



OPEN

Transcriptomic analysis of cork during seasonal growth highlights regulatory and developmental processes from phellogen to phellem formation

Sandra Fernández-Piñán, Pau Boher, Marçal Soler, Mercè Figueras & Olga Serra

The phellogen or cork cambium stem cells that divide periclinally and outwardly specify phellem or cork. Despite the vital importance of phellem in protecting the radially-growing plant organs and wounded tissues, practically only the suberin biosynthetic process has been studied molecularly so far. Since cork oak (*Quercus suber*) phellogen is seasonally activated and its proliferation and specification to phellem cells is a continuous developmental process, the differentially expressed genes during the cork seasonal growth served us to identify molecular processes embracing from phellogen to mature differentiated phellem cell. At the beginning of cork growth (April), cell cycle regulation, meristem proliferation and maintenance and processes triggering cell differentiation were upregulated, showing an enrichment of phellogenic cells from which phellem cells are specified. Instead, at maximum (June) and advanced (July) cork growth, metabolic processes paralleling the phellem cell chemical composition, such as the biosynthesis of suberin, lignin, triterpenes and soluble aromatic compounds, were upregulated. Particularly in July, polysaccharides- and lignin-related secondary cell wall processes presented a maximal expression, indicating a cell wall reinforcement in the later stages of cork formation, presumably related with the initiation of latecork development. The putative function of relevant genes identified are discussed in the context of phellem ontogeny.

Abbreviations

DEG Differentially expressed gene
SCW Secondary cell wall

Radial growth (secondary growth) of plant organs, either herbaceous or woody plants, angiosperms or gymnosperms, involves the activation of two secondary meristems: the cambium (or vascular cambium) and phellogen (or cork cambium). Both meristems constitute a bifacial stem cell population from which derivatives are formed by periclinal divisions and specified on opposing sides. Secondary growth is initiated by vascular cambium, which generates vascular tissue (xylem inwardly and phloem outwardly) and results in stem and root thickening. Due to the vascular radial growth, the primary protective tissues (epidermis and endodermis) break and a phellogen formed inwardly develops a complete periderm conferring protection to the mature organs¹. Periderm is also generated in the exposed surfaces of the abscission zones of plants, and in response to mechanical injury or parasite attack forming a wound periderm².

The periderm ontogeny is initiated by the periclinal division of the phellogen stem cells, which provide the derivatives at opposing sides that will specify to phellem (outwardly) and phellogen (inwardly). The three layers, phellogen, phellem (or cork) and phellogen constitute the periderm². The newly formed phellem cell elongates outwardly in the radial direction while its cell wall thickens as a consequence of accumulating structural components. In addition to the polysaccharide primary cell wall, the phellem cell wall deposits suberin and lignin polymers, and also some soluble secondary metabolites such as extractives (fatty-acyl derived and triterpenes) and tannins^{3,4}. This specialized cell wall provides a waterproofing layer that in turn isolate the cell itself from

Laboratori del Suro, Departament de Biologia, Universitat de Girona, Campus Montilivi, 17003 Girona, Spain. email: olga.serra@udg.edu

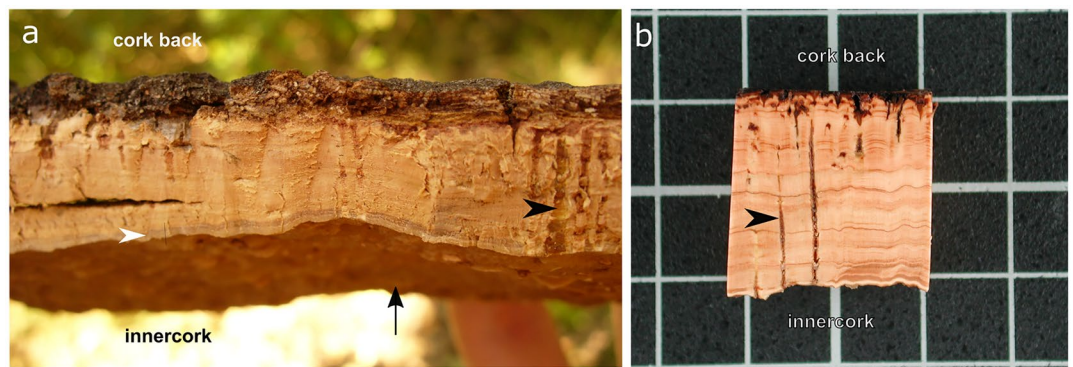


Figure 1. Anatomy of cork plank (outer bark) from *Quercus suber*. **(a)** Cork freshly-harvested at the late spring (June) in Romanyà de la Selva (Girona, Spain) showing this year produced living tissue (white arrowhead) in the innercork side, which faces the phellogen (black arrow). **(b)** A dried and polished sample is shown to better appreciate the annual cork-rings seen as lighter (earlycork) and darker brown (latecork) zones. Lenticels are highlighted with a black arrowhead. The grid scale distance in B corresponds to 1 cm. The figure was constructed using the free and open source Inkscape software 1.0 (<https://inkscape.org/>).

water and nutrients, thus dying at maturity. These phellem cells are usually forming a multilayered tissue which provide altogether the protective function².

In most woody plants, the-each-year periderm is replaced by a new periderm formed underneath the previous one. However, a few exceptions exist, notably cork oak (*Quercus suber*), which naturally produces a unique and single periderm throughout the entire tree life. The cork oak permanent and continuous phellogen cylinder presents seasonal growth, which through activation-inactivation series adheres the new formed phellem cells to the death phellem cells of the previous years⁵, hence providing a pure outer bark of thick phellem/cork. The fact that the phellogen layer forms a continuous cylinder lengthwise the tree trunk allows the cork harvesting as an entire plank by applying a moderate tensile force. This is possible when phellogen is fully active and has produced a set of phellem cells with also thin and fragile cell walls (usually June)^{6,7}. In the plank, the phellogen and new phellem cells are in the innercork side and the old and death phellem cells of previous years are extended outwardly to the cork back side (Fig. 1). The first cork produced by cork oak, known as virgin cork, is irregular in structure, density and thickness and is hard-rough. The following harvested corks, extracted sequentially every 9–12 years, the time required to allow the tree to re-grow its outer bark, are known as reproduction cork. Although all the cork planks are exploited commercially as raw material for *i.e.* agglomerates, only the third (2nd reproduction cork) and successive harvests can meet the regular structure with smooth and unblemished bark to present the quality required for stopper production. This best quality reproduction cork is termed amadia cork⁸.

The seasonal growth of cork, starting around April, produces two cork cell populations. From April to July cork expands at a high rate producing the earlycork⁹, being around June the highest growth rate. At the cellular level, it is when the phellogen and phellem are fully active: the rate of cell division and cell expansion is high and the resulting phellem cells (earlycork cells) will be numerous (40 to 200 cells per annual row), long and with thin walls. Advancing to the end of the growing season, the cell division and expansion rates decrease, resulting in the formation of few cork cells (4–8 cells per annual row) with thicker walls (latecork cells)^{4,10}. These phellem cellular differences result in the observation of annual-cork rings (Fig. 1), which includes the earlycork and latecork subsequently accumulated in successive years¹¹.

Most of the processes essential for phellem formation and development are mainly unknown. However, a few genes involved in suberin biosynthesis, transport and regulation, mostly identified as preferentially expressed in cork oak or poplar phellem^{5,12}, have been functionally characterized in potato tuber phellem (skin) or in *Arabidopsis* root endodermis and seeds¹³. This has evidenced that most suberized tissues, including phellem, use the same protein network. In the same line, phellogen and vascular cambium, both bifacial secondary meristems with seasonal growth, commonly upregulates a set of specific genes not present in other tree trunk tissues¹⁴, pointing that both meristems share molecular mechanisms for its formation and maintenance. The shared mechanisms have been evidenced very recently, since the vascular cambium master regulators WUSCHEL-RELATED HOMEBOX 4 (WOX4) and KNAT1/BREVIPEDICELLUS (BP) were also found to control the *Arabidopsis* root phellogen activity¹⁵.

In this work, we wanted to understand the molecular mechanisms expanding from the phellogen derivative formation to the phellem cell specification and differentiation. To that, we analysed the transcriptome of the newly formed cork cells (living phellem cells or phelloid) isolated from innercork planks at three different time points of the growing season: at the beginning of the cork growing season (April), at the maximum cork growth rate (June)^{6,7} and at advanced stage (July). The transcriptome seasonal variation was used as a readout of the molecular processes involved in phellem cell ontogeny. The main processes and candidate genes highlighted are involved in the phellogen stem cell identity and phellem differentiation, while others may be key players in the phellem cell wall modification.

