We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,900
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Cell Wall Proteomics as a Means to Identify Target Genes to Improve Second-Generation Biofuel Production

Maria J. Calderan-Rodrigues, Juliana G. Fonseca, Carlos A. Labate and Elisabeth Jamet

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/66189

Abstract

Second-generation biofuels (B2G) generally uses residues composed of lignocellulosic materials to produce renewable energy (potentially up to 50%), without increasing the planted areas. However, the high cost of enzymes required for cell wall disassembly prior to the saccharification makes the B2G production more expensive yet, compared to the first-generation biofuels. Designing plants with less lignin, a barrier to B2G production, or facilitating cell wall disassembly by searching for the plant mechanisms can be the way to obtain B2G feasibility. Therewith, plant cell wall proteomics provides valuable information concerning the main cell wall proteins (CWP)s involved in its biosynthesis and rearrangements. Essentially, two plants of the grass family have been studied: sugarcane as a crop amenable to second-generation ethanol (E2G) production; and *Brachypodium distachyon* as a model plant amenable to genetic transformation. Cell wall proteomics has allowed the identification of numerous CWP's as well as their fine profiling in different organs and at various developmental stages. Proteins acting on carbohydrates, mostly glycosyl hydrolases, and oxidoreductases, including class III peroxidases and laccases, can be highlighted. Both kinds of CWP's are assumed to contribute to the remodelling of cell wall polysaccharides by enzymatic or non-enzymatic mechanisms. CWP's present in growing organs could also be attractive candidates since they greatly contribute to cell wall plasticity.

**Keywords:** *Brachypodium distachyon*, cell wall protein, grass, second generation ethanol, sugarcane
1. Introduction

Second-generation biofuels (B2G) are a promising renewable alternative to supply energy demand of fossil fuels worldwide, whose advantage is mostly due to the lower emission of greenhouse gases and the possibility to increase the production without widening the planted area. However, we are still far from producing B2G at an economically competitive way and reasonable amount to replace fossil fuels. B2G uses lignocellulosic material as substrates. Since sugarcane has been considered one of the best crops to produce bioethanol, its bagasse and straw have been studied as one of the main complementary sources of C$_6$ and C$_5$ sugars for B2G. One of the main constraints to its economic feasibility relies on the rate of success of the enzymatic saccharification enabling the conversion of the plant cell wall sugars into bioethanol [1]. Saccharification of the cell wall is the process of hydrolysis by which a complex carbohydrate, such as cellulose can be broken into monosaccharides. Thus, the production requires a pre-treatment of the biomass prior to expose the wall carbohydrates to substantial amounts of expensive enzymes in the industrial process.

Several strategies have been recently used to improve saccharification, mostly using microorganism enzymes. Different enzymes with cell wall polysaccharide degradation activity have been prospected from several organisms such as seaweed [2], termite stomach [3] and fungi [4]. However, even presenting some advances [5, 6], the cost of E2G is not competitive for first-generation ethanol production from sugarcane.

New approaches are emerging from the plant’s perspective itself, which together may be the “eureka” to solve this puzzle. Presently applied research has been focusing on lowering or modifying the lignin content to allow its removal in the industrial production and thus increasing the access of carbohydrates to saccharification [7]. Indeed, lignin is frequently the major reason for biomass recalcitrance. However, several strategies that focused on diminishing the lignin content, and thus leading to improved saccharification, resulted in deleterious effect on plant development [8]. A different point-of-view based on lignin modification may be more effective, since even increased lignin content showed improved saccharification in Brachypodium distachyon [9]. Thereby, the expression of a bacterial enzyme into Arabidopsis thaliana altered lignin and improved saccharification, without lowering the lignin content [10].

Another strategy is to engineer the plant cell wall genes in order to enable the plant itself to produce easier breakable sugars. By producing cellulose with more adequate characteristics to allow a more efficient saccharification, such as crystallinity, the plant material showed to have improved saccharification efficiency in A. thaliana [11]. Genetic engineered rice and wheat also showed increased enzymatic saccharification when cell wall proteins (CWPs) acting on polysaccharides had their expression changed [1, 12].

The plant cell wall represents 50% of the organic carbon present on earth [13]. Cellulose is a major cell wall polysaccharide and the major second-generation ethanol (E2G) source. The biosynthesis of wall polymers and all the processes that occur in the plant cell wall are mediated by CWPs among which numerous enzymes. Prospective and directed studies to increase the
knowledge on CWPs both in model species and in plants of agricultural interest provide valuable information on target-proteins in order to direct the plant pathways and produce plant carbohydrates easily saccharified. Accordingly, the high potential of this research can be the key to B2G industrial production.

