University of São Paulo
"Luiz de Queiroz" College of Agriculture

Proteome characterization of sugarcane primary cell wall

Maria Juliana Calderan Rodrigues

Thesis submitted in fulfillment of the requirements for the degree of Doctor in Science. Area: Genetics and Plant Breeding

Piracicaba
2012

Maria Juliana Calderan Rodrigues
Biologist

Proteome characterization of sugarcane primary cell wall

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Adviser:
Prof. Dr. **CARLOS ALBERTO LABATE**

Thesis submitted in fulfillment of the requirements for the degree of Doctor in Science. Area: Genetics and Plant Breeding

Piracicaba
2012

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - ESALQ/USP

Rodrigues, Maria Juliana Calderan
Proteome characterization of sugarcane primary cell wall / Maria Juliana Calderan Rodrigues.- - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2012.
180 p: il.

Tese (Doutorado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2012.

1. Cana-de-açúcar 2. Etanol 3. Parede celular vegetal 4. Proteínas I. Título

CDD 634.61
R696p

DEDICATION

A meus pais, Mônica e Francisco Sérgio, dedico.

I also dedicate to professor Rafael Pont-Lezica (in memorian)

ACKNOWLEDGEMENTS

I would like to express my gratitude to everyone who supported me and this project. First of all, I would like to thank my adviser, professor Carlos Alberto Labate for his guidance and friendship for the last ten years. Also, I would like to thank him for the constant support, especially when I needed the most.

I wish to extend my thanks to professor Rafael Pont-Lezica (in memorian) who generously contributed to this research and make possible my staying in France to develop part of my project in the Université Paul Sabatier (UPS). I am also indebted to professor Elisabeth Jamet, also from UPS, for her several contributions which improved this research in numerous ways and also for her generosity. I wanted to thank MSc. Thibaut Douché for his help in the bench and in the discussions during my staying in Toulouse. I also thank the helpful comments and suggestions made by professors Helaine Carrer, José Baldin Pinheiro and Claudia Barros Monteiro Vitorello. I would like to thank especially professor Claudia who was my supervisor during the PAE project, and taught me how to be a good teacher. I also wanted to thank Dra. Maria Cristina Falco, from CTC, that kindly provided the sugarcane cultured cells and the young plants to start my project. I am thankful to Prof. Glaucia Souza and Dr. Milton Nishiyama for providing the sugarcane ESTs translated database.

I also would like to thank my friends in the laboratory for the cheerful moments: Mônica, David, Pedro, Carol, Leonardo, Janaína, Andressa, Ivan, Felipe Marques, Felipe Boaretto, Ivan, Fabrício, Camila, Larissa, Haná, Júlia and Ana Paula. Special thanks to Simone Guidetti Gonzalez and Fernanda Salvato who helped me in the sequencing of proteins and Livia Maria Franceschini who helped me in the bioinformatic analysis. Due to the long hours of study and discussions I would like to thank the friends Thiago Falda, Matheus Bonatto, Ilara Budzinski and Fernada Salvato, also for sharing their friendship. I owe special thanks to my friends of academic life Simone Guidetti Gonzalez, Danielle Rodrigues and Andréa Bogas to the good moments shared during the trips and in the lab. Special thanks also to Amanda Begossi, Laís Vaz Setem and Juliana Fonseca for their joy, being always there when I needed and helping me in the experiments.

I would like to thank God for everything. I would not achieve my goals without the support of my parents, Mônica and Francisco Sérgio who taught me how to have an ethical conduct and to pursue my goals. Thank you for your unconditional love. I also would like to thank my sister Maria Beatriz who I truly love and always supported me and also contributed

with the benchwork, as well as with insightful discussions of the results. I would like to thank the love of my husband Eduardo who was crucial to my welfare and motivated me to keep going in several moments. I also would like to thank him for the help with the tables and the references.

I also would like to thank the professors and employees of the department of Genetics for their dedication.

I acknowledge CNPQ, INCT, BIOEN and especially Fapesp (process number 2007/59327-8) for the financial support for this research.

EPIGRAPH

“Ninguém é tão pequeno que não possa ensinar
nem tão grande que não possa aprender”
Blaise Pascal

CONTENTS

RESUMO

Caracterização do proteoma da parede celular primária de cana-de-açúcar

Este estudo fornece informação para auxiliar o uso da parede celular vegetal, a partir do bagaço de cana, para a produção de etanol celulósico. Com isso, as proteínas da parede celular de folhas, colmos e células em suspensão foram identificadas. Para isso, foram utilizados diferentes protocolos. Utilizando folhas e colmos de cana-de-açúcar de dois meses de idade, as extracções foram realizadas por meio de método destrutivo, com base na trituração dos tecidos, submetendo-os a gradiente crescente de sacarose e centrifugação, sendo a parede da célula extraída e depois isolada por lavagem sobre uma rede de nylon. Depois disso, as proteínas de parede celular foram extraídas utilizando dois sais, 0,2 M de $CaCl_2$ e 2 M de LiCl. Para células em suspensão, um protocolo semelhante foi utilizado, contendo, no entanto, um passo anterior de separação da parede celular por meio de maceração e precipitação em glicerol 15%. Usando colmos da mesma idade, dois meses, um protocolo não destrutivo foi testado com base na infiltração a vácuo dos tecidos nos mesmos sais já descritos, 0,2 M de $CaCl_2$ e 2 M de LiCl, e posterior centrifugação. Duas repetições foram usadas nos experimentos com plantas de dois meses de idade, e três, no caso de células em suspensão. As amostras complexas foram digeridas, fracionadas e seqüenciadas por espectrometria de massas, utilizando o equipamento SYNAPT G2HDMS acoplado ao cromatógrafo nanoACQUITY, ambos da Waters. Os peptídeos foram processadas utilizando ProteinLynx 2,5 comparando com a base de dados de ESTs traduzidos da cana. Utilizando programas de bioinformática, como Blast2GO, foi possível encontrar a anotação e classificação de proteínas semelhantes. Apenas proteínas igualmente encontradas em todas as repetições foram consideradas na análise principal. SignalP, WolfPSORT, TargetP, TMHMM e Predotar foram softwares utilizados para prever a localização subcelular, tanto para ESTs como proteínas, e apenas as proteínas preditas para serem secretadas por dois ou mais programas foram consideradas como proteínas de parede celular. Ao todo, 157 SAS diferentes relacionados à parede celular da cana foram encontrados. Dentre eles, 101 diferentes proteínas de parede foram caracterizadas em oito classes funcionais. O método baseado na infiltração a vácuo mostrou-se o mais eficiente, uma vez que apresentou quase metade, 48,84%, das proteínas preditas para serem secretadas, o que é um bom valor quando comparado com outros estudos. A maioria das proteínas secretadas estava relacionada com o metabolismo lipídico, como proteínas de transporte de lípidos, oxido-redutases, tais como peroxidases, enzimas modificadoras da parede, como as glicosil-hidrolases, proteases, proteínas com domínios de interação, proteínas sinalizadoras, entre outras. Os resultados estão de acordo com o papel que se espera da matriz extracelular no metabolismo de polissacarídeos e fenômenos de sinalização. Portanto, este trabalho forneceu informações valiosas sobre a parede celular da cana, tornando possível a utilização desses dados em futuros estudos para otimizar a produção de etanol celulósico.

Palavras-chave: Parede celular; Proteínas; Cana-de-açúcar; Etanol celulósico

ABSTRACT

Proteome characterization of sugarcane primary cell wall

This study provides information to support the use of plant cell wall, from sugarcane bagasse, to produce cellulosic ethanol. Therewith, cell wall proteins from sugarcane cells cultures, leaves and culms were identified. To do so, different protocols were used. Using two-month-old leaves and culms, the extractions were performed using a destructive method, based on griding the tissues, submitting them to a growing gradient of succrose and centrifugation, being the cell wall extract later isolated by washing on a nylon net. After that, the cell wall proteins were extracted using two salts, 0,2 M $CaCl_2$ and 2 M LiCl. Using cultured cells, a similar protocol was used, but it had a previous step of separation of the cell wall through grinding and precipitation in glycerol 15%. Using culms of the same age, a non-destructive protocol was tested based on vacuum infiltration of the tissues in the same salts already described, 0,2 M $CaCl_2$ and 2 M LiCl, and posterior centrifugation. Two replicates were used from two-month-old plants and three in the case of suspension cells. The complex samples were digested, fractionated and sequenced through mass spectrometry, using SYNAPT G2HDMS coupled to nanoACQUITY, both from Waters. Peptides were processed using ProteinLynx 2.5 Global Server against sugarcane translated-EST database. Using bioinformatic programs, such as Blast2GO, it was possible to find the annotation and classification of similar proteins. Only proteins equally found in all repetitions were considered in the main analysis. SignalP, WolfPSORT, TargetP, TMHMM and Predotar were used to predict the subcellular location, both from ESTs and blasted proteins, and only the proteins predicted to be secreted in two or more programs were considered as cell wall proteins. Altogether, 157 different SAS related to sugarcane cell wall were found. Among these, 101 different cell wall proteins were characterized from eight functional classes. The method based on vacuum infiltration seems to be the most efficient one, since it had almost half, 48,84% of the proteins predicted to be secreted, which is a good percentage when comparing to other studies. From secreted proteins most of them were related to lipid metabolism, as lipid-transfer proteins, oxido-reductases, such as peroxidases, cell wall modifying enzymes, like glycoside-hydrolases, proteases, proteins with interacting domains, signaling proteins and several others. Results are in agreement with the expected role of the extracellular matrix in polysaccharide metabolism and signaling phenomena. Therefore, this work provided valuable information about sugarcane cell wall that can lead to future studies to enhance cellulosic ethanol production.

Keywords: Cell wall; Proteins; Sugarcane; Cellulosic ethanol

LIST OF ABBREVIATIONS

BSA	Bovine Serum Albumine
$CaCl_2$	Calcium Chloride
CWPs	Cell Wall Proteins
ER	Endoplasmic Reticulum
EST(s)	Expressed Sequence Tag(s)
Fmol	Fentomol
g	G-force
g	Grams
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
kDa	KiloDaltons
L	Liter(s)
LiCl	Lithium Chloride
µg	Microgram
µL	Microliter
µm	Micrometer
min	Minutes
mL	Milliliters
mM	Millimolar
M	Molar
MS	Mass Spectrometry
N	Normal
NH_4HCO_3	Ammonium Bicarbonate
NH_4OH	Ammonium Hydroxide
nL	NanoLiter
pH	Potential Hydrogen
*p*mol	Picomoles
psi	Pounds per Square Inch
PVPP	PolyVinylPolyPyrrolidone
rpm	Revolutions Per Minute
SAS	Sugarcane Assembled Sequences
TFA	Trifluoroacetic Acid
v/v	Volume/Volume

LIST OF SYMBOLS

α	alpha
β	beta
°C	degrees Celsius
%	percentage
®	*registered* trademark
™	trademark

1 INTRODUCTION

The current energy sources from fossil fuels are not sustainable in the long term (FOOD AND AGRICULTURE ORGANIZATION-FAO, 2008). Ethanol is probably the most widely used biofuel in the world, since it comes from different sources of biomass such as starch, sucrose, oil, among others (STICKLEN, 2008). In this context, energy from biomass has become one of the most important renewable sources in recent years (MATSUOKA et al., 2009).

A global effort to obtain sustainable energy is essential for the preservation of natural resources, and mitigation of greenhouse gases emissions, such as CO_2 (FISCHER et al., 2008). Thus, the need for a renewable energy source to replace oil as a fuel, and being introduced the Kyoto Protocol, more attention is targeted to Brazilian ethanol, because when considering all the agro-industrial processes, sugarcane ethanol generates carbon credits (SOARES; ROSSEL, 2007). Brazil harvested about 620 million tons of sugarcane during 2010/2011, from these, more than 27 billion liters of ethanol and about 37 million tons of sugar were produced (UNIÃO DAS INDÚSTRIAS DE CANA-DE-AÇÚCAR - UNICA, 2012):

In addition, bioethanol produced from sugarcane was considered the best biofuel in the world, both in price as in sustainability of its production. Soil quality, climate conditions and agricultural technology placed sugarcane as one of the most promising sources of renewable energy on the planet (LEITE et al., 2009). Recently, a production crisis pointed to need for more investment to increase the cultivated area, and Brazil had to import ethanol to supply the internal demand. This depends on the harvest, because if the price of sugar is higher than the price of ethanol in the market, producers tend to increase sugar production and consequently, decrease ethanol generation. This is another indicative that more advances must be achieved in this field, especially considering alternative raw materials to produce ethanol, such as bagasse.

The use of plant material to generate energy awakens the interest of researchers and industrialists, as it is waste, found in large quantities. Sugarcane bagasse, industrial process waste, and straw, containing the cell wall and other components, are available in a clean, concentrated way and in large amount, making them energy sources that can be burned directly or converted into chemicals such as ethanol (SOARES; ROSSEL, 2007). Thus, studies are designed to use the plant cell wall, part of sugarcane bagasse, in order to produce ethanol, eliminating waste and generating extra energy (OLIVÉRIO; PROENÇA, 2004).

Ethanol made from sucrose or starch is called the first generation biofuel (YUAN et al., 2008). The so-called second generation ethanol is based on the saccharification of cellulosic feedstock through various pre-treatments that transform cellulose into fermentable sugars, which, in turn, may ferment and produce alcohol. Currently, the major barriers in obtaining this fuel are the high cost of cellulase production in microbial reactors, and especially, the costs of pre-treatment of lignocellulosic feedstock (STICKLEN, 2008). Selection of new yeast strains, genetic breeding, and recombinant DNA technology are one of the fields that can enhance bioethanol production (AMORIM et al., 2011). As a example, genetically engineered plants or microorganisms, to reduce the need for pre-treatment through lignin modification along with alternatives such as increasing the content of plant polysaccharides can be a way to overcome these challenges (STICKLEN, 2008).

Present in the bagasse, the plant cell wall is essential for cell division, expansion and differentiation (ROBERTS, 1994), responding to biotic and abiotic factors (ELLIS et al., 2002, VOGEL et al, 2004) being a source of signals to intra and interspecies cell recognition (CARPITA; GIBEAUT, 1993; PENNELL, 1998; BROWNLEE, 2002). It is a structure composed of high molecular weight polysaccharides, proteins and lignin, the latter found only in some cell types.

In order to know the functions of plant cell wall, and how it can be manipulated, researches have been studying the identification of these proteins. For this purpose, a proteomic approach is essential, increasing the view of the proteins present in a particular compartment at a given stage of development (JAMET et al., 2006).

Understanding how the plant carries out its own process of growth and cell expansion through protein characterization is a basic study that can be practical if applied in the future. As an example of its many applications, it may help in research to manipulate this mechanism, engineering a plant that has its wall more easily degradable, thus decreasing the cost of enzymes for cellulosic ethanol production. Therefore, the characterization of plant cell wall proteins, from sugarcane, for example, can generate information about cell wall components and which processes lead to a better response to economic and environmental interests.

This work intends to characterize cell wall proteins from sugarcane, leaves and culms, and cultured cells, in order to generate information that can help in the use of the cell wall to produce ethanol.

2 DEVELOPMENT

2.1 Review

2.1.1 General aspects of sugarcane

In Brazil, sugarcane has historical, economic, social and political relevance, being the main source of raw material to produce sugar and alcohol. Sugarcane has been cultivated in Brazil since the beginning of the portuguese settlements, being introduced in 1532 by Martim Afonso, in order to implement sugar mills. This plant adapted very well to the Brazilian soil and was extensively cultivated during the colonial period, promoting an important cycle in Brazil that lasted almost two centuries. With the expansion of the sugar industry in the region of the Antilles, and the expulsion of the dutch from the northeastern Brazil, in the middle of the XVIII century, Brazilian production decreased, but it still remained as an important activity. In 1993, the creation of the Institute of Sugar and Alcohol increased the production of automotive bioethanol, expanding sugarcane production as well, mainly in the southeastern Brazil (SZMRECSÁNYI, 1979). After 2003, with the growing environmental concerns, the attentions turned to Brazilian bioethanol, since is one of the most sustainable biofuels. Besides Brazil, sugarcane has been cultivated in more than one hundred countries (FOOD AND AGRICULTURE ORGANIZATION - FAO, 2008), making part of the economy all over the world.

The rapid expansion of sugarcane production in Brazil has raised a series of questions about its consequences and sustainability. Positive impacts are the elimination of lead compounds emitted by burning gasoline and less emission of harmful gases. There is also reduction of CO_2 emissions, since ethanol requires only a small amount of fossil fuels for its production. It is also considered a renewable fuel, contributing little to global warming, since carbon dioxide, the main gas of this phenomenon, released by the combustion of alcohol in one year, is reabsorbed by the plants in the next crop (OLIVÉRIO, 2006). These positive impacts are noticeable in the air quality where mechanical harvesting has been introduced, eliminating the need for sugarcane burning.

New challenges for the environmental consequences of expanding the sugarcane production encourage the industry to be structured in order to avoid or minimize these impacts. Processes and policies have been recommended as the proper planning and risk assessment of growing in new areas, improving the use of land to prevent erosion and

nitrogen pollution, protection of coastal resources, the end of the fires and proper work conditions to sugarcane cutters (MARTINELLI; FILOSO, 2008).

Regarding international concern about areas of sugarcane plantation, the agro-ecological zoning of sugarcane was created, which provided a mapping of the areas where sugarcane can be grown and where that culture should not be found. This zoning, introduced in 2009, prohibits sugarcane growing in Pantanal, Amazon and Upper Paraguay River Basin. Still, the producer must choose areas that do not require full irrigation in order to save water and energy, don't use lands with a slope up to 12 degrees, promoting mechanized harvesting, and the use of degraded areas or pastures (EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA – EMBRAPA, 2011). Thus, it would be possible to increase sugarcane planted area without degrading new areas of important ecosystems.

Another global concern is sugarcane burning for harvest, which is about to be extinguished. An agreement signed in 2007 by the sugarcane industry and the state government of São Paulo, the Agro-Environmental Protocol, proposes the end of the fire and manual cutting by 2017 in practically all sugarcane plantations in São Paulo, almost seven years earlier than established by law (UNIÃO DAS INDÚSTRIAS DE CANA-DE-AÇÚCAR – UNICA, 2012).

Sugarcane, understood as the cultivated crop, has a very intrincated family tree, which combined a lot of species to generate what people know today. It is member of Poaceae family, *Saccharum* genus, being commonly known as *Saccharum oficinarum*. The *Saccharum* genus is divided by technologists into six species, two of them wild and four existing only in cultivation, including *S. oficinarum* (NAIDU; SREENIVASAN, 1987; ROACH; DANIELS, 1987 apud IRVINE, 1999).

Sugarcane genome is polyploid, and the size and complexity of this genome make the research harder and more time consuming comparing to the genetic tools available in model plants such as *Arabidopsis* and rice. In these plants the whole genome is sequenced and an increasing number of mutants are available for studying the connection among genes, physiology and phenotype (INMAN-BAMBER et al., 2005). For sugarcane, more than 300.000 ESTs are sequenced and available at the moment (VETTORE et al., 2003), however the whole genome of some varieties are being sequenced through a consortium among five countries (Brazil, France, United States, South Africa and Australia): the Sugarcane Genome Sequencing Initiative (SUGARCANE GENOME INITIATIVE – SUGESI, 2012). The size of sugarcane genome is 10 Gb, and along with its highly polyploid and aneuploidy genome structure, with a complete set of homo(eo)logous genes predicted to range from 8 to 10

copies, sequencing sugarcane has been a challenge. Also, sugarcane is highly polymorphic which represents another drawback in order to obtain an assembled monoploid genome (SUGARCANE GENOME INITIATIVE – SUGESI, 2012). It is no wonder that sugarcane genome will need a lot of effort to be sequenced. Therefore, despite sugarcane genome complexity some studies can be performed based on the amount of information already available in these ESTs collections.

The stalks, also called culms, of sugarcane, *Saccharum officinarum* L., consist of a series of internodes, one above another, each of which bears a leaf. The leaves are usually attached alternately to the nodes, thus forming two ranks on opposite sides, which lie in approximately the same plane (RUINARD, 1966). The main substance desired by the industry; the sucrose, is produced in the leaves, translocated in the phloem and stored in the stem (MOORE, 1995). However, physiological and morphological traits responsible for improved yield, sucrose content and resource use are yet poorly understood in sugarcane (INMAN-BAMBER et al., 2005).

One key point to increase the area of sugarcane production, is to develop new varieties adapted to poor soil, drought, and also resistant to pests and diseases (MATSUOKA et al., 2009). Another strategy to increase the production is the development of new technologies that can use residues as raw materials to produce ethanol.

2.1.1.2 The stem

Sugarcane stems are composed of joints, each comprising a node and an internode, that initiate very short at the base, and then gradually increase in length until a maximum is reached (DILLEWJIN, 1952). A node consists of a lateral bud which leaves a scar if the leaf is removed, a band containing root primordia, and a growth ring (BLACKBURN, 1984). Each joint is a separate unit, which can generate another plant, a clone, by vegetative propagation. Generally, growth rate is small in the early stages of development. The joints of the stem are at different stages of development along its length (MOORE, 1995).

The stem is solid and unbranched (BLACKBURN, 1984), and the diameter is greater in the parts that are buried in the soil, decreasing as it goes to the top (DILLEWJIN, 1952). Sugarcane development is a continuous process resulting in sucrose being accumulated at higher concentrations in an increasing gradient from top to bottom culms, until a point where they reach full maturity and a stable, high sucrose concentration. A measure called Brix, can

estimate the effect of the normal pattern of maturation on carbohydrate accumulation, both within an internode and along the culm, giving the percentage by weight of soluble solids in solution (FERNANDES; BENDA, 1985 apud MOORE, 1995). In sugarcane, Brix concentration appears spatially and temporally regulated (MOORE, 1995).

Sugarcane stem is a culm, being a complex organ composed of epidermal, vascular, meristematic and parenchyma tissues (MOORE, 1995). In the epidermis of the stem there are usually two cell types: some elongated and rectangular ones, called large cells, and other small ones. The large cells constitute the major part of the pattern found in epidermis. The small cells can be of two types, cork cells and siliceous cells. The cork cells are suberized and have thin walls; they exhibit a wide range of shapes and are usually pointed. They can appear alone or in groups with two or three. Generally, however, the cork cells are the cells together siliceous cells, forming a small group. Siliceous cells are more uniform and are rarely alone, having their stomata scarce or not present. There are differences in epidermal patterns that occur in each variety (DILLEWJIN, 1952).

Each stem has a wax-covered ring surrounding the parenchyma and interspersed with fibers. The wax prevents the loss of water, the fiber gives rigidity to the stem and the parenchyma stores the juice (BLACKBURN, 1984). The bark of the stem is composed of a cortex consisting of several layers of cells, most of them sclerenchyma. Next to the epidermis, there are two layers of thick and lignified cells, giving rigidity to the stem. Beside these cells, one or more rows of thin-walled parenchyma cells contain chlorophyll. Also, there are one or more layers of sclerenchyma tissue that surrounds the outer vascular bundles. The cortex varies in width and composition in different regions of the stem (DILLEWJIN, 1952).

Sugarcane stem has an unusually large free-space volume (WELBAUM; MEINZER 1990). The fibrovascular bundles are widely spaced in the central stem, but towards the periphery the number of the bundles increases, however their size decreases. The bundles in the periphery are so close together that they almost form a solid ring. The bundle is composed of sclerenchyma sheath surrounding the xylem and phloem (DILLEWJIN, 1952).

The stem is not only a storage organ, it is an integral part of the plant body, and it is almost not possible to analyze it completely without destroying the plant, whicht greatly restricts the types of experiments possible (MOORE, 1995).

2.1.1.3 The leaves

The leaves of sugarcane are usually attached alternately to the nodes, forming two rows on opposite sides, which are almost in the same plane. The leaf consists of two parts: the blade and sheath, separated by a joint structure. The sheath is tubular in shape but broader at the base than at the top, fitting closely to the stem and separated from the leaf blade by a ligule and two dewlaps (BLACKBURN, 1984). In contrast to expanded leaves that can reach over a meter long, the young leaves are at the top, growing until it reaches their maximum over time (DILLEWJIN, 1952). When the collar and the upper leaf sheath emerge from the leaf roll, the leaves stop expanding and then the internodes become apparent (MOORE, 1995). The collar is formed by wedge-shaped coloured areas called dewlaps (RUINARD, 1966).

Mature leaves are replaced by new ones in a continuous process (RUINARD, 1966). The number of leaves is small in younger plants and increases gradually as the sugarcane grows. During the period of development, the number of leaves per stem is about ten, depending on variety and growing conditions. The leaves are asymmetric, the midrib on its adaxial side is more white and concave, while its abaxial side is greener and convex. The leaf sheath is tubular in shape, being wider at the base and stretching toward the apex. The outermost part of the sheath is green and often has hairs, but the inside is white and glabrous. The part that overlaps the sheath can be inserted into a horizontal line (DILLEWJIN, 1952).

The leaf sheaths are tubular and thin structures that overlap the stalk at the base, but are less closely pressed against the stalk, which has no vascular bundles (JAMES, 2004). The leaf blade has vascular bundles in three sizes: large, medium and small. The bundle consists of xylem and phloem, while the fibers are found adjacent to the phloem. The epidermis of the blade displays a variety of alternating patterns at regular intervals. The small bundles are close to the lower epidermis and the medium ones are in the center of the blade. The large bundles are always flanked by two small ones, the smaller and medium bundles occur among the large ones in an alternate way (DILLEWJIN, 1952).

The leaf has a midrib, which is strong, white and concave on the upper surface and pale green and convex on the abaxial side (JAMES, 2004). In the midrib, a thick layer of parenchyma is found between the fibrovascular bundles and upper epidermis, thus all bundles in this part of the blade are restricted to the bottom of the sheet, while the outer surface, which contains no bundles, is reinforced by a solid layer of sclerenchyma. The layer of parenchyma, which is interpolated between the upper epidermis and vascular bundles does not contain

chlorophyll, which explains the whitish collor, in the opposite, the bottom is green by the presence of cells with chlorophyll. In the leaf blade each vascular bundle is surrounded by a ring of parenchyma cells that contain chlorophyll (DILLEWJIN, 1952).

The sheath has radial rows of fibrovascular bundles, but the ligule has no vascular bundles (BLACKBURN, 1984). The larger bundles are located in equal distance from the inner to the outer epidermis, and epidermis toward the outer bundles becomes smaller gradually. The bundles are embedded in parenchyma cells, and the structures of the bundles resemble the stalks. Several bundles frequently join to form a large composite bundle (DILLEWJIN, 1952). The bundle is surrounded by the sclerenchyma, similar to the stem. In the basal sheath, sclerenchyma is replaced by collenchyma, a structure that requires flexibility. The epidermis of the sheath is similar to the stem (DILLEWJIN, 1952).

The sheaths stand on the opposite side from the bud, surrounding and protecting it (BLACKBURN, 1984). The junction of blade and sheath has an inner surface, known as the throat and an outer one, called collar. The collar consists of two or more wedge-shaped area called the joint triangle. These joints are flexible due to the large presence of collenchyma, as in the case of the sheath node and the growth ring. Its form may vary according to the plant age (DILLEWJIN, 1952).

The ligule, a membranous appendage from the sheath, separates the blade from its sheath. It is a conserved organ that is few influenced by external conditions. The ligule has elongated parenchyma cells and contains no vascular bundle. According to age, ligule dries, assuming an asymmetric shape (DILLEWJIN, 1952), being an important diagnostic feature. The leaf blade broadens form the ligule and then narrows towards the tip (JAMES, 2004).

In some varieties, the older leaves die and may drop to the ground, and if retained, they impede harvesting and may shelter pests (JAMES, 2004).

2.1.1.3 Numbering sugarcane leaves and culms

The purpose of numbering sugarcane parts, such as leaves and stems, is to identify each piece individually, determining developmental patterns and associations. The method of KUIJPER (1915) apud BENDA (1971) numbers leaves and stems from the youngest to the oldest. The leaf numbered as 1, or +1, is the first that has the dewlap exposed. The internode that is below it, by association, is considered internode number 1 as well.

Leaves older than number one, placed below it, are numbered consecutively as 2, 3, 4, ... , n, and younger leaves are numbered as 0, -1, -2, -3, ... n, and so on. Although it varies with

respect to age and variety, usually the most active growth occurs in the leaves -9 to -1, and in stems, in internodes 0 to 3 (KUIJPER, 1915 apud BENDA, 1971). Growth in length of the leaf blade is completed by 0 and of the stem by 4. Thus, internodes numbered as 1, 2 and 3 would be considered as young ones, and from 5 on could be considered mature. In leaves, from -1 on can be considered as young ones and those from 1 or +1 on could be considered mature, or totally expanded in length. Figure 1 shows an illustration of the numbered leaves, nodes, internodes of sugarcane.

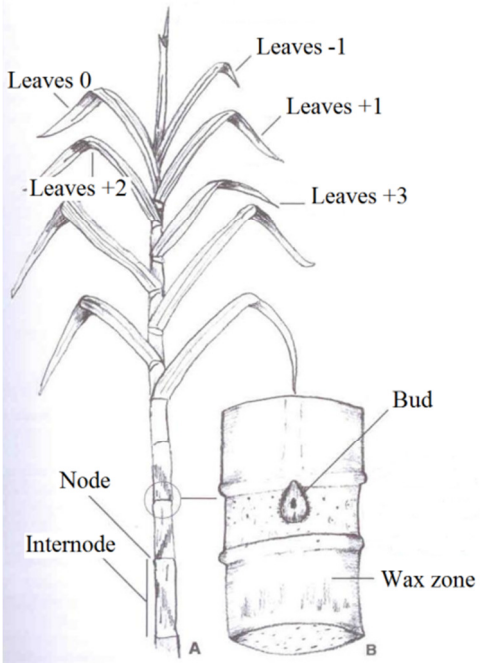

Figure 1 – Sugarcane. A. numbering of leaves by the Kuijper system. B. Node of the culm showing its structures. (Alexandre S. Pinto)

2.1.2 Sugarcane ethanol

In 2008, the world demand for energy from fossil fuels accounted for 80% of the total, while renewable energy represented around 13%, of which biomass contributed to 10% (FOOD AND AGRICULTURE ORGANIZATION - FAO, 2008). This shows that the supply of world energy needs renewable fuels to help to contain the phenomenon of global warming and rising emissions of greenhouse gases. Current technologies, together with those being developed for the conversion of cellulosic biomass into energy, can contribute to global energy needs and thereby reduce the impacts of CO_2 emissions.

Thus, sugarcane ethanol produced in Brazil is the best biofuel regarding the energy balance. The forecasts for the year 2025 show that, in accordance to improvements in sugarcane production process, not only the efficiency of ethanol generation will increase, but

there will also be an optimization in the energy balance, and reduction in the emission of greenhouse gases (MACEDO et al., 2008).

Brazil is the most competitive producer of bioethanol with a well-developed domestic market stimulated by the flex fuel cars (IBETO et al., 2011 apud AMORIM, 2011). Sugarcane bioethanol has been supported by the development of new varieties, adequate weather, fertile soils, and new technologies (LEITE et al. 2009). The big step for the industry was after the world oil crisis of 1973, when the Brazilian government launched the "ProAlcool" program, which reduced fuel prices and the dependency on oil (GOLDEMBERG, 2008). The industry has achieved good results when alcohol was used as a fuel for vehicles, encouraged by the program, created in 75/76. "ProAlcool" aimed to create an alternative source of fuel for motor vehicles based on Otto engine, fostering job creation in Brazil (SOARES; ROSSEL, 2007). In 1986, 96% of cars produced in the country had an engine fueled by ethanol (FALCO, 1998). The incorporation of innovative technologies was fundamental to increase industrial efficiency as well as to reduce the environmental impact of the distilleries. In Brazil, it was consumed more sugarcane ethanol than gasoline in 2008 (CHADDAD, 2010). A certain amount of etanol is mixed to the gasoline, a percentage that is controlled by the brazilian government as an incentive, and currently it is added 25% of ethanol in the gasoline.

Nevertheless, the "ProAlcool" program was abandoned because there was a fall in the price of the oil barrel which made the Brazilian government to remove some subsidies given to the distilleries, encouraging several of them to produce sugar instead. However, in order to overcome the low prices in the market, the industries improved their fermentation processes, developing new technologies that were transferred to distilleries allowing them to survive several crises over the last 20 years (AMORIM, 2006). So, over the years, the industries needed to overcome these drawbacks and the success largely depended on how they kept going.

Alcohol only returned to prominence after 2003 by the growing environmental concerns and also by the fear of fossil fuels depletion. Since the report on the global climate has been released, showing that it is necessary to reduce greenhouse gas emissions from burning fossil fuels, the race to partially replace gasoline, a fossil fuel, by alcohol, renewable and less polluting, has been intensified; a global concern that Brazil takes into account for nearly 40 years since the creation of "Pro-Alcool" (SOARES; ROSSEL, 2007).

Until 2004, investments in "Pro-Alcool" accumulated US$ 11 billion, and Brazil has saved up to twenty seven billion dollars in oil imports. Sugarcane ethanol production

generated 152 times more jobs than the oil industry, being also advantageous for environmental reasons (PESSOA-JR et al., 2005). Consumption grew so much that in the 2008/2009 harvest, Brazil produced more than 27 billion gallons of ethanol, increasing by nearly five billion liters compared to 2007/2008 (UNIÃO DAS INDÚSTRIAS DE CANA-DE-AÇÚCAR - UNICA, 2012).

São Paulo is the state of the largest production, growing about 60% of sugarcane and being responsible for almost 61% of ethanol production in the country (UNIÃO DAS INDÚSTRIAS DE CANA-DE-AÇÚCAR - UNICA, 2012). Brazilian government announced that until 2015 investments in the sector will reach R$ 65.000.000.000,00 (CENTRO NACIONAL DAS INDÚSTRIAS DO SETOR SUCROALCOOLEIRO E ENERGÉTICO – CEISE, 2012). Despite this, nowadays the sugar and ethanol industry is passing through a crisis, principally due to the rainfall patterns and oscillation of the prices of its main products. That is another reason to search for technologies that increase the production without increasing the planted area.

The development of the entire process for the production of cellulosic ethanol - or second-generation ethanol, may generate a striking change in the sugar industry (MATSUOKA et al., 2009). Projections show that it can double the production of ethanol (HAILING; SIMMS-BORRE, 2008), which is critical in the current scenario of reduction of greenhouse gases emissions. Furthermore, using this technology, other residues could be reused such as rice straw, in countries where sugarcane is not grown, making them consumers of biofuels and contributing to the environment (STICKLEN, 2008). However, in spite of its importance, commercial production is not feasible mainly by the high cost in order to reach and isolate the cellulose to ethanol production, and therefore alternatives have been sought to overcome this problem and reduce the costs.

The conversion of plant material in ethanol is related to the concept of biorefineries, which it is the integration of the hydrolysis with the agricultural processes, resulting in the development of an agro-industrial process. The main advantages of bio-refineries are the processing of agricultural waste into commercial products or energy source, the sharing of all technical infrastructure management, reducing operational costs, economic and financial risks, enabling flexible sources of raw materials and finished products, reducing the impact of fluctuations in international prices of commodities, reducing the impact to the environment by reducing waste generation and the possibility of its use in another activity (SOARES; ROSSEL, 2007). Thus, ethanol production, using cell wall components, can reuse the industrial process waste, having the advantages mentioned above.

The plant materials have very similar composition; however, the presence of different substances can cause significant changes in the reactions of the ethanol manufacturing process through cell wall components (SOARES; ROSSEL, 2007). An important strategy to optimize ethanol production from the bagasse is the study of genes and proteins involved in cellulose and hemicellulose synthesis, as well as other components present in the cell wall that play a key role in the manufacturing process of cellulosic ethanol. Therefore, the study of these substances in the bagasse, such as cell wall proteins, is essential to generate information for the process optimization, and the characterization and analysis of these proteinscan enable the genetic manipulation.

The search for solutions to enable the cellulosic ethanol production has become a global race, and many countries are investing in making this process possible, using advanced technology. These new technologies arising from research conducted in universities and also in industry, have been bringing promising results. Breeding combined with molecular biology, as well as biotechnology, can contribute to the formation of elite varieties by inserting characters through genetic transformation (MATSUOKA et al., 2009).

2.1.2.1 Cellulosic etanol

Cysewski and Wilke (1976) described the technique to produce cellulosic ethanol for the first time, through enzymatic hydrolysis, and then fermentation by the yeast *Saccharomyces cerevisiae*. In order to produce second generation ethanol, or cellulosic ethanol, by enzymatic hydrolysis, lignocellulosic biomass must be harvested, packed and transported to a biorefinery, where it is stored and is ready for conversion. Subsequently, the biomass is then chemically treated or pre-heated to high temperatures in order to break it into intermediates and remove the lignin, followed by detoxification, neutralization and separation of liquid and solid components. The solids are then hydrolyzed using enzymes produced in microbial reactors using bacteria or fungi. Finally, sugars are separated and fermented to produce ethanol (STICKLEN, 2008).

As the raw material for its production is cellulose, the cultures recommended and used to produce cellulosic ethanol must have high biomass concentration, as is the case of sugarcane, corn, rice, grasses, such as switchgrass and *Miscanthus* or fast-growing wood such as willow, *Populus* (KNAUF; MONIRUZZAMAN, 2004) and *Eucalyptus*. Research on linseed peel (GONZÁLEZ-GARCÍA et al., 2009), crop residue from linen fibers production in the paper industry, as well as the aerial part of cassava and peanut shells have also been

carried out (MARTIN; THOMSEN, 2007), as well as in several other species, such as food scraps and newspaper.

Ethanol advantage is not only the lower CO_2 emissions but is also its low toxicity and being quickly degradable (STICKLEN, 2008). Still, cellulosic ethanol has its particular advantages, like emitting less toxic particles, minimizing environmental pollution, and reducing the amount of fertilizers and pesticides that contaminate aquatic ecosystems (HILL et al., 2009). The high availability of cellulosic biomass compared to the limited availability of sugar and grains make possible the second generation ethanol to generate extra energy (STICKLEN, 2008).

In addition, cellulosic ethanol can be achieved either from non-food crops, as well as waste of food plants, which were previously discarded, allowing the countries that do not have the right climate for sugarcane to produce cellulosic ethanol, could produce its own biofuel, as for example, through the rice straw (STICKLEN, 2008), putting an end to the debate on using agricultural land for biofuels rather than food. Figure 2 shows a simplified system of cellulosic ethanol production.

