BMC Plant Biology

RESEARCH ARTICLE

Open Access

CrossMark

ChIP-Seq reveals that QsMYB1 directly targets genes involved in lignin and suberin biosynthesis pathways in cork oak (*Quercus suber*)

Tiago Capote[1,2,3], Pedro Barbosa[1,2], Ana Usié[1,2], António Marcos Ramos[1,2], Vera Inácio[3], Ricardo Ordás[4], Sónia Gonçalves[1,5*] and Leonor Morais-Cecílio[3]

Abstract

Background: Gene activity is largely controlled by transcriptional regulation through the action of transcription factors and other regulators. QsMYB1 is a member of the R2R3-MYB transcription factor family related to secondary growth, and in particular, with the cork development process. In order to identify the putative gene targets of QsMYB1 across the cork oak genome we developed a ChIP-Seq strategy.

Results: Results provide direct evidence that QsMY1B targets genes encoding for enzymes involved in the lignin and suberin pathways as well as gene encoding for ABCG transporters and LTPs implicated in the transport of monomeric suberin units across the cellular membrane. These results highlight the role of QsMYB1 as a regulator of lignin and suberin biosynthesis, transport and assembly.

Conclusion: To our knowledge, this work constitutes the first ChIP-Seq experiment performed in cork oak, a non-model plant species with a long-life cycle, and these results will contribute to deepen the knowledge about the molecular mechanisms of cork formation and differentiation.

Keywords: Cork oak, Chromatin-immunoprecipitation, Transcription factor, R2R3-MYB, ABCG transporters, Lipid-protein transporters

Background

Regulation of gene activity at the transcriptional level is the most common form of gene control. Regulation of transcription generally occurs via changes in the amounts and activities of transcription factors (TFs), which modulate the transcription of specific genes either activating or repressing the rate of transcription [1]. TFs interact and function in a combinatorial manner forming homo and heterodimers, assembling on control elements of DNA and recruiting cofactors [2, 3]. Cofactors in turn recruit DNA and histone modifying enzymes modulating the chromatin configuration state [1]. TFs also function in

networks in which a protein may regulate the expression of another to control directly or indirectly the expression of a particular gene or group of genes in a temporal and spatial fashion manner, allowing the unique expression of each gene in different cell types and during development [1].

A distinct characteristic of TFs is that they present a specific DNA-binding domain, which confers them the ability to bind to specific DNA regions and controlling target genes. Chromatin immunoprecipitation (ChIP) followed by high-throughput DNA sequencing (ChIP-Seq) is a method widely used to identify the binding sites of a target protein across a genome. In a ChIP assay, a TF, cofactor, or other chromatin protein is enriched by immunoprecipitation from cross-linked cells along with its associated DNA [4]. The resulted DNA is then sequenced and mapped against the species genome in order to identify the binding sites of protein of interest. This allows the subsequent identification

* Correspondence: sonia.goncalves@cebal.pt
[1]Centro de Biotecnologia Agrícola e Agro-alimentar do Alentejo (CEBAL) / Instituto Politécnico de Beja (IPBeja), Beja, Portugal
[5]Present Address: Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge CB101SA, UK
Full list of author information is available at the end of the article

Capote *et al. BMC Plant Biology* (2018) 18:198

Page 2 of 19

of their gene targets, unravelling potential regulatory networks [4].

Plant TFs are characterized by a larger number of genes and by the diversity of families. In *Arabidopsis* there are around 2000 TFs genes belonging to approximately 30 different TF families [5]. Most plant TFs as APETALA2/Ethylene Responsive element binding factor (AP2/ERF), NAC, MADS box, basic helix-loop-helix (bHLH), basic leucine zipper (bZIP) and myeloblastosis (MYB) form large domain families playing important roles in the control of plant growth and development [1]. The MYB family constitute the most abundant group of TFs in plants. In *Arabidopsis,* 198 MYB TFs were identified, of which 126 belong to the R2R3-MYB subfamily [6]. The evolution of R2R3-MYBs in plants seems to be related with a specific expansion of the subfamily giving rise to species-specific gene subgroups in certain species [7]. This expansion is mainly attributed to whole genome and segmental duplication as gene sequence and phylogenetic tree analyses confirm [7]. They have been classified into 28 subgroups according to the conserved amino acid sequence motifs present in C-terminal MYB domain [3, 8].R2R3-MYB TFs are characterized by their role in a variety of plant-specific processes, such as cell shape and morphogenesis, cellular proliferation and differentiation, hormone response, abiotic and biotic stress, and regulation of primary and secondary metabolism such as phenylpropanoid, lignin and suberin metabolism [9, 10]. However analysis of the promotor sequences in chinese pear demonstrated that transcriptional regulation of the MYB genes is variable among species [11].AtMYB41, for example, it is known to be involved although indirectly in the regulation of suberin biosynthesis, export, assembly and deposition in plants under stress conditions [12], however it not expressed in the poplar suberized tissue phellem [13]. It was recently reported a relation between AtMYB107 and AtMYB9 which synchronize the transcriptional induction of aliphatic and aromatic monomer biosynthesis as well as suberin transport and polymerization in seed outer integument layer [14]. Gou et al., [10] reported AtMYB107 as a positive regulator for seed coat suberin synthesis, highlighting the important role of MYB TFs in suberin synthesis regulation. In poplar at least 18 MYB transcription factors were reported as up-regulated in phellem with functions related with the suberin metabolic pathways [13]. Among these, Arabidopsis homologs of *AtMYB107*, *AtMYB9* and *AtMYB93* were found, which have regulatory functions in the suberization process [10, 14–16]. Also,homologs of *AtMYB96, AtMYB94, AtMYB106* and *AtMYB16* related with cuticle metabolism [17–19], homologs of *AtMYB3, AtMYB7, AtMYB63* linked to phenylpropanoid metabolism [20, 21], an homolog of root AtMYB36 which is related with

root casparian band formation [22], and *AtMYB111* with a regulatory function in the flavonoid metabolism were identified [23].

Cork oak (*Quercus suber L.)* is an evergreen broadleaved tree species native to the Mediterranean basin. It is a valuable economic resource due to the sustainable exploitation of its thick bark, the cork [24, 25]. Cork or phellem is a tissue derived from the meristematic activity of the phellogen characterized by a layered deposition of suberized death cells [26]. The high content in suberin provides cork with unique insulator and elastic properties that translates into a large variety of industry applications. Despite the importance of cork, the knowledge on the molecular mechanisms on cork formation and development, are still poorly understood.

Almeida and co-workers (2013) [27] have characterized a R2R3-MYB gene previously identified as related with cork formation and differentiation [28], which was named QsMYB1. Authors showed that QsMYB1 is mainly active in organs and tissues with secondary growth resulting from the activity of phellogen [27]. Moreover, *QsMYB1* transcripts are more abundant in cork, a highly suberized tissue, than in wood, a lignified but non-suberized tissue. The authors have also found several *cis*-acting regulatory elements related to phenylpropanoid pathway in the *QsMYB1* promotor region [29]. QsMYB1 expression was also shown to be modulated in response to heat and drought stress, which points to a function in the regulatory network of cork oak response to abiotic stress [29]. These findings led the authors to hypothesize that QsMYB1 may be regulating one or more metabolic pathways involved in cork formation, namely the biosynthesis of lignin and suberin which constitute the major biopolymers present in cork.

The present work aims to validate these results, by developing a ChIP-Seq strategy to identify the QsMYB1 target genes. Results showed that QsMYB1 acts directly on genes of the lignin and suberin metabolic pathways, confirming the specific role of QsMYB1 in regulating cork formation and development, a specific biological process with great relevance in cork oak.

Results

QsMYB1::3xFLAG somatic embryo selection for ChIP-Seq

Stable modified cork oak somatic embryo cell lines overexpressing QsMYB1::triple FLAG fused (QsMYB1::3x-FLAG) protein were produced using Ce1 line, previously generated (unpublished observation). Ce1 line, stable and with a high rate of secondary embryogenesis, was previously characterized for kanamycin resistance, allowing the determination of the somatic embryo natural resistance to grow under the presence of the antibiotic. Concentrations over 25 mg/ml of kanamycin showed to inhibit embryo proliferation setting kanamycin at this

Capote *et al. BMC Plant Biology* (2018) 18:198

Page 3 of 19

concentration as selective agent of putative transformants. After 30 days subcultures during 24 months, MYB1::3xFLAG transcripts were detected in several transformed embryo clusters by quantitative real time PCR (RT-qPCR). When comparing with the expression of *QsMYB1* in non-transformed embryos, MYB1::3x-FLAG transcripts presented a 4–12 fold change in the level of expression (Fig. 1a). Furthermore, the specificity of the anti-FLAG antibody was validated by western blot (Fig. 1b) and the MYB1::3xFLAG protein was detected in the nucleus by fluorescence microscopy (Fig. 1c), confirming the nuclear location of the protein. Due to the lack of known target DNA of QsMYB1 and consequent impossibility to control the enrichment quality of ChIPed DNA, we selected embryo clusters with fold change expression of *QsMYB1::3xFLAG* higher than 6
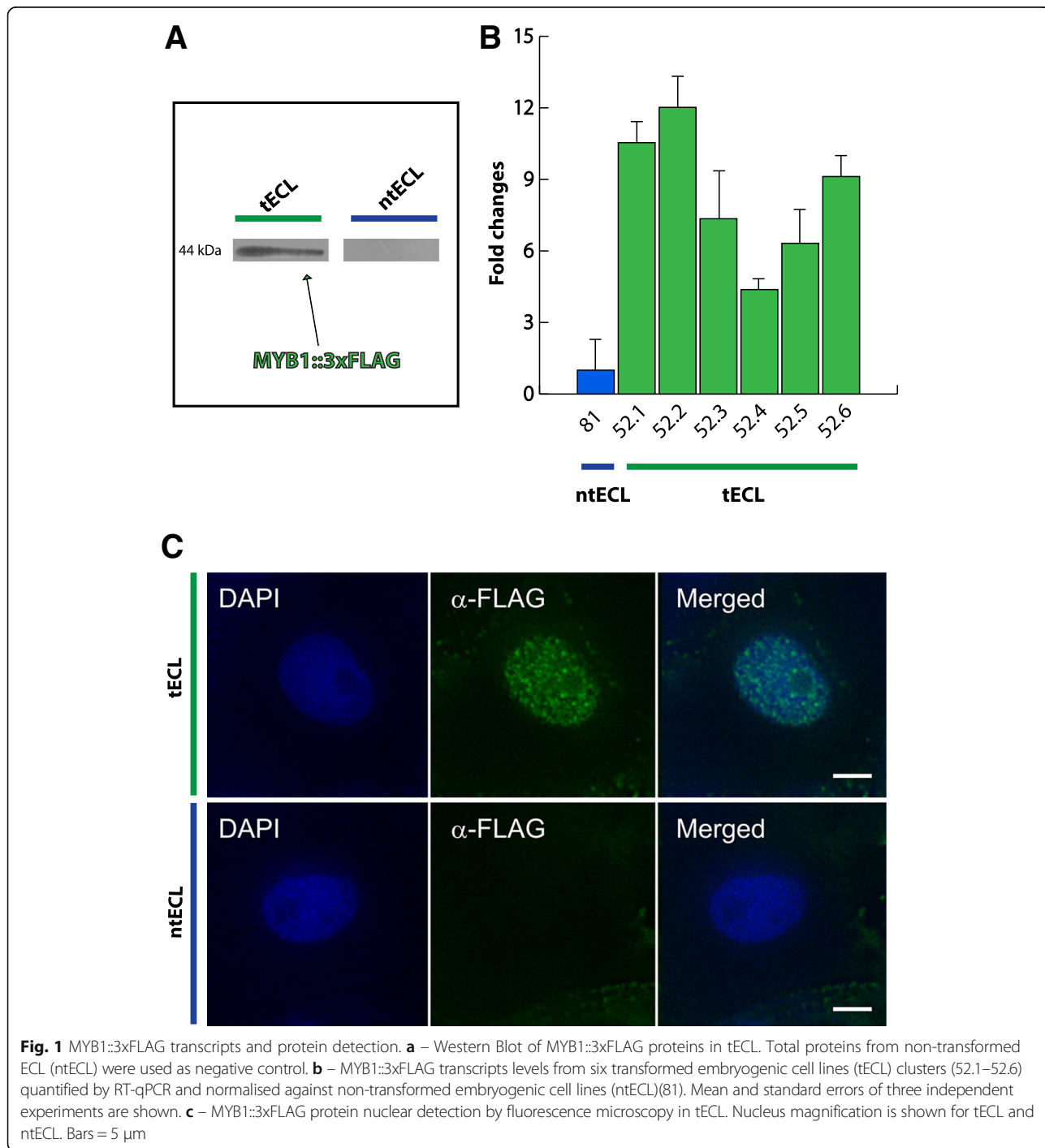

Fig. 1 MYB1::3xFLAG transcripts and protein detection. **a** – Western Blot of MYB1::3xFLAG proteins in tECL. Total proteins from non-transformed ECL (ntECL) were used as negative control. **b** – MYB1::3xFLAG transcripts levels from six transformed embryogenic cell lines (tECL) clusters (52.1–52.6) quantified by RT-qPCR and normalised against non-transformed embryogenic cell lines (ntECL)(81). Mean and standard errors of three independent experiments are shown. **c** – MYB1::3xFLAG protein nuclear detection by fluorescence microscopy in tECL. Nucleus magnification is shown for tECL and ntECL. Bars = 5 μm

Capote *et al. BMC Plant Biology* (2018) 18:198

Page 4 of 19

(Fig. 1a) in order to have enough protein suitable to perform the ChIP-Seq assay.

QsMYB1 target genes identified by ChIP-Seq

In order to identify the DNA targets of QsMYB1, ChIP-Seq was performed by immunoprecipitating QsMYB1 protein and the cross-linked DNA after chromatin fragmentation of ~ 300 bp (Additional file 1: A and B). DNA was purified and further sequenced with the Illumina HiSeq 4000 system producing ~ 70–90 million reads per sample of which ~ 123 million reads uniquely mapped to the cork oak draft genome (Table 1). One ChIP (anti-FLAG) and one mock (IgG) library were sequenced with one technical duplicate each. After reads, filtering peaks were called with the MACS2 and selected for downstream analysis according to the ENCODE ChIP-Seq guidelines (detailed description of data analysis in Methods section). A total of 18,165 putative binding sites were identified. It was firstly analysed whether the peaks in the ChIP experiment were distributed by genic regions. Based on these analyses, MYB1 binding sites are located in genic regions of 14,290 genes: 13.4% in promotor regions, 8.1% in 5′ untranslated regions (UTR), 19.2% in intron regions, 25.5% in exon regions, 6.5% in 3' UTR and 6.0% in the terminator regions (Fig. 2a). Intergenic regions represent 18.4% of the total binding sites and 2.9% of the binding sites were not annotated (unknown) (Fig. 2a). The Peak-calling analysis identified several target genes reflecting the binding of QsMYB1 to specific DNA loci. Between these genes, QsMYB1 targets other TFs, genes encoding for enzymes related with the lipid metabolism and transport, as well as enzymes from the phenylpropanoid pathway. Amongst these genes, several are related with various aspects of cork formation, as suberin and lignin biosynthesis or transport and deposition of suberin monomeric units. In order to explore the regulatory functions of QsMYB1 in cork formation and differentiation we focused the analysis in these genes, however QsMYB1 potentially targets a vaster number of genes.

