Insights into temperature modulation of the *Eucalyptus globulus* and *Eucalyptus grandis* antioxidant and lignification subproteomes

CrossMark

Marília Gabriela de Santana Costa [a], Paulo Mazzafera [b], Tiago Santana Balbuena [a, *]

[a] Department of Technology, São Paulo State University, Faculty of Agriculture and Veterinary Sciences, Jaboticabal, SP, Brazil
[b] Department of Plant Biology, Institute of Biology, University of Campinas, Campinas, SP, Brazil

ARTICLE INFO

ABSTRACT

Eucalyptus grandis and *Eucalyptus globulus* are among the most widely cultivated trees, differing in lignin composition and plantation areas, as *E. grandis* is mostly cultivated in tropical regions while *E. globulus* is preferred in temperate areas. As temperature is a key modulator in plant metabolism, a large-scale proteome analysis was carried out to investigate changes in the antioxidant system and the lignification metabolism in plantlets grown at different temperatures. Our strategy allowed the identification of 3111 stem proteins. A total of 103 antioxidant proteins were detected in the stems of both species. Hierarchical clustering revealed that alterations in the antioxidant proteins are more prominent when Eucalyptus seedlings were exposed to high temperature and that the superoxide isoforms coded by the gene Eucgr.B03930 are the most abundant antioxidant enzymes induced by thermal stimulus. Regarding the lignin biosynthesis, our proteomics approach resulted in the identification of 13 of the 17 core proteins involved in this metabolism, corroborating with gene predictions and the proposed lignin toolbox. Quantitative analyses revealed significant differences in 8 protein isoforms, including the ferulate 5-hydroxylase isoform F5H1, a key enzyme in catalyzing the synthesis of sinapyl alcohol, and the cinnamyl alcohol dehydrogenase isoform CAD2, the last enzyme in monolignol biosynthesis. Data are available via ProteomeXchange with identifier PXD005743.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

According to the Food and Agriculture Organization of the United Nations (FAO), planted forests correspond to a global area of approximately 264 mi hectares, most of them comprising Pinus and Eucalyptus species (Del Lungo et al., 2006). Within the economically important Eucalyptus plants, *Eucalyptus grandis* is probably one of the most planted and economically important species due to its fast growth rate and adaptability to different climate conditions. Recently, the landmark genome sequencing of a woody perennial organism was accomplished (Myburg et al., 2014) paving the way to explore the genetic information regarding the growth and the metabolic responses in *E. grandis* and in closely related species.

After water deficit, high and low temperatures are among the most influential stresses limiting plant productivity (Boyer, 1982;

Cramer et al., 2011). Although not representing the first line of responses against thermal stress, reactive oxygen species (ROS) generation is a common phenomenon used against plant abiotic stresses. As the production of ROS at high levels may negatively affect the cellular machinery, it has to be rapidly degraded to counteract potential damages induced by those unstable molecular species. Common enzymatic scavengers of plant ROS are well described and include the enzymes superoxide dismutase, ascorbate peroxidase, glutathione peroxidase and catalase (Apel and Hirt, 2004; Das and Roychoudhury, 2014; Foyer, 2005).

Due to their crucial impact on plant metabolism, temperature regions are a worldwide and locally limiting factor for crop plantations, including planted forests. Eucalyptus plant distribution is based on climate zones, where the species *E. grandis* and its hybrids are mostly cultivated in tropical regions and the species *Eucalyptus globulus* is preferred for cultivation in temperate areas (Potts, 2004; Dasgupta et al., 2015). As the latter is a cold tolerant hardwood tree with major economical importance, the molecular responses induced by low temperature have been explored. Using a qRT-PCR strategy, Fernández et al. (2010) observed a differential regulation

---

of the genes ELIP (early light-inducible protein), NCED (9-cis-epoxycarotenoid dioxygenase) and GS (galactinol synthase) in *E. globulus* plants exposed to cold. Gamboa et al. (2007) isolated a CBF (cold-binding factor) cDNA clone from *E. globulus* plants after exposure to low environmental temperatures, while Fernández et al. (2012) observed the upregulation of three dehydrin genes during cold acclimation. Although these targeted gene expression studies have been used to understand *E. globulus* responses to thermal stress, there are no reports on large scale technologies used to unveil this phenomenon.

In addition to the low-temperature tolerance trait, *E. globulus* plant cell walls differ from other *Eucalyptus* species as its lignin polymer presents a different composition in terms of the *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) subunits. As the ratio of S (doubly methylated, on the 3- and 5-hydroxyl moieties) to G (singly methylated, on the 3-hydroxyl group moiety) subunits dictates the degree of lignin condensation, the high S/G ratio in *E. globulus* lignin is an attractive feature for cell wall degradation during paper and pulp production (Araújo et al., 2014a, 2014b; Bhuiya and Liu, 2010; Araújo et al., 2014b). However, a deep understanding of all steps involved in Eucalyptus lignin biosynthesis is essential prior to any attempts to engineer plant cell walls in this genus. Recently, Carocha et al. (2015) carried out a deep analysis of the core genes involved in the lignin biosynthesis of the closely related species *E. grandis*, using RNA-seq high-throughput expression profiling combined with real-time quantitative PCR.

Due to its potential to provide insights into the proteins that were effectively decoded from the plant's genome, shotgun proteomics technologies were used here to study changes in the gene expression of *E. globulus* and *E. grandis* induced by cultivation in different temperatures, with a focus on profiling ROS-related proteins and lignin plant cell wall-related enzymes.

2. Results and discussion

2.1. Proteome data correlates with Eucalyptus grandis genomic architecture

Advances in large-scale nucleotide sequencing have tremendously increased the acquisition and the exploitation of plant proteome information. Although peptide *de novo* sequencing is a powerful and database-independent approach, fragment spectral mining against reference databases is still the most widely adopted strategy in the proteomics field. *Eucalyptus* genome sequencing initiatives were first launched in the early 2000s and set a milestone for forest tree biotechnology: the *E. grandis* draft genome data, a 640-megabase sequence, was assembled into 11 main chromosome linkage groups (Myburg et al., 2014). This high-quality reference is a valuable genetic resource for studies involving gene expression. As proteomics stands out from other large-scale approaches due to its capacity to target the final products of the genome information path, we firstly attempted to correlate protein abundance profile, obtained from gene encoded proteins identified herein (Tables 1S–30S), with the chromosome gene density pattern proposed by Myburg et al. (2014). The gene expression profile, in terms of the number of identified proteins, corroborated the gene density pattern for each of the 11 Eucalyptus chromosomes (Fig. 1).