Results

High-throughput gene expression profiling of innercork along the growing season. To understand the changes produced in cork transcriptome during the growing season, cork samples were collected in April, June and July from three to four different trees (biological replicates). Different trees were used for April, June and July sample collection, amounting to a total of 11 trees. The eleven corresponding RNA samples were sequenced by Illumina HiSeq2500 sequencing technology. The tissue used for RNA extraction was the newly-produced and living cork tissue from the innerbark of the plank, adjacent to the phellogen tissue (Fig. 1). A total of 568×10^6 reads with an average length of 122 bp were obtained. On average, 81% of the reads uniquely mapped against the cork oak genome (GCF_002906115.1 (CorkOak1.0))¹⁶, identifying 32,024 *Q. suber* active loci (Supplemental Table 1). For each gene, the normalized read count was calculated obtaining for the whole transcriptome the gene expression profile along the cork growing season (Supplemental Data Set 1). For functional characterization, 30,676 *Q. suber* genes were annotated to the corresponding Arabidopsis most similar gene, which has been considered the homolog throughout the manuscript (TAIR10 protein database; BLASTP E-value $< 10^{-5}$) (Supplemental Table 1 and Supplemental Data Set 1).

To validate the expression values of RNA-seq data, the \log_2 Fold-Change (FC) values obtained in this RNA-seq experiment for 29 genes representative of processes described for cork development were compared with the \log_2 FC values obtained from new and previously published RT-qPCR analyses which used the same RNA samples (Supplemental Table 2)^{7,17}. A significant Pearson correlation ($p < 0.001$) in all pair-wise comparisons (June/April, July/April and July/June) was observed (Supplemental Figure 1) validating the RNA-seq results.

The innercork samples contained transcripts from phellem and phellogen. When the cork plank is extracted, it breaks along the phellogenic tissue, therefore the RNA extracted was hypothesized to contain transcripts from living phellem cells but also phellogen. To understand at which extent phellem and phellogen were represented in the tissue collected, we seek in the cork transcriptome for those genes described to be specifically and differentially expressed in phellem and phellogen/phelloderm of *Betula pendula* (birch) tree trunk¹⁴, using the corresponding Arabidopsis gene with higher similarity. The cork transcriptome expressed the 88.17% of genes identified as differentially expressed in birch phellem, but also the 79.90% of the phellogen/phelloderm ($\text{padj} < 0.01$ and a fold change $-1 > \log_2\text{FC} > 1$) (Supplemental Table 3). These results suggest that the cork samples we used also contained phellogen cells.

Most genes involved in suberin deposition are induced in June. Since suberin is a key structural component deposited in the phellem cell wall during its development, we followed a targeted approach in which the expression pattern of suberin related genes was evaluated along the growing season. First, we identified in the cork transcriptome 68 *Q. suber* genes that corresponded to 25 Arabidopsis homologs involved in suberin and/or suberin-associated wax accumulation classified on AraLip database (<http://aralip.plantbiology.msu.edu/pathways/pathways>) or reported more recently in literature (Supplemental Table 4A). These data suggest putative duplications of these genes in *Q. suber* genome. The identified suberin-related genes, based on the enzymatic or regulatory function of the encoded protein, were depicted and classified within cell compartments and metabolic pathways/processes active in a suberizing cell, namely fatty acid elongation complex (FAE), fatty acid modification, transport to the cell wall, phenylpropanoid pathway and transcription factors (Supplemental Table 4A and Fig. 2). For each *Q. suber* transcript, a heatmap shows the relative expression along the cork growing season (April, June and July) and the total expression.

Importantly, 58.33% (21 of 36) of the genes with highest expression (> 760 normalized read counts) in cork, including the 15 of the top 16, showed a maximum of expression in June, and included genes involved in suberin biosynthesis (CYP86A1, CYP86B1, GPAT5, KCS2, KCS1 and ASFT), polymerization (GELP38, GELP51 and GELP96), transport to the cell wall (ABCG1, ABCG2, ABCG11, LTPG5 and LTPG15) and positive regulation (MYB9) (see Supplemental Table 4A for references). These results point that, although the phellem cells are committed to suberin production along the entire growing season, it is in June when there is a more intense suberization since cork tissue is enriched in phellem cells actively suberizing.

To seek for new candidate suberin genes, the complete AraLip database was compared with the cork transcriptome. We identified 207 *Q. suber* genes (110 Arabidopsis homologs) not previously related with suberin (Supplemental Table 4B), which were classified into the previously mentioned suberin functional categories and the fatty acid precursor synthesis in the plastids (Supplemental Figure 2). Based on the expression profile of suberin-known genes, the best candidates were those showing highest expression in cork tissue with a maximum in June.

Differential expression analyses highlight seasonal transcriptional patterns during cork formation. To identify the genes regulated during the growing season, the gene expression at the different cork growing stages were pair-wised (June/April, July/April and July/June) compared using the DESeq2 statistical package. DEGs were selected using a filter of $\text{padj} < 0.01$ and absolute \log_2 Fold-Change > 1 . While the June/July comparison showed only 80 DEGs, the June/April and July/April comparisons resulted in a considerably higher number of DEGs: 1,682 and 1,822, respectively (Fig. 3A). These results indicated that cork samples of June and July were transcriptionally similar and that differed substantially from those of April. In agreement, in June/April and July/April comparisons 928 were overlapping DEGs and practically all (927) were equally regulated, 529 upregulated and 398 downregulated (Fig. 3A). The expression of all DEGs identified was used to construct a hierarchical clustering of cork samples (Fig. 3B) in which the biological replicates for each month (April, June and July) grouped separately, evidencing a good replicability of the samples. Besides, the clade of June and July samples clustered together, evidencing again a greatest similarity of these months' transcriptomes compared to those of April.

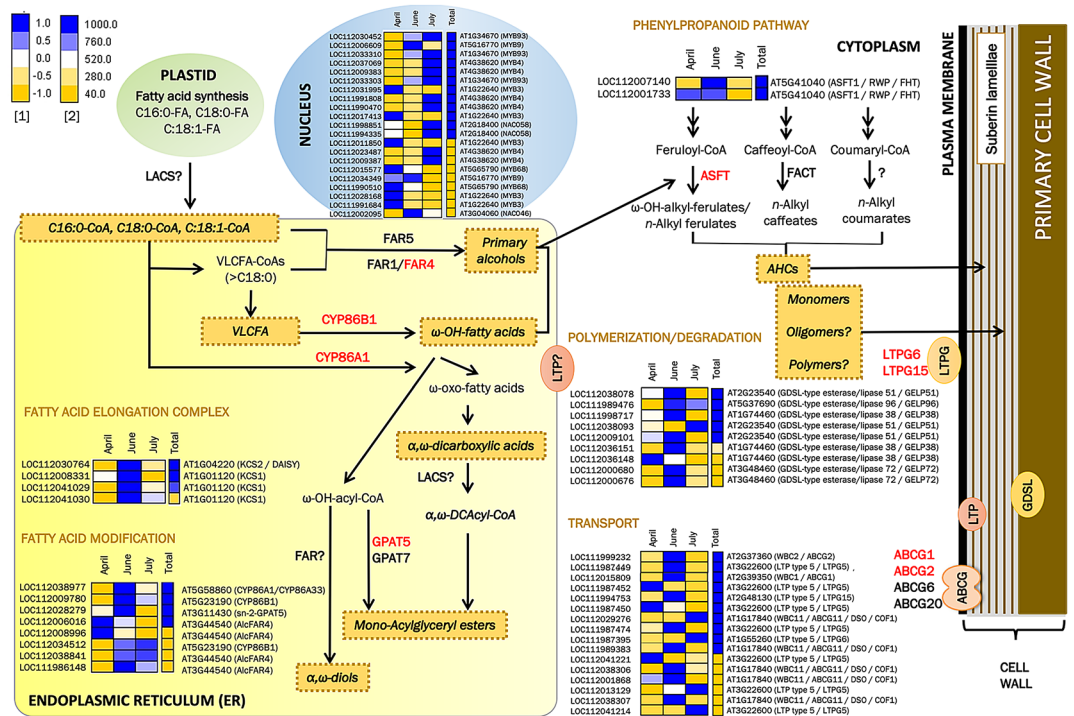


Figure 2. Expression pattern along the cork growing season of genes involved in suberin-related processes. For each gene, the relative expression pattern (April, June and July)¹ and the total expression² of normalized read counts are shown as a heatmap, in which blue and yellow colours represent high and low expression values, respectively. The genes with demonstrated function identified in the cork transcriptome are highlighted in red. AHCs = alkyl hydroxycinnamates; VLCFA = very long chain fatty acid. The heatmaps were constructed using the EXPANDERS software⁷⁵.

DEGs transcript profiles of cork samples in April, June and July were grouped in 4 clusters (Fig. 3C). Clusters 2 and 4 contained genes upregulated in April and downregulated in June and July. In contrast, clusters 1 and 3 contained genes downregulated in April and upregulated in June and July, showing cluster 3 a highest variation in gene expression between months and cluster 1 a slight increase in expression in July. Subsequently, GO enrichment analysis was performed to highlight the most important biological processes, molecular functions and cell components within each cluster (Supplemental Tables 5–8).