Binding motif analysis reveals QsMYB1 *cis*-regulatory elements

To explore the MYB1 biding motifs environment, 1-kb flanking sequences around all peaks summit were analysed by the motif discovery tool MEME-ChIP using the DAP motifs [30] and PBM motifs [31] databases (Fig. 2b).

Motifs CWHCAA (E-value = 3.5e-50), CYTCBTC (E-value = 8.8e-38) and BKTGG (E-value = 2.9e-30) were the most statistically relevant. The BKTGG motif is presented in 12,308 peaks (67.9%), while the CWHCAA and the CYTCBTC motifs were identified in 7926 (43.7%) and 2016 (11,1%) peak sequences, respectively. When the occurrence of motifs by binding locations is individually analysed, it is observed that some genic regions present more than one identity motif. Nevertheless, each binding feature has a similarity with the 3 global motifs identified (Fig. 2a).
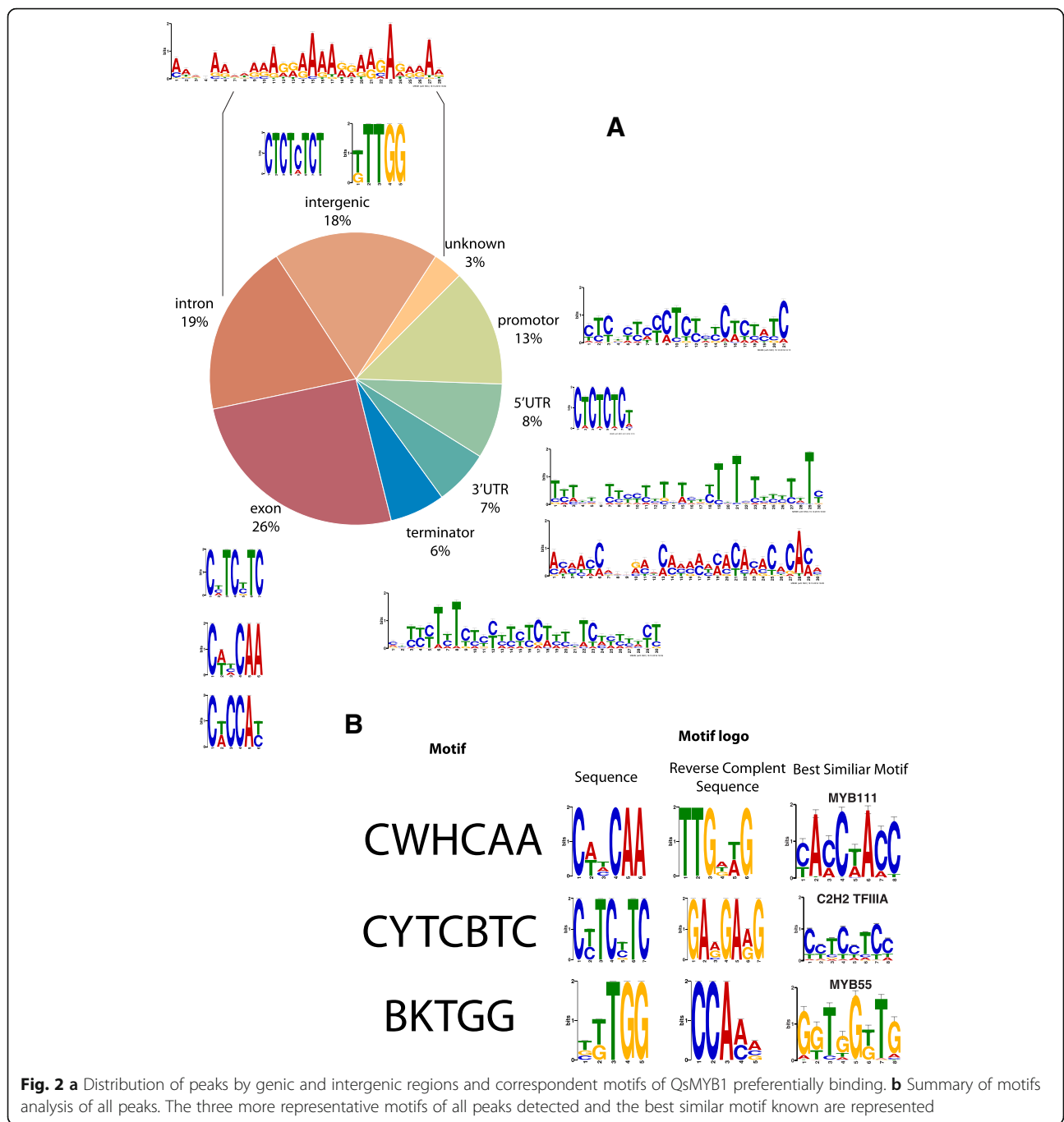
QsMYB1 directly targets other transcriptional elements

To detect genes regulated by QsMYB1 gene, the QsMYB1 DNA-targets putatively encoding for TFs, transcription regulators and chromatin regulators the genes associated with each peak were analysed with the PlantTFcat tool [32]. Results showed that 414 regulatory genes are targeted by QsMYB1: TF (42%), transcription regulators (50.5%) and chromatin regulators (7.5%) comprising different types of regulators families (Fig. 3). From these families the most representative putatively encodes for CCHC zinc finger proteins (CCHC) (192), C2H2-type zinc fingers TFs (C2H2) (43), WD40-like TFs (19), BED-type zinc finger proteins (BED-type(Zn)) (13), AP2/ERF TFs (10) and bHLH TFs (10) (Additional file 2).

Enzyme encoding genes targeted by QsMYB1

To identify the targets of QsMYB1 directly related with secondary growth and cork formation, all QsMYB1 putative target genes were submitted to the KEGG database giving focus to the ones that present a fold enrichment higher or equal to 3 (Additional file 3). With this approach, 16 genes encoding three distinct enzymes essential to the phenylpropanoid pathway (Fig. 4) were identified, namely 4-coumarate: CoA ligase (4CL - 6.2.1.12), cinnamyl alcohol dehydrogenase (CAD - 1.1.1.195) and class III plant peroxidase (PPO - 1.11.1.7). Also related with the phenylpropanoid metabolism, the results show that QsMYB1 putatively targets three distinct genes encoding for β-galoctosidase (β-GAL - 3.2.1.21).

Diverse QsMYB1 target genes encoding for key enzymes on the metabolism of lipids were identified, namely enzymes related with fatty acid biosynthesis (Fig. 5) as acetyl-CoA carboxylase (ACC – 6.4.1.2), long-chain-fatty-acid-CoA ligase (LACS – 6.2.1.3) and members of fatty acid synthase complex: enoyl-[acyl-carrier-protein] reductase (NADH) (FABI – 1.3.1.9) and acyl-[acyl-carrier-protein] desaturase (ACP - 1.14.19.2). Moreover, we found two Arabidopsis homolog genes of β-ketoacyl-CoA synthase in our data. Related with fatty acid oxidation (Fig. 6), the results revealed two genes encoding for a putative acyl-CoA oxidase (ACX – 1.3.3.6) and two genes encoding for a aldehyde dehydrogenase (NAD+) (ALDH - 1.2.1.3). Also, two distinct target
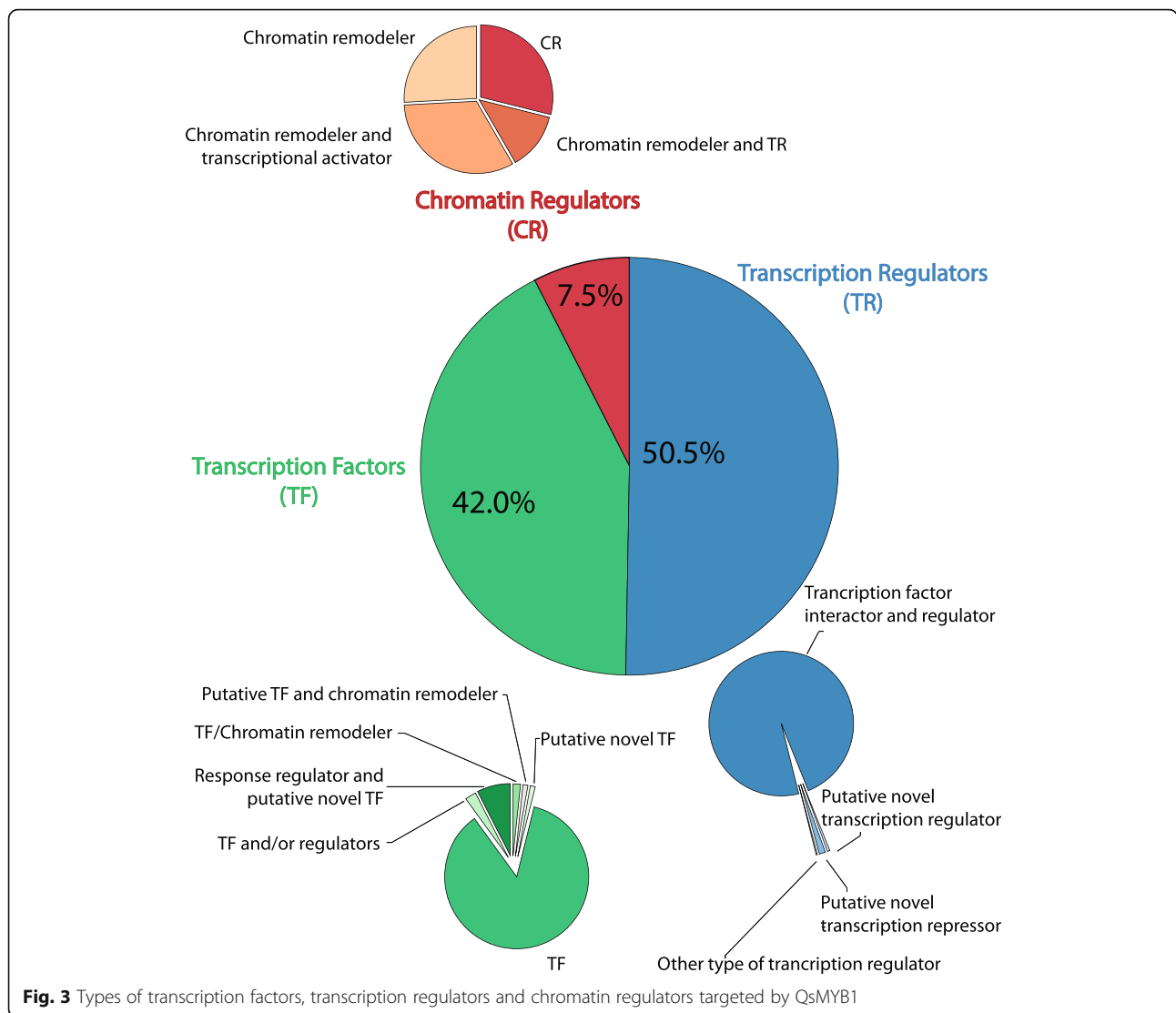
Table 1 Sequencing throughput and mapping results obtained for ChIP and mock samples

Sample	Raw reads	Pre-processed	UMR reads	% UMR
ChIP_L1	75,987,297	72,855,787	28,558,358	37.58%
ChIP_L5	66,520,992	63,594,856	24,887,132	37.41%
Mock_L1	93,072,373	89,005,882	37,335,088	40.11%
Mock_L5	80,189,676	76,512,503	32,053,662	39.97%

Capote *et al. BMC Plant Biology* (2018) 18:198

Page 5 of 19

Fig. 2 a Distribution of peaks by genic and intergenic regions and correspondent motifs of QsMYB1 preferentially binding. **b** Summary of motifs analysis of all peaks. The three more representative motifs of all peaks detected and the best similar motif known are represented

genes related the ω-hydroxylation of various fatty acids and with high homology for the Arabidopsis Cytochrome P450 (CYP) 86A8 (*AtCYP86A8*) and for the Arabidopsis CYP 96A1 (*AtCYP96A1*) were found.

The results additional show several genes encoding for enzymes involved in the glycerol metabolism (Fig. 7) as aldehyde dehydrogenase (NAD+) (ALDH - 1.2.1.3), NADP-dependent alcohol dehydrogenase (ADH – 1.2.1.1), NADPH-dependent aldehyde reductase (ADR – 1.1.1.21), glycerol 3-phosphate acyltransferase (GPAT – 2.3.1.15),

1,2-diacyl-3-β-D-galactosyl-*sn*-glycerol acylhydrolase (DGL – 3.1.1.26), triacylglycerol acylhydrolase (LIP - 3.1.1.3) and diacylglycerol O-acyltransferase (DGAT - 2.3.1.20).

Several genes encoding for enzymes related with the glycerophopspholipid metabolism (Fig. 8) were also identified: glycerol-3-phosphate dehydrogenase (NAD +) (GPDH - 1.1.1.8), glycerol-3-phosphate dehydrogenase [NAD(P)+] (GPDH - 1.1.1.94), phospholipase A1 (PLA1–3.1.1.32) and the phosphoethanolamine N-methyltransferase (NMT – 2.1.1.103).

