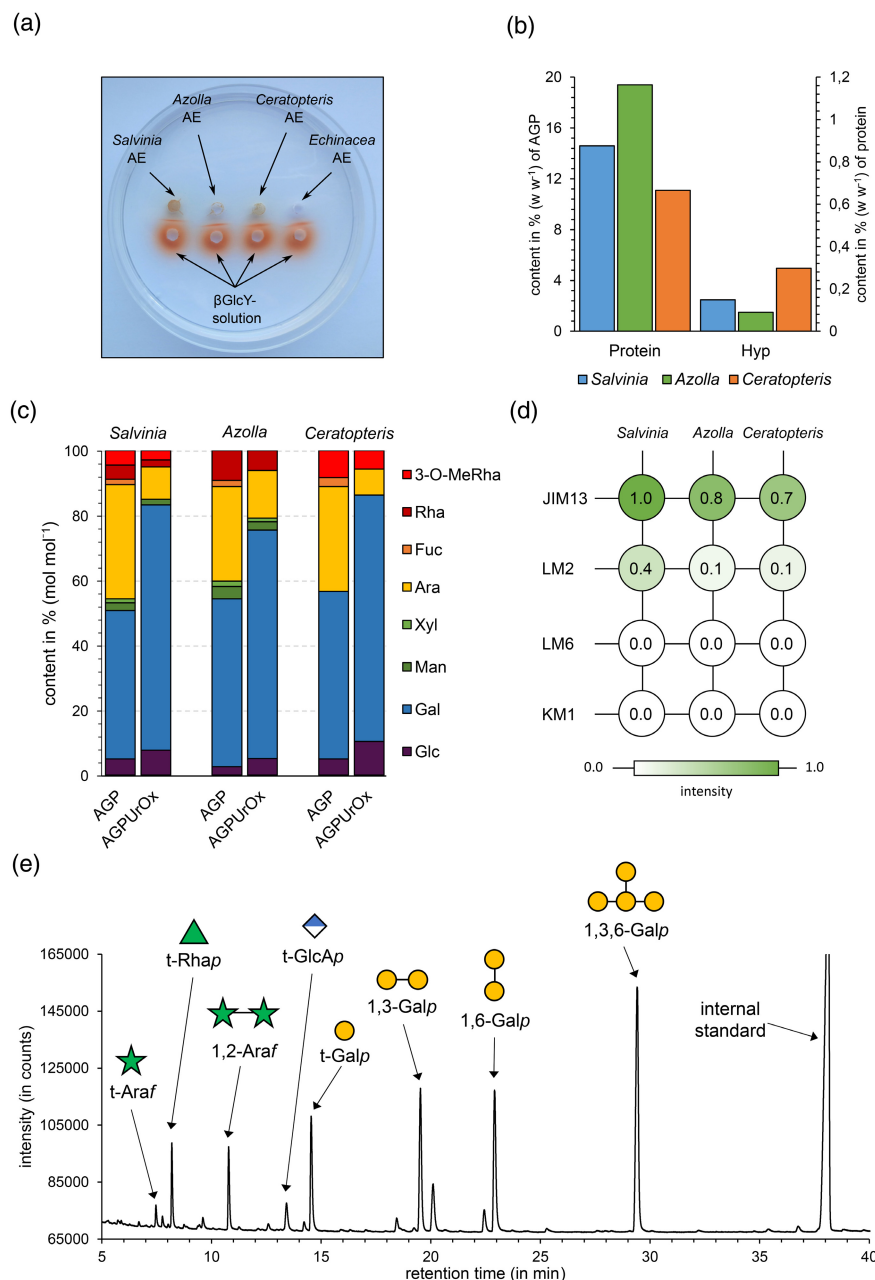


Figure 2. Features of arabinogalactan proteins (AGPs) from *Salvinia molesta*, *Azolla filiculoides*, and *Ceratopteris richardii*. (a) Gel diffusion assay with β GlcY and fractions of aqueous extract (AE) from the three ferns (100 mg ml⁻¹) compared with *Echinacea purpurea* AGP (10 mg ml⁻¹). Red precipitation line indicates presence of AGPs. (b) Protein (% w/w of AGP) and hydroxyproline (Hyp) content (% w/w of protein) in AGPs from the three species. (c) Relative monosaccharide composition of AGPs from the three ferns in the native state (AGP) and after reduction of uronic acids and partial hydrolysis with oxalic acid (AGP_{UroX}) determined by gas chromatography (% mol mol⁻¹). (d) Results of enzyme-linked immunosorbent assay (ELISA) experiments with AGPs and antibodies directed against epitopes present in AGPs (JIM13, LM2, LM6, KM1) after 10 min. For epitopes of the antibodies see Table S6. (e) Flame ionization detection chromatogram of partially methylated alditol acetates from *Salvinia* AGP_{UroX} after gas-liquid chromatography analysis.

presence of AGPs was further confirmed by isolation with β GlcY in amounts of 0.08% (*Salvinia*, $n = 3$), 0.18% (*Azolla*, $n = 2$), and 0.07% (*Ceratopteris*, $n = 1$) in relation to the dry weight of the ferns.

The amounts of nitrogen in the AGP samples were determined by elemental analyses and used to estimate the protein content of the AGPs according to Kjeldahl (factor 6.25; Figure 2b). The protein amounts of AGPs

were 14.6% for *Salvinia*, 19.4% for *Azolla*, and 11.1% for *Ceratopteris*. In AGPs, the amino acid hydroxyproline is responsible for O-glycosidic linkage of the arabinogalactan moieties to the protein. Photometric quantification of this amino acid (Stegemann & Stalder, 1967) revealed amounts of 0.36% for *Salvinia*, 0.29% for *Azolla*, and 0.55% for *Ceratopteris* AGP. This means that hydroxyproline (Hyp) accounts for 2.5% of the protein moiety in *Salvinia*, 1.5% of the protein moiety in *Azolla*, and 5.0% of the protein moiety in *Ceratopteris* AGP (Figure 2b).

As expected, Gal and Ara were the dominant monosaccharides of the Yariv-precipitated AGPs and accounted for over 80% of the neutral monosaccharides in all species (Figure 2c, Table S4) with an Ara/Gal ratio comparable with other AGPs of ferns (Bartels & Classen, 2017) and seed plants (Clarke et al., 1978). These monosaccharides were accompanied by rhamnose (Rha)/3-O-MeRha in substantial amounts between 8.3% and 9.2% (counted together). Interestingly, the ratio between both deoxyhexoses varied strongly. In *Salvinia* AGP, both were present in nearly equal amounts, whereas in *Azolla* AGP, 3-O-MeRha was present only in traces and in *Ceratopteris* AGP, the methylated Rha is strongly dominating (Figure 2c, Table S4). The monosaccharides Glc, Man, Xyl, and Fuc were present in smaller amounts, but it cannot be fully excluded that they might be part of other polysaccharides not completely separated during Yariv precipitation. Some Glc might also be the residue of the precipitating agent β GlcY. Photometric quantification of uronic acids revealed a content of 11.0% for *Salvinia*, 5.6% for *Azolla*, and 11.1% for *Ceratopteris* AGPs (Table S3).

To gain further information, uronic acids were carboxy-reduced to the corresponding neutral monosaccharides with sodium borodeuteride. Afterwards, deuterated Glc was detected, thus revealing that D-glucuronic

acid (GlcA) has been part of the native AGP. Furthermore, the samples were partially hydrolyzed with oxalic acid, a treatment known to hydrolyze mainly furanosidic Ara residues. This led to a strong decrease of the Ara content (Figure 2c, Table S4) and to further decrease of the minor monosaccharides Rha, Fuc, Xyl, and Man.

Fern AGP structural characteristics appear evolutionarily more flexible compared with angiosperms

Linkage types of monosaccharides were determined by methylation analyses before and after partial acid hydrolysis and uronic acid reduction (Table 1, Figure 2e, compare Figure S1 and S1). In the native AGPs, the typical galactan backbone known from seed plant AGPs was present with pyranosidic Gal in 1-, 1,3-, 1,6-, and 1,3,6-linkage. According to the retention time and the mass spectra, a further peak represented a mixture of pyranosidic, 1,2- and 1,4-linked hexoses (Figure S1). 1,2 and/or 1,4-linked Galp have also been detected in some moss and other fern AGPs (Bartels et al., 2017; Bartels & Classen, 2017). Furanosidic Ara, which is typically located at the periphery of the AGP molecule, was 1-, 1,2-, and 1,5-linked with dominance of the 1,2-linkage type. After partial acid hydrolysis, there was nearly complete loss of the 1,5-linked Ara_f, strong reduction of terminal Ara_f, and moderate reduction of 1,2-Ara_f. Furthermore, acid hydrolysis strongly increased the amount of 1,6-linked Galp, which is an indication that Ara is mainly bound to Gal at C-3 of 1,3,6-linked Gal. The increase in terminal Gal is a hint that some Ara or acid-labile Rha or Fuc is bound to C-6 of Gal in the native AGP. Comparable with other fern AGPs, Rhap (including 3-O-Me-Rhap) is localized terminally (Bartels & Classen, 2017). After the carboxy reduction of uronic acids to the corresponding deuterium-labeled neutral monosaccharides, only GlcA was detected as a terminal monosaccharide (AGP_{UroX}).

Table 1 Linkage type analysis of AGPs and hydrolyzed AGPs (AGP_{UroX}) from *Salvinia molesta*, *Azolla filiculoides*, and *Ceratopteris richardii* in percentages (mol mol⁻¹, n = 1)

Monosaccharide	Linkage type	<i>S. molesta</i>		<i>A. filiculoides</i>		<i>C. richardii</i>	
		AGP	AGP _{UroX}	AGP	AGP _{UroX}	AGP	AGP _{UroX}
Galp	1,3,6-	30.7	33.8	27.0	30.8	35.9	37.8
	1,6-	2.1	15.8	2.6	11.5	2.2	12.4
	1,3-	15.2	14.4	15.5	18.7	16.4	15.0
	1-	2.1	12.0	3.4	5.1	6.0	9.4
Ara _f	1,5-	3.9	—	4.4	—	7.3	—
	1,2-	17.2	6.7	16.9	10.7	11.5	3.9
	1-	9.6	1.8	8.5	2.6	6.4	1.8
Rhap	1-	11.0	5.9	9.5	6.1	7.8	8.2
Fucp	1-	—	—	1.4	—	2.2	—
GlcAp	1-	ND	3.6	ND	2.9	ND	3.2
Manp	1,6-	2.9	—	2.9	1.2	1.3	2.7
Hexp	1,4-/1,2-	5.3	6.0	7.9	9.2	3.0	5.6

AGP, arabinogalactan protein; ND, not detectable due to the method.

Antibodies directed against epitopes of the glycan part of angiosperm AGPs (see Experimental procedures) were tested for their binding capacities to the different fern AGPs to gain further information on AGP structures (Figure 2d, Figure S2). In the tested concentrations, there was no binding of the fern AGPs to KM1 and LM6, thus supporting structural differences to angiosperm AGPs. For KM1 (Classen et al., 2004), it has been shown that β -D-1,6-Galp is part of the epitope (Ruprecht et al., 2017), which is present in the fern AGPs only in small amounts. LM6 is directed against α -L-1,5-linked Araf (Verhertbruggen et al., 2009) present in arabinans but also in some AGPs. Although this Ara linkage type is present in the investigated fern AGPs, the amounts are lower compared with the 1,2-linkage type. Weak binding of LM6 to *Ceratopteris* AGP was observed after longer times of incubation (data not shown), thus confirming higher amounts of 1,5-Araf in *Ceratopteris* compared with *Salvinia* and *Azolla* AGP. LM2 showed moderate binding to the fern AGPs, probably due to presence of terminal GlcAp, which has been described as part of the epitope (Ruprecht et al., 2017). JIM13 strongly binds to the three fern AGPs, particularly to *Salvinia* AGP. JIM13 is known to recognize many AGPs but the exact epitope is under debate. The trisaccharide β -D-GlcAp (1 \rightarrow 3)- α -D-Galp-(1 \rightarrow 2)- α -L-Rhaf isolated from hydrolysate of gum karaya binds to this antibody (Yates et al., 1996). Strong binding of JIM13 to a rhamnogalactan protein from *Spirogyra* (Pfeifer et al., 2022) further suggests that terminal Rha must be an important part of the epitope of JIM13 and is possibly responsible for binding to the investigated fern AGPs. An exemplary structural proposal for AGP of *Azolla* based on the results above is given in Figure S3.

