ORIGINAL PAPER

Expression of DNA methyltransferases is involved in *Quercus suber* cork quality

**Miguel Ramos · Margarida Rocheta · Luísa Carvalho ·
Vera Inácio · José Graça · Leonor Morais-Cecilio**

Abstract Cork oak (*Quercus suber*) is an important Portuguese species, mainly due to the economic value of the cork it produces. Cork results from phellogen, a meristematic tissue, which can locally produce lenticels or have discontinuities, originating "defects": pores and nail inclusions that are detrimental to cork industrial use. Epigenetic processes control plant development and its deregulation can lead to altered phenotypes; therefore, the study of epigenetic players in the phellogen is important to understand the emergence of cork's defects. DNA methyltransferases (DNMTs) and one protein associated to MET1 (DMAP1) were characterized in *Q. suber*, and their gene expression was analyzed in phellogen and contiguous differentiating cell layers of trees producing high and low quality cork, after the evaluation of their defects by physical and image analysis methods. All classes of DNMTs (MET, DRM, and CMT) with the respective canonical motifs were identified in *Q. suber*. The expression analyses of these genes showed that *QsDRM2* was the most active methyltransferases in the cells analyzed, and that all the genes were differentially expressed in trees with distinct cork quality, with a tendency for higher expression levels in low quality producers. Interestingly, the global methylation level was higher in cells with low expression of DNA methyltransferases. A positive and significant correlation was obtained between *QsDMAP1* gene expression and the percentage of cork defects. This work provides the first evidence that cork quality in *Q. suber* is likely influenced by epigenetic mechanisms.

Keywords Cork oak · Cork quality · Phellogen activity · MET1 · MET2 · DRM · CMT3 · DMAP1

M. Ramos · M. Rocheta · L. Carvalho · V. Inácio ·
L. Morais-Cecilio (✉)
Instituto Superior de Agronomia, CBAA, Universidade Técnica de Lisboa, Tapada da Ajuda, 1349-017 Lisbon, Portugal
e-mail: lmorais@isa.utl.pt

J. Graça
Instituto Superior de Agronomia, CEF, Universidade Técnica de Lisboa, Tapada da Ajuda, 1349-017 Lisbon, Portugal

Introduction

Cork oak (*Quercus suber*) is an important forest species in the Mediterranean basin (Pereira 2007) due to its high ecological and socioeconomic impacts, since cork production and processing contributes to the maintenance of a large amount of employment and income for the economy of cork-producing countries (Leal et al. 2006). Cork oak is a tree well adapted to the Mediterranean ecosystems since it produces a thick outer bark, the cork that protects it against adverse environments such as drought and, particularly, fire (Pereira 2007; Ricardo et al. 2011; Soler et al. 2008). The production of cork is a sequential lifetime process: when the tree trunk or branches have reached enough perimeter (70 cm), cork can be harvested for the first time; afterwards, cork extraction can be repeated every 9 years, without damaging the tree, until a final age of circa 200 years old (Pereira 2007; Natividade 1950).

Cork tissue results from the activity of the phellogen, a meristematic tissue with seasonal activity, and normally with no discontinuities in *Q. suber* (Pereira 2007; Soler et al. 2008; Şen et al. 2011). Cork is formed by suberized cells, which have very thin cell walls, less than 1 μm thick, and are rich in suberin (Pereira 2007; Graça and Pereira 2004). Cork cells have the form of hexagonal prisms, growing in the radial direction, and are packed as a honeycomb structure (Pereira 2007). Several "defects" can affect cork in the perspective of its usage as an industrial material. The main defects are

"pores," which result from the lenticel channels that cross cork in the radial direction. Cork pores exist in all corks, varying in number, size, and form and are detrimental for industrial quality. The second most important defects are occasional inclusions of phloem tissue within the cork tissue, named "nails," which add to the negative effects of cork pores (Pereira et al. 1996). Cork's quality is the major contributor to its economic value, since it determines its industrial uses and the performance of the final cork products (Pereira 2007; Pereira et al. 1996; Fortes et al. 2004). Cork is used worldwide in a huge variety of products, the more relevant and symbolic being cork stoppers. Therefore, a good quality cork plank must be thick enough to allow the cutting of cork stoppers and be made of mostly homogeneous suberized tissue with as few as possible defects (Fortes et al. 2004).

The ratio of the area of pores to the total area of a defined plane is named cork "porosity." Porosity has been the characteristic widely chosen to evaluate cork quality (Pereira et al. 1996). Pores are approximately cylindrical in section and run radially across cork planks, from its innermost surface to the outside (Pereira 2007; Graça and Pereira 2004; Fortes et al. 2004). Pores are originated by the desegregation of special cells, the "filling tissue," creating the lenticel channels (Pereira 2007; Graça and Pereira 2004; Fortes et al. 2004). Because pores in the first cork are formed in the continuity of the epidermis stomata, it is thought that they may contribute to gas exchange, being present in all high and low quality corks (Pereira 2007). Large and abundant pores are unwanted in industry as they interfere in cork's insulation and mechanical properties. On the other hand, nail inclusions are even more relevant defects affecting industrial use, especially when the inclusion is oriented in the tangential plane (Gonzalez-Adrados and Pereira 1996). Inclusions are formed when the phellogen dies in small localized areas, and in the affected area, a new phellogen is formed farther inside, isolating phloemic tissue inside the suberized tissue (Pereira 2007). These nail inclusions are adverse to industry, as they are formed by lignified cells, namely phloem fibers and sclereids, with thick walls, making the cork much more dense and rigid (Pereira 2007).

Cork development involves several steps such as cell proliferation and differentiation of the phellogen derivatives, and cell expansion and extensive deposition of suberin and waxes, besides an irreversible program of senescence ending in cell death (Pereira 2007). The importance of the epigenetic regulation of plant developmental steps is widely recognized (Finnegan and Kovac 2000; Huang et al. 2010; Wada 2005). Epigenetic phenomena are associated with several chromatin modifications being DNA methylation, a heritable DNA modification, the best studied (Martienssen and Colot 2001). The DNA methylation mark is imposed through the activity of DNA methyltransferases (DNMTs) which add a methyl group to the fifth carbon of a cytosine from the cofactor S-adenosyl-L-methionine (SAM) (Cao et al. 2000;

Finnegan and Kovac 2000; Genger et al. 1999). In plants, there are three DNMT classes which play a role in different DNA sequence contexts and in different developmental situations: METHYLTRANSFERASE I (MET1), DOMAIN REARRANGED METHYLTRANSFERASE (DRM1/2), and CHROMOMETHYLTRANSFERASE 3 (CMT3). The MET class is usually the most highly expressed in all situations, as it is responsible for the maintenance of DNA methylation on symmetric CpG contexts (Huang et al. 2010; Genger et al. 1999). These enzymes have eight domains characteristics of the DNMTs (I, II, IV, VI, VII, VIII, IX, and X), the most relevant ones being domain I, for binding to SAM; domain IV, the active center of the enzyme; and domain IX, for binding DNA (Finnegan and Dennis 1993). *MET2* is present in all eukaryotic organisms and is the most highly conserved of the *MET* genes (Goll and Bestor 2005), encoding a protein with all ten domains (I to X) (Dong et al. 2001). Despite the ongoing controversy about its capacity to methylate DNA, Song et al. (2010) were able to locate this protein in the nucleus. The DRM class, which name is due to the special arrangement of its domains (VI, IX, X, I, II, III, IV, and V) (Cao et al. 2000), is responsible for the de novo methylation in any context and for the maintenance of non-CpG methylation (Huang et al. 2010; Cao et al. 2000). DRMs also have two or three ubiquitin binding domains, UBA1–3 (Cao et al. 2000). Finally, CMT3 is a class that is exclusive of plants and its members can maintain methylation in CpHpG contexts (Huang et al. 2010; Cao and Jacobsen 2002). As the other classes, CMT has eight characteristic domains of DNMTs and also a chromodomain, between the II and IV domains (Genger et al. 1999).