The highest number of translated proteins, within a 1 megabase (Mb) genomic range, was detected in chromosomes 6 and 9, corroborating the high gene density suggested for upstream positions of those chromosomes. Additionally, the 1 Mb expression profile ranges from proteins encoded by chromosomes 3 and 5 also reflected the gene distribution suggested for those large but gene-poor chromosomes. Genomic positions of the genes coding for the
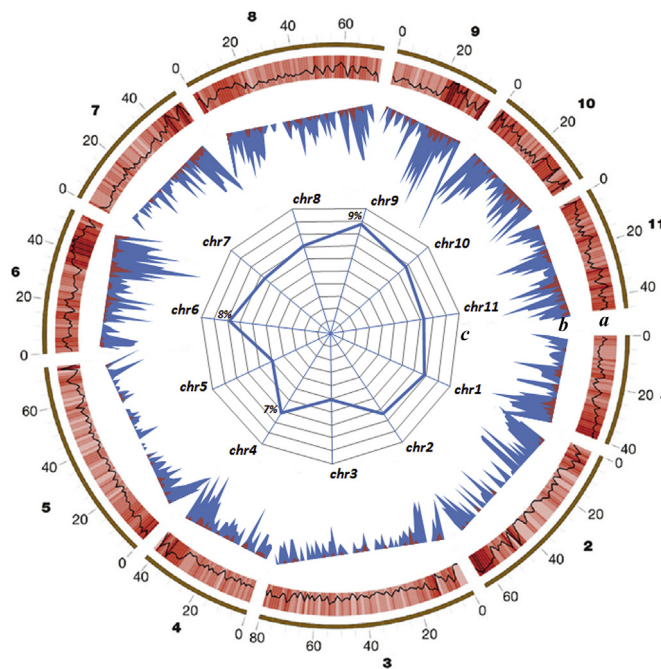

Fig. 1. Correlation between proteomics data and the *Eucalyptus grandis* genome structure as described by Myburg et al. (2014). Chromosome identities are indicated by the outermost numbers. **a**, gene density per 1 megabase (Mb) range; **b**, expressed gene density (number of identified proteins, blue peaks) and expressed gene variants (number of identified isoforms per locus, red peaks) per 1 Mb range; **c**, proteome-based genomic coverage for each proposed chromosome. Figure adapted from Myburg et al. (2014). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

protein isoforms are also in accordance with gene density patterns for all suggested chromosome linkage groups, where high gene density regions tend to express the highest number of genomic variants. Genomic coverage, in terms of identified gene products, reflected the gene distribution across *Eucalyptus* chromosomes. As a consequence of the physicochemical properties of proteins and the analytical strategies we used, the highest sequence coverage was achieved for chromosomes harboring the highest number of gene sequences (chromosomes 1, 9 and 10), while chromosomes 3 and 5 contained the lowest number of identified genes. Overall, proteomics-based gene expression analysis corroborated with the genomic architecture and the gene spatial distribution suggested by Myburg et al. (2014), providing proper support for comparative proteome analyses between the Eucalyptus species described here.

2.2. Eucalyptus antioxidant sub-proteome is regulated by thermal stimulus

Reactive oxygen species (ROS) are key elements in plant stress responses. Ascorbate peroxidase (APX), glyoxalase (GLX), glutathione peroxidase (GPX), peroxiredoxin (PRX), superoxide dismutase (SOD), thioredoxin (TRX) and catalase (CAT) are the most common antioxidant enzymes in plant cells (Foyer, 2005; Hasanuzzaman et al., 2012; Choudhury et al., 2013). In Eucalyptus species, *E. dunnii* plants grown at 4 °C showed an increase of 25% in CAT activity after 3 h of stress exposure (Liu et al., 2014). An increase in the expression of CAT was confirmed through a large-scale transcriptome analysis where the authors detected a two-fold increase in CAT-related transcripts after 24 h of treatment. Using shotgun proteomics coupled with stringent database mining against *E. grandis* genome coded proteins, we identified 103 ROS-

related proteins in *E. globulus* and *E. grandis* stems. Although not statistically significant, a combined analysis of the expression of all antioxidant-related proteins in *E. globulus* plantlets indicates a higher oxidative metabolism in plants grown at high temperature (HT) than in relation to those grown at low temperature (LT) (Fig. 2). Contrarily, no difference was found in *E. grandis* plants grown at different temperatures, suggesting that the *E. globulus* species may be more sensitive to those temperature changes than the close relative *E. grandis*. The increased expression of antioxidant enzymes detected in *E. globulus* stems grown at HT corroborates with the results of the study carried out by Macfarlane et al. (2002), from which a maximum growth rate of *E. globulus* seedlings was detected at 25 °C, a temperature in which an increase in the respiration rates is counterbalanced by a decrease in the carbon conversion efficiency.

Thioredoxin was found to be the enzyme with the highest number of identified protein isoforms (38) associated with a particular gene family in the *Eucalyptus* stem antioxidant sub-proteome (Table 31S). From all of the 38 TRX protein isoforms, 18 were identified as isoforms of only six genes, which include Eucgr.B02223, Eucgr.F02962, Eucgr.K02606, Eucgr.K03198, Eucgr.H02332, Eucgr.K03531. For the gene loci Eucgr.H02332 and Eucgr.K03531, four protein isoforms were identified; the latter was detected in all *Eucalyptus* samples and with a predicted localization in the reticulum endoplasmic compartment, which suggests a constitutive role of these proteins in the protein folding mechanism. Other proteins involved in redox processes were also identified in a large number of copies. Superoxide dismutase, an enzyme involved in reducing superoxide radicals, accounted for 18 protein species encoded by 7 different genes. The highest number of protein isoforms was identified for the gene Eucgr.K00110, from which encoded proteins were highly abundant in both Eucalyptus species and described as exclusively localized in the chloroplasts.

In silico sub-cellular predictors (*e.g.,* Predotar, WofPsort, PPDB, SUBA) indicated that 53 proteins were localized in the stem chloroplasts, a major cellular source of ROS in plants (Fig. 3). The hierarchical clustering of chloroplastic and mitochondrial antioxidant proteins suggests a closer proteome regulation profile between *E. globulus* plants exposed to LT and medium temperature (MT) than in relation to the HT sub-proteome. Although the regulation profile of the antioxidant mitochondrial proteome in *E. grandis* stems suggests a close relation across all temperature treatments, *E. grandis* plants exposed to LT and HT showed a closer sub-proteome expression profile compared to those grown at MT for the chloroplastic sub-proteome. These differential antioxidant responses between the *Eucalyptus* species are in accordance with the distinguishing use of these species in tropical or temperate plantation areas, suggesting the recruitment of characteristic antioxidant machinery when *E. globulus* and *E. grandis* plants are cultivated in higher or milder temperatures, respectively.

Superoxide dismutase (SOD) Eucgr.B03930, TRX Eucgr.K02606 and the GRXs Eucgr.I01764 and Eucgr.I01760, all genes encoding chloroplastic proteins, were those with the highest expression profiles constitutively expressed in all treatments in both *Eucalyptus* species (Fig. 3). Similarly, APX Eucgr.B02456, PRX Eucgr.B01112 and TRX Eucgr.A01555 genes, which encode mitochondrial proteins, were also found to be highly expressed in the Eucalyptus stems upon thermal treatments. As a general rule, protein isoforms of the mitochondrial and chloroplastic antioxidant system presented the same regulation profile across treatments. However, the protein SOD Eucgr.B03930.1 presented a specific abundance trend, being hierarchically clustered in a different group from the other isoforms from the gene Eucgr.B03930. While five chloroplastic isoforms derived from the gene Eucgr.B03930 were found to be highly accumulated in all thermal treatments in both *Eucalyptus* species, the variant Eucgr.B03930.1 was identified only in *E. grandis* stems cultivated at LT and HT. This unique regulation profile suggests the specific recruitment of this isoform to counterbalance superoxide damage in *E. grandis* chloroplasts after high thermal stimulus. Due to its high abundance in both *Eucalyptus* species, we carried out a dedicated comparative analysis of the Eucgr.B03930 isoforms. We found a very similar response profile for all protein isoforms, as indicated by the clustering, and significant differences in the abundance of Eucgr.B03930.2, Eucgr.B03930.3, Eucgr.B03930.4 and Eucgr.B03930.5 between *E. grandis* and *E. globulus* cultivated under HT treatment (Fig. 4), corroborating the hypothesis that *E. globulus* is more sensitive to a high temperature stimulus than *E. grandis* and that SOD gene products play an important role in the antioxidant response.

