genes

MDPI

Article

Analysis of NAC Domain Transcription Factor Genes of *Tectona grandis* L.f. Involved in Secondary Cell Wall Deposition

Fernando Manuel Matias Hurtado [1,†] , **Maísa de Siqueira Pinto** [1,†] , **Perla Novais de Oliveira** [1] , **Diego Mauricio Riaño-Pachón** [2] , **Laura Beatriz Inocente** [1] and **Helaine Carrer** [1,*]

1 Department of Biological Sciences, Luiz de Queiroz College of Agriculture (ESALQ), University of São Paulo, Av. Pádua Dias, 11, CP 9, Piracicaba, SP 13418-900, Brazil; fermathur1992@usp.br (F.M.M.H.); maisadesiqueira@gmail.com (M.d.S.P.); perla.oliveira@usp.br (P.N.d.O.); laurabinocente@gmail.com (L.B.I.)

2 Computational, Evolutionary and Systems Biology Laboratory, Center for Nuclear Energy in Agriculture (CENA), University of São Paulo. Av. Centenário 303, Piracicaba, SP 13416-000, Brazil; diego.riano@cena.usp.br

* Correspondence: hecarrer@usp.br; Tel.: +55-19-3429-4344 (ext. 21)

† These authors contributed equally to this work.

check for updates

Abstract: NAC proteins are one of the largest families of plant-specific transcription factors (TFs). They regulate diverse complex biological processes, including secondary xylem differentiation and wood formation. Recent genomic and transcriptomic studies of *Tectona grandis* L.f. (teak), one of the most valuable hardwood trees in the world, have allowed identification and analysis of developmental genes. In the present work, *T. grandis* NAC genes were identified and analyzed regarding to their evolution and expression profile during wood formation. We analyzed the recently published *T. grandis* genome, and identified 130 NAC proteins that are coded by 107 gene loci. These proteins were classified into 23 clades of the NAC family, together with *Populus*, *Eucalyptus*, and *Arabidopsis*. Data on transcript expression revealed specific temporal and spatial expression patterns for the majority of teak NAC genes. RT-PCR indicated expression of VND genes (*Tg11g04450-VND2* and *Tg15g08390-VND4*) related to secondary cell wall formation in xylem vessels of 16-year-old juvenile trees. Our findings open a way to further understanding of NAC transcription factor genes in *T. grandis* wood biosynthesis, while they are potentially useful for future studies aiming to improve biomass and wood quality using biotechnological approaches.

Keywords: tropical tree; wood formation; secondary growth

1. Introduction

NAC proteins are one of the largest families of plant-specific transcription factors (TFs), in which the DNA binding domain is highly conserved. The NAC designation is derived from the *NO APICAL MERISTEM (NAM)* gene from *Petunia hybrida* E. Vilm. and from the *Arabidopsis thaliana* L. genes *ATAF1–2* and *CUP-SHAPED COTYLEDON (CUC2)* [1]. Structures of NAC proteins are commonly divided into two regions, a conserved N-terminal DNA-binding domain and a variable C-terminal transcriptional regulatory region [2]. The N-terminal NAC domains comprise five (A–E) subdomains [2–4] located at the N-terminal. Usually, the subdomains contain 150–160 amino acid residues that are implicated in DNA binding, nuclear localization, and formation and localization of homodimers or heterodimers [2]. The C-terminal regions confer regulation of transcriptional activity [3,4]. NAC TFs have been analyzed at the genome-wide level in diverse species such as

Arabidopsis thaliana L., *Eucalyptus grandis* W.Hill, *Populus trichocarpa* L. (black cottonwood), *Nicotiana tabacum* L. (tobacco), and *Glycine max* (L.) Merr. (soybean) [3,5–8].

Paleobotanical studies of tracheids resistant to degradation indicate that lignification and secondary cell wall (SCW) differentiation evolved in land plants in the early Devonian (about 415 million years ago) [9]. SCW biosynthesis requires coordinated expression of genes for cellulose, xylan, glucomannan, and lignin biosynthesis. A transcriptional network composed of NAC and MYB transcription factors regulates SCW biosynthetic pathways [10,11]. These closely related TFs function as master transcriptional switches that regulate downstream targets [12]. Thus far, NAC TF genes were identified in all analyzed land plants, whereas eukaryotic algae species lack NAC TFs [13]. Comparative genomic and gene functional analyses have indicated that NAC TFs underwent an expansion in the lineage of vascular plants [13]. Also, tracheid evolution is associated with a burst of structural diversity [14], which is probably related to NAC genes controlling SCW biosynthesis in vascular plants, comprising seedless vascular plants, gymnosperms, and angiosperms [15]. Noticeably, plant species that are structurally less complex contain significantly fewer NAC genes, such as the model bryophyte *Physcomitrella patens* (Hedw.) Mitt. (32 NAC genes), and no NAC genes are found in Chlorophytes (*Ostreococcus lucimarinus* Palenik, *Ostreococcus tauri* C. Courties & M.-J. Chrétiennot-Dinet, *Chlamydomonas reinahrdtii* P.A.Dang.) [16]. Most likely, diversification of NAC genes began with the colonization of terrestrial environments by the ancestors of modern land plants [14]. NAC TFs have been associated with the regulation of a wide range of plant developmental processes acting as transcription activators or repressors. These plant developmental processes include ripening [17], cellular morphogenesis, signaling transduction [18], and establishment of the shoot apical meristem [4], floral organs [19], and lateral roots [20]. They are also involved in the signaling pathway induced during biotic and abiotic stresses [18,21,22], and in the regulation of leaf senescence [23]. As for secondary growth, NAC TFs have been described in the regulation of xylogenesis, fiber development, and secondary cell wall formation [24–32].

In Arabidopsis, NAC TFs represent the most upstream regulators in the transcriptional regulatory network that controls wood formation. Master regulators of this development are the VASCULAR NAC DOMAIN (VND1–VND7), NAC SECONDARY WALL THICKENINGs (NST1 and NST2), and SECONDARY WALL NAC DOMAIN TFs (SND1/NST3 and SND2) [10]. These NAC TFs regulate the expression of the MYB genes [11] during SCW formation [12]. Besides, the PROTEIN SOMBRERO (SMB) TF and its two close homologs, BEARSKIN1 (BRN1) and BRN2, are related to the VND/NST TFs in this SCW regulatory network [33]. In this same Arabidopsis regulatory network, the VND-INTERACTING 2 (VNI2) TF is described to negatively regulate VND7 [34]. *Tectona grandis* L.f. (teak), a worldwide attractive tropical tree with valuable wood of high economic importance in industrial forestry, is also a plant species of interest to better understand the SCW biosynthesis and wood formation. Teak's secondary xylem biosynthesis is a complex molecular process that involves several genes and environmental factors. Recent genomic and transcriptomic analyses of *T. grandis* have allowed to improve the understanding of the xylogenesis role in biomass increase and wood quality [35–38], even though *T. grandis* NAC transcription factors have not yet been identified. The identification and characterization of teak NAC TFs are fundamental to elucidate their regulation role in vascular tissue differentiation and wood formation, as well as future biotechnological studies aiming to improve biomass and wood quality.

Here, we report a genome-wide identification, phylogenetic profile, gene structure, chromosome localization, phylogenetic relationship, and expression profiles of *T. grandis* NAC genes related to wood formation. In addition, tissue-specific expression patterns of 13 teak NAC TF genes homologous to Arabidopsis in their response to wood formation were investigated by quantitative real-time RT-PCR (RT-qPCR). The data represent a step forward to the understanding of plant NAC transcription factors and the regulation of secondary cell wall biosynthesis and wood formation in *T. grandis*.

2. Materials and Methods

2.1. Database Search and Phylogenetic Profiles

The *Tectona grandis* NAC genes were compared to the model species *Eucalyptus grandis*, *Populus trichocarpa*, and *Arabidopsis thaliana* in which NAC genes were already identified and analyzed [3,5,6]. Deduced proteomes of *Eucalyptus grandis* [39] and *Populus trichocarpa* [40] were obtained from Phytozome v12, and The Arabidopsis Portal [41] was used for *A. thaliana* proteins. As for *T. grandis*, the proteins were accessed in the Dryad database [38]. We considered that any protein carrying the NAM domain was a member of the NAC family of TFs. The NAM domain was searched and identified using the profile Hidden Markov Model for this domain from PFAM (Accession Number PF02365). This profile was downloaded from Pfam Database (http://pfam.sanger.ac.uk/) and a hmm-search was performed with the HMMer software v3.2.1 (Cambridge, Massachusetts, United States) 'hmmscan'. Hit scores higher than 21.2 (gathering cutoff from PFAM specific for the NAM model) were considered true positives and selected for further analyses. A diagram of the phylogenetic relationship among teak and 28 species was constructed with the tree generator PhyloT (http://phylot.biobyte.de/index.html) based on the NCBI taxonomy IDs. The APGIII classification system was used to manually indicate the phylogenetic orders of the flowering plants species in the tree [42]. As for *Clorophyta*, *Bryophyta*, and *Licophyta*, the classification was done according to Novikiv and Barabas-Krasni [43].

2.2. Phylogenetic Analysis of NAC Transcriptional Factors

Non-redundant peptide sequences of NAC domain proteins (Pfam ID: NAM; PF02365) from *Tectona grandis* (116), *Populus trichocarpa* (196), *Arabidopsis thaliana* (153), and *Eucalyptus grandis* (190), previously identified (Section 2.1), were aligned using MAFFT v7.407 in the auto mode [44]. Poorly aligning regions from the multiple sequence alignment were removed using TrimAl v1.4, in the "automated1" mode [45]. Phylogenetic inference under the Maximum Likelihood approach was carried out with IQ-Tree v1.6.9 [46], with the options ""-m MFP -st AA -seed 12345 -lmap 65400 -alrt 1000 -bb 1000". The best evolutionary model for phylogenetic inference was also estimated within IQ-Tree with the option -m MFP (Model Finder Plus), by computing the log-likelihoods for many different evolutionary models against an initial parsimony tree and choosing the model that minimizes the Bayesian information criterion (BIC) [47]. Branch support was assessed with the Shimodaira–Hasegawa (SH)-like approximate likelihood ratio test [46] and ultrafast bootstrap [48], both with 1000 replicates. The multi-species phylogenetic tree of NAC TFs was reconciled with the species tree, using Notung 2.9 in order to infer the most likely root and the history of gene duplications in the NAC gene family [49]. Briefly, the species tree–gene tree reconciliation process tries to account for the differences between the species tree and the gene tree, with the gene-level processes of gene duplication and gene losses.