Coworking with DNA methyltransferases are several other proteins such as Methyltransferase 1 Associated Protein 1 (DMAP1), which was found to be co-localized with MET1 on replication *foci* (Rountree et al. 2000; Shin et al. 2010). In addition to its function helping MET1 to methylate DNA, DMAP1 has a large range of other functions such as its association with histone deacetylases (Rountree et al. 2000), cell cycle control (Shin et al. 2010), and nucleic acid reparation (Negishi et al. 2009).

Knowledge concerning the epigenetic regulation of cork oak is very limited (Ribeiro et al. 2009; Correia et al. 2013). Up until now, no epigenetic players had been identified and characterized in this species. With this work, not only genes of four putative DNA methyltransferases and one associated protein were identified but also gene activity and DNA methylation content were correlated with physical characteristics of cork quality. With this approach, we can begin to unravel the correlations between cork quality and its potential epigenetic regulation in *Q. suber*. For this, cork quality was objectively determined by image analysis and density methods, and the structure of DNA methyltransferases was characterized. The transcriptional profile of the putative MET1, MET2, CMT3

and DRM2, and one DMAP1 in *Q. suber* was evaluated in the active phellogen of trees producing high and low quality cork, together with the relative quantification of global methylation levels.

Materials and methods

Sampling

Six *Q. suber* trees were selected from a cork oak stand (*montado*) in production in the Portalegre district (Portugal) based on a visual observation of small cork samples known as *calas*. Since active phellogen is a material difficult to obtain, we choose six trees with very contrasting cork quality. From those six cork oaks, three were trees producing high quality cork (HL12, HL14, and HL16), and the other three produced cork of low quality (hl2, hl3, and hl18). To analyze gene expression, phellogen and contiguous differentiating tissue were scraped from the inner side of the cork planks immediately after their extraction from the trees, and immediately kept in liquid nitrogen. Those cork planks were kept to assess cork quality. This sampling was done in June (2011) when phellogen activity is more intense, in coincidence with the period of cork commercial harvesting.

Cork quality evaluation

The cork planks harvested were cut in 30×30 cm boards in order to assess the cork's thickness and density, before and after boiling.

Cork's thickness was measured in eight positions: two in each radial surface and two others in each transversal face. The measured positions were marked for comparison with data obtained after boiling.

For the determination of density, the volume of cork samples was assessed through the Archimedes' principle: mass variation as a consequence of water dislocation when cork was submerged.

To boil the cork, the 30×30 cm samples were submerged in water at approximately 90 °C for 1 h. During the drying process, cork was kept protected from direct light and was flattened with the action of weights. To control cork dryness, the moisture content was periodically measured.

Defects were quantified in the two transversal and two radial surfaces of 20×20 cm boards cut from the previous 30×30 cm boiled planks, and also in four strips of 20×1 cm (obtained from the same 30×30 cm samples), corresponding to transversal and radial surfaces (two strips for each surface). Each of the four faces of the 20×20 cm boards and the two faces of the 20×1 cm strips were photographed with a camera (10 mega pixels resolution). The defect areas were measured by image analysis, defining the limiting color threshold by visual inspection, using the Leica QWin software (V3.X—Leica Microsystems, Germany). Lengths were calibrated using a calibration ruler. The two types of defects were measured separately: cork porosity and nail inclusions. Defect areas were expressed as a percentage of the total area of cork under analysis.

Putative *QsDNMTs* and *QsDMAP1* characterization

Sequences of interest, putatively homologous of characterized methyltransferases from other plant species, were obtained from expressed sequence tags (ESTs) libraries belonging to the Cork Oak EST Consortium, 2010. Potential methyltransferase sequences were used to search homologies in NCBI database (http://www.ncbi.nlm.nih.gov) with algorithms blastn, blastp, and blastx (Altschul et al. 1997) in order to confirm its annotation.

The amino acid sequence of each chosen EST was used to identify the motifs present in *Q. suber* sequences, by comparison with orthologous proteins (Supplemental File 1) of other characterized species (Cao et al. 2000; Kim et al. 2007; Militello et al. 2008; Papa et al. 2001; Posfai et al. 1989). Motifs regions were aligned with ClustalW (http://www.ebi. ac.uk/Tools/msa/clustalw2/) (Larkin et al. 2007; Goujon et al. 2010) and logos were created with WebLogo algorithm (http://weblogo.berkeley.edu/) (Crooks et al. 2004). Protein mass was estimated with the algorithm ProtParam (http:// web.expasy.org/protparam) (Gasteiger et al. 2005).

QsDNMTs and QsDMAP1 were subjected to a phylogenetic analysis with orthologous sequences. All sequences were aligned with MUSCLE (http://www.phylogeny.fr/ version2_cgi/one_task.cgi?task_type=muscle) and the alignment was corrected with GBLOCKS (http://www. phylogeny.fr/version2_cgi/one_task.cgi?task_type=gblocks) (Dereeper et al. 2008). Phylogeny analysis was obtained with MEGA 5 software (Tamura et al. 2011), using the maximum likelihood method.

Nucleic acid extraction and manipulation

Total RNA was extracted from samples collected during cork extraction of three high quality cork producers (HL12, HL14, and HL16) and from three low quality cork producers (hl2, hl3, and hl18) with Spectrum™ Total RNA kit (Sigma-Aldrich, Inc, Spain), according to the manufacturer's instructions. To obtain cDNA, the RETROscript® kit (Ambion, Life Technologies, Spain) was used, following the manufacturer's protocol. RNA and cDNA concentrations were measured in a microplate reader Synergy HT (Biotek, Germany), using the software Gen5™ (Biotek, Germany) for nucleic acid quantification.

Genomic DNA was extracted from the phellogen-derived tissue of the same trees with CTAB DNA isolation protocol (Doyle and Doyle 1990). DNA was then concentrated and

desalted by precipitation. The samples were then resuspended in an appropriate volume of nuclease-free, sterile water.

Global DNA methylation

To quantify relative amounts of DNA global methylation in the six phellogen-derived samples, an ELISA-based colorimetric assay (Imprint® Methylated DNA Quantification Kit, Sigma-Aldrich, Spain) was used according to the manufacturer's instructions. Briefly, for each sample, the methylation analysis was performed in triplicate aliquots with 100 ng of genomic DNA each. Purified DNA was then immobilized to a strip well with a high affinity for it. The methylated fraction of DNA was identified using a 5-methylcytosine monoclonal antibody and quantified by an ELISA-like reaction. The optical density was read on a microplate reader (Synergy HT, Biotek, Germany) at 450 nm and the levels of methylated DNA are proportional to its values.

Real-time PCR conditions and analysis

Polymerase chain reactions were performed in optical 96-well plates with an IQ^{TM}5 Real Time PCR (Bio-Rad, Hercules, CA). The 20 µL reaction mixture was composed of cDNA, 0.5 µM gene-specific primers (Table 1), and 2× master mix (SsoFast^{TM}_EvaGreen® Supermix, Bio-Rad, Hercules, CA). Amplification of PCR products was monitored via intercalation of Eva-Green (included in the master mix). The following program was applied: initial polymerase activation, 95 °C for 2 min; then 40 cycles at 95 °C for 15 s (denaturation), 56 °C for 30 s (annealing), and 76 °C for 30 s (extension) with a single fluorescence reading taken at the end of each cycle. Each run was completed with a melting curve analysis to confirm the specificity of amplification and the lack of primer dimers.

