# Massive sequencing of *Ulmus minor*'s transcriptome provides new molecular tools for a genus under the constant threat of Dutch elm disease

Pedro Perdiguero[1,2], Martin Venturas[1], María Teresa Cervera[2,3], Luis Gil[1,2] and Carmen Collada[1,2]*

[1] Grupo de Investigación en Genética, Fisiología e Historia Forestal, Departamento de Sistemas y Recursos Naturales, Universidad Politécnica de Madrid, Madrid, Spain, [2] Unidad Mixta de Genómica y Ecofisiología Forestal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria/Universidad Politécnica de Madrid, Madrid, Spain, [3] Departamento de Ecología y Genética, Centro de Investigación Forestal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Madrid, Spain

Elms, especially *Ulmus minor* and *U. americana*, are carrying out a hard battle against Dutch elm disease (DED). This vascular wilt disease, caused by *Ophiostoma ulmi* and *O. novo-ulmi*, appeared in the twentieth century and killed millions of elms across North America and Europe. Elm breeding and conservation programmes have identified a reduced number of DED tolerant genotypes. In this study, three *U. minor* genotypes with contrasted levels of tolerance to DED were exposed to several biotic and abiotic stresses in order to (i) obtain a *de novo* assembled transcriptome of *U. minor* using 454 pyrosequencing, (ii) perform a functional annotation of the assembled transcriptome, (iii) identify genes potentially involved in the molecular response to environmental stress, and (iv) develop gene-based markers to support breeding programmes. A total of 58,429 putative unigenes were identified after assembly and filtering of the transcriptome. 32,152 of these unigenes showed homology with proteins identified in the genome from the most common plant model species. Well-known family proteins and transcription factors involved in abiotic, biotic or both stresses were identified after functional annotation. A total of 30,693 polymorphisms were identified in 7,125 isotigs, a large number of them corresponding to single nucleotide polymorphisms (SNPs; 27,359). In a subset randomly selected for validation, 87% of the SNPs were confirmed. The material generated may be valuable for future *Ulmus* gene expression, population genomics and association genetics studies, especially taking into account the scarce molecular information available for this genus and the great impact that DED has on elm populations.

Keywords: Dutch elm disease, next-generation sequencing, SNPs, transcriptome, *Ulmus*

## Introduction

Owing to their long life cycles, forest tree species need to develop evolutionary mechanisms and processes to cope with different abiotic and biotic stresses. Plant responses to simultaneous stresses are complex, and may result in antagonistic, synergistic, or additive interactions (Suzuki et al., 2014). The results will depend on the timing, nature, and severity of each stress

(Prasch and Sonnewald, 2015). Current climate change is drastically affecting Mediterranean regions and promoting the occurrence of numerous situations of combined environmental stresses. For example, reduction of water availability in the Iberian Peninsula, often accompanied by pest outbreaks, is causing important damage to its forests (Sánchez-Salguero et al., 2012).

Field elm (*Ulmus minor* Mill.) is a riparian noble hardwood tree which is affected by a multiple stress scenario like the one described above. On the one hand, climate change models predict an increase in temperature, aridity and extreme events throughout its natural distribution, composed mainly by central and southern Europe. Although the species is well adapted to a wide range of pluviometric conditions, with annual precipitation ranging from 350 to 900 mm as well as a 3- to 4.5-month summer drought period (Venturas et al., 2013b), this species is expected to face more severe and restrictive environmental conditions. On the other hand, and more importantly, Dutch elm disease (DED) is an important threat to elm populations. DED is a wilt disease caused by two fungi, *Ophiostoma ulmi* (Buisman) Nannf. (moderately pathogenic) and *O. novo-ulmi* Brasier (highly pathogenic). Two successive pandemics of DED that occurred during the twentieth century caused the death of most adult trees in many locations of Europe and North America, including those used as ornamental trees in urban areas (Brasier, 2000). The combined effects of climate change and DED, which is really difficult to control due to its high impact and fast expansion, will jeopardize the survival of the remaining elm forests during the next decades.

Tolerant elm genotype selection and breeding has been the most successful strategy for elm recovery (Santini et al., 2004; Solla et al., 2005a). Several breeding programmes for *Ulmus* sp. affected by DED have been developed since the disease appeared (Heybroek, 1993; Smalley and Guries, 1993; Santini et al., 2004). Usually Asian elms that show low susceptibility to the disease (f.i. *U. pumila* L. or *U. wallichiana* Planch.) have been used as a source of resistance in these breeding programs. The inherent adaptive capacity of *U. minor* has been investigated by the Spanish Elm Breeding and Conservation Programme (Martín et al., 2012). This programme has identified several Iberian genotypes with a high degree of tolerance to DED when they were artificially inoculated with an aggressive strain of *O. novo-ulmi* (Solla et al., 2005a; Martín et al., 2015). Genetic heritability of resistance has been shown both by crossing native and Asian elms (Smalley and Guries, 2000; Santini et al., 2011) and by crossing susceptible and tolerant Iberian *U. minor* genotypes (Venturas et al., 2014).