Twenty-one different species representing the streptophyte lineage (a streptophyte alga, three bryophytes, a lycophyte, seven ferns, three gymnosperms, and six angiosperms) were compared with regard to AGP linkage types (Table 2). Here, we only used data from our working group to ensure a comparable methodology for isolation and structural characterization (see Figure 3 for references). The comparison of AGP fine structures over the streptophyte lineage by principal components analysis (PCA; Figure 3) revealed first hints on lineage-specific structural features of AGPs. The streptophyte alga *Spirogyra praten-sis* and the lycophyte *Lycopodium annotinum* were clearly separated from all other species due to high amounts of terminal and unique 1,3-linked Rhap in *Spirogyra* AGP and unusual terminal Arap accompanied by 1,3-linked Araf in *Lycopodium* AGP. Yet, a broader sampling is needed to clarify how common these features are in the respective lineages, or whether these are species-specific characteristics. PCA reveals clustering of angiosperm AGPs due to the common structure with a Galp backbone in 1,3,6- and 1,6-linkage decorated with 1,5-linked and terminal Araf. In

contrast, gymnosperms linkage types of AGPs seem to be less conserved, although it needs to be considered that only three species from three lineages, *Ginkgo*, *Cycas*, and *Ephedra*, and no representatives from conifers were included in the analyses. For fern AGPs a clear separation to angiosperm AGPs is visible in line with their evolutionary distance. There is an overlap of fern AGPs with those of mosses and gymnosperms. A common feature of AGPs from mosses, ferns, and gymnosperms is the presence of the unusual monosaccharide 3-O-Me-Rhap not observed in any angiosperm AGP or any other angiosperm cell wall polysaccharide. Within the ferns, the only eusporangiate fern AGP from *Equisetum* is separated from the leptosporangiate AGPs. The leptosporangiate *Osmunda* AGP is located in between *Equisetum* AGP and the other leptosporangiate AGPs. This is in accordance with the phylogeny of ferns, where *Osmunda* is sister to the rest of the leptosporangiates. *Azolla* and *Salvinia* are members of the same family (Salviniaceae) and the similarities in their AGPs reflect their close phylogenetic relationship or the similarity of their aquatic lifestyle. Indeed, *Pteridium* and *Ceratopteris* belong to the same family (Pteridaceae), yet *Pteridium* AGPs cluster with those from *Dryopteris*, while those of the aquatic fern *Ceratopteris* is more similar to AGPs from *Salvinia* and *Azolla*.

Genetic framework for AGP glycosylation in ferns

Glycosylation of AGPs is very complex and involves the addition of large AG moieties to Hyp residues often arranged in characteristic dipeptide repeats: Ala-Hyp, Ser-Hyp, and Thr-Hyp. The AGs consist of β -1,3-Galp backbones with β -1,6-linked Galp side chains that are further decorated with mainly Ara but also GlcA, Rha, Fuc, or Xyl (for review, see Showalter & Basu, 2016; Silva et al., 2020; Strasser et al., 2021). The responsible glycosyltransferases (GTs) belong to different GT families that have been classified in the carbohydrate active enzymes (CAZy) database (<http://www.cazy.org>). Although several GTs responsible for AGP glycosylation have been identified in *Arabidopsis*, many enzymes are still unknown, e.g., most arabinosyltransferases and all rhamnosyltransferases (Silva et al., 2020). Responsible for galactosylation of AGPs are enzymes of the family GT31. A phylogeny of the *Arabidopsis thaliana* sequences in this family divides into three distinct groups, which were subdivided into nine subgroups I–IX (group A with clades I–IV, group B with clades V and VI, group C with clades VII–IX; Qu et al., 2008). Eleven of these enzymes belonging to clades II, III, V, and VI have already been characterized and are involved in AGP galactosylation in *A. thaliana* (Silva et al., 2020).

To elucidate whether homologs of these enzymes are present in ferns, *Arabidopsis* sequences from groups A and B were used to search for GT31 sequences in the genomes of *Salvinia*, *Azolla*, and *Ceratopteris* as well as

Table 2 Linkage type analysis of arabinogalactan proteins throughout the streptophyte lineage in percentage (mol mol⁻¹)

	1,3,6-Galp	1,2,3-Galp	1,3-Galp	1,6-Galp	t-Galp	1,2-Araf	1,3-Araf	1,5-Araf	t-Araf	t-Araf	t-Arap	t-Rhap	t-3-O-MeRhap	1,3-Rhap	1,4-/1,2-Hexp	others
Charophyte alga																
<i>Spirogyra pratensis</i>	30.7	0	9.8	5.8	0	0	0	0	1.1	0	24.7	2.0	20.0	0	5.9	
Bryophytes																
<i>Marchantia polymorpha</i>	27.1	0	19.2	0	0	0	4.3	2.7	36.3	0	1.1	1.1	0	4.3	3.9	
<i>Sphagnum</i> sp.	27.5	4.8	17.8	7.3	7.0	0	0.5	5.3	3.9	0	1.9	7.2	0	11.1	5.7	
<i>Polytrichastrum formosum</i>	28.2	6.6	14.5	2.6	14.2	0.6	3.4	1.8	9.2	0	7.4	2.5	0	6.8	2.2	
Lycophyte																
<i>Lycopodium annotinum</i>	29.2	0	19.5	1.1	1.3	0.2	15.2	1.0	11.8	13.4	0	0	0	3.1	4.2	
Ferns																
<i>Equisetum arvense</i>	26.4	0	16.3	1.8	2.4	0.5	0.2	14.3	23.1	0	0	3.4	0	5.2	6.4	
<i>Osmunda regalis</i>	35.5	0	19.7	1.3	3.4	2.9	0.8	2.4	15.9	0	1.1	5.1	0	6.3	5.6	
<i>Salvinia molesta</i>	30.7	0	15.2	2.1	2.1	17.2	0	3.9	9.6	0	5.4	5.6	0	5.3	2.9	
<i>Azolla filiculoides</i>	27.0	0	15.5	2.6	3.4	16.9	0	4.4	8.5	0	9.5	0	0	7.9	4.3	
<i>Ceratopteris richardii</i>	35.9	0	16.4	2.2	6.0	11.5	0	7.3	6.4	0	0	7.8	0	3.0	3.5	
<i>Pteridium aquilinum</i>	29.7	0	20.6	2.2	10.6	9.6	1.7	2.5	4.6	0	0.5	1.3	0	7.6	9.1	
<i>Dryopteris filix-mas</i>	23.7	0	18.5	2.9	9.0	5.0	1.7	2.9	15.2	0	0.7	1.3	0	4.8	14.3	
Gymnosperms																
<i>Cycas revoluta</i>	32.4	0	13.2	3.8	2.1	1.5	7.1	6.9	21.7	3.7	4.4	0	0	1.1	2.1	
<i>Ginkgo biloba</i>	12.6	0	2.8	1.3	2.3	1.0	2.3	13	8.4	1.0	2.1	1.2	0	47.3	4.7	
<i>Ephedra distachya</i>	39.2	0	7.3	4.0	0.9	0.7	0	3.7	27.4	0.5	2.4	8.7	0	5.2	0	
Angiosperms																
<i>Triticum aestivum</i>	37.7	0	19.0	3.0	1.2	0	0	9.0	30.1	0	0	0	0	0	0	
<i>Secale cereale</i>	36.3	0	14.2	6.4	3.3	0	0	7.7	32.1	0	0	0	0	0	0	
<i>Avena sativa</i>	36.8	0	19.6	6.1	3.6	0	0	7.3	26.6	0	0	0	0	0	0	
<i>Baptisia tinctoria</i>	24.2	0	16.7	5.8	3.6	0	1.7	9.4	33.2	1.9	1.7	0	0	1.8	0	
<i>Echinacea pallida</i>	26.9	0	16.5	6.3	2.5	0	0	14.2	29.4	0	0	0	0	0	4.2	
<i>Echinacea purpurea</i>	31.8	0	12.7	15.4	0	0	0	11.7	24.8	0	0	0	0	0	3.6	

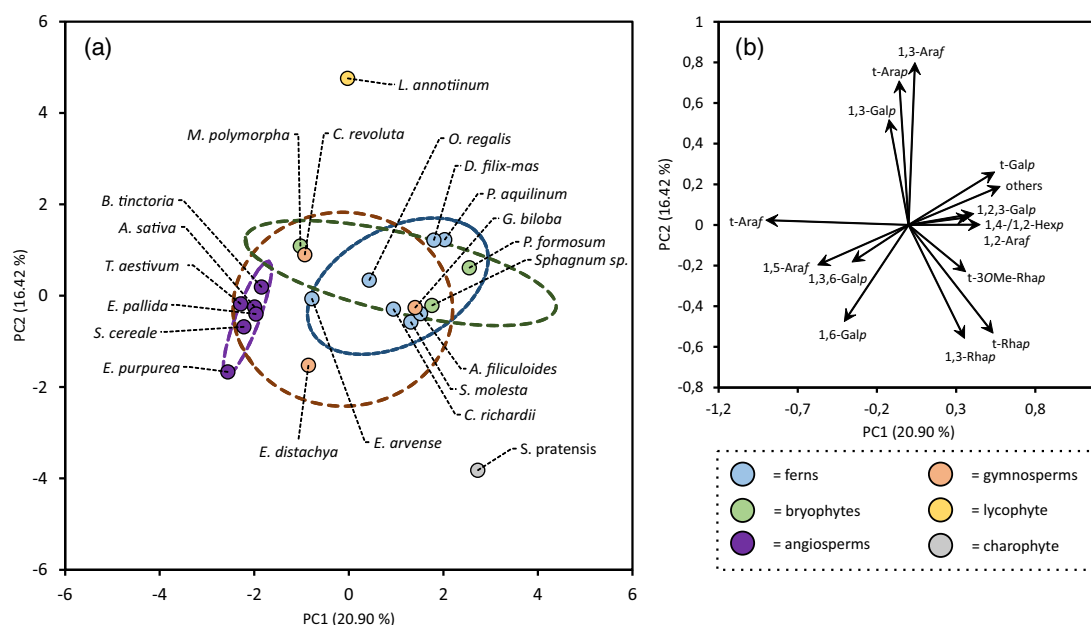


Figure 3. Results of the principal components analysis of linkage-types in arabinogalactan proteins throughout the streptophyte lineage.

(a) Scatter plot of PC1 (20.90%) against PC2 (16.42%). Different colors were used for each evolutionary group. Ellipses represent 90% confidence intervals. For full plant names refer to Bartels and Classen (2017), Bartels et al. (2017), Baumann et al. (2021), Classen et al. (2000), Goellner et al. (2010, 2011, 2013), Happ and Classen (2019), Pfeifer et al. (2022), Thude and Classen (2005), and Wack et al. (2005).

(b) Loading plot for the principal components analysis with influences of the different input variables.

in transcriptomic data from other ferns (Carpenter et al., 2019; Leebens-Mack et al., 2019; Figure 4, Table S5 and Figure S4). The first step of AGP galactosylation is the transfer of the first Gal onto Hyp by a hydroxyproline O- β -galactosyltransferase. In *Arabidopsis*, GALT2-6 (Basu, Tian, et al., 2015; Basu, Wang, et al., 2015) belonging to clade VI as well as HPGT1, HPGT2, and HPGT3 from clade III within GT31 (Ogawa-Ohnishi & Matsubayashi, 2015) have been identified to be responsible for this step of the biosynthesis. Whereas GALT2-6 contain a galectin domain, this domain is absent in HPGT1-3 (Showalter & Basu, 2016). Genomes of *Azolla*, *Salvinia*, and *Ceratopteris* and transcriptomes of all ferns contain at least one corresponding sequence in one of these GT31 families, thus supporting that the transfer of Gal onto Hyp is possible in all fern AGPs.

The *Arabidopsis* genes *At1G77810* (Qu et al., 2008), *At1G33430* (called *UPEX1* or *KNS4*; Suzuki et al., 2017) as well as GALT8 (Narciso et al., 2021) belonging to clade II and possibly CAGE1 and CAGE2 (clade I; Nibbering et al., 2022) encode β -1,3-galactosyltransferases, which function in AGP β -1,3-galactan backbone synthesis. Galactan side chains of AGPs consisting of β -1,6-linked Gal are synthesized by β -1,6-galactosyltransferases of family GALT31A (clade II, Geshi et al., 2013), although in a recent study GALT31A was found to galactosylate substituted and unsubstituted β -1,3-galactan oligosaccharides rather than β -1,6-linked galactan oligosaccharides (Ruprecht et al., 2020).

Homologs to sequences from all *Arabidopsis* enzymes from groups A and B responsible for synthesis of the galactan structure are present in *Azolla*, *Salvinia*, *Ceratopteris*, and all investigated fern transcriptomes. Clades I and II could not be phylogenetically separated (Figure S4; although clade II had also no support as an independent group in Qu et al., 2008, where the groups were originally defined). On balance, the evolution of GT31 group I and II is characterized by lineage-specific diversification (Figure 4): *Arabidopsis* appeared to have slightly more homologs (eight sequences) compared with leptosporangiate (on average 4.8 per species) and eusporangiate ferns (on average 1.8 per species). That said, the investigated fern genomes contain similar numbers of homologs as *Arabidopsis* (eight in *Azolla*, six in *Salvinia*, and 10 in *Ceratopteris*, see Table S5), indicating that the overall lower number of homologs in transcriptomes of ferns is likely an underestimate caused by technical limitations or other reasons, e.g., low expression or differential regulation. To summarize, the genetic toolkit necessary for galactosylation of Hyp and the biosynthesis of the typical AGP galactan scaffold known from seed plants is present in the fern lineage, corroborating our analytical findings (Figure S3). Overall, this suggests that Hyp galactosylation and the synthesis of typical AGP galactan is conserved and was already present in the LCA of ferns and seed plants.