To generate a baseline-subtracted plot of the logarithmic increase in fluorescence signal (ΔRn) versus cycle number, baseline data were collected between the cycles 7 and 20. All amplification plots were exported into a MS Excel 2010 workbook (Microsoft Inc., DC, USA), and the software LinRegPCR (http://www.hartfaalcentrum.nl/) (Ramakers et al. 2003) was used to obtain Cq (quantification cycle) and primer efficiency. The Normfinder algorithm (http://www.mdl.dk/publicationsnormfinder.htm) (Andersen et al. 2004) was used to assess the stability of the genes.

Cloning and sequencing of QsDNMTs and QsDMAP1 real-time PCR reactions

cDNA of *QsDNMT* and *QsDMAP1* was amplified through the following protocol: initial denaturation at 94 °C for 3 min, 29 cycles of amplification, each of which consisted of 30 s of denaturation at 94 °C; 30 s of annealing at 55 °C for *QsCMT3*, *QsDRM2*, *QsDMAP1*, and *QsMET2* and 51 °C for *QsMET1*; and 1 min of elongation at 72 °C with a final elongation step of 5 min at 72 °C. After purification with the QIAquick® PCR purification kit, the amplified fragments were cloned using pCR 2.1-TOPO vector (TOPO® TA Cloning® kit, Invitrogen, Life Technologies, Spain) and sequenced. Sequences were deposited in GenBank with the following accession numbers: [GenBank:KC121544, GenBank:KC121545, GenBank:KC121546, GenBank:KC121547, GenBank:KC121548].

Statistical analysis

Statistical analysis was performed in order to cluster data from each group of trees (hl or HL) and to evaluate the statistical differences between these two groups. ANOVAs and Welch *t* tests (two-tailed) analysis were performed on cork density, all percentage of defects analysis and also to all gene expression studies, using GraphPad Prism V5.0 software (Graph-Pad©, San Diego, CA). Standard error is provided along results (±SE). Global DNA methylation was also statistically studied with Welch *t* test (one-tailed) on GraphPad Prism V5.0 software, aiming to determine the statistical significance of the differences between samples of high and low quality cork.

Table 1 Specific primers used in qRT-PCT. qRT-PCR amplification efficiency was measured by LinRegPCR software (Ramakers et al. 2003)

Target gene	Primers	Product size (bp)	qRT-PCR eff.
Actin	5′-TGGATTCTGGTGATGGTGTGAGTC-3′	162	1.1
	5′-CAATTTCCCGTTCAGCAGTAGTGG-3′		
QsMET1	5′-GCTGATGATGGAAGTGG-3′	160	1.0
	5′-TACGGATTGAAATGAAAAC-3′		
QsMET2	5′-GTTATTATCGGTATGTGAAGG-3′	238	1.0
	5′-ATCGGAGTAAGGGAGCAA-3′		
QsDRM2	5′-TCCCTTTGAAGAATGTTG-3′	224	1.0
	5′-GATGCCTGTTGCTACCTG-3′		
QsCMT3	5′-CGTTTATGGTGGGATGAG-3′	200	1.0
	5′-ACAGGAACCGCAACAGCA-3′		
QsDMAP1	5′-TATCACTGGGTCAGGGTTGT-3′	366	1.1
	5′-GCTCGTTTGCGTTCTATCTC-3′		

Results

High quality cork is less dense

All cork is boiled before industrial processing and after boiling cork grows in size particularly in the radial direction. Density assessment, before and after boiling, was performed showing that the high quality corks (HL12, HL14, and HL16) have lower density, in opposition to low quality corks (hl2, hl3, and hl18) which present higher density (Fig. 1). The difference between the two quality groups was increased by boiling (p value<0.0001): on average, high quality cork has a density of 0.2217 g cm^{-3} (SE±0.0004) and low quality cork has a density of 0.2850 g cm^{-3} (SE±0.007).

High quality cork has fewer defects

To evaluate the percentage of defects present in all sampled corks, the Leica QWin software was used to calculate the total area of nail inclusions, porosity, and the total cork area without the back of the cork (Fig. 2). With these different areas, it was possible to quantify the percentage of nail inclusions, the percentage of porosity, and the percentage of total defects, adding the two former values (Fig. 3). In the total defects category, the differences between the two groups are highly statistically different (p value<0.0001). HL12 cork (high quality) had the lowest value, 5.44 % (SE±0.47) (Fig. 3a). The high quality group presented, in average, a percentage of defects of 6.88 % (SE±0.28). In opposition, low quality corks have a percentage of defects of 12.63 % (SE±0.41), with the maximum of 14.42 % (SE±0.46) in hl18 cork.

High and low quality corks can also be distinguished by the percentage of nail inclusions (p value<0.0001) ranging from an average of 2.60 % (SE±0.20) in high quality corks and 9.03 % (SE±0.53) in the average of low quality corks.
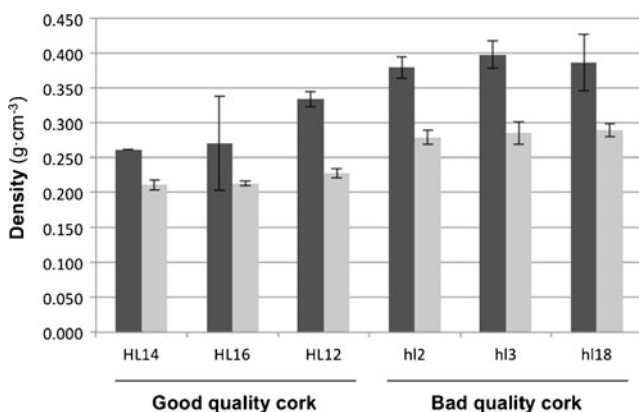

Fig. 1 Good corks are less dense than bad corks. Density of the 30× 30 cm planks, ±standard error, assessed through the Archimedes' principle. *Dark gray bars*, cork density before boiling; *light gray bars*, density of boiled cork

Concerning the study of porosity, high quality corks present the lowest values: HL16 shows the highest value of 5.49 % (SE±0.31) and HL12 the lowest of 3.21 % (SE±0.32). On average, high quality corks showed 4.28 % (SE±0.24) of the surface covered by pores, a value not significantly different (p value=0.8787) from the average of the low quality cork (5.12 %±0.21).

In order to establish the better plane to evaluate cork defects, measurements were performed in the radial and the transversal surfaces. No significant differences (p value=0.3198) were detected in the percentage of nail inclusions and porosity when both planes were considered. The total percentage of defects varied between 14.22 % (SE±0.85) and 13.94 % (SE±0.75) in the radial and transversal surfaces of hl18, respectively, and 5.52 % (SE±1.10) and 6.06 % (SE±0.76) in the radial and transversal planes of HL12 (Fig. 3b).

