OakContigDF159.1, a reference library for studying differential gene expression in *Quercus robur* during controlled biotic interactions: use for quantitative transcriptomic profiling of oak roots in ectomycorrhizal symbiosis

Mika T. Tarkka[1]*, Sylvie Herrmann[1,2]*, Tesfaye Wubet[1]*, Lasse Feldhahn[1,3]*, Sabine Recht[1]*, Florence Kurth[1]*, Sarah Mailänder[4], Markus Bönn[1,3], Maren Neef[4], Oguzhan Angay[5,6], Michael Bacht[7], Marcel Graf[8], Hazel Maboreke[9], Frank Fleischmann[5], Thorsten E. E. Grams[6], Liliane Ruess[9], Martin Schädler[2,7], Roland Brandl[7], Stefan Scheu[8], Silvia D. Schrey[4], Ivo Grosse[3] and François Buscot[1,10]

[1]Department of Soil Ecology, UFZ - Helmholtz Centre for Environmental Research, Theodor-Lieser-Str. 4, 06120 Halle/Saale, Germany; [2]Department of Community Ecology, UFZ - Helmholtz Centre for Environmental Research, Theodor-Lieser-Str. 4, 06120 Halle/Saale, Germany; [3]Institute of Computer Science, Martin-Luther University, Von-Seckendorff-Platz 1, 06120, Halle/Saale, Germany; [4]IMIT-Physiological Ecology of Plants, Auf der Morgenstelle 1, 72076, Tübingen, Germany; [5]Section Pathology of Woody Plants, Technische Universität München, Hans-Carl-von-Carlowitz-Platz 2, 85354 Freising, Germany; [6]TEEG: Ecophysiology of Plants, Technische Universität München, Hans-Carl-von-Carlowitz-Platz 2, 85354 Freising, Germany; [7]Animal Ecology, Department of Ecology, Faculty of Biology, Philipps-Universität Marburg, Karl-von-Frisch Str. 8, 35032, Marburg, Germany; [8]J.F. Blumenbach Institute of Zoology and Anthropology, Georg August University Göttingen, Berliner Str. 28, 37073 Göttingen, Germany; [9]Ecology Group, Institute of Biology, Humboldt-Universität zu Berlin, Philippstr. 13, 10115 Berlin, Germany; [10]Institute of Biology, Leipzig University, Johannisallee 21-23, 04103 Leipzig, Germany

Author for correspondence:
Mika Tarkka
Tel: +49 345 5585414
Email: mika.tarkka@ufz.de

Received: *26 February 2013*
Accepted: *2 April 2013*

New Phytologist (2013) **199:** 529–540
doi: 10.1111/nph.12317

Key words: *de novo* assembly, herbivory, multitrophic interactions, mutualism, mycorrhiza, pathogenesis, *Quercus robur*, transcriptome.

Summary

- Oaks (*Quercus* spp.), which are major forest trees in the northern hemisphere, host many biotic interactions, but molecular investigation of these interactions is limited by fragmentary genome data. To date, only 75 oak expressed sequence tags (ESTs) have been characterized in ectomycorrhizal (EM) symbioses.
- We synthesized seven beneficial and detrimental biotic interactions between microorganisms and animals and a clone (DF159) of *Quercus robur*. Sixteen 454 and eight Illumina cDNA libraries from leaves and roots were prepared and merged to establish a reference for RNA-Seq transcriptomic analysis of oak EMs with *Piloderma croceum*.
- Using the Mimicking Intelligent Read Assembly (MIRA) and Trinity assembler, the OakContigDF159.1 hybrid assembly, containing 65 712 contigs with a mean length of 1003 bp, was constructed, giving broad coverage of metabolic pathways. This allowed us to identify 3018 oak contigs that were differentially expressed in EMs, with genes encoding proline-rich cell wall proteins and ethylene signalling-related transcription factors showing up-regulation while auxin and defence-related genes were down-regulated.
- In addition to the first report of remorin expression in EMs, the extensive coverage provided by the study permitted detection of differential regulation within large gene families (nitrogen, phosphorus and sugar transporters, aquaporins). This might indicate specific mechanisms of genome regulation in oak EMs compared with other trees.

Introduction

Oaks (*Quercus* spp.) are key trees in many of the vegetation types found in the temperate and Mediterranean biomes of the Holarctic (Iverson & Prasad, 2001). The oak genus includes species and lineages with specific adaptations to a wide range of climates and habitats (Ellenberg, 2010). For instance, the pedunculate oak, *Quercus robur* L., is widely distributed across Europe in predominantly humid areas, and prefers compact, calcareous and hydromorphic grounds (Levy *et al.*, 1992).

Being long-lived and widely distributed trees, oaks harbour large communities of microorganisms and invertebrates, which interact with their host and with each other (Brändle & Brandl, 2001; Jumpponen & Jones, 2009). Most of the fine roots of oaks form ectomycorrhizas (EMs) with soil fungi, a form of mutualistic symbiosis which facilitates nutrient acquisition (Richard *et al.*, 2005; Herrmann & Buscot, 2007). Oaks are also often infected by a series of parasites which are believed to be partly responsible

*These authors contributed equally to this work.

for the decline of this tree species during recent decades (Thomas *et al.*, 2002). For instance, the pathogen species *Phytophthora ramorum* is the causative agent of sudden oak death in North America and Europe (Grünwald *et al.*, 2012), and infection by the epiphytic pathogenic fungus *Microsphaera alphitoides* leads to a decrease in the total leaf Chl content and net carbon assimilation rate (Brüggemann & Schnitzler, 2001; Hajji *et al.*, 2009). Oaks also host species-rich assemblages of herbivores and mites (Brändle & Brandl, 2001), which may decrease their growth rate and even cause mortality (Marquis & Whelan, 1994).

The genetics of the pedunculate oak and of the closely related sessile oak (*Quercus petraea*) have attracted increased attention during recent years (Barreneche *et al.*, 1998; Ueno *et al.*, 2010; Kremer *et al.*, 2012). These two sympatric oak species have become model systems for comparative analyses of physiological differentiation and speciation in forest trees (Epron & Dreyer, 1993; Abadie *et al.*, 2012). As a first step towards genomic analyses of both pedunculate and sessile oak, Ueno *et al.* (2010) developed a combined *Q. robur* and *Q. petraea* cDNA contig assembly, based on large collections of expressed sequence tags (ESTs). These collections, however, consisted mainly of leaf ESTs from the sessile oak, and included only a limited number of ESTs from oak tissues involved in biotic interactions.

Based on an experimental system using genetically identical microcuttings from pedunculate oak clone DF159 (Herrmann *et al.*, 1998), the joint experimental platform TrophinOak, 'Multitrophic Interactions with Oaks', has recently been established in order to study interactions among *Q. robur*, microorganisms and invertebrates in a soil-based culture system under controlled conditions (www.trophinoak.de). Seven representative interacting organisms are part of the platform and were used in the experiments presented in this paper (see Table 1):

- The ectomycorrhizal fungus *Piloderma croceum* J. Erikss. & Hjortst. strain 729 (DSM-4924) was selected for mycorrhizal syntheses. Mycorrhizal interaction between *Q. robur* and *P. croceum* has been intensively studied (Krüger *et al.*, 2004; Herrmann & Buscot, 2007).
- Formation of mycorrhiza is promoted by mycorrhization helper bacteria, and the strain *Streptomyces* sp. AcH 505, which promotes ectomycorrhiza formation and root branching (Maier *et al.*, 2004; Schrey *et al.*, 2005), was selected.
- Leaves of oak seedlings are particularly vulnerable to powdery mildew infections (Edwards & Ayres, 1981), and *Microsphaera alphitoides* (syn. *Erysiphe alphitoides*), the causal organism of the majority of powdery mildew infections in *Q. robur*, is the representative powdery mildew species in the project.
- The involvement of *Phytophthora quercina* in the decline of oaks in Europe has been well documented in the last two decades (Jung & Blaschke, 1996; Jönsson *et al.*, 2003), and this root pathogen was selected.
- Caterpillars of the phytophagous moth *Lymantria dispar*, which are known to feed preferentially on oaks, were selected for experimentation. Herbivory by *L. dispar* has been related to a shift in carbon allocation towards the below-ground parts of trees (Babst *et al.*, 2008).

Table 1 Treatments of pedunculate DF159 oak (*Quercus robur*) microcuttings with seven different interacting organisms

Sample	Treatment type
No treatment	No inoculation
None applied	
Ectomycorrhizal fungus	Fungal inoculum was mixed with the soil
Piloderma croceum	substrate once, at day 0
Mycorrhization helper	2.5×10^7 spores were applied to the soil
bacterium	twice, at 3 and 4.5 wk
Streptomyces sp. AcH 505	
Leaf pathogen	1.5×10^6 spores were applied to leaves
Microsphaera alphitoides	once, at 4 wk
Leaf herbivore	One caterpillar per plant was applied once,
Lymantria dispar	on the last day before harvest
Root pathogen	1.0×10^6 zoospores per plant were applied
Phytophthora quercina	to the soil once, at 5 wk
Root feeding nematode	$\approx 1.0 \times 10^4$ nematodes per plant were
Pratylenchus penetrans	applied to the soil once, at 5 wk
Rhizosphere consumer	Ninety individuals per plant were applied
Protaphorura armata	to the soil once, at 5 wk

The microcuttings were grown in Petri dish soil microcosms for 6 wk to produce the material for the contig assembly and for 8 wk for the study of differential gene expression in ectomycorrhizas. Day 0 indicates the date on which the oak microcuttings were placed in the soil microcosms.