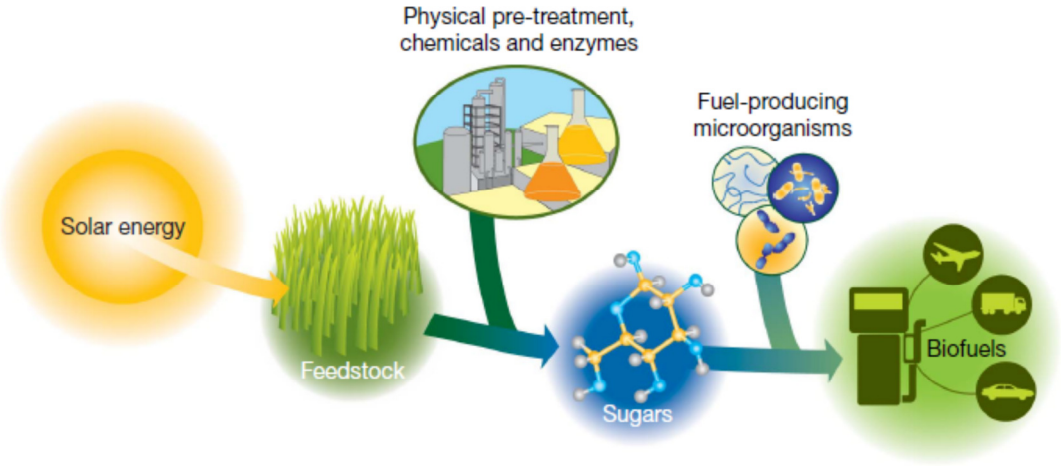

Figure 2 – A simplified system of biofuel production from cellulosic material (RUBIN, 2008)

2.1.2.2 Sugarcane bagasse ethanol

About 2/3 of the energy present in sugarcane are not used to produce ethanol, which corresponds to the portion of straw, tip and bagasse (SOARES; ROSSEL, 2007). Bagasse also can be burned and reused to generate electricity, but the income is still insignificant when compared to the production of sugar and alcohol. Estimates indicate that if this surplus biomass was reused for ethanol production, the yield would increase by 30-40% without, however, increasing the planted area. In 2010/2011, sugarcane plantations occupied about 9

million hectares, and production was approximately 27 billion liters in Brazil, increasing only about 2 billion liters when compared to the previous crop (UNIÃO DAS INDÚSTRIAS DE CANA-DE-AÇÚCAR – UNICA, 2012). Considering the estimated increase in productivity of 40% with cellulosic ethanol technologies, the same area can reach a yield of nearly 37.8 billion liters. Prospects for gains in cellulosic ethanol production may lead to the development of the so-called "energy cane," a material that produces less amount of sucrose and more quantity of fibers, being more resistant (ALEXANDER, 1985) and increasing the efficiency of obtaining second generation ethanol. However, this process is not yet completely economically viable because production is still expensive, and one of the reasons is sugarcane cell wall recalcitrance. Thus, an estimate of the future production can be expanded as new scientific and technological advances are incorporated into cellulose hydrolysis processes to obtain ethanol.

The use of technologies capable of dismantling plant cell wall requires a deeper understanding of physiology and cell wall structure. In addition, the study of enzymatic processes of microorganisms, which already involves the cell wall and therefore have specific enzymes for this purpose, may help to use the energy available in these polysaccharides.

To contribute to the understanding of sugarcane genome, expressed sequence tags (ESTs) of this crop were sequenced. The program SUCEST (THE SUGARCANE EST PROJECT – SUCEST, 2012) sequenced more than 260,000 cDNA clones, resulting in 26 standard libraries. These ESTs were grouped into 43,141 putative transcripts. The study showed that about 50% of genes are related to protein metabolism, cellular communication/signal transduction, bioenergetics, and stress response. Possibly 33,620 unique genes were found, indicating that more than 90% of the sugarcane expressed genes were sequenced (VETTORE et al., 2003). Based on this information, another work related possible sequences of sugarcane ESTs with genes expressing cell wall components. In this search, the authors found 459 potential genes that are related to the plant cell wall, about 1% of the ESTs found in SUCEST (THE SUGARCANE EST PROJECT – SUCEST, 2012). Correlation analysis indicated that the expression of genes related to wall biosynthesis is associated with hemicellulose hydrolysis, being the pectin hydrolases associated with xyloglucan hydrolases (LIMA et al., 2001). These findings contribute to understand the genetic machinery of the cell wall and provide the first information that can direct genetic engineering in the future. However, more information about which proteins are in sugarcane cell wall is required.

As previously discussed, a consortium of several countries, the Sugarcane Genome Sequencing Initiative (SUGARCANE GENOME INITIATIVE – SUGESI, 2012), is sequencing the whole sugarcane genome. The dimension of this challenge is understood when it's observed the complexity of sugarcane genome, which is derived from interspecific hybrids obtained by crossing *Saccharum officinarum* and *Saccharum spontaneum*, possessing a genome of about 10 Gb, polyploid and aneuploid. In this partnership, the Brazilian variety SP80 3280 will be sequenced, as well as several other varieties, according to the interests of each country (SUGARCANE GENOME INITIATIVE – SUGESI, 2012). With this information, it will be easier to conduct studies on gene expression, which can help new discoveries in the field of cellulosic ethanol.

2.1.2.3 Constraints on production

Cell wall polysaccharides can be converted into fermentable sugars by enzymatic hydrolysis using enzymes such as cellulases and hemicellulases. However, lignin is the main barrier to such conversion, preventing enzymes to access the polysaccharides (SOMMERVILLE, 2006). To solve this problem, pre-treatments are being developed in order to break the wall and remove the lignin, exposing the polysaccharides to the enzymes. These enzymes are produced naturally by a number of organisms, and a large number of species have been investigated for this purpose. Many enzymes have been isolated and characterized, and others are being investigated, particularly those resistant to high temperatures and pH variation during pre-treatment, two major limiting factors (STICKLEN, 2008).

A major problem in the production of cellulosic ethanol is that, if on one hand, fermentation converts about 90% of energy from simple sugars into alcohol, only 40% of the energy contained in lignocellulosic materials can be converted into alcohol with the technologies that exist. This means that cellulosic ethanol mills require more raw materials compared to the first generation ethanol to make the same amount of ethanol (SERVICE, 2010).

Another bottleneck in cellulosic biofuel production lies in the fact that only the six-carbon sugars are utilized, and there is no route to transform the five-carbon sugars in ethanol (MARQUES, 2009). To do so, it would be required the development of new yeasts that are able to transform pentoses into ethanol, further increasing its income.

2.1.1.1 Cell suspension cultures

Suspension cultures of plant cells have been used for a number of studies involving the cell wall proteins, the elucidation of defense mechanisms and many others. As an example, plant cell suspension was also used to study the secondary wall formation as well as lignification (BLEE et al., 2001). Another work showed some differences in lignin structure between primary and secondary cell wall using aspen cell culture (CHRISTIERNIN et al., 2005). A strategy chosen to understand the relation between changing the nutrient conditions and the proteome content was applied in *Arabidopsis thaliana* cells, and it was shown that the culture medium may significantly influence the expression pattern of the soluble proteins (SARRY et al., 2006).

Cell suspension cultures are formed by cells aggregates as well as disperse cells, in a liquid medium, growing by agitation. The quantity of material grows until it reaches to a maximum, when the culture can be diluted through subculturing, allowing it to be continuously propagated (STREET, 1977 apud FALCO, 1993).

Suspension culture cells generally contain only a primary cell wall, and can be used in experiments for being reproducibly grown, obtaining large quantities of biomass (SLABAS et al., 2004).

Several studies obtained sugarcane suspension cells from sugarcane calli, and particularly in the work of Falco, Mendes and Tullmann-Neto (1996), the cultures were composed predominantly of embryogenic cells, being round, uniform and densely cytoplasmic. In general, to maintain the culture, the flasks should be kept in agitation, the culture medium has to be changed periodically and the examination of the culture under a microscope can be essential to obtain a good culture (FALCO, 1993).

2.1.3 Plant cell wall

Plant cells have walls outside the plasma membrane. The wall confines the cell volume, and together with the internal osmotic pressure, reaches the mechanical stability. The wall has been considered inert, but evidence indicates that it is dynamic and active in many processes such as signaling and coordination of growth and development (REITER, 1998). The cell walls participate in processes such as growth and cell elongation (FRY et al., 1992), defense against pathogens (BRADLEY et al., 1992), and provide structural support to individual cells, tissues and the whole plant. In sugarcane, cell wall also plays a fundamental role in sucrose

distribution as well (GLASZIOU; GAYLER, 1972). However, the mechanisms that underlie a large number of the cell wall functions are not well understood.

The plant cell wall consists of a complex matrix of carbohydrates polymers, proteins, and others, and carbohydrates constitute about 90% of the mass and the proteins around 10% (WOJTASZEK, 1997). However, precisely because it is dynamic, the wall is constantly modified during growth and development (CHIVASA et al., 2002). Despite their low abundance, wall proteins comprise several hundreds of different molecules with various functions (CARPITA et al., 2001), being involved in the modifications of the components, their structure, signaling and interactions with plasma membrane proteins on the cell surface (JAMET et al., 2006). Since the polysaccharides that compose the cell wall may function as interconnected fibers or matrix, they are classified into three categories: pectins, hemicelluloses and cellulose (CARPITA; MCCANN, 2000). Figure 3 shows the main components of the plant cell wall.

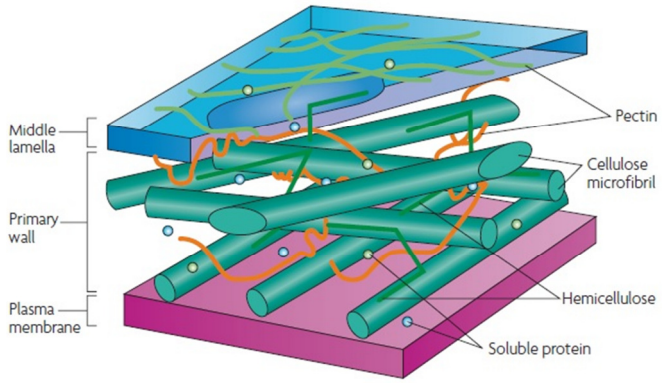

Figure 3 – Cell wall containing cellulose microfibrils, hemicellulose, pectin, lignin and proteins. (STICKLEN, 2008).

Hemicelluloses prevent the collapse of cellulose fibers, which are parallel to each other, allowing a weak interaction among them and thus forming a network. Hemicelluloses can be defined as the polysaccharides that can bind to cellulose through hydrogen bonds (ALBERSHEIM et al., 2011). The cellulose-hemicellulose complex is immersed in a domain consisting of highly branched pectins and sugars, which determine the porosity of the wall and signal the presence of pathogens and insects (BUCKERIDGE et al. 2008).

One mechanism that controls cell extension is the pectic layers that allow easy movement among the cellulose and hemicellulose layers. These pectic layers are composed of pectins, polysaccharides composed of D-galactosiluronic acid in a large proportion (> 20 mol%) (ALBERSHEIM et al., 2011).

Cellulose microfibrils are coated with xyloglucans, and have an important role in wall structure (ARIOLI et al., 1998). Cellulose is the most stable and also the toughest polysaccharide present in the cell wall. It is extremely simple, containing continuous residues of β-1,4-glucopyranosyl, and this binding requires that alternate residues be placed from 180° relative to each other. Usually, primary cell walls are about 30% cellulose, whereas secondary cell walls contain 50% cellulose. However, specialized cells can contain up to 95% cellulose, as in the case of cotton fibers (ALBERSHEIM et al., 2011).

In the case of grasses, the cell walls contain different matrix polysaccharides and protein components, when compared to dicots, even when displaying the same functions. As an example, the grasses cell walls present the microfibrils interlocked by glucuronoarabinoxylans instead of xyloglucans. And, unlike type I cells, the grasses cell walls contain a substancial portion of the non-cellulosic polymers 'wired on' the microfibrils by alkali-resistant phenolic linkages. Probably, during growth, the β-D-glucan appears during cell elongation, being the principal interlocking polymer. Then, long stretches of glucan could bind to cellulose or other glucan, and cleavage of glucosyl-units by endo-glucanases could free the cellulose microfibrils. Thus, the fundamental cross-linking role in the primary wall of growing cells is from the esterified and etherified phenolic compounds that lock the wall into place and, then stretch the microfibrils in grasses (CARPITA; GIBEAUT, 1993).

Concerning classification, cell walls can be primary, associated with young tissue in growth and differentiation, or secondary, characterized by its presence in the terminal stage of cell development, in which the cell shape and size are permanently established (LEE et al., 2004). The composition and structural organization of the primary cell wall allow it to be strong, thin, flexible and capable of undergoing plastic and elastic extension (FRY, 2004).

The last decades have shown progress in elucidating the structure and organization of the extracellular matrix, but also revealed the importance of the wall as an interface between the biotic and abiotic environment (HOSONO, 1998). In addition, research about plant biology resulted in a growing list of development processes that are influenced by molecular interactions and signaling patterns, since a substantial portion of the proteome in plants is located in the cell wall (LEE et al., 2004).

In addition to biological functions, plant cell wall attracts great interest in the economic area. The wall is used commercially for papermaking, textile manufacturing, fiber, charcoal, construction, wood products, to generate energy by burning bagasse of sugarcane or second generation ethanol.

2.1.3.1 Cell wall proteins

There are three characteristics common to classical cell wall proteins. The first is a cleavage signal-peptide on the *N*-terminus, that is responsible to target these proteins to the Endoplasmic Reticulum (RAPOPORT, 1992); the first organelle in the secretory pathway (VITALE; DENECKE, 1999). The signal peptide is not formed by a consensus amino acid sequence, however, it has a positively charged n-region on the *N*-terminus and a central hydrophobic h-region followed by a polar c-region containing the cleavage site (CHIVASA et al., 2002). Another characteristic is that cell wall proteins do not possess the KDEL and HDEL tetrapeptide on the *C*-terminus, the canonical ER retention motif that prevents the secretion of proteins that should remain in this organelle (VITALE; DENECKE, 1999**,** CROFTS et al., 1999). The thrid feature of proteins that are going to be secreted is that they do not have a transmembrane domain in their amino acid sequences. In the secretory pathway, proteins goes from ER to the Golgi complex in order to be packed into vesicles that are directed for the site of secretion to the extracellular matrix; the plasma membrane. Thus, integral plasma membrane proteins use the same features of Cell Wall Proteins - CWPs in order to be directed to the membrane. However, these membrane proteins have a hydrophobic domain that is the transmembrane domain, which makes them to become embedded into the lipid bilayer (CHIVASA et al., 2002). Therefore, sequences that contain a transmembrane helix downstream of position 40 can be assigned as transmembrane proteins (EMANUELSSON et al., 2000).

Cell wall proteins can be classified into several functional categories. The three major functional categories are: wall-modifying proteins (DARLEY et al., 2001), structural proteins (CASSAB, 1998) and defense proteins synthesized in response to biotic and abiotic stresses. The plant cell wall can even affet the cell fate, sensing stress signals and transmitting them to the cell interior (KONG et al., 2010). However, other studies can extend this functional classification, adding classes like proteinases, unknown proteins and proteins of various functions (BOUDART et al., 2005). The plant cell wall is involved in many fields of research, due to the already cited versatility when it comes to its function. That is why the term cell wall can be related not only to the structural matrix that surrounds all plant cells, but also to include the apoplast, or the extracellular environment (ROSE; LEE, 2010).

The plant cell wall contains several classes of structural proteins, classified according to their amino acids composition such as hydroxyproline-rich glycoproteins (HRGP), glycine-rich proteins (GRP) and proline-rich proteins (PRP). These proteins are regulated by

development and their amounts vary according to tissue type. An example of HRGP is extensin (CARPITA; MCCANN, 2000), which is involved in relaxation and extension of the cell wall (COSGROVE, 2003). In addition to these structural proteins, cell walls also contain arabinogalactans (AGPs), which are highly glycosylated containing mainly galactose and arabinose residues (TAIZ; ZEIGER, 2004).

On the other hand, expansins play an important role in cell expansion (COSGROVE, 2001), promoting extension and relaxation of isolated cell walls (MCQUEEN-MASON et al., 1992). Thus, expansins have significant effects on the loss of the wall, because they induce wall extension *in vitro*, promoting faster growth, and their genes are expressed at the exact time and place of cell growth control (COSGROVE, 2001).

Enzymes such as peroxidases have a role in stopping the expansion to make a rigid wall (ANDREWS et al. 2002; PEDREIRA et al., 2004, PRICE et al., 2003). The cell expansion is regulated by various stimuli such as light, gravity, water stress, hormones, and anoxia (HOSONO, 1998). Several studies describe correlations among these regulatory factors and changes in the structure of wall polysaccharides and primary expression of proteins located on the wall, or genes corresponding enzyme activity (LEE et al., 2004).

The wall contains enzymes capable of modifying the polysaccharides matrix (FRY, 1995), including endoglicanases that cleave the backbone of the polysaccharide matrix; glycosylases that remove side chains; transglycosylases, which cut the polysaccharides and then link them together; esterases that remove methyl groups of the pectins and cleave ester bonds among polysaccharide chains, and peroxidases that form or break phenolic bonds in the wall. These enzymes offer many possibilities to change the wall structure, and modulate its expansion. Enzymes that cleave specific places on the wall (CARPITA; GIBEAUT, 1993) offer great potential for mechanical and structural analysis of the cell wall.

The sequences databases are expanding, as an example is the sequencing of the genome and ESTs from sugarcane (VETTORE et al., 2001), and increasing genetic resources offer additional tools to test the role of specific genes in the structure and function of plant cell wall proteins.

Three types of wall proteins can be distinguished according to their interactions with the wall components. Loosely linked proteins do not have interaction with the polysaccharides of the wall and move freely in the intercellular space. The weakly bound proteins can be connected to the matrix by Van der Waals interactions, hydrogen bonds or ionic interactions. Weakly bound proteins represent only 5-10% in the primary cell wall of all proteins within the complete mass of the cell wall, but despite their low abundance, they are linked to

important roles such as remodeling, signaling, interactions with the plasma membrane and defense (CASSAB; VARNER, 1988). The proteins tightly bound to the wall are resistant to salt extraction, like the extensins which are connected by covalent bonds (JAMET et al., 2008). Figure 4 shows a scheme of cell wall proteins and their link to the wall components.

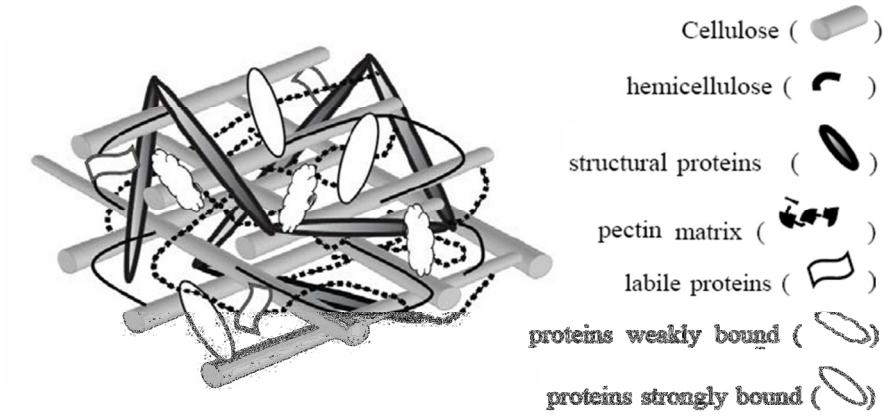

Figure 4 – The principal classes of cell wall proteins according to their link to the wall components (JAMET et al., 2008)

The composition and structure of proteins from plant cell walls are different according to their development and tissue, thus we can also compare different organs and relate them to their functions and their stage of development (JAMET et al., 2008).

Several studies have been performed in order to identify cell wall proteins in species like *Arabidopsis thaliana* (CHIVASA et al. 2002; BORDERIES et al. 2003; BOUDART et al. 2005; FEIZ et al., 2006, BAYER et al., 2006; IRSHAD et al., 2008), maize (ZHU et al., 2006), grape (NEGRI et al., 2008), rice (CHEN et al., 2009), tobacco (MILLAR et al., 2009), among others. However, most of the existing work carried out the extraction and analysis of proteins in poorly differentiated tissues and cell suspensions, calluses, rosettes, hypocotyls and primary roots, and a more appropriate and specific method is used with each species and type of tissue. Thus, a study that can isolate and identify proteins in mature tissues, as well as working with species never used before, will face a number of difficulties until obtaining the optimal protocol adapted to that particular situation.

From the approximately 400 CWPs identified in *Arabidopsis*, the most abundant proteins are enzymes for polysaccharides, such as glysoide hydrolases. The second most representative group is the exido-reductase enzymes, such as peroxidases, and the third covers proteins with interacting domains, such as protease inhibitors. Other proteins found were signaling and lipid metabolism-related proteins, as well as proteins with unknown function, among others (JAMET et al., 2008).

2.1.3.2 Trafficking of plant cell wall proteins

Cell wall biogenesis, growth, differentiation, remodeling, signaling and defense responses depend on the efficiency of the secretory pathway (DE CAROLI et al., 2011). The secretory pathway of eukaryotic cells comprises organelles that link three membranes: the plasma membrane, the vacuole and the endoplasmic reticulum (FORESTI; DENECKE, 2008). In the secretory pathway, many processes have evolved differently in plants and that is why they cannot be studied using other species cells as models (FORESTI; DENECKE, 2008).

The organelles that are best characterized are the Golgi and the prevacuolar compartment, however recent work highlighted the role of the trans Golgi network, which can be designated as a separate organelle in plants (FORESTI; DENECKE, 2008). Matrix polysaccharides, synthesized in the Golgi, and cell wall proteins, synthesized in the Endoplasmic Reticulum (ER), are packaged into vesicles that merge with the plasma membrane releasing the proteins into the cell wall through the secretory pathway (DE CAROLI et al., 2011). The Golgi apparatus and the post-Golgi organelles that control vacuolar sorting, secretion and endocytosis are intermediate organelles of the endocytic and biosynthetic routes (FORESTI; DENECKE, 2008). The synthesis and deposition of cellulose are also driven by the endomembrane system, which controls the assembly, and the export to the plasma membrane cellulose synthase rosettes, within the Golgi (SOMMERVILLE, 2006).

The pathway for secretion of plant proteins involves an N-terminal sequence, also known as the signal peptide or leader sequence, which directs the nascent protein to the endoplasmic reticulum, where the protein is cotranslationally translocated through the Sec61 complex into the ER lumen (RAPOPORT, 2007). That is when some proteins can undergo *N*-linked glysosylation. The proteins are then packaged into vesicles while passing through the endomembrane system, comprising the Golgi apparatus and the trans-Golgi network, where these vesicles migrate to, and fuse with, the plasma membrane, releasing the proteins in the cell wall, consisting the classical secretory pathway. However, several proteins can be retained in the ER or Golgi, including phases of retrogade trafficking, target to the vacuole or post-Golgi compartments (FORESTI; DENECKE, 2008).

There are growing evidences about routes to the secretion of plant proteins that are independent of the ER-Golgi pathway (ROSE; LEE, 2010), which have been already described in mammalian cells (COOPER; BARONDES, 1990; RUBARTELLI et al., 1990), seeming likely that a nonclassical secretion is common to all eukaryotes. Surely the main

pathway for secretion is the classical ER-Golgi route, and only a small part of proteins is likely to be secreted by nonclassical ways, which remains a challenging and uncharacterized area (ROSE; LEE, 2010). Recent evidences show that the Golgi, or other compartments of the secretory pathway, can retain cell wall proteins, prior to targeting to the apoplast, however is still a case in point (DAL DEGAN et al., 2001; WOLF et al., 2009). Probably these alternative pathways may coordinate spatiotemporal proteins secretion in plants (ROSE; LEE, 2010).

The classical mechanism of transport of proteins in the secretory pathway involves the presence of a signal peptide on the proteins N-terminus, responsible for targeting them to the endoplasmic reticulum - ER. The signal peptide has no consensus sequence, however it has three conserved structural determinants: the *n*-region is positively charged, a hydrophobic *h*-region is in the middle and the polar *c*-region is adjacent to the peptidase-cleavage site. Also, it has a conserved motif presenting a small and a neutral aminoacid at positions -3 and -1 from the cleavage site (von HEIJNE, 1985). Then, the ER retains some proteins possessing a KDEL/HDEL retention motif in the C-terminal, while others can proceed to the Golgi apparatus. In the Golgi complex, proteins can have target signals to the vacuole, being sent there, whereas the rest of proteins proceed to the plasma membrane. There, if the protein posses an extensive hydrophobic domain, it stays embedded in the plasma membrane, otherwise, in its absence, the protein is secreted to the extracellular matrix. Possesing this information, gives one the bioinformatic tools to find out wheter the proteins can have a cell-wall location prediction (SLABAS et al., 2004).

Therefore, determining the subcellular localization of a protein is essential to understand its function (EMANUELSSON et al., 2007).

2.1.3.2.1. Moonlighting proteins

More and more, proteins are found to be located in multiple cellular compartments. Also called ancient enzymes, several of the already known moonlighting proteins are highly conserved enzymes (HUBERTS; der KLEI, 2010). As an example, Slabas et al. (2004) showed that, probably, a citrate synthase and a luminal binding protein can be found in different organelles, such as mithocondrion and ER lumen, respectively, but also were found to have a signal-peptide predicition. This kind of bias makes it difficult to use bioinformatics tools to predict the cellular location of a protein.

Another example can be seen in an apyrase that can be involved with mRNA transport in the nucleous as well as a phosphate transporter and a disease resistance protein when it is in the extracellular matrix (DAVIES et al., 2001). Therefore, slight post-translational modifications may occur allowing multiple locations and functions to a same protein.

Intrinsically disordered proteins can be an example of protein signal moonlighting having multiple, and, sometimes, unrelated functions, like activating and deactivating the same or different partners (TOMPA et al., 2005).

Elongation factor 1 has also being seen as one of the most important multifunctional protein, displaying several roles in a large number of organisms (EJIRI, 2002). Another example is Fructose-1,6-biphosphate aldolases, that are traditionally cytoplasmic glycolytic enzymes, but were also found to be located extracellularly exhibiting no glycolytic function, despite the lack of a secretion sign, in *Neisseria meningitis*. Their role was found to be related with cell adhesion to the host cell (TUNIO et al., 2010). In the human malaria parasite *Plasmodium falciparum*, an isoform of histone 3 resides in the parasitophorus vacuole, playing a role in the host-parasite relation (LUAH et al., 2010). Other examples can be seen in histone 1, cyclooxygenase, carbinolamine dehydratase, chaperone, peroxiredoxin, ribosomal proteins, among several others (JEFFERY, 2004).

Most of these proteins were already found in cell wall proteins studies, showing that, probably, even not having the prediction of the signal-peptide to secretion, they can play a role in the extracellular matrix.

2.1.3.3 Extraction and analysis of cell wall proteins

Substantial progress has been made recently about the characterization of plants subcellular proteomes, such as chloroplasts and mitochondria. However, much less is known about the proteome of the cell wall, which probably reflects major technical challenges to be overcome (LEE et al., 2004).

The aim of the subcellular fractionation prior to proteomic analysis is to capture the most accurate representation of the proteins, and minimize contamination by proteins from other sites. Thus, efficient protocols have been developed to isolate organelles involving tissue breakdown, followed by density centrifugation, separating the spectra of organelles in highly enriched fractions. In such cases, the membrane acts to contain the proteins of organelles and exclude contaminant proteins. In contrast, any disruption of the tissue that

compromises the integrity of the plasma membrane leads to contamination of the cell wall fraction that has intracellular proteins, many of which bind to the wall matrix with high affinity (FEIZ et al., 2006). However, without disrupting the cell wall, some methodologies do not work, because it is necessary that the extraction buffers come into contact with the proteins. Figure 5 shows different strategies to perform cell wall proteins extraction.

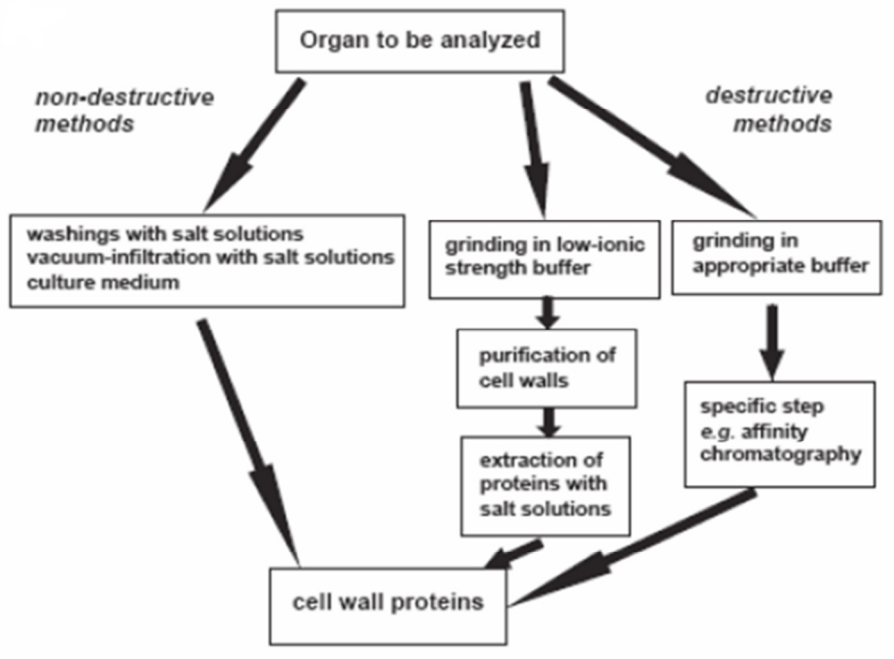

Figure 5 – Some strategies used to extract cell wall proteins (JAMET et al., 2008)

The analysis of cell wall proteins depends on the quality of sample preparation. Extraction is particularly complex since the proteins in the apoplast may be free or embedded in a matrix of polysaccharides, which are held by Van der Waals interactions, hydrogen bonding, ionic or hydrophobic interactions or covalent bonds (BORDERIES et al., 2003). Another difficulty is the lack of a limiting membrane, which can result in loss of cell wall proteins during the isolation process. As mentioned above, obstacles are the networks of cellulose, hemicellulose and pectins that form "traps" for intracellular proteins, contaminating the extract. The extraction can also be difficult because the proteins are embedded in a matrix of polysaccharides and interact in different ways with other cellular components. Moreover, the separation of proteins by 2-DE wall is not very efficient, since most of them are basic glycoproteins. A review on CWPs concluded that about 60% of them have a basic pI and around 80% can be extracted by salts (JAMET et al., 2006). After extraction, a final challenge is the identification of highly *O*-glycosylated proteins, which can be very difficult using the

mapping of the peptides masses by mass spectrometry. Thus, specific strategies have been developed to overcome these difficulties (JAMET et al., 2008).

Cell wall proteins (CWPs) have biochemical characteristics and affinity with the matrix of polysaccharides. Many are soluble with no apparent interaction with the wall polysaccharides and thus combined with the protein fraction from the cytosol during tissue homogenization, the extract is lost from the wall (LEE et al., 2004). However, other proteins, such as extensins (WOJTASZEK et al., 1995), have covalent bonds with the cell wall and in consequence are resistant to extraction. In addition, since cell wall proteins can be highly glycosylated, making the extraction more difficult, as well as isolation and identification, and other extracellular proteins have domains that are anchored in the plasma membrane, which represents a barrier to cell wall proteins extraction (BORNER et al., 2003).

Thus, there isn't just one protocol that can be used to extract all the proteins of the wall, and then techniques should be used that are appropriate to a specific subset of proteins of interest, to the tissue and to the species that one is dealing with. These techniques can be divided into two categories: disruptive, that cleaves the wall and the membrane, and not disruptive.

A strategy to obtain proteins located in the cell wall involves the use of plant cultured cells, minimizing contamination with intracellular proteins. This system was used in a study pioneered by Robertson et al. (1997). Proteins were isolated from suspension cells and intact cells were washed sequentially with a series of solvents formulated to leave the plasma membrane intact. Several families of proteins located in the cell walls have been identified. Feiz et al. (2006) evaluated different types of cell wall isolation, identifying proteins and establishing protocols containing a minimum of intracellular proteins.

Often, CWPs extraction requires the use of different salts and chelating agents that release proteins from the wall, as well as bioinformatics tools and/or molecular markers used to identify these proteins and predict their sublocation (BORDERIES et al., 2003).

The extraction of CWPs has also been reported in different species and materials, such as loosely bound protein in callus of rice (CHEN et al., 2009); cell wall proteins of the alga *Haematococcus pluvialis* (WANG et al., 2004), proteins from corn root (ZHU et al., 2006), among others. The studies cited show the variety and growing interest in this type of subcellular proteomic characterization. However, each type of tissue used and species produce specific results, as it is not possible to compare different species and the percentage of contamination, for example. The data depend largely on the anatomy and morphology of each organism, and therefore the extraction protocols should be adapted in each case.

After sequencing proteins by mass spectrometry, the next step is to identify those proteins. The identification can be performed by programs such as Mascot (MATRIX SCIENCE, 2012) or ProteinLynx, for example, based on the species of interest or a whole group. The experimental mass values are then compared to masses of peptides or fragments masses. Using an algorithm to measure score, the most appropriate combination is found. Thus, by comparison, the program selects a list of proteins that are compatible with the sample. In the analysis of cell wall proteins, there is an extra and exhausting step on the prediction of subcellular location. This analysis is performed on specific programs and sites that use different algorithms to perform this prediction. Figure 6 shows different methods to identify the cell wall proteins.

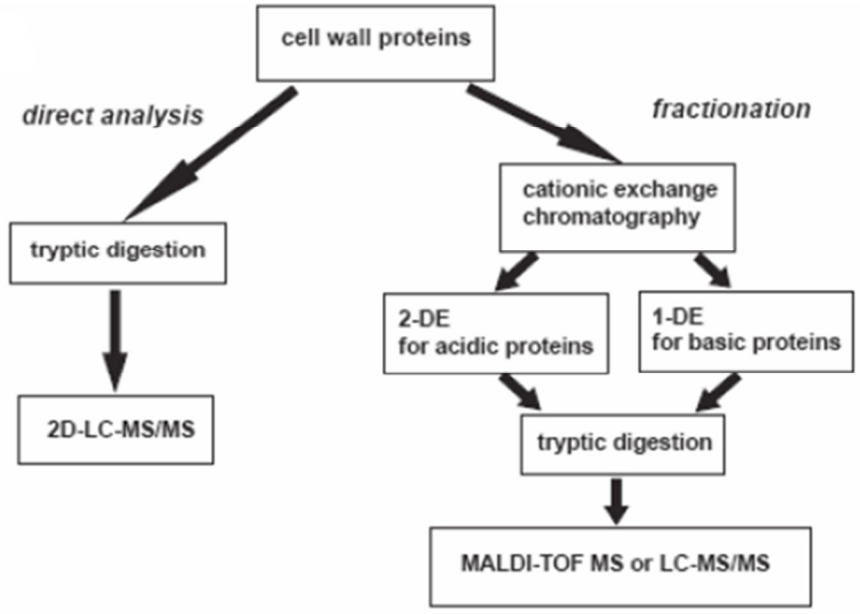

Figure 6 – Strategies used to sequence cell wall proteins (JAMET et al., 2008)

The bioinformatic tools are used in two ways in the analysis of CWPs: the identification of proteins through sequencing and analysis to predict their subcellular localization and characterize their function. To facilitate interpretation of these data, the ProtAnnDB Annotation-Protein DataBase (ANNOTATION PROTEIN DATABASE – PROTANNDB, 2012) has a database of subcellular localization and functional domains performed by programs that provide scores to the quality of these predictions. For the moment, ProtAnnDB has sequences of Arabidopsis, rice and Brachypodium (SAN CLEMENTE et al., 2009).

WoLFPSORT program (HORTON et al., 2007) predicts the subcellular localization of proteins through their amino acid sequences. This program makes predictions based on known sorting signal motifs and some correlative sequence features (HORTON et al., 2006). Thus,

proteins labeled to the secretory pathway, with a signal-peptide at the N-terminal sequence, are considered to be secreted, and consequently, CWPs. WoLFPSORT predicts the presence of the signal sequence based on PSORT method (PREDICTION OF SORTING SIGNALS AND LOCALIZATION SITES IN AMINO ACIDS SEQUENCES – PSORT, 2012), that in the other hand was developed through the method of Mcgeoh (1985) modified by Nakai and Kanehisa (1991). The method treats the N-terminal basic load (RL) and the central hydrophobic region (UR) of the signal sequences. A discriminant score is calculated with three values: length of the RH, the peak of the UR and the CR network load. These results are summarized in "McG." A large positive score means discriminating a high probability of having a signal sequence, whether or not cleavable. Subsequently, the software implements the method of von Heijne for the recognition of the signal sequence (VON HEIJNE, 1986). It's a weight matrix method and incorporates patterns information around the cleavage sites (the rule (-3, -1)) and then can be used to detect non-cleavable signal sequences. The score of "GvH" is the original score of the weight matrix subtracted by 3.5. A high score indicates a high possibility of having a signal sequence cleavage. The position of a possible cleavage site, the C-terminal portion of a signal sequence, is also shown. Experiments show that the overall prediction accuracy of WoLF PSORT is over 80% (HORTON et al., 2006).

Another program commonly used is Predotar (SMALL et al., 2004). For each protein sequence, Predotar provides an estimate of the probability to contain a targeting sequence to the mitochondria, the endoplasmic reticulum or the plastids. These estimates assume that the sequence in question was selected at random from a proteome that about 10% of the proteins are targeted to the mitochondria; 10% to the plastids and 20% for the endoplasmic reticulum. Predotar screens large batches of proteins for identifying putative targeting sequences. It has a low rate of false positives compared with similar programs, but a small drop in sensitivity (SMALL et al., 2004).

TargetP (EMANUELSSON et al., 2000) also predicts the subcellular localization of eukaryotic proteins. The location is predicted based on the presence of any pre-N-terminal sequences: the chloroplast transit peptide (CTP), the mitochondrion targeting peptide (MTP) or peptide-signal to the secretory pathway (SP). A potential cleavage site can be predicted as well (NIELSEN et al, 1997). TargetP makes the subcellular prediction similarly to SignalP 4.0 program (PETERSEN et al., 2011). SignalP incorporates a prediction of cleavage sites and a signal peptide/non-signal peptide prediction based on a combination of artificial neural networks (PETERSEN et al., 2011).