2. Plant cell wall proteomics

2.1. The plant cell wall

The plant cell wall was once considered as a static structure, but since the 1990s, it has been addressed as a dynamic part of the cell, more similar to an extracellular compartment [14]. It has to be strong and flexible at the same time to enable its several roles such as mechanical stability, osmotic control, signalling and defence against different types of stresses. Its composition varies according to the stage of development, cell types and environmental cues. As an example, epidermis cells have to be better prepared for water loss than inner cells [15].

Cell walls can be classified into two types: primary and secondary. The former is found in growing tissues, and thus extendable; and the latter type is formed after the end of cell growth. It can allow cells to resist to compression forces [16]. Cell wall composition includes cellulose, hemicelluloses, pectins, proteins [17] and lignin in some cell types [18].

Cellulose is a cell wall polysaccharide with a high molecular mass, formed by long linear chains of β-1,4-linked glucose residues forming microfibrils [19]. Primary walls contain around 20-30% cellulose, and secondary walls up to 50% [20]. Hemicelluloses are composed of β-1,4-linked monosaccharides with side chains [19]. The most present hemicelluloses in dicots and grasses are xyloglucan (XG) and β-(1,3-1,4)-mixed linked glucans, respectively. XG is probably involved in forming cross-links between cellulose microfibrils [21]. Pectic polysaccharides are formed by structures enriched by galacturonic acid with complex side chain structures [22]. Sugarcane and other grass family species cell walls present specific characteristics such as being poor in pectins and having no XG interlocking the cellulose microfibrils in dividing cells; this role is performed by glucuronoarabinoxylans (GAXs) [14]. Lignin is a phenolic polymer and confers rigidity to cellulose microfibrils, and thus, to the cell wall [23].

Cell wall biosynthesis seems to be specific for each cell type [21]. During this process, cellulose is synthesized at the level of the plasma membrane by specific protein complexes. Conversely, non-cellulosic polysaccharides, such as hemicelluloses and pectins, are synthesized in the secretion pathway and secreted to the apoplast, where they form the wall networks together with cellulose [24]. Cell expansion occurs with enzymatic or non-enzymatic cleavages of cell wall polymers and the osmotic pressure separating the microfibrils. Polymers are then deposited in the internal part of the cell wall, forming the new cross-linked network [14]. Several phytohormones are involved in cell expansion, acting specifically at the reorientation of the microtubules, which may reorient the cellulose deposition [21].

As widely known, sugarcane is the raw material for one of the largest bioethanol production. E2G production uses lignocellulosic material to convert into ethanol through the steps of
pretreatment (to expose the cell wall polysaccharides to the enzymes), hydrolysis of the cellullosic and hemicellulosic polysaccharides into monomers and finally fermentation of these sugars into ethanol [25].

Over the years, the information regarding cell wall components from the chemical point-of-view has increased, enabling us to think about strategies to modulate the cell wall structure. There is knowledge available related to cellulose and hemicelluloses biochemical properties and to the pectic polysaccharides biochemistry [26]. However, less is known about the overall architecture of the cell wall. This knowledge should be enlarged to provide clues to engineer walls. Indeed, since the cell wall is constantly being modified either to respond to internal and external stimuli, this self-regulatory mechanism could be modulated to respond to commercial interests.

2.2. The plant cell wall proteome

The concept of CWPs includes not only the proteins present inside the cell wall structure but also those present in the apoplast. CWPs are essential to the wall functions such as modification of the cell wall components, its structure, signalling, interaction with the plasma membrane and response to stresses [27]. Several factors can modify the cell wall proteome content, such as development [28–31] and biotic or abiotic stresses [32, 33].

CWPs share three common characteristics: a signal peptide to be targeted to the secretory pathway, no intracellular retention motif and the absence of hydrophobic trans-membrane domains. The signal peptide presents a positive charge at its N-terminus, a hydrophobic central region and a polar C-terminus [34]. One of the best-described intracellular retention motif is the C-terminal H/KDEL, which maintains proteins inside the endoplasmic reticulum [35]. On the contrary, other sorting determinants are more complex. For example, vacuolar targeting routes are diverse and there seems to be different types of vacuole sorting determinants [36]. Bioinformatic programs can help predicting the subcellular location of proteins through protein amino acid sequences, but they rely on experimental evidence which can be incomplete [37].

