Plant cell wall proteomics: the leadership of *Arabidopsis thaliana*

Cécile Albenne1,2, Hervé Canut1,2 and Elisabeth Jamet1,2*

1 Laboratoire de Recherche en Sciences Végétales, Université de Toulouse, UPS, UMR 5546, Castanet-Tolosan, France
2 CNRS, UMR 5546, Castanet-Tolosan, France

INTRODUCTION

Plant primary cell walls are mainly composed of polysaccharide networks such as cellulose microfibrills, hemicellulosic wrapping and interlacing cellulose microfibrills and pectins (Carpita and Gibeaut, 1993). After the end of cell growth, secondary walls which contain additional compounds such as lignins, wax or cutin, are synthesized. Cell wall proteins (CWPs) play critical roles in plant cell walls during development and adaptation to environmental cues (Fry, 2004; Passardi et al., 2004). For this reason, extensive studies leading to their identification and characterization have been undertaken. Cell wall proteomics started about 10 years ago when the first plant genome sequences became available. Nowadays, there are about 40 papers covering this field (Figure 1), half of them concerning *Arabidopsis thaliana* whose genome was available in 2000 (Arabidopsis Genome Initiative, 2000). The availability of new genome sequences such as those of *Oryza sativa* (International Rice Genome Sequencing Project, 2005), *Populus trichocarpa* (Tuskan et al., 2006) and *Solanium lycopersicum* (Tomato Genome Consortium, 2012) enlarged the range of plant proteomics studies.

For plant cell wall proteomics studies, organs or cell suspension cultures have been used as starting materials containing cells surrounded by primary and/or secondary walls. Various experimental approaches were undertaken to characterize cell wall proteomes. Five specific features of CWPs need to be emphasized to understand them. (i) CWPs represent only 5–10% of the cell wall mass (Cassab and Varner, 1988). They are embedded in a complex matrix of carbohydrate polymers, aromatic compounds, wax or cutin depending on the type of cell walls. (ii) CWPs may interact with cell wall components by non-covalent linkages (Carpin et al., 2001; Spadoni et al., 2006). They can also be covalently linked, thus forming insoluble networks, like structural proteins networks of Proline-Rich Proteins (PRPs) or extensins (Brisson et al., 1994; Brady et al., 1996). (iii) Contrary to other sub-cellular compartments, plant cell walls constitute an open space connecting the cells in a tissue. It is located between the cell plasma membrane and the cuticle in aerial organs or the suberin layer in roots conferring to the plant surface waterproof qualities and protection against biotic and abiotic stresses (Thomas et al., 2007; Javelle et al., 2010). (iv) Most CWPs are basic proteins (Jamet et al., 2008). (v) Most CWPs undergo post-translational modifications (PTMs), like hydroxylation of proline (Pro) residues converting them to hydroxyproline (Hyp), N-glycosylation, O-glycosylation or addition of a glycosylphosphatidylinositol (GPI)-anchor (Kieliszewski and Lamport, 1994; Spiro, 2002; Faye et al., 2005).

For each step of the cell wall proteomics flowchart, the specificities of CWPs must be taken into account: plant fractionation, protein extraction, protein separation, protein identification by mass spectrometry (MS), and bioinformatics. Indeed, CWPs can be tightly trapped into the extracellular matrix and escape the extraction procedure. They may not be resolved at the step of separation by two-dimensional electrophoresis (2D-E) because they are mainly basic glycoproteins (Jamet et al., 2008). Finally, the databases used for protein identification using MS data
contain no information about PTMs such as glycosylation, thus preventing the identification of some of them.

In this review, we will give a survey of plant cell wall proteomics studies with a focus on *A. thaliana* because this plant provides the best documented cell wall proteomes. The main protein families identified and the persistent issue of the non-canonical CWPs will be discussed. Finally, we will provide perspectives in the field of plant cell wall proteomics, going beyond the present data with systems biology approaches and peptidomics to decipher the roles of proteins and peptides in cell walls.

**MATERIALS AND METHODS**

All the *A. thaliana* proteins reported in this review have been analyzed with different bioinformatics software to predict their sub-cellular localization and their functional domains using ProtAnnDB ([http://www.polebio.lrvs.ups-tlse.fr/ProtAnnDB/index.php](http://www.polebio.lrvs.ups-tlse.fr/ProtAnnDB/index.php)) as previously described (San Clemente et al., 2009). Briefly, the following programs have been used for prediction of sub-cellular localization: TargetP ([http://www.cbs.dtu.dk/services/TargetP/](http://www.cbs.dtu.dk/services/TargetP/)), SignalP ([http://www.cbs.dtu.dk/services/SignalP/](http://www.cbs.dtu.dk/services/SignalP/)), Predotar ([http://urgi.versailles.inra.fr/predotar/predotar.html](http://urgi.versailles.inra.fr/predotar/predotar.html)), Aramemnon ([http://aramemnon.botanik.uni-koeln.de/](http://aramemnon.botanik.uni-koeln.de/)) and TMHMM ([http://www.cbs.dtu.dk/services/TMHMM-2.0/](http://www.cbs.dtu.dk/services/TMHMM-2.0/)). The programs used for prediction of functional domains were PROSITE ([http://prosite.expasy.org/](http://prosite.expasy.org/)), Pfam ([http://pfam.sanger.ac.uk/](http://pfam.sanger.ac.uk/)), and InterPro ([http://www.ebi.ac.uk/interpro/](http://www.ebi.ac.uk/interpro/)).

**FIGURE 1 | Occurrence of cell wall proteomics papers since 1997.** The number of papers published each year is represented with colored bars for each plant: *A. thaliana* (green), *O. sativa* (pink), *P. deltoides* (yellow), *S. lycopersicum* (red), and others (blue). The arrows show the year of the genomic sequence release for each of them. Note that the genomic sequence of *P. trichocarpa* was used for protein identification in *P. deltoides.*

**A SURVEY OF PLANT CELL WALL PROTEOMICS**

**EXTRACELLULAR PROTEOMES**

In this review, we will focus on different type of extracellular proteomes, commonly named as cell wall proteomes. For example, secretome, in which all of the secreted proteins of a cell suspension culture, roots or seedling are collected in liquid culture media.

Another type of extracellular proteome encompasses apoplastic proteomes in which proteins from the cell wall can be eluted by vacuum infiltration with various solutions. Extraction of proteins from purified cell walls with various solutions is the third category of cell wall proteome that have been used to elute loosely bound CWPs. In addition, sub-proteomes such as N-glycoproteomes and a GPI-anchored proteome have been analyzed. All the extracellular proteomes have been obtained with different plants like *A. thaliana* (see Table 1), *Cicer arietinum* (Bhushan et al., 2006, 2011), *Glycine max* (Komatsu et al., 2010), *Helianthus annuus* (Pinedo et al., 2012), *O. sativa* (Chen et al., 2008a; Jung et al., 2008; Cho et al., 2009; Zhou et al., 2011), *Medicago sativa* (Watson et al., 2004; Verdonk et al., 2012), *Nicotiana benthamiana* (Goulet et al., 2010), *Nicotiana tabacum* (Robertson et al., 1997; Dani et al., 2005; Morel et al., 2006; Delannoy et al., 2008; Millar et al., 2009), *Populus deltoides* (Pechanova et al., 2010), *S. lycopersicum* (Robertson et al., 1997; Yeats et al., 2010; Catalá et al., 2011), *Solanum tuberosum* (Fernández et al., 2012; Lim et al., 2012) and *Zea mays* (Zhu et al., 2006, 2007). Besides, several xylem sap proteomes have been analyzed and were found to be very close to cell wall proteomes (Kehr et al., 2005; Alvarez et al., 2006; Dafoe and Constabel, 2009; Ligat et al., 2011). With 20 published papers (Table 1) and 500 proteins with predicted signal peptide identified, the most studied plant is *A. thaliana*. Its genome was the first one to be sequenced, thus allowing a precise identification of the proteins. Altogether, between one fourth and one third of the expected cell wall proteome of *A. thaliana* has been identified (Jamet et al., 2006). The second most studied plant is *O. sativa* with 270 proteins with predicted signal peptide identified. When no genome information is available, protein identification relies on the availability of expressed sequenced tags (ESTs) or cDNAs (Lim et al., 2012). Alternatively, proteins are identified by sequence homology. In this case, it is not possible to obtain precise identification of proteins and to distinguish between members of multigene families like in *C. arietinum* or *H. annuus* (Bhushan et al., 2006; Pinedo et al., 2012).

**STRAATEGIES IN PLANT CELL WALL PROTEOMICS**

Many different strategies have been used to identify extracellular proteins of plants. A synopsis of the different experimental procedures is presented in Figure 2 and Table A1 in five general steps. Steps 1 and 2 lead to protein extraction. Step 3 consists in protein separation. Steps 4 and 5 lead to protein identification. The first step distinguishes: (i) studies of secretomes in which only proteins spontaneously released in culture media are analyzed (Charmont et al., 2005; Oh et al., 2005; Basu et al., 2006; Tran and Plaxton, 2008; Cheng et al., 2009); (ii) the release of proteins by non-destructive methods in which the integrity of the cell plasma membranes is preserved either by vacuum infiltration of tissues (Haslam et al., 2003; Boudart et al., 2005; Casasoli et al., 2008) or by washing of cells cultured in liquid medium (Robertson et al., 2009).
Table 1 | Cell wall proteomes of A. thaliana.