In turn, the TMHMM program (PREDICTION OF TRANSMEMBRANE HELICES IN PROTEINS – TMHMM, 2012) predicts the presence of transmembrane helix in proteins. It is considered the best prediction program of TM helix in proteins (MOLLER et al., 2001), since predicted transmembrane segments may be signal peptides (KROGH et al., 2001; SONNHAMMER et al., 1998).

2.1.4 Proteomics and mass spectrometry

Proteomics is the systematic analysis of the proteome, the proteins derived from the gene expression (CHEN; HARMON-REF1, 2006). This kind of "omic" studies the protein complement of cells, using techiniques such as identification, modification, quantification and localization (YATES et al., 2009). Its importance relies on the research areas in the biosciences that are critically dependent on the development of new and improved technologies. Proteomics can also provide information on key components of biological systems that cannot always be inferred from DNA or RNA data.Sample preparation is a challenge, with sample complexity and high dynamic range being factors that can compromise the analysis (DUNN, 2012).

There are two principal approaches in proteomics: the bottom-up and the top-down. The bottom-up proteomics is used when the proteins are proteolytically digested into peptides, prior to mass spectrometry analysis. Then, the experimental data obtained from the MS analysis are compared to already existing databases. The top-down approach uses masses of intact proteins and their fragments to compare to the databases and identify them, or, in other words, there is no digestion of the proteins into peptides prior to the analysis (YATES et al., 2009).

In plants, one of the most critical steps is the tissue disruption, which usually happens with the grinding of the materials using a mortar and a pestle, in the presence of N_2, since the presence of the plant cell wall can impose some resistance to the cell disruption. Then, in a total proteins extraction, there are three main kinds of extraction buffer: phenol, TCA or acetone, and one using all these buffers together, in order to remove secondary metabolites. The use of 2-DE technologies, as well as new gel-free protein separation processes can be a good alternative to separate the profiles of large and complex proteins mixtures (WANG et al., 2009).

Mass spectrometry is the most comprehensive and versatile tool in large-scale proteomics, and different techniques can be used to achieve proteins mass identification. Usually, a mass spectrometer consists of three parts: ion source and optics, the mass analyzer

and the data processing electronics (YATES et al., 2009). The most used methods of ionization are the ESI and the MALDI. The ESI, or ElectroSpray Ionization, produces ions from solution, driven by high voltage applied between the emitter, at the end of the separation process, and the inlet of the mass spectrometer. Then, there is the creation of electrically charged spray, Taylor cone, followed by formation and desolvation of analyte-solvent droplets. The MALDI technology, or Matrix Assisted Laser Desorption Ionization, absorbs the laser energy and transfers it to the analyte, whereas the heating causes desorption of matrix and ions of the analyte into the gas phase. The mass analyzers can store ions and separate them based on the mass-to-charge ratios. Two main categories of mass analyzers are used: the scanning and ion-beam mass spectrometers, such as TOF and Q; and the trapping mass spectrometers, such as IT, Orbitrap, and FT-ICR (YATES te al., 2009).

Bioinformatics is also an important part of proteomic analysis, involving the application of computational methods and several areas, such as machine learning and statistics (DUNN, 2012). Still, bioinformatics can help in the data analysis, either helping to study post translational modification (YATES et al., 2009), or predicting subcellular location of a protein (JAMET et al., 2006), among other uses.

Mascot is a bioinformatics software that uses mass spectrometry data to identify proteins from sequence databases. The different search methods used by Mascot are peptide mass fingerprinting, sequence query ans MS/MS ion search. Taking a small sample of the protein and digesting it, with tryosin, for example, makes the sample suitable to be analysed by mass spectrometry. The experimental mass values are then compared with calculated peptide mass or fragment ion mass values and. by using a scoring algorithm, the closest match or matches can be identified (MATRIX SCIENCE, 2012).

On the other hand, ProteinLynx Global SERVER™ (PLGS) is a fully integrated Mass-Informatics™ platform for proteomics research. It an analytical platform for Waters. PLGS accountd for the selectivity of exact mass measured and the specificity of MS^E analysis. Filtering and scoring can help to minimize the false positive results. PLGS provides an automated workflow for high-throughput data analysis (WATERS, 2012).

Another program used in many proteomic and genomics studies is named Blast2GO, a research tool designed to enable Gene Ontology (GO) based data mining on sequence data, using statistical analysis. This tool offers a good platform for functional genomics and proteomics research in both model and non-model species. It is an interactive application, allowing monitoring and comprehension of the annotation and analysis process (CONESA et al., 2005).

2.2 Material and methods

2.2.1 Material

For the extraction and the identification of proteins from sugarcane cell wall, three materials were chosen: suspension cells, leaves and stems.

2.2.1.1 Culturing sugarcane suspension cells

Active cultured cells were chosen for reasons of uniformity, since, normally, only the primary cell wall is present and there are several protocols which utilize this material (BORDERIES et al., 2003; CHIVASA et al., 2002).

The method for culturing the cell in suspension was based on Falco (1993). The cell cultures and *in vivo* material were provided by CTC, the Sugarcane Technology Center, by Dr. Maria Cristina Falco. Both cell cultures and the plants are from the sugarcane brazilian variety SP 80-3280. This variety was chosen for being sequenced by SUCEST-FAPESP project (VETTORE et al., 2004) as well as is one of the varieties chosen to be sequenced in the genome project, as already detailed in the review section. Its features include the high sucrose content and productivity, intermediate tillering and good canopy closing, due to vigorous early growth.

In order to maintain these cultures, the cells were cultivated in MS medium (MURASHIGE; SKOOG, 1962), supplemented with 50 mL/L coconut water; 3 mg/l 2,4-D and 2 g/l acid hydrolized casein, besides thiamine: 0,1 mg/l; nicotinic acid and pyridoxine: 0,5 mg/l; and 100 mg/l myo-inositol. The pH was adjusted to 5.8. The medium was sterilized in an autoclave each time before use to prevent contamination.

The suspension cells were kept in the dark at a temperature of 26 ± 1 °C under stirring at 110 rpm cultured in 50 mL of medium, in flasks of 250 mL, to provide adequate aeration. Every seven days the medium was changed to avoid the accumulation of phenolic substances, possibly toxic to the cells. The stock was monthly maintained in Petri dishes containing solid medium in order to avoid problems regarding contamination by fungi and bacteria.

Furthermore, the cultured cells were observed under an optical microscope (Olympus BX40) after seven days of cultivation in order to assess viability and anatomy, with or without

staining blue dye 70% blue austral/30% Safranine, and can be seen in the photographs (Olympus PMC35B) in Figure 8.

After seven days of subculturing, suspension cells from four different flasks were collected and used in CWPs extraction for each replicate.

2.2.1.2 Plant material

Sugarcane plants were acclimated in a greenhouse at controlled temperature (26 °C). The stems of the plants were cut in pieces and then planted in vases containing a mixture of vermiculite 1:1 substratum (Plantmax). Sugarcane plants that resprouted from these pieces were watered daily and nutrient solution (PlantProd 4g/L) was added every fifteen days. At the ages of two months, sugarcane leaves and stems were collected and their cell wall proteins were extracted. The age of 2 months was selected for presenting a soft material, since most of the already existing CWPs extraction protocols used young tissues that can present growing structures. Therefore, extracting cell wall proteins from these young tissues could show which proteins can be involved in the expansion of the cell wall, and thus, clarify the mechanisms that the own plant uses to break its own cell wall and grow.

The experimental design was completely randomized, through the use of five plants for each replicate.

2.2.2 Methods

2.2.2.1 Isolation and extraction of plant cell wall proteins

Some methods have been tested to achieve isolation of cell wall proteins aiming a high amount of proteins and trying to avoid the appearance of contaminating proteins, as the intracellular ones.

After testing several protocols, for isolation and extraction of cell wall proteins, three protocols were used in order to obtain enough quantity of proteins and enable analysis. For sugarcane culms two protocols were used based on Borderies et al. (2003) and Feiz et al., (2006), and for leaves only the last protocol was used. The one used in cells cultures CWPs extraction was based in the mixture of the two already mentioned protocols (BORDERIES et al., (2003); FEIZ et al., (2006)).

2.2.2.1.1 Isolation and extraction of cell wall proteins from cell cultures (based on Borderies et al. (2003) and Feiz et al. (2006))

About four flasks containing 200 mL of a week-old cultured cells each were collected for the experiments. This amount of cells corresponded to a volume of 10 mL when centrifuged, in average. The cells were filtered on a filter paper to discard the culture medium. The first steps of this methodology were based on the work of Borderies et al. (2003) with some modifications. The immersion of the cells in 25% glycerol was tested, but this procedure did not precipitate the cell wall, as proposed in the original protocol, probably because a less dense concentration was required in order to precipitate and isolate the cell wall. Therefore, it was used the concentration of 15% glycerol and the next steps were adapted from Feiz et al. (2006). Thus, after filtration, the cells were washed with Milique water. The cells were lightly ground and then precipitated by immersion in 15% glycerol containing 0.1% (v/v) anti-protease cocktail (P9599, Sigma) for 3 hours at 4 °C. Subsequent extractions were also performed at 4 °C.

The resulting precipitate was then added to an increasing gradient of sucrose by three successive centrifugation (Super T21, Sorvall® Centrifuges) in 500 mL of acetate buffer 5 mM, pH 4.6, plus 0.4 M, 0.6 M or 1 M sucrose, respectively. The pellet was then washed through a 20 µm Nylon net (Millipore) with 500 mL of acetate buffer containing 5mM sodium acetate, pH 4.6, to remove contaminants.

Using 200 mL of solution containing seven-day-old cultured cells as starting material, it was possible to recover a significant amount of proteins extracted from the cell wall. Figure 7 shows some photos of the extraction method used and a brief description of this process.

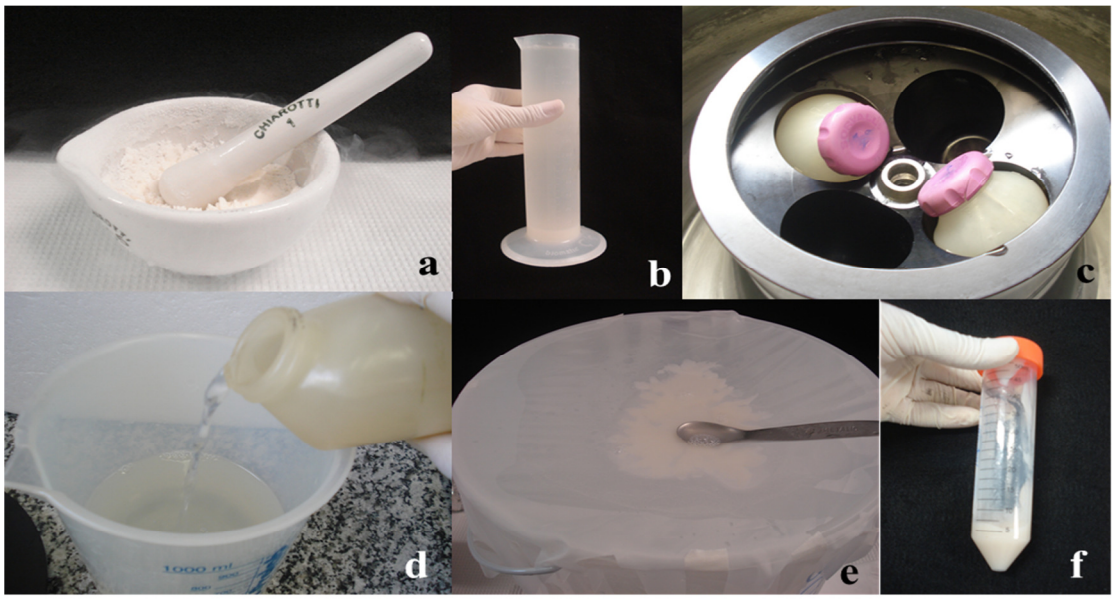

Figure 7 - Method for extraction of cell wall proteins in sugarcane suspension cells. **a.** Cells being slightly macerated in liquid nitrogen; **b.** Precipitation of the ground cells in 15% glycerol for 3 hours; **c.** and **d.** Steps of centrifuging and discarding the supernatant in an increasing gradient of sucrose; **e.** Wash of the pellet on a nylon net (pore size = 20 µm) with 5 mM sodium acetate buffer and **f.** The result of the cell wall isolation. Thereafter, this material was freeze dried for 48 hours, and the extraction of proteins with salt solutions was achieved

After washing, the cell wall was collected and freeze dried for 48 hours in a freeze drier (ModuLyod-230, Thermo Scientific) and the resultant was extracted with the following: (a) 200 mM $CaCl_2$ and (b) 2 M LiCl, both in solutions containing 5 mM sodium acetate, pH 4.6, and 0.1% (v/v) anti-protease cocktail (P9599, Sigma). In each solution, the extract containing the wall was left for 30 minutes and thereafter centrifuged for 20 minutes at 10000 *x* g in a centrifuge (Super T21, Sorvall® Centrifuges). The resulting extracts were desalted on Econo Pac® 10DG column (BIO-RAD). For each column, it was used 100 mM ammonium formate to equilibrate and elute the proteins. After desalted, the samples were freeze dried for about 48 hours to concentrate. The proteins were then solubilized in Milique water, quantified using a commercial kit (Agilent Protein 230) and about 40 µg was applied on a one-dimension gel, SDS-PAGE (LAEMMLI, 1970), as a quality test.

For this experiment three replicates were performed.

This experiment was all conducted in Piracicaba, at the Laboratory Max Feffer of Plant Genetics, ESALQ/USP.

2.2.2.1.2 Isolation and extraction of sugarcane cell wall proteins from leaves and culms

The proteins extractions of this experiment were conducted in Toulouse, France, at the Laboratoire de Recherche en Sciénces Végétales, Université Toulouse III/Paul Sabatier. The sequencing of proteins and analyses were performed in Piracicaba, at the Laboratory Max Feffer of Plant Genetics, ESALQ/USP.

2.2.2.1.2.2 Protocol based on Feiz et al. (2006)

This method is based on the study performed by Feiz et al. (2006), who evaluated the cell walls of Arabidopsis hypocotyls, as well as discussed and compared different techniques for cell wall proteins extraction. This method was used for two-month old plants, both leaves and culms. For sugarcane two month-old leaves about 25g of starting material was collected and about 50g of tissue was used for culms in each repetition.

After collected, the material was washed with Milique water and transferred to a blender containing 500 mL of a sodium acetate buffer 5 mM, pH 4.6, added with 0.4 M sucrose, PVPP (1 g per 10 g of fresh tissue) and 3.3% (v/v) anti-protease cocktail (P9599, Sigma). PVPP helps to trap phenolic compounds and the cocktail inhibits protein degradation during the manipulation (FEIZ et al., 2006). The plant tissue was ground in the blender for 8 minutes at top speed. The cell wall was separated from the soluble cytoplasmic fluid through centrifugation for 15 min, 2500 rpm at 4 °C. The resulting precipitate was then submitted to an increasing gradient of sucrose by two successive centrifugations in 500 mL of sodium acetate buffer 5 mM, pH 4.6, plus 0.6 M and 1 M sucrose, respectively. The final residue was washed with 3 L of 5 mM sodium acetate buffer, pH 4.6, on a nylon net (pore size = 20 µm) (NY20, Millipore). The resulting cell wall fraction was then ground using a mortar and a pestle, with liquid nitrogen, and after that, freeze dried for 48 hours in a freeze drier (ModuLyod-230, Thermo Scientific).

Later, the extraction was conducted by consecutive washes and centrifugations with salt solutions in the following order: two extractions with 6 mL of a solution of $CaCl_2$ (5 mM sodium acetate buffer, pH 4.6, 0.2 M $CaCl_2$ and 3.3% (v/v) anti-protease cocktail (P9599, Sigma) followed by two extractions with 6 mL of a LiCl solution (5 mM sodium acetate buffer, pH 4.6, 2 M LiCl and 3.3% (v/v) anti-protease cocktail (P9599, Sigma). The powder

containing the cell wall fraction and the subsequent pellets were resuspended in 5 mM sodium acetate buffer, pH 4.6, by stirring for 5-10 minutes at room temperature and then centrifuged for 15 min, 40000 *x g* at 4 °C. Supernatants from each wash were recovered, desalted on Econo Pac® 10DG column (BIO-RAD). For each column, it was used 100 mM ammonium formate to equilibrate and elute, according the method described in the manual. After desalted, the samples were freeze dried for about 48 hours to concentrate in a freeze drier (ModuLyod-230, Thermo Scientific). The proteins were then solubilized in Milique water, quantified by the method of Bradford (1976) using a commercial kit (UFP8640 Coo Protein Assay, Uptima). For this method two replicates were performed.

2.2.2.1.2.1 Protocol based on Boudart et al. (2005)

This extraction protocol was adapted from Boudart et al. (2005), who worked with cell wall proteins extracted from *Arabidopsis* rosettes, extracting mostly ionically bound or weakly bound CWPs.

In the present study, the materials used for the extraction were the sugarcane culms. In the extraction of two-month old plants it was used about 50 g for culms as a starting material, and all the internodes were used, from five different plants for both replicates, because young and mature culms couldn't be distinguished.

The tissues were collected, washed with Milique water, and then immersed in a beaker with buffer solution containing 5 mM sodium acetate, pH 4.6, added with 0.3 M mannitol, 0.2 M $CaCl_2$ and 0.1% (v/v) anti-protease cocktail (P9599, Sigma). Mannitol induces plasmolysis, minimizing membranous or cytosolic contamination (CHEN et al., 2009). The beaker was placed in a desiccator attached to a vacuum pump (DOA-P704, Gast) and plant materials were infiltrated under vacuum for ten minutes. The vacuum pump could reach a pressure up to 4.08 bar/80 psi. Thereafter, the infiltrated material was transferred to centrifuge tubes and subjected to centrifugation (200 *x g* for 20 min at 4 °C) in a centrifuge (Centrifuge 5810R, Eppendorf) with swinging buckets. The resulting fluids at the bottom of the tubes were collected with a micropipette. The processes of vacuum infiltration and centrifugation were repeated once and the volumes of fluids were measured. The materials were then infiltrated again and centrifuged, as in the previous step, in a solution containing, instead of calcium chloride, 2 M LiCl.

The resulting extracts were desalted on Econo Pac® 10DG column (BIO-RAD). For each column, it was used 100 mM ammonium formate to equilibrate and elute the proteins. After desalted, the samples were freeze dried for about 48 hours to concentrate in a freeze drier (ModuLyod-230, Thermo Scientific). The proteins were then solubilized in Milique water. The extracts were quantified by the method of Bradford (1976), using a commercial kit (UFP8640 Coo Protein Assay, Uptima) in the case of two-month-old plants.

2.2.2.2 Quantification of proteins

The quantification of proteins for two-month-old plants was accomplished by the method of Bradford (1976) using the max sensitivity protocol, as detailed in the Comassie based commercial kit (UFP8640 Coo Protein Assay, Uptima). The protein concentrations were determined by absorbance in a spectrophotometer (U-3300, Hitachi) and calculated by comparison with a standard curve of a protein with known concentration: BSA, Bovine Serum Albumine, provided by the same kit (UP36859A, Uptima). This measurement was performed to establish the protein concentration after the extraction, since 60 µg of proteins were applied in the gel, as a quality test, and about 50 µg of proteins were prepared to be sequenced.

For cultured cells, the quantification was performed based on a chip-based separation and microfluidic capillary gel electrophoresis using a commercial kit (Agilent Protein 230) that was able to determine the size and quantitation through the migration time and fluorescence intensity measured by the Agilent 2100 BioAnalyzer. The 230 chips were used and the protocol is described in the manual. The advantage of this method is that the chip format dramatically reduces separation time and sample consumption.

2.2.2.3 One-dimensional electrophoresis in polyacrylamide gels

In order to verify the quality of proteins extractions, proteins from two-month-old sugarcane materials and cultured cells were applied on a one-dimension SDS-PAGE, as described in Laemmli (1970) with some modifications. The stacking gel was prepared with a concentration of 4% bis-acrylamide, 3,05 mL of water, 1,25 mL of 0,5 M Tris (pH 6.8), 25 µL of ammonium persulfate and 10 µL of Temed (UltraPure™ Temed, Invitrogen). And the running gel had 12,5% Bis-acrylamide, 3,35 mL of water, 2,5 mL of 1,5 M Tris (pH 8.8), 50 µL of ammonium persulfate and 10 µL of Temed (UltraPure™ Temed, Invitrogen). The one-

dimension gel can separate the proteins based on their size, since the heavier proteins stop the migration earlier in the gel when compared to smaller ones, and can be used as a quality test.

The gels were submitted to an electrical current of 50 volts for 30 minutes and then 70 volts until the proteins had reached the end of the gel. After that, the gels were placed in a solution of Comassie Blue Staining G-250 in order to visualize the bands.

Formerly, the bands from these gels were cut and digested to be sequenced. However, after having access to a new mass spetrometer which could separate and sequence proteins in a liquid extract, the spotting procedure was no longer necessary, and all the analyses were repeated using this new equipment. Besides being faster and more practical to prepare the material and analyze the results, the new equipment has a better resolution and a higher capacity of separation, providing the identification of more proteins with a high efficiency.

2.2.2.4 Image analysis

The image of the gel of two-month old plants extractions and cultured cells, in which we can observe the proteins separated by one-dimensional electrophoresis on polyacrylamide gel and stained with Coomassie Brilliant Blue G-250, was obtained through a scanner (GE-III Image scanner, Healthcare).

2.2.2.5 Samples preparation

The extracts of CWPs were freeze dried and about 50 µg of proteins were ressuspended in 50 mL of 50 mM NH_4HCO_3. After that, 0.2% of RapiGest SF (Waters®) was added to the tubes that were slightly agitated. Then the extracts were placed at 80 °C for 15 min, and after they were submitted to a centrifuge pulse. To increase the access of proteins to alkylation and digestion, about 2.5 µL of 100 mM Dithiothreitol (DTT, BIO-RAD) was added to the tubes that were agitated and placed at 60 °C for 30 min. After cooling until they reached the room temperature, the samples were spinned and 300 mM Iodoacetamide was added to perform the alkylation step. The tubes were transferred to the dark, where they stayed for 30 min (reaction time). After that, 10 µL of Trypsin solution (V511A, Promega) was added to the tubes, which were agitated, and the digestion was performed overnight at 37 °C. In the next day, 10 µL of 5% (v/v) TFA (53102, Pierce) was added, the tubes were agitated and the samples were incubated at 37 °C for 90 min. The tubes were then centrifuged

(18000 g at 6 °C for 30 min, Eppendorf). The supernatants were transferred to vials (Waters Total Recovery Vial, Waters®). Then, the pH was increased adding 5 µL of 1N NH_4OH 25%, so that there was a good association between the samples and the first column. After, the internal standard was added: 2.5 µL de 1*p*mol/µL to the digested aliquot (80 µL), and the vials were agitated. About 17.5 µL of 20 mM Ammonium formate was added to the vials, reaching a final volume of 100 µL.

2.2.2.6 Peptides identification

For each extract, the mixture of peptides was subjected to fractionation by ultra performance liquid chromatography system efficiency in 2D nanoACQUITY UPLC (Waters ®). Separation of the peptides was accomplished using a C18 guard column (50mm x 300µm XBridge C18), 5µm column trap followed by a reverse phase C18 (Symmetry C18, 180µm x 20 mm, 5µm) and finally an analytical column (C18 BEH130, 150µm x 100mm, 1.7 mm). The configuration of these columns allows fragmentation into two dimensions using a binary system of pumps and a pH gradient to elute the peptides gradient started with 3% of B1 (100% acetonitrile) and increased to 65% B in 5 min, followed by the balance with 97% A1 (100% water, 0.1% formic acid). The peptides were then eluted in capture column with a linear gradient which started with 10% B2 (100% Acetonitrile, 0.1% formic acid) increasing to 40% for 26 min, then 85% for 10 min to elute the peptides through an analytical column. The flow rate was 400 nL.min-1 with a total run time of 45 min. The columns were re-equilibrated for the initial condition of 90% A2 (100% water, 0.1% formic acid) and 10% B2 Nano-ESI-MS/MS that was held in the Q-TOF Synapt G2 HMDS (Waters ®) using positive ionization. As a reference to the mass spectrometer, it was used 100 fmol of GFP (human (Glu1)-fibrin-Peptide) at a constant flow of 0.6µL.min-1.

2.2.2.7 Analysis of peptides sequencing

Once sequenced, the files containing the data about the proteins were analyzed by the program ProteinLynxGlobalServer (PLGS) version 2.5v, using the following parameters of tolerance: minimum of consecutive fragment ions (y^+, b^+) for a peptide equal to 2; minimum of consecutive fragments for a protein equal to 5; minimum of peptides (identified by the already described criteria) for a protein equal to 2; and only one lost cleavage allowed. The

program compared sequences already available in the SUCEST (THE SUGARCANE EST PROJECT – SUCEST, 2012) translated proteins database with the sequences obtained in the samples. These sequences will be named either EST or SAS in this text. After inserted the data file, the program presents which proteins show similarity. Thus, we obtained parameters such as the score, the percentage of coverage and the peptides that matched. As the proteins were sequenced in a complex mixture, several proteins had similarity with the accessed file. Only proteins that had a score equal or above seventy and matched two or more peptides were considered. In order to identify the proteins, the sequences of the translated ESTs database identified in the extracts were submitted to a blast analysis with the help of Blast2GO software (CONESA et al., 2005) version PRO. The database chosen to perform the blastp was the *Sorghum bicolor* proteins from Phytozome (PHYTOZOME, 2012), being selected the first five hits to each sequence and the e-value of 1.0E-3 or less, if possible. It was always considered the *Sorghum bicolor* match to each EST or when this first hit was not found, Viridiplantae database from NCBI (NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION –NCBI, 2012) was used instead. Then, several different ESTs matched the same protein, and when this happened, only one hit was considered, in the counting of the total number of proteins, as well as to consider the contaminating proteins. The sequences used to predict the subcellular localization were mainly the ones that were obtained through blastp, since the translated ESTs may not exhibit their full sequence and thus it was not possible to predict the location of the proteins. Only in the cases that the EST showed a different location when comparing to the blasted protein, and this EST contained the full length of the protein, then the EST location was considered instead.

The prediction of subcellular localization of the proteins was performed through the programs: TargetP (EMANUELSSON et al., 2000), Predotar (SMALL et al., 2004), WolfPSORT (HORTON et al., 2007), SignalP (PETERSEN et al., 2011) and TMHMM (PREDICTION OF TRANSMEMBRANE HELICES IN PROTEINS – TMHMM, 2012). Since each program predicts the localization using a different method or a different calculation, when the protein was predicted to possess a signal-peptide targeting it to the cellular secretion, it was regarded as belonging to the cell wall, only in accordance with two or more of these programs. When the protein did not present a signal-peptide prediction it was considered a contaminant, either when it did not have a clear cellular localization, being then considered from undefined location, or when it was predicted to be plastidic or mitochondrial in two or more programs, as an example. The percentage of CWPs as well as contaminating proteins was calculated in order to evaluate the efficiency of the protocols. This proportion

was obtained by dividing the number of CWPs and contaminants found in each of them by the total number of identified proteins.

The proteins considered to be secreted, and then CWPs, were divided into eight functional classes according to the annotation in Interpro (APWEILER et al., 2000) and PFAM (SONNHAMMER; EDDY; DURBIN, 1997) and a previous work (JAMET et al., 2006). The proportion of each functional class was obtained by dividing the number of proteins found in each class by the total number of CWPs.

2.3 Results and discussion

2.3.1 Cultured cells

2.3.1.1 Establishing an extraction protocol

The cultured cells were stained and observed by optical microscopy in order to access viability, as can be seen in Figure 8, together with images of the cells in liquid suspension and solid culture media for maintenance. Figure 8c and 8d shows various cell structures, demonstrating the viability of the suspension cells prior to CWPs extraction.

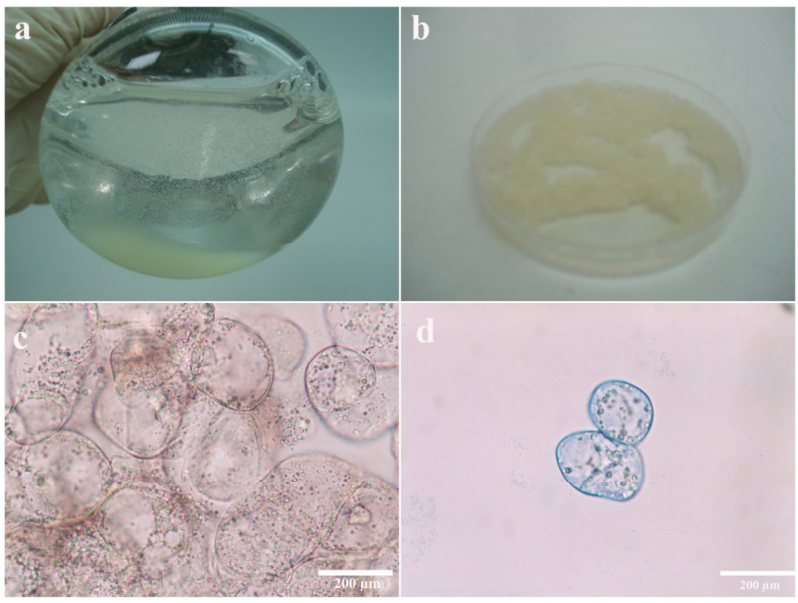

Figure 8 - Cultured cells. a. Culture medium containing the cells; b. Cultured cells in solid medium for maintenance after 30 days of culture; c. Seven-day-old cultured cells observed by an optical microscope with an increase of 40 x without dye; d. A couple of seven-day-old cultured cells observed under a microscope with an increase of 40 x with dye (70% blue austral/30% safranin) (Bar = 200 µm)

To restrain potential contamination by cytoplasmatic proteins and to achieve a significant amount of proteins that could be used in the analysis, different extraction protocols were used and adapted to create a new protocol specifically for cultured cells from sugarcane. Initially, it was used the protocol based on the work of Borderies et al. (2003), however, the immersion of the cells in 25% glycerol caused no precipitation of the sugarcane cell wall, as proposed in the protocol. Therefore, different concentrations of glycerol were tested and the most suitable was the solution containing 15%, which allowed the isolation of the cell walls. A higher percentage was not so efficient to precipitate the cell wall, probably because sugarcane cell walls were less dense than 25% glycerol, as an example, and the cell walls stayed diluted in the solution and could not be precipitated.

No immunological test was used to access the contamination before the proteins extraction. That is because, unfortunately, these methods are much less sensitive than the analysis by mass spectrometry, being the last 10 to 1000 times more sensitive than the use of specific markers (FEIZ et al., 2006), being that the case in the work of Chivasa et al. (2002), who found a very high percentage of contaminants despite the lack of contaminants in other assays. Therefore, the best way to evaluate the quality of a cell preparation is the identification of all proteins by mass spectrometry and an extensive bioinformatic analysis to predict the presence of a signal peptide and no retention signals (FEIZ et al., 2006).

After, centrifugation in a growing sucrose gradient with a low ionic strength buffer (based on the work of Feiz et al. (2006)) allowed the precipitation and isolation of the cell wall, which the cytoplasmic contaminants discarded in the supernatant. A series of sedimentations and centrifugations in a gradient of crescent gradient can act eliminating organelles and other compounds that have less density than cell wall polysaccharides (FEIZ et al., 2006). This buffer was used in order to preserve the ionic bonds while diluting the ionic strength of the cell wall itself, as well as its acidic pH could maintain the interactions between polysaccharides and proteins (FEIZ et al., 2006). The last step of the isolation was suppposed to purify the wall extract in order to wash away possible contaminants that insisted in remaining, on a nylon net.

The protein extraction was then performed in this isolated cell wall extract. The protocol was adapted from Feiz et al. (2006) as described before. So the extraction with calcium chloride could sucessfully liberate several wall-bound proteins while supposedly left the plasma membrane intact, as well as previously studies also demonstrated so (ROBERTSON et al., 1968; CHIVASA et al., 2002; BORDERIES et al., 2003; BOUDART et al., 2005; FEIZ et al., 2006). The efficiency of calcium chloride to release cell wall proteins

can rely on the fact that acidic and neutral carbohydrates strongly chelate calcium (ANGYAL et al., 1989) and then these proteins, that can be weakly bound to the cell wall polyssaccharides, can be solubilized by $CaCl_2$ through competition (JAMET et al., 2007). On the other hand, the lithium chloride was used in order to extract mostly hydroproline-rich glycoproteins (BOUDART et al., 2005). These two salts were used based on several studies that indicated that they were the ones that recovered most of the proteins, moreover, the other substances showed a lot of contaminants or did not extract many proteins or different ones from what was found in calcium and lithium extracts. The use of detergents can extract proteins stronlgy linked to cell wall carbohydrates (HE; FUJIKI; KOHORN, 1996), however, this approach was not used in this study because the same substance can extract mostly intracellular proteins, increasing the percentage of contaminants, and should be avoided in characterization studies (FEIZ et al., 2006).

Using 200 mL of solution containing seven-day-old cultured cells as starting material, it was possible to recover a significant amount of proteins extracted from the cell wall. The efficiency of the extraction can be observed in Figure 9, on the SDS-PAGE analysis.

Figure 9 shows the patterns of proteins obtained after electrophoresis and Comassie Brilhant Blue G-250 staining of the three repetitions. The SDS-PAGE (LAEMMLI, 1970) shows that this method was efficient to extract proteins and can be repeatable. However, only through the final steps of sequencing and bioinformatic analysis is possible to characterize them and establish the percentage of contaminants, as well as the real efficiency of CWPs extraction.

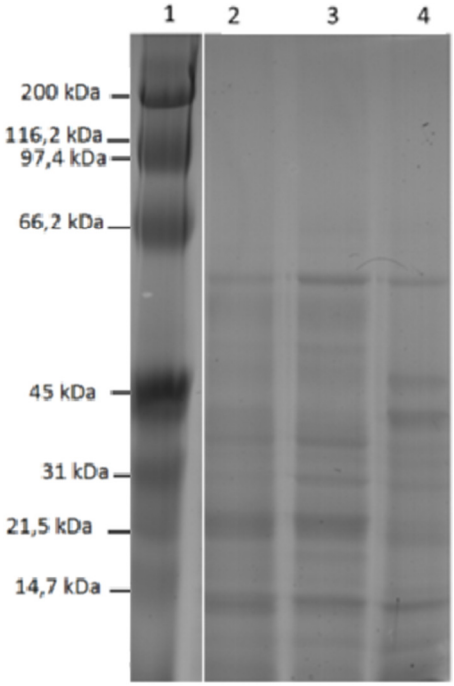

Figure 9 - SDS-PAGE analysis of CWPs from sugarcane cultured cells. (1) molecular mass marker; (2), (3) and (4) 40 µg of proteins loaded in the gel in the three repetitions: each column represents a different repetition.

2.3.1.1.2 Sequencing and analysis

In the extraction procedure using cultured cells, we were able to identify 303 different ESTs from sugarcane ESTs database (THE SUGARCANE EST PROJECT – SUCEST, 2012), when considering only the ones which were found in the three repetitions. From these ESTs, 142 different proteins were identified through blastp, since some different ESTs matched the same proteins.

Figure 10 shows a Venn diagram with the number of ESTs identified in each repetition and the ones found in all of three. It is able to see that the extractions are not very repeatable.

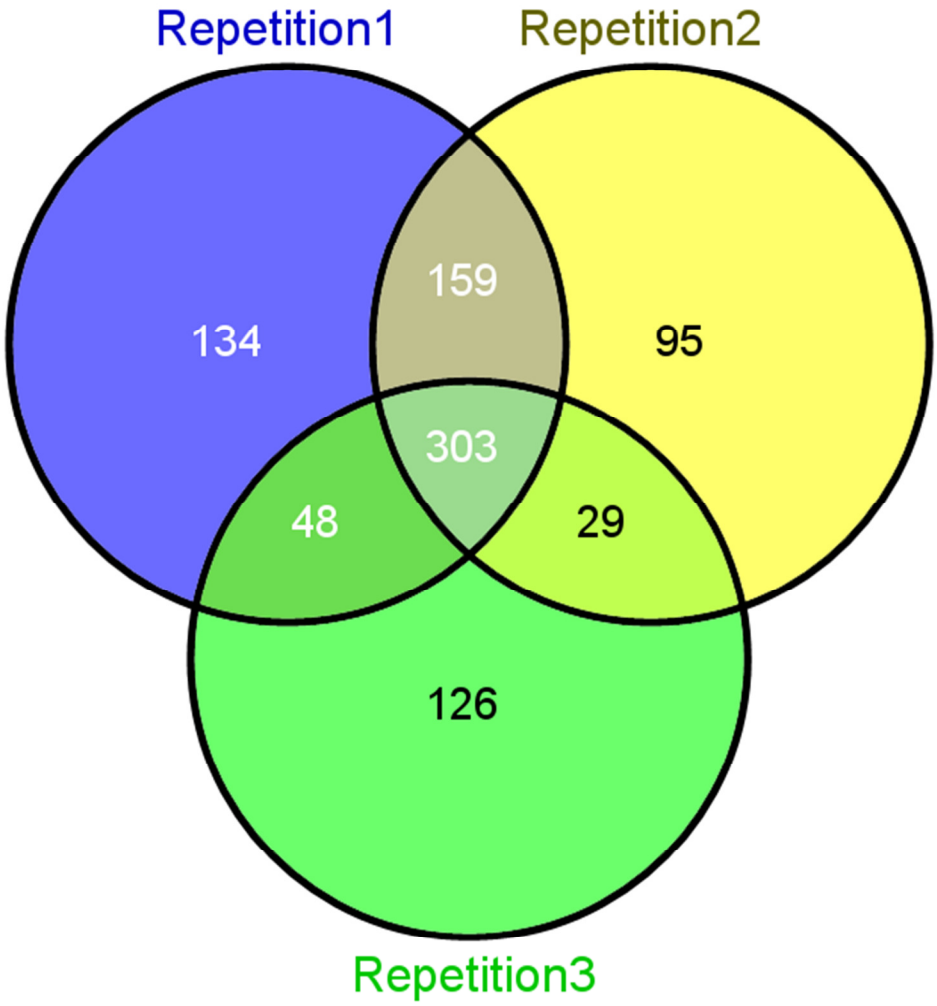

Figure 10 - Venn diagram (OLIVEROS, 2007) showing the number of ESTs in the three repetitions of extractions performed, in cultured cells.