- The nonspecific plant parasitic nematode *Pratylenchus penetrans*, which produces root lesions in broadleaved trees, was chosen as a representative root feeder (Viggars & Tarjan, 1949; Jaffee *et al.*, 1982).
- Plant rhizospheres are colonized by species of the extremely widespread collembolan genus *Protaphorura* (springtails). The chosen representative, *Protaphorura armata*, lives predominately on plant resources, presumably fine roots or root hairs (Endlweber *et al.*, 2009).

Currently, there is no full genome sequence available for any oak species. Therefore, one key objective of the TrophinOak project is to generate a reference transcriptome library, specific to the pedunculate oak clone DF159, which is comprehensive enough to enable RNA sequencing (RNA-Seq) analyses of all seven biotic interactions under investigation. To meet this objective, we performed a series of 454 sequencing runs on transcripts from roots and leaves of DF159 microcuttings interacting with the seven biotrophic organisms listed, and from noninfected control tissues (Table 1). Particular care was taken to obtain the most diverse possible collections of reads, and for this purpose, normalized cDNA libraries were prepared for 454 pyrosequencing from roots and leaves for each interaction type. In addition, to obtain a high amount of coverage of each transcript, sequences with a read length of 100 bp were obtained from paired-end libraries (average insert size 400 bp) of root and shoot tissues using Illumina sequencing technology. Both types of reads were combined to create a hybrid transcriptome assembly. After evaluating the coverage of this library by *in silico* comparisons with genome-sequenced plant species, the effect of mycorrhiza formation with *P. croceum* on the expression levels of oak genes was quantified by RNA-Seq analysis. Our objective was to gain an in-depth insight into the regulation of gene expression in EM oak roots, by greatly increasing the number of transcripts known to be differentially

expressed. Previous studies on the oak clone DF159 had identified only 51 differentially expressed transcripts in premycorrhizal roots and 75 in EM, using subtractive suppressive hybridization (SSH; Krüger *et al.*, 2004) and macroarrays (Frettinger *et al.*, 2007), respectively.

Bruns & Shefferson (2004) have pointed out that the EM symbiosis habit was acquired independently by diverse plant lineages, and that these independent acquisitions may have relied on parallel gains of morphologies and behaviours in plants and fungi. Whether the genetic background of these changes relies on gains and losses of genes, as has been shown for EM fungi (Plett & Martin, 2011), has not been analysed in plants, and gene diversification and changes in gene expression patterns could matter as well. On this basis, we hypothesized that with the help of a large reference library we might be able to detect that EM formation in oak leads to specific patterns of up- and down-regulation among different members of gene families, and that the plant genes induced in other EM associations may not be induced in oak EMs.

Materials and Methods

The experimental culture system

To obtain a homogeneous soil substrate for the experiments, 3 m^3 of the upper soil were collected from an oak forest stand at the Dölauer Heide close to Halle/Saale, Saxony Anhalt, Germany (51.51016°N, 11.91291°E). The A0 (humus, −10 cm) and A1A2 (organic, −30 cm) horizons were gathered, air-dried, sieved at 5 mm, mixed 1 : 1 (v/v), separated into 500 ml aliquots, and sterilized at 50 kGy by BGS Beta-Gamma-Service (Wiehe, Germany). The soil aliquots were stored at 8°C and their sterility was tested before use by plating on LB agar.

Micropropagation and rooting of the pedunculate oak (*Q. robur* L.) clone DF159 was done according to Herrmann *et al.* (2004), reviewed in Herrmann & Buscot (2008). To ensure the maximum possible production of microcuttings, the plant hormones indole acetic acid and 6-benzylaminopurine were continuously supplied to the cultures. The root part of each microcutting was placed into square Petri dishes (12 × 12 cm^2) filled with γ-sterilized soil, which is an adaptation of the initial cultivation system described in Herrmann *et al.* (1998). Shoots were grown outside the Petri dishes. Seven interacting organisms were introduced to the culture system either at the time of establishment (mycorrhizal fungus) or later; procedures used are listed in Table 1 and detailed in Supporting Information, Methods S1, except for the interaction with *P. croceum*, which was used for quantitative RNA-Seq analysis, and is detailed later in this paper. For all interactions, the oak microcuttings were grown at 23°C with a 16 : 8 h day : night (photosynthetic photon flux density of 180 µmol m^{-2} s^{-1}). After transfer into the Petri dish system, the plants were cultivated for 6 wk, before the tissues were harvested for RNA extraction. While the shoot tissues consisted of a mixture of leaves at different developmental stages (buds, sink and source leaves), the root tissues were exclusively lateral roots. After harvest, tissues were immediately submerged in liquid nitrogen. Material was ground in a mortar with a pestle under liquid nitrogen, divided into aliquots, and stored at −80°C.

Piloderma croceum J. Erikss. & Hjortst. Strain 729 (DSM-4924) was cultivated in Petri dishes on Melin Norkrans Modified by Marx (1969) agar medium supplemented with 0.1% (w/v) casein hydrolysate in darkness at 20°C (Herrmann *et al.*, 1998). Fungal inoculum was produced by inoculating a substrate mixture of vermiculite (675 ml), sphagnum peat (75 ml) and 300 ml Melin Norkrans modified by Marx (1969) liquid medium without carbohydrates and with 1/10 strength for P and N as described in Herrmann *et al.* (1998) with a 2-wk-old liquid fungal culture previously grown in 100 ml glass flasks at 20°C in the dark with shaking at 100 rpm. After 4 wk incubation at 20°C in the dark, the inoculum was used for mycorrhizal synthesis, mixing it 1 : 1 (v/v) with the gamma-sterilized soil. The first yellow mycorrhizal root tips were visible after 5 wk of coculture. Two sets of plants were produced with *P. croceum*. One set was harvested at 6 wk at the onset of EM formation. To obtain a larger amount of EM for quantifying differential gene expression, a second set of plants was harvested 8 wk after inoculation with *P. croceum*.

RNA extractions

Based on preliminary experiments comparing the performance of different RNA extraction methods with oak roots, the MasterPure Plant RNA Purification Kit (Epicentre, Hessisch Oldendorf, Germany) was selected for RNA extractions. Fifty milligrams of leaf or 100 mg of root material were used for each extraction. The extracted RNA was treated with DNase I (Fermentas, St Leon-Rot, Germany), and RNA quantification was carried out using NanoDrop (Thermo Scientific, Passau, Germany) and a Quant-iT RiboGreen RNA Assay Kit (Invitrogen, Darmstadt, Germany). RNA quality was checked on a Nano Chip with a Bioanalyzer 2100 (Agilent, Böblingen, Germany).

Preparation and normalisation of cDNA pools for 454 pyrosequencing and Illumina RNA-Seq

Eight leaf and eight root samples were prepared for 454 pyrosequencing, corresponding to above- and below-ground tissues of plants interacting with each of the seven organisms plus noninoculated control plants; each sample was prepared from four plants. Each of the 16 cDNA samples was prepared from 1 µg total RNA with a SMARTer PCR cDNA Synthesis Kit and amplified with Advantage DNA Polymerase (Clontech, Saint-Germain-en-Laye, France). To reduce the prevalence of high-abundance transcripts and to equalize transcript concentrations in the cDNA samples, the SMARTer amplification products (5 µg) were subjected to TRIMMER cDNA normalization (Evrogen, Heidelberg, Germany). The normalized cDNA pools were then used to prepare 454 sequencing libraries and sequenced in-house by means of a titration run followed by two picotitre plates with eight lanes each on a Roche 454 GS-FLX Titanium platform.

One sample from the total root system and another from leaves were used to produce sequences with a read length of

100 bp from paired-end libraries (average insert size 198 bp), which were sequenced using an Illumina HiSeq 2000 at the Beijing Genomics Institute, Hong Kong, China. In addition, for the transcriptome assembly as well as the transcriptomic study of EM plants, three individually selected samples of EMs and three samples of noninoculated fine roots were used to prepare sequences with a read length of 100 bp from paired-end libraries (average insert size, 400 bp), which were sequenced by Illumina HiSeq 2000 at IGA Technologies, Udine, Italy. The latter two steps resulted in eight Illumina libraries in total.

Read processing and construction of the OakContigDF159.1 hybrid assembly

The 454 reads were screened for primers and adaptors with crossmatch (P. Green, http://bozeman.mbt.washington.edu/phredphrap/phrap.html). The following steps were implemented using custom Java scripts. The 454 reads were masked, and for each read, the longest nonmasked region was extracted. Remaining primer and adaptor artefacts were also eliminated. For both 454 and Illumina reads, poly(A) tails, low complexity and low-quality sequences were removed with SeqClean (http://compbio.dfci.harvard.edu/tgi/software/). Nucleotides with quality score < 20 were removed from the ends of the reads using a custom Java script. Sequences < 50 bp were discarded, as were sequences without paired-end information after preprocessing. In order to minimize the number of contaminating reads, a decontamination procedure was introduced for both the 454 and the Illumina reads, as described (Fig. S1). A hybrid assembly approach was selected to combine 454 and Illumina reads to produce an OakContigDF159.1 reference transcriptome. This process is described in Methods S1 (see the 'Construction of OakContigDF159.1 hybrid assembly' section) and illustrated in Fig. 1.