| Organ                      | References                  | Type of proteome     | Total number of identified proteins (% without predicted signal peptide)\(^c\) | Number of proteins with predicted signal peptide\(^c\) |
|----------------------------|-----------------------------|----------------------|---------------------------------------------------------------------------------|-------------------------------------------------|
| Culture medium of roots     | Basu et al., 2006           | Secretome            | 52 (31)                                                                         | 36                                               |
| Stems                      | Minic et al., 2007          | N-glycoproteome      | 102 (12)                                                                        | 90                                               |
| Leaves                     | Haslam et al., 2003         | Apoplastic proteome  | 10 (10)                                                                          | 9                                                 |
| Leaves                     | Boudart et al., 2005        | Apoplastic proteome  | 93 (7)                                                                           | 87                                               |
| Etiolated hypocotyls (5 days) | Feiz et al., 2006          | Cell wall proteome   | 96 (13)                                                                         | 84                                               |
| Etiolated hypocotyls (5 and 11 days) | Irshad et al., 2008 | Cell wall proteome   | 173 (21)                                                                        | 137                                              |
| Etiolated hypocotyls (11 days) | Zhang et al., 2011        | Cell wall glycoproteome | 129 (2)                                                                       | 127                                              |
| Seedlings (10 days)         | Casasoli et al., 2008\(^a\) | Apoplastic proteome  | 20 (30)                                                                          | 14                                               |
| (oligogalacturonides)       |                             |                      |                                                                                  |                                                  |
| Cell suspension cultures (3 days) | Bayer et al., 2006       | Cell wall proteome   | 792 (87)                                                                        | 106                                              |
| Cell suspension cultures (5 days) | Kwon et al., 2005          | Cell wall proteome\(^b\) | 39 (13)                                                                         | 34                                               |
| Cell suspension cultures (5 days) | Chivasa et al., 2002      | Cell wall proteome   | 72 (64)                                                                          | 26                                               |
| Cell suspension cultures (5 days) | Robertson et al., 1997     | Cell wall proteome\(^b\) | 33 (46)                                                                         | 18                                               |
| Cell suspension cultures (7 days) | Borderies et al., 2003    | Cell wall proteome\(^b\) | 95 (48)                                                                         | 50                                               |
| Culture medium of calli (21 days) | Borner et al., 2003     | GPI-anchored proteome | 30 (0)                                                                          | 30                                               |
| Culture medium of cell suspension cultures (4 days) | Oh et al., 2005             | Secretome            | 45 (47)                                                                          | 24                                               |
| Cell suspension cultures (7 days) | Ndimba et al., 2003\(^a\)   | Cell wall proteome   | 6 (0)                                                                            | 6                                                 |
| (fungal elicitors)          |                             |                      |                                                                                  |                                                  |
| Culture medium of cell suspension cultures (7 days) (salicylic acid treatment) | Cheng et al., 2009        | Secretome            | 74 (47)                                                                          | 39                                               |
| Culture medium of cell suspension cultures (7 days) (phosphate deficiency) | Tran and Plaxton, 2008\(^a\) | Secretome            | 37 (43)                                                                          | 21                                               |
| Protoplasts (1 h)           | Kwon et al., 2005           | Cell wall proteome\(^b\) | 71 (65)                                                                         | 25                                               |
| Protoplasts (3 h)           | Kwon et al., 2005           | Cell wall proteome\(^b\) | 66 (65)                                                                         | 23                                               |
| Liquid cultured etiolated seedlings | Schultz et al., 2004 | AGP proteome         | 12 (0)                                                                          | 12                                               |
| Culture medium of etiolated seedlings | Charmont et al., 2005 | Secretome            | 49 (10)                                                                          | 44                                               |

\(^a\) Protein identification has been only performed on protein spots showing variation between control and treated samples.

\(^b\) These proteomes have been obtained by washing of cells or protoplasts with various salt solutions (see Table A1).

\(^c\) All the bioinformatic predictions of sub-cellular localization have been done as described in Materials and Methods to allow the comparison between the A. thaliana cell wall proteomes.

1997; Borderies et al., 2003; Kwon et al., 2005); and (iii) the release of proteins by destructive methods starting with a grinding of the tissues, thus mixing intracellular and extracellular compartments. In this case, either cell walls were purified prior to protein extraction (Chivasa et al., 2002; Ndimba et al., 2003; Feiz et al., 2006; Minic et al., 2007; Irshad et al., 2008; Zhang et al., 2011) or the tissues were ground prior to isolation of N-glycosylated proteins by lectin affinity chromatography (Minic et al., 2007). In the case of the GPI-anchored proteome, the first step consisted in the preparation of a membrane fraction followed by the cleavage of GPI-anchors by phosphatidylinositol-specific phospholipase C (Pi-PLC) (Borner et al., 2003).

The second step (Figure 2) is very diverse using different solutions to extract proteins. These solutions can be acidic or basic (Feiz et al., 2006; Casasoli et al., 2008). Their main components are: salts (NaCl, CaCl\(_2\), MgCl\(_2\), KCl, or LiCl) or osmotic agents (mannitol) (Borderies et al., 2003; Boudart et al., 2005; Kwon et al., 2005; Feiz et al., 2006); chelating agents (EDTA or CDTA) (Robertson et al., 1997; Boudart et al., 2005); detergents (SDS, Triton or CHAPS) (Chivasa et al., 2002; Borner et al., 2003); phenol (Bayer et al., 2006); and/or chaotropic agents (urea and thiourea) (Chivasa et al., 2002). The β-glucosyl Yariv reagent has been used to isolate arabinoxylan proteins (AGPs) (Schultz et al., 2004). In some cases, several salt solutions have been used successively (Chivasa et al., 2002; Borderies et al., 2003; Boudart et al., 2005; Feiz et al., 2006; Irshad et al., 2008; Zhang et al., 2011). As mentioned above, in the case of the GPI-anchored proteome, a Pi-PLC has been used (Borner et al., 2003). The methods used to extract CWPs have been previously described in detail (Feiz et al., 2006; Jamet et al., 2008).
The third possible step is protein separation (Figure 2). It can be done by chromatography (cationic exchange, lectin affinity, boronic acid), and/or by 1D- or 2D-E. Cationic exchange chromatography has been performed under physico-chemical conditions similar to those found in cell walls, that is an acidic medium at a pH around 4.5 at which basic proteins are positively charged (Boudart et al., 2005; Irshad et al., 2008). Affinity chromatography on Concanavalin A (ConA) has been artfully used to isolate N-glycoproteins from a total extract of proteins (Minic et al., 2007). As expected, most of the identified proteins were predicted to be addressed to the secretion pathway where N-glycosylation occurs. Other lectins have been used to separate proteins extracted from cell walls: *Artocarpus integrifolia* Lectin (AIL) specific for α-Gal residues, PeaNut Agglutinin (PNA) specific for β-Gal residues, and wheat germ agglutinin (WGA) specific for GlcNAc residues (Zhang et al., 2011). With regard to separation of proteins by electrophoresis, 2D-E has shown limitations due to the fact that CWPs are mainly basic glycoproteins (Jamet et al., 2008). Considering the number of identified proteins, the most efficient cell wall proteomics analyses have been performed using two steps of protein separation (Boudart et al., 2005; Minic et al., 2007; Irshad et al., 2008; Zhang et al., 2011).

Two additional steps are necessary to achieve protein identification (steps 4 and 5, Figure 2). The fourth step consists in proteolytic digestion of proteins and MS analyses of peptides, using Matrix-Assisted Laser Desorption Ionization-Time Of Flight (MALDI-TOF) MS (Boudart et al., 2005; Kwon et al., 2005; Irshad et al., 2008), liquid chromatography (LC)-MS/MS (Minic et al., 2007; Casasoli et al., 2008; Zhang et al., 2011) or 2D-LC-MS/MS (Basu et al., 2006; Bayer et al., 2006; Cheng et al., 2009). Trypsin is the most widely used protease. In a few cases, Edman N-terminal sequencing has been performed (Robertson et al., 1997; Schultz et al., 2004). When proteins are heavily glycosylated, like O-glycoproteins, it is necessary to deglycosylate them with hydrogen fluoride (HF) to get access to their polypeptide skeleton (Schultz et al., 2004). The fifth step consists in bioinformatics analyses to identify proteins, predict their
sub-cellular localization, their functional domains and eventually get information about their PTMs (San Clemente et al., 2009).

Different complementary strategies are now available to study plant cell wall proteomes. It is possible to design the most relevant flowchart for a new cell wall proteomics study and to perform it in an efficient way. The main limitation remains the availability of genomic sequences for many plants of agronomic interest.

PLANT CELL WALL GLYCOPROTEOMES

During their secretion, proteins undergo glycosylation which is one of the most common and complex PTM known to control many physiological processes (Faye et al., 2005). Glycosylation is of two main types, namely N- and O-glycosylation, depending on the nature of the amino acid bearing them. Unlike yeast and mammalian glycoproteins which are extensively studied, plant glycoproteins are still poorly characterized. Hyp-rich glycoproteins (HRGPs) undoubtedly constitute the most documented cell wall O-glycoprotein superfamiliy (Kieliszewski, 2001; Tan et al., 2012; Velasquez et al., 2012). A few N-glycoproteins have been studied in detail, e.g., a peroxidase (Lige et al., 2001), an α-mannosidase (Kimura et al., 1999) or a polygalacturonase inhibiting protein (PGIP) (Lim et al., 2009) for which glycosylation has been shown to contribute to activity. Beyond the study on targeted glycoproteins, the concept of glycoproteomics is now emerging in plants. New analytical pipelines are available (Song et al., 2011; Ruiz-May et al., 2012). They aim at detection, enrichment and MS analysis of large sets of glycoproteins.

A few systematic surveys have been carried out so far on plants. Minic et al. were the first ones to use a ConA lectin chromatography step to capture N-glycoproteins from a protein extract of A. thaliana (Minic et al., 2007). A similar approach has been developed to characterize the N-glycoproteome of S. lycopersicum (Catalá et al., 2011). Finally, Zhang et al. enlarged the coverage of the A. thaliana cell wall glycoproteome using multi-dimensional lectin chromatography and boronic acid chromatography (Zhang et al., 2011). The obtained sub-proteomes mostly corresponded to N-glycoproteomes with only few O-glycoproteins detected. Plant glycoproteomics is only at its premise and is undoubtedly a very promising approach toward an integrated study of both sugar and proteins moieties to gain new insight into the structure and function of glycoproteins.