Regulatory and abiotic stress genes are induced at the beginning of the growing season (April). Cluster 2 included the 37.0% of DEGs and was enriched for GO terms associated with (i) response to abiotic stimulus; (ii) metabolic process mainly related with nucleic acid metabolism; (iii) regulation of cell cycle; (iv) transcription factor activity; (v) nucleus compartment; (vi) structural constituent of ribosome and (vii) methyltransferase activity (Supplemental Table 5). The genes in this cluster that were related with GO terms of transcription factor activity, DNA binding and regulation of cell cycle are presented in Supplemental Data Set 2. The set of 114 transcription factors and/or DNA binding proteins were subjected to a secondary GO enrichment analysis to highlight the most important processes in which they were participating. Thirty-four genes were classified in reproductive system development and flower development, 14 in chromosome organization and chromatin organization (6 also in methylation), 9 in meristem development (6 also in meristem maintenance), 5 in cell proliferation, 28 in cell communication, 22 in cell differentiation, 6 in leaf development, 5 in plant epidermis development, 10 in root development and 49 were not included in any of these enriched categories (Supplemental Data Set 2). In Table 1, we highlight some of the candidates to regulate (i) the phellogen meristem identity and maintenance, (ii) the differentiation of phellogen derivatives to phellem cells and (iii) the cell cycle regulation. To narrow those candidates that may have a specific function in phellogen or a shared function with vascular cambium (both secondary tissues), we cross-checked the data with the secondary meristem genes that has been recently described to be upregulated in vascular cambium and/or phellogen/phellogen of birch tree (*Betula pendula*)¹⁴, with candidates of phellogen of potato (*Solanum tuberosum*) tuber¹⁸ and with the candidates for vascular cambium in Arabidopsis root¹⁹. This information was included in Supplemental Data Set 2. Selected *Q. suber* phellogen candidates discussed in Discussion section are reported in Table 1.

Cluster 4 included 11.3% of DEGs and was enriched in GOs associated with (i) response to biotic and abiotic stimulus; (ii) ADP and carbohydrate derivative binding; and (iii) cell periphery (Supplemental Table 6).

Metabolism and abiotic stress processes dominate in the cork cell transcriptome at the maximum (June) and advanced (July) seasonal growth stages. Cluster 3 included the 40.8% of DEGs

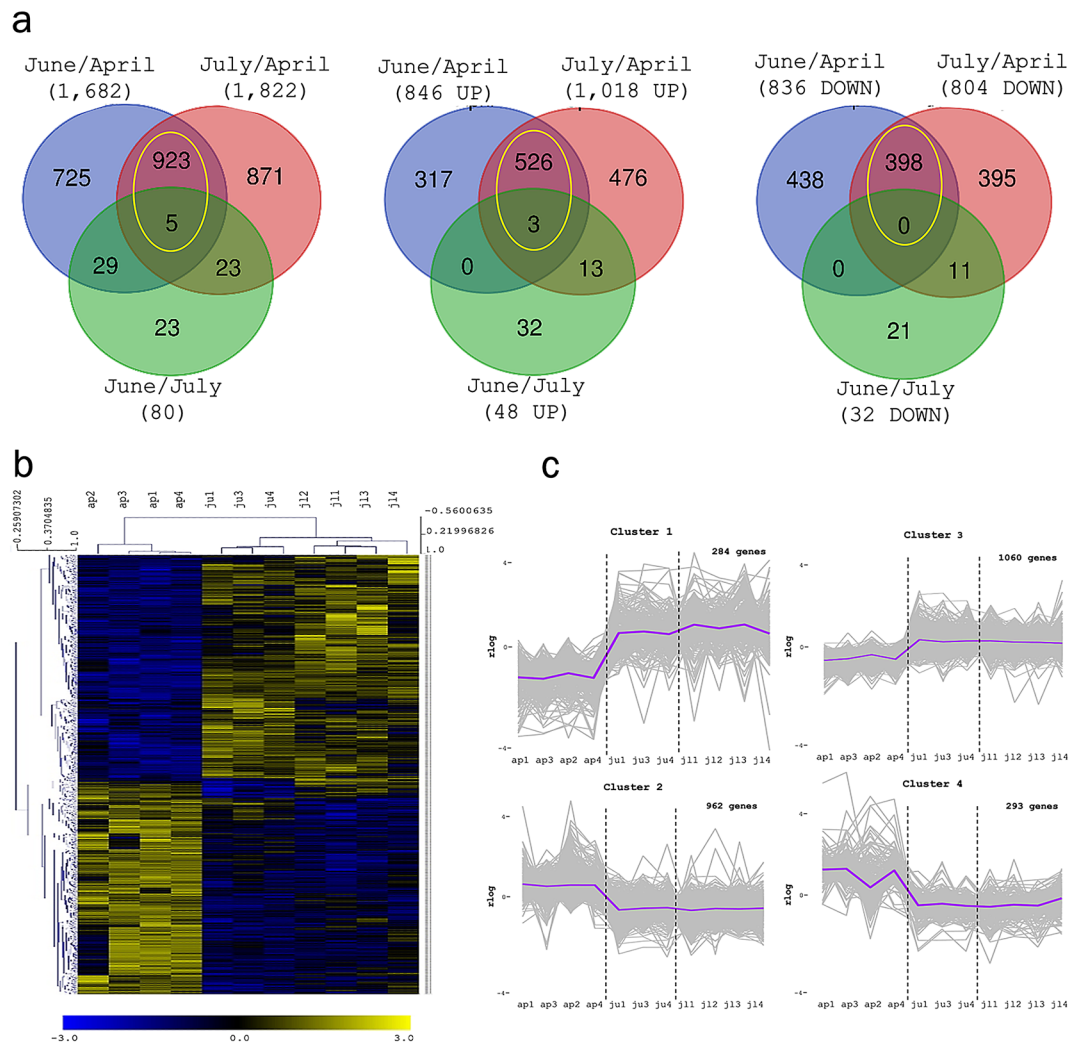


Figure 3. Differential expression along the cork growing season. **(a)** Venn diagrams showing DEGs, upregulated (UP) and downregulated (DOWN), between June/April, July/April and June/July. The yellow circle includes the genes identified in both June/April and July/April comparisons. **(b)** Hierarchical clustering of cork samples based on the expression of DEG along the growing season. **(c)** K-means clusters of co-expressed genes in April, June and July. Cluster 1 and 3 contain genes downregulated in April and upregulated in June and July, whereas cluster 2 and 4 showed the opposite pattern containing genes upregulated in April and downregulated in June and July.

and was enriched for GO terms associated with (i) general metabolic processes; (ii) secondary metabolic processes related to lipids, terpenoids, phenylpropanoids and oxylipins; (iii) response to abiotic and biotic stimulus, and in response to salicylic and jasmonic acid; (iv) oxidation–reduction process; (v) chloroplast RNA modification and chloroplast/plastid compartment; (vi) ion binding; (vii) kinase activity in relation to ATP and carbohydrate binding; (viii) membrane, especially plasma membrane, but also (ix) cytoplasm (Supplemental Table 7). Interestingly, pathways upstream of suberin, lignin and cork extractives biosynthesis, such as fatty acids, triterpenes, long chain aliphatic compounds, tannins and other aromatic compounds²⁰ were all co-upregulated in this cluster.

Cluster 1 included the 10.9% of DEGs and was enriched for GO terms associated with (i) response to stimulus such as light and heat/cold temperature -mostly containing heat shock proteins-, oxidative stress -mainly that generated by hydrogen peroxide-, cadmium, water deprivation and biotic stimulus; (ii) oxidation–reduction process and oxidoreductase activity; (iii) cell wall biogenesis, including lignin and phenylpropanoid metabolism, polysaccharide and carbohydrate-related, the later in agreement with the enriched UDP-glycosyltransferase activity necessary for carbohydrate biosynthesis; (iv) iron and other ion binding; (v) apoplast; (vi) plasma membrane; and (vii) Golgi apparatus, a very active compartment in cell wall matrix polysaccharide biosynthesis and export to the cell wall (Supplemental Table 8). Overall, most of the genes included in this cluster are related to cell wall biogenesis including lignin and polysaccharide biosynthesis and export to the cell wall. Surprisingly, almost 67% (12 from 18) of the genes included in the cell wall biogenesis category were described to contribute to the secondary cell wall thickening of the xylem cells (Table 2) despite that phellem cells, although lignified,