Three types of CWPs can be considered according to their interaction with the cell wall matrix [27]. The labile proteins have little or no interaction with the cell wall polysaccharides and circulate in the extracellular matrix. They can be recovered by vacuum infiltration of tissues [38]. The weakly bound proteins can be linked to the wall components through Van der Waals interaction, hydrogen bonds, or ionic links and can be recovered with salt solutions. Strongly bound proteins such as structural proteins (SPs) are resistant to salt extractions and can be linked together or to polysaccharides by covalent bonds [39]. Regarding functions, CWPs can be divided into nine functional classes including a class of miscellaneous proteins (MIPs) and a class of proteins yet unknown function (PUFs) [40]. As all classifications, this one has some drawbacks like the difficulty to classify proteins with dual functions such as protease possibly involved in protein turnover or in signalling, but it allows getting an overview of cell wall proteomes [41].
Proteins acting on carbohydrates (PACs) mostly comprise glycosyl hydrolases (GHs) and are involved in cell wall polysaccharides remodelling [42]. PACs belong to the most represented classes in cell wall proteomes. Cellulases and glucanases are examples of proteins that can be found in this family. These enzymes are used in enzymatic hydrolysis cocktails used in E2G production, so they could be targets for manipulation in the plant species. Oxido-reductases (ORs) are mostly class III peroxidases (Prxs). Prx activities are diverse, they can break cell wall polysaccharides in a non-enzymatic way and facilitate cell wall extension but they can also favour the cross-linking of cell wall components such as monolignols and SPs [43]. Proteins related to lipid metabolism (PLMs) are almost all lipid transfer proteins and lipases. Some of them could be involved in cell wall loosening through the bind of lipids to their hydrophobic cavity [44]. Proteases (Ps) can play roles in protein turnover, protein maturation, signalling or defence [45]. SPs, such as hydroxyproline-rich glycoproteins, proline-rich proteins and glycine-rich proteins can be cross-linked in cell walls and contribute to its architecture [46, 47]. Proteins with interaction domains with proteins or polysaccharides (PIDs) comprise lectins and enzyme inhibitors. There is a lack of knowledge regarding the role of lectins in plant cell walls [48]. Enzyme inhibitors play a critical role in the regulation of enzymatic activities. As an example, there is a subtle interplay between pectin methyltransferase and pectin methylesterase inhibitors [49]. Proteins possibly involved in signalling (PSs) include arabinogalactan proteins which have been assumed to play diverse roles during plant development, and particularly in calcium signalling [50]. The miscellaneous proteins (MPs) contain many protein families which are not numerous enough to form a distinct class. The roles of proteins with domains of unknown function (PUFs) are mostly unknown, but this functional class offers potential for future research. Among PUFs, the DUF642 proteins have been shown to interact with cellulose in vitro [51]. They could also be involved in pectin methyltransferase or in defence [52, 53].

Isolating and identifying CWPs is particularly challenging. Indeed, the difficulty begins with the extraction procedure. The cell wall is an open compartment and the polysaccharidic network can be a trap for intracellular contaminants. Either destructive (DP) or non-destructive (NDP) protocols have been used. DP’s rely on grinding the tissues to isolate cell walls prior to the extraction of proteins with salt solutions [54]. The purification of cell walls relies on the fact that it is the denser cell compartment [55]. NDPs, using vacuum infiltration of tissues with mannitol or salt solutions, do not harm the cells and allow extraction of apoplastic proteins [56]. Usually, the salts used in the extraction protocols are CaCl₂ and LiCl. CaCl₂ extract CWPs through a competition mechanism [40] since pectins strongly chelate calcium ions [57]. An illustration of the effects of CaCl₂ has been provided by plasmolysis experiments performed on leaf tissues transiently expressing a CWP fused to the fluorescent TagRFP (red fluorescent protein) [38]. The fusion protein in displaced from the cell wall to the apoplastic space upon CaCl₂ application. On the other hand, LiCl is able to extract hydroxyproline-rich glycoproteins [58]. The use of both types of protocols to extract CWPs can be a good strategy to increase the coverage of the cell wall proteome [30]. However, some CWPs still escape because they are strongly bound to cell wall components [38]. At present, the cell wall proteomes are poor in SPs such as hydroxyproline-rich glycoproteins or proline-rich proteins. In addition, since some CWPs are heavily glycosylated, these post-translational modifications can be a problem for
protein identification by mass spectrometry. Finally, proteomics studies of species that do not have a fully sequenced genome present an additional bottleneck because the precise identification of proteins cannot be achieved.

