

**ANÁLISE DA DINÂMICA DE REGULAÇÃO DE GENES  
ASSOCIADOS À PAREDE CELULAR EM DIFENTES ESTÁGIOS DE  
DESENVOLVIMENTO DE SOJA**

**SARA SANGI MIRANDA**

**UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE  
DARCY RIBEIRO – UENF**

**CAMPOS DOS GOYTACAZES - RJ  
FEVEREIRO – 2018**

ANÁLISE DA DINÂMICA DE REGULAÇÃO DE GENES  
ASSOCIADOS À PAREDE CELULAR EM DIFERENTES ESTÁGIOS DE  
DESENVOLVIMENTO DE SOJA

**SARA SANGI MIRANDA**

“Dissertação apresentada ao Centro de Ciências e Tecnologias Agropecuárias da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Mestre em Genética e Melhoramento de Plantas.”

**Orientadora: Prof<sup>a</sup>. Clícia Grativol Gaspar de Matos**

CAMPOS DOS GOYTACAZES - RJ  
FEVEREIRO - 2018

## FICHA CATALOGRÁFICA

Preparada pela Biblioteca do **CCH / UENF**

032/2018

M672 Miranda, Sara Sangi.

Análise da dinâmica de regulação de genes associados a parede celular em diferentes estágios de desenvolvimento de soja / Sara Sangi Miranda. – Campos dos Goytacazes, RJ, 2018.

157 f.

Bibliografia: f. 69 – 86.

Dissertação (Mestrado em Genética e Melhoramento de Plantas) – Universidade Estadual do Norte Fluminense Darcy Ribeiro, Centro de Ciências e Tecnologias Agropecuárias, 2018.

Orientadora: Clícia Grativol Gaspar de Matos.

1. Semente. 2. Análise Gênica. 3. Soja. 4. Germinação. I. Universidade Estadual do Norte Fluminense Darcy Ribeiro. II. Título.

CDD – 633.34

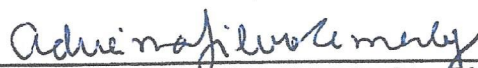
ANÁLISE DA DINÂMICA DE REGULAÇÃO DE GENES  
ASSOCIADOS À PAREDE CELULAR EM DIFERENTES ESTÁGIOS DE  
DESENVOLVIMENTO DE SOJA

**SARA SANGI MIRANDA**

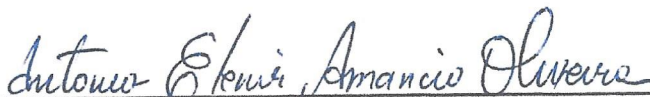
“Dissertação apresentada ao Centro de Ciências e Tecnologias Agropecuárias da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Mestre em Genética e Melhoramento de Plantas.”

Aprovada em: 07 de fevereiro de 2018.

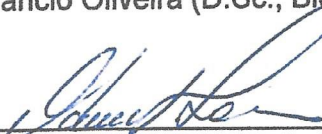
Comissão Examinadora:



Prof<sup>ª</sup>. Adriana Silva Hemerly (D.Sc., Biotecnologia) - UFRJ



Prof<sup>ª</sup>. Antônia Elenir Amâncio Oliveira (D.Sc., Biociências e Biotecnologia) - UENF



Prof<sup>ª</sup>. Valdirene Moreira Gomes (D.Sc., Ciências) – UENF



Prof<sup>ª</sup>. Clícia Grativol Gaspar de Matos (D.Sc., Química Biológica) – UENF  
(Orientadora)

## **AGRADECIMENTOS**

Aos meus pais, Luéia e Jorcy, por lutarem pela minha educação. Por acreditarem em mim e por não medirem esforços para que eu tivesse a melhor educação possível. Eu não tenho palavras que consigam expressar e agradecer tudo o que vocês fizeram por mim. Amo vocês.

Às minhas tias, Ester e Neusi, pelo apoio incondicional durante não só os meus estudos, mas em toda a minha vida. Obrigada titias, por sempre me ajudarem e orarem por mim.

À toda a minha família por torcerem pelo meu sucesso. Obrigada.

À Universidade Estadual do Norte Fluminense - UENF, por toda estrutura e oportunidades. E obrigada, Darcy Ribeiro, por fazer do seu sonho o sonho de tantas outras pessoas.

Ao Programa de Pós-Graduação em Genética e Melhoramento de Plantas e à sua coordenação, pela oportunidade de realização do mestrado e por todo empenho e zelo pelo programa.

À minha orientadora, Clícia Grativol, fundamental para a realização desse trabalho. Obrigada, por mesmo não me conhecendo, acreditar em mim. Obrigada pela

paciência, desafios, enigmas e, principalmente, obrigada por seus ensinamentos. Gratidão.

À todos os professores que fizeram parte da minha jornada na UENF, obrigada por contribuírem para minha formação.

Ao LBCT, em especial à professora Maura e à Camila, muito obrigada pela disponibilidade, paciência e colaboração.

Ao LBMP – UFRJ, em especial ao Helkin, por toda ajuda no processamento das imagens de celulose.

Ao Daniel, por sua atenção e por nunca poupar esforços para ajudar. Muito obrigada. O mundo precisa de mais pessoas como você.

À associação de alunos do corredor do LQFPP. Obrigada por fazerem mais leve essa jornada.

Aos meus companheiros de laboratório. Obrigada pelos risos, conversas e experiências trocadas.

Ao grupo da abdução, obrigada Marci, Pri e Michel, pelos choros, lamúrias e lamentações compartilhadas. E sobretudo, obrigada por mesmo de longe se fazerem presentes em minha vida.

Ao Kayan e Mariana, obrigada pelas conversas, alegrias compartilhadas e conselhos. E além de tantas outras coisas, obrigada pela amizade.

Ao Sci-Hub e a sua fundadora Alexandra Elbakyan, por sua coragem de lutar contra a desigualdade de conhecimento no mundo e por levantar a bandeira da remoção de barreiras no acesso da ciência

À FAPERJ, pela bolsa concedida mesmo em meio a tantas dificuldades impostas pela má administração pública. “Que os órgãos de fomento à pesquisa e os

representantes públicos nas três esferas de governo valorizem os pesquisadores iniciantes e priorizem a educação e a ciência como fundamentos para uma sociedade democrática."

À todas energias positivas que contribuíram para a realização desse trabalho.  
Gratidão.

*“Se fracassar, ao menos que fracasse ousando grandes feitos, de modo que a sua postura não seja nunca a dessas almas frias e tímidas que não conhecem nem a vitória nem a derrota.”*

Theodore Roosevelt

*“A vida me ensinou a nunca desistir. Nem ganhar, nem perder, mas procurar evoluir. Oriente. Se oriente. Alma. Espírito. Corpo. Mente.”*



## SUMÁRIO

RESUMO.....	xi
ABSTRACT .....	xiii
1. INTRODUÇÃO.....	1
2. REVISÃO BIBLIOGRÁFICAS .....	4
2.1.Aspectos gerais do cultivo de soja.....	4
2.2.    Aspectos fisiológicos da embriogênese e formação da semente de soja	7
2.3.    Parede Celular Vegetal.....	11
2.3.1.    Biossíntese dos componentes estruturais da parede celular .....	13
2.4.    Transcriptoma: RNA-Seq .....	21
3. OBJETIVOS.....	25
3.1.    Objetivo geral .....	25
3.2.    Objetivos específicos .....	25
4. MATERIAIS E MÉTODOS .....	27
4.1.Anotação dos genes relacionados com a parede celular no genoma de soja	27
4.1.1.    Genes COBRA.....	27
4.1.1.1.    Identificação e anotação dos genes COBRA no genoma de soja.....	28
4.1.1.2.    Análise da estrutura gênica e distribuição dos genes COBRA nos cromossomos de soja.....	28
4.1.1.3.    Análise filogenética da família COBRA .....	29
4.1.1.4.    Perfil de expressão dos genes cobra em soja.....	29

4.1.1.5. Duplicação gênica e estimativa das taxas de substituição sinônimas e não-sinônimas dos genes COBRA em soja.....	29
4.2. Obtenção das bibliotecas de transcriptomas de genes envolvidos com a parede celular em diferentes estágios de desenvolvimento da semente de soja .....	30
4.3. Análise de mapas transcricionais de parede celular durante os diferentes estágios de desenvolvimento da semente de soja .....	31
4.4. Processamento das sementes de soja e observação microscópica .....	32
4.5. Ensaio de atividade da Peroxidase.....	33
5. RESULTADOS .....	35
5.1. Identificação e anotação dos genes pertencentes às vias de síntese da parede celular	35
5.1.1. Identificação, caracterização e classificação dos membros da família COBRA em soja .....	37
5.1.2. Análise filogenética da família COBRA em soja.....	39
5.1.3. Duplicação dos genes COBRA, localização e distribuição cromossomial	41
5.1.4. Estimativa da taxa de substituição sinônima e não-sinônima ( $K_a/K_s$ ) e estrutura gênica .....	43
5.1.5. Perfil de expressão e função biológica da família COBRA em soja .....	44
5.2. Caracterização do perfil de expressão dos genes pertencentes às vias de síntese, arquitetura e montagem da parede celular em sementes de soja.....	46
5.3. Caracterização morfológico-tecidual da parede celular do eixo embrionário durante a germinação da semente de soja.....	49
5.3.1. Caracterização fenotípica do eixo embrionário durante os estágios de germinação da semente de soja.....	53
5.3.2. Caracterização tecidual da parede celular do eixo embrionário durante os primeiros estágios de germinação da semente de soja .....	54
5.3.4. Determinação da atividade peroxidase .....	57
6 DISCUSSÃO.....	58
6.1. Identificação de genes de parede celular em soja .....	59
6.2. Correlação entre a expressão gênica e dinâmica da parede celular.....	62
7 CONCLUSÕES.....	67
REFERÊNCIAS BIBLIOGRÁFICAS.....	69
APÊNDICE.....	87

## LISTA DE FIGURAS

- Figura 1.** Principais países produtores de soja entre os anos de 2013-2014. A figura apresenta a quantidade de toneladas do grão produzidas e seus respectivos países produtores entre a safra de 2013-2014 (FAO, 2017). .....6
- Figura 2.** Regiões de cultivo de soja no Brasil. O mapa mostra as áreas de cultivo de soja no Brasil na safra de 2015-2016. O Estado do Mato Grosso é o maior produtor sendo responsável por 26% da produção de soja do país. Adaptado de CONAB (2016). .....7
- Figura 3.** Eventos fisiológicos que ocorrem durante o desenvolvimento e maturação da semente de soja. A figura apresenta os diversos estágios de desenvolvimento da semente de soja juntamente com todos os processos morfológicos e fisiológicos que ocorrem no interior da mesma Adaptado de Seed Gene Network (2016). .....8
- Figura 4.** Ilustração esquemática da parte externa e interna da semente de soja. A – Imagem ilustrativa da estrutura externa da semente dando enfoque as suas principais estruturas. B – Imagem ilustrativa da estrutura interna da semente dando enfoque as suas principais estruturas. Adaptado de AGEITEC (2016). .....9
- Figura 5.** Sementes de soja nas primeiras horas de germinação. 0 – Semente quiescente. 3, 6, 12 e 24 – Horas após a embebição da semente em água. .... 11
- Figura 6.** Modelo simplificado da estrutura da parede celular da planta. (A) A estrutura consiste em três camadas principais: a lamela média e a parede primária e secundária. (B) Os principais polissacarídeos que formam a estrutura da parede. Adaptado de Hild (2014). ..... 12
- Figura 7.** Modelo esquemático da estrutura da celulose encontrada na parede celular da planta. Dezenas de cadeias lineares de glucano formam a fibra de celulose, a interação e união dessas fibras formam as microfibrilas e um conjunto de microfibrilas formam a macrofibrila estrutural da parede. Adaptado de Hild (2014). ..... 14

- Figura 8.** Modelo de como o complexo CESA sintetiza uma microfibrila de celulose. Cada proteína CESA sintetiza uma única cadeia de D-glucano ligada  $\beta$ -(1,4). A celulose é formada como uma fita cristalina que é composta por muitos desses glucanos. No modelo, 36 cadeias de  $\beta$  1,4-D-glucano formam a roseta, que é composta por um hexâmero de hexâmeros CESA. Adaptado de Doblin (2005)... 15
- Figura 9** Estrutura esquemática geral das pectinas mostrando os quatro grupos de polissacarídeos pécticos, homogalacturonano (HG), xilogalacturonano (XGA), Rhamnogalacturonan I (RG-I) e Rhamnogalacturonan II (RG-II). Adaptado de Mohnen, (2008)..... 19
- Figura 10.** Esquema dos estágios de desenvolvimento da semente de soja mostrando em quais momentos do desenvolvimento as bibliotecas de transcriptoma foram realizadas. DAF= Dias após da fertilização. WS= Semente inteira. COT= Cotilédone. HAE=Horas após a embebição. ....31
- Figura 11.** Árvore filogenética dos genes COBRA. A árvore filogenética foi gerada no MEGA7 usando o método de Neighbor-Joining. Utilizamos o total de 72 sequências codificadoras de proteínas de diferentes espécies. ....40
- Figura 12.** Distribuição do GmCOBRA nos cromossomos de soja. Somente os cromossomos que possuem genes COBRA estão representados. As linhas vermelhas indicam a duplicação dos genes parálogos. ....42
- Figura 13.** Estrutura gênica das 23 sequências COBRA em soja. Os éxons são representados em retângulos pretos, 5'UTR em retângulos cinza e introns em linhas pretas. ....43
- Figura 14.** Agrupamento hierárquico dos perfis de expressão de genes COBRA de soja em diferentes tecidos. A - Tecidos vegetativos e reprodutivos. B - Etapas da formação de sementes e germinação. HAI - Horas após embebição. WS - Semente inteira. EI - Eixo embrionário .....45
- Figura 15.** Gráficos de distribuição da frequência gênica por  $\log_2(\text{RPKM})$  dos genes de parede celular anotados. A- desenvolvimento da semente. B- Eixo embrionário após a germinação da semente. DAF – Dias após a fertilização. WS – Semente inteira. COT – Cotilédone. Dry seed – Semente quiescente. Dry – Eixo embrionário quiescente. HAI – Horas após a embebição. ....48
- Figura 16** *Heatmap* mostrando a frequência e expressão das vias de parede celular durante diferentes estágios de desenvolvimento e germinação da semente de soja. A- Frequência das vias de desenvolvimento da semente. B- Expressão das vias de desenvolvimento da semente. C- Frequência das vias de germinação. D- Expressão das vias de germinação. ....50
- Figura 17.** Expressão de genes-chave envolvidos na regulação da parede celular em eixo embrionário de sementes de soja durante a germinação. Os gráficos mostram a expressão dos genes com a descrição do seu respectivo homólogo em Arabidopsis, durante os estágios de desenvolvimento da semente e germinação do eixo embrionário.....52

- Figura 18.** Eixo embrionário durante a germinação da semente de soja. HAI – Horas após a embebição. Escala 1 cm.....53
- Figura19.** Anatomia do eixo embrionário durante 3 (acima) e 24 (abaixo) horas de germinação. A – Meristema radicular; B – Procâmbio; C- Meristema apical. Pd – Protoderme, Pc – Procâmbio, Mfm – Meristema fundamental medular, Mfc – meristema fundamental cortical. ....55
- Figura 20** Fluorescência relativa de celulose nos eixos de 3 e 24 horas após a embebição. A-C. eixo embrionário de 3h, D-F eixo embrionário de 24h. (A, D) meristema radicular; (B, E) procâmbio; (C, F) meristema apical. G. Gráfico de fluorescência relativa de celulose nos tecidos. \* diferenças entre as médias estatisticamente significativas com  $p \leq 0,05$  \*\* - diferenças entre as médias estatisticamente significativas com  $p \leq 0,01$ . ....56
- Figura 21** Atividade da enzima peroxidase determinada em membrana após a separação das proteínas por eletroforese nativa a 12%. M – Marcador de massa molecular. 0, 3, 6, 12 e 24 horas após a embebição do eixo embrionário. ....57
- Figura 22** - Principais eventos metabólicos e vias ativadas durante a germinação do eixo embrionário. Durante o período de 0 para 3h, a semente absorve água e ativa vias responsáveis pela extensão da parede celular. Entre o período de 6 a 12 horas após a absorção de água, a parede celular participa do balanço hídrico celular e tem-se início o remodelamento da parede celular para a protusão do embrião. Entre 12 e 24 horas após a embebição, pela ação do remodelamento e expansão da parede celular, o embrião se expande.....66

## RESUMO

MIRANDA, SARA SANGI; Universidade Estadual do Norte Fluminense Darcy Ribeiro; Fevereiro, 2018. Análise da dinâmica de regulação de genes associados à parede celular em diferentes estágios de desenvolvimento de sementes de soja. Orientadora: Clícia Grativol Gaspar de Matos; Conselheiros: Valdirene Moreira Gomes e Thiago Motta Venâncio.

A soja é a cultura líder do agronegócio no Brasil. A importância da soja no mercado deve-se principalmente ao alto teor de óleo e proteína de seus grãos. O desenvolvimento da semente é estritamente controlado pelo crescimento e expansão das células. A morfogênese celular é restrita e influenciada pela presença da parede celular. A parede celular define a forma da célula e é fortemente influenciada por fatores bióticos e abióticos em resposta às necessidades fisiológicas da planta. Além disso, componentes estruturais da parede celular são responsáveis pelo fornecimento de matéria-prima para algumas produções industriais como bebidas, cereais, massas, entre outros. Nesse sentido, o presente trabalho teve como objetivo caracterizar as variações de expressão dos genes envolvidos com a biossíntese da parede celular vegetal em diversos estágios do desenvolvimento de soja, com enfoque na regulação de parede celular em eixo embrionário durante a germinação. Os dados obtidos demonstram que existem 2.143 genes que controlam a biossíntese e montagem da parede celular no genoma de soja. Esses genes compreendem diversas famílias gênicas que possuem um padrão de expressão característico durante os estágios de desenvolvimento da semente e de formação do embrião durante a germinação. Entre as vias de

formação da parede celular, as vias de fenilpropanoides, peroxidases, GT família 8 e vias de interconversão de açúcar-nucleotídeos se destacaram como altamente expressas em todo o processo de desenvolvimento da semente e germinação do embrião. As análises microscópicas mostraram a expansão da parede celular do eixo embrionário e aumento do conteúdo de celulose durante a germinação, corroborando com os dados de expressão obtidos. A identificação de genes-chave regulados durante a germinação fornece uma base para uma caracterização funcional e bioquímica dos componentes da parede celular que controlam não só a biossíntese, mas também a expansão da parede em sementes de soja. Além disso, a identificação desses genes fornece uma melhor compreensão do processo de germinação da semente, o que pode contribuir para novas abordagens na produção de soja.

Palavras-chave: parede celular; transcriptoma; expressão gênica; soja

## **ABSTRACT**

MIRANDA, SARA SANGI; Universidade Estadual do Norte Fluminense Darcy Ribeiro; February, 2018. Analysis of the cell wall associated genes regulation at different stages of soybean seed development. Advisor: Clícia Grativol Gaspar de Matos; Counselors: Valdirene Moreira Gomes e Thiago Motta Venâncio.

Soybean is the leading agribusiness culture in Brazil. The importance of soybean in the market is mainly due to its high oil and protein content. The development of the seed is strictly controlled by the growth and expansion of the cells. Cell morphogenesis is restricted and influenced by the presence of the cell wall. The cell wall defines the shape of the cell and is strongly influenced by biotic and abiotic factors. In addition, structural components of the cell wall are responsible for the supply of raw material for some industrial products such as beverages, cereals, pasta, among others. The present work aimed to characterize the expression variations of genes involved in plant cell wall biosynthesis at various stages of soybean seed development, focusing on the cell wall regulation in the embryonic axis during germination. The results demonstrate that soybean genome has 2,143 genes involved with the biosynthesis and assembly of the cell wall. These genes comprise several gene families with a recognizable expression pattern during the stages of seed development and germination. Among the pathways of cell wall formation, the pathways of phenylpropanoids, peroxidases, GT family 8 and sugar nucleotide interconversion were found as highly expressed throughout the seed development and embryo germination. Microscopic analysis showed the embryonic axis cell wall expansion and cellulose content increase during germination and



corroborated the expression data. The identification of key genes regulated during germination provides a basis for a functional and biochemical characterization of cell wall biosynthesis and expansion components in soybean seeds. In addition, the identification of these genes provides a better understanding of the seed germination process, which may contribute to new approaches in soybean production.

Key words: cell wall; transcriptome; gene expression; soybean

## 1. INTRODUÇÃO

A população mundial cresce a cada ano. Em 2010, alcançamos a marca de 7 bilhões de pessoas no mundo e, em 2017, já atingimos 7,5 bilhões. Segundo a ONU (2017), em 2050 a população mundial chegará a 9,8 bilhões de pessoas. Segundo a FAO, nos próximos 30 anos a produção agrícola precisará aumentar pelo menos 50% para atender a demanda populacional (FAO, 2009). Entretanto, o ritmo atual de produção agrícola é insuficiente para atender tal demanda, além da variação climática que afeta diretamente a produção das culturas (Ray et al., 2015). No Brasil, estima-se que para o ano de 2018 a safra de grãos seja 8,9% menor em relação ao ano de 2017, o que corresponde a cerca de 20 milhões de toneladas de grãos perdidos (IBGE, 2017). Nesse sentido, o atual desafio da agricultura é aumentar o rendimento das culturas utilizando técnicas e estratégias que permitam a ampliação da qualidade e produção, onde as mesmas não agridam o meio ambiente e garantam a segurança alimentar.

A biotecnologia tem tido um papel muito importante na agricultura mundial. Com o avanço desta área, por intermédio das técnicas de biologia molecular e de bioinformática, tem sido possível identificar e selecionar genes responsáveis pela expressão de características agronômicas de interesse, aumentando a produção e/ou desenvolvendo variedades mais resistentes. Um exemplo desse tipo de estudo é o desenvolvimento da variedade de soja resistente à ferrugem asiática. Nesta variedade, o gene *Rpp5*, que confere resistência a doença, foi identificado e selecionado para desenvolvimento de cultivares resistentes (Garcia et al., 2008).

Além disso, várias tecnologias como, por exemplo, RT-PCR (*reverse transcription polimerase chain reaction*), microarranjos de DNA e análise serial da expressão gênica (SAGE) foram desenvolvidas a fim de caracterizar parte ou todo o transcriptoma de organismos potencialmente úteis na medicina e agricultura. (Lowe et al., 2017).

Uma das tecnologias mais recentes na análise de transcriptomas é o RNA-seq, uma ferramenta altamente sensível para medir a expressão do transcriptoma em diferentes tecidos e estágios de desenvolvimento. Essa técnica é capaz de detectar genes expressos em baixos níveis e, também, os genes altamente expressos, como fatores de transcrição (Wang et al., 2010). Por meio da tecnologia do RNA-seq, pode-se melhorar significativamente a anotação dos genes e fornecer recursos valiosos para estudos funcionais dos mesmos (Wang et al., 2014). O RNA-seq tem sido utilizado em vários estudos de análise da expressão gênica de plantas de interesse econômico como o milho (Kakumanu et al., 2012), cana-de-açúcar (Manners e Casu, 2011), trigo (Duan et al., 2012) e soja (Severin et al., 2010).

A soja é uma das principais fontes mundiais de óleo e proteína. O Brasil ocupa uma posição de destaque frente ao mercado mundial do grão, sendo o segundo maior produtor. Na safra de 2016/2017 foram produzidos no Brasil cerca de 114.075,3 milhões de toneladas do grão em uma área de cerca de 33.914,9 milhões de hectares (CONAB, 2017). Alguns estudos utilizando a tecnologia RNA-seq já foram realizados em soja, desde a análise da planta sob condições de estresse hídrico (Rodrigues et al., 2015), condições de seca (Marcolino-Gomes et al., 2015), respostas a patógenos (Zhang e Song, 2017), até genes responsáveis pela formação da semente (Jones e Vodkin, 2013; Kour et al., 2014; Du et al., 2017) e germinação (Bellieny-Rabelo et al., 2016).

As sementes são estruturas biológicas complexas que possuem a capacidade de tolerar a dessecação, manter a viabilidade metabólica, garantir o sucesso da germinação e, conseqüentemente, o estabelecimento das plântulas (Bewley et al., 2013). Após a fertilização e embriogênese, as sementes acumulam reservas durante o seu desenvolvimento. Durante as etapas de formação da semente, o armazenamento de polissacarídeos constituintes da parede celular e proteínas de reserva é intensificado. Esses polissacarídeos e proteínas de reserva possuem um importante papel durante o processo de germinação da semente, onde, são degradados e seus produtos utilizados como fonte de carbono e energia

para o crescimento inicial das plântulas. Ao longo da germinação a semente se expande, as enzimas se tornam hidratadas e ativas, reiniciando o metabolismo para produzir a energia que será usada para completar a germinação e proporcionar o crescimento da plântula (Bewley et al., 2013). Além de servir como reserva, os polissacarídeos da parede celular apresentam outras funções como o controle da embebição e controle da expansão celular dos cotilédones (Buckeridge et al., 2000).

A parede celular vegetal possui uma grande diversidade de funções durante a vida da planta: proporciona robustez mecânica à célula, protege a célula contra organismos potencialmente patogênicos, controla a expansão celular e desempenha um papel central na qualidade dos alimentos de origem vegetal para o consumo humano e animal (Evert e Eichhorn, 2014). Mudanças morfológicas e fisiológicas que ocorrem durante o crescimento e expansão da planta resultam em grande parte da divisão celular que é controlada juntamente com a modificação estrutural e reorganização dos componentes da parede celular (Nicol e Höfte, 1998). Contudo, os fatores que regulam essas modificações da parede celular ainda são pouco conhecidos em sementes de soja.

Apesar da essencial importância da parede celular ao longo do ciclo de vida da planta, a regulação gênica envolvida na síntese da parede celular durante as etapas de desenvolvimento da semente e germinação da mesma permanece, em grande parte, inexplorada. Os recentes avanços nos campos da bioquímica, biologia molecular e bioinformática podem contribuir significativamente para a identificação, análise e aprofundamento do conhecimento acerca do grande grupo de genes relacionados à parede celular em soja.

## 2. REVISÃO BIBLIOGRÁFICAS

### 2.1. Aspectos gerais do cultivo de soja

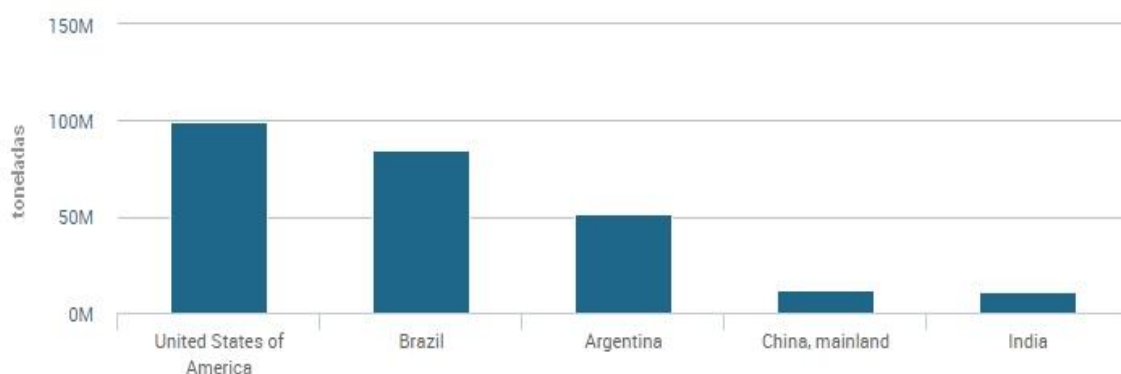
A soja comumente cultivada pertence ao reino Plantae, filo Magnoliophyta, classe Rosidea, ordem Fabales, família Fabaceae, subfamília Faboideae, gênero *Glycine* e espécie *Glycine max* (The Angiosperm Phylogeny Group III, 2009). A espécie possui  $2n=40$  cromossomos (Schmutz et al., 2010). Essa cultura foi primeiramente descrita há cerca de 5.000 mil anos na costa leste da Ásia, essencialmente na China, onde era utilizada como forragem. Posteriormente, a cultura foi utilizada como alimento para os animais, onde obteve grande sucesso. A partir da China, através das rotas mercantis, a soja se dispersou geograficamente e, hoje, é cultivada em grande parte do planeta (Miyasaka e Medina, 1981).

No ocidente, a soja foi provavelmente introduzida em 1765 por colonos nos Estados Unidos (NCSOY, 2016). Em 1879, ainda nos Estados Unidos, foram divulgados os primeiros estudos científicos realizados sobre a espécie (Federizzi, 2017). Por volta de 1904, um químico norte-americano, George Washington Carver, realizando suas pesquisas, constatou que a soja é uma importante fonte de proteína e óleo para a alimentação humana. Na década de 40, durante a Segunda Guerra Mundial, a procura por óleos, plásticos e outros produtos aumentaram consideravelmente, o que acarretou em uma maior produção da cultura (NCSOY, 2016).

No Brasil, as primeiras sementes de soja chegaram no Nordeste do país vindas dos Estados Unidos através do professor da Escola de Agronomia da Bahia, Gustavo Dutra, por volta de 1882. Porém, a cultura não obteve muito sucesso por conta do fotoperíodo e clima da região serem muito discrepantes daqueles encontrados nos Estados Unidos. Por volta de 1900, as sementes foram cultivadas em escolas de São Paulo e Rio Grande do Sul onde encontraram efetivas condições para o seu desenvolvimento e se estabeleceram (Bois, 2008). Entretanto, a grande expansão da cultura da soja no país iniciou-se nas décadas de 80 e 90 apoiada pelos programas de melhoramento genético da soja, onde foram desenvolvidos cultivares adaptados à região dos Cerrados, que hoje é a maior região produtora dos grãos no Brasil. Atualmente, a soja é a cultura com a maior área plantada no Brasil (Embrapa, 2017).

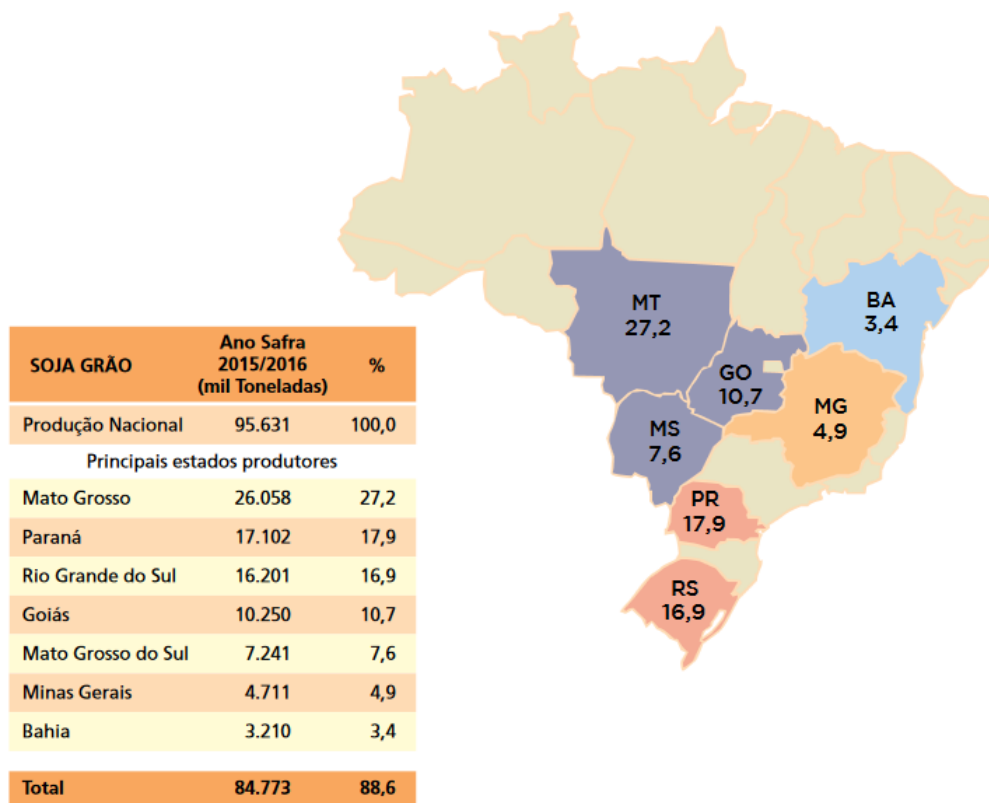
A primeira característica que possibilitou a expansão da soja para as regiões do centro-oeste do país foi o período juvenil longo, onde a planta não floresce durante o período juvenil, mesmo que as condições climáticas existam, permitindo o crescimento e desenvolvimento do cultivar (Kiihl e Garcia, 1989). Nos dias de hoje, são produzidas 317.253 milhões de toneladas do grão de soja em todos os continentes e em uma área de 118.135 milhões de hectares, sendo então, um dos 10 alimentos mais produzidos no planeta (FAOSTAT, 2017).

De acordo com a Food and Agriculture Organization of the United Nations (FAO, 2016), os maiores produtores de soja são os Estados Unidos, produzindo na safra de 2013/2014 cerca de 108.014 milhões de toneladas do grão, seguido do Brasil, com produção de 95,07 milhões de toneladas e, em terceiro lugar, a Argentina, produzindo 59,50 milhões de toneladas. Juntos, Estados Unidos, Brasil e Argentina produzem 82% da produção de soja mundial (Figura 1).



**Figura 1.** Principais países produtores de soja entre os anos de 2013-2014. A figura apresenta a quantidade de toneladas do grão produzidas e seus respectivos países produtores entre a safra de 2013-2014 (FAO, 2017).

No Brasil, a soja é a cultura líder do agronegócio e é produzida por todas as regiões do país. O estado do Mato Grosso é o maior produtor, correspondendo a cerca de 30% do total produzido no país. Em seguida, os estados do Paraná e Rio Grande do Sul com 19% e 13% do total, respectivamente (Figura 2). O Brasil vem sendo considerado a grande promessa no fortalecimento da produção mundial do grão (CONAB, 2017). Segundo a Companhia Nacional de Abastecimento (CONAB), a produção de soja no Brasil deve chegar a 109.183,4 milhões de toneladas na safra de 2017/2018 (CONAB, 2017). E ainda, segundo uma previsão realizada pela FAO, a produção brasileira aumentará 52,8% nos próximos 10 anos. A estimativa é que o Brasil seja o maior produtor do grão em 2024 (OECD-FAO, 2015). Essa previsão de aumento expressivo da produtividade da soja no Brasil está intimamente associada aos avanços tecnológicos, ao manejo, à eficiência dos produtores e à pesquisa.



**Figura 2.** Regiões de cultivo de soja no Brasil. O mapa mostra as áreas de cultivo de soja no Brasil na safra de 2015-2016. O Estado do Mato Grosso é o maior produtor sendo responsável por 26% da produção de soja do país. Adaptado de CONAB (2016).

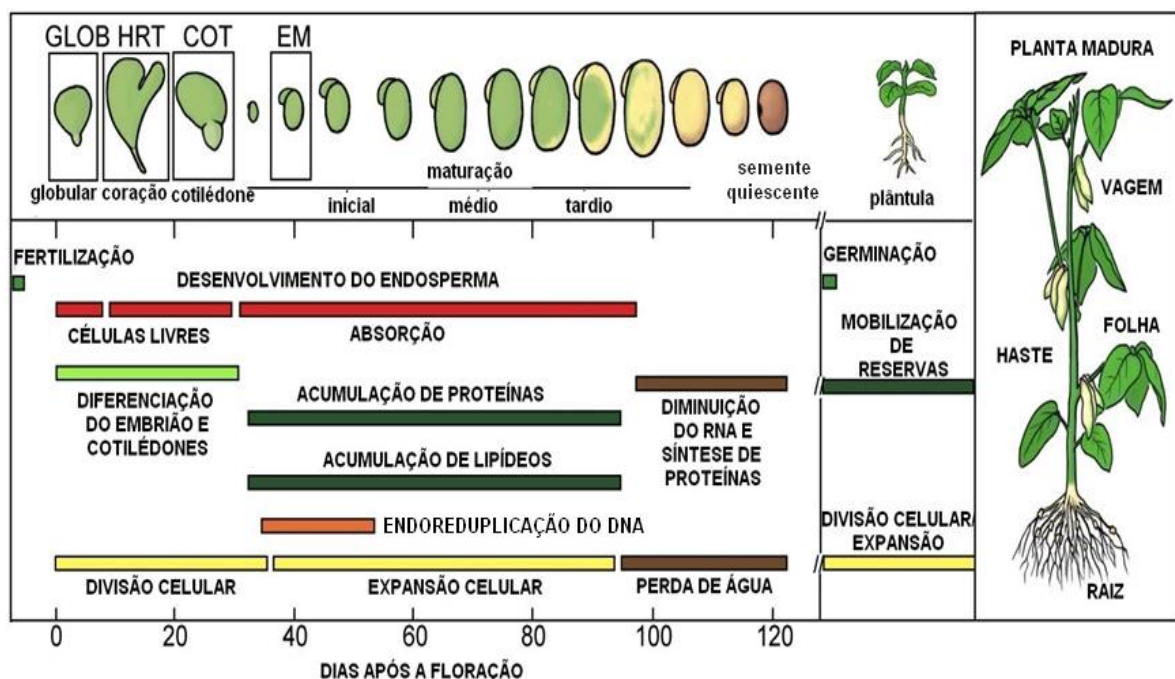
## 2.2. Aspectos fisiológicos da embriogênese e formação da semente de soja

Em angiospermas, a formação da semente depende do processo de reprodução sexual. Em plantas, o processo de reprodução sexual é denominado de dupla fertilização, que ocorre a partir da fusão dos dois núcleos reprodutivos haploides do grão de pólen com o óvulo (Evert e Eichhorn, 2014). O processo inicia-se no momento da polinização, quando o grão de pólen se aloja no estigma da flor. O grão de pólen germina no estigma emitindo um tubo polínico que cresce até o ovário. O tubo polínico infiltra-se no óvulo e libera seus dois núcleos reprodutivos. Um dos núcleos reprodutivos do grão de pólen se funde com a oosfera, um dos núcleos do óvulo, gerando um zigoto diploide (2n) que futuramente formará o embrião da semente. O outro núcleo reprodutivo se funde com os dois núcleos



polares do óvulo, formando uma célula triploide ( $3n$ ) a partir da qual formara o futuro endosperma da semente. O embrião da semente formará a nova planta enquanto que o endosperma fornecerá nutrientes para o desenvolvimento do embrião e para a futura plântula (Bewley et al., 2013).

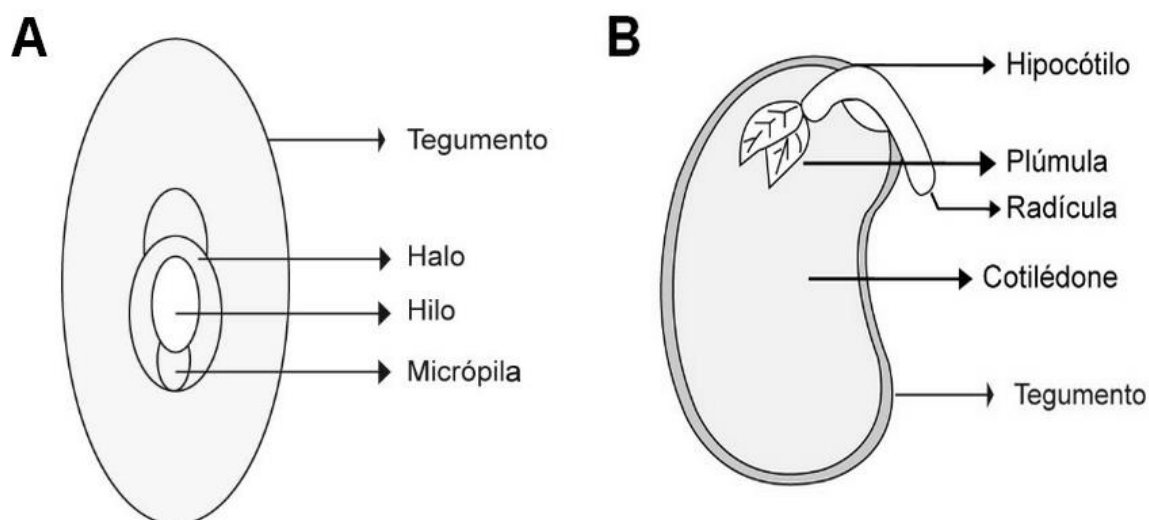
Depois da fertilização, inicia-se o processo de formação do embrião, a embriogênese. Esse processo envolve inúmeros eventos fisiológicos, desde a divisão, expansão e diferenciação celular, deposição de reservas até o dessecamento da semente recém-formada (Figura 3). A primeira fase da embriogênese é chamada de fase globular, onde a célula apical do zigoto inicia uma série ordenada de divisões, gerando um embrião globular com um endosperma acelular em intensa divisão e um sistema vascular começando o seu desenvolvimento (Goldberg et al., 1994). Em soja, a fase de coração inicia-se de 8 a 10 dias após a fertilização. Nessa fase, com o endosperma quase inteiramente celular, é estabelecida a bilateralidade dos embriões com a formação dos cotilédones (Ritchie et al., 1985).



**Figura 3.** Eventos fisiológicos que ocorrem durante o desenvolvimento e maturação da semente de soja. A figura apresenta os diversos estágios de desenvolvimento da semente de soja juntamente com todos os processos morfológicos e fisiológicos que ocorrem no interior da mesma Adaptado de Seed Gene Network (2016).

Ao longo do tempo, devido à intensa divisão mitótica, os cotilédones estabelecidos anteriormente alongam-se. De 30 a 50 dias, as proteínas e os lipídios começam a acumular-se nesses cotilédones, é o início do processo de maturação da semente. A fase intermediária de maturação ocorre entre 60 a 80 dias após a fertilização. Nessa fase, a semente aumenta de tamanho devido, principalmente, à expansão das células e à deposição de reservas proteicas e lipídicas. Por volta de 90 a 120 dias após a fertilização, o alargamento do embrião cessa, o grau de umidade que permaneceu constante durante todo o processo decresce, o que acarreta na diminuição do metabolismo da semente recém-formada e o embrião passa para um estado de metabolismo mínimo ou estado de quiescência. Nesse estágio, a semente está completamente madura e preparada para resistir a condições adversas (Bewley et al., 2013).

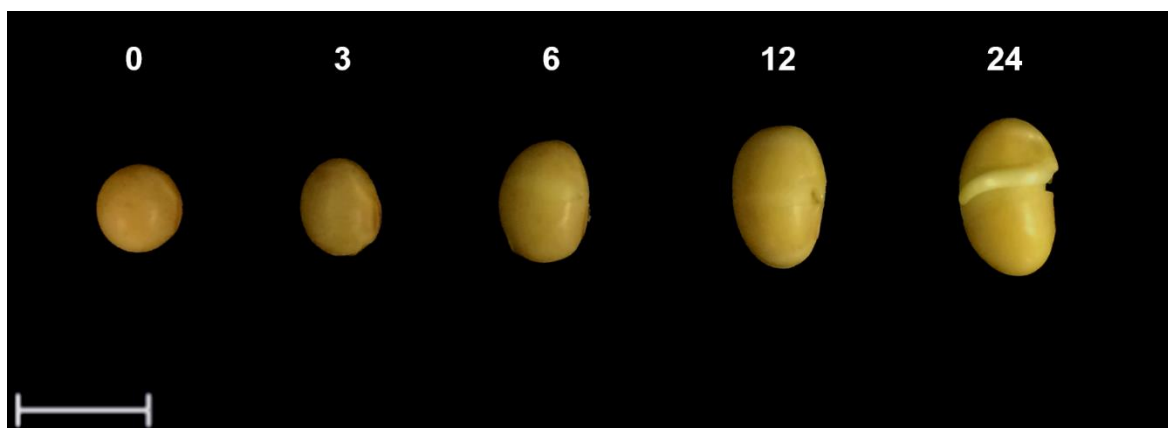
Na maturidade, a semente (óvulo maduro) é composta basicamente por uma cobertura protetora (tegumento) que delimita a semente, um tecido meristemático (eixo embrionário) e um ou dois cotilédones, dependendo da biologia da planta, como pode ser observado na figura 4.



**Figura 4.** Ilustração esquemática da parte externa e interna da semente de soja. A – Imagem ilustrativa da estrutura externa da semente dando enfoque as suas principais estruturas. B – Imagem ilustrativa da estrutura interna da semente dando enfoque as suas principais estruturas. Adaptado de AGEITEC (2016).

O tegumento possui função protetora, serve como barreira à entrada de microrganismos, protege as partes internas da semente contra choques e, ainda, regula a velocidade das trocas gasosas e a reidratação da semente durante o processo de germinação (Mohamed-Yasseen et al., 1994). O eixo embrionário é a parte vital da semente, contém tecidos meristemáticos que apresentam capacidade de divisão. A partir dele se desenvolverão o caule, as raízes e folhas da futura planta. Os cotilédones acumulam reservas que serão usadas durante e após a germinação da semente (Goldberg et al., 1994). Em leguminosas, como a soja, as proteínas, óleos e amidos que são armazenados nos cotilédones são as principais fontes de alimentos fornecidos por sementes para consumo humano e animal (Grau et al., 2004). A composição química da semente depende basicamente das características genéticas e das condições climáticas. A semente de soja possui, em matéria seca, aproximadamente 40% de proteínas, 20% de óleos, 35% de carboidratos e 5% de minerais (Messina, 1999).