Peroxidases are also known to play an important role in the antioxidant metabolism. In the present work, we identified 30 peroxidase isoforms in the *Eucalyptus* stems: 14 APXs, 09 GPXs and 07 TPXs/PRXs. The importance of the antioxidant metabolism in *E. grandis* and *E. globulus* stems may be exemplified by the expression redundancy of the 14 APX isoforms found to be coded by only 07 genes (Eucgr.A01180, Eucgr.B02456, Eucgr.E00760, Eucgr.F00373, Eucgr.F04344, Eucgr.I01408 and Eucgr.J00965). Although most APX isoforms were identified in both species after all thermal treatments, isoform Eucgr.E00760.1 was only detected in *E. globulus* grown in high temperature. The isoform Eucgr.I01408.2 also showed a preferential expression in *E. globulus* stems but only in plants grown in the low temperature environment. Similarly, two GPX isoforms of the gene Eucgr.C02602 were not identified in *E. grandis* plants and were only identified in *E. globulus* stems after low temperature treatment.

Thioredoxins are small-sized proteins with the ability to reduce other protein species due to the presence of two cysteine residues in the canonical active site motif (Martí et al., 2009; Meyer et al., 2005; Pfalz et al., 2012; Goodstein et al., 2012). Interestingly, the Eucgr.C00774.1 and Eucgr.H00536.1 proteins were identified in both stems of *E. grandis* and *E. globulus* plants and annotated as nuclei TRX (Table 31S). Although the former did not show expression differences across the assayed thermal treatments and between the Eucalyptus species, the latter was over-induced in stems of *E. grandis* grown at LT and HT in relation to those cultivated at MT. Nuclei localization of thioredoxins was already observed in other plant species (Martí et al., 2009; Serrato et al., 2001; Pulido et al., 2009). However, the exact role of this of protein family within the
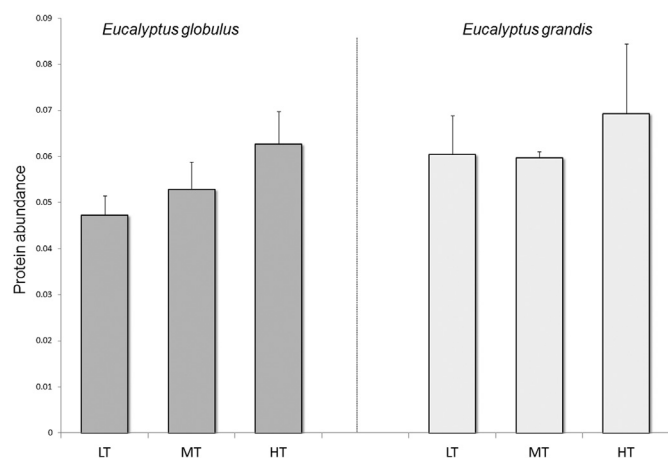

Fig. 2. Protein abundance (in terms of NSAF) of *Eucalyptus globulus* and *Eucalyptus grandis* antioxidant-related proteins.
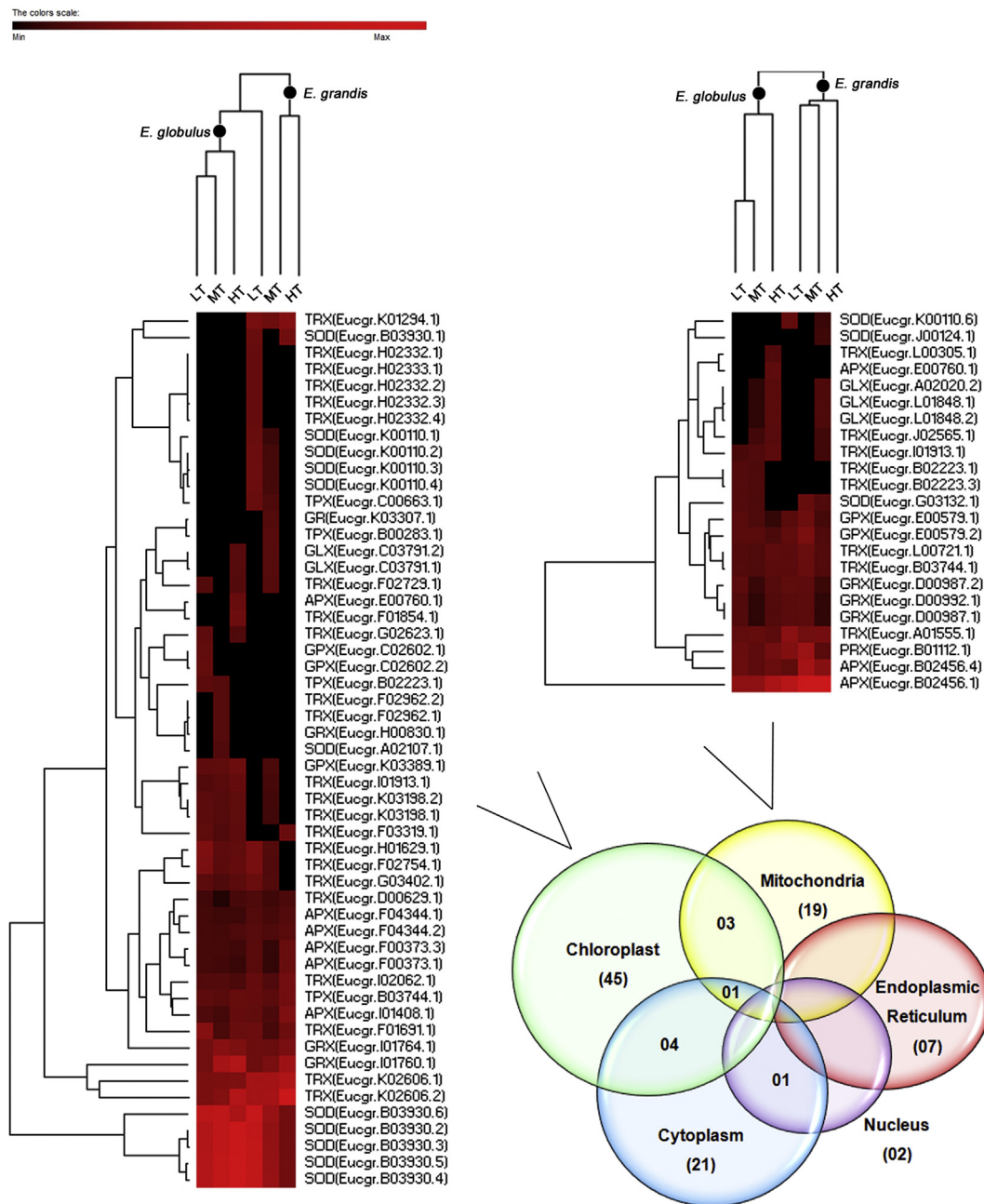
Fig. 3. Mitochondrial and chloroplastic antioxidant enzymes identified in *Eucalyptus grandis* and *Eucalyptus globulus* seedlings grown at LT, MT and HT as described in the Materials and methods section. Identified proteins were clustered according to their relative expression and their predicted sub-cellular localization. *In silico* predictors were used to identify the sub-cellular location of identified proteins.