Even carefully performing all these protocols, the identification of proteins that are not secreted through the classical secretory pathway has been reported. These proteins can be predicted to belong to different cell compartments such as cytoplasm, nucleus, mitochondria, chloroplasts or vacuoles. The question of the existence of alternative routes of secretion is still a matter of debate [41].

3. A focus on *B. distachyon* and sugarcane cell wall proteomes

After designing several protocols to analyse the cell wall proteome of *A. thaliana* as a test case, around 700 CWPs have been identified in different organs such as leaves, stems, roots and etiolated hypocotyls as well as in cell suspension cultures, i.e. about one-third of the expected total number [59]. In order to widen the knowledge regarding CWPs targeted to find candidate routes to improve E2G production from the plant perspective, two additional species were studied: (i) *B. distachyon* as a model for grass species from temperate areas, amenable to genetic transformation and having a fully sequenced genome [60]; and (ii) sugarcane, only having a large EST collection, but being one of the major sources for E2G production.

3.1. Plant material

For *B. distachyon*, three types of organs were used: leaves, internodes and grains (Figures 1A, B). Two-month-old plants were used and the CWP extractions were performed in young or

![Figure 1. *B. distachyon* and sugarcane plants used for proteomics studies: 2-month-old sugarcane plants (A), 4-month-old sugarcane plants (B, C), and 2-month-old *B. distachyon* plants (D, E). f (young leaves), g (mature leaves), h (apical internodes), and i (basal internodes).](image-url)
mature leaves and apical or basal internodes [29]. These organs were studied in order to compare the differences between organs and to look for proteins possibly involved in cell wall extension and growth arrest. Grains were collected at different times after flowering (9, 13 or 19 days) [31, 61]. The aim of the study was to understand the modifications of cell wall polysaccharides during grain development and filling because they are key determinant of the size and mass of the grain.

In the case of sugarcane, three types of materials have been studied (Figures 1C–E): 11-day-old cell suspension cultures [62], 2-month-old stems [30], and 4-month-old young or mature leaves and apical or basal internodes [63]. The aim was to identify among CWPs possible targets for cell wall modification in order to facilitate E2G production.

3.2. Methods

3.2.1. Extraction procedures

In these experiments, different extraction techniques were used. For B. distachyon, a DP was used for all the materials [54]. It started with mixing the tissue in a 5 mM sodium acetate buffer, pH 4.6, 0.4 M sucrose and protease inhibitor cocktail. After that, the mixture had to be ground in a blender at full speed for about 15 min. PVPP was added to the homogenate, and it was stirred for 30 min at 4°C. To isolate cell walls, the mixture was submitted to several successive centrifugations (1000×g) in a solution of increasing sucrose concentration (0.6-1.0 M). The pellet was then extensively washed through a Nylon net (25 µm) to remove sucrose. The cell wall fraction was ground in liquid nitrogen. Then, proteins were extracted by different salt buffers prepared in 5 mM sodium acetate, pH 4.6: twice in 0.2 M CaCl₂, followed by twice in 2 M LiCl. Cell walls were resuspended in these buffers and centrifuged at high speed (40,000×g/15 min/4°C). The four supernatants were pooled.

The same DP with minor modifications was used for sugarcane cell suspension cultures and 2-month-old stems [30, 62]. Another extraction method was tested with young or mature leaves and basal or apical internodes. This method was based on vacuum infiltration [56], which is a NDP requiring working with fresh material only. The plant organs were cut to fit in a beaker and completely immersed in a solution of 3.0 M mannitol and 0.2 M CaCl₂ in a dessicator connected to a vacuum pump. The tissues were vacuum-infiltrated for 5 min. Plant organs were centrifuged in a swinging bucket rotor (200×g/15 min/20°C). The apoplastic fluids (released at the bottom of the tube) were collected and stored at low temperature. This procedure was repeated once with the same solution. Additional two rounds of vacuum infiltration were performed in a solution with 2 M LiCl instead of 0.2 M CaCl₂. All four extracts were pooled.

Samples resulted from DP and NDP were desalted, freeze-dried to concentrate proteins and then used in 1D-electrophoresis (1D-E) to check the quality of the protein extracts.