Q. suber DNMTs and DMAP1 proteins present all expected domains

The deduced amino acid sequences derived from the putative translation of the five ESTs studied revealed the following: *QsMET1* encodes a putative partial protein with an estimated number of 243 amino acids, where only the replication *foci* domain can be detected (Fig. 4); *QsMET2* encodes a putative protein with 387 amino acids and a predicted mass of 43,960 kDa, which contains all ten domains described for MET2 proteins: I, II, III, IV, V, VI, VII, VIII, IX, and X, from N to C terminal. All domains seem much conserved in the angiosperms used in this study, except domain VII (see Supplemental file 1 for list of angiosperms used and Supplemental File 2 for logos); *QsDRM2* is expected to contain 590 amino acids and have a molecular weight of 66,291 kDa. As other DRMs, *QsDRM2* contains two UBA domains, and all the eight DRM class domains in the canonical order: VI, IX, X, I, II, III, IV, and V (Fig. 4). All these eight *QSDRM2* domains seem to be conserved along angiosperms (Supplemental Files 1 and 2); *QsCMT3* isolated sequence should codify 446 amino acids which contain a chromodomain and six of the eight expected domains: IV, VI, VII, VIII, IX, X (Fig. 4). The two domains not identified in this work (domains I and II) may be localized near the N terminal region, which was not present in the EST library. Nevertheless, the identified domains of *QsCMT3* are very well conserved when compared with other angiosperms (Supplemental Files 1 and 2); *QsDMAP1* is putatively complete and estimated to be composed by 450 amino acids, with an estimated mass of 50,275 kDa and containing the expected SANT (SWI3, ADA2, N-CoR and NFIIIB DNA-binding domain) domain (Fig. 4).

Fig. 2 Example of defects on cork. The *yellow arrows* point to pores and the *red arrow* points to a nail inclusion. **a** Cork sample strip before the delimitation of the defects area. **b** All defects marked by Leica QWin software (V3.X—Leica microsystems) in order to evaluate general defects on this cork strip
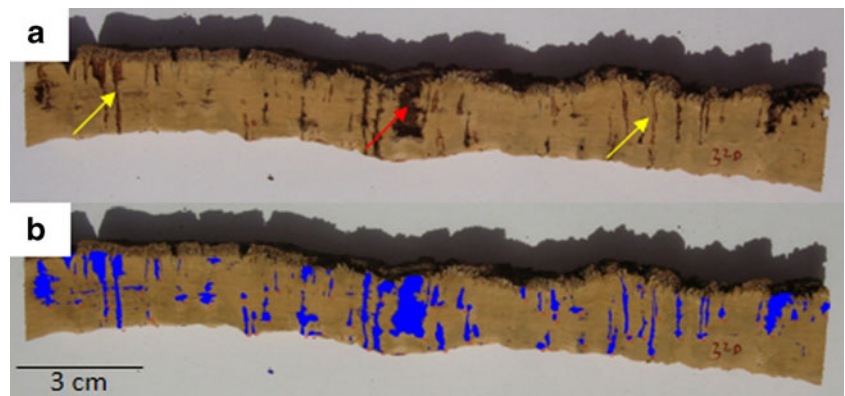
**DNMTs and DMAP1 proteins group well with their orthologous**

To determine the relationships between *Q. suber* methyltransferases and other known methyltransferases, we performed alignments with the conserved catalytic motifs against orthologous sequences (Supplemental File 1), which were then used to generate a phylogenetic tree. All the proteins are well individualized and each *Q. suber*

sequence grouped with high bootstrap values to the clades of each type of proteins (Fig. 5).

DNMTs gene expression on *Q. suber* active phellogen and derivative cells

qRT-PCR efficiencies (Table 1) and melting curves (data not shown) reveled correct amplifications to all genes. The stability of *actin* and the other *Q. suber* genes was

Fig. 3 Percentage of defects evaluation on cork sample. **a** Percentage of total defects, nail inclusions, and porosity on sampled corks. *Light gray bars*, good corks; *dark gray bars*, bad corks. For each kind of defect the best cluster is provided. **b** Comparison of the percentage of defects according to the considered plane. *Light gray bars*, radial plane; *dark gray bars*, transversal plane. Significant differences were observed for *p* value<0.05
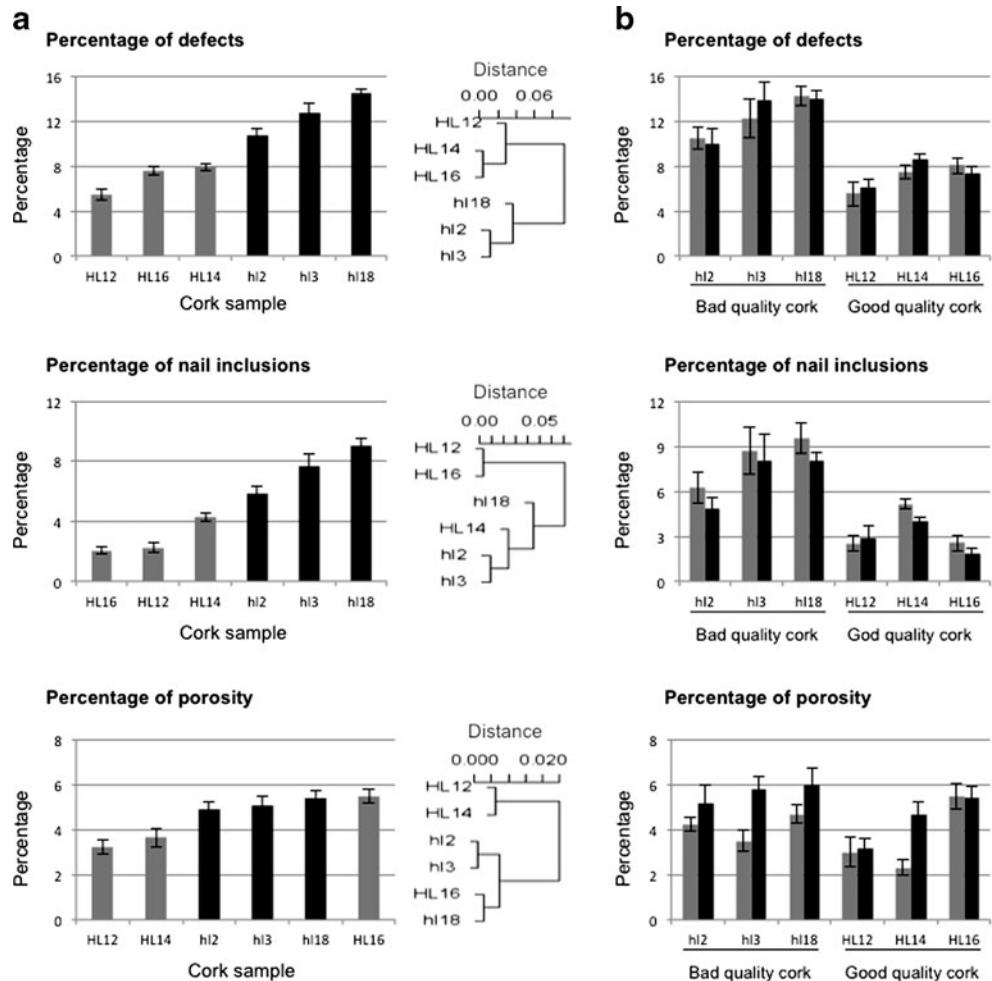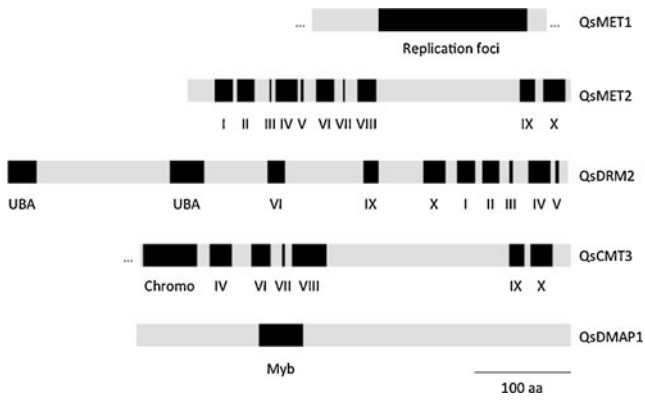
Fig. 4 Conserved domains present in the amino acid sequences of QsDNMTs and QsDMAP1. Protein QsMET1 is not complete in both ends and QsCMT3 is not complete in the N terminal

evaluated using the algorithm Normfinder (Andersen et al. 2004) and gene expression was normalized to the average

gene expression of trees producing low quality cork (hl) (Fig. 6a). *Q. suber actin* showed a stability coefficient of 0.741 which is within the appropriate range for a reference gene (<1.5). *QsDRM2* was the gene with the most stable expression, with a coefficient of 0.589. In opposition, *QsMET1* showed the highest level of instability, with a coefficient of 1.754.