AGI code	Gene name	Gene function	Differentially expressed <i>Q. suber</i> identifier	Normalized read counts mean (BaseMean)	Log ₂ FC June/April (DEG = in bold)	Log ₂ FC July/April (DEG = in bold)
Meristem development						
AT4G37750	AINTEGUMENTA (ANT)	Promotion of stem cell proliferation ¹	LOC111985236	5976.93	-1.67	-0.73
AT4G28190	ULTRAPETALA1 (ULT1)	Meristem-restricting factor	LOC112024131	5162.97	-1.34	-1.39
AT4G08150	KNAT1/BREVIPEDI-CELLUS (BP)	Master regulator of cambial activity ¹ and phellogen regulator	LOC112015571 ³	3784.91	-0.98	-0.84
AT1G46480	WUSCHEL related homeobox 4 (WOX4)	Master regulator of cambial activity ¹ and phellogen regulator	LOC111988759	2676.09	-1.29	-0.68
AT1G62360	SHOOT MERISTEM-LESS (STM)	Prevents stem cells differentiation ²	LOC111993990	1942.25	-1.64	-1.56
AT4G22950	AGAMOUS-like 19 (AGL19)	Regulation of timing of transition from vegetative to reproductive phase	LOC111989313	1622.10	-1.13	-1.45
AT1G16060	WRINKLED 3 (WRI3)	Positive regulator of cutin and fatty acid biosynthesis ²	LOC112006985	659.92	-0.51	-1.28
AT1G63100	SCARECROW-LIKE 28 (SCL28)	Required for cambial activity ¹	LOC111999471	590.00	-1.59	-0.64
AT5G10510	AINTEGUMENTA-like 6 (AIL6)	Promotes of stem cell proliferation	LOC112018242	529.45	-1.18	-0.90
AT4G36920	APETALA 2 (AP2)	Promotes the expression of WUS	LOC111995790	528.69	-1.46	-1.38
AT2G34710	PHABULOSA (PHB)	Stem cell identity maintenance	LOC112031480	485.93	-1.54	-1.10
AT2G43000	NAC domain containing protein 42 (NAC042)	Negative regulator of leaf senescence ²	LOC112006470	211.59	-1.71	-0.89
AT2G17950	WUSCHEL (WUS)	Stem cell identity maintenance	LOC111987351	39.35	-1.85	-1.12
AT1G31320	LOB domain-containing protein 4 (LBD4)	Cambial regulator of vascular patterning ¹	LOC112038775	13.65	-2.08	-2.87
Cell differentiation						
AT5G57620	Myb domain protein 36 (MYB36)	Promotion of Casparian Strip lignification and endodermal suberization	LOC112023410	1943.17	-0.63	-1.06
			LOC112023896	1875.43	-0.86	-1.73
AT3G01140	Myb domain protein 106 (MYB106)	Regulation of cuticular lipid-deposits in epidermal cells	LOC111998628	1329.04	-0.65	-1.54
AT4G38620	Myb domain protein 4 (MYB4)	Repressor of phenylpropanoid biosynthesis (lignin and suberin)	LOC111991808	823.27	-1.10	-0.43
AT2G23380	CURLY LEAF (CLF)	Polycomb-group protein with repressive activity ⁴	LOC111988803	784.64	-1.36	-1.27
AT1G30970	SUPPRESSOR of FRI 4 (SUF4)	Negative regulation of flower development ⁴	LOC111995697	519.78	-1.10	-0.69
AT1G68320	Myb domain protein 62 (MYB62)	Unknown	LOC112008278	368.82	-2.54	-1.86
AT5G41315	GLABROUS 3 (GL3)	Promotion of epidermis differentiation (root hair and nonhair cell fate)	LOC111986568	350.05	-1.41	-0.71
AT1G18800	NAP1-related protein 2 (NRP2)	Promotion of cell cycle progression at G2/M and cellular organization, regulation of lateral root formation ⁴	LOC111996871	323.02	-2.13	-1.50
AT3G06740	GATA transcription factor 15 (GATA15)	Unknown	LOC112028517	263.08	-1.13	-0.49
AT4G00150	LOST MERISTEMS 3 (LM3)/HAIRY MERISTEM 3 (HAM3)	Promotion of cell differentiation at the periphery of shoot meristems and help to maintain their polar organization.	LOC111988127	199.06	-1.31	-0.74
AT4G18710	BRASSINOSTEROID INSENSITIVE 2 (BIN2)	GSK3 targeting the positive regulator of xylem cell differentiation BES1	LOC111992186 ⁵	121.03	-0.96	0.40
AT5G14750	Myb domain protein 66 (MYB66)/WEREWOLF	Promotion of epidermis differentiation (root hair and nonhair cell fate)	LOC111999826	107.31	-1.96	-0.13
Continued						

AGI code	Gene name	Gene function	Differentially expressed <i>Q. suber</i> identifier	Normalized read counts mean (BaseMean)	Log ₂ FC June/April (DEG = in bold)	Log ₂ FC July/April (DEG = in bold)
AT2G46410	CAPRICE (CPC)	Promotion of epidermis differentiation (root hair and nonhair cell fate)	LOC111988432	106.42	-2.02	-0.69
AT4G18390	TEOSINTE BRANCHED 1, cycloidea and PCF transcription factor 2 (TCP2)	Regulation of leaf cell differentiation and secondary shoot formation	LOC112024464	49.96	-2.32	-0.63
AT1G22640	Myb domain protein 3 (MYB3)	Repressor of phenylpropanoid biosynthesis (lignin and suberin)	LOC112011850	47.10	-2.08	-0.60
AT4G25560	Myb domain protein 18 (MYB18)/LONG AFTER FAR-RED LIGHT 1 (LAF1)	Positive regulation of the PhyA photoresponse	LOC111998979	24.91	-3.22	-3.33
AT2G31180	Myb domain protein 14 (MYB14)	Unknown	LOC112001589	23.89	-1.34	-1.82
AT1G43700	VIRE2-interacting protein 1 (VIP1)/SULPHATE UTILIZATION EFFICIENCY 3 (SUE3)	Negative regulation of cell differentiation, regulation of touch response (thigmotropism)	LOC112021096	14.96	-2.60	-2.94
Cell cycle regulation						
AT4G34160	CYCLIN D3;1 (CYCD3;1)	Involved in the switch from cell proliferation to the final stages of differentiation ¹	LOC112022222	583.60	-1.87	-0.33

Table 1. Selection of candidate genes for phellogen (cork cambium) regulation classified into Meristem development, Cell differentiation and Cell cycle regulation categories that were upregulated at the beginning of the cork growing season (April, in cluster 2). They are sorted by highest expression (Normalized read counts mean) within each category. The complete list is presented in Supplemental Data Set 2. ¹Involved in vascular cambium development. ²Highly expressed in vascular cambium (Zhang et al., 2019). ³Not differentially expressed and thus not included in cluster 2, but being at the limit of the threshold (absolute log₂Fold Change > 1) to be considered differentially expressed. ⁴Related with chromatin/chromosome organization. ⁵Included in cluster 2 but only DEG in June/July.

does not show the typical secondary cell wall thickenings⁴. The putative involvement of the cell wall candidate genes in the phellem tissue is discussed in the [Discussion](#) section.

Discussion

The phellem development is a continuous process in which phellogen cells produces phellogen derivatives outwardly to form fully-developed phellem cells. Due to the seasonal nature of the phellogen growth, the transcriptome profiling along the cork growing season highlighted key molecular processes for the progression of the phellem cell development and overall the formation of cork tissue. The results provide evidences that cell proliferation and cell differentiation is enriched at the beginning of the season while cell wall biogenesis and secondary metabolism are predominant in later stages of cork growth (June and July), and that abiotic stress signalling is a constant factor (Fig. 4).

Regarding phellogen cell proliferation, many genes involved in cell cycle control, such as cyclins, were upregulated in April (cluster 2) (Supplemental Data Set 2), hence indicating that April cork samples were enriched with phellogen cells actively dividing. D-type cyclin CYCD3;1 (Table 1), which is required for vascular cambium-derived secondary thickening, stands out for regulating the cell cycle of phellogen meristematic cells since its promoter is also activated in phellogen compatible cells in Arabidopsis roots²¹, as well as it is highly expressed in potato tuber phellogen¹⁸.

Cluster 2 was also enriched with genes showing DNA binding capacity, most of them transcription factors (Supplemental Data Set 2 and Table 1). The Meristem development category included genes whose homologs positively regulate the maintenance of the shoot and/or floral meristem and thus they might function similarly in phellogen meristem. From them, WUSCHEL (WUS) maintains stem cell identity²² and SHOOT MERISTEMLESS (STM) prevents stem cells from undergoing differentiation²³. APETALA 2 (AP2) has been described to promote the expression of WUS in floral meristems²⁴ and PHABULOSA (PHB), a member of HD-ZIP III family of transcription factors, also maintains stem cells²⁵. In floral meristems, AINTEGUMENTA (ANT) and AINTEGUMENTA-LIKE 6 (AIL6) activities have also been suggested to promote continued proliferation of stem cells²⁶. Overall, these regulatory genes may converge to functionally maintain phellogen cells in an undifferentiated state. Conversely, Ultrapetala 1 (ULT1), which negatively regulates WUS and acts as a meristem-restricting factor²⁷, may contribute to limit the excessive accumulation of cells in phellogen meristem. Interestingly, STM has been suggested to be a potential epigenetic target of the ULT1, that may counteract, as a trithorax-group protein factor, the Polycomb-group repressive activity of the histone methyltransferase CURLY LEAF (CRL), also found in cork transcriptome. The presence of genes involved in Chromosome/Chromatin organization in cluster 2 (Supplemental Data Set 2) reinforce the role of the epigenetic mechanisms in cork development in April which may synchronize the local climate with the transition from dormant to active phellogen growth.