It should be mentioned that all the experiments have been repeated twice or thrice to take into account biological variation. Only CWPs identified in at least two biological replicates have been validated. A detailed description of these protocols can be found in Refs. [29–31, 61–63].

http://dx.doi.org/10.5772/66189

Cell Wall Proteomics as a Means to Identify Target Genes to Improve Second-Generation Biofuel Production
3.2.2. Identification of proteins by mass spectrometry and bioinformatic analyses

Then, proteins were identified by mass spectrometry (LC-MS/MS) and bioinformatics after tryptic digestion performed at 4°C, after separation by 1D-E or in solution. A detailed description of the parameters used for MS analysis can be found in [29–31, 61–63]. For *B. distachyon*, the genomic sequence data were used [64, 65]. For sugarcane, the SUCEST translated EST database was used [66]. The amino acid sequences of the identified proteins were systematically compared to those of *Sorghum bicolor*, the closest related species having a fully sequenced genome [64]. In case of partial EST sequence, this comparison allowed the bioinformatics prediction of sub-cellular localization and functional domains.

For both plant species, the bioinformatics analysis of the identified proteins was carried out *de novo* in the same way regarding the prediction of their subcellular localization and of functional domains using the ProtAnnDB annotation pipeline [67, 68]. All the experimental data were collected in the WallProtDB database [59, 69]. The Venn diagrams used in this chapter were made with the Venny online software [70].

3.2.3. A comparative survey of *B. distachyon* and sugarcane cell wall proteomes

As a key indicator of the quality of the protein extract, the percentage of proteins predicted to be secreted and not retained in an intracellular compartment can be calculated (Figure 2). The other proteins can be considered as intracellular contaminants. The highest proportion of proteins predicted to be intracellular has been found in sugarcane cell suspension cultures (82%). This could be explained by two facts: a DP was used thus increasing the chance for intracellular proteins to be trapped in the cell wall polysaccharidic matrix; and/or cell suspension cultures contain a certain proportion of dead cells whose content is released in the culture medium, so that intracellular proteins can interact with the cell walls of living cells. Such result has also been obtained with cell suspension cultures of *A. thaliana* [71]. Apart from this sample, the proportion of proteins predicted to be intracellular is above 40%. The highest proportion of CWPs was obtained with basal internodes of *B. distachyon*. In that case, we noticed that the

![Figure 2. Percentage of CWPs and proteins predicted to be intracellular in each proteome. *B. distachyon* proteomes are in black and white, whereas sugarcane proteomes are in green and white. AI: apical internodes; BI: basal internodes; C: cell suspension cultures; G: grains; ML: mature leaves; YL: young leaves; 2MS: 2-month-old stems.](image)
sedimentation of cell wall fragments were particularly easy for this sample, thus facilitating its purification [29].

Altogether, 567 and 273 different CWPs were identified in all mentioned experiments for *B. distachyon* and sugarcane, respectively. At present, these species, together with *Oryza sativa* (270 CWPs), have the largest cell wall proteomes among monocots [59].

The specific proteins found in each experiment, and the common ones are shown in Figure 3 for both species. A first comparison can be made between the cell wall proteomes of the aerial parts of *B. distachyon* and sugarcane, the most amenable to E2G production. Sixty-three out of the 314 CWPs (20.1%) identified in *B. distachyon* leaves and internodes were common to both organs taken at two different stages of development (Figure 3A). The percentage of common proteins two by two was also homogenous, varying from 27% to 39%. This proportion was very different for sugarcane cell wall proteomes, with only 3.0% of the proteins common to all samples, i.e. 6 of 201 CWPs (Figure 3C). The comparison two by two reached a result similar to that obtained with *B. distachyon* only for CWPs present in apical and basal internodes (37.4%). The other duos have between 4.0% and 14.0% of common CWPs. This is probably related to the smaller size of the sugarcane cell wall proteomes of compared to those of *B. distachyon* and to the very different number of CWPs identified in leaves in comparison to stems for sugarcane. Using 2-month-old leaves, the difficulty in extracting proteins from cell walls was also observed (unpublished results). This might be inherent to the leather type of sugarcane leaves requiring a different extraction strategy. Another explanation could rely on the hexa- to octaploid genetic basis of sugarcane [72], which could lead to the expression of different sets of multigene family members at different developmental stages and in different organs.