In spite of these results, the molecular knowledge of *Ulmus* sp. is scarce. The major information available is an EST database with 52,823 unique transcripts including defense response genes activated in *U. minor* leaves after egg laying by the elm leaf beetle (*Xanthogaleruca luteola* Mull.) or the application of methyl jasmonates (Büchel et al., 2012). Only a few genes from *U. americana* L. upregulated during *O. novo-ulmi* infection have been identified and analyzed using *in vivo* or *in vitro* interaction (Nasmith et al., 2008; Aoun et al., 2010). Noteworthy, these studies did not use tolerant genotypes. Regarding molecular

markers, a collection of nuclear microsatellites (Simple Sequence Repeats, SSRs) have been developed for *Ulmus* sp. (Whiteley et al., 2003; Collada et al., 2004; Zalapa et al., 2008) and have been commonly used in different diversity studies (Zalapa et al., 2009; Brunet et al., 2013; Venturas et al., 2013a; Fuentes-Utrilla et al., 2014a,b). In contrast, no single nucleotide polymorphisms (SNPs) have yet been described for these species.

Knowledge of a species transcriptome is very useful for researchers who work with non-model plant species (Garg and Jain, 2013). Next-generation sequencing (NGS) technologies have greatly facilitated transcriptome sequencing of species with scarce or without prior gene sequence information (Ekblom and Galindo, 2010; Neale and Kremer, 2011). These NGS techniques have been used by forest researchers for gene discovery and annotation, large-scale expression analysis and marker discovery (Novaes et al., 2008; Bräutigam and Gowik, 2010; Parchman et al., 2010; Ge et al., 2011; Pinosio et al., 2014). Taking into account that SNPs are the most abundant genetic variations in genomes, this information is highly useful in breeding and conservation programs. Thus, SNP-based markers are replacing or being combined with SSRs and other molecular markers for construction of dense genetic maps of trees and other plants (Pavy et al., 2008; Vezzulli et al., 2008; Chancerel et al., 2011; Gaur et al., 2012) allowing a more effective dissection of traits of interest.

The available *Ulmus* sp. molecular resources are clearly insufficient to preserve and manage the threatened elm populations (Bernier et al., 2015). Large-scale development of SNP markers would allow researchers to use more efficient and powerful tools in the management and conservation of *U. minor* genetic resources and those of other Ulmaceae species. In this study we use GS-FLX 454 pyrosequencing to study elm transcriptome responses to abiotic and biotic stresses. We designed this study to (i) construct a non-normalized library covering responses to the most important biotic and abiotic stresses suffered by three *U. minor* genotypes with different degrees of tolerance to DED, (ii) perform the assembly and functional annotation of the transcriptome, (iii) identify those genes potentially involved in abiotic and biotic stress responses, and (iv) identify potential functional gene-based markers (SNPs). The three genotypes were inoculated with *O. novo-ulmi* and *O. ulmi*, the causal agents of DED, as well as with the endophytic fungus *Daldinia concentrica* (Bolton) Cesati & de Notaris. These genotypes were also exposed to water stress. In order to obtain the molecular response of all plant organs, the different tissues (roots, stems, and leaves) were collected separately, and their RNA extracted. Afterward, RNA of all tissues was mixed in equal proportions and all samples pooled for sequencing.

## Materials and Methods

### Plant Material

Clonal material from three Iberian *U. minor* genotypes with different degrees of tolerance to DED was used for this study. Two of these genotypes, M-DV4/5 from Dehesa de la Villa

(Madrid) and J-CA2 from Sierra de Cazorla (Jaén), showed a moderate to high degree of tolerance to DED, and the third, TO-AL1 from Algodor (Toledo), was highly susceptible to the disease. These genotypes have been widely used as parents in controlled crossed carried out within the Spanish Elm Breeding and Conservation Programme (Venturas et al., 2014). Plants were propagated by stem cuttings and grown in 2 L pots (1:3 perlite:peat substrate, v/v) at Puerta de Hierro nursery (Madrid). Plants were well watered. One-year-old plants were subjected to four different treatments: drought and three different inoculations with *O. novo-ulmi* (strain CU-HU), *O. ulmi* (PM-VA2) or *D. concentrica* (X23-1).

