Published March 16, 2017

ORIGINAL RESEARCH

Transcriptomic Analysis of *Olea europaea* L. Roots during the *Verticillium dahliae* Early Infection Process

Jaime Jiménez-Ruiz, María de la O Leyva-Pérez, Elisabetta Schilirò, Juan Bautista Barroso, Aureliano Bombarely, Lukas Mueller, Jesús Mercado-Blanco, and Francisco Luque*

Abstract

Olive cultivation is affected by a wide range of biotic constraints. Verticillium wilt of olive is one of the most devastating diseases affecting this woody crop, inflicting major economic losses in many areas, particularly within the Mediterranean Basin. Little is known about gene-expression changes during plant infection by *Verticillium dahliae* of woody plants such as olive. A complete RNA-seq transcriptomic analysis of olive tree roots was made. Trinity assembler proved to be the best option to assemble the olive and *V. dahliae* transcriptomes. The olive transcriptome (Oleup) consisted of 68,259 unigenes (254,252 isoforms/transcripts), and the *V. dahliae* transcriptome (Vedah) consisted of 37,425 unigenes (52,119 isoforms/transcripts). Most unigenes of the Oleup transcriptome corresponded to cellular processes (12,339), metabolic processes (10,974), single-organism processes (7263), and responses to stimuli (5114). As for the Vedah transcriptome, most unigenes correspond to metabolic processes (25,372), cellular processes (23,718), localization (6385), and biological regulation (4801). Differential gene-expression analysis of both transcriptomes was made at 2 and 7 d post-infection. The induced genes of both organisms during the plant-pathogen interaction were clustered in six subclusters, depending on the expression patterns during the infection. Subclusters A to C correspond to plant genes, and subcluster D to F correspond to *V. dahliae* genes. A relevant finding was that the differentially expressed gene (DEGs) included in subclusters B and C were highly enriched in proteolysis as well as protein-folding and biosynthesis genes. In addition, a reactive oxygen species (ROS) defense was induced first in the pathogen and later in the plant roots.

Core Ideas

- A transcriptomic RNA-seq analysis was conducted to study the olive–*V. dahliae* interaction.
- The transcriptomes of olive roots and *V. dahliae* were compiled at an early stage of infection.
- A number of putative genes involved in the plant defense were found.
- Most of the induced genes in response to the infection are related to protein turnover.
- An ROS stress-defense response is induced first in the pathogen and later in the plant.

OLIVE (*Olea europaea* L.) was one of the first tree species to be domesticated and cultivated. Wild and cultivated olives are diploid ($2n = 46$) and have a genome size of approximately 1800 Mb (De la Rosa et al., 2003). Today, olive cultivation has spread worldwide and has far-reaching economic, social, and ecological implications within the Mediterranean Basin. In fact, according to International Olive Oil Council data (http://www.

J. Jiménez-Ruiz, M.O. Leyva-Pérez, J.B. Barroso, and F. Luque, Center for Advanced Studies in Olive Grove and Olive Oils, Dep. of Experimental Biology, Univ. Jaén, 23071-Jaén, Spain; E. Schilirò and J. Mercado-Blanco, Dep. de Protección de Cultivos, Institute for Sustainable Agriculture (CSIC), 14004-Córdoba, Spain; A. Bombarely, Virginia Polytechnic Institute and State Univ., Blacksburg, VA, 24061; L. Mueller, Boyce Thompson Institute for Plant Research, Ithaca, NY 14853-1801. Received 6 July 2016. Accepted 13 Nov. 2016. *Corresponding author (fjluque@ujaen.es). Assigned to Associate Editor Stephen Moose.

Abbreviations: D, defoliating; DEG, differentially expressed gene; GO, Gene Ontology; NGS, next-generation sequencing; qPCR, quantitative polymerase chain reaction; ROS, reactive oxygen species; SNP, single-nucleotide polymorphism; SOD, superoxide dismutase; VWO, verticillium wilt of olive.

Published in Plant Genome
Volume 10. doi: 10.3835/plantgenome2016.07.0060

internationaloliveoil.org/documents/viewfile/4248-production3-ang/), more than 90% of the global cultivation area and an equal percentage of olive-oil production are located in this region. Virgin olive oil, the main product from olive trees and the principal component of the so-called Mediterranean diet, is consumed worldwide for its potential health and nutritional benefits and its exceptional organoleptic properties (Donaire et al., 2011).

Verticillium wilts are diseases caused by the pathogenic soilborne fungi *Verticillium* spp. and occur on a wide range of susceptible host plants throughout temperate and subtropical regions (Faino et al., 2012). *Verticillium dahliae* Kleb., one of the most damaging species within this genus, can infect more than 200 plant species, including high-value annual and perennial crop plants as well as wild, fruit, and ornamental trees and shrubs (Klosterman et al., 2009). *V. dahliae* is difficult to control due to a range of factors including lack of host specificity; the production of resistant structures (i.e., microsclerotia), which can survive for years in the soil; and the systemic nature of the infection (Pegg and Brady, 2002).

Olive cultivation is affected by a wide array of biotic constraints. Verticillium wilt of olive (VWO) is one of the most devastating diseases affecting this woody crop, inflicting heavy economic losses in many areas, particularly within the Mediterranean Basin. The disease is caused by *V. dahliae* and was first described in Italy (Ruggieri, 1946). Since then, it has been detected in almost all regions where olive is cultivated, causing serious concern to growers, nursery companies, and the olive-oil industry (López-Escudero and Mercado-Blanco, 2011). In addition, the steady spread and increasing prevalence of a highly virulent defoliating (D) pathotype in some geographical areas has increased the threat posed by *V. dahliae* in olive-growing regions (Dervis et al., 2010; López-Escudero et al., 2010). Although some cultivars of olive tree have been defined as resistant or tolerant to the pathogen, most of the currently cultivated olive cultivars are susceptible to the disease (Sesli et al., 2010; Erten and Yildiz, 2011; Trapero et al., 2013; García-Ruiz et al., 2015). Moreover, the susceptible and extremely susceptible cultivars are the economically and historically important ones in their cultivation areas: 'Hojiblanca' and 'Picual' in Spain, 'Konservolia' and 'Kalamon' in Greece, and the worldwide-cultivated 'Manzanillo' (López-Escudero and Mercado-Blanco, 2011). General aspects of Verticillium wilts and their control have been widely reviewed (Fradin and Thomma, 2006; Klosterman et al., 2009) and host-pathogen interactions in VWO have been amply studied and well summarized in recent years by Jiménez-Díaz et al. (2012) and López-Escudero and Mercado-Blanco (2011).

Different efforts have been implemented to evaluate olive-tree resistance to VWO (Markakis et al., 2009), to control the infection process with resistant rootstocks (Bubici and Cirulli, 2012), and to study resistance under field and greenhouse conditions (Trapero et al., 2013; Garcia-Ruiz et al., 2014). Recently, host resistance has been tested using a large number of olive genotypes (Arias-Calderón et al., 2015a, 2015b; Trapero et al., 2015). Other aspects of the infection have been investigated, including the response to stem-puncture inoculation (López-Escudero et al., 2007), the colonization process (Prieto et al., 2009), natural recovery as a control strategy (Bubici and Cirulli, 2014), the effect of soil temperature in olive response to VWO (Calderón et al., 2014), systemic responses of olive potentially related to VWO resistance (Gómez-Lama Cabanás et al., 2015), and *V. dahliae* genetic structure and cell-wall-degrading enzymes as pathogenicity factors (Gharbi et al., 2015a, 2015b).

During the last decade, the development of next-generation sequencing (NGS), together with whole-genome sequencing and advances in proteomics, has propelled a large number of current researchers into the omics technologies. These methods offer unprecedented opportunities to increase our understanding of the function and dynamics of biological systems, from cells to ecosystems. In particular, transcriptomic studies have been boosted by implementing this technology. Today, RNA-seq, that is, the sequencing of cDNA derived from an RNA population using NGS, is one of the most current methodologies. Indeed, other methods of studying gene expression at a large scale, such as microarrays and serial analysis of gene expression, are being replaced by RNA-seq. This approach can show the repertoire of expressed sequences found in a particular tissue at a specific time, including rare transcripts, due to the great depth of sequencing (Strickler et al., 2012), and it offers a nearly complete picture of transcriptomic events in a biological sample. Furthermore, no previous genome-sequence knowledge is necessary, because RNA-seq data sets themselves can be used to create the sequence assemblies for subsequent mapping of reads, together with the potential for detecting exon/exon boundaries, alternative splicing, and novel transcribed regions in a single sequencing run (Martin et al., 2013). The data can be used to characterize genes (Alves-Carvalho et al., 2015), improve genome annotation (Warren et al., 2015), identify and characterize novel noncoding RNAs (Wang et al., 2009), reveal information on novel transcripts (Hübner et al., 2015), assess gene expression (Mata-Pérez et al., 2015), and check for single-nucleotide polymorphisms (SNPs) (Shearman et al., 2015) and/or alternative splicing and structural variation. Moreover, RNA-seq is extremely practical in non-model species for which sequence data and resources are limited or nonexistent (Leyva-Pérez et al., 2014), since the focus of sequence is restricted to the coding region rather than to the entire genome.

Various genomic/transcriptomic studies have been undertaken in olive, with diverse aims: identifying small RNAs (Donaire et al., 2011; Yanik et al., 2013); SNP discovery (Kaya et al., 2013); studying juvenile-to-adult transition (García-López et al., 2014); analyzing systemic defense response in the cultivar Frantoio in *V. dahliae* infection (Gómez-Lama Cabanás et al., 2015); using RNA-seq technologies to assemble an olive transcriptome from different tissues and sequencing platforms,

and 454 and Sanger sequencing (Muñoz-Mérida et al., 2013); constructing a database with reproductive information (Carmona et al., 2015); and studying cold acclimation in olive leaves (Leyva-Pérez et al., 2014). Changes in gene expression induced by *V. dahliae* infection have been studied using transcriptional profiling in several plant–*V. dahliae* interactions (Gayoso et al., 2010; Derksen et al., 2013; Zhang et al., 2013; Daayf, 2015). In addition, RNA-seq has been used to broadly analyze gene expression in the microsclerotia (Duressa et al., 2013; Hu et al., 2014; Xiong et al., 2014) and to study pathogen-host interaction in *Verticillium* disease extensively in cotton (Xu et al., 2011; Sun et al., 2013; Zhang et al., 2013; Chen et al., 2015), tobacco (Faino et al., 2012), and tomato (Tan et al., 2015). However, little is known about gene-expression changes during the infection and the plant–*V. dahliae* interaction in woody plants such as olive, in which the disease is more problematic because they are perennial plants that cannot be grown as an annual to help control pests. For this reason, this study, using RNA-seq technology, is aimed at making a complete transcriptomic analysis of the plant-pathogen interaction between olive tree roots and the soilborne fungus *V. dahliae*.