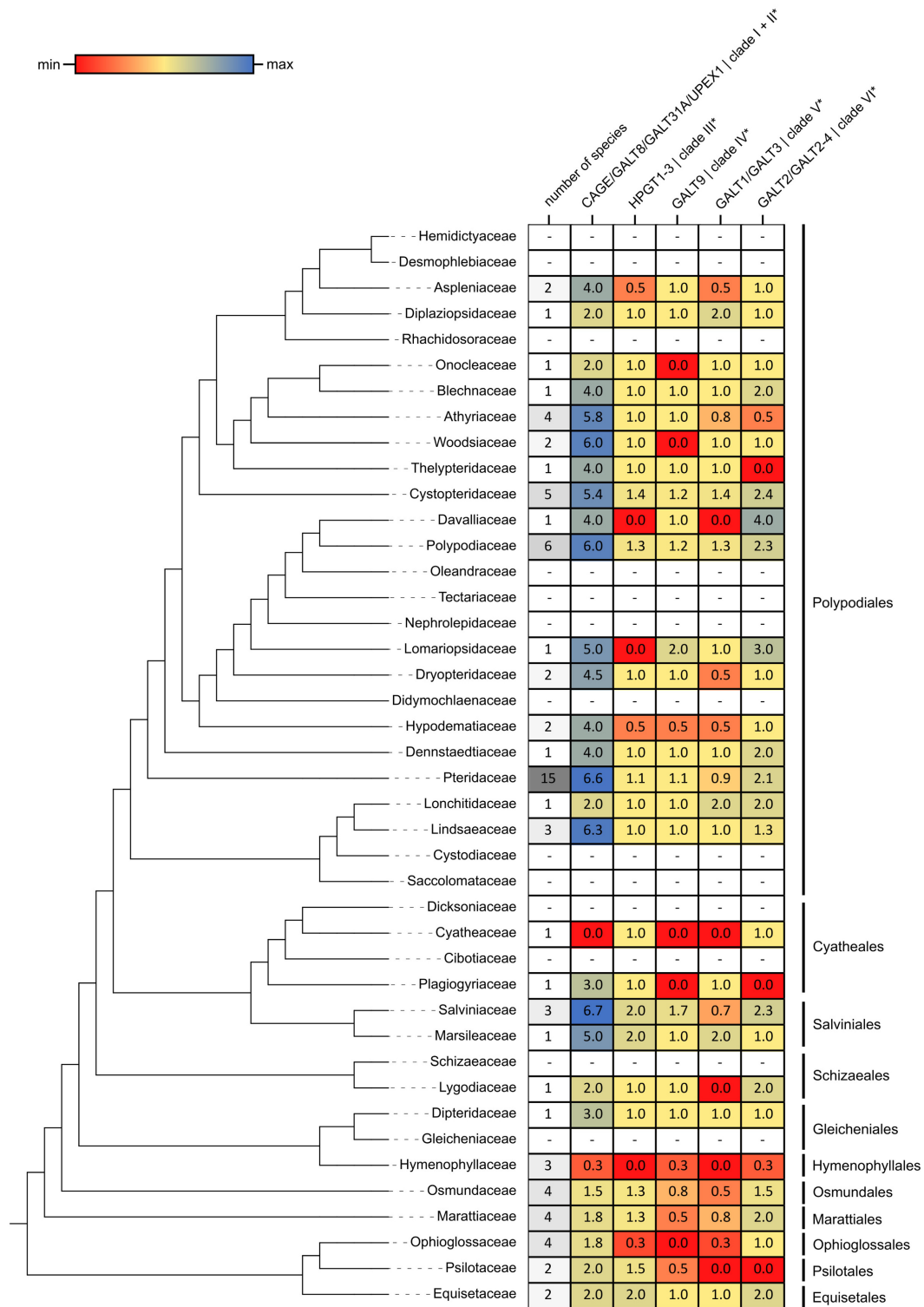


Figure 4. Overview of GT31 homologs from fern transcriptomes and genomes.

Fern phylogeny on the left side is based on the dataset of Nitta et al. (2022) and was visualized with iTOL (v6.6, Letunic & Bork, 2021). Heatmap on the right side is based on the average homolog number in the respective clades of the phylogenetic tree (Figure S4). White fields with dashes were not covered by the dataset. Clade annotations marked with asterisks are used according to Qu et al. (2008).

Identification of AGP and other HRGP protein sequences in fern genomes

Sequences for classical AGPs and other HRGPs (extensins and PRPs) as well as hybrid sequences with motifs from different HRGPs were identified (Figure 5) using the established motif and amino acid bias “MAAB” classification

system (Johnson, Andrew, et al., 2017; Johnson, Cassin, et al., 2017). Class 1 and 4 comprise classical AGPs with and without the glycosylphosphatidylinositol (GPI) anchor respectively, class 2 and 9 are extensins (cross-linking or with GPI anchor) and class 3 contain PRPs. All other classes consist of hybrid HRGP sequences. For the

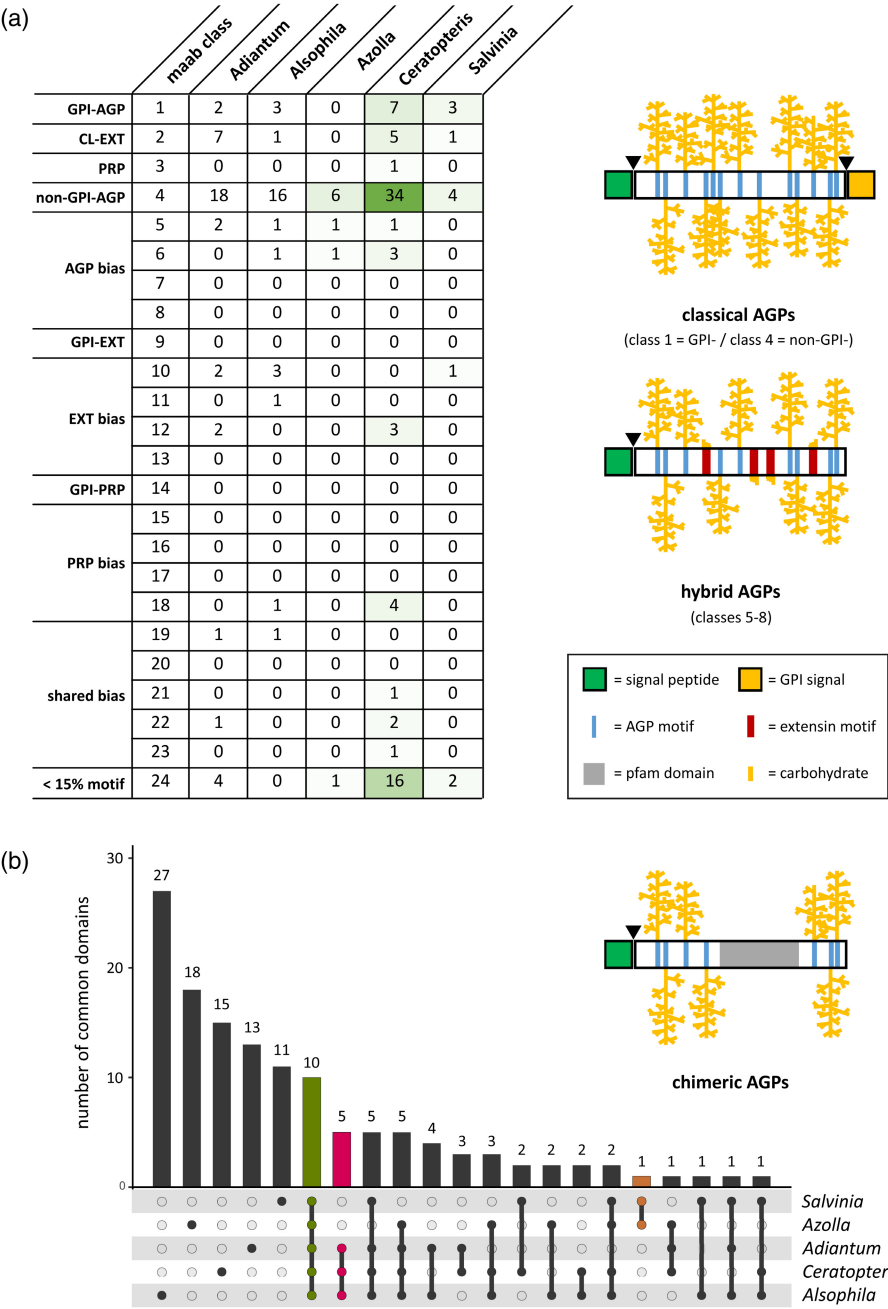


Figure 5. Results of the bioinformatic analyses on arabinogalactan protein (AGP) backbone sequences from the five fern genomes. (a) MAAB classification of all hydroxyproline-rich-glycoproteins within a possible 24 distinct classes. Schematics on the right side highlight the predicted post-translational modifications on classical, hybrid, and chimeric AGPs. CL, crosslinked; EXT, extensins; GPI, glycosylphosphatidylinositol; PRP, proline-rich proteins. (b) Chimeric AGPs are visualized as UpSet plot. Number of common domains of all (green), of all aquatic heterosporous (orange) and all homosporous (magenta) ferns are highlighted. For detailed knowledge of domains, see Figure S5.

heterosporous ferns *Azolla* and *Salvinia*, the AGP sequences identified were similar in number with six for *Azolla* (none with GPI anchor) and seven for *Salvinia* (three with GPI anchor), whereas for homosporous ferns *Adiantum*, *Alsophila*, and *Ceratopteris*, numbers were much higher with 19–41 AGP sequences (two to seven with GPI anchor). The same trend was found for extensins with no sequence identified in *Azolla*, only one sequence in *Salvinia* and higher numbers of sequences in *Adiantum* and *Ceratopteris*. PRPs were absent in all investigated species and only one sequence was detected in the *Ceratopteris* genome. In *Azolla* and *Salvinia* only one or two hybrid HRGPs were present, whereas in the homosporous ferns, numerous sequences were found, particularly in classes 6 (AGP bias), 10 and 12 (extension bias), and 18 (PRP bias). Sequences belonging to class 24 have been detected in all species with the exception of *Alsophila* but have low percentages of HRGP motifs and possibly comprise likely either non-HRGPs or unknown HRGPs (Johnson, Cassin, et al., 2017).

Chimeric AGPs are characterized by AGP motifs, a signal peptide and one or more known functional protein domains. They were identified by screening the translated genome for sequences, which meet the criteria (i) signal peptide (SignalP 5.0), (ii) predicted AG regions, and (iii) at least one detectable protein domain (Pfam). One hundred and eight chimeric AGPs were present in *Adiantum*, 257 in *Alsophila*, 75 in *Azolla*, 184 in *Ceratopteris*, and 71 in *Salvinia* (Figure S5 and Figure 5b).

By evaluation of the intersections between chimeric AGPs from all five species, shared domains were detected (highlighted in Figure S5 and Figure 5b). Ten domains were present in all ferns. Interestingly, they comprise between 47.1% and 54.9% of all chimeric AGPs in ferns (Figure S5). The most dominant members are protein kinase-like AGPs, plastocyanin-like AGPs (Cu bind-like domain, pfam02298), and fasciclin-like AGPs. Another shared group of all ferns are xylogen-like AGPs (containing the LTP-2 domain, pfam14368), which occurs in lower numbers in the small aquatic ferns (each two for *Azolla* and *Salvinia*) but higher numbers in *Adiantum* (three sequences), *Alsophila* (four sequences), and *Ceratopteris* (seven sequences).

DISCUSSION

Polysaccharides of fern cell walls

Plant terrestrialization was accompanied by drought stress and led to diversification of land plants through innovative strategies to reduce dependence on water availability. Structural innovations of tracheophytes include those that enhance transport of water and solutes, as well as taller and turgor-stabilized upright stems for improved spore dispersal and more efficient light capture (Bateman et al., 1998; Eeckhout et al., 2014). Cell wall modifications

are important at all levels of plant growth and development and it is likely that the evolutionary pressures brought about the differences in cell wall composition that occurred during the colonization of land and the emergence of the tracheophytes. With increasing focus on the molecular processes of plant evolution, different groups have contributed to the knowledge of specific cell wall characteristics in ferns using a variety of methods. Often, immunocytochemical workflows (Carafa et al., 2005; Chernova et al., 2020; Eeckhout et al., 2014; Harholt et al., 2012; Leroux et al., 2011; Leroux et al., 2015; Leroux, Leroux, et al., 2013; Popper, 2006) were applied using various monoclonal antibodies against glycan epitopes (Marcus et al., 2008; Marcus et al., 2010; McCartney et al., 2005; Meikle et al., 1994; Pedersen et al., 2012; Puhmann et al., 1994). Furthermore, enzyme-assisted digestion with (partial) characterization of the released oligosaccharides was used and shed light on fern xyloglucans (Hsieh & Harris, 2012; Peña et al., 2008), Man (Franková & Fry, 2011; Popper, 2006; Silva et al., 2011) as well as mixed-linkage glucans (Fry et al., 2008; Popper & Fry, 2003; Sørensen et al., 2008; Xue and Fry, 2012). All these studies fostered the conclusion that fern cell walls share basic features with those of seed plants, such as the occurrence of cellulose, hemicelluloses, and pectic polysaccharides (Leroux, Eeckhout, et al., 2013; Matsunaga et al., 2004; Popper, 2008; Popper & Fry, 2004). Yet they differ from them, for example, in their composition of lignin (Weng et al., 2008), suggesting that lineage-specific cell wall properties have evolved after the split of the fern and seed plant lineages. These lineage-specific differences seem to occur mainly in the fine structure of the cell wall, e.g., through variation of side-chain structures. Our focus on AGPs in ferns using methods ranging from carbo-analytical, immunocytochemical to bioinformatic approaches provide detailed insights into the evolutionary history of AGPs in tracheophytes.

AGPs abound in fern cell walls

Already in the 1970s, the occurrence of AGPs in several leptosporangiate ferns was shown by gel-diffusion assay with Yariv's reagent (Clarke et al., 1978). Monoclonal antibodies directed against arabinogalactan epitopes (e.g., JIM4, JIM8, JIM13, KM1, LM2, LM6, or MAC207; for epitopes and references see Classen et al., 2019) were used to verify the presence of AGPs in combination with glycan microarrays, ELISA, and microscopy to detect AGPs in different monilophytes (Bartels & Classen, 2017; Eeckhout et al., 2014; Leroux et al., 2015; Lopez & Renzaglia, 2014; Lopez & Renzaglia, 2016; Moore et al., 2013; Popper, 2006). However, reliability of these discoveries are sometimes questionable because the exact structures of epitopes of the antibodies are not always known and cross-reactions with other polysaccharides might occur. For example, LM6

is directed against 1,5-linked Araf and binds to arabinans, yet it can also detect AGPs.