Table 1 shows the identified ESTs/SAS and proteins considered to be secreted showing parameters as hit species, functional classification, PFAM and InterPro families. Also, in this table is possible to see the multiple ESTs and the correspondent proteins that matched them through blast.

Table 1 – Identified ESTs and proteins in sugarcane cultured cells showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCCCCL4008H04	0.0	Sb01g016600.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCCCCL5071F02	1,66E-101						
SCEQRT1028H09	0.0	Sb06g033840.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCSFCL6068C11	6,80E-102						
SCCCCL7037A10	0.0	Sb02g027330.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCCCLB1002D05	0.0	Sb10g028500.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCCCLB1004B09	0.0	Sb10g027490.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCEQRT2030A04	5,31E-103						

Table 1 – Identified ESTs and proteins in sugarcane cultured cells showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCCCLR1C05G08	0.0	Sb03g024460.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCEPRZ1011A06	0.0	Sb03g010250.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCEQRT1026F09	0.0	Sb03g013210.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCSGRT2063H01	0.0	Sb06g033870.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCACSB1038A04	2,68E-93	Sb03g011590.1	*Sorghum bicolor*	Oxido-reductases	Germin	Cupin_1 (PF00190)	IPR006045
SCBGRT1047A04	3,50E-73						
SCEZHR1049D07	1,36E-94						
SCMCRT2089H04	9,81E-122						
SCVPRZ2036E02	1,39E-130						

Table 1 – Identified ESTs and proteins in sugarcane cultured cells showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCCCLR1C01A02	0.0	Sb09g004370.1	*Sorghum bicolor*	Miscellaneous	Isomerase	Thioredoxin (PF00085)	IPR013766
						Thioredoxin (PF00085)	IPR013766
						ERp29 (PF07749)	IPR011679
SCEQRT1026G10	2,64E-07	Sb08g016690.2	*Sorghum bicolor*	Miscellaneous	Metallophosphoesterase	Metallophos (PF00149)	-
						Metallophos_C (PF14008)	-
SCVPRZ2043H01	0.0	Sb08g016690. 1	*Sorghum bicolor*	Miscellaneous	Metallophosphoesterase	Metallophos (PF00149)	IPR004843
						Metallophos_C (PF14008)	-

Table 1 – Identified ESTs and proteins in sugarcane cultured cells showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCEZRZ1014C04	2,77E-139	Sb03g039330.1	*Sorghum bicolor*	Miscellaneous	Thaumatin-like protein	Thaumatin (PF00314)	IPR001938
SCJLLR1054F05	0.0	Sb04g001140.1	*Sorghum bicolor*	Miscellaneous	Heat shock protein	HSP70 (PF00012)	IPR013126
SCSFHR1046G01	3,14E-80	Sb02g029060.1	*Sorghum bicolor*	Miscellaneous	Isomerase	FKBP_C (PF00254)	IPR001179
SCCCRZ3002G10 SCEZRT2018F03	1,76E-48 1,82E-52	Sb01g031470.1	*Sorghum bicolor*	Proteins with interacting domains (with proteins or polysaccharides)	Protease inhibitor	Cystatin (PF00031)	IPR000010
SCEPLR1051C09	9,85E-136	Sb03g010730.1	*Sorghum bicolor*	Proteins with interacting domains (with proteins or polysaccharides)	Protease inhibitor	Cystatin (PF00031) Cystatin (PF00031)	IPR000010 IPR000010

Table 1 – Identified ESTs and proteins in sugarcane cultured cells showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCEQRT1025E03	6,99E-59	Sb03g037370.2	*Sorghum bicolor*	Proteins with interacting domains (with proteins or polysaccharides)	Cystatin	Cystatin (PF00031)	IPR000010
SCMCLV1031F09	9,00E-41	Sb04g021920.1	*Sorghum bicolor*	Proteins with interacting domains (with proteins or polysaccharides)	PMEI	PMEI (PF04043)	IPR006501
SCVPRZ2040D09	6,71E-108						
SCCCCL4003F09	0.0	Sb02g004660.1	*Sorghum bicolor*	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_18 (PF00704)	IPR001223
SCJFRT1062D12.b	7,02E-69						
SCRLLV1024C09.b	4,13E-92						

Table 1 – Identified ESTs and proteins in sugarcane cultured cells showing parameters of the identification together with the functional classification

(concludes)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCEQLR1093F09	0.0	Sb01g008040.3	*Sorghum bicolor*	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_3 (PF00933)	IPR001764
SCQSLR1089A04	1,93E-152					Glyco_hydro_3_C (PF01915)	IPR002772
SCVPRT2073A02	4,33E-86						
SCCCLR1022B11	0.0	Sb06g030800.1	*Sorghum bicolor*	Proteases	Cys proteases	Inhibitor_I29 (PF08246)	IPR013201
						Peptidase_C1 (PF00112)	IPR000668
						Granulin (PF00396)	IPR000118
SCACLR2029H09	1,72E-40	Sb01g033830.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	-
SCCCCL3001E03.b	9,87E-40						
SCCCCL4009G04	2,76E-143	Sb01g004270.1	*Sorghum bicolor*	Unknown	-	DUF642 (PF04862)	IPR006946
SCSGLR1084A12	1,76E-109						

Therefore, from 142 different proteins found, 40 ESTs corresponding to 25 proteins had a signal-peptide prediction in two or more prediction programs, being considered as secreted proteins and thus CWPs, shown in Table 1 together with the functional classification, based on Interpro (APWEILER et al., 2000) and PFAM (SONNHAMMER; EDDY; DURBIN, 1997).

When comparing the total number of proteins identified, 142, it is possible to see that only 17,61% had a prediction for secretion, being considered CWPs, as shown in Figure 11. Moreover, about 82,39% of the total number of proteins were considered contaminants, being intracellular proteins, directed to other organelles or considered from undefined location.

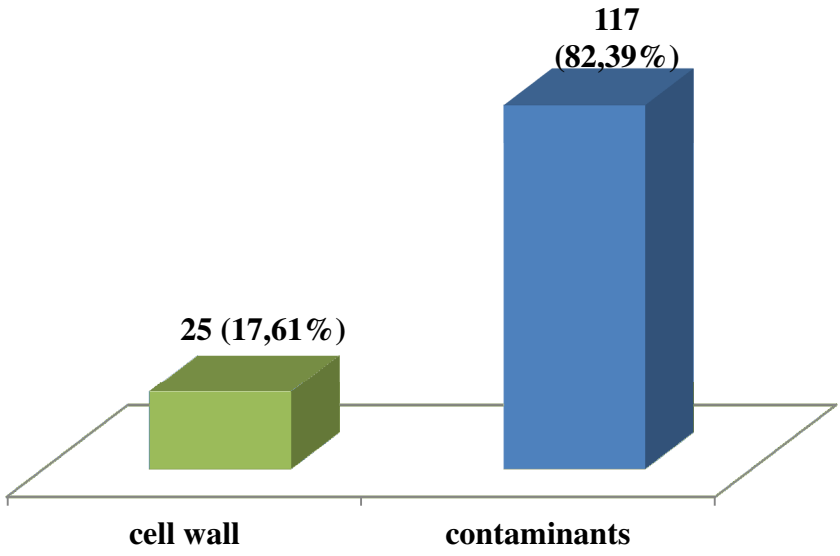

Figure 11 – Number of identified proteins that were predicted to be secreted and the contaminants, in sugarcane cultured cells.

Using cultured cells as a material, it was able to identify a good number of proteins, even because it was used a translated EST database, and probably if we had the whole genome of sugarcane already available this number could be increased. As an example, in rice calli, which is a species that has its genome fully sequenced, it was possible to identify 292 proteins using the MudPIT technology (CHEN et al., 2008). But this is not so reasonable, since in a study that aimed to characterize *Brassica oleracea* xylem sap proteome, they also could identify a good number of proteins, 189, achieving a very low percentage of contamination, about 13%, even the species not having its genome fully sequenced (LIGAT et al., 2011). However, they considered ESTs macthing to same proteins as different proteins, unlike our study. This possibility has to be considered, since different ESTs that correspond to the same protein can actually be different proteins because sugarcane is a hexaploid/octaploid

genome, but this is remote chance since these ESTs are grouped into clusters. However, comparing to several other studies (CHIVASA et al., 2002; ZHU et al., 2006; SLABAS et al., 2004), the number of CWPs was equal or lower when comparing to the present study. In the study of Bayer et al. (2006), only 15% of the proteins had a signal-peptide, and maybe the MudPIT technology used could be responsible to resolve a larger number of proteins, such as the contaminant ones.

The presence of these contaminant proteins can have several explanations. These proteins can be actual contamination, since this method is based on grinding the tissue, and this step can release intracellular proteins that can be trapped in the cell wall matrix and somehow can't be washed away. Given the acidic nature of the extracellular matrix, the cytoplasmic proteins can bind to walls through ionic interactions (BAYER et al., 2006) and then could be carried with the wall. Moreover, maybe the number of released intracellular proteins increased along several extractions, as proposed by Borderies et al. (2003). Also, when observing the results in two-month-old plants, in which only two repetitions were used, the number of contaminants can decrease since the contaminants usually are the most present and each time the number of replicates is increased the number of contaminants it is as well. Thus, it is suggested that experiments working with cell wall proteins use only two replicates, even because most of them are purely descriptive. As an example, it is possible to observe that when considering the proteins found only in two replicates, in cultured cells, the percentage of cell wall proteins almost double, being 21,43%, accounting for 45 ESTs and 36 identified proteins (Supplementary data – APPENDIX A).

In the model proposed by Carpita and Gibeaut (1993), the microfibrils in the cell wall of grasses are interlocked by glucuronoarabionoxylans, instead of xyloglucans, the latter found in dicotyledons. In contrast to xyloglucans, the xylans are substituted with arabinosyl units that block hydrogen bonding, being able to bind on either face to cellulose or to each other. Still, unlike walls from most plants, in grasses, a substantial portion of the non-cellulosic polymers are connected to the microfibrils by alkali-resistant phenolic linkages (CARPITA; GIBEAUT, 1993). Perhaps the unique characteristics of grasses cell wall can increase the trapping of these proteins, since polysaccharide networks of cellulose, hemicellulose and pectins form potential traps (FEIZ et al., 2006).

Furhtermore, these proteins considered to be contaminants can have high identity to known intracellular proteins but with distinct functions, or they may have multiple-targeting, displaying more than one function, localized in multiple cellular compartments (ROSE; LEE, 2010).

Another possibility, is the existence of an alternative secretory pathway, independent from the ER-Golgi mechanism in a way that secreted proteins may not possess the signal-peptide (ROSE; LEE, 2010). Examples can be seen in the proteins GAPDH, enolase, peroxiredoxin, and elongation fator 1 alpha, which are known to be secreted in spite of the absence of a signal peptide, in several species (CHIVASA et al., 2002). This unusual way of secretion has already been proved to exist in mammals (COOPER; BARONDES, 1990; RUBARTELLI et al., 1990), and may be related to stress conditions. Although this hypothesis has been discussed by several authors, it has not been proved to exist in plants yet and more experiments such as in planta localisation of the proteins using antibodies or fluorescent fusion proteins should be performed in order to elucidate this mechanism. Thus, in this study, it is considered CWPs only the ones that present a signal-peptide in two or more prediction programs, as already stated.

The percentage of contaminants was much higher than the cell wall proteins. In one of the original protocols, described by Feiz et al. (2006), when working with *Arabidopsis* hypocotyls, about 73% of the proteins were considered to be secreted, showing one of the lowest percentage of contamination, as well as Chen et al. (2009), who described almost the same percentage in rice calli. Boudart et al. (2003), showed that when extracting CWPs from *Arabidopsis* cultured cells, even when some steps of the extraction damaged the plasma membrane, it was possible to identify 50 CWPs, presenting about 35-40% of contaminants. However, despite some studies describing lower percentage of contamination, the total number of identified CWPs was small. When working with roots from maize, that is a monocotyledon like sugarcane, Zhu et al. (2006), even finding that about 40% and 16% of the proteins predicted to have a signal peptide in the N-terminal sequence, in one fraction and total soluble proteins, respectively, they only identified 22 CWPs, at the total. Even working with a different method to perform the extraction, in the present study it was possible to identify more CWPs, which means that this method was sucessful to extract proteins from the cell wall of sugarcane cells, even though together with the CWPs, a lot of intracellular and contaminants were also extracted. Working with a similar material, Chivasa et al. (2002) also found only 24 proteins predicted to be secreted in cultured cells from *Arabidopsis*. Comparing a similar method of extraction, also 25 CWPs were identified in *Medicago sativa* (WATSON et al., 2004), however, the percentage of contamination in one of the extracts was much lower, 50%. Nevertheless, the present study showed that even having a higher percentage of contaminants, more CWPs were identified, and only these proteins were considered in the analysis.

Regarding the contaminants, different histones, mostly h1, h2b and h4, were found, together with many ubiquitins, 40 S and 60 S ribossomal proteins, malate dehydrogenase, methyonine synthase, heat shock protein, calmodulin, chaperonin, glyceraldehyde 3-phosphate dehydrogenase, splicing factors, elongation factor 1 alpha, enolase, fructose-bisphosphate aldolase, luminal binding proteins, and some hypothetical and unknown proteins, among others (data not shown). Some of them were well-known plastidic, mitochondrial or nuclear proteins, and a great number was considered from undefined location. The great number of histones or mitochondrial and chloroplastidial proteins can be explained since they can be trapped with the membranes that circumvent these organelles and be mistakenly extracted. If considering the existence of an alternative secretory pathway that does not involve a signal-peptide, the number of CWPs could be higher. As an example, the elongation factor 1 alpha is a GTP-binding subunit of the translation elongation complex that can act as an actin-binding protein, which was immunolocalised in tobacco cell walls where it may function as a cell adhesion fator (ZHU et al., 1994). Still, this protein has a sequence almost identical to vitronectin-like adhesion protein, PVN1, which can also be an explanation to this finding (EJIRI, 2002). Another example is Histone 1, that was implicated with receptor activity, found on the cell surface (BRIX et al., 1998). Also, GAPDH was immunolocalised in *Candida albicans* cell wall, being involved in cell adhesion (GOZALBO et al., 1998). Luminal binding proteins and enolase were also found in CWPs extract from *Arabidopsis* cells, suggesting that it can be targeted to the wall (SLABAS et al., 2004), being the last one immunolocalised in *Saccharomyces cerevisiae* cell wall as well (EDWARDS; BRALEY, CHAFFIN, 1999). Even malate dehydrogenase was reported to be exported to the extracellular compartment *via* a nonclassical pathway (CHIVASA et al., 2002) and could be involved regenerating NAD(P)H needed by cell wall peroxidases for free radical generation associated with lignin polymerization. Besides, the existence of of alternate gene splicing, that can be tissue-specific, can explain the presence of these contaminants, because an alternate copy could contain a signal-peptide (CHIVASA et al., 2002). Another possibility is that such proteins originate from different cells that are released after cell death in the medium and could be identified together the healthy cells.

As shown in Figure 12, proteins predicted to be secreted were divided into seven functional classes. The biggest proportion was classified as oxido-reductases contributing to 40%. The second more abundant class is composed of miscellaneous proteins, 24%, represented by several proteins. It was also found proteins with interacting domains (with proteins or polysaccharides), representing 16%. About 8% were considered proteins acting on

cell wall polysaccharides. Proteins related to lipid metabolism accounted for 4% from the total, as well as proteases, unknown and membrane proteins. It was not found any signaling protein in this experiment.

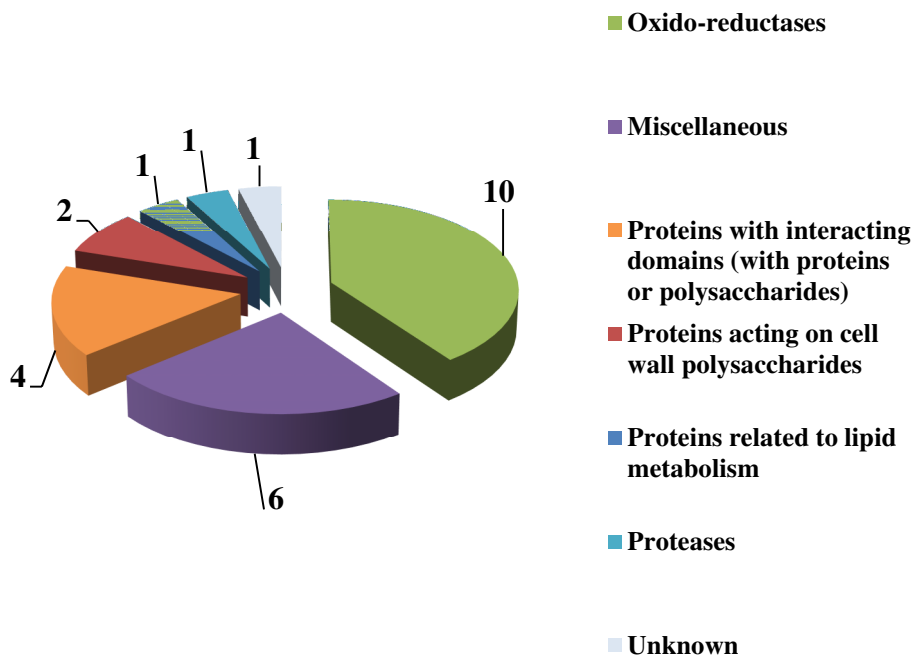

Figure 12 – Number of cell wall proteins identified and their functional classification, in cultured cells.

The most representative class, the oxido-reductases, was divided into nine peroxidases and one germin-like protein, as can be seen in Table 1. Peroxidases are well-known cell wall enzymes, identified in several studies (SLABAS et al., 2004; BAYER et al., 2006; KONG et al., 2010, BORDERIES et al., 2003), mainly plant peroxidase class III, involved with cell defense (ALMAGRO et al., 2008). Plant peroxidases are mainly localised in the extracellular matrix and vacuole and in growing plant cell walls, they are involved in the oxidative cross-linking of cell wall polymers (BAYER et al., 2006). The peroxidases found in this extraction are members of the family PF00141, according to PFAM (PFAM DATABASE - PFAM, 2012). Their activity is very versatile, during the hydroxylic cycle, peroxidases can produce reactive oxygen species that can break cell wall polysaccharides in a non-enzymatic way, favoring cell wall extension, and during the peroxidative cycle, peroxidases can also promote cross-linking of the cell wall components such as structural proteins or lignins, as well as play house-keeping roles (IRSHAD et al., 2008). The high number of peroxidases can reflect the vulnerability of the cultured cells, since their habitat is a liquid medium and they can produce this kind of proteins to deffend themselves against possible attacks and pathogens that can be

present even if the flask was previously sterilized. Germin-like proteins were also commonly found in cell wall extracts (BORDERIES et al., 2003; BOUDART et al., 2005; BAYER et al., 2006; KONG et al., 2010), being related to the defense mechanism and plant development (SEGARRA et al., 2003; BERNIER; BERNA, 2001) and appear to be a superfamily functionally diverse of proteins. The cupin family has a conserved β-barrel fold and can have eighteen different functions, from bacterial enzymes involved in the modification of cell wall carbohydrates, to two-domain bicupins such as the desiccation-tolerant seed storage globulins (DUNWELL et al., 2001). So, it can be even from a bacterial contamination in the medium as well as displaying one of its many functions in sugarcane cells. The oxido-reductases can be interesting when studying defense response against pathogens and cell growth mechanisms.

In the miscellaneous class, an isomerase member of the thioredoxin family was found. Thioredoxins can be involved in several processes, such as photosynthesis, growth, flowering, development and germination of seeds. But, also, it has been found to play a role in cell-to-cell communication (MENG et al., 2010). This kind of protein can be found in other cell wall protein studies (BORDERIES et al., 2003). A protein disulphide isomerase can be secreted in animal cells by an unkown mechanism even when this protein has the ER retention motif (XIAO et al., 1999). Probably, the role in cultured cells involves the growth of the cells that is continuously happening as well as cell to cell communication which has to be very intense since there are thousands and/or millions of cells spreaded in the medium and they need to communicate themselves through signals, which are mainly proteins. Fk506-binding protein is a protein folding chaperone for proteins containing proline residues and can be found from yeast to humans (BALBACH; SCHMID, 2000). Thaumatin-like proteins were also found in other studies, such as in Kong et al. (2010). They appear to be pathogenesis-related proteins that can be activated by pathogens or reagents, being involved in systematically acquired resistance and stress response, although their precise role is unknown (HERRERA-ESTRELLA et al., 1992). Probably, in cultured cells it is activated by the salts present in the medium, being related to osmotic regulation. Despite the surprise, a heat shock protein, a luminal binding protein, was found in this class. The Hsp70 plays many roles, but in cells probably it aids in transmembrane transport of proteins, stabilizing them in a folded state and still can act to protect cells from thermal or oxidative stress (WEGELE; MULLER; BUCHNER, 2004). A luminal binding protein is a classical chaperone located in the ER lumen, however, in *Arabidopsis* cells it was also found to be secreted since it also showed a signal-peptide prediction (SLABAS et al., 2004). Other proteins found in miscellaneous class were the phosphoesterases from Metallophos family. It is suggested that this kind of protein

function as acid phosphatases, which role is to hydrolyze several substrates but its specific function is still unknown (SZALEWICZ et al., 2003).

The proteins with interacting domains found in this experiment, are formed mostly by protease inhibitors such as cystatin and PMEI. Cystatin, which is a cystein proteinase, was also found in several studies (BOUDART et al., 2005; CHEN et al., 2009; BORDERIES et al., 2003; LIGAT et al., 2011), and shows a anti-herbivorous activity (SIQUEIRA-JÚNIOR et al., 2002). On the other hand, plant invertase/PME Inhibitor proteins inhibits pectin methylesterases and invertases and it has been implicated in the fruit development, carbohydrate metabolism and cell wall extension and also may be involved in inhibiting microbial pathogens (CAMARDELLA et al., 2000). In cells, it can inhibit the sugar metabolism enzymes, and be involved in various steps of plant development, including the inhibition of microrganisms that can threaten the cells in the liquid medium. If the goal is to look for proteins related to cell extension and growth, this protein is a good example to work with.

Curiously, only two glycoside hydrolases were found, in contrast with some works that showed that the cell wall modifying enzymes was the most represented group for CWPs (BAYER et al., 2006; BOUDART et al., 2005). The role of such proteins in cell walls points out to the rearrangements of polysaccharides during development (CASSAB; VARNER, 1988; FRESHOUR et al., 1996; NICOL; HOFTE, 1998; ROSE; BENNETT, 1999). These proteins can show the importance of polysaccharide metabolism within the extracellular matrix (BOUDART et al., 2005), especially in young tissues. In sugarcane cultured cells it can be linked to the rearrangement of the wall carbohydrates, leaving a more easily degradable wall making possible the cells to grow. However, because in these cells the walls are much less developed in comparison with adult plants, maybe the lower content of wall modifying enzymes can point to its small complexity of carbohydrates.

Only one protein related to lipid metabolism was listed, a lipid transfer protein. This protein is relatively common in cell wall protein studies (CHEN et al., 2009; BOUDART et al., 2005), and can act on binding lipids to their hydrophobic cavity, which can be essential for the cell wall–loosening activity, facilitating wall extension (NIEUWLAND et al., 2005). Maybe the small content of this kind of protein can be related to the physiological priorities of the cells, such as the basal metabolism.

One cysteine protease represents the Proteases class. Cell wall proteases also play a role in cell defense against pathogen attacks (BAYER et al., 2006), and also play roles in maturation of enzymes, signaling and protein turnover (ALBENNE et al., 2009), which makes

sense since the cells have to be prepared to respond to any attack since they are very vulnerable.

One protein having a domain of an unknown function, member of the family DUF642, appeared in the sugarcane cells, however, it is a member of a clan, the galactose-binding domain-like superfamily that is linked in carbohydrate recognition and probably the protein can have a similar function (PFAM database, 2012).

Interestingly, one gibberellin-regulated protein (XP_002438216) was found, belonging to the GASA family, but only in the proteins identified in a single repetition (APPENDIX B). This kind of protein was also identified in rice calli (CHEN et al., 2009), but not in *Arabidopsis* CWPs extractions, and the expression of this protein is up-regulated by gibberellin, showing some correlation with the plant development (AUBERT et al., 1998).

Also, in a single experiment, this study identified an alfa glucosidase (XP_002438844), the same kind of protein found in rice (CHEN et al., 2009). Alpha-glucosidase catalyzes the first step in N-glycan trimming and might participate in cellulose biosynthesis (CHEN et al., 2009).

Despite the large number of proteins not predicted to be secreted, this first experiment could identify 25 different cell wall proteins related to the sugarcane cultured cells. It is reasonable that a poorly differentiated material, such as cultured cells, could present a small number of identified proteins, which certainly is not the case, since it was possible to list 142 proteins present in the three replicates. However, just a few of them were considered CWPs, according to already described criteria. This shows that maybe the method, involving tissue grinding and $CaCl_2$, a salt that makes the membrane more permeable, being used in genetic transformation (JAMET et al., 2008), can increase the contamination, and still has to be adapted in order to generate lower contamination. But it was possible to find a good number of CWPs, even higher in comparison to other studies, and probably, the cell walls that are present in cultured cells are much less developed than in whole plants. This could also be a plausible explanation to the relative low number of CWPs in this extract, since the cell wall would be a much smaller portion of the cell. Logically, it is not possible to extract all the CWPs, especially structural ones, that very rarely can be extracted in CWPs studies (BORDERIES et al., 2003). The proteins found in this study are in agreement with the expected cell wall roles. Therefore, this experiment could generate information about the CWPs from sugarcane, characterizing some of them and pointing to some proteins that can be a target to genetic manipulation, especially those linked to carbohydrate modification, cell expansion and cell growth.

2.3.2 Extraction of CWPs from two-month-old sugarcane leaves and culms

Plants were submitted to different methods of extraction, as described previously. The protocol based on Feiz et al. (2006), that will be named method 1 and the one based on Boudart et al. (2005), that it will be named method 2.

Figure 13 shows the SDS-PAGE of sugarcane CWPs from leaves and culms, in both methods. The extractions recovered proteins and the gel bands have a good resolution. Moreover, even in this dimension, it is possible to observe some differentiated group of proteins in each tissue or method of extraction.

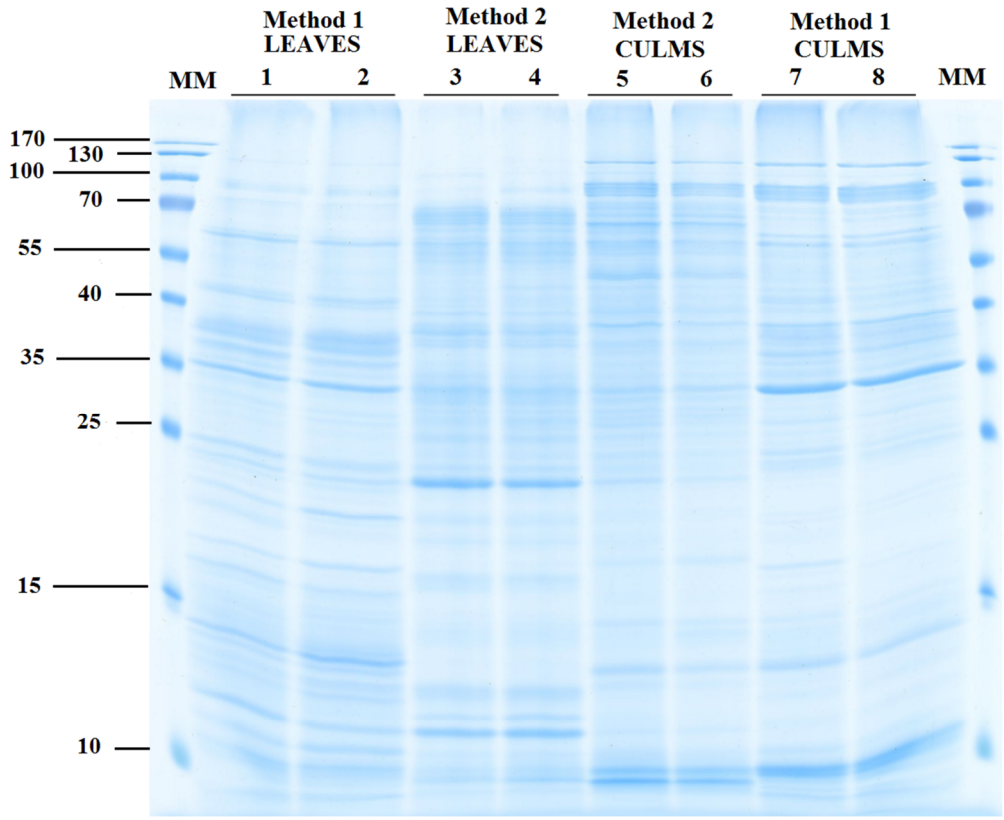

Figure 13 - SDS-PAGE of CWPs from 2-month-old sugarcane leaves and culms in methods 1 and 2. The MM columns shows the molecular mass marker, in kDa; (1) and (2) Both repetitions from leaves extraction using method 1; (3) and (4) Both repetitions from leaves infiltration, method 2; (5) and (6) Both repetitions from culms infiltration, method 2; (7) and (8) Both repetitions from culms extraction using method 1. In each channel 60 µg of proteins' extract were loaded

2.3.2.1 Sequencing and analysis

2.3.2.1.1 Culms from method 1 (based on Feiz et al. (2006))

In the extraction procedure (Method 1), using two-month old culms, it was possible to identify 220 different ESTs from sugarcane ESTs database (THE SUGARCANE EST PROJECT – SUCEST, 2012) when considering only the ones which were found in both repetitions. From these ESTs, we could identify 77 different proteins through blastp, since some different ESTs matched the same proteins.

Figure 14 shows a Venn diagram with the number of ESTs identified in each repetition and the ones found in both of them.

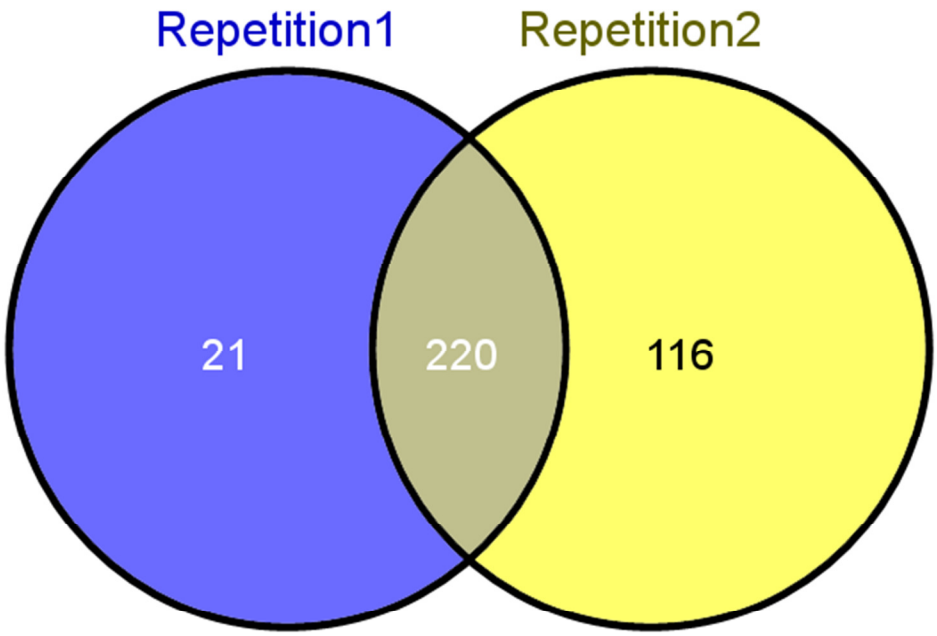

Figure 14 - Venn diagram showing the number of ESTs in the two repetitions of extractions performed, in culms extraction from method 1.

In Table 2, it is possible to see the identified ESTs and proteins predicted to be secreted showing parameters as hit species together with the functional classification, based on Interpro (APWEILER et al., 2000) and PFAM (SONNHAMMER; EDDY; DURBIN, 1997).

Table 2 – Identified ESTs and proteins in sugarcane culms, method 1, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCACSB1035A05 SCCCRZ1002B03	1,78E-127 0,00E+00	Sb01g041770.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCBFAD1091A01 SCCCRT1001G12	1,23E-72 0	Sb04g008590.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCCCLB1004B09 SCEQRT2030A04	0 5.36E-103	Sb10g027490.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases (Haem)	Peroxidase (PF00141)	IPR002016
SCCCLR1C03A09	0	Sb09g004650.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCCCLR1C05G08	0	Sb03g024460.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCEPRZ1011A06	0	Sb03g010250.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016

Table 2 – Identified ESTs and proteins in sugarcane culms, method 1, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCJFRT1061G12	2,76E-88	Sb10g021650.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCRURT3064C05	1,67E-93	Sb06g016610.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCVPLB1020D03	3,85E-132	Sb03g046760.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCJFRT1010F10	1,61E-70						
SCRLAD1042E05	8.18E-73	Sb09g002740.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCVPRZ2035F03	0						
SCBFFL4112F05	6,79E-61	Sb06g018350.1	*Sorghum bicolor*	Oxido-reductases	Blue copper binding proteins	Cu_bind_like (PF02298)	IPR003245

Table 2 – Identified ESTs and proteins in sugarcane culms, method 1, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCRFHR1006G03	3,69E-67	Sb01g010510.1	*Sorghum bicolor*	Oxido-reductases	Blue copper binding proteins	Cu_bind_like (PF02298)	IPR003245
SCRLST3166D09	3,87E-98	Sb03g005150.1	*Sorghum bicolor*	Oxido-reductases	Germin-like	Cupin_1 (PF00190)	IPR006045
SCSFAD1115E03	3,92E-113						
SCUTLR1037C05	3.19E-121						
SCBFLR1046E09	6,21E-48	Sb03g038280.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	-
SCCCAM2002F12	1.95E-53						
SCVPRZ2039B03	5.50E-53						
SCVPRZ2041C11	5.50E-53						

Table 2 – Identified ESTs and proteins in sugarcane culms, method 1, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCBGLR1114E07	2,78E-35	Sb08g002690.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	Tryp_alpha_amyl (PF00234)	IPR003612
SCCCCL3004H07.b	4,55E-19						
SCCCLR1048F06	2.13E-42						
SCCCLR1C08B02	5.19E-36						
SCUTST3131G03	2.14E-39						
SCVPHR1092G06	3,98E-17						
SCCCLR1024C05	2,46E-50	Sb08g002660.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	Tryp_alpha_amyl (PF00234)	IPR003612
SCCCLR1076D05	2,86E-51						
SCEPLB1044H11	2,13E-50						
SCCCLR1072C06	1,00E-39	Sb08g002700.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	-
SCRFLR1012A10	1.00E-39						

Table 2 – Identified ESTs and proteins in sugarcane culms, method 1, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCCCLR2C03F01	1,07E-40	Sb08g002670.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	Tryp_alpha_amyl (PF00234)	IPR003612
SCEZLB1006F09	4.07E-34						
SCEPLB1044H04	7,90E-29	Sb01g049830.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	-
SCBGLR1023G11	1,40E-93	Sb04g029670.1	*Sorghum bicolor*	Proteases	Asp proteases	Asp (PF00026)	IPR001461
SCBGLR1097G03	9,98E-71						
SCMCLR1123H12	3,32E-99	Sb05g027510.1	*Sorghum bicolor*	Proteases	Asp proteases	Asp (PF00026)	IPR001461
SCQGST1032H01	1,20E-139						
SCEZLB1013B06	2,88E-136	Sb10g001440.1	*Sorghum bicolor*	Proteases	Asp proteases	Asp (PF00026)	IPR001461

Table 2 – Identified ESTs and proteins in sugarcane culms, method 1, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCVPLR2012E01	4,89E-85	Sb01g044790.1	*Sorghum bicolor*	Proteases	Asp proteases	Asp (PF00026)	IPR001461
SCVPRZ2038B09	4,89E-85					Asp (PF00026)	IPR001461
SCAGFL1089C05	1,49E-45	Sb08g018710.1	*Sorghum bicolor*	Miscellaneous	Aminoacid Biosynthesis DAHP synthase	DAHP_synth_1 (PF00793)	IPR006218
SCCCCL3140H01	7.17E-88						
SCRUFL4024B08.b	3.91E-36						
SCSFST1066G10	2.32E-90						
SCCCCL3005A02.b	2,71E-66	Sb03g039330.1	*Sorghum bicolor*	Miscellaneous	Thaumatin-like protein	Thaumatin (PF00314)	IPR001938
SCEZRZ1014C04	2,78E-139						
SCQSLB1049A01	2.19E-30						
SCCCRZ1C01H06	0	Sb08g001950.1	*Sorghum bicolor*	Miscellaneous	Nucleoside Phosphatase	GDA1_CD39 (PF01150)	IPR000407
SCJFRZ2029D05	7.96E-161						

Table 2 – Identified ESTs and proteins in sugarcane culms, method 1, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCCCRZ2003A02	1,00E-31	AAL79776.1	Saccharum hybrid cultivar CP65-357	Miscellaneous	PARP	Trypan_PARP (PF05887) Trypan_PARP (PF05887)	IPR008882 IPR008882
SCCCRZ2004B02 SCEQLR1050H09 SCRFFL4006E01	2,73E-57 6.59E-57 2.41E-64	Sb03g000700.1	*Sorghum bicolor*	Miscellaneous	Drepp protein	DREPP (PF05558)	-
SCJLRT3078H06	1,06E-70	Sb05g025670.1	*Sorghum bicolor*	Miscellaneous	Defense-related	Dirigent (PF03018)	IPR004265
SCCCRZ3002G10 SCEZRT2018F03	1,76E-48 1.83E-52	Sb01g031470.1	*Sorghum bicolor*	Proteins with interacting domains (with proteins or polysaccharides)	Protease inhibitor	Cystatin (PF00031)	IPR000010
SCJFLR1013A04 SCRUFL3062D08	4,24E-26 4.11E-26	Sb05g026650.1	*Sorghum bicolor*	Proteins with interacting domains (with proteins or polysaccharides)	Protease inhibitor	Bowman-Birk_leg (PF00228)	IPR000877

Table 2 – Identified ESTs and proteins in sugarcane culms, method 1, showing parameters of the identification together with the functional classification

(concludes)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCSGLR1025E03	0	Sb02g042780.1	*Sorghum bicolor*	Proteins with interacting domains (with proteins or polysaccharides)	PMEI	PMEI (PF04043) Pectinesterase (PF01095)	IPR006501 IPR000070
SCEQLR1093F09	0	Sb01g008040.3	*Sorghum bicolor*	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_3 (PF00933)	IPR001764
SCVPRT2073A02	4.37E-86					Glyco_hydro_3_C (PF01915)	IPR002772
SCCCCL4009G04	2,77E-143	Sb01g004270.1	*Sorghum bicolor*	Unknown	-	DUF642 (PF04862)	IPR006946
SCSGLR1084A12	1.76E-109						
SCCCLB1001G04	2,63E-158	Sb03g027650.1	*Sorghum bicolor*	Unknown	-	DUF642 (PF04862)	IPR006946

Therefore, from 77 different proteins found, 35 had a signal-peptide prediction in two or more prediction programs, being considered as secreted proteins and thus CWPs.