Capote *et al. BMC Plant Biology* (2018) 18:198

Page 6 of 19

Fig. 3 Types of transcription factors, transcription regulators and chromatin regulators targeted by QsMYB1

MYB1 target genes related with lipid transporters

Our ChIP-Seq data reveals that QsMYB1 is targeting several ATP-binding cassette proteins (ABC) or ABC-like transporters. Of these, 26 present high homology for the G family of ABC (ABCG) genes (Additional file 4). Five of these genes have high similarity with *AtABCG11*, six of them with *AtABCG37* and six with the *AtABCG40*. In addition, we identified five genes encoding for lipid-transfer proteins (LTPs) (Additional file 5), three of them with similarity to protease inhibitor/seed storage/LTPs and two of them putative LTPs genes in *Q. suber*.
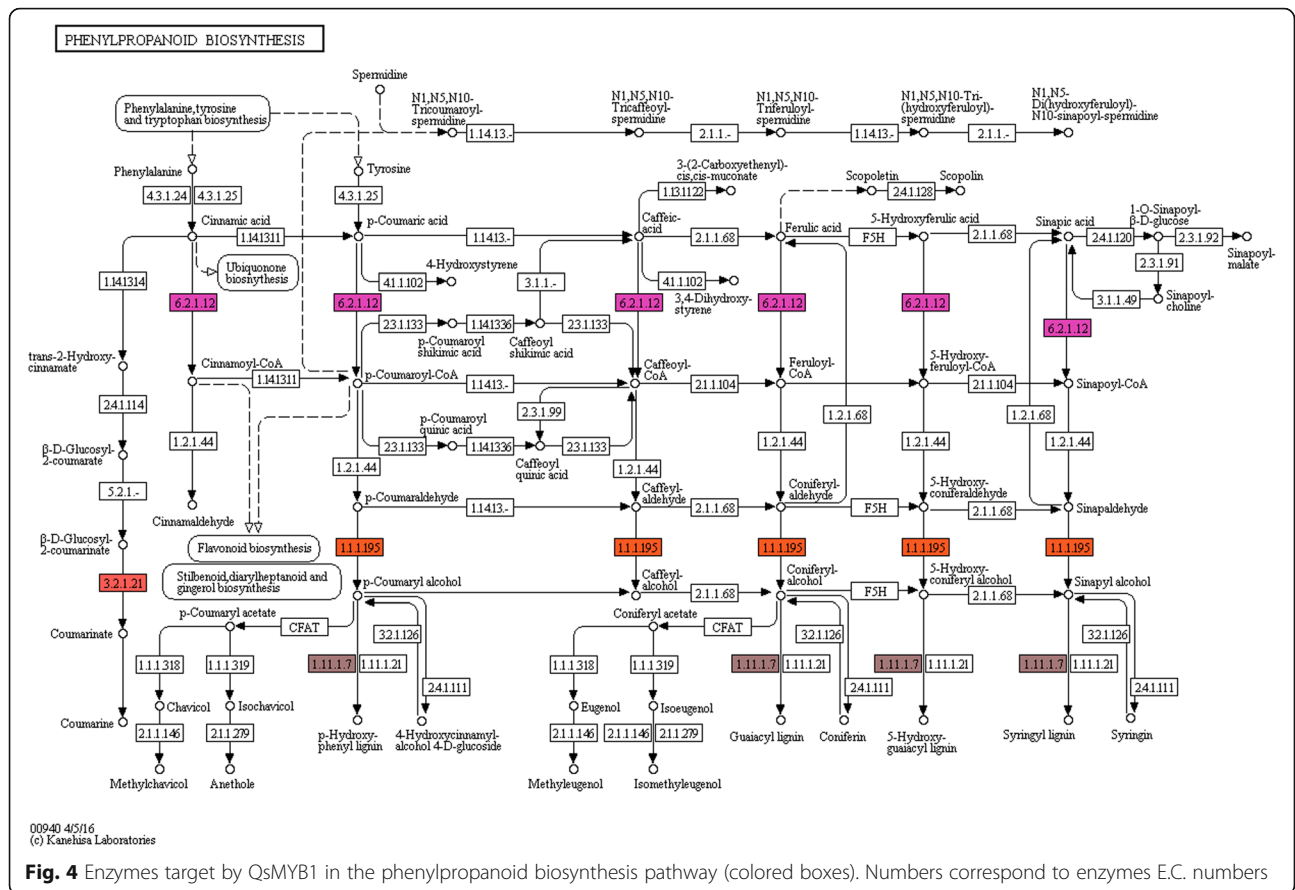
Relative expression of QsMYB1 targeted genes

The expression of five putative genes targeted by QsMYB1: *QsGPAT, Qsb-GLU, QsCAD, QsABCG11* and *QsPOX* were evaluated by RT-qPCR on transformed embryos overexpressing QsMYB1 (tECL) and non-transformed embryos

(ntECL)(Fig. 9). Results shows relative increased expression when genes are targeted in 5'UTR (*QsGPAT, QsABCG11* and *POX*) and in the promotor region (*Qsβ-GLU*). In contrast, *QsCAD* which is targeted in an exonic region exhibited a relative decreased expression in embryos overexpressing QsMYB1::3xFLAG.

Discussion

Quercus suber MYB1 (QsMYB1) is a TF that is mainly expressed in organs and tissues with secondary growth resulting from the activity of the phellogen [27]. This points to a putative regulation of QsMYB1 in the lignin and suberization processes and therefore possibly in cork formation and differentiation. In fact, *Arabidopsis* mutants generated by gene trap insertion in *AtMYB68*, the *QsMYB1* orthologue, produce increased biomass and lignin levels relatively to wild type. In addition its closest

Fig. 4 Enzymes target by QsMYB1 in the phenylpropanoid biosynthesis pathway (colored boxes). Numbers correspond to enzymes E.C. numbers
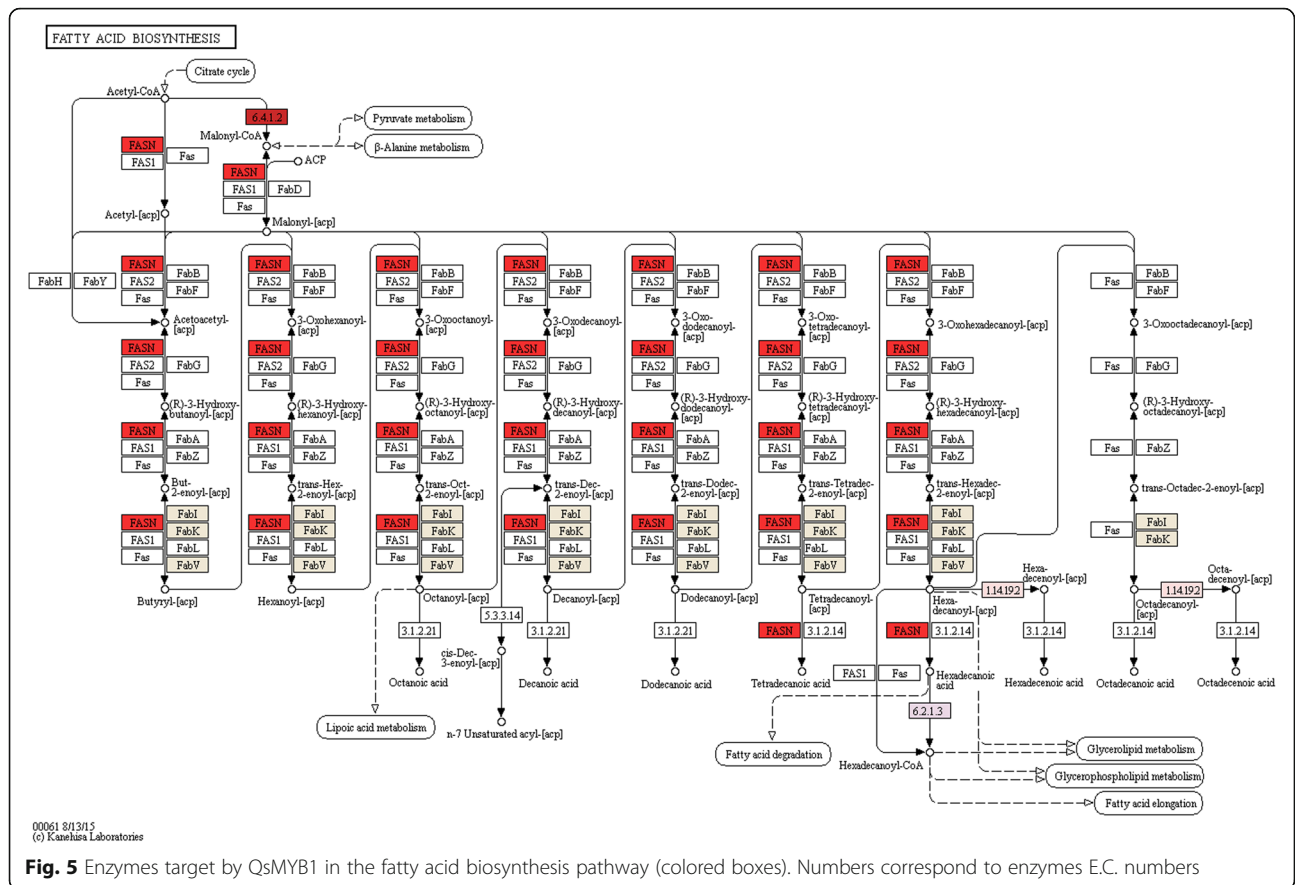
homolog, *AtMYB84*, exhibit an overlapping expression suggesting a partly redundant function [33].

In order to identify the genes regulated by QsMYB1 in a genome-wide scale, a cork oak embryo model was established, allowing protein overexpression, isolation and easy access to genomic material. In this context, genetic modified cork oak embryos overexpressing QsMYB1 were produced and used in a ChIP-Seq experiment allowing for the identification of the QsMYB1 putative binding sites. Although this is a proxy model, as embryos do not present neither the developmental stage nor the secondary growth tissues where *QsMYB1* is naturally up-regulated, it allowed for the identification of several QsMYB1 target genes with specific and defined DNA binding domains related with several aspects of tissue lignification, suberization and consequently with cork formation, which are further discussed. Furthermore QsMYB1 may not target all the binding sites as at the secondary tissues where it has been found to be expressed [27]. However, the overexpression system used in this study, which forces the production of QsMYB1, is an attempt to mimic the up-regulation and DNA-binding status observed in other tissues or organs [27]. The results showed that for five QsMYB1 gene targets analysed by RT-qPCR there is a pronounced gene

expression modulation in the embryos overexpressing QsMYB1. These findings confirm that QsMYB1 regulates its expression. Moreover, it was observed that, for the tested genes, when QsMYB1 targets a gene region upstream of the coding sequence (e.g. 5'UTR and promoter) the gene expression is up-regulated whereas when the target region is an exon, the targeted gene expression is down-regulated (Fig. 9). Thus, QsMYB1 seems to have a dual opposite regulatory role either up-regulating or down-regulating gene expression.

QsMYB1 targets phenylpropanoid pathway genes regulating the biosynthesis of essential lignin and suberin precursors

The results of our study clearly show that QsMYB1 is targeting genes coding for key enzymes responsible for the biosynthesis of phenylpropanoids, which constitute phenolic components of suberin and lignin polymers. Members of the R2R3-MYB TF family are well known for their regulatory function in diverse metabolisms as the phenylpropanoid [34], lignin [35, 36], lipid [37] and flavonoid metabolism [38] among others. Phenylpropanoids are phenolic derivatives which contain a phenyl ring and a C_3 side chain that comprise a multitude of plant secondary metabolites and cell wall components as

Capote *et al. BMC Plant Biology* (2018) 18:198

Page 8 of 19

Fig. 5 Enzymes target by QsMYB1 in the fatty acid biosynthesis pathway (colored boxes). Numbers correspond to enzymes E.C. numbers

lignin and suberin [39, 40]. The *Arabidopsis MYB41*, phylogenetically close to *AtMYB68* (the Arabidopsis homolog of *QsMYB1*), when overexpressed has an up-regulatory effect on phenylpropanoid, lignin and suberin synthesis and related-genes expression [12]. Like AtMYB41, our results indicate an involvement of QsMYB1 in the phenylpropanoid metabolism. QsMYB1 targets 4-coumarate:CoA ligase which is involved in the conversion of ferrulic, caffeic and coumaric acids to ω-OH-alkyl-ferulates/n-alkyl ferulates, n-alkyl caffeates and n-alkyl coumarates, that in turn are linked to fatty alcohols to produce alkyl hydroxycinnamates, an essential constituent in suberin-associated waxes [41]. In addition, QsMYB1 targets cinnamyl alcohol dehydrogenase and class III plant peroxidase, which together with 4-coumarate:CoA are involved in the biosynthesis of the monomeric precursor of lignin *p*-hidroxy-cinnamyl alcohols, the so-called monolignols. Furthermore class III plant peroxidases are also involved in polymerization of suberin monomers in the apoplast [42]. A recent work demonstrated that various genes for enzymes of the phenylpropanoid pathway are up-regulated in the poplar phellem tissue, namely 4-coumarate:CoA ligase and cinnamyl alcohol dehydrogenase [13]. Together, the results indicate an involvement of QsMYB1 in the synthesis of

phenolic compounds from the phenylpropanoid pathway evidencing an earlier regulatory action in the synthesis of essential lignin and suberin precursors.