One plant per genotype and treatment was collected at every sampling point. For the water stress treatment, plants were watered to field capacity at the beginning of experiment and were collected after 48 h and 7 days without irrigation. For the analysis of elm response to *O. ulmi* and *O. novo-ulmi*, trees were inoculated with $10^6$ spores ml$^{-1}$ according to Solla et al. (2005a) and plant material was collected 24 h and 7 days post-inoculation. These plants showed wilting symptoms after inoculation, especially the susceptible clone, but this process was not quantified. For the analysis of elm response to *D. concentrica*, inoculation was conducted removing a small piece of bark (1 cm$^2$) from the main stem and inserting *D. concentrica* mycelium in the cambium. Immediately after inoculation the wound was sealed with parafilm. Plant material of this treatment was also collected 24 h and 7 days after inoculation. Unstressed plants of the three genotypes were harvested at the end of treatments. The plants were dissected into leaves, stems, and roots. Tissues were immediately frozen in liquid nitrogen and stored at −80°C.

## RNA Extraction and cDNA Synthesis

Total RNA for each sampling point was extracted separately from roots, stems, and leaves of each sampled plant following the CTAB–LiCl precipitation method (Chang et al., 1993). Equal amounts of total RNA from roots, stems, and leaves from all the plants were mixed to form one RNA pool to maximize the diversity of transcripts obtained. RNA was DNase treated to remove genomic contamination and purified using the Qiagen RNeasy kit (Qiagen, CA, USA). cDNA was synthesized from 5 μg de RNA using SMART approach (Zhu et al., 2001). Non-normalized cDNA was purified using QIAquick PCR Purification Kit (Qiagen, CA, USA).

## Roche 454 Sequencing and Assembly

Five μg of non-normalized cDNA were used for library construction and sequencing using GS-FLX Titanium System (Roche). The sample was nebulized, adaptor-ligated, and pyrosequenced (1 plate). Sequencing was performed by the Lifesequencing service (Valencia University, Spain). Sequences were *de novo* assembled with Newbler version 2.3 (Roche). Using this software reads were filtered by successive steps. In the first step low quality reads were rejected. In a second step the remaining reads were trimmed and classified. Those that were too short (reads lower than 100 pb) and outliers (problematic reads such as chimeras) were rejected, whereas singletons (no

significant overlapping with other reads) were considered unique transcripts at this point. Finally, the remaining repeated, partially assembled and assembled reads were used to construct isotigs. The minimum match length for overlapping was fixed in 40 pb and the minimum identity percentage in 90% maintaining the rest of parameters as default.

## Sequence Analysis and Annotation

The transcriptome assembly (isotigs from assembled reads and singletons) was used as query for a BLASTx in two steps. Initially all sequences were compared with RefSeq proteins from highly annotated genus (*Arabidopsis*, *Populus*, *Oryza*, *Solanum*, *Glycine*, *Vitis*, *Medicago*, *Ricinus*, and *Malus*) using an E-value threshold of $10^{-5}$. Sequences without blast results were compared in a second step with the remaining plant RefSeq protein database. In order to estimate the number of different Unigenes identified, two independent steps of filtering were performed to remove any possible source of contamination from the final transcriptome and to avoid redundancy as far as possible. On the one hand, sequences without blast results were analyzed using SeqTrimNext software (advanced version of SeqTrim, Falgueras et al., 2007) with a specific template customized for Roche 454 GS-FLX plant transcriptomic data. To identify and reject sequences from inoculated organisms an additional comparison was performed with a database containing sequences from *O. novo-ulmi* and *O. ulmi* genomes (Forgetta et al., 2013; Khoshraftar et al., 2013), from *D. concentrica* (48 sequences in GenBank database) and sequences from *D. eschscholzii* genome (Ng et al., 2012). By this analysis a set of sequences that present homology with inoculated organisms were identified and analyzed with fungus RefSeq protein database, and also with protein dataset from the last annotation available for *O. novo-ulmi* (Comeau et al., 2015). Contaminant sequences from other common microorganisms and putative artifacts were identified and rejected. On the other hand, sequences sharing the best blast hit were considered the parts of same Unigene.

BLASTx annotations were conducted with version V.2.7.2 of Blas2GO for the Gene Ontology (GO) assignment. GO terms coming from the three different GO ontologies (biological process, molecular function and cellular component) were analyzed separately. Annotation results were filtered according to Plant GO-Slim, a reduced version of GO adapted for plants. The main goal of this analysis is to provide a more general and homogeneous perspective of the GO Terms found in a sample. Enzyme codes (ECs) and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways annotation were generated in Blast2GO from the direct mapping of the GO terms to their EC equivalents. Annotation inferred by homology with protein families, domains, and functional sites from secondary databases were obtained using InterProScan software implemented in Blast2GO.

## SNP Identification and Validation

To detect SNPs, we used the consensus assembly generated from all sequencing runs as a reference sequence to which individual reads were aligned using GS Mapper. GS Reference Mapper

only scores polymorphisms where two or more reads contain the variant allele. Only SNPs detected differences that were classified as "High Confidence" were considered for the analysis. These differences are characterized by the following features: (1) there must be at least three non-duplicate reads with the difference; (2) there must be both forward and reverse reads showing the difference, unless there are at least five reads with quality scores over 20 (or 30 if the difference involves a 5-mer or higher).