In order to evaluate *QsDNMTs* gene expression and the relative contribution of each DNA methyltransferase in the active phellogen-derived cells of *Q. suber*, gene expression was estimated in all the trees of this study, with values normalized to *actin* and referenced to *QsDRM2*. All the genes in study revealed to have lower expression when compared to *QsDRM2*, ranging from 4 % in *QsMET1* to circa 40 % in *QsMET2* and *QsCMT3*. *QsDMAP1* showed only 15 % of the value of the reference (Fig. 6b).
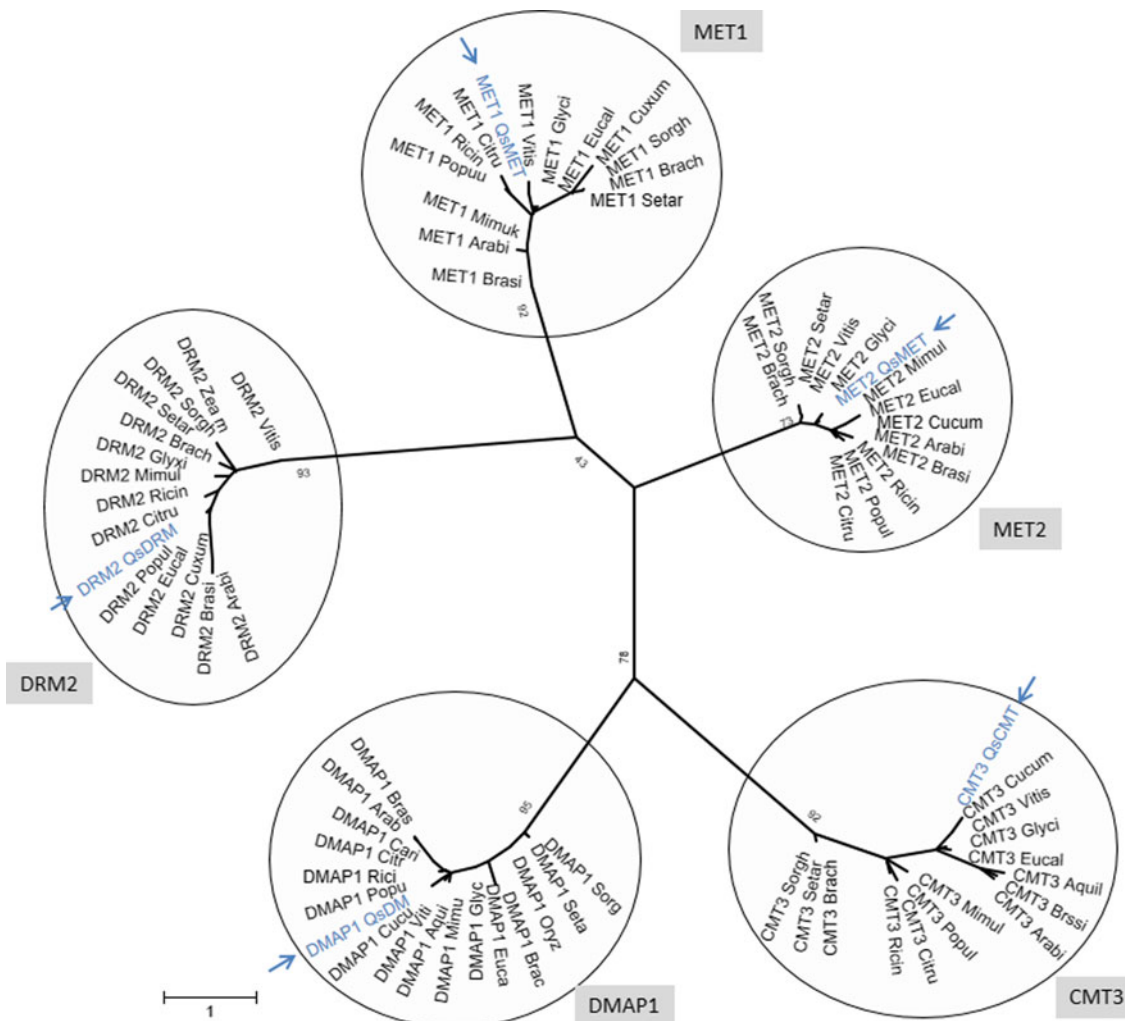

Fig. 5 Phylogenetic relation between QsDNMTs or QsDMAP1 and their paralogous sequences. Phylogenetic tree created on MEGA 5 software (http://www.megasoftware.net/), with the maximum likelihood method and a bootstrap of 1,000. All sequences were alignment in MUSCLE (http://www.phylogeny.fr/version2_cgi/one_task.cgi?task_type=muscle) and the aliment was corrected with GBLOCKS (http://www.phylogeny.fr/version2_cgi/one_task.cgi?task_type=gblocks). Sequences used are listed in Supplemental File 1. *Arrows* indicate *Quercus suber* sequences
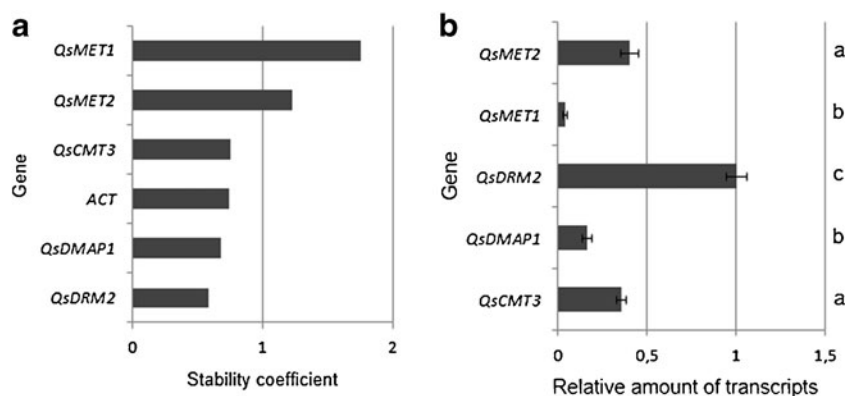
Fig. 6 General gene stability and expression analysis on *Q. suber* phellogen-derived tissue. **a** Gene stability on *Q. suber* phellogen-derived tissue assessed with Normfinder algorithm (http://www.mdl.dk/publicationsnormfinder.htm) with data normalized to the bad cork quality producers (hl). **b** Relative gene expression on *Q. suber* phellogen-derived tissue considering all good and bad cork producers, in relation to *QsDRM2* expression. Statistical analysis comparing each good cork sample with the respective control, *different letters* correspond to significant differences for *p* value<0.05

Aiming to compare the differential gene expression of DNMTs in trees producing different quality cork, gene expression was measured individually in each tree, in relation to the average of bad cork producers (hl). Exceptionally, HL14 showed the highest levels of expression of *QsCMT3* and *QsDRM2* among all the trees studied (with relative amounts of transcript of 1.53 ± 0.030 and 1.24 ± 0.137, respectively). This tree also presented the highest expression of *QsMET1* (0.72 ± 0.204) and *QsDMAP1* (0.70 ± 0.023). On the other hand, all high quality cork producers consistently showed low levels of expression of *QsMET1*, *QsMET2*, and *QsDMAP1*. HL16 showed the lowest relative expression with 0.08 (±0.038) relative amount of *QsMET1* transcripts, 0.27 (±0.027) of *QsMET2*, and 0.35 (±0.112) of *QsDMAP1*. *QsMET2* showed the highest expression in HL12, with relative amounts of transcript of 0.62 (±0.112) (Fig. 7). Concluding, in average, high quality cork trees showed lower levels of DNMTs gene expression than low quality producers.