2.3. Chromosomal Location

The teak NAC genes were mapped on chromosomes in accordance with the whole genome of this species (https://datadryad.org/stash/dataset/doi:10.5061/dryad.77b2422) [38]. Of the 18 teak chromosomes, 17 near-complete pseudomolecules were generated with one chromosome present as two chromosome arm scaffolds [38]. Chromosomal locations of the identified teak NAC genes were extracted from the general feature format (GFF) file provided with the genome sequence and visualized in the Integrative Genomics Viewer (IGV v2.7.0) [50]. Note that for the purpose of the visualization of NAC genes on teak chromosomes, the pseudomolecules 18 and 19 were joined into a single chromosome.

2.4. Gene Structure and Conserved Motifs

Exon–intron structures of *T. grandis* NAC genes were analyzed and illustrated with the Gene Structure Display Server (GSDS) (http://gsds.gao-lab.org/index.php) by comparison of coding sequence (CDS) regions with genomic DNA sequences of this species [38]. The software MEME v5.0.5

(MEME—http://meme-suite.org/index.html) was employed for the detection of conserved motifs with the following parameters: Distribution of motif occurrences, zero or one per sequence; minimum width, 6, maximum width, 50; maximum number of motifs, 10; and optimum motif width, ≥6 and ≤116 [51]. The subcellular localization of the identified NAC proteins was predicted by using an online analysis tool from Molecular Bioinformatics Center v2.5 (http://cello.life.nctu.edu.tw/) and BUSCA (http://busca.biocomp.unibo.it)

2.5. In Silico Gene Expression Profiling

The RAW RNASeq data from *T. grandis* were accessed in the Sequence Read Archive (SRA-NCBI) under the accession number SRP059970. These data were previously generated by our group [36] and comprise various plant tissues in three plant developmental stages. The teak RAW RNASeq data were processed with BBDuk2 [52] in order to remove low-quality regions from the reads, remainders of adapter sequences, and ribosomal RNA. Salmon v1.0.0 [53] was employed to estimate expression values, as transcripts per million (TPM), using as reference the predicted cDNAs from the *T. grandis* genome [38]. Expression values were imported into R [54] with the tximport package [55] and summarized to gene level. Gene expression values for each gene were transformed using Z-scores, i.e., the expression value of a gene in a given condition was subtracted from the average gene expression and expressed as number of standard deviation from the mean, and then visualized as heatmaps using the pheatmap package.

2.6. Expression Analysis of Marker Genes for Secondary Cell Wall Formation by RT-qPCR

Expression patterns of genes involved in wood formation were analyzed in samples of teak stem collected from 4-year-old plants grown in the greenhouse, and sapwood samples collected from 16- and 64-year-old trees grown in a field in Piracicaba, São Paulo State, Brazil (Latitude: 22°42′23″S, Longitude: 47°37′7″W, 650 m above sea level). This is the same population of trees used for the previously reported transcriptional profile mentioned in Section 2.5 [36]. In this expression analysis, new tissue samples were collected with a Pressler borer at DBH, 4 years apart of the first sampling [36]. Frozen tissue samples (0.6 g) were ground to a fine powder in liquid nitrogen using a sterilized mortar and pestle. Total RNA was extracted following the Trizol (Invitrogen, Carlsbad, California, USA) protocol. RNA was quantified in a NanoDrop 2000 spectrophotometer (Thermo Scientific, Carlsbad, California, USA) and RNA integrity was examined by gel electrophoresis. Total RNA was treated with DNAse I (Promega, Madison, Wisconsin, USA) and then used for cDNA synthesis using SuperScript™ III First-Strand Synthesis System for RT-PCR (Invitrogen, Carlsbad, California, USA), according to the manufacturer's instructions. Quantitative RT-PCR (qPCR) reactions were conducted in 12.5 μL total volume using a Platinum Sybr Green Supermix (Invitrogen, Carlsbad, California, USA), and ran in an ABI 7500 qPCR thermocycler (Applied Biosystems, Foster City, California, USA). Expression data were normalized using the 2-ΔCt method. The constitutive *Elongation Factor-1 alpha* (*EF-1 alpha*) housekeeping gene was used as internal control [56]. We used three biological repetitions and two technical repetitions. The statistical analysis was performed using Prism v8 software (GraphPad, San Diego, California, USA). Normality of data was confirmed by the Shapiro–Wilk test, which indicated that the data follow normal distribution. Statistical significance was determined by one-way ANOVA, with Tukey post-hoc analysis ($p > 0.1$). The analyzed genes are key regulators of SCW formation in Arabidopsis (Table S1).

3. Results and Discussion

3.1. Identification and Evolution Analyses of NAC Family Transcription Factors

The analysis of teak NAC TF encoding genes identified 130 NAC proteins that are coded by 107 loci. We analyzed, with CD-HIT, the 130 teak NAC proteins and identified 14 that were 100% identical to another protein in the original set. Thus, we generated a set of non-redundant proteins (at

100% identity), which has 116 proteins, and used that set for phylogenetic analysis. This procedure was also carried out for each species in the phylogenetic analyses. Each gene was annotated as TgNAC0XX, where Tg refers to the initials of *Tectona grandis* and XX refers to the position of the gene in the chromosomes. Detailed information of teak's NAC family genes, including gene locus, given code, accession numbers, and similarities to their Arabidopsis orthologues, is listed in Table S2.

NAC TF genes have been identified in all analyzed land plants, but so far, they have not been found in eukaryotic algae species [14] (Figure S1). The NAC multigene family is highly variable among plant species. Presumably, NAC TFs first appeared in bryophytes such as *Physcomitrella patens* and, most likely, they were responsible for the expansion of the NAC genes named VNS (VND-NST/SND-SMB-related proteins) in land plants [13,57]. *Populus* (Rosideae) contains large numbers of NAC genes in a relatively small genome size compared to *Eucalyptus grandis* and *Glycine max*, both containing large numbers of NAC genes in larger genome sizes. Expansion of NAC genes in various plant species is probably the result of multiple events of gene duplication. In the *Populus* lineage, for example, whole genome duplication and multiple segmental and tandem duplication events may have contributed to the expansion of the NAC family [40]. As for *T. grandis*, the small number of duplication events could be one of the reasons for the relatively small number of identified NAC genes (107). As depicted in Figure S2, only 36 gene duplication events were inferred for *T. grandis*, whereas they were 77 in Arabidopsis, 114 in Eucalyptus, and 98 in *Populus*. It is also curious that only 31 NAC genes were found in the coniferous *Pinus taeda*, a species with a large genome, possibly due to the high content of pseudogenes in conifer genomes [58]. Besides, only two genes of the VASCULAR NAC DOMAIN (VND) family were identified in *Pinus taeda*. This finding could be interpreted as evidence of the importance of co-option and expansion of the VND gene family during the evolution of angiosperms [59].

3.2. Phylogenetic Analysis of the NAC Gene Family

Evolutionary relationships were examined among non-redundant NAC protein sequences of *T.grandis* (116), *Populus trichocarpa* (196), *Eucalyptus grandis* (190), and *Arabidopsis thaliana* (153). A rooted tree was constructed and maximum-likelihood algorithm separated the NAC family proteins into 23 distinct clades (NAC-a–NAC-w), as represented in Figure 1. *T. grandis* NAC proteins were identified in 14 of the 23 clades (NAC-h–NAC-s, NAC-u, and NAC-w).

Clade NAC-s clustered the largest group of conserved genes that consisted of *T. grandis* (16 genes), *Eucalyptus* (19 genes), *Populus* (34 genes), and Arabidopsis (34 genes) sequences (Figure 1). Clades NAC-d, -e, and -f comprised exclusively *Populus* NAC genes and that is consistent with the evidence of major gene duplication events in this species [40]. In general, NAC members were interspersed among the majority of the clades, which indicates expansion of NAC genes previously to the evolutionary divergence of *T. grandis*, *Populus*, Arabidopsis, and *Eucalyptus*.

Noticeably, NAC genes with the same functions tended to fall into a same clade, as previously reported [60]. For instance, clade NAC-p grouped SCW biosynthesis genes involved in xylem vessels biosynthesis (*VND4* (*Tg15g08390, Tg16g07170*), *VND1* (*Tg09g04510, TgUn296g00020*), *VND7* (*Tg03g10560, Tg03g10970, Tg18g00740*)) and in biosynthesis of fibers (*NSTI* (*Tg05g19210, Tg05g19290, Tg11g07410*) and *SMB* (*Tg02g10510*)). As for VNI genes, both VNI1 (*Tg09g02550*) and VNI2 (*Tg10g07670, Tg11g14670, Tg16g09090, Tg15g04300, Tg16g09080*) genes were located in clade NAC-n, with the exception of *Tg12g02970* (VNI2), that was placed in clade NAC-w. The SND2 genes (*Tg02g15550, Tg17g02630*) were clustered in clade NAC-q (Figure 1).