In order to validate a sample of detected SNPs, we designed primers to amplify 13 isotigs randomly selected. Genomic DNA was used as template, this was isolated according to Doyle (1990) from leaves of three different genotypes included in library constructions. Specific primers for PCR amplifications were designed using Primer 3 software (Rozen and Skaletsky, 1999; Supplementary Table S1). Sequencing chromatograms were aligned and analyzed with BioEdit software (Hall, 1999). SNPs were identified as overlapping nucleotide peaks and validated.

## Data Accessibility

Files containing 454 reads are available from National Center for Biotecnology Information (NCBI) Short Read Archive [GenBank: accession number SRR1687227]. The assembled transcriptome, annotations, and SNP identification are available from Dryad entry doi: 10.5061/dryad.ps837.

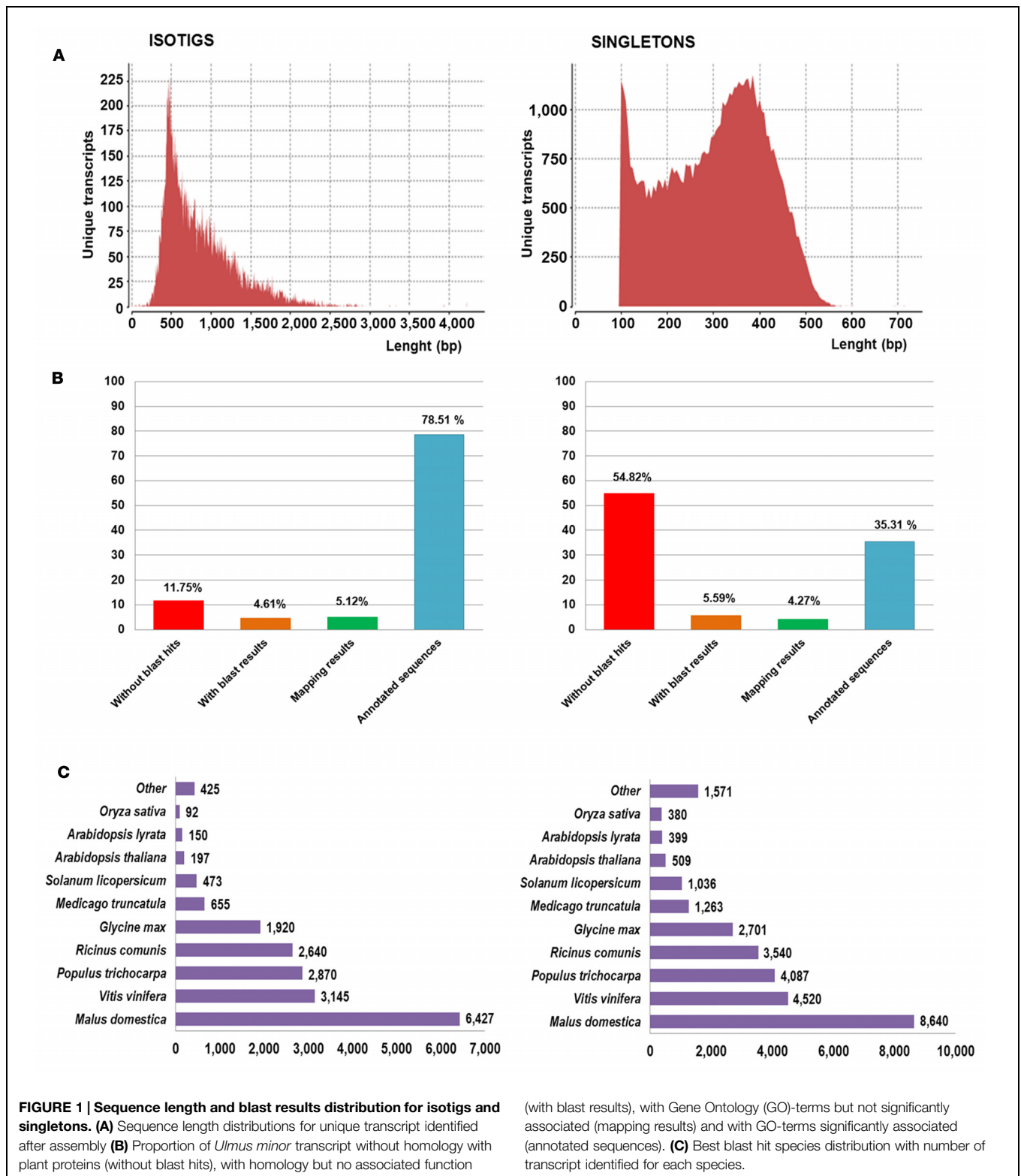## Results

### Sequencing and Assembly

The non-normalized library prepared using RNA from different genotypes, organs, and treatments included a total of 971,002 reads. During alignment, Newbler assembler trimmed the reads, and a total of 22,963 reads, corresponding to outlier or too short reads, were deleted. Before this filter, high quality reads were processed in two groups: (a) reads with significant overlapping between them (>40 bp and >90% identity) were *de novo* assembled resulting in 21,524 isotigs; the average length for all isotigs was 861 pb with a 30% of sequences ranging from 400 to 600 bp, and the N50 value for isotigs was 1,000 pb. (b) A set of 63,405 reads over 100 pb after being trimmed with Newbler did not show sufficient overlap with any other read and they were classified as singletons; the average length of these sequences was 303 pb (**Figure 1A**). Altogether, a total of 84,929 unique transcripts were obtained covering 37.77 Mbp.

