miRNA profiling in leaf and cork tissues of *Quercus suber* reveals novel miRNAs and tissue-specific expression patterns

Inês Chaves[1,2], Yao-Cheng Lin[3,4], C. Pinto-Ricardo[2], Yves Van de Peer[3,4], Célia Miguel[1,2]

Affiliations

[1] iBET - Instituto de Biologia Experimental e Tecnológica, Apartado 12, 2780-901 Oeiras, Portugal

[2] Instituto de Tecnologia Química e Biológica-Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal

[3] Department of Plant Systems Biology, VIB, Technologiepark 927, 9052 Ghent, Belgium

[4] Department of Plant Biotechnology and Bioinformatics, Ghent University, Technologiepark 927, 9052 Ghent, Belgium

Abstract

The differentiation of cork (phellem) cells from the phellogen (cork cambium) is a secondary growth process observed in the cork oak tree conferring a unique ability to produce a thick layer of cork. At present, the molecular regulators of phellem differentiation are unknown. The previously documented involvement of microRNAs in the regulation of developmental processes, including secondary growth, motivated the search for these regulators in cork oak tissues.

We performed deep-sequencing of the small-RNA fraction obtained from cork oak leaves and differentiating phellem. RNA sequences with lengths of 19-25 nt derived from the two libraries were analysed, leading to the identification of 41 families of conserved miRNAs, of which the most abundant were miR167, miR165/166, miR396 and miR159. Thirty novel miRNA candidates were also unveiled, 11 of which were unique to leaves and 13 to phellem. Northern blot detection of a set of conserved and novel-miRNAs confirmed their differential expression profile. Prediction and analysis of putative miRNA target genes provided clues regarding processes taking place in leaf and phellem tissues but further experimental work will be needed for functional characterization. In conclusion, we here provide a first characterization of the miRNA population in a Fagacea species and the comparative analysis of miRNA expression in leaf and phellem libraries represents an important step to uncovering specific regulatory networks controlling phellem differentiation.

Keywords (5/6): small RNA; miRNA; *Quercus suber*; phellem; tree; secondary growth

Introduction

Cork oak is a long-lived evergreen tree endemic to the western-Mediterranean, where it is found both in native forests and in the human-shaped ecosystem of high socio-economic and ecological value, known as "montado" (Bugalho et al. 2011; Varela 1995). The economic value of the cork oak tree comes from cork, the out-bark of cork oak, which the tree continuously produces and is removed from the tree every nine years. Due to properties such as elasticity, thermal insulation and impermeability, conferred by its characteristic chemical composition (Pereira 2007), cork has several commercial applications of which cork stoppers are probably best known. Most importantly, the cork layer allows the tree to withstand adverse environmental conditions and confers protection against the periodic fires that are common throughout the Mediterranean (Pausas 1997). This rare feature among plant species results from the secondary growth sustained by the activity of cork cambium or phellogen, the lateral meristem that envelopes the tree trunk, giving rise each year to a new cork ring that is not shed naturally. Through periclinal divisions, the pluripotent cork cambium gives rise to daughter cells that differentiate either as phelloderm (inwards) or phellem (outwards). Differentiation of phellem or cork cells involves the deposition of a high amount of suberin in the cell walls (Pereira 1988) culminating in cell death.

The first genomic approach to suberin biosynthesis and cork differentiation used a suppression subtractive hybridization (SSH) strategy (Soler et al. 2007). In that work, a list of candidate genes isolated from cork tissues were identified including genes for the synthesis, transport, and polymerization of suberin monomers as well as a number of regulatory genes including members of the NO APICAL MERISTEM (NAM), MYB and WRKY transcription factor families with putative functions in meristem identity and cork differentiation. Since then, a number of studies have been published characterizing the role of a few selected candidate orthologues in potato tuber suberin formation by reverse genetics approaches (Serra et al. 2010a; Serra et al. 2010b; Serra et al. 2009a; Serra et al. 2009b) and identifying additional candidates in potato tuber periderm (Barel and Ginzberg 2008; Chaves et al. 2009; Ginzberg et al. 2009; Soler et al. 2011). Despite the remarkable ability of cork oak to produce cork, potato has been the preferred model to study these processes since it is more amenable to *in vitro* manipulation and transformation having, in addition, its genome fully sequenced (Xu et al. 2011). However, it is becoming increasingly evident that species-specific regulators are crucial in the regulation of a wide range of processes including developmental processes (Osorio et al. 2012 ; Sunkar et al. 2007).

MicroRNAs (miRNAs) represent an important class of plant small RNAs regulating gene expression at post-transcriptional level during growth and development (Bartel 2004) and in response to abiotic stress and pathogens (Liu and Vance 2010 ; Sunkar 2010). These non-coding RNAs range in size from 20 to 24 nucleotides and negatively regulate gene expression by degrading target mRNAs or inhibiting their translation. There are a number of plant miRNAs that are highly conserved between species and play crucial roles in the regulation of large transcriptional networks, but the majority are family- or species-specific (Cuperus et al. 2011). In the latest version of the miRBase (release 19, http://www.mirbase.org) with data from 67 plant species (15 eudicotyledoneous families), no miRNAs are present for *Quercus* species or other members in the Fagaceae family.

Because cork oak is probably the species where the most extensive process of phellem differentiation is observed, it represents an attractive system to address the miRNA-mediated regulatory processes controlling phellem differentiation.

Although the similarities between the processes that occur as a result of the activity of vascular cambium and cork cambium suggest that conserved regulatory mechanisms may exist, there is much less knowledge about the molecular regulation of phellem differentiation than xylem differentiation, which has been the subject of intensive research due to its role in wood formation. During xylem differentiation a few miRNAs are implicated in the process. By comparing miRNAs from developing xylem in *Populus trichocarpa* stems and in *Arabidopsis*, Lu et al. (2005) found that conserved miRNAs exhibited species-specific developmental expression patterns. In this work, Lu et al further suggest that even conserved miRNAs may have different regulatory roles in different species and possible roles in tree-specific processes (Lu et al. 2005) such as secondary cell wall synthesis and deposition in developing xylem cells. A microarray analysis comparing primary and secondary xylem growth in poplar revealed a dominant expression of miR164, miR162a and miR168 in the secondary development zone (Dharmawardhana et al. 2010) reinforcing the role of miRNAs during secondary growth. In a recent study using *Acacia mangium* as lignification model system, the conserved miR166 was shown to be strongly down-regulated in xylem when compared to phloem, suggesting its potential role in xylem development by indirect regulation of the genes involved in lignin biosynthesis (Ong and Wickneswari 2012). Transgenic *Populus* expressing a miR166 resistant target gene have developed phenotypic abnormalities affecting the primary and secondary growth, including the abnormal formation of cambia within cortical parenchyma that can produce secondary vascular tissues in reverse polarity (Robischon et al. 2011). More recently, Puzey et al (2012) identified several xylem-enriched miRNAs predicted to target genes known to be important in secondary growth, including xyloglucan endo-transglycosylase/hydrolase and vascular-related transcription factors (Puzey et al. 2012).

In this work we aimed to identify novel and conserved miRNAs in cork oak and look for leaf and cork tissue-specific patterns of miRNA expression. Since no whole-genome sequence data is available for cork oak yet, we have conducted deep-sequencing of small RNA libraries and used available EST data for miRNA profiling and prediction of miRNA target genes in the sampled tissues. To our knowledge, this represents the first report on small RNA analysis in *Quercus* and Fagaceae species and provides novel resources to address the molecular regulation of phellem differentiation.

Material and methods

Small RNA isolation and sequencing

The first fully-expanded leaves were collected from 2 year-old plants. The phellem tissues were harvested from adult cork oak trees by scraping the inner surface of the removed cork plank with a scalpel. All tissues were immediately frozen in liquid nitrogen and stored at -80ºC until RNA isolation.

A protocol allowing the isolation of good quality low molecular weight RNA from leaves and phellem was established based on the CTAB method described by Chang (Chang et al. 1993), with following modifications:

0.5 ml of pre-warmed (65ºC) extraction buffer (300 mM Tris HCl pH 8.0, 25 mM EDTA, 2 M NaCl, 2% CTAB, 2% PVP10, 0.05% spermidine trihydrochloride, 2% β-mercaptoethanol) was added to 100 mg grinded tissue in 1.5 ml tube, followed by incubation at 65ºC for 10 min; samples were extracted twice with 1ml of chloroform:isoamyl alcohol (24:1) and centrifuged at 15000 xg for 10 min at 4ºC; 0.1 vol 3M NaOAc (pH 5.2) and 1 vol isopropanol was added to the supernatant and tubes were kept at -80ºC for 3h; nucleic acids were recovered by centrifugation at 15000 xg (30 min at 4ºC); the nucleic acid pellet was resuspended in 375 µl TE (pH7.5) and the high molecular weight (HMW) RNA was precipitated overnight at 4ºC with 140 µl of 8M

LiCl; after 30 min centrifugation at 15000 xg at 4ºC, the pellet for HMW RNA recovery can be kept if needed and the low molecular weight (LMW) RNA fraction is precipitated from the supernatant with 1 vol isopropanol and 0.1% SDS at -80ºC for 3 h; after centrifugation at 15000 xg, for 30 min at 4ºC, the pellet was washed with ice-cold ethanol (1ml), resuspended in 10 µl RNAse-free water and treated with DNAse Turbo (Ambion® TURBO™ DNase, Invitrogen-Life Technologies).

RNA quality and yield was checked in all samples by measuring the UV absorbance (A) at 230, 260 and 280 nm using the ND-1000 Spectrophotometer (NanoDrop Technologies Inc., USA). The RNA integrity of both the HMW and the LMW RNA fractions was also evaluated on the Agilent 2100 Bioanalyzer.

The purified LMW RNA fraction was prepared for sequencing by Illumina HiSeq 2000 (Fasteris, Switzerland). The smallRNA sequences obtained for leaf and cork are deposited in the NCBI GenBank database under the accession numbers SRR 988108 and SRR988109 respectively.

Small RNA sequence processing

After removal of the sequencing adapter, reads were further processed using the UEA sRNA toolkit (Moxon et al. 2008b). 'Filter' function and only read length ranging from 15 to 25 bp with no match with the plant rRNA/tRNA were kept for further analysis. We searched for the conserved mature miRNAs in the miRBase database release 18 by miRProf allowing two mismatches. The identified conserved miRNA were further grouped by match signature and combined by organism name and the miRNA family members. The number of reads of the conserved miRNAs was normalized with the total number of sequenced reads in each library.