Low quality cork is less methylated than high quality cork

In order to establish a relationship between *QsDNMTs* gene expression and the amount of methylation in the differentiating cells from phellogen, the relative global DNA methylation levels were evaluated in samples belonging to all the six trees analyzed. One sample of high cork quality, HL14, and one of low cork quality, hl3, were not considered as there was no consistency between replicates. The low quality cork samples presented 15 % less global methylation than high cork quality samples (*p* value<0.01) (Fig. 8).

QsDMAP1 may be involved in cork quality

In order to assess any relationship between the gene expression of DNMTs or associated proteins and the cork quality, a

correlation between the two groups of results was estimated. From the five genes studied, only *QsDMAP1* and *QsMET2* allowed us to group producers of high quality cork independently of the group of low cork quality producers. The expression of *QsDMAP1* revealed a positive and significant correlation to all cork quality parameters ($0.50<R^2<0.59$), with the exception of porosity (Supplemental File 3).
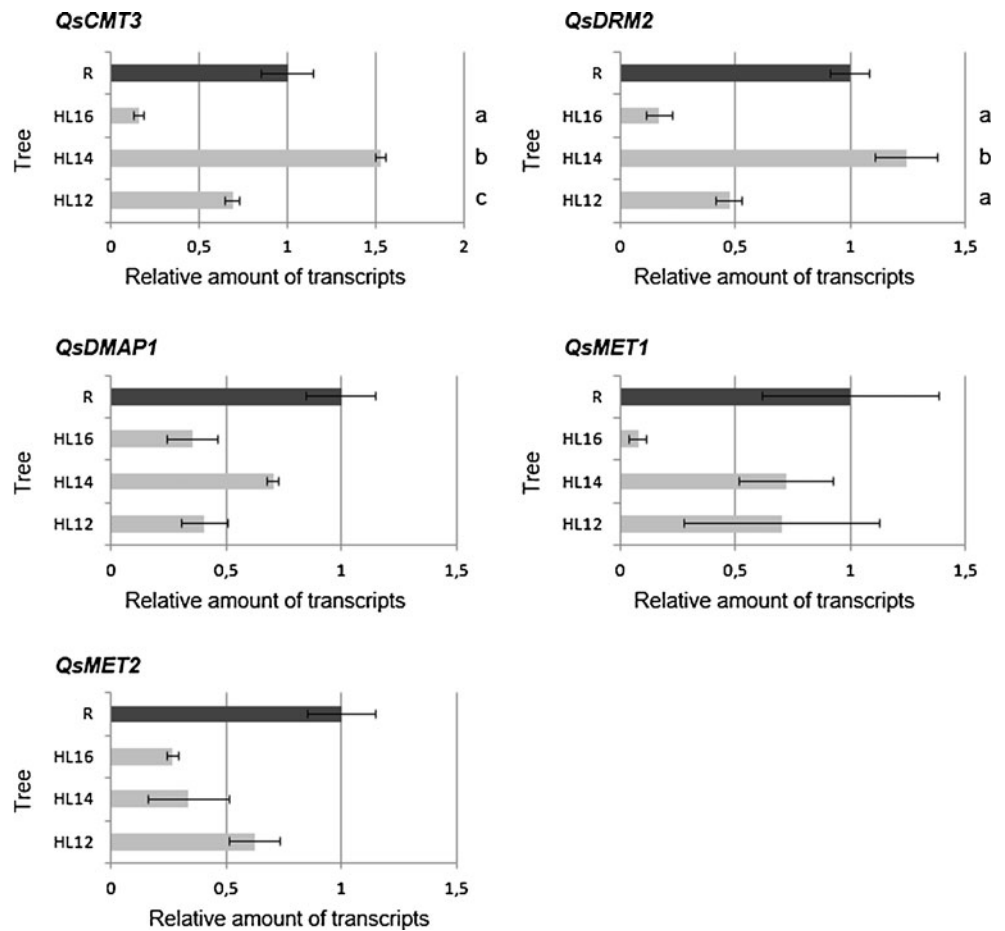
Discussion

The relationship between cork quality and DNA methylation was assessed for the first time in active phellogen differentiation tissue of two groups of trees producers of high and low quality cork.

High quality corks generally have less percentage of defects and are less dense

Analysis of cork quality revealed that high quality corks are less dense and have lower percentage of defects. Density increases with nail inclusions, as these inclusions are formed by lignified cells arising from secondary phloem fibers, which are harder and denser than cork (Pereira 2007). Due to its hardness, nail inclusions are the most relevant defects affecting the industrial use of cork (Gonzalez-Adrados and Pereira 1996). Data from porosity seems to confirm these results, since a high quality cork producer tree was the one with the highest percentage of pores. Pores play a fundamental role in cork trees due to their function in gas exchange (Pereira 2007); therefore, they cannot be avoided and should be considered as an unwanted characteristic but not a real "defect." On the other hand, the percentage of porosity was obtained after boiling and flattening of the planks, which may have led to the collapse of pores, increasing and masking the cork's original porosity. Nevertheless, boiling is a basic

Fig. 7 Gene expression with data normalized to the average of bad cork quality producers. Relative gene expression variation in each cork oak tree. *Dark gray bars*, reference value (average of bad cork quality producers—hl); *light gray bars*, trees producing good cork quality. Statistical analysis comparing each good cork sample with the respective control, *different letters* correspond to significant differences for p value<0.05 between trees producing good cork quality. Absence of letter means no significant differences

Fig. 8 Global levels of cytosine methylation of good and bad quality corks. The results are expressed as fold induction of optical density values at 450 nm relative to good cork samples. Data are shown as mean±SD

industrial treatment applied to all the reproduction cork used, which improves its properties and quality, as we have shown through density evaluation before and after boiling.

Cork is formed by the activity of the phellogen and involves several steps, from the proliferation and reprogramming of the phellogen derivatives, through extensive deposition of suberin and waxes, towards an irreversible program of cell death (Pereira 2007). Appropriate DNA methylation patterns which are established and maintained by DNA methyltransferases (DNMTs) are essential for correct development and tissue differentiation (Cao et al. 2000; Finnegan and Kovac 2000; Huang et al. 2010; Wada 2005). To study the involvement of DNA methylation in this process, we first characterized the structure of the four types of DNA methyltransferases (QsDNMTs) and one associated protein, DMAP1, in *Q. suber*.