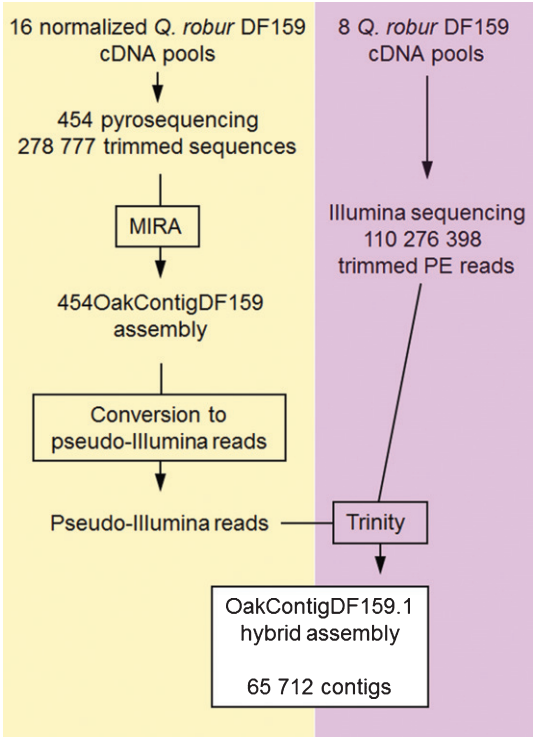

Fig. 1 Pedunculate oak DF159 (*Quercus robur*) hybrid assembly pipeline for Roche 454 and Illumina reads. 454 reads are assembled by Mimicking Intelligent Read Assembly (MIRA) and converted into overlapping 100 bp single-end reads. Single-end and 100 bp paired-end (PE) Illumina reads are assembled by Trinity.

Analysis of differential expression in EMs by Illumina RNA-Seq

Illumina libraries from EMs and from fine roots were used to quantify gene expression. The Illumina reads were aligned against the OakContigDF159.1 hybrid assembly by bowtie (Langmead *et al.*, 2009) and quantified by RSEM (Li & Dewey, 2011) and the significance of differences in gene expression was measured using the DESeq (Anders & Huber, 2010) function of the Bioconductor package (Gentleman *et al.*, 2004) in R (R core group, http://www.r-project.org/). The tools used for transcript annotation and for metabolic pathway analyses, and the quantitative reverse transcription polymerase chain reaction (qRT-PCR) methodology, are described (Table S1).

Results

Generation of a hybrid OakContigDF159.1 reference transcriptome

Root and shoot material from successfully established interactions between oak microcuttings and seven representative organisms, and from control plants, were used to generate a total of 821 534 reads from TRIMMER-normalized cDNA pools using a Roche 454 FLX instrument with Titanium chemistry (Table S2). Most 454 reads were either unique or present in low numbers in the normalized cDNA pools. The 454 reads with homology to genes known to be expressed at a low level were differentially represented in the individual 454 libraries (Fig. S2). For instance, only two cDNA pools included reads homologous to the transcriptional suppressor gene LHP1 of *Arabidopsis thaliana*.

Additional Illumina RNA-Seq of eight cDNA pools, four from roots, three from EMs and one from leaves, allowed a greater depth of sequencing for the pedunculate oak clone DF159 transcriptome assembly. Depending on the sample, the libraries yielded 21–62 mio 100 bp paired-end reads with a Q20 percentage (percentage of sequences with predicted sequencing error rate lower than 1%) of over 93% (Table S3).

Contaminating reads originating from oak-interacting organisms were, as far as possible, eliminated from all sequence libraries by BLASTx searching against reference datasets (see Methods S1 and Fig. S1 for details). A pedunculate oak DF159 reference transcriptome was produced from the decontaminated reads using a combination of overlap layout consensus (OLC) and short read assemblers (Fig. 1). In the first step, the Mimicking Intelligent Read Assembly (MIRA) OLC assembler was implemented to generate contigs from 454 reads. MIRA contigs and singletons (reads which were not incorporated into MIRA

(a)

	MIRA 454	Trinity Illumina	Trinity 454/Illumina
Contigs and single reads	73 151*	57 737	65 712
Amount of nt	36 349 699	58 854 648	65 913 455
Mean length (bp)	496	1019	1003
Number of contigs with complete CDS	20 040	29 755	33 871

*31 607 contigs and 41 544 un-assembled reads (singletons)

(b)

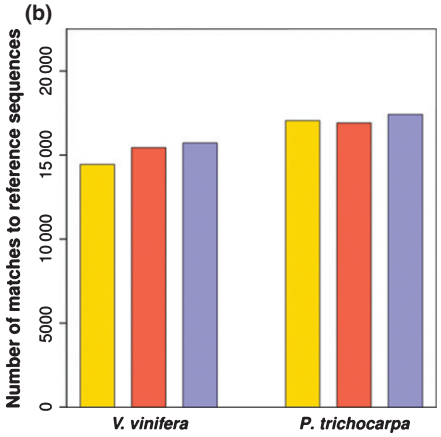

Fig. 2 Characteristics of the assemblies generated by the Mimicking Intelligent Read Assembly (MIRA) and Trinity assembly programs. (a) Basic assembly metrics. Values are shown from MIRA assembly of 454 reads, Trinity assembly of Illumina reads only, and Trinity assembly of Illumina reads and MIRA contigs, as well as unassembled single reads converted into overlapping 100 bp single-end reads. (b) Numbers of BLASTx matches of the contigs against *Vitis vinifera* and *Populus trichocarpa* RefSeq protein databases at an e-value cut-off of 1.0e–20. MIRA 454, yellow bars; Trinity Illumina, red bars; Trinity 454/Illumina, purple bars. CDS, polypeptide coding sequence.

contigs) were then converted into overlapping 100 bp single-end reads and assembled with the Illumina reads using the Trinity short read assembler. The 454/Illumina hybrid assembly generated more contigs, which encoded a larger number of predicted full-length polypeptide coding sequences than the 454 or Illumina read assemblies alone (Fig. 2a). BLASTx searches against *Vitis vinifera* and *Populus trichocarpa* protein indices showed that the numbers of matches to the reference sequences were highest for 454/Illumina hybrid assembly contigs at e-values > 1e–50 and highest for sequences in Illumina-only assemblies at e-values < 1e–50 (Fig. S3). Comparable numbers of matches in the two Trinity assemblies occurred at 1e–50 (Fig. 2b). Cross-comparison of the MIRA and Trinity assemblies by BLASTn with the threshold 1e–50 showed that 71 305 of 73 161 (97%) MIRA contigs and single reads are homologous to Trinity 454/Illumina contigs, and 69 057 (94%) are homologous to Trinity Illumina contigs. On the basis of the slightly higher number of matches to reference sequences, the 454/Illumina hybrid assembly was selected as being the most comprehensive.

The OakContigDF159.1 reference transcriptome comprises 65 712 contigs with a mean length of 1003 bp, totalling 65 913 455 bp. Contig lengths in this transcriptome range from 200 to 15 438 bp. More than 57% of the contigs have a length of over 500 bp and > 36% are over 1000 bp. As expected, the Trinity contigs of the OakContigDF159.1 reference transcriptome show the highest degree of homology with sequences from higher plants (Fig. S4). The contigs were classified using the Gene Ontology (GO) terminology with Blast2GO and a range of diverse functions could be assigned to them (Fig. S5; Table S4). On the basis of the Kyoto Encyclopedia of Genes and Genomes (KEGG) global metabolic pathway annotation, the distributions of metabolic pathway-related accessions in the OakContigDF159.1 assembly and in the *A. thaliana* proteome were highly comparable (Fig. S6). The results of these analyses demonstrated that the OakContigDF159.1 assembly is comprehensive and adequate for the analysis of oak gene expression at the transcriptome level.

Differential oak gene expression induced by mycorrhiza formation

In total, 3018 contigs of the OakContigDF159.1 reference transcriptome were differentially expressed, of which 1399 were up-regulated and 1619 down-regulated in oak EMs with *P. croceum* (Fig. 3). Differential expression levels of 14 contigs were confirmed by qRT-PCR analysis (Fig. 4). On the one hand, GO enrichment analysis using DAVID detected significantly enriched GO terms containing the words ribosome, vacuole, response to stimulus, generation of precursor metabolites and energy, starch metabolic process and transporter activity among genes up-regulated in EMs, and enriched KEGG terms included ribosome and

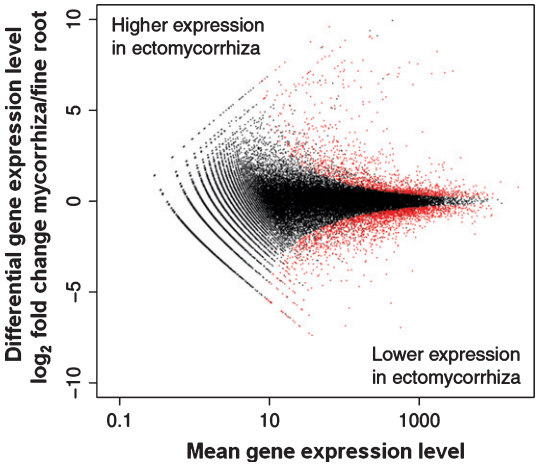

Fig. 3 RNA-Seq based comparison of gene expression levels in fine roots and ectomycorrhizas (EMs) of pedunculate oak DF159 (*Quercus robur*) with *Piloderma croceum*. Raw read counts were generated by quantification using RSEM, and differentially expressed contigs were detected by DESeq. Red dots mark contigs detected as being significantly differentially expressed at a 10% false discovery rate with Benjamini–Hochberg multiple testing adjustments ($P < 0.01$). In EMs, 3018 contigs were differentially expressed, of which 1399 were up-regulated and 1619 were down-regulated.