3.3. Gene Structure and Conserved Motifs of Teak NAC Genes

Variations in gene and protein structure are the basis of evolution in multigene families [61]. A structural analysis of teak NAC genes and proteins was performed in order to gain information on structural diversity among them. In the phylogenetic tree, the analyzed proteins were clustered according to the presence of conserved motifs, as shown in Figure 2A. The color of protein names refers to

the clades in Figure 1. Exon–intron structure of TgNAC coding sequences was also analyzed individually (Figure 2B). NAC members of a same clade shared similar gene length and exon–intron structure.
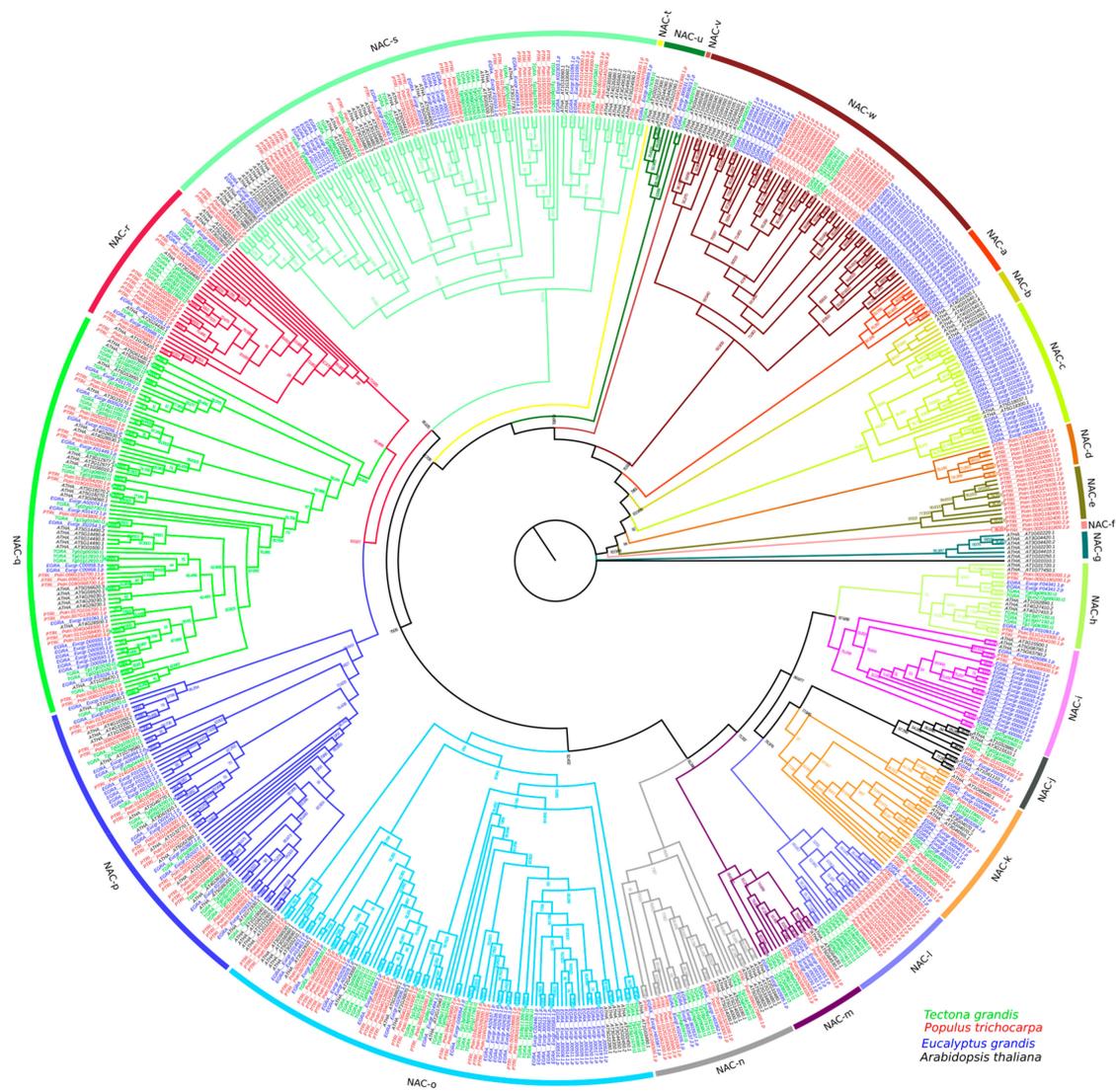

Figure 1. Maximum-likelihood phylogenetic tree rooted of NAC domain proteins of *T. grandis*, *Populus*, Arabidopsis, and *Eucalyptus*. The NAC family proteins were clustered in 23 distinct clades (NAC-a–NAC-w). Members of *T. grandis* (116), *Populus* (196), Arabidopsis (153), and *Eucalyptus* (190) NAC protein family were denoted with green, brown, black, and blue letters, respectively. Bootstrap values are described in the branch points.

In total, 10 divergent motifs were localized and named as motifs 1–10 (Figure 2C, Table S3). The motifs 1–3 and 5–7 were recognized as NAC subdomains in *Eucalyptus*, as described previously [5,62]. Although motifs 4 and 8 have already been described in *Eucalyptus* [5] and *Populus* [6], respectively, they were not recognized as NAC. The motifs 9 and 10 were not described until this study. As expected, most phylogeny-based groups displayed common motifs with the same alignment and position. Thus, NAC proteins with similar gene structures and motifs tended to cluster in the same clade. Clade NAC-q contains the proteins VNS, VNDs (VND1, VND4, and VND7), NST1, and SMB that have very conserved structures (Figure 2A,C). The gene SMB is among the groups of NAC genes present in early land plants, such as the bryophyte *Physcomitrella patens* [63]. Proteins that were clustered in the clades NAC-o, NAC-s, and NAC-q also shared high similarity in their structures, as shown in Figure 2C.

The shortest teak TgNAC gene (TgUn720g00010) is 544 bp long, and the two longest genes (Tg12g03010 and Tg09g02550) are 6 kb in size. Most *T. grandis* genes contained introns in their sequences, except for *Tg10g05820*, *Tg11g02730*, *Tg12g02970,* and *TgUn296g00020* (Figure 2B). The subcellular localization of the identified NAC proteins was predicted by using the online analysis tools Molecular Bioinformatics Center v2.5 (http://cello.life.nctu.edu.tw/) and BUSCA (http://busca.biocomp.unibo.it). *T. grandis* NAC proteins were located in the nucleus, cytoplasm, chloroplast, plasma membrane, endomembrane system, and mitochondria. Such subcellular localization is similar to those of various plant species, such as *Zea mays* and *Cucumin sativus,* in which NAC TFs are also located in diverse cell compartments (Table S4) [64,65].
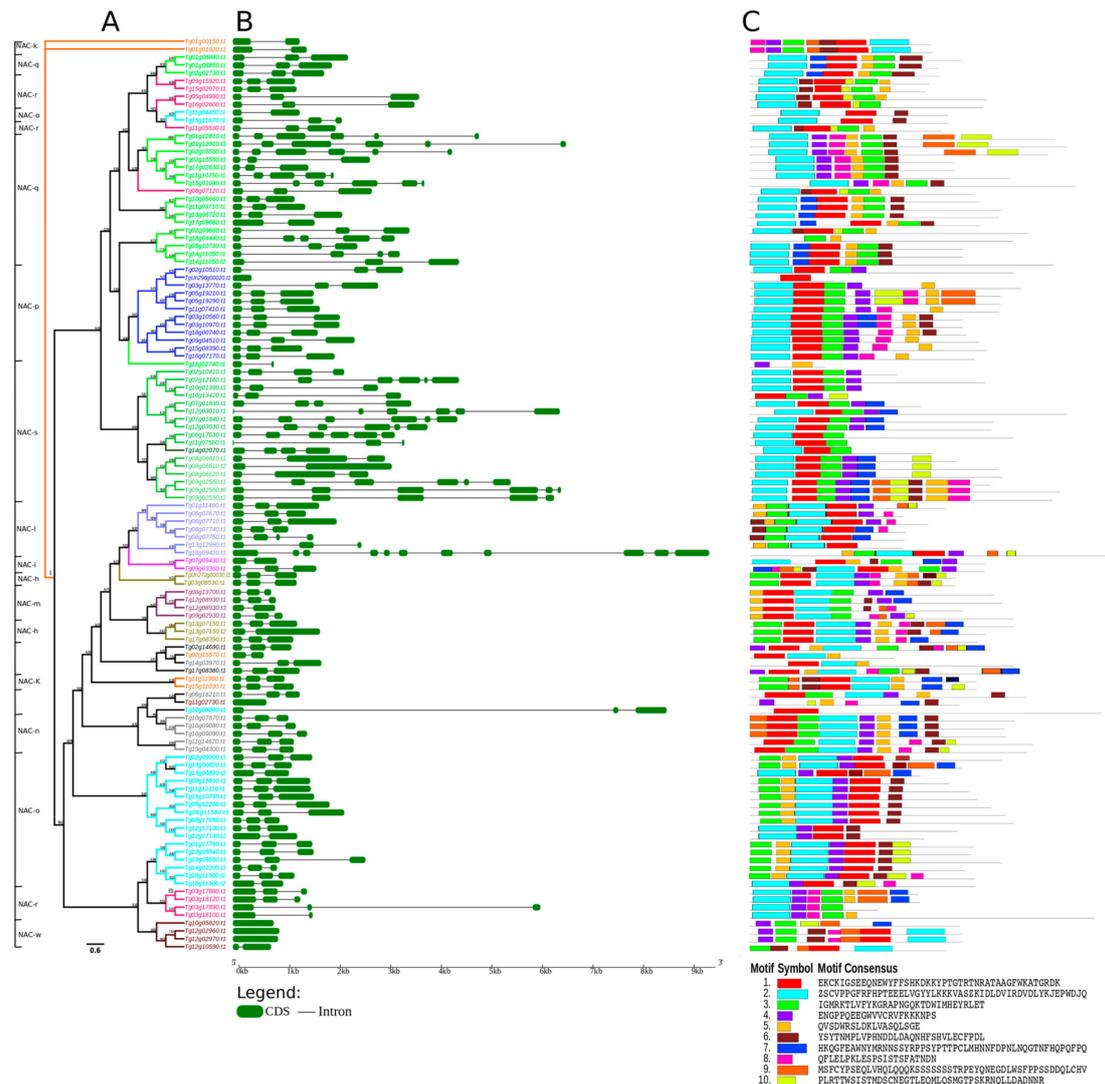
Figure 2. Phylogenetic relationship, gene structure, and motif composition of *T. grandis* NAC genes. (**A**) Maximum-likelihood phylogenetic tree rooted of 116 teak NAC proteins was constructed using Q-Tree v1.6.9 and ultrafast bootstrap, with 1000 replicates. (**B**) Exon–intron structure represented in kilobase (kb) scale of 107 TgNAC genes. (**C**) Schematic representation of the conserved motifs in the NAC proteins from teak. Green boxes and black lines represent exons and introns, respectively.