Due to the lack of a reference cork oak genome we collected available genome and transcriptome sequence information to identify the miRNA precursors. To identify novel miRNA, we combined the Oak GeneIndex EST (http://compbio.dfci.harvard.edu/cgi-bin/tgi/gimain.pl?gudb=oak), Oak EST from NCBI (http://www.ncbi.nlm.nih.gov/dbEST/index.htm), Oak Genome Survey Sequence (GSS, http://www.ncbi.nlm.nih.gov/projects/dbGSS/), an in-house cork oak transcript database including three EST leaf libraries (http://www.corkoakdb.org/) (Pereira-Leal et al, submitted) and two phellem libraries from SRA (NCBI accessions SRX022024 and SRX022025) subsequently assembled by Newbler, as a customised enlarged Oak reference database (CEOR DB), and *Populus trichocarpa* genome. We applied the miRCat tool in the UEA package to identify the novel miRNAs with follow setting: minimum free energy (MFE): -10; maximum number of genome hits: - 5; maximum distance between consecutive hits on the genome: – 100; minimum length of the miRNA: 19; maximum length of the miRNA: 24. Overview of the analysis workflow is summarized in Figure 1.

Northern blot hybridization

For experimental validation of conserved and non-conserved miRNAs, northern blot hybridization was performed using LMW RNA from cork oak phellem and leaves. LMW RNA enriched samples (1 µg) were separated in a 17 % acrylamide gel containing 8M urea in TBE buffer. Gels were run at 180V in TBE buffer and RNA was subsequently transferred to Hybond N^+ membrane (Amersham Biosciences) in TBE buffer at 300 mA for 1 h (mini-protean® 3 BIO-RAD). After UV crosslinking the membrane was pre-hybridized in 10 ml of PerfectHyb^{TM} (Sigma) hybridization solution for 1 h at 42ºC. Afterwards, each specifically designed miRNA probe was generated using DIG labeling (DIG oligonucleotide 3-end labeling kit, 2nd generation, Roche) following the instructions of the manufacturer. Five pmol of probe were added and the membrane was hybridized overnight at 42ºC. The membrane was rinsed twice in 4xSSC for 20 min followed by 2xSSC/0.1% SDS for 15 min and incubated for 30 min in blocking solution (1 g blocking reagent (Roche) in 0.1 M maleic

acid buffer pH 7.5) followed by 1 h in antibody solution (anti-digoxigenin-AP, Roche). After four washes in washing buffer (0.1 M maleic acid buffer, 0.3% Tween 20, pH 7.5) for 15 min, the membrane was equilibrated in detection buffer (0.1 M Tris-HCl, 0.1 M NaCl pH 9.5), then incubated in CDP-Star (Ambion) working solution (1:100 dilution in detection buffer) and finally exposed to X-ray film. U6 DNA probe was used as sample loading control. The X-ray film images were acquired using the ChemiDoc™ XRS+ System and the intensity of the bands detected in the northern blot were measured using the Image Lab™ Software.

Prediction of miRNA target genes

Target prediction was performed for all predicted non-conserved miRNAs. An in-house cork oak transcript database containing information from cork oak (CEORT DB) was used. Due to the lack of directional information in the Cork Oak EST database and since the prediction algorithm only searches the target site on the forward strand, the transcript database was constructed to contain both the forward and the reverse complement sequences.

The target gene search was performed using the TAPIR software (Bonnet et al. 2010) with the follow settings: TAPIR hybrid (combining FASTA and RNAhybrid); maximum internal loop size: 5, number of hits per target: 10.

Results

Small RNA isolation and deep sequencing

To uncover regulatory mechanisms involving small RNA pathways that are specifically functioning in phellem differentiation, we have sequenced the small RNA transcriptome of cork oak developing phellem and leaves and performed expression analysis of the miRNAs present in both tissues.

To this end, we have used a modified CTAB extraction method which is effective for tissues with high polyphenol content such as cork oak phellem. The buffer described in this method stabilizes the polyphenols allowing the efficient isolation of total RNA. To increase the effectiveness of the extraction and recovery of high quality low molecular weight (LMW) RNA, we performed additional precipitation steps adapted from Zhu et al (2008) (Zhu et al. 2008). The A(260)/A(230) and A(260)/A(280) RNA absorbance ratios of the samples varied from 1.7-2.2 and 1.9-2.1, respectively, what evidences the absence of polysaccharide and protein contaminants. Additionally, the obtained yield was in the range of 10-15 µg high molecular weight (HMW) RNA and 1-1.5 µg LMW RNA per 100 mg of tissue. The electropherograms obtained for the HMW RNA fraction from phellem tissues showed a profile with a RNA Integrity Number (RIN) above 8 which is indicative of high quality RNA with no degradation. The electropherograms of the LMW RNA fraction obtained from the same tissue samples showed peaks corresponding to expected LMW RNA species (small ribosomal RNA (rRNA), transfer RNA (tRNA) and miRNAs) (Suplementary material **SM1_Electropherogram of RNA fractions**).

A total of 19,620,275 and 12,731,020 sequence reads were obtained from leaves and phellem of cork oak, respectively (Table 1). After removing reads shorter than 19-nt and longer than 25-nt and with an abundance of 3 or lower, 85.1% (16,692,453) and 68.9% (8,775,468) of the reads from leaves and phellem, respectively, remained for subsequent analysis.

The distribution by sequence length of the reads between 19 and 25-nt showed that the 24-nt length sequences were expressed at higher levels followed by the 21-nt length (Figure 2). When considering the total number of small RNA reads, the 21-nt class was as abundant in leaves (12.8%; 2,519,032 reads) as in

phellem (12.7%; 1,616,326 reads) but, interestingly, the 24-nt class was approximately twice more expressed in leaves (60%; 11,766,735 reads) than in phellem (36.8%; 4,687,046 reads) (Figure 2A).

The other classes of small RNAs such as those derived from ribosomal RNA (rRNA) and transfer RNA (tRNA) identified in the libraries were afterwards discarded leading to a decrease of 2 % of the reads in leaves and 21 % in the phellem (Table 1). The filtering of the rRNA/tRNA had a significant effect on the size class profile (Figure 2B). While the 21-nt class remained similar in both tissues, decreasing approximately 1% in leaves (from 12.8% to 11.3%) and 4% in phellem (from 12.7% to 8.7%), the number of reads in the 24-nt class decreased more than 2-fold in leaves (from 60.0% to 28.0%) and over 3-fold in phellem (from 36.8% to 10.6%). The removal of rRNA/tRNA resulted in almost identical levels of the 21-nt and the 24-nt size classes in the phellem, but in the leaves the 24-nt size class was about twice more expressed. However, when comparing the levels of the non-redundant sequences according to the size (Figure 2C), the 24-nt size class was much more diverse in both tissues, with the maximum number of non-redundant reads being detected in the leaves. The level of non-redundant reads was very low in the other size classes evidencing the fact that the 21-nt class is much less diverse that the 24-nt class in the analysed cork oak tissues.

Conserved miRNAs and expression profile

Although a diverse set of miRNAs has been already identified in several plant species, miRNAs from *Quercus* species have been up to now absent from the miRBase repository. Many miRNAs families are greatly conserved among plant species allowing their identification through BLAST searches using the annotated miRNA deposited on miRBase (Griffiths-Jones 2006). After filtering the reads according to size, abundance and rRNA/tRNA, as described above, the conserved miRNAs in the leaves and phellem of cork oak were identified by BLAST searches of the small RNA reads against miRBase release 18, which includes miRNAs from 52 plant species. The search for the conserved mature miRNAs led to the identification of members of 34 and 35 conserved miRNAs in the cork oak leaves and phellem, respectively, out of a total of 41 miRNA families (Table 2). It was also possible to identify precursor sequences for 22 conserved miRNA from leaves and 21 from phellem (Supplementary material - **SM2_Leaf_miRCat_results.xls** and **SM3_cork-tissues_miRCat_results.xls**).

The most highly expressed families in both tissues were miR167, miR165/166, miR396 and miR159. Although high similarity was observed regarding the conserved families of miRNA present in leaves and phellem, their expression level was, in some cases, quite different. The main difference was found for miR164 which was approximately 109 times more expressed in leaves than in phellem, followed by miR395 (85 x), mir156/157 (67 x) and miR399 (58 x). The miRNAs whose expression was more different between tissues had a maximum expression level below 500 normalized weighted counts. The differences between tissues were smaller for the miRNA families that presented higher expression in phellem, which was the case of miR390 and miR168, and were 49 and 15 times more expressed in phellem than in leaves, respectively.

The analysis of our data revealed that 6 miRNA families were leaf-specific (miR163, miR2089, miR2950, miR477, miR529, and miR394) while seven miRNA families were phellem-specific (miR1140, miR158, miR1863, miR2916, miR479, miR5083 and miR530). It should be noted that the miRNAs with a tissue specific pattern generally showed low expression levels.

Based on recent reports (Devers et al. 2011; Todesco et al. 2012; Zhou et al. 2010), referring the putative importance of the miRNA* form, we have also performed the analysis of these forms in our data. There was a general agreement between the relative abundance of the mature miRNA and its star form but exceptions were found for 4 miRNA* which showed different profiles when compared to the mature miRNA: miR165*/166*, miR172*, miR393* and miR396* (Table 2). In leaves, the miR165/166 mature form was expressed at higher levels than the star form, but in phellem the miR165*/166* was twice the expression level of the mature form; the miR172* was expressed at higher levels than miR172 in both tissues; the miR393 and miR393* had the same expression level in leaves but miR393* was expressed at higher levels in phellem; and the miR396 and miR396* were expressed at similar levels in the phellem but the miR396* was expressed at higher levels in leaves.

Novel miRNAs in *Quercus suber*

In order to perform the prediction of the stem-loop secondary structure characteristic of the pre-miRNA we have taken into account the previously described structure and minimum free-energy criteria (Jones-Rhoades et al. 2006). Since a reference cork oak genome is not yet available we customised an enlarged oak reference database: CEOR DB (see Material and Method for detail) for the prediction of the stem-loop secondary structures using miRCat from UEA package (Moxon et al. 2008b) (Suplementary material **SM2_Leaf_miRCat_results.xls, SM3_cork-tissues_miRCat_results.xls**).

Using this approach a total of 30 novel miRNA candidates were identified (Table 3) of which 11 were unique to leaves and 13 were unique to phellem. The average expression level of the novel miRNAs was approximately 362 reads. However, the tissue-specific novel miRNAs showed lower average expression levels that ranged from 71 to 162 in leaves and phellem, respectively.

An additional criteria proposed by (Meyers et al. 2008) to confirm the annotation of a novel miRNA is the isolation in the same library of the miRNA* form with an extensive base-pairing with the miRNA. Of the 30 novel miRNAs, 8 had the corresponding star form present in the same library. The expression level of the new miRNAs was lower in comparison with the levels observed for the conserved miRNA and the new miRNAs were more abundant in phellem than in leaves. The secondary structure for phellem specific novel miRNA is shown in Table 4.

Validation of identified miRNAs by northern blot analysis

In order to confirm the deep-sequencing data and the differential expression of the miRNA reads, 19 miRNAs were selected for further analysis by northern blot (Figure 3), including 13 conserved miRNAs (miR156/157, miR159, miR164, miR165/166, miR166*, miR168, miR171, miR319, miR390, miR395, miR399, miR482 and miR894) and 4 novel miRNAs (miR-A, miR-B, miR-3P and miR-13P). We designed DNA probes complementary to the miRNA sequence (Supplementary material **SM4_northernblot_probes.pdf**) and detected signal for the miRNAs under validation, even for those with low expression levels. Remarkably, we validated 4 non-conserved miRNAs despite their low expression levels, 2 in both leaf and phellem (miR-A and miR-B) and the other 2 specifically in phellem (miR-3P and miR-13P).