Material and Methods

Pathogen Isolation, Culture Conditions, and Inoculum Production

Olive *V. dahliae* isolate V937I, which is representative of the highly virulent D pathotype, was used in this study. This isolate has been genetically, molecularly, and pathogenically characterized in previous studies (Collado-Romero et al., 2006; Prieto et al., 2009; Maldonado-González et al., 2015) and is deposited in the culture collection of the Department of Crop Protection, Institute for Sustainable Agriculture, Córdoba, Spain. The isolate is stored both in liquid-paraffin-covered cultures of plum-extract agar (Talboys, 1960) at 4°C in the dark and is also cryopreserved in glycerol (30%) at −80°C. Active culture of the isolate was first performed on chlorotetracycline-amended (30 mg L^{-1}) water agar and then further subcultured on potato dextrose agar for 7 d at 24°C in the dark. The inoculum (conidia suspension) used in the bioassay was made as follows. Four 500-mL Erlenmeyer flasks, each containing 250 mL of potato dextrose broth, were inoculated with three or four potato-dextrose agar plugs harboring actively growing mycelium of V937I. Cultures were then grown on an orbital shaker (175 rpm) for 7 d at 25°C in the dark. A conidia suspension was made by filtering the whole volume through two layers of sterile cheesecloth. The final density of the working inoculum was adjusted by counting conidia with a hemocytometer.

Verticillium dahliae–Olive Bioassay

A bioassay was made to study the *V. dahliae*–olive interaction at the transcriptomic level using non-gnotobiotic conditions. Four-month-old potted olive plants of the cultivar Picual were purchased from a commercial nursery located in province of Córdoba, southern Spain. At the nursery, plants were propagated by rooting leafy stem cuttings under mist conditions in plastic tunnels. Picual has been characterized as highly susceptible to *V. dahliae* infection, particularly from D isolates (Mercado-Blanco et al., 2003; López-Escudero and Mercado-Blanco, 2011). Before starting the experiment, the plants were acclimated for 6 wk under the same controlled conditions used after *V. dahliae* inoculation (see below). The isolate V937I was inoculated as follows. Infected plants ($N = 40$) were gently uprooted from the original substrate, their roots thoroughly washed in tap water, avoiding intentional wounding, and dipped for 30 min in a conidia suspension (10^7 conidia mL^{-1}) of V937I prepared as indicated above. For the control treatment, another group of 40 Picual plants was treated similarly, except that their root systems were dipped in sterile water. This group was considered as the control of any potential, unintentional damage that root systems could undergo during the manipulation and inoculation procedures. Plants were then individually transplanted into 7- × 7- × 8-cm polypropylene pots filled with an autoclaved (121°C, 1 h, twice on consecutive days) sandy substrate prepared ad hoc as described by Prieto and Mercado-Blanco (2008). Plants were randomly distributed over a 2-m^2 clean surface and incubated in a growth chamber adjusted to 24°C (day)/21°C (night), 60% relative humidity, and a 14-h photoperiod of fluorescent light (360 μE m^{-2} s^{-1}) for 15 d. To ease plant stress after manipulation, inoculation, and transplanting, we progressively lengthened the photoperiod during the first 5 d after transplanting. Both aerial tissues and root systems of each plant were harvested at 0, 8, 24, and 48 h and at 3, 4, 5, 6, 7, 10, 13, and 15 d (three plants/time point) after *V. dahliae* inoculation. Tissue samples were immediately frozen in liquid nitrogen and kept at −80°C until extraction of total RNA. An additional group of 24 plants (12 *V. dahliae*-inoculated and 12 control) was kept for up to 2 mo after inoculation as a reference for the development of VWO symptoms. Plants were selected until 15 d, because at that time the root-infection process can be considered complete (Mercado-Blanco et al., 2003; Gómez-Lama Cabanás et al., 2015), and samples from 0, 2, and 7 d were used for the RNA-seq analysis.

RNA Sample Preparation and Next-Generation Sequencing

To reduce plant-to-plant variability, we defined three groups of three randomly selected plants within each treatment condition. Total RNA samples were extracted from leaves and roots of control, 48-h- and 7-d-infected, and control uninoculated plants, with a Spectrum Plant Total RNA kit (Sigma-Aldrich, St. Louis, MO) according to the manufacturer's instructions. Any DNA contamination was removed by DNase I treatment in a column (Roche, Basel, Switzerland). The RNA quality tests were performed with the Agilent 2100 bioanalyzer (Agilent Technologies, Santa Clara, CA) using an RNA 6000 Pico

assay kit (Agilent Technologies). Equimolar amounts of RNA from each tissue and group were pooled. Then cDNA libraries were prepared, and NGS sequencing was made by GeneSystems (Valencia, Spain) with an Illumina HiSeq 1000 sequencer. Two replicates per sample were sequenced on different lanes in the flow cell.

Real-Time Quantitative Polymerase-Chain-Reaction Analysis

First-strand cDNA was synthesized from 1 μg of total RNA primed with 60 μM of random hexamer primer and Transcriptor reverse transcriptase, using the Transcriptor First Strand cDNA Synthesis Kit (Roche, Basel, Switzerland) following the manufacturer's instructions. The real-time quantitative polymerase chain reaction (qPCR) was performed in a Bio-Rad CFX96 and CFX384 PCR system with master mix SsoFast EvaGreen Supermix (Bio-Rad Laboratories, Hercules, CA) in 10 μL of reaction mixture containing 10 ng of cDNA. Amplifications were performed under the following conditions: initial polymerase activation at 95°C for 30 s, then 40 cycles at 95°C for 3 s, and at 60°C for 7 s, followed by a melting step from 65 to 95°C. An internal control of constitutive olive actin, previously selected as the most constant in expression after comparing several genes and primer pairs on different olive tissues, was used for the normalization (García-López et al., 2014). Each PCR reaction was performed three times and a pool of three different trees were analyzed at each time point. The oligonucleotides used for the amplifications are listed in Supplemental Table S1.

Table 1. Transcriptome assembly samples and accessions numbers.

Samples	Biosample accession number	SRA accession number	No. of reads after trimming
Control plant roots	SAMN02937426	SRR1525051	18,803,949
		SRR1525052	18,924,764
Roots damaged 8h	SAMN02937427	SRR1525231	45,343,122
		SRR1525237	45,239,794
Roots damaged 24h	SAMN02937428	SRR1524947	26,566,849
		SRR1524948	26,376,003
Roots damaged 48h	SAMN02937429	SRR1524949	22,342,392
		SRR1524950	22,413,949
Roots damaged 7 d	SAMN02937430	SRR1524951	30,730,174
		SRR1524952	30,827,818
Roots infected by *Verticillium dahliae* 48h	SAMN02937431	SRR1525086	26,952,334
		SRR1525087	26,886,101
Roots infected by *V. dahliae* 7 d	SAMN02937432	SRR1525113	27,437,498
		SRR1525114	27,136,149
Roots infected by *V. dahliae* 15 d	SAMN02937433	SRR1525213	22,019,596
		SRR1525114	21,755,642
Control plant leaves	SAMN02937434	SRR1525224	28,615,386
		SRR1525226	28,663,124
Cold-stressed leaves 24h	SAMN02937435	SRR1525284	34,718,359
		SRR1525285	34,272,790
Cold-stressed leaves 10 d	SAMN02937436	SRR1525286	11,056,074
		SRR1525287	10,999,707
Leaves of plants with damaged roots 15 d	SAMN02937437	SRR1525415	47,203,490
		SRR1525416	26,174,566
Leaves of plants infected by *V. dahliae* 15 d	SAMN02937438	SRR1525436	37,162,192
		SRR1525437	36,663,230

Data Preprocessing

The raw Illumina RNA-seq reads were first preprocessed using Fastqmcf (Aronesty, 2011) by discarding primers and reads with adaptors, unknown nucleotides, and poor-quality or short-length reads, increasing the Q score to more than 30 for all the libraries and lengths greater than 50 bp (Q30L50). A thorough quality control of sequencing was performed twice using FastQC software (version 0.10.1, http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc; Schmieder and Edwards, 2011) to provide a summary and compare statistics files before and after preprocessing. Part of the Illumina reads coming from the study of different transcriptional responses during cold acclimation in olive trees (Leyva-Pérez et al., 2014) and additional ones were added to build an exhaustive olive transcriptome of diverse abiotic and biotic stresses in which the pathogens were widely represented. About 31% of the total reads were from infected samples. A summary is presented in Table 1.

Transcriptome Assembly and Quality Evaluation

De novo assembly based on the de Bruijn graph was performed with different assemblers for comparison. After trimming and cleaning, reads were de novo assembled on three different platforms: SOAPdenovo-Trans (Xie et al., 2014), Trans-ABySS (Simpson et al., 2009), and Trinity software (Grabherr et al., 2011), which were then evaluated for the one that produced the best results. Assemblies were performed according to the developers' instructions, adjusting *k*-mer size in SOAPdenovo-Trans (k = 63, 55, 47, 39, 31, 23, and 19) and using k = 64 in Trans-ABySS and k = 25 (default) in Trinity, for which an in silico read normalization was also conducted, as advised. For all de novo assemblies and downstream analyses a server with 1 terabyte of RAM, 64 cores(central processing units), and Ubuntu as the operating system was used. Primitive assemblies were first filtered for sequences longer than 200 bp, assuming that shorter contigs were insufficiently assembled reads. The basic quality parameters of each assembly were calculated with GenoToolBox (https://github.com/aubombarely/GenoToolBox), and a collection of scripts served to manipulate genomic data. Full-LengtherNext 0.0.8 was used for sequence-structure analysis (Claros et al., 2012). Plots were generated using R software.

Transcriptome Clean-Up

A specific nucleotide database was created that comprised sequences from phylogenetically related plant species, the pathogen genome, and cDNA from the most representative contaminant species plus more specific ones obtained from a BLASTX of the longest assembled sequences with public databases. The complete assembly generated was

Table 2. Organisms included in the database for the BLASTN.

Organism	Treated as	Download	No. of sequences
Mimulus guttatus	Plant	phytozome.net	28,282
Solanum lycopersicum		solgenomics.net	34,727
Vitis vinifera		phytozome.net	26,346
Arabidopsis thaliana		phytozome.net	35,386
Oryza sativa		phytozome.net	49,061
V. dahliae	Pathogen	broadinstitute.org	395†
Dictyostelium discoideum	Contaminant	dictybase.org	11‡
Homo sapiens		ensemble.org	191,495
Mus musculus		ensemble.org	83,328
Xenopus tropicalis		xenbase.org	22,878
Saccharomyces pombe		pombase.org	4‡
Saccharomyces cerevisiae		yeastgenome.org	17‡
Escherichia coli		ncbi.nlm.nih.gov	4,321
O. europaea chloroplast		ncbi.nlm.nih.gov	1§

†Sequence of complete chromosomes and unplaced sequences.
‡Complete chromosomes.
§Added as a contaminant to remove from the transcriptome.

divided into three parts: plantlike contigs, *V. dahliae*–like contigs, and contaminant-like contigs. Selected cDNA or genomes were downloaded specifically from each organism, and a BLASTN was performed with a 1×10^{-5} Expect (E) value to crudely separate contigs that were plantlike, *V. dahliae*–like, and those considered to be contaminant-like (Table 2). The complete genome from *V. dahliae* was used to improve the subtraction of most pathogen sequences. Furthermore, aligned contigs were tagged to differentiate them as belonging either to olive or *V. dahliae*. After first being debugged, unaligned sequences were then matched using the BLASTX algorithm against the UniProtKB/SwissProt *Viridiplantae* current database with a minimum E value of 1×10^{-10} to recover probable plantlike contigs, and were also labeled. The final de novo assembly included all the plant and pathogen filtered contigs as a final reference for downstream analyses.