3.4. Chromosomal Location

Of the 18 teak chromosomes, 17 were generated as near-complete pseudomolecules, and one of them as two chromosome arm scaffolds [38]. Integrative Genomics Viewer (IGV v2.7.0 - Cambridge, Massachusetts, USA) was used for assembling the 18 chromosomes (Chr) and to localize the 107

NAC genes, as represented in Figure 3. Three of the 107 genes (*TgUn272g00030*, *TgUn296g00020*, *TgUn720g00010*), identified by using the published *T. grandis* genome [38], were located in scaffolds that have not been placed within a chromosome; therefore, they were represented in a box at the bottom right-hand corner of Figure 3. The largest number of TgNAC genes was located on Chr 11 and consisted of 11 TgNAC genes. Other TgNAC genes were distributed among 16 of the remaining chromosomes, as follows: 9 on Chr2, 8 on Chr3, 7 on chromosomes 1, 6, 12, and 15, and in lower numbers among the other teak chromosomes. No NAC gene was identified on Chr4. Likewise, unequal distributions of NAC TF genes across chromosomes and highest gene density in distal regions have been reported in barley and wheat [66,67].
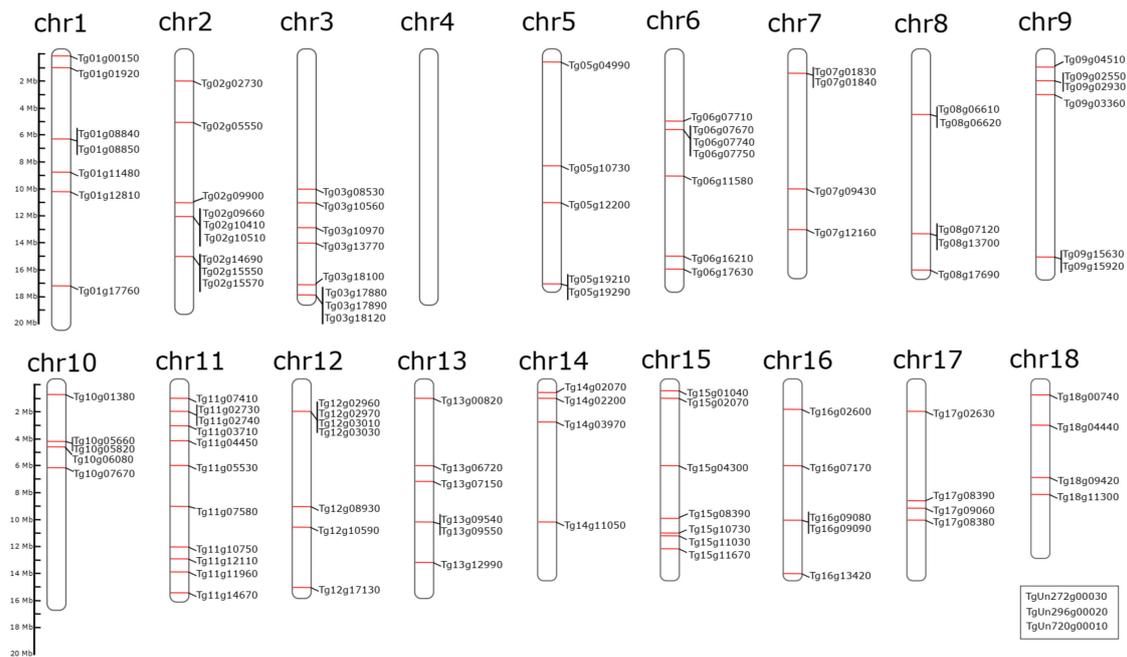
Figure 3. Location and distribution of 107 TgNAC genes on 18 *T. grandis* chromosomes. The box at the bottom right-hand side represents three TgNAC genes that were not placed within any of the chromosomes and were present in orphan scaffolds. TgNAC genes were mapped according to their genomic position in the pseudomolecules (Table S5). The scale on the left-hand side is represented by megabase (Mb).

Many of the teak NAC genes were located adjacent to each other (less than 10,000 bp apart), as, for instance, *Tg01g08840* and *Tg01g08850* on Chr1, *Tg03g17880* and *Tg03g17890* on Chr3, *Tg06g07740* and *Tg06g07750* on Chr6, *Tg07g01830* and *Tg07g01840* on Chr7, *Tg12g02960* and *Tg12g02970* on Chr12, and *Tg16g09080* and *Tg16g09090* on Chr16 (Figure 3, Table S5). Most likely, these closely adjacent genes could have resulted from duplication events. In wheat, single loci are found more frequently in proximal regions of the chromosomes, while gene duplications often occur in the distal regions [67].

3.5. In Silico Gene Expression Profiles of Teak NAC Genes

Expression profiles of 107 TgNAC genes were assessed in order to gain insights into their transcript accumulations. Transcript abundance analyses were carried out for diverse tissues of teak, including primary (leaves, flowers, roots, seedling) and secondary (branch and stem secondary xylem of 12- and 60-year-old plants) tissues [36]. Of the 107 analyzed TgNAC genes, 94 were expressed in the tissues and conditions represented in Figure 4. The heatmap based on hierarchical clustering of expression levels revealed 10 putative NAC groups (A–J).
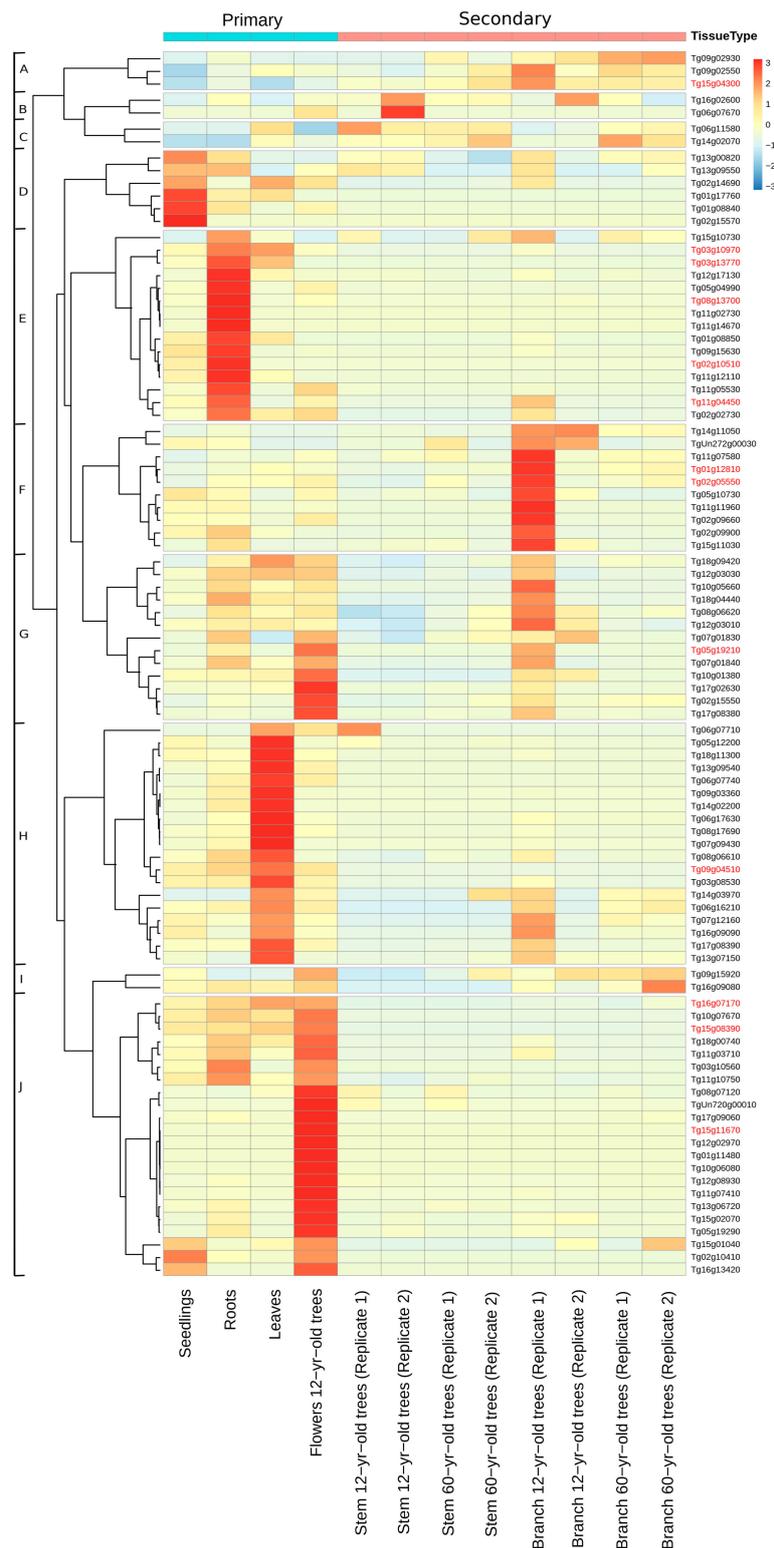
Figure 4. Hierarchical clustering of expression profiles of TgNAC genes in diverse *T. grandis* tissues. The 94 genes were clustered into 10 groups (**A–J**). Plant tissues (branch of 12- and 60-year-old trees; stem of 12- and 60-year-old trees, flowers, leaves, roots, and seedlings) were clustered according to the type tissue. The color scale on the left-hand side represents transcript per million (TPM). Gene expression values for each gene were transformed using Z-scores. Higher expression levels are represented in red, and lower expression levels are denoted in blue (Table S6). The TgNAC TF genes marked in red were selected for expression analysis.