Estimation of the relative expression levels between leaves and phellem by comparing the intensity of hybridization signal obtained by northern blot was consistent with the deep-sequencing data. For instance, the identified novel miRNAs miR-A and the miR-B were less expressed in the leaves while the miR-3P and

the miR-13P were only expressed in the phellem. Considering the expression level observed both by deep-sequencing and northern blot analysis, miR-13P is the best miRNA phellem-specific candidate (Table 3, Figure 3). However, it is likely that the conserved miRNAs miR166*, miR168 and miR390 play important roles in the differentiating phellem as judged by analysis of their expression patterns by northern blot hybridization (Table 2, Figure 3).

Target prediction for the tissue specific novel miRNAs

To better understand the putative functions of the conserved and novel miRNAs in the analysed tissues, their corresponding target genes were predicted using TAPIR, a bioinformatic tool developed specifically for plants (Bonnet et al. 2010). Since complete genome sequence information is not available for oaks, we have constructed a database called CEORT DB (see Material and Methods for details) that we have used to predict the miRNA target genes. Considering that the miRNA and its target gene should coexist in the same tissue for interaction to take place, only the predicted target genes found in the same tissue as the corresponding miRNA were considered as putative targets (**SM5_ non-conserved miRNAs_targets.xls, SM6_target_contigs_leaf.fasta, SM7_target_contigs_cork.fasta)**. Using these criteria, we have found that several miRNA have more than one predicted target gene but some of them had the same annotation (Table 3).

To obtain further insights into the regulatory mechanisms involving miRNA pathways and their relevance concerning tissue-specificity, all the specific targets for each tissue (72 for leaves and 219 for phellem) were annotated using Blast2GO (Conesa and Götz 2008; Conesa et al. 2005; Götz et al. 2011; Götz et al. 2008). The genes with functional annotation (54 for leaves and 156 for phellem) were grouped by gene ontology using Blast2GO. We further assigned the GO terms based on the plant GOslim classification with the combined graph category (Figure 4A and 4B).

Discussion

MiRNAs regulate development, metabolism, adaptation and evolution of plants but, despite their importance, relatively few studies were hitherto performed in woody plants (Sun et al. 2012). The control of secondary growth in trees and the formation of highly specialized types of secondary cell walls in wood and cork cells, likely involve several small RNA-mediated networks which are still largely unknown. In this work we report for the first time the sequencing and analysis of the small RNA transcriptome of a *Quercus* species and, to our knowledge, the first in the Fagaceae. Moreover, a comparative analysis between the small RNA transcriptome of leaves and phellem of cork oak was conducted aiming to identify and characterize the putative involvement of miRNAs in phellem differentiation. Over 25 million small RNA sequences with 19-25 nt length derived from two libraries representing leaf and phellem tissues were analysed.

A higher number of total small RNA sequences was obtained from leaves when compared to phellem. This was not surprising considering that more cell types are present in whole leaves than in the isolated phellem tissues which are expected to contain mostly differentiating phellem cells. Therefore the higher number of small RNAs in leaves probably reflects a wider range of regulatory functions being performed in leaves. The 24-nt length class of small RNAs was the most abundant followed by the 21-nt class in both analysed tissues but the 24-nt/21-nt proportion was higher in leaves than in phellem. It is well established that the relative

abundance of the 24-nt and the 21-nt length classes may vary depending on the species and the tissue type, among other factors. In fact, previous reports refer that the 24-nt sRNAs is the predominant size class, for instance in *Arabidopsis thaliana*, Nic*othiana benthamiana*, *Solano lycopersicum*, *Oriza sativa* and *Olea europea* (Morin et al. 2008; Pantaleo et al. 2010; Donaire et al. 2011; Kasschau et al. 2007), while the 21-nt class is more abundant in species such as *Eschscholzia californica*, *Hordeum vulgare*, *Vitis vinifera*, *Pinus contorta* and *Panax gynseng* (Pantaleo et al. 2010; Morin et al. 2008; Schreiber et al. 2011; Wu et al. ; Barakat et al. 2007). The tissue-dependent variation in the pools of 24 and 21-nt classes has also been observed in other plant species (Pantaleo et al. 2010) and, in a few cases, striking variation has been detected (Slotkin et al. 2009) reflecting the sRNA biogenesis pathways operating in such tissues. The determined level of non-redundancy further showed that in both cork oak tissue types the 21-nt class is less diverse than the 24-nt class, which is in agreement with previous reports in other species (Kasschau et al. 2007; Martínez et al. 2011; Pantaleo et al. 2010; Rajagopalan et al. 2006).

The rRNA/tRNA-derived small RNAs were mainly present in the 24-nt length class and their number was higher in the phellem. Interestingly, recent studies report that these elements may contribute to gene regulation, being cleaved by DICER RNAse to produce stable RNA products rather than being just randomly generated degradation products (Haussecker et al. 2010; Cole et al. 2009; Sobala and Hutvagner 2011). Their functions are still largely unknown, however, a recent report from Li *et al* (2012) based on human and mouse constitutively expressed non-coding RNAs suggests a role for these products in the regulation of unwarranted expression of endogenous viruses through the RNA interference pathway (Li et al. 2012). The presence of tRNA-derived RNA fragments (Lee et al. 2009) has been detected also in plants (Chen et al. 2011; Hackenberg et al. 2012; Hsieh et al. 2009) as well as their putative target genes (Guilherme Loss-Morais et al. 2013).

Following the commonly used guidelines for miRNA annotation (Meyers et al. 2008) we have found 41 families of conserved miRNA and 30 novel miRNA families, some of which were tissue-specific. The prediction of precursor sequences for about 60% of the conserved miRNAs in our analysis was high when compared to previous reports in other species without an available genome sequence such as *Arachis hypogaea* L., for which precursors of nine conserved miRNA families out of 75 were identified (Zhao et al. 2010). These results suggest that the database used in this analysis adequately covers the transcriptome of the studied tissues.

The most abundant miRNAs were those of the families miR167, followed by miR165/166, miR396, and miR159 which are described as the most widespread miRNA families in plants (Barakat et al. 2007; Sun 2012). MiR167 is one of the miRNA families involved in the regulation of auxin signalling pathways. It targets members of the AUXIN RESPONSE FACTOR (ARF) family of transcription factors, namely ARF6/8, involved in early auxin response through activation of auxin responsive genes (Wu et al. 2006), as well as IAA-Ala Resistant3 (IAR3) which hydrolyzes an inactive form of auxin (indole-3-acetic acid [IAA]-alanine) and releases bioactive auxin (IAA) (Kinoshitaa et al. 2012). The regulation of auxin signalling pathways by miR167 and its targets has been related to male and female flower development (Wu et al. 2006) and

osmotic stress–induced root architecture changes (Kinoshitaa et al. 2012). In contrast, miR160 which also targets members of the ARF family, showed a very low expression level, especially in the phellem. The targets of miR160 are the ARF10/16/17 involved in auxin-mediated developmental processes including root growth, blade outgrowth and fruit development (Hendelman et al. 2012; Khan et al. 2011; Xing et al. 2011). We also have found another miRNA mainly present in phellem, miR390, which directs cleavage of tasiRNA locus 3 (TAS3) leading to the production of ta-siRNAs that target ARF2/3/4 (Allen et al. 2005; Marin et al. 2010) involved in lateral root initiation (Marin et al. 2010). Although the regulatory function of these miRNAs in cork oak phellem needs to be studied in detail, these results suggest that auxin signaling-mediated events play a relevant role in phellem differentiation.

When comparing the level of expression of conserved miRNAs in the leaves and phellem we observed that the families of miRNAs expressed at higher levels in leaves are miR156, miR160, miR164, miR395 and miR399, while miR166*, miR168, miR171and miR390, are expressed at higher levels in phellem. These miRNAs had already been described in other woody plants including *Populus thricocarpa* (Tuskan et al. 2006) and *Vitis vinifera* (Mica et al. 2009) and their targets are known (Khraiwesh et al. 2012). Based on the description of the role of conserved plant miRNAs by Khraiwesh et al (2012), the miR156, miR160, miR395 and miR399 are controlling *adaptive responses to stress* in the leaves, whereas *adaptive responses to stress* (miR168, miR171 and miR319) and *leaf polarity* (miR166, miR168 and miR390) are the two dominant roles identified in the phellem. We have also identified some miRNAs partially conserved across vascular plants such as miR403 only present in Eudicots clade and miR482 from Core rosids clade (Cuperus et al. 2011). From an evolutionary point of view the ancient miRNA have roles in growth, development and differentiation while young miRNA have a more specific set of functions and are usually less abundant (Sun et al. 2012). Interestingly, we have found one "young" miRNA, miR482, more expressed in phellem which was found to be also involved in poplar xylem differentiation (Lu et al. 2005).
It has been generally considered that the miRNA* is expressed in lower abundance (< 10 %) than the mature form of a miRNA (Rajagopalan et al. 2006) but the idea that the miRNA* is an inactive form with no functional relevance has been questioned (Devers et al. 2011). For instance, in human plasmacytoid dendritic cells, there is evidence for a cooperative action of the star-form and the mature form of a miRNA (Zhou et al. 2010). Interestingly, in this study we have found that the miR166 was expressed at higher levels in the leaves compared to its star form, but inversely, miR166* was expressed at higher levels in phellem. The predicted target for the miR166* was a putative beta-1,6-N-acetylglucosaminyltransferase coding transcript. These enzymes are known to catalyze attachment of oligosaccharide side chains to glycoproteins (Siddiqui et al. 2005). Interestingly, a recent report by Zalepa-King and Citovsky (2013) (Zalepa-King and Citovsky 2013) describe the location of one of these enzymes in plasmodesmata suggesting that it plays a role in these structures. Also, it has been shown that the number of plasmodesmata might significantly increase in the cambial zone of tomato stems (Ehlers and van Bel 2010). Establishing a parallel between the differentiation process of xylem cells from vascular cambium and phellem cell differentiation from cork cambium derivatives, it is expected that rapid cell expansion occurs before active cell wall deposition in the phellem cells. It is therefore tempting to speculate that the number of plasmodesmata might also significantly increase in differentiating phellem cells and that the regulation of this process may require the function of specific miRNA molecules.