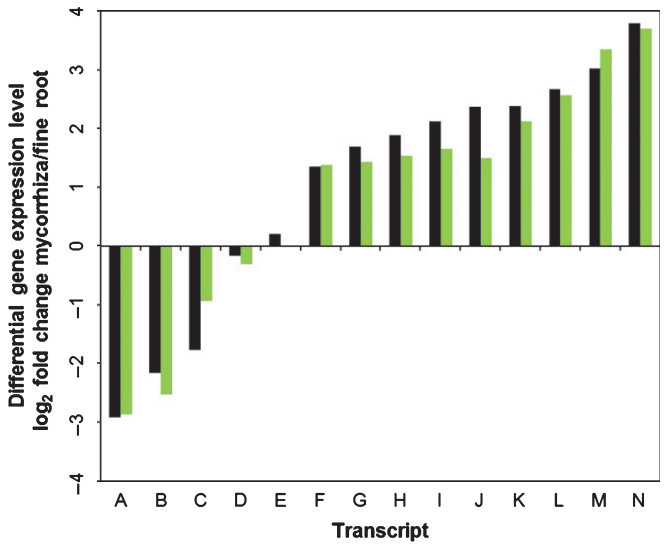

Fig. 4 Real-time quantitative reverse transcription polymerase chain reaction (qRT-PCR) confirmation of 14 differentially expressed genes in ectomycorrhizas (EMs) synthesized between oak microcuttings DF159 and *Piloderma croceum* in comparison to noninfected lateral fine roots. RNA-Seq results (black bars) represent means of three biological replicates. qRT-PCR results (green bars) represent means of three biological and two technical replicates, normalized with respect to an *18S rRNA* gene. The coefficient of variation was < 6.0 for all qRT-PCR reactions. The transcripts analysed were predicted to encode the following proteins by BLASTx searches against the nr database at an e-value cut-off of 1.0e–20: A, extensin; B, sieve element occlusion protein; C, plasma membrane H+-ATPase; D, endo-1,4-beta-glucanase; E, endomembrane transport protein; F, 1-aminocyclopropane-1-carboxylate oxidase; G, glucose-1-phosphate adenylyltransferase; H, nucleoredoxin; I late embryogenesis abundant protein 5; J, proline-rich protein PRP1; K, inositol transporter; L, calcium-binding protein; M, aspartic proteinase; N, galactinol synthase.

spliceosome (Table S5). On the other hand, GO terms that were depleted in EMs included root growth, cytoskeleton, auxin-mediated signalling pathway and auxin polar transport, laccase activity and phenylpropanoid metabolism (Table S5).

Highly significant up-regulation of gene expression was observed for contigs encoding, for example, galactinol synthase, inositol transporter, and remorin (Table 2). Other up-regulated contigs encoded sucrose and SWEET1 sugar transporters. The RNA-Seq analysis also revealed a general up-regulation of contig family members. For instance, seven predicted ethylene response transcription factors, eight predicted proline-rich proteins, and six predicted that aquaporin contigs had higher expression levels in EMs (Table S6). The expression levels of contigs associated with the starch metabolic pathway also increased in EMs (Fig. S7).

Contigs encoding pumilio RNA binding protein and sieve element-occlusion protein were the two most strongly down-regulated in EMs (Table 3). In agreement with the results of GO enrichment analysis, nine auxin-related contigs were down-regulated in EMs (Table S6). Cell wall protein, ammonium and phosphate transporter contig families included contigs that were both up- and down-regulated in EMs (Table S6). Overall, the high resolution of RNA-Seq enabled the identification of numerous EM-regulated genes and the visualization of coregulated contig families.

Discussion

In this study, RNA-Seq enabled the generation of the first specific reference transcriptome for the pedunculate oak clone DF159 under a range of biotic interactions; the study of global transcriptional responses in *P. croceum* ectomycorrhizal roots despite the lack of reference genome sequence or array platform.

Hybrid assembly of 454 and Illumina reads to produce a reference transcriptome

Mimicking Intelligent Read Assembly (MIRA) was chosen for the preassembly of 454 reads from cDNA of leaves and roots of pedunculate oaks involved in seven types of interactions plus a noninfected control, since it proved to be the most robust of the assemblers tested. By contrast, the Illumina reads generated from EMs and noninfected roots and leaves were assembled well by Trinity. Numerous studies suggest that hybrid 454/Illumina assembly is superior in quality to assemblies from 454 or Illumina reads alone (Blythe *et al.*, 2010; Sandmann *et al.*, 2011; Hornett & Wheat, 2012). Following this advice, we constructed a hybrid assembly pipeline for the pedunculate oak reads. The hybrid assembly approach generated more contigs than the Illumina-only assembly, and included sequence information from the majority of MIRA contigs and singletons. Furthermore, the number of unique contigs was noticeably larger in the hybrid assembly than in the Illumina-only assembly. High representation of global KEGG biochemical pathways among the contigs indicates that the OakContigDF159.1 reference transcriptome provides extensive coverage, even though it does not cover the whole-genome sequence.

Differential gene expression in pedunculate oak EMs

Developmental reprogramming has been observed previously in both roots and fungal hyphae upon formation of EMs (Johansson *et al.*, 2004; Duplessis *et al.*, 2005; Martin *et al.*, 2007). However, the authors of these papers noted a much greater magnitude of change in gene expression in the mycelium (up to 20% of the analysed transcripts) than in the root cells (2% of the transcripts). Our RNA-Seq analysis of plant gene expression in mature pedunculate oak EMs found a > twofold change (4.6% of the plant contigs were differentially expressed in EMs at a significance level of $P < 0.01$). In total we found 3018 differentially expressed plant genes in oak EMs, which increases the number identified by previous SSH (Krüger *et al.*, 2004) or macroarray (Frettinger *et al.*, 2007) approaches applied to the same experimental system by a factor of 40. In addition, the quantification was confirmed by qRT-PCR analyses for selected genes. This indicates that the strategy adopted in the present study provided comprehensive and accurate coverage of gene expression changes.

Previous analyses of *Eucalyptus–Pisolithus*, *Betula–Paxillus* and *Quercus–Piloderma* symbioses did not indicate expression of EM-specific plant genes, but showed rather subtle changes in the level of gene expression (Voiblet *et al.*, 2001; Johansson *et al.*, 2004; Duplessis *et al.*, 2005; Le Quéré *et al.*, 2005; Frettinger *et al.*,

Table 2 The 20 most significantly up-regulated contigs in ectomycorrhizas synthesized between oak DF159 (*Quercus robur*) microcuttings and *Piloderma croceum*

Genes up-regulated in mycorrhiza							
Contig no.	Raw read counts in Mycorrhiza RNA-Seq	Raw read counts in fine roots RNA-Seq	*P*-value RNA-Seq	Myc/FR (log2 fold change) RNA-Seq	Myc/FR (log2 fold change) qRT-PCR	Predicted function	Alignment e-value, organism giving the best BLASTx match
43090_0_2	2058.39	147.45	4.60e–182	3.80	3.70***	Galactinol synthase	1e–93, *Populus trichocarpa*
36915_0_2	5973.13	1146.97	4.71e–150	2.38	2.13***	Inositol transporter 1	3e–114, *Glycine max*
35872_0_1	4171.80	1229.97	1.01e–149	1.76		No match	
42280_0_1	3225.65	884.76	1.98e–147	1.86		Hypothetical protein	0, *Populus trichocarpa*
29157_0_2	639.69	13.27	4.76e–139	5.59		Protein phosphatase 2c	1e–65, *Populus trichocarpa*
36374_0_1	3219.85	992.25	6.17e–124	1.69	1.43	Glucose-1-phosphate adenylyltransferase	0, *Populus trichocarpa*
29927_0_1	1204.42	147.21	7.95e–123	3.03	3.35***	Aspartyl protease	6e–174, *Ricinus communis*
550515_0_1	1736.25	11.49	2.50e–108	8.23		No match	
38461_0_5	534.23	18.19	5.85e–106	4.87		Pantothenate kinase 2	0, *Vitis vinifera*
43090_0_1	701.43	44.69	4.27e–101	3.97		Galactinol synthase	4e–157, *Populus trichocarpa*
36915_0_1	2941.45	725.17	4.06e–96	2.02		Inositol transporter 1	0, *Glycine max*
28563_0_1	2521.02	802.43	2.31e–94	1.65		Remorin	3e–76, *Jatropha curcas*
40696_0_2	2352.19	575.33	2.26e–91	1.09		Expansin b1	4e–103, *Ricinus communis*
36836_0_1	422.87	9.68	4.78e–88	5.44		Farnesylated protein	1e–51, *Vitis vinifera*
21193_0_1	2043.99	817.43	9.54e–83	2.00		Lipid binding protein	1e–23, *Ricinus communis*
42096_3_1	730.62	80.61	2.27e–81	3.17		Hypothetical protein	1e–71, *Populus trichocarpa*
32514_0_1	1503.68	402.99	2.47e–75	1.89		Nucleoredoxin 2	2e–174, *Vitis vinifera*
42599_0_1	7358.51	4082.83	1.31e–70	0.84		Granule-bound starch synthase	0, *Prunus persica*
33802_0_1	3493.05	801.10	5.66e–62	2.12	1.66*	Late embryogenesis abundant protein	1e–20, *Citrus sinensis*
33859_0_1	3548.20	1705.6	1.40e–59	1.05		Formate dehydrogenase	0, *Quercus robur*