Transcript abundance of the teak NAC genes in group A (*Tg09g02930-XND1*, *Tg09g02550-VNI1*, and *Tg15g04300-VNI2)* was higher in secondary tissues. These TgNAC genes are homologous to Arabidopsis NAC genes involved in the SCW biosynthesis (Figure 4). In group F, TgNAC genes were preferentially transcribed in branches of 12-year-old trees. Only the genes *Tg01g12810-ANAC075* and *Tg02g05550-SND2* were homologous to the Arabidopsis NAC genes recognized to be involved in secondary tissue formation. As for group G, there are 13 NAC genes highly expressed in secondary tissues, from which only two (*Tg0519210-NST1* and *Tg02g15550-SND2)* are homologous to Arabidopsis genes involved in SCW formation. Most genes in the remaining groups were preferentially expressed in primary tissues (Figure 4). These data reveal several *T. grandis* NAC genes that seem to be involved in secondary growth and that will be better characterized and understood in further studies.

In Figure 4, the NAC genes marked in red were selected for qRT-PCR expression analysis due to their homology to Arabidopsis NAC genes involved in SCW formation. They are (*VND1-Tg09g04510*; *VND2-Tg11g04450*; *VND4-Tg15g08390*; *VND4-Tg15g11670*; *VND4-Tg16g07170*; *VND7-Tg03g10970*; *NST1-Tg05g19210*; *VNI2-Tg15g04300*; *XND1-Tg08g13700*; *SND2-Tg02g15550*; *ANAC075-Tg01g12810*; *ANAC33-Tg02g10510*; and *ANAC70-Tg03g13770*). Several of these genes showed high transcript abundance in teak primary tissues (roots, leaves, and flowers), suggesting function divergence from the split between both lineages.

3.6. Expression Analysis of Marker Genes for Secondary Cell Wall Formation at Different Tree Ages

The expression pattern of teak NAC genes that are homologous to Arabidopsis genes involved in SCW biosynthesis was analyzed to verify their functionality during teak wood formation in tissues of trees of different ages. Tissue samples were analyzed by quantitative real-time RT-PCR (RT-qPCR) in order to verify the expression of 13 selected TgNAC TF genes in stem from 4-year-old teak plants and in sapwood of 16- and 64-year-old teak trees (Figure 5). These 13 genes were selected based on their regulatory activity in the SCW formation pathway and programed cell death, both important events during wood formation (Figure 6) [68].

The process of wood formation comprises a series of consecutive events, starting with cell division in the vascular cambium and then followed by cell expansion and secondary cell wall deposition. The wood formation process ends in programmed cell death and heartwood formation [68]. Several transcription factors regulate the entire wood formation process, mainly the ones of the NAC and MYB families [10]. Among the NAC transcription factors regulating SCW deposition, VASCULAR-RELATED NAC-DOMAIN1-7 (VND1–VND7 proteins) are the main regulators of xylem vessel cell differentiation (Figure 6) [69]. In the present work, we analyzed the expression of four homologs of this family in *T. grandis* (*VND1-Tg09g04510*; *VND2-Tg11g04450*; *VND4-Tg15g08390*, *Tg15g11670*, *Tg16g07170*; and *VND7-Tg03g10970*). The relative expression mean for these VND genes followed a similar pattern. Of all VND genes analyzed, the highest expression was observed in 16-year-old trees, followed by the 64-year-old ones (Figure 5a–f). The 4-year-old plants presented the lowest expression among the age groups of trees analyzed. However, only the relative expression of the genes *VND2-Tg11g04450* and *VND4-Tg15g08390* showed statistically significant differences among the samples of teak trees (Figure 5b,c). These data may reflect the role of VND transcription factors during wood formation in *T. grandis* trees during diverse developmental stages, which are consistent with other analyzed species of trees. For instance, the SCW deposition in *Populus* developing wood is first seen in the vessel elements and in contact cells, and only occurs in fibers later on [70]. Likewise, in *Eucalyptus,* the wall thickness of fibers is increased and vessel frequencies and numbers decrease to supply the mechanical and physiological requirements during the tree development [71,72]. In *Pinus radiate,* cell wall-related transcripts are more abundant during its fast-growing phase, which occurs in the early-growth stages (9 years old), when compared to mature-growth (30 years old) [73]. The highest expression of VND genes was found in 16-year-old trees, which indicates the time of high secondary cell wall deposition in the ages analyzed.
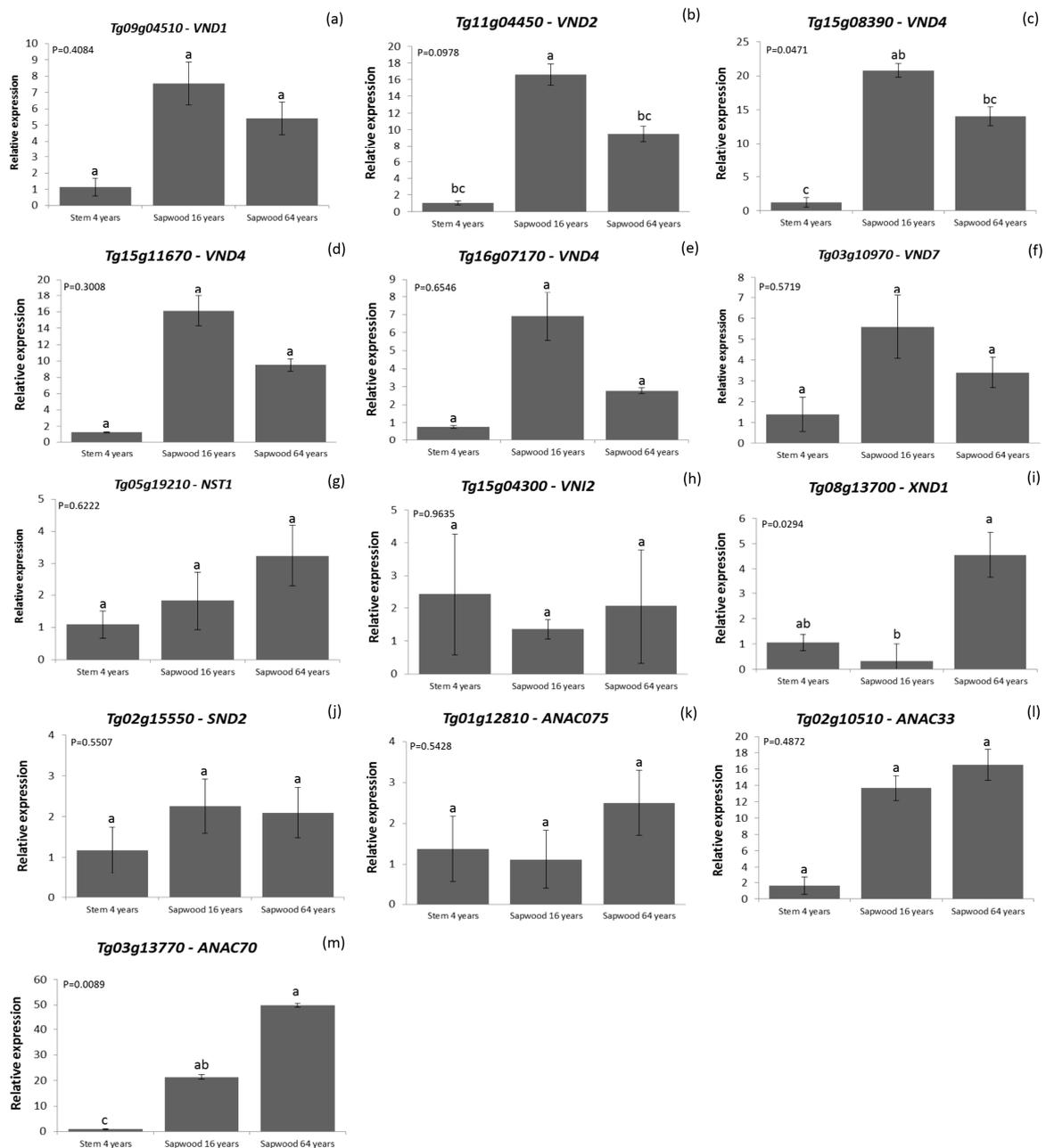
Figure 5. Expression levels of selected TgNAC genes using RT-qPCR. The relative expression of 13 selected NAC genes of *T. grandis* was normalized to the reference gene TgEF-1α in stem of 4-year-old trees and sapwood of 16- and 64-year-old trees. Bars represent standard errors (SEs) of three biological replicates and two technical replicates. The *p*-values are shown, and different letters represent statistically different means. Relative expression of the genes: (**a**) *VND1-Tg09g04510*, (**b**) *VND2-Tg11g04450*, (**c**) *VND4-Tg15g08390*, (**d**) *VND4-Tg15g11670*, (**e**) *VND4-Tg16g07170*, (**f**) *VND7-Tg03g10970*, (**g**) *NST1-Tg05g19210*, (**h**) *VNI2-Tg15g04300*, (**i**) *XND1-Tg08g13700*, (**j**) *SND2-Tg02g15550*, (**k**) *ANAC075-Tg01g12810*, (**l**) *ANAC33-Tg02g10510*, (**m**) *ANAC70-Tg03g13770*.