### Annotation

*Ulmus minor* transcriptome (isotigs and singletons) was compared with GenBank RefSeq protein database. Blast results were found for a total of 47,640 unique transcripts. 88.14% of isotigs showed homology with proteins from Viridiplantae; for 18,707 sequences blast results were found within most common model species, whereas results were assigned for 287 isotigs in second round. Regarding singletons, blast results were obtained for 27,231 and 1,415 at the first and second round, respectively,
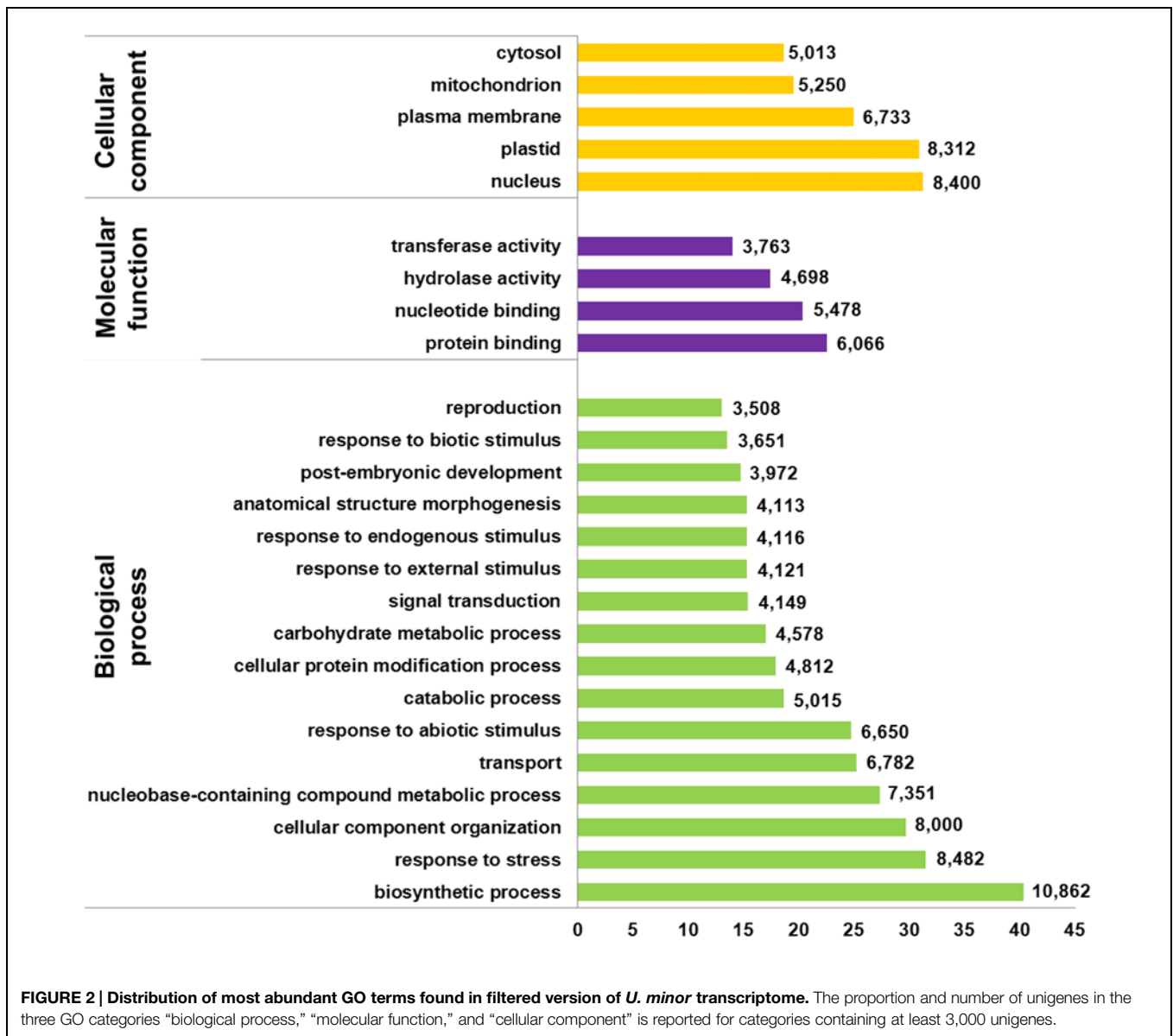
reaching the 45.18% (**Figure 1B**). Putative isoforms or different regions from the same gene were grouped in unigenes according to blast results. The 47,640 unique transcripts with blast results identified a total of 32,152 different proteins as best blast hit which were considered different unigenes. Sequences with no blast hits were analyzed using SeqTrimNext. From 37,289 unique transcripts without homology, a total of 26,277 were accepted by SeqTrimNext and considered Unigenes without homology. On the contrary 11,012 unique transcripts (10,827 singletons and 185 isotigs) were identified as artifacts or contaminations and were rejected. Noteworthy, a little group of 181 unique transcripts within these sequences showed similarities with regions from genome of inoculated organisms (Supplementary Table S2). Specifically, 146 unique transcripts found a best blast hit with *O. ulmi* genome, 6 showed the highest homology with *O. novo-ulmi*, and the remaining 29 sequences aligned best with *D. eschscholzii* genome. Commonly its homology resulted in a partial alignment with a low percentage of coverage, lower than 25% in 86 sequences. Fifty-six sequences out of 181 (31%) showed homology with proteins from fungi, out of which 20 unique transcripts showed homology with *O. novo-ulmi* proteins from Comeau et al. (2015). In summary, regarding *U. minor* transcriptome, a total of 58,429 unigenes were identified. This data represents 1.3 to 1.7% of the *U. minor* genome, considering a genome size of 2,078.25 Mbp/1C (Loureiro et al., 2007). **Figure 1C** shows the species with high number of best blast results.