QsMYB1 is targeting genes encoding enzymes responsible for suberin biosynthesis

Suberin may be defined as a complex glycerol-based polymer consisting of a polyaliphatic polyester linked with phenolic components and embedded waxes [43–45]. Besides suberin composition vary between developmental stages, tissue and plant species, in *Q. suber* depolymerisation of cork suberin include monoacylglycerols of ω-hydroxy acids and α,ω-dicarboxylic acids, ferrulic acid linked to ω-hydroxi acids, trimeric diesters of glycerol linked to α,ω-dicarboxylic acids and ω-hydroxi acids linked to ferrulic acid and glycerol [46, 47]. Our results reveal that QsMYB1 is targeting genes that encode enzymes involved in synthesis of the suberin polyester building blocks; one *QsLACS* and two *QsKCS* predicted genes, constituting evidence of genes associated with suberin monomers synthesis in *Q. suber*. LACS catalyses the acyl activation reaction of free fatty acids to fatty thiosteres [48]. In poplar *LACS1* and *LACS2* are up-regulated in phellem [13]. In *Arabidopsis LACS1* and *LACS2* have overlapping roles activating fatty acids in cutin and wax
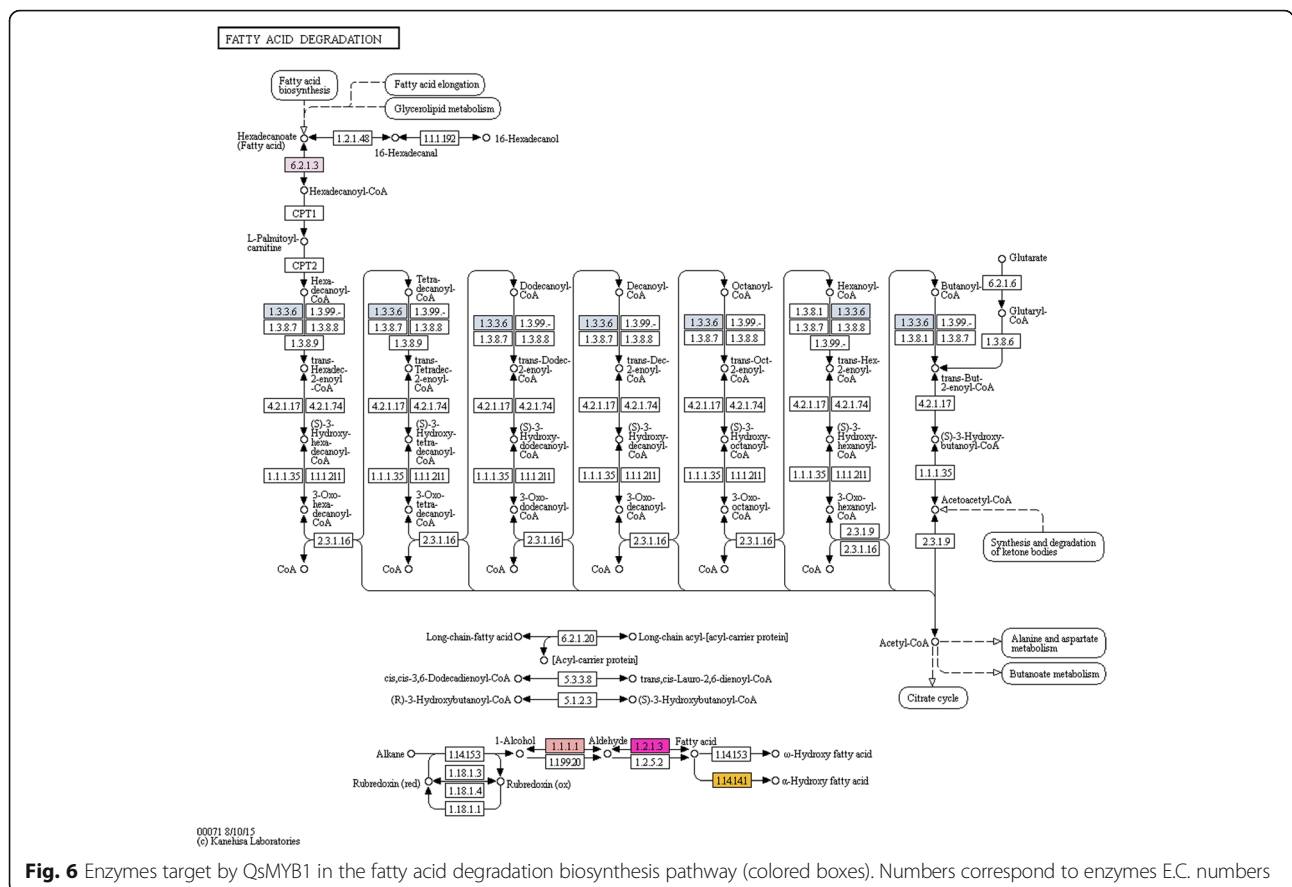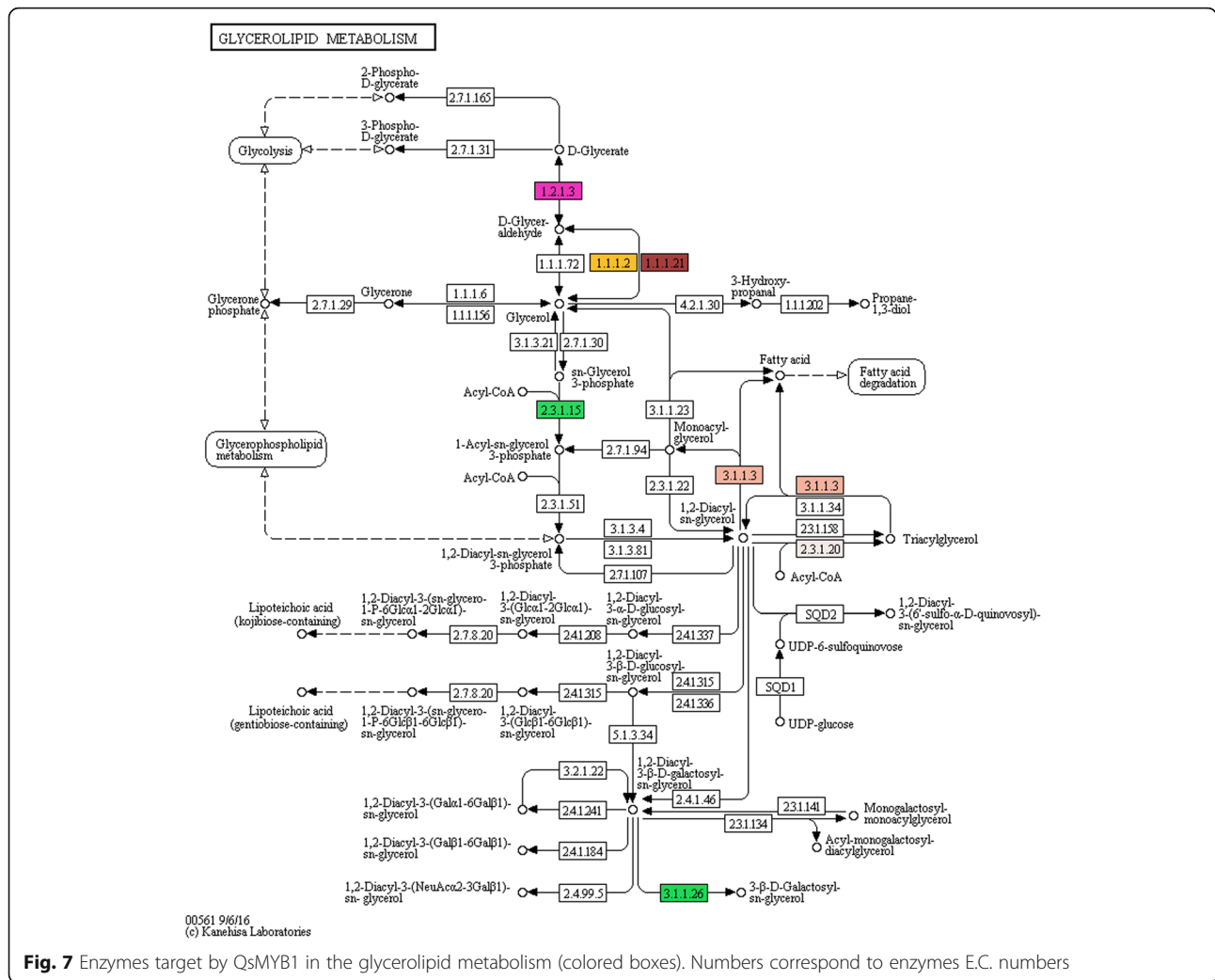
Capote *et al. BMC Plant Biology* (2018) 18:198

Page 9 of 19

Fig. 6 Enzymes target by QsMYB1 in the fatty acid degradation biosynthesis pathway (colored boxes). Numbers correspond to enzymes E.C. numbers

synthetic pathways. In addition, *LACS2* is involved in suberin formation besides the capability of LACS enzymes may also act on modified fatty acids as ω-hydroxy acids and α,ω-dicarboxilic acids before esterification to glycerol [49]. In turn, β-ketoacyl-CoA synthase (KCS) is responsible for controlling the extension of elongation of the long-chain fatty acyl-CoAs [50]. In *Arabidopsis*, two β-ketoacyl-CoA synthases, *AtKCS2* and *AtKCS20* are involved in elongation of C_{20} acyl chain suberin precursors [51, 52]. In potato, silencing of *StKCS6* gene resulted in reduction of suberin monomers with chain lengths of C_{28} and higher in the tuber periderm indicating that StKCS6 is involved in elongation of suberin precursors to C_{28} or higher chain lengths [53].

Several CYP86 family members were identified in our data, namely homologues of *AtCYP86A1* and of *AtCYP86A8*. CYP86A8 monooxigenase are capable of ω-hydroxylating saturated and unsaturated fatty acids with chain lengths from C_{12} to C_{18} and is likely to be involved in cutin biosynthesis [54]. Two major monomers of the suberin polyester are ω-hydroxy acids and α,ω-dicarboxylic acids, which in plants are formed by the hydroxiliation of the terminal methyl group (ω-position) catalysed by enzymes of the CYP86 subfamily of cytochrome P450 monooxygenases [55]. The *Arabidopsis*

CYP86A1 and CYP86B1 are involved in root suberin synthesis [55–57]. Similarly, in potato, the StCYP86A33 orthologue of AtCYP86A1 is important for the production of ω-functionalized suberin monomers of the tuber periderm resulting in significant reduction in 18:1 ω-hydroxy acids and α,ω-dicarboxilic acids when the gene is silenced [58].

Our results reveal a *QsGPAT* homolog target by QsMYB1 indicating a regulatory role of this TF in the production of monoacylglycerols, which may be considered as the initial step in the aliphatic suberin biosynthesis. In poplar, several *GPATs* are up-regulated in phellem, namely *GPAT5*, *GPAT6* and *GPAT8* [13]. In *Arabidopsis*, GPAT4, 6 and 8 are bifunctional enzymes with *sn*-2 acyltransferase and phosphatase activity, catalysing the formation of *sn*-2 monoacylglycerol products. GPAT5 in turn is involved in the transfer of very long-chain aliphatics to a glycerol based acceptor in *Arabidopsis* root and seed coat suberin [59]. *Arabidopsis* mutants of *GPAT5* show a large reduction in long-chain unsubstituted fatty acids, ω-hydroxy acids and α,ω-dicarboxilic acids in root and seed coat suberin with a total reduction of 50% in total suberin content [59]. In the case of GPAT7, the *GPAT7* gene is induced by wounding producing soluble suberin-like precursors in overexpressing lines [60]. These results
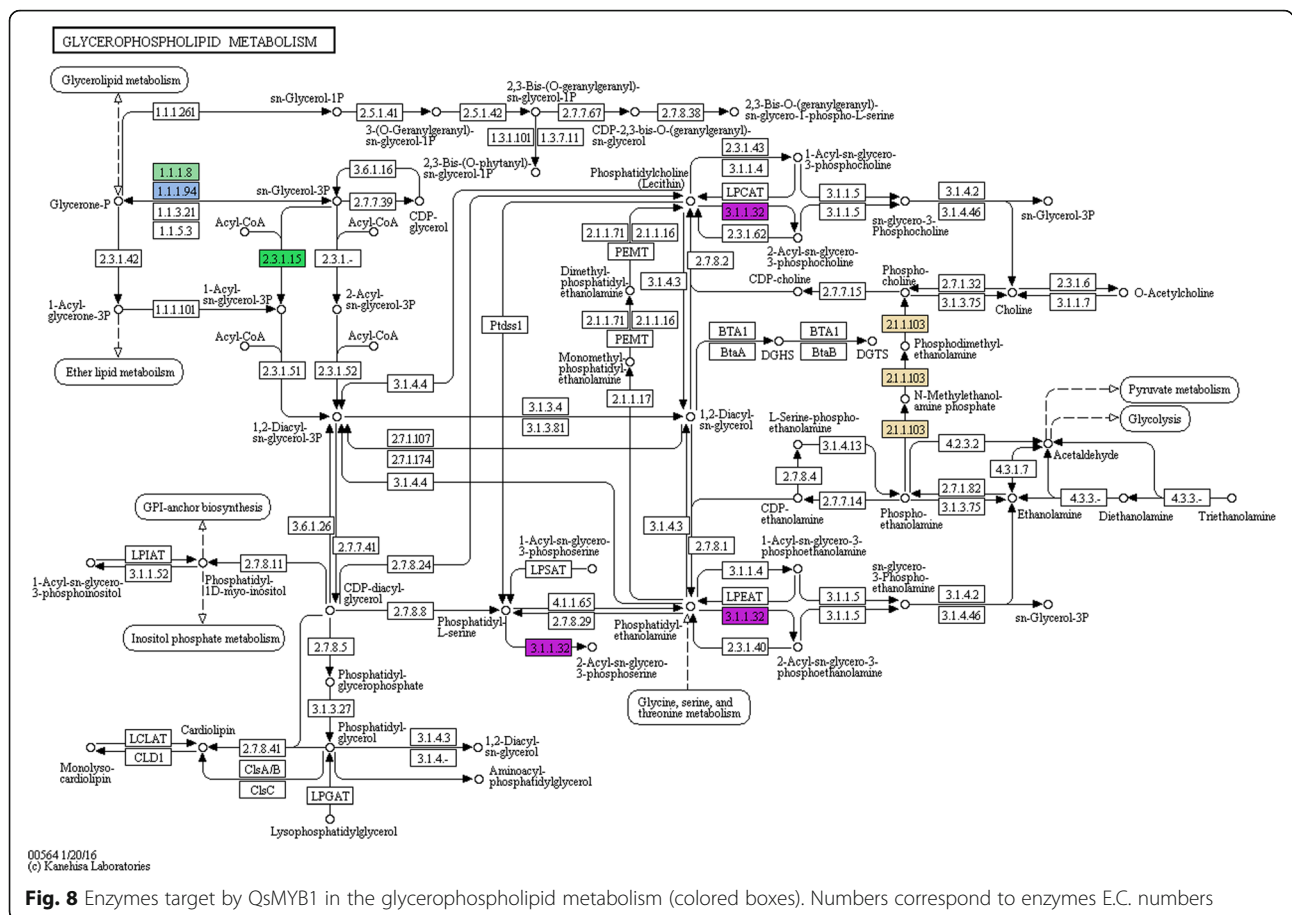
Capote *et al. BMC Plant Biology* (2018) 18:198

Page 10 of 19

Fig. 7 Enzymes target by QsMYB1 in the glycerolipid metabolism (colored boxes). Numbers correspond to enzymes E.C. numbers

clearly indicate that QsMYB1 is targeting genes encoding for enzymes that are responsible for the production of aliphatic domain of suberin.

QsMYB1 targets putative transporters of suberin-lipid components

By directly targeting genes coding for 4-coumarate:CoA, cinnamyl alcohol dehydrogenase, LACSs, KCSs, GPATs amongst others, QsMYB1 is likely to regulate the phenylpropanoid and the suberin biosynthesis pathways. In addition, QsMYB1 may have a regulatory effect in the expression of suberin transporters.