Gene Expression and Differentially Expressed Genes

Gene expression was performed with RNA-Seq by expectation-maximization software (Li and Dewey, 2011) implemented in the Trinity pipelines. EdgeR, a Bioconductor package (Robinson et al., 2010) was used to identify and cluster DEGs comparing different samples using R statistical software.

Functional Annotation

Functional annotation and Gene Ontology (GO) analysis was performed comparing both free and open sources: Blast2GO (https://www.blast2go.com; Conesa et al., 2005) and Trinotate pipeline (https://trinotate.github.io/). The GO terms associated were loaded in the Blast2GO interface, and GO-term-enrichment statistical analysis was performed. Blast2GO integrated the Gossip package for statistical assessment of differences in GO-term abundance between two sets of sequences (Blüthgen et al., 2005), using Fisher's exact test and corrections for multiple testing. A one-tailed Fisher's exact test was performed with a False Discovery Rate with a filter value of less than 0.01. Results were saved in a Microsoft Excel data sheet, and charts were generated.

Availability of Data and Materials

The RNaseq sequence data set supporting the results of this study is available at NCBI with accession numbers SRR1525051, SRR1525052, SRR1524949, SRR1524950, SRR1524951, SRR1524952, SRR1525086, SRR1525087, SRR1525113, SRR1525114) SRR1525231, SRR1525237, SRR1524947, SRR1524948, SRR1525213, SRR1525114, SRR1525224, SRR1525226, SRR1525284, SRR1525285, SRR1525286, SRR1525287, SRR1525415, SRR1525416, SRR1525436, and SRR1525437.

Results

De novo Assembly and Characterization of Olive and *V. dahliae* Transcriptomes

To have a complete plant-pathogen (Picual–*V. dahliae* D pathotype) transcriptome, we extracted RNA samples from *V. dahliae*–inoculated (infected), control treatment (uninoculated), and control, untreated Picual 4-mo-old olive plants. The mRNA-seq constructed libraries were sequenced on different lanes in the flow cell with Illumina HiSeq 1000 paired-end technology. We generated a total of 735 million paired-end reads after cleaning and trimming (139.79 Gb), thus improving quality of raw data to Q30L50, which ranged from 11 to 45 million reads across the 26 libraries (Table 1).

To assess the quality of each assembly and to compare them preliminarily, we calculated the basic statistics and quality parameters (number of contigs, contig mean size, longest contig, N50, L50, and proportion of contigs longer than 1 kb) with GenoToolBox (Table 3). The quality of de novo transcriptome assemblies was evaluated by comparing them with data from genomes of the most phylogenetically related species. Within the Asterids subclass, which includes the olive tree, there are three accessible genomes: two from the genus *Solanum*, which is within the Solanaceae family (*Solanum lycopersicum* L., tomato, and *S. tuberosum* L., potato) and one from the family Phrymaceae (*Mimulus guttatus* Fisch. ex DC., monkey flower), within the order Lamiales, in which the olive tree is also included (Goodstein et al., 2012). For this purpose we chose *M. guttatus* as most closely related species and *S. lycopersicum* as the representative species from the Asterids subclass. Total cDNA was downloaded from the respective genomes, and statistical analyses were performed in all de novo transcriptomes and the aforementioned related species cDNA data set, which was first filtered for sequences longer than 200 bp to normalize them. The comparison by the software used in our study showed better results with Trinity than with SOAPdenovo-Trans and Trans-ABySS through the entire

Table 3. Summary of basics statistics of the primary assemblies generated with Trinity (Trinity_k25), SOAPdenovo-Trans (Soap_k63–19) and ABySS (ABySS_k64) software and tomato (ITAG2.3) and monkey flower (Mg v1.1) cDNA datasheets.

Assembler	Contigs†	Contig mean size	Longest contig	N50‡	L50§	contigs > 1kb
		bp			bp	%
Trinity_k25	691,424	1064	17,990	120,580	1919	34.98
Soap_k63	698,404	887	23,468	109,315	1614	18.69
Soap_k55	819,853	865	20,482	131,744	1511	16.27
Soap_k47	929,366	827	27,733	154,273	1380	15.17
Soap_k39	1022,968	777	18,533	174,065	1254	11.70
Soap_k31	1087,478	725	20,187	194,065	1103	6.61
Soap_k23	1080,452	709	19,399	197,832	1044	7.33
Soap_k19	787,608	688	19,865	153,739	960	6.40
ABySS_k64	761,563	546	17,974	150,902	692	0.52
ITAG2.3	32,518	1280	23,220	15,536	1693	53.3
Mg v1.1	28,261	1368	15,339	8,029	1658	62.3

†Total number assembled.
‡Minimum number of contigs representing 50% of the assembly.
§Minimum contig length representing 50% of the assembly.

range of different conditions provided; Trans-ABySS exhibited the lowest quality of the assemblies (Fig. 1). The *k*-mer number was adjusted to include representative lengths, from 19 to 63 with SOAPdenovo-Trans, and better results were found with *k*-mer 63, which returned a higher L50 and higher percentages of sequences longer than 1 kb. Trans-ABySS was used with the best *k*-mer for SOAPdenovo-Trans to save computational time, assuming that the results would be similar. The Trinity assembler was used with default *k*-mer 25, yielding a lower number of contigs assembled and higher mean size, higher L50, and the highest percentage of transcript longer than 1 kb. Regarding the statistical data of related species, Trinity produced a similar mean size and L50, although a higher number of assembled contigs. Consequently, in accordance with our data, Trinity performed the most accurate assembly because of its resemblance to the tomato and monkey flower cDNA data sets and outperformed the other assemblers used.

Transcriptome Debug

The primary plant-pathogen transcriptome had a size of 736 Mb, in which 422,124 unigenes were identified, yielding a total of 691,424 isoforms/transcripts. As an initial step in the transcriptome debug, filtered sequences longer than 1 kb of the transcriptome were aligned using the BLASTX algorithm against the current UniProtKB database and indicated that 26.2% (44,194) and 19.7% (33,178) map with plant and fungal species, respectively. The transcriptome, which was expected to contain mostly contigs of both plant and pathogen, was mapped with plant and fungal species, respectively. The transcriptome, which was expected to be composed mostly of olive and *V. dahliae* contigs, was mapped also with amoeba (3%), human (2.7%), mouse (2.5%), rat (1.2%), and other predictable contaminants and most representative species found in the database (Strickler et al., 2012; Lusk, 2014). Finally, 39.1% of the sequences remained unaligned (65,962; Supplemental Fig. S1). These results allowed the generation of a local specific nucleotide-sequence database consisting of three main parts: cDNA from phylogenetically related plant species, the *V. dahliae* genome, and the most representative contaminant species (Table 3). For the plant set, *M. guttatus* and *S. lycopersicum* were chosen as the most closely related species. Moreover, *Vitis vinifera* L., *Arabidopsis thaliana* (L.) Heynh., and *Oryza sativa* L. were selected as the most represented plant species at the first BLASTX alignment. The pathogen set was completely constituted by the whole set of *V. dahliae* genome sequences present in http://www.broadinstitute.org/ftp/pub/annotation/fungi/verticillium_dahliae/. The contaminant set was constructed from cDNA or complete genomes (depending on size) of the most displayed contaminant species in the previous BLASTX alignment, and also by adding the olive tree chloroplast sequence. A BLASTN alignment of the whole transcriptome against the previous database was used to separate the sequences into three subsets: plant, pathogen, and contaminant sets (Table 4). A total of 258,411 (37%) isoforms/transcripts from the original transcriptome were aligned to any of the three data sets, while 433,013 (63%) remained unaligned. Only 37% of the unaligned sequences were longer than 500 bp. As a means of recovering most of plantlike sequences from the unaligned set, a BLASTX (1×10^{-10}) with the sequences longer than 500 bp was performed with the current Viridiplantae database from UniProtKB/SwissProt. This allowed us to retrieve a total of 58,995 isoform/transcripts. The joining of the two sets of plant sequences configured the olive transcriptome (Oleup), consisting of 68,259 unigenes (254,252 isoforms/transcripts). In addition, a *V. dahliae* transcriptome consisting of 37,425 unigenes (52,119 isoforms/transcripts) was also configured (Vedah). The discarded sequences consisted of the contaminant set, which contained 11,035 (1.6%) isoform/transcripts, and of the unaligned sequences, which were 310,864 (45%) isoform/transcripts. Trinity assembler enabled also differentiation between unigenes and transcript/isoforms.