When comparing the total number of proteins identified, 77, it's possible to see that 45,45% were predicted to be secreted, being considered CWPs, as shown in Figure 15. Moreover, 54,55% of the total number of proteins were considered contaminants, being intracellular proteins, directed to other organelles or considered from undefined location.

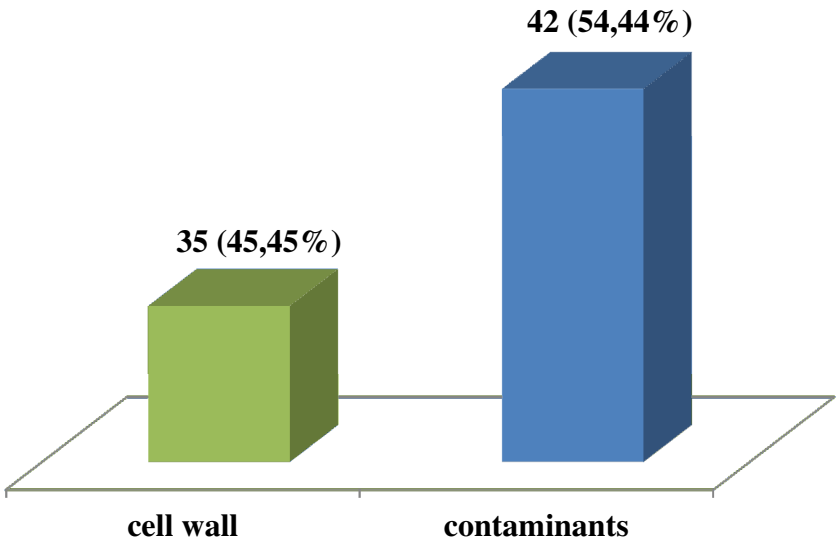

Figure 15 – Number of identified proteins that were predicted to be secreted and the contaminant ones, in sugarcane culms, method 1

Using this method, it was possible to identify a good number of proteins, compared to several other studies (CHIVASA et al., 2002; ZHU et al., 2006; SLABAS et al., 2004), the number of identified proteins was higher in the present study.

The presence of these contaminant proteins can have several explanations, as already discussed in the suspension cells extraction, since both protocols are very similar. It can be actual contamination, since this method is based on grinding the tissue and this releases and traps intracellular proteins; particularly because the sugarcane cell wall, as other grasses, are interlocked by glucuronoarabionoxylans, instead of xyloglucans (CARPITA; GIBEAUT, 1993), which forms a tighter net comparing to dicotyledons, perhaps explaining why the monocots can show more contaminants, as in Brachypodium (personal communication). Furhtermore, contaminants can have assemblance to intracellular proteins but distinct functions, multiple-targeting, or there can be an alternative secretory pathway in each they can be actually secreted (ROSE; LEE, 2010), as already discussed. However, in this study, it's considered CWPs only the ones that present a signal-peptide in two or more prediction programs, as already stated.

The percentage of contaminants was somewhat higher than the cell wall proteins. In the original protocol described by Feiz et al. (2006), when working with *Arabidopsis* hypocotyls, about 73% , as well as Chen et al. (2009), who described almost the same percentage in rice calli. However, the proportion of CWPs was higher when compared to other studies, and in some of them, even when they described lower percentage of contamination, the total number of identified CWPs was much lower. Zhu et al. (2006), found about 40% and 16% of CWPs accounting only for 22 proteins in maize roots. Even working with a different method of extraction, in the present study it was possible to identify more CWPs, 35, which means that this method was sucessful to extract proteins from the cell wall of sugarcane young culms. Comparing the same tissue, stems, as well as a similar method of extraction, only 25 CWPs were identified in *Medicago sativa* (WATSON et al., 2004), however, the percentage of contamination in one of the extracts was lower, 50%. Nevertheless, this study showed that even having a higher percentage of contaminants, more CWPs were identified, and we only considered these proteins in the analysis.

Regarding the contaminants, different histones, mostly 2 and 4, were found, together with many ubiquitins, ribossomal proteins, elongation factor 1 alpha, superoxide dismutase, photosystem reaction center proteins, sedoheptulose, methyl binding domain proteins, oxygen-evolving enhancer protein 1, drepp4 proteins, methionine synthase and some hypothetical proteins (data not shown). Some of them were well-known plastidic or nuclear proteins, and a great number was considered from undefined location. The great number of histones or mitochondrial and chloroplastidial proteins can be explained since they can be trapped with the CWPs. The methionine synthase was also found in other works dealing with cell wall proteins (LIGAT et al., 2011). Considering the existence of an alternative secretory pathway that does not involve a signal-peptide, the number of CWPs could be higher, as an example, the elongation factor 1 alpha was already found in the cell wall (ZHU et al., 1994).

As shown in Figure 16, proteins predicted to be secreted were divided into seven functional classes. The largest proportion was classified as oxido-reductases contributing to 37,14%. The second more abundant class is composed of proteins related to lipid metabolism, 17,14%. The same percentage, about 17,14%, were considered from the miscellaneous class. Proteases were also found in this extract (11,46%). Proteins with interacting domains (with proteins or polysaccharides), accounted for 8,57% from the total. And the unknown proteins corresponded to 5,72%. Surprisingly, proteins acting on cell wall polysaccharides contributed only to 2,86%, and only one protein was found: a glycoside hydrolase. We haven’t found any signaling protein in this experiment.

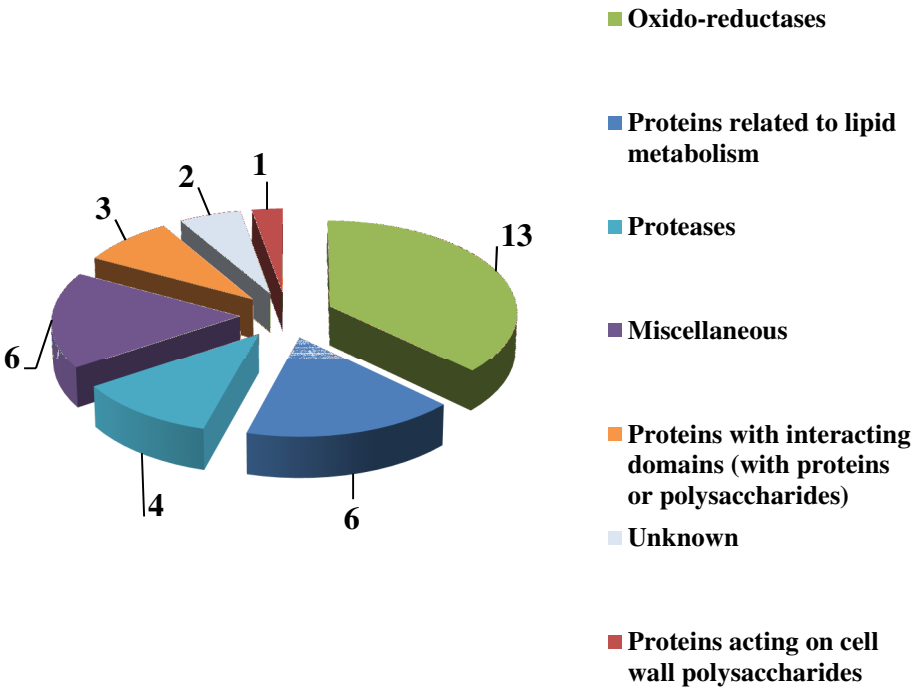

Figure 16 – Number of cell wall proteins identified and their functional classification, in sugarcane culms, method 1

The functional class that presented more proteins was the oxido-reductases, being divided into 10 peroxidases, 2 blue copper binding proteins and 1 germin-like proteins. Peroxidases are well-known cell wall enzymes, identified in several studies (SLABAS et al., 2004; BAYER et al., 2006; KONG et al., 2010), mainly plant peroxidase class III, involved with cell defense (ALMAGRO et al., 2008), as already showed. The peroxidases found in this extraction are members of the family PF00141, according to PFAM (PFAM DATABASE – PFAM, 2012). Perhaps sugarcane culms are very exposed to the action of bugs and microrganisms, since they store the largest amount of sugar in the plant, and that is why it was possible to find so many of them. The number encountered in this extract was somewhat similar to the one found in cells, possibly because peroxidases are very abundant proteins found in all tissues of the plant since they display a fundamental role in the defense mechanism. Being a very young plant, only two-month-old, it is quite reasonable that the plant produce large amounts of peroxidases since the tissues are soft and somehow more vulnerable.

Chemocyanin precursors, the blue copper binding proteins, were also already found in the cell wall proteome (CHEN et al., 2009; LIGAT et al., 2011), and may be related to the formation of the pollen tube in lilly (KIM et al., 2003), as well as they have been associated to redox processes such as electron transfer proteins with small molecular weight compounds (NERSISSIAN: SHIPP, 2002). However, this kind of protein was not found in sugarcane suspension cells. On the other hand, germin-like proteins can be related to the defense mechanism and plant development (SEGARRA et al., 2003; BERNIER; BERNA, 2001) and are relatively common in cell wall proteins studies (BORDERIES et al., 2003; BOUDART et al., 2005; BAYER et al., 2006; KONG et al., 2010). Recently, the germin-like proteins were proved to be related to the enhancement of cross-linking of the cell wall components in the stem tissues of a transgenic tobacco plant (BANERJEE et al., 2010), and studying these proteins we could try to do the opposite, decreases the linking in the cell wall so we achieved a plant with an easily degradable wall in order to produce ethanol from the cellulose.

Proteins related to lipid metabolism were all considered lipid or phospholipid transfer proteins. They are relatively common in cell wall proteins studies (CHEN et al., 2009; BOUDART et al., 2005), and can act on binding lipids to their hydrophobic cavity, which can be essential for the cell wall–loosening activity, facilitating wall extension (NIEUWLAND et al., 2005), and since these culms are formed by growing tissues, it explains the large amount of proteins found in this class, even because sugarcane has a very active growth, much higher when compared to plants like *Arabidopsis*, as an example, so these proteins should be expressed in a large amount and this could be one explanation of why this class of proteins is more abundant in the present study. Besides growth in height, sugarcane culms also increase their width during their development, filling the culms with storage sugar.

A high number of proteases was found, when comparing to cultured cells. Cell wall proteases also play a role in cell defense against pathogen attacks (BAYER et al., 2006). Moreover, aspartic proteinase nepenthesin proteins are members of a subfamily of proteases which physiological role is largely unknown and can be found in several tissues (TAKAHASHI et al., 2008).

The miscellaneous proteins are represented by thaumatin-like proteins, found also in other studies (ZHU et al., 2006; ROBERTSON, 1997; MILLAR et al., 2009), a nucleoside phosphatase, a disease resistance response protein, member of the dirigent family, a DAHP synthase protein, a PARP and a drepp protein. Osmotin protein, the thaumatin-like protein is an antifungal protein and specific interactions of the osmotin with the plasma membrane are responsible for cell death signaling. However, because the cell wall governs access of osmotin

to the plasma membrane, the differences in cell wall composition may account for the differential osmotin sensitivity in *S. cerevisiae* strains (LEE et al., 2010). The family of disease resistance proteins is formed by proteins that are induced during disease response in plants. Members of this family are involved in lignification as well (DAVIN; LEWIS, 2005), suggesting that the formation of a secondary cell wall is developing in these culms. On the other hand, DAHP synthase (EC 2.5.1.54) is the first enzyme in the shikimate pathway, responsible for the biosynthesis of the amino acids phenylalanine, tyrosine, and tryptophan. Thus, it controls the amount of carbon entering the pathway, through transcriptional control, in plants. But its primary function is to catalyze the reaction of phosphoenolpyruvate and D-erythrose 4-phosphate to DAHP and phosphate (HERMANN; ENTUS, 2001), and this enzyme has also been related in stress adaptability (SHARMA et al., 1999).

The proteins with interacting domains found in this study, are formed mostly by protease inhibitors such as cystatin, bowman-birk type trypsin inhibitor and PMEI, and were also described in sugarcane cells. Cystatin, which is a cystein proteinase, was also found in several studies (BOUDART et al., 2005; CHEN et al., 2009; BORDERIES et al., 2003), showing an anti-herbivorous activity (SIQUEIRA-JÚNIOR et al., 2002). Bowman-birk trypsin inhibitor is associated with the regulation of endogenous seed proteinases, storage of sulfur amino acids, and defense against pathogens and insect attack (MELLO et al., 2002). On the other hand, pectinesterase proteins performs the de-esterification of the pectin (PRESSEY, 1984), which can result in changes in the properties of the interactions between cell wall polymers associated with plant growth (YAMAOKA; CHIBA, 1983). And because the culms are in constant growth at this age, the presence of these proteins are in agreement with sugarcane developmental process.

Only one glycoside hydrolase was found, in contrast to some works that showed that the cell wall modifying enzymes was the most represented group for CWPs (BAYER et al., 2006; BOUDART et al., 2005). The role of such proteins in cell walls points out to the rearrangements of polysaccharides during development (CASSAB; VARNER, 1988; FRESHOUR et al., 1996; NICOL; HOFTE, 1998; ROSE; BENNETT, 1999). These proteins can show the importance of polysaccharide metabolism within the extracellular matrix (BOUDART et al., 2005), especially in young tissues like the two-month-old culms. Still, this class of proteins was more abundant in elongating hypocotyls than in fully-grown ones (IRSHAD et al., 2008), so it would be expected that in a young tissue, as sugarcane two-month-old culms, these proteins would be more abundant. Surprisingly they are part of large multigenic families, and their number should be much higher, especially because they are

involved in cell wall carbohydrates modification, a process that should be very active in such young plants. One hypothesis relies on the fact that perhaps these proteins are linked to the wall in a non tight manner, being lost during the tissue washing.

Some unknown proteins were found, both members of DUF642 family. The hypothetical proteins have domains of unknown function. This family represents a conserved region found in a number of uncharacterized plant proteins (BAYER et al., 2006), and was also found in suspension cells. Attention should be driven to this kind of protein, since they are probably abundant in the wall and display an important but still mysterious role.

In this experiment, we were able to generate information about the CWPs found in sugarcane stems that will be compared to the proteins found in method 2.

2.3.2.1.2 Culms from method 2 (based on Boudart et al. (2005))

This method was based on vacuum infiltration followed by centrifugation. This protocol extracts proteins mostly from the apoplastic fluid, being a non-disruptive technique. In order to preserve the integrity of the cells from the damage caused by the combination of vacuum and centrifugation, all the infiltrations were performed in solutions containing mannitol, along with the previously described salt solution.

In the infiltration procedure, using two-month old culms, it was able to identify 223 different ESTs from sugarcane ESTs database (THE SUGARCANE EST PROJECT – SUCEST, 2012), when considering only the ones which were found in both repetitions. From these ESTs, we could identify 91 different proteins through blast, since some different ESTs matched the same proteins. Nine ESTs did not match to any protein, being discarded in the counting of percentage of contaminants.

Figure 17 shows a Venn diagram with the number of ESTs identified in each repetition and the ones found in both of them.

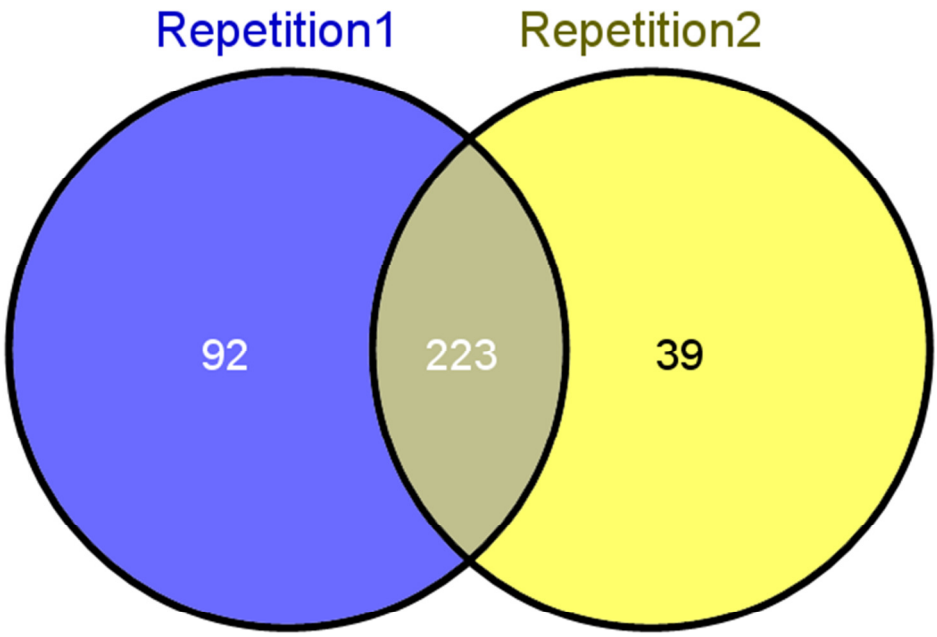

Figure 17 - Venn diagram showing the number of ESTs in the two repetitions of extractions performed, in culms infiltration from method 2.

Therefore, from 91 different proteins found, 58 had a signal-peptide prediction in two or more prediction programs, being considered as secreted proteins and thus CWPs. In Table 3 is possible to see the identified ESTs and proteins predicted to be secreted showing parameters as hit species, together with the functional classification, based on Interpro (APWEILER et al., 2000) and PFAM (SONNHAMMER; EDDY; DURBIN, 1997). Also, in this table is possible to see the multiple ESTs and the correspondent proteins that matched them through blast.

Table 3 – Identified ESTs and proteins in sugarcane culms, method 2, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCCCCL3001B10.b SCJFLR1017E03	0 1.91E-159	Sb01g010840.1	*Sorghum bicolor*	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_1 (PF00232)	IPR001360
SCEQHR1082B01	0	Sb02g028400.1	*Sorghum bicolor*	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_1 (PF00232)	IPR001360
SCVPHR1089A09	1,84E-17	Sb01g008030.5	*Sorghum bicolor*	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_3 (PF00933)	IPR001764
SCEQLR1093F09	0	Sb01g008040.3	*Sorghum bicolor*	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_3 (PF00933)	IPR001764
SCQSLR1089A04	1.90E-152						
SCVPRT2073A02	4.30E-86					Glyco_hydro_3_C (PF01915)	IPR002772

Table 3 – Identified ESTs and proteins in sugarcane culms, method 2, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCEZLB1007A09	0	Sb01g008030.2	*Sorghum bicolor*	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_3 (PF00933)	IPR001764
						Glyco_hydro_3_C (PF01915)	IPR002772
SCCCCL4009F05	0	Sb06g030270.1	*Sorghum bicolor*	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_3 (PF00933)	IPR001764
SCQSAM1030G04	1.26E-65					Glyco_hydro_3_C (PF01915)	IPR002772
						Fn3-like (PF14310)	-
SCQSRT2031D12	0	Sb03g045490.1	*Sorghum bicolor*	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_17 (PF00332)	IPR000490

Table 3 – Identified ESTs and proteins in sugarcane culms, method 2, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCVPRZ3029G05	4,77E-72	Sb03g040600.1	*Sorghum bicolor*	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_18 (PF00704)	IPR001223
SCEZRZ3015E11	2,53E-158	Sb01g048140.1	*Sorghum bicolor*	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_19 (PF00182)	IPR000726
SCBFSD1037A08	1.53E-79	Sb06g021220.1	*Sorghum bicolor*	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Chitin_bind_1 (PF00187)	IPR001002
SCJLLB2076C12	3.17E-158					Glyco_hydro_19 (PF00182)	IPR000726
SCCCCL5004G07	0	Sb10g000660.1	*Sorghum bicolor*	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_28 (PF00295)	IPR000743
SCJFRT1007G04	6.23E-107						

Table 3 – Identified ESTs and proteins in sugarcane culms, method 2, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCCCCL6004H07	0	Sb01g040750.1	*Sorghum bicolor*	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_35 (PF01301) Gal_Lectin (PF02140)	IPR001944 IPR000922
SCVPRZ3029F03	1,18E-150	Sb03g029700.1	*Sorghum bicolor*	Proteins acting on cell wall polysaccharides	Carbohydrate esterase (CE)	PMR5N (PF14416) PC-Esterase (PF13839)	- -
SCBFAD1091A01 SCCCRT1001G12	1,21E-71 0	Sb04g008590.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCCCAD1001B08 SCEQRT1024D03 SCJFRZ2013F04 SCJLRT1019B02	3.10E-19 0 3.25E-154 6.36E-67	Sb03g010740.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016

Table 3 – Identified ESTs and proteins in sugarcane culms, method 2, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCCCAD1001C08	4,43E-84	Sb02g042860.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCEPRZ1011A06	0	Sb03g010250.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCJFRT1010F10	1,58E-70	Sb09g002740.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCRLAD1042E05	8.04E-73						
SCVPRZ2035F03	0						
SCJFRT1061G12	4,00E-100	Sb10g021650.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCQSST3114C09	3,25E-159	Sb01g031740.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCBGRT1047G10	4,19E-111	Sb02g004500.1	*Sorghum bicolor*	Oxido-reductases	Germin	Cupin_1 (PF00190)	IPR006045

Table 3 – Identified ESTs and proteins in sugarcane culms, method 2, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCCCLR2C02D04	4,54E-128	Sb09g004970.1	*Sorghum bicolor*	Oxido-reductases	Germin	Cupin_1 (PF00190)	IPR006045
SCJLLR1104H07	4,23E-60	Sb07g011870.1	*Sorghum bicolor*	Oxido-reductases	Copper ion binding	Cu_bind_like (PF02298)	IPR003245
						Cu-oxidase_3 (PF07732)	IPR001117
SCEPAM1021H07	0	Sb10g027270.1	*Sorghum bicolor*	Oxido-reductases	Copper ion binding	Cu-oxidase (PF00394)	IPR001117
						Cu-oxidase_2 (PF07731)	IPR001117
SCACLR2029H09	1,69E-40	Sb01g033830.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	-
SCCCCL3001E03.b	9.71E-40						

Table 3 – Identified ESTs and proteins in sugarcane culms, method 2, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCCCAD1001B02	1.18E-41						
SCCCLR1072C06	9.87E-40						
SCEPRT2047G01	1.19E-37						
SCEZLR1031G07	1.03E-38						
SCJFRZ2033G07	6.13E-40	Sb08g002700.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	-
SCRFLR1012A10	9.87E-40						
SCRUFL4024B04	1.64E-39						
SCRUSB1064D08	3.39E-41						
SCCCRZ1001H02	5,04E-63	Sb03g039880.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	-

Table 3 – Identified ESTs and proteins in sugarcane culms, method 2, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCCCRZ2002G09 SCQSFL3039E08.b	3,15E-40 1.54E-26	Sb06g016170.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	-
SCEPLB1044H04	7,76E-29	Sb01g049830.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	-
SCBFLR1046E09 SCCCAM2002F12 SCVPRZ2039B03 SCVPRZ2041C11	5.40E-53 1.92E-53 5.40E-53 5.40E-53	Sb03g038280.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368) Fasciclin (PF02469)	- IPR000782

Table 3 – Identified ESTs and proteins in sugarcane culms, method 2, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCBGLR1114E07	2.73E-35	Sb08g002690.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	Tryp_alpha_amyl (PF00234)	IPR003612
SCCCCL3004H07.b	4,47E-19						
SCCCLR1048F06	2.09E-42						
SCCCLR1C08B02	5.10E-36						
SCUTST3131G03	2.10E-39						
SCVPHR1092G06	3,91E-17						
SCCCLR1024C05	2,42E-51	Sb08g002660.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	Tryp_alpha_amyl (PF00234)	IPR003612
SCCCLR1076D05	2,81E-50						
SCEPLB1044H11	2,09E-50						
SCCCLR2C03F01	1,05E-40	Sb08g002670.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	Tryp_alpha_amyl (PF00234)	IPR003612
SCEZLB1006F09	4.00E-32						

Table 3 – Identified ESTs and proteins in sugarcane culms, method 2, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCCCRT1003B03	0	Sb10g003930.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipase	Lipase_GDSL (PF00657)	IPR001087
SCCCCL3005A02.b	2,66E-66	Sb03g039330.1	*Sorghum bicolor*	Miscellaneous	Thaumatin-like protein	Thaumatin (PF00314)	IPR001938
SCEZRZ1014C04	2,73E-139						
SCQSLB1049A01	2.16E-30						
SCCCLR2003G06	6,33E-101	Sb08g018720.1	*Sorghum bicolor*	Miscellaneous	Thaumatin-like protein	Thaumatin (PF00314)	IPR001938
SCUTLR1037F02	3.04E-100						
SCCCSD1003E02	1,21E-106	Sb08g022410.1	*Sorghum bicolor*	Miscellaneous	Thaumatin-like protein	Thaumatin (PF00314) Thaumatin (PF00314)	IPR001938
SCRUHR1076B06	4.30E-107						

Table 3 – Identified ESTs and proteins in sugarcane culms, method 2, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCVPRT2073B04	2,84E-87	Sb08g022420.1	*Sorghum bicolor*	Miscellaneous	Thaumatin-like protein	Thaumatin (PF00314) Thaumatin (PF00314)	IPR001938 IPR001938
SCAGFL1089C05	1.46E-45						
SCCCCL3140H01	7.05E-88						
SCJLFL1049F05	8.86E-17	Sb08g018710.1	*Sorghum bicolor*	Miscellaneous	Aminoacid Biosynthesis DAHP synthase	DAHP_synth_1 (PF00793)	IPR006218
SCRUFL4024B08.b	3.84E-37						
SCSFST1066G10	2.28E-90						
SCAGLR2011E04	4,14E-42						
SCEPAM2057B02	4.14E-42	Sb08g003040.1	*Sorghum bicolor*	Miscellaneous	Stress responsive/Unknown function	Dabb (PF07876)	IPR013097
SCEPLR1051E09	2.11E-46						
SCCCRZ1C01H06	0	Sb08g001950.1	*Sorghum bicolor*	Miscellaneous	Nucleoside Phosphatase	GDA1_CD39 (PF01150)	IPR000407
SCJFRZ2029D05	7.82E-160						

Table 3 – Identified ESTs and proteins in sugarcane culms, method 2, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCVPRT2073B08	3,04E-92	Sb10g001940.1	*Sorghum bicolor*	Miscellaneous	Cysteine-rich secretory proteins	CAP (PF00188)	IPR014044
SCCCRZ2004B02	2,69E-57	Sb03g000700.1	*Sorghum bicolor*	Miscellaneous	Drepp	DREPP (PF05558)	-
SCEQLR1050H09	6.47E-57						
SCRFFL4006E01	2.37E-62						
SCEZRZ3127C02	6,45E-73	Sb02g041760.1	*Sorghum bicolor*	Proteases	Asp proteases	Asp (PF00026)	IPR001461
SCQGSB1083B11	1,47E-119						
SCRLRZ3042B09	2,16E-140	Sb03g026970.1	*Sorghum bicolor*	Proteases	Asp proteases	Asp (PF00026)	IPR001461
SCVPLR2012E01	4,80E-84	Sb01g044790.1	*Sorghum bicolor*	Proteases	Asp proteases	Asp (PF00026)	IPR001461
SCVPRZ2038B09	9,92E-89					Asp (PF00026)	IPR001461

Table 3 – Identified ESTs and proteins in sugarcane culms, method 2, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCMCLR1123H12	4,15E-114	Sb05g027510.1	*Sorghum bicolor*	Proteases	Asp proteases	No PFAM families.	-
SCCCST1004B07	0	Sb01g013970.1	*Sorghum bicolor*	Proteases	Ser proteases	Inhibitor_I9 (PF05922)	IPR010259
						Peptidase_S8 (PF00082)	IPR000209
						PA (PF02225)	IPR003137
SCJFRZ2011B07	5,76E-162	Sb06g016860.1	*Sorghum bicolor*	Proteases	Ser proteases	Inhibitor_I9 (PF05922)	IPR010259
						Peptidase_S8 (PF00082)	IPR000209
						PA (PF02225)	IPR003137

Table 3 – Identified ESTs and proteins in sugarcane culms, method 2, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCCCLR1022B11	0	Sb06g030800.1	*Sorghum bicolor*	Proteases	Cys proteases	Inhibitor_I29 (PF08246)	IPR013201
						Peptidase_C1 (PF00112)	IPR000668
						Granulin (PF00396)	IPR000118
SCAGLB1071B11	2.24E-10	Sb05g026650.1	*Sorghum bicolor*	Proteins with interacting domains (with proteins or polysaccharides)	Protease inhibitor	Bowman-Birk_leg (PF00228)	IPR000877
SCJFLR1013A04	4.17E-26						
SCJFRZ1007E04	1.21E-24						
SCRUFL3062D08	4.03E-26						
SCCCRZ3002G10	1,73E-48	Sb01g031470.1	*Sorghum bicolor*	Proteins with interacting domains (with proteins or polysaccharides)	Protease inhibitor	Cystatin (PF00031)	IPR000010
SCEZRT2018F03	1.80E-52						
SCRUAD1063C06	4,16E-161	Sb09g000430.1	*Sorghum bicolor*	Proteins with interacting domains (with proteins or polysaccharides)	LRR protein	LRRNT_2 (PF08263)	IPR013210
						LRR_4 (PF12799)	-

Table 3 – Identified ESTs and proteins in sugarcane culms, method 2, showing parameters of the identification together with the functional classification

(concludes)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCSGLR1025E03	0	Sb02g042780.1	*Sorghum bicolor*	Proteins with interacting domains (with proteins or polysaccharides)	PMEI	PMEI (PF04043) Pectinesterase (PF01095)	IPR006501 IPR000070
SCCCRZ2002A12	2,80E-97	Sb09g028480.1	*Sorghum bicolor*	Signaling	Fasciclin-like	Fasciclin (PF02469)	IPR000782
SCRFLR1034D04	2,79E-92	Sb03g030690.1	*Sorghum bicolor*	Signaling	Fasciclin-like	Fasciclin (PF02469)	IPR000782
SCCCCL4009G04 SCSGLR1084A12	2.72E-144 1.73E-110	Sb01g004270.1	*Sorghum bicolor*	Unknown	-	DUF642 (PF04862)	IPR006946
SCVPLR2027A11	1,55E-77	Sb07g026630.1	*Sorghum bicolor*	Unknown	Auxin-responsive protein AIR12 (DUF568)	DUF568 (PF04526)	IPR017214

When comparing the total number of proteins identified, 91, we see that 63,74% had a prediction for secretion, being considered CWPs, as shown in Figure 18. Moreover, 36,26% of the total number of proteins were considered contaminants, being intracellular proteins, directed to other organelles or considered from undefined location.

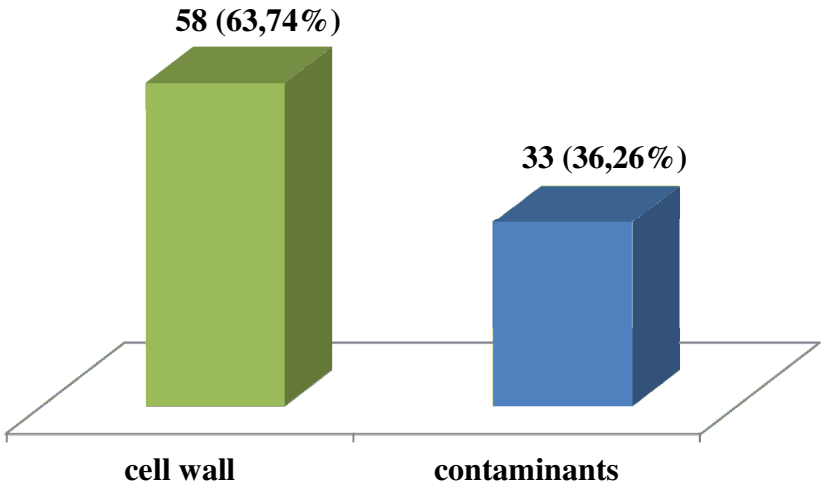

Figure 18 – Number of identified proteins that were predicted to be secreted and the contaminant ones in sugarcane culms, method 2

Using this method, it was able to identify more CWPs, when comparing to Method 1. Thus, comparing to several other studies (CHIVASA et al., 2002; ZHU et al., 2006; SLABAS et al., 2004), this number was also higher.

The presence of these contaminant proteins can have several explanations. Although this method is called a non-disruptive technique, maybe the unique characteristics of grasses cell wall can increase the trapping of these proteins, even using vacuum infiltration. Also, since sugarcane culms cannot fit a centrifuge tube without cutting them, this could also increase the percentage of contaminants. In the original protocol, Boudart et al. (2005) used whole plants, without cutting them to perform the extractions, and this was not possible when working with sugarcane, since the culms were cut from each 7 cm, and this should have released some contaminant proteins because of the tissue disrupture.

Other possibilities already discussed could be proteins considered to be contaminants that have high high identity to intracellular ones but distinct functions, double-targeting (ROSE; LEE, 2010) as already mentioned, the existence of an alternative secretory pathway (ROSE; LEE, 2010), among others. Thus, we're gonna consider CWPs only the ones that present a signal-peptide in two or more prediction programs, as already stated.

The percentage of contaminants is lower, compared to the cell wall proteins. In the original protocol described by Boudart et al. (2005), when working with *Arabidopsis* rosettes,

about 94% of the proteins were considered to be secreted, showing one of the lowest percentage of contamination. However, Boudart et al. (2005) study identified 93 proteins in the total, despite of 87 being CWPs. A similar percentage of contamination was found in *Medicago sativa* stems, in comparison to the present study, but a much lower number of CWPs were identified: only 25 (WATSON et al., 2004). Maybe when working with harder materials, such as sugarcane culms, the percentage of contamination can be increased. Moreover, The proportion of CWPs was higher when compared to several studies (ZHU et al., 2006; CHIVASA et al., 2002) Nevertheless, this study showed that even having a higher percentage of contaminants, 58 CWPs were identified, which represents a good number, and only these proteins were considered in the analysis.

Regarding the contaminants, several ubiquitins and polyubiquitins, histones, hypothetical proteins, drepp4 proteins, ribossomal proteins, methyl binding domain106, polyphenol oxidase, oxygen-evolving enhancer protein 3-1, pyruvate dehydrogenase, superoxide dismutase, 5-methyltetrahydropteroyltriglutamate-homocysteine expressed, elongation factor 1 alpha, c3hl domain class transcription factor, translationally-controlled tumor protein, dna-binding protein s1fa2, L-ascorbate oxidase precursor, remorin-like isoform 1, udp-glucose pyrophosphorylase, cytosolic ascorbate peroxidase, methionine synthase, were identified, among others. Some of them were well-known plastidic, chloroplastidic or nuclear proteins, and a great number was considered from undefined location. If considering the existence of an alternative secretory pathway that doen'st involve a signal-peptide, the number of CWPs could be increased, as the case of the elongation factor 1 alpha, as already described.

As shown in Figure 19, the proteins predicted to be secreted were divided into eight functional classes. The largest proportion was classified as proteins acting on cell wall polysaccharides contributing to 22,41%. The second more abundant class is composed of oxido-reductases contributing to 18,96%. It was also found proteins related to lipid metabolism (17,24%), and miscellaneous (15,52%). The proteases added 12,07% of the total and proteins with interacting domains contributed to 6,90%. About 3,45% were considered unknown proteins and the same percentage was shown in the signaling proteins.

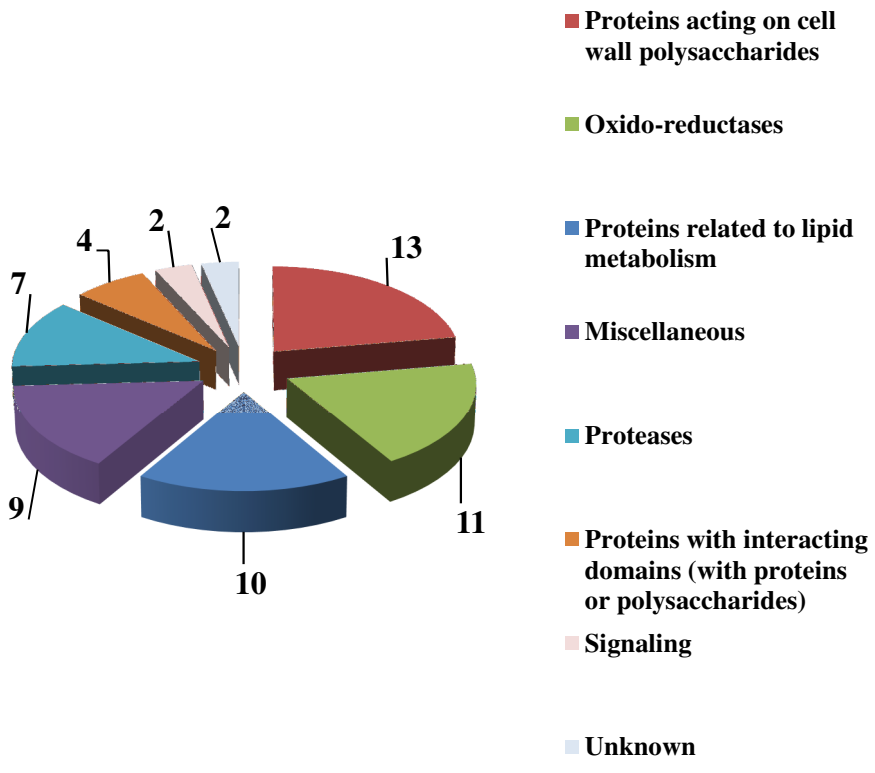

Figure 19 – Number of cell wall proteins identified and their functional classification, in sugarcane culms, method 2

The cell wall modifying enzymes were represented by twelve glycoside hydrolases and one carbohydrate esterase. In accordance, some works also showed that the cell wall modifying enzymes was the most represented group for CWPs (BAYER et al., 2006; BOUDART et al., 2005). Differently from method 1, several types of glycoside hydrolases could be identified, probably due to the type of extraction used that could release these proteins. Perhaps these proteins were lost in the washes from method 1 and that is why in the method 2 they appear to be more abundant. The role of such proteins in cell walls points out to the rearrangements of polysaccharides during development (CASSAB; VARNER, 1988; FRESHOUR et al., 1996; NICOL; HOFTE, 1998; ROSE; BENNETT, 1999). These proteins can show the importance of polysaccharide metabolism within the extracellular matrix (BOUDART et al., 2005), especially in young tissues like the two-month-old culms. In this class there were β-galactosidases, β-d-xylosidase, glucosidases, endochitinases, polygalacturonase and lustrin a-like proteins. The polygalacturonase, only found in this experiment, is related to the replacement of pectins (MILLAR et al., 2009). Curiously, the lustrin proteins are multifunctional proteins, and probably are members of a multiprotein

family, involved in the adhesion of the nacreous layer of the shell and pearl of abalone (SHEN et al., 1997). Probably, the lustrin-like proteins have a similar function in the adhesion of the cell wall layers, since it was found in this analysis and were not found in any other extraction, being a interesting protein to be studied in order to elucidate the mechanics of the cell wall.