Depois de formada, a semente pode germinar, dependendo das condições ambientais. O processo de germinação da semente retoma o desenvolvimento do embrião. A germinação inicia-se com a embebição, a quantidade de água absorvida pela semente deve ser o suficiente para iniciar e continuar todo o processo de germinação. A absorção de água acarreta na ativação dos processos bioquímicos preparatórios, onde ocorre a digestão, translocação e assimilação das reservas acumuladas ao longo da formação da semente. O processo de digestão das reservas consiste em uma intensa atividade de várias enzimas, incluindo as lipases, proteases e peptidases. No caso das lipases, os lipídeos são digeridos para formação de ATP e de ácidos graxos para a formação de glucídios e, posteriormente, sacarose e ATP. Depois de digeridas, as reservas são translocadas para o eixo embrionário. No embrião, as reservas são assimiladas e usadas como fonte de energia para modificação dos tecidos existentes e formação de novos o que culmina na expansão celular e emergência da plântula (Figura 5). O processo pode estender-se de horas a dias, dependendo da espécie (Nonogaki et al., 2010).



**Figura 5.** Sementes de soja nas primeiras horas de germinação. 0 – Semente quiescente. 3, 6, 12 e 24 – Horas após a embebição da semente em água.

### 2.3. Parede Celular Vegetal

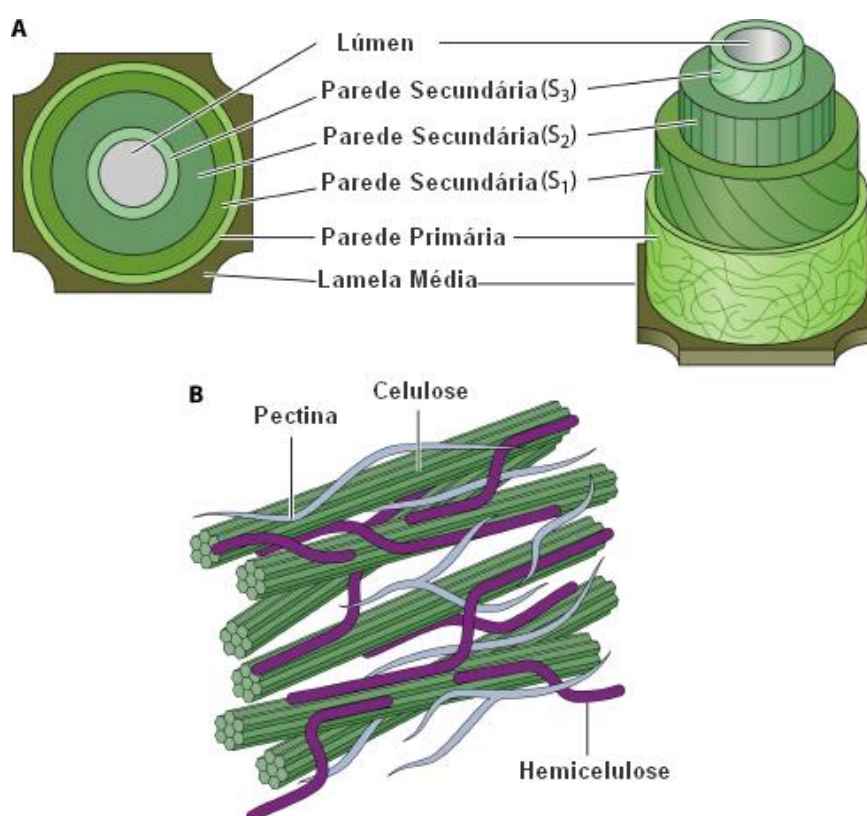
Por muito tempo persistiu a ideia de que a parede celular vegetal possui uma estrutura química inalterável ao longo da vida da planta (Wilder e Albersheim, 1973; Fry, 1988). Contudo, hoje já é sabido que os componentes da parede celular se alteram durante os diversos estágios de desenvolvimento da planta e podem apresentar composição distinta dependendo do tecido. Por isso, a parede celular pode ser considerada um compartimento metabólico ativo (Brown, 2004; Lashbrook e Cai, 2008; Fincher, 2009; Krzesłowska, 2011; Tenhaken, 2015).

Dentre os diversos aspectos das plantas, a parede celular define a rigidez, morfologia e tamanho da célula, controle da expansão celular, transporte e comunicação intercelular e confere proteção à uma enorme quantidade de patógenos (Cosgrove, 2005). Em adição a essas funções biológicas, a parede celular também é importante para economia humana, sendo utilizada na forma de papéis, fibras, géis, revestimentos, linhos, entre outros produtos (Buckeridge e Tiné, 2001).

A parede celular vegetal fica localizada na matriz extracelular que é composta basicamente por água, um dos determinantes das propriedades físicas da parede, e compostos solúveis. Morfologicamente, a parede celular é composta por três diferentes estruturas: a lamela média, a parede primária e a parede secundária (Figura 6A). A estrutura e composição de cada componente da parede

celular pode variar de acordo com o tipo celular e conforme o desenvolvimento da célula.

A lamela média, constituída de substâncias pécicas, é uma membrana elástica e permeável que conecta as células entre si (Zamil et al., 2014). A parede primária é composta por: microfibrilas de celulose, estruturas rígidas que ficam embebidas em uma matriz de polissacarídeos e fornecem força, resistência e flexibilidade a célula; hemiceluloses, polissacarídeos flexíveis que se ligam a superfície das microfibrilas de celulose formando uma rede forte e flexível; pectinas, que agem como um material de preenchimento para evitar o colapso da rede de celulose; e proteínas estruturais que podem fornecer resistência mecânica à parede (Figura 6B). Quantidades pequenas de minerais e ésteres também estão presentes. Como a parede primária é formada por células em divisão, ela pode se alongar em um crescimento difuso ou orientado (Cosgrove, 2005).



**Figura 6.** Modelo simplificado da estrutura da parede celular da planta. (A) A estrutura consiste em três camadas principais: a lamela média e a parede primária e secundária. (B) Os principais polissacarídeos que formam a estrutura da parede. Adaptado de Hild (2014).

Quando o crescimento celular cessa tem-se início o processo de diferenciação celular. Juntamente com esse processo inicia-se a formação da parede celular secundária, onde ocorre um aumento na síntese de pectina e um grande aumento na síntese de celulose, hemiceluloses e lignina (Carpita e Gibeaut, 1993a) . Dependendo do tipo celular, as células vegetais podem ou não possuir uma parede secundária. Células em crescimento possuem apenas a parede primária (Buckeridge, 2010). A parede celular secundária é formada por três camadas distintas: S1, S2 e S3 (Figura 6A) que se diferenciam em espessura, composição e orientação das microfibrilas de celulose. A parede secundária se diferencia da parede primária por possuir uma quantidade maior de cadeias de celulose em cada microfibrila (Cosgrove, 2005).

### **2.3.1. Biossíntese dos componentes estruturais da parede celular**

Os polissacarídeos estruturais da parede celular possuem um importante papel durante o processo de germinação da semente, onde são degradados e seus produtos utilizados como fonte de carbono e energia para o crescimento inicial das plântulas (Buckeridge et al., 2000). Além dessa função vital na vida da planta, os polissacarídeos da parede celular apresentam outras diversas funções essenciais durante todo o ciclo de vida das plantas. Entre elas, podemos citar a participação no controle da expansão celular, suporte estrutural da célula e proteção (Buckeridge et al., 2000).

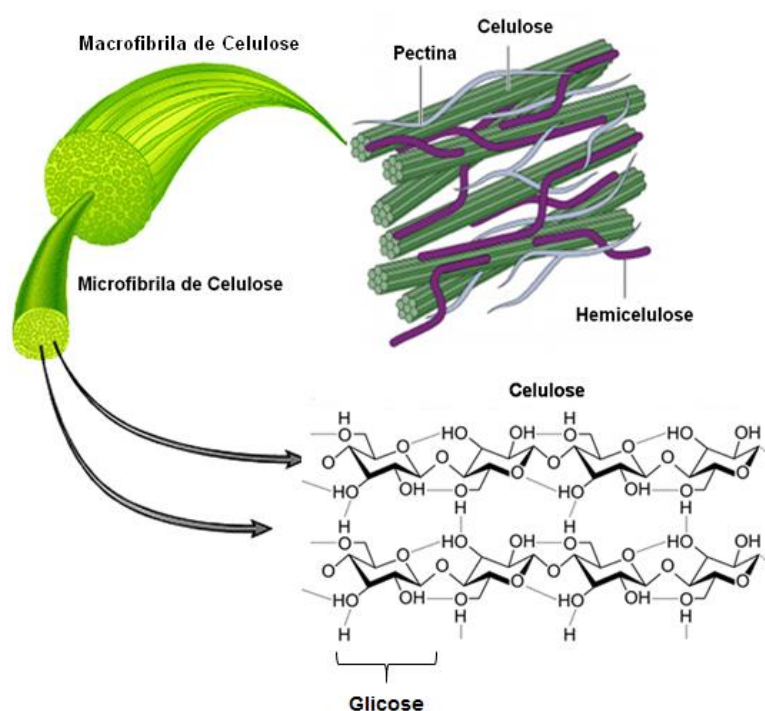
#### **- Celulose**

A celulose é o mais conhecido de todos os polissacarídeos da parede celular. Correspondendo a cerca de 30%-40% da massa seca da parede, ela é responsável por controlar o tamanho e a forma das células. Além de ser o material de construção primário da parede celular, a celulose possui usos industriais como, por exemplo, na fabricação de papéis, filmes, plásticos, entre outros (Bemiller e Lafayette, 2001).

A formação da celulose ocorre a partir da construção e cristalização de dezenas de cadeias lineares de glucano unidas por ligações tipo  $\beta$ -(1,4) (Figura 7).

Entre e dentro dessas cadeias formam-se ligações de hidrogênio, o que garante uma maior interação e resistência. A interação entre as cadeias de celulose forma uma estrutura denominada microfibrila de celulose. O conjunto de microfibrilas é denominado de macrofibrila. O agrupamento das macrofibrilas formam o feixe de celulose.

A síntese das microfibrilas de celulose ocorre em um complexo denominado Celulose Sintase. Esse complexo é organizado em arranjos hexaméricos denominado rosetas (Figura 8). Cada roseta possui seis proteínas celulose-sintase (CESA) que ficam embebidas na membrana plasmática (KIMURA et al., 1999). A CESA forma um complexo com o substrato uridina-difosfato-glicose (UDP-Glicose) e catalisa a transferência de glicose para a polimerização da cadeia linear de glucano crescente (Burton, 2006).

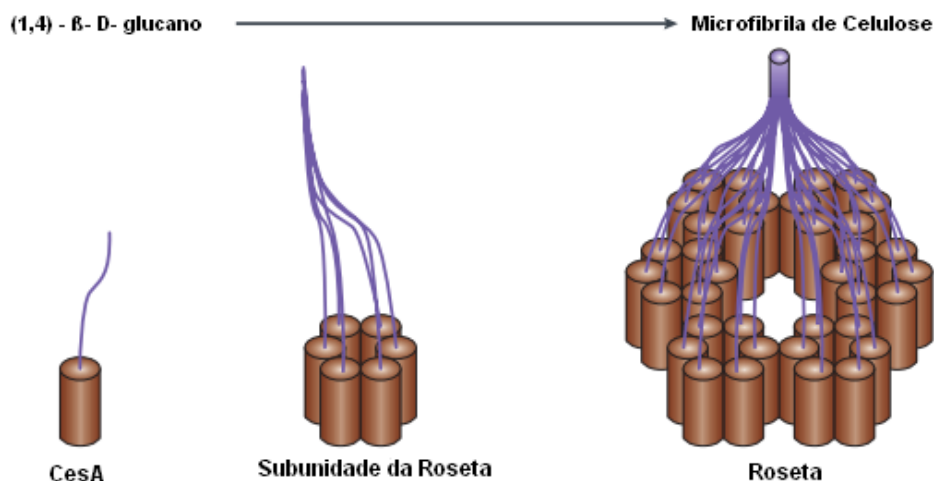


**Figura 7.** Modelo esquemático da estrutura da celulose encontrada na parede celular da planta. Dezenas de cadeias lineares de glucano formam a fibra de celulose, a interação e união dessas fibras formam as microfibrilas e um conjunto de microfibrilas formam a macrofibrila estrutural da parede. Adaptado de Hild (2014).

A proteína CESA é codificada por genes da superfamília celulose sintase (CesA). Os genes CesA já foram caracterizados em espécies como *Populus* (Suzuki et al., 2006), *Physcomitrella* (Roberts e Bushoven, 2007) e soja (Nawaz et al., 2017). A espécie melhor caracterizada é *Arabidopsis thaliana*, onde existem duas subfamílias CesA com funções distintas. As CesA 1, 3 e 6 são responsáveis pela formação de celulose em paredes primárias e as CesA 4, 7 e 8 são responsáveis pela formação de celulose em paredes secundárias (Endler e Persson, 2011) .

### - Hemicelulose

Na parede celular, a celulose liga-se por meio de ligações de hidrogênio a outro grupo de polissacarídeos, as hemiceluloses. As hemiceluloses compreendem um conjunto de polissacarídeos, caracterizados por possuir uma cadeia de glicose, manose ou xilose com ligações do tipo  $\beta$ -(1 $\rightarrow$ 4) e algumas ramificações laterais. A função central da hemicelulose é fortificar a parede celular pela interação com a celulose.. Os grupos de hemiceluloses podem ser divididos em: xiloglucanos, xilanos, mananos, glicomananos e  $\beta$ -(1,3-1,4)-glucanos (Scheller e Ulvskov, 2010).



**Figura 8.** Modelo de como o complexo CESA sintetiza uma microfibrila de celulose. Cada proteína CESA sintetiza uma única cadeia de D-glucano ligada  $\beta$ -(1,4). A celulose é formada como uma fita cristalina que é composta por muitos desses glucanos. No modelo, 36 cadeias de  $\beta$  1,4-D-glucano formam a roseta, que é composta por um hexâmero de hexâmeros CESA. Adaptado de Doblin (2005).



Os xiloglucanos (XyG) controlam a extensibilidade da parede e, portanto, são componentes indispensáveis para a estrutura da mesma (Park e Cosgrove, 2015). Grande parte dos XyG ocupam o espaço entre as microfibrilas de celulose, podendo separá-las, o que promove a extensão da parede. Além disso, os XyG ligam-se por ligações de hidrogênio às microfibrilas de celulose, fornecendo força de tensão para a parede. Em células em crescimento, os XyG estão envolvidos na reestruturação da parede. Neste caso, os XyG são quebrados por endoglicosiltransferases e remontados em diferentes pontos da célula, de modo que ocorra a expansão celular sem comprometer a resistência mecânica da parede (Darley et al., 2001; Cosgrove, 2016b).

Os XyGs ocorrem em maior proporção na parede celular primária e estão presentes na parede celular de todas as plantas terrestres (Peña et al., 2008). Os XyGs compreendem um conjunto de polissacarídeos que possuem uma cadeia principal de glucano com ligações tipo  $\beta$ -1,4 que pode ser substituído por diversos resíduos de glicosil e xylosil. O tipo e a ordem dos resíduos substituintes da cadeia do XyG depende da espécie, do tipo de célula, do tipo de tecido e do estágio de desenvolvimento do tecido (Schultink et al., 2014; Leroux et al., 2015; Pauly e Keegstra, 2016). O substituinte mais regular da cadeia XyG são resíduos de  $\alpha$ -D-xilopiranosose ligados na posição O-6 ao esqueleto glucano (Pauly e Keegstra, 2016).

A biossíntese de XyG ocorre no complexo de Golgi, onde muitas enzimas participam da sua síntese. A cadeia central do XyG é sintetizada por membros de proteínas semelhantes à família CESA, a Cellulose Synthase – Like (CSL). Essa semelhança se deve à similaridade da cadeia central de ambos. Arabidopsis possui cinco membros da família CSL (Bernal et al., 2008; Liepman e Cavalier, 2012). As ramificações na cadeia central de XyG são adicionadas pelas glicosil-transferases (Família GT). Os membros das famílias GT podem adicionar resíduos de xilose, galactose e fucose dependendo do tipo de substrato de açúcar e do tipo de família GT (Lairson et al., 2008).

Outro componente estrutural das hemiceluloses são os xilanos. Os xilanos aumentam a recalcitração da parede ajudando na defesa contra herbívoros e patógenos (Rennie e Scheller, 2014; Miedes et al., 2014). A estrutura do xilano consiste em um polímero linear de resíduos de xilose com ligações do tipo  $\beta$ -1,4. Nessa estrutura podem ocorrer ramificações por cadeias laterais com resíduos de ácido 4-O-metil glicurônico (glicuronoxilano, GX), arabinose (arabinoxilano), ou

uma combinação dos dois (glicuronoarabinoxilano GAX). Existem variações nas estruturas do xilano entre diferentes espécies e entre os diferentes tecidos na mesma espécie. Nas dicotiledôneas, o xilano é a hemicelulose predominante nas paredes celulares secundárias e pouco encontrado nas paredes primárias (Rennie e Scheller, 2014). Diferentemente dos xiloglucanos, a estrutura dos xilanos não é repetitiva (Scheller & Ulvskov, 2010).

Devido à semelhança estrutural do xilano com as cadeias das outras hemiceluloses, tem sido amplamente assumido que a sua biossíntese envolve membros das famílias CSL (Bernal et al., 2007). Porém, os estudos das proteínas CSL não forneceram forte evidência do seu envolvimento na biossíntese dos xilanos (Carpita, 2011). Em vez disso, a caracterização de alguns mutantes indicou que os GT pertencentes às famílias GT43 e GT47 são responsáveis pelo alongamento da cadeia dos xilanos (Brown et al., 2007; Scheller e Ulvskov, 2010)

Outro grupo de polissacarídeos com ligação do tipo  $\beta$ -1,4 são os mananos e glicomananos. Sua cadeia principal pode ser composta apenas por resíduos de manose, sendo chamados de mananos ou galactomananos (ramificações com resíduos de galactose). Quando composta por manose e glicose em um padrão não repetitivo são denominados de glicomananos e galactomananos. Os glicomananos são frequentemente acetilados. A biossíntese dos membros do grupo de mananos e glicomananos ocorre por membros do subgrupo A da família CSL. O substrato utilizado para a síntese são os precursores GDP-manose e GDP-glicose (Carpita e McCann, 2002; Silva et al., 2011). Os mananos exercem um papel importante como compostos de armazenamento de sementes (Scheller e Ulvskov, 2010).

Os  $\beta$ -(1,3-1,4) –glucanos, outro grupo de hemiceluloses, são bem conhecidos em gramíneas (Henry, 1987; Carpita, 1996). Porém, não foram encontrados em dicotiledôneas (Carpita e Gibeaut, 1993b) e, por isso, não serão descritos aqui.

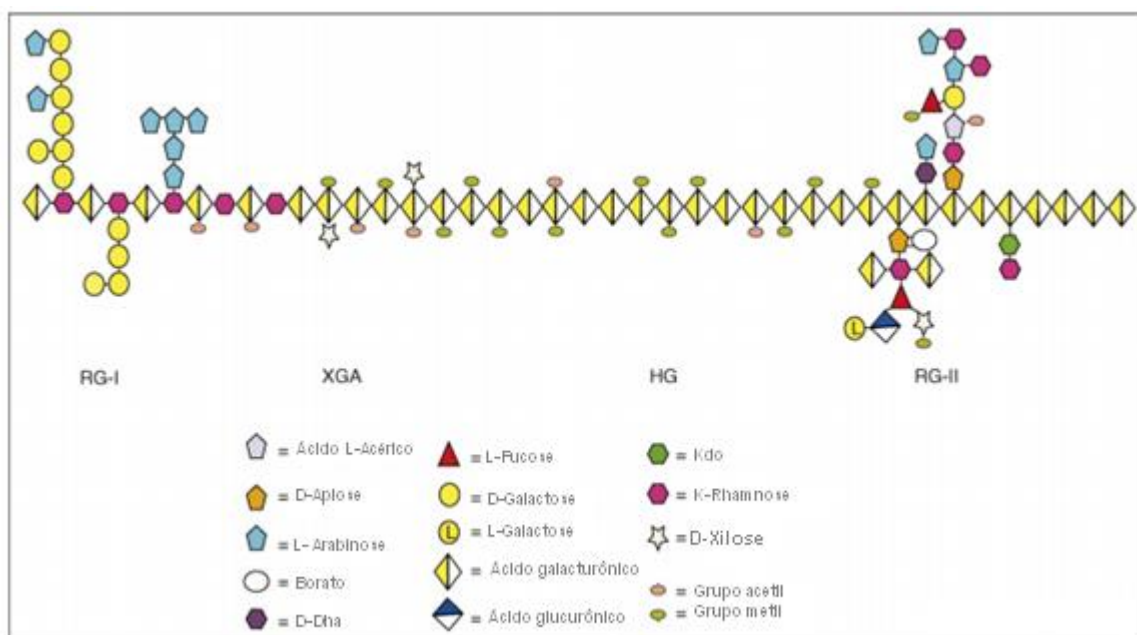
#### - Pectinas

As pectinas são um grupo de polissacarídeos ligados covalentemente, altamente hidratados, que ficam mergulhados na matriz da parede. Várias evidências indicam que as pectinas exercem um importante papel no crescimento, desenvolvimento, morfogênese, defesa, adesão célula-célula, estrutura da parede,

sinalização, expansão celular, porosidade da parede, ligação de íons, hidratação de sementes, abscisão de folhas e desenvolvimento de frutos (Mohnen, 2008; Gwanpua et al., 2016; Zega e D'Ovidio, 2016; Wallace e Williams, 2017). A pectina também é usada como agente gelificante e estabilizante nas indústrias de alimentos e cosméticos (Willats et al., 2001; Domozych et al., 2007; Mohnen, 2008; Cosgrove, 2016a). Todos os polissacarídeos pécticos contêm ácido galacturônico ligado na posição O-1 e O-4 e podem ser divididos em ramnogalacturanos I e II (RGI e RGII), homogalacturanos (HG) e xilogalacturanas (XGA) (Figura 9).

O polissacarídeo péctico mais abundante na parede celular é o homogalacturonano (HG), um homopolímero linear de ácido galacturônico com ligações do tipo  $\alpha$ -1,4. A XGA é um HG substituído na posição O-3 por uma xilose. A XGA é mais prevalente nos tecidos reprodutivos (Zandleven et al., 2007). O RG-I representa 20-35% da pectina, sua estrutura é composta por dissacarídeos formados por resíduos de ácido galacturônico e ramnose, que pode conter diversas ramificações. O RG-II representa cerca de 10% da pectina e é a pectina mais estruturalmente complexa e ramificada. (Mohnen, 2008).

A biossíntese dos polissacarídeos pécticos ocorre no aparelho de Golgi e está associada as GTs, que transferem resíduos glicosílicos de açúcares nucleotídicos (*nucleotide sugar*) (Seifert, 2004). Alguns genes que codificam os membros pécticos já foram identificados. Porém, existem poucos resultados conclusivos a respeito da atividade enzimática das proteínas identificadas (Sterling et al., 2006; Egelund et al., 2006; Mohnen, 2008).



**Figura 9** Estrutura esquemática geral das pectinas mostrando os quatro grupos de polissacarídeos pécnicos, homogalacturonano (HG), xilogalacturonano (XGA), Rhamnogalacturonan I (RG-I) e Rhamnogalacturonan II (RG-II). Adaptado de Mohnen, (2008).

#### - Proteínas estruturais e Proteínas modificadoras de parede celular

Além dos polissacarídeos descritos, a parede celular também é composta por proteínas estruturais e modificadoras da parede. As proteínas estruturais podem ser agrupadas em quatro classes: glicoproteínas ricas em hidroxiprolina (HRGPs – Extensinas); proteínas ricas em arabinogalactanos (AGPs); proteínas ricas em glicina (GRPs) e proteínas ricas em prolina (PRPs). As principais proteínas modificadoras da parede são as pectinases, esterases, expansinas e as xiloglicano endotransglicosilase/hidrolase (XTHs) (Showalter, 1993; Cassab, 1998).

As extensinas ou glicoproteínas ricas em hidroxiprolina (HRGPs) são pentapeptídeos com repetições de serina e hidroxiprolina. Muitas das funções das HRGPs são desconhecidas. Em *Arabidopsis*, as HRGPs são essenciais para o correto desenvolvimento do embrião (Hall e Cannon, 2002). Em soja, os genes codificadores das HRGPs já foram identificados e alguns foram descritos como essenciais para a maturação das raízes primárias (Hong et al., 1990; Hoon Ahn J et al., 1998).

As proteínas ricas em arabinogalactanos (AGPs) são caracterizados por possuírem um teor proteico inferior a 10%, sendo constituídas por 90% de carboidratos. A fração de proteína é rica em resíduos de aminoácidos Hyp, Ala, Thr, Gly e Ser (Clarke et al., 1979). Muitas funções diferentes foram atribuídas as AGPs. Possuem propriedades de agregação e adesividade, fornecem precursores de carboidratos para a parede crescente do tubo de pólen e também podem participar do controle do balanço hídrico na célula (Majewska-sawka e Nothnagel, 2000; Showalter, 2001). Em arábida, os genes que codificam as AGPs já foram identificados e participam da regeneração das raízes, germinação e desenvolvimento da semente, além de promoverem a embriogênese (Schultz, 2000; Van Hengel et al., 2002; Van Hengel e Roberts, 2003). Em soja, já foi descrito que alguns membros das AGPs exercem um papel essencial no desenvolvimento das raízes (Cassab, 1986).

Os GRPs são outra classe de proteínas estruturais de parede. Essas proteínas são compostas por cerca de 68% de Gly e 12% de Ser e são ligeiramente glicosiladas. Além disso, esse grupo de proteínas possui uma estrutura primária repetitiva de Gly (Mousavi e Hotta, 2005; Mangeon et al., 2010). Essas proteínas estão relacionadas à adaptação durante estresses como o frio, adaptação ao ritmo circadiano e controle da alongação das raízes laterais (Carpenter et al., 1994; Kim et al., 2010; Boron et al., 2014). A expressão das GRPs aumenta significativamente em células epidérmicas e nas regiões alongadas e maduras do hipocótilo (Wyatt et al., 1992).

As proteínas ricas em prolina (PRPs) possuem em sua estrutura um motivo repetitivo de -Pro-Hyp-Val-Tyr-Lys- e são levemente glicosiladas (Datta et al., 1989; Marcus et al., 1991). Vários estudos indicam que os PRPs estão envolvidos em vários aspectos do desenvolvimento, que vão desde a germinação até a formação de vagens, até os estágios iniciais da nodulação (Höfte e Voxeur, 2017). E, ainda, o padrão de localização de PRPs é semelhante ao da maioria das GRPs. Em soja, alguns genes da família PRP foram expressos em resposta ao ácido salicílico, estresse por sais e seca e adaptação ao ritmo circadiano (He et al., 2002). Além disso, essas proteínas são capazes de afetar a cor da semente de alguns genótipos de soja (Lindstrom e Vodkin, 1991).

Além das proteínas estruturais, também se encontram na parede celular outras proteínas e compostos capazes de alterar os polissacarídeos da parede ou

conferindo proteção contra patógenos, como por exemplo, pectinases, esterases, endoglucanases, xiloglicano endotransglicosilase, quitinases e  $\beta$ -glucanases. Existem também aqueles compostos capazes de alterar, converter ou modificar os substratos na parede, como por exemplo, fosfatases, invertases, desidrogenases e as peroxidases. As peroxidases são oxidoreduções que catalisam a oxidação de substratos usando peróxido de hidrogênio como doador de elétrons. Além disso, a peroxidase está envolvida em diversas reações como catalisador das ligações dos polissacarídeos, oxidação do ácido indol-3-acético, ligações de monômeros, oxidação de fenóis, regulação da elongação celular, dentre outros (Passardi et al., 2004; Francoz et al., 2015)

Uma classe de proteínas modificadoras da parede bastante estudadas são as expansinas. A expansina age rompendo as ligações de hidrogênio entre as fibras de celulose e rompendo a interação entre a celulose e hemicelulose. Assim, esse grupo de proteínas consegue expandir os feixes da parede celular (Cosgrove, 2000, 2015).

Outra família de genes que participam na regulação da expansão da parede celular e é coexpressa com o complexo CESA, é a família COBRA (Brady et al., 2007). A família COBRA está localizada na membrana plasmática e participa na modulação, montagem e cristalização das cadeias celulósicas individuais (Fernandez et al., 2005; Liu et al., 2013). O gene COBRA codifica uma proteína com uma âncora glicosilfosfatidilinositol (GPI), com um domínio CCVS, um potencial local de N-glicosilação e um possível local de ligação de celulose (Roudier, 2002).

Os componentes da parede celular puderam ser descritos até hoje quanto a sua estrutura química, processo de biossíntese e como essas estruturas interagem entre si formando todo o compartimento da parede celular. Porém, pouco se sabe sobre a regulação e montagem desses compartimentos nas células em desenvolvimento.

#### **2.4. Transcriptoma: RNA-Seq**

O produto inicial da expressão do genoma é o transcriptoma, uma coleção de todos os genes expressos em determinado estágio de desenvolvimento. O

transcriptoma de um organismo varia de acordo com a fase de desenvolvimento e as condições ambientais. Através da análise do transcriptoma é possível determinar quando e onde cada gene é expresso em determinado organismo e ainda, é possível quantificar essa expressão (Rapaport et al., 2013).

Várias tecnologias como, por exemplo, RT-PCR (*reverse transcription polimerase chain reaction*), microarranjos de DNA e análise serial da expressão gênica (SAGE) foram desenvolvidas a fim de caracterizar parte ou todo o transcriptoma de organismos potencialmente úteis na medicina e agricultura. (Lowe et al., 2017) Uma das tecnologias mais recentes é o RNA-seq, uma ferramenta altamente sensível para medir a expressão do transcriptoma em diferentes tecidos e estágios de desenvolvimento. Essa técnica também é capaz de detectar genes expressos em baixos níveis e fatores de transcrição com alta expressão (Wang et al., 2010).

A metodologia do RNA-seq consiste em diversas etapas iniciando com a preparação da biblioteca, onde o mRNA será extraído, isolado e fragmentado. Em seguida, ocorre a síntese de cDNA, onde os adaptadores são ligados por complementaridade de bases em ambas as extremidades dos fragmentos da amostra. Após a preparação da biblioteca, são realizados os ciclos de amplificação em ponte para que ocorra a síntese da segunda fita do fragmento. Os fragmentos amplificados, *amplicons*, são incorporados no suporte e submetidos aos ciclos de desnaturação para a separação dos dúplex formados durante a amplificação. Durante esse ciclo as fitas são separadas e linearizadas. Estes passos são repetidos sucessivamente e, assim, são geradas cerca de mil cópias de cada fragmento que formam grupos ou *clusters* de sequenciamento de determinado fragmento. Após a formação dos *clusters*, iniciam-se os ciclos automatizados. Nos próximos ciclos de amplificação automatizada, nucleotídeos terminadores marcados são introduzidos na reação (Nagalakshmi et al., 2010). Atualmente, o equipamento NextSeq series produz até 400 milhões de leituras (*reads*) por corrida (Illumina, 2017). Após o sequenciamento, os *reads* passam por uma avaliação de qualidade, utilizando métodos computacionais. Em seguida, são alinhados em um genoma ou transcriptoma de referência ou realizando a montagem "*de novo*".

O RNA-seq tem sido utilizado em vários estudos de análise da expressão gênica de plantas de interesse econômico, como o milho (Kakumanu et al., 2012), cana-de-açúcar (Manners e Casu, 2011) e trigo (Duan et al., 2012). Em soja, alguns

estudos utilizando esta tecnologia já foram realizados. Lambirth *et al.*, (2015) realizou uma comparação de alguns cultivares de sementes transgênicas de soja em relação às selvagens, analisando o perfil de ambas utilizando o RNA-seq e evidenciando as mudanças na expressão gênica das diferentes sementes. A análise do transcriptoma também pode ser usada para evidenciar quais genes estão relacionados às condições de estresse nas plantas. Segundo o trabalho realizado por Rodrigues *et al.*, 2015, durante o estresse hídrico, as redes de expressão gênica podem ser reguladas de forma dinâmica para aclimatar as plantas às condições de estresse. Os autores identificaram genes que são expressos sob condições normais de desenvolvimento e genes cuja expressão oscila sob estresse. Já Marcolino-gomes *et al.*, (2015) identificaram na base de dados do transcriptoma da soja genes de referência para análise da resposta da planta à seca e encontraram genes mais estáveis em condições de seca ao longo do dia. E, ainda, Le *et al.*, (2012) analisaram o transcriptoma de soja e encontraram uma expressão gênica diferencial de folhas em estágio final de desenvolvimento sob estresse hídrico. Os autores também identificaram genes candidatos a resistência à seca para um futuro desenvolvimento de cultura de soja tolerante a esse estresse.

Além da análise de expressão gênica, a tecnologia do RNA-seq pode melhorar significativamente a anotação dos genes e fornecer recursos valiosos para estudos funcionais. Um guia do transcriptoma de soja, que reúne um conjunto de dados de RNA-seq de diferentes tecidos, foi desenvolvido para facilitar a compreensão do genoma da espécie, permitindo, assim, que os pesquisadores gerem avanços tecnológicos à cultura (Severin *et al.*, 2010).

Em parede celular, foi realizado um estudo examinando o transcriptoma do revestimento da semente em três fases de desenvolvimento da semente com fenótipo normal e fenótipo defeituoso. O estudo revelou 364 genes expressos diferencialmente entre as duas variedades. Os genes responsáveis por processos referentes à parede celular compreenderam 19% dos genes expressos. Dentro desses, os genes *proline-rich* (PRP) e *glycine-rich* (GRP), constituintes da parede celular, foram altamente diferenciais. Esses dados revelam que genes de parede celular podem alterar o tegumento de sementes de soja durante o seu desenvolvimento (Kour *et al.*, 2014).



Em relação ao desenvolvimento da semente de soja, Jones e Vodkin (2013) fizeram uma análise do desenvolvimento da semente de soja a partir de alguns dias após a fertilização da semente até a sua maturação. Foram utilizadas sete fases diferentes do desenvolvimento da semente produzindo de 12 milhões a 78 milhões de transcritos sequenciados. Vários genes foram identificados com expressão exclusiva nas primeiras fases de desenvolvimento da semente. Genes que codificam proteínas de armazenamento obtiveram maior expressão na fase de maior peso fresco, confirmando que esses produtos se acumulam nas fases iniciais de desenvolvimento da semente. Outros genes mostraram uma expressão elevada em sementes secas, indicando a preparação das vias necessárias para os estágios iniciais de germinação.

Durante o processo de germinação e o desenvolvimento do eixo embrionário, órgão vital para a futura plântula, já foram relatadas informações sobre a regulação bioquímica que ocorre desde a embebição da semente até a sua germinação. Entre esses processos ocorre a glicólise e o remodelamento da parede celular (Bellieny-Rabelo et al., 2016). Apesar da sua inerente importância, trabalhos relacionando a análise transcriptômica com a rede de regulação da parede celular em soja ainda são escassos.

### **3. OBJETIVOS**

#### **3.1. Objetivo geral**

Identificar e caracterizar as variações de expressão dos genes envolvidos com a biossíntese da parede celular vegetal em diversos estágios do desenvolvimento de soja e na regulação de parede celular em eixo embrionário durante a germinação de sementes de soja.

#### **3.2. Objetivos específicos**

-Identificar e anotar os genes relacionados com a parede celular no genoma de soja.

-Analisar mapas transcricionais de genes envolvidos com a parede celular quanto a sua função durante os diferentes estágios do desenvolvimento das sementes de soja.

-Avaliar a expressão de genes-chave de parede celular em sementes de soja durante as etapas de desenvolvimento e germinação.

-Caracterizar morfológico-tecidualmente a parede celular de eixo embrionário de sementes de soja durante a etapa de germinação.

-Analisar a atividade da enzima peroxidase envolvida com o metabolismo da parede celular em eixo embrionário de soja durante a germinação.

## 4. MATERIAIS E MÉTODOS

### 4.1. Anotação dos genes relacionados com a parede celular no genoma de soja

Para anotar os genes relacionados com a parede celular no genoma de soja, foi realizada uma busca de genes relacionados a parede celular no genoma de soja disponível no banco de dados Phytozome9 (<http://www.phytozome.net/>) com um total de 73.320 transcritos e 54.154 genes. Utilizamos como referência sequências gênicas relacionadas à parede celular já anotadas no organismo modelo, *Arabidopsis thaliana*, no banco de dados Cell Wall Genomics (<https://cellwall.genomics.purdue.edu>) (Yong et al., 2005), que fornece uma caracterização dos fenótipos e vias regulatórias das paredes celulares de mutantes de parede celular de *Arabidopsis*. As sequências gênicas de parede celular em *Arabidopsis* compreendem um total de 1.080 genes.

As sequências gênicas obtidas no genoma de soja foram então anotadas manualmente de acordo com a sua via metabólica de parede celular correspondente, a fim de construir uma lista de genes de parede celular de soja.

#### 4.1.1. Genes COBRA

Além de anotar os genes referentes às vias gerais de parede celular no genoma de soja, outro objetivo do trabalho foi analisar uma família gênica de biossíntese da parede celular. Nessa busca, escolhemos a família COBRA por ser

co-expressa com o complexo celulose sintase, por ser descrita na literatura como essencial para a formação de vários órgãos das plantas e, por ainda, não existir a identificação dessa família em soja.

#### **4.1.1.1. Identificação e anotação dos genes COBRA no genoma de soja**

Para identificar e anotar os genes COBRA em soja, aqui denominados de GmCOBRA, foi utilizada a versão v2.0 do genoma de *Glycine max* disponível no *Phytozome11* (<https://phytozome.jgi.doe.gov/>). Foram utilizados os genes COBRA já identificados em *Arabidopsis* como *query* (sequência a ser comparada). Para tal, foi utilizado o programa de alinhamento local *BLASTp* (<http://blast.ncbi.nlm.nih.gov/blast.cgi>) para comparar por homologia as sequências de *Arabidopsis* com as sequências de soja. O valor de expectativa (*e-value*) de  $1e-10$  foi utilizado como parâmetro de corte.

As sequências que obtiveram *hits* (sequências que alinharam com a *query*) superiores ao *e-value* utilizado como limite de corte foram selecionadas. Todas as informações relacionadas as sequências, incluindo a sequência de genes, sequências de proteínas e anotações foram adquiridas no *Phytozome v11*.

A ferramenta *Pfam* (<http://pfam.xfam.org/>) foi utilizada para confirmar a presença do domínio COBRA nas sequências. As sequências sem o domínio foram desconsideradas nas análises posteriores. O peptídeo sinal, típico de proteínas COBRA, foi verificado com a ferramenta SignalP 3.0 (<http://www.cbs.dtu.dk/services/SignalP/>) e as possíveis modificações GPI, típicas de proteínas ancoradas na membrana celular, foram verificadas com a ferramenta BigPI ([http://mendel.imp.ac.at/gpi/gpi\\_server.html](http://mendel.imp.ac.at/gpi/gpi_server.html)).

#### **4.1.1.2. Análise da estrutura gênica e distribuição dos genes COBRA nos cromossomos de soja**

A estrutura gênica dos genes COBRA anteriormente identificados foi construída com base na sequência de codificação, na fase de íntrons, no comprimento e no número do éxons utilizando a ferramenta *Gene Structure Display Server* (GSDS) (Hu et al., 2015).

Todos os GmCOBRA foram mapeados nos cromossomos de soja de

acordo com as informações de distribuição, tamanho inicial e terminal dos genes e comprimento dos cromossomos depositados no banco de dados Phytozome11. Nesse processo, foi utilizado a ferramenta *PhenoGram Plot* (<http://visualization.ritchielab.psu.edu/phenograms/plot>).

#### 4.1.1.3. Análise filogenética da família COBRA

Para investigar as relações filogenéticas e a evolução molecular da família COBRA em soja, foi realizado um alinhamento com um total de 72 sequências de proteínas COBRA de *Arabidopsis thaliana*, *Vitis vinifera*, *Medicago truncatula*, *Phaseolus vulgaris* e *G. max* utilizando a ferramenta de alinhamento de sequências, *MUSCLE*, no *software Molecular Evolutionary Genetics Analysis 7.0* (MEGA 7) (Kumar et al., 2016). Foi gerada uma árvore filogenética do modelo evolutivo Neighbor-Joining (BioNJ) com 1.000 repetições, no MEGA7. A árvore filogenética final foi visualizada e editada no *software FigTree v1.4.3* (<http://tree.bio.ed.ac.uk/software/figtree/>).

#### 4.1.1.4. Perfil de expressão dos genes cobra em soja

A fim de analisar o padrão de expressão dos GmCOBRA identificados, foi utilizada a ferramenta *Phytomine* (<https://phytozome.jgi.doe.gov/phytomine/begin.do>), que possui uma coleção de dados de expressão de RNA-seq em diversos tecidos de soja. Os valores de FPKM (Fragmentos por Kilobase por Milhão de Reads Mapeados) para cada gene em todas as amostras de tecidos foi utilizado para gerar o *heatmap*. A ferramenta de análise SEA com teste de significância de Fisher com multicorreção FDR <0.05 foi utilizada para análise de enriquecimento de ontologia gênica na plataforma agriGO (<http://bioinfo.cau.edu.cn/agriGO/index.php>) (Du et al., 2010).

#### 4.1.1.5. Duplicação gênica e estimativa das taxas de substituição sinônimas e não-sinônimas dos genes COBRA em soja

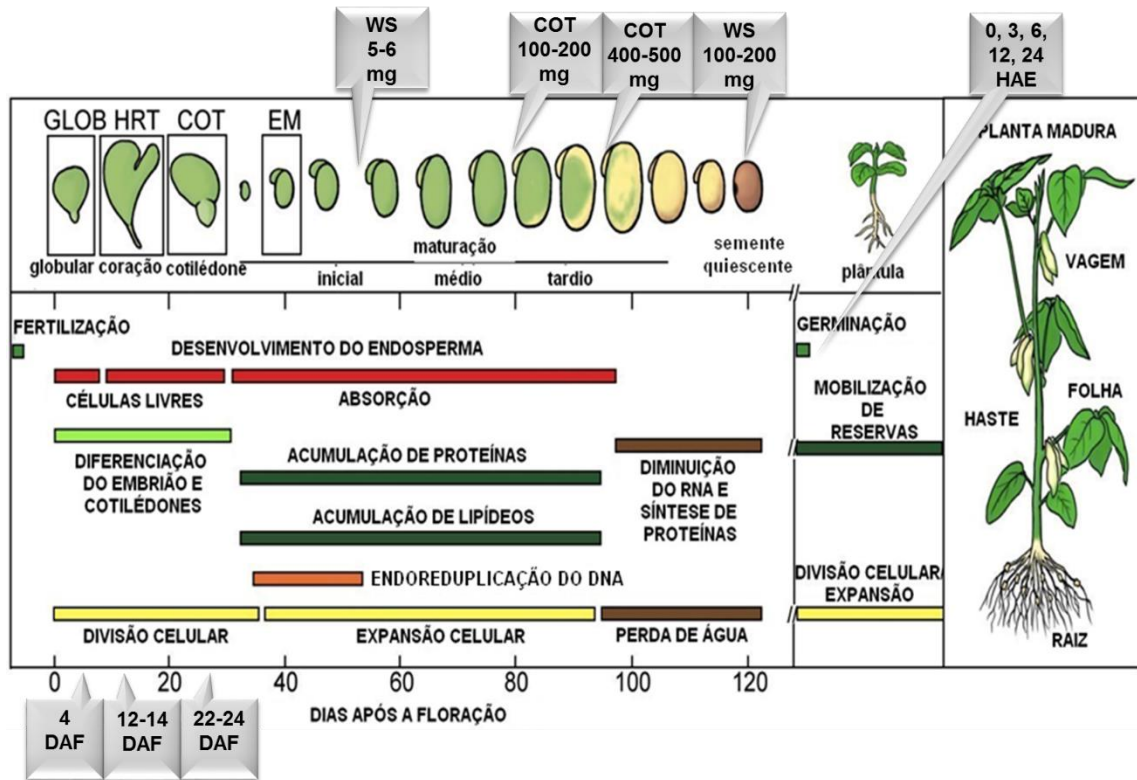
A relação das taxas de substituição sinônimas (Ka) e não-sinônimas (Ks) durante a evolução foi usada para inferir os mecanismos de evolução da duplicação de genes em soja, após a divergência de seus ancestrais. Um valor de  $Ka/Ks = 1$  sugere seleção neutra, um valor  $Ka/Ks < 1$  sugere seleção negativa e um valor  $Ka/Ks > 1$  sugere uma seleção positiva.