A difficult challenge in this type of studies is the identification of species-specific miRNAs. The approach usually consists in the identification of the secondary stem-loop structure characteristic of the precursor of miRNAs (pre-miRNA) by computational prediction. This prediction is best performed using the full genome information, but when this information is lacking the ESTs can also be used to perform this analysis. We have been able to predict 30 novel miRNA, some of them tissue specific. The level of expression of these new miRNAs was lower in comparison with the levels observed for the conserved miRNA which is consistent with previous studies reporting that non-conserved miRNAs were expressed at lower levels and were usually characteristic of a tissue or developmental stage (Fahlgren et al. 2007; Moxon et al. 2008a; Rajagopalan et al. 2006; Yao et al. 2007). Despite their lower expression, 4 new miRNAs were validated by northern blotting, two of them phellem-specific (miR-3P and miR-13P). The difficulties in finding the miRNA* for many of the novel miRNA candidates can be explained by the low expression levels making it less likely to identify the star form, usually less expressed than the corresponding miRNA. Since the miRNA and its true target should coexist in the cell, we excluded the predicted targets that were not present in the same tissue as the corresponding novel miRNA in order to increase the confidence in the prediction. In the phellem, miRNA target genes related to *response to stimulus*, *biological regulation* and *signalling* were predominant when comparing to leaves, while the Biological Process *growth* was only present in phellem. Moreover the Molecular Function categories *receptor activity* and *molecular transducer activity* were only found for target genes present in the phellem. We have predicted two targets related to chromatin modifications, a histone deacetylase and a histone acetyltransferase, for the newly identified phellem specific miRNA Qsu-miR-2P and Qsu-miR-7P, respectively. These two enzymes have opposite roles in the acetylation of the histones, where acetylation is usually associated to activation of transcription (Shahbazian and Grunstein 2007). According to the expression levels of the miRNAs targeting these genes, it could be hypothesized that histone deacetylases are being repressed resulting in transcription activation. However, complex interactions exist between different chromatin modifications to regulate gene transcription and further studies would be necessary to clarify the roles of these miRNAs. It was shown by Conde et al (2013) that winter dormancy in poplar stems is epigenetically controlled depending on the level of histone acetylation (Conde et al. 2013). Moreover, genes related with histone acetylation such as histone acetyltransferases and histone deacetylases are differentially expressed during regeneration of *Populus tomentosa* secondary vascular tissue (Zhang et al. 2011). Since phellem formation was highly increased at the time of sample collection, the regulation of the gene transcriptional activity should be crucial to control tissue differentiation, and histone acetylation may play an important role during this process. When comparing the putative proteins encoded by the predicted target genes of phellem-specific miRNAs with the proteins from the same tissues identified in a previous study (Ricardo et al. 2011), we found some overlapping, notably in the case of enolases and laccases. When considering the Biological Process, overlap was found for heat shock proteins, ATPases and proteasome complex proteins.

Overall, these results suggest that tightly controlled regulatory mechanisms are operating in phellem and may be relevant for the differentiation process. The identification of conserved and novel miRNAs of cork oak as well as of their target genes performed in this work complements the recent effort to sequence the coding transcriptome (Pereira-Leal et al. submitted) and provides new tools to dissect relevant regulatory mechanisms operating in cork oak phellem.

ACKNOWLEDGEMENTS

The authors thank Sónia Gonçalves from CEBAL (Beja, Portugal) for providing the tissue samples. This research was supported by Fundação para a Ciência e Tecnologia (FCT), through projects PEst-OE/EQB/LA0004/2011 and PTDC/AGR-GPL/098369/2008, and grant SFRH/BPD/20833/2004 (to I. Chaves).

Data archiving statement

The smallRNA sequences obtained for leaf and cork are deposited in the NCBI GenBank database under the accession numbers SRR 988108 and SRR988109 respectively.

Supplementary material

SM1_Electropherogram of RNA fractions.pdf: Electropherogram of HMW RNA and LMW RNA enriched fractions from *Q. suber* cork tissues.

SM2_Leaf_miRCat_results.xls: miRCat results for cork oak leaf tissues

SM3_cork-tissues_miRCat_results.xls: miRCat results for cork oak phellem tissues

SM4_northernblot_probes.pdf: Northern blot DNA probes

SM5_non-conserved miRNAs_targets.xls: Target genes for non-conserved miRNAs

SM6_target_contigs_leaf.fasta: Target gene contigs in leaf tissues

SM7_target_contigs_cork.fasta: Target gene contigs in cork tissues

References

Allen E, Xie Z, Gustafson AM, Carrington JC (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell 121 (2):207 - 221

Barakat A, Wall K, Leebens-Mack J, Wang YJ, Carlson JE, DePamphilis CW (2007) Large-scale identification of microRNAs from a basal eudicot (*Eschscholzia californica*) and conservation in flowering plants. Plant J 51 (6):991-1003

Barel G, Ginzberg I (2008) Potato skin proteome is enriched with plant defence components. J Exp Bot 59 (12):3347-3357

Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116 (2):281-297

Bonnet E, He Y, Billiau K, Van de Peer Y (2010) TAPIR, a web server for the prediction of plant microRNA targets, including target mimics. Bioinformatics 26 (12):1566-1568

Bugalho MN, Caldeira MC, Pereira JS, Aronson J, Pausas JG (2011) Mediterranean cork oak savannas require human use to sustain biodiversity and ecosystem services. Frontiers in Ecology and the Environment 9:278-286

Chang S, Puryear J, Cairney J (1993) A simple and efficient method for isolating RNA from pine trees. Plant Molecular Biology 11:114-117

Chaves I, Pinheiro C, Paiva JA, Planchon S, Sergeant K, Renaut J, Graca JA, Costa G, Coelho AV, Ricardo CP (2009) Proteomic evaluation of wound-healing processes in potato (Solanum tuberosum L.) tuber tissue. Proteomics 9 (17):4154-4175

Chen CJ, Liu Q, Zhang YC, Qu LH, Chen YQ, Gautheret D (2011) Genome-wide discovery and analysis of microRNAs and other small RNAs from rice embryogeniccallus. RNA Biol 8:538-547

Cole C, Sobala A, Lu C, Thatcher SR, Bowman A, Brown JW, Green PJ, Barton GJ, Hutvagner G (2009) Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. RNA 15 (12):2147-2160

Conde D, González-Melendi P, Allona I (2013) Poplar stems show opposite epigenetic patterns during winter dormancy and vegetative growth. Trees 27 (1):311-320. doi:10.1007/s00468-012-0800-x

Conesa A, Götz S (2008) Blast2GO: A Comprehensive Suite for Functional Analysis in Plant Genomics. International Journal of Plant Genomics 2008:1-13

Conesa A, Götz S, Garcia-Gomez JM, Terol J, Talon M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674-3676

Cuperus JT, Fahlgren N, Carrington JC (2011) Evolution and functional diversification of miRNA genes. The Plant Cell 23 (2):431 - 442

Devers EA, Branscheid A, May P, Krajinski F (2011) Stars and symbiosis: microRNA- and microRNA*-mediated transcript cleavage involved in arbuscular mycorrhizal symbiosis. Plant Physiol 156 (4):1990-2010
Dharmawardhana P, Brunner AM, Strauss SH (2010) Genome-wide transcriptome analysis of the transition from primary to secondary stem development in Populus trichocarpa. BMC Genomics 11:150
Donaire L, Pedrola L, de la Rosa Rl, Llave Cs (2011) High-Throughput Sequencing of RNA Silencing-Associated Small RNAs in Olive (*Olea europaea* L.). PLoS ONE 6 (11):e27916
Ehlers K, van Bel AJ (2010) Dynamics of plasmodesmal connectivity in successive interfaces of the cambial zone. Planta 231 (2):371-385. doi:10.1007/s00425-009-1046-8.
Fahlgren N, Howell MD, Kasschau KD, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Law TF, Grant SR, Dangl JL, Carrington JC (2007) High-throughput sequencing of arabidopsis microRNAs: evidence for frequent birth and death of miRNA genes. PLoS One 2:e219
Ginzberg I, Barel G, Ophir R, Tzin E, Tanami Z, Muddarangappa T, de Jong W, Fogelman E (2009) Transcriptomic profiling of heat-stress response in potato periderm. J Exp Bot 60 (15):4411-4421
Götz S, Arnold R, Sebastián-León P, Martín-Rodríguez S, Tischler P, Jehl M-A, Dopazo J, Rattei T, Conesa A (2011) B2G-FAR, a species centered GO annotation repository. Bioinformatics 27 (7):919-924
Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A (2008) High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Research 36:3420-3435
Griffiths-Jones S (2006) miRBase: the microRNA sequence database. Methods Mol Biol 342:129-138
Guilherme Loss-Morais, Waterhouse PM, Margis R (2013) Description of plant tRNA-derived RNA fragments (tRFs) associated with argonaute and identification of their putative targets. Biology Direct 8:6
Hackenberg M, Huang PJ, Huang CY, Shi BJ, Gustafson P, Langridge P (2012) A Comprehensive Expression Profile of MicroRNAs and Other Classes of Non-Coding Small RNAs in Barley Under Phosphorous-Deficient and -Sufficient Conditions. DNA Research 19:1-17
Haussecker D, Huang Y, Lau A, Parameswaran P, Fire AZ, Kay MA (2010) Human tRNA-derived small RNAs in the global regulation of RNA silencing. RNA 16:673-695
Hendelman A, Buxdorf K, Stav R, Kravchik M, Arazi T (2012) Inhibition of lamina outgrowth following Solanum lycopersicum AUXIN RESPONSE FACTOR 10 (SlARF10) derepression. Plant Molecular Biology 78 (6):561-576
Hsieh LC, Lin SI, Shih ACC, Chen JW, Lin WY, Tseng CY, Li WH, Chiou TJ (2009) Uncovering small RNA-mediated responses to phosphate deficiency in Arabidopsis by deep sequencing. Plant Physiology 151:2120 - 2132
Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAS and their regulatory roles in plants. Annu Rev Plant Biol 57:19-53
Kasschau KD, Fahlgren N, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Carrington JC (2007) Genome-wide profiling and analysis of Arabidopsis siRNAs. PLoS Biol 5 (3):e57
Khan GA, Declerck M, Sorin C, Hartmann C, Crespi M, Lelandais-Briere C (2011) MicroRNAs as regulators of root development and architecture. Plant Molecular Biology 77 (1-2):47-58. doi:10.1007/s11103-011-9793-x
Khraiwesh B, Zhu JK, Zhu J (2012) Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. Biochim Biophys Acta 1819 (2):137-148
Kinoshitaa N, Wanga H, Kasaharac H, Liua J, MacPhersona C, Machidab Y, Kamiyac Y, Hannahe MA, Chuaa N-H (2012) IAA-Ala Resistant3, an Evolutionarily Conserved Target of miR167, Mediates Arabidopsis Root Architecture Changes during High Osmotic Stress. The Plant Cell 24 (9):3590-3602
Lee YS, Shibata Y, Malhotra A, Dutta A (2009) A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). Genes Dev 23:2639-2049
Li Z, Ender C, Meister G, Moore PS, Chang Y, John B (2012) Extensive terminal and asymmetric processing of small RNAs from rRNAs, snoRNAs, snRNAs, and tRNAs. Nucleic Acids Res 40 (14):6787-6799
Liu J, Vance CP (2010) Crucial roles of sucrose and miRNA399 in systemic signaling of P deficiency - A tale of two team players? Plant Signaling and Behaviour 5 (12):1-5
Lu S, Sun Y-H, Shi R, Clark C, Li L, Chiang VL (2005) Novel and Mechanical Stress–Responsive MicroRNAs in *Populus trichocarpa* That Are Absent from Arabidopsis. Plant Cell 17:2186–2203
Marin E, Jouannet V, Herz A, Lokerse AS, Weijers D, Vaucheret H, Nussaume L, Crespi MD, Maizel A (2010) miR390, Arabidopsis TAS3 tasiRNAs, and their AUXIN RESPONSE FACTOR targets define an autoregulatory network quantitatively regulating lateral root growth. The Plant Cell 22 (4):1104-1117
Martínez G, Forment J, Llave C, Pallás V, Gómez G (2011) High-Throughput Sequencing, Characterization and Detection of New and Conserved Cucumber miRNAs. PLoS ONE 6 (5):e19523
Meyers BC, Axtell MJ, Bartel B, Bartel DP, Baulcombe D, Bowman JL, Cao X, Carrington JC, Chen X, Green PJ, Griffiths-Jones S, Jacobsen SE, Mallory AC, Martienssen RA, Poethig RS, Qi Y, Vaucheret H,