AGI code	Gene name	Gene function	Involved in secondary cell wall biogenesis	Differentially Expressed (july/april) <i>Q. suber</i> identifier	Normalized read count mean (BaseMean)	Log2FC July/April
AT4G18780	Cellulose synthase 8 (IRX1)	Cellulose biosynthesis	Yes	LOC112002106	2071.36	2.69
AT5G17420	Cellulose synthase 7 (IRX3)	Cellulose biosynthesis	Yes Yes	LOC112026131	2513.41	2.56
AT5G44030	Cellulose synthase 4 (IRX5)	Cellulose biosynthesis		LOC112023957	871.09	2.45
				LOC111984736	423.38	2.90
AT1G19300	Nucleotide-diphospho-sugar transferases superfamily protein (GALACTURONOSYL-TRANSFERASE-LIKE 1, PARVUS)	Glucuronoxylan biosynthesis (hemicellulose). The mutant has thinner secondary cell wall and a reduction of glucuronoxylan	Yes	LOC112011992	659.24	2.21
AT1G27440	Exostosin family protein (IRX10)	Glucuronoxylan biosynthesis (hemicellulose)	Yes	LOC111985053	1245.77	2.05
AT3G18660	Glucuronic acid substitution of xylan 1 (GUX1), Plant glycogenin-like starch initiation protein 1 (PGSIP1)	Glucuronyltransferase adding GlcA residues onto xylan (hemicellulose) and for secondary wall deposition. Also starch biosynthesis	Yes	LOC112008671	278.01	2.86
AT5G54690	Galacturonosyltransferase 12 (IRX8)	The mutant displays reduced xylan (hemicellulose) and cellulose	Yes	LOC112027387	991.26	2.50
AT5G57560	Xyloglucan endotransglucosylase/hydrolase 22 (ATXTH22); Touch 4 (TCH4)	Cell-wall rearrangement; transglycosylation between xyloglucan and xyloglucan-oligosaccharide and analogous reactions (hemicellulose)	-	LOC112028317	1109.07	2.95
AT4G01070	UDP-Glycosyltransferase superfamily protein	Xylosyltransferase	-	LOC112012259 ¹	334.20	2.18
				LOC112014189 ¹	6907.00	1.88
AT2G38080	Laccase/Diphenol oxidase family protein (IRX12)	Lignin biosynthesis	Yes	LOC112011756	377.32	3.30
				LOC112018093	4900.57	2.65
AT2G29130	Laccase 2	Lignin biosynthesis	-	LOC111992001	1530.80	2.76
AT5G60020	Laccase 17	Lignin biosynthesis, G units	Yes	LOC112004816	91.09	2.67
				LOC112025438	2827.40	3.08
				LOC112025440	322.56	2.30
AT5G09360	Laccase 14	Predicted lignin biosynthesis	-	LOC112039873	7035.98	2.29
AT2G21100	Disease resistance-responsive (dirigent-like protein) family protein	Lignan biosynthetic process	-	LOC112026141 ¹	1151.89	3.73
AT2G46770	NAC (No Apical Meristem) domain transcriptional regulator superfamily protein 43 (NST1)	Regulation of secondary wall thickening	Yes	LOC112008346	478.29	2.15
AT5G03170	FASCICLIN-like arabinogalactan-protein 11 (FLA11) (IRX13)	Integrity and elasticity of the plant cell wall matrix	Yes	LOC112031083	1379.34	2.70
AT5G15630	COBRA-like 4, extracellular glycosyl-phosphatidylinositol (GPI)-anchored protein family (IRX6)	The mutant displays reduced cellulose content, cellulose of lower crystallinity, and thinner secondary cell walls	Yes	LOC112004749	637.38	2.29
AT5G16490	ROP-interactive CRIB motif-containing protein 4	Exocytosis and pollen tube growth	-	LOC111985119	187.41	2.68

Table 2. *Q. suber* genes in cluster 1 (upregulated in July and/or June) classified within the enriched GOs related with cell wall biogenesis based on the most homologous Arabidopsis gene. ¹Also differentially expressed in June/April comparison.

Overall, the fact that that meristem-promoting and meristem-restricting factors are enriched in April suggests that overlapping networks may co-exist and adjust phellogen meristem activity and phellogen derivative differentiation in a tight regulatory manner.

Vascular cambium and phellogen both constitute a bifacial stem cell population from which derivatives are specified on opposing sides, in vascular cambium by positional signals²⁸. In agreement, some of the genes required for cambium meristematic activity were also upregulated in this cluster 2: AINTEGUMENTA (ANT), WUSCHEL-RELATED HOMEBOX 4 (WOX4) and SCARECROW-LIKE 28 (SCL28)^{19,21,29,30} (Table 1). Also,

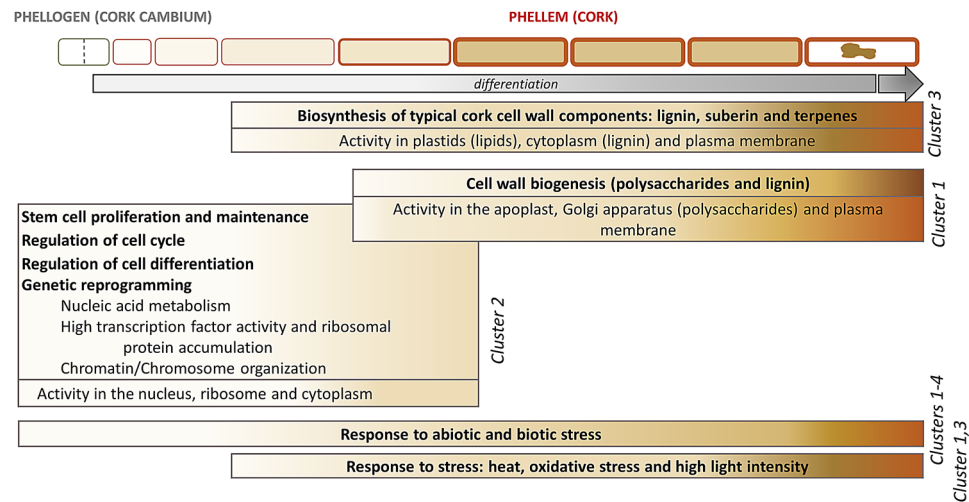


Figure 4. Summary of biological and molecular processes occurring during phellem cell ontogeny, from phellogen stem cell periclinal division to fully differentiated phellem cell, based on upregulated genes and processes in cork samples harvested in April, June and July.

the master regulator of vascular cambium activity *KNAT1/BREVIPEDICELLUS (BP)*¹⁹ shows high expression with a maximum in April (June/April $\log_2FC = -0.98$). In agreement with our data and meaningfully, very recently *WOX4* and *BP* has been demonstrated to control phellogen activity in *Arabidopsis* root standing as periderm-positive regulators¹⁵, suggesting that phellogen in herbaceous and woody plants may be regulated similarly. Other identified genes have also been observed to be highly expressed in vascular cambium: *STM*, *WRINKLED 3 (WRI3)*, *ANAC042* and *LOB DOMAIN-CONTAINING PROTEIN 4 (LBD4)*, the latter also showing expression in phellogen and involved in vascular patterning¹⁹.

Many regulators to trigger phellem cell differentiation were also identified upregulated in April (Table 1). We identified *LOST MERISTEMS 3 (LM3)*, also known as *HAIRY MERISTEM 3/HAM3* which in shoot apical meristem promotes the differentiation at the periphery³¹, suggesting that in cork it may act similarly by promoting the differentiation of phellem cells outwardly. Also, nine *MYB* transcription factors were also upregulated in April (Table 1), indicating the importance of this gene family in contributing to phellem cell differentiation. From them, *AtMYB106* was described to regulate cuticular lipid-deposits (wax and cutin) in epidermal cells³² and three *MYBs* were related with lignification and/or suberization: *AtMYB36* promotes the lignification of the Casparian strip and suberization of endodermis^{33,34}, and *AtMYB3* and *AtMYB4* repress phenylpropanoid biosynthesis^{35–37} and when overexpressed lead to dysfunctional Casparian strip lignification and lower suberin deposition in endodermis and phellem³⁸. Overall, these *MYB* transcription factors are candidates to early regulate in phellogen derivatives to phellem cells the phenylpropanoid and fatty acyl biosynthetic processes that lead to the final accumulation of soluble or insoluble phenylpropanoids (such as condensed tannins or lignin) or fatty-acyl derived compounds (such as waxes or suberin). Also, cluster 2 included a *GSK3* protein named *BRASSINOSTEROID INSENSITIVE 2 (BIN2; AT4G18710)* which regulates xylem cell differentiation³⁹. Moreover, three genes upregulated in April were homologs to those which define cell fates in *Arabidopsis* root epidermis: *WEREWOLF/MYB66*, *GLABROUS 3* and *CAPRICE (CPC)*. Their contribution to phellem formation remains elusive.