O *software* KaKs\_calculator (Zhang et al., 2006) foi utilizado para investigar a divergência dos GmCOBRA e estimar as substituições não-sinônimas (Ka) e sinônimas (Ks). Foi utilizada a fórmula  $T = Ks / 2\lambda$  ( $\lambda = 6,5 \times 10^{-9}$ ) x  $10^{-6}$  Mya (milhões de anos atrás) (Conery, 2000) para calcular o tempo de divergência (T) de cada gene.

A duplicação genômica do GmCOBRA foi consultada nos bancos de dados Plaza 3.0 ([http://bioinformatics.psb.ugent.be/plaza/versions/plaza\\_v3\\_dicots/](http://bioinformatics.psb.ugent.be/plaza/versions/plaza_v3_dicots/)), Plant Genome Duplication Database (<http://chibba.agtec.uga.edu/duplicação/>) e Soybase (<http://www.soybase.org/>).

#### **4.2. Obtenção das bibliotecas de transcriptomas de genes envolvidos com a parede celular em diferentes estágios de desenvolvimento da semente de soja**

Depois de identificados os genes envolvidos na síntese da parede celular em soja, foram utilizadas bibliotecas do transcriptoma de sementes de soja em vários estágios de desenvolvimento, desde a fertilização até a maturação da semente, depositadas no banco GEO/NCBI com número de série de acesso GSE42871 (Figura 10). Foram utilizadas também, bibliotecas do transcriptoma de eixo embrionário de sementes de soja durante a germinação (Bellieny-Rabelo et al., 2016) cedidas gentilmente pelo Prof. Dr. Thiago Motta Venâncio. No total, foram utilizadas 13 bibliotecas de transcriptomas (Figura 10). As bibliotecas foram mapeadas aos genes anteriormente anotados. A caracterização da expressão de cada um dos genes envolvidos com parede celular foi feita através do valor RPKM (*Reads* por Kilobase por Milhão de *Reads* Mapeados).



**Figura 10.** Esquema dos estágios de desenvolvimento da semente de soja mostrando em quais momentos do desenvolvimento as bibliotecas de transcriptoma foram realizadas. DAF= Dias após da fertilização. WS= Semente inteira. COT= Cotilédone. HAE=Horas após a embebição.

#### 4.3. Análise de mapas transcricionais de parede celular durante os diferentes estágios de desenvolvimento da semente de soja

Os genes identificados foram analisados tanto quanto a sua função, considerando o seu homólogo em *Arabidopsis*, quanto pela sua frequência e expressão. Após isso, foi construído um mapa de regulação transcricional nas diversas etapas de formação da semente quiescente e, após a germinação. Os valores de RPKM para cada gene em todas as amostras de tecidos foram utilizados para gerar dois *heatmaps*. Foram definidas duas fases de desenvolvimento da semente: semente inteira (abrangendo as fases de embriogênese e cotilédone) e desenvolvimento do eixo embrionário durante a germinação. A frequência de genes de parede celular e sua expressão em cada tecido foi plotada utilizando o pacote R. Os valores de RPKM foram normalizados para log2 para melhor visualização dos dados. Valores de RPKM  $\geq 1$  foram utilizados como corte para a análise da



expressão de genes. O nível de expressão dos genes em cada etapa do desenvolvimento das sementes de soja foi representado em um *heatmap* plotado utilizando a função *heatmap2* do pacote R. Da mesma forma, o número de genes anotados em determinada via envolvida com a parede celular também foi plotado no *heatmap*. Genes-chave envolvidos com a regulação da parede celular em sementes de soja foram selecionados para avaliação da sua expressão. Os gráficos foram plotados no programa GraphPad Prism versão 7.

#### **4.4. Processamento das sementes de soja e observação microscópica**

Para análise microscópica, as sementes de soja cultivar BRS284, cedidas pela Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), previamente desinfestadas, foram germinadas em placas de Petri contendo 3ml de água esterilizada e 2g de algodão. As placas foram colocadas em estufa BOD a 28° C com 12/12h de fotoperíodo. Durante os estágios de desenvolvimento de 3, 6, 12 e 24 horas após a embebição os eixos embrionários das sementes foram coletados.

Depois de coletado, cada eixo embrionário foi colocado em um microtubo contendo a solução fixadora de Karnovsky (Karnovsky, 1964). Após o processo de fixação, o material foi lavado em tampão cacodilato de sódio 0,05 M a pH 7,2 e iniciada uma série alcoólica para a desidratação do material. Depois de desidratado, o material foi infiltrado com historesina (Historesin Leica, preparada conforme instruções do fabricante) e emblocado em resina pura.

Os blocos foram cortados em seções, com uma espessura de 5 µm, em micrótomo rotativo de avanço automático (RM 2155, Leica). Os cortes foram aderidos em lâminas histológicas contendo água destilada e secados.

Foram utilizados dois métodos de coloração:

- Azul de toluidina, 1% pH 4, onde, se obtém uma imagem geral da qualidade das amostras e da estrutura dos tecidos. As paredes primárias foram coradas de azul a púrpura. Os cortes foram colocados durante 5 minutos nesse corante. Em seguida, foram lavados em água destilada para remoção do excesso de corante e observados ao microscópio óptico de campo claro (Axioplan ZEISS) acoplado a câmera Canon Powershot A640 e através do *software* Analysis® - LINK/ISIS/ZEISS (Oxford, Reino Unido).

- Calcofluor White M2R, 0,1mg/mL diluído em tampão fosfato pH7. Primeiramente, os cortes foram reidratados e submetidos a uma série alcóolica para a remoção da resina. Em seguida, os cortes foram colocados por 15 minutos no corante e posteriormente lavados com água destilada para a remoção de qualquer excesso de corante. Os cortes foram excitados no comprimento de onda de ultravioleta (340 nm) e a emissão coletada no azul (450 a 500 nm). Em seguida, foram observados ao microscópio confocal Leica TCS- SPE, utilizando o software Leica LAS interactive a fim de tomar os dados quantitativos.

Nas observações de celulose, a fluorescência relativa foi calculada com a seguinte equação:

$$F_{relative} = \frac{f1: \textit{fluorescencia emitida}}{f0: \textit{fluorescencia background}}$$

A fluorescência emitida é a fluorescência emitida pelo tecido da amostra e a fluorescência background foi definida no software do microscópio como autofluorescência do tecido.

#### **4.5. Ensaio de atividade da Peroxidase**

Para analisar a atividade enzimática das peroxidases no eixo embrionário de sementes de soja, primeiramente, foi realizada a extração das proteínas totais. Amostras de 40 eixos embrionários de sementes de soja quiescente (0h) e em 3, 6, 12 e 24 horas após a embebição com água, foram coletados e macerados com nitrogênio líquido. Após a maceração, foram pesados 7mg de cada amostra e adicionados 250  $\mu$ L de tampão fosfato de sódio monobásico (100mM); NaCl (150 mM) pH7,6. Em seguida, as amostras foram homogeneizadas no vórtex durante 1h. Depois de homogeneizadas, as amostras foram centrifugadas por 10 minutos a -4 °C e a 10.000 rpm. Depois de centrifugadas, o sobrenadante de cada amostra foi coletado e armazenado em tubos eppendorf.

Depois da extração das proteínas totais de cada amostra, dois géis de poliacrilamida a 12% sem dodecil sulfato de sódio (SDS) foram preparados para a

eletroforese em gel não-desnaturante (gel nativo) (BN-PAGE). Vinte microlitros de cada amostra foram aplicados em cada um dos géis. A eletroforese foi conduzida pelo método vertical com placas de 20x20 imersas em solução tampão Tris 0,1 M e Glicina 0,9 M e a 140V por cerca de 2h e 30min. Um dos géis foi corado com Coomassie Brilliant Blue (CBB) para verificar a qualidade da corrida e o outro gel foi usado para eletrotransferência para a membrana de nitrocelulose.

Após a corrida de eletroforese, as bandas do gel foram então transferidas para uma membrana por aproximadamente 3h usando o equipamento TRANSBLOT da BIORAD. Depois da transferência, a membrana foi imersa em uma solução contendo 100 µl de Tris HCL, 2 M, pH 7,5; 300 µl imidazol 0,1 M; 4,9 mL H<sub>2</sub>O; 5 mg DAB (3,3 – Diaminobenzidínea); 5 µl H<sub>2</sub>O<sub>2</sub> e colocada no escuro para a revelação da atividade da enzima peroxidase nas amostras.

## 5. RESULTADOS

### 5.1. Identificação e anotação dos genes pertencentes às vias de síntese da parede celular

Utilizando como referência as sequências gênicas de vias de síntese de parede celular de *Arabidopsis* disponíveis no banco de dados *Cell Wall Genomics*, foram anotados um total de 2.143 genes referentes à parede celular, o que corresponde a 3,95% do genoma de soja (Tabela 1). Os genes foram agrupados de acordo com a sua via de síntese de parede celular. As vias que tiveram o maior número de genes foram as vias de biossíntese de fenilpropanoides (220), Peroxidases (188), Pectinas esterases (132), Poligalacturanas (108), família GT 8 (86) e vias de interconversão de açúcar-nucleotídeos (83).

A partir da lista de genes foi possível observar que os genes de soja parecem ter sofrido ao menos uma duplicação em comparação aos genes de *Arabidopsis*. Todos os genes identificados no genoma de soja foram listados de acordo com o seu homólogo em *Arabidopsis* e a sua via de parede celular correspondente (Apêndice- Tabela S1).

**Tabela 1.** Número de genes envolvidos na síntese de parede celular em soja e Arabidopsis.

Vias de síntese da parede celular		Número de genes	
		Arabidopsis	Soja
Vias de geração de substrato	Vias de interconversão de açúcar-nucleotídeos	36	83
	Vias de biossíntese de fenilpropanoide	81	220
Síntese de Polissacarídeos e glicosiltransferases	Celulose sintase	10	36
	Celulose like sintase	29	60
	GT família 8	41	86
	GT família 47	38	80
	GT família 37	10	10
	GT família 34	8	8
	GT família 31	33	68
	Genes de síntese de calose	12	32
Vias de secreção	Vesicle trafficking	10	21
	Endocytosis	23	59
Montagem, arquitetura e crescimento	Expansinas	35	65
	Yieldins	10	27
	Xyloglucanos endotransglucosilase/hidrolase	33	81
	Família 35 $\beta$ -Galactosidase	18	55
	Família 9 Glicosido- hidrolase	25	42
	Família 17 Glicosido- hidrolase	47	84
	Poligalacturonases - Pgases	66	108
	Pectatos e pectinas liases	27	42
	Ramnogalacturonas - liases	7	13
	Pectinas esterases	66	132
	Pectinas acetilesterases	11	26
	Glicoproteínas ricas em hidroxiprolina (HRGPs)	13	16
	Proteínas Arabinogalactanas (AGPs)	40	82
	Prolil-4-hidroxilase	11	24
Síntese da Parede Secundária	Lacases	17	61
	Peroxidases	74	188
Vias de mecanismos e sinalização de resposta a estímulos	Receptores like-kinases e outros ligantes	32	10
	Glicosilfosfatidilinositol (GPI)-ancoragem	17	23
	Nodulinas-like de estímulo rápido	18	35
	Uclacianin-like	8	20
	Stelacianin-like	4	14
	Proteínas relacionadas a Extensinas	7	3
	Receptores-like	16	31
	HIPL	3	2
	Proteases	18	53
	LPTL	24	44
	GPDL	6	9
	Família COBRA	11	25
	Família SKU5	3	11
	Família de ancoragem de peptídeos GPI (GAPEP)	8	2
	Proteínas desconhecidas com âncora GPI	43	52
	Total	1049	2143

### 5.1.1. Identificação, caracterização e classificação dos membros da família COBRA em soja

Na anotação de genes da parede celular foi possível verificar a presença de uma família gênica ainda pouco conhecida em soja, a família COBRA. Em *Arabidopsis*, mutantes COBRA resultam em níveis reduzidos de celulose e na de posição desorganizada das microfibrilas, causando defeitos na parede celular (Fernandez et al., 2005). Em arroz, mutantes *BC1*, que codificam proteínas COBRA, possuem uma redução no teor de celulose e na espessura da parede celular (Li et al., 2003; Liu et al., 2013). Esses dados sugerem que as proteínas COBRA estão possivelmente envolvidas no desenvolvimento de vários tecidos e possuem um importante papel na montagem da estrutura primária da parede celular. Considerando essa importante função na biossíntese da parede celular, os genes COBRA de soja foram identificados e caracterizados.

No genoma de soja foram identificados um total de 32 genes candidatos na primeira busca por genes COBRA homólogos à *Arabidopsis*. Após remover as sequências redundantes e os transcritos, 24 genes COBRA putativos foram identificados no genoma de soja. Os genes putativos foram verificados na ferramenta *Pfam* para confirmar a presença do domínio característico da família COBRA. Em apenas um candidato não foi encontrado o domínio COBRA e, por isso, foi removido das análises subsequentes. Com a remoção dessa sequência, o número total de genes COBRA no genoma de soja anotados foram 23 (Tabela 2).

Os genes COBRA possuem tamanho entre 2.172 a 4.805 pb e proteínas com 209 a 671 resíduos de aminoácidos. Todos os genes COBRA em soja foram denominados GmCOBRA e nomeados de acordo com a ordem de seus homólogo em *Arabidopsis*. Os parálogos de GmCOBRA foram nomeados com base na sua localização no cromossomo em seu subgrupo.

**Tabela 2.** Genes COBRA identificados no genoma de soja e suas características.

Gene ID	Anotação	Tamanho da CDS (bp)	CCVS domain	Peptídeo sinal N-terminal	Potencial local- $\omega$	Número de éxons	Chr
Glyma.04G006100	GmCOBL1a	1344	sim	sim	P(418)	6	4
Glyma.06G005800	GmCOBL1b	1338	sim	sim	P(419)	6	6
Glyma.1G003300	GmCOBL10a	573	sim	sim	P(623)	2	11
Glyma.12G213400	GmCOBL10b	2016	sim	não	P(647)	2	12
Glyma.13G288300	GmCOBL10c	2001	sim	não	P(642)	2	13
Glyma.07G243200	GmCOBL7a	1932	sim	sim	P(621)	2	7
Glyma.09G039900	GmCOBL7b	1953	sim	sim	P(624)	2	9
Glyma.17G030700	GmCOBL7c	1935	sim	sim	P(620)	2	17
Glyma.01G240200	GmCOBL11	1800	sim	sim	no	4	1
Glyma.04G160000	GmCOBL4a	1377	sim	sim	P(429)	6	4
Glyma.06G205500	GmCOBL4b	1377	sim	sim	P(429)	6	6
Glyma.08G249800	GmCOBL4c	1296	sim	sim	P(408)	6	8
Glyma.18G272100	GmCOBL4d	1296	sim	sim	P(408)	6	18
Glyma.19G033600	GmCOBL4e	1296	sim	sim	P(408)	6	19
Glyma.02G196100	GmCOBa	1110	sim	não	no	4	2
Glyma.04G160100	GmCOBb	1371	sim	sim	P(429)	6	4
Glyma.05G019100	GmCOBc	630	sim	sim	no	5	5
Glyma.06G205400	GmCOBd	1371	sim	sim	P(429)	6	6
Glyma.08G249700	GmCOBe	1347	sim	sim	P(422)	6	8
Glyma.13G053300	GmCOBf	1377	sim	sim	P(432)	6	13
Glyma.17G080600	GmCOBg	1365	sim	sim	P(429)	6	17
Glyma.18G271900	GmCOBh	1263	não	sim	no	6	18
Glyma.19G033500	GmCOBi	1365	sim	sim	P(428)	6	19

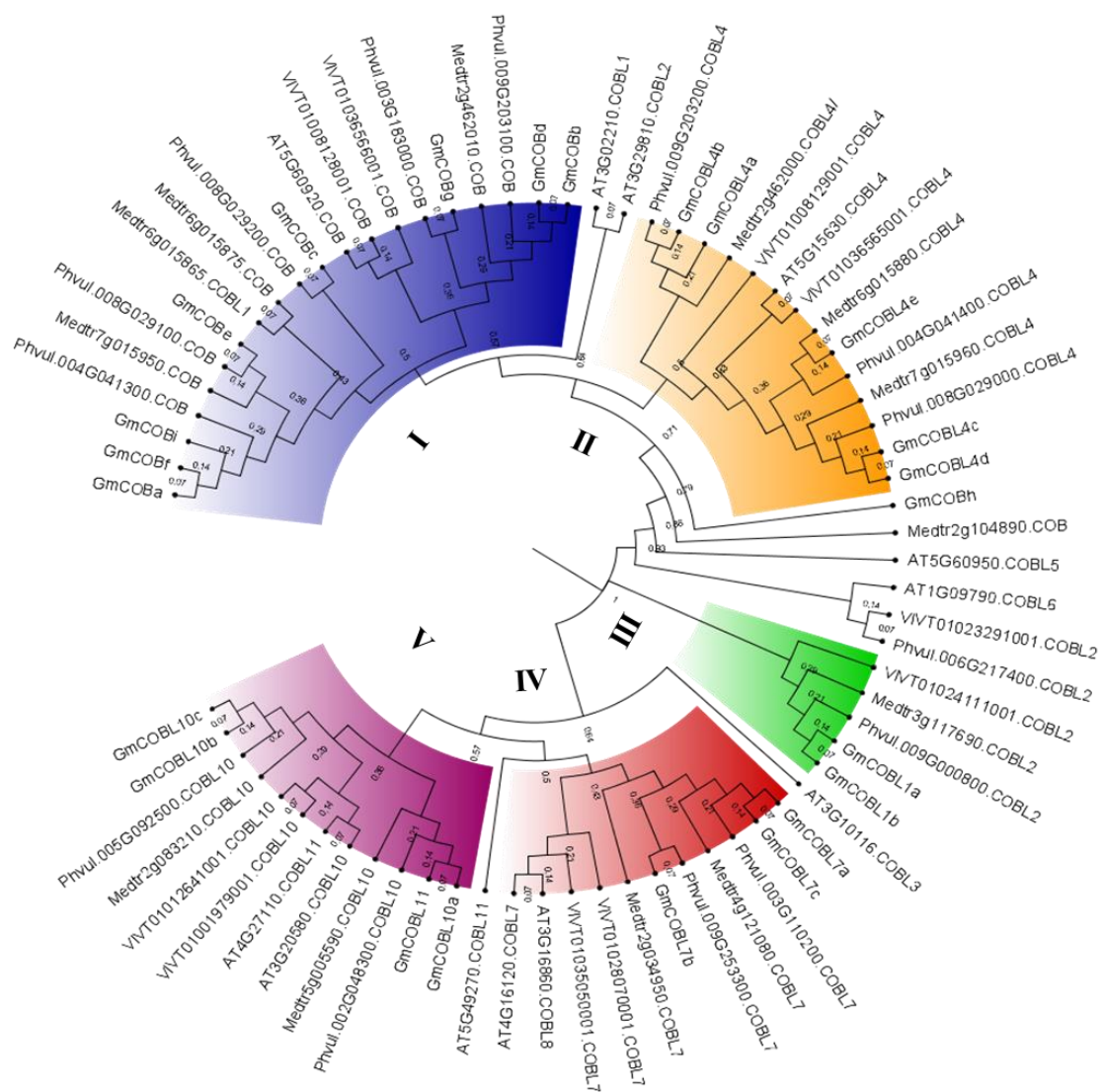
### 5.1.2. Análise filogenética da família COBRA em soja

Para investigar as relações evolutivas entre os membros da família COBRA de soja com demais leguminosas e eudicotiledôneas, foi construída uma árvore filogenética a partir do alinhamento de 72 sequências de proteínas COBRA de soja, *Arabidopsis thaliana*, *Vitis vinifera*, *Medicago truncatula* e *Phaseolus vulgaris*.

A análise da árvore filogenética obtida permitiu identificar que os homólogos COBRA pertencentes as diferentes espécies podem ser divididos em 5 grupos (I-V) (Figura 11). Os grupos I e II foram os maiores em número de genes com 20 e 14 membros, respectivamente. O grupo menor foi o III contendo apenas 5 membros.

No genoma de *Arabidopsis* estão anotados 11 genes COBRA. Seis membros dessa família possuem representantes em soja (Tabela 2). Os GmCOBRA foram agrupados no mesmo clado do seu homólogo em *Arabidopsis* (Figura 11). A maioria dos genes COBRA de *Arabidopsis* possuem mais de um homólogo em soja. Como exemplo, o gene denominado COB em *Arabidopsis* foi o que apresentou maior número de representantes em soja, sendo um total de nove (Tabela 2). Todos esses representantes foram classificados no grupo I com as demais espécies (Figura 11). Esse fato sugere que a maioria dos GmCOBRA foram duplicados após a divergência entre essas espécies gerando novos parálogos em soja.





**Figura 11.** Árvore filogenética dos genes COBRA. A árvore filogenética foi gerada no MEGA7 usando o método de Neighbor-Joining. Utilizamos o total de 72 seqüências codificadoras de proteínas de diferentes espécies.

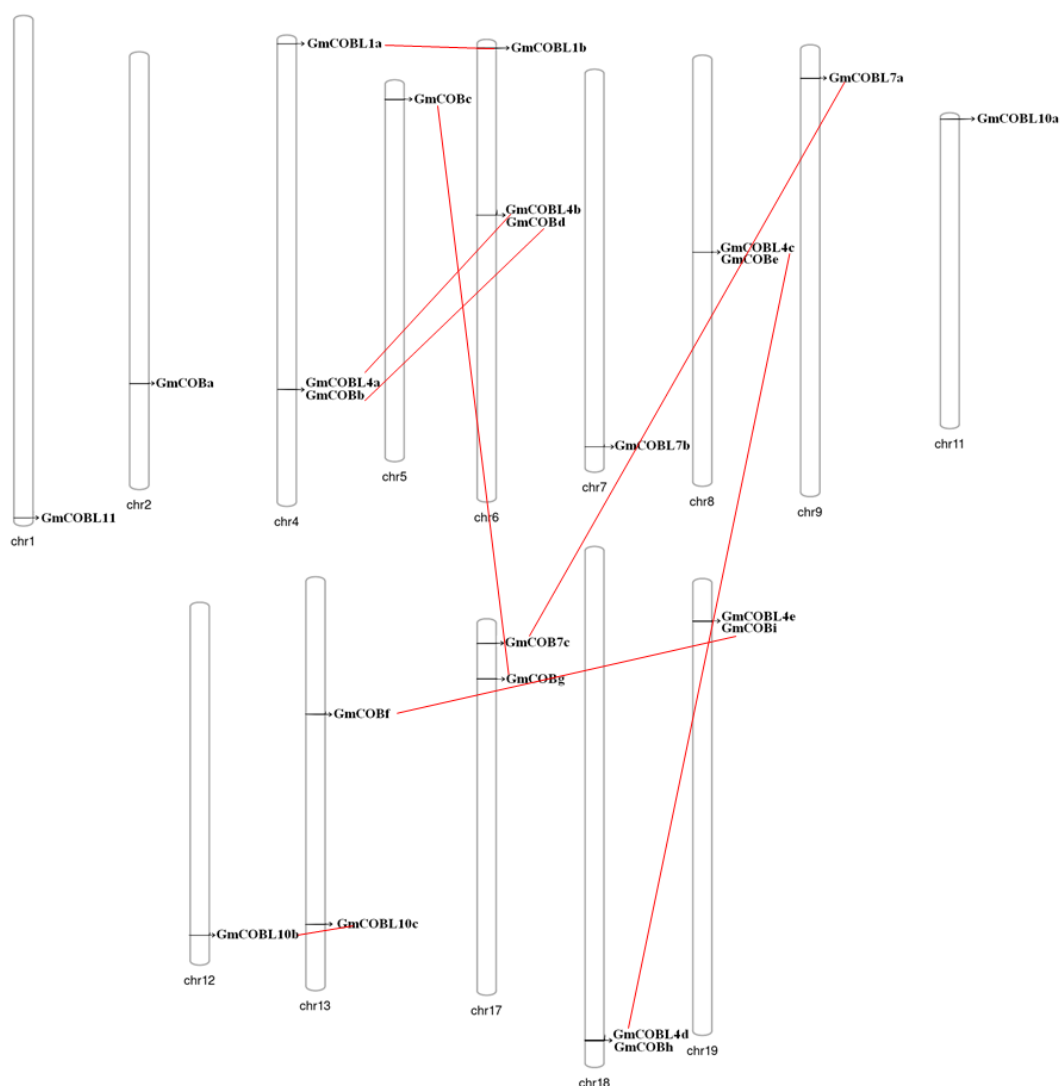
### 5.1.3. Duplicação dos genes COBRA, localização e distribuição cromossomial

Em comparação com *Arabidopsis*, a soja sofreu ao menos duas duplicações completas do genoma ao longo da sua história evolutiva (Schmutz et al., 2010). Para identificar os pares de genes duplicados, compreender a duplicação e expansão dos genes COBRA em soja, analisamos os eventos de duplicação de COBRA em todo o genoma de soja. Todos os GmCOBRA foram distribuídos preferencialmente em duplicação em blocos (Tabela 3). A análise indicou que cerca de metade do GmCOBRA sofreu duplicação em *tandem*, o que provavelmente contribuiu para a sua expansão no genoma de soja (Tabela 3).

**Tabela 3.** Lista dos eventos de duplicação dos genes COBRA em soja.

<b>Gene</b>	<b>Tipo de Duplicação</b>
GmCOBL1a	Duplicação Segmentar e em Bloco
GmCOBL1b	Duplicação Segmentar e em Bloco
GmCOBL10a	Duplicação Segmentar e em Bloco
GmCOBL10b	Duplicação Segmentar e em Bloco
GmCOBL10c	Duplicação Segmentar e em Bloco
GmCOBL7a	Duplicação Segmentar e em Bloco
GmCOBL7b	Duplicação Segmentar e em Bloco
GmCOBL7c	Duplicação Segmentar e em Bloco
GmCOBL11a	Duplicação Segmentar e em Bloco
GmCOBL4a	Duplicação em Tandem e em Bloco
GmCOBL4b	Duplicação em Tandem e em Bloco
GmCOBL4c	Duplicação em Tandem e em Bloco
GmCOBL4d	Duplicação em Tandem e em Bloco
GmCOBL4e	Duplicação em Tandem e em Bloco
GmCOBa	Sem informações
GmCOBb	Duplicação Segmentar e em Bloco
GmCOBc	Duplicação em Tandem e em Bloco
GmCOBd	Duplicação em Tandem e em Bloco
GmCOBe	Duplicação em Tandem e em Bloco
GmCOBf	Duplicação em Tandem e em Bloco
GmCOBg	Duplicação em Tandem e em Bloco
GmCOBh	Sem informações
GmCOBi	Duplicação em Tandem e em Bloco

O mapeamento de 23 genes GmCOBRA nos 20 cromossomos de soja indicou uma distribuição heterogênea (Figura 12). Os genes foram distribuídos em 14 dos 20 cromossomos de soja. Os cromossomos 3, 10, 14, 15, 16 e 20 não apresentaram genes. Todos os genes pertencentes à subfamília GmCOB e GmCOB4 sofreram duplicação em tandem e em bloco e, ainda, apresentaram colinearidade. Estes genes são provavelmente par de parálogos que divergiram ao longo do tempo em soja.



**Figura 12.** Distribuição do GmCOBRA nos cromossomos de soja. Somente os cromossomos que possuem genes COBRA estão representados. As linhas vermelhas indicam a duplicação dos genes parálogos.

#### 5.1.4. Estimativa da taxa de substituição sinônima e não-sinônima (Ka/Ks) e estrutura gênica

Para mensurar a divergência entre GmCOBRA parálogos no genoma de soja, foram analisados os índices Ka/Ks para todos os pares de genes parálogos. Os índices sugerem que os genes estão sob forte pressão negativa, uma vez que os valores de Ka/Ks são inferiores a 1 (Apêndice – Tabela 2). Todos os pares de genes parálogos obtiveram valores de Ks entre 0,073 a 0,22. A transformação desses valores Ks em tempo de divergência mostrou que os GmCOBRA parálogos divergiram entre 5,64 e 16,92 Mya, com uma média de 10,34 Mya. O último evento de duplicação total do genoma (WGD) de soja foi há aproximadamente 13 Mya. Este resultado sugere que a duplicação desses parálogos ocorreu após a última duplicação completa do genoma soja.

Os genes duplicados podem apresentar diferenças estruturais que muitas vezes levam a geração de parálogos com diferentes funções. Para investigar a diversidade estrutural dos GmCOBRA e seus possíveis mecanismos de evolução estrutural, analisamos a estrutura e a organização dos éxons/íntrons na sequência de codificação. Ocorreu uma mudança no número de éxons (2 a 6) entre as subfamílias COBRA. A maioria dos parálogos compartilharam estrutura parcialmente semelhante (Figura 13).

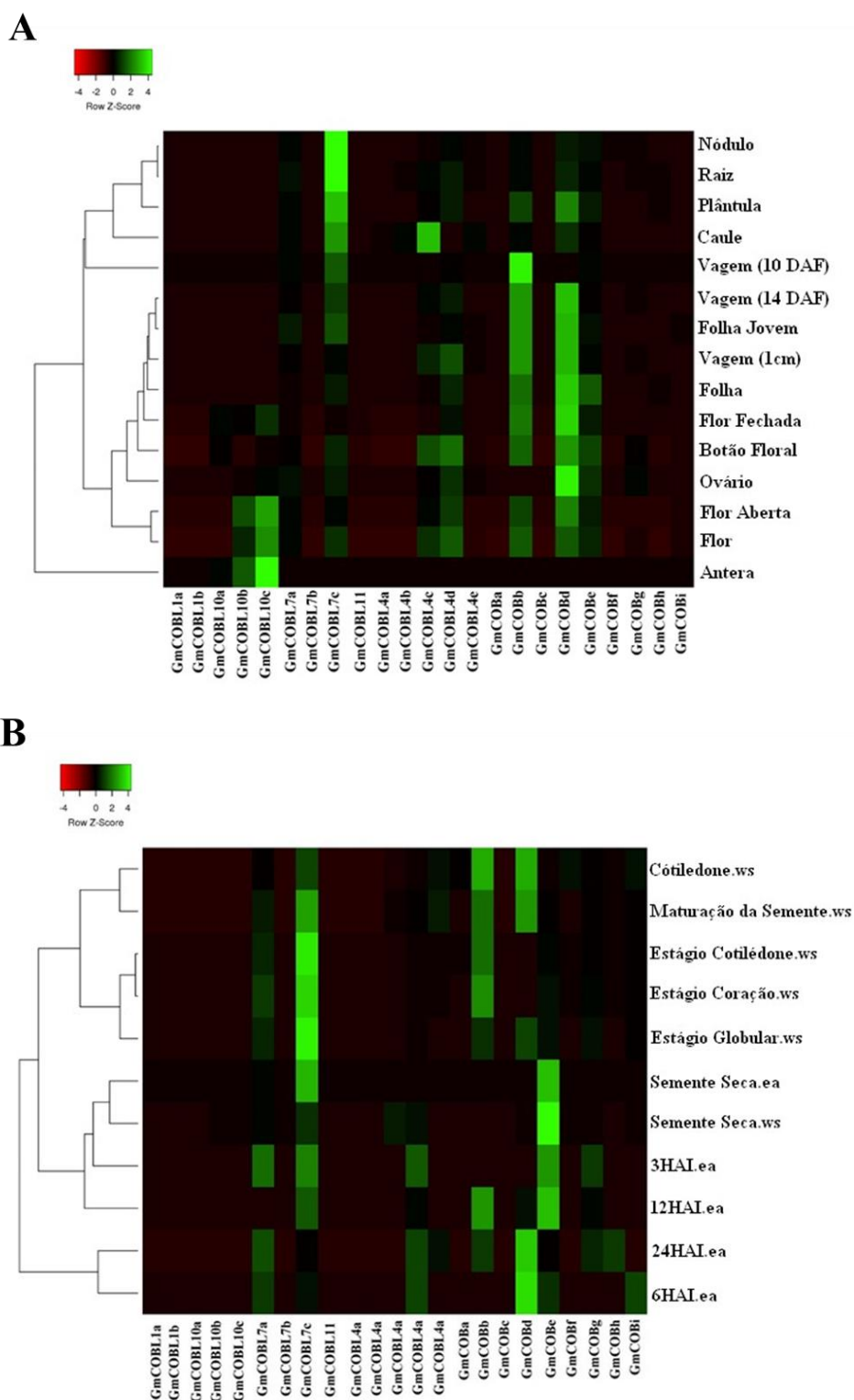


**Figura 13.** Estrutura gênica das 23 sequências COBRA em soja. Os éxons são representados em retângulos pretos, 5'UTR em retângulos cinza e íntrons em linhas pretas.

### 5.1.5. Perfil de expressão e função biológica da família COBRA em soja

A análise do perfil de expressão de genes parálogos pode fornecer importantes evidências sobre uma possível especialização funcional. A fim de analisar o padrão de expressão dos GmCOBRA, foram construídos *heatmaps* com dados de bibliotecas de RNA-seq geradas a partir de amostras de soja em diferentes etapas do desenvolvimento. O perfil de expressão obtido mostrou que a maior parte dos GmCOBRA são pouco expressos (Figura 14A e B). Como exemplo, dentre os GmCOBRA 7, somente o gene GmCOBL7c apresentou expressão na maioria dos tecidos. O GmCOBL7a parece ter uma maior expressão somente em tecidos de sementes em desenvolvimento e germinação (Figura 14B), assim como os GmCOBg, h e i durante a germinação. Já os GmCOBL10b e c parecem possuir uma maior expressão em tecidos reprodutivos como antera, flor, flor aberta e não aberta (Figura 14A). Enquanto que os genes GmCOBL10a, GmCOBf, e GmCOBL1a-b não tiveram expressão detectada nos tecidos investigados.

Pares de genes parálogos apresentaram diferentes níveis de expressão. Por exemplo, os GmCOBb mostraram maior expressão na maioria dos tecidos investigados, enquanto que os GmCOBa e GmCOBc mostraram baixíssima expressão nos mesmos. Esses dados sugerem que os genes parálogos GmCOB e GmCOBL adquiriram funções diferentes após o evento de duplicação sofrido no genoma de soja. Além disso, a análise de ontologia gênica dos GmCOBRA de soja, utilizando a ferramenta AgriGO, mostrou que estes genes estão relacionados a duas principais funções biológicas (Apêndice – Figura S1). Um grupo de genes foi descrito como relacionado à montagem, organização e regulação biológica da parede celular. Enquanto outro grupo está envolvido na biogênese e no crescimento celular (Apêndice – Figura S1), funções ainda não descritas para esta família gênica.



**Figura 14.** Agrupamento hierárquico dos perfis de expressão de genes COBRA de soja em diferentes tecidos. A - Tecidos vegetativos e reprodutivos. B - Etapas da formação de sementes e germinação. HAI - Horas após embebição. WS - Semente inteira. EI - Eixo embrionário

## 5.2. Caracterização do perfil de expressão dos genes pertencentes às vias de síntese, arquitetura e montagem da parede celular em sementes de soja

Após a identificação e anotação dos genes relacionados à síntese da parede celular em soja, as 13 bibliotecas de RNA-seq relacionadas aos estágios de desenvolvimento de sementes de soja foram mapeadas nos mesmos. A caracterização da expressão de cada um dos genes envolvidos com a parede celular foi feita através do valor RPKM e plotados os gráficos referentes à frequência e expressão dos genes das diferentes vias de síntese e regulação da parede celular.

Nos gráficos de frequência (Figura 15A e B) é possível analisar o número de genes que estão sendo expressos em função dos níveis de expressão dados em RPKM durante os diferentes estágios de desenvolvimento da semente. Com isso, observou-se que existe uma maior frequência de genes de parede celular sendo expressos com o RPKM ( $\log_2$ ) acima de 1, com picos de expressão em torno de 2. No estágio de embriogênese da semente, na fase de 22-24 DAF existe um pico de frequência gênica com RPKM ( $\log_2$ ) em relação às fases 4 e 12-14 DAF, ou seja, existe uma maior quantidade de genes envolvidos com parede celular sendo expressos nessa fase em comparação com as outras fases do seu grupo (Figura 15A). No estágio de desenvolvimento, existe uma expressão decrescente dos genes da parede celular a partir do primeiro estágio de cotilédone (5-6mg COT), culminando na menor quantidade de genes expressos em semente seca (Dry seed). As fases de 4 DAF, 12-14 DAF e 5-6 mg WS tiveram o mesmo padrão de quantidade de genes sendo expressos. Já durante a germinação, a quantidade de genes expressos envolvidos com a parede celular aumenta gradativamente em eixo embrionário a partir da semente seca (Figura 15B). Entre 12 e 24HAI não ocorrem grandes variações no número de genes de parede celular e na sua expressão.

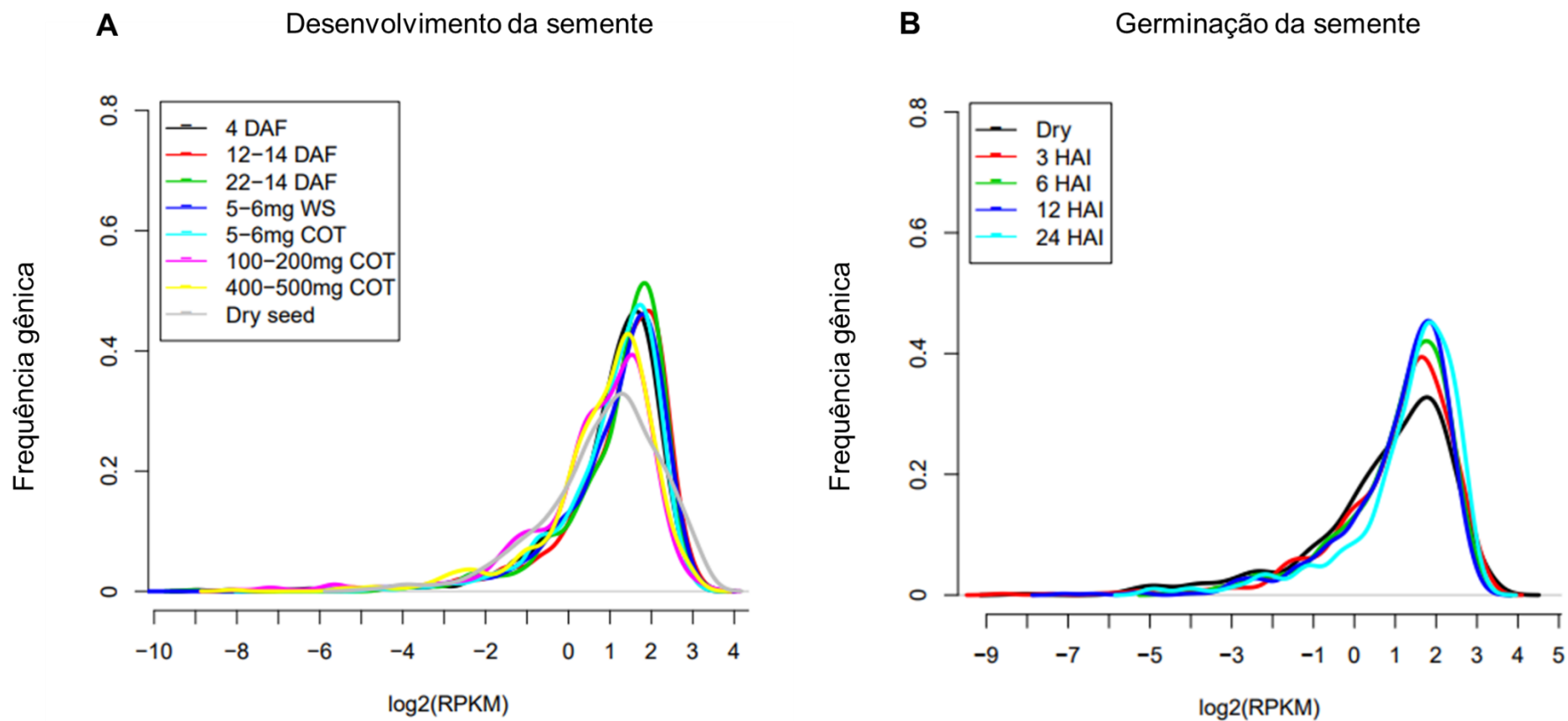
A fim de elucidar quais vias estavam sendo enriquecidas ao longo do desenvolvimento da semente tanto em número de genes, quanto ao nível de expressão, foram analisados os *heatmaps* obtidos a partir da anotação dos genes de parede celular em cada via de biossíntese e também de sua expressão. Os *heatmaps* apresentam a frequência de genes e sua expressão em RPKM de acordo

com as suas vias de síntese, arquitetura e montagem da parede celular correspondentes exibidas na Tabela 1. Para uma melhor visualização, os dados foram divididos em dois grupos: o primeiro grupo consistiu-se nos estágios de desenvolvimento da semente (desde a embriogênese até a maturação da semente) e o segundo grupo abrange os estágios germinativos do eixo embrionário após a embebição da semente.

No desenvolvimento da semente, as vias que apresentaram um maior enriquecimento em número de genes ao longo do desenvolvimento foram as vias de geração de substrato (açúcar nucleotídeos e biossíntese de fenilpropanóides), endocitose e família GT 8 (Figura 16A). Em consonância, o nível de expressão dessas mesmas vias aumentou ao longo do desenvolvimento da semente (Figura 16B). Em contraste, as vias de peroxidase, AGP's, HRGP's, xiloglucanos foram altamente expressas ao longo de todo o desenvolvimento da semente, mesmo não ocorrendo um aumento no número de genes. Isto sugere que, independentemente da quantidade de genes, algumas vias podem ter expressão aumentada a partir da contribuição de apenas alguns membros da respectiva família gênica. Este fato também foi mostrado anteriormente neste trabalho para a família de genes COBRA. Ainda no grupo de desenvolvimento da semente, muitas vias apresentaram pouco enriquecimento em número de genes, sendo algumas delas: extensinas, lacases, stelacianinas, glicosilfosfatidilinositol (GPI) – ancoragem, receptores like-kinases e ramnogalacturanas. Em consonância, a expressão de genes dessas mesmas vias também foi baixa.

No grupo abrangendo o estágio germinativo do eixo embrionário ocorreu um enriquecimento em número de genes das mesmas vias que estavam sendo enriquecidas no grupo de desenvolvimento da semente (Figura 16C). No entanto, algumas vias apresentaram expressões contrastantes ao número de genes. Assim, a expressão das vias foi dividida em três clusters hierárquicos: alta, baixa e média expressão ao longo de todos os estágios de desenvolvimento do eixo embrionário (Figura 16D). Entre as vias que apresentaram alta expressão, algumas se destacaram, tais como: geração de substrato, endocitose, proteínas com âncoras GPI, família glicosil hidrolases 17, arabinogalactanas, expansinas, peroxidases, xyloclucanas, proteases, famílias GT 8,31 e 47, HRGPs, UgE5 e celulose sintase.





**Figura 15.** Gráficos de distribuição da frequência gênica por  $\log_2(\text{RPKM})$  dos genes de parede celular anotados. A- desenvolvimento da semente. B- Eixo embrionário após a germinação da semente. DAF – Dias após a fertilização. WS – Semente inteira. COT – Cotilédone. Dry seed – Semente quiescente. Dry – Eixo embrionário quiescente. HAI – Horas após a embebição.

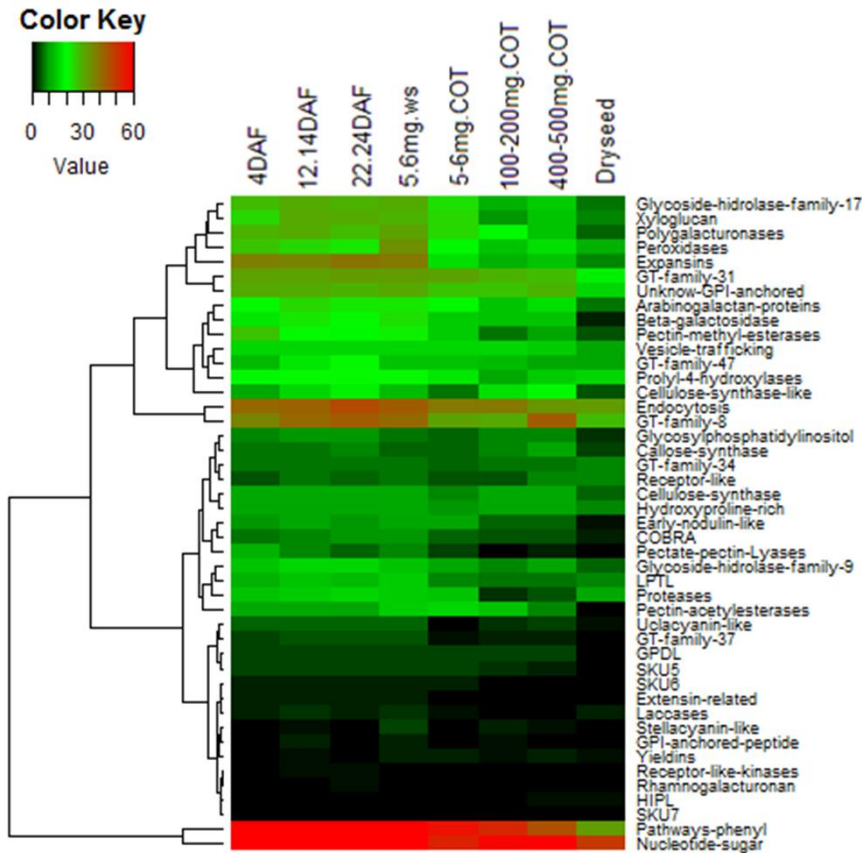
Além disso, a via de celulose sintase-like aumentou a sua expressão longo do desenvolvimento do eixo, atingindo o seu pico de expressão em 24 horas após a germinação. O mesmo aconteceu para as vias de nodulinas de estímulo rápido (ENOD), pectinas esterases, LPTL e família 9 de glicosil hidrolases. Já as vias de SKU 6 e 7 apresentaram uma expressão mínima ao longo da germinação.