Information on the structure of AGPs from seed plants is extensive, while limited data are available for monilophytes and other seedless plants. Such studies are, to the best of our knowledge, currently limited to seven moss (Fu et al., 2007; Geddes & Wilkie, 1971; Happ & Classen, 2019; Kremer et al., 2004; Lee et al., 2005) and four monilophyte species (Bartels & Classen, 2017). To gain further information on fern AGPs and insight into AGP evolution, we isolated AGPs from *Salvinia*, *Azolla*, and *Ceratopteris* by specific precipitation with β GlcY. The amount of *Azolla* AGP was comparable with AGP content of the mosses *Sphagnum* and *Physcomitrium* (Bartels et al., 2017) and slightly higher compared with other monilophytes (in mean 0.12–0.15%; Bartels & Classen, 2017) whereas AGP amounts of *Salvinia* and *Ceratopteris* were slightly lower and comparable with AGP recovery in the moss *Polytrichastrum* (Bartels et al., 2017). However, variation in AGP amounts is observed at different times of harvest (Bartels & Classen, 2017); thus, it cannot be excluded that some of the observed variation may stem from different times of harvest.

With regard to the AGP protein moieties (high protein and low Hyp content), AGPs from the three fern species investigated here are more similar to some moss AGPs (Bartels et al., 2017; Happ & Classen, 2019) than to those from other ferns (Bartels & Classen, 2017). This may reflect similarities in habitat of the mosses and the water ferns we investigated, including characteristics such as high moisture.

The carbohydrate backbones of AGPs mainly consist of linear chains of β 1,3-linked Galp attached to Hyp, and branched with β 1,6-linked Galp side chains (Ma et al., 2018). Our results reveal that AGPs of the three ferns contain this typical galactan core with Galp in 1-, 1,3-, 1,6-, and 1,3,6-linkage known from seed plants and from other fern genera (Bartels & Classen, 2017). A proposal for the structure of the carbohydrate moiety of AGP from *Azolla* is shown in Figure S3. The lower amounts of 1,6-linked Galp, indicating a highly branched structure, may be a special feature of fern AGPs. For example, in AGP from the angiosperm *Echinacea*, 1,6-linked Galp was present in high amounts (Classen et al., 2000). KM1, an antibody directed against 1,6-linked Galp (Ruprecht et al., 2017), has been generated against *Echinacea* AGP (Classen et al., 2004). This antibody showed no reactivity with the fern AGPs tested here (Figure 2d), thus supporting low amounts of 1,6-linked Galp in the three species.

We observed several differences to seed plants with regard to the monosaccharides present at the periphery of the fern AGPs. The unusual monosaccharide 3-O-Me-Rha is present as terminal monosaccharide in *Salvinia*, *Azolla*, and *Ceratopteris* AGPs. 3-O-Me-Rha has been detected in

cell walls of streptophyte algae, bryophytes, lycophytes, ferns, and gymnosperms before (Anderson & Munro, 1969; Matsunaga et al., 2004; Pfeifer et al., 2022; Popper et al., 2004; Popper & Fry, 2003), but has never been found in angiosperm cell walls. In *Lycopodium* and different ferns, 3-O-Me-Rha occurs in rhamnogalacturonan II (Matsunaga et al., 2004). Several studies now show that 3-O-Me-Rha is also part of many moss, fern, and gymnosperm AGPs (Bartels et al., 2017; Bartels & Classen, 2017; Baumann et al., 2021; Fu et al., 2007; Happ & Classen, 2019), but it was not found in *Lycopodium* AGP (Bartels & Classen, 2017).

Furanosidic Ara is present in all typical AGPs in substantial amounts, but linkage types differ. The terminal linkage is the preferred one while Araf in 1,2-, 1,3-, and/or 1,5-linkage occurs less often. AGPs from *Salvinia*, *Azolla*, and *Ceratopteris* had high and comparable amounts of 1,2-linked Araf to other leptosporangiate fern AGPs (except the early diverging species *Osmunda regalis*, Bartels & Classen, 2017). This linkage type is, however, have not been described for angiosperm AGPs so far. It is noteworthy that linkage types can also differ within one species. In AGPs from the medicinal plant, *Echinacea purpurea*, Araf was present in the 1,5-linkage, but secreted AGPs from cell cultures of the same species predominantly formed a 1,3-Araf linkage (Classen, 2007). The latter is not limited only to cell culture AGPs but is also the predominant type in AGPs from *Arabidopsis* leaves (Tryfona et al., 2012) and *Zostera* (Pfeifer et al., 2020).

Structure of fern AGPs in light of evolution

The comparison of AGP fine structures over the streptophyte lineage by PCA revealed first hints on lineage-specific structural features of AGPs. Generally, angiosperms are most likely more stable regarding the type of side-chain linkages while the other major (land) plant lineages show more flexibility in this direction. The set of angiosperms included in our PCA is composed of members of the monocot order of Poales as well as the two dicot orders, Asterales and Fabales. While this dataset does not cover all of angiosperm diversity, the inclusion of monocots and dicots covers the entire evolutionary distance of angiosperms; thus, ruling out that the clustering of the PCA can stem from a sampling bias. Clustering of the aquatic ferns *Ceratopteris*, *Salvinia*, and *Azolla* seems to correlate with their aquatic habitat. Overall, the data suggest that both a close phylogenetic relationship and a very divergent lifestyle can drive the evolution of AGP linkage types.

AGP GTs, AGP as well as other HRGP proteins have diversified in ferns

Transfer of the first Gal on to Hyp as well as biosynthesis of the AGP galactan consisting of a β -1,3-linked Gal with

β -1,6-linked Gal side chains is performed by members of GT31. We find homologs of *Arabidopsis* Hyp-O-galactosyltransferases, β -1,3-galactosyltransferases, and a β -1,6-galactosyltransferase in the genomes of the three investigated ferns and the transcriptomes of other leptosporangiate and eusporangiate ferns. These data agree with our analytical investigations on fern AGP structure and support that the seed plant-typical AGP galactosylation is present in ferns, and was likely present in the LCA of the two lineages.

The occurrence of terminal Rha, partially present as 3-O-MeRha, is the most striking difference of fern to seed plant AGPs. Up to now, the only rhamnosyltransferase characterized with regard to cell wall biosynthesis is involved in pectic rhamnogalacturonan I synthesis (Takanaka et al., 2018). Future characterizations of rhamnosyltransferases acting on AGPs will be illuminating.

Protein backbones of classical AGPs consist of an N-terminal signal peptide, a protein sequence rich in Pro, Ala, Ser, and Thr residues and may have an additional C-terminal GPI anchor sequence (Schultz et al., 2000). *In silico* analyses revealed several homologs to classical AGPs, extensins, and hybrid HRGPs in ferns such as *Adiantum*, *Alsophila*, and *Ceratopteris* in similar amounts as found in seed plants and slightly lower numbers for the ferns *Azolla* and *Salvinia* (Johnson, Andrew, et al., 2017; Johnson, Cassin, et al., 2017; Ma et al., 2017). The latter two have small genomes compared with other ferns (Li et al., 2018), which could explain our results. In all ferns, the chimeric AGPs comprise mainly protein kinase-like, phytocyanin-like, fasciclin-like, and xylogen-like AGPs, which are well described chimeric AGPs in seed plants (Basu et al., 2016; Costa et al., 2019; He et al., 2019; Johnson et al., 2003; Ma et al., 2022; Shafee et al., 2020). Conservation of AG-attachment sites close to these protein domains underlines the necessity of glycosylation for the correct function of the respective domains (Dragičević et al., 2020). The group of xylogen-like AGPs is of interest with regard to the evolution of vascular tissues (Kobayashi et al., 2011; Motose et al., 2004). The characteristic pfam domain is among the top 10 protein domains conserved in all investigated fern genomes. However, the domain was less abundant in the genomes of the heterosporous ferns, *Azolla* and *Salvinia*. Such universal occurrence across fern species hints towards a fundamental contribution of AGPs to the process of vascular tissue differentiation (see below); lower numbers in some of the water ferns might be connected to their aqueous habitat. While AGP protein sequence numbers of ferns and their galactosylation are similar to that of seed plants, their oligosaccharide structures at the periphery of the AGP molecules differ between ferns and seed plants. This might result in deviating functions and explain lineage-specific differences in cell wall structures of ferns and seed plants.

On possible functions of AGPs in ferns

Functional roles of AGPs in seed plants among others include cell division, growth, programmed cell death, pattern formation, and plant-microbe interactions (for reviews see Ellis et al., 2010; Hromadová et al., 2021; Ma et al., 2018; Seifert & Roberts, 2007). The interaction of GPI-anchored AGPs with plasma membrane-bound transmembrane proteins of the same cell or soluble AGPs with receptors in adjacent cells have been proposed (Ellis et al., 2010). Signaling caused by AGPs might also occur by oligosaccharides enzymatically cleaved from the AG glycan moieties or via Ca^{2+} ions (Lampert et al., 2014). Knowledge on the function of fern AGPs is limited, but their roles in gamete and embryo development (Lopez & Renzaglia, 2014, 2016) and desiccation tolerance (Moore et al., 2013) have been described previously. In *Ceratopteris richardii*, eggs and multiflagellated sperm cells (Lopez & Renzaglia, 2014, 2016) are covered by an extraprotoplasmic AGP matrix. Given that seed plant AGPs bind and release Ca^{2+} at the plasmalemma (Lampert & Varnai, 2013), Lopez and Renzaglia (2016) proposed that AGPs from *C. richardii* could facilitate gamete fusion through Ca^{2+} oscillations. These data together with the well-documented role of seed plant AGPs in these processes (for review see Leszczuk et al., 2019; Ma et al., 2018) suggest that AGPs have likely facilitated reproduction and embryogenesis already in the LCA of vascular plants. Comparison of hydrated and desiccated leaf material of the resurrection fern *Mohria caffrorum* revealed that Ara-rich polysaccharides (arabinan-rich pectins and AGPs) protect cell walls against desiccation (Moore et al., 2013). The Ara-containing polysaccharides act as plasticizers ensuring the maintenance of hydration during drought stress. When the first land plants evolved from an aquatic ancestor, adaptation to water deficiency was a big challenge; the evolution of different kinds of water-transporting mechanisms was likely key for land plants to establish themselves on land and diversify (Woudenberg et al., 2022). Involvement of AGPs in water transport has been suggested for many lineages across the land plant tree of life: antibodies directed against AGP epitopes labeled water-conducting cells of different liverworts and mosses (Ligrone et al., 2002). In addition, xylogen-type genes are present in the moss *Physcomitrium* as well as the lycophyte *Selaginella* (Kobayashi et al., 2011). The AGP “xylogen” induces xylem differentiation in suspension cultures of the angiosperm *Zinnia* (Motose et al., 2004). In *Arabidopsis* AGP epitopes mark the initial cell that gives rise to protoxylem and metaxylem (Dolan et al., 1995). In roots, stems, and leaf stalks of the angiosperm *E. purpurea*, labeling of AGPs is most dominant in the area of the xylem (Bossy et al., 2009; Goellner et al., 2013). The recurrent use of AGPs in water transporting tissues in bryophytes, lycophytes, and

angiosperms suggest that this function is also conserved in other lineages, including ferns; more data are however needed to support their involvement in development of water-transporting tissues in ferns.

Because of similarities in overall AGP structure, some functions of AGPs might be conserved in bryophytes and tracheophytes. However, differences in fine structure, particularly the arabinogalactan moiety, can also cause functional differences. Here we show that *Salvinia*, *Azolla*, and *Ceratopteris* reveal high content of Rha and methylated Rha at the periphery of their AGPs. This feature is similar to the structure of AGPs of the bryophytes *Physcomitrium*, *Sphagnum*, and *Polytrichastrum* (Bartels et al., 2017; Fu et al., 2007). The cell walls of the streptophyte algae *Spirogyra* contain rhamnogalactan proteins that are structurally similar to AGPs with nearly complete replacement of Ara by Rha (Pfeifer et al., 2022). Rha and methylated Rha are less polar compared with Ara. This suggests a change from a less hydrophilic AGP surface with Rha and methylated Rha as terminal groups to a more hydrophilic surface dominated by terminal Ara in vascular plants. Given the less hydrophilic features of AGPs in algae, bryophytes, and the water ferns it is logical to assume that these AGPs have alternative functions with regard to hydrophobic interactions (Fu et al., 2007). Arabinosylation of existing cell wall polymers has been proposed to be an evolutionary strategy to prevent polymer aggregation during water loss (Moore et al., 2013). If true, one may assume that the challenges of a life on land have selected for more hydrophilic AGPs with higher amounts of Ara, allowing for desiccation tolerance in tracheophytes. The modifications observed here for the aquatic ferns may be an adaptation to their reversion to aquatic life.

CONCLUSION

During plant terrestrialization, plants acquired molecular adaptations to cope with the terrestrial environment. Tracheophyte innovations such as xylem and stabilized upright stems needed modifications in cell wall composition, either through elaboration of ancestral polymers or through synthesis of new components. Our investigations show that AGPs, which are important cell wall components with signaling functions, are present in the three fern species investigated and their basic molecular features are similar to those known from seed plants.