Annotation and Comparison with Other Transcriptomes

Functional annotation assigns to each transcript a putative function through a group of bioinformatic tools based mainly on DNA- or protein-sequence alignment scores. Transcriptomes were annotated using two software programs: Blas2GO (Conesa et al., 2005) and Trinotate pipeline. Blast2GO retrieves information from multiple public databases, among which UniProtKB and TAIR (Lamesch et al., 2012) are the most represented. For improved annotation, protein domains were examined with InterPro database (Hunter et al., 2012), and the

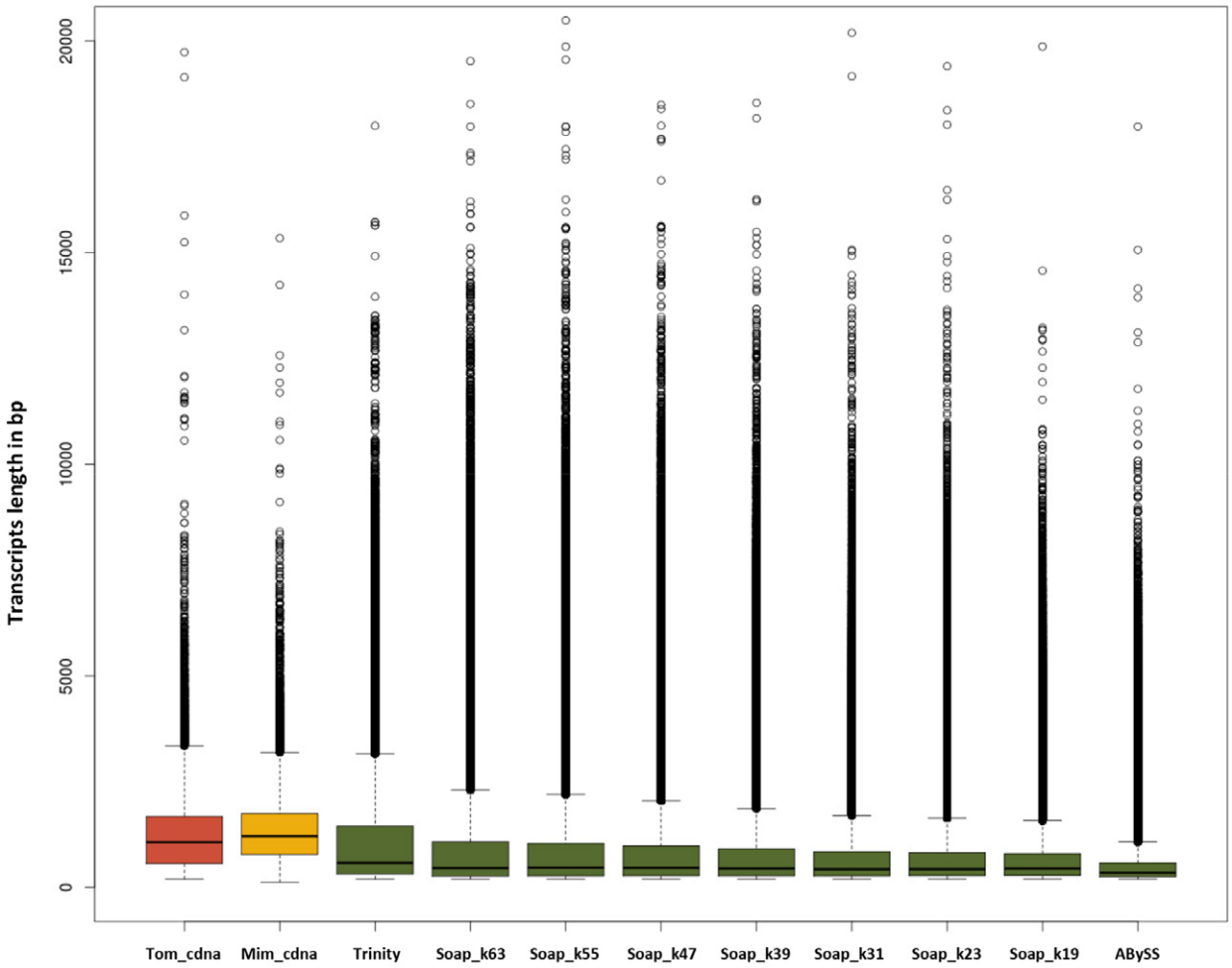

Figure 1. Boxplot comparisons of reference cDNA data sets and de novo assembled transcript length distribution using Trinity, SOAPdenovo-trans, and TransABySS software. First (ITAG 2.3) and second columns (Mg v1.1) indicate tomato and *M. guttatus* full CDS transcriptome, third column represents assembly using default *k*-mer set as 25 with Trinity. The 4th to 11th columns represent assembly generated with SOAPdenovo-trans with *k*-mers set from 63 to 19, respectively. The last column represents assembly generated with Trans-ABySS.

Table 4. Summary of BLASTN results of transcriptome assembled against local nucleotide database.

Set	Organism	No. of Contigs
Plantlike	*M. guttatus*	106,722
	S. lycopersicum	43,935
	V. vinifera	36,568
	A. thaliana	4,963
	O. sativa	3,069
		Total: 195,257
Pathogen-like	*V. dahliae*	52,119
		Total: 52,119
Contaminant-like	*D. discoideum*	1,674
	H. sapiens	1,559
	M. musculus	1,424
	X. tropicalis	1,147
	S. pombe	1,118
	S. cerevisiae	1,110
	E. coli	960
	O. europaea chloroplast	2,043
		Total: 11, 035

data were merged. On the other hand, the Trinotate pipeline includes protein domain identification (HMMER/PFAM), protein-signal prediction (SignalP/tmHMM), and comparison with currently curated annotation databases (EMBL/UniProtKB/eggnog/GO Pathways databases), and all functional data derived from the analyses were integrated into a SPQLite database, providing a useful tool to search for specific or desired terms. The Blast2GO complete annotation process gave 71 and 75% of the total sequences annotated from Oleup and Vedah, respectively, while the Trinotate pipeline annotated 79% of both transcriptomes. The three species with the most BLAST hits were *A. thaliana*, *Populus trichocarpa* Torr. & A. Gray, and *Ricinus communis* L. in Oleup, and *V. dahliae*, *Nectria hematococca* Berk. & Broome, and *Verticillium albo-atrum* Reinke & Berthold in Vedah. GO terms were used to classify functions of the assembled unigenes, and 187,171 total GO term annotations were obtained for the Oleup proposed transcriptome and 150,625 for the Vedah transcriptome. Most unigenes of the Oleup transcriptome corresponded to cellular

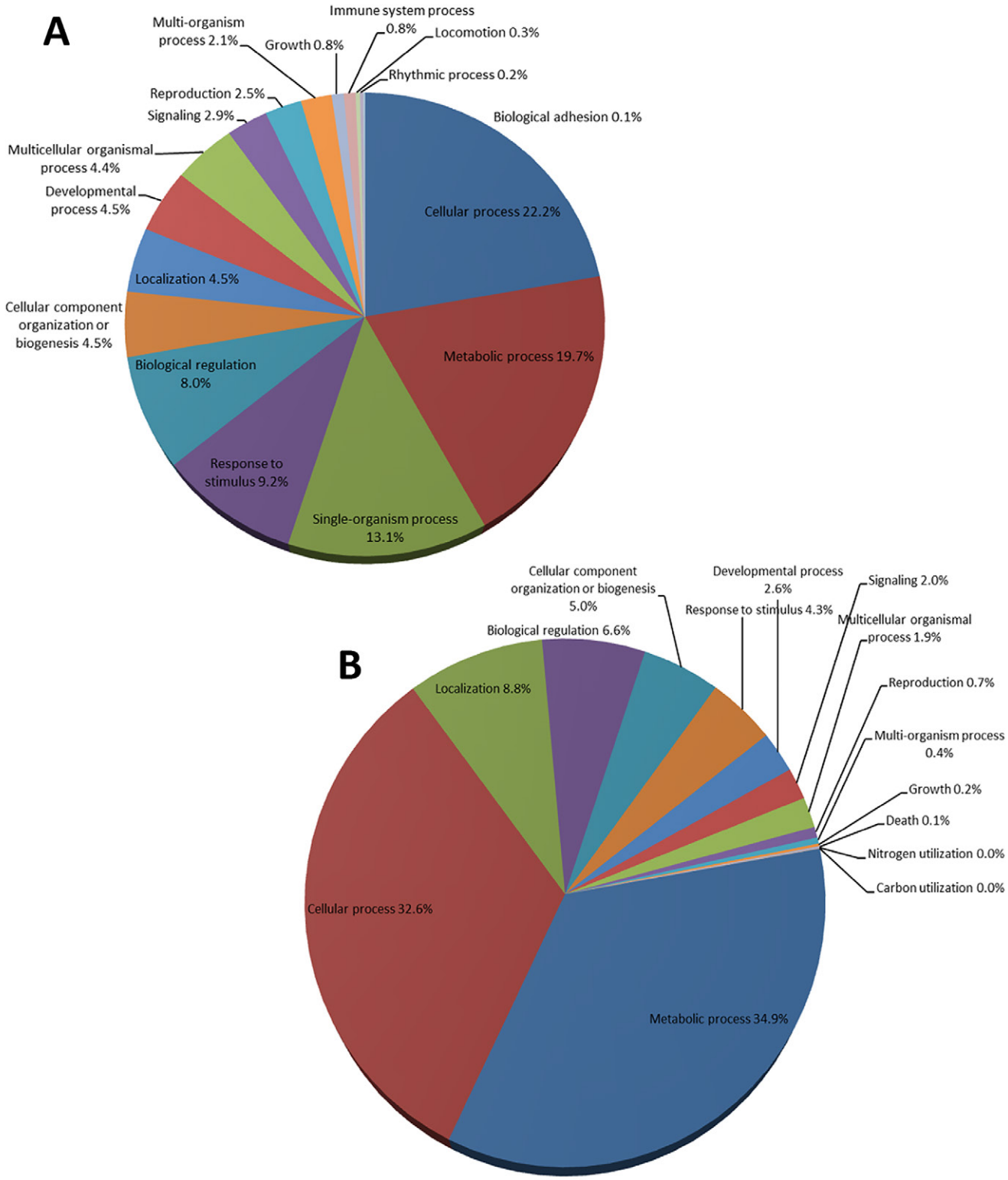

Figure 2. Genetic Ontology–term of biological processes. The Oleup (A) and Vedah (B) trancriptomes are represented.

processes (12,339), metabolic processes (10,974), single-organism processes (7263), and responses to stimuli (5114). As for the Vedah transcriptome concerns, most unigenes of corresponded to metabolic processes (25,372), cellular processes (23,718), localization (6385), and biological regulation (4801) (Fig. 2).

Differential Gene-Expression Analysis

To determine the number of olive and *V. dahliae* DEGs during the infection of the roots, a pairwise matrix comparison was made with both Oleup and Vedah transcriptomes (Table 5). The analysis of the Oleup transcriptome showed a response to unintentional root damage at 2 d post-treatment (459 DEGs) and a higher number

Table 5. Pairwise matrix comparison of differentially expressed genes of control plant roots (CTR), control-damage plant roots (DAM), and *Verticillium* infected plant roots (VER) at two different times.