A identificação de genes-chave que sirvam como marcadores moleculares para uma característica ou via de interesse tem sido uma das principais ferramentas biotecnológicas para o melhoramento vegetal nos últimos anos. Durante a análise global da expressão dos genes de parede celular durante o desenvolvimento e germinação da semente de soja, foi possível observar que alguns desses genes parecem ter uma maior contribuição na biossíntese de parede celular, principalmente no eixo embrionário durante a germinação. Esta última é uma etapa crucial no desenvolvimento da cultura de soja, uma vez que todo o sucesso da cultura perpassa pelo sucesso germinativo da semente. Os genes-chave envolvidos com a regulação da parede celular que apresentaram maior expressão e, portanto, contribuíram mais para o enriquecimento da via estão apresentados na Figura 17. Dentre eles estão genes de açúcar nucleotídeos, família GT 8, expansinas, celulose sintase-like, arabinogalactanas, celulose sintase, peroxidases, COBRA, hidroxiprolinas e xiloglucanos.

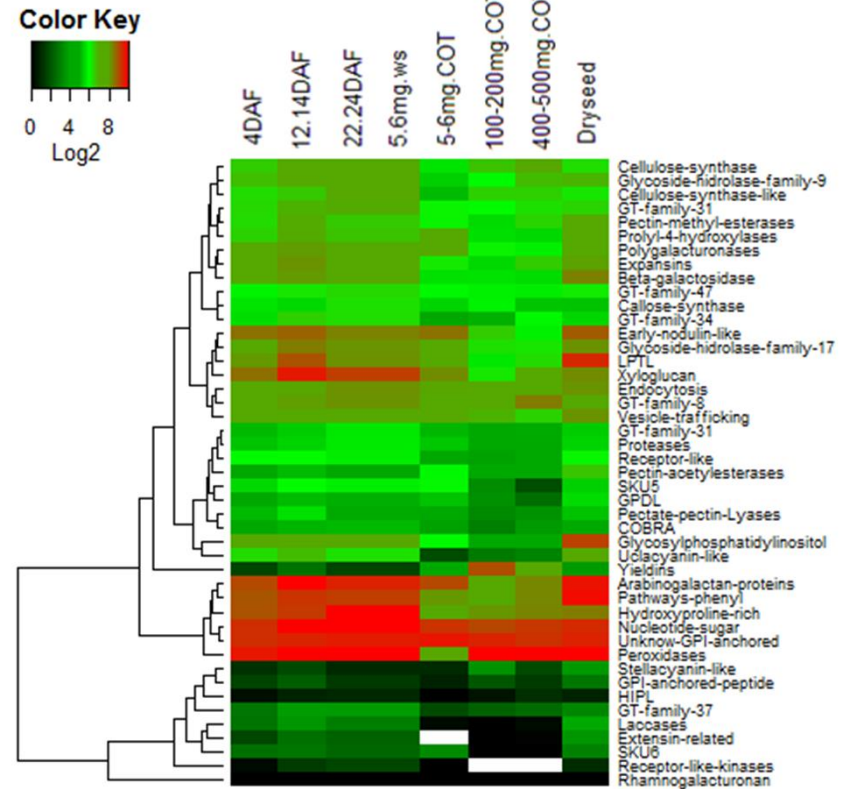
### **5.3. Caracterização morfológico-tecidual da parede celular do eixo embrionário durante a germinação da semente de soja**

A fim de reforçar a compreensão das atividades metabólicas que estão ocorrendo na parede celular do eixo embrionário durante o processo de germinação e compará-las com os dados moleculares das bibliotecas de transcriptoma do eixo embrionário, foram realizadas análises fenotípicas tanto macro quanto microscópicas do eixo embrionário das sementes de soja durante a germinação. As análises foram realizadas em condições semelhantes às das bibliotecas de transcriptoma, sendo elas: eixo embrionário quiescente (0h de germinação), eixo embrionário em 3, 6, 12 e 24 horas após a embebição da semente.

A

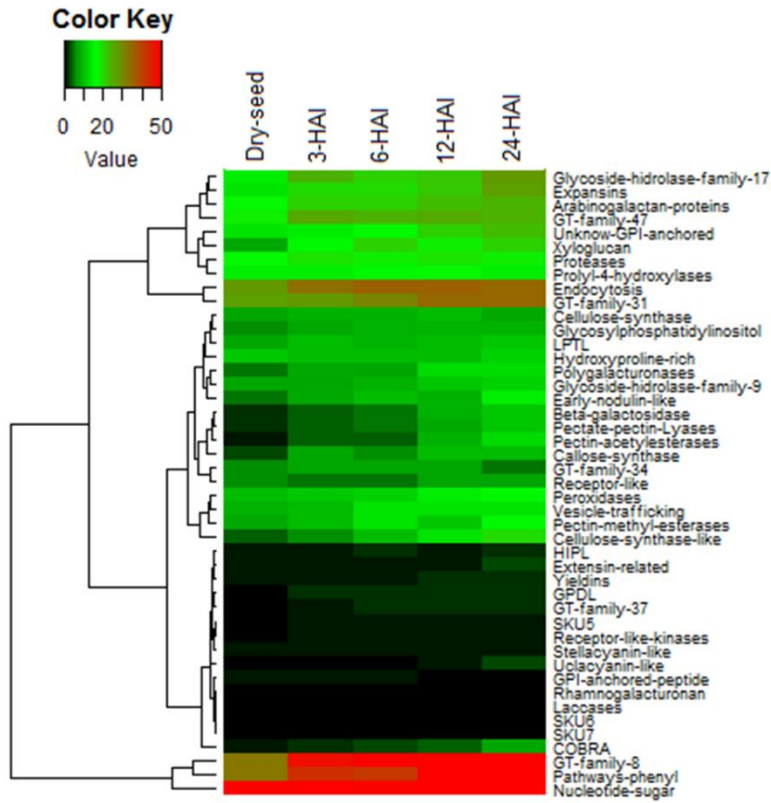


B



**Figura 16** Heatmap mostrando a frequência e expressão das vias de parede celular durante diferentes estágios de desenvolvimento e germinação da semente de soja. A- Frequência das vias de desenvolvimento da semente. B- Expressão das vias de desenvolvimento da semente. C- Frequência das vias de germinação. D- Expressão das vias de germinação.

C



D

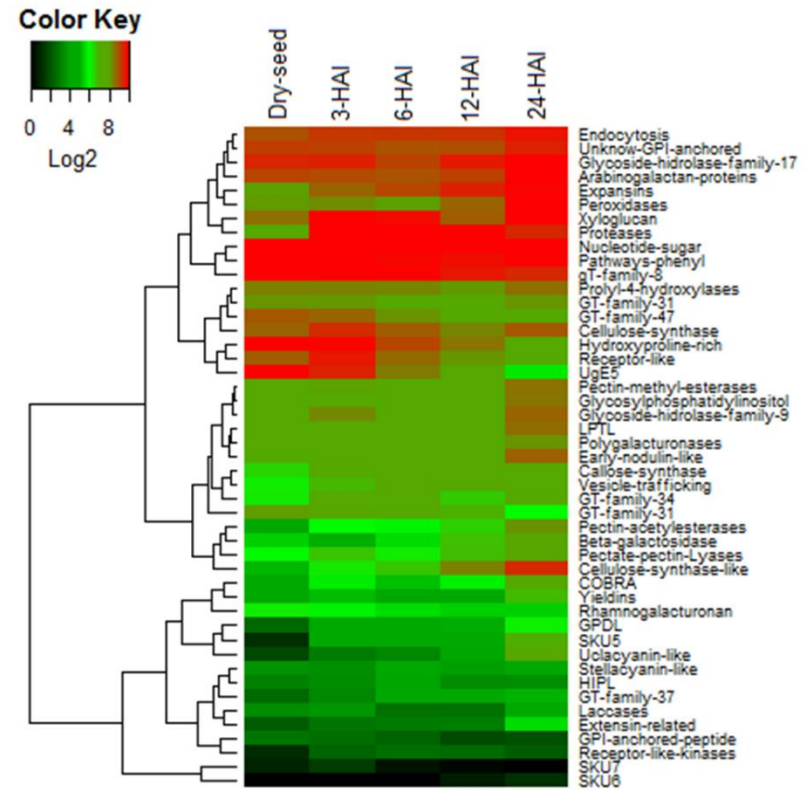
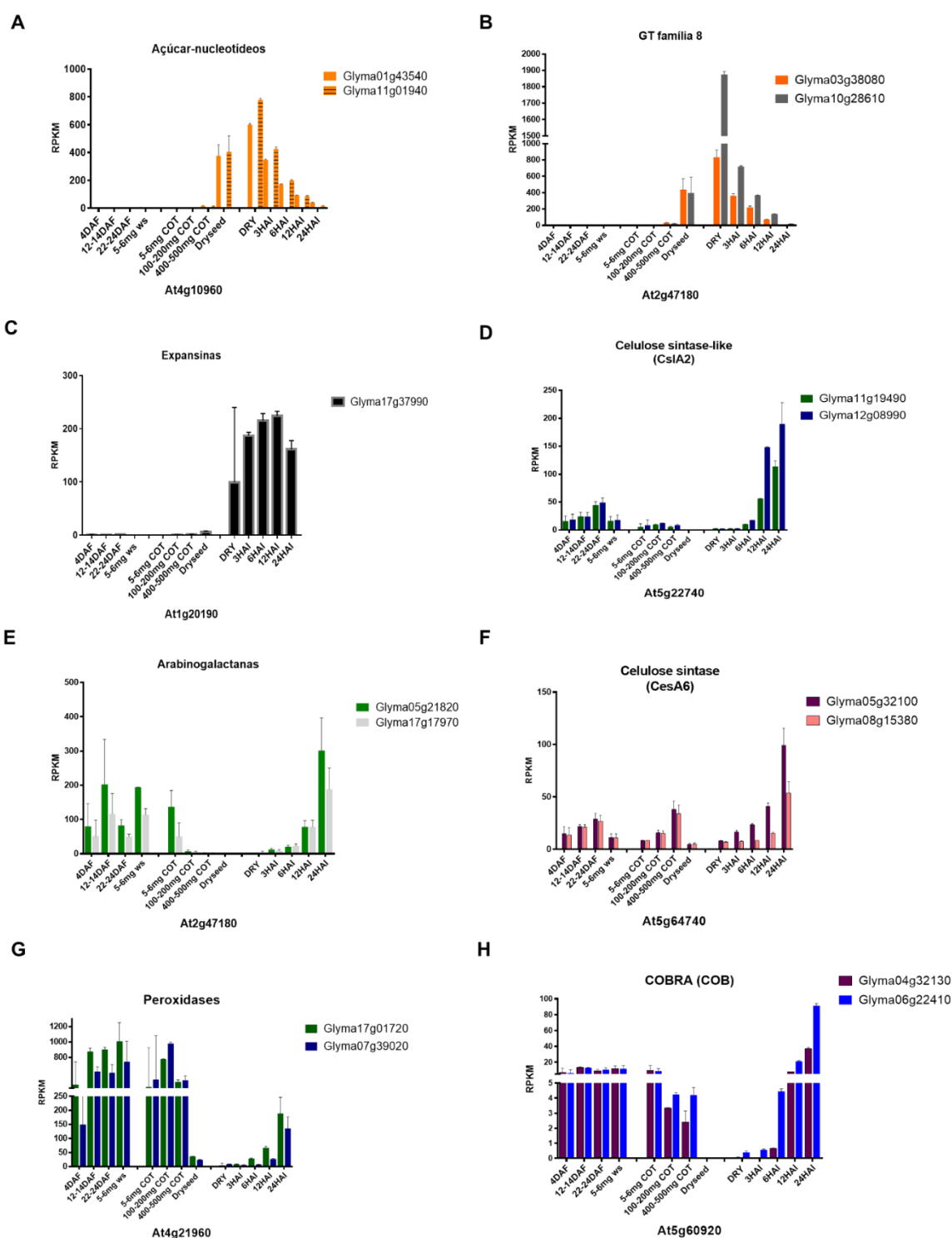
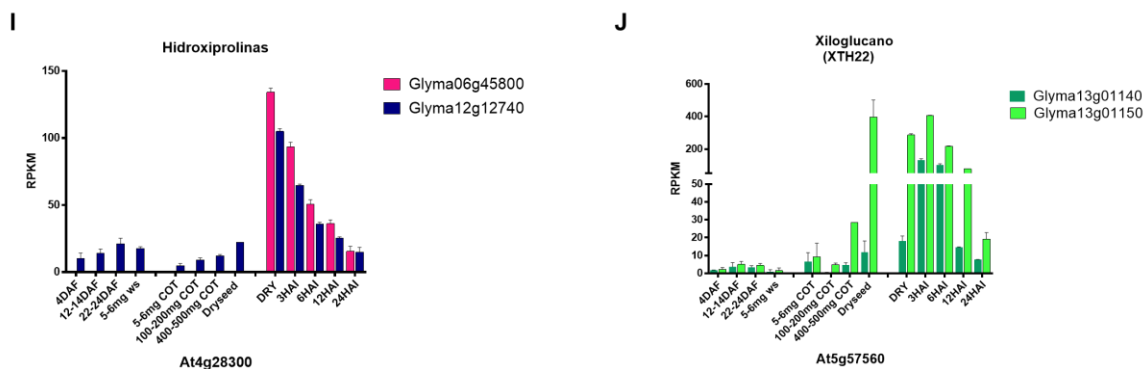


Figura 16 – Cont.



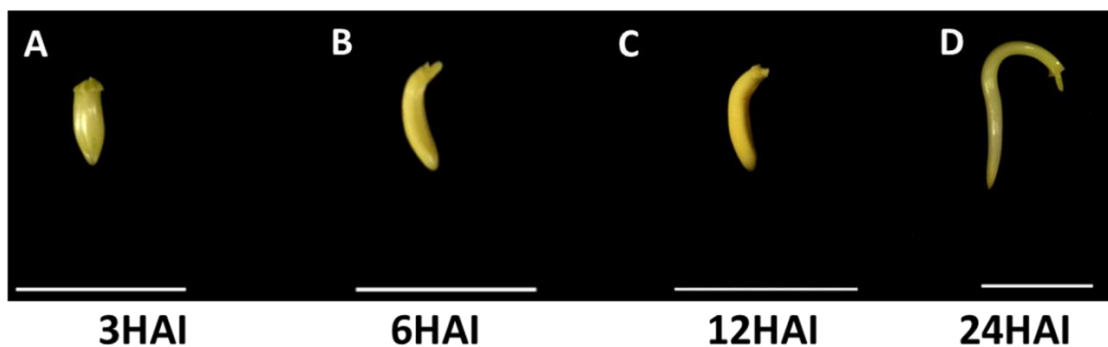
**Figura 17.** Expressão de genes-chave envolvidos na regulação da parede celular em eixo embrionário de sementes de soja durante a germinação. Os gráficos mostram a expressão dos genes com a descrição do seu respectivo homólogo em *Arabidopsis*, durante os estágios de desenvolvimento da semente e germinação do eixo embrionário.



**Figura 17.** Continuação

### 5.3.1. Caracterização fenotípica do eixo embrionário durante os estágios de germinação da semente de soja

Primeiramente, foi realizada uma análise fenotípica macroscópica do eixo embrionário de 3, 6, 12 e 24 horas após a embebição (Figura 18). A partir das imagens é possível observar o aumento crescente do eixo embrionário ao longo da germinação, culminando em seu maior tamanho no ponto de 24 HAI, quando a germinação já está completa. O crescimento do eixo embrionário durante a germinação parece estar relacionado com a alta expressão de genes das vias de geração de substrato, expansinas, peroxidases, xyloclucanos, arabinogalactanos e síntese de celulose (Figura 16). Essas vias são estritamente necessárias para a expansão, alongamento e crescimento celular.



**Figura 18.** Eixo embrionário durante a germinação da semente de soja. HAI – Horas após a embebição. Escala 1 cm.

### **5.3.2. Caracterização tecidual da parede celular do eixo embrionário durante os primeiros estágios de germinação da semente de soja**

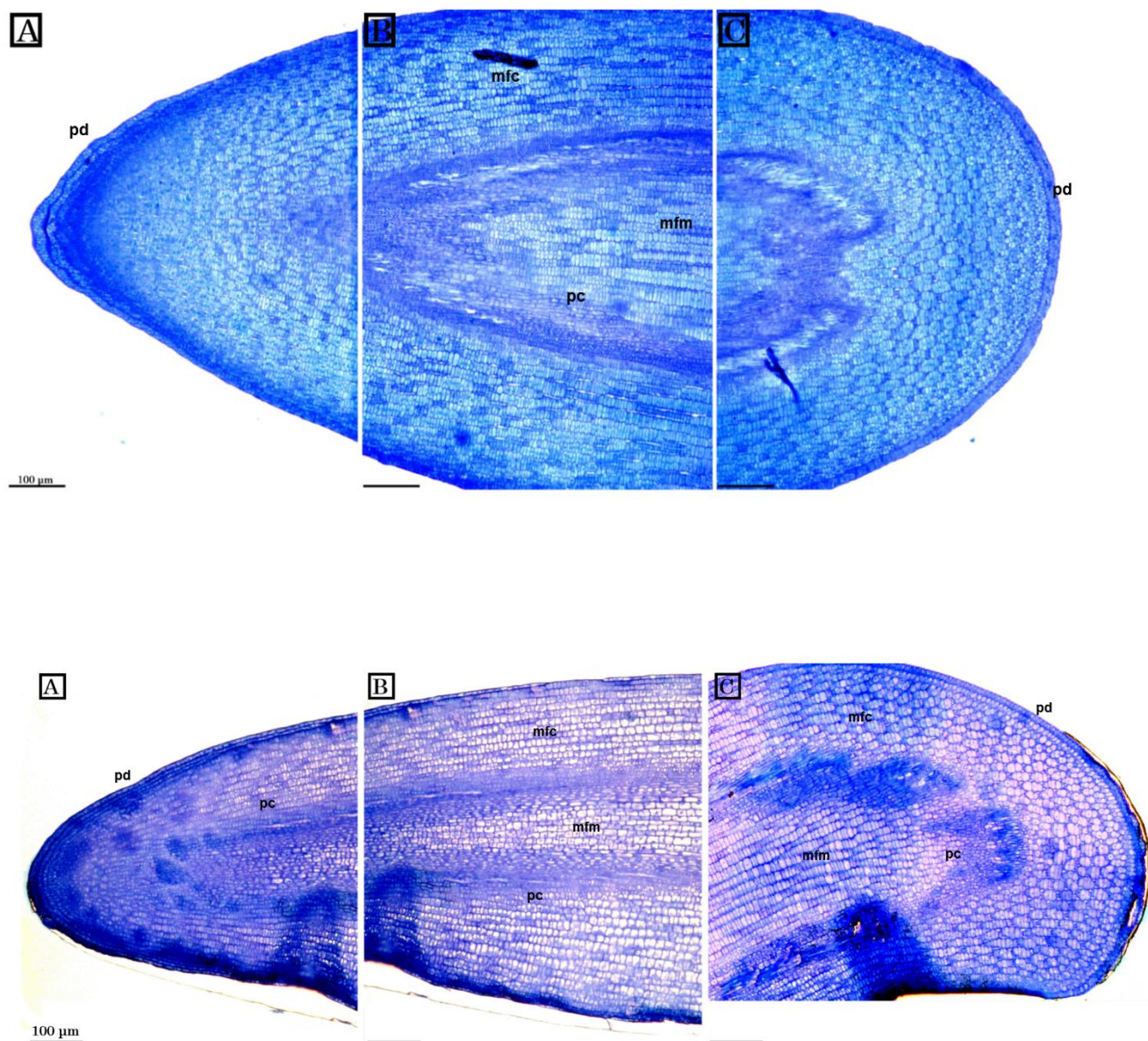
Para a análise microscópica do eixo embrionário durante a germinação, foram feitos cortes longitudinais ao longo do eixo embrionário de 3 e 24 HAI. Para uma melhor visualização, o material foi dividido em três áreas: meristema radicular, procâmbio e meristema apical.

Nas observações microscópicas foi possível analisar a anatomia básica do eixo embrionário (Figura 19). As imagens apresentaram uma anatomia básica de tecidos meristemáticos. A protoderme é glabra com células isodiamétricas. O procâmbio é constituído por células alongadas envolto pelos meristemas fundamentais medular e cortical. Pode-se evidenciar a diferença entre os eixos de 3h para 24h, onde, devido ao crescimento primário e expansão celular, as células do eixo embrionário de 24h apresentam-se maiores tanto em extensão quanto em diâmetro. Além disso, as células do eixo de 24 horas são mais túrgidas em relação ao eixo de 3 horas e o procâmbio parece alongar-se por toda extensão do eixo.

### **5.3.3 Medição de fluorescência relativa de celulose da parede celular do eixo embrionário durante a germinação da semente de soja**

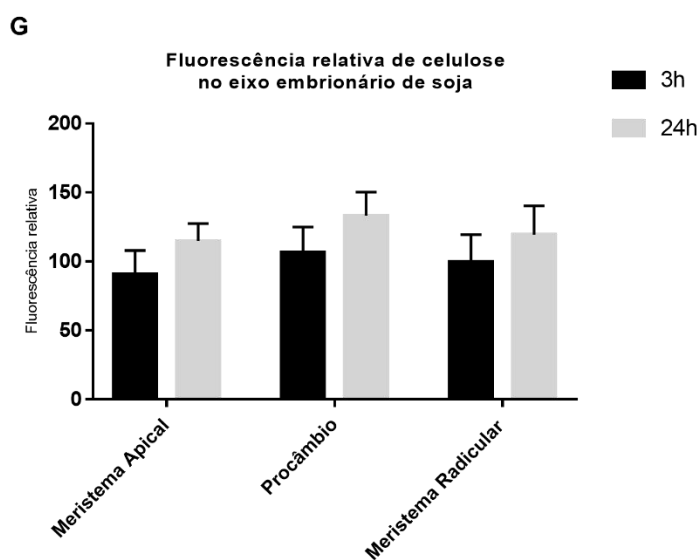
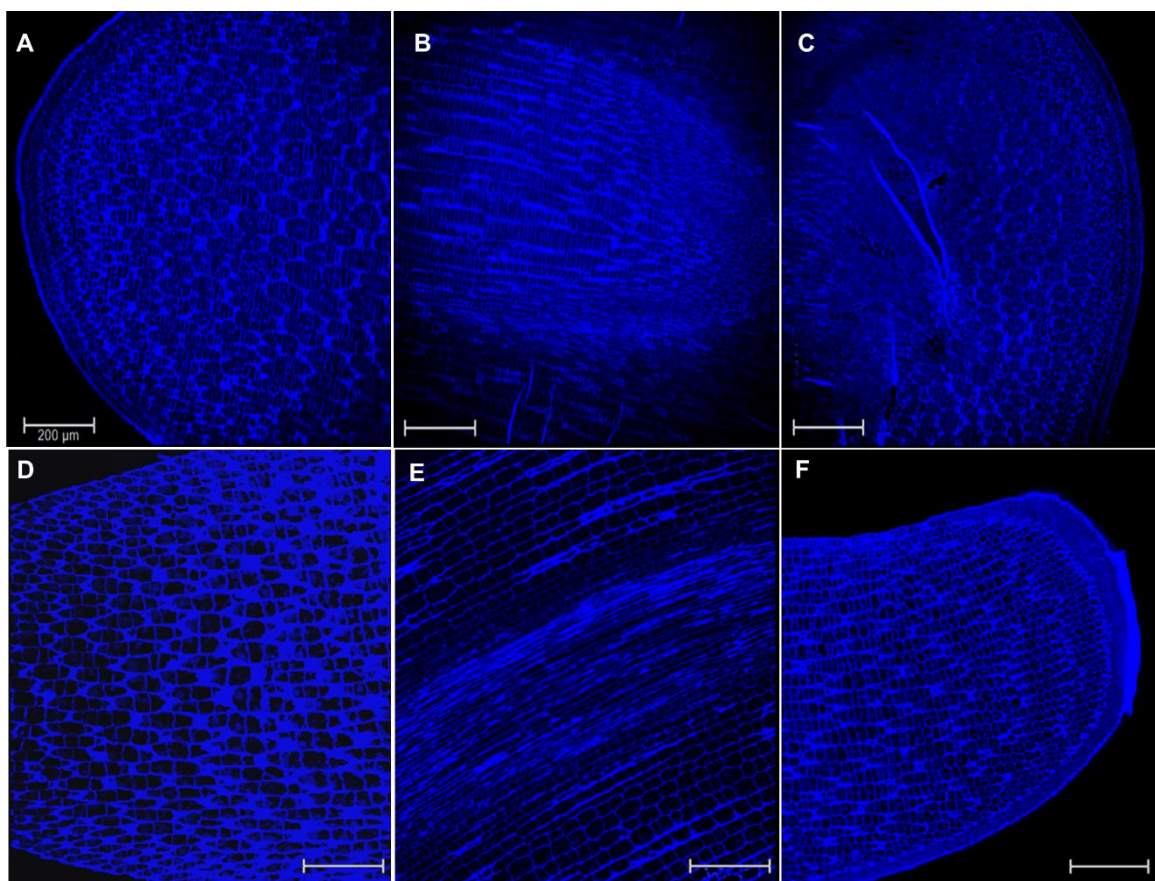
Considerando que genes-chave para a formação das microfibrilas de celulose (Figura 17) possuem aumento da expressão em eixo embrionário durante a germinação de sementes de soja, a medida do conteúdo de celulose a partir da intensidade de fluorescência pode ser usada na validação dessa modulação da biossíntese de parede celular ao longo da germinação. Após a medição da fluorescência relativa de celulose dos tecidos, os resultados mostraram que o eixo embrionário de 24h após a embebição possui uma maior fluorescência de celulose em relação ao eixo de 3h em todas as três áreas - meristema radicular, procâmbio e meristema apical (Figura 20). Esse dado reforça os dados de expressão, onde os genes de celulose sintase aumentaram sua expressão ao longo do desenvolvimento do eixo embrionário durante a germinação da semente (Figura 17F).





**Figura19.** Anatomia do eixo embrionário durante 3 (acima) e 24 (abaixo) horas de germinação. A – Meristema radicular; B – Procâmbio; C- Meristema apical. Pd – Protoderme, Pc – Procâmbio, Mfm – Meristema fundamental medular, Mfc – meristema fundamental cortical.



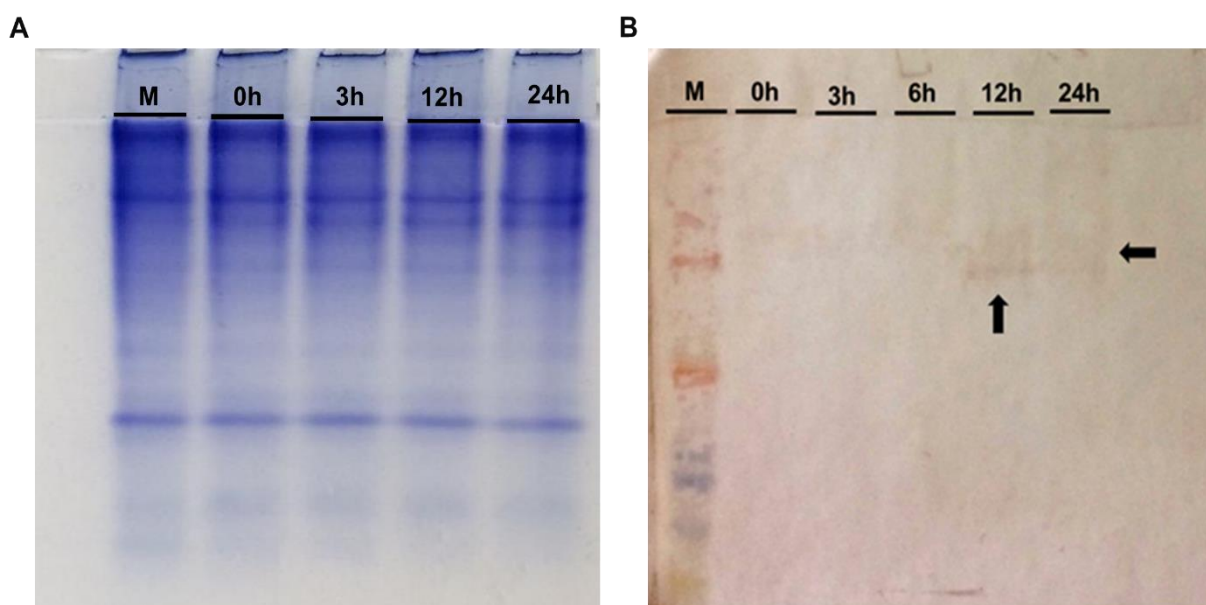


**Figura 20** Fluorescência relativa de celulose nos eixos de 3 e 24 horas após a embebição. A-C. eixo embrionário de 3h, D-F eixo embrionário de 24h. (A, D) meristema radicular; (B, E) procâmbio; (C, F) meristema apical. G. Gráfico de fluorescência relativa de celulose nos tecidos. \* diferenças entre as médias estatisticamente significativas com  $p \leq 0,05$  \*\* - diferenças entre as médias estatisticamente significativas com  $p \leq 0,01$ .

### 5.3.4. Determinação da atividade peroxidase

A peroxidase apresentou alta expressão ao longo do desenvolvimento tanto da semente quanto do eixo embrionário. Para reforçar a compreensão das atividades metabólicas de parede celular que estão ocorrendo ao longo do desenvolvimento do eixo embrionário durante a germinação, realizamos a determinação da atividade da enzima peroxidase em membrana após a separação em gel.

Após a eletroforese e posterior transferência das proteínas do gel para membrana e revelação da atividade, foi possível observar duas bandas majoritárias tanto em 12 quanto em 24 horas após a embebição (Figura 21). De acordo com esta análise, a peroxidase apresenta atividade reduzida em eixo embrionário nas primeiras horas de embebição, aumentando a sua atividade ao longo da germinação da semente, o que coincide com os dados de expressão gênica.



**Figura 21** Atividade da enzima peroxidase determinada em membrana após a separação das proteínas por eletroforese nativa a 12%. M – Marcador de massa molecular. 0, 3, 6, 12 e 24 horas após a embebição do eixo embrionário.

## 6 DISCUSSÃO

O Brasil é o segundo maior produtor global de soja, que por sua vez, é uma das mais importantes fontes de óleo comestível e proteína. A soja se tornou a primeira espécie de leguminosa a ter seu genoma totalmente sequenciado (Schmutz et al., 2010), abrindo novas possibilidades de estudos genômicos com esta espécie. Junto a este avanço, o desenvolvimento de novas ferramentas de análise gênica como o RNA-Seq, que é capaz de fornecer o perfil do transcriptoma em diferentes tecidos e fases do desenvolvimento, permite a realização de pesquisas comparativas de processos importantes para o desenvolvimento celular (Wang et al., 2010).

O embrião da semente, órgão vital para a futura plântula, inicia sua formação no momento da fertilização do óvulo e desenvolve-se durante o processo de maturação da semente até que o grau de umidade diminua, o que impossibilita o seu crescimento. O processo de germinação inclui uma combinação de processos biofísicos, bioquímicos e fisiológicos (Bradford; Goldberg et al., 1994). Durante todo esse processo, a parede celular desempenha papel crucial no crescimento e desenvolvimento celular, garantindo o sucesso da formação da futura plântula (Reid, 1985; Ben-Tov et al., 2015).

A germinação inicia-se com a embebição, onde a quantidade de água absorvida pela semente deve ser suficiente para retomar e continuar o processo de desenvolvimento do embrião. A absorção de água pelas células resulta no afrouxamento da parede celular, o que facilita a expansão celular e a protusão do

eixo (Bewley et al., 2013). Com a entrada de água, iniciam-se os processos bioquímicos preparatórios, como a digestão, translocação e assimilação das reservas acumuladas ao longo da formação da semente para o correto desenvolvimento do embrião. Os fatores que limitam a expansão do embrião devem ser regulados de maneira apropriada em sementes. Para tal, uma série de genes devem ser finamente regulados (Bewley et al., 2013). A compreensão dos mecanismos moleculares que regem o desenvolvimento da semente e do embrião podem fornecer valiosas informações a respeito do controle do desenvolvimento e germinação das sementes.

Assim, o presente estudo visou caracterizar as variações de expressão dos genes envolvidos com a biossíntese da parede celular vegetal em diversos estágios do desenvolvimento de sementes de soja, com enfoque na regulação de parede celular em eixo embrionário durante a germinação. Para tal, realizamos uma identificação e análise global das vias que estavam sendo enriquecidas, tanto em frequência gênica quanto em expressão durante os estágios de desenvolvimento da semente e do eixo embrionário durante a germinação. Além disso, realizamos a análise de uma família gênica de parede celular e também a caracterização morfológico-tecidual do eixo embrionário durante a germinação de sementes de soja.

### **6.1. Identificação de genes de parede celular em soja**

Pela importância das contribuições da parede celular e de seus componentes para inúmeros processos biológicos da planta e para processos industriais, como produção de papel (Ververis et al., 2004), qualidade e textura de alimentos (Brummell e Harpster, 2001), produção de bioetanol (Sticklen, 2008), entre outros, cresce as tentativas de manipulação genética e técnicas de produção com o intuito de modificar tanto a qualidade quanto a quantidade dos componentes da parede.

No entanto, a manipulação dos principais componentes da parede, os polissacarídeos, tem sido dificultada pela falta de conhecimento sobre os mecanismos de controle das etapas de biossíntese dos componentes juntamente com a pouca compreensão das interações físicas e químicas entre os mesmos.

Contudo, a disponibilidade de sequências da parede celular de *Arabidopsis* (Building, 2010) e os avanços das abordagens genômicas têm facilitado a identificação e descrição dos processos genéticos envolvidos na biossíntese dos componentes da parede, como os relacionados à celulose (Arioli et al., 1998; Appenzeller et al., 2004; Persson et al., 2005; Paredez, 2006; Endler e Persson, 2011; McFarlane et al., 2014; Kaur et al., 2016), hemiceluloses (Burton, 2006; Bernal et al., 2008; Liepman e Cavalier, 2012), pectinas (Willats et al., 2001; Gwanpua et al., 2016; Zega e D'Ovidio, 2016; Lin et al., 2017; Wallace e Williams, 2017) e de proteínas relacionadas a parede celular (Lindstrom e Vodkin, 1991; Roudier, 2002; Draeger et al., 2015; Cosgrove, 2016b; Tanaka et al., 2017).

Um dos maiores desafios na identificação dos genes responsáveis pela formação da parede celular é o grande número de genes e de famílias gênicas envolvidas em todo o processo. Além disso, as funções de muitos desses genes ainda são desconhecidas. No presente estudo, usando *Arabidopsis* como referência, conseguimos identificar o total de 2.143 genes responsáveis pela biossíntese e montagem da parede celular em soja. As vias que tiveram o maior número de genes foram as de biossíntese de fenilpropanoides (220), Peroxidasas (188), Pectinas esterases (132), Poligalacturanas (108), GT família 8 (86) e vias de interconversão de açúcar-nucleotídeos (83).

Dentre as proteínas estruturais e modificadores da parede celular, foi analisada mais detalhadamente neste trabalho a família COBRA em soja. A família COBRA é tida como necessária para correta expansão celular (Schindelman et al., 2001; Roudier, 2002), mediação do crescimento do tubo polínico (Li et al., 2013), rendimento de grãos (Hochholdinger et al., 2008) e na interferência direta no conteúdo de celulose da parede celular (Thumma et al., 2009; Liu et al., 2013), sendo requerido para a síntese da mesma (Persson et al., 2005).

Nesta ocasião, identificamos um total de 23 genes COBRA em todo o genoma de soja, o que corresponde a mais que o dobro de genes COBRA em *Arabidopsis* (11) (Tabela 2). Esse aumento no número de genes COBRA em soja parece ser devido às duplicações sofridas no genoma da soja (Schmutz et al., 2010), também analisadas nesta mesma ocasião. Todos os genes COBRA de *Arabidopsis*, milho e arroz, exceto AtCOBL1, AtCOBL4 e AtCOBL5, foram confirmados como proteínas com âncoras GPI (Borner et al., 2002; Roudier, 2002; Li et al., 2003; Brady et al., 2007). Uma âncora GPI indica que a proteína está

localizada na superfície das membranas biológicas no espaço extracelular (Eisenhaber et al., 1998, 1999). A nossa anotação dos genes COBRA reforçaram a ideia do seu potencial como proteínas com âncoras GPI em soja (Tabela 2). Alguns autores sugerem que árvores filogenéticas podem ser usadas para determinar a ontologia de genes em famílias gênicas (Eisen, 1998; Remm et al., 2001; Wu et al., 2006). No presente estudo, a análise filogenética mostrou que os 23 membros da família COBRA em soja podem ser divididos em cinco clados: um com similaridade com COB, outro com similaridade com COBL7, outro semelhante ao COBL4 e os outros dois grupos com similaridade com COBL10 e 11 e com COBL2 e 1 (Figura 11). A diferença na estrutura gênica quanto ao número de éxons e íntrons oferece um suporte à essa divisão em cinco clados (Figura 13).

Após a análise da sequência gênica dos membros da família COBRA em soja, focamos em prever a função de todos os membros identificados. Alguns membros da família COBRA já possuem sua função bem conhecida em outros organismos. A subfamília COB já foi predita em *Populus* como expressa em resposta a estresses e hormônios (Ye et al., 2009). Em *Arabidopsis*, já foi visto que essa subfamília é responsável pela expansão celular (Schindelman et al., 2001) e em arroz, COB é expresso principalmente no desenvolvimento de células do esclerênquima, em feixes vasculares e na montagem das microfibrilas de celulose (Li et al., 2003; Liu et al., 2013). Membros da subfamília COB em algodão obtiveram correlações significativas com características da qualidade da fibra (Niu et al., 2015) e em tomate a subfamília COB desempenha um papel importante na arquitetura da parede celular e fornece uma potencial ferramenta genética para prolongar a vida útil do fruto (Cao et al., 2012). No presente estudo, os membros da subfamília COB de soja apresentaram expressão em diferentes órgãos e tecidos, o que pode indicar que os membros dessa subfamília são requeridos para a correta arquitetura da parede celular, ancorando as microfibrilas de celulose.

Membros da subfamília COBL4 já foram preditos como reguladores da deposição da parede celular secundária (Sato et al., 2010), mostrando expressão preferencial nos troncos e caules (Niu et al., 2015). Em soja, alguns membros da subfamília COBL4 não apresentaram expressão. No entanto, o gene GmCOBL4c apresentou expressão no caule (Figura 14). Já os membros da subfamília COBL10 parecem expressar-se apenas nos tecidos florais de soja. Em *Arabidopsis*, foi relatado que mutações no gene COBL10 causam esterilidade gametófito devido ao

crescimento reduzido do tubo polínico (Li et al., 2013), o que poderia explicar a expressão preferencial deste gene em tecidos reprodutivos de soja.

No geral, na análise de expressão gênica, demonstra-se que os subgrupos da família COBRA possuem padrões tecido-específico e funções distintas em soja, o que pode sugerir que os membros da família COBRA são expressos em tecidos específicos para facilitar a expansão das células vegetais.

## **6.2. Correlação entre a expressão gênica e dinâmica da parede celular**

A parede celular da planta é uma estrutura altamente organizada e dinâmica, que fornece não só suporte mecânico para a célula, mas também responde a inúmeros fatores bióticos e abióticos. Durante os estágios de desenvolvimento e germinação das sementes de soja, os genes da parede celular contribuem para o enriquecimento das vias de geração de substrato (açúcar nucleotídeos e biossíntese de fenilpropanóide), endocitose e família GT 8 (Figura 16).

O metabolismo de fenilpropanóides, maior classe de componentes encontrado na identificação dos genes, compreende uma série de reações bioquímicas que fornecem a planta importantes compostos fenólicos como flavonóides, antocianinas e taninos (Weisshaar e Jenkinst, 1998). Dentre as muitas funções desses compostos estão inclusas as de componentes estruturais, produção de pigmentos, proteção contra estresses e sinalização (Weisshaar e Jenkinst, 1998). Em sementes de soja, já foi descrita anteriormente a biossíntese e acumulação de isoflavonóides durante todo o desenvolvimento da semente (Dhaubhadel et al., 2003), o que corrobora com a alta expressão dessa via durante o processo de formação e desenvolvimento da semente.

As GTs estão entre os maiores grupos de enzimas e podem ser divididas em mais de 50 famílias em plantas (Keegstra e Raikhel, 2001). No geral, a homologia dentro das famílias GT é baixa, o que dificulta a definição de função baseada na similaridade de sequências (Breton e Imberty, 1999; Rosén et al., 2004). No entanto, a maioria das famílias dessas enzimas catalisam a formação de ligações glicosídicas, utilizando doadores de açúcar que possuem um grupo de

fosfato (Lairson et al., 2008). Essa atividade seria importante para a montagem da parede celular nos tecidos em desenvolvimento, como os de sementes.

As vias de interconversão de açúcar nucleotídeos representam uma série de reações enzimáticas pelas quais as plantas sintetizam monossacarídeos ativados para a formação dos polissacarídeos da parede celular (Seifert, 2004). As glicosiltransferases utilizam os açúcares nucleotídeos como doadores de monossacarídeos na síntese da parede celular (Gibeaut, 2000; Seifert, 2004; Reiter, 2008). Esse fato explica o enriquecimento dessas vias durante o processo de desenvolvimento da semente e germinação do eixo embrionário. Provavelmente, essas vias são requeridas durante esses estágios para a formação e remodelamento da parede celular.

Outra via que também se manteve em alta expressão durante o desenvolvimento tanto da semente quanto na germinação do eixo embrionário foi a via de peroxidase. As peroxidases de parede celular possuem uma dupla atividade. Elas são capazes de regular os níveis de espécies reativas de oxigênio geradas no metabolismo celular como o peróxido de hidrogênio, desempenhando um papel fundamental no crescimento celular, controlando a síntese/fortalecimento e o afrouxamento da parede (Kawano, 2003; Passardi et al., 2004; Francoz et al., 2015). No presente trabalho, além da análise *in silico* da expressão das vias de peroxidase durante a germinação do eixo embrionário, determinamos a atividade da peroxidase no decorrer das primeiras horas de germinação. Nesta análise, a peroxidase apresentou aumento da sua atividade nos pontos de 12 e 24h de germinação do eixo embrionário (Figura 22). Nas primeiras horas de contato da semente quiescente com a água, ocorre a entrada de água nas células e a consequente expansão celular. A expressão crescente de expansinas a partir do ponto de 3h de embebição está fortemente associada ao afrouxamento da parede celular. A peroxidase é capaz de contribuir no afrouxamento da parede celular regulando a concentração local de H<sub>2</sub>O<sub>2</sub> ou gerando radicais de oxigênio que podem quebrar as ligações covalentes nos polímeros da parede celular (Francoz et al., 2015), facilitando o alongamento do eixo embrionário.