Expression was also analyzed for the homolog gene of *NAC SECONDARY WALL THICKENING PROMOTING FACTOR 1* (*NST1-Tg05g19210*). These NST1 NAC TFs have been associated with secondary cell wall deposition in xylem fibers of Arabidopsis [24,31]. However, the analysis of the expression of *Tg05g19210* showed no statistically significant difference among the secondary tissues

analyzed (Figure 5g). Therefore, it can be inferred that the *NST1–Tg05g19210* gene does not have a significative influence in SCW formation in teak.
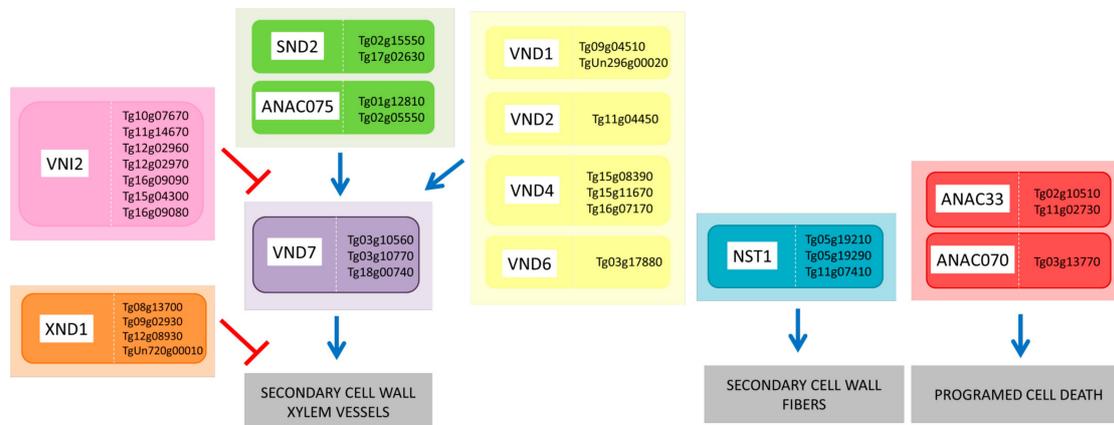

Figure 6. Schematic model of the transcriptional regulatory network controlling secondary cell wall biosynthesis in teak based in homologous genes of Arabidopsis. The NAC TFs VNDs (yellow and purple rectangles), SND2 (green rectangle), and ANAC075 (green rectangle) positively regulate SCW deposition in xylem vessels. TFs VNI2 (pink rectangle) and XND1 (orange rectangle) negatively regulate VND7 expression and differentiation of xylem vessels, respectively. NSTI (blue rectangle) is a transcriptional activator of SCW deposition in fibers, while ANAC33 and ANAC070 (red rectangles) are involved in programmed cell death. Blue and red arrows denote positive and negative regulation, respectively. Arabidopsis genes are represented in the white boxes and the putative homolog genes of teak are together in the same colored box.

In plant cells, the majority of NAC TFs are transcriptional activators, although transcriptional repressors are also present in the NAC gene family [74]. VND-INTERACTING2 (VNI2) is described in Arabidopsis as a transcriptional repressor that regulates differentiation of xylem cells by interacting with VND proteins and possibly with other NAC domain proteins [34]. We analyzed the expression of the *T. grandis Tg15g04300* gene, which is homologous to the *VNI2* in Arabidopsis (Figure 5h). There were no significant differences in the *Tg15g04300* expression among the teak trees of different ages. The VNI2 transcription factor represses the expression of genes regulated by VND7, which is a master regulator of xylem vessel differentiation [34]. In the analysis, also the expression of *VND7* homologue (*Tg03g10970*) was not altered during teak developmental stages, which indicates that perhaps VNI2 does not act as a repressor, or the repression may be occurring in all the teak stages analyzed.

Another repressor of secondary cell wall deposition is the XYLEM NAC DOMAIN1 (XND1) transcription factor (Figure 6). The gene of this TF is highly expressed in xylem and inhibits secondary cell wall deposition and autolysis in xylem vessels [75]. XND1 negatively regulates expression of genes involved in both programmed cell death and lignocellulose synthesis during secondary cell wall formation [29,76]. Interestingly, the expression of the teak repressor *Tg08g13700-XND1* (Figure 5i) was low in 4- and 16-year-old teak plants, but it was significantly higher in 64-year-old trees. These results corroborate the hypothesis that the secondary cell wall deposition is lower in older trees than in faster-growing younger trees. Apparently, the XND1 molecular mechanism of action is not related to VND TFs [77]. The NAC protein named SECONDARY WALL-ASSOCIATED NAC DOMAIN2 (SND2) is involved in the regulation of cellulose and hemicellulose biosynthesis (Figure 6). SND2 seems to occupy a subordinate place in the central layer of the transcriptional network of secondary cell wall formation [78]. There was no significant difference among the teak plants analyzed for the expression of the *Tg02g15550-SND2* gene (Figure 5j) among the teak plants analyzed, which could be acting in a secondary role in the secondary cell wall formation [78].

Another NAC transcription factor, the ARABIDOPSIS NAC DOMAIN CONTAINING PROTEIN 75 (ANAC075) is also involved in secondary cell wall biosynthesis and it is an upstream regulator of

VND7 [75]. *ANAC075* overexpression induced ectopic secondary cell wall formation in Arabidopsis tissues through increasing expression of *VND7* [75,79]. Similarly, in *Populus*, higher expression of *PNAC127* (homologous of *ANAC075*) was observed in differentiating xylem [6]. No difference in the expression of the teak *ANAC075* homolog *Tg01g12810* (Figure 5k) was observed in our analysis. Probably, the other NAC transcription factors have a stronger influence than *Tg01g12810* during the SCW formation in the *T. grandis* tissues analyzed.

We also analyzed TFs ANAC033 and ANAC070 that are known to regulate programmed cell death, the last step of secondary growth. The highest expression of the *ANAC70* teak homologue (*Tg03g13770)* was observed in 16- and 64-year-old trees (Figure 5m). This result indicates that this gene might have a key function in programed cell death of *T. grandis* secondary tissues. While, the evaluation of *ANAC33 (Tg02g10510)* homologue expression, showed no statistically significant difference between the three ages analyzed (Figure 5l). Apparently, these two Arabidopsis homologous proteins have a more prominent role in regulating lateral root cap formation than in regulation of secondary cell wall deposition [80].

4. Conclusions

Here we presented a genome-wide analysis of *T. grandis* NAC domain genes with identification and characterization of NAC transcriptional factors homologous to Arabidopsis and that are related to SCW formation. Analysis of teak NAC TFs encoding genes identified 130 NAC proteins, from which 116 represented non-redundant proteins coded by 107 loci. These identified 116 *T. grandis* NAC proteins were phylogenetically clustered into 23 distinct clades in evolutionary relationships with Arabidopsis, *Populus,* and *Eucalyptus* trees. Many of the TgNAC genes were located adjacent to each other, indicating that they might have resulted from duplication events. Most genes in groups A and F were preferentially expressed in secondary tissues, where they presented higher transcript abundance. As for the remaining groups, most NAC genes were highly expressed in flowers, roots, leaves, and seedlings. Considering the regulation function of the NAC genes in SCW biosynthesis, 13 NAC genes were analyzed for their expression in various tissues of three developmental stages of teak trees. Data from RT-qPCR revealed highest expression of VND genes *VND2-Tg11g04450* and *VND4-Tg15g08390* in tissues of 16-year-old trees. VND TFs are related to xylem vessels formation, indicating that these two VND genes (*Tg11g04450* and *Tg15g08390*) possibly participate in the regulation of *T. grandis* SCW deposition. In addition, the NAC homologous repressor *XND1* (*Tg08g13700*) presented highest expression in 64-year-old tissues and gave support to the hypothesis of lower secondary cell wall deposition in older trees. Data reported in the present work contribute to the understanding of structure and functionality of NAC transcription factor genes in *T. grandis* and they are potentially useful for future studies aiming to improve biomass and wood quality using biotechnological approaches.

References

1. Aida, M.; Ishida, T.; Fukaki, H.; Fujisawa, H.; Tasaka, M. Genes involved in organ separation in Arabidopsis: An analysis of the cup-shaped cotyledon mutant. *Plant Cell* **1997**, *9*, 841–857. [CrossRef]
2. Olsen, A.N.; Ernst, H.A.; Leggio, L.L.; Skriver, K. DNA-binding specificity and molecular functions of NAC transcription factors. *Plant Sci.* **2005**, *169*, 785–797. [CrossRef]
3. Jensen, M.K.; Kjaersgaard, T.; Nielsen, M.M.; Galberg, P.; Petersen, K.; O'Shea, C.; Skriver, K. The *Arabidopsis thaliana* NAC transcription factor family: Structure function relationships and determinants of ANAC019 stress signalling. *Biochem. J.* **2010**, *426*, 183–196. [CrossRef]
4. Duval, M.; Hsieh, T.F.; Kim, S.; Thomas, T.L. Molecular characterization of AtNAM: A member of the Arabidopsis NAC domain superfamily. *Plant Mol. Biol.* **2002**, *50*, 237–248. [CrossRef]
5. Hussey, S.G.; Saïdi, M.N.; Hefer, C.A.; Myburg, A.A.; Grima-Pettenati, J. Structural, evolutionary and functional analysis of the NAC domain protein family in Eucalyptus. *New Phytol.* **2015**, *206*, 1337–1350. [CrossRef] [PubMed]
6. Hu, R.; Qi, G.; Kong, Y.; Kong, D.; Gao, Q.; Zhou, G. Comprehensive Analysis of NAC Domain Transcription Factor Gene Family in *Populus trichocarpa*. *BMC Plant Biol.* **2010**, *10*, 145. [CrossRef] [PubMed]
7. Li, W.; Li, X.; Chao, J.; Zhang, Z.; Wang, W.; Guo, Y. NAC Family Transcription Factors in Tobacco and Their Potential Role in Regulating Leaf Senescence. *Front. Plant Sci.* **2018**, *9*, 1900. [CrossRef] [PubMed]
8. Hao, Y.J.; Wei, W.; Song, Q.X.; Chen, H.W.; Zhang, Y.Q.; Wang, F.; Zou, H.F.; Lei, G.; Tian, A.G.; Zhang, W.K.; et al. Soybean NAC transcription factors promote abiotic stress tolerance and lateral root formation in transgenic plants. *Plant J.* **2011**, *68*, 302–313. [CrossRef]
9. Friedman, W.E.; Cook, M.E. The origin and early evolution of tracheids in vascular plants: Integration of palaeobotanical and neobotanical data. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* **2000**, *355*, 857–868. [CrossRef]
10. Nakano, Y.; Yamaguchi, M.; Endo, H.; Rejab, N.A.; Ohtani, M. NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. *Front. Plant Sci.* **2015**, *6*, 288. [CrossRef]
11. Ko, J.H.; Jeon, H.W.; Kim, W.C.; Han, K.H. The MYB46/MYB83-mediated transcriptional regulatory programme is a gatekeeper of secondary wall biosynthesis. *Ann. Bot.* **2014**, *114*, 1099–1107. [CrossRef] [PubMed]
12. McCahill, I.W.; Hazen, S.P. Regulation of cell wall thickening by a medley of mechanisms. *Trends Plant Sci.* **2019**, *24*, 853–866. [CrossRef] [PubMed]
13. Lang, D.; Weiche, B.; Timmerhaus, G.; Richardt, S.; Riaño-Pachón, D.M.; Corrêa, L.G.; Reski, R.; Mueller-Roeber, B.; Rensing, S.A. Genome-wide phylogenetic comparative analysis of plant transcriptional regulation: A timeline of loss, gain, expansion, and correlation with complexity. *Genome Biol. Evol.* **2010**, *2*, 488–503. [CrossRef] [PubMed]
14. Popper, Z.A. Evolution and diversity of green plant cell walls. *Curr. Opin. Plant Biol.* **2008**, *11*, 286–292. [CrossRef]
15. Karanja, B.K.; Xu, L.; Wang, Y.; Muleke, E.M.; Jabir, B.M.; Xie, Y.; Zhu, X.; Cheng, W.; Liu, L. Genome-wide characterization and expression profiling of NAC transcription factor genes under abiotic stresses in radish (*Raphanus sativus* L.). *PeerJ* **2017**, *5*, e4172. [CrossRef] [PubMed]
16. Pérez-Rodríguez, P.; Riaño-Pachón, D.M.; Corrêa, L.G.; Rensing, S.A.; Kersten, B.; Mueller-Roeber, B. PlnTFDB: Updated content and new features of the plant transcription factor database. *Nucleic Acids Res.* **2010**, *38*, D822–D827. [CrossRef] [PubMed]
17. Zhang, G.; Li, T.; Zhang, L.; Dong, W.; Wang, A. Expression analysis of NAC genes during the growth and ripening of apples. *Hortscience* **2018**, *45*, 1–10. [CrossRef]
18. Kimotho, R.N.; Baillo, E.H.; Zhang, Z. Transcription factors involved in abiotic stress responses in Maize (*Zea mays* L.) and their roles in enhanced productivity in the post genomics era. *PeerJ* **2019**, *7*, e7211. [CrossRef]