The oxido-reductases class was divided into 7 peroxidases, 2 germin-like proteins and 2 Copper ion binding proteins. Peroxidases are well-known cell wall enzymes, identified in several studies (SLABAS et al., 2004; BAYER et al., 2006; KONG et al., 2010), mainly peroxidase class III, as well as in method 1. Plant peroxidases are mainly localised in the extracellular matrix and vacuole and in growing plant cell walls, being are involved in the oxidative cross-linking of cell wall polymers (BAYER et al., 2006), and in the mechanism of cell defense (ALMAGRO et al., 2008). In the original protocol (BOUDART et al., 2005), no peroxidases were identified, even this group is extremely abundant in the wall, and one hypothesis was that maybe the method couldn't extract this proteins. However, several peroxidases were identified in sugarcane culms, what shows that is possible to extract these proteins using vacuum infiltration. Despite no chemocyanin proteins were found in this method, it seems that rhicadhesin proteins can have a high protein identity to germin-like proteins (SEBASTIAN et al., 2007), and thus, can show correlated functions. Phytocyanins are plant-specific blue copper proteins involved in electron transport, and some of them are known to be chimeric arabinogalactan proteins (MA et al., 2011).

The proteins related to lipid metabolism were all considered lipid transfer proteins except for a lipase. Lipid transfer proteins are relatively common in cell wall proteins studies (CHEN et al., 2009; BOUDART et al., 2005), and can act on binding lipids to their hydrophobic cavity, which can be essential for the cell wall–loosening activity, facilitating wall extension (NIEUWLAND et al., 2005). Some lipid transfer proteins can be involved in binding fatty acids and to transfer phospholipids among membranes (KADER, 1997), another one, LTP2, was found to be up-regulated in the epidermis of stems contributing to cuticle formation during elongation (SUH et al., 2005). Also, the LTPs were proposed to induce systemic resistance signaling , an ability to promote cell wall expansion through the bind to a hydrophobic partner in cell walls and activation of a polygalacturonase (NIEUWLAND et al., 2009). In growing tissues is reasonable to think that we encountered such group of proteins in a large amount. On the other hand, the GDSL Lipase family can act on cutin or suberin lipids (IRSHAD et al., 2008).

The miscellaneous proteins were represented by pop3 peptide, thaumatin-like proteins, also found in other studies (BOUDART et al., 2005; BAYER et al., 2006; KONG et al.,

2010), as well as aminoacid biosynthesis proteins, a nucleoside phosphatase, a cysteine-rich secretory protein and a Drep protein. Thaumatin-like proteins are often found and their roles have been already discussed previously in this work. Already found in other CWPs studies (BAYER et al., 2006), the family of disease resistance proteins is formed by glycoproteins that are induced during disease response in plants, being the secretory-rich proteins a large multigenic family involved in several functions. The major group of CAP superfamily proteins in plants, Pr-1 family, are the most highly up-regulated proteins in plants after pathogen attack (van LOON et al., 2006), and, in culms, it is reasonable to think that these proteins are involved in plant defense. Thus, the hypersensitive response, that leads to the isolation and death of infected cells, faces a dramatic up-regulation of the Pr proteins (GIBBS; ROELANTS, O, BRIAN, 2008). The Pr-1 proteins can correspond to10% of the total protein within the infected leaf (CORNELISSEN et al., 1986). It seems that this family may possess an antioomycete activity through negative regulation of β-(1–3)-glucanase (RIVIERE et al., 2008). This family of proteins can occur in noninfected parts of the plants, as well as in healthy ones (FRASER, 1981), suggesting that can be involved in plant systemic acquired resistance response (RYALS et al., 1996), analogous to acquired immunity in mammals (GIBBS; ROELANTS, O, BRIAN, 2008). Pop3 proteins can be linked to pollen tube guidance (WILHELMI; PREUSS, 1998), and its presence surprises us.

A large number of proteases were found in culms infiltration. Besides aspartyl proteases, it was also found two serine proteases and one cysteine protease. Cell wall proteases can play a role in cell defense against pathogen attacks (BAYER et al., 2006). Moreover, aspartic proteinase nepenthesin proteins are members of a subfamily of proteases which physiological role is largely unknown and can be found in several tissues (TAKAHASHI et al., 2008). On the other hand, plant cysteine-proteases regulates physiological functions, like senescence and seed germination, and defense roles, representing a well-characterized type of proteolytic enzymes (MARTÍNEZ et al., 2011). Subtilisin-like proteins, also called subtilases are serine-dependent enzymes and play degradative roles, and may also display specific functions in plant development and signalling (RAUTENGARTEN et al., 2005; SCHALLER, 2004). In *Arabidopsis*, subtilases were shown to be involved in the generation of signals responsible for stomata density regulation (VON GROLL; BERGER; ALTMAN, 2002) and cuticle formation and epidermal differentiation during *Arabidopsis* embryo development (TANAKA et al., 2001). Serine proteases were only found in this method, maybe because the vacuum infiltration could extract them more easily.

The proteins with interacting domains found in this study, are formed mostly by bowman-birk type trypsin inhibitor, cystatin, PMEI and a LRR protein. Almost the same proteins found in culms extraction from method 1, with the exception of LRR protein, as already showed. It is known that LRR proteins are involved in cell defense, however, the biochemical or biologicalfunctions of these proteins are still unknown, being also found in the original protocol (BOUDART et al., 2005).

Two fasciclin-like proteins were identified and are related to the signaling mechanism, as well as cell adhesion. Fasciclin-like proteins are one of the most common signaling proteins found in CWPs extracts.

Like the extraction from method 1, we have found a member of DUF642 family, , as well as a member of DUF 568 family, called auxin-responsive protein AIR12. This family represents a conserved region found in a number of uncharacterised plant proteins (BAYER et al., 2006). Evidence suggests that AIR12 may interact with other redox partners within the plasma membrane to constitute a redox link between cytoplasm and apoplast (PREGER et al., 2009).

This method based on vacuum infiltration seems adequate to isolate CWPs, since showed a smaller percentage of contaminants, comparing to the method 1. The proteins found in the same tissue, culms, were quite similar, but some different proteins were found in method 2 indicating that is a more successful method to be used.

The proteins described here can contribute to the comprehension of the cell wall processes, and can be used as information to try to elucidate some unknown mechanisms.

2.3.2.1.3 Leaves from method 1 (based on Feiz et al. (2006))

In the extraction procedure, using two-month old leaves, it was able to identify 177 different ESTs from sugarcane EST database (THE SUGARCANE EST PROJECT – SUCEST, 2012), when considering only the ones which were found in both repetitions. From these ESTs, 67 different proteins were identified through blastp, since some different ESTs matched the same proteins. All the ESTs matched proteins.

Figure 20 shows a Venn diagram with the number of ESTs identified in each repetition and the ones found in both of them.

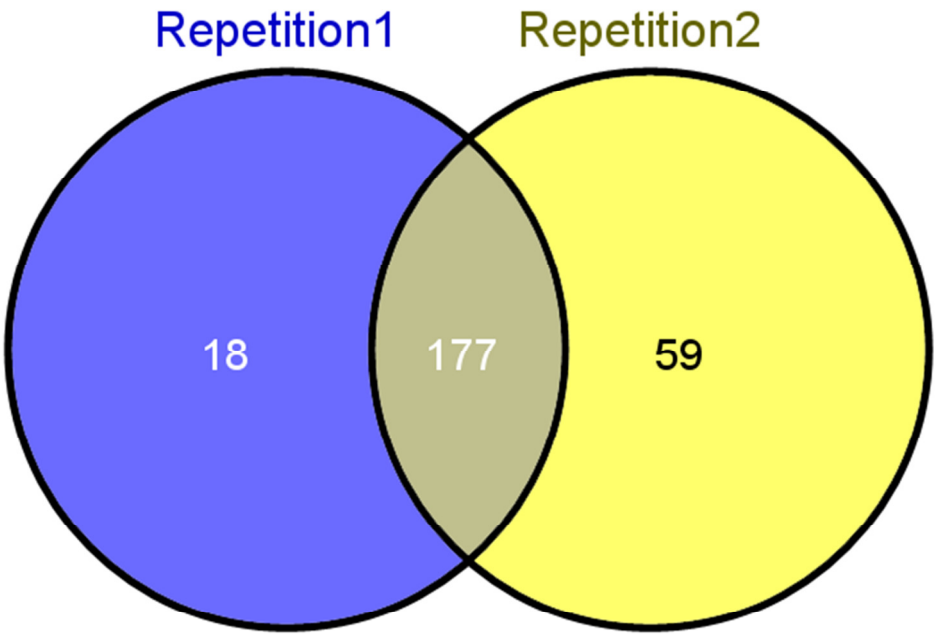

Figure 20 - Venn diagram showing the number of ESTs in the two repetitions of extractions performed, in leaves extraction from method 1.

Therefore, from 67 different proteins found, only 15 had a signal-peptide prediction in two or more prediction programs, being considered as secreted proteins and thus CWPs. In Table 4, it is shown the identified ESTs and proteins predicted to be secreted showing parameters as hit species together with the functional classification, based on Interpro (APWEILER et al., 2000) and PFAM (SONNHAMMER; EDDY; DURBIN, 1997). Also, in this table is possible to see the multiple ESTs and the correspondent proteins that matched them through blast.

Table 4 – Identified ESTs and proteins in sugarcane leaves, method 1, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCBGLR1114E07	2,78E-36	Sb08g002690.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	Tryp_alpha_amyl (PF00234)	IPR003612
SCCCCL3004H07.b	4,55E-19						
SCCCLR1048F06	2.13E-43						
SCCCLR1C08B02	5.19E-36						
SCUTST3131G03	2.14E-39						
SCVPHR1092G06	3,98E-17						
SCCCLR1024C05	2,46E-50	Sb08g002660.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	Tryp_alpha_amyl (PF00234)	IPR003612
SCCCLR1076D05	2,86E-50						
SCEPLB1044H11	2,13E-50						
SCCCLR2C03F01	1,07E-40	Sb08g002670.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	Tryp_alpha_amyl (PF00234)	IPR003612

Table 4 – Identified ESTs and proteins in sugarcane leaves, method 1, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCEPLB1044H04	7,90E-29	Sb01g049830.1	*Sorghum bicolor*	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	-
SCCCCL3002E11.b	0	Sb03g013200.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCCCCL7037A10	0	Sb02g027330.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCCCLR1C03A09	0	Sb09g004650.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016
SCJFLR1035D05	3.87E-140						
SCVPLB1020D03	3,85E-132	Sb03g046760.1	*Sorghum bicolor*	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	IPR002016

Table 4 – Identified ESTs and proteins in sugarcane leaves, method 1, showing parameters of the identification together with the functional classification

(continued)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCCCRZ1C01H06 SCJFRZ2029D05	0 7.96E-160	Sb08g001950.1	*Sorghum bicolor*	Miscellaneous	Nucleoside Phosphatase	GDA1_CD39 (PF01150)	IPR000407
SCCCSD1089G04 SCJLLR1103A09	4,16E-44 4,16E-44	Sb07g000600.1	*Sorghum bicolor*	Miscellaneous	Ferredoxin	Fer2 (PF00111)	IPR001041
SCCCRZ2004B02 SCEQLR1050H09 SCRFFL4006E01	2,74E-57 6.59E-57 2.41E-62	Sb03g000700.1	*Sorghum bicolor*	Miscellaneous	Drepp	DREPP (PF05558) DREPP (PF05558) Drepp (PF05558)	-
SCEZRZ3015E11	2,57E-158	Sb01g048140.1	*Sorghum bicolor*	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_19 (PF00182)	IPR000726

Table 4 – Identified ESTs and proteins in sugarcane leaves, method 1, showing parameters of the identification together with the functional classification

(concludes)

SAS SUCEST	E-value	Hit	Species	Functional class	Functional class II	PFAM (Family)	InterPro
SCQSRT2031D12	0	Sb03g045490.1	*Sorghum bicolor*	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_17 (PF00332)	IPR000490
SCBGLR1097G03 SCMCLR1123H12	9,98E-71 5,58E-121	Sb05g027510.1	*Sorghum bicolor*	Proteases	Asp proteases	Asp (PF00026)	IPR001461
SCEZLB1013B06	2,88E-136	Sb10g001440.1	*Sorghum bicolor*	Proteases	Asp proteases	Asp (PF00026)	IPR001461

Therefore, from 67 different proteins found, only 15 had a signal-peptide prediction in two or more prediction programs, being considered as secreted proteins and thus CWPs.

When comparing the total number of proteins identified, 67, we see that only 22,39% had a prediction for secretion, being considered CWPs, as shown in Figure 21. Moreover, 77,61% of the total number of proteins were considered contaminants, being intracellular proteins, directed to other organelles or considered from undefined location.

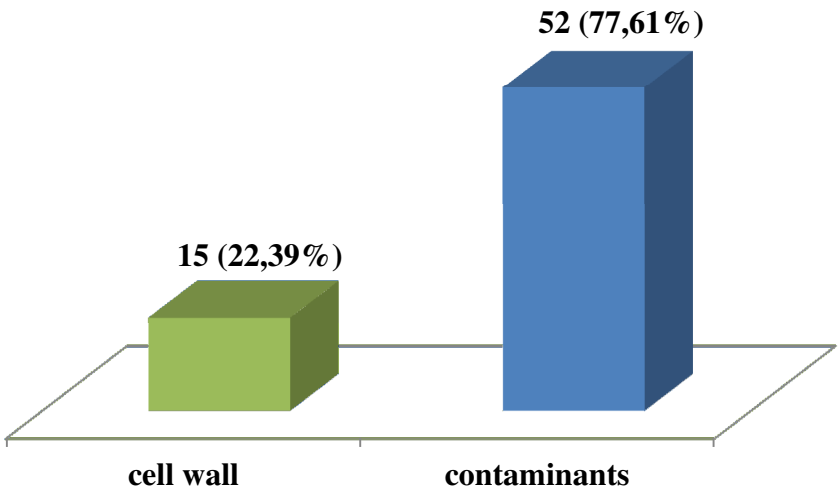

Figure 21 – Number of identified proteins that were predicted to be secreted and the contaminant ones in sugarcane leaves, method 1

Using this method, it was able to identify a small number of proteins and the percentage of contaminants was high. The presence of these contaminant proteins can have several explanations, as already discussed previously in the extraction from the culms using this same method. But when comparing leaves and culms, the percentage of contamination is much higher in the first. These proteins can be actual contamination, beacuse the method can release intracellular proteins (BORDERIES et al., 2003). Perhaps the unique characteristics of grasses cell wall (CARPITA; GIBEAUT, 1993) can increase the trapping of these proteins. Moreover, the stem and leaves have different characteristics; as an example, in barley, comparing stem and leaves through NMR, the stem fraction showed more lignin with a higher proportion of ether-linked syringyl residues and the leaf fraction contained more phenolic acids and cutin (LOVE; SNAPE; JARVIS, 1998). In other words, probably the stem is more rigid and withstands mechanical stress, what makes more difficult to break the structure of the cell walls. When breaking the tissues, it is possible that more contaminant proteins are

released, increasing then the percentage of contamination of the leaves. On the other hand, because the leaves are formed by softer tissues, the same amount of effort to disrupt the cell can break it more efficiently, and, thus, release more contaminants, when comparing to the stem.

Furhtermore, these proteins considered to be contaminants can have high identity to known intracellular proteins but distinct functions, or they may display more than one function, in multiple cellular compartments (ROSE; LEE, 2010).

Another possibility, is the existence of an alternative secretory pathway, as already discussed (ROSE; LEE, 2010). Thus, it is considered CWPs only the ones that present a signal-peptide in two or more prediction programs, as already stated.

The percentage of contaminants was much higher than the cell wall proteins. In the original protocol described by Feiz et al. (2006), a much higher percentage of CWPs was found. However, when working with roots from maize, that is a monocotyledon like sugarcane, Zhu et al. (2006), found about 40% and 16% of proteins that were predicted to have a signal peptide in the N-terminal sequence, in one fraction and total soluble proteins, respectively, accounting for 22 CWPs, at the total. Thus, the percentage of CWPs in one of these extracts was smaller than found in sugarcane leaves. Chen et al. (2009) also found that the lack of studies regarding CWPs from rice leaves was due to their rigid and hydrophobic surfaces, like sugarcane.

In addition, Chivasa et al. (2002), as well as Watson et al. (2004), found a similar number of CWPs when comparing to the present study . Nevertheless, this experiment showed the higher percentage of contaminants comparing to all the experiments already discussed.

Regarding the contaminants, all the types of proteins displayed on the previous experiments were found, plus ribosome reciclying factor, malate dehydrogenase, peroxiredoxins, pyruvate dehydrogenase, ribulose-bisphosphate carboxylase oxygenase, thioredoxin m- and f- type, cytochrome c, glyoxalase I, ferredoxin--nadp enzyme, ribose-5-phosphate isomerase, malate glyoxysomal, nadp-dependent malic enzyme, unknown and hypothethical proteins. Some of them were well-known plastidic, chloroplastidic or nuclear proteins, and a great number was considered from undefined location. The great number of histones or mitochondrial and chloroplastidial proteins can be explained since they can be trapped with the membranes that circumvent these organelles and be mistakenly extracted. If we consider the existence of an alternative secretory pathway that does not involve a signal-peptide, the number of CWPs could be increased. For example, a peroxiredoxin was

immunolocalised to the extracellular surface of a potato nemotode despite the lack of a signal-peptide for secretion (ROBERTSON et al., 2000) and could also play a role as a phospholipase (CHEN, 2000).

As shown in Figure 22, the proteins predicted to be secreted were divided into five functional classes. The largest proportion was classified as proteins related to lipid metabolism as well as oxido-reductases contributing to 26,67% each. The third more abundant class was the miscellaneous (20%). Proteins acting on cell wall polysaccharides (with proteins or polysaccharides) accounted for 13,33%, as well as proteases. No proteins were classified into the classes: unknown, signaling and proteins with interacting domains.

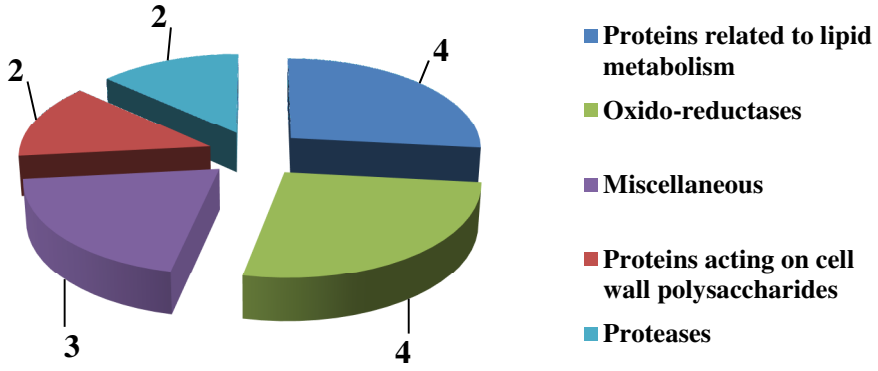

Figure 22 – Number of cell wall proteins identified and their functional classification, in sugarcane leaves, method 1

Proteins related to lipid metabolism were all considered lipid or phospholipid transfer proteins. These type of proteins can be essential for the cell wall–loosening activity, facilitating wall extension (NIEUWLAND et al., 2005). In two-month-old leaves, formed by growing tissues, is reasonable to think that we encountered such group of proteins in a large amount.

The oxido-reductases class was represented only by peroxidases. Especially class III peroxidases are among the proteins induced during the host plant defense. The peroxidases belong to a large multigene family, and participate in lignin and suberin formation, cross-linking of cell wall components, and synthesis of phytoalexins (ALMAGRO et al., 2008). Since the leaves can suffer from pathogen attacks, the number of peroxidases found in this study make sense in order to avoid damage and protect the plant.

Three proteins were classified as being from the miscellaneous class, a nucleoside phosphatase, a drepp, and a ferredoxin-precursor. Using the same method of extraction, the same nucleoside phosphatase was found in culms. Curiously, a traditionally chloroplastidic

protein was found as well, and, more surprisingly, it showed a signal-peptide prediction for secretion. This is probably a case of moonlighting protein, that can be directed to both chloroplast, as well as to the cell wall. In *Arabidopsis*, a ferredoxin directed to the apoplast was able to confer resistance against bacteria (YI-HSIEN et al., 2010). Perhaps, the ferredoxin found in the present study was also performing this kind of resistance, in leaves. This kind of protein was also found in the cell wall extract of alfafa stems (WATSON et al., 2004).

The cell wall modifying enzymes class was represented by glycoside hydrolases: an endochitinase precursor and a beta-glucanase precursor. Both were also found in culms extraction, but from method 2. Endochitinase is involved in the defense mechanism, especially against fungi (GIRHEPUJE; SHINDE, 2011), among other functions. On the other hand, in *Arabidopsis*, beta-glucanase is found to be associated with plant growth, xylem development and cell wall thickening (SHANI et al., 2006), being a good

Only two proteases, aspartyl-proteases, were found in this analysis. Exactly the same types of proteins found in culms extraction, from method 1.

No proteins with interacting domains, signaling and unknown proteins were found in this experiment. The lower content of CWPs found in leaves can be related either to the high percentage of contamination and/or to a less complex cell wall structure, when comparing to culms that present the functions of succrose accumulation in addition. Another possibility, is that culms are much thicker tissues in comparison to leaves, and this can implicate to a more intricate system of proteins expression related to the cell wall. The proteins found in this study are in agreement with the cell wall expected role.

2.3.2.1.4 All extracts

Considering all extracts, both common cell wall proteins as well as rare ones were found and from them we can highlight the ones related to cell expansion, growth, signaling and carbohydrate modification. These proteins can be represented by proteins acting on cell wall polysaccharydes, such as hydrolases, lipid transfer proteins, among others. When trying to manipulate sugarcane wall carbohydrates, attention should be driven to the proteins encountered in this study, especially glycoside-hydrolases, chitinases, beta-glucaneses, that are involved in the rearrangement of the wall, germin-like proteins that are linked to carbohydrates modification, lipid-transfer proteins, lustrin-like proteins, peroxidases and PMEI, that are related to the wall extension and also can be involved in cell adhesion. We

could not forget the unknown proteins, especially DUF642, found in all extracts, and that can play a still misterious but important role in the cell wall. No structural protein was found, just like other studies (BOUDART et al., 2005; MINIC et al., 2007; LIGAT et al., 2011), and that is probably because these proteins are difficult to extract when they are cross-linked, and also are hard to identify since they have numerous post-translational modifications (IRSHAD et al., 2008).

Found only in one replicate of cultured cells CWPs extraction (APPENDIX B), a glycosyl transferase may be involved with hemicellulose biosynthesis (MILLAR et al., 2009). This was also the case for three beta-expansins (XP_002464945, XP_002464944, XP_002464951), which regulates cell growth by breaking hydrogen bonds between cellulose and xyloglucan (COSGROVE, 2000).

Also in only one extraction (APPENDIX D), an alfa-L-arabinofuranosidase (XP_002442705) was found in culms from method 2, which acts in major polysaccharydes mainly present on type II cell walls, that is the type of sugarcane and most of the monocots cell wall (ZHU et al., 2006).

Some glycoside hydrolases, specific from monocots, were found in this study, and we suggest that this kind of protein can be used in future research to manipulate the sugarcane cell wall in order to produce a plant that doesn't need so many enzymes in order to produce cellulosic ethanol in a viable way.

These findings corroborate the information that cell wall can differ from species to species as well as from different tissues, but at the same time, they all share several common characteristics, since various proteins were found in all extracts.

Regarding the contaminants, superoxide dismutase is a recorrent protein found in the experiments, as well as in other works (CHEN et al., 2009), being triggered by environmental stresses. Others like histones, ribossomal proteins, malate dehydrogenase, methyonine synthase, heat shock proteins, calmodulin, chaperonin, elongation factor 1 alpha, enolase, fructose-bisphosphate aldolase, luminal binding proteins, and some hypothetical and unknown proteins, were also found and many of them, despite the lack of a signal-peptide for secretion, were already immunolocalized in the apoplast (CHIVASA et al., 2002). Thus, the number of real contaminants could be decreased, since this study also corroborates to provide evidence for the occurrence of a non traditional secretory pathway, independent from the presence of a signal-peptide, similar than found in mammals (COOPER; BARONDES, 1990; RUBARTELLI et al., 1990).

3 CONCLUSION

The plant cell wall is a compartment that is the interface between the intra and extracellular matrices, and, thus, take part of a large number of activities. From these, we can highlight cell wall stress and disease response, signaling, secretion of proteins, cell-to-cell communication, cell growth and expansion, among many others. However, all the processes that cell walls can be involved are not completely known and more information need to be searched about its components and their function in cell wall dynamics. Proteomic analysis provides a new approach to discover new components and pathways that can try to elucidate and understand cell wall functions. This lack of information can reflect the technical problems and the difficulty that is the work with cell wall proteins. Since the cell wall is not an isolate organelle, rather being an intricate network embedded in a matrix, it is a major hurdle to extract the proteins that are strongly linked to it. And by doing so, some intracellular proteins can be found in the extract as well.

In the present study, it was possible to identify different cell wall proteins using three different tissues from sugarcane. To our knowledge, this is the first report of characterization of cell wall proteins from sugarcane using an extraction protocol especifically for this subset of proteins. Besides the low percentage of cell wall proteins found in two-month-old leaves and cultured cells, which could be due to the destructive processes in which the material were submitted and the technical challenges, among other reasons, this study could establish protocols that can be used in cell wall proteins extraction sucessfully similar to other studies. One possibility that cannot be discarded is the existence of an alternative secretory pathway, which could increase the number of secreted proteins, but has not been proved to exist in plant cells and more experiments such as in planta localisation of the proteins using antibodies or fluorescent fusion proteins should be performed in order to elucidate this mechanism.

Altogether, the present study could identify 157 different ESTs clusters related to sugarcane cell wall. Considering that in a previous study, Lima et al (2001) identified 459 sugarcane clusters related to the cell wall in several tissues and different ages, in the present study, performing extractions with only three kinds of material: cells, leaves and culms, it was possible to account for almost one-third of this amount. Among these, 75 different cell wall proteins were characterized in eight functional classes. Regarding the cell wall protein extraction of sugarcane culms, a lower percentage of contaminants was found in both methods. However, the contamination in the infiltration protocol was notably smaller, and thus, this could be a good strategy to be used when working with monocots, since perhaps the

more intricated cell wall of the grasses can interfere negatively in the use of destructive protocols such as method 1. Even knowing that we can not reach one hundred percent of the cell wall proteins, we could highlight the vacuum infiltration protocol as a more suitable protocol to be used and generate new information regarding some proteins with unexpected location.

The data generated in this study can be used in future research to modify the cell wall components by genetic transformation in order to reach a more productive plant. As an example, research can be done to generate a plant with its wall more easily degradable, reducing the need of enzymes, and thus, eliminating costs in the cellulosic ethanol production. Some questions remains in the low percentage of glycoside hydrolases found in some extracts, being a large multigenic families, and the high content of lipid transfer proteins in the extracts, in general.

Therefore, this study could collaborate in three main aspects: to test different protocols as well as to establish a suitable one for cell wall proteins from sugarcane; to characterize the cell wall proteins from sugarcane in cells and in two-month-old plants showing which proteins could be found and their function; and to generate information for future research that can manipulate the wall in order to make a more viable cellulosic ethanol.

REFERENCES

ALBENNE, C.; CANUT, H.; BOUDART, G.; ZHANG, Y.; SAN CLEMENTE, H.; PONT-LEZICA, R.; JAMET, E. Plant cell wall proteomics: mass spectrometry data, a trove for research on protein structure/function relationships. **Molecular Plant**, v. 2, n. 5, p. 977-989, Sep. 2009.

ALBERSHEIM, P.; DARVILL, A.; ROBERTS, K.; SEDEROFF, R.; STAEHELIN, A. **Plant Cell Walls**. New York: Garland Science, 2011. 430 p.

ALEXANDER, A. G. **The energy cane alternative**. Amsterdam: Elsevier, 1985. 509 p.

ALMAGRO, L.; ROS, L.V.G.; BELCHI-NAVARRO, S.; BRU, R.; BARCELO, A.R.; PEDRENO, M.A. Class III peroxidases in plant defence reactions. **Journal of Experimental Botany**, Oxford, v. 60, n. 2, p. 377-390, Feb. 2009.

AMORIM, H. V.; LOPES, M.L.; OLIVEIRA, J.V.D.; BUCKERIDGE, M.S.; GOLDMAN, G.H. Scientific challenges of bioethanol production in Brazil. **Applied Microbiology and Biotechnology**, Berlin, v. 91, n. 5, p. 1267-1275, Sep. 2011.

ANDREWS, J.; ADAMS, S.R.; BURTON, K.S.; EVERED, C.E. Subcellular localization of peroxidase in tomato fruit skin and the possible implications for the regulation of fruit growth. **Journal of Experimental Botany**, Oxford, v. 53, n. 378, p. 2185-2191, Nov. 2002.

ANGYAL S. J.; CRAIG, D.C.; KUSZMANN, J. The composition and conformation of D-Threo-3,4-Hexodiulose in solution, and the X-ray crystal-structure of its beta-anomer. **Carbohydrate Research**, Amsterdam, v. 194, p. 21-29, Dec. 1989.

ANNOTATION PROTEIN DATABASE – PROTANNDB. Disponível em: <http://www.polebio.scsv.ups-tlse.fr/ProtAnnDB/>. Acesso em: 05 ago. 2012.

APWEILER, R.; ATTWOOD, T.K.; BAIROCH, A.; BATEMAN, A.; BIRNEY, E.; BISWAS, M.; BUCHER, P.; CERUTTI, L.; CORPET, F.; CRONING, M.D.R.; DURBIN, R.; FALQUET, L.; FLEISCHMANN, W.; GOUZY, J.; HERMJAKOB, H.; HULO, N.; JONASSEN, I.; KAHN, D.; KANAPIN, A.; KARAVIDOPOULOU, Y.; LOPEZ, R.; MARX, B.; MULDER, N.J.; OINN, T.M.; PAGNI, M.; SERVANT, F.; SIGRIST, C.J.A.; ZDOBNOV, E.M. InterPro - an integrated documentation resource for protein families, domains and functional sites. **Bioinformatics**, Oxford , v. 16, n. 12, p. 1145-1150, Dec. 2000.

ARIOLI, T; PENG, L.; BETZNER, A. S.; BURN, J.; WITTKE, W.; HERTH, W.; CAMILLERI, C.; HÖFTE, H.; PLANZINSKI, J.; BIRCH, R.; CORK, A.; GLOVER, J.; REDMOND, J.; WILLIAMSON, R. E. Molecular analysis of cellulose biosynthesis in *Arabidopsis*. **Science**, New York, v. 279, n. 5351, p. 717-720, Jan. 1998.

ASHBURNER, M. BALL, C.A.; BLAKE, J.A.; BOTSTEIN, D.; BUTLER, H.; CHERRY, J.M.; DAVIS, A.P.; DOLINSKI, K.; DWIGHT, S.S.; EPPIG, J.T.; HARRIS, M.A.; HILL, D.P.; ISSEL-TARVER, L.; KASARSKIS, A.; LEWIS, S.; MATESE, J.C.; RICHARDSON, J.E.; RINGWALD, M.; RUBIN, G.M.; SHERLOCK, G. Gene Ontology: tool for the unification of biology. **Nature Genetics**, New York, v. 25, n. 1, p. 25-29, May. 2000.

AUBERT, D.; CHEVILLARD, M.; DORNE, A. M.; ARLAUD, G.; HERZOG, M. Expression patterns of GASA genes in *Arabidopsis* thaliana: the GASA4 gene is up-regulated by gibberelins in meristematic regions. **Plant Molecular Biology**, v. 38, p. 871-883, 1998.

BALBACH, J.; SCHMID, F. X. Proline isomerization and its catalysis in protein folding In: PAIN, R. H. **Mechanisms of protein folding**. New York: Oxford University Press, 2000. 249p.

BANERJEE, J.; DAS, N.; DEY, P.; MITI, M. K. Transgenically expressed rice germin-like protein1 in tobacco causes hyper-accumulation of H_2O_2 and reinforcement of the cell wall components. **Biochemical and Biophysical Research Communications**, London, v. 402, n. 04, p. 637-643, Nov. 2010.

BAYER, E. M.; BOTTRILL, A. R.; WALSHAW, J.; VIGOUROUX, M.; NALDRETT, M. J.; THOMAS, C. L.; MAULE, A. J. *Arabidopsis* cell wall proteome defined using multidimensional protein identification technology. **Proteomics**, Weinheim, v. 6, n. 1, p. 301-311, Jan. 2006.

BENDA, G. T. A. Numbering sugarcane leaves and shoots. In: CONGRESS OF THE INTERNATIONAL SOCIETY OF SUGAR CANE TECHNOLOGISTS, 1971, Louisiana. **Annais…** Lousiana: CNS , 1971. p.67-78.

BERNIER, F.; BERNA, A. Germins and germin-like proteins: Plant do-all proteins. But what do they do exactly? **Plant Physiology and Biochemistry**, Paris, v. 39, n. 7-8, p. 545-554, Jul.-Aug. 2001.

BLEE, K. A.; WHEATLEY, E.R. BONHAM, V.A.; MITCHELL, G.P.; ROBERTSON, D.; SLABAS, A.R.; BURRELL, M.M.; WOJTASZEK, P.; BOLWELL, G.P. Proteomic analysis reveals a novel set of cell wall proteins in a transformed tobacco cell culture that synthesises secondary walls as determined by biochemical and morphological parameters **Planta**, New York, v. 212, n. 3, p. 404-415, Feb. 2001.

BORDERIES, G.; JAMET, E.; LAFITTE, C.; ROSSIGNOL, M.; JAUNEAU, A.; BOUDART, G.; MONSARRAT, B.; ESQUERRÉ-TUGAYÉ, M-T.; BOUDET, A.; PONT-LEZICA, R. Proteomics of loosely bound cell wall proteins of *Arabidopsis* thaliana cell suspension cultures: A critical analyses. **Electrophoresis,** Weinheim, v. 24, n. 19-20, p. 3421-3432, Oct. 2003.

BORNER, G. H. H.; LILLEY, K. S.; STEVENS, T. J.; DUPREE, P. Identification of glycosylphosphatidylinositol-ancored proteins in *Arabidopsis*. A proteomic and genomic analysis. **Plant Physiology**, Minneapolis, v. 132, n. 2, p. 568-577, Jun. 2003.

BOUDART, G.; JAMET, E.; ROSSIGNOL, M.; LAFITTE, C.; BORDERIES, G.; JAUNEAU, A.; ESQUERRÉ-TUGAYÉ, M.-T.; PONT-LEZICA, R. Cell wall proteins in apoplastic fluids of *Arabidopsis thaliana* rosettes: Identification by mass spectrometry and bioinformatics. **Proteomics**, Weinheim, v. 5, n. 1, p. 212-221, Jan. 2005.

BRADFORD, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. **Analytical Biochemistry**, New York, v. 72, n. 1-2, p. 248-254, 1976.

BRADLEY, D. J.; KJELBOM, P.; LAMB, C. J. Elicitor-induced and wound-induced oxidative cross-linking of a proline-rich plant-cell wall protein - a novel, rapid, defense response. **Cell**, Cambridge, v. 70, n. 1, p. 21-30, Jul. 1992.

BROWNLEE, C. Role of the extracellular matrix in cell-cell signaling: paracrine paradigms. **Current Opinion in Plant Biology**, London**,** v. 5, n. 5, p. 396-401, Oct. 2002.

BUCKERIDGE, M.S.; SILVA, G.B.; CAVALARI, A.A. Parede Celular. In: KERBAUY, G.B. (Ed). **Fisiologia Vegetal**, 2nd ed. Rio de Janeiro: Guanabara Koogan, 2008. p. 165-181.

CARPITA, N. C.; GIBEAUT, D. M. Structural models of primary-cell walls in flowering plants - consistency of molecular-structure with the physical-properties of the walls during growth. **Plant Journal**, Oxford, v. 3, n. 1, p. 1-30, Jan. 1993.

CARPITA, N. C.; MCCANN, M. The Cell Wall. In: BUCHANAN, B.B.; GRUISSEM, W.; JONES, R. L. (Ed). **Biochemistry & Molecular Biology of Plants**. Rockville: American Society of Plant Physiologists, 2000. v. 2, p. 52-108.

CARPITA, N.; TIERNEY, M.; CAMPBELL, M. Molecular biology of the plant cell wall: searching for the genes that define structure, architecture and dynamics. **Plant Molecular Biology**, Dordrecht,v. 47, n. 1-2, p. 1-5, Jan. 2001.

CASSAB, G. I. Plant cell wall proteins. **Annual Review of Plant Physiology and Plant Molecular Biology**, Palo Alto, v. 49, p. 281-309, Jun. 1998.

CASSAB, G.; VARNER, J. E. Cell wall proteins. **Annual Review of Plant Physiology and Plant Molecular Biology**, Palo Alto, v. 39, p. 321-353, Jun. 1988.

CENTRO NACIONAL DAS INDÚSTRIAS DO SETOR SUCROENERGÉTICO E BIOCOMBUSTÍVEIS – CEISE. Disponível em: <http://www.ceisebr.com/site/index.php/conteudo/visualizar/etanol-governo-anuncia-r-65-bi-ao-setor-de-cana-ate-2015>. Acesso em: 03 mar. 2012.