O conteúdo de celulose durante a germinação do eixo embrionário de sementes de soja também foi analisado. A celulose é o polímero mais abundante na parede celular. A síntese de celulose afeta diretamente o desenvolvimento tanto da semente quanto do eixo embrionário, pois um desenvolvimento bem-sucedido



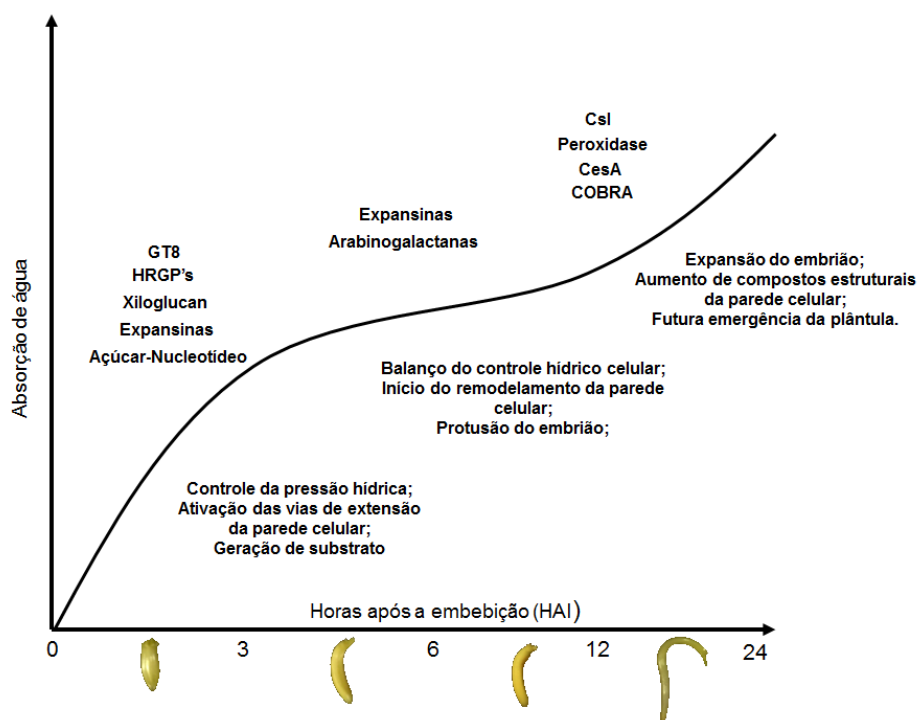
está intimamente relacionado à integridade da parede. A extensão da parede celular começa com o relaxamento da mesma, o que permite que as células absorvam água. Nesse processo também são atuantes as expansinas que induzem o relaxamento e extensão da parede rompendo as ligações não covalentes entre as microfibrilas de celulose (Cosgrove, 2005; Paredez, 2006; Li et al., 2014; McFarlane et al., 2014; Nawaz et al., 2017). Como mostrado anteriormente, a expressão da via de expansinas aumenta durante o processo de germinação do eixo embrionário (Figura 16D). De fato, a partir do momento em que a semente entra em contato com a água deve ocorrer um relaxamento da parede celular, a fim de permitir a extensão da mesma. Além disso, nossos resultados demonstraram que o conteúdo de celulose das células aumenta no decorrer da germinação (Figura 20), o que corrobora com a ideia de que além das células se expandirem nos primeiros estágios da germinação, a síntese de celulose aumenta gradativamente ao longo da germinação para promover a biossíntese de parede celular.

A partir dos dados de expressão gênica e análise morfológica-tecidual, propõe-se um esquema representativo dos principais eventos e vias da parede celular ativadas no eixo embrionário durante o processo de germinação de sementes de soja (Figura 22). No início da embebição da semente, ocorre a absorção de água pelas células do embrião. Nesse estágio, ocorre a reidratação dos tecidos que sofreram desidratação durante os estágios de desenvolvimento da semente. Processos metabólicos, como respiração e digestão, devem ser reiniciados. A fonte inicial de energia para a glicólise vem da sacarose e de pequenos oligossacarídeos já presentes na semente. As vias de geração de substrato, açúcares nucleotídeos e família GT8 começam a ser ativadas (Figura 17A e B). Nessa etapa, a parede celular precisa ser suficientemente forte para controlar a pressão hidráulica da célula ao mesmo tempo em que precisa relaxar para que ocorra a protusão da radícula. Assim, encontra-se altamente expressos genes responsáveis pela extensão da parede como as HRGP's (extensinas) (Figura 17I) que, em sua maioria, encontram-se associadas aos componentes pécticos da parede e as expansinas (Figura 17C). Essas últimas rompem as ligações não covalentes entre as microfibrilas de celulose, permitindo o relaxamento da mesma. Nessa etapa, também se encontram altamente expressos os genes relacionados ao metabolismo dos xiloglucanos (Figura 17J), que vão diminuindo a sua expressão ao longo da germinação do embrião. Esse grupo de hemiceluloses que estavam

armazenados na semente, podem auxiliar no controle da expansão celular através do enfraquecimento das ligações entre microfibrilas de celulose na parede celular (Sechet et al., 2016) facilitando a protusão do embrião.

Com o passar das horas de germinação, a semente é totalmente hidratada e as células encontram-se em turgor. Nesse momento, a absorção de água é limitada pela parede celular. As vias ativadas inicialmente vão diminuindo sua expressão, com exceção das expansinas. Novas vias aumentam sua expressão, como a de arabinogalactanos (Figura 17E). As arabinogalactanas podem estar auxiliando no balanço do controle hídrico na célula, impedindo que a célula absorva mais água que o necessário (Houston et al., 2016). Com a participação de arabinogalactanas e expansinas, o crescimento do embrião torna-se contínuo e a absorção de água é retomada, dando início a protusão do embrião.

Entre 12 e 24 horas de embebição, ocorre a ativação das vias de peroxidase, celulose sintase, celulose sintase-like e COBRA (Figura 17D, F, G e H). As vias de peroxidases regulam a concentração local de  $H_2O_2$  gerado pelo metabolismo celular (Francoz et al., 2015). Durante o crescimento e desenvolvimento celular é necessário um aumento extensivo dos materiais que compõem a estrutura da parede celular. A ativação das vias de celulose e celulose sintase-like contribuem para formação da estrutura da parede celular no embrião em crescimento. Nesse estágio, os genes COBRA também são mais expressos (Figura 17H), o que contribui para a deposição de celulose na parede celular. Ao final da germinação as vias da parede celular responsáveis pela formação de uma parede primária mais complexa com a presença de celulose e hemicelulose, além da interação dessas com proteínas COBRA, parecem preparar o embrião para uma série de divisões celulares que irão propiciar o desenvolvimento da futura plântula.



**Figura 22** - Principais eventos metabólicos e vias ativadas durante a germinação do eixo embrionário. Durante o período de 0 para 3h, a semente absorve água e ativa vias responsáveis pela extensão da parede celular. Entre o período de 6 a 12 horas após a absorção de água, a parede celular participa do balanço hídrico celular e tem-se início o remodelamento da parede celular para a protusão do embrião. Entre 12 e 24 horas após a embebição, pela ação do remodelamento e expansão da parede celular, o embrião se expande.

## 7 CONCLUSÕES

Nosso trabalho é o primeiro a utilizar dados de transcriptômica para identificar e elucidar as vias de biossíntese da parede celular durante os estágios de desenvolvimento da semente e germinação do eixo embrionário em soja. Nossos resultados fornecem uma base para uma caracterização funcional e bioquímica dos componentes da parede celular que controlam não só a biossíntese, mas também a expansão da parede. Dentre estes resultados, destacamos:

- A identificação de um total de 2.143 genes relacionados a parede celular no genoma de soja. Todos os genes foram divididos em suas respectivas vias de biossíntese e montagem da parede celular. Estes dados podem servir como base para a seleção de genes relacionados a estrutura da parede celular em soja com vistas à manipulação genética.

- A caracterização de 23 genes COBRA em soja e a descoberta da expansão da família COBRA, onde genes parálogos podem ter adquirido funções diferentes após o último evento de duplicação do genoma de soja.

- Durante o desenvolvimento e germinação de sementes de soja, o número de genes expressos envolvidos nas vias de geração de substrato (açúcar nucleotídeos e biossíntese de fenilpropanóides), endocitose e família GT 8 foi muito similar. No entanto, somente durante a germinação as vias xiloglucanos,

expansinas, proteases, famílias GT 8, HRGPs, e celulose sintase foram altamente expressas.

- As análises morfológicas-teciduals mostraram que as células do eixo embrionário se expandem ao longo da germinação. Quanto ao conteúdo relativo de celulose nessas células, também ocorre um aumento significativo em todo o eixo embrionário no decorrer da germinação, corroborando os dados de aumento da expressão de genes de celulose sintase ao longo do desenvolvimento do eixo embrionário durante a germinação da semente.

- A enzima peroxidase apresenta atividade reduzida em eixo embrionário nas primeiras horas do evento de germinação, aumentando a sua atividade nos pontos de 12 e 24 horas de embebição da semente, o que coincide com os dados de expressão gênica para essa enzima. Esta atividade parece ser importante para o afrouxamento da parede celular para facilitar o alongamento do eixo embrionário.

- A embebição da semente de soja em água modula a expressão de genes associados à estrutura da parede celular, ativando vias responsáveis pela sua extensão logo nas primeiras horas de germinação. Após isso, a parede celular participa do balanço hídrico celular e também do remodelamento da parede celular para a protusão do embrião. Nas últimas 12 horas do evento de germinação, a biossíntese da parede celular parece ser aumentada para propiciar o crescimento do embrião.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Appenzeller, L., Doblin, M., Barreiro, R., Wang, H., Niu, X., Kollipara, K., Carrigan, L., Tomes, D., Chapman, M., Dhugga, K.S. (2004) Cellulose synthesis in maize: isolation and expression analysis of the cellulose synthase (CesA) gene family. *Cellulose* 11: 287–299.
- Arioli, T., Peng, L., Bethner, S.A., Burn, J., Wittke, W., Herth, W., Camilleri, C., Hfte, H., Platzinski, J., Birch, R., Cork, A., Glover, J., Richmond, J., R.E, W. (1998) Molecular Analysis of Cellulose Biosynthesis in Arabidopsis. *Science* (80- ) 717–720.
- Belliény-Rabelo, D., de Oliveira, E.A.G., Ribeiro, E. da S., Costa, E.P., Oliveira, A.E.A., Venancio, T.M. (2016) Transcriptome analysis uncovers key regulatory and metabolic aspects of soybean embryonic axes during germination. *Sci Rep* 6: 36009.
- Bemiller, J.N., Lafayette, W. (2001) Plant Cell Walls: Economic Significance. *Science* (80- ) 1–5.
- Ben-Tov, D., Abraham, Y., Stav, S., Thompson, K., Loraine, A., Elbaum, R., de Souza, A., Pauly, M., Kieber, J.J., Harpaz-Saad, S. (2015) COBRA-LIKE2, a member of the glycosylphosphatidylinositol-anchored COBRA-LIKE family, plays a role in cellulose deposition in arabidopsis seed coat mucilage secretory cells. *Plant Physiol* 167: 711–24.

- Bernal, A.J., Jensen, J.K., Harholt, J., Sørensen, S., Møller, I., Blaukopf, C., Johansen, B., De Lotto, R., Pauly, M., Scheller, H.V., Willats, W.G.T. (2007) Disruption of ATCSLD5 results in reduced growth, reduced xylan and homogalacturonan synthase activity and altered xylan occurrence in *Arabidopsis*. *Plant J* 52: 791–802.
- Bernal, A.J., Yoo, C.-M., Mutwil, M., Jensen, J.K., Hou, G., Blaukopf, C., Sørensen, I., Blancaflor, E.B., Scheller, H.V., Willats, W.G.T. (2008) Functional Analysis of the Cellulose Synthase-Like Genes CSLD1, CSLD2, and CSLD4 in Tip-Growing *Arabidopsis* Cells. *Plant Physiol* 148: 1238–1253.
- Bewley, J.D., Bradford, K., Hilhorst, H., Nonogaki, H. (2013) *Seeds*. 3rd.ed. New York Heidelberg Dordrecht London: Springer, 381p.
- Bois. (2008) *The world of soy*. 337p.
- Borner, G.H.H., Sherrier, D.J., Stevens, T.J., Arkin, I.T., Dupree, P. (2002) Prediction of Glycosylphosphatidylinositol-Anchored Proteins in *Arabidopsis*. A Genomic Analysis 1. 129: 486–499.
- Boron, A.K., Van Orden, J., Nektarios Markakis, M., Mouille, G., Adriaensen, D., Verbelen, J.-P., Höfte, H., Vissenberg, K. (2014) Proline-rich protein-like PRPL1 controls elongation of root hairs in *Arabidopsis thaliana*. *J Exp Bot* 65: 5485–5495.
- Bradford, K.J. Seed Development , Dormancy and Germination Seed Development , Dormancy Edited by.
- Brady, S.M., Song, S., Dhugga, K.S., Rafalski, J.A., Benfey, P.N. (2007) Combining Expression and Comparative Evolutionary Analysis. The *COBRA* Gene Family. *Plant Physiol* 143: 172–187.
- Breton, C., Imberty, A. (1999) Structure/function studies of glycosyltransferases. *Curr Opin Struct Biol* 9: 563–571.
- Brown, D.M., Goubet, F., Wong, V.W., Goodacre, R., Stephens, E., Dupree, P., Turner, S.R. (2007) Comparison of five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis. *Plant J* 52: 1154–1168.

- Brown, R.M. (2004) Cellulose Structure and Biosynthesis: What is in Store for the 21st Century?. *J Polym Sci Part A Polym Chem* 42: 487–495.
- Brummell, D.A., Harpster, M.H. (2001) Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants. *Plant Mol Biol* 47: 311–339.
- Buckeridge, M.S. (2010) Seed cell wall storage polysaccharides: Models to understand cell wall biosynthesis and degradation. *Plant Physiol* 154: 1017–1023.
- Buckeridge, M.S., Tiné, M.A.S. (2001) Composição polissacarídica: estrutura da parede celular e fibra alimentar. \_\_\_\_\_ *Fibra dietética Iberoamérica Tecnol y salud São Paulo Sarvier* 27–38.
- Buckeridge, M.S., Tiné, M.A.S., SANTOS, H.P. dos., LIMA, D.U. de., dos Santos, H.P., de Lima, D.U. (2000) Polissacarídeos de reserva de parede celular em sementes. Estrutura, metabolismo, funções e aspectos ecológicos. *Rev Bras Fisiol Veg* 12: 137–162.
- Building, M.J. (2010) Arabidopsis – a powerful model system for plant cell wall. 1107–1121.
- Burton, R.A. (2006) Cellulose Synthase-Like CslF Genes Mediate the Synthesis of Cell Wall (1,3;1,4)-D-Glucans. *Science* (80- ) 311: 1940–1942.
- Cao, Y., Tang, X., Giovannoni, J., Xiao, F., Liu, Y. (2012) Functional characterization of a tomato COBRA-like gene functioning in fruit development and ripening. *BMC Plant Biol* 12: 211.
- Carpenter, C., Kreps, J.A., Simon, A.E. (1994) Genes encoding glycine-rich Arabidopsis thaliana proteins with RNA-binding motifs are influenced by cold treatment and an endogenous circadian rhythm. *Plant Physiol* 104: 1015–1025.
- Carpita, N.C. (2011) Update on Mechanisms of Plant Cell Wall Biosynthesis: How Plants Make Cellulose and Other (1->4)-D-Glycans. *Plant Physiol* 155: 171–184.
- Carpita, N.C. (1996) Structure and biogenesis of the cell walls of grasses. *Annu Rev*



Plant Physiol Plant Mol Biol. doi: 10.1146/annurev.arplant.47.1.445

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

Carpita, N.C., Gibeaut, D.M. (1993b) Structural models of primary cell walls in flowering land plants – consistency of molecular structure with the physical properties of the walls during growth. *Plant J.* doi: 10.1111/j.1365-313X.1993.tb00007.x

Carpita, N.C., McCann, M.C. (2002) The functions of cell wall polysaccharides in composition and architecture revealed through mutations. *Plant Soil.* doi: 10.1023/A:1021115300942

Cassab, G.I. (1986) Arabinogalactan proteins during the development of soybean root nodules. *Planta* 168: 441–446.

Cassab, G.I. (1998) Plant Cell Wall Proteins. *Stand Annu Rev Plant Physiol Plant Mol Biol* 49: 281–309.

Clarke, A.E., Anderson, R.L., Stone, B.A. (1979) Form and function of arabinogalactans and arabinogalactan-proteins. *Phytochemistry* 18: 521–540.

CONAB: Companhia Nacional de Abastecimento. (2017) Acompanhamento da safra brasileira: grãos. *Monit Agric Safra 2017* 4: 1–98.

Conery, J.S. (2000) The Evolutionary Fate and Consequences of Duplicate Genes.

Cosgrove, D.J. (2005) Growth of the plant cell wall. *Nat Rev Mol Cell Biol* 6: 850–861.

Cosgrove, D.J. (2016a) Catalysts of plant cell wall loosening. *F1000Research* 5: 1–13.

Cosgrove, D.J. (2000) Loosening of plant cell walls by expansins. *Nature* 407: 321–326.

Cosgrove, D.J. (2016b) Plant cell wall extensibility: Connecting plant cell growth with

- cell wall structure, mechanics, and the action of wall-modifying enzymes. *J Exp Bot* 67: 463–476.
- Cosgrove, D.J. (2015) Plant expansins: Diversity and interactions with plant cell walls. *Curr Opin Plant Biol* 25: 162–172.
- Darley, C.P., Forrester, A.M., McQueen-Mason, S.J. (2001) The molecular basis of plant cell wall extension. *Plant Mol Biol* 47: 179–195.
- Datta, K., Schmidt, A., Marcus, A. (1989) Characterization of two soybean repetitive proline-rich proteins and a cognate cDNA from germinated axes. *Plant Cell* 1: 945–952.
- Dhaubhadel, S., McGarvey, B.D., Williams, R., Gijzen, M. (2003) Isoflavonoid biosynthesis and accumulation in developing soybean seeds. *Plant Mol Biol* 53: 733–743.
- Domozych, D.S., Serfis, A., Kiemle, S.N., Gretz, M.R. (2007) The structure and biochemistry of charophycean cell walls: I. Pectins of *Penium margaritaceum*. *Protoplasma*. doi: 10.1007/s00709-006-0197-8
- Draeger, C., Ndinyanka Fabrice, T., Gineau, E., Mouille, G., Kuhn, B.M., Moller, I., Abdou, M.-T., Frey, B., Pauly, M., Bacic, A., Ringli, C. (2015) Arabidopsis leucine-rich repeat extensin (LRX) proteins modify cell wall composition and influence plant growth. *BMC Plant Biol* 15: 155.
- Du, J., Wang, S., He, C., Zhou, B., Ruan, Y., Shou, H. (2017) Identification of regulatory networks and hub genes controlling soybean seed set and size using RNA sequencing analysis. 68: 1955–1972.
- Du, Z., Zhou, X., Ling, Y., Zhang, Z., Su, Z. (2010) agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Res* 38: W64-70.
- Duan, J., Xia, C., Zhao, G., Jia, J., Kong, X. (2012) Optimizing de novo common wheat transcriptome assembly using short-read RNA-Seq data. *BMC Genomics* 13: 392.
- Egelund, J., Petersen, B.L., Motawia, M.S., Damager, I., Faik, A., Olsen, C.E., Ishii,

- T., Clausen, H., Ulvskov, P., Geshi, N. (2006) *Arabidopsis thaliana* RGXT1 and RGXT2 Encode Golgi-Localized (1,3)- $\alpha$ -D-Xylosyltransferases Involved in the Synthesis of Pectic Rhamnogalacturonan-II. *Plant Cell* 18: 2593–2607.
- Eisen, J.A. (1998) Phylogenomics: Improving Functional Predictions for Uncharacterized Genes by Evolutionary Analysis. 163–167.
- Eisenhaber, B., Bork, P., Eisenhaber, F. (1998) Sequence properties of GPI-anchored proteins near the omega-site: constraints for the polypeptide binding site of the putative transamidase. *Protein Eng Des Sel* 11: 1155–1161.
- Eisenhaber, B., Bork, P., Eisenhaber, F. (1999) Prediction of Potential GPI-modification Sites in Proprotein Sequences. *J Mol Biol* 292: 741–758.
- Embrapa. (2017) História da Soja. <https://www.embrapa.br/soja/cultivos/soja1/historia>. Accessed 12 Dec 2017
- Endler, A., Persson, S. (2011) Cellulose synthases and synthesis in *Arabidopsis*. *Mol Plant* 4: 199–211.
- Evert, R.F., Eichhorn, S.E. (2014) *Biologia Vegetal*. 628-712p.
- FAO. (2009) Global agriculture towards 2050. In: Glob. Agric. Towar. 2050.
- FAOSTAT. (2017) Food and Agriculture Organization of the United Nations. <http://www.fao.org/faostat/en/#data/QC/visualize>.
- Federizzi, L. (2017) *A SOJA COMO FATOR DE COMPETITIVIDADE NO MERCOSUL : HISTÓRICO, PRODUÇÃO E PERSPECTIVAS FUTURAS*.
- Fernandez, A.G., Fujita, M., Himmelpach, R., Borner, G.H.H., Schindelman, G., Song, S., Baskin, T.I., Dupree, P., Wasteneys, G.O., Benfey, P.N. (2005) COBRA, an *Arabidopsis* Extracellular Glycosyl-Phosphatidyl Inositol-Anchored Protein, Specifically Controls Highly Anisotropic Expansion through Its Involvement in Cellulose Microfibril Orientation. 17: 1749–1763.
- Fincher, G.B. (2009) Revolutionary Times in Our Understanding of Cell Wall Biosynthesis and Remodeling in the Grasses. *Plant Physiol* 149: 27–37.

- Francoz, E., Ranocha, P., Nguyen-Kim, H., Jamet, E., Burlat, V., Dunand, C. (2015) Roles of cell wall peroxidases in plant development. *Phytochemistry* 112: 15–21.
- Fry, S.C. (1988) *The growing plant cell wall: chemical and metabolic analysis*. Harlow: Longman Group Limited, 333 pp.p.
- Garcia, A., Calvo, É.S., De Souza Kiihl, R.A., Harada, A., Hiromoto, D.M., Vieira, L.G.E. (2008) Molecular mapping of soybean rust (*Phakopsora pachyrhizi*) resistance genes: Discovery of a novel locus and alleles. *Theor Appl Genet* 117: 545–553.
- Gibeaut, D.M. (2000) Nucleotide sugars and glycosyltransferases for synthesis of cell wall matrix polysaccharides. *Plant Physiol Biochem* 38: 69–80.
- Goldberg, R.B., Paiva, G. De., Yadegari, R. (1994) Plant Embryogenesis: Zygote to Seed.
- Grau, C., Dorrance, A., Bond, J., Russin, J. (2004) *SOYBEANS: Improvement, Production, and Uses*. 679-744p.
- Gwanpua, S.G., Mellidou, I., Boeckx, J., Kyomugasho, C., Bessemans, N., Verlinden, B.E., Hertog, M.L.A.T.M., Hendrickx, M., Nicolai, B.M., Geeraerd, A.H. (2016) Expression analysis of candidate cell wall-related genes associated with changes in pectin biochemistry during postharvest apple softening. *Postharvest Biol Technol* 112: 176–185.
- Hall, Q., Cannon, M.C. (2002) The cell wall hydroxyproline-rich glycoprotein RSH is essential for normal embryo development in *Arabidopsis*. *Plant Cell* 14: 1161–1172.
- He, C.-Y., Zhang, J.-S., Chen, S.-Y. (2002) A soybean gene encoding a proline-rich protein is regulated by salicylic acid, an endogenous circadian rhythm and by various stresses. *Theor Appl Genet* 104: 1125–1131.
- Henry, R.J. (1987) Pentosan and (1 → 3),(1 → 4)- $\beta$ -Glucan concentrations in endosperm and wholegrain of wheat, barley, oats and rye. *J Cereal Sci* 6: 253–258.

- Hochholdinger, F., Wen, T.J., Zimmermann, R., Chimot-Marolle, P., Da Costa E Silva, O., Bruce, W., Lamkey, K.R., Wienand, U., Schnable, P.S. (2008) The maize (*Zea mays* L.) roothairless3 gene encodes a putative GPI-anchored, monocot-specific, COBRA-like protein that significantly affects grain yield. *Plant J* 54: 888–898.
- Höfte, H., Voxeur, A. (2017) Plant cell walls. *Curr Biol* 27: R865–R870.
- Hong, J.C., Nagao, R.T., Key, J.L. (1990) Characterization of a proline-rich cell wall protein gene family of soybean. A comparative analysis. *J Biol Chem* 265: 2470–2475.
- Hoon Ahn J., Choi, Y., Kim, S., Myung Kwon Y., Do Choi Y., Seob Lee J. (1998) Expression of a Soybean Hydroxyproline-Rich Glycoprotein Gene Is Correlated with Maturation of Roots. *Plant Physiol* 116: 671–9.
- Houston, K., Tucker, M.R., Chowdhury, J., Shirley, N., Little, A. (2016) The Plant Cell Wall: A Complex and Dynamic Structure As Revealed by the Responses of Genes under Stress Conditions. *Front Plant Sci* 7: 984.
- Hu, B., Jin, J., Guo, A.Y., Zhang, H., Luo, J., Gao, G. (2015) GSDS 2.0: An upgraded gene feature visualization server. *Bioinformatics* 31: 1296–1297.
- IBGE. (2017) No Title. In: IBGE prevê safra grãos 8,9% menor em 2018. <https://agenciadenoticias.ibge.gov.br/2013-agencia-de-noticias/releases/17901-ibge-preve-safra-de-graos-8-9-menor-em-2018.html>.
- Illumina. (2017) NextSeq Series | High-throughput benchtop sequencers for everyday genomics.
- Jones, S.I., Vodkin, L.O. (2013) Using RNA-Seq to Profile Soybean Seed Development from Fertilization to Maturity. doi: 10.1371/journal.pone.0059270
- Kakumanu, A., Ambavaram, M.M.R., Klumas, C., Krishnan, A., Batlang, U., Myers, E., Grene, R., Pereira, A. (2012) Effects of Drought on Gene Expression in Maize Reproductive and Leaf Meristem Tissue Revealed by RNA-Seq. *Plant Physiol* 160: 846–867.

- Karnovsky, M. (1964) *A Formaldehyde-Glutaraldehyde Fixative of High Osmolality for Use in Electron Microscopy*.
- Kaur, S., Dhugga, K.S., Gill, K., Singh, J. (2016) Novel Structural and Functional Motifs in cellulose synthase (CesA) Genes of Bread Wheat (*Triticum aestivum*, L.). *PLoS One* 11: e0147046.
- Kawano, T. (2003) Roles of the reactive oxygen species-generating peroxidase reactions in plant defense and growth induction. *Plant Cell Rep* 21: 829–837.
- Keegstra, K., Raikhel, N. (2001) Plant glycosyltransferases. *Curr Opin Plant Biol* 4: 219–224.
- Kiihl, R.A.S., Garcia, A. (1989) The use of long juvenile trait in breeding soybean cultivars. *In: World Soybean Research Conference*.
- Kim, J.Y., Kim, W.Y., Kwak, K.J., Oh, S.H., Han, Y.S., Kang, H. (2010) Glycine-rich RNA-binding proteins are functionally conserved in *Arabidopsis thaliana* and *Oryza sativa* during cold adaptation process. *J Exp Bot* 61: 2317–2325.
- Kour, A., Boone, A.M., Vodkin, L.O. (2014) RNA-Seq Profiling of a Defective Seed Coat Mutation in *Glycine max* Reveals Differential Expression of Proline- Rich and Other Cell Wall Protein Transcripts. doi: 10.1371/journal.pone.0096342
- Krzyszowska, M. (2011) The cell wall in plant cell response to trace metals: polysaccharide remodeling and its role in defense strategy. *Acta Physiol Plant* 33: 35–51.
- Kumar, S., Stecher, G., Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 33: 1870–1874.
- Lairson, L.L., Henrissat, B., Davies, G.J., Withers, S.G. (2008) Glycosyltransferases: Structures, Functions, and Mechanisms. *Annu Rev Biochem* 77: 521–555.
- Lambirth, K.C., Whaley, A.M., Blakley, I.C., Schlueter, J.A., Bost, K.L., Loraine, A.E., Piller, K.J. (2015) A Comparison of transgenic and wild type soybean seeds :

- analysis of transcriptome profiles using RNA-Seq. *BMC Biotechnol* 1–17.
- Lashbrook, C.C., Cai, S. (2008) Cell wall remodeling in Arabidopsis stamen abscission zones. *Plant Signal Behav* 3: 733–736.
- Le, D.T., Nishiyama, R., Watanabe, Y., Tanaka, M., Seki, M., Ham, L.H., Yamaguchi-Shinozaki, K., Shinozaki, K., Tran, L.S.P. (2012) Differential Gene Expression in Soybean Leaf Tissues at Late Developmental Stages under Drought Stress Revealed by Genome-Wide Transcriptome Analysis. *PLoS One*. doi: 10.1371/journal.pone.0049522
- Leroux, O., Sørensen, I., Marcus, S.E., Viane, R.L.L., Willats, W.G.T., Knox, J.P. (2015) Antibody-based screening of cell wall matrix glycans in ferns reveals taxon, tissue and cell-type specific distribution patterns. *BMC Plant Biol*. doi: 10.1186/s12870-014-0362-8
- Li, S., Bashline, L., Lei, L., Gu, Y. (2014) Cellulose Synthesis and Its Regulation. *Arab B* 12: e0169.
- Li, S., Ge, F.R., Xu, M., Zhao, X.Y., Huang, G.Q., Zhou, L.Z., Wang, J.G., Kombrink, A., McCormick, S., Zhang, X.S., Zhang, Y. (2013) Arabidopsis COBRA-LIKE 10, a GPI-anchored protein, mediates directional growth of pollen tubes. *Plant J* 74: 486–497.
- Li, Y., Qian, Q., Zhou, Y., Yan, M., Sun, L., Zhang, M., Fu, Z., Wang, Y., Han, B., Pang, X., Chen, M., Li, J. (2003) *BRITTLE CULM1*, which encodes a COBRA-like protein, affects the mechanical properties of rice plants. *Plant Cell* 15: 2020–2031.
- Liepman, A.H., Cavalier, D.M. (2012) The CELLULOSE SYNTHASE-LIKE A and CELLULOSE SYNTHASE-LIKE C families: recent advances and future perspectives. *Front Plant Sci* 3: 109.
- Lin, S., Huang, L., Yu, X., Xiong, X., Yue, X., Liu, T., Liang, Y., Lv, M., Cao, J. (2017) Characterization of BcMF23a and BcMF23b, two putative pectin methylesterase genes related to pollen development in *Brassica campestris* ssp. *chinensis*. *Mol Biol Rep* 44: 139–148.

- Lindstrom, J.T., Vodkin, L.O. (1991) A soybean cell wall protein is affected by seed color genotype. *Plant Cell* 3: 561–571.
- Liu, L., Shang-Guan, K., Zhang, B., Liu, X., Yan, M., Zhang, L., Shi, Y., Zhang, M., Qian, Q., Li, J., Zhou, Y. (2013) Brittle Culm1, a COBRA-Like Protein, Functions in Cellulose Assembly through Binding Cellulose Microfibrils. *PLoS Genet.* doi: 10.1371/journal.pgen.1003704
- Lowe, R., Shirley, N., Bleackley, M., Dolan, S., Shafee, T. (2017) Transcriptomics technologies. *PLoS Comput Biol* 13: 1–23.
- Majewska-sawka, A., Nothnagel, E.A. (2000) Update on Extracellular Matrix The Multiple Roles of Arabinogalactan Proteins in Plant Development 1 COMPLEX STRUCTURES AMENABLE TO. 122: 3–9.
- Mangeon, A., Junqueira, R.M., Sachetto-Martins, G. (2010) Functional diversity of the plant glycine-rich proteins superfamily. *Plant Signal Behav* 5: 99–104.
- Manners, J.M., Casu, R.E. (2011) Transcriptome Analysis and Functional Genomics of Sugarcane. *Trop Plant Biol* 4: 9–21.
- Marcolino-Gomes, J., Rodrigues, F.A., Fuganti-Pagliarini, R., Nakayama, T.J., Reis, R.R., Farias, J.R.B., Harmon, F.G., Molinari, H.B.C., Molinari, M.D.C., Nepomuceno, A. (2015) Transcriptome-wide identification of reference genes for expression analysis of soybean responses to drought stress along the day. *PLoS One* 10: 1–16.
- Marcus, A., Greedberg, J., Averyhart-Fullard, V. (1991) Repetitive proline-rich proteins in the extracellular matrix of the plant cell. *Physiol Plant* 81: 273–279.
- McFarlane, H.E., Döring, A., Persson, S. (2014) The Cell Biology of Cellulose Synthesis. *Annu Rev Plant Biol* 65: 69–94.
- Messina, M.J. (1999) Legumes and soybeans: Overview of their nutritional profiles and health effects.
- Miedes, E., Vanholme, R., Boerjan, W., Molina, A. (2014) The role of the secondary cell wall in plant resistance to pathogens. *Front Plant Sci* 5: 358.



- Miyasaka, S., Medina, J.C. (1981) A soja no Brasil.
- Mohamed-Yasseen, Y., Barringer, S.A., Splittstoesser, W.E., Costanza, S. (1994) The role of seed coats in seed viability. *Bot Rev* 60: 426–439.
- Mohnen, D. (2008) Pectin structure and biosynthesis. *Curr Opin Plant Biol* 11: 266–277.
- Mousavi, A., Hotta, Y. (2005) Glycine-rich proteins. *Appl Biochem Biotechnol* 120: 169–174.
- Nagalakshmi, U., Waern, K., Snyder, M. (2010) RNA-Seq: A Method for Comprehensive Transcriptome Analysis. 1–13.
- Nawaz, M.A., Rehman, H.M., Baloch, F.S., Ijaz, B., Ali, M.A., Khan, I.A., Lee, J.D., Chung, G., Yang, S.H. (2017) Genome and transcriptome-wide analyses of cellulose synthase gene superfamily in soybean. *J Plant Physiol* 215: 163–175.
- NCSOY. (2016) History of Soybeans - North Carolina Soybeans. <http://ncsoy.org/media-resources/history-of-soybeans/>. Accessed 8 Jun 2017
- Nicol, F., Höfte, H. (1998) Plant cell expansion: scaling the wall. *Curr Opin Plant Biol* 1: 12–17.
- Niu, E., Shang, X., Cheng, C., Bao, J., Zeng, Y., Cai, C., Du, X., Guo, W. (2015) Comprehensive analysis of the COBRA-like (COBL) gene family in *Gossypium* identifies two COBLs potentially associated with fiber quality. *PLoS One* 10: 1–21.
- Nonogaki, H., Bassel, G.W., Bewley, J.D. (2010) Germination-still a mystery. *Plant Sci* 179: 574–581.
- OECD-FAO. (2015) Agricultural Outlook 2015-2024: Special Feature – Brazil:Prospects and Challenges.
- Paredez, A.R. (2006) Visualization of Cellulose Synthase Demonstrates Functional Association with Microtubules. *Science (80- )* 312: 1491–1495.
- Park, Y.B., Cosgrove, D.J. (2015) Xyloglucan and its interactions with other

- components of the growing cell wall. *Plant Cell Physiol* 56: 180–194.
- Passardi, F., Penel, C., Dunand, C. (2004) Performing the paradoxical: How plant peroxidases modify the cell wall. *Trends Plant Sci* 9: 534–540.
- Pauly, M., Keegstra, K. (2016) Biosynthesis of the Plant Cell Wall Matrix Polysaccharide. doi: 10.1146/annurev-arplant-043015-112222
- Peña, M.J., Darvill, A.G., Eberhard, S., York, W.S., O'Neill, M.A. (2008) Moss and liverwort xyloglucans contain galacturonic acid and are structurally distinct from the xyloglucans synthesized by hornworts and vascular plants\*. *Glycobiology* 18: 891–904.
- Persson, S., Wei, H., Milne, J., Page, G.P., Somerville, C.R. (2005) Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc Natl Acad Sci U S A* 102: 8633–8.
- Rapaport, F., Khanin, R., Liang, Y., Pirun, M., Krek, A., Zumbo, P., Mason, C.E., Socci, N.D., Betel, D. (2013) Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. *Genome Biol* 14: R95.
- Ray, D.K., Gerber, J.S., MacDonald, G.K., West, P.C. (2015) Climate variation explains a third of global crop yield variability. *Nat Commun* 6: 5989.
- Reid, J.S.G. (1985) Cell Wall Storage Carbohydrates in Seeds—Biochemistry of the Seed “Gums” and “Hemicelluloses.” *In*: Callow, J.A., Woolhouse, H.W.B.T.-A. in B.R. (ed). Academic Press, p. 125–155
- Reiter, W.D. (2008) Biochemical genetics of nucleotide sugar interconversion reactions. *Curr Opin Plant Biol* 11: 236–243.
- Remm, M., Storm, C.E. V., Sonnhammer, E.L.L. (2001) Automatic clustering of orthologs and in-paralogs from pairwise species comparisons<sup>11</sup> Edited by F. Cohen. *J Mol Biol* 314: 1041–1052.
- Rennie, E.A., Scheller, H.V. (2014) Xylan biosynthesis. *Curr Opin Biotechnol* 26: 100–107.
- Ritchie, S., Hanway, J., Thompson, H.. (1985) How a soybean plant develops. *Iowa*

*State Univ Sci Technol Coop Ext Serv 20.*

- Roberts, A.W., Bushoven, J.T. (2007) The cellulose synthase (CESA) gene superfamily of the moss *Physcomitrella patens*. *Plant Mol Biol* 63: 207–219.
- Rodrigues, F.A., Fuganti-Pagliarini, R., Marcolino-Gomes, J., Nakayama, T.J., Molinari, H.B.C., Lobo, F.P., Harmon, F.G., Nepomuceno, A.L. (2015) Daytime soybean transcriptome fluctuations during water deficit stress. *BMC Genomics* 16: 505.
- Rosén, M.L., Edman, M., Sjöström, M., Wieslander, Å. (2004) Recognition of fold and sugar linkage for glycosyltransferases by multivariate sequence analysis. *J Biol Chem* 279: 38683–38692.
- Roudier, F. (2002) The COBRA Family of Putative GPI-Anchored Proteins in Arabidopsis. A New Fellowship in Expansion. *Plant Physiol* 130: 538–548.
- Sato, K., Suzuki, R., Nishikubo, N., Takenouchi, S., Ito, S., Nakano, Y., Nakaba, S., Sano, Y., Funada, R., Kajita, S., Kitano, H., Katayama, Y. (2010) Isolation of a novel cell wall architecture mutant of rice with defective Arabidopsis COBL4 ortholog BC1 required for regulated deposition of secondary cell wall components. 257–270.
- Scheller, H.V., Ulvskov, P. (2010) Hemicelluloses. *Annu Rev Plant Biol* 61: 263–289.
- Schindelman, G., Morikami, A., Jung, J., Baskin, T.I., Carpita, N.C., Derbyshire, P., Mccann, M.C., Benfey, P.N. (2001) COBRA encodes a putative GPI-anchored protein, which is polarly localized and necessary for oriented cell expansion in Arabidopsis. 1115–1127.
- Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., Cheng, J., Xu, D., Hellsten, U., May, G.D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S., Goodstein, D., Barry, K., Futrell-griggs, M., Abernathy, B., Du, J., Tian, Z., Zhu, L., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X., Shinozaki, K., Nguyen, H.T., Wing, R.A., Cregan,

- P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R.C., Jackson, S.A. (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463: 178–183.
- Schultink, A., Liu, L., Zhu, L., Pauly, M. (2014) Structural Diversity and Function of Xyloglucan Sidechain Substituents. *Plants* 3: 526–542.
- Schultz, C.J. (2000) The Classical Arabinogalactan Protein Gene Family of Arabidopsis. *Plant Cell Online* 12: 1751–1768.
- Sechet, J., Frey, A., Cuzzi, D., Berger, A., Perreau, F., Cueff, G., Charif, D., Rajjou, L., Mouille, G., North, H.M., Marion-Poll, A. (2016) Xyloglucan metabolism differentially impacts the cell wall characteristics of the endosperm and embryo during Arabidopsis seed germination. *Plant Physiol* 170: pp.01312.2015.
- Seifert, G.J. (2004) Nucleotide sugar interconversions and cell wall biosynthesis: how to bring the inside to the outside. *Curr Opin Plant Biol* 7: 277–284.
- Severin, A.J., Woody, J.L., Bolon, Y., Joseph, B., Diers, B.W., Farmer, A.D., Muehlbauer, G.J., Nelson, R.T., Grant, D., Specht, J.E., Graham, M.A., Cannon, S.B., May, G.D., Vance, C.P., Shoemaker, R.C. (2010) RNA-Seq Atlas of Glycine max : A guide to the soybean transcriptome.
- Showalter, A.M. (2001) Arabinogalactan-proteins: structure, expression and function. *Cell Mol Life Sci C* 58: 1399–1417.
- Showalter, A.M. (1993) Structure and function of plant cell wall proteins. *Plant Cell* 5: 9–23.
- Silva, G.B., Ionashiro, M., Carrara, T.B., Crivellari, A.C., Tiné, M.A.S., Prado, J., Carpita, N.C., Buckeridge, M.S. (2011) Cell wall polysaccharides from fern leaves: evidence for a mannan-rich Type III cell wall in *Adiantum raddianum*. *Phytochem.* doi: 10.1016/j.phytochem.2011.08.020
- Sterling, J.D., Atmodjo, M.A., Inwood, S.E., Kumar Kolli, V.S., Quigley, H.F., Hahn, M.G., Mohnen, D. (2006) Functional identification of an Arabidopsis pectin biosynthetic homogalacturonan galacturonosyltransferase. *Proc Natl Acad Sci U S A* 103: 5236–5241.

- Sticklen, M.B. (2008) Plant genetic engineering for biofuel production: Towards affordable cellulosic ethanol. *Nat Rev Genet* 9: 433–443.
- Suzuki, S., Li, L., Sun, Y.-H., Chiang, V.L. (2006) The Cellulose Synthase Gene Superfamily and Biochemical Functions of Xylem-Specific Cellulose Synthase-Like Genes in *Populus trichocarpa*. *Plant Physiol* 142: 1233–1245.
- Tanaka, N., Uno, H., Okuda, S., Gunji, S., Ferjani, A., Aoyama, T., Maeshima, M. (2017) SRPP, a cell-wall protein is involved in development and protection of seeds and root hairs in *Arabidopsis thaliana*. *Plant Cell Physiol* 13: pcx008.
- Tenhaken, R. (2015) Cell wall remodeling under abiotic stress .
- The Angiosperm Phylogeny Group III. (2009) An update of the Angiosperm Phylogeny Group Classification for the orders and families of flowering plants: APG III. *Bot J Linn Soc* 161: 105–121.
- Thumma, B.R., Matheson, B.A., Zhang, D., Meeske, C., Meder, R., Downes, G.M., Southerton, S.G. (2009) Identification of a cis-acting regulatory polymorphism in a eucalypt COBRA-like gene affecting cellulose content. *Genetics* 183: 1153–1164.
- Van Hengel, A.J., Roberts, K. (2003) AtAGP30, an arabinogalactan-protein in the cell walls of the primary root, plays a role in root regeneration and seed germination. *Plant J* 36: 256–270.
- Van Hengel, A.J., Van Kammen, A., De Vries, S.C. (2002) A relationship between seed development, Arabinogalactan-proteins (AGPS) and the AGP mediated promotion of somatic embryogenesis. *Physiol Plant* 114: 637–644.
- Ververis, C., Georghiou, K., Christodoulakis, N., Santas, P., Santas, R. (2004) Fiber dimensions, lignin and cellulose content of various plant materials and their suitability for paper production. *Ind Crops Prod* 19: 245–254.
- Wallace, S., Williams, J.H. (2017) Evolutionary origins of pectin methylesterase genes associated with novel aspects of angiosperm pollen tube walls. *Biochem Biophys Res Commun* 487: 509–516.