During phellem cell development, cells are devoted to synthesize the components that will be deposited in an accurate organization within their cell walls. The major cell wall components of cork (in mass) are: suberin (43%, a glycerol and fatty-acyl-derived polyester including ferulic acid), lignin (22%, a polymer of phenylpropane aromatic units), celluloses and hemicellulose (20%, the most important hemicelluloses in cork are xylans) and extractives (16%, including triterpene and phenolic compounds)^{40–42}. During the last fifteen years, an intensive work in characterizing biosynthetic genes of suberin, and also extractives, has been carried out in potato tuber phellem but also in other suberizing tissues such as *Arabidopsis* root endodermis and seed coat, and recently also root phellem (see Supplemental Table 4A for citations). From them, the most highly expressed genes showed a peak of expression in June (Fig. 2), an expression profile used to identify new molecular players that participate in the deposition of suberin and associated waxes within the cell wall (Supplemental Figure 2). Recently, three cork oxidosqualene cyclases involved in most cork triterpenoid formation (i.e. friedelin) were biochemically characterized, and the friedelin synthase also showed a maximum expression in June⁴³. This maximum expression in June coincides with the highest cork growth rate^{7,17} and overall aligns with the upregulation of lipid, terpenoid and phenylpropanoid metabolisms in June and July (cluster 3) (Supplemental Table 7).

In contrast to suberin, the molecular processes and functional relevance of lignin and cell wall polysaccharides in phellem, as well as in other suberized cells, is still unclear. Cell wall genes related with lignin and polysaccharide metabolism were enriched in Cluster 1 showing higher expression in June and July but only differentially expressed in July/April comparison (Table 2). Most of these genes were involved in secondary cell wall (SCW) production and included 8 homologous genes (from the 15 identified to date⁴⁴) that when mutated displayed an

irregular xylem (irx) phenotype, characterized for the disruption of the xylem cells, especially vessel cells⁴⁵. The IRX genes correspond to the three cellulose synthase (CESA4/IRX5, CESA7/IRX3 and CESA8/IRX1) required for normal cellulose synthesis in the SCW, three xylan-related hemicellulose biosynthetic genes (*IRX8*, *IRX10* and *IRX15*) specific from SCW production, a lignin (laccase) biosynthetic gene IRX12 and the IRX6/COBRA-LIKE4, a lipid anchored protein with a putative cellulose-binding domain^{46–53}. Other genes from Cluster 1 acting on SCW were also involved in xylan biosynthesis (PARVUS and GUX1/PGSIP1) and regulation of wall thickening (NST1/ANAC043)^{54,55}. Several lignin-polymerization enzymes specially laccases were also included in this Cluster 1 (Table 2). The identification of genes specific and related with SCW formation was surprising since the phellem or cork cells do not develop a SCW, instead they develop a primary cell wall which inwardly deposits the suberin lamellae and a tertiary cell wall. Possibly, the upregulation of typical SCW genes in July could reproduce the cork-ring anatomy (Fig. 1) since these genes may contribute to the transition from earlycork to latecork cells, which have thicker walls⁴. Hence, we point that typical SCW cellulose, xylan and lignin may participate in the cell wall thickening and biogenesis of latecork formation, providing higher strength, rigidity and hydrophobicity as reported for SCW^{56,57}. It is also worth to mention that xylan has been described to be an anchor for lignin deposition, and, in grasses, lignin is often linked to xylan chains via ferulic acids^{58,59}, that remarkably is part of the suberin polyester³.

Response to abiotic stress is found in all clusters which is in agreement with the ability of ABA to induce suberin accumulation^{60,61}. Genes with higher expression in both June and July (cluster 1 and cluster 3) presented specific enrichment in responses to specific stresses, such as response to light and heat and oxidative stress. The response to light and heat aligns with higher temperature conditions typical of Mediterranean summers, which are reported for the year and area where the cork oaks used for these analyses were growing (Supplemental Figure 3). Oxidative stress in cork samples, and other ligno/suberizing tissues, has been documented⁶² and associated with oxidative coupling of polyphenolic components through a peroxidase/H₂O₂ free radical formation⁶³. This is in agreement with an enrichment of lignin/suberin accumulation in June and July (Fig. 4).

Conclusion

The main goal of the current study was to determine the molecular mechanisms and genes that are involved in phellogen to phellem formation. Taking advantage that phellem formation is a continuous developmental process and that cork follows seasonal growth, the comparison of the transcriptome of cork samples harvested at the beginning, maximum and advanced stage of cork formation have highlighted robustly the molecular mechanisms involved in phellogen stem cell proliferation and maintenance, regulation of phellem cell differentiation and phellem cell wall biogenesis. The fact that some of the genes identified have been already characterized in the root periderm of the model plant *Arabidopsis* reinforces our data, suggests putative shared molecular mechanisms in periderm of both herbaceous and woody species and provide valuable information to assist in the understanding of the general mechanisms controlling phellem protective tissue function.

Methods

Plant material and tissue harvesting. The virgin cork (phellem) tissue used was the same used and described by Soler et al. (2008)⁷. Cork samples were obtained from trees of 15–20 years old, located in a typical cork oak (*Quercus suber*) forest in Romanyà de la Selva (Girona, Spain) (41°51'42.5" N, 3°27'9" E; UTM X = 502,951; Y = 4,634,516.2) at a density of 1178 trees/ha. The tree mean diameter was 17.3 ± 5.4 cm and the mean height 8.2 m. Cork samples were collected during the growth period of 2005 (April 26, June 17 and July 19). Climatic data for the 2005 growing season were obtained from the nearby weather station of Castell d'Aro (UTM X = 502,731; Y = 4,628,777), belonging to the Servei Meteorològic de Catalunya (SMC) and are depicted in Supplemental Figure 3. The cork was the first cork stripped from tree trunk (virgin cork), which transcriptome is very similar to that of amadia cork⁶⁴. The tissue harvested was the newly generated and active cork tissue from the belly side of the plank, adjacent to the phellogen tissue. The samples were individually collected from eleven independent trees: four in April, three in June and four in July, immediately frozen in liquid nitrogen and stored at –80 °C until RNA was extracted.

The virgin cork collection complies with relevant institutional, national, and international guidelines and legislation. For the collection of cork, we obtained the appropriate permission of the Mr. J. Arnedo, the owner of the land where the cork oak forest was located.

RNA extraction and mRNA purification

Total RNA was extracted from phellem as reported by Soler et al. (2008)⁷ using a procedure modified by Chang et al. (1995)⁶⁵. Briefly, 5 g of grinded frozen cork tissue was mixed with extraction buffer (2% CTAB, 2% PVP, 10 mM Tris–HCl pH 8.0, 25 mM EDTA, 2 M NaCl and 2.67% 2-mercaptoethanol) previously heated at 65 °C and the mixture was incubated at 65 °C for 10 min. Then, two consecutive extractions using 1 volume of chloroform:isoamyl alcohol 24:1 (v/v) were performed, and the RNA in the aqueous phase was precipitated overnight adding 0.25 volumes of 10 M LiCl at 4 °C. The precipitate was collected by centrifuging at 12,000 g for 20 min, and the pellet was resuspended in buffer containing 1 M NaCl, 0.5% SDS, 10 mM Tris–HCl pH 8.0 and 1 mM EDTA and incubated at 60 °C for 5 min. Then, one acid phenol:chloroform:isoamyl alcohol (125:24:1 (v/v)) and two 24:1 (v/v) chloroform:isoamyl alcohol extractions were performed. The RNA in the upper phase was precipitated with 2 volumes of 100% ethanol and the precipitate collected by centrifugation. The pellet was washed twice with 70% (v/v) ethanol and resuspended in 50 µl of RNase-free water. RNA extracts were cleaned with the RNeasy MinElute Cleanup (Qiagen) and genomic DNA was on-column digested with DNase I (Qiagen). The RNA samples sequenced had a RIN (RNA integrity number) value of 8 or higher.

RNA-seq and bioinformatics analyses

Cork oak cDNA libraries were constructed using the Truseq v4 kit and sequencing was performed by the IlluminaHiSeq2500 platform (paired end reads of 125 bp) at the Ultrasequencing Unit of Center for Genomic Regulation (CRG, Barcelona). In total, 11 samples were sequenced; 3 biological replicates for June and 4 for April and July each. The bioinformatics treatment of raw sequences was performed at the CRG Bioinformatic Unit. The quality of reads obtained was evaluated using the FastQC software v0.11.5 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>) and the adapters were eliminated using the Skewer software version 0.2.2⁶⁶. To remove the sequences of polluting ribosomal RNA, the riboPicker version 0.4.3⁶⁷ was used. Then, transcriptome alignment and mapping was made using the STAR_2.5.3a program⁶⁸ against the reference genome GCF_002906115.1 (CorkOak1.0). SAM data was changed to BAM data with samtools 1.4.1⁶⁹. The alignment quality was evaluated using QualiMap v2.2.1 program⁷⁰ and to compare genomic features we used Bedtools v2.26.0⁷¹. Due to the natural genetic variation of samples, the mapping on cork oak genome was critical to annotate all the corresponding polymorphic alleles of a particular gene of different trees ($n = 11$) as a unique gene/locus. This allowed to create a matrix of expression at gene/locus level and prevented that specific polymorphic alleles differentially represented among samples were identified as differentially expressed transcripts. For Arabidopsis annotation we used the Blastp and the TAIR10 library from Ensembl, with the options num_alignments 1 and evaluate 1E-5.