plant nuclei remains unknown. Nuclei TRX may be directly involved in gene expression regulation or indirectly involved in the expression of specific genes by controlling ROS levels inside plant nuclei. Serrato and Cejudo (2003) observed a predominant nuclei subcellular localization of thioredoxins in wheat seeds stimulated by oxidative stress, while Pulido et al. (2009) identified TRX involvement in the regulation of transcription factors upon oxidative stress. Marchal et al. (2014) identified a novel nuclei TRX system involved in pollen fertility in *Arabidopsis* plants. Induction of nuclei TRX in *E. grandis* plants may be related to its higher adaptation to different climate conditions compared to the closely related species *E. globulus*. However, the nuclei signaling cascade induced by temperature stimulus and the target genes that may be regulated by this abiotic stress still deserve further study in Eucalyptus species.

2.3. Differences in the Eucalyptus lignification metabolism revealed by thermal stimulus

After cellulose, lignin is the most abundant organic compound on Earth (Boerjan et al., 2003). Lignin is synthesized by the oxidative polymerization of three precursors (*p*-coumaryl, coniferyl and sinapyl alcohols), which are also referred as lignin monomers or monolignols. Once these precursors are released in the cell wall, they become oxidized by laccases and peroxidases and then undergo oxidative combinatorial coupling (Cesarino et al., 2012).
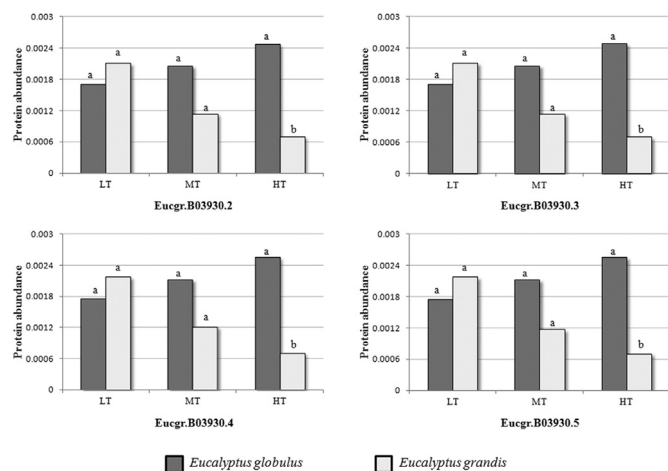
Fig. 4. The differential abundance of the superoxide dismutase Eucgr.B03930 isoforms identified in between *Eucalyptus globulus* and *Eucalyptus grandis* stems. Significant differences were determined according to the Kruskal-Wallis test at 1% confidence. Treatments: LT, low temperature; MT middle temperature; HT, high temperature.

Upon incorporation into the lignin backbone, these precursors are then referred to as *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) units. The lignin monomers differ in the degree of methoxylation in their aromatic ring (Boerjan et al., 2003), and, thus, depending on the lignin composition, the recovery of cellulose may be affected by its native composition.

Using the proteomics strategy described here, we identified and quantitatively profiled 27 protein isoforms potentially involved in the lignin biosynthesis in the target species. Recently, Carocha et al. (2015) carried out a genome-wide survey of the genes involved in

the monolignol biosynthesis and identified 38 *E. grandis* genes belonging to *bona fide* lignin clades. Additionally, by using different treatments with the goal of inducing lignification and studying changes in the expression of those genes, they suggested the direct involvement of 17 gene products in Eucalyptus xylem lignification. This set of genes was labeled by the authors as the lignification toolbox. Our proteomics approach resulted in the identification of 13 of the 17 lignification-related genes (Figs. 5 and 6). The only proteins not identified were those coded by the genes *EgrC4H2*, *EgrC3'H4*, *EgrCAD3* and *EgrCCR1*, a unique representative of the cinnamoyl CoA reductase (CCR) clade. Those missing identifications could be a consequence of the technical challenges faced by the proteomics community to successfully extract, solubilize and acquire mass spectrometer data from non-model organisms with the goal of achieving a deep coverage of the complex and dynamic cellular proteome in a single shot. Additionally, we need to emphasize the potential existence of variants not previously described for the studied species, which contain amino acid polymorphisms not predicted in the genome used here for database searching.

From the nine phenylalanine ammonia-lyase (PALs) isoforms suggested in the 38 *bona fide* genes, we identified eight in our proteome analysis, including the isoforms PAL3 and PAL9 that were suggested as directly related to lignin biosynthesis (Fig. 5). The protein abundance of both isoforms was increased in the high and low temperature treatments for *E. globulus* and *E. grandis* species; those differences were more prominent for *E. globulus* stems grown in the LT environment. As the enzyme PAL plays a dual role in the lignin and other phenylpropanoid biosyntheses, changes in the abundance of PAL isoforms, as shown here (Figs. 5 and 6), cannot be categorically determined as involved in the lignification process. However, it is possible to affirm that those isoforms may be
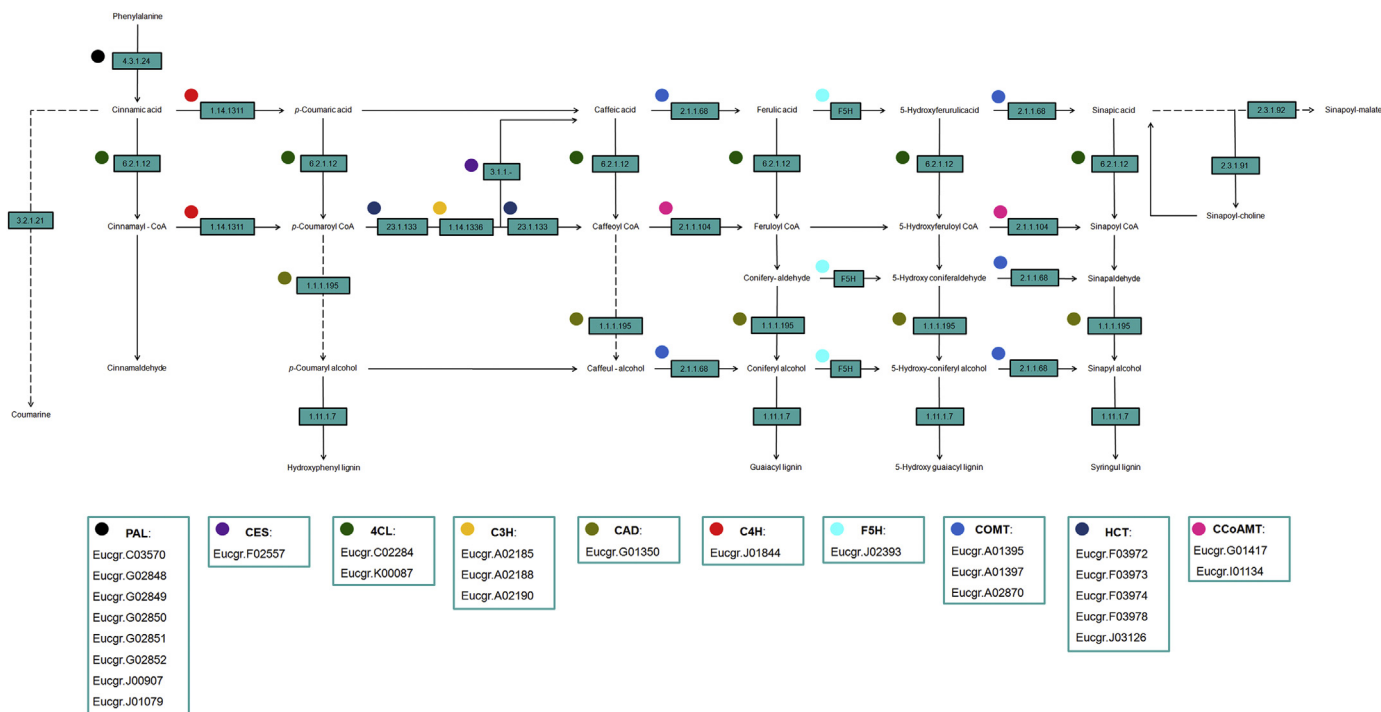