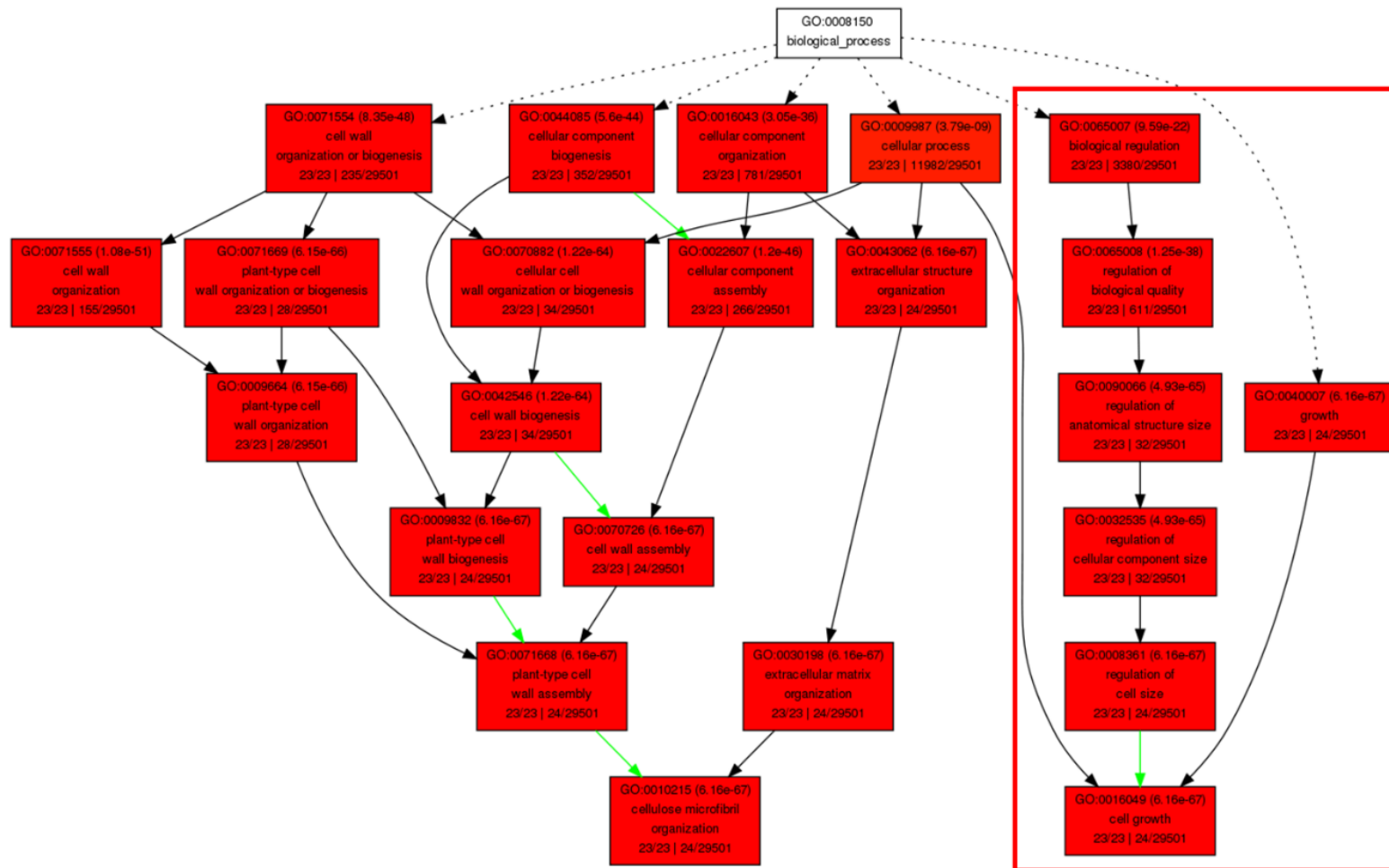
- Wang, L., Cao, C., Ma, Q., Zeng, Q., Wang, H., Cheng, Z., Zhu, G., Qi, J., Ma, H., Nian, H., Wang, Y. (2014) RNA-seq analyses of multiple meristems of soybean: novel and alternative transcripts, evolutionary and functional implications. *BMC Plant Biol* 14: 169.
- Wang, Z., Gerstein, M., Snyder, M. (2010) RNA-Seq: a revolutionary tool for transcriptomics. *10*: 57–63.
- Weisshaar, B., Jenkinst, G.I. (1998) Phenylpropanoid biosynthesis and its regulation. *Curr Opin Plant Biol* 1: 251–257.
- Wilder, B., Albersheim, P. (1973) The Structure of Plant Cell Walls. *Plant Physiol* 51: 889–893.
- Willats, W.G.T., Orfila, C., Limberg, G., Buchholt, H.C., Alebeek, G.-. J.W.M., Voragen, A.G.J., Marcus, S.E., Christensen, T.M.I.E., Mikkelsen, J.D., Murray, B.S., Knox, J.P. (2001) Modulation of the degree and pattern of methyl-esterification of pectic homogalacturonan in plant cell walls: implications for pectin methyl esterase action, matrix properties and cell adhesion. *J Biol Chem*. doi: 10.1074/jbc.M011242200
- Wu, F., Mueller, L.A., Crouzillat, D., Pétiard, V., Tanksley, S.D. (2006) Combining bioinformatics and phylogenetics to identify large sets of single-copy orthologous genes (COSII) for comparative, evolutionary and systematic studies: A test case in the euasterid plant clade. *Genetics* 174: 1407–1420.
- Wyatt, R.E., Nagao, R.T., Key, J.L. (1992) Patterns of soybean proline-rich protein gene expression. *Plant Cell* 4: 99–110.
- Ye, X., Kang, B.G., Osburn, L.D., Cheng, Z.M. (2009) The COBRA gene family in *Populus* and gene expression in vegetative organs and in response to hormones and environmental stresses. *Plant Growth Regul* 58: 211–223.
- Yong, W., Link, B., O'Malley, R., Tewari, J., Hunter, C.T., Lu, C.A., Li, X., Bleecker, A.B., Koch, K.E., McCann, M.C., McCarty, D.R., Patterson, S.E., Reiter, W.D., Staiger, C., Thomas, S.R., Vermerris, W., Carpita, N.C. (2005) Genomics of plant cell wall biogenesis. *Planta* 221: 747–751.

- Zamil, M.S., Yi, H., Puri, V.M. (2014) Mechanical characterization of outer epidermal middle lamella of onion under tensile loading. *Am J Bot* 101: 778–787.
- Zandleven, J., Sørensen, S.O., Harholt, J., Beldman, G., Schols, H.A., Scheller, H. V., Voragen, A.J. (2007) Xylogalacturonan exists in cell walls from various tissues of *Arabidopsis thaliana*. *Phytochemistry* 68: 1219–1226.
- Zega, A., D'Ovidio, R. (2016) Genome-wide characterization of pectin methyl esterase genes reveals members differentially expressed in tolerant and susceptible wheats in response to *Fusarium graminearum*. *Plant Physiol Biochem* 108: 1–11.
- Zhang, H., Song, B. (2017) Genomics Data RNA-seq data comparisons of wild soybean genotypes in response to soybean cyst nematode ( *Heterodera glycines* ). *Genomics Data* 14: 36–39.
- Zhang, Z., Li, J., Zhao, X.Q., Wang, J., Wong, G.K.S., Yu, J. (2006) KaKs\_Calculator: Calculating Ka and Ks Through Model Selection and Model Averaging. *Genomics, Proteomics Bioinforma* 4: 259–263.

## APÊNDICE



**Figura S1.** Distribuição de categorias de processos funcionais e biológicos dos genes COBRA em soja. Caixa vermelha representa os processos relacionados a regulação da parede celular.



**Tabela S1.** Todos os 2143 genes relacionados a parede celular identificados no genoma de soja de acordo com o seu homólogo em Arabidopsis e a sua via de parede celular correspondente.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma10g28610	At2g47180	GT-family-8
Glyma03g38080	At2g47180	GT-family-8
Glyma11g01940	At4g10960	Nucleotide-sugar
Glyma01g43540	At4g10960	Nucleotide-sugar
Glyma01g00730	At4g26220	Pathways-phenyl
Glyma05g37085	At2g46330	Arabinogalactan-proteins
Glyma20g28440	At3g22440	Hydroxyproline-rich
Glyma19g41550	At1g56600	GT-family-8
Glyma10g39330	At3g22440	Hydroxyproline-rich
Glyma17g08000	At5g17420	Cellulose-synthase
Glyma07g39950	At4g17180	Glycoside-hidrolase-family-17
Glyma13g01150	At5g57560	Xyloglucan
Glyma11g08240	At4g36010	Receptor-like
Glyma05g38120	At1g12780	Nucleotide-sugar
Glyma10g38710	At5g62220	GT-family-47
Glyma01g37040	At4g36010	Receptor-like
Glyma09g27870	At5g62200	Unknow-GPI-anchored
Glyma17g02430	At5g18900	Prolyl-4-hydroxylases
Glyma10g07146	At1g63180	Nucleotide-sugar
Glyma11g10080	At4g16260	Glycoside-hidrolase-family-17
Glyma03g02250	At3g61130	GT-family-8
Glyma11g19550	At1g08200	Nucleotide-sugar
Glyma06g45800	At4g28300	Hydroxyproline-rich
Glyma09g02160	At5g49720	Glycoside-hidrolase-family-9
Glyma06g20030	At5g62620	GT-family-31
Glyma06g46230	At1g05170	GT-family-31
Glyma19g43400	At5g28840	Nucleotide-sugar
Glyma19g43410	At5g28840	Nucleotide-sugar
Glyma12g12740	At4g28300	Hydroxyproline-rich
Glyma17g37990	At1g20190	Expansins
Glyma13g38780	At3g44100	Unknow-GPI-anchored
Glyma18g00505	At3g23820	Nucleotide-sugar
Glyma14g39840	At1g20510	Pathways-phenyl
Glyma08g14900	At4g36220	Pathways-phenyl
Glyma07g38330	At5g18900	Prolyl-4-hydroxylases
Glyma08g26520	At5g15490	Nucleotide-sugar
Glyma03g39860	At1g54860	Unknow-GPI-anchored
Glyma13g03810	At3g43720	LPTL
Glyma12g04020	At5g15320	Glycosylphosphatidylinositol
Glyma04g34620	At5g62620	GT-family-31
Glyma18g49740	At3g29090	Pectin-methyl-esterases

**Tabela S1. Cont.**

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma10g36680	At1g71695	Peroxidasas
Glyma19g42420	At1g54860	Unknow-GPI-anchored
Glyma11g08530	At2g17760	Proteases
Glyma19g40680	At2g47180	GT-family-8
Glyma20g36740	At5g28840	Nucleotide-sugar
Glyma05g35360	At3g22440	Hydroxyproline-rich
Glyma20g22700	At2g47180	GT-family-8
Glyma02g14090	At1g68850	Peroxidasas
Glyma17g07740	At4g30440	Nucleotide-sugar
Glyma09g35651	At5g15320	Glycosylphosphatidylinositol
Glyma18g14640	At4g23820	Polygalacturonases
Glyma11g00610	At3g06035	Unknow-GPI-anchored
Glyma07g40090	At5g49720	Glycoside-hidrolase-family-9
Glyma12g08930	At1g08200	Nucleotide-sugar
Glyma09g36950	At3g29090	Pectin-methyl-esterases
Glyma08g11510	At3g23820	Nucleotide-sugar
Glyma05g26840	At4g39350	Cellulose-synthase
Glyma13g39261	At2g27500	Glycoside-hidrolase-family-17
Glyma13g19670	At3g55140	Pectate-pectin-Lyases
Glyma03g14210	At5g01210	Pathways-phenyl
Glyma12g10520	At1g05170	GT-family-31
Glyma13g22290	At5g58170	Glycosylphosphatidylinositol
Glyma08g15840	At2g33160	Polygalacturonases
Glyma04g36250	At4g20360	Endocytosis
Glyma10g31640	At1g64640	Early-nodulin-like
Glyma02g37020	At4g30440	Nucleotide-sugar
Glyma08g47670	At5g13000	Callose-synthase
Glyma01g01350	At4g19010	Pathways-phenyl
Glyma08g11840	At1g55260	LPTL
Glyma12g10160	At3g44100	Unknow-GPI-anchored
Glyma06g09760	At1g09910	Rhamnogalacturonan
Glyma06g09770	At1g09890	Rhamnogalacturonan
Glyma08g04360	At3g22440	Hydroxyproline-rich
Glyma02g15990	At1g06780	GT-family-8
Glyma09g38480	At1g05840	Proteases
Glyma03g40720	At5g28840	Nucleotide-sugar
Glyma10g29150	At4g02330	Pectin-methyl-esterases
Glyma01g04660	At5g62620	GT-family-31
Glyma02g36720	At5g17420	Cellulose-synthase
Glyma11g10091	At4g16260	Glycoside-hidrolase-family-17
Glyma09g05010	At4g02500	GT-family-34
Glyma10g03770	At1g06780	GT-family-8

**Tabela S1. Cont.**

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma10g36230	At1g34270	GT-family-47
Glyma17g08110	At4g31840	Early-nodulin-like
Glyma18g50000	At5g15490	Nucleotide-sugar
Glyma11g08020	At5g66280	Nucleotide-sugar
Glyma06g14270	At1g34510	Peroxidases
Glyma19g34420	At1g06780	GT-family-8
Glyma02g04280	At1g60780	Endocytosis
Glyma10g44480	At2g43080	Prolyl-4-hydroxylases
Glyma15g13120	At2g02400	Pathways-phenyl
Glyma08g22220	At4g14746	Unknow-GPI-anchored
Glyma07g02990	At2g33590	Pathways-phenyl
Glyma05g02670	At4g20360	Endocytosis
Glyma08g11670	At2g32810	Beta-galactosidase
Glyma06g18640	At4g20360	Endocytosis
Glyma03g40450	At5g17540	Pathways-phenyl
Glyma01g09650	At1g68850	Peroxidases
Glyma10g13210	At4g20360	Endocytosis
Glyma09g03490	At1g30620	Nucleotide-sugar
Glyma10g40870	At3g19450	Pathways-phenyl
Glyma04g04080	At4g33330	GT-family-8
Glyma20g29030	At5g62220	GT-family-47
Glyma08g38740	At4g40090	Arabinogalactan-proteins
Glyma01g27810	At5g01210	Pathways-phenyl
Glyma17g10950	At2g03090	Expansins
Glyma18g06063	At2g40610	Expansins
Glyma02g40290	At2g30490	Pathways-phenyl
Glyma05g04630	At4g16600	GT-family-8
Glyma10g01960	At3g62660	GT-family-8
Glyma17g02880	At4g22580	GT-family-47
Glyma10g05260	At3g53520	Nucleotide-sugar
Glyma17g00710	At5g49720	Glycoside-hidrolase-family-9
Glyma04g07220	At4g32410	Cellulose-synthase
Glyma07g34570	At2g35100	GT-family-47
Glyma02g44580	At1g08500	Early-nodulin-like
Glyma10g44351	At2g43160	Endocytosis
Glyma01g36740	At5g66060	Prolyl-4-hydroxylases
Glyma17g35000	At5g11710	Endocytosis
Glyma08g44210	At5g55730	Arabinogalactan-proteins
Glyma06g12280	At1g21660	Endocytosis
Glyma20g31360	At1g34270	GT-family-47
Glyma04g42501	At1g21660	Endocytosis
Glyma10g35230	At3g54840	Endocytosis

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma03g31590	At1g06780	GT-family-8
Glyma01g03330	At1g60780	Endocytosis
Glyma15g13080	At5g49720	Glycoside-hidrolase-family-9
Glyma10g31550	At1g32860	Glycoside-hidrolase-family-17
Glyma02g02900	At5g62620	GT-family-31
Glyma14g38580	At2g30490	Pathways-phenyl
Glyma01g05860	At1g26810	GT-family-31
Glyma18g08530	At5g55730	Arabinogalactan-proteins
Glyma16g17190	At4g19420	Pectin-acetylerases
Glyma14g40430	At1g47830	Endocytosis
Glyma05g28510	At3g23820	Nucleotide-sugar
Glyma05g03830	At4g00110	Nucleotide-sugar
Glyma13g04590	At3g28200	Peroxidases
Glyma13g01140	At5g57560	Xyloglucan
Glyma14g00790	At3g07680	Vesicle-trafficking
Glyma10g38690	At5g62200	Unknow-GPI-anchored
Glyma06g02420	At1g20190	Expansins
Glyma01g36770	At2g17760	Proteases
Glyma11g20020	At4g05160	Pathways-phenyl
Glyma19g29020	At4g38040	GT-family-47
Glyma13g17170	At1g19170	Polygalacturonases
Glyma08g39080	At1g60780	Endocytosis
Glyma11g11550	At5g16890	GT-family-47
Glyma09g36740	At1g02000	Nucleotide-sugar
Glyma02g37700	At4g24550	Endocytosis
Glyma07g03420	At3g15800	Glycoside-hidrolase-family-17
Glyma14g08200	At4g34480	Glycoside-hidrolase-family-17
Glyma14g10510	At5g11710	Endocytosis
Glyma04g31770	At1g18580	GT-family-8
Glyma11g37610	At1g67750	Pectate-pectin-Lyases
Glyma11g15010	At2g28760	Nucleotide-sugar
Glyma11g36160	At5g10080	Proteases
Glyma03g28841	At3g57270	Glycoside-hidrolase-family-17
Glyma05g15220	At5g63380	Pathways-phenyl
Glyma06g17140	At5g61840	GT-family-47
Glyma06g03560	At5g42740	Glycosylphosphatidylinositol
Glyma01g37020	At4g38660	Receptor-like
Glyma12g02550	At3g45970	Expansins
Glyma01g45110	At1g11580	Pectin-methyl-esterases
Glyma08g02480	At2g46330	Arabinogalactan-proteins
Glyma14g40140	At1g20190	Expansins
Glyma12g31980	At1g05170	GT-family-31
Glyma11g00250	At4g34150	Hydroxyproline-rich

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma16g23120	At4g35880	Proteases
Glyma16g01640	At1g53840	Pectin-methyl-esterases
Glyma17g14170	At1g63000	Nucleotide-sugar
Glyma02g01880	At3g62660	GT-family-8
Glyma06g20840	At1g74680	GT-family-47
Glyma15g15820	At3g62720	GT-family-34
Glyma03g38910	At1g56600	GT-family-8
Glyma18g47070	At4g00300	GT-family-31
Glyma18g51920	At1g05840	Proteases
Glyma02g12870	At3g29360	Nucleotide-sugar
Glyma04g36725	At2g36850	Callose-synthase
Glyma13g10800	At4g35410	Endocytosis
Glyma13g01110	At4g25810	Xyloglucan
Glyma17g27550	At5g25820	GT-family-47
Glyma04g02380	At1g20190	Expansins
Glyma02g40790	At3g03220	Expansins
Glyma03g10046	At3g23730	Xyloglucan
Glyma12g20960	At3g53520	Nucleotide-sugar
Glyma20g32320	At3g54840	Endocytosis
Glyma04g43340	At1g77810	GT-family-31
Glyma08g29040	At1g05840	Proteases
Glyma02g02170	At3g62830	Nucleotide-sugar
Glyma14g06340	At3g08530	Endocytosis
Glyma08g41530	At4g23820	Polygalacturonases
Glyma08g07490	At5g64310	Arabinogalactan-proteins
Glyma12g34280	At5g54690	GT-family-8
Glyma11g01010	At3g22600	LPTL
Glyma02g42560	At3g08530	Endocytosis
Glyma20g29040	At5g62200	Unknow-GPI-anchored
Glyma01g00860	At2g19790	Endocytosis
Glyma12g26693	At1g12780	Nucleotide-sugar
Glyma08g02690	At4g23920	Nucleotide-sugar
Glyma08g27560	At5g60920	COBRA
Glyma11g37010	At1g27950	LPTL
Glyma19g05060	At3g02350	GT-family-8
Glyma04g03490	At5g42740	Glycosylphosphatidylinositol
Glyma06g20970	At2g03090	Expansins
Glyma17g14340	At4g00110	Nucleotide-sugar
Glyma14g39120	At3g03220	Expansins
Glyma17g37730	At1g47830	Endocytosis
Glyma08g47740	At2g43160	Endocytosis
Glyma13g04940	At3g28480	Prolyl-4-hydroxylases
Glyma04g12465	At5g56170	Glycosylphosphatidylinositol

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma02g48115	At1g21630	Endocytosis
Glyma11g17920	At3g44100	Unknow-GPI-anchored
Glyma12g06990	At3g46440	Nucleotide-sugar
Glyma06g33880	At1g53290	GT-family-31
Glyma07g39020	At4g21960	Peroxidases
Glyma06g11710	At1g09580	Vesicle-trafficking
Glyma11g36600	At3g23820	Nucleotide-sugar
Glyma10g44150	At3g07160	Callose-synthase
Glyma07g39070	At1g05170	GT-family-31
Glyma05g30410	At1g30620	Nucleotide-sugar
Glyma06g22730	At1g18580	GT-family-8
Glyma17g17810	At4g33910	Prolyl-4-hydroxylases
Glyma18g53750	At2g43160	Endocytosis
Glyma20g03930	At1g26690	Vesicle-trafficking
Glyma03g21730	At3g07330	Cellulose-synthase-like
Glyma13g10460	At1g64640	Early-nodulin-like
Glyma07g08550	At2g36870	Xyloglucan
Glyma17g13361	At4g20360	Endocytosis
Glyma02g47820	At3g07680	Vesicle-trafficking
Glyma07g40150	At1g12780	Nucleotide-sugar
Glyma20g09170	At1g53290	GT-family-31
Glyma03g40980	At5g18480	GT-family-8
Glyma06g07320	At4g32410	Cellulose-synthase
Glyma01g04460	At1g70090	GT-family-8
Glyma09g39246	At4g00300	GT-family-31
Glyma14g13080	At3g53520	Nucleotide-sugar
Glyma01g37270	At3g51160	Nucleotide-sugar
Glyma09g40610	At2g38650	GT-family-8
Glyma04g34600	At1g74790	HIPL
Glyma05g36850	At4g10960	Nucleotide-sugar
Glyma05g32100	At5g64740	Cellulose-synthase
Glyma09g41650	At1g17890	Nucleotide-sugar
Glyma02g07610	At5g13870	Xyloglucan
Glyma09g40420	At3g60900	Arabinogalactan-proteins
Glyma11g05020	At2g22900	GT-family-34
Glyma07g15160	At2g19790	Endocytosis
Glyma19g03460	At3g01040	GT-family-8
Glyma06g24870	At1g08750	Glycosylphosphatidylinositol
Glyma11g04460	At4g18010	Endocytosis
Glyma10g02290	At3g62830	Nucleotide-sugar
Glyma17g32180	At4g32120	GT-family-31
Glyma13g38500	At1g05170	GT-family-31
Glyma17g11060	At5g07720	GT-family-34
Glyma15g43040	At5g05170	Cellulose-synthase

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma17g11060	At5g07720	GT-family-34
Glyma15g43040	At5g05170	Cellulose-synthase
Glyma19g02110	At3g28480	Prolyl-4-hydroxylases
Glyma07g05460	At5g54160	Pathways-phenyl
Glyma07g05470	At5g54160	Pathways-phenyl
Glyma02g12030	At1g26810	GT-family-31
Glyma10g01230	At1g05260	Peroxidases
Glyma02g06350	At3g50860	Endocytosis
Glyma07g40011	At5g47780	GT-family-8
Glyma08g07890	At5g64080	LPTL
Glyma11g02130	At2g46330	Arabinogalactan-proteins
Glyma11g10240	At3g45970	Expansins
Glyma14g03710	At4g23820	Polygalacturonases
Glyma19g01620	At5g40150	Peroxidases
Glyma20g38860	At3g07160	Callose-synthase
Glyma06g18030	At5g63380	Pathways-phenyl
Glyma14g00360	At1g21630	Endocytosis
Glyma07g35490	At1g26690	Vesicle-trafficking
Glyma11g08250	At4g38660	Receptor-like
Glyma08g15380	At5g64740	Cellulose-synthase
Glyma13g30550	At1g28680	Pathways-phenyl
Glyma04g37920	At5g61840	GT-family-47
Glyma05g22270	At4g33910	Prolyl-4-hydroxylases
Glyma05g27950	At1g67410	GT-family-47
Glyma15g25060	At4g19810	Yieldins
Glyma16g09760	At3g07390	Unknow-GPI-anchored
Glyma08g46210	At3g25140	GT-family-8
Glyma03g37150	At3g62720	GT-family-34
Glyma20g02340	At2g35100	GT-family-47
Glyma16g22488	At5g14650	Polygalacturonases
Glyma18g15230	At5g24090	Yieldins
Glyma14g06070	At2g40370	Laccases
Glyma04g08881	At5g25820	GT-family-47
Glyma11g35460	At3g11130	Endocytosis
Glyma14g35990	At4g24550	Endocytosis
Glyma08g13470	At1g09580	Vesicle-trafficking
Glyma04g24410	At1g08750	Glycosylphosphatidylinositol
Glyma08g27050	At5g54160	Pathways-phenyl
Glyma18g45226	At2g38650	GT-family-8
Glyma08g12020	At3g13560	Glycoside-hidrolase-family-17
Glyma09g36130	At2g30700	Unknow-GPI-anchored
Glyma06g15940	At4g33440	Polygalacturonases
Glyma13g02190	At5g10080	Proteases



**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma14g40170	At4g39330	Pathways-phenyl
Glyma20g30910	At1g71695	Peroxidases
Glyma17g11060	At5g07720	GT-family-34
Glyma15g43040	At5g05170	Cellulose-synthase
Glyma19g02110	At3g28480	Prolyl-4-hydroxylases
Glyma07g05460	At5g54160	Pathways-phenyl
Glyma07g05470	At5g54160	Pathways-phenyl
Glyma02g12030	At1g26810	GT-family-31
Glyma10g01230	At1g05260	Peroxidases
Glyma02g06350	At3g50860	Endocytosis
Glyma07g40011	At5g47780	GT-family-8
Glyma08g07890	At5g64080	LPTL
Glyma11g02130	At2g46330	Arabinogalactan-proteins
Glyma11g10240	At3g45970	Expansins
Glyma14g03710	At4g23820	Polygalacturonases
Glyma19g01620	At5g40150	Peroxidases
Glyma20g38860	At3g07160	Callose-synthase
Glyma06g18030	At5g63380	Pathways-phenyl
Glyma14g00360	At1g21630	Endocytosis
Glyma07g35490	At1g26690	Vesicle-trafficking
Glyma11g08250	At4g38660	Receptor-like
Glyma08g15380	At5g64740	Cellulose-synthase
Glyma13g30550	At1g28680	Pathways-phenyl
Glyma04g37920	At5g61840	GT-family-47
Glyma05g22270	At4g33910	Prolyl-4-hydroxylases
Glyma05g27950	At1g67410	GT-family-47
Glyma15g25060	At4g19810	Yieldins
Glyma16g09760	At3g07390	Unknow-GPI-anchored
Glyma08g46210	At3g25140	GT-family-8
Glyma03g37150	At3g62720	GT-family-34
Glyma20g02340	At2g35100	GT-family-47
Glyma16g22488	At5g14650	Polygalacturonases
Glyma18g15230	At5g24090	Yieldins
Glyma14g06070	At2g40370	Laccases
Glyma04g08881	At5g25820	GT-family-47
Glyma11g35460	At3g11130	Endocytosis
Glyma14g35990	At4g24550	Endocytosis
Glyma08g13470	At1g09580	Vesicle-trafficking
Glyma04g24410	At1g08750	Glycosylphosphatidylinositol
Glyma08g27050	At5g54160	Pathways-phenyl
Glyma18g45226	At2g38650	GT-family-8
Glyma08g12020	At3g13560	Glycoside-hidrolase-family-17

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma09g36130	At2g30700	Unknow-GPI-anchored
Glyma06g15940	At4g33440	Polygalacturonases
Glyma13g02190	At5g10080	Proteases
Glyma14g40170	At4g39330	Pathways-phenyl
Glyma20g30910	At1g71695	Peroxidases
Glyma07g05145	At3g14310	Pectin-methyl-esterases
Glyma13g36280	At5g54690	GT-family-8
Glyma01g44940	At5g26330	Stellacyanin-like
Glyma13g19640	At3g53520	Nucleotide-sugar
Glyma06g01500	At4g34480	Glycoside-hidrolase-family-17
Glyma20g37000	At5g18480	GT-family-8
Glyma04g41810	At5g53340	GT-family-31
Glyma12g11556	At4g28100	Unknow-GPI-anchored
Glyma02g47880	At4g12730	Arabinogalactan-proteins
Glyma04g43030	At1g09580	Vesicle-trafficking
Glyma18g00630	At3g23730	Xyloglucan
Glyma09g33330	At4g38040	GT-family-47
Glyma05g00860	At5g07720	GT-family-34
Glyma18g53823	At5g13000	Callose-synthase
Glyma02g03210	At1g59970	Proteases
Glyma02g03230	At1g59970	Proteases
Glyma05g36210	At4g26220	Pathways-phenyl
Glyma14g03110	At3g58790	GT-family-8
Glyma17g01720	At4g21960	Peroxidases
Glyma16g05960	At5g40960	GPI-anchored-peptide
Glyma06g02760	At1g75680	Glycoside-hidrolase-family-9
Glyma02g47790	At2g04780	Arabinogalactan-proteins
Glyma18g04020	At3g11420	GT-family-31
Glyma15g14433	At1g30620	Nucleotide-sugar
Glyma18g47840	At3g02740	Proteases
Glyma05g06090	At4g18010	Endocytosis
Glyma05g24730	At5g64080	LPTL
Glyma04g02740	At1g75680	Glycoside-hidrolase-family-9
Glyma02g03090	At1g70090	GT-family-8
Glyma04g02300	At1g20270	Prolyl-4-hydroxylases
Glyma02g46590	At1g66970	GPDL
Glyma01g06970	At3g29360	Nucleotide-sugar
Glyma02g39630	At5g58490	Pathways-phenyl
Glyma15g08690	At1g28680	Pathways-phenyl
Glyma14g33700	At1g77810	GT-family-31
Glyma05g26420	At2g17120	Receptor-like
Glyma17g00782	At5g47780	GT-family-8

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma17g05550	At1g19170	Polygalacturonases
Glyma09g40300	At1g08500	Early-nodulin-like
Glyma06g11330	At1g77810	GT-family-31
Glyma12g10920	At4g15240	GT-family-31
Glyma18g17380	At1g70170	Proteases
Glyma02g43230	At2g40230	Pathways-phenyl
Glyma15g00600	At1g15950	Pathways-phenyl
Glyma02g07840	At1g64760	Glycoside-hidrolase-family-17
Glyma03g03180	At4g00110	Nucleotide-sugar
Glyma14g14000	At4g32120	GT-family-31
Glyma07g22970	At1g23040	Extensin-related
Glyma10g21710	At3g57630	GT-family-47
Glyma14g02350	At5g55180	Glycoside-hidrolase-family-17
Glyma13g20800	At2g37040	Pathways-phenyl
Glyma07g06760	At5g24090	Yieldins
Glyma12g37060	At2g22420	Peroxidases
Glyma08g02610	At1g64390	Glycoside-hidrolase-family-9
Glyma15g03250	At2g43480	Peroxidases
Glyma11g05470	At4g37450	Arabinogalactan-proteins
Glyma08g18740	At3g62000	Pathways-phenyl
Glyma04g38580	At5g63800	Beta-galactosidase
Glyma16g04680	At1g64760	Glycoside-hidrolase-family-17
Glyma07g09600	At5g53250	Arabinogalactan-proteins
Glyma01g45560	At4g34150	Hydroxyproline-rich
Glyma04g33170	At2g03220	GT-family-37
Glyma06g06450	At2g25060	Early-nodulin-like
Glyma06g13400	At5g53370	Pectin-methyl-esterases
Glyma04g06295	At4g31760	Peroxidases
Glyma08g03670	At5g56590	Glycoside-hidrolase-family-17
Glyma12g01200	At2g30700	Unknow-GPI-anchored
Glyma08g09940	At1g14720	Xyloglucan
Glyma10g30700	At5g18480	GT-family-8
Glyma02g11530	At4g34050	Pathways-phenyl
Glyma08g09331	At2g17120	Receptor-like
Glyma15g38936	At1g32860	Glycoside-hidrolase-family-17
Glyma06g42110	At3g20570	Early-nodulin-like
Glyma06g26610	At5g20230	Stellacyanin-like
Glyma19g43630	At5g18480	GT-family-8
Glyma16g10680	At3g07330	Cellulose-synthase-like

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma20g06200	At4g35410	Endocytosis
Glyma18g07330	At2g29660	Unknow-GPI-anchored
Glyma20g35630	At4g35160	Pathways-phenyl
Glyma20g35640	At4g35160	Pathways-phenyl
Glyma06g18216	At2g36850	Callose-synthase
Glyma13g02420	At1g77810	GT-family-31
Glyma18g44040	At1g17890	Nucleotide-sugar
Glyma09g32716	At5g19670	GT-family-47
Glyma05g01570	At5g62620	GT-family-31
Glyma02g03320	At1g59970	Proteases
Glyma18g45420	At2g45470	Arabinogalactan-proteins
Glyma06g12970	At5g53340	GT-family-31
Glyma13g39770	At4g05160	Pathways-phenyl
Glyma17g17970	At4g37450	Arabinogalactan-proteins
Glyma12g33150	At4g28300	Hydroxyproline-rich
Glyma19g40170	At1g02730	Cellulose-synthase-like
Glyma18g02960	At3g11130	Endocytosis
Glyma08g10920	At1g67410	GT-family-47
Glyma08g10930	At1g67980	Pathways-phenyl
Glyma16g04471	At5g56540	Arabinogalactan-proteins
Glyma17g04940	At3g43270	Pectin-methyl-esterases
Glyma11g19490	At5g22740	Cellulose-synthase-like
Glyma08g15420	At4g39330	Pathways-phenyl
Glyma04g33350	At2g03090	Expansins
Glyma19g40100	At3g16850	Polygalacturonases
Glyma01g33650	At4g00110	Nucleotide-sugar
Glyma12g30490	At1g08200	Nucleotide-sugar
Glyma07g08910	At3g61130	GT-family-8
Glyma10g33300	At4g24000	Cellulose-synthase-like
Glyma04g33600	At1g74680	GT-family-47
Glyma13g40870	At3g46060	Endocytosis
Glyma06g02360	At1g20270	Prolyl-4-hydroxylases
Glyma11g03550	At2g35710	GT-family-8
Glyma18g00926	At1g27950	LPTL
Glyma08g40300	At1g70170	Proteases
Glyma19g03020	At1g26690	Vesicle-trafficking
Glyma06g03280	At4g24550	Endocytosis
Glyma15g06460	At1g72680	Pathways-phenyl
Glyma19g41030	At3g62060	Pectin-acetylsterases

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma11g08560	At5g66060	Prolyl-4-hydroxylases
Glyma03g19690	At5g60020	Laccases
Glyma06g14850	At2g40610	Expansins
Glyma04g38590	At5g63810	Beta-galactosidase
Glyma12g08990	At5g22740	Cellulose-synthase-like
Glyma12g30125	At4g05160	Pathways-phenyl
Glyma05g03640	At1g63000	Nucleotide-sugar
Glyma13g01060	At5g57500	GT-family-31
Glyma18g14150	At5g03760	Cellulose-synthase-like
Glyma08g23120	At2g33590	Pathways-phenyl
Glyma13g06990	At3g02350	GT-family-8
Glyma07g07360	At5g41460	GT-family-31
Glyma12g16340	At3g20570	Early-nodulin-like
Glyma11g26240	At2g40610	Expansins
Glyma15g05770	At5g64080	LPTL
Glyma05g36930	At1g64390	Glycoside-hidrolase-family-9
Glyma07g15737	At3g62060	Pectin-acetylerases
Glyma09g00480	At2g22420	Peroxidases
Glyma02g06640	At3g50760	GT-family-8
Glyma04g05271	At5g11740	Arabinogalactan-proteins
Glyma08g46110	At3g07320	Glycoside-hidrolase-family-17
Glyma12g36570	At5g05170	Cellulose-synthase
Glyma13g42560	At1g72990	Beta-galactosidase
Glyma09g05311	At3g46550	Arabinogalactan-proteins
Glyma15g10150	At1g29980	Unknow-GPI-anchored
Glyma02g42751	At5g05340	Peroxidases
Glyma07g35620	At2g03090	Expansins
Glyma09g01910	At4g17180	Glycoside-hidrolase-family-17
Glyma07g02690	At1g15950	Pathways-phenyl
Glyma05g09200	At4g38270	GT-family-8
Glyma09g07560	At1g70170	Proteases
Glyma16g25400	At3g50860	Endocytosis
Glyma02g41865	At4g38040	GT-family-47
Glyma18g17450	At4g12420	SKU5
Glyma15g10670	At5g18900	Prolyl-4-hydroxylases
Glyma04g12480	At5g56170	Glycosylphosphatidylinositol
Glyma08g26480	At3g01040	GT-family-8
Glyma18g33210	At3g25140	GT-family-8
Glyma08g23310	At1g15950	Pathways-phenyl
Glyma03g34660	At5g03800	GT-family-47
Glyma08g18891	At5g64310	Arabinogalactan-proteins
Glyma10g06600	At2g37040	Pathways-phenyl

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma01g41750	At4g16640	Proteases
Glyma13g23810	At3g55250	Pectate-pectin-Lyases
Glyma20g30450	At5g24090	Yieldins
Glyma05g21820	At4g37450	Arabinogalactan-proteins
Glyma01g44010	At1g33250	GT-family-31
Glyma14g23030	At1g09580	Vesicle-trafficking
Glyma06g16420	At5g63810	Beta-galactosidase
Glyma04g01450	At4g34480	Glycoside-hidrolase-family-17
Glyma09g24900	At5g23940	Pathways-phenyl
Glyma18g49960	At3g01040	GT-family-8
Glyma05g33420	At5g61840	GT-family-47
Glyma14g00830	At2g04780	Arabinogalactan-proteins
Glyma07g10051	At1g73890	LPTL
Glyma19g26200	At5g40960	GPI-anchored-peptide
Glyma18g50760	At5g60920	COBRA
Glyma04g03220	At4g24550	Endocytosis
Glyma13g27250	At5g05170	Cellulose-synthase
Glyma13g09707	At5g51480	SKU7
Glyma17g15256	At4g16745	GT-family-47
Glyma16g04360	At2g19070	Pathways-phenyl
Glyma03g38650	At5g26670	Pectin-acetylerases
Glyma09g15620	At5g05170	Cellulose-synthase
Glyma15g03091	At2g28950	Expansins
Glyma14g37680	At5g58490	Pathways-phenyl
Glyma13g05610	At1g26690	Vesicle-trafficking
Glyma06g21090	At2g03220	GT-family-37
Glyma12g09660	At5g22430	Unknow-GPI-anchored
Glyma07g15910	At2g37640	Expansins
Glyma19g01220	At3g55250	Pectate-pectin-Lyases
Glyma08g44300	At3g07680	Vesicle-trafficking
Glyma15g38670	At2g19070	Pathways-phenyl
Glyma11g36460	At1g55260	LPTL
Glyma18g45710	At1g08500	Early-nodulin-like
Glyma12g07240	At2g28410	Glycosylphosphatidylinositol
Glyma03g29431	At2g41850	Polygalacturonases
Glyma06g44770	At4g03550	Callose-synthase
Glyma08g04020	At4g28850	Xyloglucan
Glyma07g37610	At3g62830	Nucleotide-sugar
Glyma08g42280	At3g58790	GT-family-8
Glyma09g34431	At4g19010	Pathways-phenyl
Glyma18g47610	At1g65380	Receptor-like-kinases
Glyma10g27960	At1g57590	Pectin-acetylerases

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma13g23010	At5g20260	GT-family-47
Glyma20g00760	At1g68470	GT-family-47
Glyma17g04230	At3g06440	GT-family-31
Glyma12g06980	At2g28760	Nucleotide-sugar
Glyma01g45050	At3g06035	Unknow-GPI-anchored
Glyma08g16660	At2g13680	Callose-synthase
Glyma18g12870	At1g06490	Callose-synthase
Glyma14g00720	At4g12730	Arabinogalactan-proteins
Glyma10g32010	At4g35160	Pathways-phenyl
Glyma10g39530	At5g26330	Stellacyanin-like
Glyma12g12750	At4g03550	Callose-synthase
Glyma17g10330	At5g62620	GT-family-31
Glyma02g41590	At2g39700	Expansins
Glyma12g31060	At2g27500	Glycoside-hidrolase-family-17
Glyma11g36730	At3g23730	Xyloglucan
Glyma08g16715	At5g13000	Callose-synthase
Glyma08g16730	At2g31960	Callose-synthase
Glyma19g43061	At5g17540	Pathways-phenyl
Glyma12g36290	At1g78570	Nucleotide-sugar
Glyma15g05650	At4g25980	Peroxidases
Glyma19g16971	At3g23730	Xyloglucan
Glyma07g13840	At3g60280	Uclacyanin-like
Glyma12g31810	At2g32540	Cellulose-synthase-like
Glyma15g16650	At3g46550	Arabinogalactan-proteins
Glyma10g05290	At3g55140	Pectate-pectin-Lyases
Glyma13g04780	At1g70090	GT-family-8
Glyma16g03680	At3g61490	Polygalacturonases
Glyma13g02890	At1g09580	Vesicle-trafficking
Glyma18g20720	At1g60780	Endocytosis
Glyma05g28310	At3g23730	Xyloglucan
Glyma17g12170	At2g32300	Uclacyanin-like
Glyma17g12180	At2g26600	Glycoside-hidrolase-family-17
Glyma01g24455	At5g03800	GT-family-47
Glyma17g34040	At4g31840	Early-nodulin-like
Glyma05g30330	At1g09580	Vesicle-trafficking
Glyma13g06050	At3g29360	Nucleotide-sugar
Glyma17g29760	At4g31140	Glycoside-hidrolase-family-17
Glyma01g38520	At3g50760	GT-family-8
Glyma03g22260	At3g07390	Unknow-GPI-anchored
Glyma16g04950	At5g13870	Xyloglucan
Glyma18g25160	At2g39700	Expansins

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma12g32590	At4g28100	Unknow-GPI-anchored
Glyma05g28720	At1g55260	LPTL
Glyma15g01250	At1g80170	Polygalacturonases
Glyma12g32820	At2g20810	GT-family-8
Glyma07g37210	At4g16120	COBRA
Glyma13g05950	At3g01040	GT-family-8
Glyma01g02630	At4g38040	GT-family-47
Glyma13g23020	At5g33290	GT-family-47
Glyma13g37300	At4g28300	Hydroxyproline-rich
Glyma05g32130	At4g39330	Pathways-phenyl
Glyma13g22640	At2g26600	Glycoside-hidrolase-family-17
Glyma03g33880	At2g37040	Pathways-phenyl
Glyma17g03390	At4g16120	COBRA
Glyma11g35510	At2g40230	Pathways-phenyl
Glyma02g40480	At1g77130	GT-family-8
Glyma05g32380	At5g09760	Pectin-methyl-esterases
Glyma13g33960	At1g78570	Nucleotide-sugar
Glyma14g38290	At1g21480	GT-family-47
Glyma08g15650	At5g09760	Pectin-methyl-esterases
Glyma07g02460	At5g48930	Pathways-phenyl
Glyma06g01060	At4g39010	Glycoside-hidrolase-family-9
Glyma08g23560	At5g48930	Pathways-phenyl
Glyma16g26860	At1g64760	Glycoside-hidrolase-family-17
Glyma20g36121	At1g05840	Proteases
Glyma13g41160	At2g39700	Expansins
Glyma06g08960	At5g25820	GT-family-47
Glyma13g01120	At4g25810	Xyloglucan
Glyma06g36590	At5g53870	Early-nodulin-like
Glyma09g01980	At5g47780	GT-family-8
Glyma06g48040	At5g56170	Glycosylphosphatidylinositol
Glyma17g07270	At4g25810	Xyloglucan
Glyma13g37650	At2g20810	GT-family-8
Glyma01g40270	At2g22900	GT-family-34
Glyma11g10061	At4g16260	Glycoside-hidrolase-family-17
Glyma20g38490	At3g55140	Pectate-pectin-Lyases
Glyma14g06280	At2g40230	Pathways-phenyl
Glyma04g05100	At4g07960	Cellulose-synthase-like
Glyma17g08350	At5g58050	GPDL
Glyma11g33190	At3g03220	Expansins
Glyma08g39340	At1g60590	Polygalacturonases
Glyma10g39010	At4g14746	Unknow-GPI-anchored



**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma03g34060	At5g03760	Cellulose-synthase-like
Glyma11g03630	At4g16640	Proteases
Glyma15g12850	At4g17180	Glycoside-hidrolase-family-17
Glyma09g08270	At1g19170	Polygalacturonases
Glyma02g47076	At1g55850	Cellulose-synthase-like
Glyma17g04770	At5g26670	Pectin-acetylerases
Glyma19g39760	At3g62720	GT-family-34
Glyma17g09341	At2g14620	Xyloglucan
Glyma20g10310	At3g43720	LPTL
Glyma13g42140	At5g24070	Peroxidases
Glyma09g39200	At3g61490	Polygalacturonases
Glyma10g24080	At1g65680	Expansins
Glyma13g34630	At3g14960	GT-family-31
Glyma08g01480	At1g12780	Nucleotide-sugar
Glyma13g09200	At2g40230	Pathways-phenyl
Glyma07g02500	At5g15350	Early-nodulin-like
Glyma08g42150	At1g06490	Callose-synthase
Glyma14g08430	At3g50760	GT-family-8
Glyma05g34930	At2g01630	Glycoside-hidrolase-family-17
Glyma07g37751	At4g22580	GT-family-47
Glyma05g37490	At4g23500	Polygalacturonases
Glyma16g27880	At1g71695	Peroxidases
Glyma16g23640	At4g38660	Receptor-like
Glyma14g16830	At4g31140	Glycoside-hidrolase-family-17
Glyma01g04380	At1g24140	Proteases
Glyma07g27990	At1g14720	Xyloglucan
Glyma03g03360	At3g60730	Pectin-methyl-esterases
Glyma10g38520	At5g51890	Peroxidases
Glyma08g03440	At4g26220	Pathways-phenyl
Glyma13g24190	At3g23770	Glycoside-hidrolase-family-17
Glyma09g09050	At3g43270	Pectin-methyl-esterases
Glyma02g42940	At2g40370	Laccases
Glyma15g20550	At3g43270	Pectin-methyl-esterases
Glyma05g27610	At4g21060	GT-family-31
Glyma12g12350	At4g28250	Expansins
Glyma19g02031	At5g42740	Glycosylphosphatidylinositol
Glyma08g10590	At4g21060	GT-family-31
Glyma08g41160	At5g24090	Yieldins
Glyma09g38720	At1g65380	Receptor-like-kinases
Glyma18g50770	At5g15630	COBRA