Phylogenetic analysis of relevant GT31 enzymes confirmed that AGP protein backbones and homologs of the basic enzymes responsible for biosynthesis of the galactan core of AGPs are present in ferns likely underpinning the conservation of AGP structures in tracheophytes. An unusual attribute of the three fern AGPs was the presence of 1,2-linked Araf in high amounts, which has also been detected in other leptosporangiate fern AGPs, but not in AGPs from other taxa. Interestingly, fern AGPs share a

special feature with moss and also gymnosperm AGPs, the occurrence of terminal 3-O-MeRha not present in angiosperms. Identification of enzymes that synthesize AG glycans in non-flowering plants, particularly rhamnosyl-transferases, is a challenge for the future.

More phylodiverse fern genomes are needed to identify AGP protein backbones and GTs involved in polysaccharide biosynthesis and AGP glycosylation in this lineage of tracheophytes. Future studies should also include detailed structural analyses of AGPs from higher numbers of bryophytes and monilophytes. Generation of AGP mutants, e.g., by CRISPR/Cas9 technique, will help to gather new information on AGP functions in the streptophyte lineage. As AGPs are possibly involved in the water balance of plants, they might be interesting targets to modulate desiccation tolerance of plants in times of climate change.

EXPERIMENTAL PROCEDURES

Plant material, growth conditions, and sampling

The ferns *Salvinia* × *molesta* D. S. Mitchell (*Salvinia*, *S. molesta*), *Salvinia auriculata* Aubl., *S. cucullata* Roxb., *Salvinia natans* (L.) All., *A. filiculoides* Lam. (*Azolla*), and *C. richardii* Brongn. (*Ceratopteris*) were cultivated in the greenhouse in the Garden of the Pharmaceutical Institute of the Christian-Albrechts-University of Kiel between August and October 2018 (*Salvinia*), September and January 2019/2020 (*Azolla*), and July and August 2021 (*Ceratopteris*); see Figure 1. Spores of *Ceratopteris* RN3 strain were kind gifts of Péter Szövényi (University of Zurich). For spore germination, gametophyte development and sporophyte culture on C-fern medium, the protocol of Plackett et al. (2015) was used. Plant material was cleaned with water and freeze-dried.

Isolation of AGPs

The freeze-dried and ground plant material was extracted two times (2 and 21 h) with 70% acetone solution (v/v) in a ratio of 1:20 (w/v) for *A. filiculoides* and *C. richardii* and in a ratio of 1:30 (w/v) for *S. molesta* under constant stirring at 4°C to remove phenolic compounds. Afterwards, the air-dried plant residue was extracted with double-distilled water (ddH₂O) in a ratio of 1:20 (w/v) for *A. filiculoides* and *C. richardii* and in a ratio of 1:30 (w/v) for *S. molesta*. After constant stirring for 21 h at 4°C, the AE was separated through a tincture press from the remaining plant material.

The AE was heated in a water bath at 90–95°C for 10 min to denature proteins. These proteins were removed by centrifugation (4122 g, 20 min, 4°C), and the extract was concentrated with a rotary evaporator at 40°C and 0.010 mbar to approximately one-tenth of its volume. To precipitate polysaccharides and AGPs, the AE was poured into ethanol at 4°C at a ratio of 1:4 (v/v) and stored overnight at 4°C. The precipitated material was isolated by centrifugation (19 000 g, 4°C, 20 min) and freeze-dried (AE).

Following the procedure of Classen et al. (2005), AGP was isolated from the AE with the β-Glc-Yariv reagent (βGlcY). After dissolving AE in ddH₂O and βGlcY (1 mg ml^{−1}) in an equal volume of sodium chloride solution (0.3 mol L^{−1}), the AGP was precipitated overnight at 4°C after addition of the βGlcY solution. The AGP-βGlcY-complex was separated by centrifugation (19 000 g, 4°C, 20 min) and re-dissolved in ddH₂O. After heating to 50°C,

sodium hydrosulfite was added until the red color disappeared. To purify the aqueous AGP solution, it was dialyzed against demineralized water for 4 days at 4°C (molecular weight cutoffs 12–14 kDa) and then freeze-dried.

Analysis of monosaccharides

For determination of the neutral monosaccharide composition, a modified method of Blakeney et al. (1983) was used. To 1–10 mg of sample, 1.0 ml trifluoroacetic acid (2 mol L⁻¹) and 50 µl internal standard *myo*-inositol (10 g L⁻¹) were added, incubated for 1 h at 121°C, washed three times with 5 ml ddH₂O, and then evaporated. The samples were reduced with 200 µl ammonium hydroxide (1 mol L⁻¹) and 1.0 ml sodium borohydride in dimethyl sulfoxide (1 g 50 ml⁻¹). After incubation for 1.5 h at 40°C the reaction was stopped with 100 µl glacial acetic acid. For the acetylation, 200 µl of 1-methylimidazole and 2.0 ml of acetic anhydride were added, incubated for 20 min at room temperature and stopped by addition of 10 ml ddH₂O. The samples were acidified with 1.0 ml aqueous sulfuric acid (0.1 mol L⁻¹) and liquid–liquid extraction was performed with an equivalent volume of dichloromethane. The identification and quantification of the neutral monosaccharides was performed by gas chromatography with flame ionization detection and mass spectrometry detection: gas chromatography + flame ionization detection: 7890B; Agilent Technologies, Santa Clara, CA, USA; MS: 5977B mass spectrometry detection; Agilent Technologies; column: Optima-225; Macherey-Nagel, Düren, Germany; 25 m, 250 µm, 0.25 µm; helium flow rate: 1 ml min⁻¹; split ratio 30:1. A temperature gradient was used to achieve peak separation (initial temperature 200°C, subsequent holding time of 3 min; final temperature 243°C with a gradient of 2°C min⁻¹). 3-*O*-MeRha was identified by retention time and mass spectrum (see Happ & Classen, 2019).

According to a modified method of Blumenkrantz and Asboe-Hansen (1973), the uronic acid amount was determined photometrically using a linear calibration generated with a mixture of glucuronic and galacturonic acid (1:1, w/w). Two hundred µl of a 500 µg L⁻¹ sample dissolved in sulfuric acid (4%, v/v) was mixed with 1.2 ml of sodium tetraborate in sulfuric acid (75 mmol L⁻¹) and incubated at 100°C in a water bath for 20 min. After cooling for another 10 min on ice, 20 µl of a solution of *meta*-hydroxydiphenyl (0.15%, w/v) in NaOH (0.5%, w/v) was added to the sample. After incubation for 15 min at room temperature, the absorbance was determined at 525 nm. For each sample, a threefold determination was performed and the absorbance of an additional blank sample, containing NaOH (0.5%, w/v) instead of the color reagent, was subtracted.

Reduction of uronic acids and partial degradation of AGPs

The AGP fraction was first subjected to uronic acid reduction followed by oxalic acid partial hydrolysis. A modified method of Taylor and Conrad (1972) was used to perform a carboxy-reduction of the uronic acids. After dissolving 20–30 mg AGP in 20 ml ddH₂O, 216 mg of *N*-cyclohexyl-*N*-[2-(*N*-methylmorpholino)ethyl]-carbodiimide-4-toluolsulfonate was added slowly under constant stirring until it was completely dissolved. An autotitrator adjusted the pH to 4.75 with 0.01 M HCl and maintained it for another 2 h. The uronic acids were then reduced by the dropwise addition of sodium borodeuteride solutions in increasing concentrations (2.0 ml of 1 mol L⁻¹; 2.5 ml of 2 mol L⁻¹; 2.5 ml of 4 mol L⁻¹). To avoid strong foaming during the reaction, one to two drops of 1-octanol were added beforehand. In addition, the pH value was set to 7.00 with 2 M HCl by an autotitrator and maintained for another 2 h after the reduction agent was fully added.

Finally, the pH was adjusted to 6.5 with glacial acetic acid, the solution was dialyzed for 3 days at 4°C against demineralized water (molecular weight cutoffs 12–14 kDa) and freeze-dried.

For the oxalic acid hydrolysis, a modified method of Gleeson and Clarke (1979) was used. Ten to 15 mg of the uronic acid reduced sample was dissolved in 2.0 ml of oxalic acid (12.5 mmol L⁻¹) and incubated for 5 h at 100°C. After cooling to room temperature, the hydrolysate was precipitated in ethanol at a final concentration of 80% (v/v). The precipitate was separated by centrifugation (19 000 g, 10 min, 4°C), washed twice with 80% ethanol (v/v), dissolved in 2 ml ddH₂O, and then freeze-dried.

Structural characterization of arabinogalactan moiety of the AGP

Structure elucidation of AGPs was performed by a modified method according to Harris et al. (1984) with potassium methylsulfanyl carbanion (KCA) and iodomethane (IM) in dimethyl sulfoxide. The sample (1–5 mg freeze-dried AGP) was treated stepwise with the KCA and IM (1. 100 µl KCA for 10 min; 2. 80 µl IM for 5 min; 3. 200 µl KCA for 30 min; 4. 150 µl IM for 30 min; 5. 500 µl KCA for 30 min; 6. 400 µl IM for 60 min). The permethylated sample was dissolved in 3.0 ml dichloromethane/methanol in a ratio of 2:1 (v/v), washed three times with 2.0 ml water and centrifuged (5 min, 2500 g). Two ml of 2,2-dimethoxy propane and 20 µl glacial acid were added to the sample, heated to 90°C, and then dried under a stream of nitrogen. The samples were hydrolyzed with trifluoroacetic acid (2.0 mol L⁻¹) at 121°C for 1 h, mixed with 2.0 ml water, and then freeze-dried. Afterwards, the partially methylated monosaccharides were reduced by addition of sodiumborohydride (0.5 mol L⁻¹), dissolved in ammonium hydroxide (2.0 mol L⁻¹), incubated at 60°C for 1 h, and the reaction stopped by 0.5 ml acetone. Acetylation was performed by addition of 200 µl glacial acid, 1.0 ml ethyl acetate, 3.0 ml acetic anhydride and 100 µl perchloric acid and stopped by demineralized water and 100 µl 1-methylimidazole. The permethylated alditol acetates were extracted with dichloromethane and separated and detected by gas–liquid chromatography–mass spectrometry as described above (instrumentation see above: “Analysis of monosaccharides”; column: Optima-1701, 25 m, 250 µm, 0.25 µm; helium flow rate: 1 ml min⁻¹; initial temperature: 170°C; hold time 2 min; rate 1°C min⁻¹ until 210°C was reached; rate: 30°C min⁻¹ until 250°C was reached; final hold time 10 min).

PCA

Based on the resulting biochemical data of the methylation analyses, a PCA was performed with the mean values after standardization by using SigmaPlot (build 14.5.0.101) with default settings. The proportion of the terminal Rha (unmethylated) to the terminal 3-*O*-Me-Rha was calculated from the compositional data of the acetylation analyses. See Figure 1 for full plant names and phylogenetic position as well as Figure 3 for references.

Determination of the protein moiety and the Hyp content

According to the method of Kjeldahl (1883), the protein content was calculated from the nitrogen content (factor 6.25). The nitrogen content was measured by elemental analysis in the Chemistry Department of CAU Kiel University, Kiel, Germany (HEKAtech CHNS Analyzer).

For the quantification of Hyp, the methodology of Stegemann & Stalder (1967) was used with slight modifications. AGPs (10 mg ml⁻¹) were hydrolyzed in a 6 M HCl solution at 110°C for 22 h. After cooling to room temperature, the hydrolysate was

centrifuged at 25 000 *g* for 10 min. One hundred and sixty μ l of the supernatant were diluted with 3840 μ l ddH₂O and 0.6 ml of this dilution was measured in triplicate, as well as a blank value. The blank value was mixed with 0.3 ml of oxidation buffer (pH 6.8; 2.6 g citric acid monohydrate; 7.8 g sodium acetate anhydrous; 1.4 g sodium hydroxide; 25 ml 1-propanol filled to 100 ml ddH₂O) and the samples with 0.3 ml of an oxidating solution (210 mg chloramine T in 15 ml oxidation buffer pH 6.8) and incubated for 20 min at room temperature in the dark. After addition of 0.3 ml color reagent (1.5 g 4-dimethylaminobenzaldehyde; 5.25 ml perchloric acid, 60% (v/v); 9.25 ml 2-propanol) to all samples, they were incubated for 15 min in a water bath at 60°C. After cooling under running tap water for 3 min, incubation followed for 30 min at room temperature, before the absorbance was measured at 558 nm with an ultraviolet/Visual-spectrophotometer. To determine the Hyp content, a linear regression analysis was performed, using 4-hydroxy-L-proline as a standard.

Indirect ELISA

Ninety-six-well plates were coated with 100 μ l per well of the sample in the concentrations, 12.5, 25, and 50 μ g ml⁻¹, in ddH₂O and incubated at 37.5°C for 3 days with open cover. After washing the plates three times with 100 μ l phosphate-buffered saline (PBS)-T (pH 7.4, 0.05% Tween® 20) per well, they were blocked with 200 μ l of bovine serum albumin (1% w/v) in PBS per well and incubated again at 37.5°C for 1 h. The plates were again washed three times with 100 μ l of PBS-T per well. After addition of 100 μ l primary antibody solution (JIM13, KM1, LM2, LM6, LM10, LM15, LM19, LM20, LM25) per well in 1:20 (v/v) PBS 7.4 dilution, the plates were incubated for 1 h at 37.5°C and washed again three times with 100 μ l PBS-T per well. This was repeated with the secondary antibody (antimouse-IgG (KM1) or antirat-IgG (others) conjugated with alkaline phosphatase, produced in goat; Sigma-Aldrich Chemie GmbH, Taufkirchen, Germany) in a dilution of 1:500 (v/v) in PBS 7.4. Finally, 100 μ l of the substrate *p*-nitro-phenylphosphate was added per well and the plates were incubated in the dark at room temperature. The absorbance was measured at 405 nm with a plate reader after 10 min. The samples were analyzed in triplicate. For visualization, the absorbance of the control was set 0 and the highest signal in the dataset was set 1.0. Epitopes of the antibodies and key references are listed in Table S6.