Fig. 5. Lignin biosynthetic pathway. Enzymes (E.C. numbers) and corresponding isoforms (accession numbers) identified in the *Eucalyptus globulus* and *Eucalyptus grandis* stems are indicated as colored circles. Enzyme abbreviations: PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate:CoA ligase; HCT, p-hydroxycinnamoyl-CoA:quinate/shikimate p-hydroxycinnamoyltransferase; C3H, p-coumarate 3-hydroxylase; CSE, caffeoyl shikimate esterase; CCoAOMT, caffeoyl-CoA O-methyltransferase; CCR, cinnamoyl-CoA reductase; CAD, cinnamyl alcohol dehydrogenase; COMT, caffeic acid O-methyltransferase; F5H, ferulate 5-hydroxylase. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
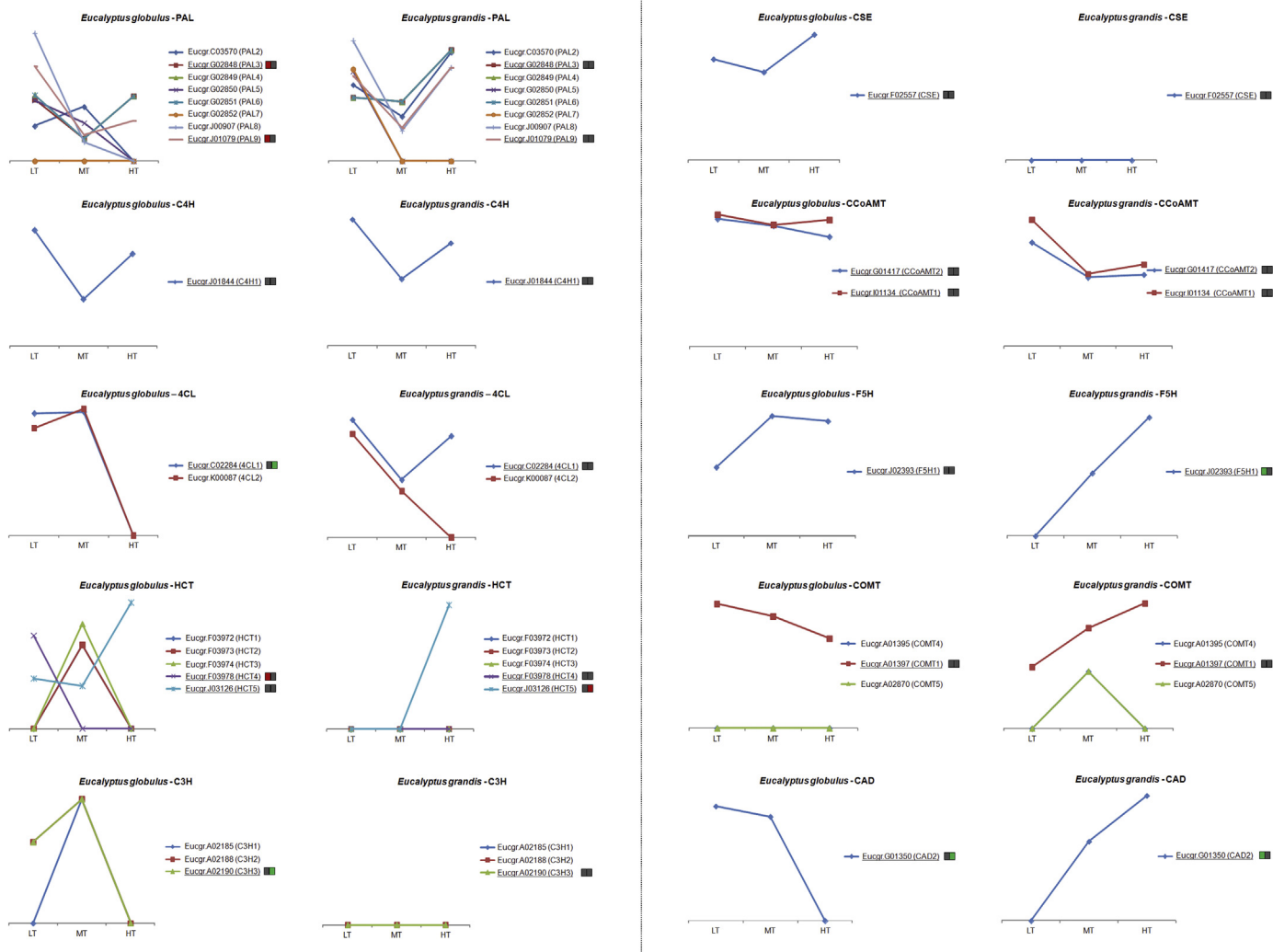
Fig. 6. The relative abundance of the enzymes involved in Eucalyptus lignin biosynthesis. Abundance data are shown in terms of NSAF. Treatments: LT, low temperature; MT: middle temperature; HT: high temperature. Enzyme abbreviations: PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate:CoA ligase; HCT, p-hydroxycinnamoyl-CoA:quinate/shikimate p-hydroxycinnamoyltransferase; C3H, p-coumarate 3-hydroxylase; CSE, caffeoyl shikimate esterase; CCoAOMT, caffeoyl-CoA O-methyltransferase; CCR, cinnamoyl-CoA reductase; CAD, cinnamyl alcohol dehydrogenase; COMT, caffeic acid O-methyltransferase; F5H, ferulate 5-hydroxylase. Underlined protein isoforms denote the gene products indicated by Carocha et al. (2015) as those directly involved in the Eucalyptus lignin biosynthesis. Red and green squares represent significant up- and down-regulation, respectively, according to the Kruskal-Wallis ($p < 0.01$). Gray squares indicate the absence of significant difference according to the Kruskal-Wallis ($p < 0.01$). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

involved in temperature responses in stems of *E. globulus* and *E. grandis*. In accordance with the regulation data obtained for the PAL isoforms, quantitative data of the cinnamate 4-hydroxylase (C4H) enzyme indicates a stimulation of the lignin/phenyl-propanoid biosynthetic path when plants are cultivated at 10 or 32 °C.