The contigs most significantly up-regulated according to the test statistic implemented in DESeq are listed. The RNA-Seq-based gene expression levels in mycorrhiza (Myc) and in fine roots (FR) are means of three biological replicates. The mean number of reads that map to the respective contigs is given. *P* values represent the probability of no difference between treatments with Benjamini–Hochberg multiple testing adjustment. Putative gene functions were predicted by BLASTx searching against the nr database. The expected value of the sequence with the best BLASTx hit and its source organism are given in each case. Quantitative reverse transcription polymerase chain reaction (qRT-PCR) results represent means of three biological and two technical replicates, normalized with respect to an *18S rRNA* gene. The coefficient of variation was < 6.0 for all qRT-PCR reactions. Asterisks indicate significant differences according to a randomization test: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

2007). This suggests that the development and metabolism of plant EM tissues are driven by differential regulation of transcriptional regulators, signal transduction, and metabolic pathways, rather than by expression of symbiosis-specific genes (Duplessis *et al.*, 2005; Martin *et al.*, 2007). Our data confirm these findings.

Down-regulation of plant defence-related genes

Early plant response to mycorrhizal fungi involves nonspecific, broad-spectrum defences, including increased chitinase and peroxidase activities during hyphal penetration into the apoplastic space of the root cortex. However, this pattern of overexpression is only transient and it is attenuated in mature EMs (Sauter & Hager, 1989; Albrecht *et al.*, 1994; Münzenberger *et al.*, 1997). In agreement with these observations, we found chitinase and laccase contig families, as well as phenylalanine ammonia lyase contigs, to be down-regulated in oak mycorrhizal roots (Table S6), and GO enrichment analysis identified genes related to phenylpropanoid metabolism as being depleted (Table S5). This confirms that plant defences were attenuated in the mature oak EMs examined here, while roots of oak clone DF159 at the premycorrhizal stage of association with *P. croceum* (Frettinger *et al.*, 2006) overexpressed one class III chitinase. The down-regulation of chitinase that we found in the mature EM confirms the transitory nature of induction of defence-related genes during EM formation on oak.

Plants experiencing abiotic environmental stresses produce elevated concentrations of the phytohormone ABA and generate stress resistance responses through ABA signal transduction. As EM formation attenuates plant stress (Smith & Read, 2008), down-regulation of ABA-induced genes is to be expected in EMs. In accordance with this hypothesis, we detected the down-regulation of two contigs encoding putative ABA receptors in mature oak EMs, confirming the previous analysis of Voiblet *et al.* (2001), who were the first to show the down-regulation of a gene encoding an ABA-induced protein in EMs of eucalyptus.

Enhanced expression of ethylene-related contigs

In the *Quercus–Piloderma* symbiosis, we detected enhanced ethylene signalling (Table S6), which has not been previously reported (Voiblet *et al.*, 2001; Johansson *et al.*, 2004; Duplessis *et al.*, 2005; Frettinger *et al.*, 2007). The gaseous phytohormone ethylene inhibits root elongation and regulates transcription of numerous cell wall-related genes (Sanchez-Rodriguez *et al.*, 2010). When Arabidopsis roots engage in symbiosis with the generalist

Table 3 The 20 most significantly down-regulated contigs in ectomycorrhizas synthesized between oak DF159 (*Quercus robur*) microcuttings and *Piloderma croceum*

Genes down-regulated in mycorrhiza							
Contig no.	Raw read counts in Mycorrhiza RNA-Seq	Raw read c ounts in fine roots RNA-Seq	*P*-value RNA-Seq	Myc/FR (log2 fold change) RNA-Seq	Myc/FR (log2 fold change) qRT-PCR	Predicted function	Alignment e-value, organism with the best BLASTx match
42518_1_2	8.99	1123.58	5.77e–237	−6.96		Pumilio RNA binding protein	0, *Vitis vinifera*
40371_0_1	2289.14	5126.08	4.36e–95	−1.16		Sieve element-occlusion protein	0, *Malus* x domestica
42634_0_1	3033.65	7169.74	1.13e–96	−1.24	−0.66	MDR type ABC transporter	0, *Vitis vinifera*
43602_1_1	5226.82	9598.42	4.15e–90	−0.87		Beta-glucosidase 24	5e–180, *Sorghum bicolor*
32110_0_1	6552.97	11553.25	1.25e–87	−0.81		Sucrose synthase	0, *Manihot esculenta*
39154_0_1	6915.03	11639.13	9.58e–77	−0.75		Ent-kaurenoic acid oxidase	4e–135, *Medicago truncatula*
43332_0_2	5906.27	7538.81	8.46e–68	−0.81		Cytochrome p450	3e–159, *Populus trichocarpa*
43934_0_1	2585.79	4894.94	9.36e–61	−0.92		U-box domain protein 20	2e–162, *Populus trichocarpa*
42460_1_2	392.97	1290.31	1.76e–54	−1.71		Metal transporter	1e–156, *Vitis vinifera*
23289_0_1	653.02	1731.4	2.13e–51	−1.40		Serine threonine protein kinase	0, *Populus trichocarpa*
35773_0_2	318.11	1063.76	2.34e–46	−1.74		Trehalose phosphate synthase	0, *Vitis vinifera*
44296_0_1	563.66	1469.16	2.28e–43	−1.38		Hypothetical protein	3e–164, *Populus trichocarpa*
36775_0_1	287.34	982.41	1.20e–41	−1.77	−0.93	Plasma membrane H+ ATPase	0, *Cucumis sativus*
43667_2_1	201.98	788.65	1.30e–41	−1.96		Hypothetical protein	0, *Populus trichocarpa*
42185_1_1	1885.13	3395.16	1.63e–39	−0.84		ATP binding protein	0, *Ricinus communis*
42363_1_1	4379.53	6831.20	1.65e–39	−0.64		Phenylalanine ammonia-lyase	0, *Quercus suber*
37455_0_1	33.92	348.41	2.48e–39	−3.36		Hypothetical protein	2e–95, *Ricinus communis*
38461_0_1	339.44	1059.08	1.10e–38	−1.64		Pantothenate kinase 2-like	0, *Ricinus communis*
40709_0_2	56.13	489.57	7.93e–37	−3.12		Translation initiation factor eif-4f	1e–152, *Carica papaya*
41450_0_1	1575.09	2892.91	2.43e–36	−0.87		Cytochrome P450	0, *Ricinus communis*

The contigs most significantly down-regulated according to the test statistic implemented in DESeq are listed. The RNA-Seq-based gene expression levels in mycorrhiza (Myc) and in fine roots (FR) are means of three biological replicates. The mean number of reads which map to the respective contigs is given. *P* values represent the probability of no difference between treatments with Benjamini–Hochberg multiple testing adjustment. Putative gene functions were predicted by BLASTx searching against the nr database. The expected value of the sequence with the best BLASTx hit and its source organism are given in each case. Quantitative reverse transcription polymerase chain reaction (qRT-PCR) results represent means of three biological and two technical replicates, normalized with respect to an *18S rRNA* gene. The coefficient of variation was < 6.0 for all RT-qPCR reactions. Asterisks indicate significant differences according to a randomization test: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

endophytic fungus *Piriformospora indica*, ethylene biosynthesis is induced. Moreover, if ethylene signalling is impaired in Arabidopsis, this results in reduced root colonization by the fungus. This suggests that the hormone has a role in symbiotic root colonization by *P. indica* (Khatabi *et al.*, 2012). EM fungi produce ethylene in pure culture, and ethylene production is enhanced in symbiosis with tree roots (Graham & Linderman, 1980). Moreover, a role for ethylene in the dichotomous branching of the short root tips in pines has been established (Kaska *et al.*, 1999). The up-regulation of the ethylene-related transcription factor family in oak EMs indicates that the ethylene signalling may play a role in suppressing root elongation and regulating the morphogenetic program of the symbiotic roots.

Differential expression of auxin-related contigs

Previous research has shown that auxin signalling is central to the regulation of EM root development (Tagu *et al.*, 2002). Herrmann *et al.* (2004) showed that addition of IAA to the *Quercus–Piloderma* culture system stimulates EM formation. In the *Hebeloma cylindrosporum–Pinus pinaster* symbiosis, an auxin-overproducing mutant strain of *H. cylindrosporum* developed EMs with a thicker fungal mantle and multilayered Hartig net (Gea *et al.*, 1994), suggesting that this phytohormone controls EM morphogenesis. Felten *et al.* (2009) observed that before EM development, exudates of *Laccaria bicolor* stimulate lateral root formation in poplar, concomitantly with an up-regulation of multiple auxin-related genes, for example, components of polar auxin transport and auxin signalling. The present study revealed that, in the mature oak EM, auxin signalling genes are differentially expressed (Table S6) and that the expression levels of most putative auxin transporters, and, in particular, of many contigs encoding transcription factors, decreases. As many auxin-related genes are down-regulated in mature oak EMs, this indicates that auxin signalling is central to the early mycorrhizal phase, and less important in regulating processes in the mature symbiotic roots.