Oleur	CTR	DAM_2d	DAM_7d	VER_2d	VER_7d
Vedah					
CTR	0	459	2858	918	6871
DAM_2d	—	0	954	776	3215
DAM_7d	—	—	0	2320	2750
VER_2d	2058	—	—	0	1960
VER_7d	2915	—	—	351	0

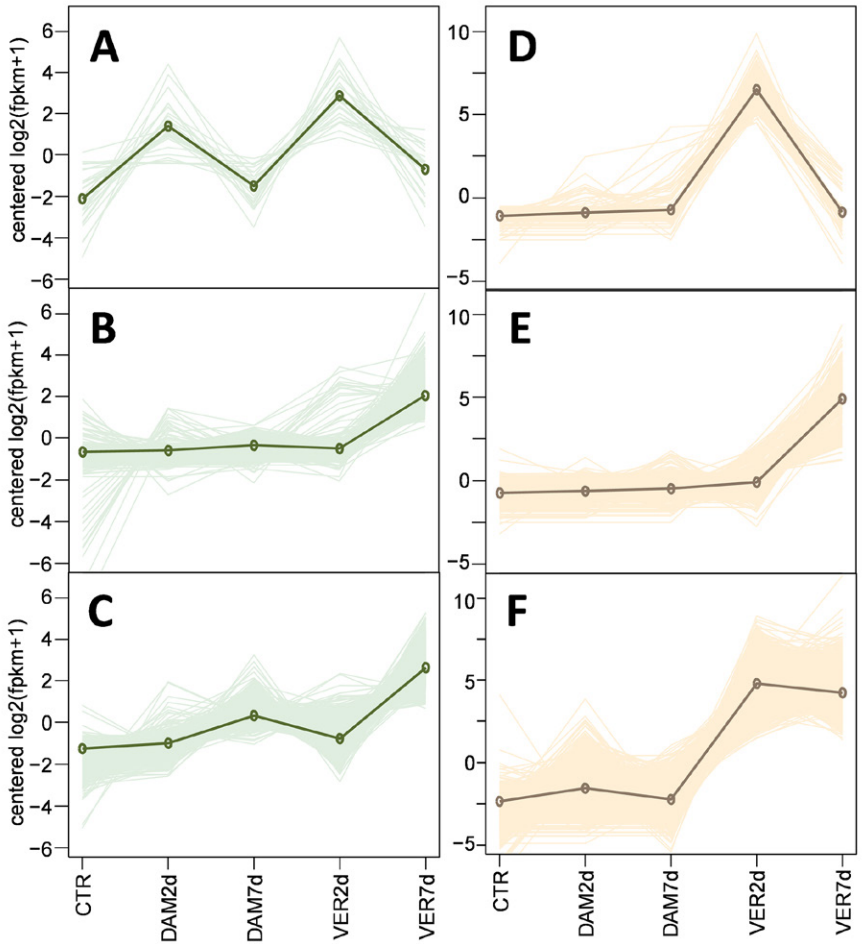

Figure 3. Subclusters with differently upregulated transcript genes. In all panels (A-F) soft/pale color lines indicate individual gene-expression level and the dark line indicates a consensus within a specific subcluster. A-C subclusters belong to the Oleup transcriptome (olive genes) and D-F to the Vedah transcriptome (*Verticillium dahliae* genes). CTR = control (untreated) time 0; DAM 2d = Control (uninoculated) roots 2 d; DAM 7d = Control (uninoculated) roots 7 d; VER 2d = Infected roots sampled at 2 d and VER 7d = Infected roots sampled at 7 d. In all cases the average of two independent replicates are represented.

of DEGs at 7 d post-treatment (2858 DEGs). However, this response was stronger in roots inoculated with *V. dahliae*. Moreover, the number of DEGs also increased with time: 918 DEGs at 2 d and 6871 DEGs at 7 d post-infection. When the infected plants were compared with the control uninoculated plants at each time, the genes that responded to unintentional root damage were subtracted, leaving the DEGs corresponding to genes that specifically responded to the *V. dahliae* infection. In this case, DEGs in the plant roots also increased from 2 d post-infection (776 DEGs) to 7 d post-infection (2750 DEGs). The number of *V. dahliae* DEGs in response to the plant infection was again higher at 7 d post-infection (2915 DEGs) than at 2 d post-infection (2058 DEGs), but in this case at 2 d the number of DEGs was already quite high and not much lower than at 7 d.

Clusters of Genes with Induced Expression Patterns during Infection

When searching in the Oleup transcriptome for olive upregulated DEGs during the infection, 2227 genes where found and grouped in three subclusters (Fig. 3, A–C). Group A consists of 27 DEGs transiently overexpressed in response to both unintentional root damage

inflicted during treatment and *V. dahliae* infection. Group B contains 1131 DEGs upregulated only after 7 d of infection and no induction at all in response to the root damage. Finally, group C, which consists of 1069 DEGs upregulated after 7 d post-infection, was like the group B but shows a low induction at 7 d in the control uninoculated root plants, so that group C was induced in response to *V. dahliae* infection and moderately upregulated in response to unintentional root damage. An analysis after normalization of the fungal Vedah transcriptome of genes overexpressed in olive roots infected with *V. dahliae* also identified three subclusters. Group D corresponds to 183 early and transiently upregulated DEGs at 2 d of infection, group E includes 302 DEGs that were late upregulated at 7 d post-infection, and group F consists of 3902 early and sustained overexpressed DEGs that were highly expressed at both time points (2 and 7 d post-infection; Fig. 3, D–F). Supplemental Table S2 shows a list of expression values per gene normalized by FPKM (fragments per kilobase of transcript per million mapped reads) and the data time points of each subcluster. The pattern of expression of the six subclusters was confirmed by quantitative real-time qPCR of 12 Oleup and 6 Vedah randomly selected DEGs (Fig. 4).

The results obtained after the most specific GO-term-enriched analysis of biological-process DEGs of the six subclusters involved quite general processes for subclusters A, D, E, and F (not shown), but in the case of subclusters B and C, the main biological process found was proteolysis, followed by protein folding and other processes related to protein biosynthesis (Fig. 5). Reactive oxygen species (ROS) are frequently used by plants as defenses against pathogens. In our study, the subcluster C included an overexpressed putative Cu/Zn superoxide dismutase (Cu/Zn-SOD; Oleup_comp247260_c0), which catalyzes the production of hydrogen peroxide from ROS. In the same subcluster a putative glutathione S-transferase (Oleup_comp248616_c0) and in subcluster B three putative glutathione peroxidases (Oleup_comp442709_c0; Oleup_comp268153_c0; Oleup_comp415514_c0), each of which catalyzes the conversion of hydrogen peroxide to water by the oxidation of glutathione, were also overexpressed. An induction of the ROS defense mechanism was also observed in the pathogen; however, this effect took place earlier in *V. dahliae* than in olive roots. Thus, four transcripts were induced in *V. dahliae* at very early times during the infection process. These potentially coded for (i) a putative Mn-SOD (Vedah_comp294808_c0, belonging to subcluster D); (ii) two putative peroxidases: a catalase (Vedah_comp451243_c1) and a glutathione S-transferase (Vedah_comp418642_c0), both included in subcluster F; and (iii) a putative glutathione reductase (Vedah_comp329313_c0; subcluster E). In addition to the ROS stress, another plant-defense mechanism against fungal infection may be attacks on the fungus wall by chitinases; a subcluster C endochitinase PR4 (Oleup_comp465700_c2) was overexpressed by the plant roots (Supplemental Fig. S2).Alternatively, the pathogen could have induced the expression of effector genes that might be relevant for its virulence. Thus, DEGs coding for chitin-binding proteins were found in group F (Vedah_comp485742_c0; Vedah_comp544404_c0; Vedah_comp461761_c0) and in group E (Vedah_comp236634_c0), all of which may be used to inactivate the plant endochitinase PR4 protein. Also in group F there were two *Ace1*-like DEGs (Vedah_comp438107_c0; Vedah_comp472974_c2) (Supplemental Fig. S3). The *mcm1* gene, which has been described as a virulence factor in *V. dahliae* (Xiong et al., 2016), was found to be expressed at a very low level after 7 d post-inoculation. In addition, we found that cinnamyl alcohol dehydrogenase was induced in the infected plants (group B), as they are in Sea Island cotton, in response to *V. dahliae* infection (Sun et al., 2013).

Discussion

Verticillium wilt is a major economic problem for olive cultivation, and most olive cultivars, including those of economic and agricultural value, are either moderately or highly susceptible to the pathogen. In fact, no real resistance has yet been found among olive cultivars, and only some of them show some level of tolerance to VWO, although they can also be infected by *V. dahliae* (Mercado-Blanco et al., 2003; Markakis et al., 2009, 2010; Gómez-Lama Cabanás et al., 2015). During the infection process, the pathogen invades the roots, enters the vascular system, and spreads throughout the entire olive plant (Prieto et al., 2009). To know how the plant and the fungus interact at the genetic level, it is crucial to a gain a more complete understanding of the infection and pathogenic processes. Moreover, elucidating the underlying mechanisms of these processes may help clarify the molecular basis of olive tolerance and resistance to *V. dahliae* and to design novel tools for the effective control of this disease. For this reason, in this study we sought mainly to determine the transcriptomes of the plant-pathogen interaction during the early stages of the infection until the infection is established and to finally identify the genes whose expression is modified during the infection. To achieve this, we studied olive plants that were artificially infected with a highly virulent D isolate of *V. dahliae* and analyzed the gene expression of both partners when they actively interacted: the olive root tissue and the invading pathogen, by RNA-seq at 0, 2, and 7 d post-infection. A 2-d time interval represents an early stage in the infection process, while at 7 d the infection is well established in the roots but it is too soon to detect visible symptoms (Mercado-Blanco et al., 2003). Therefore, the changes observed in gene expression should be due mostly to the plant-pathogen interaction rather than to a response to heavy root damage and consequent severe plant deterioration.

Transcriptome assembly was performed using different software, and in our study Trinity showed better