The first differences in the abundance of the lignification-related enzymes between *E. globulus* and *E. grandis* were detected for the enzymes 4-coumarate:CoA ligase (4CL) and *p*-coumarate 3-hydroxylase (C3H). In *E. globulus* stems, 4CL1 was down-regulated when the plants were exposed to high temperature. In the same manner, the C3H3 isoform was also found to be down-regulated when *E. globulus* plants were exposed to high temperature treatment. Oppositely, 4CL1 was up-regulated in *E. grandis* stems cultivated in the same conditions while no C3H mass spectrometric data were acquired for this isoform (Fig. 6). It is important to note that the absence of identification is not an evidence of the lack of gene expression but a suggestion that this gene product was present in

such a low number of copies that it hampered a confident identification by the analytical technique used here.

One of the first enzymes involved in the lignin biosynthesis, a branch from the general phenylpropanoid pathway, is the enzyme hydroxycinnamoyl CoA:shikimate hydroxycinnamoyl transferase (HCT). In the present work, we identified both HCT isoforms directly involved in the lignification process. The abundance of the HCT5 isoform was increased when both plant species were cultivated at high temperature. Although we did not detect the isoform HCT4 in the *E. grandis* stems, this isoform showed an opposite profile, in relation to the other protein variant, as an up-regulation was observed in *E. globulus* stems only when plants were cultivated at low temperature. Methylation in the Eucalyptus lignification pathway is catalyzed by the enzymes CCoAOMT1, CCoAOMT2 and COMT1. From our proteomics data, similar regulation profiles were observed for those isoforms in *E. globulus* and *E. grandis* plants after growth in different thermal conditions and, thus, from our data, no biological significance could be drawn based on the abundance of

these enzymes.

Historically, the caffeoyl shikimate esterase (CSE) is the latest enzyme identified in lignin biosynthesis (Vanholme et al., 2013). The importance of CSE was recently highlighted for lignification in all plant species since loss of function mutants of *Medicago truncatula* were profoundly affected during growth, while no CSE orthologs were found in the grasses *Brachypodium distachyon* and *Zea mays* (Ha et al., 2016). CSE catalyzes the synthesis of caffeoyl shikimic acid to caffeic acid, which is then converted to caffeoyl-CoA by 4CL through a bypass of the HCT activity. The CSE enzyme was initially found in *cse* deficient mutants of *Arabidopsis thaliana*, and it was suggested as an alternative pathway for caffeoyl-CoA. Our data confirm the expression of the CSE in a Eucalyptus organism and indicate a higher abundance in the species *E. globulus* in relation to *E. grandis*.

Due to their similar structures, monolignol biosynthesis depends mostly on the activity of the same set of enzymes. However, the ferulate 5-hydroxylase (F5H) is a key enzyme in catalyzing the synthesis of sinapyl alcohol and, consequently, S lignin moieties. Due to its specific position in the lignin biosynthesis, changes in the abundance or activity of this enzyme may result in increased biosynthesis of S or G monolignols. Our proteome data shows a clear distinction in the abundance of the F5H1 isoform between *E. globulus* and *E. grandis* stems. The *E. globulus* plants grown under controlled conditions (MT) show a constitutive higher abundance of F5H1 in relation to *E. grandis* stems. Such a difference corroborates with the reported high S/G ratio in *E. globulus* in relation to the *E. grandis* lignin matrix. It is interesting to note that changes in the abundance of this isoform are pronounced in *E. grandis* stems depending on the applied thermal stimulus. High temperature treatment resulted in an expressive increase in the abundance of this variant (Fig. 6). Although the up-regulation of a single enzyme may not be sufficient to result in a phenotypic alteration, F5H1 may be considered a key target for genetic engineering aiming at an increase in the recovery of *E. grandis* cellulose from the lignocellulose matrix.

Finally, the cinnamyl alcohol dehydrogenase (CAD), an enzyme involved in the reduction of hydroxycinnamyl aldehydes to their corresponding alcohols and the last known enzyme in the monolignol biosynthesis, showed a contrasting regulation profile between the studied species. While the biosynthesis of CAD2 was repressed in *E. globulus* stems from plants grown at high temperature; abundance of this protein variant was increased in *E. grandis* stems from plants cultivated at high temperatures. The exact reason why CAD2 showed this contrasting regulation profile cannot be concluded based on our proteomics experiment. However, the mandatory and ubiquitous role of CAD2 in monolignol biosynthesis suggests a higher lignin accumulation in *E. globulus* plants grown at low temperature than in plants grown at high temperature. Contrarily, *E. grandis* plants should potentially accumulate more lignin in stems when grown in high temperature climates. Although we did not carry out a lignin assay to measure the abundance of this metabolite in the evaluated species under the conditions proposed in the present work, preferential usage of one of those species in tropical or temperature climate regions corroborates our findings regarding the CAD2 abundance profile.

3. Concluding remarks

To our knowledge, this study is the most comprehensive proteomics study carried out in the Eucalyptus genus. The results presented here suggest differential antioxidant response mechanisms between *E. globulus* and *E. grandis* species and that those differences are more prominent when plants are exposed to a high temperature stimulus. Identification and quantification of most of

the core lignification toolbox allowed us to pinpoint changes induced by the abiotic stress evaluated here. Considering all quantification data from the proteins described in the protein dataset, it can be inferred that the lignin biosynthetic path is stimulated in *E. globulus* plants grown at low temperature, while the *E. grandis* lignin biosynthetic path is induced upon growth in high temperature environments. Finally, it is important to mention that protein identities described here, as specific gene products, can be used to monitor or engineer plants for the purpose of obtaining a higher lignocellulose yield.

An Excel file containing all protein identifications and mass spectrometer data is available for download (Tables 1S–30S). The mass spectrometry proteomics data have also been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD005743.

4. Experimental procedures

4.1. Plant material and growth conditions

Eucalyptus globulus and *Eucalyptus grandis* seeds were obtained from a local nursery, germinated in a mixture of sand and soil (1:1, v/v) and grown in 0.5 L pots under regular greenhouse conditions until they grew 5–6 fully expanded leaves. Then, they were transferred to temperature-controlled chambers where they were cultivated at 10 °C, 22 °C and 32 °C for five weeks, designated henceforth as low temperature (LT), medium temperature (MT) and high temperature (HT), respectively. During this period, plants were watered every other day. Chambers were adjusted to receive 10 h of light, which was provided by a set of blue and red light-emitting diodes (LED, 250 μmol photons $m^2.s^{-1}$). After five weeks of treatment, leaves were removed and the stems were quickly chopped with scissors before being frozen in liquid nitrogen. Plant materials were stored at −80 °C prior to sample processing.