Overexpressed remorin contig

One of the most significantly up-regulated contigs showed homology to remorins, and, to our knowledge, this is the first report of EM-induced remorin expression. Remorins act as scaffolding proteins in signalling complexes, and they have crucial functions in plant–microbe interactions. For instance, a remorin protein interacts with symbiotic receptors and regulates bacterial infection in legume root nodule symbiosis (Lefebvre *et al.*, 2010), and induction of a remorin gene takes place upon the establishment of arbuscular mycorrhizal symbiosis (Kistner *et al.*, 2005).

In the present study, only one of the 15 remorin contigs contained in our OakContigDF159.1 library was up-regulated in EMs, suggesting that this member of the family may play a specific role in the oak–*Piloderma* interaction.

Specific up-regulation of proline-rich protein contigs

In EMs of broadleaved trees such as pedunculate oaks, fungal colonization induces dramatic changes in root epidermal cells, which are stimulated to enlarge radially and to loosen their cell wall structure (Peterson & Farquhar, 1994). Our results suggest that a network of plant cell wall proteins, particularly proline-rich proteins (Table S6), participates in the remodelling of cell walls of symbiotic roots. The proline-rich protein (PRP) and extensin subfamilies belong to the ubiquitous plant protein family commonly known as hydroxyproline-rich glycoproteins (Newman & Cooper, 2011). The PRPs have been related to plant development, biotic interactions and environmental stresses (van de Wiel *et al.*, 1990; Newman & Cooper, 2011). Previous analysis performed on the oak clone DF159 detected one PRP transcript which was up-regulated in both premycorrhizal roots and mature EMs (Frettinger *et al.*, 2007). Using the oak contig assembly, the expression pattern of the PRP family in EM oak roots was shown to be tightly regulated, confirming the crucial role played by these proteins in determining the extracellular matrix of EM root cells. Extensins are joined to each other and to cell wall components by cell wall peroxidases (Schnabelrauch *et al.*, 1996), increasing the tensile strength of the primary cell wall (Lamport *et al.*, 2011). Two extensin contigs and several peroxidase-encoding contigs were down-regulated in EMs, indicating that there is reduced potential for cross-linking of cell wall components in EM roots. This hypothesis was supported by the up-regulation of an expansin-encoding contig, since expansins have the capacity to induce extensibility and stress relaxation in plant cell walls (Sanchez-Rodriguez *et al.*, 2010). Cell wall extensibility is further modulated by the xyloglucan endotransglucosylase/hydrolases (XTH) (Sanchez-Rodriguez *et al.*, 2010), but most of the XTH contigs were down-regulated in the oak EMs (Table S6). Overall, the striking and specific up-regulation of the PRP contig family in oak indicates the importance of these proteins in mycorrhiza-related cell wall reorganization.

Expression of genes associated with metabolic pathways

Oak GO terms related to metabolic pathways were altered upon EM formation. The oak data corroborate those from previous analyses of aspen EM (Larsen *et al.*, 2011), as, in both cases, enrichment for GO terms related to starch metabolism and transporter activity was detected. These changes are central to the physiology of EM tissue, as it acts as a strong carbon sink and is the site of intensive sugar and nutrient transport (Nehls, 2008). In the poplar-fly agaric symbiosis, the host plant supplies the fungal partner with hexoses by converting apoplastic sucrose to glucose and fructose by means of plant invertase (Nehls, 2008). In the present study, up-regulation of a sucrose transporter contig was observed, but the invertase encoding contig family was constitutively expressed. Whereas enhanced expression of three monosaccharide transporter genes takes place in poplar EMs (Nehls, 2008), from the oak transcriptome, none of the 12 contigs similar to the poplar monosaccharide transporter genes was up-regulated. This confirms our second hypothesis, that some of the EM-related genes of other systems are not affected in oak. More recently, plant SWEET genes have been shown to be implicated in sugar efflux targeted to plant pathogens and symbionts, and the SWEET1 protein of Arabidopsis expresses glucose transporter activity (Chen *et al.*, 2010). In the present oak EM study, one putative bidirectional glucose transporter of the SWEET1 family was shown to be up-regulated. Although transporter activity has yet to be confirmed for the predicted oak SWEET1 protein, the result could indicate direct export of hexose into the plant apoplast to support the fungus and may suggest the existence of a complementary sugar exchange mechanism in oak EMs.

External EM fungal hyphae transport nutrients, particularly ammonium and phosphorus, to plant roots (Selle *et al.*, 2005; Loth-Pereda *et al.*, 2011). EM formation with *Amanita muscaria* results in up-regulation of three poplar ammonium transporter (AMT) genes (Selle *et al.*, 2005). In oak EMs, one AMT contig was up-regulated and three were down-regulated, indicating a lower induction of plant AMT expression in oak EM than in poplar EM. This result is in accordance with our first hypothesis, which postulates that EM formation in oak leads to specific patterns of up- and down-regulation among the different members of gene families. In line with our first hypothesis, we also observed the up-regulation of one, and down-regulation of two, pedunculate oak phosphate transporter family 1 genes. In poplar, two phosphate transporters of the same family were up-regulated and two were down-regulated in EMs (Loth-Pereda *et al.*, 2011), suggesting that specific EM-related phosphate transporting proteins exist in both systems.

Marjanovic *et al.* (2005) reported that four poplar aquaporin genes encoding members of the plasma membrane intrinsic protein family were up-regulated in the poplar-fly agaric symbiosis, and here, six oak aquaporin contigs of the same family were found to be up-regulated. These proteins are potentially involved in cell turgor regulation in EM tissues.

Philippe *et al.* (2010) observed the induction of poplar galactinol and raffinose synthase contigs and increased raffinose concentrations as a systemic response to herbivory, and suggested that raffinose might be involved in plant responses to biotic interactions. In oak EMs, genes of the raffinose pathway were up-regulated, and a galactinol synthase contig was one of those that were most significantly overexpressed in EMs. However, induction of raffinose during EM formation has not yet been confirmed by metabolite analysis.

Taken together, our data support and confirm the view that instead of a general reprogramming of metabolic networks or transporter families, gene families are precisely regulated to adjust the plant metabolism to mycorrhizal symbiosis. The greatly increased capacity offered by our reference transcriptome for identification of differential gene expression in oak EMs enabled

us not only to identify single genes but also to analyse regulation within whole gene families. This degree of precision enabled us to reveal several traits important for the function of EM symbiosis in oaks (regulation of ethylene or remorin encoding genes), which had not been detected in other host plants investigated to date, such as poplar, eucalypt or birch. In the oak model system, different up- and down-regulation patterns were found in genes and gene families already observed to be involved in EM symbiosis on other host plant models (invertase, transporters of monosaccharides, ammonium and phosphorus, and aquaporins). Confirmation that these traits are really oak-specific, however, requires analysing at a similar depth the gene expression of these host plants when inoculated with *P. croceum*. In addition, elucidation of the pedunculate oak whole-genome sequence and supporting functional analysis will further facilitate comparisons between host responses in different EM systems.

Conclusions

Deep next-generation sequencing was successfully implemented to generate a more complete oak transcriptome. The reference transcriptome of the pedunculate oak clone DF159 thus produced is a valuable addition to previously existing oak genomic resources, including the sessile and pedunculate oak contig transcriptome assembly (Ueno *et al.*, 2010). It is also supporting an ongoing pedunculate oak genome sequencing project (Faivre-Rampant *et al.*, 2011; Kremer *et al.*, 2012), as the reference transcriptome will help in achieving a better understanding of interactions between host and associated organisms, allow development of new reagents sets for 'omic approaches, and assist the experimental annotation of the pedunculate oak genome. Of immediate significance is the ability to use the assembly for RNA-Seq analyses to look at global changes in oak gene expression. Here we have shown the power of this strategy by identifying an extensive transcriptional program associated with EMs on oak roots. In the context of the TrophinOak project, we will use this resource to analyse the responses of the oak clone DF159 to a wide range of beneficial and detrimental interacting organisms, in relation to plant development and under variable environmental conditions.

Acknowledgements

We thank I. Krieg and B. Krause for oak micropropagation, and M. Gawlich, K. Hommel, S. Härtling and B. Schnabel for their assistance in 454 sequencing. Support from the following grants by the German Science Foundation (DFG) is gratefully acknowledged: BU 941/20-1, GR 1881/3-1, RU 780/5-1, SCHA 1242/4-1, SCHE 376/26-1, SCHR 1257/1-1, TA 290/4-1.

References

Abadie P, Roussel G, Dencausse B, Bonnet C, Bertocchi E, Louvet JM, Kremer A, Garnier-Géré P. 2012. Strength, diversity and plasticity of postmating reproductive barriers between two hybridizing oak species (*Quercus robur* L. and *Quercus petraea* (Matt) Liebl.). *Journal of Evolutionary Biology* **25**: 157–173.