The various suberin lipidic precursors are generically transported from the endoplasmic reticulum where they are synthesized to and across the plasma membrane, and then polymerized in the apoplast to form the suberin barrier [61]. We identified several members of the ABCG gene family targeted by QsMYB1, namely homologs of *AtABCG11* and *AtABCG6*, which implicates directly QsMYB1 in regulation of lipid transport and suberin

formation. Transport of aliphatic lipids across the plasma membrane involves plasma membrane-localized transporters of the ABCG family which consists of half-transporters that oligomerise to form the functional transporter [62]. Rains et al. [13], recently reported up-regulated ABCG transporter genes in poplar phellem, namely homologs of *AtABCG11* and *AtABCG2*. In *Arabidopsis*, the ABCG transporter ABCG11 is implicated with cutin and suberin biosynthesis [63–65] while ABCG12 and ABCG13 are involved in export of cuticular wax monomers [65, 66]. Because homo- and heterodimers of ABCG11 and ABCG12 are reported as important for wax export and ABCG13 seems to be required for cutin deposition in flowers [67], it is widely accepted that tissue-dependent combinations of different half-size ABCG proteins involved in the transport of wax, cutin and suberin components are possible to occur [61]. Furthermore, ABCG11 can form dimers with ABCG9 and ABCG14 which links it to a function in lipid homeostasis regulation in vascular organs as phloem [68]. Additionally, a recent study reported that

Capote *et al. BMC Plant Biology* (2018) 18:198

Page 11 of 19

Fig. 8 Enzymes target by QsMYB1 in the glycerophospholipid metabolism (colored boxes). Numbers correspond to enzymes E.C. numbers

half-size ABCG2, ABCG6 and ABCG20 are involved in the formation of suberin layers in seed coats and roots where *abcg2/6/20* triple mutants had increased permeability, altered suberin structure and reduce aliphatic components [69]. Another type of proteins involved in lipid excretion into the extracellular space are LTPs. Evidence of LTPs involvement in cuticular wax, suberin and sporopollenin assembly is increasing in recent literature [70, 71]. Our data also identified five genes coding for LTPs reinforcing the involvement of QsMYB1 in the regulation of genes involved in lipid traffic across the cellular membrane and possibly in suberin assembly and deposition.

Putative regulation of glycerolipids metabolism by QsMYB1

Glycerolipids are fatty acid esters of glycerol. Fatty acids are constituents of membrane lipids of every cell. The first step in *the novo* synthesis of fatty acids is the carboxylation of acetyl-CoA to malonyl-CoA by acetyl-CoA carboxylase [72]. In *Arabidopsis* acetyl-CoA carboxylase is required for very long chain fatty acid elongation [73]. Our data suggests that QsMYB1 is regulating a homolog gene of acetyl-CoA carboxylase, which indicates direct regulation on a fundamental key step of fatty acid

synthesis and elongation pathways. The second enzyme responsible for fatty acids biosynthesis is the fatty acid synthase (FAS) which in plants is a type II FAS consisting of a multienzyme complex of acyl carrier proteins [72]. The results also revealed that at least 2 genes with homology for *Arabidopsis* enoyl-[acyl-carrier-protein] reductases (NADH) are QsMYB1 gene targets. One acyl-[acyl-carrier-protein] desaturase was also identified in our data. This type of desaturases are responsible for fatty acid modifications introducing *cis*-double ligation in the acyl chain acyl-[acyl-carrier-protein] and ultimately in the correspondent fatty acid [72]. Therefore, QsMYB1 may regulate the production of unsaturated fatty acids, which in turn is used as precursors for the suberin monomeric units synthesis.

Suberin, in cork oak, presents monoacylglycerols, diglycerol diesteres and diacylglycerols in is chemical composition [40, 46]. The availability of glycerol-derived molecules may constitute a critical limitation step in glycerol-derived monomers synthesis of suberin [59]. In this study, were identified two genes coding for aldehyde dehydrogenases which are responsible for the production of D-glyceraldehyde, which in turn is converted to glycerol by aldehyde reductase and alcohol dehydrogenase
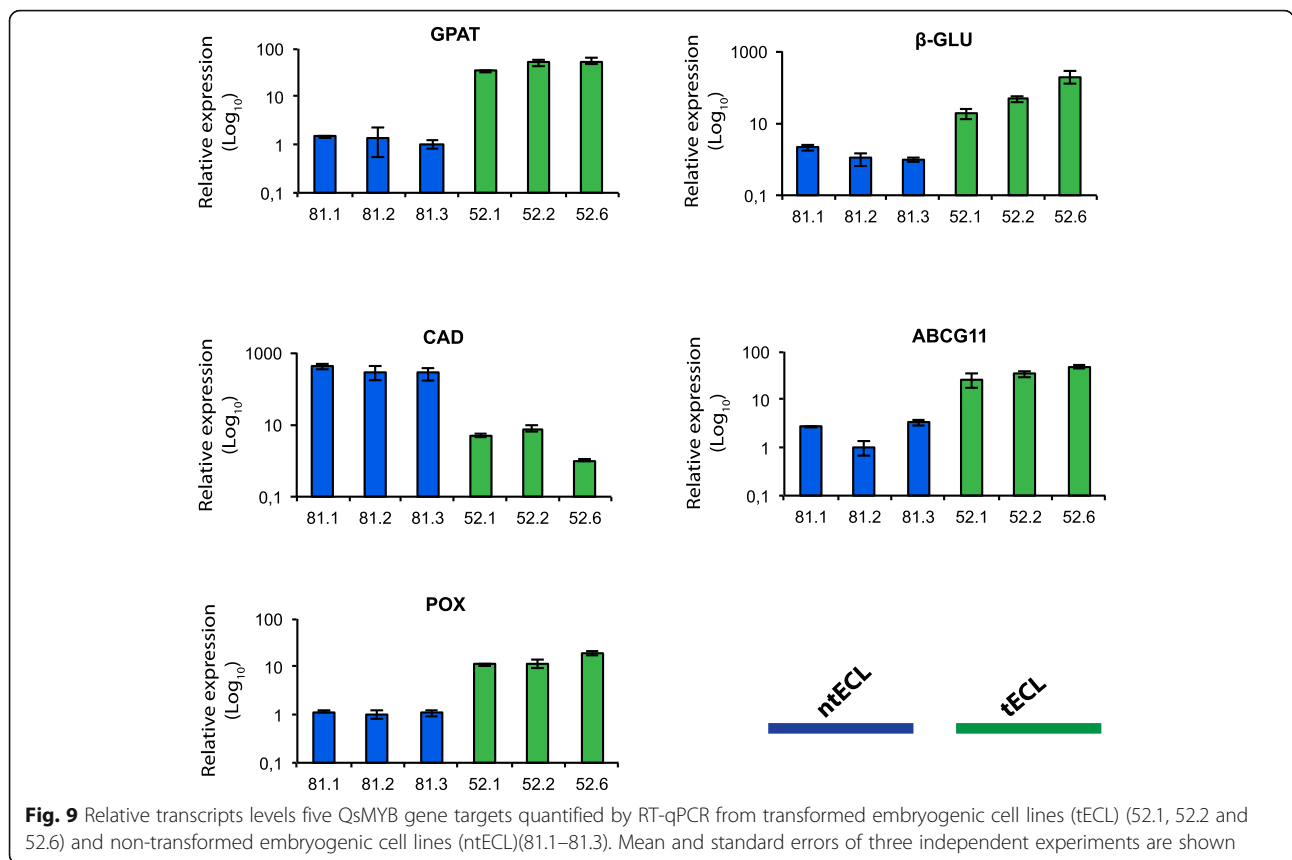
Capote *et al. BMC Plant Biology* (2018) 18:198

Page 12 of 19

Fig. 9 Relative transcripts levels five QsMYB gene targets quantified by RT-qPCR from transformed embryogenic cell lines (tECL) (52.1, 52.2 and 52.6) and non-transformed embryogenic cell lines (ntECL)(81.1–81.3). Mean and standard errors of three independent experiments are shown

(both also identified in our data) to form glycerol in plants [72]. These results indicate that QsMYB1 is acting on transcription regulation of genes involved in the synthesis of glycerol. In addition, the results revealed that QsMYB1 is exerting its action on a gene encoding for diacylglycerol O-acyltransferase which is responsible for the conversion of diacylglycerols to triacylglycerols [74]. Interestingly, QsMYB1 also seems to act on the opposite conversion reaction of triacylglycerols to the correspondent diacylglycerols by targeting the homolog gene of a triacylglycerol acylhydrolase. Triacylglycerol acylhydrolases are also described to have the ability for catalysing the reaction of diacylglycerols to monoacylglycerols [75, 76]. This finding may constitute an evidence for the transcription regulation of diacylglycerols and monoacylglycerols synthesis in *Q. suber*, however further studies have to be performed in order to confirm this hypothesis.

QsMyb1 may be involved in abiotic stress responses through the regulation of the glycerophospholid pathway

Glycerophospholipids are major components of cellular membranes which are synthesized from glycerol-3-phosphate and have functions as second messengers in plant growth regulation and cellular response to environmental change or stress [72]. Lipid membrane composition, remodelling and activation of a variety of

phospholipid-based pathways is a strategy developed by plants to survive and adapt to osmotic stress [77].

Almeida et al. [29] reported a putative function of QsMYB1 in the regulatory network of cork oak response to heat and drought stresses. Accordingly, we found that QsMYB1 is targeting a gene coding for phospholipase A1, a *sn*-1-specific phospholipase that releases free fatty acids from phospholipids and can act on the cellular membrane galactolipids and triacylglycerols [78]. In *Arabidopsis*, the *DEFECTIVE IN ATHER DEHISCENCE1 (DAD1)* gene has been shown to code phospholipase A1 catalysing the initial step of jasmonic acid biosynthesis [79]. The *Arabidopsis* phospholipase coded by the At2g42690 gene, which is UV-B inducible, was also reported as having properties that suggest a possible involvement in the biosynthesis of jasmonic acid [80]. These studies point to involvement of phospholipases A in hormonal response to stress and may indicate evidence of a relation between QsMYB1 and response to stress mediated by hormones. Furthermore, our results revealed that QsMYB1 is targeting a *phosphoethanolamine N-methyltransferase* homolog gene of *Arabidopsis* which catalyses key steps in choline biosynthesis, namely the N-methylation of phosphoethanolamine [81]. Choline is well reported as an important player in plant grow and development [82]. Choline is also oxidized to

Capote *et al. BMC Plant Biology* (2018) 18:198

Page 13 of 19

glycinebetaine which is a strong osmoprotectant that confers tolerance to salinity, drought and other stress [83, 84]. In *Arabidopsis*, the silencing *of phosphoethanolamine N-methyltransferase* results in temperature-sensitive male sterility and salt hypersensitivity, which demonstrates that choline biosynthesis also contributes to tolerance to environmental stress [85]. Our results suggest an involvement of QsMYB1 response to drought stress specifically involving a key choline biosynthesis enzyme supporting the hypothesis that QsMYB1 expression is modulated in response to heat and drought stresses [29].

QsMYB1 targets other transcriptional elements expanding its regulatory action

Three motifs ([CWHCAA], [CYTCBTC] and [BKTGG] were identified as the principal motifs characteristic of the QsMYB1-DNA binding. The [CWHCAA] and [BKTGG] motifs present high homology with already described R2R3-MYB binding site motifs [31], while [CYTCBTC] has a similarity with zinc finger C2H2-type motif of TF IIIA, which in *Arabidopsis* encodes a protein required for transcription of 5S ribosomal RNA [86]. However, the presence of secondary motifs signatures is a constant in the dataset. Secondary motifs result from two events: two TFs binding to neighbouring sites, or one TF binding to another TF that in turn binds to DNA. Taking into account that members of the MYB family are well known by acting as homo- or heterodimers [87], the data suggest that QsMYB1 is either binding to DNA or to other TFs resulting in an extensive and complex transcriptional regulation of genes. Therefore, QsMYB1 may be involved in a complex regulatory network targeting both transcriptional elements and a group of specific genes with direct function in several biological processes. We found that QsMYB1 is targeting other transcription elements, which supports the idea that QsMYB1 directly regulates the expression of genes and also exerts its effect through other TFs, transcription regulators and chromatin regulators. Amongst these, the more abundant gene targets of QsMYB1 are involved: in fine tuning transcription, translation and posttranscriptional modifications as CCHC zinger finger proteins (CCHC) family, which specifically interacts with single-stranded DNA or RNA oligonucleotides [88] or C2H2-type zinc fingers TFs (C2H2) which besides DNA binding also provides protein-protein and RNA-protein interactions [89] or BED-type zinc finger proteins, which function as either transcription activators or repressors by modifying local chromatin structure on binding to GC-rich sequences [90, 91]; in regulation of cell division, cell-fate determination, transmembranar signaling and vesicle fusion as the WD40-like TFs; and in integrating gene regulatory networks which controls the metabolic,

hormonal and environmental signals in plant growth, development and response to plant abiotic and biotic stresses as the AP2/ERF TF [92, 93] and the bHLH TF families [87].

Together with MYB, bHLH represents a large proportion of the TFs in *Arabidopsis* both regulating multiple biological process [94]. Members of the two families are described by acting as homodimers or heterodimers consisting of proteins from the same family or in complexes with TFs from other families including the WD40-like TFs previously referred [94]. In many cases, MYB/bHLH complexes have been described and associated with the developmental and metabolic plasticity present in plants [94]. The regulation of pathways controlled by these two types of TFs as well as the mechanism regulating their activity is therefore linked, which is in line with our results. Thus, by forming heterodimers with other transcriptional elements, QsMYB1 may trigger additional regulatory mechanisms at chromatin, transcription and post-transcriptional level.

Conclusions

This study has contributed to increase the knowledge about the molecular mechanisms behind cork formation and differentiation. To address the role of QsMYB1 we performed a genome-wide analysis of the DNA targets of this TF using ChIP-Seq. QsMYB1 was shown to be involved in lignin and suberin biosynthesis and assembly, acting through a complex regulatory network directly by targeting enzyme genes, and indirectly by targeting genes encoding for other transcriptional regulatory elements. The complex QsMYB1 regulatory mechanisms that we disclosed may represent an example of how *Q. suber* has optimised its developmental regulatory systems during evolution. Furthermore, we found QsMYB1 as a master regulatory factor of cork formation and differentiation since it acts on genes belonging to the lignin and suberin biosynthetic pathways, the two main components of cork. In addition, QsMYB1 has the capability to modulate the process of lignin and suberin synthesis through the regulation of specific genes of the phenylpropanoid pathway. Further studies are required to confirm and reveal the specific role of the reported genes and to clarify the exact function of QsMYB1 in the process. Even so, the generated data is a starting point to an in-depth understanding of suberin biosynthesis and deposition. By starting to unravel the biosynthetic pathways involved in cork formation we provide important knowledge to support the development of breeding programs based e.g., on the design of cork metabolic engineering strategies.

Capote *et al. BMC Plant Biology* (2018) 18:198

Page 14 of 19

Methods

Plant material and growth conditions

No specific permission was needed for the development of this study. *Quercus suber* is protected in Portugal against logging. Sample collections did not involve endangered or other protected species.

A stable *Q. suber* somatic embryogenesis line, Ce1, was obtained by somatic embryogenesis induction from adult tree leaves as described by Hernandéz et al. [95].