MAIN CWP FAMILIES

GROUPING OF CWPs: PRINCIPLES AND DRAWBACKS

A major challenge is to interpret cell wall proteomics data, in other words, to get a biological message from a list of proteins. In a first effort, it is tempting to group proteins to get an overview of the extracellular proteome, to highlight specific proteins or protein families in the physiological context of interest, or to identify house-keeping proteins. This is a difficult exercise because most of the identified proteins have no experimentally defined function. Another difficulty is that two series of proteins can be distinguished in all cell wall proteomes: those having a signal peptide predicted with at least two bioinformatics programs, and those having no predicted signal peptide or having a motif that addresses them to an intracellular compartment. Only those having a bona fide predicted signal peptide are named CWPs in this review. This point will be discussed below.

From the fifth step of the proteomics flowchart (Figure 2), bioinformatic analyses lead to group proteins in families either by sequence comparison to proteins present in databases and already annotated, or by search for functional domains as defined in domain repertoires like PROSITE, Pfam, or InterProScan. Two types of classification have then been proposed. Both of them have drawbacks and suffer from ambiguity. The first type is based on the physiological processes in which the proteins are assumed to be involved, like growth and development, stress or defense against pathogens. The drawback is that it can be difficult to sort the proteins. For example, glycoside hydrolases (GHs) could be involved in both plant development and defense (Kasprzewska, 2003). In the same way, proteases could be involved in protein turnover or in signaling by releasing biologically active peptides (Berger and Altmann, 2000; Hunt et al., 2010; Leasure and He, 2012). The second type of classification is based on predicted functional domains and possible partners or targets (i.e., polysaccharides, lipids or proteins) of CWPs in cell walls. The drawbacks are the followings: (i) all the proteins do not have a predicted biochemical activity; (ii) it is difficult to assign proteins with several functional domains to a class as exemplified below. Of course, all these classifications need to evolve to take into account new experimental results demonstrating protein functions.

AN EXAMPLE OF FUNCTIONAL CLASSIFICATION OF CWPs

In this review, we present a functional classification taking into account predicted functional domains as well as possible partners or targets in cell walls. A. thaliana CWPs are taken as an example. Nine functional classes listed in Table 2 have been proposed (Janet et al., 2008). The most populated functional class is that of proteins acting on carbohydrates. It represents about one fourth of the proteomes (25.6%) and it includes GHs, carbohydrate esterases (CEs), polysaccharide lyases (PLs) and expansins. The importance of such proteins is not surprising since polysaccharides constitute the largest fraction of cell walls and are constantly submitted to remodeling during plant development or in response to environmental cues (Fry, 2004; Cosgrove, 2005). The second most predominant class of CWPs is that of oxido-reductases (14.6%), like peroxidases, multicopper oxidases, blue copper binding proteins, and berberine bridge enzymes. Again, the importance of this class was expected because many oxidation reactions occur in the extracellular matrix to modify polymer networks involving carbohydrates, aromatic compounds, or structural proteins (Passardi et al., 2004). However, the biochemical functions of proteins homologous to berberine bridge enzymes and of blue copper binding proteins in cell walls are not known (Nersissian and Shipp, 2002). Then, numerous proteases have been found in cell wall proteomes (11.2%). Until recently, the roles of such enzymes have probably been under-estimated in cell walls. They could be involved in protein turnover, protein maturation or release of biologically active peptides (Van Der Hoorn, 2008). Nothing is known about CWP turnover. The maturation of enzymes having N- or C-terminal pro-peptides or N-terminal inhibitory domains has
Table 2 | A classification of *A. thaliana* CWPs in 9 functional classes according to their predicted functional domains and their possible partners in cell walls.

| Functional class | Protein families | Possible partners in cell walls | % *A. thaliana* proteomes |
|------------------|------------------|-------------------------------|--------------------------|
| Proteins acting on carbohydrates | Glycoside hydrolases (GHs) | Carbohydrates | 19.4 |
| | Carbohydrate esterases (CEs) | Pectins | 2.4 |
| | Polysaccharide lyases (PLs) | Pectins | 0.6 |
| | Expansins | Cellulose/hemicellulose | 2.4 |
| Oxido-reductases | Peroxidases | Carbohydrates/Structural proteins/Lignins | 6.2 |
| | Multicopper oxidases | | 2.0 |
| | Blue copper binding proteins | | 2.2 |
| | Berberine bridge enzymes | | 2.2 |
| Proteases | Asp proteases | Proteins | 2.8 |
| | Cys proteases | Proteins | 1.6 |
| | Ser proteases | Proteins | 3.6 |
| | Ser carboxypeptidases | Proteins | 2.2 |
| Proteins with interaction domains | Lectins | Carbohydrates | 2.4 |
| | Leucine-rich-repeat domains (LRRs) | Proteins | 2.6 |
| | Enzyme inhibitors | Enzymatic proteins | 4.0 |
| Proteins possibly involved in signaling | Arabinogalactan proteins (AGPs) | | 2.4 |
| | Fasciclin-AGPs (FLAs) | | 1.2 |
| | Receptors | | 2.4 |
| Structural proteins | Glycine-rich proteins (GRPs) | | 0.2 |
| | Leucine-rich-repeat extensins (LRXs) | Proteins | 0.8 |
| | Proline-rich proteins (PRPs) | Proteins | 0.6 |
| Proteins related to lipid metabolism | Lipase GDSL | Lipids | 2.4 |
| | Lipid transfer proteins (LTPs) | Lipids | 1.8 |
| Miscellaneous proteins | Purple acid phosphatases (PAPs) | | 1.2 |
| | Phosphate-induced (phi) proteins | | 1.0 |
| | Germin | | 1.6 |
| Unknown function | Domains of unknown function (DUFs) | | 5.0 |
| | No clue | | 4.4 |

The annotation of proteins is based on the presence of functional domains as defined in the PROSITE, Pfam, and InterPro bioinformatics programs. Only the major protein families are mentioned in each functional class.
been demonstrated only in a few cases, such as type I pectin methylesterases (PMEs) (Wolf et al., 2009) and some GHs (Lee et al., 2003; Minic et al., 2004; Albenne et al., 2009). It was also shown that the AtSBT1.7 Ser protease plays a role in mucilage release from A. thaliana seed coat (Rautengarten et al., 2008). It is assumed that this protease contributes to the degradation or the maturation of cell wall modifying enzymes. The class of CWPs possibly involved in signaling (6.6%) is a difficult one. It comprises proteins like AGPs and proteins with transmembrane domains which are predicted to be plasma membrane receptors having extracellular domains. The roles of AGPs are not completely understood. Besides their roles in signaling (Seifert and Roberts, 2007), the identification of receptors mostly relies on peptides located in the extracellular domain, but they do not really belong to cell wall proteomes. Finally a protease like stomatal density and distribution 1 (SDD1) could also be included in the class of signaling proteins because it is assumed to generate an extracellular signal to control the stomatal pattern (Von Groll et al., 2002). The next class of CWPs is that of proteins predicted to be related to lipid metabolism which is unexpectedly populated (5.8%). Such proteins are assumed to be involved in cuticle formation. However, the cuticle does not represent a major part of the organs analyzed. Other roles might be possible for these proteins. For example, a lipid transfer protein (LTP) has been assumed to be involved in cell wall extension by interacting with the cellulose/xyloglucan network of tobacco cell walls (Nieuwland et al., 2005). The class of structural proteins (1.6%) only groups a few proteins like Glycine-Rich Proteins (GRPs), PRPs, and Leucine-Rich Repeat Extensiins (LRXs). The problem of CWP classification appears again with LRXs which could also be involved in signaling or be classified among proteins having predicted interaction domains (Baumberger et al., 2001; Leiber et al., 2010). No extensin has been identified in the published cell wall proteomes probably because they are covalently cross-linked (Wilson and Fry, 1986). The strategies presently used for cell wall proteomics fail to efficiently isolate such proteins. The class of CWPs with interaction domains (11.0%) presently groups proteins having predicted carbohydrate binding domains, Leucine-Rich-Repeat (LRR) domains assumed to be involved in protein–protein interactions and enzyme inhibitors. This class can be split according to these three categories of CWPs (Catalá et al., 2011). A better knowledge of the function of proteins interacting with polysaccharides will also contribute to a more precise classification. The group of miscellaneous proteins (11.0%) is the Achilles’heel of the classification since it comprises all the proteins which cannot be put elsewhere. A few protein families emerge from this group like purple acid phosphatases (Wang et al., 2011), phosphate-induced proteins (Farrar et al., 2003) and germins (Membré et al., 2000). Finally, about one eighth of the cell wall proteomes correspond to proteins of yet unknown function with no predicted functional domain or a predicted domain of unknown function (DUF). This is a puzzling class of proteins which will probably reveal new functions in cell walls. It is expected to disappear when these proteins are characterized.

**WallProtDB, A CELL WALL PROTEOMICS DATABASE**

The classification of proteins described above has been used to build up the WallProtDB database (http://www.polebio.scv.ups-tlse.fr/WallProtDB/) (Pont-Lezica et al., 2010; Ligat et al., 2011). The 20 published A. thaliana cell wall proteomes listed in Table 1 (described in more details in Table A1) have been subjected to the same bioinformatics software pipeline (ProtAnnDB) in order to compare them more accurately. The number of CWPs identified in these proteomes is very variable, ranging from 6 to 137. The less populated proteomes are a leaf apoplast proteome (Haslam et al., 2003), the AGP proteome (Schultz et al., 2004) and those focused on proteins which level of accumulation changes in response to a treatment (Ndimba et al., 2003; Casasoli et al., 2008; Tran and Plaxton, 2008). On the contrary, the most populated proteomes are those relying on efficient CWP extraction (Boudart et al., 2005) or separation (Minic et al., 2007; Irshad et al., 2008; Zhang et al., 2011), or on the most sensitive MS techniques (Bayer et al., 2006). In addition to the A. thaliana cell wall proteomes (500 CWPs), the cell wall proteomes of O. sativa and a B. oleracea xylem sap proteome have been included in WallProtDB, thus representing about 1000 CWPs.