Figure 3. Venn diagrams showing common and specific CWPs for each experiment performed with *B. distachyon* (A and B) or sugarcane (C and D). AI: apical internodes; BI: basal internodes; C: cell suspension cultures; G: grains; ML: mature leaves; YL: young leaves; 2MS: 2-month-old stems.
Including the cell wall proteomes of *B. distachyon* grains, 25% of the CWPs were common to all organs (Figure 3B). It should be noted that the largest cell wall proteome was that of grains comprising 481 CWPs and that 45% of its CWPs were specific to this organ.

Now, looking at all the known cell wall proteomes of sugarcane, cell suspension cultures, leaves, 2- and 4-month-old stems only showed two common CWPs (Figure 3D). Eighty two of 273 CWPs (30.4%) were specific to 4-month-old basal and apical internodes.

These comparisons are of special interest because they allow identifying both CWPs specific to organ or developmental stages and CWPs common to all organs which may belong to a set of housekeeping CWPs essential for cell wall maintenance. For example, the set of proteins common to the 8 cell wall proteomes of *B. distachyon* comprises 42 CWPs among which 10 GHs, 4 Prxs, 8 proteases, 1 lipid transfer protein (LTP), 2 GDSL lipases and 1 DUF642 protein. In sugarcane, six CWPs were found to be common to 4-month-old leaves and internodes (Figure 3C): one GH, two Prxs, two proteinase inhibitors, and one subtilisin, whereas two CWPs were common to all six cell wall proteomes (Figure 3D): a protein of unknown function and a cys-protease. These CWPs would deserve functional studies to better understand their functions. The case of sugarcane seems more complex than that of *B. distachyon* with less putative housekeeping CWPs identified up to now.

Now, cell wall proteomes can be considered from the functional point of view. As explained above, it is possible to group proteins according to the prediction of functional domains [27, 56]. Table 1 shows the distribution of *B. distachyon* and sugarcane CWPs into functional classes in the different cell wall proteomes. Some specific features can be noticed in *B. distachyon*: (i) PACs are less represented in basal internodes; (ii) ORs are more represented in internodes; (iii) PLMs are less represented in mature leaves; (iv) Ps are more represented in leaves; and (v) PIDs are less represented in mature leaves. Finally, SPs have been only found in grains with two leucine-rich extensins identified. In sugarcane, the situation is very different: (i) PACs are less represented in cell suspension cultures and in leaves; (ii) ORs are more represented in cell suspension cultures and in mature leaves; (iii) PLMs are less represented in cell suspension cultures and in internodes of 4-month-old plants; (iv) Ps are less represented in cell suspension cultures, but more in 4-month-old stems; (v) PIDs are poorly represented in 2-month-old stems, but more represented in cell suspension cultures and in mature leaves; and (vi) PSs are less represented in cell suspension cultures, 2-month-old stems, and in young leaves. In both plants, there are also variations in the contribution of MPs and PFUs to all cell wall proteomes.

This overview allows getting a profiling of the cell wall proteomes and to focus on specific functional classes of CWPs. Because of the variations observed in the contribution of each functional class to the whole proteomes, it also shows that each plant and each organ has to be studied in detail before choosing a strategy to modify its cell walls. For example, ORs includes mostly Prxs, but also blue copper-binding proteins, and multicopper oxidases. Prxs are involved in diverse physiological processes, such as signalling [43], lignification [73], and cross-linking of SPs [74]. Their roles in cell wall polysaccharide and protein network rearrangements could be the reason why they are more represented in *B. distachyon* stems. Curiously, the sugarcane cell wall proteomes exhibit the highest proportions of ORs compared
to other plants. Such CWPs are interesting targets whose genes could be engineered for E2G production optimization.