Trees were located at Cercal do Alentejo, Portugal, growing in a *montado* ecosystem. Embryo clusters were grown under long day conditions (16 h light per 8 h dark cycle) at 25 °C on MSSH medium, containing Murashige and Skoog (MS) micronutrients [96], Schenk and Hildebrandt (SH) macronutrients [97] and supplemented with MS vitamins [96], 3% (*w/v*) sucrose and 1% (w/v) agar.

Quercus suber genetic modified cell lines

A binary vector for overexpression of QsMYB1::triple FLAG epitope fusion protein was constructed using Gateway™ technology (Gateway® Recombination Cloning, Invitrogen, Lifetechnologies Corporation, CA, USA). QsMYB1 coding sequence (accession number JF970262) was amplified with primers MYB1ns_attB1F and MYB1ns_attB2R (Additional file 6: Table S1) using attB1 and attB2 adapters and recombined with the pDONR221 vector to generate a QsMYB1 ENTRY vector (pENTRY_QsMYB1). Triple FLAG epitope (3xFLAG) nucleotide sequence was generated by annealing the 3xFLAG_F_Olig and the 3xFLAG_R_Olig oligonucleotides (Additional file 6: Table S1) and recombined with the pDONR221 vector in order to produce a 3xFLAG ENTRY vector (pENTRY_3xFLAG). A PCR-fusion/Gateway procedure for gene fusion [98] was applied to generate an overexpression pK7WG2,0 vector carrying the QsMYB1::3xFLAG fused coding sequence (pK7MYB1::3xFLAG). A diagram of the procedure is showed in (Additional file 6: Figure S1). The resulting binary vector was mobilized into *A. tumefaciens* AGL1 by triparental mating [99] using *E. coli* MC1012 harbouring the mobilizing plasmid pRK2013. After a 10 days sub-culture period, embryos were transformed as described by Álvarez and Ordás (2007) [100].

RT-qPCR validation of QsMYB1::3xFLAG expression

Genetic transformed embryo clusters were grown by repetitive embryogenesis in kanamycin selective medium during 24 months with regular sub-culturing every 30 days, to allow for transformed embryos selection. After 24 months of embryo transformation, aleatory independent embryo clusters were selected for RT-qPCR analysis in order to screen the expression of the QsMYB1::3xFLAG transcript. Specific primers qPCRQsMYB1F and SR(C) (Additional file 6: Table S2) were used in the RT-qPCR experiments.

One microgram of total RNA was used for cDNA synthesis using the QuantiTect Reverse Transcription kit (Qiagen, Valencia, CA, USA), which includes an additional genomic DNA elimination step and uses a mix of oligo(dT) and random hexamer primers. RT-qPCR experiments were carried out in a iCycler iQ5 Instrument (Bio-Rad Laboratories, Hercules, CA, USA) using the Sso Advanced Universal SYBR Green Master mix (Bio-Rad Laboratories, Hercules, CA, USA) in 96-well plates. Three replicates were performed in reaction mixtures of 20 μL containing 10 μL of 2X Sso Advanced Universal SYBR Green Master mix, 400 nM of each specific primer pair (Forward/Reverse) and 1 μL of cDNA with a dilution of 1:100 as template. Normalization between samples was performed using two reference genes: ACTIN (ACT) and CLATHRIN ADAPTOR COMPLEXES (CACs) shown to be constitutively expressed (data not shown). Normalized relative quantities (NRQ) were calculated by $NRQ = \dfrac{E_{goi}^{\Delta Ct,goi}}{\sqrt[f]{\prod_o^f E_{ref_o}^{\Delta Ct,\ ref_o}}}$ where E is the amplification efficiency for each primer pair, f the number of reference genes used to normalize the data, goi the gene of interest, ref the reference gene and ΔCt is the Ct of the sample with higher Ct across samples minus the Ct value of the sample in test [101].

RT-qPCR of QsMYB1 target genes

Three random transformed embryo clusters overexpressing QsMYB1::3xFLAG and three random non-transformed embryo clusters were selected for assessment of the relative expression of 5 QsMYB1 putative target genes: *QsGPAT, Qsβ-GLU, QsCAD, QsABCG11* and *QsPOX* genes. Specific primers for each gene (Additional file 7) were used in the RT-qPCR experiments. RNA and cDNA were prepared as for RT-qPCR validation of QsMYB1::3xFLAG expression. RT-qPCR experiment were performed by the same described conditions, using as reference genes ELONGATION FACTOR 1α (EF1-α) and β-TUBILIN (β-TUB). Normalized relative quantities (NRQ) were calculated as described in the previous section.

Detection of QsMYB1::3xFLAG fused protein in *Q. suber* embryos

Expression of full-length QsMYB1 protein fused with 3xFLAG tag was confirmed by Western Blot. Total protein content of genetic modified and non-genetic modified embryos was extracted using TCA-acetone method [102]. Supernatants were collected, and total protein concentrations quantified according to Lowry method [103], using BSA as protein standard. Protein components of cell lysates (25–40 μg protein) were separated

Capote *et al. BMC Plant Biology* (2018) 18:198

Page 15 of 19

on sodium dodecyl sulfate 10% polyacrylamide gel, and then transferred onto polyvinylidene difluoride (PVDF) membranes (Amersham Biosciences, Buckinghamshire, UK). PVDF membranes were blocked with 5% (*w/v*) of non-fat dry milk at room temperature for 1 h, and incubated overnight at 4 °C with a mouse monoclonal ANTI-FLAG® M2 (Sigma-Aldrich, St Louis, MO) primary antibody (1:2000). Bands were visualized by chemiluminescence using anti-mouse horseradish peroxidase-conjugated secondary antibodies, and developed with ECL reagents (Amersham Biosciences, Buckinghamshire, UK), according to manufacturer's instructions.

Immunodetection

Quercus suber embryos were fixed in a freshly prepared solution of 4% paraformaldehyde in PBS (phosphate buffered saline: 137 mM NaCl; 0.27 mM KCl; 1 mM phosphate buffer, pH 7.4). The fixed embryos were sectioned (10- to 20-μm-thick) under water using a vibratome. The sections were kept at 4 °C in a 0.1% solution of formaldehyde in PBS until further use. Subsequently, sections were incubated with blocking solution (5% (w/v) Bovine Serum Albumin (BSA) in 1× PBST – PBS supplemented with 0.5% (*v/v*) Tween 20), and after washing in 1xPBST were incubated with a mouse monoclonal ANTI-FLAG® M2 (Sigma-Aldrich, St Louis, MO) primary antibody (1:200) supplemented with 1% BSA for 16 h at 4 °C. The primary antibody was detected with a goat polyclonal secondary antibody to mouse - Alexa Fluor® 488 IgG - H&L (ab150113, Abcam Cambridge, UK) (1:200 dilution in PBST supplemented with 1% BSA) for 1 h at 37 °C. Sections were mounted with Vectashield mounting medium with DAPI (Vector Laboratories, USA). Images were acquired on an epifluorescence microscope Axio Imager.Z1.

ChIP-Seq assay

ChIP-Seq assay of *Q. suber* embryos expressing QsMYB1::3xFLAG was performed by cross-linking and isolating chromatin embryos using the Abcam Plant Chromatin Extraction Kit (ab156906, Abcam, Cambridge, UK) protocol, with slight modifications to the cross-linking procedure: starting with 3 g of fresh embryos, 1 volume of Plant RNA Isolation Aid (Thermo Fisher Scientific, Waltham, MA, USA) per unit mass of embryo tissue (mL/mg) was added to 1% formaldehyde solution and vacuum infiltration step was performed during 20 min at 4 °C. Chromatin was sheared with a Bioruptor® Plus (Diagenode): high intensity pulses, 5 rounds of 10 cycles, 30 s ON / 30 s OFF, with a sample volume of 150 μL to generate ~ 300-bp average fragment sizes, as determined by agarose gel and capillary electrophoresis (Agilent 2100 Bioanalyzer) using the High Sensitivity DNA ChIP kit (Agilent, 5067–4626).

ChIP was performed followed by the protocol developed by Li and collaborators [104] using mouse monoclonal ANTI-FLAG® M2 antibody (Sigma-Aldrich, St Louis, MO) and mouse IgG serum (Sigma-Aldrich, St Louis, MO), generating ChIP and Mock samples, respectively. Mock samples were generated in the same experimental conditions, except the immunoprecipitation step that was performed with IgG serum instead of the ANTI-FLAG® M2 specific antibody. After reverse cross-link and DNA purification, the precipitated DNA was quantified and analysed by capillary electrophoresis (Agilent 2100 Bioanalyzer) using the High Sensitivity DNA ChIP kit (Agilent, 5067–4626).

Library preparation and Illumina sequencing

DNA from ChIP of 3xFLAG (ChIP) and from ChIP of IgG serum (Mock) was used to produce sequencing libraries using the KAPA Hyper Prep Kit (KAPA Biosystems, Wilmington, MA). A total of four single-end libraries were prepared, one for each of the two replicates of the ChIP experiment and two of the control (Mock), and sequenced using the HiSeq 4000 system (Illumina, San Diego, CA), in two independent sequencing lanes (L1 and L5), with a read length of 50 bp.

Sequencing data pre-processing and read mapping

FastQC [105] was used as the first data control tool to check the quality of the sequencing reads. Raw reads were preprocessed with Sickle [106] using 20 for the quality threshold and 40 for the read length. Mapping of the reads against a draft version of the cork oak genome (Additional file 8) was performed with Bowtie2 [107] and unique mapped reads were extracted using the mapping quality value of 255. Library complexity was measured by the non-redundant fraction (NRF) and PCR Bottlenecking Coefficients (PBC). NRF was calculated as described by Nakato and Shirahige [108] by setting the threshold for non-redundant reads to 2. Non-redundant reads were extracted using the output of MACS2 [109] filterdup utility. PBC metrics were obtained using BEDtools [110] genomecov. Phantompeakqualtools [4] was applied for the cross-correlation analysis of the ChIP and Mock libraries.

Peak calling and downstream processing

An irreproducible discovery rate (IDR) analysis was employed to find the most consistent set of peaks across replicates. Self-pseudoreplicates and pooled-pseudoreplicates were also generated and peaks were called with Model Based Analysis for ChIP-Seq data (MACS2) [109], using a q-value < 0.05, based on a Poisson distribution comparing the ChIP and Mock sequenced samples. A maximum number of 8 duplicates tags were defined, as it seemed the appropriate value for which the number of peaks reached a plateau phase using

the fragment length estimated by Phantompeakqualtools (Additional file 9). IDR v2.0.2 [111] was used to measure the reproducibility of replicates ranking the peaks by their *p*-value with an IDR threshold of 0.05. Peaks passing the IDR threshold by comparing true replicates were selected for downstream analysis according to the ENCODE (Encyclopedia of DNA elements) ChIP-Seq guidelines [4].

Based on the structural annotation generated for the genome draft (Additional file 8), custom python scripts were created to assign a gene to each IDR peak and its respective location regarding gene boundaries. Promotor regions were defined to be up to 2000 bp upstream of the beginning of each gene and terminator regions were set to be up to 1000 bp downstream of the end of each gene.

Analysis of targets related with transcriptional elements
Identification of target genes encoding for TFs, transcription regulators and chromatin regulators were analyzed with the PlantTFcat tool [32] submitting the gene sequences correspondent of each peak to the web platform analysis.

Motif analysis
A motif discovery analysis was performed using MEME-ChIP software [112]. MEME-ChIP analysis was done with the gene models available for the cork oak draft genome (Additional file 8) and selected based on the location of detected peaks. Sequences with 250 bp from both sides of peaks summits were retrieved and these 501 bp sequences were given as input in MEME-ChIP software to identify common motifs. Once motifs were identified, the motifs comparison tool TOMTOM [113] was used to compare with already identified and annotated motifs as the Arabidopsis motifs detected in protein microarray (PBM) [31] or Arabidopsis motifs identified by DNA affinity purification sequencing (DAP-Seq) [30].

Metabolic pathway analysis
Genomic sequences with 500 bp from both sides of peak summits were checked against the NCBI non-redundant protein (Nr) database, annotated with gene ontology plant-terms, enriched and refined using ANNEX function and used for metabolic pathway analysis using through the Blast2GO software [114]. Metabolic pathway analysis was then performed using the Blast2GO interface to access Kyoto Encyclopedia of Genes and Genomes (KEGG) database [115]. Pathway maps were downloaded and inspected manually focusing in the identified enzymes of pathways related with suberin and lignin metabolism, namely phenylpropanoid biosynthesis, fatty acid biosynthesis and degradation, glycerolipid metabolism and glycerophospholipid metabolism.

Additional files

Additional file 1: Fragmented DNA. DNA fragments after chromatin fragmentation visualized and quantified by agarose gel (A) and Bioanalyzer (B). (PDF 2085 kb)

Additional file 2: Types of transcriptional regulators. Families and number of transcriptional regulators targeted by QsMYB1. (XLSX 52 kb)

Additional file 3: Enzyme peaks. Detected peaks and annotation related with enzyme coding genes. (XLSX 15 kb)

Additional file 4: ABCG peaks. Detected peaks and annotation related with ABCG genes. (XLSX 11 kb)

Additional file 5: LTPs peaks. Detected peaks and annotation related with LTPs genes. (XLSX 40 kb)

Additional file 6: Gene construction strategy and primers. **Figure S1.** Gene construction strategy used for QsMYB::triple FLAG epitope fusion protein production. **Table S1.** Primers used to generate the overexpression destination vector pK7MYB1::3xFLAG. **Table S2.** Primers used to confirm the integration of the foreign DNA delivered by the destination vector plasmid and to quantify the QsMYB1::3xFLAG transcript by RT-qPCR. (DOCX 101 kb)

Additional file 7: List of primers used in RT-qPCR experiments of *QsGPAT, Qsβ-GLU, QsCAD, QsABCG11* and *QsPOX* genes. (DOCX 12 kb)

Additional file 8: Cork oak genome. Description of cork oak genome draft generation. (DOCX 26 kb)

Additional file 9: IDR analysis. Description of Irreproducible discovery rate (IDR) analysis. (DOCX 40 kb)

Abbreviations
AP2/ERF: APETALA2/−Ethylene Responsive element binding factor; BED-type (Zn): BED-type zinc finger proteins; bHLH: Basic helix-loop-helix; bZIP: Basic leucine zipper; C2H2: C2H2-type zinc fingers TFs; CCHC: CCHC zinc finger proteins; ChIP: Chromatin immunoprecipitation; ChIP-Seq: Chromatin immunoprecipitation followed by high-throughput DNA sequencing; FAS: Fatty acid synthase; LTPs: Lipid-transfer proteins; MYB: Myeloblastosis; RT-qPCR: Quantitative real time PCR; TF(s): Transcription factor(s); UTR: Untranslated regions

Availability of data and materials
Raw data of ChIP-Seq is available at SRA archive trough BioSample accessions SAMN09531222, SAMN09531223, SAMN09531224, SAMN09531225. All data supporting the findings is contained in the manuscript and its supplementary files.