Albrecht C, Asselin A, Piché Y, Lapeyrie F. 1994. Chitinase activities are induced in *Eucalyptus globulus* roots by ectomycorrizal or pathogenic fungi during early colonization. *Physiologia Plantarum* **91**: 104–110.

Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biology* **11**: R06.

Babst BA, Ferrieri RA, Thorpe MR, Orians CM. 2008. *Lymantria dispar* herbivory induces rapid changes in carbon transport and partitioning in *Populus nigra*. *Entomologia Experimentalis et Applicata* **128**: 117–125.

Barreneche T, Bodenes C, Lexer C, Trontin JF, Fluch S, Streiff R, Plomion C, Roussel G, Steinkellner H, Burg K *et al.* 1998. A genetic linkage map of *Quercus robur* L. (pedunculate oak) based on RAPD, SCAR, microsatellite, minisatellite, isozyme and 5S rDNA markers. *Theoretical and Applied Genetics* **97**: 1090–1103.

Blythe MJ, Kao D, Malla S, Rowsell J, Wilson R, Evans D, Jowett J, Hall A, Lemay V, Lam S *et al.* 2010. A dual platform approach to transcript discovery for the planarian *Schmidtea mediterranea* to establish RNA-Seq for stem cell and regeneration biology. *PLoS ONE* **5**: e15617.

Brändle M, Brandl R. 2001. Species richness of insects and mites on trees: expanding Southwood. *Journal of Animal Ecology* **70**: 491–504.

Brüggemann N, Schnitzler JP. 2001. Influence of powdery mildew (*Microsphaera alphitoides*) on isoprene biosynthesis and emission of pedunculate oak (*Quercus robur* L.) leaves. *Journal of Applied Botany-Angewandte Botanik* **75**: 91–96.

Bruns TD, Shefferson RP. 2004. Evolutionary studies of ectomycorrhizal fungi: recent advances and future directions. *Canadian Journal of Botany-Revue Canadienne De Botanique* **82**: 1122–1132.

Chen LQ, Hou BH, Lalonde S, Takanaga H, Hartung ML, Qu XQ, Guo WJ, Kim JG, Underwood W, Chaudhuri B *et al.* 2010. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* **468**: 527–532.

Duplessis S, Courty PE, Tagu D, Martin F. 2005. Transcript patterns associated with ectomycorrhiza development in *Eucalyptus globulus* and *Pisolithus microcarpus*. *New Phytologist* **165**: 599–611.

Edwards MC, Ayres PG. 1981. Cell death and cell wall papillae in the resistance of oak species to powdery mildew disease. *New Phytologist* **89**: 411–418.

Ellenberg H. 2010. *Vegetation Mitteleuropas mit den Alpen: in ökologischer, dynamischer und historischer Sicht*. Stuttgart, Germany: Ulmer.

Endlweber K, Ruess L, Scheu S. 2009. Collembola switch diet in presence of plant roots thereby functioning as herbivores. *Soil Biology & Biochemistry* **41**: 1151–1154.

Epron D, Dreyer E. 1993. Long-term effects of drought on photosynthesis of adult oak trees *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L. in a natural stand. *New Phytologist* **125**: 381–389.

Faivre-Rampant PF, Lesur I, Boussardon C, Bitton F, Martin-Magniette ML, Bodenes C, Le Provost G, Berges H, Fluch S, Kremer A *et al.* 2011. Analysis of BAC end sequences in oak, a keystone forest tree species, providing insight into the composition of its genome. *BMC Genomics* **12**: 292.

Felten J, Kohler A, Morin E, Bhalerao RP, Palme K, Martin F, Ditengou FA, Legué V. 2009. The ectomycorrhizal fungus *Laccaria bicolor* stimulates lateral root formation in poplar and *Arabidopsis* through auxin transport and signaling. *Plant Physiology* **151**: 1991–2005.

Frettinger P, Derory J, Herrmann S, Plomion C, Lapeyrie F, Oelmüller R, Martin F, Buscot F. 2007. Transcriptional changes in two types of pre-mycorrhizal roots and in ectomycorrhizas of oak microcuttings inoculated with *Piloderma croceum*. *Planta* **225**: 331–340.

Frettinger P, Herrmann S, Lapeyrie F, Oelmüller R, Buscot F. 2006. Differential expression of two class III chitinases in two types of roots of *Quercus robur* during pre-mycorrhizal interactions with *Piloderma croceum*. *Mycorrhiza* **16**: 219–223.

Gea L, Normand L, Vian B, Gay G. 1994. Structural aspects of ectomycorrhiza of *Pinus pinaster* (Ait.) Sol. formed by an IAA-overproducer mutant of *Hebeloma cylindrosporum* Romagnesi. *New Phytologist* **128**: 659–670.

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge YC, Gentry J *et al.* 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology* **5**: R80.

Graham JH, Linderman RG. 1980. Ethylene production by ectomycorrhizal fungi, *Fusarium oxysporum* f. sp. pini, and by aseptically synthesized

ectomycorrhizae and *Fusarium*-infected Douglas fir roots. *Canadian Journal of Microbiology* **26**: 1340–1357.

Grünwald NJ, Garbelotto M, Goss EM, Heungens K, Prospero S. 2012. Emergence of the sudden oak death pathogen *Phytophthora ramorum*. *Trends in Microbiology* **20**: 131–138.

Hajji M, Dreyer E, Marcais B. 2009. Impact of *Erysiphe alphitoides* on transpiration and photosynthesis in *Quercus robur* leaves. *European Journal of Plant Pathology* **125**: 63–72.

Herrmann S, Buscot F. 2007. Cross talks at the morphogenetic, physiological and gene regulation levels between the mycobiont *Piloderma croceum* and oak microcuttings (*Quercus robur*) during formation of ectomycorrhizas. *Phytochemistry* **68**: 52–67.

Herrmann S, Buscot F. 2008. Why and how using micropropagated trees rather than germinations for controlled synthesis of ectomycorrhizal associations. In: Varma A, ed. *Mycorrhiza: state of the art, genetics and molecular biology, eco-function, biotechnology, eco-physiology, structure and systematics*. Berlin, Germany: Springer, 439–466.

Herrmann S, Munch JC, Buscot F. 1998. A gnotobiotic culture system with oak microcuttings to study specific effects of mycobionts on plant morphology before, and in the early phase of, ectomycorrhiza formation by *Paxillus involutus* and *Piloderma croceum*. *New Phytologist* **138**: 203–212.

Herrmann S, Oelmuller R, Buscot F. 2004. Manipulation of the, onset of ectomycorrhiza formation by indole-3-acetic acid, activated charcoal or relative humidity in the massociation between oak microcuttings and *Piloderma croceum*: influence on plant development and photosynthesis. *Journal of Plant Physiology* **161**: 509–517.

Hornett EA, Wheat CW. 2012. Quantitative RNA-Seq analysis in non-model species: assessing transcriptome assemblies as a scaffold and the utility of evolutionary divergent genomic reference species. *BMC Genomics* **13**: 361.

Iverson L, Prasad A. 2001. Potential changes in tree species richness and forest community types following climate change. *Ecosystems* **4**: 186–199.

Jaffee BA, Abawi GS, Mai WF. 1982. Role of soil microflora and *Pratylenchus penetrans* in an apple replant disease. *Phytopathology* **72**: 247–251.

Johansson T, Le Quéré A, Ahren D, Söderström B, Erlandsson R, Lundeberg J, Uhlen M, Tunlid A. 2004. Transcriptional responses of *Paxillus involutus* and *Betula pendula* during formation of ectomycorrhizal root tissue. *Molecular Plant-Microbe Interactions* **17**: 202–215.

Jönsson U, Jung T, Rosengren U, Nihlgard B, Sonesson K. 2003. Pathogenicity of Swedish isolates of *Phytophthora quercina* to *Quercus robur* in two different soils. *New Phytologist* **158**: 355–364.

Jumpponen A, Jones KL. 2009. Massively parallel 454 sequencing indicates hyperdiverse fungal communities in temperate *Quercus macrocarpa* phyllosphere. *New Phytologist* **184**: 438–448.

Jung T, Blaschke H. 1996. *Phytophthora* root rot in declining forest trees. *Phyton-Annales Rei Botanicae* **36**: 95–101.

Kaska DD, Myllylä R, Cooper JB. 1999. Auxin transport inhibitors act through ethylene to regulate dichotomous branching of lateral root meristems in pine. *New Phytologist* **142**: 49–58.

Khatabi B, Molitor A, Lindermayr C, Pfiffi S, Durner J, von Wettstein D, Kogel KH, Schafer P. 2012. Ethylene supports colonization of plant roots by the mutualistic fungus *Piriformospora indica*. *PLoS ONE* **7**: e35502.

Kistner C, Winzer T, Pitzschke A, Mulder L, Sato S, Kaneko T, Tabata S, Sandal N, Stougaard J, Webb KJ *et al.* 2005. Seven *Lotus japonicus* genes required for transcriptional reprogramming of the root during fungal and bacterial symbiosis. *Plant Cell* **17**: 2217–2229.

Kremer A, Abbott AG, Carlson JE, Manos PS, Plomion C, Sisco P, Staton ME, Ueno S, Vendramin GG. 2012. Genomics of *Fagaceae*. *Tree Genetics & Genomes* **8**: 583–610.