**THE CASE OF NON-CANONICAL CWPs**

Apart from the proteins having predicted signal peptides, all the cell wall proteomes contain proteins which are not predicted to be secreted and proteins predicted to be endoplasmic reticulum resident proteins. The proportion of these non-canonical CWPs varies from none in the case of the AGP and GPI-anchored proteomes (Borner et al., 2003; Schultz et al., 2004) to 87% (Bayer et al., 2006), the average being 30% (Table 1). The cell wall proteomes containing the lowest proportion of non-canonical CWPs have been obtained with the following strategies: secretome analysis (Charmont et al., 2005), extraction of apoplastic fluids with salt solutions (Boudart et al., 2005), affinity chromatography on lectins to isolate glycoproteins (Minic et al., 2007; Catalá et al., 2011; Zhang et al., 2011) and cell wall purification with an adapted protocol followed by extraction of proteins with salt solutions (Feiz et al., 2006). Apart from the limitations of bioinformatics programs to predict sub-cellular localization (Imai and Nakai, 2010), the difficulties mentioned above to preserve membrane integrity or to purify cell walls have to be taken into account to understand these contrasting results. Moreover, the facts that the percentage of non-canonical CWPs varies between experiments and that these proteins are not always the same, indicate that most of them are probably intracellular contaminants. However, it cannot be excluded that some of them are present in cell walls. This point has been recently reviewed (Rose and Lee, 2010).

Several authors used prediction of sub-cellular localization with SecretomeP which performs *ab initio* predictions of non-classical, i.e., not signal peptide-triggered, protein secretion for mammalian proteins (http://www.cbs.dtu.dk/services/SecretomeP/) (Bendtsen et al., 2004). However, this software is not well-adapted to plant proteins since it has been designed for mammalian proteins. Moreover, only a small proportion of the non-canonical proteins identified in cell wall proteomics studies
gave a score above threshold (Jamet et al., 2008; Pechanova et al., 2010; Bhushan et al., 2011; Fernández et al., 2012).

Unfortunately, experimental data are too scarce to confirm the localization of all the non-canonical CWPs. In animal cells, several alternative mechanisms of protein secretion have been proposed and partly demonstrated (Nickel and Seedorf, 2008). Unconventional secretory proteins seem to share several common features like (i) no leader sequence, (ii) absence of PTMs specific for ER or Golgi apparatus, and (iii) secretion not affected by brefeldin A which blocks the classical ER/Golgi-dependent secretion pathway. A jacalin-related lectin has been first identified in extracellular fluids of sunflower seedlings and then demonstrated to be extracellular by immunolocalization (Pinedo et al., 2012). This is the only case where the three criteria defined for animal unconventional secretory proteins described above were met, leading to the assumption of the release of exosomes to the extracellular matrix (Regente et al., 2012). In addition, a few moonlighting proteins were described like a rice α-amylase (GH13 family) which was shown to be present in both cell walls and plastids (Chen et al., 2004).

There is an urgent need for systematic localization of plant proteins by (i) biochemical strategies like isotope tagging successfully used for the membrane organelle proteome of A. thaliana (Dunkley et al., 2006), (ii) immunolocalization to get a reliable protein atlas as done in the Human Protein Atlas project (Lundberg and Uhlén, 2010), (iii) green fluorescent protein (GFP) tagging (Heazlewood et al., 2005) or (iv) a yeast secretion trap assay (Lee and Rose, 2012). An interesting tool is SUBA3 (SUBcellular localization database for Arabidopsis proteins) which collects bioinformatics and experimental data of sub-cellular localization (http://suba.plantenergy.uwa.edu.au/) (Heazlewood et al., 2007). The precise identification of secreted proteins devoid of predicted signal peptide will allow the demonstration of the existence of alternative secretion pathways in plant cells and the design of bioinformatics software able to predict non-classical secretion of plant proteins.

PRESENT CHALLENGES: OVERCOMING TECHNOLOGICAL BOTTLENECKS

Plant cell wall proteomics has been a very active research area during the last 10 years, and is rapidly expanding with the availability of new genome sequences. However, knowledge on plant CWPs will gain new insight thanks to new methodological and technological developments aiming at the identification of low-abundant proteins, the characterization of protein–protein complexes and the description of PTMs.

TOWARD COMPLETE CELL WALL PROTEOMES

Proteomics studies aim at providing a global description of proteins present in a biological extract. However, the complexity of protein samples renders difficult their exhaustive analysis since (i) a few highly-abundant proteins can mask low-abundant proteins, and (ii) the dynamic range of proteins can be very broad, i.e., up to 12 orders of magnitude (Corthals et al., 2000). To overcome these limitations, new separation techniques have been developed, namely depletion and equalization methods. These methods have first proven their efficiency for mammalian and microbial systems and are now emerging for plants. Plant depletion methods described so far mostly concern the depletion of storage proteins or ribulose-1,5-biphosphate carboxylase/oxygenase (RUBISCO). A fast and simple fractionation technique to precipitate legume seed storage proteins has been developed, allowing the detection of 541 low-abundant proteins of G. max seeds after a 2D-E separation (Krishnan et al., 2009). A similar approach has been carried out to precipitate RUBISCO from soybean leaf soluble protein extract, permitting the detection of 230 new protein spots (Krishnan and Natarajan, 2009). Another RUBISCO depletion method based on immunocapture has been successfully performed to detect low-abundant proteins differentially regulated during A. thaliana defense (Widijaya et al., 2009). Even if storage proteins are not CWPs, they can be found as contaminants in specific cell wall proteomes, like seed cell wall proteomes (Merah et al., unpublished data). These depletion methods should then be useful to remove such major contaminants and improve the identification of low-abundant CWPs. Even more interesting and relevant for cell wall proteomics studies, is the now equalization technology based on the use of combinatorial hexapeptide ligand libraries (CPLLs) to reduce the dynamic range of protein concentrations (Fröhlich and Lindermayr, 2011). CPLLs consist in 64 millions different peptides fixed to a single bead commercially available. Specific binding of proteins depends on the physico-chemical properties of each protein. Highly-abundant proteins quickly saturate their ligands whereas all low-abundant ones are bound, resulting after elution, in a narrower dynamic range of all the proteins initially present. First successful results have been obtained for proteomics of spinach leaves (Fasoli et al., 2011), leaf extracts of A. thaliana, and phloem exudates of pumpkin (Fröhlich et al., 2012). Protein extraction is a critical step since native conformation of proteins is required for interaction with CPLLs. Notwithstanding the different limitations of this technique (Fröhlich and Lindermayr, 2011), it could be applied to the study of plant cell wall proteomes providing sufficient amounts of proteins are obtained. This approach would undoubtedly permit to identify low-abundant proteins.

PROTEIN–PROTEIN INTERACTIONS IN CELL WALLS

Many protein/protein interactions are expected in plant cell walls. Indeed, enzymes and their inhibitors like proteases and protease inhibitors, PMEs and PME inhibitors (PMEIs), or proteins with LRR domains have been detected in cell wall proteomes (Jamet et al., 2008). However, the present knowledge on plant CWPs suffers from the lack of data on protein–protein interaction mapping since most of the protein extraction methods used did not preserve supramolecular assemblies. One of the future challenges in plant cell wall proteomics will consist in developing extraction and capture methods to analyze CWP complexes. Concerning the purification of protein complexes, tandem affinity purification (TAP), in combination with MS, has become the method of choice to explore in vivo protein interactions (Xu et al., 2010). This method is based on the expression of a target protein fused to a double affinity tag. The first successful study of nuclear and cytoplasmic plant protein complexes using the TAP
method has been carried out in a transient expression system of *N. benthamiana* (Rohila et al., 2004). The method has been optimized for use in plants. Since this first report, only a limited number of plant complexes through TAP have been reported from *A. thaliana* and *O. sativa* (Rohila et al., 2009; Andrès et al., 2011). It would be of special interest to carry out this method to analyze CWPs with predicted protein interaction domains, thus permitting to identify their partners. Further optimization will be necessary for (i) CWP extraction, possibly associated with protein cross-linking treatment, and (ii) protein complex capture which will require the design of a new TAP tag to preserve the level of accumulation of CWPs as well as their localization, stability, and function. Alternatively, the analysis of intact CWP assemblies could be conducted by applying low energy MS methods preserving non-covalent interactions developed in the frame of the analysis of mammalian or microbial protein complexes (Stengel et al., 2012). Such approaches should provide a more detailed description of plant CWP complex architecture.

**NEW MS TOOLS TO IMPROVE CELL WALL PROTEOME DESCRIPTION**

Overcoming the future challenges in plant cell wall proteomics including analysis of low-abundance proteins, PTMs, protein–protein interactions, and quantitative proteomics will be facilitated by significant advances in MS technologies (Thelen and Miernyk, 2012). MS instrumentation evolves very quickly and impressive improvement in sensitivity, mass accuracy and fragmentation has been achieved in recent years. Instruments like Fourier Transformed-Ion Cyclotron Resonance (FT-ICR) are capable of mass accuracy of less than 2 ppm and have a high resolution (above 10⁶). Sensitivity of new generation MS instruments reaches the femtomole or the attomole range. New fragmentation methods such as electron capture dissociation (ECD) and electron transfer dissociation (ETD) (Bond and Kohler, 2007) are also very promising. They will provide new insight into the structure of CWPs, as recently achieved for the AGP31, an *A. thaliana* cell wall O-glycoprotein (Hijazi et al., 2012). Finally, progresses in bioinformatics will be very helpful to characterize cell wall glycoproteins. Several computer programs like GlycoMod (Cooper et al., 2001), GlycodelG™ (Joshi et al., 2004), GlycoMiner (Ozohonakis et al., 2008), or Peptoonist (Goldberg et al., 2007) have been developed, but most of them do not consider plant glycan specificities. The ProTerNyc software has been developed in this purpose and efficiently used to predict N-glycan motifs on cell wall glycoproteins (Albenne et al., 2009; Zhang et al., 2011). However, additional bioinformatics tools should be developed to improve automatic data interpretation.