Voinnet O, Watanabe Y, Weigel D, Zhu JK (2008) Criteria for annotation of plant MicroRNAs. The Plant Cell 20 (12):3186-3190
Mica E, Piccolo V, Delledonne M, Ferrarini A, Pezzotti M, Casati C, Del Fabbro C, Valle G, Policriti A, Morgante M, Pesole G, Pe ME, Horner D (2009) High throughput approaches reveal splicing of primary microRNA transcripts and tissue specific expression of mature microRNAs in Vitis vinifera. BMC Genomics 10 (1):558
Morin RD, Aksay G, Dolgosheina E, Ebhardt HA, Magrini V, Mardis ER, Sahinalp SC, Unrau PJ (2008) Comparative analysis of the small RNA transcriptomes of Pinus contorta and Oryza sativa. Genome Res 18 (4):571-584
Moxon S, Jing R, Szittya G, Schwach F, Rusholme Pilcher RL, Moulton V, Dalmay T (2008a) Deep sequencing of tomato short RNAs identifies microRNAs targeting genes involved in fruit ripening. Genome Res 18 (10):1602-1609
Moxon S, Schwach F, Dalmay T, MacLean D, Studholme DJ, Moulton V (2008b) A toolkit for analysing large-scale plant small RNA datasets. Bioinformatics 24 (19):2252-2253. doi:10.1093/bioinformatics/btn428
Ong SS, Wickneswari R (2012) Characterization of microRNAs Expressed during Secondary Wall Biosynthesis in *Acacia mangium*. PLoS One 7 (11):e49662
Osorio S, Alba R, Nikoloski Z, Kochevenko A, Fernie AR, Giovannoni JJ (2012) Integrative Comparative Analyses of Transcript and Metabolite Profiles from Pepper and Tomato Ripening and Development Stages Uncovers Species-Specific Patterns of Network Regulatory Behavior. Plant Physiol 159 (4):1713-1729
Pantaleo V, Szittya G, Moxon S, Miozzi L, Moulton V, Dalmay T, Burgyan J (2010) Identification of grapevine microRNAs and their targets using high-throughput sequencing and degradome analysis. Plant J 62 (6):960-976
Pausas JG (1997) Resprouting of Quercus suber in NE Spain after fire. J Veg Sci 8:703-706
Pereira H (1988) Chemical-Composition and Variability of Cork from Quercus-Suber L. Wood Science and Technology 22 (3):211-218
Pereira H (2007) Cork: biology, production and uses. Elsevier, Amsterdam, Netherlands
Puzey JR, Karger A, Axtell M, Kramer EM (2012) Deep Annotation of Populus trichocarpa microRNAs from Diverse Tissue Sets. PLoS One 7 (3):e33034
Rajagopalan R, Vaucheret H, Trejo J, Bartel DP (2006) A diverse and evolutionarily fluid set of microRNAs in Arabidopsis thaliana. Genes Dev 20 (24):3407-3425
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
Robischon M, Du J, Miura E, Groover A (2011) The Populus Class III HD ZIP, popREVOLUTA, Influences Cambium Initiation and Patterning of Woody Stems. Plant Physiology 155 (3):1214-1225
Schreiber AW, Shi BJ, Huang CY, Langridge P, Baumann U (2011) Discovery of barley miRNAs through deep sequencing of short reads. BMC Genomics 12:129
Serra O, Figueras M, Franke R, Prat S, Molinas M (2010a) Unraveling ferulate role in suberin and periderm biology by reverse genetics. Plant Signal Behav 5 (8):953-958
Serra O, Hohn C, Franke R, Prat S, Molinas M, Figueras M (2010b) A feruloyl transferase involved in the biosynthesis of suberin and suberin-associated wax is required for maturation and sealing properties of potato periderm. Plant J 62 (2):277-290
Serra O, Soler M, Hohn C, Franke R, Schreiber L, Prat S, Molinas M, Figueras M (2009a) Silencing of StKCS6 in potato periderm leads to reduced chain lengths of suberin and wax compounds and increased peridermal transpiration. J Exp Bot 60 (2):697-707
Serra O, Soler M, Hohn C, Sauveplane V, Pinot F, Franke R, Schreiber L, Prat S, Molinas M, Figueras M (2009b) CYP86A33-targeted gene silencing in potato tuber alters suberin composition, distorts suberin lamellae, and impairs the periderm's water barrier function. Plant Physiol 149 (2):1050-1060
Shahbazian MD, Grunstein M (2007) Functions of site-specific histone acetylation and deacetylation. Annual Review of Biochemistry 76:75-100. doi:10.1146/annurev.biochem.76.052705.162114
Siddiqui S, Pawelek J, Handerson T, Lin C, Dickson R, al. e (2005) Coexpression of beta-1,6-N-acetylglucosaminyltransferase V glycoprotein substrates defines aggressive breast cancers with poor outcome. Cancer Epidemiol Biomarkers Prev 14:2517-2523. doi:doi: 10.1158/1055-9965.EPI-05-0464.
Slotkin RK, Vaughn M, Borges F, Tanurdzic M, Becker JD, Feijo JA, Martienssen RA (2009) Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell 136 (3):461-472
Sobala A, Hutvagner G (2011) Transfer RNA-derived fragments: origins, processing, and functions. Wiley Interdiscip Rev RNA 2 (6):853-862
Soler M, Serra O, Fluch S, Molinas M, Figueras M (2011) A potato skin SSH library yields new candidate genes for suberin biosynthesis and periderm formation. Planta 233 (5):933-945

Soler M, Serra O, Molinas M, Huguet G, Fluch S, Figueras M (2007) A genomic approach to suberin biosynthesis and cork differentiation. Plant Physiol 144 (1):419-431
Sun G (2012) MicroRNAs and their diverse functions in plants. Plant Molecular Biology 80 (1):17-36
Sun YH, Shi R, Zhang XH, Chiang VL, Sederoff RR (2012) MicroRNAs in trees. Plant Molecular Biology 80 (1):37-53
Sunkar R (2010) MicroRNAs with macro-effects on plant stress responses. Seminars in Cell & Developmental Biology 21 (8):805-811
Sunkar R, Chinnusamy V, Zhu J, Zhu J-K (2007) Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. Trends in Plant Science 12 (7):301-309
Todesco M, Balasubramanian S, Cao J, Ott F, Sureshkumar S, Schneeberger K, Meyer RC, Altmann T, Weigel D (2012) Natural variation in biogenesis efficiency of individual Arabidopsis thaliana microRNAs. Curr Biol 22 (2):166-170
Tuskan G, Difazio S, S J, J B, I G, Hellsten U PN, Ralph S, Rombauts S, Salamov A, Schein J, Sterck L, Aerts A, Bhalerao RR, Bhalerao RP, Blaudez D, Boerjan W, Brun A, Brunner A, Busov V, Campbell M, Carlson J, Chalot M, Chapman J, Chen GL, Cooper D, Coutinho PM, Couturier J, Covert S, Cronk Q, Cunningham R, Davis J, Degroeve S, Déjardin A, Depamphilis C, Detter J, Dirks B, Dubchak I, Duplessis S, Ehlting J, Ellis B, Gendler K, Goodstein D, Gribskov M, Grimwood J, Groover A, Gunter L, Hamberger B, Heinze B, Helariutta Y, Henrissat B, Holligan D, Holt R, Huang W, Islam-Faridi N, Jones S, Jones-Rhoades M, Jorgensen R, Joshi C, Kangasjärvi J, Karlsson J, Kelleher C, Kirkpatrick R, Kirst M, Kohler A, Kalluri U, Larimer F, Leebens-Mack J, Leplé JC, Locascio P, Lou Y, Lucas S, Martin F, Montanini B, Napoli C, Nelson DR, Nelson C, Nieminen K, Nilsson O, Pereda V, Peter G, Philippe R, Pilate G, Poliakov A, Razumovskaya J, Richardson P, Rinaldi C, Ritland K, Rouzé P, Ryaboy D, Schmutz J, Schrader J, Segerman B, Shin H, Siddiqui A, Sterky F, Terry A, Tsai CJ, Uberbacher E, Unneberg P, Vahala J, Wall K, Wessler S, Yang G, Yin T, Douglas C, Marra M, Sandberg G, Van de Peer Y, Rokhsar D. (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313 (5793):1596-1604
Varela MC (1995) Consevation of genetic resources of Quercus suber in Portugal. European Forest Resources Programme (EUFORGEN)
Wu B, Wang M, Ma Y, Yuan L, Lu S (2012) High-Throughput Sequencing and Characterization of the Small RNA Transcriptome Reveal Features of Novel and Conserved MicroRNAs in *Panax ginseng*. PLoS ONE 7 (9):e44385
Wu MF, Tian Q, Reed JW (2006) Arabidopsis microRNA167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction. Development 133 (21):4211-4218
Xing H, Pudake R, Guo G, Xing G, Hu Z, Zhang Y, Sun Q, Ni Z (2011) Genome-wide identification and expression profiling of auxin response factor (ARF) gene family in maize. BMC Genomics 12 (1):178
Xu X, Pan S, Cheng S, Zhang B, Mu D, Ni P, Zhang G, Yang S, Li R, Wang J, Orjeda G, Guzman F, Torres M, Lozano R, Ponce O, Martinez D, De la Cruz G, Chakrabarti SK, Patil VU, Skryabin KG, Kuznetsov BB, Ravin NV, Kolganova TV, Beletsky AV, Mardanov AV, Di Genova A, Bolser DM, Martin DM, Li G, Yang Y, Kuang H, Hu Q, Xiong X, Bishop GJ, Sagredo B, Mejia N, Zagorski W, Gromadka R, Gawor J, Szczesny P, Huang S, Zhang Z, Liang C, He J, Li Y, He Y, Xu J, Zhang Y, Xie B, Du Y, Qu D, Bonierbale M, Ghislain M, Herrera Mdel R, Giuliano G, Pietrella M, Perrotta G, Facella P, O'Brien K, Feingold SE, Barreiro LE, Massa GA, Diambra L, Whitty BR, Vaillancourt B, Lin H, Massa AN, Geoffroy M, Lundback S, DellaPenna D, Buell CR, Sharma SK, Marshall DF, Waugh R, Bryan GJ, Destefanis M, Nagy I, Milbourne D, Thomson SJ, Fiers M, Jacobs JM, Nielsen KL, Sonderkaer M, Iovene M, Torres GA, Jiang J, Veilleux RE, Bachem CW, de Boer J, Borm T, Kloosterman B, van Eck H, Datema E, Hekkert B, Goverse A, van Ham RC, Visser RG (2011) Genome sequence and analysis of the tuber crop potato. Nature 475 (7355):189-195
Yao Y, Guo G, Ni Z, Sunkar R, Du J, Zhu JK, Sun Q (2007) Cloning and characterization of microRNAs from wheat (Triticum aestivum L.). The Plant Cell 8 (6):R96
Zalepa-King L, Citovsky V (2013) A Plasmodesmal Glycosyltransferase-Like Protein. PLoS ONE 8 (2):e58025. doi:doi:10.1371/journal.pone.0058025
Zhang J, Gao G, Chen JJ, Taylor G, Cui KM, He XQ (2011) Molecular features of secondary vascular tissue regeneration after bark girdling in Populus. The New phytologist 192 (4):869-884. doi:10.1111/j.1469-8137.2011.03855.x
Zhao C-Z, Xia H, Frazier T, Yao Y-Y, Bi Y-P, Li A-Q, Li M-J, Li C-S, Zhang B-H, Wang X-J (2010) Deep sequencing identifies novel and conserved microRNAs in peanuts (Arachis hypogaea L.). BMC Plant Biol 10 (1):3
Zhou H, Huang X, Cui H, Luo X, Tang Y, Chen S, Wu L, Shen N (2010) miR-155 and its star-form partner miR-155* cooperatively regulate type I interferon production by human plasmacytoid dendritic cells. Blood 116 (26):5885-5894
Zhu QH, Spriggs A, Matthew L, Fan L, Kennedy G, Gubler F, Helliwell C (2008) A diverse set of microRNAs and microRNA-like small RNAs in developing rice grains. Genome Res 18 (9):1456-1465