Krüger A, Peskan-Berghofer T, Frettinger P, Herrmann S, Buscot F, Oelmüller R. 2004. Identification of premycorrhiza-related plant genes in the association between *Quercus robur* and *Piloderma croceum*. *New Phytologist* **163**: 149–157.

Lamport DTA, Kieliszewski MJ, Chen YN, Cannon MC. 2011. Role of the extensin superfamily in primary cell wall architecture. *Plant Physiology* **156**: 11–19.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**: R25.

Larsen PE, Sreedasyam A, Trivedi G, Podila GK, Cseke LJ, Collart FR. 2011. Using next generation transcriptome sequencing to predict an ectomycorrhizal metabolome. *BMC Systems Biology* **5**: 70.

Le Quéré A, Wright DP, Soderstrom B, Tunlid A, Johansson T. 2005. Global patterns of gene regulation associated with the development of ectomycorrhiza between birch (*Betula pendula* Roth.) and *Paxillus involutus* (Batsch) fr. *Molecular Plant-Microbe Interactions* **18**: 659–673.

Lefebvre B, Timmers T, Mbengue M, Moreau S, Herve C, Toth K, Bittencourt-Silvestre J, Klaus D, Deslandes L, Godiard L *et al.* 2010. A remorin protein interacts with symbiotic receptors and regulates bacterial infection. *Proceedings of the National Academy of Sciences, USA* **107**: 2343–2348.

Levy G, Becker M, Duhamel D. 1992. A comparison of the ecology of pedunculate and sessile oaks: radial growth in the center and northwest of France. *Forest Ecology and Management* **55**: 51–63.

Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**: 323.

Loth-Pereda V, Orsini E, Courty PE, Lota F, Kohler A, Diss L, Blaudez D, Chalot M, Nehls U, Bucher M *et al.* 2011. Structure and expression profile of the phosphate Pht1 transporter gene family in mycorrhizal *Populus trichocarpa*. *Plant Physiology* **156**: 2141–2154.

Maier A, Riedlinger J, Fiedler H-P, Hampp R. 2004. *Actinomycetales* bacteria from a spruce stand: characterization and effects on growth of root symbiotic, and plant parasitic soil fungi in dual culture. *Mycological Progress* **3**: 129–136.

Marjanovic Z, Uehlein N, Kaldenhoff R, Zwiazek JJ, Weiss M, Hampp R, Nehls U. 2005. Aquaporins in poplar: what a difference a symbiont makes!. *Planta* **222**: 258–268.

Marquis RJ, Whelan CJ. 1994. Insectivorous birds increase growth of white oak through consumption of leaf-chewing insects. *Ecology* **75**: 2007–2014.

Martin F, Kohler A, Duplessis S. 2007. Living in harmony in the wood underground: ectomycorrhizal genomics. *Current Opinion in Plant Biology* **10**: 204–210.

Marx DH. 1969. Influence of ectotrophic mycorrhizal fungi on resistance of pine roots to pathogenic infections. I. Antagonism of mycorrhizal fungi to root pathogenic fungi and soil bacteria. *Phytopathology* **59**: 153–163.

Münzenberger B, Otter T, Wüstrich D, Polle A. 1997. Peroxidase and laccase activities in mycorrhizal and non-mycorrhizal fine roots of Norway spruce (*Picea abies*) and larch (*Larix decidua*). *Canadian Journal of Botany-Revue Canadienne De Botanique* **75**: 932–938.

Nehls U. 2008. Mastering ectomycorrhizal symbiosis: the impact of carbohydrates. *Journal of Experimental Botany* **59**: 1097–1108.

Newman AM, Cooper JB. 2011. Global analysis of proline-rich tandem repeat proteins reveals broad phylogenetic diversity in plant secretomes. *PLoS ONE* **6**: e23167.

Peterson RL, Farquhar ML. 1994. Mycorrhizas: integrated development between roots and fungi. *Mycologia* **86**: 311–326.

Philippe RN, Ralph SG, Mansfield SD, Bohlmann J. 2010. Transcriptome profiles of hybrid poplar (*Populus trichocarpa x deltoides*) reveal rapid changes in undamaged, systemic sink leaves after simulated feeding by forest tent caterpillar (*Malacosoma disstria*). *New Phytologist* **188**: 787–802.

Plett JM, Martin F. 2011. Blurred boundaries: lifestyle lessons from ectomycorrhizal fungal genomes. *Trends in Genetics* **27**: 14–22.

Richard F, Millot S, Gardes M, Selosse MA. 2005. Diversity and specificity of ectomycorrhizal fungi retrieved from an old-growth Mediterranean forest dominated by *Quercus ilex*. *New Phytologist* **166**: 1011–1023.

Sanchez-Rodriguez C, Rubio-Somoza I, Sibout R, Persson S. 2010. Phytohormones and the cell wall in Arabidopsis during seedling growth. *Trends in Plant Science* **15**: 291–301.

Sandmann T, Vogg MC, Owlarn S, Boutros M, Bartscherer K. 2011. The head-regeneration transcriptome of the planarian *Schmidtea mediterranea*. *Genome Biology* **12**: R76.

Sauter M, Hager A. 1989. The mycorrhizal fungus *Amanita muscaria* induces chitinase activity in roots and in suspension-cultured cells of its host *Picea abies*. *Planta* **179**: 61–66.

Schnabelrauch LS, Kieliszewski M, Upham BL, Alizedeh H, Lamport DTA. 1996. Isolation of pl 4.6 extensin peroxidase from tomato cell suspension cultures and identification of Val-Tyr-Lys as putative intermolecular cross-link site. *Plant Journal* **9**: 477–489.

Schrey SD, Schellhammer M, Ecke M, Hampp R, Tarkka MT. 2005. Mycorrhiza helper bacterium *Streptomyces* AcH 505 induces differential gene expression in the ectomycorrhizal fungus *Amanita muscaria*. *New Phytologist* **168**: 205–216.

Selle A, Willmann M, Grunze N, Gessler A, Weiss M, Nehls U. 2005. The high-affinity poplar ammonium importer PttAMT1.2 and its role in ectomycorrhizal symbiosis. *New Phytologist* **168**: 697–706.

Smith SE, Read D. 2008. *Mycorrhizal symbiosis*. London, UK: Academic Press.

Tagu D, Lapeyrie F, Martin F. 2002. The ectomycorrhizal symbiosis: genetics and development. *Plant and Soil* **244**: 97–105.

Thomas FM, Blank R, Hartmann G. 2002. Abiotic and biotic factors and their interactions as causes of oak decline in Central Europe. *Forest Pathology* **32**: 277–307.

Ueno S, Le Provost G, Leger V, Klopp C, Noirot C, Frigerio JM, Salin F, Salse J, Abrouk M, Murat F *et al.* 2010. Bioinformatic analysis of ESTs collected by Sanger and pyrosequencing methods for a keystone forest tree species: oak. *BMC Genomics* **11**: 650.

Viggars RM, Tarjan AC. 1949. A new root disease of pin oaks possibly caused by the nematode *Hoplolaimus coronatus* Cobb. *Plant Diseases Reporter* **33**: 132–133.

Voiblet C, Duplessis S, Encelot N, Martin F. 2001. Identification of symbiosis-regulated genes in *Eucalyptus globulus-Pisolithus tinctorius* ectomycorrhiza by differential hybridization of arrayed cDNAs. *Plant Journal* **25**: 181–191.

van de Wiel C, Scheres B, Franssen H, Vanlierop MJ, Vanlammeren A, Vankammen A, Bisseling T. 1990. The early nodulin transcript ENOD2 is located in the nodule parenchyma (inner cortex) of pea and soybean root nodules. *Embo Journal* **9**: 1–7.

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 Strategy for removing reads from oak-interacting organisms.

Fig. S2 Representation of reads with homology to *Arabidopsis* accessions in 16 pedunculate oak Roche 454 transcript libraries.

Fig. S3 Comparison of the oak assemblies generated by the MIRA and Trinity assembly programs with reference databases.

Fig. S4 Relatedness of the contigs in the OakContigDF159.1 reference transcriptome to sequences in the GenBank nr database.

Fig. S5 Classification of contigs in the OakContigDF159.1 reference transcriptome by Gene Ontology terms.

Fig. S6 Comparison of coverage of global KEGG metabolic pathways by the OakContigDF159.1 reference transcriptome relative to the *Arabidopsis thaliana* proteome.

Fig. S7 Mycorrhiza formation on oak roots leads to increased transcript abundances of contigs associated with starch metabolism in comparison with the abundance in fine roots.

Table S1 Quantitative polymerase chain reaction primers

Table S2 Numbers and lengths of oak transcripts obtained by Roche 454 sequencing

Table S3 Transcripts in roots and leaves of oak DF159 microcuttings and in EMs synthesized with *Piloderma croceum* as revealed by Illumina sequencing

Table S4 GO annotation of contigs in the OakContigDF159.1 reference transcriptome

Table S5 GO enrichment analysis of EMs synthesized between pedunculate oak DF159 microcuttings and *Piloderma croceum*

Table S6 Differentially expressed transcripts in EMs synthesized between microcuttings of the pedunculate oak clone DF159 and *Piloderma croceum*

Methods S1 Supporting methods.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.