Gel diffusion assay

In an agarose gel (Tris-HCl, 10 mmol L⁻¹; CaCl₂, 1 mmol L⁻¹; NaCl, 0.9% w/v; agarose, 1% w/v), several cavities were stamped. The cavities in the first row were filled with dilutions (100 mg ml⁻¹) of each sample and one of them with the positive control, AGP from *E. purpurea* (10 mg ml⁻¹). In the second line, the cavities were filled with the β GlcY solution (1 mg ml⁻¹). After incubation in the dark at room temperature for 22 h, a red precipitation line appears if the sample contains AGPs.

Phylogenetic analysis of GT31

Putative GT31 homologs of ferns were identified by using the BLASTp interface (E-value of 0.01) at the 1KP database (Carpenter et al., 2019; Leebens-Mack et al., 2019) webserver as well as blasting the translated genomes of *A. filiculoides* (v1.1), *C. richardii* (v2.1), and *S. cucullata* (v1.2) (Table S7). AGP-relevant GT31 protein sequences from group A and B of *A. thaliana* were used as query sequences against all accessible leptosporangiate and eusporangiate ferns (accessed at June 23, 2022). The resulting candidate sequences as well as the query sequences were aligned with L-INS-I using MAFFT (version 7.490; Katoh & Standley, 2013).

All sequences with a minimum of 200 amino acids were included in the dataset. The phylogenetic tree was constructed by using IQ-Tree using model JTT + I + G4 identified by model finder in-built in IQ-TREE (multicore version 1.5.5; Letunic & Bork, 2021). One hundred bootstrap replicates were performed.

AGP backbone sequence identification and classification

The annotated protein sequences from the ferns *A. filiculoides* (v1.1), *Adiantum capillus-veneris*, *Alsophila spinulosa* (version 5), *C. richardii* (v2.1), and *S. cucullata* (v1.2) were downloaded from the respective web resources (Table S7) and imported into the R environment. By using the R-package “ragp” (Dragičević et al., 2020), the N-terminal signal sequences were identified and the maab-pipeline (Johnson, Andrew, et al., 2017) as implemented in the R-package was used to classify the proteins based on their motifs and amino acid biases. All small (<90 amino acids), highly similar (≥ 0.95 similarity) and chimeric sequences (containing any Pfam domain, E-value <1e-5) were excluded from the classification. To differentiate the classes 1 and 4 the bigPI algorithm (Eisenhaber et al., 2003) was used to predict GPI-anchor signals.

Chimeric AGP sequences were identified by detecting (STAGVIP) by using the implementation in the “ragp”-package with the strict mode (at least four dipeptides in a maximal distance of four amino acids) used by Paunović et al. (2021). All dipeptides in an annotated pfam domain (by querying the Pfam database via the CDD webtool, Lu et al., 2020) or in the signal sequence were not count. The candidate sequences were length-filtered (≥ 90 amino acids) and highly-similar sequences (> 0.95 similarity) were discarded.

For each translated genome, the chimeric AGPs were grouped by their largest domain and the number of sequences was plotted as bars in Figure S5. An analysis for common domains in the different ferns was performed and visualized by an UpSet plot (Lex et al., 2014) with the R-package “UpSetR” (Conway et al., 2017).

AUTHORS CONTRIBUTIONS

BC and LP planned and designed the research. KM, LP, and LS cultivated the plant material, in case of *Ceratopteris* from spores delivered by PS. KM, LP, and LS performed the extractions as well as carbohydrate and ELISA experiments. LP, SdV, JdV, and KJ performed bioinformatic searches. LP created all main text figures. All authors analyzed the data (BC, KM, LS, and LP: carbohydrate analysis; LP, SdV, and JdV: bioinformatics search for glycosyltransferases; LP and KJ: bioinformatics search for AGP protein backbones). BC wrote the draft manuscript with help of KM and LP; all authors revised the manuscript, and read and approved the final manuscript.

ACKNOWLEDGEMENTS

LP and BC (project-number 440046237; CL448/3-1), JdV (project-number 440231723; VR132/4-1), and PSZ (project-number 440370236; PSJ1111/1) are grateful for funding within the framework of MAdLand (<http://madland.science>), priority program 2237 of the German Research Foundation (DFG). JdV further thanks the European Research Council for funding under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 852725; ERC-StG “TerreStriAL”). Additional funding was received from the Swiss National Science Foundation (grant nos 160004, 184826, and 212509 to PS); project funding

through the University Research Priority Program “Evolution in Action” of the University of Zurich to PS; a Georges and Antoine Claraz Foundation grant to PS. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests.

DATA AVAILABILITY STATEMENT

All relevant experimental data can be found within the manuscript and its supporting materials. Datasets for bio-informatic analyses can be found here: <https://figshare.com/s/50d87902e1fba3ad98c2>

SUPPORTING INFORMATION

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

Table S1. Yields of water-soluble polysaccharides (AE) from *S. molesta*, *A. filiculoides*, and *C. richardii* in percentage of dry plant material (w/w).

Table S2. Neutral monosaccharide composition of water-soluble polysaccharides from *A. filiculoides*, *S. molesta*, and *C. richardii* in percentage (mol mol⁻¹).

Table S3. Colorimetric determination of the content of uronic acids in the water-soluble polysaccharides and the AGPs of *S. molesta*, *A. filiculoides*, and *C. richardii* in percentage (w/w).

Table S4. Neutral monosaccharide composition of AGPs and partially hydrolyzed AGPs_{UroX} from *S. molesta*, *A. filiculoides*, and *C. richardii* in percentage (mol mol⁻¹).

Table S5. Galactosyltransferase sequences of family GT31 present in genomes (*Azolla*, *Salvinia*, *Ceratopteris*) compared with transcriptomes of other ferns.

Table S6. Antibodies tested for binding to fern cell wall AGPs.

Table S7. Accessed resources for the analysis of translated fern genomes.

Figure S1. FID chromatograms of *Salvinia* AGPs after GLC analysis.

Figure S2. Results of the ELISA experiments with the fern AGPs in comparison with the positive control *Echinacea purpurea* AGP.

Figure S3. Proposed structure for the carbohydrate moiety of AGPs from *Azolla filiculoides*. The proposal was derived from the compositional, immunocytochemical, and linkage-type analyses of AGP and hydrolyzed AGPs.

Figure S4. Phylogeny of GT31 enzymes throughout the different groups of ferns.

Figure S5. Chimeric AGPs of all investigated fern genomes ranked by decreasing numbers.

REFERENCES

- Akiyama, T., Tanaka, K. & Yamamoto, S. (1987) An arabinogalactan-rich protein containing 3-O-methylrhamnose (Acofiose) in young plants of *Osmunda japonica*. *Agricultural and Biological Chemistry*, **51**, 2599–2600.
- Anderson, D.M.W. & Munro, A.C. (1969) The presence of 3-O-Methylrhamnose in *araucaria* resinous exudates. *Phytochemistry*, **8**, 633–634.
- Banks, J.A., Nishiyama, T., Hasebe, M., Bowman, J.L., Gribskov, M., de Pamphilis, C. et al. (2011) The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science*, **332**, 960–963.
- Bartels, D., Baumann, A., Maeder, M., Geske, T., Heise, E.M., von Schwartzenberg, K. et al. (2017) Evolution of plant cell wall: arabinogalactan-proteins from three moss genera show structural differences compared to seed plants. *Carbohydrate Polymers*, **163**, 227–235.
- Bartels, D. & Classen, B. (2017) Structural investigations on arabinogalactan-proteins from a lycophyte and different monilophytes (ferns) in the evolutionary context. *Carbohydrate Polymers*, **172**, 342–351.
- Basu, D., Tian, L., Debrosse, T., Poirier, E., Emch, K., Herock, H. et al. (2016) Glycosylation of a fasciclin-like arabinogalactan-protein (SOS5) mediates root growth and seed mucilage adherence via a cell wall receptor-like kinase (FEI1/FEI2) pathway in *Arabidopsis*. *PLoS One*, **11**, e0145092. Available from: <https://doi.org/10.1371/journal.pone.0145092>
- Basu, D., Tian, L., Wang, W., Bobbs, S., Herock, H., Travers, A. et al. (2015) A small multigene hydroxyproline-O-galactosyltransferase family functions in arabinogalactan-protein glycosylation, growth and development in *Arabidopsis*. *BMC Plant Biology*, **15**, 295.
- Basu, D., Wang, W., Ma, S., DeBrosse, T., Poirier, E., Emch, K. et al. (2015) Two hydroxyproline galactosyltransferases, GALT5 and GALT2, function in arabinogalactan-protein glycosylation, growth and development in *Arabidopsis*. *PLoS One*, **10**, e0125624. Available from: <https://doi.org/10.1371/journal.pone.0125624>
- Bateman, R.M., Crane, P.R., DiMichele, W.A., Kenrick, P.R., Rowe, N.P., Speck, T. et al. (1998) Early evolution of land plants: phylogeny, physiology, and ecology of the primary terrestrial radiation. *Annual Review of Ecology and Systematics*, **29**, 263–292.
- Baumann, A., Pfeifer, L. & Classen, B. (2021) Arabinogalactan-proteins from non-coniferous gymnosperms have unusual structural features. *Carbohydrate Polymers*, **261**, 11783.
- Becker, B. & Marin, B. (2009) Streptophyte algae and the origin of embryophytes. *Annals of Botany*, **103**, 999–1004.
- Berry, E.A., Tran, M.L., Dimos, C.S., Budziszek, M.J., Jr., Scavuzzo-Duggan, T.R. & Roberts, A.W. (2016) Immuno and affinity cytochemical analysis of cell wall composition in the moss *Physcomitrella patens*. *Frontiers in Plant Science*, **7**, 248.
- Blakeney, A.B., Harris, P.J., Henry, R.J. & Stone, B.A. (1983) A simple and rapid preparation of alditol acetates for monosaccharide analysis. *Carbohydrate Research*, **113**, 291–299.
- Blumenkrantz, N. & Asboe-Hansen, G. (1973) New method for quantitative determination of uronic acids. *Analytical Biochemistry*, **34**, 484–489.
- Bossy, A., Blaschek, W. & Classen, B. (2009) Characterization and immuno-localization of arabinogalactan-proteins in roots of *Echinacea purpurea*. *Planta Medica*, **75**, 1526–1533.
- Bowman, J.L. (2013) Walkabout on the long branches of plant evolution. *Current Opinion in Plant Biology*, **16**, 70–77.
- Bowman, J.L., Kohchi, T., Yamato, K.T., Jenkins, J., Shu, S., Ishizaki, K. et al. (2017) Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell*, **171**, 287–304.e15. Available from: <https://doi.org/10.1016/j.cell.2017.09.030>
- Brinkhuis, H., Schouten, S., Collinson, M.E., Slujs, A., Sinninghe-Damsté, J.S., Dickens, G. et al. (2006) Episodic fresh surface waters in the Eocene Arctic Ocean. *Nature*, **441**, 606–609.
- Carafa, A., Duckett, J.G., Knox, J.P. & Ligrone, R. (2005) Distribution of cell-wall xylans in bryophytes and tracheophytes: new insights into basal interrelationships of land plants. *New Phytologist*, **168**, 231–240.
- Carpenter, E.J., Matasci, N., Ayyampalayam, S., Wu, S., Sun, J., Yu, J. et al. (2019) Access to RNA-sequencing data from 1,173 plant species: the 1000 plant transcriptomes initiative (1KP). *GigaScience*, **8**, 1–7.
- Chernova, T., Ageeva, M., Mikshina, P., Trofimova, O., Koslova, L., Lev-Yadun, S. et al. (2020) The living fossil *Psilotum nudum* has cortical fibers with mannan-based cell wall matrix. *Frontiers in Plant Science*, **11**, 488.
- Clarke, A.E., Gleeson, P.A., Jermyn, M.A. & Knox, R.B. (1978) Characterization and localization of β -lectins in lower and higher plants. *Australian Journal of Plant Physiology*, **5**, 707–722.
- Classen, B. (2007) Characterization of an arabinogalactan-protein from suspension culture of *Echinacea purpurea*. *Plant Cell, Tissue and Organ Culture*, **88**, 267–275.
- Classen, B., Baumann, A. & Utermohlen, J. (2019) Arabinogalactan-proteins in spore-producing land plants. *Carbohydrate Polymers*, **210**, 215–224.
- Classen, B., Csávas, M., Borbás, A., Dingermann, T. & Zündorf, I. (2004) Monoclonal antibodies against an arabinogalactan-protein from pressed juice of *Echinacea purpurea*. *Planta Medica*, **70**, 861–865.