Using Blast2GO gene ontology terms were assigned to unique transcripts for which a BLASTx match was obtained. A total of 26,903 unigenes were annotated with 266,843 GO terms. Most of the assignments (153,693; 58%) belonged to the biological process category, while cellular component category reached a 25% (67,419 GO terms) and molecular function was represented at a lower extent with 45,730 GO terms (17%). The distribution of most abundant GO annotations in the different functional categories evidenced a representative transcriptome of *U. minor* (**Figure 2**). All terms related to responses to stimulus are among the most abundant GO terms. Six-thousand six-hundred and fifty unigenes were associated to "response to abiotic stimulus" (GO: 0009628) whereas 3,651 unigenes were annotated for "response to biotic stimulus" (GO: 0009607). In both groups specific unigenes commonly described in response to different stresses were found; for instance, numerous *MYB* transcription factors, dehydration-responsive element-binding proteins, heat stress transcription factors, or late embryogenesis abundant proteins were identified specifically in "response to abiotic stress" group whereas leucine-rich repeat protein kinases, thaumatin-like proteins, disease resistance response proteins or wall-associated receptor kinases were commonly found in "response to biotic stress" group. One-thousand eight-hundred and thirty of these unigenes shared both GO terms which include numerous transcription factors families such as *NAC*, *ERF*, or *WRKY* as well as genes involved in common signal transduction pathways as calcium-dependent or serine threonine-protein kinases (**Figure 3**). Twenty-one isotigs were constructed with more than 1,000 reads. Among them 12 were associated with response to abiotic or biotic stimulus (**Table 1**).

**FIGURE 1 | Sequence length and blast results distribution for isotigs and singletons. (A)** Sequence length distributions for unique transcript identified after assembly **(B)** Proportion of *Ulmus minor* transcript without homology with plant proteins (without blast hits), with homology but no associated function (with blast results), with Gene Ontology (GO)-terms but not significantly associated (mapping results) and with GO-terms significantly associated (annotated sequences). **(C)** Best blast hit species distribution with number of transcript identified for each species.

Enzyme codes and KEGGs were also assigned to unigenes using Blast2GO. Nine-hundred and seven different ECs were identified and were included in 146 KEGGs pathways. Transferases and oxidoreductases were the most represented enzymes (33 and 27%, respectively), followed by hydrolases (17%). "Purine metabolism" with 55 ECs and "starch and sucrose metabolism" with 41 different ECs were the most represented pathways. Some pathways identified are relevant