**BEYOND CELL WALL PROTEOMICS**

**IDENTIFICATION OF CANDIDATE PROTEINS AND SEARCH FOR FUNCTION**

In addition to the basic work of protein identification resulting in lists of proteins, cell wall proteomics has become a new tool to identify candidate proteins involved in developmental processes or in response to environmental cues. Some examples involve quantitative analyses. Up to now, label-free techniques have been favored like quantification of stained spots on polyacrylamide gels (Ndimba et al., 2003; Oh et al., 2005; Tran and Plaxton, 2008), spectral counting (Irshad et al., 2008) or calculation of area under the curve (AUC) (Cheng et al., 2009). Only one study has been performed with the difference in gel electrophoresis (DIGE) technique which requires the labeling of proteins with fluorescent dyes prior to electrophoresis (Casasoli et al., 2008). Quantifications performed on stained spots are difficult to interpret since some proteins are present in different spots for different reasons such as presence of PTMs, degradation or maturation of proteins. In addition, contrary to staining with fluorescent molecules like Sypro® Ruby, staining with Coomassie blue or silver nitrate has a narrow dynamic range, i.e., about two orders of magnitude (Moritz and Meyer, 2003). Only a few of these proteomics studies has given rise to functional or structural studies of proteins. A protein containing a GDSL motif lipase/hydrolase (GLIP1) has been identified as one of the salicylic acid (SA) responsive proteins secreted by *A. thaliana* cell suspension cultures (Oh et al., 2005). The increase in protein level was calculated to be three-fold after comparison of proteins extracted from control and from SA-treated cells separated by 2D-E and stained with silver nitrate. Two *glp1* T-DNA insertion mutants have been found to be more resistant to the *Alternaria brassicicola* necrotrophic fungus. It has also been shown that the recombinant GLIP1 protein has a lipase activity and an antimicrobial activity able to disrupt the integrity of fungal spores. The AGP AGP31 has been identified as a major protein in the cell wall proteome of etiolated hypocotyls of *A. thaliana* (Irshad et al., 2008). AGP31 is a multi-domain protein having a N-terminal AGP, a central Pro-rich and a C-terminal Cys-rich domains. The combination of several MS technologies has allowed the first description of the Pro hydroxylation and O-glycosylation patterns of its Pro-rich domain (Hijazi et al., 2012). Finally, the *N. tabacum* NtSCP1 serine carboxypeptidase III identified in leaf intercellular fluids has been later shown to be involved in cell elongation (Delannoy et al., 2008; Bienert et al., 2012). The protease activity of NtSCP1 has been demonstrated in *vitro* and its cell wall localization has been confirmed by expression of the protein as a GFP fusion protein in *vivo*. Over-expression of NtSCP1 has led to reduce flower length due to decrease in cell size and to etiolated seedlings with short hypocotyls.

In addition to proteomics data, it would be interesting to consider other data to identify proteins of interest such as transcriptomics or gene regulatory networks. Such data are available online (e.g., http://bar.utoronto.ca/welcome.htm, https://www. geneinvestigator.com/gv/, http://aranet.mpimp-golm.mpg.de, http://atted.jp/). The feeding of new portals like MASCP Gator which aims at unifying the *A. thaliana* proteomics resources in a single interface for the research community is also essential (Joshi et al., 2011). This systems biology approach would allow a better understanding of gene regulation, from gene transcription to protein synthesis and even PTMs, allowing the description of protein active forms. Indeed, several studies have shown that proteomics and transcriptomics data are complementary and do not give exactly the same picture of a physiological situation depending on the level of regulation of gene expression (Jamet et al., 2009; Minic et al., 2009).
FROM PROTEOMICS TO PEPTIDOMICS

During the last decade, it has become evident that secreted peptides function as signaling molecules in cell-to-cell communication in plants. Recently, they have been recognized as hormones that coordinate and specify cellular functions in complex developmental processes (Shinohara and Matsubayashi, 2010; Murphy et al., 2012). The secreted signaling peptides identified so far can be categorized in two groups: (i) the small post-translationally modified peptides are less than 20 amino acids and undergo extensive proteolytic processing from a longer precursor and PTMs such as Tyr sulfation, Pro hydroxylation and arabinosylation on Hyp; (ii) the Cys-rich peptides are larger (~160 amino acids), cationic at the extracellular pH and have multiple intramolecular disulfide bonds. All of them have a predicted signal peptide. Peptidomics is becoming a stimulating field especially because of the description of the active forms of the signaling peptides (Shinohara and Matsubayashi, 2010; Murphy et al., 2012). Indeed, the PTMs are essential for their biological activity. For instance, STOMAGEN, a Cys-rich peptide that positively regulates stomatal density in *A. thaliana*, is active at nanomolar concentrations when forming disulphide bonds. When Cys residues were replaced by Ser residues, STOMAGEN was unable to increase stomatal density even at very high (10 μM) concentrations (Ohki et al., 2011). Tyr sulfation and arabinosylation were also required for the full activity of small post-translationally modified peptides (Ohyama et al., 2009; Matsuzaki et al., 2010). Like cell wall proteomics, peptidomics has the potential to reveal new secreted signaling peptides as well as new functions of plant cell walls.

Today, the description of cell wall peptidomes, or secreted peptidomes, defined as a set of peptides present in cell walls at a specified physiological state, is lacking. Two main reasons can explain such a gap. First, the signaling peptides are believed to be present in very low quantity in plant tissues, they are active at nanomolar concentrations and their transcripts have been found to be transiently expressed (Ito et al., 2006; Chen et al., 2008b). Most of the well-characterized signaling peptides have been identified by genetics and in silico approaches (Murphy et al., 2012). In order to fully characterize the structure of their mature forms, they have been produced by cells or plants over-expressing the corresponding genes to obtain sufficient amounts of peptides amenable to LC-MS-based structure analysis (Amano et al., 2007; Ohyama et al., 2008, 2009; Sugano et al., 2010) or to in situ MALDI-TOF-MS analysis (Kondo et al., 2006). The latter study used *A. thaliana* plants over-expressing CLV3, a gene encoding a 96 amino acid propeptide containing a signal peptide. CLV3 has been shown to play a role in the control of the size of the shoot apical meristem. The identified mature peptide contained 12 amino acids from Arg70 to His84 in CLV3 in which two of three Pro residues were modified to Hyp. Nevertheless, a number of studies that employed LC purification and Edman sequencing or MS identification have been developed and applied successfully to the analysis of native peptide sources (Pearce et al., 1991; Matsubayashi and Sakagami, 1996; Ito et al., 2006; Chen et al., 2008b). When these studies used high amounts of plants as starting material, they have allowed the identification of mature signaling peptides active at a concentration of 10^{-11} M in the case of TDIF (Tracheary element Differentiation Inhibitory Factor) (Ito et al., 2006). The corresponding cDNA of *Zinnia elegans* encodes a protein of 132 amino acids, but only 12, from His^{120} to Asn^{131}, match the TDIF sequence with two Hyp residues (Hyp^{123} and Hyp^{126}).

Second, peptide-encoding genes are frequently overlooked during the annotation of genomes. Indeed, gene prediction programs hardly distinguish between short, often intronless peptide-encoding genes and random small open reading frames (ORFs). To minimize incorrect gene predictions, it is common that small ORFs are rejected (Olsen et al., 2002). To overcome such a deficiency, a bioinformatics approach has been undertaken to identify candidate peptide-encoding genes in the *A. thaliana* genome (Lease and Walker, 2006). It has led to an unannotated secreted peptide database containing 33,809 ORFs. The identified peptides have been characterized by the presence of a predicted N-terminal signal peptide and by the absence of transmembrane domains and ER retention sequences. Since the expression of some ORFs has been detected by RT-PCR, it is suggested that the number and diversity of plant peptides is broader than currently assumed (Lease and Walker, 2006). The secreted peptide database will permit the necessary retrieval of information required for the identification of *A. thaliana* signaling peptides. Together with the progress of MS sensitivity, cell wall peptidomics is now a reachable objective.

CONCLUSION: UP AND COMING OF CELL WALL PROTEOMICS

Within the last 10 years, cell wall proteomics studies have received full credit among the OMICS strategies. They have allowed not only the precise identification of proteins in particular physiological conditions, but also their quantification and the characterization of their PTMs. Proteomics could also provide information about the dynamics of CWPs by kinetics analysis to follow the de novo synthesis of proteins or their degradation during plant development or in response to environmental cues. All the knowledge presently available on cell wall proteomics contributes to a better understanding of CWP structures and functions in cell walls. However, it is not yet possible to distinguish proteomes of primary and secondary walls notably because it is difficult to separate the cells surrounded by either of them. Micro-dissection of tissues should help solving this problem providing enough material can be obtained, but the extraction of proteins from the intricate macromolecular networks of secondary walls remains a great challenge. Next development will take advantage of cutting-edge MS technologies for a better coverage of cell wall proteomes, a more precise description of protein forms and protein complexes and for an insight into cell wall peptidomics.

ACKNOWLEDGMENTS

The authors are thankful to Université Paul Sabatier (Toulouse, France) and CNRS for supporting their research work. The authors also wish to thank Hélène San Clemente for her contribution to the bioinformatics work. This work has been done at LRSV, part of the “Laboratoire d’Excellence” (LABEX) entitled TULIP (ANR -10-LABX-41). The authors wish to apologize for the papers not quoted in this review, due to lack of space.
proteomics of chickpea genotypes with contrasting tolerance. J. Proteome Res. 10, 2027–2046.

Bhushan, D., Pandey, A., Choudhary, A., Choudhary, M. K., Chakraborty, S., Datta, A., et al. (2006). Extracellular matrix proteome of chickpea (Cicer arietinum L.) illustrates pathway abundance, novel protein functions and evolutionary perspective. J. Proteome Res. 5, 1710–1720.

Bienert, M. D., Delannoy, M., Navarre, C., and Boudart, G. (2012). NCSP1 from tobacco is an extracellular serine carboxypeptidase III that has an impact on cell elongation. Plant Physiol. 158, 1220–1229.

Bond, M. R., and Kohler, J. J. (2007). Chemical methods for glycoprotein discovery. Curr. Opin. Chem. Biol. 11, 52–58.

Borders, G., Jamet, E., Lafitte, C., Rossignon, M., Jeanneau, A., Boudart, G., et al. (2003). Proteomes of loosely bound cell wall proteins of Arabidopsis thaliana cell suspension cultures: a critical analysis. Electrophoresis 24, 3421–3432.