CHEN, S.; HARMON, A.C.; Advances in Plant Proteomics. **Proteomics**, Weinheim, v. 6, p. 5504-5516, May. 2006.

CHEN, X. Y.; KIM, S. T.; CHO, W. K.; RIM, Y.; KIM, S.; KIM, S. W.; KANG, K. Y.; PARK, Z. Y.; KIM, J. Y. Proteomics of weakly bound cell wall proteins in rice calli. **Journal of Plant Physiology**, Stuttgart, v. 166, n. 7, p. 675-685, May 2009.

CHIVASA, S.; NDIMBA, B. K.; SIMON, W. J.; ROBERTSON, D.; YU, X. L.; KNOX, J. P.; BOLWELL, P. SLABAS, A. R. Proteomic analysis of the *Arabidopsis thaliana* cell wall. **Electrophoresis**, Weinheim, v. 23, n. 11, p. 1754-1765, Jun. 2002.

CHRITIERNIN, M.; Ohlsson, A.B.; Berglund, T.; Henriksson, G. Lignin isolated from primary walls of hybrid aspen cell cultures indicates significant differences in lignin structure between primary and secondary cell wall. **Plant Physiology and Biochemistry**, Oxford, v.43, n.8, p.777-785, Aug 2005.

CONESA, A.; GOTZ, GARCIA-GOMEZ, J. M.; TEROL, J.; TALON, M.; ROBLES, M.; Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. **Bioinformatics**, Oxford, v.21, n. 18, p. 3674-3676, Sep. 2005.

COOPER, D. N. W.; BARONDES, S. H. Evidence for export of a muscle lectin from cytosol to extracellular matrix and for a novel secretory mechanism. **Journal of Cell Biology,** New York, v. 110, n. 5, p. 1681-1691, May 1990.

COSGROVE, D. J. Wall structure and wall losening. A look backwards and forwards. **Plant Physiology,** Minneapolis, v. 125, n. 1, p. 131-134, Jan. 2001.

COSGROVE, D. J. Expansion of the plant cell wall.In: ROSE, J. K. C. (Ed.), **The Plant Cell Wall**. Oxford: Blackwell, 2003. chap. 8, p. 237-263.

CROFTS, A. J.; LEBORGNE-CASTEL, N.; HILLMER, S.; ROBINSON, D.G.; PHILLIPSON, B.; CARLSSON, L.E.; ASHFORD, D.A.; DENECKE, J. Saturation of the endoplasmic reticulum retention machinery reveals anterograde bulk flow. **Plant Cell**, Rockville, v. 11, n. 11, p. 2233-2247, Nov. 1999.

CYSEWSKI, G. R.; WILKE, C. R. Utilization of cellulosic materials through enzymatic hidrolysis. I. Fermentation of hydrolisate to ethanol and single-cell. **Biotechnology and Bioengineering**, New York, v. 18, n. 9, p. 1297-1313, 1976.

DAL DEGAN F.; CHILD, R.; SVENDSEN, I.; ULVSKOV, P. The cleavable N-terminal domain of plant endopolygalacturonases from clade B may be involved in a regulated secretion mechanism. **Journal of Biological Chemistry**, Rockville, v. 276, n. 38, p. 35297-35304, Sep. 2001.

DARLEY, C. P.; FORRESTER, A.; MCQUEEN-MASON, S. The molecular basis of plant cell wall extension. **Plant Molecular Biology**, Dordrecht, v. 47, n. 1-2, p. 179-195, 2001.

DAVIES, E.; STANKOVIC, B.; AZAMA, K.; SHIBATA, K.; ABE, S. Novel components of the plant cytoskeleton: a beginning to plant “cytomics”. **Plant Science**. Oxford, v. 160, n. 2, 2001, 160, p. 185-196, Jan. 2001.

DAVIN, L. B.; LEWIS, N. G. Lignin primary structures and dirigent sites. **Current Opinion Biotechnology**. London, v. 16, p. 407-415, 2005.

DE CAROLI, M.; LENUCCI, M. S.; DI SANSEBASTIANO, G.P.; DALESSANDRO, G.; DE LORENZO, G.; PIRO, G. Protein trafficking to the cell wall occurs through mechanisms distinguishable from default sorting in tobacco. **Plant Journal**, Malden, v. 65, n. 2, p. 295-308, Jan 2011.

DE LORENZO, G.; FERRARI, S. Polygalacturonase-inhibiting proteins in defense against phytopathogenic fungi. **Current Opinion in Plant Biology**, London, v. 5, n. 4, p. 295-299, Aug. 2002.

DILLEWJIN, C. V. **Botany of sugarcane**. Waltham: The Chronica Botanica Co. 1952. 371p.

DUNWELL, J. M.; CULHAM, A.; CARTER, C. E.; SOSA-AGUIRRE, C.; GOODENOUGH, C. W. Evolution of functional diverse cupin superfamily. **Trends in Biochemical Science**, London, v. 26, n. 12, p. 740-746, Dec. 2001.

EDWARDS, S. R.; BRALEY, R.; CHAFFIN, W. L. Enolase id present in the cell wall of Saccharomyces cerevisiae. **FEMS Microbioogy Letters**, v. 177, n. 2, p. 211-216, Aug. 1999.

ELLIS, C. KARAFYLLIDIS, I.; WASTERNACK, C.; TURNER, J. G. The *Arabidopsis* mutant *cev1* links cell wall signaling to jasmonate and ethylene responses. **Plant Cell**, Rockville, v. 14, n. 7, p. 1557-1566, Jul 2002.

EMANUELSSON, O.; NIELSEN, H.; BRUNAK, S.; VON HEIJNE, G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. **Journal of Molecular Biology**, London, v. 300, n. 4, p. 1005-1016, Jul. 2000.

EMANUELSSON, O.; BRUNAK, S.; VON HEIJNE, G.; NIELSEN, H. Locating proteins in the cell using TargetP, SignalP and related tools. **Nature Protocols**, London, v. 2, n. 4, p. 953-971, Nov. 2007.

EMPRESA BRASILEIRA DE AGROPECUÁRIA. Disponível em: <www.cnpae.embrapa.br>. Acesso em: 20 mar. 2011.

FALCO, M. C. **Cultura de células em suspensão e protoplastos em cana-de-açúcar.** 1994. 131 p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1994.

FALCO, M. C. Cell suspension culture of sugarcane: growth, management and plant regeneration. **Revista Brasileira de Fisiologia Vegetal**, v. 8, n. 1, p.1-6, 1996.

FALCO, M. C. **Transformação de cana-de-açúcar para resistência ao herbicida glufosinato de amônio, via biolística.** 1998. 129 p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1998.

FEIZ, L.; IRSHAD, M.; PONT-LEZICA, R.F.; CANUT, H.; JAMET, E. Evaluation of cell wall preparations for proteomics: a new procedure for purifying cell walls from *Arabidopsis* hypocotyls. **Plant Methods**, London, v. 2, p. 10-23, 2006.

FISCHER, G.; TEIXEIRA, E.; HIZSNYIK, E. T.; VELTHUIZEN, H. Land use dynamics and sugarcane production. In: ZUURBIER, P.; VAN DE VOOREN, J. (Ed.), **Sugarcane ethanol: contribution to climate change mitigation and the environment**. Wageningen: Wageningen Academic, 2008. chap. 2, p. 29-62.

FOOD AND AGRICULTURE ORGANIZATION – FAO. Disponível em: <http://www.fao.org/docrep/013/i2050e/i2050e00.htm>. Acesso em: 10 dez. 2008.

FORESTI, O.; DENECKE, J. Intermediate organelles of the plant secretory pathway: Identity and function. **Traffic**, Malden, v. 9, n. 10, p. 1599-1612, Oct. 2008.

FRESHOUR, G.; CLAY, R.P.; FULLER, M.S.; ALBERSHEIM, P.; DARVILL, A.G.; HAHN, M.G. Developmental and tissue-specific structural alterations of the cell-wall polysaccharides of *Arabidopsis thaliana* roots. **Plant Physiology**, Rockville, v. 110, n. 4, p. 1413-1429, Apr. 1996.

FRY, S. C.; SMITH, R. C.; RENWICK, K. F.; MARTIN, D. J.; HODGE, S. K.; MATTHEWS, K.. J. Xyloglucan endotransglycosylase, a new wall-loosening enzyme-activity from plants. **Biochemical Journal**, London, v. 282, p. 821-828, Mar 1992.

FRY, S. C. Polysaccharide-modifying enzymes in the plant cell wall. **Annual Review of Plant Physiology and Plant Molecular Biology**, Palo Alto, v. 46, p. 497-520, 1995.

FRY, S. C. Primary cell wall metabolism: tracking the carrers of the wall polymers in living plant cells. **New Phytologist,** Oxford, v. 161, n. 3, p. 641-675, Mar. 2004.

GIRHEPUJE, P. V.; SHINDE, G. B. Transgenic tomato plants expressing a wheat endochitinase gene demonstrate enhanced resistance to Fusarium oxysporum f. sp lycopersici. **Plant Cell Tissue and Organ Culture,** Dordrecht, v. 105, n. 2, p. 243-251, May. 2011.

GOLDEMBERG, J.; COELHO, S. T.; GUARDABASSI, P. The sustainability of ethanol production in sugarcane. **Science**, Oxford, v. 36, n.6, p. 2086-2097, Jun. 2008.

GONZÁLEZ-GARCÍA, S.; LUO, L.; MOREIRA, M. T.; FEIJOO, G.; HUPPES, G. Life cycle assessment of flax shives derived second generation ethanol fueled automobiles in Spain. **Renewable and Sustainable Energy Reviews**, Oxford, v. 13, p. 1922-1933, Oct. 2009.

GLASZIOU, K. T.; GAYLER, K. R. Sugar accumulation in sugarcane-role of cell walls in sucrose transport. **Plant Physiology**, Minneapolis, v. 49, n. 6, p. 912-921, 1972.

HAILING, P.; SIMMS-BORRE, P. Overview of lignocellulosic feedstock conversion into ethanol-focus on sugarcane bagasse. **International Sugar Journal**, Glamorgan, v. 110, n. 1311, p. 191-202, Mar 2008.

GOZALBO, D.; GIL-NAVARRO, I.; AZORIN, I.; RENAU-PEQUERAS, J.; MARTINEZ, J. P.; GIL, M. L. **The cell wall-associated glyceraldehyde-3-phosphate dehydrogenase of Candida albicans is also a fibronectin and laminin binding protein**. **Infection and Immunity**, v. 66, n. 5, p. 2052-2059, May, 1998.

HE, Z. H.; FUJIKI, M.; KOHORN, B. D. A cell wall-associated, receptor-like protein kinase. **Journal of Biological Chemistry**, Rockville, v. 271, n. 33, p. 19789-19793, Aug 1996.

HERRERA-ESTRELLA, L.; RUIZ-MEDRANO, R.; JIMENEZ-MORAILA, B.; RIVERA-BUSTAMANTE, R. F. Nucleotide sequence of an osmotin-like cDNA induced in tomato during viroid infection. **Plant Molecular Biology**, London, v. 20, n. 6, p. 1199–1202. Jan. 1992.

HERRMANN, K.; ENTUS, R. Skikimate pathway: Aromatic Amino Acids and Beyond. **Encyclopedia of Life Sciences**, London, v. 1, p. 99-110, Nov. 2001.

HILL, J.; POLASKY, S.; NELSON, E.; TILMAN, D.; HUO, H.; LUDWIG, L.; NEUMANN, J.; ZHENG, H.; BONTA, D. Climate change and health costs of air emissions from biofuels and gasoline. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 106, n. 6, p. 2077-2082, Feb. 2009.

HORTON, P.; PARK, K. J.; OBAYASHI, T.; FUJITA, N.; HARADA, H.; ADAMS-COLLIER, C. J.; NAKAI, K. WoLF PSORT: protein localization predictor. **Nucleic Acids Rese**arch, Oxford, v. 35, p. W585-W587, Jul. 2007.

HOSON, T. Apoplast as the site of response to environmental signals. **Journal of Plant Research**, Tokyo, v. 111, n. 1101, p. 167-177, Mar. 1998.

HUBERTS, D.E.W.; VAN DER KLEI, I.J. Moonlighting proteins: An intriguing mode of multitasking. **Biochimica et Biophysica Acta**, Amsterdam, v. 1803, n. 4, p. 520-525, Apr. 2010.

INMAN-BAMBER, N. G.; BONNETT, G.D.; SMITH, D.M.; THORBURN, P. J. Sugarcane physiology: Integrating from cell to crop to advance sugarcane production. **Field Crops Research**, Amsterdam, v. 92, n. 2-3, p. 115-117, Jun. 2005.

IRSHAD, M.; CANUT, H.; BORDERIES, G.; PONT-LEZICA, R.; JAMET, E. A new picture of cell wall protein dynamics in elongating cells of *Arabidopsis thaliana*: Confirmed actors and newcomers. **Bmc Plant Biology**, London, v. 8, p. 16-32, Sep. 2008.

IRVINE, J. E. Saccharum species as horticultural classes. **Theoretical and Applied Genetics**, Berlin, v. 98, n. 2, p. 186-194, Feb. 1999.

JAMET, E.; CANUT, H.; BOUDART, G.; PONT-LEZICA; R. F. Cell wall proteins: a new insight through proteomics. **Trends in Plant Science**, Oxford, v. 11, n. 1, p. 33-39, Jan. 2006.

JAMET, E.; CANUT, H.; ALBENNE, C.; BOUDART, G.; PONT-LEZICA, R. Cell wall. In: Plant Proteomics: Technologies, Strategies, and Applications. G.K. Agrawal and R. Rakwal, eds. John Wiley & Sons, Inc. pp. 293-307 (2008).

JAMET, E.; ALBENNE, C.; BOUDART, G.; IRSHAD, M.; CANUT, H.; PONT-LEZICA, R. Recent Advances in plant cell wall proteomics. **Proteomics**, Weinheim, v. 8, n. 4, p. 893-908, Feb 2008.

JAMET, E.; PONT-LEZICA, R. The puzzle of protein location in plant proteomics. In: RANCOURT, G., (Ed), **Proteomics:** Methods, Applications and Limitations. Hauppauge: Nova Science Publishers, 2010. chap. 1, p. 1-13.

KADER, J.-C. Lipid transfer proteins: a puzzling family of plant proteins. **Trends in Plant Science**, v. 2, p. 66-70, 1997.

KIM, S.; MOLLET, J.C.; DONG, J.; ZHANG, K.L.; PARK, S.Y.; LORD, E.M. Chemocyanin, a small basic protein from the lily stigma, induces pollen tube chemotropism. **Proceedings of the National Academy of Sciences of the United States of America,** Washington, v. 100, n. 26, p. 16125-16130, Dec 2003.

KNAUF, M.; MONIRUZZAMAN, M. Lignocellulosic biomass processing: a perspective. **International Sugar Journal**, Glamorgan, v. 106, n. 1263, p. 147-150, Mar 2004.

KONG, F.-J.; OYANAGI, A.; KOMATSU, S. Cell wall proteome of wheat roots under flooding stress using gel-based and LC MS/MS-based proteomics approaches. **Biochimica et Biophysica Acta**, Amsterdam, v. 1804, p.124-136, 2010.

KROGH, A.; LARSSON, B.; VON HEIJNE, G.; SONNHAMMER, E. L. L. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. **Journal of Molecular Biology**, London, v. 305, n. 3, p. 567-580, Jan 2001.

LAEMMLI, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature**, London, v. 227, n. 5259, p. 680-693, 1970.

LEE, S-J.; SARAVANAN, R. S.; DAMASCENO, C. M. B.; YAMANE, H.; KIM, B-D; ROSE, J. K. C. Digging deeper into cell wall proteome. **Plant Physiology and Biochemistry**, Paris, v. 42, n. 12, p. 979-988, Dec 2004.

LEE, H.; DAMSZ, B.; WOLOSHUK, C. P.; BRESSAN, R. A.; NARASIMHAN, M. L. Use of the Plant Defense Protein Osmotin To Identify Fusarium oxysporum Genes That Control Cell Wall Properties. **Eukaryotic Cell**, Washington, v. 9, n. 4, p. 558-568, Apr 2010.

LEITE, R. C. D.; LEAL, M.R.L.V.; CORTEZ, L.A.B.; GRIFFIN, W.M.; SCANDIFFIO, M.I.G. Can Brazil replace 5% of the 2025 gasoline world demand with ethanol? **Energy**, Oxford, v. 34, n. 5, p. 655-661, May 2009.

LIMA, D. U.; SANTOS, H. P.; TINÉ, M. A.; MOLLE, F. D.; BUCKERIDGE, M. S. Patterns of expression of cell wall related genes in sugarcane. **Genetics and Molecular Biology**, Ribeirão Preto, v. 24, n. 1-4, p. 191-198, Mar. 2001.

LOVE, G. D.; SNAPE, C. E.; JARVIS, M. C. Comparison of leaf and stem cell-wall components in barley straw by solid-state C-13 NMR. **Phytochemistry**, New York , v. 49, n. 5, p. 1191-1194, Nov. 1998.

LUAH, Y. H.; CHAAL, B.K.; ONG, E.Z.; BOZDECH, Z. A Moonlighting Function of Plasmodium falciparum Histone 3, Mono-Methylated at Lysine 9? **Plos One**, v. 5, n. 4, APR. 2010.

MACEDO, I. C.; SEABRA, J. E. A.; SILVA, J. E. A. R. Green house gases emissions in the production and use of ethanol from sugarcane in Brazil: The 2005/2006 averages and a prediction for 2020. **Biomass & Bioenergy**. Oxford, v. 32, n.7, p. 582-595. Jul. 2008.

MARQUES, F. O alvo é o bagaço. **Revista Pesquisa Fapesp**. São Paulo, v. 11, n.163, p. 30-33, sep. 2009.

MARTÍN, C.; THOMSEN, A. B. Wet oxidation pretreatment of lignocellulosic residues of sugarcane, rice, cassava and peanuts for ethanol production. **Journal of Chemical Technology and Biotechnology**, Oxford, v. 82, n. 2, p. 174-181, Feb. 2007.

MARTINELLI, L. A.; FILOSO, S. Expansion of sugarcane ethanol production in Brazil: environmental and social challenges. **Ecological Applications**, Tempe, v. 18, n. 4, p. 885-898, Jun. 2008.

MARTINEZ, M.; CAMBRA, I.; GONZALEZ-MELENDI, P.; SANTAMARIA, M.E.; DIAZ, I. C1A cysteine-proteases and their inhibitors in plants. **Physiologia Plantarum**, Malden, v. 145, n. 1, p. 85-94, May 2012.

MASCOT SOFTWARE. Disponível em: <www.matrixscience.com>. Acesso em: 20 set. 2012.

MATSUOKA, S.; FERRO, J.; ARRUDA, P. The Brazilian Experience of Sugarcane Ethanol Industry. **In Vitro Cellular & Developmental Biology-Plant**, Columbia, v. 45, n. 3, p. 372-381, Jun. 2009.

MENG, L.; WONH, J. H.; FELDMAN, L. J.; PEGGY, L. G. BUCHANAN, B. B. A membrane-associated thioredoxin required for plant growth moves from cell to cell, suggestive of a role in intercellular communication. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 107, n. 08, p. 227-238, Dec. 2009.

MCGEOH, D. J. On the predictive recognition of signal peptide sequences. **Virus Research**, Amsterdam, v.3, n.3, p. 271-286. 1985.

MCQUEEN-MASON, S.; DURACHKO, D. M.; COSGROVE, D. J. Two endogenous proteins that induce cell wall expansion in plants. **Plant Cell**, Rockville, v. 4, p. 1425-1433, 1992.

MELLO, M. O.; TANAKA, A. S.; SILVA-FILHO, M. C. Molecular evolution of Bowman-Birk type proteinase inhibitors in flowering plants. **Molecular Phylogenetics and Evolution**, Orlando, v. 27, n. 1, p. 103-112, Apr. 2003.

MILLAR, D. J.; WHITELEGGE, J. P.; BINDSCHEDLER, L. V.; RAYON, C.; BOUDET, A.-M.; ROSSIGNOL, M.; BORDERIES, G.; BOLWELL, G. P. The cell wall and secretory proteome of a tobacco cell line synthesising secondary wall. **Proteomics**, Weinheim, v. 9, n. 9, p. 2355-2372, May 2009.

MINIC, Z.; JAMET, E.; NEGRONI, L.; DER GARABEDIAN, P. A. A subproteome of *Arabidopsis* thaliana trapped on Concanavalian A is enriched in cell wall glycoside hydrolases**. Journal of Experimental Botany**, v. 58, n. 10, p. 2503-2512, Feb. 2007.

MOLLER, S.; CRONING, M. D. R.; APWEILER, R. Evaluation of methods for the prediction of membrane spanning regions. **Bioinformatics**, Oxford, v. 17, n. 7, p. 646-653, Jul 2001.

MOORE, P. J.; STAFSTROM, J. P.; STAEHELIN, L. A. Immunocytochemical localization of xyloglucan rhamnogalacturonan-I and extensin in plant cell walls. **Journal of Cellular Biochemistry**, Atlanta, v. 10C, p.17-17 1986.

MURASHIGE, T.; SKOOG, F. A revised medium for rapid growth and bioassays with tobacco tissue cultures. **Plant Physiology**, Minneapolis, v. 15, n. 3, p. 473-497, 1962.

NAKAI, K.; KANEHISA, M. Expert System for Predicting Protein Localization Sites in Gram-Negative Bacteria. **Proteins-Structure Function and Genetics**, New York, v. 11, n. 2, p. 95-110, 1991.

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION –NCBI. Disponível em: <http://www.ncbi.nlm.nih.gov/>. Acesso em: 05 set. 2012.

NEGRI, A. S.; PRINSI, B.; SCIENZA, A.; MORGUTTI, S.; COCUCCI, M.; ESPEN, L. Analysis of grape berry cell wall proteome: A comparative evaluation of extraction methods. **Journal of Plant Physiology**, Stuttgart, v. 165, n. 13, p. 1379-1389, Jan. 2008.

NEMIR, A. S. Alcohol fuels-the Brazilian experience and its implications for the United States. **Sugar Journal**, Glamorgan, v. 45, n. 12, p. 10-13, 1983.

NERSISSIAN, A. M.; SHIPP, E. L. Blue copper-binding domains. Adv. Protein Chem. 2002, 60, 271-340.
NICOL, F.; HOFTE, H. Plant cell expansion: scaling the wall. **Current Opinion in Plant Biology**, London, v. 1, n. 1, p. 12-17, Feb. 1998.

NIELSEN, M. N.; SORENSEN, J.; FELS, J.; PEDERSEN, H.C. Secondary metabolite- and endochitinase-dependent antagonism toward plant-pathogenic microfungi of Pseudomonas fluorescens isolates from sugar beet rhizosphere. **Applied and Environmental Microbiology**, Washington, v. 64, n. 10, p. 3563-3569, Oct. 1998.

NIEUWLAND, J.; FERON, R.; HUISMAN, B.A.H.; FASOLINO, A.; HILBERS, C.W.; DERKSEN, J.; MARIANI, C. Lipid transfer proteins enhance cell wall extension in tobacco. **Plant Cell**, Rockville, v. 17, n. 7, p. 2009-2019, Jul. 2005.

OLIVÉRIO, J.L.; PROENÇA, A.G.H. DHR – Dedini Hidrólise Rápida: revolucionary process for producing alcohol from sugar cane bagasse". **International Sugar Journal**, Glamorgan, v. 106, p. 168-172, 2004.

OLIVÉRIO, J. L. Tecnological evolution of the Brazilian Sugar and Alcohol Sector: Dedini´s contribution". **International Sugar Jornal**, Glamorgan, v. 4, p. 101-108, 2006.

PEDREIRA, J.; SANZ, N.; PENA, M. J.; SANCHEZ, M.; QUEJEIRO, E.; REVILLA, G.; ZARRA, I. Role of the apoplastic ascorbate and hydrogen peroxide in the control of cell growth in pine hypocotyls. **Plant and Cell Physiology**, Kyoto, v. 45, n. 5, p. 530-534, Sep. 2004.

PENNELL, R. Cell walls: structures and signals. **Current Opinion in Plant Biology**, London, v. 1, n. 6, p. 504-510, Dec. 1998.

PESSOA-JR, A.; CONCEIÇÃO, I. R.; MENOSSI, M.; DOS SANTOS, R. R; FILHO, S. O.; PENNA, T. C. V. Perspectives on bioenergy and biotechnology in Brazil. **Biochemistry and Biotechnology**, Duluth, v. 121, p. 59-70, Spr 2005.

PETERSEN, T. N.; BRUNAK, S.; VON HEIJNE, G.; NIELSEN, H. SignalP 4.0: discriminating signal peptides from transmembrane regions, **Nature Methods**, New York, v. 8, p. 785-786, Oct 2011.

PFAM DATABASE – PFAM. Disponível em: <http://pfam.sanger.ac.uk/>. Acesso em: 24 set. 2012.

PREDICTION OF SORTING SIGNALS AND LOCALIZATION SITES IN AMINO ACIDS SEQUENCES – PSORT. Disponível em: http://psort.hgc.jp/. Acesso em: 24 set. 2012.

PREDICTION OF TRANSMEMBRANE HELICES IN PROTEINS – TMHMM. Disponível em: <http://www.cbs.dtu.dk/services/TMHMM/>. Acesso em: 16 jul. 2011.

PREGER, V.; TANGO, N.; MARCHAND, C.; LEMAIRE, S.D.; CARBONERA, D.; DI VALENTIN, M.; COSTA, A.; PUPILLO, P.; TROST, P. Auxin-Responsive Genes AIR12 Code for a New Family of Plasma Membrane b-Type Cytochromes Specific to Flowering Plants. **Plant Physiology**, Minneapolis, v. 150, n. 2, p. 606-620, Jun 2009.

PRESSEY, R. ROLE OF PECTINESTERASE IN PH-DEPENDENT INTERACTIONS BETWEEN PEA CELL-WALL POLYMERS. **Plant Physiology**, Minneapolis, v. 76, n. 2, p. 547-549, 1984.

PRICE, N. J.; PINHEIRO, C.; SOARES, C. M.; ASHFORD, D. A.; RICARDO, C. P.; JACKSON, P. A. A biochemical and molecular characterization of LEP1, an extensin peroxidase from lupin. **Journal of Biological Chemistry**, Rockville, v. 278, n. 42, p. 41389-41399, Oct 2003.

PROJETO GENOMA DA CANA. Disponível em: <http://cnrgv.toulouse.inra.fr/Projects/The-Sugarcane-Genome-Sequencing-Initiative-SUGESI-Strategies-for-Sequencing-a-Highly-Complex-Genome>.Acesso em: 20 mar. 2011.

RAPOPORT, T. A. Transport of proteins across the endoplasmic reticulum membrane. **Science**, Washington, v. 258, n. 5084, p. 931-936, Nov. 1992.

RAPOPORT, T. A. Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. **Nature**, London, v. 450, n. 7170, p. 663-669, Nov. 2007.

REITER, W.D. The molecular analysis of cell wall components. **Trends in Plant Science**, Oxford, v.3, n.1, p. 27-32. Jan. 1998.

ROBERTS, K. The plant extracellular matrix. **Current Opinion in Cell Biology**, London, v. 1, p. 1020-1027, Oct. 1994.

ROBERTSON, D.; MITCHELL, G. P.; GILROY, J. S.; GERRISH, C.; BOLWELL, G. P.; SLABAS, A. R. Differential extraction and protein sequencing reveals major differences in patterns of primary cell wall proteins from plants. **Journal of Biological Chemistry**, Rockville, v. 272, n. 25, p. 15841-15848, Jun. 1997.

ROBERTSON, L.; ROBERTSON, W. M.; SOBCZAK, M.; HELDER, J.; TETAUD, E.; ARIYANAYAGAM, M. R.; FERGUSON, M. A.; FAIRLAMB, A.; JONES, J. T.; Mol. Biochem. Parasitol. 2000, 111, 41-49.

ROSE, J. K. C.; LEE, S. J. Straying off the Highway: Trafficking of Secreted Plant Proteins and Complexity in the Plant Cell Wall Proteome. **Plant Physiology**, Minneapolis, v. 153, n. 2, p. 433-436, Jun 2010.

RUBIN, E. M. Genomics of cellulosic biofuels. **Nature**, London, v. 454, n. 14, p. 841-845, Aug. 2008.

RUINARD, J. Varietal differences in leaf-spindle structure and growth distribution in leaves and stems of sugarcane. **Annals of Botany**, London, v. 30, n. 117, p. 1-&, 1966.

SARRY, J.E.; KUHN, L.; LE LAY, P.; GARIN, J.; BOURGUIGNON, J. Dynamics of Arabidopsis thaliana soluble proteome in response to different nutrient culture conditions. **ELlectrophoresis**, New York, v. 27, p. 495-507, Feb. 2006.

SAN CLEMENTE, H.; PONT-LEZICA, R.; JAMET, E. Bioinformatics as a tool for assessing the quality of sub-cellular proteomic strategies and inferring functions of proteins: plant cell wall proteomics as a test case. **Bioinform Biol Insights**, London, v. 3, p. 15-28, Mar. 2009.

SCHALLER, A. A cut above the rest: the regulatory function of plant proteases. **Planta**, Berlin, v. 220, n. 2, p. 183-197, Dec. 2004.

SEGARRA, C. I.; CASALONGUE, C.A.; PINEDO, M.L.; RONCHI, V.P.; CONDE, R.D. A germin-like protein of wheat leaf apoplast inhibits serine proteases. **Journal of Experimental Botany**, Oxford, v. 54, n. 386, p. 1335-1341, May 2003.

SERVICE, R. F. Is there a road ahead for cellulosic ethanol? **Science**, Washington, v. 329, n. 5993, p. 784-785, Aug. 2010.

SHANI, Z.; DEKEL, M.; ROIZ, L.; HOROWITZ, M.; KOLOSOVSKI, N.; LAPIDOT, S.; ALKAN, S.; KOLTAI, H.; TSABARY, G.; GOREN, R.; SHOSEYOV, O. Expression of endo-1,4-beta-glucanase (cel1) in *Arabidopsis thaliana* is associated with plant growth, xylem development and cell wall thickening. **Plant Cell Reports**, New York, v. 25, n. 10, p. 1067-1074, Oct. 2006.

SHARMA, R.; JAIN, M. BHATNAGAR, R.N.; BHALLA-SARIN, N. Differential expression of DAHP synthase and chorismate mutase in various organs of *Brassica juncea* and the effect of external factors on enzyme activity. **Physiologia Plantarum**, Malden, v. 105, n. 4, p. 739-745, Jan. 2002.

SIQUEIRA-JUNIOR, C. L.; FERNANDES, K.V.S.; MACHADO, O.L.T.; DA CUNHA, M.; GOMES, V.M.; MOURA, D.; JACINTO, T. 87 kDa tomato cystatin exhibits properties of a

defense protein and forms protein crystals in prosystemin overexpressing transgenic plants. **Plant Physiology and Biochemistry**, Paris, v. 40, n. 3, p. 247-254, Mar. 2002.

SLABAS, A. R.; NDIMBA, B.; SIMON, W. J.; CHIVASA, S. Proteomic analysis of the *Arabidopsis* cell wall reveals unexpected proteins with new cellular locations. **Biochemical Society Transactions**, London, v. 32, n. 3, p. 524-528, Jun. 2004.

SMALL, I.; PEETERS, N.; LEGEAI, F.; LURIN, C. Predotar: a tool for rapidly screening proteomes for N-terminal targeting sequences. **Proteomics**, Weinheim, v. 4, n. 6, p. 1581-1590, Jun. 2004.

SOARES, P. A.; ROSSEL, C. E. V. **NAIPPE:** o setor sucroalcooleiro e o domínio tecnológico. Rio de Janeiro: Nova Série, 2007. 107 p.

SOMMERVILLE, C. S. The billion-ton biofuels vision. **Science**, Washington, v. 2, p. 312-1277, Jan. 2006.

SONNHAMMER, E. L. L.; VON HEIJNE, G.; KROGH, A. A hidden Markov model for predicting transmembrane helices in protein sequences. In: INTERNATIONAL CONFERENCE ON INTELLIGENT SYSTEMS FOR MOLECULAR BIOLOGY, 1998, Menlo Park. **Anais…**Menlo Park: AAAI Press, 1998. p. 175-182.

STICKLEN, M.B. Plant Genetic Engineering for biofuel production: towards affordable cellulosic ethanol. **Nature Reviews Genetics**, London, v. 9, p. 433-443, Jul. 2008.

SUGARCANE GENOME INITIATIVE – SUGESI. Disponível em: < http://cnrgv.toulouse.inra.fr/fr/Projets/Canne-a-sucre/The-Sugarcane-Genome-Sequencing-Initiative-SUGESI-Strategies-for-Sequencing-a-Highly-Complex-Genome>. Acesso em nov. 2012.

SUH, M. C.; SAMUELS, A. L.; JETTER, R.; KUNST, L.; POLLARD, M.; OHLROGGE, J.; BEISSON, F. Cuticular lipid composition, surface structure, and gene expression in *Arabidopsis* stem epidermis. **Plant Physiology**, Minneapolis, v. 139, p. 1649-1655, 2005.

TAIZ, L.; ZEIGER, E. **Fisiologia Vegetal**. Porto Alegre: Artmed, 2004. 364 p.

TAKAHASHI, K.; NIWA, H.; YOKOTA, N.; KUBOTA, K.; INOUE, H. Widespread tissue expression of nepenthesin-like aspartic protease genes in *Arabidopsis thaliana*. **Plant Physiology and Biochemistry**, Paris, v. 46, n. 7, p. 724-729, Jul. 2008.

TANAKA, H.; ONOUCHI, H.; KONDO, M.; HARA-NISHIMURA, I.; NISHIMURA, M.; MACHIDA, C.; MACHIDA, Y. A subtilisin-like serine protease is required for epidermal surface formation in *Arabidopsis* embryos and juvenile plants. **Development**, Cambridge, v. 128, n. 23, p. 4681-4689, Dec. 2001.

THE SUGARCANE EST PROJECT – SUCEST. Disponível em: <http://sucest-fun.org/>. Acesso em: 07 ago. 2012.

TOMPA, P.; SZASZ, C.; BUDAY, L. Structural disorder throws new light on moonlighting. **Trends in Biochemical Sciences**, New York, v. 30, n. 9, p. 484-489, Sep. 2005.

TUNIO, S.A.; OLDFIELD, N.J.; BERRY, A.; ALA'ALDEEN, D.A.A.; WOOLDRIDGE, K.G.; TURNER, D.P.J. The moonlighting protein fructose-1, 6-bisphosphate aldolase of Neisseria meningitidis: surface localization and role in host cell adhesion. **Molecular Microbiology**, New York, v. 76, n. 3, p. 605-615, May. 2010.

UNIÃO DOS PRODUTORES DE CANA-DE-AÇÚCAR. Disponível em: <http://www.unicadata.com.br/historico-de-producao-e-moagem.php>. Acesso em: 24 set. 2012.

VETTORE, A. L.; DA SILVA, F. R.; KEMPER, E. L.; ARRUDA, P. The libraries that made SUCEST. **Genetics and Molecular Biology**, Ribeirão Preto, v. 24, n. 1-4, p. 1-7, Mar. 2001.

VETTORE, A. L. Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. **Genome Research**, Cold Spring Harbor, v. 13, n. 12, p. 2725-2735, Dec 2003.

VITALE, A.; DENECKE, J. The endoplasmic reticulum - Gateway of the secretory pathway. **Plant Cell**, Rockville, v. 11, n. 4, p. 615-628, Apr. 1999.

VOGEL, J. P. RAAB, T.K., SOMERVILLE, S.C. Mutations in PMR5 result in powdery mildew resistance and altered cell wall composition. **Plant Journal**, Oxford, v. 40, p. 968-978, Sep. 2004.

VAN LOON, L.C.; REP, M.; PIETERSE, C.M. Significance of inducible defense-related proteins in infected plants. **Annual Review of Phytopathology**, New York, v. 44, p.135-162, Apr. 1999.

VON HEIJNE, G. A new method for predicting signal sequence cleavage sites. **Nucleic Acids Research**, Oxford, v. 14, n. 11, p. 4683-4690, 1986.

VON GROLL, U.; BERGER, D.; ALTMANN, T. The subtilisin-like serine protease SDD1 mediates cell-to-cell signaling during *Arabidopsis* stomatal development. **Plant Cell**, Rockville, v. 14, p. 1527-1539, 2002.

XIAO, G.; CHUNG, T.-F.; PYUN, H. Y.; FINE, R. E.; JOHNSON, R. J. Calreticulin is transported to the surface of NG108-15 cells where it forms surface patches and is partially degraded in an acidic compartment . **Molecular Brain Research**, Canda, v. 72, P. 121-128. Dec.1999.

WANG, S. B.; CHEN, F.; SOMMERFELD, M.; HU, Q. Proteomic analysis of molecular response to oxidative stress by the green alga Haematococcus pluvialis (Chlorophyceae). **Planta**, Berlin, v. 220, n. 1, p. 17-29, 2004.

WATERS. Disponível em: <http://www.waters.com/waters/home.htm>. Acesso em: 03 mar. 2012.

WATSON, B. S.; LEI, Z.T.; DIXON, R.A.; SUMNER, L.W. Proteomics of Medicago sativa cell walls. **Phytochemistry**, New York, v. 65, n. 12, p. 1709-1720, Jun. 2004.

WEGELE, H.; MULLER, L.; BUCHNER, J. Hsp70 and Hsp90 – A relay team for protein folding. Reviews of physiology, biochemistry and pharmacology. **Reviews of Physiology, Biochemistry and Pharmacology**, New York, v. 151, p. 1-44, 2004.

WELBAUM, G. E.; MEINZER, F. C. Compartimentation of solutes and water in developing sugarcane stalk tissue. **Plant Physiology**, Minneapolis, v. 93, n. 3, p. 1147-1153, Jul. 1990.

WILHELMI, L.; PREUSS, D. Pollen tube guidance in *Arabidopsis*: Pop2 and pop3 gene function in cell-autonomous gamete behavior. **Molecular Biology of the Cell**, Bethesda, v. 9, p. 402A-402A, Nov. 1998.

WOJTASZEK, P.; TRETHOWAN, J.; BOLWELL, G. P. Specificity of cell-wall proteins in response to different elicitor molecules in suspension-cultured cells of French bean (Phaseolus vulgaris L.). **Plant Molecular Biology**, Dordrecht, v. 28, n. 6, p. 1075-1087, Sep. 1995.