LEGENDS

Figure 1. Workflow of the bioinformatics analysis of small RNAs from cork oak leaves and phellem tissues.

Figure 2. Size distribution of small RNA sequences obtained from cork oak leaves and phellem: (A) redundant reads; (B) after removal of rRNA/tRNA exact matches; (C) non-redundant reads. Bars representing numbers of reads are in grey (leaves) and black (phellem).

Figure 3. Expression analysis by northern blot hybridization of selected conserved miRNAs in phellem and leaves of cork oak. The numbers below blots represent relative abundance of each miRNA normalized to U6 expression.

Figure 4. Gene Ontology of the predicted target genes for the novel miRNAs considering the number of genes obtained for Biological Process (A) and Molecular Function (B) and their relative value for the Biological Process (C) and Molecular function (D).

Table 1. General statistics for cork oak leaf and phellem small RNA sequence reads processing.

Table 2. Conserved miRNA in cork oak.

Table 3. Novel miRNA in cork oak and the predicted target genes

Table 4. Predicted pre-miRNA secondary structure for phellem-specific novel miRNA in cork oak**.** The hairpin structures were plotted based on miRCat prediction using the RNA fold software. The miRNA is highlighted in blue and the miRNA* in red. The adjusted Minimum Free Energy (MFE) and the number of mismatches are included.

Figure 1.

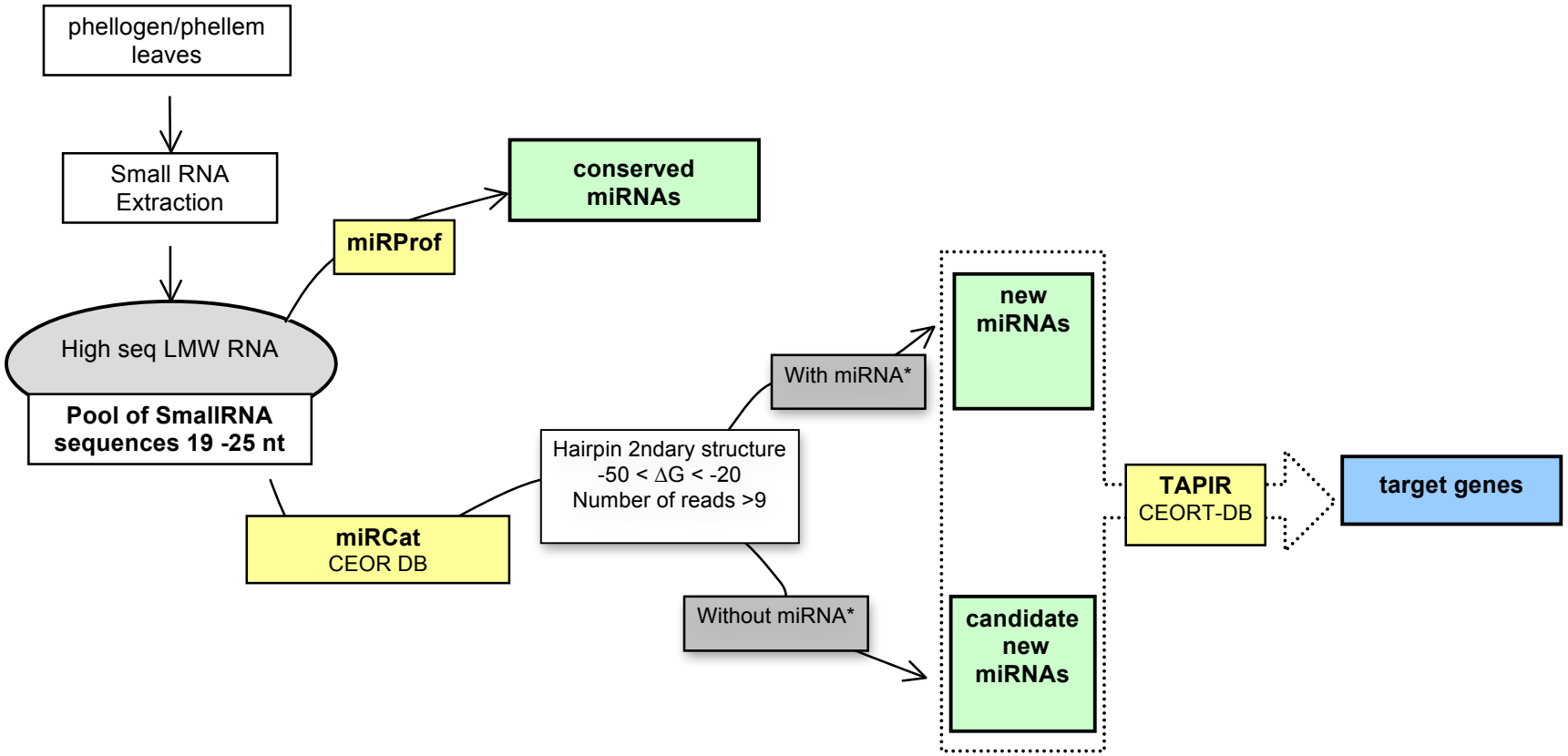

phellogen/phellem leaves
Small RNA Extraction
High seq LMW RNA
Pool of SmallRNA sequences 19 -25 nt
miRProf
conserved miRNAs
miRCat
CEOR DB
Hairpin 2ndary structure
-50 < ΔG < -20
Number of reads >9
With miRNA*
new miRNAs
Without miRNA*
candidate new miRNAs
TAPIR
CEORT-DB
target genes

Figure 2.

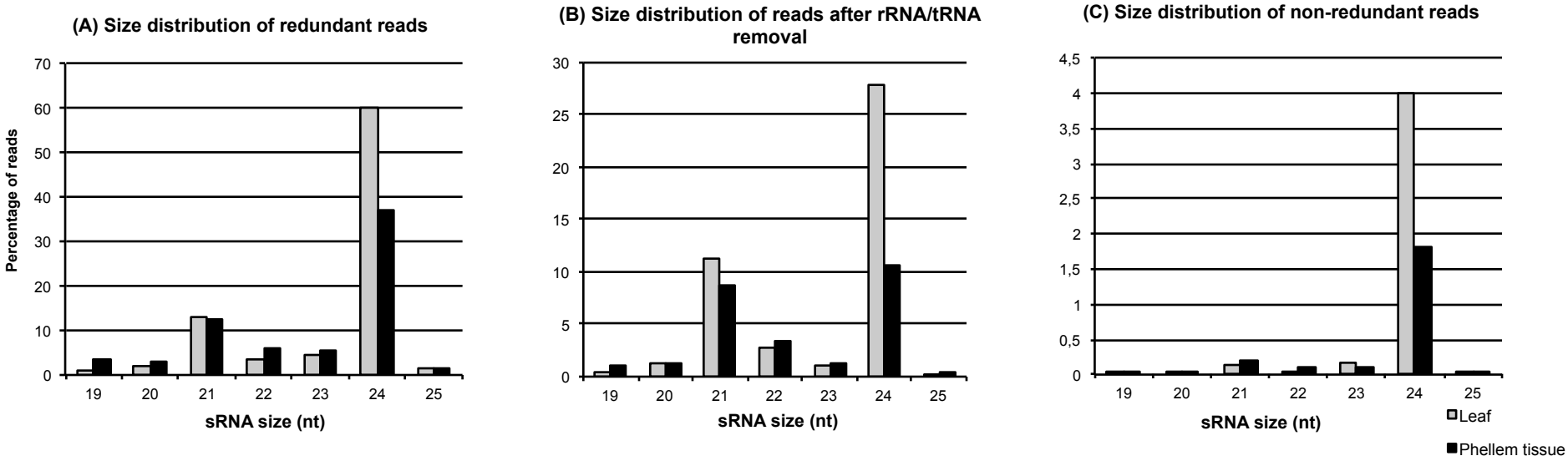

(A) Size distribution of redundant reads
Percentage of reads
70
60
50
40
30
20
10
0
19
20
21
22
23
24
25
sRNA size (nt)
(B) Size distribution of reads after rRNA/tRNA removal
30
25
20
15
10
5
0
19
20
21
22
23
24
25
sRNA size (nt)
(C) Size distribution of non-redundant reads
4,5
4
3,5
3
2,5
2
1,5
1
0,5
0
19
20
21
22
23
24
25
sRNA size (nt)
Leaf
Phellem tissue

Figure 3. Expression analysis by northern blot hybridization of selected conserved miRNAs in phellem and leaves of cork oak. The numbers below blots represent relative abundance of each miRNA normalized to U6 expression.