- Classen, B., Mau, S.-L. & Bacic, A. (2005) The arabinogalactan-proteins from pressed juice of *Echinacea purpurea* belong to the hybrid class of hydroxyproline-rich glycoproteins. *Planta Medica*, **71**, 59–66.
- Classen, B., Witthohn, K. & Blaschek, W. (2000) Characterization of an arabinogalactan-protein isolated from pressed juice of *Echinacea purpurea* by precipitation with the β -glucosyl Yariv reagent. *Carbohydrate Research*, **327**, 497–504.
- Conway, J.R., Lex, A. & Gehlenborg, N. (2017) UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics*, **33**, 2938–2940.
- Costa, M., Pereira, A.M., Pinto, S.C., Silva, J., Pereira, L.G. & Coimbra, S. (2019) In silico and expression analyses of fasciclin-like arabinogalactan proteins reveal functional conservation during embryo and seed development. *Plant Reproduction*, **32**, 353–370.
- de Vries, J. & Archibald, J.M. (2018) Plant evolution: landmarks on the path to terrestrial life. *New Phytologist*, **217**, 1428–1434.
- de Vries, S. & de Vries, J. (2022) Evolutionary genomic insights into cyanobacterial symbioses in plants. *Quantitative Plant Biology*, **3**, e16.
- de Vries, S., de Vries, J., Teschke, H., von Dahlen, J., Rosem, L.E. & Gould, S.B. (2018) Jasmonic and salicylic acid response in the fern *Azolla filiculoides* and its cyanobiont. *Plant, Cell & Environment*, **41**, 2530–2548.
- Delwiche, C.F. & Cooper, E.D. (2015) The evolutionary origin of a terrestrial flora. *Current Biology*, **25**, R899–R910.
- Dolan, L., Linstead, P. & Roberts, K. (1995) An AGP epitope distinguishes a central metaxylem initial from other vascular initials in the *Arabidopsis* root. *Protoplasma*, **189**, 149–155.
- Dragičević, M.B., Paunović, D.M., Bogdanović, M.D., Todorović, S.I. & Simonović, A.D. (2020) Rapp: pipeline for mining of plant hydroxyproline-rich glycoproteins with implementation in R. *Glycobiology*, **30**, 19–35.
- Eeckhout, S., Leroux, O., Willats, W.G.T., Popper, Z.A. & Viane, R.L.L. (2014) Comparative glycan profiling of *Ceratopteris richardii* 'C-Fern' gametophytes and sporophytes links cell-wall composition to functional specialization. *Annals of Botany*, **114**, 1295–1307.
- Eisenhaber, B., Wildpaner, M., Schultz, C.J., Borner, G.H.H., Dupree, P. & Eisenhaber, F. (2003) Glycosylphosphatidylinositol lipid anchoring of plant proteins. Sensitive prediction from sequence- and genome-wide studies for *Arabidopsis* and *Rice*. *Plant Physiology*, **133**, 1691–1701.
- Ellis, M., Egelund, J., Schultz, C.J. & Bacic, A. (2010) Arabinogalactan-proteins: key regulators at the cell surface? *Plant Physiology*, **153**, 403–419.
- Fang, Y., Qin, X., Liao, Q., Du, R., Luo, X., Zhou, Q. et al. (2022) The genome of homosporous maidenhair fern sheds light on the euphyllophyte evolution and defences. *Nature Plants*, **8**, 1024–1037.
- Franková, L. & Fry, S.C. (2011) Phylogenetic variation in glycosidases and glycanases acting on plant cell wall polysaccharides, and the detection of transglycosidase and trans- β -xylanase activities. *The Plant Journal*, **67**, 662–681.
- Fry, S.C., Nesselrode, B.H.W.A., Miller, J.G. & Mewburn, B.R. (2008) Mixed-linkage (1–3, 1–4)- β -D-glucan is a major hemicellulose of *equisetum* (horsetail) cell walls. *New Phytologist*, **179**, 104–115.
- Fu, H., Yadav, M.P. & Nothnagel, E.A. (2007) *Physcomitrella patens* arabinogalactan proteins contain abundant terminal 3-O-methyl-L-rhamnosyl residues not found in angiosperms. *Planta*, **226**, 1511–1524.
- Galtier, J. & Scott, A.C. (1985) Diversification of early ferns. *Proceedings of the Royal Society of Edinburgh, Section B: Biological Sciences*, **86**, 289–301.
- Geddes, D.S. & Wilkie, K.C.B. (1971) Hemicelluloses from the stem tissues of the aquatic moss *Fontinalis antipyretica*. *Carbohydrate Research*, **18**, 333–335.
- Geshi, N., Johansen, J.N., Dilokpimol, A., Rolland, A., Belcram, K., Stephane Verger, S. et al. (2013) A galactosyltransferase acting on arabinogalactan protein glycans is essential for embryo development in *Arabidopsis*. *The Plant Journal*, **76**, 128–137.
- Gleeson, P.A. & Clarke, A.E. (1979) Structural studies on the major component of gladiolus tissue mucilage, an arabinogalactan-protein. *Biochemical Journal*, **181**, 607–621.
- Goellner, E.M., Blaschek, W. & Classen, B. (2010) Structural investigations on arabinogalactan-protein from wheat, isolated with Yariv reagent. *Journal of Agricultural and Food Chemistry*, **58**, 3621–3626.
- Goellner, E.M., Gramann, J.C. & Classen, B. (2013) Antibodies against Yariv's reagent for immunolocalization of arabinogalactan-proteins in aerial parts of *Echinacea purpurea*. *Planta Medica*, **79**, 175–180.
- Goellner, E.M., Ichinose, H., Kaneko, S., Blaschek, W. & Classen, B. (2011) An arabinogalactan-protein from whole grain of *Avena sativa* L. belongs to the wattle-blossom type of arabinogalactan-proteins. *Journal of Cereal Science*, **53**, 244–249.
- Happ, K. & Classen, B. (2019) Arabinogalactan-proteins from the liverwort *Marchantia polymorpha* L., a member of a basal land plant lineage, are structurally different to those of angiosperm. *Plants*, **8**, 460.
- Harholt, J., Sørensen, I., Fangel, J., Roberts, A., Willats, W.G.T., Scheller, H.V. et al. (2012) The glycosyltransferase repertoire of the spikemoss *Selaginella moellendorffii* and a comparative study of its cell wall. *PLoS One*, **7**, e35846.
- Harris, P.J., Henry, R.J., Blakeney, A.B. & Stone, B.A. (1984) An improved procedure for the methylation analysis of oligosaccharides and polysaccharides. *Carbohydrate Research*, **127**, 59–73.
- Harrison, J.C. (2017) Development and genetics in the evolution of land plant body plans. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, **372**, 20150490. Available from: <https://doi.org/10.1098/rstb.2015.0490>
- He, J., Zhao, H., Cheng, Z., Ke, Y., Liu, J. & Ma, H. (2019) Evolution analysis of the fasciclin-like arabinogalactan proteins in plants shows variable fasciclin-AGP domain constitutions. *International Journal of Molecular Sciences*, **20**, 1945.
- Hromádová, D., Soukup, A. & Tylová, E. (2021) Arabinogalactan proteins in plant roots – an update on possible functions. *Frontiers in Plant Science*, **12**, 674010.
- Hsieh, Y.S.Y. & Harris, P.J. (2012) Structures of xyloglucans in primary cell walls of gymnosperms, monilophytes (ferns sensu lato) and lycophytes. *Phytochemistry*, **79**, 87–101.
- Huang, X., Wang, W., Gong, T., Wickell, D., Kuo, L.-Y., Zhang, X. et al. (2022) The flying spider-monkey tree fern genome provides insights into fern evolution and arborescence. *Nature Plants*, **8**, 500–512.
- Johnson, K.L., Andrew, M., Cassin, A.M., Lonsdale, A., Bacic, A., Doblin, M.S. et al. (2017) A motif and amino acid bias bioinformatics pipeline to identify hydroxyproline-rich glycoproteins. *Plant Physiology*, **174**, 886–903.
- Johnson, K.L., Cassin, A.M., Lonsdale, A., Ka-Shu Wong, G., Soltis, D.E., Miles, N.W. et al. (2017) Insights into the evolution of hydroxyproline-rich glycoproteins from 1000 plant transcriptomes. *Plant Physiology*, **174**, 904–921.
- Johnson, K.L., Jones, B.J., Bacic, A. & Schultz, C.J. (2003) The fasciclin-like arabinogalactan proteins of *Arabidopsis*. A multigene family of putative cell adhesion molecules. *Plant Physiology*, **133**, 1911–1925.
- Johnson, K.L., Jones, B.J., Schultz, C.J. & Bacic, A. (2018) Non-enzymic cell wall (glyco)proteins. In: Roberts, J.A. (Ed.) *Annual plant reviews online*. Hoboken, NJ, USA: John Wiley & Sons, Corp, pp. 111–154. Available from: <https://doi.org/10.1002/9781119312994.apr0070>
- Katoh, K. & Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, **30**, 772–780.
- Kenrick, P. & Crane, P.R. (1997) The origin and early evolution of plants on land. *Nature*, **389**, 33–39.
- Kjeldahl, J. (1883) Neue Methode zur Bestimmung des Stickstoffs in organischen Körpern. *Zeitschrift für Analytische Chemie*, **22**, 366–382.
- Kobayashi, Y., Motose, H., Iwamoto, K. & Fukuda, H. (2011) Expression and genome-wide analysis of the xylogen-type gene family. *Plant & Cell Physiology*, **52**, 1095–1106.
- Kremer, C., Pettolino, F., Bacic, A. & Drinnan, A. (2004) Distribution of cell wall components in *sphagnum* hyaline cells and in liverwort and hornwort elaters. *Planta*, **219**, 1023–1035.
- Lampert, D.T.A. & Varnai, P. (2013) Periplasmic arabinogalactan glycoproteins act as a calcium capacitor that regulates plant growth and development. *New Phytologist*, **197**, 58–64.
- Lampert, D.T.A., Varnai, P. & Seal, C.E. (2014) Back to the future with the AGP-Ca²⁺ flux capacitor. *Annals of Botany*, **114**, 1069–1085.
- Lee, K.J.D., Sakata, Y., Mau, S.-L., Pettolino, F., Bacic, A., Quatrano, R.S. et al. (2005) Arabinogalactan proteins are required for apical cell extension in the moss *Physcomitrella patens*. *The Plant Cell*, **17**, 3051–3065.
- Leebens-Mack, J.H., Barker, M.S., Carpenter, E.J., Deyholos, M.K., Gitzendanner, M.A., Graham, S.W. et al. (2019) One thousand plant transcriptomes and the phylogenomics of green plants. *Nature*, **574**, 679–685.