Authors' contributions
TC designed and performed all the molecular experiments, undertook bioinformatics data analysis, interpreted and discussed results and conceived the original manuscript. TC and RO performed the generation of genetic modified cell lines. TC, PB, AU, and AMR performed bioinformatics analyses. TC and VI performed the immunodetection. SG conceived the study, contributed to securing funding, discussed and interpreted the results, and coordinated the

Capote *et al. BMC Plant Biology* (2018) 18:198

Page 17 of 19

Author details
[1]Centro de Biotecnologia Agrícola e Agro-alimentar do Alentejo (CEBAL) / Instituto Politécnico de Beja (IPBeja), Beja, Portugal. [2]Instituto de Ciências Agrárias e Ambientais Mediterrânicas (ICAAM), Universidade de Évora, Évora, Portugal. [3]Linking Landscape, Environment, Agriculture and Food (LEAF) Instituto Superior de Agronomia, University of Lisbon, Lisboa, Portugal. [4]Departamento BOS, Escuela Politécnica de Mieres, Oviedo University, Oviedo, Spain. [5]Present Address: Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge CB101SA, UK.

References
1. Hong JC. General Aspects of Plant Transcription Factor Families. In: DHB G, editor. Plant Transcr. Factors Evol. Struct. Funct. Asp. Boston: Academic Press; 2016. p. 35–56. Available from https://doi.org/10.1016/B978-0-12-800854-6.00003-8.
2. Funnell APW, Crossley M. Homo- and Heterodimerization in Transcriptional Regulation. In: Matthews JM, editor. Protein Dimerization Oligomerization Biol. New York: Springer New York; 2012. p. 105–21. Available from: https://doi.org/10.1007/978-1-4614-3229-6_7.
3. Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L. MYB transcription factors in Arabidopsis. Trends Plant Sci. 2010;15:573–81.
4. Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, et al. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res. 2012;22:1813–31. Available from: http://genome.cshlp.org/content/22/9/1813.full.
5. Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, et al. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science. 2000;290:2105–10.
6. Yanhui C, Xiaoyuan Y, Kun H, Meihua L, Jigang L, Zhaofeng G, et al. The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. Plant Mol. Biol Netherlands. 2006;60:107–24.
7. Du H, Liang Z, Zhao S, Nan MG, Tran LSP, Lu K, et al. The evolutionary history of R2R3-myb proteins across 50 eukaryotes: new insights into subfamily classification and expansion. Sci Rep. 2015;5:11037.
8. Stracke R, Werber M, Weisshaar B. The R2R3-MYB gene family in Arabidopsis thaliana. Curr Opin Plant Biol. 2001;4:447–56.
9. Ambawat S, Sharma P, Yadav NR, Yadav RC. MYB transcription factor genes as regulators for plant responses: an overview. Physiol Mol Biol Plants. 2013;19:307–21.
10. Gou M, Hou G, Yang H, Zhang X, Cai Y, Kai G, et al. The MYB107 Transcription Factor Positively Regulates Suberin Biosynthesis. Plant Physiol. 2017;173:1045–58. Available from: https://doi.org/10.1104/pp.16.01614.
11. Cao Y, Han Y, Li D, Lin Y, Cai Y. MYB transcription factors in Chinese pear (Pyrus bretschneideri Rehd.): genome-wide identification, classification, and expression profiling during fruit development. Front Plant Sci. 2016;7 Available from: https://doi.org/10.3389/fpls.2016.00577 .
12. Kosma DK, Murmu J, Razeq FM, Santos P, Bourgault R, Molina I, et al. AtMYB41 activates ectopic suberin synthesis and assembly in multiple plant species and cell types. Plant J. 2014;80:216–29. Oxford: BlackWell Publishing. Available from: https://doi.org/10.1111/tpj.12624 .
13. Rains MK, Gardiyehewa de Silva ND, Molina I. Reconstructing the suberin pathway in poplar by chemical and transcriptomic analysis of bark tissues. Tree Physiol. 2018;38:340–61. Available from: https://doi.org/10.1093/treephys/tpx060.
14. Lashbrooke J, Cohen H, Levy-Samocha D, Tzfadia O, Panizel I, Zeisler V, et al. MYB107 and MYB9 homologs regulate Suberin deposition in angiosperms. Plant Cell. 2016;28:2097–116. Available from: https://doi.org/10.1105/tpc.16.00490.
15. Legay S, Guerriero G, Deleruelle A, Lateur M, Evers D, André CM, et al. Apple russeting as seen through the RNA-seq lens: strong alterations in the exocarp cell wall. Plant Mol Biol. 2015;88:21–40. Available from: https://doi.org/10.1007/s11103-015-0303-4.
16. Legay S, Guerriero G, André C, Guignard C, Cocco E, Charton S, et al. MdMyb93 is a regulator of suberin deposition in russeted apple fruit skins. New Phytol. 2016;212:977–91. Available from: https://doi.org/10.1111/nph.14170.
17. Seo PJ, Lee SB, Suh MC, Park M-J, Go YS, Park C-M. The MYB96 transcription factor regulates Cuticular wax biosynthesis under drought conditions in Arabidopsis. Plant Cell. 2011;23:1138–52. Available from: https://doi.org/10.1105/tpc.111.083485.
18. Lee SB, Suh MC. Cuticular wax biosynthesis is up-regulated by the MYB94 transcription factor in arabidopsis. Plant Cell Physiol. 2015;56:48–60. Available from: https://doi.org/10.1093/pcp/pcu142.
19. Oshima Y, Shikata M, Koyama T, Ohtsubo N, Mitsuda N, Ohme-Takagi M. MIXTA-like transcription factors and WAX INDUCER1/SHINE1 coordinately regulate cuticle development in Arabidopsis and Torenia fournieri. Plant Cell. 2013;25:1609–24. Available from: https://doi.org/10.1105/tpc.113.110783.
20. Zhou J, Lee C, Zhong R, Ye Z-H. MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in Arabidopsis. Plant Cell. 2009;21:248–66. Available from: https://doi.org/10.1105/tpc.108.063321.
21. Zhou M, Sun Z, Wang C, Zhang X, Tang Y, Zhu X, et al. Changing a conserved amino acid in R2R3-MYB transcription repressors results in cytoplasmic accumulation and abolishes their repressive activity in Arabidopsis. Plant J. 2015;84:395–403. Available from: https://doi.org/10.1111/tpj.13008.
22. Kamiya T, Borghi M, Wang P, Danku JMC, Kalmbach L, Hosmani PS, et al. The MYB36 transcription factor orchestrates Casparian strip formation. Proc Natl Acad Sci. 2015;112:10533–8. Available from: https://doi.org/10.1073/pnas.1507691112.
23. Stracke R, Jahns O, Keck M, Tohge T, Niehaus K, Fernie AR, et al. Analysis of production of flavonol glycosides-dependent flavonol glycoside accumulation in Arabidopsis thaliana plants reveals MYB11-, MYB12- and MYB111-independent flavonol glycoside accumulation. New Phytol. 2010;188:985–1000. Available from: https://doi.org/10.1111/j.1469-8137.2010.03421.x.
24. Aronson J, Pereira JS, Pausas JG. In: Aronson J, Pereira JS, Pausas JG, editors. Cork oak woodlands on the edge: ecology, adaptive management, and restoration. Washington: Island Press; 2009. Available from: https://doi.org/10.1080/00207230903097578.
25. Bugalho MN, Caldeira MC, Pereira JS, Aronson J, Pausas JG. Mediterranean cork oak savannas require human use to sustain biodiversity and ecosystem services. Front. Ecol. Environ. Ecol Soc Am. 2011;9:278–86. Available from: https://doi.org/10.1890/100084.
26. Graça J, Pereira H. The periderm development in Quercus suber. IAWA J. 2004;25:325–35.
27. Almeida T, Menéndez E, Capote T, Ribeiro T, Santos C, Gonçalves S. Molecular characterization of Quercus suber MYB1, a transcription factor up-regulated in cork tissues. J Plant Physiol. 2013;170:172–8. Available from: https://doi.org/10.1016/j.jplph.2012.08.023.
28. Soler M, Serra O, Molinas M, Huguet G, Fluch S, Figueras M. A genomic approach to suberin biosynthesis and cork differentiation. Plant Physiol. 2007;144:419–31. Available from: https://doi.org/10.1104/pp.106.094227
29. Almeida T, Pinto G, Correia B, Santos C, Gonçalves S. QsMYB1 expression is modulated in response to heat and drought stresses and during plant recovery in Quercus suber. Plant Physiol Biochem. 2013;73:274–81. Available from: https://doi.org/10.1016/j.plaphy.2013.10.007.
30. O'Malley RC, Huang SC, Song L, Lewsey MG, Bartlett A, Nery JR, et al. Cistrome and Epicistrome features shape the regulatory DNA landscape. Cell. 2017;165:1280–92. Elsevier. Available from: https://doi.org/10.1016/j.cell.2016.04.038

Capote *et al. BMC Plant Biology* (2018) 18:198

Page 18 of 19

31. Franco-Zorrilla JM, López-Vidriero I, Carrasco JL, Godoy M, Vera P, Solano R. DNA-binding specificities of plant transcription factors and their potential to define target genes. Proc. Natl. Acad. Sci. U. S. A. 2014;111:2367–72. Available from: https://doi.org/10.1073/pnas.1316278111.

32. Dai X, Sinharoy S, Udvardi M, Zhao PX. PlantTFcat: an online plant transcription factor and transcriptional regulator categorization and analysis tool. BMC Bioinformatics. 2013;14:321. Available from: https://doi.org/10.1186/1471-2105-14-321.

33. Feng C, Andreasson E, Maslak A, Mock HP, Mattsson O, Mundy J. Arabidopsis MYB68 in development and responses to environmental cues. Plant Sci. 2004;167:1099–107.

34. Liu J, Osbourn A, Ma P. MYB transcription factors as regulators of phenylpropanoid metabolism in plants. Mol Plant. 2015;8:689–708.

35. Li C, Wang X, Lu W, Liu R, Tian Q, Sun Y, et al. A poplar R2R3-MYB transcription factor, PtrMYB152, is involved in regulation of lignin biosynthesis during secondary cell wall formation. Plant Cell Tissue Organ Cult. 2014;119:553–63. Available from: https://doi.org/10.1007/s11240-014-0555-8

36. Tian Q, Wang X, Li C, Lu W, Yang L, Jiang Y, et al. Functional characterization of the poplar R2R3-MYB transcription factor PtoMYB216 involved in the regulation of lignin biosynthesis during wood formation. PLoS One. 2013;8:76369.

37. Liu Y-F, Li Q-T, Lu X, Song Q-X, Lam S-M, Zhang W-K, et al. Soybean GmMYB73 promotes lipid accumulation in transgenic plants. BMC Plant Biol. 2014;14:73. Available from: https://doi.org/10.1186/1471-2229-14-73

38. Czemmel S, Heppel SC, Bogs J. R2R3 MYB Transcription factors: key regulators of the flavonoid biosynthetic pathway in grapevine. Protoplasma. 2012;249:109–18.

39. Vanholme R, Demedts B, Morreel K, Ralph J, Boerjan W. Lignin Biosynthesis and Structure. PLANT Physiol. 2010;153:895–905. Available from: https://doi.org/10.1104/pp.110.155119.

40. Graça J, Santos S. Suberin: a biopolyester of plants' skin. Macromol Biosci. 2007;7:128–35. Available from: https://doi.org/10.1002/mabi.200600218

41. Vishwanath SJ, Delude C, Domergue F, Rowland O. Suberin: biosynthesis, regulation, and polymer assembly of a protective extracellular barrier. Plant Cell Rep. 2015;34:573–86.

42. Arrieta-Baez D, Stark RE. Modeling suberization with peroxidase-catalyzed polymerization of hydroxycinnamic acids: cross-coupling and dimerization reactions. Phytochemistry England. 2006;67:743–53.

43. Pollard M, Beisson F, Li Y, Ohlrogge JB. Building lipid barriers: biosynthesis of cutin and suberin. Trends Plant Sci. 2008;13:236–46.

44. Bernards MA. Demystifying suberin. Can J Bot. 2002;80:227–40. NRC Research Press. Available from: https://doi.org/10.1139/b02-017

45. Franke R, Schreiber L. Suberin-a biopolyester forming apoplastic plant interfaces. Curr Opin Plant Biol England. 2007;10:252–9.

46. Graça J, Santos S. Glycerol-derived ester oligomers from cork suberin. Chem Phys Lipids. 2006;144:96–107.

47. Graça J, Pereira H. Methanolysis of bark suberins: analysis of glycerol and acid monomers. Phytochem Anal. 2000;11:45–51.

48. Shockey JM, Fulda MS, Browse JA. Arabidopsis contains nine long-chain acyl-coenzyme a synthetase genes that participate in fatty acid and glycerolipid metabolism. Plant Physiol. 2002;129:1710–22. United States

49. Li-Beisson Y, Shorrosh B, Beisson F, Andersson MX, Arondel V, Bates PD, et al. Acyl-Lipid Metabolism. Arabidopsis Book. Am Soc Plant Biol. 2013;11: e0161. Available from: https://doi.org/10.1199/tab.0161.

50. a a M, Kunst L. Very-long-chain fatty acid biosynthesis is controlled through the expression and specificity of the condensing enzyme. Plant J. 1997;12:121–31.

51. Franke R, Hofer R, Briesen I, Emsermann M, Efremova N, Yephremov A, et al. The DAISY gene from Arabidopsis encodes a fatty acid elongase condensing enzyme involved in the biosynthesis of aliphatic suberin in roots and the chalaza-micropyle region of seeds. Plant J. England. 2009;57:80–95.

52. Lee S-B, Jung S-J, Go Y-S, Kim H-U, Kim J-K, Cho H-J, et al. Two Arabidopsis 3-ketoacyl CoA synthase genes, KCS20 and KCS2/DAISY, are functionally redundant in cuticular wax and root suberin biosynthesis, but differentially controlled by osmotic stress. Plant J. England. 2009;60:462–75.

53. Serra O, Soler M, Hohn C, Franke R, Schreiber L, Prat S, et al. Silencing of StKCS6 in potato periderm leads to reduced chain lengths of suberin and wax compounds and increased peridermal transpiration. J Exp Bot. 2009;60: 697–707. Oxford University Press. Available from: https://doi.org/10.1093/jxb/ern314.

54. Wellesen K, Durst F, Pinot F, Benveniste I, Nettesheim K, Wisman E, et al. Functional analysis of the LACERATA gene of Arabidopsis provides evidence for different roles of fatty acid omega -hydroxylation in development. Proc. Natl. Acad. Sci. U. S. A. 2001;98:9694–9. Available from: https://doi.org/10.1073/pnas.171285998.