Borner, G. H., Lilley, K. S., Stevens, T. J., and Dupree, P. (2003). Identification of glycosylphosphatidylinositol-anchored proteins in Arabidopsis. A proteomic and genomic analysis. Plant Physiol. 132, 568–577.

Boudart, G., Jamet, E., Rossignon, M., Lafitte, C., Borders, G., Jeanneau, A., et al. (2005). Cell wall proteins in apoplastic fluids of Arabidopsis thaliana rosettes: identification by mass spectrometry and bioinformatics. Proteomics 5, 212–221.

Bradly, J. D., Siddler, I. H., and Fry, S. C. (1996). Di-isodityrosine, a novel tetramer derivative of tyrosine in plant cell wall proteins: a new potential cross-link. Biochem. J. 315, 323–327.

Brisson, L. F., Tenhaken, R., and Lamb, C. (1994). Function of oxidative cross-linking of cell wall structural proteins in plant disease resistance. Plant Cell 6, 1703–1712.

Carpin, S., Crevecoeur, M., De Meyer, M., Simon, P., Greppin, H., and Penel, C. (2001). Identification of a Ca(2+)-pectate binding site on an apoplastic peroxidase. Plant Cell 13, 511–520.

Carpita, N. C., and Gibeaut, D. M. (1993). Structural models of primary cell walls in flowering plants, consistency of molecular structure with the physical properties of the walls during growth. Plant J. 3, 1–30.

Casamoli, S., Spadoni, S., Lilley, K. S., Cervone, F., De Lorenzo, G., and Mattei, B. (2008). Identification by 2-D DIGE of apoplastic proteins regulated by oligogalacturonides in Arabidopsis thaliana. Proteomics 8, 1042–1054.

Cassab, G. L., and Varner, J. E. (1988). Cell wall proteins. Annu. Rev. Plant Physiol. Plant Biol. 39, 321–353.

Catalá, C., Howe, K. J., Hucko, S., Rose, J. K., and Thannhauser, T. W. (2011). Towards characterization of the glycoproteome of tomato (Solanum lycoerum) fruit using Concanavalin A lectin affinity chromatography and LC-MALDI-MS/MS analysis. Proteomics 11, 1530–1544.

Charmont, S., Jamet, E., Pont-Lezica, R., and Canut, H. (2005). Proteomic analysis of secreted proteins from Arabidopsis thaliana seedlings: improved recovery following removal of phenolic compounds. Phytochemistry 66, 453–461.

Chen, M. H., Huang, L. F., Li, H. M., Chen, Y. R., and Yu, S. M. (2004). Signal peptide-dependent targeting of a rice alpha-amylase and cargo proteins to plastids and extracellular compartments of plant cells. Plant Physiol. Biochem. 135, 1367–1377.

Chen, X. Y., Kim, S. T., Cho, W. K., Rim, Y., Kim, S., Kim, S. W., et al. (2008a). Proteomics of weakly bound cell wall proteins in rice cells. J. Plant Physiol. 166, 665–685.

Chen, Y. C., Siems, W. F., Pearce, G., and Ryan, C. A. (2008b). Six peptide wound signals derived from a single precursor protein in Ipomoea batatas leaves activate the expression of the defense gene sporamin. J. Biol. Chem. 283, 11469–11476.

Cheng, F.-Y., Blackburn, K., Lin-Y.-M., Goshe, M. B., and Willamson, J. D. (2009). Absolute protein quantification by LC/MS for global analysis of salicylic acid-induced plant protein secretion responses. J. Proteome Res. 8, 82–93.

Chivasa, S., Ndima, B. K., Simon, W. J., Robertson, D., Yu, X.-L., Knox, J. A., et al. (2006). Dehydration-responsive reversible and irreversible changes in the extracellular matrix: comparative proteomics research. Electrophoresis 21, 1104–1115.

Coggrove, D. J. (2005). Growth of the plant cell wall. Nat. Rev. Mol. Cell. Biol. 6, 825–861.

Dafne, N. I., and Constabel, P. (2009). Proteomic analysis of hybrid poplar xylem sap. Phytochemistry 70, 856–863.

Dani, V., Simon, W. J., Duranti, M., and Croy, R. R. (2005). Changes in the tobacco leaf apoplastic proteome in response to salt stress. Proteomics 5, 737–745.

Delauney, M., Alves, G., Vertommen, D., Ma, J., Bountray, M., and Navarre, C. (2008). Identification of peptides in Nicotiana tabacum leaf intercellular fluid. Proteomics 8, 2285–2298.

Dunkley, T. P., Hester, S., Shafroth, I. P., Runicions, J., Weimar, T., Hanton, S. L., et al. (2006). Mapping the Arabidopsis organelle proteome. Plant Physiol. Biochem. U.S.A. 103, 6518–6523.

Farrar, K., Evans, I. M., Topping, J. F., Souter, M. A., Nielsen, J. E., and Lindsey, K. (2003). EXORDIUM—a gene expressed in proliferating cells and with a role in meristem function, identified by promoter trapping in Arabidopsis. Plant J. 33, 61–73.

Fasel, E., D’Amato, A., Kravchuk, A. V., Boschetti, E., Bachi, A., and Righetti, P. G. (2011). Popeye strikes again: the deep proteome of spinach leaves. J. Proteomics 74, 127–136.

Faye, L., Boulaflous, A., Benchabane, M., Gomord, V., and Michaud, D. (2005). Protein modifications in the plant secretory pathway: current status and practical implications in molecular pharming. Vaccine 23, 1770–1778.

Feiz, L., Irshad, M., Pont-Lezica, R. F., Canut, H., and Jamet, E. (2006). Evaluation of cell wall preparations for proteomics: a new procedure for purifying cell walls from Arabidopsis hypocotyls. Plant Methods 2, 10.

Fernández, M. B., Pagano, M. R., Dáleo, G. R., and Guevara, M. G. (2012). Hydrophobic proteins secreted into the apoplasm may contribute to resistance against Phytophthora infestans in potato. Plant Physiol. Biochem. 60, 59–66.

Fröhlich, A., Gaupels, F., Sarioglu, H., Holzmeister, C., Spannagl, M., Duran, J. E., et al. (2012). Looking deep inside: detection of low-abundance proteins in leaf extracts of Arabidopsis and phloem exudates of pumpkin. Plant Physiol. 159, 902–914.
Frohlich, A., and Lindermayr, C. (2011). Deep insights into the plant proteome by pretreatment with combinatorial hexapeptide ligand libraries. J. Proteomics 74, 1182–1189.

Fry, S. C. (2004). Primary cell wall metabolism: tracking the careers of wall polymers in living plant cells. New Phytol. 161, 641–675.

Goldberg, D., Bern, M., Parry, S., Sutton-Smith, M., Panico, M., Morris, H. R., et al. (2007). Automated N-glycopeptide identification using a combination of single- and tandem-MS. J. Proteome Res. 6, 3995–4005.

Goulet, C., Goulet, C., Goulet, M.-C., Frontiers in Plant Science (2010). International Rice Genome Sequencing Project. (2005). The map-based reference maps. Plant Physiol. 139, 598–609.

Hadianto, P., Drees, A. L., Raventon, M., Gancedo, K., Job, D., Pellet, K. E., et al. (2003). The assessment of enriched apoplastic extracts using proteome approaches. Annu. Appl. Biol. 143, 81–91.

Heazlewood, J., Verboom, R., Tonti-Filippini, I., Small, I., and Millar, A. (2007). SUBA: the Arabidopsis sub-cellular database. Nucleic Acids Res. 35, D213–D218.

Heazlewood, J. L., Tonti-Filippini, I., Verboom, R. E., and Millar, A. H. (2005). Combining experimental and predicted datasets for determination of the subcellular location of proteins in Arabidopsis. Plant Physiol. 139, 598–609.

Hijazi, M., Durand, J., Pichereaux, C., Pont, E., Jamet, E., and Albenne, C. (2012). Characterization of the Arabinogalactan Protein 31 (AGP31) of Arabidopsis thaliana: new advances on the Hyp-O-glycosylation of the Pro-rich domain. J. Biol. Chem. 287, 9623–9632.

Hunt, L., Bailey, K. J., and Gray, J. E. (2010). The signalling peptide EFL9 is a positive regulator of stomatal development. New Phytol. 183, 609–614.

Imai, K., and Nakai, K. (2010). Prediction of subcellular locations of proteins: where to proceed? Proteomics 10, 3970–3983.

International Rice Genome Sequencing Project. (2005). The map-based sequence of the rice genome. Nature 436, 793–805.

Ishrd, M., Canut, H., Borderics, G., Pont-Lezica, R., and Jamet, E. (2008). A new picture of cell wall protein dynamics in elongating cells of Arabidopsis thaliana: confirmed actors and newcomers. BMC Plant Biol. 8:94. doi: 10.1186/1471-2229-8-94

Ito, Y., Nakamoto, Y., Motose, H., Iwamoto, K., Sawa, S., Doimae, N., et al. (2006). Dodeca-CLE peptides as suppressors of plant stem differentiation. Science 313, 8842–8845.

Janet, E., Albenne, C., Boudart, G., Irshad, M., Canut, H., and Pont-Lezica, R. (2008). Recent advances in plant cell wall proteomics. Proteomics 8, 895–908.

Janet, E., Canut, H., Boudart, G., and Pont-Lezica, R. F. (2006). Cell wall proteins: a new insight through proteomics. Trends Plant Sci. 11, 33–39.

Kondo, T., Sawa, S., Kinoshita, A., Mizuno, S., Kakimoto, T., Fukuda, H., et al. (2006). A plant peptide encoded by CLV3 identified by in situ MALDI-TOF MS analysis. Science 313, 845–848.

Krishnan, H. B., and Natarajan, S. S. (2009). A rapid method for depletion of Rubisco from soybean (Glycine max) leaf for proteomic analysis of lower abundance proteins. Phytochemistry 70, 1958–1966.

Kumarasamy, S., Nanao, Y., and Furukawa, K. (2011). A new picture of the site-directed removal of N-glycosylation from cationic peanut peroxidase on its function. J. Biochem. Biophys. 86, 17–24.