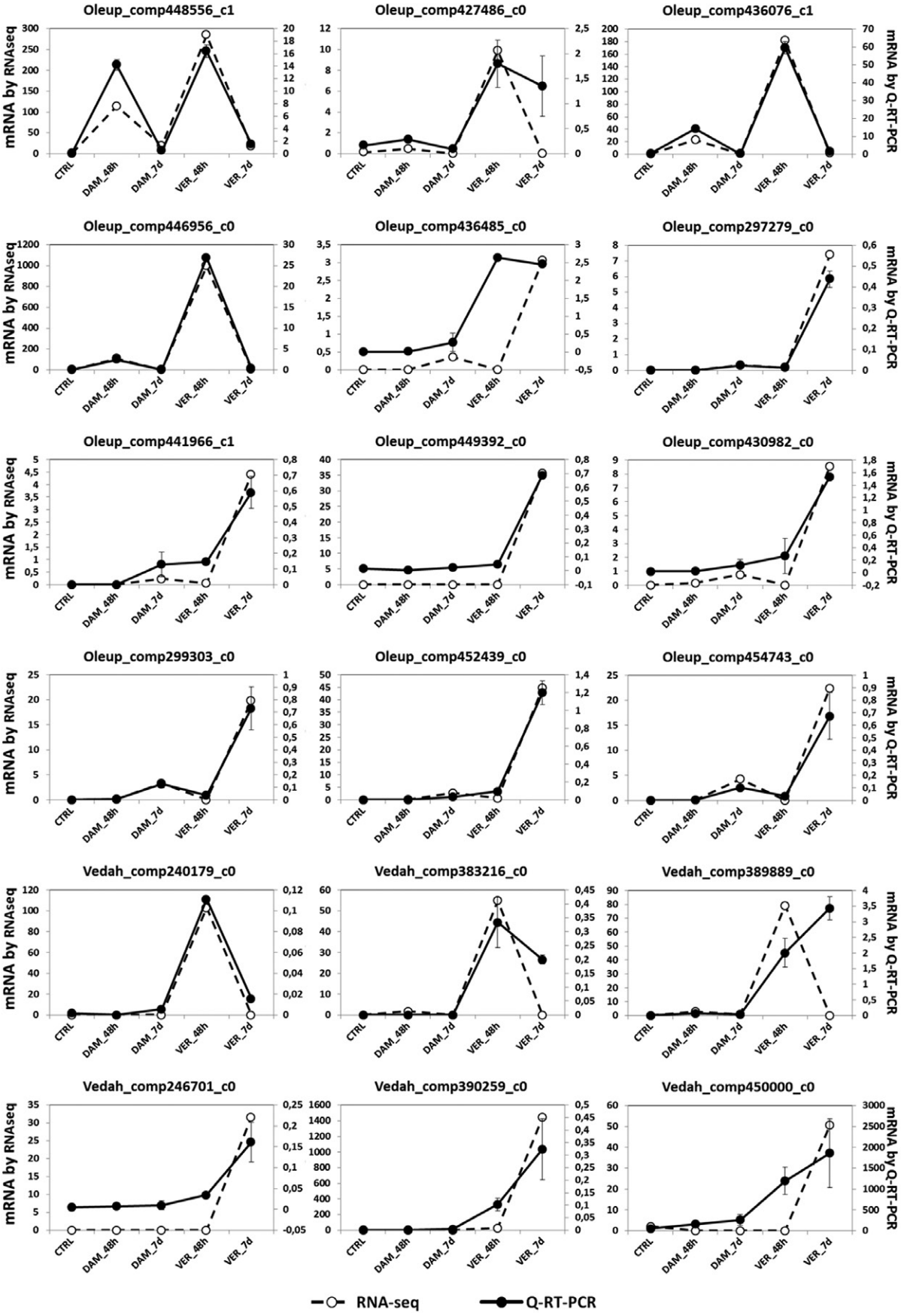

Figure 4. The mRNA expression level of 18 randomly selected genes analyzed by real-time quantitative PCR (qPCR) and RNA-seq. The real-time qPCR results are the average of three samples and are represented in arbitrary expression level units.

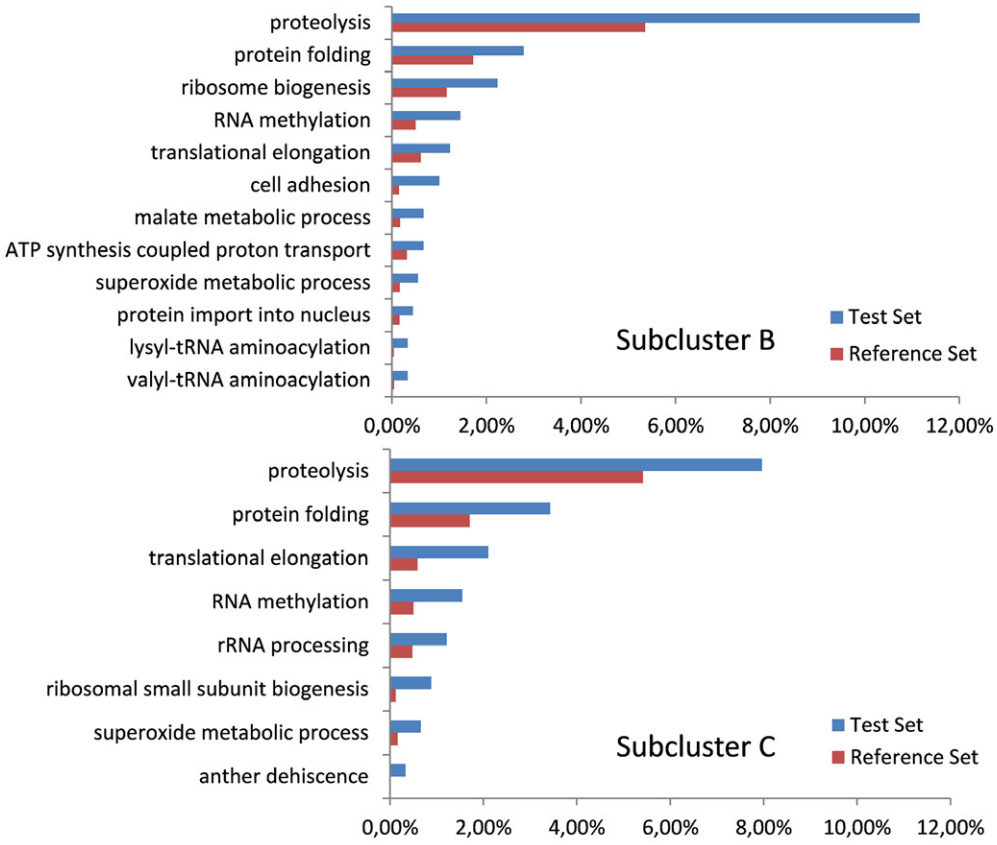

Figure 5. A graph of the most-specific Genetic Ontology–term-enriched biological-process DEGs in subclusters B and C. A filter node of p-value $\leq$ 0.01 was used.

results than SOAPdenovo-Trans and Trans-ABySS across the entire range of different conditions, with Trans-ABySS having the lowest quality of the assemblies (Fig. 1). Trinity produced the most similar mean size and L50 to the reference transcriptomes of tomato and monkey flower, and consequently it was chosen as the best transcriptome assembly to carry out further data analysis. Because the olive roots were inoculated with *V. dahliae*, the original transcriptome was aligned to plants, *V. dahliae*, and possible contaminants. Contaminants were removed and two transcriptomes resulted: a plant transcriptome corresponding to olive genes (Oleup) and a *V. dahliae* transcriptome (Vedah). In an initial transcriptomic study by suppression subtractive hybridization involving the olive–*V. dahliae* interaction, 18 genes were identified as being induced in the aerial tissues of 'Frantoio', an olive cultivar tolerant to *V. dahliae* (Gómez-Lama Cabanás et al., 2015). In this work, two new transcriptomes, one from olive and one from the pathogen, resulted from a potent RNA-seq approach during the infection process. This strategy represents a quite novel and valuable way to study the plant-pathogen interaction: it allows identification of DEGs and, more important, of the genes that are induced and the timing of induction simultaneously on both sides of the infection process.

The transcriptional analysis of the RNaseq results produced three patterns (subclusters A, B, and C) of induced DEGs in olive and other three patterns (subclusters D, E, and F) of highly expressed DEGs in *V. dahliae* during infection (Fig. 3). Subcluster A contains genes transiently overexpressed after the root wounding during the pathogen-inoculation process. These genes were overexpressed at 2 d after inoculation in both infected and control uninoculated roots and returned to a basal expression level after 7 d post-inoculation. Subclusters B and C represent groups of genes that were induced later, at 7 d post-inoculation, and without (subcluster B) or with (subcluster C) some induction in control uninoculated roots. Therefore, the DEGs included in subclusters B and C correspond to genes that were induced mainly in response to the *V. dahliae* infection. The most specific GO-term-enriched analysis of the biological-process DEGs included in subclusters B and C showed a high enrichment of proteolysis as well as protein folding and biosynthesis. This finding is indicative of a sharp change in protein abundance (Fig. 5). A proteomics study found similar changes in cotton during the *V. dahliae* infection (Xie et al., 2013), although in that case some lesions were already evident in cotton leaves but not yet in our olive samples. Therefore, these marked changes in the protein content of the roots appeared prior to any observable plant damage and are probably a consequence of the plant response to the fungal infection. In addition, a very strong overexpression of a subcluster C, chitinase class I (PR4) was noted, this probably being a frontline defense protein against the fungus. In

recent work, Gómez-Lama Cabanás et al. (2015) studied a subtractive cDNA library of a different olive cultivar that is far more tolerant to the pathogen and different tissues than the one in this work. So it is not surprising that there are differences between the results of the two studies; however, a few of the genes described as induced in response to the pathogen infection are the same or related. Thus, in our work, several β-glucosidases, lipoxygenases, cytochrome P450 subunits, thaumatin-like proteins, alcohol dehydrogenases, and methyltransferases are also found to be induced in response to the infection.

Subclusters D–F correspond to *V. dahliae* DEGs and mainly represent genes for general metabolism. Subcluster D includes genes that were transiently expressed at 2 d post-infection but that returned to a basal or undetectable level 7 d after the inoculation. Subcluster E contains DEGs were induced late in the infection, that is, at 7 d post-inoculation; and subcluster F is a large group of fungal DEGs that were overexpressed early in the infection and in a sustained manner. Effector genes were found among DEGs of the F and E groups. Interestingly, four DEGs coding for chitin-binding proteins were found and could be induced to inactivate the plant endochitinase PR4 protein. Also two *Ace1*-like DEGs might mediate avirulence indirectly by their involvement in the biosynthesis of an unknown secondary metabolite (Stergiopoulos and de Wit, 2009) (Supplemental Fig. S3). However, we did not find significant expression of the pathogen's potential virulence genes nor of many of the plant's resistance genes. We speculate that the reason for this observation is that this is a compatible host-pathogen interaction and the resistance response of the plant is very insufficient. In addition, a *V. dahliae* virulence factor, the *mcm1* gene (Xiong et al., 2016), was expressed only at a very low level after 7 d post-inoculation. But most genes described in other plants for resistance to *V. dahliae* are not overexpressed in this susceptible cultivar in response to the infection. In addition, cinnamyl alcohol dehydrogenase, which was induced in the infected plants (group B), has also been found to be upregulated in Sea Island cotton in response to *V. dahliae* infection (Sun et al., 2013), although is yet not clear what its involvement in resistance is.

Some of the induced DEGs in olive code for proteins directly involved in ROS protection. Thus, a Cu/Zn-SOD, three glutathione-peroxidases, and a glutathione S-transferase were found distributed between subclusters B and C. The ROS protective response was also detected in the pathogen. In that case, a Mn-SOD, a glutathione S-transferase, a catalase, and a glutathione reductase were highly expressed. It is noticeable that the response in *V. dahliae* response took place earlier (2 d after inoculation) than in olive roots, where the protective response against ROS was found later, at 7 d after inoculation. Therefore, once the pathogen enters the roots, a ROS protective response in the hyphae is rapidly induced, while the plant roots require more time to develop a response to the ROS stress and then also to significantly alter gene expression and protein abundance. It is important to note that the experiment was performed using a susceptible cultivar (Picual) and that plants equally and simultaneously infected died a few weeks later. Comparison of the transcriptome changes in response to the infection between a susceptible and a tolerant cultivar is be a future goal for understanding the genetic basis of the tolerance to *V. dahliae* in the cultivated olive tree.

Supplemental Information Available

Supplemental information is available with the online version of this article.

Acknowledgments

This work was supported by Grant AGR-5948 from Junta de Andalucía (Consejería de Economía, Innovación y Ciencia) and Ministerio de Economía y Competitividad. Technical and human support provided by CICT of Universidad de Jaén (UJA, MINECO, Junta de Andalucía, FEDER) is gratefully acknowledged.