55. Höfer R, Briesen I, Beck M, Pinot F, Schreiber L, Franke R. The Arabidopsis cytochrome P450 CYP86A1 encodes a fatty acid ω-hydroxylase involved in suberin monomer biosynthesis. J Exp Bot. 2008;59:2347–60. Oxford University Press. Available from: https://doi.org/10.1093/jxb/ern101 .

56. Li Y, Beisson F, Koo AJK, Molina I, Pollard M, Ohlrogge J. Identification of acyltransferases required for cutin biosynthesis and production of cutin with suberin-like monomers. Proc Natl Acad Sci U S A. 2007;104:18339–44. Available from: https://doi.org/10.1073/pnas.0706984104.

57. Compagnon V, Diehl P, Benveniste I, Meyer D, Schaller H, Schreiber L, et al. CYP86B1 Is Required for Very Long Chain ω-Hydroxyacid and α,ω-Dicarboxylic Acid Synthesis in Root and Seed Suberin Polyester. Plant Physiol. 2009;150:1831–43. American Society of Plant Biologists. Available from: https://doi.org/10.1104/pp.109.141408.

58. Serra O, Soler M, Hohn C, Sauveplane V, Pinot F, Franke R, et al. CYP86A33-Targeted Gene Silencing in Potato Tuber Alters Suberin Composition, Distorts Suberin Lamellae, and Impairs the Periderm's Water Barrier Function. Plant Physiol. 2009;149:1050–60. American Society of Plant Biologists. Available from: https://doi.org/10.1104/pp.108.127183.

59. Beisson F, Li Y, Bonaventure G, Pollard M, Ohlrogge JB. The Acyltransferase GPAT5 Is Required for the Synthesis of Suberin in Seed Coat and Root of Arabidopsis. Plant Cell. 2007;19:351–68. American Society of Plant Biologists. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1820950/

60. Yang W, Simpson JP, Li-Beisson Y, Beisson F, Pollard M, Ohlrogge JB. A Land-Plant-Specific Glycerol-3-Phosphate Acyltransferase Family in Arabidopsis: Substrate Specificity, sn-2 Preference, and Evolution. Plant Physiol. 2012;160:638–52. American Society of Plant Biologists. Available from: https://doi.org/10.1104/pp.112.201996.

61. Li N, Xu C, Li-Beisson Y, Philippar K. Fatty acid and lipid transport in plant cells. Trends plant Sci. 2016;21:145–58. Elsevier Ltd. Available from: https://doi.org/10.1016/j.tplants.2015.10.011

62. McFarlane HE, Shin JJH, Bird DA, Samuels AL. Arabidopsis ABCG transporters, which are required for export of diverse cuticular lipids, dimerize in different combinations. Plant Cell. 2010;22:3066–75. Available from: https://doi.org/10.1105/tpc.110.077974.

63. Bird D, Beisson F, Brigham A, Shin J, Greer S, Jetter R, et al. Characterization of Arabidopsis ABCG11/WBC11, an ATP binding cassette (ABC) transporter that is required for cuticular lipid secretion. Plant J England. 2007;52:485–98.

64. Panikashvili D, Savaldi-Goldstein S, Mandel T, Yifhar T, Franke RB, Hofer R, et al. The Arabidopsis DESPERADO/AtWBC11 transporter is required for cutin and wax secretion. Plant Physiol. 2007;145:1345–60. United States

65. Panikashvili D, Shi JX, Bocobza S, Franke RB, Schreiber L, Aharoni A. The Arabidopsis DSO/ABCG11 transporter affects cutin metabolism in reproductive organs and suberin in roots. Mol Plant. 2010;3:563–75. England

66. Panikashvili D, Aharoni A. ABC-type transporters and cuticle assembly: linking function to polarity in epidermis cells. Plant Signal Behav. 2008;3: 806–9.

67. Panikashvili D, Shi JX, Schreiber L, Aharoni A. The Arabidopsis ABCG13 transporter is required for flower cuticle secretion and patterning of the petal epidermis. New Phytol. 2011;190:113–24. England

68. Le Hir R, Sorin C, Chakraborti D, Moritz T, Schaller H, Tellier F, et al. ABCG9, ABCG11 and ABCG14 ABC transporters are required for vascular development in Arabidopsis. Plant J. 2013;76:811–24. Available from: https://doi.org/10.1111/tpj.12334

69. Yadav V, Molina I, Ranathunge K, Castillo IQ, Rothstein SJ, Reed JW. ABCG Transporters Are Required for Suberin and Pollen Wall Extracellular Barriers in Arabidopsis. Plant Cell. 2014;26:3569–88. Available from: https://doi.org/10.1105/tpc.114.129049.

70. Salminen TA, Blomqvist K, Edqvist J. Lipid transfer proteins: classification, nomenclature, structure, and function. Planta. 2016;244:971–97. Available from: https://doi.org/10.1007/s00425-016-2585-4

71. Deeken R, Saupe S, Klinkenberg J, Riedel M, Leide J, Hedrich R, et al. The nonspecific lipid transfer protein AtLtpI-4 is involved in Suberin formation of Arabidopsis thaliana crown galls. Plant Physiol. 2016;172:1911–27. American Society of Plant Biologists. Available from: https://doi.org/10.1104/pp.16.01486.

72. Heldt H-W, Piechulla B. Lipids are membrane constituents and function as carbon stores. Plant Biochem. 2011:359–98. San Diego: Academic Press.

Capote *et al. BMC Plant Biology* (2018) 18:198

Page 19 of 19

73. Baud S, Guyon V, Kronenberger J, Wuillème S, Miquel M, Caboche M, et al. Multifunctional acetyl-CoA carboxylase 1 is essential for very long chain fatty acid elongation and embryo development in Arabidopsis. Plant J. 2003;33:75–86.

74. Li R, Yu K, Hildebrand DF. DGAT1, DGAT2 and PDAT expression in seeds and other tissues of epoxy and hydroxy fatty acid accumulating plants. Lipids Germany. 2010;45:145–57.

75. Seo YS, Kim EY, Kim JH, Kim WT. Enzymatic characterization of class I DAD1-like acylhydrolase members targeted to chloroplast in Arabidopsis. FEBS Lett. 2009;583:2301–7. England

76. El-Kouhen K, Blangy S, Ortiz E, Gardies A-M, Ferte N, Arondel V. Identification and characterization of a triacylglycerol lipase in Arabidopsis homologous to mammalian acid lipases. FEBS Lett. 2005;579:6067–73. England

77. Munnik T, Meijer HJG. Osmotic stress activates distinct lipid and MAPK signalling pathways in plants. FEBS Lett. 2001;498:172–8.

78. Wang G, Ryu S, Wang X. Plant phospholipases: an overview. Methods Mol Biol. 2012;861:123–37. United States

79. Ellinger D, Stingl N, Kubigsteltig II, Bals T, Juenger M, Pollmann S, et al. DONGLE and DEFECTIVE IN ANTHER DEHISCENCE1 lipases are not essential for wound- and pathogen-induced jasmonate biosynthesis: redundant lipases contribute to jasmonate formation. Plant Physiol. 2010;153:114–27. United States

80. Lo M, Taylor C, Wang L, Nowack L, Wang T-W, Thompson J. Characterization of an ultraviolet B-induced lipase in Arabidopsis. Plant Physiol. 2004;135: 947–58. United States

81. Cruz-Ramirez A, Lopez-Bucio J, Ramirez-Pimentel G, Zurita-Silva A, Sanchez-Calderon L, Ramirez-Chavez E, et al. The xipotl mutant of Arabidopsis reveals a critical role for phospholipid metabolism in root system development and epidermal cell integrity. Plant Cell. 2004;16:2020–34. United States

82. Lin Y-C, Liu Y, Nakamura Y. The choline/ethanolamine kinase family in Arabidopsis: essential role of CEK4 in phospholipid biosynthesis and embryo development. Plant Cell. 2015;27:1497–511. Available from: https://doi.org/10.1105/tpc.15.00207.

83. Waditee R, MNH B, Rai V, Aoki K, Tanaka Y, Hibino T, et al. Genes for direct methylation of glycine provide high levels of glycinebetaine and abiotic-stress tolerance in Synechococcus and Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. 2005;102:1318–23. National Academy of Sciences. Available from: https://doi.org/10.1073/pnas.0409017102.

84. Shirasawa K, Takabe T, Takabe T, Kishitani S. Accumulation of glycinebetaine in rice plants that overexpress choline monooxygenase from spinach and evaluation of their tolerance to abiotic stress. Ann Bot. 2006;98:565–71.

85. Mou Z, Wang X, Fu Z, Dai Y, Han C, Ouyang J, et al. Silencing of Phosphoethanolamine N-methyltransferase results in temperature-sensitive male sterility and salt hypersensitivity in Arabidopsis. Plant cell. 2002;14: 2031–43. Available from: https://doi.org/10.1105/tpc.001701.

86. Layat E, Cotterell S, Vaillant I, Yukawa Y, Tutois S, Tourmente S. Transcript levels, alternative splicing and proteolytic cleavage of TFIIIA control 5S rRNA accumulation during Arabidopsis thaliana development. Plant J. 2012;71:35–44. England

87. Pireyre M, Burow M. Regulation of MYB and bHLH transcription factors: a glance at the protein level. Mol. Plant. 2015;8:378–88. Elsevier Ltd. Available from: https://doi.org/10.1016/j.molp.2014.11.022

88. Brown RS. Zinc finger proteins: getting a grip on RNA. Curr Opin Struct Biol. 2005;15:94–8.

89. Razin SV, Borunova VV, Maksimenko OG, Kantidze OL. Cys2His2 zinc finger protein family: classification, functions, and major members. Biochemistry. (Mosc). 2012;77:217–26. United States

90. Aravind L. The BED finger, a novel DNA-binding domain in chromatin-boundary-element-binding proteins and transposases. Trends Biochem. Sci. 2017;25:421–3. Elsevier. Available from: https://doi.org/10.1016/S0968-0004(00)01620-0.

91. Saghizadeh M, Akhmedov NB, Yamashita CK, Gribanova Y, Theendakara V, Mendoza E, et al. ZBED4, a BED-type zinc-finger protein in the cones of the human retina. Invest Ophthalmol Vis Sci. 2009;50:3580–8. Available from: https://doi.org/10.1167/iovs.08-2751.

92. Dietz K-J, Vogel MO, Viehhauser A. AP2/EREBP transcription factors are part of gene regulatory networks and integrate metabolic, hormonal and environmental signals in stress acclimation and retrograde signalling. Protoplasma Austria. 2010;245:3–14.

93. Chen L, Han J, Deng X, Tan S, Li L, Li L, et al. Expansion and stress responses of AP2/EREBP superfamily in Brachypodium distachyon. Sci. Rep. 2016;6: 21623. England

94. Feller A, Machemer K, Braun EL, Grotewold E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. Plant J. 2011;66:94–116. Blackwell Publishing Ltd. Available from: https://doi.org/10.1111/j.1365-313X.2010.04459.x

95. Hernández I, Celestino C, Alegre J, Toribio M. Vegetative propagation of Quercus suber L. by somatic embryogenesis. II. Plant regeneration from selected cork oak trees. Plant Cell Rep. 2003;21:765–70.

96. Murashige T, Skoog F. A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol Plant. 1962;15:473–97. Blackwell Publishing Ltd. Available from: https://doi.org/10.1111/j.1399-3054.1962.tb08052.x

97. Schenk RU, Hildebrandt AC. Medium and techniques for induction and growth of monocotyledonous and dicotyledonous plant cell cultures. Can J Bot. 1972;50:199–204. NRC Research Press. Available from: https://doi.org/10.1139/b72-026

98. Atanassov II, Atanassov II, Etchells JP, Turner SR. A simple, flexible and efficient PCR-fusion/gateway cloning procedure for gene fusion, site-directed mutagenesis, short sequence insertion and domain deletions and swaps. Plant Methods. 2009;5:14.

99. Ditta G, Stanfield S, Corbin D, Helinski DR. Broad host range DNA cloning system for gram-negative bacteria: construction of a gene bank of rhizobium meliloti. Proc Natl Acad Sci U S A. 1980;77:7347–51.

100. Álvarez R, Ordás RJ. Improved genetic transformation protocol for cork oak (Quercus suber L.). Plant Cell Tissue Organ Cult. 2007;91:45–52.

101. Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biol. 2007;8: R19. Available from: https://doi.org/10.1186/gb-2007-8-2-r19.

102. Méchin V, Damerval C, Zivy M. Total protein extraction with TCA-acetone. Methods Mol Biol. 2007;355:1–8.

103. LOWRY OH, ROSEBROUGH NJ, FARR AL, RANDALL RJ. Protein measurement with the Folin phenol reagent. J Biol Chem. 1951;193:265–75.

104. Li W, Lin Y-C, Li Q, Shi R, Lin C-Y, Chen H, et al. A robust chromatin immunoprecipitation protocol for studying transcription factor-DNA interactions and histone modifications in wood-forming tissue. Nat Protoc. 2014;9:2180–93. Available from: https://doi.org/10.1038/nprot.2014.146.

105. Andrews S. FastQC: A quality control tool for high throughput sequence data. Babraham Bioinforma. 2010. Available from: http://www.bioinformatics.babraham.ac.uk/projects/.

106. Joshi NA FJN. Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files. 2011. Available from: https://github.com/najoshi/sickle

107. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. Nat Methods. 2012;9:357–9. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved. Available from: https://doi.org/10.1038/nmeth.1923

108. Nakato R, Shirahige K. Recent advances in ChIP-seq analysis: from quality management to whole-genome annotation. Brief Bioinform. 2017;18:279–90.

109. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9:R137. BioMed Central. Available from: https://doi.org/10.1186/gb-2008-9-9-r137.

110. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26:841–2. Available from: https://doi.org/10.1093/bioinformatics/btq033.

111. Li Q, Brown JB, Huang H, Bickel PJ. Measuring reproducibility of high-throughput experiments. Ann Appl Stat. 2011;5:1752–79. Institute of Mathematical Statistics. Available from: http://projecteuclid.org/euclid.aoas/1318514284

112. Machanick P, Bailey TL. MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics. 2011;27:1696–7.

113. Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble W. Quantifying similarity between motifs. Genome Biol. 2007;8:R24. Available from: https://doi.org/10.1186/gb-2007-8-2-r24.

114. Conesa A, Gotz S. Blast2GO: a comprehensive suite for functional analysis in plant genomics. Int J Plant Genomics. 2008;2008:619832. United States

115. Kanehisa M, Goto S. Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28:27–30. Available from http://www.genome.jp/kegg/