Lim, J.-M., Aoki, K., Angel, P., Garrison, D., King, D., Tiemeyer, M., et al. (2009). Mapping glycans onto specific N-linked glycosylation sites of Pyrus communis PGIIP redresses the interface for EPG-PGIIP interactions. J. Proteome Res. 8, 673–680.

Lundberg, E., and Uhlen, M. (2010). Creation of an antibody-based sub-cellular protein atlas. Proteomics 10, 3984–3996.

Matsubayashi, Y., and Sagakami, Y. (1996). Phytoallolofoline, sulfated peptides that induce the proliferation of single mesophyll cells of Asparagus officinalis L. Proc. Natl. Acad. Sci. U.S.A. 93, 7623–7627.

Matsuzaki, Y., Ogawa-Ohmihi, M., Mori, A., and Matsubayashi, Y. (2010). Secreted peptide signals required for maintenance of root stem cell niche in Arabidopsis. Science 329, 1065–1067.

Menéndez, N., Bernier, E., Stäger, D., and Berna, A. (2000). Arabidopsis thaliana germin-like proteins: common and specific features point to a variety of functions. Planta 211, 345–354.

Millar, D. J., Whitelegge, J. P., Bindschedler, L. V., Rayon, C., Boudet, A. M., Rossignol, M., et al. (2009). The cell wall and secretory proteome of a tobacco cell line synthesising secondary wall. Proteomics 9, 2355–2372.

Minic, Z., Jamet, E., Negroni, L., Der Garabedian, P. A., Zivy, M., and Jouanin, L. (2007). A sub-proteome of Arabidopsis thaliana trapped on Concanavalin A is enriched in cell wall glycoside hydrolases. J. Exp. Bot. 58, 2503–2512.

Minic, Z., Jamet, E., Albenne, C., Sen Clemente, H., Valot, B., Zivy, M., et al. (2011). Analysis of the xylem sap proteome of Brassica oleracea reveals a high content in secreted proteins. Proteomics 11, 1798–1813.

Ligat, L., Lauber, E., Albenne, C., San Clemente, H., Valot, B., Zivy, M., et al. (2011). Analysis of the xylem sap proteome of Brassica oleracea reveals a high content in secreted proteins. Proteomics 11, 1798–1813.

Lige, B., Shengwu, M., and Van Huystee, R. B. (2001). The effects of the site-directed removal of N-glycosylation from cationic peanut peroxidase on its function. Arch. Biochem. Biophys. 386, 17–24.

Lee, S.-J., and Rose, J. K. C. (2012). Extensin: repetitive domain. Annu. Rev. Cell Dev. Biol. 28, 507–530.

Leasure, C. D., and He, Z. H. (2012). A new picture of lectin: sequence, structure, and evolution. Annu. Rev. Plant Biol. 63, 489–518.

Leslie, C. R., John, E., Verberbruggen, Y., Diet, A., Knox, J. P., and Ringli, C. (2010). The TOR pathway modulates the structure of cell walls in Arabidopsis. Plant Cell 22, 1898–1908.
genes. BMC Plant Biol. 9:6. doi: 10.1186/1471-2229-9-6

Minic, Z., Rihouey, C., Do, C. T., Lerouve, P., and Jouanin, L. (2004). Purification and characterization of enzymes exhibiting δ-D-xylodiase activities in stem tissues of Arabidopsis. Plant Physiol. 135, 867–878.

Moritz, B., and Meyer, H. E. (2003). Analysis of the secretory pathway of Arabidopsis thaliana. In Silico Biol. 1, 1–10.

Morel, J., Claverol, S., Mongrand, S., Mor, C., and Gascuel, O. (2005). Secretome analysis reveals that extracellular matrix proteins interact with pectin through glycosylphosphatidylinositol-anchor motifs. J. Proteomics 68, 131–140.

Ohki, S., Takeuchi, M., and Mori, M. (2011). Identification of a biologically active small, secreted peptide in Arabidopsis thaliana. Nat. Chem. Biol. 5, 578–580.

Olsen, A. N., Mundy, J., and Skriver, K. (2002). Peptidomics, identification of novel cationic Arabidopsis peptides with conserved sequence motifs. In Silico Biol. 2, 441–451.

Ozonoehan, O., Krenzayc, L., Ludányi, K., Polreisz, F., Vékey, K., and Drahos, L. (2008). GlycoMiner: a new software tool to elucidate glycoprotein composition. Rapid Commun. Mass Spectrom. 22, 3245–3254.

Passardi, F., Penel, C., and Dunand, C. (2004). Performing the paradoxical: how plant peroxidases modify the cell wall. Trends Plant Sci. 9, 534–540.

Pearce, G., Stridom, D., Johnson, S., and Ryan, C. A. (1991). A polypeptide from tomato leaves induces wound-inducible proteinase inhibitor proteins. Science 253, 895–898.

Pechanová, O., Hsu, C.-Y., Adams, J. P., Pechan, T., Vandervelde, L., Drnevich, J., et al. (2010). Aplastop protein reveals that extracellular matrix contributes to multistress response in poplar. BMC Genomics 11:212. doi: 10.1186/1471-2164-11-212.

Pinedo, M., Regente, M., Elizalde, M., Quiroga, I. Y., Pagnussat, L., and Estevez, J. S., Gerrish, C., Bolwell, G. P., and Fromm, M. (2004). Improved purification and characterization of secreted plant proteins and complexity in the plant cell wall proteome. Plant Physiol. Biochem. 42, 343–436.

Ruiz-May, E., Thanhauser, T. W., Zhang, S., and Rose, J. K. C. (2012). Analytical technologies for characterization of the plant N-glycoproteome. Front. Plant Sci. 3:150. doi: 10.3389/fpls.2012.00150.

San Clemente, H., Pont-Lezica, R., and Jamet, E. (2009). Bioinformatics as a tool for assessing the quality of sub-cellular proteomic strategies and inferring functions of protein: plant cell wall proteomics as a test case. Bioinformatics. Biol. Insights 3, 13–28.

Schulz, C. J., Ferguson, K. L., Lahnstein, J., and Bacic, A. (2004). Post-translational modifications of arabinogalactan-peptides of Arabidopsis thaliana. Endoplasmic reticulum and glycosylphosphatidylinositol-anchor signal cleavage sites and hydroxylation of proline. J. Biol. Chem. 279, 45503–45511.

Seifert, G. J., and Blaukopf, C. (2010). Irritable walls: the plant extracellular matrix and signaling. Plant Physiol. 153, 467–478.

Shinozaki, H., and Matsubayashi, Y. (2010). Arabinosylated glycoprotein-tide hormones: new insights into CLAVATA3 structure. Curr. Opin. Plant Biol. 13, 515–519.

Song, W., Henquet, M. G., Mentink, R. A., Van Dijk, A. J., Cordewener, J. H., Bosch, D., et al. (2011). N-glycoproteomics in plants: perspectives and challenges. J. Proteomics 74, 1643–1674.

Spadoni, S., Zbotina, O., Di Matteo, A., Mikkelsen, J. D., Cervone, F., and Lorenzo, G., et al. (2006). Polygalacturonase-inhibiting protein interacts with pectin through a binding site formed by four clustered residues of arginine and lysine. Plant Physiol. 141, 557–564.

Suzuki, M., and Nihei, K. (2002). Plant glycoproteins as cell surface recognition molecules: new insights into glycoprotein. J. Biol. Chem. 277, 15841–15848.

Ohyama, K., Ohnishi, M., and Matsubayashi, Y. (2009). A glycoprotein regulating stem cell fate in Arabidopsis thaliana. Nat. Chem. Biol. 5, 578–580.

Ohyama, K., and Ohnishi, M. (2011). Identification of a biologically active small, secreted peptide in Arabidopsis by in silico gene screening. followed by LC-MS-based structure analysis. Plant J. 55, 152–160.

Ohyama, K., Shinohara, H., Ogawa-Ohashi, M., and Matsubayashi, Y. (2009). A glycoprotein regulating clusternet library of Residue and lysine. Plant Physiol. 141, 557–564.

Spirito, R. G. (2002). Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycoprotein bonds. Glycobiology 12, 438–568.

Stangel, F., Aebersold, R., and Robinson, C. V. (2012). Joining forces: integrating proteomics and cross-linking with the mass spectrometry of intact complexes. Mol. Cell. Proteomics 11:R111.014027. doi: 10.1074/mcp.R111.014027.

Tak, S., Showalter, A. M., Gellert, J., Hernandez-Sanchez, A., Doblin, M. S., and Bacic, A. (2012). Arabinogalactan-proteins and the research challenges for these enigmatic plant cell surface proteoglycans. Front. Plant Sci. 3:140. doi: 10.3389/fpls.2012.00140.

Tak, S., Showalter, A. M., Gellert, J., Hernandez-Sanchez, A., Doblin, M. S., and Bacic, A. (2012). Arabinogalactan-proteins and the research challenges for these enigmatic plant cell surface proteoglycans. Front. Plant Sci. 3:140. doi: 10.3389/fpls.2012.00140.

Thelen, J. I., and Miemczyk, J. A. (2012). The proteomic future: where mass spectrometry should be taking us. Biochem. J. 444, 169–181.

Thomas, R., Fang, X., Ranathunge, K., Anderson, T. R., Peterson, C. A., and Bernards, M. A. (2007). Soybean root suberin: anatomical distribution, chemical composition, and relationship to partial resistance to Phytophthora sojae. Plant Physiol. 144, 299–311.

Tomato Genome Consortium. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485, 635–641.

Van Der Hoorn, R. A. (2008). Plant proteases: from phenotypes to molecular mechanisms. Annu. Rev. Plant Biol. 59, 191–223.

Velasquez, M., Salter, J. S., Dorosz, J. G., Petersen, B. L., and Estevez, J. M. (2012). Recent advances on the posttranslational modifications of EF-Ts and their roles in plant cell walls. Front. Plant Sci. 3:93. doi: 10.3389/fpls.2012.00093.

Verdonk, J., Hatfield, R., and Sullivan, M. (2012). Proteomics analysis of cell walls of two developmental
stages of alfalfa stems. *Front. Plant Sci.* 3:279. doi: 10.3389/fpls.2012.00279

Von Groll, U., Berger, D., and Altmann, T. (2002). The subtilisin-like serine protease SDD1 mediates cell-to-cell signaling during Arabidopsis stomatal development. *Plant Cell* 14, 1527–1539.