**FIGURE 2 | Distribution of most abundant GO terms found in filtered version of _U. minor_ transcriptome.** The proportion and number of unigenes in the three GO categories "biological process," "molecular function," and "cellular component" is reported for categories containing at least 3,000 unigenes.
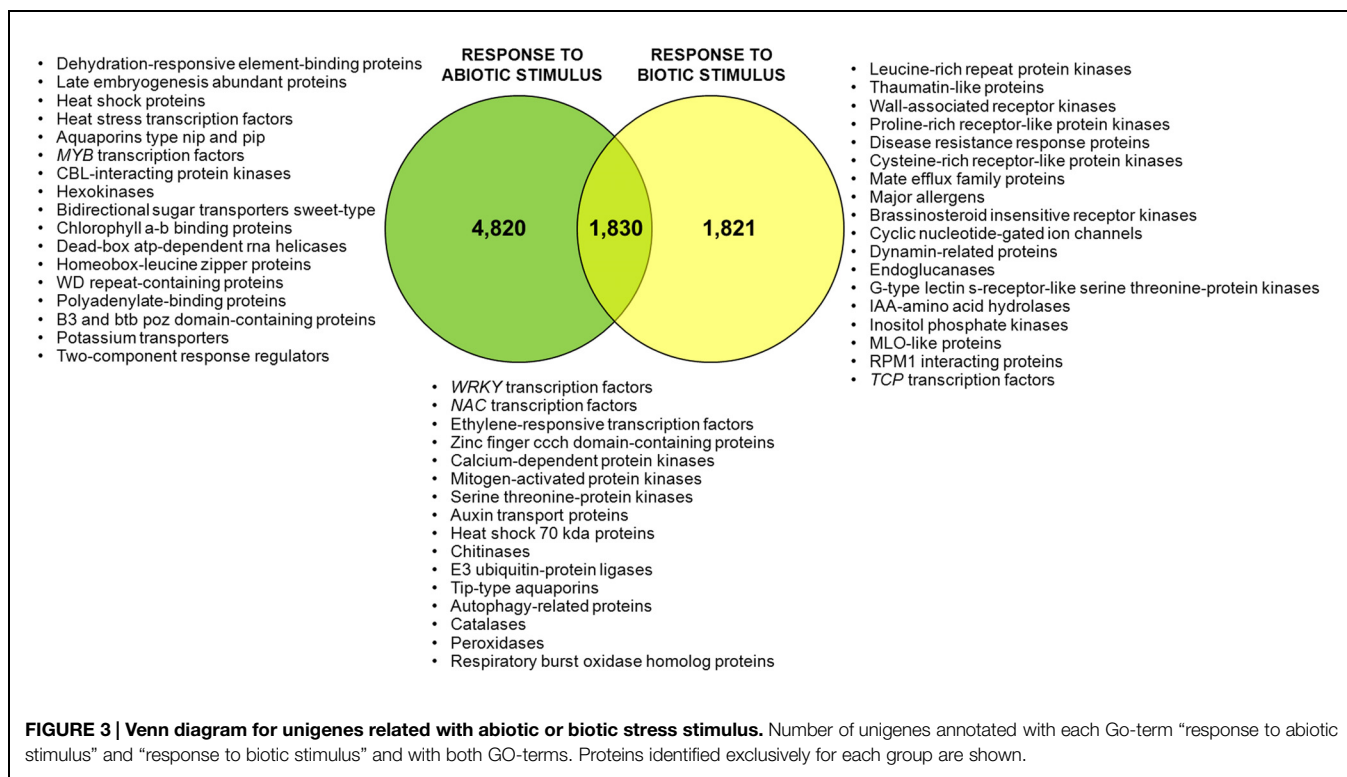
according to their implications in response to pathogens; the most significant pathways were "Phenylalanine metabolism" (represented by 25 ECs), alpha-linolenic acid metabolism (10 ECs) which results in jasmonate biosynthesis, cutin, suberine, and wax (4 ECs), flavonoids biosynthesis pathways (15 ECs), and 2 ECs (EC:2.7.10.2; protein-tyrosine kinase and EC:3.1.3.16 phosphatase) assigned to several genes involved in T-cell receptor signaling pathway.

## SNP Identification

Taking into account the use of three different genotypes for transcriptome construction a calling focused in identification of genetic variation was performed. GS Reference Mapper (454 Life Science) software was used to identify polymorphisms among ESTs by aligning individual reads against contigs from the assembly. Considering only high confidence variation, 7,125

different isotigs show a total of 30,693 polymorphic positions. Among these polymorphic variations 27,359 correspond to SNPs, 2,970 were classified as indels, and other type of variations were identified in the remaining 364 polymorphic sites. Transitions were the most common variation among SNPs, founding similar percentage regarding the type of substitution. The number of transversions identified was lower although also important (**Table 2**). To validate SNPs a sample of 13 isotigs randomly selected were PCR amplified from genomic DNA from the three genotypes used in transcriptome analysis. Each amplicon was sequenced bidirectionally (forward and reverse) using standard dideoxy-based sequencing. Long introns were observed in some genes, which were excluded from analysis during validation. From 142 identified SNPs, 123 were confirmed (87%), whereas 19 were non-validated. Also, an important number of not expected SNPs were observed

**FIGURE 3 | Venn diagram for unigenes related with abiotic or biotic stress stimulus.** Number of unigenes annotated with each Go-term "response to abiotic stimulus" and "response to biotic stimulus" and with both GO-terms. Proteins identified exclusively for each group are shown.

during validation step (14 SNPs; **Table 3**, Supplementary Table S3).