<table>
<thead>
<tr>
<th>Functional class</th>
<th>PACs</th>
<th>Ors</th>
<th>PLMs</th>
<th>Ps</th>
<th>PIDs</th>
<th>PSs</th>
<th>SPs</th>
<th>MPs</th>
<th>PUFs</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. distachyon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All proteomes</td>
<td>24.2</td>
<td>13.6</td>
<td>10.8</td>
<td>13.8</td>
<td>7.1</td>
<td>5.5</td>
<td>0.3</td>
<td>12.5</td>
<td>12.3</td>
</tr>
<tr>
<td>YL</td>
<td>21.8</td>
<td>11.2</td>
<td>11.2</td>
<td>18.2</td>
<td>3.5</td>
<td>5.9</td>
<td>11.8</td>
<td>16.5</td>
<td></td>
</tr>
<tr>
<td>ML</td>
<td>26.5</td>
<td>15.1</td>
<td>7.8</td>
<td>16.3</td>
<td>4.8</td>
<td>1.8</td>
<td>13.9</td>
<td>13.9</td>
<td></td>
</tr>
<tr>
<td>AI</td>
<td>23.5</td>
<td>18.0</td>
<td>10.4</td>
<td>13.7</td>
<td>2.7</td>
<td>4.9</td>
<td>12.0</td>
<td>14.8</td>
<td></td>
</tr>
<tr>
<td>BI</td>
<td>19.4</td>
<td>21.2</td>
<td>10.0</td>
<td>12.9</td>
<td>5.3</td>
<td>4.7</td>
<td>10.0</td>
<td>16.5</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>24.1</td>
<td>11.6</td>
<td>10.6</td>
<td>14.5</td>
<td>7.3</td>
<td>5.8</td>
<td>0.4</td>
<td>12.5</td>
<td>12.3</td>
</tr>
<tr>
<td>Sugarcane</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All proteomes</td>
<td>20.5</td>
<td>20.9</td>
<td>13.2</td>
<td>12.8</td>
<td>5.9</td>
<td>4.0</td>
<td>16.1</td>
<td>6.6</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>11.6</td>
<td>30.4</td>
<td>4.3</td>
<td>8.7</td>
<td>11.6</td>
<td>1.4</td>
<td>20.3</td>
<td>11.6</td>
<td></td>
</tr>
<tr>
<td>2MS</td>
<td>20.2</td>
<td>21.4</td>
<td>16.7</td>
<td>13.1</td>
<td>1.2</td>
<td>1.2</td>
<td>11.9</td>
<td>14.3</td>
<td></td>
</tr>
<tr>
<td>YL</td>
<td>8.5</td>
<td>25.4</td>
<td>20.3</td>
<td>13.6</td>
<td>6.8</td>
<td>1.7</td>
<td>18.6</td>
<td>5.1</td>
<td></td>
</tr>
<tr>
<td>ML</td>
<td>8.3</td>
<td>33.3</td>
<td>19.4</td>
<td>11.1</td>
<td>13.9</td>
<td>0.0</td>
<td>11.1</td>
<td>2.8</td>
<td></td>
</tr>
<tr>
<td>AI</td>
<td>25.6</td>
<td>20.8</td>
<td>4.8</td>
<td>16.8</td>
<td>6.4</td>
<td>5.6</td>
<td>14.4</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>BI</td>
<td>24.2</td>
<td>20.8</td>
<td>5.0</td>
<td>20</td>
<td>5.8</td>
<td>5.0</td>
<td>12.5</td>
<td>6.7</td>
<td></td>
</tr>
</tbody>
</table>

Results are expressed as percentages of the number of CWPs identified in each proteome. Values in bold are average values calculated with all proteomes data and values much different from these average values.

MPs: miscellaneous proteins; PLMs: proteins related to lipid metabolism; ORs: oxidoreductases; PACs: proteins acting on carbohydrates; PID: proteins with interaction domains; Ps: proteases; PSs: proteins involved in signalling; PUFs: proteins of unknown function; SPs: structural proteins; AI: apical internodes; BI: basal internodes; C: cell suspension cultures; G: grains; ML: mature leaves of 4-month-old plants; YL: young leaves of 4-month-old plants; 2MS: 2-month-old stems.

Table 1. Distribution of the CWPs found in each cell wall proteome of B. distachyon and sugarcane into functional classes.

PLMs are mostly represented by LTPs and GDSL lipases. LTPs exact biological roles are yet unknown, but they have been related to cell wall loosening and extension [44], pathogen response, and cutin assembly [75]. Since sugarcane at young developmental stages are similar to rolled leaves, this may explain the high proportion of LTPs, probably playing roles in cutin assembly of both sides leather-like leaves. Nevertheless, the better understanding of the mechanisms under this protein class may lead to the design of new strategies to increase biomass production.

The low percentage of PACs in sugarcane cell suspension cultures and leaves is also puzzling. PACs mostly include GHs, such as β-xylosidase, β-galactosidase and have been associated with cell wall loosening and expansion [76]. GH3, GH35, GH27, and GH51 can be of special interest since they show homology to enzymes of interest used for E2G production [3].

The two studied plant species, B. distachyon and sugarcane, appear to be complementary to identify CWPs and look for their functions. Although both plants are monocots and have similar cell wall composition, they seem to have different strategies to modulate cell wall
structure during development. Combining genetics and biochemical approaches should allow getting insight in those mechanisms.