Wang, L., Li, Z., Qian, W., Guo, W., Gao, X., Huang, L., et al. (2011). The Arabidopsis purple acid phosphatase AtPAP10 is predominantly associated with the root surface and plays an important role in plant tolerance to phosphate limitation. *Plant Physiol.* 157, 1283–1299.

Watson, B. S., Lei, Z., Dixon, R. A., and Sumner, L. W. (2004). Proteomics of *Medicago sativa* cell walls. *Phytochemistry* 65, 1709–1720.

Widjaja, I., Naumann, K., Roth, U., Wolf, N., Mackey, D., Dangl, J. L., et al. (2009). Combining subproteome enrichment and Rubisco depletion enables identification of low abundance proteins differentially regulated during plant defense. *Proteomics* 9, 138–147.

Wilson, L. G., and Fry, J. C. (1986). Extenssin, a major cell wall glycoprotein. *Plant Cell Environ.* 9, 239–260.

Wolf, S., Rausch, T., and Greiner, S. (2009). The N-terminal pro region mediates retention of unprocessed type-I PME in the Golgi apparatus. *Plant J.* 58, 361–375.

Xu, X., Song, Y. R., Li, Y., Chang, J., Zhang, H., and An, L. (2010). The tandem affinity purification method: an efficient system for protein complex purification and protein interaction identification. *Protein Expr. Purif.* 72, 149–156.

Yeats, T. H., Howe, K. J., Matas, A. J., Buda, G., Thannhauser, T. W., and Rose, J. K. (2010). Mining the surface proteome of tomato (*Solanum lycopersicum*) fruit for proteins associated with cuticle biogenesis. *J. Exp. Bot.* 61, 3759–3771.

Zhang, Z., Giboulot, A., Zivy, M., Valot, B., Jamet, E., and Albenne, C. (2011). Combining various strategies to increase the coverage of the plant cell wall glycoproteome. *Phytochemistry* 72, 1109–1123.

Zhou, L., Bokhari, S. A., Dong, C.-H., and Liu, J.-Y. (2011). Comparative proteomics analysis of the root apoplasts of rice seedlings in response to hydrogen peroxide. *PLoS ONE* 6:e16723. doi: 10.1371/journal.pone.0016723

Zhu, J., Alvarez, S., Marsh, E. L., Lenoble, M. E., Cho, I.-J., Sivaguru, M., et al. (2007). Cell wall proteome in the maize primary root elongation zone. II. Region-specific changes in water soluble and lightly ionically bound proteins under water deficit. *Plant Physiol.* 145, 1533–1548.

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

Received: 19 December 2012; paper pending published: 05 March 2013; accepted: 10 April 2013; published online: 01 May 2013.

Citation: Albenne C, Canut H and Jamet E (2013) Plant cell wall proteomics: the leadership of Arabidopsis thaliana. *Front. Plant Sci.* 4:111. doi: 10.3389/fpls.2013.00111

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Albenne, Canut and Jamet. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.
### APPENDIX

Table A1 | Cell wall proteomes and secretomes of *A. thaliana*: detailed information.

| Organ                  | References          | Destructive or non-destructive method | Type of proteome                  | Buffer for protein extraction | Protein separation | Protein identification | Total number of identified proteins | Number of proteins with predicted target signal and no predicted intracellular retention signal |
|------------------------|---------------------|--------------------------------------|------------------------------------|-------------------------------|--------------------|------------------------|------------------------------------|-----------------------------------------------------------------|
| Culture medium of roots| Basu et al. (2006)  | Secretome                            | 2D-LC-MS/MS                        | 52                            | 36                 |                        | 0                    |                                                                 |
| Stems                  | Minic et al. (2007) | Destructive (grinding)                | N-glycoproteome (ConA affinity chromatography) | Salt solution (200 mM CaCl2)  | 2D-E               | LC-MS/MS               | 102                               | 90                                                              |
| Leaves                 | Haslam et al. (2003)| Non-destructive (vacuum infiltration) | Apoplastic proteome                | low ionic strength buffer      | 2D-E               | MALDI-TOF MS           | 10                                | 9                                                               |
| Leaves                 | Boudart et al. (2005)| Non-destructive (vacuum infiltration)| Apoplastic proteome               | 0.3 M mannitol or salt solutions (1 M NaCl, 200 mM CaCl2, 2 M LiCl, 50 mM CDTA), acidic pH | 1D-E or 2D-E | MALDI-TOF MS           | 93                                | 87                                                              |
| Etiolated hypocotyls (5 days) | Feiz et al. (2006) | Destructive (grinding in low ionic strength acidic buffer), cell wall purification | Cell wall proteome | Salt solutions (200 mM CaCl2, 2 M LiCl), acidic pH | 1D-E               | MALDI-TOF MS           | 96                                | 84                                                              |
| Etiolated hypocotyls (5 and 11 days) | Irshad et al. (2008)| Destructive (grinding in low ionic strength acidic buffer), cell wall purification | Cell wall proteome | Salt solutions (200 mM CaCl2, 2 M LiCl), acidic pH | Cationic exchange chromatography followed by 1D-E | MALDI-TOF MS           | 173                               | 137                                                             |
| Etiolated hypocotyls (11 days) | Zhang et al. (2011) | Destructive (grinding in low ionic strength acidic buffer), cell wall purification | Cell wall glycoproteome (lectin affinity chromatography, boronic acid chromatography) | Salt solutions (200 mM CaCl2, 2 M LiCl), acidic pH | 1D-E               | MALDI-TOF MS and LC-MS/MS | 129                               | 127                                                             |

(Continued)
| Organ | References | Destructive or non-destructive method | Type of proteome | Buffer for protein extraction | Protein separation | Protein identification | Total number of identified proteins | Number of proteins with predicted target signal and no predicted intracellular retention signal |
|-------|------------|------------------------------------|-----------------|-----------------------------|-------------------|----------------------|-----------------------------|----------------------------------|
| Seedlings (10 days) (oligogalacturonides) | Casasoli et al. (2008) | Non-destructive | Apoplastic proteome | Salt solution (150 mM MgCl₂, 50 mM EDTA), basic pH | 2D-E MALDI-TOF MS and LC-MS/MS | 72 | 20 | 14 |
| Cell suspension cultures (3 days) | Bayer et al. (2006) | Destructive (grinding in salt solution) | Cell wall proteome | Phenol buffer | 2D-LC-MS/MS | 39 | 39 | 34 |
| Cell suspension cultures (5 days) | Kwon et al. (2005) | Non-destructive (washing of cells) | Cell wall proteome | Salt solution (1 M KCl) | 2D-E MALDI-TOF MS | 72 | 72 | 26 |
| Cell suspension cultures (5 days) | Chivasa et al. (2002) | Destructive (grinding) | Cell wall proteome | Salt solution, chaotropic agent (urea and thiourea) and detergent (CHAPS) | 2D-E MALDI-TOF MS | 33 | 33 | 18 |
| Cell suspension cultures (5 days) | Robertson et al. (1997) | Non-destructive (washing of cells) | Cell wall proteome | Salt solutions (0.2 M CaCl₂, 50 mM CDTA, 1 M NaCl, 0.2 M borate) | 1D-E Edman N-terminal sequencing | 95 | 95 | 50 |
| Cell suspension cultures (7 days) | Borderies et al. (2003) | Non-destructive (washing of cells) | Cell wall proteome | Glycerol solution (50%) or salt solutions (successively 0.15 M NaCl, 1 M NaCl, 50 mM EDTA, 2 M LiCl in 50% glycerol) | 1D-E or 2D-E MALDI-TOF MS | 30 | 30 | 30 |
| Culture medium of calli (21 days) | Borner et al. (2003) | Destructive (purification of a membrane fraction with detergent) | GPI-anchored proteome | Pi-PLC treatment, detergent (Triton X-114) | 1D-E or 2D-E LC-MS/MS | 45 | 45 | 24 |
| Culture medium of cell suspension cultures (4 days) | Oh et al. (2005) | Secretome | 2D-E MALDI-TOF MS | 45 | 24 |

(Continued)
### Table A1 | Continued

| Organ | References | Destructive or non-destructive method | Type of proteome | Buffer for protein extraction | Protein separation | Protein identification | Total number of identified proteins | Number of proteins with predicted target signal and no predicted intracellular retention signal |
|-------|------------|--------------------------------------|------------------|-------------------------------|-------------------|-----------------------|---------------------------------|------------------------------------------------|
| Cell suspension cultures (7 days) (fungal elicitors) | Ndimba et al. (2003) | Destructive (grinding) | Cell wall proteome | Salt solution (200 mM CaCl2) | 2D-E | MALDI-TOF MS | 6 | 6 |
| Culture medium of cell suspension cultures (7 days) (salicylic acid treatment) | Cheng et al. (2009) | | Secretome | | | 2D-LC-MS/MS | 74 | 39 |
| Culture medium of cell suspension cultures (7 days) (phosphate deficiency) | Tran and Plaxton (2008) | | Secretome | | 2D-E | MALDI-TOF MS | 37 | 21 |
| Protoplasts (1 h) | Kwon et al. (2005) | Non-destructive (washing of cells) | Cell wall proteome | Salt solution (1 M KCl) | 2D-E | MALDI-TOF MS | 71 | 25 |
| Protoplasts (3 h) | Kwon et al. (2005) | Non-destructive (washing of cells) | Cell wall proteome | Salt solution (1 M KCl) | 2D-E | MALDI-TOF MS | 66 | 23 |
| Liquid cultured etiolated seedlings | Schultz et al. (2004) | Destructive (grinding) | Arabinogalactan proteins (AGPs) proteome | Detergent-containing buffer ( Triton X-100), Yariv precipitation, HF deglycosylation | HPLC | Edman N-terminal sequencing | 12 | 12 |
| Culture medium of etiolated seedlings | Charmont et al. (2005) | | Secretome | | 2D-E | MALDI-TOF MS | 49 | 44 |