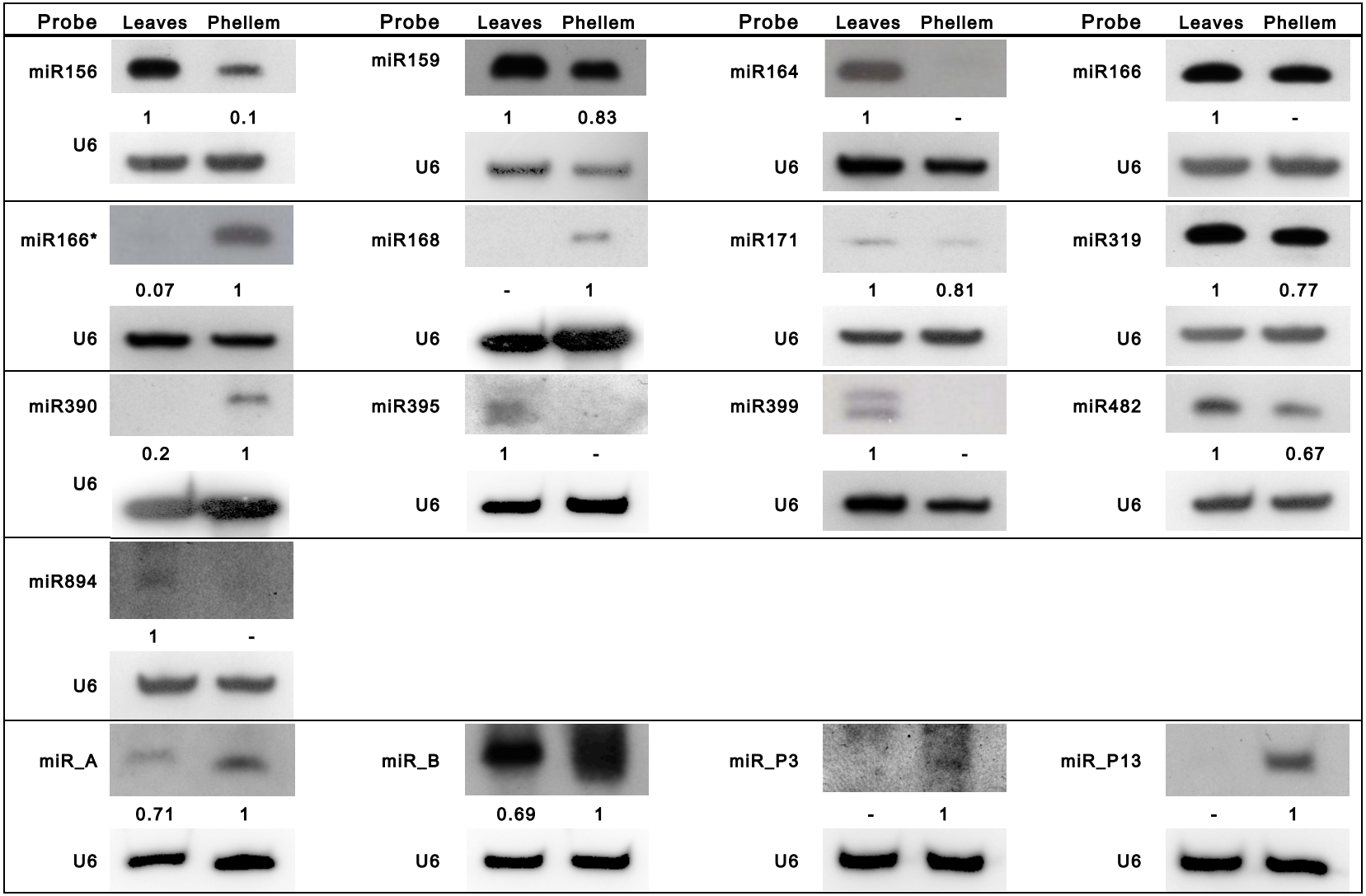


Table 1. General statistics for cork oak leaf and phellem small RNA sequence reads processing.

Sample type	***Leaves***	***Phellem***
Raw reads	19620275	12731020
Reads with 19 < sequence length < 25 nt	16684977	8771865
Reads with abundance >= 3	8988329	4352306
After filtering rRNA/tRNA	8820356	3426924
Non-redundant reads	869583 (4.4%)	298116 (2.3%)

Table 2. Conserved miRNA in cork oak.

miR FAMILY	normalised weighted count				Fold change	
	Leaf		phellem		leaf/phellem	
	miRNA	miRNA*	miRNA	miRNA*	miRNA	miRNA*
miR1140	0	0	2.24	0	-	-
miR156/157	2382.54	238.06	29.85	9.21	79.8	25.8
miR158	0	0	3.55	0	-	-
miR159	8637.86	11.76	7200.74	48.53	1.2	-4.1
miR160	30.84	1.28	3.55	2.89	8.7	-
miR162	389.11	0.97	394.83	3.03	-1.0	-
miR163	0.55	0	0	0	-	-
miR164	353.71	4.14	3.29	0.26	107.5	-
miR165/166	10749.60	683.34	6664.95	14900.00	1.6	-21.8
miR167	15936.72	4.63	20940.64	9.73	-1.3	-2.1
miR168	553.75	10.48	8068.14	196.36	-14.6	-18.7
miR169	73.85	6.64	30.78	2.63	2.4	2.5
miR171	51.00	15.47	379.70	29.41	-7.4	-
miR172	14.44	74.40	18.41	152.57	-1.3	-2.0
miR1863	0	0	0.66	0	-	-
miR2089	0.43	0	0	0	-	-
miR2911	0.85	0	3.81	0	-4.5	-
miR2916	0	0	1.71	0	-	-
miR2950	1.71	0	0	0	-	-
miR319	168.11	1.16	302.63	15.13	-1.8	-13.0
miR3627	0.24	0	0.53	0	-2.2	-
miR390	10.78	0.73	531.48	26.96	-49.3	
miR393	24.01	24.68	4.73	38.01	5.1	-1.5
miR394	6.7	0	0	0	-	-
miR395	488.36	2.86	5.79	0	84.3	-
miR396	7720.49	12600.00	5470.56	4502.55	1.4	2.8
miR397	121.01	0	419.43	0	-3.5	-
miR398	767.19	0.61	463.75	1.18	1.6	-
miR399	434.56	0	7.37	0	59.0	-
miR403	93.71	0	30.12	2.24	3.1	-
miR408	26.02	21.51	42.61	48.14	-1.6	-
miR472	1.46	0	6.84	0	-4.7	-
miR477	0.67	0	0	0	-	-
miR479	0	0	2.76	0	-	-
miR482	1067.52	0.91	2425.67	7.23	-2.3	-
miR894	1.65	0	24.85	0	-15.1	-
miR5072	14.01	0	156.78	0	-11.2	-
miR5083	0	0	1.58	0	-	-
miR5139	1.04	0	6.58	0	-6.3	-
miR529	0.43	0	0	0	-	-
miR530	0	0	0.79	0	-	-

Normalised count: count of matching sequence reads normalised to the total number of reads after last filtering step search miRBase database.

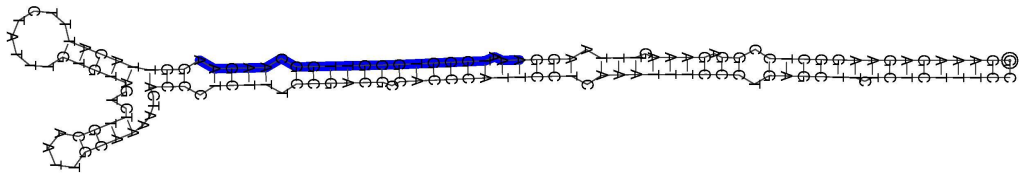
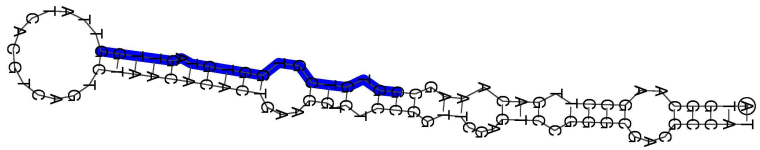
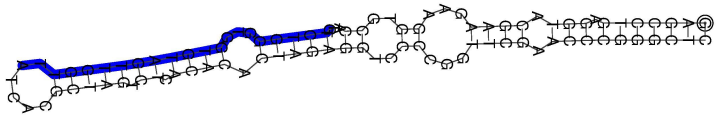
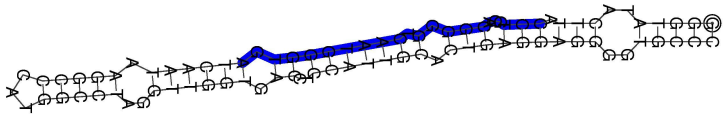
Table 3. Novel miRNA in cork oak and the predicted target genes

miRNA Sequence	length	miRNA*	normalized weight count		Predicted Target proteins
			leaf	phellem	
Qsu-miR-A AGTGGAAGGATTGGAAAGACA	21	Yes	1951.26	4634.11	- rcc1 domain-containing protein - (TF) tubby-like f-box protein 5-like - subunit of NADH dehydrogenase
Qsu-miR-B CAGCCCTGTGTCGCTTCGATTCGT	24	Yes	328.32	3123.12	- binding protein/chloroplast thylakoid membrane/signal peptide - na
Qsu-miR-D GTCGTTGTAGTATAGTGGTG	20	No	14.58	122.33	LEAVES - Protein disulfide-isomerase - Squalene monooxygenase - Phytol kinase chloroplast PHELLEM - Sieve element-occluding protein - Protein disulfide-isomerase - gpn-loop gtpase
Qsu-miR-E TCGGAAGAACAGCAGCCACCG	21	No	95.22	70.85	- bzip protein - na
Qsu-miR-F ATGATGGCCTAGGAATTTGAA	21	Yes	14.82	148.07	- Dihydrolipoyl dehydrogenase 1. mitochondrial; - Beta-amylase 1. chloroplastic; - Probable protein phosphatase 2C 52
Qsu-miR-G GTTTTGTCAAAGGATTAGCCAATT	24	No	6.42	13.91	- Ubiquitin-like-specific protease ESD4; - Sentrin-specific protease 2
Leaves specific					
Qsu-miR-1L GTTGGTTAGGATACTCGGCTCT	22	No	182.34		- 60S ribosomal protein L4
Qsu-miR-2L TCGACAAAATGCTTGTGGCGA	21	Yes	171.84		- Cytochrome P450 - Trafficking protein complex - fbd associated f-box protein
Qsu-miR-3L TGAAGCTGCCAGCATGATCTTA	22	No	165.12		- Auxin response factor; - Quinone oxidoreductase-like protein
Qsu-miR-4L TGCCGGAGTGCGGGACGATGCGGG	24	No	78.30		- Actin
Qsu-miR-5L CTTGATGATGCTGCATCGGCA	21	No	42.60		- Aminoacid binding - 4-coumarate--CoA ligase 2; - Tropinone reductase
Qsu-miR-6L TGCCAAAGGAGAATTGCCCT	20	No	39.96		- Leucoanthocyanidin reductase - glycolate oxidase - peptidase family m48
Qsu-miR-7L AGGGGAGTAGAGTAGAGTTGGCT	23	No	37.26		- Related lipid transfer protein
Qsu-miR-8L CGGGGCGTGGACCGATGCG	19	No	24.96		- gata transcription factor 9-like
Qsu-miR-9L TTTGAGCCAAGGATGACTTGC	21	No	20.58		- f-box protein - RCD1-like - 12-oxophytodienoate reductase
Qsu-miR-10L TTGGTGGAACAAAAAGTGGTA	21	No	12.30		- cinnamoyl reductase; - Dehydratation induced protein - Elongation factor
Qsu-miR-11L ATTTTAGGAAACTTTTTATGGAA	23	No	6.48		- NADHP hc-toxin reductase
Phellem/phellogen specific					
Qsu-miR-13P AAATGGGTGCGTTGGCAAGAA	21	No		1053.52	- Fumarylacetoacetase - early nodulin-like protein
Qsu-miR-1P GGTGTCGTGGTGTAGTTGG	19	No		273.26	- Oligopeptide transporter; - Probable 26S proteasome non-ATPase regulatory subunit 3; - Ferredoxin-dependent glutamate synthase
Qsu-miR-2P GTCTGGGTGGTGTAGTTGGTTA	22	No		226.59	- receptor serine-threonine protein - histone deacetylase
Qsu-miR-3P CCTGGACGGGGTCAATGGGTGA	22	No		189.67	- gdsl esterase lipase actvity
Qsu-miR-4P GTTGAGATGGCCGAGTTGG	19	No		102.70	- putative copper-transporting ATPase; - Heat shock protein 90; - Probable rhamnose biosynthetic enzyme
Qsu-miR-12P	19	No		85.80	- integrin-linked protein kinase family