3.3. Perspectives and targets for E2G production

Changes in lignin composition have led to a subtle improved saccharification with no relevant deleterious effect [77]. However, for the cell wall polysaccharides, the challenge is still bigger since there is less knowledge regarding their synthesis. The main players able to modify cell wall polysaccharides are (i) the transcription factors that control the initial steps of gene expression and (ii) the enzymes and proteins involved in the biosynthesis of cell wall components and in their modifications in muro [78]. By altering transcription factors in A. thaliana, it was possible both to increase cellulose and decrease lignin content [79] and improve secondary cell wall synthesis in fibre cells [80]. In addition, the golden pot may be near; transgenic A. thaliana expressing microbial hydrolases showed no visible changes in phenotype and increased wall degradability [81]. An alternative to decrease the transgenic debate and perhaps optimize efficiency could be altering the expression of the own plant enzymes generating a genetically modified plant, but not a transgenic one. Besides hydrolases, another possibility is to consider the potential of the plant cell wall as a sensor to perceive changes and direct cell wall polysaccharides synthesis, such as in microorganisms [78]. Then, attention should be paid to the fasciclin arabinogalactan proteins, wall-associated kinases and other membrane proteins. Expressing carbohydrate-binding proteins such as expansins could facilitate cell loosening, and it may be a possibility to improve saccharification as well [82].

As can be seen, modulation of CWPs expression offers a wide range of possibilities to achieve a plant cell wall more cost-effective in terms of E2G production. Since some CWPs have been reported to act on cell wall remodelling or expansion, and we observed a different proportion of them in the several organs and developmental stages, we suggest focusing studies on some CWP families such as Prxs, GHs and LTPs, mostly those found in young and growing organs. By targeting the level of expression of these proteins or their spatial distribution, it may be possible to design plants with cell walls easily saccharified to E2G production. In order to achieve this goal, it is recommended to use tissue-specific and spatial regulation of gene expression using precise gene promotors, so that there will be no deleterious effect to the living plant. Notwithstanding, we highlight that more information on the modifications occurring on cell wall polysaccharides has to be collected in order to provide the basis for applied results.

Acknowledgements

This research was partially funded by USP-COFECUB (Proc. 10.1.1947.11.9), BIOEN/PRONEX FAPESP (Proc. 2008/56100-5), INCT do Bioetanol, CNPQ (Proc.: 142784/2007-9), and FAPESP (Proc. 2007/59327-8 and 2012/1212521-2). The authors are thankful to Dr. R Pont-Lezica who was at the origin of the projects. They also thank Dr MC Falco and CTC for providing the sugarcane plants and cells. They thank Dr GM Souza and Dr M Nishiyama for providing the SUCEST sequences. MJR, JGF, and CAL are thankful to the Max Feffer Laboratory of Plant
Genetics team, especially TR Cataldi and S Guidetti for handling the mass spectrometer and FE Morais and LM Franceschini for the bioinformatics assistance. EJ wishes to thank the CNRS, the Paul Sabatier-Toulouse III University and ANR (Grant Génoplante/PCS-08-KBBE-003/CELLWALL) for supporting her research.

**Nomenclatures**

B2G: Second-generation biofuel  
CWP: Cell wall protein  
DP: Destructive protocol  
E2G: Second-generation ethanol  
GAX: Glucuronoarabinoxylan  
GH: Glycosyl hydrolase  
MP: Miscellaneous protein  
NDP: Non-destructive protocol  
OR: Oxidoreductase  
PAC: Protein acting on carbohydrates  
PID: Protein with interaction domains with proteins or polysaccharides  
PLM: Protein related to lipid metabolism  
P: Protease  
Prx: class III peroxidase  
PUF: Protein of unknown function  
SP: Structural protein  
XG: Xyloglucan

**Author details**

Maria J. Calderan-Rodrigues¹, Juliana G. Fonseca², Carlos A. Labate² and Elisabeth Jamet³*

*Address all correspondence to: jamet@lrsv.ups-tlse.fr

1 Brazilian Bioethanol Science and Technology Laboratory (CTBE)/Brazilian Center of Research in Energy and Materials (CNPEM), Campinas, Brazil

2 Department of Genetics/Laboratory Max Feffer of Plant Genetics/Higher School of Agriculture “Luiz de Queiroz”/University of São Paulo, Piracicaba, Brazil

3 Plant Science Research Laboratory (LRSV)/University of Toulouse/CNRS/UPS, Auzeville, France
References