- Leroux, O., Eeckhout, S., Viane, R.L.L. & Popper, Z.A. (2013) *Ceratopteris richardii* (C-fern): A model for investigating adaptive modification of vascular plant cell walls. *Frontiers in Plant Science*, **4**, 367.
- Leroux, O., Knox, J.P., Masschaele, B., Bagniewska-Zadworna, A., Marcus, S.E., Claeys, M. et al. (2011) An extensin-rich matrix lines the carinal canals in *Equisetum ramosissimum*, which may function as water-conducting channels. *Annals of Botany*, **108**, 307–319.
- Leroux, O., Leroux, F., Mastroberti, A.A., Santos-Silva, F., Van Loo, D., Bagniewska-Zadworna, A. et al. (2013) Heterogeneity of silica and glycan-epitope distribution in epidermal idioblast cell walls in *Adiantum raddianum* laminae. *Planta*, **237**, 1453–1464.
- Leroux, O., Sorensen, I., Marcus, S.E., Viane, R.L., Willats, W.G. & Knox, J.P. (2015) Antibody-based screening of cell wall matrix glycans in ferns reveals taxon, tissue and cell-type specific distribution patterns. *BMC Plant Biology*, **15**, 56.
- Leszczuk, A., Szczuka, E. & Zdunek, A. (2019) Arabinogalactan proteins: distribution during the development of male and female gametophytes. *Plant Physiology and Biochemistry*, **135**, 9–18.
- Letunic, I. & Bork, P. (2021) Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, **49**, W293–W296.
- Lex, A., Gehlenborg, N., Strobelt, H., Vuilleumot, R. & Pfister, H. (2014) UpSet: visualization of intersecting sets. *IEEE Transactions on Visualization and Computer Graphics*, **20**, 1983–1992.
- Li, F.-W., Brouwer, P., Carretero-Paulet, L., Cheng, S., de Vries, J., Delaux, P.-M. et al. (2018) Fern genomes elucidate land plant evolution and cyanobacterial symbioses. *Nature Plants*, **4**, 460–472.
- Ligrone, R., Vaughn, K.C., Renzaglia, K.S., Knox, J.P. & Duckett, J.G. (2002) Diversity in the distribution of polysaccharide and glycoprotein epitopes in the cell walls of bryophytes: new evidence for the multiple evolution of water-conducting cells. *New Phytologist*, **156**, 491–508.
- Lopez, R.A. & Renzaglia, K.S. (2014) Multiflagellated sperm cells of *Ceratopteris richardii* are bathed in arabinogalactan proteins throughout development. *American Journal of Botany*, **101**, 2052–2061.
- Lopez, R.A. & Renzaglia, K.S. (2016) Arabinogalactan proteins and arabinan pectins abound in the specialized matrices surrounding female gametes of the fern *Ceratopteris richardii*. *Planta*, **243**, 947–957.
- Lu, S., Wang, J., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R. et al. (2020) CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Research*, **48**, D265–D268.
- Ma, Y., MacMillan, C.P., de Vries, L., Mansfield, S., Hao, P., Ratcliffe, J. et al. (2022) FLA11 and FLA12 glycoproteins fine-tune stem secondary wall properties in response to mechanical stresses. *New Phytologist*, **233**, 1750–1767.
- Ma, Y., Yan, C., Li, H., Wu, W., Liu, Y., Wang, Y. et al. (2017) Bioinformatics prediction and evolution analysis of arabinogalactan proteins in the plant kingdom. *Frontiers in Plant Science*, **8**, 66.
- Ma, Y., Zeng, W., Bacic, A. & Johnson, K. (2018) AGPs through time and space. In: Roberts, J.A. (Ed.) *Annual plant reviews online*. Hoboken, NJ, USA: John Wiley & Sons, Corp. pp. 1–38. Available from: <https://doi.org/10.1002/9781119312994.apr0608>
- Marchant, D.B., Chen, G., Cai, S., Chen, F., Schafran, P., Jenkins, J. et al. (2022) Dynamic genome evolution in a model fern. *Nature Plants*, **8**, 1038–1051.
- Marcus, S.E., Blake, A.W., Benians, T.A.S., Lee, K.J.D., Poyser, C., Donaldson, L. et al. (2010) Restricted access of proteins to mannan polysaccharides in intact plant cell walls. *The Plant Journal*, **64**, 191–203.
- Marcus, S.E., Verherbruggen, Y., Hervé, C., Ordaz-Ortiz, J.J., Farkas, V., Pedersen, H.L. et al. (2008) Pectic homogalacturonan masks abundant sets of xyloglucan epitopes in plant cell walls. *BMC Plant Biology*, **8**, 60.
- Matsunaga, T., Ishii, T., Matsumoto, S., Higuchi, M., Darvill, A., Albersheim, P. et al. (2004) Occurrence of the primary cell wall polysaccharide rhamnogalacturonan II in pteridophytes, lycophytes, and bryophytes. Implications for the evolution of vascular plants. *Plant Physiology*, **134**, 339–351.
- McCartney, L., Marcus, S.E. & Knox, J.P. (2005) Monoclonal antibodies to plant cell wall xylans and arabinoxylans. *The Journal of Histochemistry and Cytochemistry*, **53**, 543–546.
- Meikle, P.J., Hoogenraad, N.J., Bonig, I., Clarke, A.E. & Stone, B.A. (1994) A (1→3,1→4)- β -glucan-specific monoclonal antibody and its use in the quantitation and immunocytochemical location of (1→3,1→4)- β -glucans. *The Plant Journal*, **4**, 1–9.
- Moore, J.P., Nguema-Ona, E.E., Vitré-Gibouin, M., Sørensen, I., Willats, W.G.T., Driouch, A. et al. (2013) Arabinose-rich polymers as an evolutionary strategy to plasticize resurrection plant cell walls against desiccation. *Planta*, **237**, 739–754.
- Motose, H., Sugiyama, M. & Fukuda, H. (2004) A proteoglycan mediates inductive interaction during plant vascular development. *Nature*, **429**, 873–878.
- Narciso, J.O., Zeng, W., Ford, K., Lampugnani, E.R., Humphries, J., Austarheim, I. et al. (2021) Biochemical and functional characterization of GALT8, an *Arabidopsis* GT31 β -(1,3)-galactosyltransferase that influences seedling development. *Frontiers in Plant Science*, **12**, 678564.
- Nibbering, P., Castilleux, R., Wingsle, G. & Niittylä, T. (2022) CAGEs are golgi-localized GT31 enzymes involved in cellulose biosynthesis in *Arabidopsis*. *The Plant Journal*, **110**, 1271–1285.
- Nitta, J.H., Schuettelpelz, E., Ramírez-Barahona, S. & Iwasaki, W. (2022) An open and continuously updated fern tree of life. *Frontiers in Plant Science*, **13**, 909768.
- Ogawa-Ohnishi, M. & Matsubayashi, Y. (2015) Identification of three potent hydroxyproline O-galactosyltransferases in *Arabidopsis*. *The Plant Journal*, **81**, 736–746.
- Paunović, D.M., Čuković, K.B., Bogdanović, M.D., Todorović, S.I., Trifunović-Momčilov, M.M., Subotić, A.R. et al. (2021) The arabinogalactan protein family of *Centaurea erythraea* Rafn. *Plants*, **10**, 1870.
- Pedersen, H.L., Fangel, J.U., McCleary, B., Ruzanski, C., Rydahl, M.G., Ralet, M.-C. et al. (2012) Versatile high-resolution oligosaccharide microarrays for plant glycobiology and cell wall research. *Journal of Biological Chemistry*, **287**, 39429–39438.
- Peña, M.J., Darvill, S.G., Eberhard, S., York, W.S. & O'Neill, M.A. (2008) Moss and liverwort xyloglucans contain galacturonic acid and are structurally distinct from the xyloglucans synthesized by hornworts and vascular plants. *Glycobiology*, **18**, 891–904.
- Pfeifer, L., Shafee, T., Johnson, K.L., Bacic, A. & Classen, B. (2020) Arabinogalactan-proteins of *Zostera marina* L. contain unique glycan structures and provide insight into adaption processes to saline environments. *Scientific Reports*, **10**, 8232.
- Pfeifer, L., Utermöhlen, J., Happ, K., Permann, C., Holzinger, A., von Schwarzenberg, K. et al. (2022) Search for evolutionary roots of land plant arabinogalactan-proteins in charophytes: presence of a rhamnogalactan-protein in *Spirogyra pratensis* (Zygnematophyceae). *The Plant Journal*, **109**, 568–584.
- Plackett, A.R.G., Rabinowitsch, E.H. & Langdale, J.A. (2015) Protocol: genetic transformation of the fern *Ceratopteris richardii* through micro-particle bombardment. *Plant Methods*, **11**, 37.
- Popper, Z.A. (2006) The cell walls of pteridophytes and other green plants – a review. *Fern Gazette*, **17**, 247–257.
- Popper, Z.A. (2008) Evolution and diversity of green plant cell walls. *Current Opinion in Plant Biology*, **11**, 286–292.
- Popper, Z.A. & Fry, S.C. (2003) Primary cell wall composition of bryophytes and charophytes. *Annals of Botany*, **91**, 1–12.
- Popper, Z.A. & Fry, S.C. (2004) Primary cell wall composition of pteridophytes and spermatophytes. *New Phytologist*, **164**, 165–174.
- Popper, Z.A., Sadler, I.H. & Fry, S.C. (2004) 3-O-Methylrhamnose in lower land plant primary cell walls. *Biochemical Systematics and Ecology*, **32**, 279–289.
- Pryer, K.M., Schuettelpelz, E., Wolf, P.G., Schneider, H., Smith, A.R. & Cranfill, R. (2004) Phylogeny and evolution of ferns (monilophytes) with a focus on the early leptosporangiate divergences. *American Journal of Botany*, **91**, 1582–1598.
- Pteridophyte Phylogeny Group I (PPG). (2016) A community-derived classification for extant lycophytes and ferns. *Journal of Systematics and Evolution*, **54**, 563–603. Available from: <https://doi.org/10.1111/jse.12229>
- Puhlmann, J., Bucheli, E., Swain, M.J., Dunning, N., Albersheim, P., Darvill, A.G. et al. (1994) Generation of monoclonal antibodies against plant cell wall polysaccharides. I. Characterization of a monoclonal antibody to a terminal α -(1→2)-linked fucosyl-containing epitope. *Plant Physiology*, **104**, 699–710.
- Puttick, M.N., Morris, J.L., Williams, T.A., Cox, C.J., Edwards, D., Kenrick, P. et al. (2018) The interrelationship of land plants and the nature of the ancestral embryophyte. *Current Biology*, **28**, 733–745.

- Qu, Y., Egelund, J., Gilson, P.R., Houghton, F., Gleeson, P.A., Schultz, C.J. et al. (2008) Identification of a novel group of putative *Arabidopsis thaliana* β -(1,3)-galactosyltransferases. *Plant Molecular Biology*, **68**, 43–59.
- Rensing, S.A. (2017) Why we need more non-seed plant models. *New Phytologist*, **216**, 355–360.
- Renzaglia, K.S. & Warne, T.R. (1995) *Ceratopteris*: an ideal model system for teaching plant biology. *International Journal of Plant Sciences*, **156**, 385–392.
- Ruprecht, C., Bartetzko, M.P., Senf, D., Dallabernadina, P., Boos, I., Andersen, M.C.F. et al. (2017) A synthetic glycan microarray enables epitope mapping of plant cell wall glycan-directed antibodies. *Plant Physiology*, **175**, 1094–1104.
- Ruprecht, C., Bartetzko, M.P., Senf, D., Lakhina, A., Smith, P.J., Soto, M.J. et al. (2020) A glycan array-based assay for the identification and characterization of plant glycosyltransferases. *Angewandte Chemie International Edition*, **59**, 12493–12498.
- Schultz, C.J., Johnson, K.L., Currie, G. & Bacic, A. (2000) The classical arabinogalactan protein gene family of *Arabidopsis*. *The Plant Cell*, **12**, 1751–1767.
- Seifert, G.J. & Roberts, K. (2007) The biology of arabinogalactan proteins. *Annual Review of Plant Biology*, **58**, 137–161.
- Shafee, T., Bacic, A. & Johnson, K. (2020) Evolution of sequence-diverse disordered regions in a protein family: order within the chaos. *Molecular Biology and Evolution*, **37**, 2155–2172.
- Showalter, A.M. & Basu, D. (2016) Glycosylation of arabinogalactan-proteins essential for development in *Arabidopsis*. *Communicative and Integrative Biology*, **5**, e1177687.
- Silva, G.B., Ionashiro, M., Carrara, T.B., Crivellari, A.C., Tine, M.A.S., Prado, J. et al. (2011) Cell wall polysaccharides from fern leaves: evidence for a mannan-rich type III cell wall in *Adiantum raddianum*. *Phytochemistry*, **72**, 2352–2360.
- Silva, J., Ferraz, R., Dupree, P., Showalter, A.M. & Coimbra, S. (2020) Three decades of advances in arabinogalactan-protein biosynthesis. *Frontiers in Plant Science*, **11**, 610377.
- Smallwood, M., Yates, E.A., Willats, W.G.T., Martin, H. & Knox, J.P. (1996) Immunochemical comparison of membrane-associated and secreted arabinogalactan-proteins in rice and carrot. *Planta*, **198**, 452–459.
- Sørensen, I., Pettolino, F.A., Bacic, A., Ralph, J., Lu, F., O'Neill, M.A. et al. (2011) The charophycean green algae provide insights into the early origins of plant cell walls. *The Plant Journal*, **68**, 201–211.
- Sørensen, I., Pettolino, F.A., Wilson, S.M., Doblin, M.S., Johansen, B., Bacic, A. et al. (2008) Mixed-linkage (1→3), (1→4)- β -D-glucan is not unique to the Poales and is an abundant component of *Equisetum arvense* cell walls. *The Plant Journal*, **54**, 510–521.
- Speelman, E.N., Van Kempen, M.M.L., Barke, J., Brinkhuis, H., Reichart, G.J., Smolders, A.J.P. et al. (2009) The eocene erctic *Azolla* bloom: environmental conditions, productivity and carbon drawdown. *Geobiology*, **7**, 155–170.
- Stegemann, H. & Stalder, K. (1967) Determination of hydroxyproline. *Clinica Chimica Acta*, **18**, 267–273.
- Strasser, R., Seifert, G., Doblin, M.S., Johnson, K.L., Ruprecht, C., Pfengle, F. et al. (2021) Cracking the “sugar code”: a snapshot of N- and O-glycosylation pathways and functions in plants. *Frontiers in Plant Science*, **12**, 640919.
- Suzuki, T., Narciso, J.O., Zeng, W., van de Meene, A., Yasutomi, M., Take-mura, S. et al. (2017) KNS4/UPEX1: a type II arabinogalactan β -(1,3)-galactosyltransferase required for pollen exine development. *Plant Physiology*, **173**, 183–205.
- Takenaka, Y., Kato, K., Ogawa-Ohnishi, M., Tsuruhama, K., Kajiuira, H., Yagyu, K. et al. (2018) Pectin RG-I rhamnosyltransferases represent a novel plant-specific glycosyltransferase family. *Nature Plants*, **4**, 669–676.
- Taylor, R.L. & Conrad, H.E. (1972) Stoichiometric depolymerization of polyuronides and glycosaminoglycans to monosaccharides following reduction of their carbodiimide-activated carboxyl groups. *Biochemistry*, **11**, 1383–1388.
- Thude, S. & Classen, B. (2005) High molecular weight constituents from roots of *Echinacea pallida*: an arabinogalactan-protein and an arabinan. *Phytochemistry*, **66**, 1026–1032.
- Tryfona, T., Liang, H.-C., Kotake, T., Tsumuraya, Y., Stephens, E. & Dupree, P. (2012) Structural characterization of *Arabidopsis* leaf arabinogalactan polysaccharides. *Plant Physiology*, **160**, 653–666.
- Verhertbruggen, Y., Marcus, S.E., Haeger, A., Verhoef, R., Schols, H.A., McCleary, B.V. et al. (2009) Developmental complexity of arabinan polysaccharides and their processing in plant cell walls. *The Plant Journal*, **59**, 413–425.
- Wack, M., Classen, B. & Blaschek, W. (2005) An acidic arabinogalactan-protein from the roots of *Baptisia tinctoria* (L.) R. Brown. *Planta Medica*, **71**, 814–818.
- Weng, J.-K., Li, X., Bonawitz, N.D. & Chapple, C. (2008) Emerging strategies of lignin engineering and degradation for cellulosic biofuel production. *Current Opinion in Biotechnology*, **19**, 166–172.
- Woudenberg, S., Renema, J., Tomescu, A.M.F., De Rybel, B. & Weijers, D. (2022) Deep origin and gradual evolution of transporting tissues: perspectives from across the land plants. *Plant Physiology*, **190**, 85–99. Available from: <https://doi.org/10.1093/plphys/kiac304>
- Xue, X. & Fry, S.C. (2012) Evolution of mixed-linkage (1→3, 1→4)- β -D-glucan (MLG) and xyloglucan in *equisetum* (horsetails) and other monilophytes. *Annals of Botany*, **109**, 873–886.
- Yates, E.A., Valdor, J.F., Haslam, S.M., Morris, H.R., Dell, A., Mackie, W. et al. (1996) Characterization of carbohydrate structural features recognized by anti-arabinogalactan-protein monoclonal antibodies. *Glycobiology*, **6**, 131–139.