DNMTs are well conserved in *Q. suber*

The predicted proteins of the five *Q. suber* ESTs, *QsMET1*, *QsMET2*, *QsCMT3*, *QsDRM2*, and *QsDMAP1*, show high similarity with plant homologues (Goll and Bestor 2005). All described sequences have the expected domains of DNMTs and DMAP1 although in *QsMET1* and *QsCMT3* some domains were not detected, since the EST sequences for both genes are incomplete, and the lacking domains are located in the missing N-terminal. In summary, the replication *foci* domain is present in *QsMET1*; all ten domains are present in *QsMET2*; the eight expected characteristic domains of *QsDRM2* are also present, in the same order as described in other species (Cao et al. 2000), together with two UBA
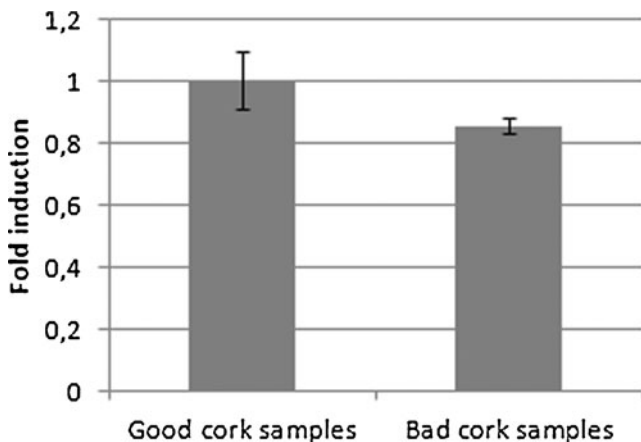
domains; and seven of the nine expected domains of *QsCMT3* are present, including a chromodomain before domain IV, characteristic of the CMT class (Genger et al. 1999). Furthermore, *Q. suber* domains showed high sequence similarity with the same domains of other angiosperms (Goll and Bestor 2005) revealing a high degree of conservation, presumably to perform the same functions.

The phylogenic analysis of *QsDNMTs* and *QsDMAP1* putative amino acid sequences supports the idea that these five sequences are cork oak orthologues of each DNMT class and DMAP1 in angiosperms and are able to methylate DNA in different sequence contexts and developmental situations.

De novo methylation is essential during cork differentiation

The expression of the characterized genes was also evaluated in the phellogen derivatives. Our data revealed that *QsDRM2* is the most stable gene in the active phellogen and derived cells of *Q. suber*, with a stability coefficient lower than *actin* which is currently used to normalize gene expression in qRT-PCR analysis in cork oak (Marum et al. 2012). *DRM* stability had also been referred by Rival et al. (2008) who found that this gene is relatively stable over time and within plant tissues, in *Elaeis guineensis*. *QsDRM2* was also the gene with the highest expression in derived cells from active phellogen. Although *DRM2* is known to be active in all of the major tissues of the plants (Cao et al. 2000), its high activity in the phellogen-differentiated cells is not surprising, owing to the active differentiation program of this tissue (Pereira 2007; Şen et al. 2011; Soler et al. 2008). These cells are in a unique developmental process where a specific epigenetic reprogramming including both maintenance and de novo methylation is likely to occur.

The putative pseudo gene *QsMET1*, responsible for maintenance of DNA methylation, was the gene with the lowest levels of expression, which contrasts with the high expression of *QsDRM2*. However, *QsMET2*, homologous of the enigmatic *DNMT2* (Dong et al. 2001), presented the second highest level of expression, after *QsDRM2*. This fact seems to indicate that QsMET2 instead of QsMET1 can achieve the maintenance of DNA methylation in *Q. suber* analyzed tissue. Although there is a lack of DNA methylation activity associated with DNMT2 (review in Schaefer and Lyko 2010), *QsMET2* has the ENV and PCQ peptides in motifs VI and IV, respectively (Supplemental File 2), that are necessary for DNA methyltransferase activity (Schaefer and Lyko 2010). In addition, DNMT2 is the only DNA methyltransferase present in all organisms (Goll and Bestor 2005) being the unique DNMT present in *Drosophila* (Kunert et al. 2003) and *Dictyostelium* (Katoh et al. 2006), where DNA methylation has been associated with retroelements (Phalke et al. 2009). The ubiquitous presence of MET2 in all eukaryotes points to an essential role of this protein, which, in *Q. suber* could be the fundamental maintenance of DNA methylation.

Relative global DNA methylation is not associated with the expression level of *QsDNMT* genes

The quantification of *DNMT* gene expression in cells of cork oaks producing cork of high or low quality allowed us to find a significant down regulation of all the *DNMTs* and related genes analyzed in high quality producers. Unexpectedly, we found a negative correlation between the relative levels of expression of *DNMTs* and global DNA methylation in the phellogen-derived cells of both groups of cork producers. However, this paradox was also observed in the oil palm tree *E. guineensis* where it was shown that the hypomethylation earlier described in the "mantle" tissue cannot be explained by a decrease in the expression of these *DNMTs* (Rival et al. 2008). Also, this negative correlation is well known in tumor cells where a global genome hypomethylation is observed together with hypermethylation of regions rich in CpG islands (review in Robertson, 2001).

Our data indicate that the global level of DNA methylation quantified in the phellogen and derivative tissue is not directly related with the levels of *DNMT* transcription. Global genome DNA methylation should result from a balance between the activity of DNA methyltransferases and DNA demethylases. One hypothesis to explain our results would be the overexpression of DNA demethylases in low quality producers that would lead to a global DNA hypomethylation despite the level of expression of *DNMTs*. In *Arabidopsis*, Penterman et al. (2007) showed that DRM2 is required for the DNA demethylase *ROS1* expression and therefore for DNA demethylation, and also that *ROS1* is downregulated in plants with mutations in the *DRM2*, resulting in overmethylation of specific sequences.

QsDNMTs and *QsDMAP1* are differentially expressed in relation to cork quality

A correlation between gene expression and cork quality revealed a tendency for higher levels of gene expression in low quality cork producers. This is quite evident in the case of *QsDMAP1* which has a positive and significant correlation with nail inclusions ($R^2 \approx 0.6$), the most relevant defect in the industry of cork (Gonzalez-Adrados and Pereira 1996). DMAP1 proteins are described as having several cellular functions such as the involvement in the repression or activation of transcription while interacting with the N-terminal of DNMT1 or as components of the nucleosome acetyltransferase of the H4 complex (Doyon et al. 2004), processes which are involved in epigenetic regulation.

The involvement of DMAP1 in DNA repair and its action against genomic instability has also been reported (Negishi et al. 2009). Nail inclusions are formed by the death of small

portions of phellogen (Pereira 2007); therefore, during cork formation in these regions, the amount of DNA damages should be higher than in "normal" tissues and *QsDMAP1* expression may be increased in order to repair the DNA and therefore avoid genomic instability. This situation may account for the highest *QsDMAP1* expression detected in low cork quality producers in which cork presents higher amounts of nail inclusions. A regression analysis with *QsDMAP1* and *QsMET1* expression failed to reveal a significant correlation between the expression of both genes which suggests that *QsDMAP1* role in this cork oak tissue exceeds the function of interacting directly with DNMT1 during the S phase (Rountree et al. 2000; Shin et al. 2010).

Conclusions

The expression analysis of several DNMTs and the associated protein DMAP1 revealed that DNA methylation should be an important process that regulates cork formation, being de novo methylation, performed by *QsDRM2*, the most notorious type of methylation during cork differentiation.

Cork quality in *Q. suber* seems to be influenced by the expression profile of *QsDMAP1* which, therefore, makes this gene a good candidate to understand the molecular processes underlying cork quality.

Further studies on genes involved in developmental epigenetic programs such as DNA demethylases, members of RNAi pathways, and post-translational histone modifiers will shed light into the epigenetic regulation of cork formation.

Data archiving statement Gene sequences were archived on GenBank.