References

Alves-Carvalho, S., G. Aubert, S. Carrère, C. Cruaud, A.L. Brochot, F. Jacquin, et al. 2015. Full-length *de novo* assembly of RNA-seq data in pea (*Pisum sativum* L.) provides a gene expression atlas and gives insights into root nodulation in this species. Plant J. 84:1–19. doi:10.1111/tpj.12967

Arias-Calderón, R., L. León, J. Bejarano-Alcázar, A. Belaj, R. de la Rosa, and D. Rodríguez-Jurado. 2015a. Resistance to Verticillium wilt in olive progenies from open-pollination. Sci. Hortic. 185:34–42. doi:10.1016/j.scienta.2015.01.015

Arias-Calderón, R., D. Rodríguez-Jurado, L. León, J. Bejarano-Alcázar, R. de la Rosa, and A. Belaj. 2015b. Pre-breeding for resistance to Verticillium wilt in olive: Fishing in the wild relative gene pool. Crop Prot. 75:25–33. doi:10.1016/j.cropro.2015.05.006

Aronesty, E. 2011. ea-utils: Command-line tools for processing biological sequencing data. Expression Analysis, Durham, NC. http://github.com/ExpressionAnalysis/ea-utils (accessed 11 Jan. 2017)

Blüthgen, N., K. Brand, B. Cajavec, M. Swat, H. Herzel, and D. Beule. 2005. Biological profiling of gene groups utilizing Gene Ontology. Genome Inf. 16:106–115.

Bubici, G., and M. Cirulli. 2012. Control of Verticillium wilt of olive by resistant rootstocks. Plant Soil 352:363–376. doi:10.1007/s11104-011-1002-9

Bubici, G., and M. Cirulli. 2014. Natural recovery from Verticillium wilt in olive: Can it be exploited in a control strategy? Plant Soil 381:85–94. doi:10.1007/s11104-014-2112-y

Calderón, R., C. Lucena, J.L. Trapero-Casas, P.J. Zarco-Tejada, and J.A. Navas-Cortés. 2014. Soil temperature determines the reaction of olive cultivars to *Verticillium dahliae* pathotypes. PLoS ONE 9(10):e110664. doi:10.1371/journal.pone.0110664

Carmona, R., A. Zafra, P. Seoane, A.J. Castro, D. Guerrero-Fernández, T. Castillo-Castillo, et al. 2015. ReprOlive: A database with linked data for the olive tree (*Olea europaea* L.) reproductive transcriptome. Front. Plant Sci. 6:625. doi:10.3389/fpls.2015.00625

Chen, J., J. Huang, N. Li, X. Ma, J.W. Wang, C.M. Liu, et al. 2015. Genome-wide analysis of the gene families of resistance gene analogues in cotton and their response to Verticillium wilt. BMC Plant Biol. 15:148. doi:10.1186/s12870-015-0508-3

Claros, M.G., R. Bautista, D. Guerrero-Fernández, H. Benzerki, P. Seoane, and N. Fernández-Pozo. 2012. Why assembling plant genome sequences is so challenging. Biology (Basel) 1:439–459. doi:10.3390/biology1020439

Collado-Romero, M., J. Mercado-Blanco, C. Olivares-García, A. Valverde-Corredor, and R.M. Jiménez-Díaz. 2006. Molecular variability within and among *Verticillium dahliae* vegetative compatibility groups determined by fluorescent amplified fragment length polymorphism

and polymerase chain reaction markers. Phytopathology 96:485–495. doi:10.1094/PHYTO-96-0485
Conesa, A., S. Götz, J.M. Garcia-Gomez, J. Terol, M. Talon, and M. Robles. 2005. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674–3676. doi:10.1093/bioinformatics/bti610
Daayf, F. 2015. Verticillium wilts in crop plants: Pathogen invasion and host defense responses. Can. J. Plant Pathol. 37:8–20. doi:10.1080/07060661.2014.989908
De la Rosa, R., A. Angiolillo, C. Guerrero, M. Pellegrini, L. Rallo, G. Besnard, et al. 2003. A first linkage map of olive (*Olea europaea* L.) cultivars using RAPD, AFLP, RFLP and SSR markers. Theor. Appl. Genet. 106:1273–1282. doi:10.1007/s00122-002-1189-5
Derksen, H., M. Badawi, M.A. Henriquez, Z. Yao, A.F. El-Bebany, and F. Daayf. 2013. Differential expression of potato defence genes associated with the salicylic acid defence signalling pathway in response to weakly and highly aggressive isolates of *Verticillium dahliae*. J. Phytopathol. 161:142–153. doi:10.1111/jph.12038
Dervis, S., J. Mercado-Blanco, L. Erten, A. Valverde-Corredor, and E. Pérez-Artés. 2010. Verticillium wilt of olive in Turkey: A survey on disease importance, pathogen diversity and susceptibility of relevant olive cultivars. Eur. J. Plant Pathol. 127:287–301. doi:10.1007/s10658-010-9595-z
Donaire, L., L. Pedrola, R. de la Rosa, and C. Llave. 2011. High-throughput sequencing of RNA silencing-associated small RNAs in olive (*Olea europaea* L.). PLoS One 6(11):e27916. doi:10.1371/journal.pone.0027916
Duressa, D., A. Anchieta, D.M. Chen, A. Klimes, M.D. Garcia-Pedrajas, K.F. Dobinson, and S.J. Klosterman. 2013. RNA-seq analyses of gene expression in the microsclerotia of *Verticillium dahliae*. BMC Genomics 14:607. doi:10.1186/1471-2164-14-607
Erten, L., and M. Yildiz. 2011. Screening for resistance of Turkish olive cultivars and clonal rootstocks to Verticillium wilt. Phytoparasitica 39:83–92. doi:10.1007/s12600-010-0136-2
Faino, L., R. de Jonge, and B.P.H.J. Thomma. 2012. The transcriptome of *Verticillium dahliae*-infected *Nicotiana benthamiana* determined by deep RNA sequencing. Plant Signal. Behav. 7(9):1065–1069. doi:10.4161/psb.21014
Fradin, E.F., and B.P.H.J. Thomma. 2006. Physiology and molecular aspects of Verticillium wilt diseases caused by *V. dahliae* and *V. albo-atrum*. Mol. Plant Pathol. 7:71–86. doi:10.1111/j.1364-3703.2006.00323.x
García-López, M.C., I. Vidoy, J. Jiménez-Ruiz, A. Muñoz-Mérida, A. Fernández-Ocaña, R. de la Rosa, et al. 2014. Genetic changes involved in the juvenile-to-adult transition in the shoot apex of *Olea europaea* L. occur years before the first flowering. Tree Genet. Genomes 10:585–603. doi:10.1007/s11295-014-0706-4
Garcia-Ruiz, G.M., C. Trapero, C. del Rio, and F.J. Lopez-Escudero. 2014. Evaluation of resistance of Spanish olive cultivars to *Verticillium dahliae* in inoculations conducted in greenhouse. Phytoparasitica 42:205–212. doi:10.1007/s12600-013-0353-6
García-Ruiz, G.M., C. Trapero, A. Varo-Suárez, A. Trapero, and F.J. López-Escudero. 2015. Identifying resistance to Verticillium wilt in local Spanish olive cultivars. Phytopathol. Mediterr. 54:453–460. doi:10.14601/Phytopathol_Mediterr-15130
Gayoso C., F. Pomar, E. Novo-Uzal, F. Merino, and O. Martínez de Ilárduya. 2010. The *Ve*-mediated resistance response of the tomato to *Verticillium dahliae* involves H_2O_2, peroxidase and lignins and drives *PAL*-gene expression. BMC Plant Biol. 10:232. doi:10.1186/1471-2229-10-232
Gharbi, Y., H. Alkher, M.A. Triki, M. Barkallah, B. Emna, R. Trabelsi, et al. 2015a. Comparative expression of genes controlling cell wall-degrading enzymes in *Verticillium dahliae* isolates from olive, potato and sunflower. Physiol. Mol. Plant Pathol. 91. doi:10.1016/j.pmpp.2015.05.006
Gharbi, Y., M.A. Triki, R. Trabelsi, I. Fandri, F. Daayf, and R. Gdoura. 2015b. Genetic structure of *Verticillium dahliae* isolates infecting olive trees in Tunisia using AFLP, pathogenicity and PCR markers. Plant Pathol. 64:871–879. doi:10.1111/ppa.12323
Gómez-Lama Cabanás, C., E. Schilirò, A. Valverde-Corredor, and J. Mercado-Blanco. 2015. Systemic responses in a tolerant olive (*Olea europaea* L.) cultivar upon root colonization by the vascular pathogen *Verticillium dahliae*. Front. Microbiol. 6:928. doi:10.3389/fmicb.2015.00928
Goodstein, D.M., S. Shu, R. Howson, R. Neupane, R.D. Hayes, J. Fazo, et al. 2012. Phytozome: A comparative platform for green plant genomics. Nucleic Acids Res. 40:D1178–D1186. doi:10.1093/nar/gkr944
Grabherr, M.G., B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson, I. Amit, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29:644–652. doi:10.1038/nbt.1883
Hu, D., C. Wang, F. Tao, Q. Cui, X. Xu, W. Shang, and X. Hu. 2014. Whole genome wide expression profiles on germination of *Verticillium dahliae* microsclerotia. PLoS One 9(6):e100046. doi:10.1371/journal.pone.0100046
Hübner, S., A.B. Korol, and K.J. Schmid. 2015. RNA-Seq analysis identifies genes associated with differential reproductive success under drought-stress in accessions of wild barley *Hordeum spontaneum*. BMC Plant Biol. 15:134. doi:10.1186/s12870-015-0528-z
Hunter, S., P. Jones, A. Mitchell, R. Apweiler, T.K. Attwood, A. Bateman, et al. 2012. InterPro in 2011: New developments in the family and domain prediction database. Nucleic Acids Res. 40(D1):D306–D312. doi:10.1093/nar/gkr948
Jiménez-Díaz, R.M., M. Cirulli, G. Bubici, M.M. Jiménez-Gasco, P.P. Antoniou, and E.C. Tjamos. 2012. Verticillium wilt: A major threat to olive production. Current status and future prospects for its management. Plant Dis. 96:304–329. doi:10.1094/PDIS-06-11-0496
Kaya, H.B., O. Cetin, H. Kaya, M. Sahin, F. Sefer, A. Kahraman, and B. Tanyolac. 2013. SNP discovery by Illumina-based transcriptome sequencing of the olive and the genetic characterization of Turkish olive genotypes revealed by AFLP, SSR and SNP markers. PLoS One 8(9):e73674. doi:10.1371/journal.pone.0073674
Klosterman, S.J., Z.K. Atallah, G.E. Vallad, and K.V. Subbarao. 2009. Diversity, pathogenicity, and management of Verticillium species. Annu. Rev. Phytopathol. 47:39–62. doi:10.1146/annurev-phyto-080508-081748
Lamesch, P., T.Z. Berardini, D. Li, D. Swarbreck, C. Wilks, R. Sasidharan, et al. 2012. The Arabidopsis Information Resource (TAIR): Improved gene annotation and new tools. Nucleic Acids Res. 40(D1):D1202–D1210. doi:10.1093/nar/gkr1090
Leyva-Pérez, M.O., A. Valverde-Corredor, R. Valderrama, J. Jiménez-Ruiz, A. Muñoz-Merida, O. Trelles, et al. 2014. Early and delayed long-term transcriptional changes and short-term transient responses during cold acclimation in olive leaves. DNA Res. 22(1):1–11. doi:10.1093/dnares/dsu033
Li, B., and C. Dewey. 2011. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12:323. doi:10.1186/1471-2105-12-323
López-Escudero, F.J., M.A. Blanco-López, C.D.R. Rincón, and J.M.C. Reig. 2007. Response of olive cultivars to stem puncture inoculation with a defoliating pathotype of *Verticillium dahliae*. HortScience 42:294–298.
López-Escudero, F.J., and J. Mercado-Blanco. 2011. Verticillium wilt of olive: A case study to implement an integrated strategy to control a soil-borne pathogen. Plant Soil 344:1–50. doi:10.1007/s11104-010-0629-2
López-Escudero, F.J., J. Mercado-Blanco, J.M. Roca, A. Valverde-Corredor, and M.A. Blanco-López. 2010. Verticillium wilt of olive in the Guadalquivir Valley (southern Spain): Relations with some agronomical factors and spread of *Verticillium dahliae*. Phytopathol. Mediterr. 49:370–380. doi:10.14601/Phytopathol_Mediterr-3154
Lusk, R.W. 2014. Diverse and widespread contamination evident in the unmapped depths of high throughput sequencing data. PLoS One 9(10):e110808. doi:10.1371/journal.pone.0110808
Maldonado-González, M.M., P.A.H.M. Bakker, P. Prieto, and J. Mercado-Blanco. 2015. *Arabidopsis thaliana* as a tool to identify traits involved in *Verticillium dahliae* biocontrol by the olive root endophyte *Pseudomonas fluorescens* PICF7. Front. Microbiol. 6:266. doi:10.3389/fmicb.2015.00266
Markakis, E.A., S.E. Tjamos, P.P. Antoniou, E.J. Paplomatas, and E.C. Tjamos. 2009. Symptom development, pathogen isolation and Real Time QPCR quantification as factors for evaluating the resistance of olive cultivars to *Verticillium* pathotypes. Eur. J. Plant Pathol. 124:603–611. doi:10.1007/s10658-009-9447-x
Markakis, E.A., S.E. Tjamos, P.P. Antoniou, P.A. Roussos, E.J. Paplomatas, and E.C. Tjamos. 2010. Phenolic responses of resistant and susceptible olive cultivars induced by defoliating and nondefoliating