Raw counts were pre-filtered for low count genes (read count > 1 in every sample) before running the DESeq2 functions⁷² for differential expression analysis between June/April, July/April and June/July comparisons. DEGs were those with an adjusted pvalue ≤ 0.01 and absolute \log_2 Fold Change > 1. The normalized read count values for DEGs in at least one comparison were hierarchical clustered using the Pearson correlation through the MeV program and default parameters⁷³. Normalized read counts were log transformed (rlog) using DESeq2 to be used for exploratory analysis such as gene expression clustering. The gene clustering was performed using K means clustering in R by imposing 4 clusters, maximum number of iterations = 100,000 and using 10,000 random initial configurations. Expression profiles for each cluster were generated using R (package ggplot2). The grey rows are the expression of every gene, as rlog, while the coloured line is their mean.

Gene ontology (GO) was carried out with AgriGO v.2 (FDR ≤ 0.05)⁷⁴ using the Arabidopsis best homologous gene (TAIR10). The functional GO categories were manually compared, and those with similar GO term description and containing mostly the same set of genes were manually and carefully grouped for a better integration of the data.

The heatmaps of suberin-related genes were constructed with EXPANDERS⁷⁵. To view the expression patterns of the different genes in the same scale the data was standardized as follows: for each gene the mean of the normalized read counts of month's replicates was \log_2 transformed and the expression pattern (April, June, July) was normalized to have a mean of 0 and a variance of 1. To show an absolute expression value, the sum of normalized read counts for the 11 samples was calculated and expressed as total expression.

For the comparison of the cork transcriptome from cork oak with that of the phellem and phellogen/phellogen/derm from birch, we used the DEGs in F1/F8 and F2/F8 comparisons from Alonso-Serra et al. (2019)¹⁴, in which F1 corresponded to phellem, F2 to phellogen and phellogen/derm and F8 to xylem. For the correspondence, the corresponding Arabidopsis annotations (hits with higher similarity in the BlastP) were used and the duplications were removed from the comparative list.

Validation of RNA-seq data

RNA-seq results were validated for a subset of 29 representative genes (Supplemental Table 2) by RT-qPCR data generated in previous works using the same samples (22 genes)^{7,17} and 7 new genes, for which the primer sequences are presented in Supplemental Table 9. For RNA-seq and RT-qPCR, the mean of normalized read counts and the mean of relative transcript abundances (RTA) were used, respectively, to calculate the corresponding \log_2 Ratio for each comparison (June/April, July/April and July/June). The correlation between RNA-seq and RT-qPCR was assessed by a Pearson correlation analysis.

Accession numbers

For a comprehensive list of accession numbers of cork oak and their Arabidopsis putative ortholog genes mentioned in the article (mainly discussion), please see Fig. 2, Table 1 and Table 2, as well as Supplemental Data Set 1 and Supplemental Data Set 2.

Data accessibility

The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus⁷⁶ and are accessible through GEO Series accession number GSE155544 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE155544>).

Received: 4 January 2021; Accepted: 19 May 2021

Published online: 08 June 2021

References

1. Tonn, N. & Greb, T. Radial plant growth. *Curr. Biol.* **27**, R878–R882 (2017).
2. Evert, R. F. Periderm in *Esaus's Plant Anatomy*. 427–445 (John Wiley & Sons, 2006).
3. Graça, J. & Santos, S. Suberin: a biopolyester of plants' skin. *Macromol. Biosci.* **7**, 128–135 (2007).
4. Pereira, H. *Cork: Biology, Production and Uses* (Elsevier, 2007).
5. Soler, M. et al. A genomic approach to suberin biosynthesis and cork differentiation. *Plant Physiol.* **144**, 419–431 (2007).
6. Caritat, A., Molinas, M. & Gutierrez, E. Annual cork-ring width variability of *Quercus suber* L. in relation to temperature and precipitation (Extremadura, southwestern Spain). *For. Ecol. Manag.* **86**, 113–120 (1996).

7. Soler, M. *et al.* Seasonal variation in transcript abundance in cork tissue analyzed by real time RT-PCR. *Tree Physiol.* **743–751** (2008).
8. Silva, S. P. *et al.* Cork: properties, capabilities and applications. *Int. Mater. Rev.* **53**, 345–365 (2005).
9. Costa, A., Pereira, H. & Oliveira, A. Influence of climate on the seasonality of radial growth of cork oak during a cork production cycle. *Ann. For. Sci.* **59**, 429–437 (2002).
10. Pereira, H., Graça, J. & Baptista, C. The Effect of Growth Rate on the Structure and Compressive Properties of Cork. *IAWA J.* **13**, 389–396 (1992).
11. Caritat, A., Gutiérrez, E. & Molinas, M. Influence of weather on cork-ring width. *Tree Physiol.* **20**, 893–900 (2000).
12. Rains, M. K., Gardiyehewa de Silva, N. D. & Molina, I. Reconstructing the suberin pathway in poplar by chemical and transcriptomic analysis of bark tissues. *Tree Physiol.* **38**, 340–361 (2018).
13. Ranathunge, K., Schreiber, L. & Franke, R. Suberin research in the genomics era—New interest for an old polymer. *Plant Sci.* **180**, 339–413 (2011).
14. Alonso-Serra, J. *et al.* Tissue-specific study across the stem reveals the chemistry and transcriptome dynamics of birch bark. *New Phytol.* **222**, 1816–1831 (2019).
15. Xiao, W. *et al.* Pluripotent Pericycle Cells Trigger Different Growth Outputs by Integrating Developmental Cues into Distinct Regulatory Modules. *Curr. Biol.* **30**, 4384–4398.e5 (2020).
16. Ramos, A. M. *et al.* The draft genome sequence of cork oak. *Sci. Data* **5**, 180069 (2018).
17. Boher, P. *et al.* A comparative transcriptomic approach to understanding the formation of cork. *Plant Mol. Biol.* **96**, 103–118 (2018).
18. Vulavala, V. K. R., Fogelman, E., Faigenboim, A., Shoseyov, O. & Ginzberg, I. The transcriptome of potato tuber phellogen reveals cellular functions of cork cambium and genes involved in periderm formation and maturation. *Sci. Rep.* **9**, 10216 (2019).
19. Zhang, J. *et al.* Transcriptional regulatory framework for vascular cambium development in Arabidopsis roots. *Nat. Plants* **5**, 1033–1042 (2019).
20. Pereira, H. Chemical composition and variability of cork from *Quercus suber* L. *Wood Sci. Technol.* **22**, 211–218 (1988).
21. Randall, R. S. *et al.* AINTEGUMENTA and the D-type cyclin CYCD3;1 regulate root secondary growth and respond to cytokinins. *Biol. Open* **4**, 1229–1236 (2015).
22. Laux, T., Mayer, K. F. X., Berger, J. & Jürgens, G. The WUSCHEL gene is required for shoot and floral meristem integrity in Arabidopsis. *Development* **122**, 87–96 (1996).
23. Clark, S. E., Jacobsen, S. E., Levin, J. Z. & Meyerowitz, E. M. The CLAVATA and SHOOT MERISTEMLESS loci competitively regulate meristem activity in Arabidopsis. *Development* **122**, 1567–1575 (1996).
24. Huang, Z. *et al.* APETALA2 antagonizes the transcriptional activity of AGAMOUS in regulating floral stem cells in Arabidopsis thaliana. *New Phytol.* **215**, 1197–1209 (2017).
25. Pierre-Jerome, E., Drapek, C. & Benfey, P. N. Regulation of Division and Differentiation of Plant Stem Cells. *Annu. Rev. Cell Dev. Biol.* **34**, 289–310 (2018).
26. Krizek, B. A. AINTEGUMENTA and AINTEGUMENTA-LIKE6 act redundantly to regulate arabidopsis floral growth and patterning. *Plant Physiol.* **150**, 1916–1929 (2009).
27. Carles, C. C., Lertpiriyapong, K., Reville, K. & Fletcher, J. C. The ULTRAPETALA1 gene functions early in Arabidopsis development to restrict shoot apical meristem activity and acts through WUSCHEL to regulate floral meristem determinacy. *Genetics* **167**, 1893–1903 (2004).
28. Smit, M. E. *et al.* A PXY-mediated transcriptional network integrates signaling mechanisms to control vascular development in Arabidopsis. *Plant Cell* **32**, 319–335 (2020).
29. Ji, J. *et al.* WOX4 promotes procambial development. *Plant Physiol.* **152**, 1346–1356 (2010).
30. Suer, S., Agustí, J., Sanchez, P., Schwarz, M. & Greb, T. WOX4 imparts auxin responsiveness to cambium cells in Arabidopsis. *Plant Cell* **23**, 3247–3259 (2011).
31. Schulze, S., Schäfer, B. N., Parizotto, E. A., Voinnet, O. & Theres, K. LOST MERISTEMS genes regulate cell differentiation of central zone descendants in Arabidopsis shoot meristems. *Plant J.* **64**, 668–678 (2010).
32. Oshima, Y. *et al.* MIXTA-like transcription factors and WAX INDUCER1/SHINE1 coordinately regulate cuticle development in Arabidopsis and *Torenia fournieri*. *Plant Cell* **25**, 1609–1624 (2013).
33. Kamiya, T. *et al.* The MYB36 transcription factor orchestrates Casparian strip formation. *Proc. Natl. Acad. Sci. USA* **112**, 10533–10538 (2015).
34. Wang, C. *et al.* Developmental programs interact with abscisic acid to coordinate root suberization in Arabidopsis. *Plant J.* **104**, 241–251 (2020).
35. Jin, H. Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. *EMBO J.* **19**, 6150–6161 (2000).
36. Dubos, C. *et al.* MYB transcription factors in Arabidopsis. *Trends Plant Sci.* **15**, 573–581 (2010).
37. Fornalé, S. *et al.* AtMYB7, a New Player in the Regulation of UV-Sunscreens in Arabidopsis thaliana. *Plant Cell Physiol.* **55**, 507–516 (2014).
38. Andersen, T. G. *et al.* Tissue-autonomous phenylpropanoid production is essential for establishment of root barriers. *Curr. Biol.* **31**, 965–977 (2021).
39. Kondo, Y. *et al.* Plant GSK3 proteins regulate xylem cell differentiation downstream of TDIF-TDR signalling. *Nat. Commun.* **5**, 1–11 (2014).
40. Conde, E., Cadahía, E., García-Vallejo, M. C. & González-Adrados, J. R. Chemical Characterization of Reproduction Cork from Spanish *Quercus Suber*. *J. Wood Chem. Technol.* **18**, 447–469 (1998).
41. Conde, E., Cadahía, E., García-Vallejo, M. C., Fernández de Simón, B. & González Adrados, J. R. Low molecular weight polyphenols in cork of *Quercus suber*. *J. Agric. Food Chem.* **45**, 2695–2700 (1997).
42. Pereira, H. Variability of the chemical composition of cork. *BioResources* **8**, 2246–2256 (2013).
43. Busta, L. *et al.* Oxidosqualene cyclases involved in the biosynthesis of triterpenoids in *Quercus suber* cork. *Sci. Rep.* **10**, 1–12 (2020).
44. Hao, Z. & Mohnen, D. A review of xylan and lignin biosynthesis: Foundation for studying Arabidopsis irregular xylem mutants with pleiotropic phenotypes. *Crit. Rev. Biochem. Mol. Biol.* **49**, 212–241 (2014).
45. Turner, S. R. & Somerville, C. R. Collapsed xylem phenotype of arabidopsis identifies mutants deficient in cellulose deposition in the secondary cell wall. *Plant Cell* **9**, 689–701 (1997).
46. Liu, L. *et al.* Brittle Culm1, a COBRA-Like Protein, Functions in Cellulose Assembly through Binding Cellulose Microfibrils. *PLoS Genet.* **9**, e1003704 (2013).
47. Persson, S., Wei, H., Milne, J., Page, G. P. & Somerville, C. R. Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc. Natl. Acad. Sci. USA* **102**, 8633–8638 (2005).
48. Brown, D. M., Zeef, L. A. H., Ellis, J., Goodacre, R. & Turner, S. R. Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell* **17**, 2281–2295 (2005).
49. Brown, D. M., Zhang, Z., Stephens, E., Dupree, P. & Turner, S. R. Characterization of IRX10 and IRX10-like reveals an essential role in glucuronoxylan biosynthesis in Arabidopsis. *Plant J.* **57**, 732–746 (2009).
50. Brown, D. *et al.* Arabidopsis genes IRREGULAR XYLEM (IRX15) and IRX15L encode DUF579-containing proteins that are essential for normal xylan deposition in the secondary cell wall. *Plant J.* **66**, 401–413 (2011).