19. Souer, E.; Von Houwelingen, A.; Kloos, D.; Mol, J.; Koes, R. The No Apical Meristem gene of petunia is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries. *Cell* **1996**, *85*, 159–170. [CrossRef]

20. He, X.J.; Mu, R.L.; Cao, W.H.; Zhang, Z.G.; Zhang, J.S.; Chen, S.Y. AtNAC2, a transcription factor downstream of ethylene and auxin signaling pathways, is involved in salt stress response and lateral root development. *Plant J.* **2005**, *44*, 903–916. [CrossRef]

21. Nakashima, K.; Takasaki, H.; Mizoi, J.; Shinozaki, K.; Yamaguchi-Shinozaki, K. NAC transcription factors in plant abiotic stress responses. *Biochim. Biophys. Acta (BBA)-Gene Regul. Mech.* **2012**, *1819*, 97–103. [CrossRef] [PubMed]

22. Hong, Y.; Zhang, H.; Huang, L.; Li, D.; Song, F. Overexpression of a stress-responsive NAC transcription factor gene ONAC022 improves drought and salt tolerance in rice. *Front. Plant Sci.* **2016**, *7*, 4. [CrossRef] [PubMed]

23. Podzimska-Sroka, D.; O'Shea, C.; Gregersen, P.; Skriver, K. NAC transcription factors in senescence: From molecular structure to function in crops. *Plants* **2015**, *4*, 412–448. [CrossRef] [PubMed]

24. Mitsuda, N.; Iwase, A.; Yamamoto, H.; Yoshida, M.; Seki, M.; Shinozaki, K.; Ohme-Takagi, M. NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of Arabidopsis. *Plant Cell* **2007**, *19*, 270–280. [CrossRef] [PubMed]

25. Mitsuda, N.; Ohme-Takagi, M. NAC transcription factors NST1 and NST3 regulate pod shattering in a partially redundant manner by promoting secondary wall formation after the establishment of tissue identity. *Plant J.* **2008**, *56*, 768–778. [CrossRef] [PubMed]

26. Li, Q.; Lin, Y.C.; Sun, Y.H.; Song, J.; Chen, H.; Zhang, X.H. Splice variant of the SND1 transcription factor is a dominant negative of SND1 members and their regulation in *Populus trichocarpa*. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 14699–14704. [CrossRef] [PubMed]

27. Zhang, J.; Huang, G.Q.; Zou, D.; Yan, J.Q.; Li, Y.; Hu, S. The cotton (*Gossypium hirsutum*) NAC transcription factor (FSN1) as a positive regulator participates in controlling secondary cell wall biosynthesis and modification of fibers. *New Phytol.* **2018**, *217*, 625–640. [CrossRef]

28. Chai, M.; Bellizzi, M.; Wan, C.; Cui, Z.; Li, Y.; Wang, G.-L. The NAC transcription factor OsSWN1 regulates secondary cell wall development in *Oryza sativa*. *J. Plant Biol.* **2015**, *58*, 44–51. [CrossRef]

29. Zhao, C.; Avci, U.; Grant, E.H.; Haigler, C.H.; Beers, E.P. XND1, a member of the NAC domain family in *Arabidopsis thaliana*, negatively regulates lignocellulose synthesis and programmed cell death in xylem. *Plant J.* **2008**, *53*, 425–436. [CrossRef]

30. Zhao, C.; Craig, J.C.; Petzold, H.E.; Dickerman, A.W.; Beers, E.P. The xylem and phloem transcriptomes from secondary tissues of the Arabidopsis root-hypocotyl. *Plant Physiol.* **2005**, *138*, 803–818. [CrossRef]

31. Zhong, R.; Demura, T.; Ye, Z.H. SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of Arabidopsis. *Plant Cell* **2006**, *18*, 3158–3170. [CrossRef] [PubMed]

32. Mitsuda, N.; Seki, M.; Shinozaki, K.; Ohme-Takagi, M. The NAC transcription factors NST1 and NST2 of Arabidopsis regulate secondary wall thickenings and are required for anther dehiscence. *Plant Cell* **2005**, *17*, 2993–3006. [CrossRef] [PubMed]

33. Bennett, T.; Toorn, A.V.D.; Sanchez-Perez, G.F.; Campilho, A.; Willemsen, V.; Snel, B.; Scheres, B. SOMBRERO, BEARSKIN1, and BEARSKIN2 regulate root cap maturation in Arabidopsis. *Plant Cell* **2010**, *22*, 640–654. [CrossRef] [PubMed]

34. Yamaguchi, M.; Ohtani, M.; Mitsuda, N.; Kubo, M.; Ohme-Takagi, M.; Fukuda, H.; Demura, T. VND-INTERACTING2, a NAC domain transcription factor negatively regulates xylem vessel formation in Arabidopsis. *Plant Cell* **2010**, *22*, 1249–1263. [CrossRef]

35. Diningrat, D.S.; Widiyanto, S.M.; Pancoro, A.; Shim, D.; Panchangam, B.; Zembower, N.; Carlson, J.E. Transcriptome of teak (*Tectona grandis* L.f) in vegetative to generative stage development. *Int. J. Innov. Appl. Stud.* **2014**, *9*, 1416–1427. [CrossRef]

36. Galeano, E.; Vasconcelos, T.S.; Vidal, M.; Mejia-Guerra, M.K.; Carrer, H. Large-scale transcriptional profiling of lignified tissues in *Tectona grandis*. *BMC Plant Biol.* **2015**, *15*, 221. [CrossRef]

37. Yasodha, R.; Vasudeva, R.; Balakrishnan, S.; Sakthi, A.R.; Abel, N.; Binai, N.; Rajashekar, B.; Bachpai, V.K.W.; Pillai, C.; Dev, S.A. Draft genome of a high value tropical timber tree, Teak (*Tectona grandis* L. f): Insights into SSR diversity, phylogeny and conservation. *DNA Res.* **2018**, *25*, 409–419. [CrossRef]

38. Zhao, D.; Hamilton, J.P.; Bhat, W.W.; Johnson, S.R.; Godden, G.T.; Kinser, T.J.; Boachon, B.; Dudareva, N.; Soltis, D.E.; Soltis, P.S.; et al. A chromosomal-scale genome assembly of *Tectona grandis* reveals the importance of tandem gene duplication and enables discovery of genes in natural product biosynthetic pathways. *GigaScience* **2019**, *8*, giz005. [CrossRef]

39. Myburg, A.A.; Grattapaglia, D.; Tuskan, G.A.; Hellsten, U.; Hayes, R.D.; Grimwood, J.; Jenkins, J.; Lindquist, E.; Tice, H.; Bauer, D.; et al. The genome of *Eucalyptus grandis*. *Nature* **2014**, *150*, 356–362. [CrossRef]

40. Tuskan, G.A.; Difazio, S.; Jansson, S.; Bohlmann, J.; Grigoriev, I.; Hellsten, U.; Putnam, N.; Ralph, S.; Rombauts, S.; Salamov, A.; et al. The genome of black cottonwood, Populus trichocarpa (Torr. *& Gray). Science* **2006**, *313*, 1596–1604. [CrossRef]

41. The Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **2000**, *408*, 796–815. [CrossRef] [PubMed]

42. Angiosperm Phylogeny Group. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot. J. Linn. Soc.* **2009**, *161*, 105–121. [CrossRef]

43. Novíkov, A.; Barabaš-Krasni, B. *Modern Plant Systematics*; Liga-Pres: Lviv, Ukraine, 2015; ISBN 078-966-397-276-3. [CrossRef]

44. Katoh, K.; Standley, D.M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef] [PubMed]

45. Capella-Gutiérrez, S.; Silla-Martínez, J.M.; Gabaldón, T. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **2009**, *25*, 1972–1973. [CrossRef]

46. Guindon, S.; Dufayard, J.-F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascue, O. New algorithms and methods to estimate Maximum-Likelihood Phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [CrossRef] [PubMed]