WOJTASZEK, P. Oxidative burst: an early plant response to pathogen infection. **Biochemical Journal**, London, v. 322, p. 681-692, Mar. 1997.

WOLF, S.; RAUSCH, T.; GREINER, S. The N-terminal pro region mediates retention of unprocessed type-I PME in the Golgi apparatus. **Plant Journal**, Oxford, v. 58, n. 3, p. 361-375, May 2009.

YAMAOKA, T.; CHIBA, N. Changes in the coagulating ability of pectin during the growth of soybean hypocotyls. **Plant and Cell Physiology**, Kyoto, v. 24, n. 7, p. 1281-1290, 1983.

YATES, J.R.; RUSE, C.I.; NAKORCHEVSKY, A. Proteomics by Mass Spectrometry: Approaches, Advances, and Applications**. Annual Review of Biomedical Engineering,** New York, v. 11, p. 49-79, Apr. 2009.

YI-HSIEN, L.; HSIANG-EN, H.; FANG-SHENG, W. Plant-ferredoxin like protein (PFLP) outside chloroplast in Arabidopsis enhances disease resistance against bacterial pathogens. **Plant Science**, Oxford, v. 179, n.5, p.450-458, Nov. 2010.

YUAN, J. S.; TILLER, K. H.; AL-AHMAD, H.; STEWART, N. R.; STEWART, C. N. Plants to power: Bioenergy to fuel the future. **Trends in Plant Science**, Oxford, v. 13, n. 8, p. 421-429, Aug. 2008.

ZHU, J. K.; DAMSZ, B.; KONONOWICZ, A. K.; BRESSAN, R.A.; HASEGAWA, P.M. A higher plant extracellular vitronectin-like adhesion protein is related to the translational elongation factor-1-alpha. **Plant Cell**, Rockville, v. 6, n. 3, p. 393-404, Mar. 1994.

ZHU, J.; CHEN, S.; ALVAREZ, S.; ASIRVATHAM, V. S.; SCHACHTMAN, D. P.; WU, Y.; SHARP, R. E. Cell wall proteome in the maize primary root elongation zone. I. Extraction and identification of water-soluble and lightly ionically bound proteins. **Plant Physiology**, Minneapolis, v. 140, n. 1, p. 311-325, Jan. 2006.

APPENDICES

APPENDIX A - CELL WALL PROTEINS FROM SUGARCANE CELLS

Table 5 – Identified ESTs and proteins in sugarcane cultured cells found only in two repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCCCLR1077B12	pdil2-2 - zea mays protein disulfide isomerase	XP_002455594	Miscellaneous	Thioredoxin	Thioredoxin (PF00085)	Calcineurin (CL0163)	IPR013766
SCRFFL1031C11					Thioredoxin (PF00085)	Thioredoxin (CL0172)	IPR013766
					ERp29 (PF07749)	-	-
SCCCLR1C03A11	60s ribosomal protein l23a	NP_001150284	Miscellaneous	-	Ribosomal_L23eN (PF03939)	-	IPR005633
SCEQLR1091A10					Ribosomal_L23 (PF00276)	RRM (CL0221)	IPR013025
SCCCRZ2C04F06 SCSFFL3092A10	phi-1-like phosphate-induced protein precursor	XP_002454241	Miscellaneous	Phosphate-induced protein	Phi_1 (PF04674)	-	IPR006766
SCEZRZ3127F12	phi-1 precursor	XP_002444839	Miscellaneous	Phosphate-induced protein	Phi_1 (PF04674)	-	IPR006766
SCMCRT2108F02	protein disulfide isomerase7 precursor	XP_002440636	Miscellaneous	Thioredoxin	Thioredoxin (PF00085)	Thioredoxin (CL0172)	IPR013766
					Thioredoxin (PF00085)	Thioredoxin (CL0172)	IPR013766
					ERp29 (PF07749)	-	IPR011679

Table 5 – Identified ESTs and proteins in sugarcane cultured cells found only in two repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCRUAD1134G06	aldose 1-epimerase precursor	XP_002446606	Miscellaneous	Aldose epimerase	Aldose_epim (PF01263)	Gal_mutarotase (CL0103)	IPR008183
SCRULR1020A04	peptidyl-prolyl cis-trans isomerase cyp19-4 precursor	XP_002437567	Miscellaneous	-	Pro_isomerase (PF00160)	Cyclophil-like (CL0475)	IPR002130
SCVPFL3047F02.b	purple acid phosphatase precursor	ACF84615	Miscellaneous	Metallophosphoesterase	Metallophos (PF00149) Metallophos_C (PF14008)	Calcineurin (CL0163) -	IPR004843 -
SCBGFL5082E07	af464738_3 transposase	NP_001169638	Miscellaneous	Translocon-associated protein (TRAP)	TRAP_alpha (PF03896)	-	IPR005595
SCBGLR1023C09 SCCCST3006A03	acidic ribosomal protein p2a-2	XP_002447780	Miscellaneous	-	Ribosomal_60s (PF00428)	-	IPR001813
SCCCLR1022B08 SCQGLR1019G02	udp-glucose 6- expressed	NP_001146018	Miscellaneous	UDP-glucose dehydrogenase	UDPG_MGDP_dh_N (PF03721) UDPG_MGDP_dh (PF00984) UDPG_MGDP_dh_C (PF03720)	NADP_Rossmann (CL0063) 6PGD_C (CL0106) -	IPR001732 IPR014026 IPR014027

Table 5 – Identified ESTs and proteins in sugarcane cultured cells found only in two repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCBGRT3013B02	tpa: class iii peroxidase 107 precursor	XP_002461100	Oxido-reductases	Peroxidases	peroxidase (PF00141)	-	IPR002016
SCCCCL3002E11.b	tpa: class iii peroxidase 14 precursor	NP_001141196	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCEPFL3089C05	blue copper-binding	XP_002459222	Oxido-reductases	Blue copper binding proteins	Cu_bind_like (PF02298)	CU_oxidase (CL0026)	IPR003245
SCJFAD1014C09.b SCJFRT1010F10 SCRLAD1042E05 SCVPRZ2035F03	peroxidase 2-like	XP_002439231	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCJFLR1035D02	peroxidase 16 precursor	XP_002448675	Oxido-reductases	Peroxidases	peroxidase (PF00141)	-	IPR002016
SCJLLR1107D08	peroxidase	AAS75395	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCSGFL4C02D07	peroxidase atp6a	XP_002455760	Oxido-reductases	Peroxidases	peroxidase (PF00141)	-	IPR002016

Table 5 – Identified ESTs and proteins in sugarcane cultured cells found only in two repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCCCCL2001D02.b	peroxidase 1 precursor	XP_002450360	Oxido-reductases	Peroxidases	peroxidase (PF00141)	-	IPR002016
SCQSLB1052E11	peroxidase 72 precursor	XP_002455406	Oxido-reductases	Peroxidases	peroxidase (PF00141)	-	IPR002016
SCQSLB1049G04	ml domain protein	XP_002459364	Proteins related to lipid metabolism	lipase	E1_DerP2_DerF2 (PF02221)	-	IPR003172
SCSFRT2072C08	ml domain protein	NP_001141311	Proteins related to lipid metabolism	lipase	E1_DerP2_DerF2 (PF02221)	-	IPR003172
SCCCLR1024C05	nonspecific lipid-transfer protein 3 precursor	NP_001142300	Proteins related to lipid metabolism	Lipid transfer protein	Tryp_alpha_amyl (PF00234)	-	IPR003612
SCEPLB1044H11	nonspecific lipid-transfer protein 3 precursor	NP_001150972	Proteins related to lipid metabolism	Lipid transfer protein	Tryp_alpha_amyl (PF00234)	-	IPR003612
SCEQRT1029E06	exoglucanase1 precursor	NP_001130296	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_3 (PF00933) Glyco_hydro_3_C (PF01915)	Glyco_hydro_tim (CL0058) -	IPR001764 IPR002772

Table 5 – Identified ESTs and proteins in sugarcane cultured cells found only in two repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCJFRZ2033D06	glucan endo- -beta-glucosidase a6 precursor	XP_002462618	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_17 (PF00332)	Glyco_hydro_tim (CL0058)	IPR000490
					X8 (PF07983)	-	IPR012946
SCJFLR1013E04	udp-glucose 6-dehydrogenase	AAM47595	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	UDPG_MGDP_dh_N (PF03721)	NADP_Rossmann (CL0063)	IPR001732
					UDPG_MGDP_dh (PF00984)	6PGD_C (CL0106)	IPR014026
					UDPG_MGDP_dh_C (PF03720)	-	IPR014027
SCJFLR1013H10	udp-glucose dehydrogenase	XP_002468295	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	UDPG_MGDP_dh_N (PF03721)	NADP_Rossmann (CL0063)	IPR001732
					UDPG_MGDP_dh (PF00984)	6PGD_C (CL0106)	IPR014026
					UDPG_MGDP_dh_C (PF03720)	-	IPR014027
SCCCCL4009B01	receptor-like protein kinase precursor	XP_002440709	Proteins with interacting domains (with proteins or polysaccharides)	LRR proteins	LRR_8 (PF13855)	LRR (CL0022)	-

Table 5 – Identified ESTs and proteins in sugarcane cultured cells found only in two repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCSFAD1125C08	polygalacturonase inhibitor 1 precursor	XP_002463048	Proteins with interacting domains (with proteins or polysaccharides)	LRR proteins	LRRNT_2 (PF08263)	-	IPR013210
SCVPLR2019B03					LRR_4 (PF12799)	LRR (CL0022)	-
					LRR_4 (PF12799)	LRR (CL0022)	-
SCSFLR2024B05	hypothetical protein SORBIDRAFT_06g000550 [Sorghum bicolor]	XP_002447409	Proteins with interacting domains (with proteins or polysaccharides)	PMEI	PMEI (PF04043)	-	IPR006501
SCBFRZ2045E03	invertase inhibitor-like	XP_002452218	Proteins with interacting domains (with proteins or polysaccharides)	PMEI	PMEI (PF04043)	-	IPR006501
SCCCCL4006G06	cysteine protease 1 precursor	XP_002448736	Proteases	Cys proteases	Inhibitor_I29 (PF08246)	-	IPR013201
					Peptidase_C1 (PF00112)	Peptidase_CA (CL0125)	IPR000668
					Granulin (PF00396)	-	IPR000118
SCCCHR1002F08	prolyl carboxypeptidase like protein precursor	XP_002466952	Proteases	Ser proteases	Peptidase_S28 (PF05577)	AB_hydrolase (CL0028)	IPR008758

Table 5 – Identified ESTs and proteins in sugarcane cultured cells found only in two repetitions showing parameters of the identification together with the functional classification

(concludes)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCCCLR1C05E11	atypical receptor-like kinase mark precursor	NP_001105207	Signaling	LRK	LRRNT_2 (PF08263) Pkinase (PF00069)	- PKinase (CL0016)	IPR013210 IPR017442
SCQGLR2032D10	lgc1 precursor	XP_002446313	Unknown	DUF	DUF2782 (PF11191)	-	IPR021357

APPENDIX B - CELL WALL PROTEINS FROM SUGARCANE CELLS II

Table 6 – Identified ESTs and proteins in sugarcane cultured cells found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCCCCL3005A02.b SCQSLB1049A01	osmotin-like protein precursor	XP_002458737	Miscellaneous	Thaumatin-like protein	Thaumatin (PF00314)	-	IPR001938
SCVPRT2073B04	pathogenesis-related protein 5	XP_002443620	Miscellaneous	Thaumatin-like protein	Thaumatin (PF00314) Thaumatin (PF00314)	- -	IPR001938 IPR001938
SCCCHR1003C05	60s acidic ribosomal protein p2a	ACR34136	Miscellaneous	-	Ribosomal_60s (PF00428)	-	IPR001813
SCCCLR2001B08 SCQSLR1089F07	acidic ribosomal protein p2a-2	XP_002448737	Miscellaneous	-	Ribosomal_60s (PF00428)	-	IPR001813
SCEQLB1068G05	60s acidic ribosomal protein p2a	XP_002441563	Miscellaneous	-	Ribosomal_60s (PF00428)	-	IPR001813
SCEQRT2098F10	nucleotide pyrophosphatase phosphodiesterase	XP_002446231	Miscellaneous	Metallophosphoesterase	Metallophos (PF00149) Metallophos_C (PF14008)	Calcineurin (CL0163) -	IPR004843

Table 6 – Identified ESTs and proteins in sugarcane cultured cells found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCSGFL4036A04	tartrate-resistant acid phosphatase type 5 precursor	XP_002450916	Miscellaneous	Metallophosphoesterase	Metallophos (PF00149)	Calcineurin (CL0163)	IPR004843
SCCCRT2002D03	type-1 pathogenesis-related protein	XP_002465112	Miscellaneous	Pathogenesis-related protein	CAP (PF00188)	-	IPR014044
SCCCRT3007D01	ribonuclease 1 precursor	XP_002461070	Miscellaneous	-	Ribonuclease_T2 (PF00445)	-	IPR001568
SCBFRZ2017C04 SCJLAM2093F08 SCRFFL1030B07	phi-1 precursor	XP_002444839	Miscellaneous	Phosphate-induced protein	Phi_1 (PF04674)	-	IPR006766
SCEZAM1080E11	anthocyanidin -o-glucosyltransferase	XP_002436527	Miscellaneous	glycosyltransferase	UDPGT (PF00201)	GT-B (CL0113)	IPR002213
SCEZRT3069B05	arabinogalactan protein precursor	NP_001147175	Miscellaneous	Pollen Ole e 1 allergen	Pollen_Ole_e_I (PF01190)	-	IPR006041
SCJFLR1073H10 SCQGFL3054F07 SCUTLR2030A01	gpi-anchored protein	XP_002440062	Miscellaneous	X8	X8 (PF07983)	-	IPR012946
SCJFRZ2014F05	gast1 protein precursor	XP_002438216	Miscellaneous	Cysteine rich protein	GASA (PF02704)	-	IPR003854

Table 6 – Identified ESTs and proteins in sugarcane cultured cells found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCRULB1059B03	creg1 protein precursor	NP_001148752	Miscellaneous	Phosphate oxidase	Pyrid_oxidase_2 (PF13883)	FMN-binding (CL0336)	-
SCSBRZ3118E10	udp-glucose 6-dehydrogenase	AAM47595	Miscellaneous	UDP-glucose/GDP-mannose dehydrogenase	UDPG_MGDP_dh_N (PF03721)	NADP_Rossmann (CL0063)	IPR001732
SCSFSB1100C03	transferring glycosyl groups precursor	XP_002466674	Miscellaneous	Glycosyl transferase	Glyco_transf_8 (PF01501)	GT-A (CL0110)	IPR002495
SCBFRZ2017C11	per1_maize ame: full=peroxidase 1 ame: full=plasma membrane-bound peroxidase 1 short=pmpox1 flags: precursor	XP_002455565	Oxido-reductases	Peroxidases	peroxidase (PF00141)	-	IPR002016
SCCCAD1001C08	peroxidase 42 precursor	XP_002461210	Oxido-reductases	Peroxidases	peroxidase (PF00141)	-	IPR002016
SCJFLR1035D05	peroxidase	XP_002440652	Oxido-reductases	Peroxidases	peroxidase (PF00141)	-	IPR002016
SCQGFL4075C03	peroxidase 72 precursor	XP_002455405	Oxido-reductases	Peroxidases	peroxidase (PF00141)	-	IPR002016
SCQSLR1090E07	peroxidase 65 precursor	ACG24487	Oxido-reductases	Peroxidases	peroxidase (PF00141)	-	IPR002016

Table 6 – Identified ESTs and proteins in sugarcane cultured cells found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCRURT3064C05	tpa: class iii peroxidase 54 precursor	ACF88307	Oxido-reductases	Peroxidases	peroxidase (PF00141)	-	IPR002016
SCBGLR1023A04	blue copper-binding	XP_002459222	Oxido-reductases	Blue copper binding proteins	Cu_bind_like (PF02298)	CU_oxidase (CL0026)	IPR003245
SCRFHR1006G03	chemocyanin precursor	XP_002464008	Oxido-reductases	Blue copper binding proteins	Cu_bind_like (PF02298)	CU_oxidase (CL0026)	IPR003245
SCACLR2007H02	lipid binding protein precursor	XP_002459511	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	Prolamin (CL0482)	-
SCCCRT1C06C05	lipid binding protein	NP_001132005	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	Prolamin (CL0482)	-
SCCCCL3120E05	beta-expansin 1a precursor	XP_002464945	Proteins related to lipid metabolism	Lipid transfer protein	DPBB_1 (PF03330) Pollen_allerg_1 (PF01357)	DPBB (CL0199) -	IPR009009 IPR007117
SCCCLB1023H08	beta-expansin 1a precursor	XP_002464944	Proteins related to lipid metabolism	Lipid transfer protein	DPBB_1 (PF03330) Pollen_allerg_1 (PF01357)	DPBB (CL0199) -	IPR009009

Table 6 – Identified ESTs and proteins in sugarcane cultured cells found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCCCLR1065A10	beta-expansin 1a precursor	XP_002464951	Proteins related to lipid metabolism	Lipid transfer protein	DPBB_1 (PF03330)	DPBB (CL0199)	IPR009009
					Pollen_allerg_1 (PF01357)	-	IPR007117
SCBFST3134G03	cysteine protease 1 precursor	NP_001149658	Proteases	Cys proteases	Inhibitor_I29 (PF08246)	-	IPR013201
					Peptidase_C1 (PF00112)	Peptidase_CA (CL0125)	IPR000668
					Granulin (PF00396)	-	IPR000118
SCEZRZ1012G04	serine carboxypeptidase precursor	XP_002446348	Proteases	Ser proteases	Peptidase_S10 (PF00450)	AB_hydrolase (CL0028)	IPR001563
SCJLRT1019E02	probable serine protease eda2-like	NP_001146300	Proteases	Ser proteases	Peptidase_S28 (PF05577)	AB_hydrolase (CL0028)	IPR008758
SCJLRT1022A07	nucellin-like aspartic protease	XP_002449121	Proteases	Asp protease	Asp (PF00026)	Peptidase_AA (CL0129)	IPR001461
SCJLST1022B10	aspartic proteinase nepenthesin i	XP_002457394	Proteases	Asp protease	Asp (PF00026)	Peptidase_AA (CL0129)	IPR001461
SCEQRT2026H06	receptor protein kinase-like	XP_002452193	Proteins with interacting domains (with proteins or polysaccharides)	Lectin	B_lectin (PF01453)	-	IPR001480
					S_locus_glycop (PF00954)	-	
					PAN_2 (PF08276)	PAN (CL0168)	

Table 6 – Identified ESTs and proteins in sugarcane cultured cells found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCSGAD1009C09	beta- -galactosyltransferase	XP_002436753	Proteins with interacting domains (with proteins or polysaccharides)	Lectin	Gal-bind_lectin (PF00337) Galactosyl_T (PF01762)	Concanavalin (CL0004) GT-A (CL0110)	IPR001079 IPR002659
SCJFRZ2027H04	ripening-related protein	XP_002438999	Proteins with interacting domains (with proteins or polysaccharides)	PMEI	PMEI (PF04043)	-	IPR006501
SCRURT2005E11	hypothetical protein SORBIDRAFT_10g029260 [Sorghum bicolor]	XP_002437554	Proteins with interacting domains (with proteins or polysaccharides)	LRR protein	LRR_8 (PF13855)	LRR (CL0022)	-
SCSBFL4070E06	pectinesterase 31-like	XP_002468434	Proteins with interacting domains (with proteins or polysaccharides)	Pectinesterase	Pectinesterase (PF01095)	Pec_lyase (CL0268)	IPR000070
SCACLB1046D04 SCSGAD1009D10.b	cysteine protease 1 precursor	XP_002448736	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_18 (PF00704)	Glyco_hydro_tim (CL0058)	IPR001223

Table 6 – Identified ESTs and proteins in sugarcane cultured cells found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCMCRT2108A04	hypothetical protein SORBIDRAFT_05g023700 [Sorghum bicolor]	XP_002451067	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_18 (PF00704)	Glyco_hydro_tim (CL0058)	IPR001223
SCVPRZ3029G05	basic class iii chitinase hib3b	XP_002456672	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_18 (PF00704)	Glyco_hydro_tim (CL0058)	IPR001223
SCSGCL6072E02	aglu_orysj ame: full=probable alpha-glucosidase os06g0675700 ame: full=maltase flags: precursor	XP_002438844	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Gal_mutarotas_2 (PF13802) Glyco_hydro_31 (PF01055)	Gal_mutarotase (CL0103) Glyco_hydro_tim (CL0058)	- IPR000322
SCCCCL1001B08.b	hypothetical protein SORBIDRAFT_01g011440 [Sorghum bicolor]	XP_002464056	Unknown	DUF	DUF538 (PF04398)	-	IPR007493
SCCCFL4091H01	hypothetical protein SORBIDRAFT_03g027650 [Sorghum bicolor]	XP_002455939	Unknown	DUF	DUF642 (PF04862)	GBD (CL0202)	IPR006946
SCRFLB1055D11	mir-interacting saposin-like protein precursor	ACG32670	Unknown	DUF	DUF3456 (PF11938)	-	IPR021852
SCVPRT2080H05	bowman-birk serine protease inhibitor precursor	XP_002451266	Unknown	-	-	-	-

Table 6 – Identified ESTs and proteins in sugarcane cultured cells found only in one of the repetitions showing parameters of the identification together with the functional classification

(concludes)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCCCLR1065B12	hypothetical protein SORBIDRAFT_05g002910 [Sorghum bicolor]	XP_002450270	Signaling	LRR-recpetor kinase	LRR_8 (PF13855) Pkinase (PF00069)	LRR (CL0022) PKinase (CL0016)	- IPR017442

APPENDIX C - CELL WALL PROTEINS FROM SUGARCANE CULMS METHOD 1

Table 7 – Identified ESTs and proteins in sugarcane culms, method 1, found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCJLLR1107D08	peroxidase	AAS75395	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCQGFL4075C03	peroxidase 72 precursor	XP_002455405	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCCCCL3002E11.b	tpa: class iii peroxidase 14 precursor	NP_001141196	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCEQRT1026F09	peroxidase 2 precursor	NP_001151423.1	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCQSLB1052E11	peroxidase 72 precursor	XP_002455406	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCSGFL4C02D07	peroxidase atp6a	XP_002455760	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCJFLR1035D05	peroxidase	XP_002440652	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCBGRT1047G10	rhicadhesin receptor precursor	XP_002459427	Oxido-reductases	Germin-like protein	Cupin_1 (PF00190)	Cupin (CL0029)	IPR006045

Table 7 – Identified ESTs and proteins in sugarcane culms, method 1, found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCCCRZ1002H10	loc100283229 precursor	XP_002463056	Oxido-reductases	Blue copper binding protein	Cu_bind_like (PF02298)	CU_oxidase (CL0026)	IPR003245
SCBGHR1058A06	thaumatin-like protein precursor	XP_002442348	Miscellaneous	Thaumatin-like protein	Thaumatin (PF00314)	-	IPR001938
SCCCLR2003G06 SCJFRZ2025E05 SCUTLR1037F02 SCVPLR1028H01	thaumatin-like cytokinin-binding expressed	XP_002443390	Miscellaneous	Thaumatin-like protein	Thaumatin (PF00314)	-	IPR001938
SCCCRZ3097D07	hypothetical protein SORBIDRAFT_02g030940 [Sorghum bicolor]	XP_002460570	Miscellaneous	Phosphatase	Metallophos (PF00149)	Calcineurin (CL0163)	IPR004843
SCEQRT1033D01	loc100284565 precursor	NP_001150932	Miscellaneous	Acid Phosphatase	Acid_phosphat_B (PF03767)	HAD (CL0137)	IPR005519
SCSFAD1106E11	lectin-like protein kinase	XP_002445881	Miscellaneous	Lectin/Kinase	Lectin_legB (PF00139) Pkinase_Tyr (PF07714)	Concanavalin (CL0004) PKinase (CL0016)	IPR001220 IPR001245
SCJLFL1049F05	hypothetical protein SORBIDRAFT_08g018710 [Sorghum bicolor]	XP_002442350	Miscellaneous	DAHP_synthase	DAHP_synth_1 (PF00793)	TIM_barrel (CL0036)	IPR006218

Table 7 – Identified ESTs and proteins in sugarcane culms, method 1, found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCACLR2029H09 SCCCCL3001E03.b	xylogen protein 1	ACF83947	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	Prolamin (CL0482)	-
SCJFRT1009E11	lipid binding protein	XP_002441288	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	Prolamin (CL0482)	-
SCRULB1057C01	nonspecific lipid-transfer protein precursor	XP_002466001	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	Prolamin (CL0482)	-
SCCCAD1001B02 SCEPRT2047G01 SCEZLR1031G07 SCJFRZ2033G07 SCRUFL4024B04 SCRUSB1064D08	nonspecific lipid-transfer protein 4 precursor	XP_002442777	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	Prolamin (CL0482)	-
SCSFRT2072C08	ml domain protein	NP_001141311	Proteins related to lipid metabolism	Lipase	E1_DerP2_DerF2 (PF02221)	-	IPR003172

Table 7 – Identified ESTs and proteins in sugarcane culms, method 1, found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCQGLR2025F03	loc100285096 precursor	XP_002463844	Proteases	Ser proteases	Inhibitor_I9 (PF05922) Peptidase_S8 (PF00082) PA (PF02225)	- - -	IPR010259 IPR000209 IPR003137
SCQGSB1083B11	aspartic proteinase nepenthesin-1 precursor	ACG35309.1	Proteases	Asp protease	Asp (PF00026)	Peptidase_AA (CL0129)	IPR001461
SCEPLR1051B04	fasciclin-like arabinogalactan-protein	NP_001147121	Signaling	Fasciclin-like	Fasciclin (PF02469)	-	IPR000782
SCEQLR1093E12	arabinogalactan protein	XP_002440662	Signaling	Fasciclin-like	Fasciclin (PF02469)	-	IPR000782
SCCCCL5004G07 SCJFRT1007G04	glycoside family 28 precursor	XP_002437680	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_28 (PF00295)	Pec_lyase (CL0268)	IPR000743
SCAGLB1071B11 SCJFRZ1007E04	bowman-birk type trypsin inhibitor	XP_002451265	Proteins with interacting domains (with proteins or polysaccharides)	Protease inhibitor	Bowman-Birk_leg (PF00228)	-	IPR000877
SCVPLR2027A11	air12 precursor	XP_002445837	Unknown	Auxin-responsive protein AIR12 (DUF568)	DUF568 (PF04526)	-	IPR017214

Table 7 – Identified ESTs and proteins in sugarcane culms, method 1, found only in one of the repetitions showing parameters of the identification together with the functional classification

(concludes)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCEZHR1088E04	aspartic proteinase nepenthesin-1-like	ACR36490	Unknown	-	No PFAM families.	-	-

APPENDIX D - CELL WALL PROTEINS FROM SUGARCANE CULMS METHOD 2

Table 8 – Identified ESTs and proteins in sugarcane culms, method 2, found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCCCLB1004B09 SCEQRT2030A04	peroxidase 3	XP_002437459	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCCCLR1C05G08	peroxidase 72	XP_002455760	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCQGFL4075C03	peroxidase 72-like	XP_002455405	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCQGLR2032G02	hypothetical protein	XP_002464826	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCRUFL1119B06.b	peroxidase 3-like	NP_001042703.1	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCCCCL3002E11.b	tpa: class iii peroxidase 14 precursor	NP_001141196	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCCCLR1C03A09	peroxidase	XP_002440652	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCSGRT2062D07	loc100280824 precursor	XP_002437128	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016

Table 8 – Identified ESTs and proteins in sugarcane culms, method 2, found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCQSLB1052E11	peroxidase 72-like	XP_002455406	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCRLST3166D09	germin-like protein	ABG46238	Oxido-reductases	Germin-like	Cupin_1 (PF00190)	Cupin (CL0029)	IPR006045
SCSFAD1115E03 SCUTLR1037C05	germin-like protein 5-1-like	XP_002457300	Oxido-reductases	Germin-like	Cupin_1 (PF00190)	Cupin (CL0029)	IPR006045
SCCCRZ1002H10	hypothethical protein	XP_002463056	Oxido-reductases	Blue copper binding proteins	Cu_bind_like (PF02298)	CU_oxidase (CL0026)	IPR003245
SCBFFL4112F05	chemocyanin precursor	XP_002447932	Oxido-reductases	Blue copper binding proteins	Cu_bind_like (PF02298)	CU_oxidase (CL0026)	IPR003245
SCBFAD1045D12	glycosyl hydrolase family 1 expressed	NP_001151737	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_1 (PF00232)	Glyco_hydro_tim (CL0058)	IPR001360
SCCCSD1089A09	bgl12_orysj ame: full=beta-glucosidase 12 short=os4bglu12 flags: precursor	XP_002448027	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_1 (PF00232)	Glyco_hydro_tim (CL0058)	IPR001360

Table 8 – Identified ESTs and proteins in sugarcane culms, method 2, found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCPIRT3022G07	glycosyl hydrolase family 1 expressed	NP_001142124	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_1 (PF00232)	Glyco_hydro_tim (CL0058)	IPR001360
SCBFLR1039C04	exo- -beta-glucanase	AAR14129.1	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_3 (PF00933) Glyco_hydro_3_C (PF01915)	Glyco_hydro_tim (CL0058) -	IPR001764 IPR002772
SCEQRT1029E06	exoglucanase1 precursor	NP_001130296	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_3 (PF00933) Glyco_hydro_3_C (PF01915)	Glyco_hydro_tim (CL0058) -	IPR001764 IPR002772
SCCCSB1003H06	probable beta-d-xylosidase 7-like	NP_001130324	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_3 (PF00933) Glyco_hydro_3_C (PF01915) Fn3-like (PF14310)	Glyco_hydro_tim (CL0058) - -	IPR001764 IPR002772 -
SCQSRT1034D03	alpha-l-arabinofuranosidase c-terminus family expressed	XP_002442705	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Alpha-L-AF_C (PF06964)	-	IPR010720

Table 8 – Identified ESTs and proteins in sugarcane culms, method 2, found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCJFLR1013E04	udp-glucose 6-dehydrogenase	ACG33823.1	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	UDPG_MGDP_dh_N (PF03721) UDPG_MGDP_dh (PF00984) UDPG_MGDP_dh_C (PF03720)	- 6PGD_C (CL0106) -	IPR001732 IPR014026 IPR014027
SCCCFL6001E11	pectin methylesterase	XP_002458861	Proteins acting on cell wall polysaccharides	Carbohydrate esterases (CE)	Pectinesterase (PF01095)	Pec_lyase (CL0268)	IPR000070
SCEPRT2044F09	beta-galactosidase 8-like isoform 1	XP_002465536	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_35 (PF01301) Gal_Lectin (PF02140)	Glyco_hydro_tim (CL0058) -	IPR001944 IPR000922
SCBGHR1058A06 SCSGFL4C02G01 SCVPRT2080B07	thaumatin-like protein precursor	XP_002442348	Miscellaneous	Thaumatin-like protein	Thaumatin (PF00314)	-	IPR001938
SCJFRZ2025E05 SCVPLR1028H01	thaumatin-like protein precursor	XP_002443390	Miscellaneous	Thaumatin-like protein	Thaumatin (PF00314)	-	IPR001938
SCCCCL2001B01.b	gda1 cd39 family expressed	XP_002448918	Miscellaneous	Nucleoside Phosphatase	GDA1_CD39 (PF01150)	Actin_ATPase (CL0108)	IPR000407

Table 8 – Identified ESTs and proteins in sugarcane culms, method 2, found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCCCRZ3097D07	probable inactive purple acid phosphatase 29-like	XP_002460570	Miscellaneous	Phosphatase	Metallophos (PF00149)	Calcineurin (CL0163)	IPR004843
SCEQLB1067F01 SCSGRT2064D08	bark storage protein a-like	XP_002454890	Miscellaneous	Phosphorylase	PNP_UDP_1 (PF01048)	PUP (CL0408)	IPR000845
SCEQRT1033D01	stem 28 kda glycoprotein precursor	NP_001150932	Miscellaneous	Acid Phosphatase	Acid_phosphat_B (PF03767)	HAD (CL0137)	IPR005519
SCJLRT3078H06	disease resistance response protein 206-like	XP_002449918	Miscellaneous	Defense-related	Dirigent (PF03018)	-	IPR004265
SCCCRT3003E05	pathogenesis-related maize seed protein	XP_002436405	Miscellaneous	CAP	CAP (PF00188)	-	IPR014044
SCCCLR1C01H05 SCUTFL1064A11	serine carboxypeptidase precursor	XP_002457308	Proteases	Ser proteases	Peptidase_S10 (PF00450)	AB_hydrolase (CL0028)	IPR001563
SCVPRT2080D02	serine carboxypeptidase precursor	NP_001150676	Proteases	Ser proteases	Peptidase_S10 (PF00450)	AB_hydrolase (CL0028)	IPR001563

Table 8 – Identified ESTs and proteins in sugarcane culms, method 2, found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCBGLR1097G03	nucleoid dna-binding protein cnd41	XP_002451310	Proteases	Asp proteases	Asp (PF00026)	Peptidase_AA (CL0129)	IPR001461
SCQGLR2025F03	loc100285096 precursor	XP_002463844	Proteases	SER proteases	Inhibitor_I9 (PF05922)	-	IPR010259
					Peptidase_S8 (PF00082)	-	IPR000209
					PA (PF02225)	-	IPR003137
SCCCLR2004B05	lipid transfer protein	NP_001150566	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	Prolamin (CL0482)	-
SCQSLR1089H10	lipid binding protein	XP_002446445	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	Prolamin (CL0482)	-
SCJLFL3013H12 SCSFRT2071C05 SCSGFL1078B08 SCVPFL1066B10	pvr3-like protein	ACG42309.1	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	Prolamin (CL0482)	-
SCCCCL7001H09	alpha-l-fucosidase 2 precursor	XP_002437861	Proteins related to lipid metabolism	Lipase	Lipase_GDSL (PF00657)	SGNH_hydrolase (CL0264)	IPR001087

Table 8 – Identified ESTs and proteins in sugarcane culms, method 2, found only in one of the repetitions showing parameters of the identification together with the functional classification

(concludes)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCCCCL7001H04	fasciclin-like arabinogalactan protein 10 precursor	NP_001151356	Signaling	Fasciclin-like	Fasciclin (PF02469)	-	IPR000782
SCCCRZ2001E02	fasciclin-like arabinogalactan protein 11-like	XP_002458287	Signaling	Fasciclin-like	Fasciclin (PF02469)	-	IPR000782
SCRUAD1133E06	fasciclin-like arabinogalactan protein 11-like	XP_002458287	Signaling	Fasciclin-like	Fasciclin (PF02469)	-	IPR000782
SCQSSB1077D06	receptor-like protein kinase haiku2-like	XP_002442601	Proteins with interacting domains (with proteins or polysaccharides)	LRR protein	LRRNT_2 (PF08263) LRR_4 (PF12799) LRR_8 (PF13855) LRR_8 (PF13855)	- LRR (CL0022) LRR (CL0022) LRR (CL0022)	IPR013210 - - -
SCEZLB1013B06	aspartic proteinase nepenthesin-1-like	ACR36490	Unknown	-	no PFAM families	-	-

APPENDIX E - CELL WALL PROTEINS FROM SUGARCANE LEAVES METHOD 1

Table 9 – Identified ESTs and proteins in sugarcane leaves, method 1, found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCCCLR1C05G08	peroxidase atp6a	XP_002455760	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCEQRT1026F09	peroxidase 2 precursor	NP_001151423.1	Oxido-reductases	Peroxidases	Peroxidase (PF00141)	-	IPR002016
SCJLLR1106C06	phytocyanin pup2	NP_001141121	Oxido-reductases	Blue copper binding proteins	Cu_bind_like (PF02298)	CU_oxidase (CL0026)	IPR003245
SCEQLB1065E04	hypothetical protein SORBIDRAFT_01g015590 [Sorghum bicolor]	XP_002464281	Miscellaneous	nucleotide-diphospho-sugar transferases	Nucleotid_trans (PF03407)	GT-A (CL0110)	IPR005069
SCVPHR1095E09	serine threonine kinase-like protein	XP_002465466	Miscellaneous	-	Stress-antifung (PF01657) Stress-antifung (PF01657)	- -	IPR002902 IPR002902
SCVPRT2080B07	thaumatin-like protein precursor	XP_002442348	Miscellaneous	Thaumatin-like protein	Thaumatin (PF00314)	-	IPR001938

Table 9 – Identified ESTs and proteins in sugarcane leaves, method 1, found only in one of the repetitions showing parameters of the identification together with the functional classification

(continued)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCEQHR1082B01	bgl29_orysj ame: full=beta-glucosidase 29 short=os9bglu29 flags: precursor	XP_002462569	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_1 (PF00232)	Glyco_hydro_tim (CL0058)	IPR001360
SCSBRT3039A04 SCVPRZ3029G05	hevamine-a precursor	XP_002456672	Proteins acting on cell wall polysaccharides	Glycoside hydrolases (GH)	Glyco_hydro_18 (PF00704)	Glyco_hydro_tim (CL0058)	IPR001223
SCCCLR1022B11	cysteine protease 1 precursor	NP_001148706	Proteases	Cys proteases	Inhibitor_I29 (PF08246) Peptidase_C1 (PF00112) Granulin (PF00396)	- Peptidase_CA (CL0125) -	IPR013201 IPR000668 IPR000118
SCCCCL3001E03.b	xylogen protein 1	ACF83947	Proteins related to lipid metabolism	Lipid transfer protein	LTP_2 (PF14368)	Prolamin (CL0482)	-
SCEQRT1025E03	cysteine proteinase inhibitor	CAA60634	Proteins with interacting domains (with proteins or polysaccharides)	Cystatin	Cystatin (PF00031)	Cystatin (CL0121)	IPR000010
SCBGLR1112D07	probable inactive receptor kinase at1g48480-like	XP_002447974	Signaling	LRR receptor kinase	LRRNT_2 (PF08263) Pkinase (PF00069)	- PKinase (CL0016)	IPR013210 IPR017442

Table 9 – Identified ESTs and proteins in sugarcane leaves, method 1, found only in one of the repetitions showing parameters of the identification together with the functional classification

(concludes)

SUCEST	Blasted proteins	Hit	Functional class	Functional class II	PFAM (Family)	PFAM (Clan)	InterPro
SCBGFL5076F07	carboxyl-terminal peptidase	XP_002466249	Unknown	-	DUF4409 (PF14365) DUF239 (PF03080)	- -	- IPR004314