51. Peña, M. J. *et al.* Arabidopsis irregular xylem8 and irregular xylem9: Implications for the complexity of glucuronoxylan biosynthesis. *Plant Cell* **19**, 549–563 (2007).
52. Jensen, J. K. *et al.* The DUF579 domain containing proteins IRX15 and IRX15-L affect xylan synthesis in Arabidopsis. *Plant J.* **66**, 387–400 (2011).
53. Wu, A. M. *et al.* The Arabidopsis IRX10 and IRX10-LIKE glycosyltransferases are critical for glucuronoxylan biosynthesis during secondary cell wall formation. *Plant J.* **57**, 718–731 (2009).
54. Lee, C. *et al.* The PARVUS Gene is Expressed in Cells Undergoing Secondary Wall Thickening and is Essential for Glucuronoxylan Biosynthesis. *Plant Cell Physiol.* **48**, 1659–1672 (2007).
55. Mortimer, J. C. *et al.* Absence of branches from xylan in Arabidopsis gux mutants reveals potential for simplification of lignocellulosic biomass. *Proc. Natl. Acad. Sci. USA* **107**, 17409–17414 (2010).
56. Scheller, H. V. & Ulvskov, P. Hemicelluloses. *Annu. Rev. Plant Biol.* **61**, 263–289 (2010).
57. Vanholme, R., Demedts, B., Morreel, K., Ralph, J. & Boerjan, W. Lignin biosynthesis and structure. *Plant Physiol.* **153**, 895–905 (2010).
58. Grabber, J. H., Ralph, J. & Hatfield, R. D. Model studies of ferulate—Coniferyl alcohol cross-product formation in primary maize walls: Implications for lignification in grasses. *J. Agric. Food Chem.* **50**, 6008–6016 (2002).
59. Gregory, A. C. E., Smith, C., Kerry, M. E., Wheatley, E. R. & Bolwell, G. P. Comparative subcellular immunolocalization of polypeptides associated with xylan and callose synthases in French bean (*Phaseolus vulgaris*) during secondary wall formation. *Phytochemistry* **59**, 249–259 (2002).
60. Barberon, M. *et al.* Adaptation of root function by nutrient-induced plasticity of endodermal differentiation. *Cell* **164**, 1–13 (2016).
61. Lulai, E. C., Suttle, J. C. & Pederson, S. M. Regulatory involvement of abscisic acid in potato tuber wound-healing. *J. Exp. Bot.* **6**, 1175–1186 (2008).
62. Pla, M. *et al.* Stress proteins co-expressed in suberized and lignified cells and in apical meristems. *Plant Sci.* **139**, 49–57 (1998).
63. Razem, F. A. & Bernards, M. A. Reactive oxygen species production in association with suberization: evidence for an NADPH-dependent oxidase. *J. Exp. Bot.* **54**, 935–941 (2003).
64. Lopes, S. T. *et al.* Phellem versus xylem: Genome-wide transcriptomic analysis reveals novel regulators of cork formation in cork oak. *Tree Physiol.* **40**, 129–141 (2020).
65. Chang, S., Puryear, J. & Cairney, J. A simple and efficient method for isolating RNA from pine trees. *Plant Mol. Biol. Report.* **11**, 113–116 (1993).
66. Jiang, H., Lei, R., Ding, S.-W. & Zhu, S. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* **15**, 182 (2014).
67. Schmieder, R., Lim, Y. W. & Edwards, R. Identification and removal of ribosomal RNA sequences from metatranscriptomes. *Bioinformatics* **28**, 433–435 (2012).
68. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
69. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
70. Garcia-Alcalde, F. *et al.* Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* **28**, 2678–2679 (2012).
71. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
72. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
73. Howe, E. *et al.* MeV: MultiExperiment Viewer in *Biomedical Informatics for Cancer Research* 267–277 (Springer, 2010).
74. Tian, T. *et al.* agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* **45**, W122–W129 (2017).
75. Ulitsky, I. *et al.* Expander: from expression microarrays to networks and functions. *Nat. Protoc.* **5**, 303–322 (2010).
76. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207–210 (2002).

Acknowledgements

We thank Prof. M. Molinas for insightful discussion about the cork biology and the cellular processes leading to cork formation. The authors are grateful to Mr. J. Arnedo and Mr. J. Casellas for providing the cork material, to Dr. G Mir and Mr. E. Juliol for their help during cork harvesting and to Dr. L. Cozzuto from Bioinformatics Unit in Centre for Genomic Regulation (CRG, Barcelona) for the bioinformatics support. This work was supported by the Ministerio de Economía y Competitividad [AGL2012-36725; AGL2015-67495-C2-1-R] and Ministerio de Ciencia e Innovación [PID2019-110330GB-C21] and The European Regional Development Fund (ERDF). S. Fernández-Piñán was supported by a University of Girona PhD fellowship.

Author contributions

M.F. and O.S. conceived the research. M.S., S.F.-P. and P.B. performed the experimental work and S.F.-P. analysed NGS data. S.F.-P. and O.S. interpreted the data, which was discussed with M.F. and M.S.. O.S. and S.F.-P. prepared the figures and wrote the manuscript, and all authors revised the final manuscript form.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-90938-5>.

Correspondence and requests for materials should be addressed to O.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021