47. Sullivan, J.; Joyce, P. Model selection in phylogenetics. *Annu. Rev. Ecol. Evol Syst.* **2005**, *36*, 445–446. [CrossRef]

48. Hoang, D.T.; Chernomor, O.; Haeseler, A.; Minh, B.Q.; Vinh, L.S. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* **2017**, *35*, 518–522. [CrossRef] [PubMed]

49. Durand, D.; Hoberman, R. Diagnosing duplications—Can it be done? *Trends Genet.* **2006**, *22*, 156–164. [CrossRef]

50. Robinson, J.T.; Thorvaldsdóttir, H.; Wenger, A.M.; Zehir, A.; Mesirov, J.P. Variant Review with the Integrative Genomics Viewer (IGV). *Cancer Res.* **2017**, *77*, 31–34. [CrossRef]

51. Bailey, T.L.; Johnson, J.; Grant, C.E.; Noble, W.S. The MEME suite. *Nucleic Acids Res* **2015**, *43*, W39–W49. [CrossRef]

52. Bushnell, B.; Rood, J.; Singer, E. BBMerge—Accurate paired shotgun read merging via overlap. *PLoS ONE* **2017**, *12*, e0185056. [CrossRef] [PubMed]

53. Patro, R.; Duggal, G.; Love, M.I.; Irizarry, R.A.; Kingsford, C. *Salmon*: Fast and bias-aware quantification of transcript expression using dual-phase inference. *Nat. Methods* **2017**, *14*, 417–419. [CrossRef] [PubMed]

54. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2014; Available online: http://www.R-project.org/ (accessed on 18 December 2019).

55. Soneson, C.; Love, M.I.; Robinson, M.D. Differential analyses for RNA-seq: Transcript-level estimates improve gene-level inferences. *F1000Research* **2015**, *4*, 1521. [CrossRef] [PubMed]

56. Galeano, E.; Vasconcelos, T.S.; Ramiro, D.A.; De Martin, V.D.F.; Carrer, H. Identification and validation of quantitative real-time reverse transcription PCR reference genes for gene expression analysis in teak (*Tectona grandis* L.f.). *BMC Res. Notes* **2014**, *7*, 464. [CrossRef] [PubMed]

57. Xu, B.; Ohtani, M.; Yamaguchi, M.; Toyooka, K.; Wakazaki, M.; Sato, M.; Kubo, M.; Nakano, Y.; Sano, R.; Hiwatashi, Y.; et al. Contribution of NAC transcription factors to plant adaptation to land. *Science* **2014**, *343*, 1505–1508. [CrossRef] [PubMed]

58. Kovach, A.; Wegrzyn, J.L.; Parra, G.; Holt, C.; Bruening, G.E.; Loopstra, C.A.; Hartigan, J.; Yandell, M.; Langley, C.H.; Korf, I. The *Pinus taeda* genome is characterized by diverse and highly diverged repetitive sequences. *BMC Genom.* **2010**, *11*, 420:1–420:14. [CrossRef] [PubMed]

59. Nystedt, B.; Street, N.R.; Wetterbom, A.; Zuccolo, A.; Lin, Y.-C.; Scofield, D.G.; Vezzi, F.; Delhomme, N.; Giacomello, S.; Alexeyenko, A.; et al. The Norway spruce genome sequence and conifer genome evolution. *Nature* **2013**, *497*, 579–584. [CrossRef]

60. Shen, H.; Yin, Y.B.; Chen, F.; Xu, Y.; Dixon, R.A. A Bioinformatic analysis of NAC genes for plant cell wall development in relation to lignocellulosic bioenergy production. *BioEnergy Res.* **2009**, *2*, 217–232. [CrossRef]

61. Mittal, S.; Banduni, P.; Mallikarjuna, M.G.; Rao, A.R.; Jain, P.A.; Dash, P.K.; Thirunavukkarasu, N. Structural, functional, and evolutionary characterization of major drought transcription factors families in maize. *Front. Chem.* **2018**, *6*, 177. [CrossRef]

62. Ooka, H.; Satoh, K.; Doi, K.; Nagata, T.; Otomo, Y.; Murakami, K.; Matsubara, K.; Osato, N.; Kawai, J.; Carninci, P.; et al. Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*. *DNA Res.* **2003**, *10*, 239–247. [CrossRef]

63. Zhu, T.; Nevo, E.; Sun, D.; Peng, J. Phylogenetic analyses unravel the evolutionary history of NAC proteins in plants. *Evolution* **2012**, *66*, 1833–1848. [CrossRef] [PubMed]

64. Liu, X.; Wang, T.; Bartholomew, E.; Black, K.; Dong, M.; Zhang, Y.; Yang, S.; Cai, Y.; Xue, S.; Weng, Y.; et al. Comprehensive analysis of NAC transcription factors and their expression during fruit spine development in cucumber (*Cucumis sativus* L.). *Hortic. Res.* **2018**, *5*, 31. [CrossRef] [PubMed]

65. Pereira-Santana, A.; Alcaraz, L.D.; Castaño, E.; Sanchez-Calderon, L.; Sanchez-Teyer, F.; Rodriguez-Zapata, L. Comparative genomics of NAC transcriptional factors in angiosperms: Implications for the adaptation and diversification of flowering plants. *PLoS ONE* **2015**, *10*, e0141866. [CrossRef] [PubMed]

66. Mascher, M.; Gundlach, H.; Himmelbach, A.; Beier, S.; Twardziok, S.O.; Wicker, T.; Radchuk, V.; Dockter, C.; Hedley, P.E.; Russell, J.; et al. A chromosome conformation capture ordered sequence of the barley genome. *Nature* **2017**, *544*, 427–433. [CrossRef]

67. Akhunov, E.D.; Goodyear, A.W.; Geng, S.; Qi, L.L.; Echalier, B.; Gill, B.S.; Gustafson, J.P.; Lazo, G.; Chao, S.; Anderson, O.D.; et al. The organization and rate of evolution of wheat genomes are correlated with recombination rates along chromosome arms. *Genome Res.* **2003**, *13*, 753–763. [CrossRef]

68. De'jardin, A.; Laurans, F.; Arnaud, D.; Breton, C.; Pilate, G.; Leple, J.-C. Wood formation in Angiosperms. *C.R. Biol.* **2010**, *333*, 325–334. [CrossRef]

69. Kubo, M.; Udagawa, M.; Nishikubo, N.; Horiguchi, G.; Yamaguchi, M.; Ito, J.; Mimura, T.; Fukuda, H.; Demura, T. Transcription switches for protoxylem and metaxylem vessel formation. *Genes Dev.* **2005**, *19*, 1855–1860. [CrossRef]

70. Murakami, Y.; Funada, R.; Sano, Y.; Ohtani, J. The differentiation of contact cells and isolation cells in the xylem ray parenchyma of *Populus maximowiczii*. *Ann. Bot.* **1999**, *84*, 429–435. [CrossRef]

71. Bao, F.C.; Jiang, Z.H.; Jiang, X.M.; Lu, X.X.; Luo, X.Q.; Zhang, S.Y. Differences in wood properties between juvenile wood and mature wood in 10 species grown in China. *Wood Sci. Technol.* **2001**, *35*, 363–375. [CrossRef]

72. Sette Jr, C.R.; Oliveira, I.R.; Tomazello Filho, M.; Yamaji, F.M.; Laclau, J.P. Effect of age and sampling position on density and anatomical characteristics of wood in eucalyptus trees. *Rev. Árvore* **2012**, *36*, 1183–1190. [CrossRef]

73. Li, X.; Wu, H.X.; Southerton, S.G. Seasonal reorganization of the xylem transcriptome at different tree ages reveals novel insights into wood formation in *Pinus radiate*. *New Phytol.* **2010**, *187*, 764–776. [CrossRef] [PubMed]

74. Ko, J.-H.; Yang, S.H.; Park, A.H.; Lerouxel, O.; Han, K.-H. ANAC012, a member of the plant-specific NAC transcription factor family, negatively regulates xylary fiber development in *Arabidopsis thaliana*. *Plant J.* **2007**, *50*, 1035–1048. [CrossRef] [PubMed]

75. Endo, H.; Yamaguchi, M.; Tamura, T.; Nakano, Y.; Nishikubo, N.; Yoneda, A.; Kato, K.; Kubo, M.; Kajita, S.; Katayama, Y.; et al. Multiple classes of transcription factors regulate the expression of *VASCULAR-RELATED NAC-DOMAIN7*, a master switch of xylem vessel differentiation. *Plant Cell Physiol.* **2015**, *56*, 242–254. [CrossRef] [PubMed]

76. Yamaguchi, M.; Demura, T. Transcriptional regulation of secondary wall formation controlled by NAC domain proteins. *Plant Biotechnol.* **2010**, *27*, 237–242. [CrossRef]

77. Zhao, C.; Lasses, T.; Bako, L.; Kong, D.; Zhao, B.; Chanda, B.; Bombarely, A.; Cruz-Ramirez, A.; Scheres, B.; Brunner, A.M.; et al. XYLEM NAC DOMAIN1, an angiosperm NAC transcription factor, inhibits xylem differentiation through conserved motifs that interact with RETINOBLASTOMA-RELATED. *New Phytol.* **2017**, *216*, 76–89. [CrossRef]

78. Hussey, S.G.; Mizrachi, E.; Spokevicius, A.V.; Bossinger, G.; Berger, D.K.; Myburg, A.A. SND2, a NAC transcription factor gene, regulates genes involved in secondary cell wall development in Arabidopsis fibres and increases fibre cell area in Eucalyptus. *BMC Plant Biol.* **2011**, *11*, 173. [CrossRef]

79. Sakamoto, S.; Mitsuda, N. Reconstitution of a secondary cell wall in a secondary cell wall-deficient Arabidopsis mutant. *Plant Cell Physiol.* **2015**, *56*, 299–310. [CrossRef]

80. Zhong, R.; Lee, C.; Ye, Z.H. Functional characterization of Poplar wood-associated NAC domain transcription factors. *Plant Physiol.* **2010**, *152*, 1044–1055. [CrossRef]