CCTGGACGGGGTCAATGGG					- phosphatise protein - na
Qsu-miR-5P TTCCCCAGCGGAGTCGCCA	19	No		35.23	- Calcium-transporting ATPase; - 4-coumarate--CoA ligase; - enolase (chloroplastic-like)
Qsu-miR-6P ATCCCCAGCGGAGTCGCCA	19	No		34.45	- Calcium-transporting ATPase - gdp-mannose-dehydratase - Enolase
Qsu-miR-7P CATAGGTTTGGGTCCGAGGA	20	No		27.56	- Histone acetyl transferase
Qsu-miR-8P GATCACCCATGACTTCTGTGTA	22	Yes		23.79	- ATP-synthase alpha-subunit - Glycerol-3-phosphate dehydrogenase - Ubiquitin-conjugating enzyme E2
Qsu-miR-9P CATTGAGTGCAGCGTTGATGA	21	No		23.53	- Calcium-transporting ATPase - Laccase - transposase protein
Qsu-miR-10P ATCCCCAGTGGAGTCGCCA	19	No		18.85	- Enolase - gdsl esterase lipase 5-like - Serine palmitoyltransferase
Qsu-miR-11P GCGGGAGGTCTTGAGTTCG	19	No		13.26	- Lupeol synthase; - ATPase related DNA repair protein - Syntaxin-like protein

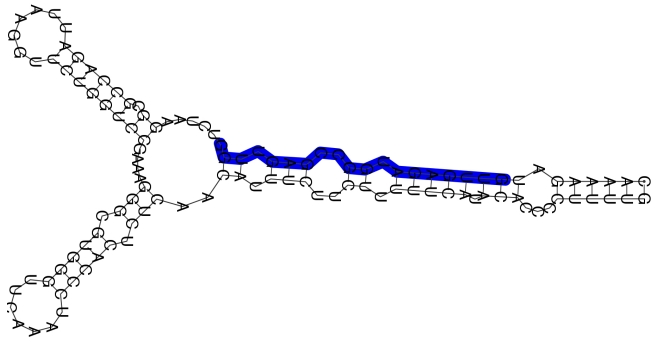
Normalised count: count of matching sequence reads normalised to total number of reads after last filtering step (see table below). Only picked up the top 3 best blastx hit with Oak transcripts: 1. all predicted target to Poplar genes were removed (though they are in the initial target site detection); 2. all homologous proteins were collapsed, ie. if there are transporter1, transporter2, named as transporter and search proteins with different function. 3. Manually searched the list from the smallest e-value (but for sure all blastx e-value < 1e-10), and the blast hit match with plant

Table 4. Predicted pre-miRNA secondary structure for phellem-specific novel miRNA in cork oak**.** The hairpin structures were plotted based on miRCat prediction using the RNA fold software. The miRNA is highlighted in blue and the miRNA* in red. The adjusted Minimum Free Energy (MFE) and the number of mismatches are included.

Novel miRNA	**Hairpin structure** **Hairpin structure in Vienna format**
Qsu-miR- 13P AAATGGGTGCGTTGGCAAGAA **MFE** -45.75 **mismatches** 3	 ((((((((((((((.((.((((.(((.(((<<-<<<<<<<<<<<<<-<<<<-(((((((((.....)))))...((((....))))....))).)))).)))))).)))))))))))).)))))))).))))))))))
Qsu-miR-1P GGTGTCGTGGTGTAGTTGG **MFE** -32.76 **mismatches** 4	 (((((..((((.(((.((.(<<-<<<--<<<<<-<<<<<.....)))))))))...)).).))).))...))).)))...)))))
Qsu-miR-2P GTCTGGGTGGTGTAGTTGGTTA **MFE** -40.83 **mismatches** 6	 (((((((.(((.((((...((..((..-<<<<<---<<<<<<<<<<<--....)))))).)).)))).)))))..))...)))).))))))))))
Qsu-miR-3P CCTGGACGGGGTCAATGGGTGA **MFE** -40.67 **mismatches** 6	 (((((..(((.<<<---<<<-<-<<<<<<<<-<(((((.(((((...)))))..)))))).))..)))))))).)))))))..)))))

Qsu-miR-4P
GTTGAGATGGCCGAGTTGG

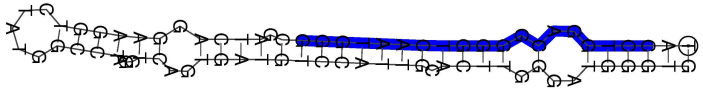
MFE -29.48
mismatches 5



(((((((.(<<<<<<<-<<--<<<-<<-.....((.((((((.......)))))))....(((.(((((.......)))))..)))..)).)))..)).)))))).)))...))))))

Qsu-miR-12P
CCTGGACGGGGTCAATGGG

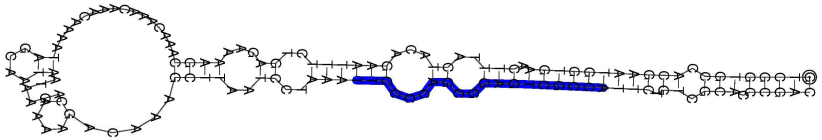
MFE -40.65
mismatches 3



((<<<<--<-<<<<<<<<<<<(.((((.(((((((.....))))...)).)).)))))))))))))).)))))..)..))))))

Qsu-miR-5P
TTCCCCAGCGGAGTCGCCA

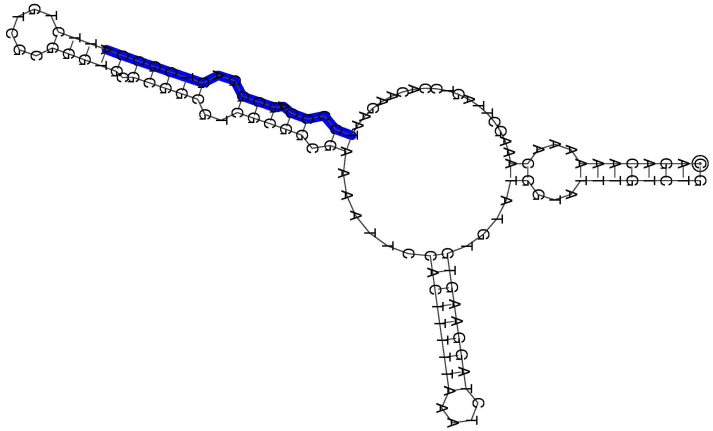
MFE -26.41
mismatches 5



(((((((((.((((((((((((.(((..((...(((((((..((...((((.................((....)).....(....).....))))..))..)))>>>--->>-->>>>>>>>>>))))).)).)))).)))))

Qsu-miR-6P
ATCCCCAGCGGAGTCGCCA

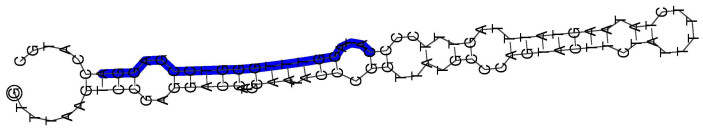
MFE -30.89
mismatches 5



(((((((((...((.................-<<-<<-<<<--<<<<<<<(((......))))).)))))..)))))).))......(((((((((....)))))))))....))...)))))))))

Qsu-miR-7P
CATAGGTTTGGGTCCGAGGA

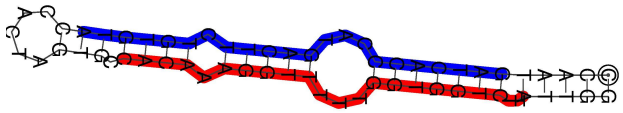
MFE -25.35
mismatches 5



((....((((..((((((...(((.(((.((..(.((..(((((((.((.....)).)))))))..)).)..)>--->>>>>>>>>>>-->>>)...))

Qsu-miR-8P
GATCACCCATGACTTCTGTGTA

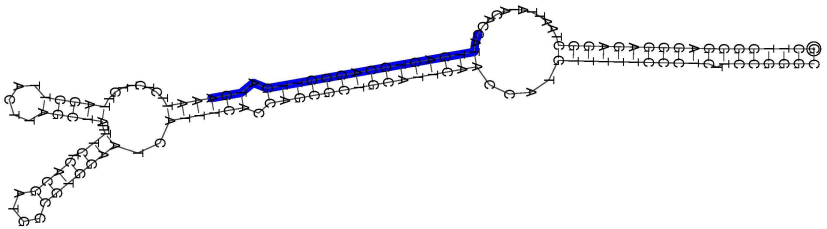
MFE -36.25
mismatches 4



(((((<<<<<<<---<<<<<-<<<<<<(.....))).}}}}}=}}}}}===}}}}}}}==}})))))

Qsu-miR-9P
CATTGAGTGCAGCGTTGATGA

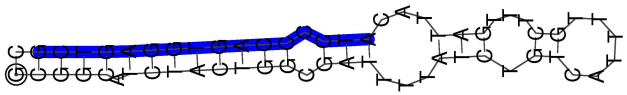
MFE -43.77
mismatches 3



((((((((((((((((((.....--<<<<<<<<<<<<<<<-<<<(((........(((((.....)))))..(((.((((....)))))))..)))))).)))))))))))....)))))))))).)))))))

Qsu-miR-10P
ATCCCCAGTGGAGTCGCCA

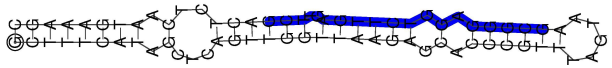
MFE -37.54
mismatches 1



(((((.(((((((.(((...(((.((.....)))...)))...>>>->>>>>>>>>>>>>

Qsu-miR-11P
GCGGGAGGTCTTGAGTTCG

MFE -38.64
mismatches 4



(((((((.(...((((((((((((.(.((((((.....)>>>>>->->>>>>>-->>>)))..).)))))))))