4.2. Protein extraction

Stem proteins were extracted according to the protocol originally proposed by Hurkman and Tanaka (1986). Aliquots of 1 g were transferred to clean tubes and 4 mL of extraction buffer (500 mM Tris, 50 mM EDTA, 700 mM sucrose, 100 mM KCl and 1% beta-mercaptoethanol, pH 8.0) were added to the material. After a 1 min incubation, an equal volume of equilibrated phenol solution (Sigma-Aldrich) was added to the extract and homogenized for 5 min at 4 °C. Organic and aqueous phases were separated by 15 min of centrifugation at 3,200g. Proteins were recovered by adding 10 mL of 0.1 M ammonium acetate in methanol into the transferred upper phase. Overnight precipitation was carried out at −20 °C and the pellets were obtained by centrifugation at 3,200g for 15 min at 4 °C. Prior to resuspension, protein pellets were washed twice with 2 mL of cold acetone, followed by 3,200g centrifugation for 15 min. Finally, proteins were resuspended in 100 μL of sample buffer (125 mM Tris pH 6.8, 20% glycerol, 1% SDS and 1% DTT) prior to Bradford assay (Bradford, 1976).

4.3. Polyacrylamide gel electrophoresis and protein digestion

Aliquots of 50 μg of stem proteins were separated under denaturing conditions through gel electrophoresis. For this separation, an equal volume of loading buffer (125 mM Tris pH 6.8, 20% glycerol, 1% SDS, 1% DTT) was added to the sample buffer and the mixture was boiled for 5 min. Discontinuous gel electrophoresis was carried out at 20 mA per gel for 2 h. Protein bands were visualized by Coomassie G-250 staining. In-gel digestions of proteins separated by SDS-PAGE were carried out by slicing the entire

gel lane into five segments and digesting each segment separately according to the protocol proposed by Shevchenko et al. (2006). Protein digestion was carried out for 16 h at 37 °C by adding 2 μg of trypsin in each tube. Tryptic peptides were eluted from the gel matrix for 2 h using a 5% formic acid solution in 100% ACN (1:2, v/v) and then dried down under vacuum centrifugation.

4.4. Mass spectrometer analyses

Tryptic peptides were separated by reverse phase liquid chromatography using a 15 cm in-house nanocolumn (75 μm i.d. packed with Magic 18AQ particles, 5 μm, 200 Å). A 60 min gradient, from 5% to 80% of 0.1% formic acid in acetonitrile, was used for peptide separation at a constant flow rate of 300 nL/min. Mass spectrometer analyses were carried out using a Q-Exactive (Thermo Fisher Scientific) instrument operated in the positive ion mode. Mass spectra acquisition was performed in a data-dependent mode with cycles consisting of one full scan (400–2000 m/z) at 70,000 (FWHM) followed by ten data-dependent scans at 35,000. Peptide fragmentation was obtained by HCD fragmentation using a stepped collision energy equal to 35 (±5%). Singly charged ions were excluded from fragmentation events and peptide ions selected for fragmentation were dynamically excluded for 60 s.

4.5. Protein identification and data analyses

Protein identification was carried out using the spectral correlation approach. For this approach, *Eucalyptus grandis* protein sequences were downloaded from the Phytozome website (Goodstein et al., 2012) and concatenated with reverse sequences to obtain a searchable database comprising 85,874 entries. Stringent-database searches were carried out by the Comet search engine (Eng et al., 2013) using the following amino-acid mass modifications: cysteine carbamidomethylation and methionine oxidation as static and dynamic modifications, respectively. Precursor mass tolerance of 10 ppm was used during stringent database searches. All suggested peptide-spectrum matches were filtered using the software SePro (Carvalho et al., 2012), adjusted for a 1% false-discovery rate. Primary and secondary scores, delta CN and the number of peaks matched were using in the calculation of the Bayesian score. A minimum length of six amino acids was required for peptide identifications and for the proteins described by only one identified peptide a Bayesian score of 2.5 was mandatory. Proteins sharing common peptides were grouped according to the principle of maximum parsimony with the care of avoiding double counting the PSMs for the shared peptides. Protein abundance within each group was evaluated using the NSAF approach (Paoletti et al., 2006). Differential regulation was only considered for proteins presenting a significant abundance difference for the Kruskal-Wallis statistical test at a confidence level of 1%. For each treatment and species, five biological replicates were used, each comprising five young plant stems.

Acknowledgments

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.phytochem.2017.01.017.

References

Apel, K., Hirt, H., 2004. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu. Rev. Plant Biol. 55, 373–399.

Araújo, P., Cesarino, I., Mayer, J.L.S., Ferrari, I.F., Kiyota, E., Sawaya, A.C.H.F., Leme, A.F.P., Mazzafera, P., 2014a. A model system to study the lignification process in *Eucalyptus globulus*. Physiol. Plant 152, 17–31.

Araújo, P., Ferreira, M.S., de Oliveira, D.N., Pereira, L., Sawaya, A.C.H.F., Catharino, R.R., Mazzafera, P., 2014b. Mass spectrometry imaging: an expeditious and powerful technique for fast in situ lignin assessment in Eucalyptus. Anal. Chem. 86, 3415–3419.

Bhuiya, M.W., Liu, C.-J., 2010. Engineering monolignol 4-O-Methyltransferases to modulate lignin biosynthesis. J. Biol. Chem. 285, 277–285.

Boerjan, W., Ralph, J., Baucher, M., 2003. Lignin biosynthesis. Annu. Rev. Plant Biol. 54, 519–546.

Boyer, J.S., 1982. Plant productivity and environment. Science 218, 443–448.

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248–254.

Carocha, V., Hefer, C., Cassan-Wang, H., Fevereiro, P., Alexander, A., Paiva, J.A.P., Grima-Pettenati, J., 2015. Genome-wide analysis of the lignin toolbox of *Eucalyptus grandis*. New Phytol. 4, 1297–1313.

Carvalho, P.C., Fischer, J.S., Xu, T., Yates, J.R., Barbosa, V.C., 2012. PatternLab: from mass spectra to label-free differential shotgun proteomics. Curr. Protoc. Bioinforma. 13, 13–19.

Cesarino, I., Araújo, P., Domingues Júnior, A.P., Mazzafera, P., 2012. An overview of lignin metabolism and its effect on biomass recalcitrance. Braz. J. Bot. 35, 303–311.

Choudhury, S., Panda, P., Sahoo, L., Panda, S.K., 2013. Reactive oxygen species signaling in plants under abiotic stress. Plant Signal. Behav. 8, e23681.

Cramer, G.R., Urano, K., Delrot, S., Pezzotti, M., Shinozaki, K., 2011. Effects of abiotic stress on plants: a systems biology perspective. BMC Plant Biol. 11, 163.

Das, K., Roychoudhury, A., 2014. Reactive oxygen species (ROS) and response of antioxidants as ROS-scavengers during environmental stress in plants. Front. Environ. Sci. 2, 53.

Dasgupta, M.G., Dharanishanthi, V., Agarwal, I., Krutovsky, K.V., 2015. Development of genetic markers in Eucalyptus species by target enrichment and exome sequencing. PLoS One 10, e0116528.