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma15g01160	At5g62200	Unknow-GPI-anchored
Glyma04g40580	At5g54160	Pathways-phenyl
Glyma04g40000	At2g40610	Expansins
Glyma03g38770	At4g26830	Glycoside-hidrolase-family-17
Glyma17g36650	At3g50760	GT-family-8
Glyma13g28380	At5g18900	Prolyl-4-hydroxylases
Glyma05g34850	At1g70710	Glycoside-hidrolase-family-9
Glyma18g32840	At3g07320	Glycoside-hidrolase-family-17
Glyma12g08460	At4g05160	Pathways-phenyl
Glyma08g13540	At1g30620	Nucleotide-sugar
Glyma13g31310	At2g13680	Callose-synthase
Glyma17g08830	At5g60920	COBRA
Glyma13g05810	At5g15350	Early-nodulin-like
Glyma15g04500	At3g46440	Nucleotide-sugar
Glyma20g01710	At1g70985	Unknow-GPI-anchored
Glyma17g08910	At1g18580	GT-family-8
Glyma18g05040	At3g03220	Expansins
Glyma09g02821	At2g19440	Glycoside-hidrolase-family-17
Glyma01g44530	At3g22600	LPTL
Glyma20g30460	At5g24090	Yieldins
Glyma05g07791	At1g48940	Early-nodulin-like
Glyma17g07150	At5g57500	GT-family-31
Glyma15g01780	At4g19640	Endocytosis
Glyma08g27570	At5g15630	COBRA
Glyma03g28850	At3g57270	Glycoside-hidrolase-family-17
Glyma10g03375	At2g48140	LPTL
Glyma02g39360	At3g11130	Endocytosis
Glyma19g22790	At3g60730	Pectin-methyl-esterases
Glyma04g41460	At5g53370	Pectin-methyl-esterases
Glyma06g15711	At5g09760	Pectin-methyl-esterases
Glyma11g21190	At4g24010	Cellulose-synthase-like
Glyma01g01780	At3g03050	Cellulose-synthase-like
Glyma09g40260	At3g61130	GT-family-8
Glyma17g07280	At4g25810	Xyloglucan
Glyma08g25020	At2g44300	LPTL
Glyma06g47670	At4g25240	SKU6
Glyma17g33730	At2g24800	Peroxidases
Glyma06g05190	At4g07960	Cellulose-synthase-like
Glyma13g24270	At4g24000	Cellulose-synthase-like
Glyma05g37540	At4g23490	GT-family-31

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma13g32830	At1g72680	Pathways-phenyl
Glyma14g34100	At5g10080	Proteases
Glyma17g14620	At1g62790	LPTL
Glyma02g05050	At4g35880	Proteases
Glyma06g45043	At5g54160	Pathways-phenyl
Glyma15g08020	At4g04970	Callose-synthase
Glyma01g39900	At4g33910	Prolyl-4-hydroxylases
Glyma10g28551	At1g02460	Polygalacturonases
Glyma02g46330	At5g55180	Glycoside-hidrolase-family-17
Glyma13g29280	At4g21060	GT-family-31
Glyma05g29690	At3g23730	Xyloglucan
Glyma05g28870	At3g13560	Glycoside-hidrolase-family-17
Glyma02g01230	At1g02460	Polygalacturonases
Glyma06g20070	At1g74790	HIPL
Glyma04g36950	At5g63380	Pathways-phenyl
Glyma01g30560	At4g16790	Hydroxyproline-rich
Glyma07g12010	At5g20710	Beta-galactosidase
Glyma12g31830	At2g32530	Cellulose-synthase-like
Glyma13g40200	At2g28470	Beta-galactosidase
Glyma12g34100	At3g20570	Early-nodulin-like
Glyma17g12150	At5g26330	Stellacyanin-like
Glyma13g40960	At3g46440	Nucleotide-sugar
Glyma13g44270	At1g80080	Glycosylphosphatidylinositol
Glyma14g00801	At5g54160	Pathways-phenyl
Glyma02g12140	At1g69530	Expansins
Glyma01g39810	At4g37450	Arabinogalactan-proteins
Glyma06g48050	At5g56170	Glycosylphosphatidylinositol
Glyma17g18800	At4g34050	Pathways-phenyl
Glyma16g24341	At4g36220	Pathways-phenyl
Glyma07g11300	At3g18050	Unknow-GPI-anchored
Glyma01g40880	At4g18010	Endocytosis
Glyma06g22410	At5g60920	COBRA
Glyma11g01230	At3g03050	Cellulose-synthase-like
Glyma08g47161	At4g13710	Pectate-pectin-Lyases
Glyma10g36810	At1g21090	Early-nodulin-like
Glyma06g45475	At4g28100	Unknow-GPI-anchored
Glyma04g15960	At5g09760	Pectin-methyl-esterases
Glyma01g42441	At1g63000	Nucleotide-sugar
Glyma16g01650	At3g14310	Pectin-methyl-esterases
Glyma19g41950	At3g05620	Pectin-methyl-esterases
Glyma04g06410	At2g25060	Early-nodulin-like
Glyma10g24630	At4g13710	Pectate-pectin-Lyases

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma02g40950	At4g12730	Arabinogalactan-proteins
Glyma05g13660	At4g14130	Xyloglucan
Glyma15g12900	At5g47780	GT-family-8
Glyma15g42330	At2g31960	Callose-synthase
Glyma17g07250	At4g25810	Xyloglucan
Glyma20g28070	At3g06035	Unknow-GPI-anchored
Glyma05g07410	At1g18580	GT-family-8
Glyma06g46450	At2g32540	Cellulose-synthase-like
Glyma11g15020	At3g46440	Nucleotide-sugar
Glyma14g13333	At4g19810	Yieldins
Glyma02g00340	At5g17540	Pathways-phenyl
Glyma10g07050	At1g20030	Receptor-like
Glyma20g24670	At5g04310	Pectate-pectin-Lyases
Glyma04g09680	At1g09890	Rhamnogalacturonan
Glyma19g37325	At5g03800	GT-family-47
Glyma20g04490	At2g03090	Expansins
Glyma06g13060	At2g46330	Arabinogalactan-proteins
Glyma19g36620	At2g37040	Pathways-phenyl
Glyma06g48260	At4g24010	Cellulose-synthase-like
Glyma01g36780	At5g42180	Peroxidases
Glyma13g44770	At2g03220	GT-family-37
Glyma17g37960	At4g39330	Pathways-phenyl
Glyma20g01700	At1g23040	Extensin-related
Glyma17g05350	At5g22740	Cellulose-synthase-like
Glyma20g23271	At4g38140	Unknow-GPI-anchored
Glyma20g39250	At2g43080	Prolyl-4-hydroxylases
Glyma12g29660	At2g28470	Beta-galactosidase
Glyma03g34800	At5g03760	Cellulose-synthase-like
Glyma11g04471	At2g35380	Peroxidases
Glyma06g08973	At5g19670	GT-family-47
Glyma09g30930	At3g18050	Unknow-GPI-anchored
Glyma13g24110	At2g34060	Peroxidases
Glyma17g03030	At3g62830	Nucleotide-sugar
Glyma02g47940	At2g37040	Pathways-phenyl
Glyma16g03740	At5g41460	GT-family-31
Glyma03g04390	At5g05290	Expansins
Glyma06g12150	At1g77410	Beta-galactosidase
Glyma13g36441	At3g20570	Early-nodulin-like
Glyma11g20606	At2g28410	Glycosylphosphatidylinositol
Glyma18g49720	At3g29090	Pectin-methyl-esterases
Glyma06g03770	At1g19300	GT-family-8
Glyma08g17300	At1g30870	Peroxidases
Glyma16g27900	At1g71695	Peroxidases

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma03g01941	At2g06850	Xyloglucan
Glyma08g47400	At5g60020	Laccases
Glyma15g09810	At4g21060	GT-family-31
Glyma17g33190	At4g19810	Yieldins
Glyma18g00381	At1g55260	LPTL
Glyma02g01910	At3g16850	Polygalacturonases
Glyma13g44950	At3g21240	Pathways-phenyl
Glyma01g35480	At4g19420	Pectin-acetylerases
Glyma16g24440	At4g36360	Beta-galactosidase
Glyma17g11170	At4g19810	Yieldins
Glyma06g03160	At4g36360	Beta-galactosidase
Glyma13g44240	At3g15800	Glycoside-hidrolase-family-17
Glyma11g34500	At2g44790	Uclacyanin-like
Glyma06g03290	At1g27620	Pathways-phenyl
Glyma10g27840	At3g62110	Polygalacturonases
Glyma12g10610	At1g61900	Unknow-GPI-anchored
Glyma11g08520	At5g42180	Peroxidases
Glyma13g38150	At2g32300	Uclacyanin-like
Glyma15g41320	At5g14180	Proteases
Glyma01g44270	At1g65060	Pathways-phenyl
Glyma01g06720	At5g24090	Yieldins
Glyma08g46820	At5g60020	Laccases
Glyma19g41240	At5g26670	Pectin-acetylerases
Glyma02g45720	At3g58790	GT-family-8
Glyma01g00771	At3g61640	Arabinogalactan-proteins
Glyma09g01250	At2g32300	Uclacyanin-like
Glyma02g43680	At2g32990	Glycoside-hidrolase-family-9
Glyma09g39233	At5g41460	GT-family-31
Glyma02g08920	At5g64740	Cellulose-synthase
Glyma10g27457	At3g48950	Polygalacturonases
Glyma14g04140	At1g08500	Early-nodulin-like
Glyma13g17740	At5g26670	Pectin-acetylerases
Glyma06g11820	At1g21880	Receptor-like
Glyma11g19910	At2g28110	GT-family-47
Glyma11g20730	At2g28470	Beta-galactosidase
Glyma15g41860	At1g30870	Peroxidases
Glyma06g14220	At5g54160	Pathways-phenyl
Glyma05g04180	At1g62790	LPTL
Glyma03g41660	At3g61640	Arabinogalactan-proteins
Glyma12g00700	At3g47400	Pectin-methyl-esterases
Glyma03g33890	At2g37040	Pathways-phenyl
Glyma01g41820	At5g47500	Pectin-methyl-esterases

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma14g07700	At1g45130	Beta-galactosidase
Glyma12g02410	At4g16260	Glycoside-hidrolase-family-17
Glyma05g28030	At1g27950	LPTL
Glyma11g36410	At4g35160	Pathways-phenyl
Glyma19g37180	At2g36710	Pectin-methyl-esterases
Glyma04g01030	At4g39010	Glycoside-hidrolase-family-9
Glyma07g01250	At3g13750	Beta-galactosidase
Glyma07g34050	At1g23040	Extensin-related
Glyma03g38430	At2g46930	Pectin-acetylerases
Glyma16g31008	At2g44300	LPTL
Glyma08g44320	At1g55850	Cellulose-synthase-like
Glyma10g39760	At1g11545	Xyloglucan
Glyma08g11300	At3g23730	Xyloglucan
Glyma06g47420	At5g17420	Cellulose-synthase
Glyma06g15671	At5g41290	Receptor-like
Glyma02g17060	At1g05260	Peroxidases
Glyma10g39110	At5g60490	Arabinogalactan-proteins
Glyma06g15240	At5g64790	Glycoside-hidrolase-family-17
Glyma13g22595	At2g13680	Callose-synthase
Glyma20g26440	At3g19450	Pathways-phenyl
Glyma20g02741	At4g34480	Glycoside-hidrolase-family-17
Glyma17g15390	At4g16790	Hydroxyproline-rich
Glyma08g19340	At4g25980	Peroxidases
Glyma11g07760	At4g36360	Beta-galactosidase
Glyma02g41190	At1g66250	Glycoside-hidrolase-family-17
Glyma02g01310	At5g19730	Pectin-methyl-esterases
Glyma20g01520	At1g14720	Xyloglucan
Glyma08g42110	At1g06490	Callose-synthase
Glyma07g38430	At3g06260	GT-family-8
Glyma16g28080	At5g64740	Cellulose-synthase
Glyma15g18430	At4g26140	Beta-galactosidase
Glyma08g02050	At4g23500	Polygalacturonases
Glyma17g01660	At1g05170	GT-family-31
Glyma02g03301	At1g59970	Proteases
Glyma05g35660	At4g28850	Xyloglucan
Glyma05g35730	At5g19670	GT-family-47
Glyma06g41300	At3g07160	Callose-synthase
Glyma18g39850	At2g39700	Expansins
Glyma11g02170	At5g53340	GT-family-31

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma02g42110	At3g55780	Glycoside-hidrolase-family-17
Glyma04g40530	At1g05260	Peroxidases
Glyma13g44830	At5g48930	Pathways-phenyl
Glyma02g44235	At4g35410	Endocytosis
Glyma18g01560	At1g67750	Pectate-pectin-Lyases
Glyma18g02280	At5g10080	Proteases
Glyma07g11822	At5g16890	GT-family-47
Glyma04g03120	At4g36360	Beta-galactosidase
Glyma11g29410	At5g58480	Glycoside-hidrolase-family-17
Glyma03g27740	At2g40890	Pathways-phenyl
Glyma19g28220	At5g13870	Xyloglucan
Glyma12g34140	At3g20580	COBRA
Glyma01g34990	At5g19670	GT-family-47
Glyma10g44141	At5g13000	Callose-synthase
Glyma11g11910	At4g39010	Glycoside-hidrolase-family-9
Glyma13g21440	At5g03760	Cellulose-synthase-like
Glyma18g12604	At3g58790	GT-family-8
Glyma09g02600	At2g38380	Peroxidases
Glyma12g02015	At3g03050	Cellulose-synthase-like
Glyma02g00730	At4g14746	Unknow-GPI-anchored
Glyma14g36500	At2g30490	Pathways-phenyl
Glyma08g23820	At4g31590	Cellulose-synthase-like
Glyma10g42390	At5g04310	Pectate-pectin-Lyases
Glyma09g07100	At4g26140	Beta-galactosidase
Glyma18g39570	At3g09410	Pectin-acetylesterases
Glyma12g31840	At2g32540	Cellulose-synthase-like
Glyma12g08530	At2g28110	GT-family-47
Glyma15g04240	At2g39700	Expansins
Glyma15g15690	At3g16850	Polygalacturonases
Glyma06g45860	At4g03210	Xyloglucan
Glyma18g08400	At3g07680	Vesicle-trafficking
Glyma06g08180	At4g31590	Cellulose-synthase-like
Glyma10g31430	At1g05840	Proteases
Glyma12g25740	At5g24090	Yieldins
Glyma11g34040	At2g37640	Expansins
Glyma20g28210	At5g26330	Stellacyanin-like
Glyma07g14930	At5g19730	Pectin-methyl-esterases
Glyma17g11840	At5g33290	GT-family-47
Glyma06g12910	At5g49720	Glycoside-hidrolase-family-9
Glyma02g03250	At1g59970	Proteases
Glyma08g04780	At2g01630	Glycoside-hidrolase-family-17

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma01g22480	At1g13250	GT-family-8
Glyma05g23170	At5g65730	Xyloglucan
Glyma04g07820	At4g31140	Glycoside-hidrolase-family-17
Glyma08g42270	At1g78570	Nucleotide-sugar
Glyma14g11760	At4g31840	Early-nodulin-like
Glyma20g35610	At4g35160	Pathways-phenyl
Glyma01g37920	At5g03260	Laccases
Glyma17g02330	At3g06260	GT-family-8
Glyma15g02750	At3g13750	Beta-galactosidase
Glyma17g11870	At5g33290	GT-family-47
Glyma13g23040	At5g20260	GT-family-47
Glyma11g36640	At4g13990	GT-family-47
Glyma11g11336	At4g38660	Receptor-like
Glyma14g02070	At1g66970	GPDL
Glyma16g29780	At3g61490	Polygalacturonases
Glyma03g37550	At1g02730	Cellulose-synthase-like
Glyma13g42680	At3g13750	Beta-galactosidase
Glyma13g01131	At4g25810	Xyloglucan
Glyma16g29960	At5g23940	Pathways-phenyl
Glyma08g17820	At5g14180	Proteases
Glyma12g10930	At1g05280	GT-family-31
Glyma02g31340	At3g57630	GT-family-47
Glyma04g03690	At3g50760	GT-family-8
Glyma11g07850	At4g36220	Pathways-phenyl
Glyma02g40020	At5g05340	Peroxidases
Glyma09g34140	At1g32170	Xyloglucan
Glyma05g22180	At5g67400	Peroxidases
Glyma08g02030	At4g23490	GT-family-31
Glyma20g19200	At4g13710	Pectate-pectin-Lyases
Glyma10g01290	At1g02460	Polygalacturonases
Glyma17g29320	At4g30170	Peroxidases
Glyma15g05190	At1g09580	Vesicle-trafficking
Glyma12g09510	At2g27500	Glycoside-hidrolase-family-17
Glyma14g18140	At4g31590	Cellulose-synthase-like
Glyma02g43480	At5g14180	Proteases
Glyma07g15341	At1g67980	Pathways-phenyl
Glyma02g45280	At1g32910	Pathways-phenyl
Glyma05g32500	At2g13680	Callose-synthase
Glyma20g37680	At1g54860	Unknow-GPI-anchored
Glyma04g42930	At1g21880	Receptor-like
Glyma13g36420	At3g20580	COBRA
Glyma17g04030	At4g16270	Peroxidases



**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma14g38301	At1g21480	GT-family-47
Glyma18g13690	At1g78990	Pathways-phenyl
Glyma17g07220	At5g57560	Xyloglucan
Glyma02g12660	At5g24090	Yieldins
Glyma05g04270	At4g12420	SKU5
Glyma16g26480	At5g13900	LPTL
Glyma02g47720	At3g07840	Polygalacturonases
Glyma02g00930	At3g05910	Pectin-acetylerases
Glyma15g27510	At1g78570	Nucleotide-sugar
Glyma02g07730	At2g01630	Glycoside-hidrolase-family-17
Glyma18g38430	At3g07010	Pectate-pectin-Lyases
Glyma07g12060	At5g20710	Beta-galactosidase
Glyma15g39420	At3g14570	Callose-synthase
Glyma10g28230	At3g05910	Pectin-acetylerases
Glyma01g40870	At2g35380	Peroxidases
Glyma19g30600	At2g40890	Pathways-phenyl
Glyma18g13105	At1g06490	Callose-synthase
Glyma16g30130	At4g25600	Prolyl-4-hydroxylases
Glyma18g44391	At1g63180	Nucleotide-sugar
Glyma08g15680	At1g78570	Nucleotide-sugar
Glyma12g03510	At4g38660	Receptor-like
Glyma18g49870	At5g54160	Pathways-phenyl
Glyma19g28640	At1g11545	Xyloglucan
Glyma07g33780	At4g34050	Pathways-phenyl
Glyma08g03430	At4g26220	Pathways-phenyl
Glyma10g01180	At3g05610	Pectin-methyl-esterases
Glyma09g25173	At4g25600	Prolyl-4-hydroxylases
Glyma08g04840	At1g70710	Glycoside-hidrolase-family-9
Glyma01g44281	At3g03050	Cellulose-synthase-like
Glyma09g35050	At4g19420	Pectin-acetylerases
Glyma15g01820	At5g24090	Yieldins
Glyma18g06660	At2g19070	Pathways-phenyl
Glyma05g03720	At1g12560	Expansins
Glyma08g43650	At5g55480	GPLD
Glyma04g13670	At4g25240	SKU6
Glyma12g10960	At4g03210	Xyloglucan
Glyma11g29620	At2g30210	Laccases
Glyma06g08966	At5g25820	GT-family-47
Glyma15g40610	At1g20030	Receptor-like
Glyma18g12280	At2g19070	Pathways-phenyl
Glyma13g17600	At5g20870	Glycoside-hidrolase-family-17
Glyma17g37240	At1g44970	Peroxidases

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma08g23480	At2g03220	GT-family-37
Glyma14g01660	At1g55850	Cellulose-synthase-like
Glyma06g11390	At1g77780	Glycoside-hidrolase-family-17
Glyma17g15060	At4g16600	GT-family-8
Glyma20g04415	At4g18010	Endocytosis
Glyma19g36810	At5g03760	Cellulose-synthase-like
Glyma02g03190	At1g24140	Proteases
Glyma10g29620	At3g06035	Unknow-GPI-anchored
Glyma03g26060	At2g32300	Uclacyanin-like
Glyma02g39930	At4g35160	Pathways-phenyl
Glyma10g36690	At1g71695	Peroxidases
Glyma09g37090	At1g69530	Expansins
Glyma13g28890	At1g29980	Unknow-GPI-anchored
Glyma09g37691	At4g16790	Hydroxyproline-rich
Glyma15g11420	At1g61900	Unknow-GPI-anchored
Glyma13g17570	At3g43270	Pectin-methyl-esterases
Glyma14g38430	At2g40610	Expansins
Glyma13g22350	At4g19810	Yieldins
Glyma02g05790	At4g36360	Beta-galactosidase
Glyma15g12080	At2g32300	Uclacyanin-like
Glyma02g46890	At5g55590	Pectin-methyl-esterases
Glyma05g00590	At3g29360	Nucleotide-sugar
Glyma01g38580	At3g11130	Endocytosis
Glyma16g17150	At4g19420	Pectin-acetylerases
Glyma16g32740	At5g62200	Unknow-GPI-anchored
Glyma02g11100	At1g13250	GT-family-8
Glyma18g00560	At4g13990	GT-family-47
Glyma09g04560	At3g16850	Polygalacturonases
Glyma02g45080	At4g23820	Polygalacturonases
Glyma02g39750	At2g30210	Laccases
Glyma17g17730	At5g67400	Peroxidases
Glyma08g42450	At2g19070	Pathways-phenyl
Glyma16g28970	At5g24090	Yieldins
Glyma18g16870	At5g62620	GT-family-31
Glyma18g16870	At5g62620	GT-family-31
Glyma07g15281	At3g61640	Arabinogalactan-proteins
Glyma11g20720	At5g03170	Arabinogalactan-proteins
Glyma02g02460	At1g14420	Pectate-pectin-Lyases
Glyma12g35411	At5g53870	Early-nodulin-like
Glyma01g01770	At1g32170	Xyloglucan
Glyma08g22680	At5g20230	Stellacyanin-like
Glyma16g30681	At2g44300	LPTL

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma19g03500	At3g29360	Nucleotide-sugar
Glyma13g44760	At2g03210	GT-family-37
Glyma08g37431	At3g19450	Pathways-phenyl
Glyma16g26630	At5g13870	Xyloglucan
Glyma20g27970	At1g11545	Xyloglucan
Glyma16g04350	At2g19070	Pathways-phenyl
Glyma16g31540	At2g44300	LPTL
Glyma08g12800	At3g23730	Xyloglucan
Glyma02g08382	At5g48100	Laccases
Glyma18g47130	At3g61490	Polygalacturonases
Glyma03g41510	At5g24090	Yieldins
Glyma08g17850	At4g33870	Peroxidases
Glyma08g44310	At1g55850	Cellulose-synthase-like
Glyma17g07260	At4g25810	Xyloglucan
Glyma10g34550	At5g03760	Cellulose-synthase-like
Glyma11g05300	At4g37530	Peroxidases
Glyma06g10500	At1g72230	Uclacyanin-like
Glyma06g15030	At5g05340	Peroxidases
Glyma08g20650	At3g13750	Beta-galactosidase
Glyma16g23030	At1g20270	Prolyl-4-hydroxylases
Glyma12g32270	At2g32300	Uclacyanin-like
Glyma19g40020	At4g02300	Pectin-methyl-esterases
Glyma06g14200	At5g54160	Pathways-phenyl
Glyma15g35290	At3g10720	Pectin-methyl-esterases
Glyma02g07961	At5g56540	Arabinogalactan-proteins
Glyma04g22130	At1g03390	Pathways-phenyl
Glyma16g27890	At1g71695	Peroxidases
Glyma14g05850	At5g05340	Peroxidases
Glyma07g34500	At4g18340	Glycoside-hidrolase-family-17
Glyma09g10985	At1g20510	Pathways-phenyl
Glyma19g41430	At1g02460	Polygalacturonases
Glyma18g27090	At2g22620	Rhamnogalacturonan
Glyma10g36790	At5g64740	Cellulose-synthase
Glyma01g06030	At1g69530	Expansins
Glyma08g41930	At1g78990	Pathways-phenyl
Glyma04g32130	At5g60920	COBRA
Glyma01g42831	At1g63450	GT-family-47
Glyma06g44930	At4g28250	Expansins
Glyma14g38210	At5g58400	Peroxidases
Glyma08g40570	At5g62620	GT-family-31
Glyma14g38170	At5g05340	Peroxidases
Glyma12g31780	At2g32540	Cellulose-synthase-like

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma10g01250	At1g05260	Peroxidases
Glyma06g44010	At4g35160	Pathways-phenyl
Glyma12g07381	At2g28470	Beta-galactosidase
Glyma11g01480	At1g33250	GT-family-31
Glyma01g04350	At1g59970	Proteases
Glyma11g20791	At5g60490	Arabinogalactan-proteins
Glyma15g02160	At1g79800	Early-nodulin-like
Glyma15g13491	At5g06730	Peroxidases
Glyma04g41720	At2g46330	Arabinogalactan-proteins
Glyma16g23680	At4g36010	Receptor-like
Glyma20g34351	At4g24010	Cellulose-synthase-like
Glyma10g36380	At1g14550	Peroxidases
Glyma18g19670	At1g60590	Polygalacturonases
Glyma14g10230	At4g07960	Cellulose-synthase-like
Glyma13g38310	At1g05260	Peroxidases
Glyma13g00280	At4g25810	Xyloglucan
Glyma09g36660	At3g47400	Pectin-methyl-esterases
Glyma09g07550	At5g19890	Peroxidases
Glyma12g32390	At4g03210	Xyloglucan
Glyma19g22460	At5g63380	Pathways-phenyl
Glyma03g37280	At3g62830	Nucleotide-sugar
Glyma09g41850	At5g54160	Pathways-phenyl
Glyma19g04220	At5g15630	COBRA
Glyma19g28200	At5g13870	Xyloglucan
Glyma16g32690	At5g62200	Unknow-GPI-anchored
Glyma06g16430	At5g63800	Beta-galactosidase
Glyma12g12230	At5g54160	Pathways-phenyl
Glyma13g01080	At3g21240	Pathways-phenyl
Glyma15g41280	At4g33870	Peroxidases
Glyma10g00620	At4g14746	Unknow-GPI-anchored
Glyma19g37480	At5g03760	Cellulose-synthase-like
Glyma13g18711	At4g28560	Receptor-like
Glyma10g22015	At1g45130	Beta-galactosidase
Glyma12g32160	At1g05260	Peroxidases
Glyma10g42230	At2g30490	Pathways-phenyl
Glyma08g03921	At5g19670	GT-family-47
Glyma12g09350	At1g08200	Nucleotide-sugar
Glyma02g01990	At4g02290	Glycoside-hidrolase-family-9
Glyma09g36750	At1g02790	Polygalacturonases
Glyma19g41350	At5g27870	Pectin-methyl-esterases
Glyma18g39440	At2g29130	Laccases
Glyma04g38280	At3g07620	GT-family-47
Glyma10g00220	At5g17540	Pathways-phenyl

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma02g01050	At3g62110	Polygalacturonases
Glyma13g35100	At5g53870	Early-nodulin-like
Glyma20g33101	At5g48100	Laccases
Glyma15g14540	At1g48100	Polygalacturonases
Glyma16g04391	At4g38040	GT-family-47
Glyma07g00590	At4g31590	Cellulose-synthase-like
Glyma10g40040	At2g14620	Xyloglucan
Glyma16g26965	At5g56540	Arabinogalactan-proteins
Glyma19g07631	At3g18590	Early-nodulin-like
Glyma11g37620	At1g67750	Pectate-pectin-Lyases
Glyma01g02570	At4g37980	Pathways-phenyl
Glyma16g06230	At5g14180	Proteases
Glyma03g34170	At2g36870	Xyloglucan
Glyma14g38101	At4g35150	Pathways-phenyl
Glyma14g38111	At4g35160	Pathways-phenyl
Glyma03g39360	At5g27870	Pectin-methyl-esterases
Glyma10g35381	At3g53260	Pathways-phenyl
Glyma11g03560	At5g47500	Pectin-methyl-esterases
Glyma06g38100	At3g15370	Expansins
Glyma10g42840	At2g31050	Stellacyanin-like
Glyma01g37430	At4g36220	Pathways-phenyl
Glyma19g29730	At3g45400	GT-family-47
Glyma11g06720	At3g11130	Endocytosis
Glyma19g25980	At5g14130	Peroxidases
Glyma06g43940	At4g35150	Pathways-phenyl
Glyma17g08400	At4g19810	Yieldins
Glyma17g29060	At4g31590	Cellulose-synthase-like
Glyma12g33530	At1g03870	Arabinogalactan-proteins
Glyma20g35680	At4g33420	Peroxidases
Glyma06g15690	At3g18660	GT-family-8
Glyma18g41886	At5g01050	Laccases
Glyma11g01240	At1g65060	Pathways-phenyl
Glyma13g29000	At4g18340	Glycoside-hidrolase-family-17
Glyma09g08920	At2g26440	Pectin-methyl-esterases
Glyma20g27380	At2g14620	Xyloglucan
Glyma05g00920	At2g03220	GT-family-37
Glyma09g34131	At3g03050	Cellulose-synthase-like
Glyma14g37510	At3g11130	Endocytosis
Glyma11g29460	At5g58490	Pathways-phenyl
Glyma18g07240	At2g38080	Laccases
Glyma07g39290	At3g17070	Peroxidases
Glyma09g21970	At5g20710	Beta-galactosidase
Glyma14g07360	At2g39700	Expansins

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma19g40940	At3g62110	Polygalacturonases
Glyma10g02030	At3g16850	Polygalacturonases
Glyma13g32950	At4g38040	GT-family-47
Glyma16g17120	At4g19420	Pectin-acetylerases
Glyma17g04950	At2g26440	Pectin-methyl-esterases
Glyma19g40030	At4g02290	Glycoside-hidrolase-family-9
Glyma06g02430	At1g20190	Expansins
Glyma08g37510	At4g39330	Pathways-phenyl
Glyma14g01830	At5g55590	Pectin-methyl-esterases
Glyma17g14730	At4g12420	SKU5
Glyma15g17620	At2g18980	Peroxidases
Glyma19g28950	At5g56540	Arabinogalactan-proteins
Glyma12g29575	At5g27530	Polygalacturonases
Glyma02g02020	At4g02330	Pectin-methyl-esterases
Glyma18g01570	At1g67750	Pectate-pectin-Lyases
Glyma09g23136	At5g24090	Yieldins
Glyma19g39990	At4g02330	Pectin-methyl-esterases
Glyma02g45620	At3g14310	Pectin-methyl-esterases
Glyma02g05510	At1g19940	Glycoside-hidrolase-family-9
Glyma16g26800	At2g01630	Glycoside-hidrolase-family-17
Glyma05g27331	At2g42800	Receptor-like
Glyma12g30210	At2g28110	GT-family-47
Glyma19g25570	At3g27200	Uclacyanin-like
Glyma16g32490	At5g51890	Peroxidases
Glyma19g04210	At5g60920	COBRA
Glyma11g34300	At3g11420	GT-family-31
Glyma20g30790	At1g21090	Early-nodulin-like
Glyma15g00390	At3g21240	Pathways-phenyl
Glyma12g27381	At1g11545	Xyloglucan
Glyma14g00930	At3g07840	Polygalacturonases
Glyma19g41080	At4g01630	Expansins
Glyma08g18390	At1g20030	Receptor-like
Glyma05g34020	At3g18050	Unknow-GPI-anchored
Glyma06g07890	At4g31140	Glycoside-hidrolase-family-17
Glyma15g01030	At3g15800	Glycoside-hidrolase-family-17
Glyma14g39510	At1g66250	Glycoside-hidrolase-family-17
Glyma10g07075	At4g38660	Receptor-like
Glyma19g02150	At4g36220	Pathways-phenyl
Glyma13g36930	At2g20520	Arabinogalactan-proteins
Glyma18g45750	At3g61130	GT-family-8

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma09g06350	At2g18980	Peroxidases
Glyma06g48140	At1g71380	Glycoside-hidrolase-family-9
Glyma20g24810	At2g30490	Pathways-phenyl
Glyma01g04370	At1g59970	Proteases
Glyma04g13600	At5g51500	Pectin-methyl-esterases
Glyma10g06990	At2g19070	Pathways-phenyl
Glyma13g38220	At1g05280	GT-family-31
Glyma10g06880	At1g20030	Receptor-like
Glyma03g37410	At1g02810	Pectin-methyl-esterases
Glyma08g03560	At2g19790	Endocytosis
Glyma01g45080	At4g27110	COBRA
Glyma10g03360	At2g48130	LPTL
Glyma13g33560	At3g14570	Callose-synthase
Glyma02g12351	At1g26690	Vesicle-trafficking
Glyma06g43701	At2g38080	Laccases
Glyma11g16010	At2g28470	Beta-galactosidase
Glyma13g04480	At3g28180	Cellulose-synthase-like
Glyma03g37391	At4g02330	Pectin-methyl-esterases
Glyma03g38350	At3g62110	Polygalacturonases
Glyma19g22490	At1g20500	Pathways-phenyl
Glyma12g23200	At3g15370	Expansins
Glyma20g04560	At1g08750	Glycosylphosphatidylinositol
Glyma10g29570	At1g54940	GT-family-8
Glyma15g18360	At4g25810	Xyloglucan
Glyma10g07560	At5g03760	Cellulose-synthase-like
Glyma05g36060	At2g19790	Endocytosis
Glyma10g29160	At3g05610	Pectin-methyl-esterases
Glyma08g03700	At5g19730	Pectin-methyl-esterases
Glyma13g38040	At4g03210	Xyloglucan
Glyma20g31190	At1g14550	Peroxidases
Glyma13g05650	At3g29090	Pectin-methyl-esterases
Glyma04g08100	At4g31590	Cellulose-synthase-like
Glyma10g37200	At5g24090	Yieldins
Glyma10g24120	At1g65680	Expansins
Glyma01g39990	At4g37530	Peroxidases
Glyma11g10760	At1g48930	Glycoside-hidrolase-family-9
Glyma10g36320	At5g09360	Laccases
Glyma10g38740	At5g51680	Hydroxyproline-rich
Glyma03g24030	At2g43870	Polygalacturonases
Glyma14g07820	At1g27620	Pathways-phenyl
Glyma20g38851	At5g13000	Callose-synthase
Glyma16g24640	At5g66390	Peroxidases
Glyma06g00810	At3g29810	COBRA

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma08g43640	At5g55480	GPDL
Glyma12g06730	At2g39700	Expansins
Glyma07g07280	At3g61490	Polygalacturonases
Glyma06g45920	At1g05260	Peroxidases
Glyma14g12170	At2g24800	Peroxidases
Glyma06g45880	At1g05280	GT-family-31
Glyma10g37890	At5g23870	Pectin-acetylesterases
Glyma20g29000	At5g51680	Hydroxyproline-rich
Glyma06g06350	At2g24800	Peroxidases
Glyma07g34060	At1g70985	Unknow-GPI-anchored
Glyma08g23470	At2g03220	GT-family-37
Glyma17g37270	At1g45130	Beta-galactosidase
Glyma02g05320	At4g38660	Receptor-like
Glyma11g05380	At4g33910	Prolyl-4-hydroxylases
Glyma15g13540	At5g06720	Peroxidases
Glyma15g13560	At5g06730	Peroxidases
Glyma15g15200	At4g16260	Glycoside-hidrolase-family-17
Glyma14g36201	At1g09890	Rhamnogalacturonan
Glyma09g33400	At2g22620	Rhamnogalacturonan
Glyma18g40350	At2g37730	GT-family-31
Glyma20g35960	At1g64640	Early-nodulin-like
Glyma15g13500	At2g38380	Peroxidases
Glyma19g45000	At5g54160	Pathways-phenyl
Glyma01g01010	At5g19730	Pectin-methyl-esterases
Glyma12g00530	At1g33430	GT-family-31
Glyma02g38050	At4g24430	Rhamnogalacturonan
Glyma08g43630	At4g26690	GPDL
Glyma03g03460	At2g45220	Pectin-methyl-esterases
Glyma06g15860	At2g13680	Callose-synthase
Glyma14g22780	At4g32790	GT-family-47
Glyma02g05060	At2g17760	Proteases
Glyma12g31800	At2g32530	Cellulose-synthase-like
Glyma05g29440	At5g03170	Arabinogalactan-proteins
Glyma05g32370	At3g18660	GT-family-8
Glyma18g49570	At1g69530	Expansins
Glyma13g20890	At3g53190	Pectate-pectin-Lyases
Glyma13g20450	At2g36870	Xyloglucan
Glyma08g41900	At1g78990	Pathways-phenyl
Glyma04g32120	At5g15630	COBRA
Glyma17g07170	At3g21240	Pathways-phenyl
Glyma17g07190	At3g21240	Pathways-phenyl



**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma18g06570	At5g58480	Glycoside-hidrolase-family-17
Glyma02g38061	At4g24430	Rhamnogalacturonan
Glyma05g31860	At5g64790	Glycoside-hidrolase-family-17
Glyma18g49090	At5g15110	Pectate-pectin-Lyases
Glyma08g12400	At4g18780	Cellulose-synthase
Glyma09g35870	At3g07970	Polygalacturonases
Glyma18g12180	At2g19070	Pathways-phenyl
Glyma18g12230	At2g19070	Pathways-phenyl
Glyma08g37420	At2g22620	Rhamnogalacturonan
Glyma18g02690	At2g40370	Laccases
Glyma17g11850	At3g42180	GT-family-47
Glyma03g15800	At3g09220	Laccases
Glyma20g15980	At3g07620	GT-family-47
Glyma16g23140	At2g17760	Proteases
Glyma07g16060	At5g60020	Laccases
Glyma19g41970	At5g27870	Pectin-methyl-esterases
Glyma15g19820	At1g19170	Polygalacturonases
Glyma13g27540	At1g61900	Unknow-GPI-anchored
Glyma15g10050	At4g18340	Glycoside-hidrolase-family-17
Glyma11g07430	At5g03260	Laccases
Glyma08g09350	At5g44030	Cellulose-synthase
Glyma13g17561	At2g26440	Pectin-methyl-esterases
Glyma10g06870	At2g19070	Pathways-phenyl
Glyma18g12210	At2g19070	Pathways-phenyl
Glyma09g05630	At5g44030	Cellulose-synthase
Glyma16g05770	At5g41040	Pathways-phenyl
Glyma05g08730	At5g14650	Polygalacturonases
Glyma18g06450	At2g30210	Laccases
Glyma14g03490	At1g32910	Pathways-phenyl
Glyma05g15376	At4g31590	Cellulose-synthase-like
Glyma14g16630	At3g13560	Glycoside-hidrolase-family-17
Glyma08g44190	At1g62940	Pathways-phenyl
Glyma03g34260	At3g53190	Pectate-pectin-Lyases
Glyma16g08240	At4g19420	Pectin-acetylerases
Glyma09g04720	At5g27870	Pectin-methyl-esterases
Glyma11g02440	At5g41460	GT-family-31
Glyma03g37480	At3g16850	Polygalacturonases
Glyma01g26750	At3g09220	Laccases
Glyma0214s00200	At1g77130	GT-family-8
Glyma08g47380	At5g60020	Laccases

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma09g21980	At1g31740	Beta-galactosidase
Glyma02g38990	At2g38080	Laccases
Glyma10g02160	At4g02330	Pectin-methyl-esterases
Glyma01g43050	At5g41460	GT-family-31
Glyma20g32135	At3g53260	Pathways-phenyl
Glyma04g39240	At3g18660	GT-family-8
Glyma13g37290	At4g03550	Callose-synthase
Glyma19g01560	At3g28180	Cellulose-synthase-like
Glyma12g03650	At2g16730	Beta-galactosidase
Glyma18g09533	At4g26690	GPDL
Glyma01g37541	At4g36360	Beta-galactosidase
Glyma07g05980	At2g46570	Laccases
Glyma04g06780	At4g18780	Cellulose-synthase
Glyma0023s00670	At1g57620	Vesicle-trafficking
Glyma01g02560	At2g22620	Rhamnogalacturonan
Glyma01g02580	At4g37990	Pathways-phenyl
Glyma01g03310	At4g08770	Peroxidases
Glyma01g03400	At3g26610	Polygalacturonases
Glyma01g04361	At1g59970	Proteases
Glyma01g04970	At1g14420	Pectate-pectin-Lyases
Glyma01g07060	At3g45400	GT-family-47
Glyma01g08690	At1g69940	Pectin-methyl-esterases
Glyma01g08730	At1g69940	Pectin-methyl-esterases
Glyma01g08760	At1g69940	Pectin-methyl-esterases
Glyma01g09350	At1g69940	Pectin-methyl-esterases
Glyma01g09540	At5g28840	Nucleotide-sugar
Glyma01g12310	At4g36360	Beta-galactosidase
Glyma01g14500	At4g35670	Polygalacturonases
Glyma01g16140	At1g65680	Expansins
Glyma01g18520	At2g43890	Polygalacturonases
Glyma01g20565	At1g11545	Xyloglucan
Glyma01g24850	At2g37730	GT-family-31
Glyma01g26310	At1g11545	Xyloglucan
Glyma01g26766	At1g61900	Unknow-GPI-anchored
Glyma01g26783	At3g09220	Laccases
Glyma01g26800	At3g09220	Laccases
Glyma01g27230	At4g24550	Endocytosis
Glyma01g27260	At3g05610	Pectin-methyl-esterases
Glyma01g27710	At2g38080	Laccases
Glyma01g30900	At4g24550	Endocytosis
Glyma01g32270	At5g05340	Peroxidases
Glyma01g32310	At5g05340	Peroxidases
Glyma01g33440	At2g45220	Pectin-methyl-esterases

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma01g33480	At2g45220	Pectin-methyl-esterases
Glyma01g33500	At2g45220	Pectin-methyl-esterases
Glyma01g34605	At2g01850	Xyloglucan
Glyma01g34770	At4g28850	Xyloglucan
Glyma01g35530	At5g17540	Pathways-phenyl
Glyma01g37060	At4g36010	Receptor-like
Glyma01g37630	At5g66390	Peroxidases
Glyma01g37930	At5g03260	Laccases
Glyma01g38540	At5g09280	Pectate-pectin-Lyases
Glyma01g39080	At1g49570	Peroxidases
Glyma01g39460	At4g34050	Pathways-phenyl
Glyma01g40460	At5g65730	Xyloglucan
Glyma01g41830	At4g16600	GT-family-8
Glyma01g42370	At1g12560	Expansins
Glyma01g44550	At4g14805	LPTL
Glyma0248s00200	At5g27870	Pectin-methyl-esterases
Glyma0248s00220	At3g05610	Pectin-methyl-esterases
Glyma02g01130	At3g62170	Pectin-methyl-esterases
Glyma02g01140	At3g05610	Pectin-methyl-esterases
Glyma02g01980	At1g02790	Polygalacturonases
Glyma02g02000	At4g02320	Pectin-methyl-esterases
Glyma02g03170	At1g59970	Proteases
Glyma02g03200	At1g70170	Proteases
Glyma02g03265	At1g59970	Proteases
Glyma02g03335	At1g70170	Proteases
Glyma02g04230	At3g26610	Polygalacturonases
Glyma02g04290	At4g08770	Peroxidases
Glyma02g04990	At5g66060	Prolyl-4-hydroxylases
Glyma02g05930	At5g66390	Peroxidases
Glyma02g05940	At5g66390	Peroxidases
Glyma02g07410	At2g19070	Pathways-phenyl
Glyma02g07740	At5g20710	Beta-galactosidase
Glyma02g07770	At5g20710	Beta-galactosidase
Glyma02g08374	At1g18140	Laccases
Glyma02g08780	At1g71695	Peroxidases
Glyma02g09535	At3g24130	Pectin-methyl-esterases
Glyma02g09540	At3g24130	Pectin-methyl-esterases
Glyma02g12920	At3g45400	GT-family-47
Glyma02g13820	At1g69940	Pectin-methyl-esterases
Glyma02g15280	At1g49570	Peroxidases
Glyma02g15290	At1g49570	Peroxidases
Glyma02g16460	At2g48140	LPTL
Glyma02g27131	At3g26610	Polygalacturonases

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma02g28880	At5g06720	Peroxidases
Glyma02g31540	At2g41850	Polygalacturonases
Glyma02g35400	At5g60920	COBRA
Glyma02g37210	At1g72230	Uclacyanin-like
Glyma02g37990	At1g70170	Proteases
Glyma02g38980	At5g48140	Polygalacturonases
Glyma02g40000	At5g05340	Peroxidases
Glyma02g40010	At5g05340	Peroxidases
Glyma02g40040	At5g58400	Peroxidases
Glyma02g40230	At2g40610	Expansins
Glyma02g42730	At5g05340	Peroxidases
Glyma02g45560	At4g38190	Cellulose-synthase-like
Glyma02g45670	At3g23730	Xyloglucan
Glyma02g46320	At2g32990	Glycoside-hidrolase-family-9
Glyma02g46400	At2g47280	Pectin-methyl-esterases
Glyma02g46880	At5g55590	Pectin-methyl-esterases
Glyma02g47083	At1g55850	Cellulose-synthase-like
Glyma02g47410	At3g25140	GT-family-8
Glyma0335s00200	At4g35160	Pathways-phenyl
Glyma03g00910	At3g45400	GT-family-47
Glyma03g01010	At4g26010	Peroxidases
Glyma03g01020	At4g26010	Peroxidases
Glyma03g03390	At2g45220	Pectin-methyl-esterases
Glyma03g03400	At2g45220	Pectin-methyl-esterases
Glyma03g03410	At2g45220	Pectin-methyl-esterases
Glyma03g03980	At1g65680	Expansins
Glyma03g04660	At5g05340	Peroxidases
Glyma03g04670	At5g05340	Peroxidases
Glyma03g04700	At5g05340	Peroxidases
Glyma03g04710	At5g05340	Peroxidases
Glyma03g04720	At5g05340	Peroxidases
Glyma03g04740	At5g05340	Peroxidases
Glyma03g04750	At5g05340	Peroxidases
Glyma03g04760	At5g05340	Peroxidases
Glyma03g04870	At5g05340	Peroxidases
Glyma03g04880	At5g05340	Peroxidases
Glyma03g08040	At1g75680	Glycoside-hidrolase-family-9
Glyma03g08190	At4g36360	Beta-galactosidase
Glyma03g08544	At3g03050	Cellulose-synthase-like
Glyma03g08748	At3g11420	GT-family-31
Glyma03g14450	At2g38080	Laccases
Glyma03g15805	At3g09220	Laccases
Glyma03g23680	At5g17200	Polygalacturonases