Verticillium dahliae pathotypes. Plant Dis. 94:1156–1162. doi:10.1094/PDIS-94-9-1156

Martin, L.B.B., Z. Fei, J.J. Giovannonim, and J.K.C. Rose. 2013. Catalyzing plant science research with RNA-seq. Front. Plant Sci. 4:66. doi:10.3389/fpls.2013.00066

Mata-Pérez, C., B. Sánchez-Calvo, J.C. Begara-Morales, F. Luque, J. Jiménez-Ruiz, M.N. Padilla, et al. 2015. Transcriptomic profiling of linolenic acid-responsive genes in ROS signaling from RNA-seq data in Arabidopsis. Front. Plant Sci. 6:122. doi:10.3389/fpls.2015.00122

Mercado-Blanco, J., M. Collado-Romero, S. Parrilla-Araujo, D. Rodrıìguez-Jurado, and R.M. Jiménez-Díaz. 2003. Quantitative monitoring of colonization of olive genotypes by *Verticillium dahliae* pathotypes with real-time polymerase chain reaction. Physiol. Mol. Plant Pathol. 63:91–105. doi:10.1016/j.pmpp.2003.10.001

Muñoz-Mérida, A., J.J. González-Plaza, A.M. Blanco, M.C. García-López, J.M. Rodríguez, L. Pedrola, et al. 2013. De novo assembly and functional annotation of the olive (*Olea europaea*) transcriptome. DNA Res. 20:93–108. doi:10.1093/dnares/dss036

Pegg, G.F., and B.L. Brady. 2002. Verticillium wilts. CAB Int., Wallingford, UK. doi:10.1079/9780851995298.0000

Prieto, P., and J. Mercado-Blanco. 2008. Endophytic colonization of olive roots by the biocontrol strain *Pseudomonas fluorescens* PICF7. FEMS Microbiol. Ecol. 64:297–306. doi:10.1111/j.1574-6941.2008.00450.x

Prieto, P., C. Navarro-Raya, A. Valverde-Corredor, S.G. Amyotte, K.F. Dobinson, and J. Mercado-Blanco. 2009. Colonization process of olive tissues by *Verticillium dahliae* and its in planta interaction with the biocontrol root endophyte *Pseudomonas fluorescens* PICF7. Microb. Biotechnol. 2:499–511. doi:10.1111/j.1751-7915.2009.00105.x

Robinson, M.D., D.J. McCarthy, and G.K. Smyth. 2010. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140. doi:10.1093/bioinformatics/btp616

Ruggieri, G. 1946. Una nuova malattia dell'olivo. L'Italia Agricola. 83:369–372.

Schmieder, R., and R. Edwards. 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27:863–864. doi:10.1093/bioinformatics/btr026

Sesli, M., E. Onan, S. Oden, H. Yener, and E.D. Yegenoglu. 2010. Resistance of olive cultivars to *Verticillium dahliae*. Sci. Res. Essays 5:1561–1565.

Shearman, J.R., D. Sangsrakru, N. Jomchai, P. Ruang-Areerate, C. Sonthirod, C. Naktang, et al. 2015. SNP identification from RNA sequencing and linkage map construction of rubber tree for anchoring the draft genome. PLoS One 10(4):e0121961. doi:10.1371/journal.pone.0121961

Simpson, J.T., K. Wong, S.D. Jackman, J.E. Schein, S.J.M. Jones, and I. Birol. 2009. ABySS: A parallel assembler for short read sequence data. Genome Res. 19:1117–1123. doi:10.1101/gr.089532.108

Stergiopoulos, I., and P.J.G.M. de Wit. 2009. Fungal effector proteins. Annu. Rev. Phytopathol. 47:233–263. doi:10.1146/annurev.phyto.112408.132637

Strickler, S.R., A. Bombarely, and L.A. Mueller. 2012. Designing a transcriptome next-generation sequencing project for a nonmodel plant species. Am. J. Bot. 99:257–266. doi:10.3732/ajb.1100292

Sun, Q., H. Jiang, X. Zhu, W. Wang, X. He, Y. Shi, et al. 2013. Analysis of sea-island cotton and upland cotton in response to *Verticillium dahliae* infection by RNA sequencing. BMC Genomics 14:852. doi:10.1186/1471-2164-14-852

Talboys, P.W. 1960. A culture-medium aiding the identification of *Verticillium albo-atrum* and *V. dahliae*. Plant Pathol. 9:57–58. doi:10.1111/j.1365-3059.1960.tb01147.x

Tan, G., K. Liu, J. Kang, K. Xu, Y. Zhang, L. Hu, et al. 2015. Transcriptome analysis of the compatible interaction of tomato with *Verticillium dahliae* using RNA-sequencing. Front. Plant Sci. 6:428. doi:10.3389/fpls.2015.00428

Trapero, C., L. Rallo, F.J. López-Escudero, D. Barranco, and C.M. Díez. 2015. Variability and selection of Verticillium wilt resistant genotypes in cultivated olive and in the *Olea* genus. Plant Pathol. 64:890–900. doi:10.1111/ppa.12330

Trapero, C., N. Serrano, O. Arquero, C. del Río, A. Trapero, and F.J. López-Escudero. 2013. Field resistance to Verticillium wilt in selected olive cultivars grown in two naturally infested soils. Plant Dis. 97:668–674. doi:10.1094/PDIS-07-12-0654-RE

Wang, Z., M. Gerstein, and M. Snyder. 2009. RNA-Seq: A revolutionary tool for transcriptomics. Nat. Rev. Genet. 10:57–63. doi:10.1038/nrg2484

Warren, R.L., C.I. Keeling, M.M.S. Yuen, A. Raymond, G.A. Taylor, B.P. Vandervalk, et al. 2015. Improved white spruce (*Picea glauca*) genome assemblies and annotation of large gene families of conifer terpenoid and phenolic defense metabolism. Plant J. 83:189–212. doi:10.1111/tpj.12886

Xie, C., C. Wang, X. Wang, and X. Yang. 2013. Proteomics-based analysis reveals that *Verticillium dahliae* toxin induces cell death by modifying the synthesis of host proteins. J. Gen. Plant Pathol. 79:335–345. doi:10.1007/s10327-013-0467-1

Xie, Y., G. Wu, J. Tang, R. Luo, J. Patterson, S. Liu, et al. 2014. SOAPdenovo-Trans: De novo transcriptome assembly with short RNA-Seq reads. Bioinformatics 30:1660–1666. doi:10.1093/bioinformatics/btu077

Xiong, D., Y. Wang, J. Ma, S.J. Klosterman, S. Xiao, and C. Tian. 2014. Deep mRNA sequencing reveals stage-specific transcriptome alterations during microsclerotia development in the smoke tree vascular wilt pathogen, *Verticillium dahliae*. BMC Genomics 15:324. doi:10.1186/1471-2164-15-324

Xiong, D., Y. Wang, L. Tian, and C. Tian. 2016. MADS-Box transcription factor *VdMcm1* regulates conidiation, microsclerotia formation, pathogenicity, and secondary metabolism of *Verticillium dahliae*. Front. Microbiol. 7:1192. doi:10.3389/fmicb.2016.01192

Xu, L., L. Zhu, L. Tu, L. Liu, D. Yuan, L. Jin, et al. 2011. Lignin metabolism has a central role in the resistance of cotton to the wilt fungus *Verticillium dahliae* as revealed by RNA-Seq-dependent transcriptional analysis and histochemistry. J. Exp. Bot. 62:5607–5621. doi:10.1093/jxb/err245

Yanik, H., M. Turktas, E. Dundar, P. Hernandez, G. Dorado, and T. Unver. 2013. Genome-wide identification of alternate bearing-associated microRNAs (miRNAs) in olive (*Olea europaea* L.). BMC Plant Biol. 13:10. doi:10.1186/1471-2229-13-10

Zhang, Y., X.F. Wang, Z.G. Ding, Q. Ma, G.R. Zhang, S.L. Zhang, et al. 2013. Transcriptome profiling of *Gossypium barbadense* inoculated with *Verticillium dahliae* provides a resource for cotton improvement. BMC Genomics 14:637. doi:10.1186/1471-2164-14-637