Del Lungo, A., Ball, J., Carle, J.B., 2006. Global Planted Forests Thematic Study: Results and Analysis, first ed. FAO, Rome, p. 178.

Eng, J.K., Jahan, T.A., Hoopmann, M.R., 2013. Comet: an open-source MS/MS sequence database search tool. Proteomics 13, 22–24.

Fernández, M., Valenzuela, S., Barraza, H., Latorre, J., Neira, V., 2012. Photoperiod, temperature and water deficit differentially regulate the expression of four dehydrin genes from *Eucalyptus globulus*. Trees 26, 1483–1493.

Fernández, M., Villarroel, C., Balbontín, C., Valenzuela, S., 2010. Validation of reference genes for real-time qRT-PCR normalization during cold acclimation in *Eucalyptus globulus*. Trees 24, 1109–1116.

Foyer, C.H., 2005. Redox homeostasis and antioxidant signaling: a metabolic interface between stress perception and physiological responses. Plant Cell Online 17, 1866–1875.

Gamboa, M.C., Rasmussen-Poblete, S., Valenzuela, P.D.T., Krauskopf, E., 2007. Isolation and characterization of a cDNA encoding a CBF transcription factor from *Eucalyptus globulus*. Plant Physiol. Biochem. 45, 1–5.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., Rokhsar, D.S., 2012. Phytozome: a comparative platform for green plant genomic. Nucleic Acids Res. 40, D1178–D1186.

Ha, C.M., Escamilla-Trevino, L., Serrani-Yarce, J.C., Kim, H., Ralph, J., Chen, F., Dixon, R.A., 2016. An essential role of caffeoyl shikimate esterase in monolignol biosynthesis in *Medicago truncatula*. Plant J. 86, 363–375.

Hasanuzzaman, M., Nahar, K., Fujita, M., 2012. Extreme temperature responses, oxidative stress and antioxidant defense in plants. In: Vahdati, K., Leslie, C. (Eds.), Abiotic Stress - Plant Responses and Applications in Agriculture. InTech, Rijeka, pp. 169–205.

Hurkman, W.J., Tanaka, C.K., 1986. Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis. Plant Physiol. 81, 802–806.

Liu, Y., Jiang, Y., Lan, J., Zou, Y., Gao, J., 2014. Comparative transcriptomic analysis of the response to cold acclimation in Eucalyptus dunnii. PLoS One 9, e113091.

Macfarlane, C., Adams, M.A., Hansen, L.D., 2002. Application of an enthalpy balance model of the relation between growth and respiration to temperature acclimation of *Eucalyptus globulus* seedlings. Proc. Biol. Sci. 269, 1499–1507.

Marchal, C., Delorme-Hinoux, V., Bariat, L., Siala, W., Belin, C., Saez-Vasquez, J., Riondet, C., Reichheld, J.-P., 2014. NTR/NRX define a new thioredoxin system in the nucleus of *Arabidopsis thaliana* cells. Mol. Plant 7, 30–44.

Martí, M.C., Olmos, E., Calvete, J.J., Díaz, I., Barranco-Medina, S., Whelan, J., Lázaro, J.J., Sevilla, F., Jiménez, A., 2009. Mitochondrial and nuclear localization of a novel pea thioredoxin: identification of its mitochondrial target proteins. Plant Physiol. 150, 646–657.

Meyer, Y., Reichheld, J.P., Vignols, F., 2005. Thioredoxins in Arabidopsis and other plants. Photosynth. Res. 86, 419–433.

Myburg, A.A., Grattapaglia, D., Tuskan, G.A., Hellsten, U., Hayes, R.D., Grimwood, J., Jenkins, J., Lindquist, E., Tice, H., Bauer, D., Goodstein, D.M., Dubchak, I., Poliakov, A., Mizrachi, E., Kullan, A.R.K., Hussey, S.G., Pinard, D., Merwe, K., Singh, P., Jaarsveld, I., Silva-Junior, O.B., Togawa, R.C., Pappas, M.R., Faria, D.A.,

Sansaloni, C.P., Petroli, C.D., Yang, X., Ranjan, P., Tschaplinski, T.J., Ye, C.-Y., Li, T., Sterck, L., Vanneste, K., Murat, F., Soler, M.M., Clemente, H.S., Saidi, N., Cassan-Wang, H., Dunand, C., Hefer, C.A., Bornberg-Bauer, E., Kersting, A.R., Vining, K., Amarasinghe, V., Ranik, M., Naithani, S., Elser, J., Boyd, A.E., Liston, A., Spatafora, J.W., Dharmwardhana, P., Raja, R., Sullivan, C., Romanel, E., Alves-Ferreira, M., Lheim, C.K., Foley, W., Carocha, V., Paiva, J., Kudrna, D., Brommonschenkel, S.H., Pasquali, G., Byrne, M., Rigault, P., Tibbits, J., Spokevicius, A., Jones, R.C., Steane, D.A., Vaillancourt, R.E., Potts, B.M., Joubert, F., Barry, K., Pappas Jr., G.J., Strauss, S.H., Jaiswal, P., Grima-Pettenati, J., Salse, J.J., Van de Peer, Y., Rokhsar, D.S., Schmutz, J., 2014. The genome of *Eucalyptus grandis*. Nature 510, 356—362.

Paoletti, A.C., Parmely, T.J., Tomomori-Sato, C., Sato, S., Zhu, D., Conaway, R.C., Conaway, J.W., Florens, L., Washburn, M.P., 2006. Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. Proc. Natl. Acad. Sci. U. S. A. 103, 18928—18933.

Pfalz, J., Liebers, M., Hirth, M., Grübler, B., Holtzegel, U., Schröter, Y., Dietzel, L., Pfannschmidt, T., 2012. Environmental control of plant nuclear gene expression by chloroplast redox signals. Front. Plant Sci. 3, 1—9.

Potts, B.M., 2004. Genetic improvement of eucalypt. In: Burley, J., Evans, J., Youngquist, J.A. (Eds.), Encyclopedia of Forest Science. Elsevier Science, Oxford, pp. 1480—1490.

Pulido, P., Cazalis, R., Cejudo, F.J., 2009. An antioxidant redox system in the nucleus of wheat seed cells suffering oxidative stress. Plant J. 57, 132—145.

Serrato, A.J., Cejudo, F.J., 2003. Type-h thioredoxins accumulate in the nucleus of developing wheat seed tissues suffering oxidative stress. Planta 217, 392—399.

Serrato, A.J., Crespo, J.L., Florencio, F.J., Cejudo, F.J., 2001. Characterization of two thioredoxins h with predominant localization in the nucleus of aleurone and scutellum cells of germinating wheat seeds. Plant Mol. Biol. 46, 361—371.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V., Mann, M., 2006. In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat. Protoc. 1, 2856—2860.

Vanholme, R., Cesarino, I., Rataj, K., Xiao, Y., Sundin, L., Goeminne, G., Kim, H., Cross, J., Morreel, K., Araujo, P., Welsh, L., Haustraete, J., McClellan, C., Vanholme, B., Ralph, J., Simpson, G.G., Halpin, C., Boerjan, W., 2013. Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in Arabidopsis. Science 341, 1103—1106.