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma03g23701	At5g17200	Polygalacturonases
Glyma03g23881	At3g15720	Polygalacturonases
Glyma03g27210	At2g45040	Proteases
Glyma03g28870	At3g57270	Glycoside-hidrolase-family-17
Glyma03g29420	At2g41850	Polygalacturonases
Glyma03g29570	At3g57630	GT-family-47
Glyma03g30180	At5g06720	Peroxidases
Glyma03g30923	At3g21900	Receptor-like-kinases
Glyma03g31240	At3g22600	LPTL
Glyma03g31251	At2g48140	LPTL
Glyma03g33720	At5g03170	Arabinogalactan-proteins
Glyma03g33730	At5g03170	Arabinogalactan-proteins
Glyma03g36610	At4g11290	Peroxidases
Glyma03g36620	At1g05260	Peroxidases
Glyma03g37400	At4g02330	Pectin-methyl-esterases
Glyma03g37420	At4g02290	Glycoside-hidrolase-family-9
Glyma03g37560	At3g62660	GT-family-8
Glyma03g38140	At1g02460	Polygalacturonases
Glyma03g38230	At5g27870	Pectin-methyl-esterases
Glyma03g38480	At4g01630	Expansins
Glyma03g38750	At3g05610	Pectin-methyl-esterases
Glyma03g40420	At5g17540	Pathways-phenyl
Glyma03g40430	At5g17540	Pathways-phenyl
Glyma0420s50	At1g70170	Proteases
Glyma04g00520	At2g16730	Beta-galactosidase
Glyma04g00800	At3g29810	COBRA
Glyma04g07890	At2g45470	Arabinogalactan-proteins
Glyma04g08870	At4g32790	GT-family-47
Glyma04g09660	At1g09910	Rhamnogalacturonan
Glyma04g09670	At1g09890	Rhamnogalacturonan
Glyma04g10670	At1g72230	Uclacyanin-like
Glyma04g10761	At4g21960	Peroxidases
Glyma04g12290	At1g71380	Glycoside-hidrolase-family-9
Glyma04g13613	At5g51490	Pectin-methyl-esterases
Glyma04g13626	At2g45220	Pectin-methyl-esterases
Glyma04g14301	At4g36360	Beta-galactosidase
Glyma04g14460	At5g11710	Endocytosis
Glyma04g15190	At4g26140	Beta-galactosidase
Glyma04g15200	At2g04780	Arabinogalactan-proteins
Glyma04g22190	At3g24330	Glycoside-hidrolase-family-17
Glyma04g23530	At5g17420	Cellulose-synthase

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma04g24793	At1g20510	Pathways-phenyl
Glyma04g30870	At3g07840	Polygalacturonases
Glyma04g30920	At3g07840	Polygalacturonases
Glyma04g30950	At3g07840	Polygalacturonases
Glyma04g33056	At4g23490	GT-family-31
Glyma04g33780	At1g45130	Beta-galactosidase
Glyma04g33861	At2g36710	Pectin-methyl-esterases
Glyma04g34241	At5g64740	Cellulose-synthase
Glyma04g34260	At5g15110	Pectate-pectin-Lyases
Glyma04g36520	At5g63180	Pectate-pectin-Lyases
Glyma04g37470	At5g41040	Pathways-phenyl
Glyma04g39120	At2g13680	Callose-synthase
Glyma04g39641	At5g64790	Glycoside-hidrolase-family-17
Glyma04g39860	At5g05340	Peroxidases
Glyma04g40591	At5g54160	Pathways-phenyl
Glyma04g41891	At5g49720	Glycoside-hidrolase-family-9
Glyma04g42620	At1g77410	Beta-galactosidase
Glyma04g43470	At4g24000	Cellulose-synthase-like
Glyma05g00950	At1g69530	Expansins
Glyma05g01051	At1g74680	GT-family-47
Glyma05g02890	At5g63180	Pectate-pectin-Lyases
Glyma05g04770	At3g18660	GT-family-8
Glyma05g04822	At1g78570	Nucleotide-sugar
Glyma05g07193	At5g24090	Yieldins
Glyma05g135	At5g60920	COBRA
Glyma05g14800	At3g27200	Uclacyanin-like
Glyma05g16140	At2g19790	Endocytosis
Glyma05g21081	At3g53190	Pectate-pectin-Lyases
Glyma05g23950	At3g27400	Pectate-pectin-Lyases
Glyma05g26390	At1g48100	Polygalacturonases
Glyma05g27960	At1g67980	Pathways-phenyl
Glyma05g29240	At4g18780	Cellulose-synthase
Glyma05g29430	At5g60490	Arabinogalactan-proteins
Glyma05g32841	At5g63800	Beta-galactosidase
Glyma05g34800	At1g23200	Pectin-methyl-esterases
Glyma05g34810	At1g23200	Pectin-methyl-esterases
Glyma05g35930	At5g19730	Pectin-methyl-esterases
Glyma05g35955	At5g56590	Glycoside-hidrolase-family-17
Glyma05g36220	At4g26220	Pathways-phenyl
Glyma05g37110	At2g25060	Early-nodulin-like
Glyma05g38290	At5g41040	Pathways-phenyl

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma06g03466	At4g19010	Pathways-phenyl
Glyma06g05950	At3g43860	Glycoside-hidrolase-family-9
Glyma06g06870	At4g18780	Cellulose-synthase
Glyma06g07040	At5g25310	GT-family-47
Glyma06g07940	At2g45470	Arabinogalactan-proteins
Glyma06g14210	At5g54160	Pathways-phenyl
Glyma06g17590	At5g41040	Pathways-phenyl
Glyma06g18370	At5g63180	Pectate-pectin-Lyases
Glyma06g20280	At4g13710	Pectate-pectin-Lyases
Glyma06g22030	At3g07820	Polygalacturonases
Glyma06g22110	At3g05910	Pectin-acetylerases
Glyma06g22430	At5g15630	COBRA
Glyma06g22890	At3g07840	Polygalacturonases
Glyma06g23474	At3g24330	Glycoside-hidrolase-family-17
Glyma06g23530	At1g03390	Pathways-phenyl
Glyma06g28630	At1g75680	Glycoside-hidrolase-family-9
Glyma06g28650	At5g20230	Stellacyanin-like
Glyma06g28890	At2g41480	Peroxidases
Glyma06g30850	At5g17420	Cellulose-synthase
Glyma06g30860	At5g17420	Cellulose-synthase
Glyma06g36860	At3g03050	Cellulose-synthase-like
Glyma06g41461	At3g07160	Callose-synthase
Glyma06g41630	At5g54690	GT-family-8
Glyma06g42850	At5g05340	Peroxidases
Glyma06g43010	At1g23210	Glycoside-hidrolase-family-9
Glyma06g43020	At1g70710	Glycoside-hidrolase-family-9
Glyma06g43961	At5g03760	Cellulose-synthase-like
Glyma06g43970	At4g35160	Pathways-phenyl
Glyma06g44550	At5g53870	Early-nodulin-like
Glyma06g44940	At4g38210	Expansins
Glyma06g45840	At4g28280	Glycosylphosphatidylinositol
Glyma06g45910	At1g05260	Peroxidases
Glyma06g46161	At1g61900	Unknow-GPI-anchored
Glyma06g47190	At3g10710	Pectin-methyl-esterases
Glyma06g47200	At5g04970	Pectin-methyl-esterases
Glyma06g47690	At2g45220	Pectin-methyl-esterases
Glyma06g47710	At5g51500	Pectin-methyl-esterases
Glyma07g02540	At2g03220	GT-family-37
Glyma07g02550	At2g03220	GT-family-37
Glyma07g02750	At5g27870	Pectin-methyl-esterases
Glyma07g02780	At5g27870	Pectin-methyl-esterases
Glyma07g02790	At5g27870	Pectin-methyl-esterases
Glyma07g03010	At5g27870	Pectin-methyl-esterases

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma07g05480	At5g54160	Pathways-phenyl
Glyma07g05970	At2g46570	Laccases
Glyma07g06773	At5g24090	Yieldins
Glyma07g07180	At4g25810	Xyloglucan
Glyma07g07290	At3g61490	Polygalacturonases
Glyma07g08660	At1g09580	Vesicle-trafficking
Glyma07g12321	At3g15720	Polygalacturonases
Glyma07g14234	At1g65060	Pathways-phenyl
Glyma07g14670	At2g45040	Proteases
Glyma07g16080	At5g60020	Laccases
Glyma07g17140	At3g09220	Laccases
Glyma07g17150	At5g01050	Laccases
Glyma07g17170	At5g01050	Laccases
Glyma07g17646	At5g01190	Laccases
Glyma07g19647	At5g66060	Prolyl-4-hydroxylases
Glyma07g26640	At2g31960	Callose-synthase
Glyma07g27450	At3g24130	Pectin-methyl-esterases
Glyma07g28040	At4g37990	Pathways-phenyl
Glyma07g32283	At4g24010	Cellulose-synthase-like
Glyma07g32471	At2g34060	Peroxidases
Glyma07g33180	At1g49570	Peroxidases
Glyma07g33761	At5g05170	Cellulose-synthase
Glyma07g34990	At2g43870	Polygalacturonases
Glyma07g36580	At4g16270	Peroxidases
Glyma07g37320	At3g06770	Polygalacturonases
Glyma07g37440	At4g18180	Polygalacturonases
Glyma07g37460	At3g05610	Pectin-methyl-esterases
Glyma08g00470	At5g63800	Beta-galactosidase
Glyma08g01360	At5g41040	Pathways-phenyl
Glyma08g04560	At1g77630	Receptor-like
Glyma08g04880	At1g23200	Pectin-methyl-esterases
Glyma08g05660	At3g18050	Unknow-GPI-anchored
Glyma08g09300	At1g48100	Polygalacturonases
Glyma08g10300	At2g42800	Receptor-like
Glyma08g12580	At5g60490	Arabinogalactan-proteins
Glyma08g12590	At5g60490	Arabinogalactan-proteins
Glyma08g12600	At5g03170	Arabinogalactan-proteins
Glyma08g15641	At3g18660	GT-family-8
Glyma08g19170	At5g64120	Peroxidases
Glyma08g19180	At2g41480	Peroxidases
Glyma08g19190	At5g64110	Peroxidases
Glyma08g19840	At1g09580	Vesicle-trafficking
Glyma08g25901	At5g17200	Polygalacturonases



**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma08g26540	At5g39270	Expansins
Glyma08g27070	At4g35160	Pathways-phenyl
Glyma08g27090	At4g35150	Pathways-phenyl
Glyma08g27110	At4g35150	Pathways-phenyl
Glyma08g27260	At4g35160	Pathways-phenyl
Glyma08g29110	At4g13990	GT-family-47
Glyma08g35016	At4g16790	Hydroxyproline-rich
Glyma08g37410	At2g22620	Rhamnogalacturonan
Glyma08g38430	At4g37980	Pathways-phenyl
Glyma08g40280	At5g40150	Peroxidases
Glyma08g40820	At1g14420	Pectate-pectin-Lyases
Glyma08g41100	At5g24090	Yieldins
Glyma08g42250	At4g14130	Xyloglucan
Glyma08g42440	At2g19070	Pathways-phenyl
Glyma08g42490	At2g19070	Pathways-phenyl
Glyma08g42500	At2g19070	Pathways-phenyl
Glyma08g46450	At4g13090	Xyloglucan
Glyma08g46495	At4g13090	Xyloglucan
Glyma08g47410	At5g60020	Laccases
Glyma09g00620	At2g47280	Pectin-methyl-esterases
Glyma09g00841	At1g11920	Pectate-pectin-Lyases
Glyma09g00850	At1g11920	Pectate-pectin-Lyases
Glyma09g02460	At2g43870	Polygalacturonases
Glyma09g02590	At5g06720	Peroxidases
Glyma09g02610	At5g06730	Peroxidases
Glyma09g02641	At2g38380	Peroxidases
Glyma09g02650	At5g06730	Peroxidases
Glyma09g02670	At5g06720	Peroxidases
Glyma09g02680	At2g38380	Peroxidases
Glyma09g03620	At1g48100	Polygalacturonases
Glyma09g03960	At3g17060	Pectin-methyl-esterases
Glyma09g04191	At4g16260	Glycoside-hidrolase-family-17
Glyma09g04200	At4g16260	Glycoside-hidrolase-family-17
Glyma09g04460	At4g16120	COBRA
Glyma09g04640	At1g02790	Polygalacturonases
Glyma09g04730	At5g27870	Pectin-methyl-esterases
Glyma09g05340	At1g30870	Peroxidases
Glyma09g07070	At4g25810	Xyloglucan
Glyma09g07281	At1g11545	Xyloglucan
Glyma09g08620	At5g28840	Nucleotide-sugar
Glyma09g08722	At4g18010	Endocytosis
Glyma09g08900	At5g20860	Pectin-methyl-esterases

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma09g08910	At2g26450	Pectin-methyl-esterases
Glyma09g08960	At3g43270	Pectin-methyl-esterases
Glyma09g10500	At2g43890	Polygalacturonases
Glyma09g11615	At2g28410	Glycosylphosphatidylinositol
Glyma09g12440	At4g35160	Pathways-phenyl
Glyma09g12480	At4g35150	Pathways-phenyl
Glyma09g15350	At5g28840	Nucleotide-sugar
Glyma09g16810	At5g06720	Peroxidases
Glyma09g21100	At2g33100	Cellulose-synthase-like
Glyma09g24470	At3g61490	Polygalacturonases
Glyma09g24590	At4g12420	SKU5
Glyma09g26905	At5g14650	Polygalacturonases
Glyma09g27390	At5g51890	Peroxidases
Glyma09g27804	At4g23310	Receptor-like-kinases
Glyma09g28460	At4g33420	Peroxidases
Glyma09g29566	At2g32300	Uclacyanin-like
Glyma09g30470	At1g18280	LPTL
Glyma09g31795	At4g36220	Pathways-phenyl
Glyma09g31850	At4g36220	Pathways-phenyl
Glyma09g32630	At4g28850	Xyloglucan
Glyma09g33371	At4g37990	Pathways-phenyl
Glyma09g33390	At4g39330	Pathways-phenyl
Glyma09g36620	At1g70710	Glycoside-hidrolase-family-9
Glyma09g36830	At1g33430	GT-family-31
Glyma09g37550	At3g01270	Pectate-pectin-Lyases
Glyma09g41440	At5g05340	Peroxidases
Glyma09g41450	At5g05340	Peroxidases
Glyma09g41840	At5g54160	Pathways-phenyl
Glyma09g42130	At4g11290	Peroxidases
Glyma09g42160	At4g11290	Peroxidases
Glyma10g01355	At5g19730	Pectin-methyl-esterases
Glyma10g02120	At1g02790	Polygalacturonases
Glyma10g02130	At4g02290	Glycoside-hidrolase-family-9
Glyma10g02140	At4g02320	Pectin-methyl-esterases
Glyma10g02730	At1g05260	Peroxidases
Glyma10g04450	At4g28560	Receptor-like
Glyma10g04530	At5g44030	Cellulose-synthase
Glyma10g05151	At4g12730	Arabinogalactan-proteins
Glyma10g05800	At2g37130	Peroxidases
Glyma10g06140	At2g36870	Xyloglucan
Glyma10g06200	At1g21090	Early-nodulin-like
Glyma10g06690	At1g67750	Pectate-pectin-Lyases
Glyma10g06830	At1g20030	Receptor-like

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma10g06940	At1g20030	Receptor-like
Glyma10g07320	At5g51500	Pectin-methyl-esterases
Glyma10g11480	At3g07970	Polygalacturonases
Glyma10g17550	At2g41850	Polygalacturonases
Glyma10g27700	At3g05610	Pectin-methyl-esterases
Glyma10g27710	At2g47030	Pectin-methyl-esterases
Glyma10g28040	At4g01630	Expansins
Glyma10g28753	At2g47280	Pectin-methyl-esterases
Glyma10g28921	At1g56320	Unknow-GPI-anchored
Glyma10g30110	At5g17540	Pathways-phenyl
Glyma10g30340	At4g14746	Unknow-GPI-anchored
Glyma10g31991	At4g35160	Pathways-phenyl
Glyma10g32020	At4g35160	Pathways-phenyl
Glyma10g32030	At4g35160	Pathways-phenyl
Glyma10g33520	At4g11290	Peroxidases
Glyma10g33720	At5g07475	Uclacyanin-like
Glyma10g33930	At2g31050	Stellacyanin-like
Glyma10g34161	At1g20510	Pathways-phenyl
Glyma10g34170	At1g20510	Pathways-phenyl
Glyma10g34190	At5g47000	Peroxidases
Glyma10g35980	At5g54160	Pathways-phenyl
Glyma10g36310	At5g48100	Laccases
Glyma10g37530	At3g61490	Polygalacturonases
Glyma10g37540	At3g48950	Polygalacturonases
Glyma10g39121	At4g26140	Beta-galactosidase
Glyma10g39200	At3g22600	LPTL
Glyma10g39690	At3g06035	Unknow-GPI-anchored
Glyma10g40520	At4g33910	Prolyl-4-hydroxylases
Glyma10g43518	At4g38140	Unknow-GPI-anchored
Glyma11g00560	At3g20580	COBRA
Glyma11g00700	At5g26330	Stellacyanin-like
Glyma11g02350	At4g23560	Glycoside-hidrolase-family-9
Glyma11g02640	At1g63450	GT-family-47
Glyma11g03000	At1g12560	Expansins
Glyma11g03310	At1g62790	LPTL
Glyma11g04820	At5g65730	Xyloglucan
Glyma11g05800	At4g34050	Pathways-phenyl
Glyma11g06180	At1g49570	Peroxidases
Glyma11g06761	At5g09280	Pectate-pectin-Lyases
Glyma11g07420	At5g03260	Laccases
Glyma11g07670	At5g66390	Peroxidases
Glyma11g08580	At4g35970	Peroxidases

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma11g09710	At1g65060	Pathways-phenyl
Glyma11g10070	At4g16260	Glycoside-hidrolase-family-17
Glyma11g10750	At1g14550	Peroxidases
Glyma11g11500	At2g16730	Beta-galactosidase
Glyma11g13901	At1g20480	Pathways-phenyl
Glyma11g14600	At2g29130	Laccases
Glyma11g14800	At2g39700	Expansins
Glyma11g15960	At5g03170	Arabinogalactan-proteins
Glyma11g15975	At5g03170	Arabinogalactan-proteins
Glyma11g15990	At5g03170	Arabinogalactan-proteins
Glyma11g16000	At5g03170	Arabinogalactan-proteins
Glyma11g16430	At4g35670	Polygalacturonases
Glyma11g17160	At1g65680	Expansins
Glyma11g18970	At2g27500	Glycoside-hidrolase-family-17
Glyma11g19090	At1g08200	Nucleotide-sugar
Glyma11g19920	At2g36870	Xyloglucan
Glyma11g20760	At5g03170	Arabinogalactan-proteins
Glyma11g20770	At5g03170	Arabinogalactan-proteins
Glyma11g20780	At5g03170	Arabinogalactan-proteins
Glyma11g20800	At5g03170	Arabinogalactan-proteins
Glyma11g20810	At5g03170	Arabinogalactan-proteins
Glyma11g20820	At5g60490	Arabinogalactan-proteins
Glyma11g21080	At5g54160	Pathways-phenyl
Glyma11g21790	At5g26330	Stellacyanin-like
Glyma11g29060	At2g19070	Pathways-phenyl
Glyma11g29070	At2g19070	Pathways-phenyl
Glyma11g29890	At5g05340	Peroxidases
Glyma11g29920	At5g05340	Peroxidases
Glyma11g30010	At5g58390	Peroxidases
Glyma11g31050	At2g35380	Peroxidases
Glyma11g33650	At1g66250	Glycoside-hidrolase-family-17
Glyma11g35700	At2g40370	Laccases
Glyma11g36058	At5g09360	Laccases
Glyma12g00630	At3g15720	Polygalacturonases
Glyma12g00740	At1g70710	Glycoside-hidrolase-family-9
Glyma12g00750	At1g70710	Glycoside-hidrolase-family-9
Glyma12g01480	At3g07970	Polygalacturonases
Glyma12g03050	At1g48930	Glycoside-hidrolase-family-9
Glyma12g04800	At2g16230	Glycoside-hidrolase-family-17
Glyma12g05856	At1g20480	Pathways-phenyl

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma12g06480	At2g29130	Laccases
Glyma12g07370	At5g03170	Arabinogalactan-proteins
Glyma12g07400	At5g03170	Arabinogalactan-proteins
Glyma12g07411	At5g03170	Arabinogalactan-proteins
Glyma12g07420	At5g03170	Arabinogalactan-proteins
Glyma12g07430	At5g03170	Arabinogalactan-proteins
Glyma12g07440	At5g03170	Arabinogalactan-proteins
Glyma12g07450	At5g03170	Arabinogalactan-proteins
Glyma12g07460	At5g03170	Arabinogalactan-proteins
Glyma12g07490	At5g03170	Arabinogalactan-proteins
Glyma12g07500	At2g28470	Beta-galactosidase
Glyma12g08520	At2g36870	Xyloglucan
Glyma12g10300	At2g32530	Cellulose-synthase-like
Glyma12g10841	At1g05260	Peroxidases
Glyma12g10850	At1g05260	Peroxidases
Glyma12g12340	At4g38210	Expansins
Glyma12g12650	At4g28280	Glycosylphosphatidylinositol
Glyma12g13130	At5g53870	Early-nodulin-like
Glyma12g13980	At4g35160	Pathways-phenyl
Glyma12g14230	At2g38080	Laccases
Glyma12g15073	At3g08530	Endocytosis
Glyma12g15460	At5g05340	Peroxidases
Glyma12g16120	At5g05340	Peroxidases
Glyma12g16470	At4g12420	SKU5
Glyma12g16550	At5g54690	GT-family-8
Glyma12g17730	At5g17420	Cellulose-synthase
Glyma12g20860	At5g57500	GT-family-31
Glyma12g22745	At2g28470	Beta-galactosidase
Glyma12g27050	At3g05910	Pectin-acetylsterases
Glyma12g27175	At2g27860	Nucleotide-sugar
Glyma12g29530	At2g28410	Glycosylphosphatidylinositol
Glyma12g29670	At5g03170	Arabinogalactan-proteins
Glyma12g30200	At2g36870	Xyloglucan
Glyma12g32170	At4g11290	Peroxidases
Glyma12g33070	At4g28250	Expansins
Glyma12g33940	At5g05340	Peroxidases
Glyma13g00790	At2g18980	Peroxidases
Glyma13g03185	At1g21880	Receptor-like
Glyma13g03730	At5g26670	Pectin-acetylsterases
Glyma13g04221	At2g19070	Pathways-phenyl
Glyma13g05150	At1g14420	Pectate-pectin-Lyases
Glyma13g06660	At5g60920	COBRA
Glyma13g06671	At5g15630	COBRA

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma13g08400	At2g46330	Arabinogalactan-proteins
Glyma13g10600	At5g14180	Proteases
Glyma13g10610	At5g14180	Proteases
Glyma13g13383	At1g05020	Endocytosis
Glyma13g16590	At5g19890	Peroxidases
Glyma13g17240	At5g20710	Beta-galactosidase
Glyma13g17390	At5g61680	Pectin-methyl-esterases
Glyma13g17550	At2g26450	Pectin-methyl-esterases
Glyma13g18780	At5g44030	Cellulose-synthase
Glyma13g20170	At2g37130	Peroxidases
Glyma13g21880	At5g09280	Pectate-pectin-Lyases
Glyma13g23000	At5g20260	GT-family-47
Glyma13g23031	At3g42180	GT-family-47
Glyma13g23620	At2g41480	Peroxidases
Glyma13g24210	At4g35160	Pathways-phenyl
Glyma13g24225	At1g76790	Pathways-phenyl
Glyma13g25550	At3g10720	Pectin-methyl-esterases
Glyma13g25560	At3g10710	Pectin-methyl-esterases
Glyma13g29080	At2g42800	Receptor-like
Glyma13g29790	At5g03170	Arabinogalactan-proteins
Glyma13g29800	At5g03170	Arabinogalactan-proteins
Glyma13g37160	At4g28365	Early-nodulin-like
Glyma13g37395	At4g28250	Expansins
Glyma13g38200	At1g05280	GT-family-31
Glyma13g38230	At4g15240	GT-family-31
Glyma13g38300	At4g11290	Peroxidases
Glyma13g38650	At2g32530	Cellulose-synthase-like
Glyma13g39700	At2g28110	GT-family-47
Glyma13g39710	At2g36870	Xyloglucan
Glyma13g40110	At2g28410	Glycosylphosphatidylinositol
Glyma13g40210	At5g03170	Arabinogalactan-proteins
Glyma13g40220	At5g03170	Arabinogalactan-proteins
Glyma13g40540	At1g07850	GT-family-31
Glyma13g40920	At4g39350	Cellulose-synthase
Glyma13g41310	At2g29130	Laccases
Glyma13g41761	At1g20480	Pathways-phenyl
Glyma13g43190	At1g79800	Early-nodulin-like
Glyma13g44060	At5g62200	Unknow-GPI-anchored
Glyma13g44140	At5g17200	Polygalacturonases
Glyma13g44700	At1g80820	Pathways-phenyl
Glyma13g44921	At2g24450	Arabinogalactan-proteins
Glyma14g01211	At3g25140	GT-family-8
Glyma14g01670	At1g55850	Cellulose-synthase-like

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma14g01670	At1g55850	Cellulose-synthase-like
Glyma14g01820	At5g55590	Pectin-methyl-esterases
Glyma14g02340	At2g32990	Glycoside-hidrolase-family-9
Glyma14g02385	At2g47280	Pectin-methyl-esterases
Glyma14g03140	At3g23730	Xyloglucan
Glyma14g03310	At4g38190	Cellulose-synthase-like
Glyma14g04555	At1g20030	Receptor-like
Glyma14g04850	At3g15720	Polygalacturonases
Glyma14g05200	At2g32990	Glycoside-hidrolase-family-9
Glyma14g05510	At5g14180	Proteases
Glyma14g05840	At5g05340	Peroxidases
Glyma14g06760	At2g38080	Laccases
Glyma14g07730	At1g44970	Peroxidases
Glyma14g09070	At4g33330	GT-family-8
Glyma14g12560	At3g13750	Beta-galactosidase
Glyma14g14025	At5g25310	GT-family-47
Glyma14g14161	At3g29360	Nucleotide-sugar
Glyma14g17200	At2g24450	Arabinogalactan-proteins
Glyma14g17880	At1g78570	Nucleotide-sugar
Glyma14g23621	At4g18180	Polygalacturonases
Glyma14g24151	At1g02460	Polygalacturonases
Glyma14g28370	At1g77130	GT-family-8
Glyma14g31650	At2g46330	Arabinogalactan-proteins
Glyma14g35530	At1g72230	Uclacyanin-like
Glyma14g37030	At3g07820	Polygalacturonases
Glyma14g37040	At2g38080	Laccases
Glyma14g37810	At2g30210	Laccases
Glyma14g38080	At5g54160	Pathways-phenyl
Glyma14g38090	At4g35160	Pathways-phenyl
Glyma14g38150	At5g05340	Peroxidases
Glyma14g38593	At1g56320	Unknow-GPI-anchored
Glyma14g40150	At5g42180	Peroxidases
Glyma15g00400	At3g05620	Pectin-methyl-esterases
Glyma15g00490	At5g48930	Pathways-phenyl
Glyma15g01166	At5g17200	Polygalacturonases
Glyma15g03641	At1g20480	Pathways-phenyl
Glyma15g04860	At1g07850	GT-family-31
Glyma15g05810	At2g41480	Peroxidases
Glyma15g05820	At2g41480	Peroxidases
Glyma15g05831	At5g64120	Peroxidases
Glyma15g06131	At5g64310	Arabinogalactan-proteins
Glyma15g06366	At4g38040	GT-family-47
Glyma15g06373	At4g38040	GT-family-47

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma15g09240	At5g60490	Arabinogalactan-proteins
Glyma15g09250	At5g03170	Arabinogalactan-proteins
Glyma15g09970	At2g42800	Receptor-like
Glyma15g11700	At1g11920	Pectate-pectin-Lyases
Glyma15g13360	At2g43870	Polygalacturonases
Glyma15g13510	At5g06730	Peroxidases
Glyma15g13550	At2g38380	Peroxidases
Glyma15g15211	At4g16260	Glycoside-hidrolase-family-17
Glyma15g16141	At5g49180	Pectin-methyl-esterases
Glyma15g16240	At1g02790	Polygalacturonases
Glyma15g16251	At1g02790	Polygalacturonases
Glyma15g16710	At1g30870	Peroxidases
Glyma15g16900	At5g44030	Cellulose-synthase
Glyma15g20460	At4g33230	Pectin-methyl-esterases
Glyma15g20470	At2g26440	Pectin-methyl-esterases
Glyma15g20500	At2g26440	Pectin-methyl-esterases
Glyma15g20530	At3g43270	Pectin-methyl-esterases
Glyma15g21150	At4g36360	Beta-galactosidase
Glyma15g21530	At3g28200	Peroxidases
Glyma15g23310	At2g43880	Polygalacturonases
Glyma15g25450	At3g08530	Endocytosis
Glyma15g26350	At1g05170	GT-family-31
Glyma15g34690	At1g05260	Peroxidases
Glyma15g35390	At3g10710	Pectin-methyl-esterases
Glyma15g35931	At2g32810	Beta-galactosidase
Glyma15g37021	At2g13680	Callose-synthase
Glyma15g38520	At5g47780	GT-family-8
Glyma15g39210	At1g30870	Peroxidases
Glyma15g39573	At1g26690	Vesicle-trafficking
Glyma15g40151	At3g62000	Pathways-phenyl
Glyma15g40431	At5g20230	Stellacyanin-like
Glyma15g42420	At2g33160	Polygalacturonases
Glyma15g43080	At3g07970	Polygalacturonases
Glyma1655s00200	At2g41480	Peroxidases
Glyma16g04960	At5g13870	Xyloglucan
Glyma16g05320	At5g20710	Beta-galactosidase
Glyma16g05930	At5g63500	Glycosylphosphatidylinositol
Glyma16g06030	At5g14130	Peroxidases
Glyma16g06241	At3g27200	Uclacyanin-like
Glyma16g07421	At3g60730	Pectin-methyl-esterases
Glyma16g08230	At4g19420	Pectin-acetylerases
Glyma16g08975	At5g09870	Cellulose-synthase



**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma16g09405	At2g38650	GT-family-8
Glyma16g09471	At5g19730	Pectin-methyl-esterases
Glyma16g09490	At5g20710	Beta-galactosidase
Glyma16g17810	At3g05910	Pectin-acetylerases
Glyma16g19790	At4g37990	Pathways-phenyl
Glyma16g21150	At2g21770	Cellulose-synthase
Glyma16g21640	At4g16260	Glycoside-hidrolase-family-17
Glyma16g21700	At4g16260	Glycoside-hidrolase-family-17
Glyma16g21710	At4g16260	Glycoside-hidrolase-family-17
Glyma16g21740	At4g16260	Glycoside-hidrolase-family-17
Glyma16g23930	At1g75680	Glycoside-hidrolase-family-9
Glyma16g24610	At5g66390	Peroxidases
Glyma16g25710	At5g09280	Pectate-pectin-Lyases
Glyma16g26400	At2g19070	Pathways-phenyl
Glyma16g27480	At5g48100	Laccases
Glyma16g28090	At1g21090	Early-nodulin-like
Glyma16g29720	At1g07850	GT-family-31
Glyma16g30420	At2g44300	LPTL
Glyma16g30531	At2g44300	LPTL
Glyma16g30616	At2g44300	LPTL
Glyma16g30845	At2g44300	LPTL
Glyma16g30901	At2g44300	LPTL
Glyma16g31040	At1g55260	LPTL
Glyma16g31220	At2g44300	LPTL
Glyma16g31571	At1g55260	LPTL
Glyma16g31780	At2g44300	LPTL
Glyma16g31840	At2g44300	LPTL
Glyma16g32670	At5g17540	Pathways-phenyl
Glyma16g32680	At4g23310	Receptor-like-kinases
Glyma16g32715	At5g62200	Unknow-GPI-anchored
Glyma16g32720	At5g17540	Pathways-phenyl
Glyma16g32726	At4g23310	Receptor-like-kinases
Glyma16g33250	At4g33420	Peroxidases
Glyma16g34140	At2g32300	Uclacyanin-like
Glyma17g01440	At3g17070	Peroxidases
Glyma17g01470	At1g11920	Pectate-pectin-Lyases
Glyma17g03170	At5g27870	Pectin-methyl-esterases
Glyma17g03191	At3g07820	Polygalacturonases
Glyma17g03300	At3g06770	Polygalacturonases

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma17g04900	At5g20870	Glycoside-hidrolase-family-17
Glyma17g04960	At2g26450	Pectin-methyl-esterases
Glyma17g05250	At5g20710	Beta-galactosidase
Glyma17g06080	At5g19890	Peroxidases
Glyma17g06090	At5g19890	Peroxidases
Glyma17g06280	At3g13750	Beta-galactosidase
Glyma17g06350	At4g25810	Xyloglucan
Glyma17g06850	At2g19070	Pathways-phenyl
Glyma17g06860	At2g19070	Pathways-phenyl
Glyma17g06890	At2g18980	Peroxidases
Glyma17g07180	At3g21240	Pathways-phenyl
Glyma17g07240	At4g25810	Xyloglucan
Glyma17g08491	At3g29360	Nucleotide-sugar
Glyma17g10840	At1g74680	GT-family-47
Glyma17g10980	At2g03220	GT-family-37
Glyma17g11860	At5g33290	GT-family-47
Glyma17g11880	At5g20260	GT-family-47
Glyma17g12160	At5g53870	Early-nodulin-like
Glyma17g13550	At5g63180	Pectate-pectin-Lyases
Glyma17g14230	At1g12560	Expansins
Glyma17g15070	At5g47500	Pectin-methyl-esterases
Glyma17g16380	At4g18010	Endocytosis
Glyma17g16795	At5g54160	Pathways-phenyl
Glyma17g16890	At5g65730	Xyloglucan
Glyma17g17960	At4g12420	SKU5
Glyma17g20450	At1g49570	Peroxidases
Glyma17g24815	At5g04960	Pectin-methyl-esterases
Glyma17g29120	At1g78570	Nucleotide-sugar
Glyma17g29710	At2g24450	Arabinogalactan-proteins
Glyma17g29820	At3g13560	Glycoside-hidrolase-family-17
Glyma17g31721	At4g35670	Polygalacturonases
Glyma17g32061	At4g38140	Unknow-GPI-anchored
Glyma17g32140	At5g25310	GT-family-47
Glyma17g36100	At4g33330	GT-family-8
Glyma17g37980	At5g42180	Peroxidases
Glyma18g02520	At5g58400	Peroxidases
Glyma18g03470	At1g02800	Glycoside-hidrolase-family-9
Glyma18g04260	At2g37640	Expansins
Glyma18g04560	At1g66250	Glycoside-hidrolase-family-17
Glyma18g04921	At5g03800	GT-family-47
Glyma18g06210	At5g58390	Peroxidases

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma18g06220	At5g05340	Peroxidases
Glyma18g06230	At5g05340	Peroxidases
Glyma18g06250	At5g05340	Peroxidases
Glyma18g06510	At5g58490	Pathways-phenyl
Glyma18g08550	At1g62940	Pathways-phenyl
Glyma18g09516	At5g55480	GPDL
Glyma18g11381	At4g39350	Cellulose-synthase
Glyma18g12320	At2g19070	Pathways-phenyl
Glyma18g12660	At1g78570	Nucleotide-sugar
Glyma18g12690	At4g14130	Xyloglucan
Glyma18g13840	At2g19070	Pathways-phenyl
Glyma18g15310	At5g24090	Yieldins
Glyma18g15320	At5g24090	Yieldins
Glyma18g15330	At5g24090	Yieldins
Glyma18g15580	At5g17420	Cellulose-synthase
Glyma18g15606	At4g11350	GT-family-31
Glyma18g17202	At1g11545	Xyloglucan
Glyma18g17410	At5g40150	Peroxidases
Glyma18g18895	At3g07820	Polygalacturonases
Glyma18g18931	At1g10550	Xyloglucan
Glyma18g19586	At1g23460	Polygalacturonases
Glyma18g21870	At1g12780	Nucleotide-sugar
Glyma18g22430	At3g07840	Polygalacturonases
Glyma18g35720	At4g13090	Xyloglucan
Glyma18g35851	At4g26690	GPDL
Glyma18g38661	At5g60020	Laccases
Glyma18g38670	At4g37990	Pathways-phenyl
Glyma18g38690	At5g60020	Laccases
Glyma18g38700	At5g60020	Laccases
Glyma18g38710	At5g60020	Laccases
Glyma18g40050	At5g60020	Laccases
Glyma18g40070	At5g60020	Laccases
Glyma18g40150	At5g26670	Pectin-acetylerases
Glyma18g41846	At5g11710	Endocytosis
Glyma18g41860	At3g09220	Laccases
Glyma18g41870	At5g01040	Laccases
Glyma18g42520	At2g38080	Laccases
Glyma18g44310	At5g05340	Peroxidases
Glyma18g44320	At5g05340	Peroxidases
Glyma18g45175	At4g23310	Receptor-like-kinases
Glyma18g45190	At4g23310	Receptor-like-kinases
Glyma18g47081	At4g23490	GT-family-31
Glyma18g49700	At1g26690	Vesicle-trafficking

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma18g50030	At5g39280	Expansins
Glyma18g50260	At4g35150	Pathways-phenyl
Glyma18g50280	At4g35150	Pathways-phenyl
Glyma18g50290	At4g35150	Pathways-phenyl
Glyma18g50470	At4g35150	Pathways-phenyl
Glyma18g50750	At5g60920	COBRA
Glyma18g52000	At4g13990	GT-family-47
Glyma18g52010	At2g29040	GT-family-47
Glyma19g00230	At5g14650	Polygalacturonases
Glyma19g01720	At1g59970	Proteases
Glyma19g01730	At1g59970	Proteases
Glyma19g01740	At1g59970	Proteases
Glyma19g01760	At1g59970	Proteases
Glyma19g01910	At1g70090	GT-family-8
Glyma19g02390	At1g14420	Pectate-pectin-Lyases
Glyma19g02810	At2g03090	Expansins
Glyma19g03050	At3g29090	Pectin-methyl-esterases
Glyma19g07460	At5g14380	Arabinogalactan-proteins
Glyma19g07490	At3g27400	Pectate-pectin-Lyases
Glyma19g07512	At4g12420	SKU5
Glyma19g16960	At4g26010	Peroxidases
Glyma19g17373	At1g51990	Pathways-phenyl
Glyma19g18774	At3g27200	Uclacyanin-like
Glyma19g18820	At5g14180	Proteases
Glyma19g22510	At5g24090	Yieldins
Glyma19g24335	At2g13680	Callose-synthase
Glyma19g26240	At5g63500	Glycosylphosphatidylinositol
Glyma19g26660	At5g41040	Pathways-phenyl
Glyma19g28600	At2g19440	Glycoside-hidrolase-family-17
Glyma19g29621	At5g17820	Peroxidases
Glyma19g31580	At3g57270	Glycoside-hidrolase-family-17
Glyma19g31590	At3g57270	Glycoside-hidrolase-family-17
Glyma19g32241	At2g41850	Polygalacturonases
Glyma19g32760	At1g05310	Pectin-methyl-esterases
Glyma19g33080	At5g06720	Peroxidases
Glyma19g33750	At3g21900	Receptor-like-kinases
Glyma19g34090	At3g22600	LPTL
Glyma19g34100	At2g48140	LPTL
Glyma19g36470	At5g03170	Arabinogalactan-proteins
Glyma19g36870	At2g36870	Xyloglucan
Glyma19g37060	At3g55500	Expansins

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma19g39270	At2g39040	Peroxidases
Glyma19g39870	At3g62830	Nucleotide-sugar
Glyma19g40000	At4g02330	Pectin-methyl-esterases
Glyma19g40010	At1g02810	Pectin-methyl-esterases
Glyma19g40180	At3g62660	GT-family-8
Glyma19g40740	At1g02460	Polygalacturonases
Glyma19g40840	At2g47030	Pectin-methyl-esterases
Glyma19g41371	At4g26830	Glycoside-hidrolase-family-17
Glyma19g41830	At4g13090	Xyloglucan
Glyma19g41960	At4g02330	Pectin-methyl-esterases
Glyma19g42380	At1g54940	GT-family-8
Glyma19g43110	At5g17540	Pathways-phenyl
Glyma20g00330	At4g11290	Peroxidases
Glyma20g00540	At2g44540	Glycoside-hidrolase-family-9
Glyma20g00590	At5g54160	Pathways-phenyl
Glyma20g00604	At5g54160	Pathways-phenyl
Glyma20g02240	At4g18340	Glycoside-hidrolase-family-17
Glyma20g02840	At2g43870	Polygalacturonases
Glyma20g03221	At5g39320	Nucleotide-sugar
Glyma20g04430	At2g35380	Peroxidases
Glyma20g04586	At1g05840	Proteases
Glyma20g06820	At1g65610	Glycoside-hidrolase-family-9
Glyma20g08830	At2g19070	Pathways-phenyl
Glyma20g15800	At3g43860	Glycoside-hidrolase-family-9
Glyma20g16130	At5g14180	Proteases
Glyma20g16490	At1g64640	Early-nodulin-like
Glyma20g16521	At1g11545	Xyloglucan
Glyma20g18910	At1g20270	Prolyl-4-hydroxylases
Glyma20g19611	At5g20230	Stellacyanin-like
Glyma20g22050	At4g01630	Expansins
Glyma20g22210	At5g26670	Pectin-acetylerases
Glyma20g25251	At3g45860	Receptor-like-kinases
Glyma20g26800	At4g33910	Prolyl-4-hydroxylases
Glyma20g28490	At3g22600	LPTL
Glyma20g29930	At5g23870	Pectin-acetylerases
Glyma20g30240	At3g48950	Polygalacturonases
Glyma20g30480	At5g24090	Yieldins
Glyma20g31270	At5g48100	Laccases
Glyma20g31280	At5g48100	Laccases
Glyma20g31600	At5g54160	Pathways-phenyl
Glyma20g31700	At5g54160	Pathways-phenyl

**Tabela S1.** Cont.

GENE_ID	Homólogo/ Arabidopsis	Via de parede celular
Glyma20g32990	At5g03760	Cellulose-synthase-like
Glyma20g33340	At5g47000	Peroxidases
Glyma20g33361	At1g20510	Pathways-phenyl
Glyma20g33370	At1g20510	Pathways-phenyl
Glyma20g33460	At4g12420	SKU5
Glyma20g33670	At2g31050	Stellacyanin-like
Glyma20g33870	At5g07475	Uclacyanin-like
Glyma20g35620	At4g35160	Pathways-phenyl
Glyma20g36700	At4g14746	Unknow-GPI-anchored
Glyma20g38160	At3g05610	Pectin-methyl-esterases
Glyma20g38171	At1g02810	Pectin-methyl-esterases
Glyma20g38590	At5g05340	Peroxidases

**Tabela S2.** Os índices Ka Ks e o tempo estimado de divergência dos pares de genes COBRA parálogos em soja.

Gene	Duplicação	Ka	Ks	Ka/Ks	Mya
GmCOBL1a	GmCOBL1b	0,03	0,11	0,272	8,46
GmCOBL10b	GmCOBL10c	0,04	0,17	0,23	11,75
GmCOBL7a	GmCOBL7c	0,02	0,14	0,142	10,76
GmCOBL4a	GmCOBL4b	0,021	0,14	0,151	10,77
GmCOBL4c	GmCOBL4d	0,008	0,073	0,121	5,64
GmCOBb	GmCOBd	0,018	0,099	0,184	7,67
GmCOBc	GmCOBg	0,18	0,22	0,818	16,92
GmCOBf	GmCOBi	0,03	0,14	0,214	10,76