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25(17):3389–3402

Andersen CL, Jensen JL, Orntoft TF (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res 64(15):5245–5250

Cao X, Jacobsen SE (2002) Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. Proc Natl Acad Sci U S A 99(Suppl 4):16491–16498

Cao X, Springer NM, Muszynski MG, Phillips RL, Kaeppler S, Jacobsen SE (2000) Conserved plant genes with similarity to mammalian de novo DNA methyltransferases. Proc Natl Acad Sci U S A 97(9):4979–4984

Correia B, Valledor L, Meijon M, Rodriguez JL, Dias MC, Santos C, Canal MJ, Rodriguez R, Pinto G (2013) Is the interplay between epigenetic markers related to the acclimation of cork oak plants to high temperatures? PLoS One 8(1):e53543

Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. Genome Res 14(6):1188–1190

Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 36 (Web Server issue):W465-469

Dong A, Yoder JA, Zhang X, Zhou L, Bestor TH, Cheng X (2001) Structure of human DNMT2, an enigmatic DNA methyltransferase homolog that displays denaturant-resistant binding to DNA. Nucleic Acids Res 29(2):439–448

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Doyon Y, Selleck W, Lane WS, Tan S, Cote J (2004) Structural and functional conservation of the NuA4 histone acetyltransferase complex from yeast to humans. Mol Cell Biol 24(5):1884–1896

Finnegan EJ, Dennis ES (1993) Isolation and identification by sequence homology of a putative cytosine methyltransferase from *Arabidopsis thaliana*. Nucleic Acids Res 21(10):2383–2388

Finnegan EJ, Kovac KA (2000) Plant DNA methyltransferases. Plant Mol Biol 43(2–3):189–201

Fortes MA, Rosa ME, Pereira H (2004) A cortiça. IST, Lisbon

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed) The proteomics protocols handbook, 1st edn. Humana, Hoboken, pp 571–607

Genger RK, Kovac KA, Dennis ES, Peacock WJ, Finnegan EJ (1999) Multiple DNA methyltransferase genes in *Arabidopsis thaliana*. Plant Mol Biol 41(2):269–278

Goll MG, Bestor TH (2005) Eukaryotic cytosine methyltransferases. Annu Rev Biochem 74:481–514

Gonzalez-Adrados JR, Pereira H (1996) Classification of defects in cork planks using image analysis. Wood Sci Technol 30(3):201–215

Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R (2010) A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res 38 (Web Server issue):W695-699

Graça J, Pereira H (2004) The periderm development in *Quercus suber*. IAWA J 25:325–335

Huang J, Wand H, Xie X, Zhang D, Liu Y, Guo G (2010) Roles of DNA methyltransferase in *Arabidopsis* development. Afr J Biotechnol 9(50):8506–8514

Katoh M, Curk T, Xu Q, Zupan B, Kuspa A, Shaulsky G (2006) Developmentally regulated DNA methylation in *Dictyostelium discoideum*. Eukaryot Cell 5(1):18–25

Kim HJ, Yano A, Wada Y, Sano H (2007) Properties of a tobacco DNA methyltransferase, NtMET1 and its involvement in chromatin movement during cell division. Ann Bot 99(5):845–856

Kunert N, Marhold J, Stanke J, Stach D, Lyko F (2003) A Dnmt2-like protein mediates DNA methylation in *Drosophila*. Development 130(21):5083–5090

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23(21):2947–2948

Leal S, Sousa V, Pereira H (2006) Within and between-tree variation in the biometry of wood rays and fibres in cork oak (*Quercus suber* L.). Wood Sci Technol 40(7):585–597

Martienssen RA, Colot V (2001) DNA methylation and epigenetic inheritance in plants and filamentous fungi. Science 293(5532):1070–1074

Marum L, Miguel A, Ricardo CP, Miguel C (2012) Reference gene selection for quantitative real-time PCR normalization in *Quercus suber*. PLoS One 7(4):e35113

Militello KT, Wang P, Jayakar SK, Pietrasik RL, Dupont CD, Dodd K, King AM, Valenti PR (2008) African trypanosomes contain 5-methylcytosine in nuclear DNA. Eukaryot Cell 7(11):2012–2016

Natividade JV (1950) Subericultura, 1st edn. Direcção Geral das Florestas, Lisbon

Negishi M, Chiba T, Saraya A, Miyagi S, Iwama A (2009) Dmap1 plays an essential role in the maintenance of genome integrity through the DNA repair process. Genes Cells 14(11):1347–1357

Papa CM, Springer NM, Muszynski MG, Meeley R, Kaeppler SM (2001) Maize chromomethylase *Zea methyltransferase2* is required for CpNpG methylation. Plant Cell 13(8):1919–1928

Penterman J, Uzawa R, Fischer RL (2007) Genetic interactions between DNA demethylation and methylation in *Arabidopsis*. Plant Physiol 145(4):1549–1557

Pereira H (2007) Cork: biology, production and uses, 1st edn. Elsevier, Amsterdam

Pereira H, Lopes F, Graça J (1996) The evaluation of the quality of cork planks by image analysis. Holzforschung 50(2):111–115

Phalke S, Nickel O, Walluscheck D, Hortig F, Onorati MC, Reuter G (2009) Retrotransposon silencing and telomere integrity in somatic cells of *Drosophila* depends on the cytosine-5 methyltransferase DNMT2. Nat Genet 41(6):696–702

Posfai J, Bhagwat AS, Posfai G, Roberts RJ (1989) Predictive motifs derived from cytosine methyltransferases. Nucleic Acids Res 17(7):2421–2435

Ramakers C, Ruijter JM, Deprez RH, Moorman AF (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci Lett 339(1):62–66

Ribeiro T, Viegas W, Morais-Cecilio L (2009) Epigenetic marks in the mature pollen of *Quercus suber* L. (Fagaceae). Sex Plant Reprod 22(1):1–7

Ricardo CP, Martins I, Francisco R, Sergeant K, Pinheiro C, Campos A, Renaut J, Fevereiro P (2011) Proteins associated with cork formation in *Quercus suber* L. stem tissues. J Proteomics 74(8):1266–1278

Rival A, Jaligot E, Beule T, Finnegan EJ (2008) Isolation and expression analysis of genes encoding MET, CMT, and DRM methyltransferases in oil palm (*Elaeis guineensis* Jacq.) in relation to the 'mantled' somaclonal variation. J Exp Bot 59(12):3271–3281

Rountree MR, Bachman KE, Baylin SB (2000) DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. Nat Genet 25(3):269–277

Schaefer M, Lyko F (2010) Solving the Dnmt2 enigma. Chromosoma 119(1):35–40

Şen A, Quilhó T, Pereira H (2011) The cellular structure of cork from *Quercus cerris* var. cerris bark in a materials' perspective. Ind Crop Prod 34(4):929–936

Shin JH, Kang HC, Park YY, Ha DH, Choi YH, Eum HY, Kang BG, Chae JH, Shin I, Lee JH, Kim CG (2010) Corepressor MMTR/DMAP1 is an intrinsic negative regulator of CAK kinase to regulate cell cycle progression. Biochem Bioph Res Co 402(1):110–115

Soler M, Serra O, Molinas M, Garcia-Berthou E, Caritat A, Figueras M (2008) Seasonal variation in transcript abundance in cork tissue analyzed by real time RT-PCR. Tree Physiol 28(5):743–751

Song Y, Wu K, Dhaubhadel S, An L, Tian L (2010) *Arabidopsis* DNA methyltransferase AtDNMT2 associates with histone deacetylase AtHD2s activity. Biochem Biophys Res Commun 396(2):187–192

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28(10):2731–2739

Wada Y (2005) Physiological functions of plant DNA methyltransferases. Plant Biotechnol 22(2):71–80