## Discussion

The aim of this work was to obtain the first transcriptome from *U. minor* in response to abiotic and biotic stress responses emphasizing on DED. Elm genotypes with different degrees of tolerance to DED were subjected to a drought treatment or inoculated with *O. ulmi*, *O. novo-ulmi*, or *D. concentrica* in order to induce transcription in response to stress. Whole plants were used in the research taking into account that spores of DED fungi germinate within stem, where hyphae grow and spread through xylem vessels inducing their blockage and cavitation. This results in foliar wilting, and in many occasions, the subsequent tree death. Additionally, DED fungi are also capable of spreading from tree to tree via root anastomosis (Neely and Himelick, 1963; Ghelardini and Santini, 2009). Therefore, we were interested in studying the transcriptome response of roots, stems and leaves.

To date no full genome from an elm species has been sequenced, and the genetic basis for tolerance to DED of *Ulmus* is practically unknown. Most of the public available information was produced to analyse *U. minor* molecular response to *Xanthogaleruca luteola* using 454 pyrosequencing methodologies to sequence different ESTs libraries (Büchel et al., 2012). This beetle can cause important damage to elm trees by defoliating them; however, the consequences of this pest are less devastating than those produced by DED which is capable of killing full elm populations (Brasier, 2000). Several works have analyzed DED

from a physiological, metabolic or anatomical perspective (e.g., Martín et al., 2005b, 2009; Solla et al., 2005b; Ďurkovič et al., 2013; Venturas et al., 2014), however few studies have investigated this disease from a molecular point of view. The published works focused on molecular response to DED used northern blots, SSH, or RT-PCR to identify and study the expression pattern of a reduced number of genes mainly during *in vitro* interaction between elms callus cells and *O. novo-ulmi* (Nasmith et al., 2008; Aoun et al., 2009, 2010).

### Sequencing and Assembly

The *U. minor* transcriptome of this study represents the largest molecular source of information from *Ulmus* genus obtained up to date. This work has significantly contributed to increasing the number of high quality reads using 454 pyrosequencing from 361,196 previously obtained by Büchel et al. (2012) to 971,002. After assembly we obtained 84,929 unique transcripts with 74.7% singletons which are consistent with previous works based on non-normalized library sequencing with GS-FLX (Wall et al., 2009). Once sequences from contaminant organisms and redundancy were removed, a total of 58,429 unigenes were identified. The high number of unigenes here obtained from plants submitted to stress increase the scarce information available about molecular response to DED. To date only 314 unisequences were isolated from *U. americana* calli inoculated with *O. novo-ulmi* (Aoun et al., 2010). It is to be noted that 88% of these unisequences have been identified in the *U. minor* transcriptome here generated. Therefore, this transcriptome represents an important contribution to advance molecular knowledge on *Ulmus* response to DED.

**TABLE 1 | Sequences with more than 1,000 reads after assembly.**

| Sequence | N° reads | Description | Protein | Species | E-value | Abiotic or biotic |
|---|---|---|---|---|---|---|
| isotig08529 | 19,048 | Senescence-associated protein | XP_003614399 | *Medicago truncatula* | 2.77E-20 | |
| isotig13329 | 15,020 | Metallothionein-like protein type 2 | XP_002299873 | *Populus trichocarpa* | 1.47E-10 | |
| isotig07130 | 2,395 | Pectin methylesterase 5 family protein | XP_002304257 | *P. trichocarpa* | 2.71E-156 | Biotic |
| isotig13408 | 2,265 | No hits found | – | – | | |
| isotig21287 | 1,951 | Hypothetical protein | XP_002865071 | *Arabidopsis lyrata* | 8.24E-07 | Abiotic |
| isotig13084 | 1,832 | Kda class i heat shock protein | XP_002298362 | *P. trichocarpa* | 6.40E-20 | Abiotic |
| isotig13342 | 1,766 | No hits found_Rejected SeqTrim | – | | | |
| isotig02731 | 1,686 | Chlorophyll a–b binding protein | XP_004248217 | *Solanum lycopersicum* | 2.03E-139 | Abiotic |
| isotig21296 | 1,605 | 60s ribosomal protein l29-1 | XP_002519200 | *Ricinus communis* | 4.85E-32 | |
| contig07984 | 1,478 | Mitochondrial protein, putative | XP_003588355 | *M. truncatula* | 1.37E-76 | |
| contig06910 | 1,475 | No hits found | – | – | | |
| contig23572 | 1,426 | Phosphoprotein ecpp44-like | XP_004238411 | *S. lycopersicum* | 4.71E-08 | Abiotic |
| isotig07047 | 1,402 | Bifunctional epoxide hydrolase 2-like | XP_003550367 | *Glycine max* | 1.34E-44 | |
| contig22192 | 1,301 | Catalase isozyme 1 | NP_001268098 | *Vitis vinifera* | 0 | Abiotic and biotic |
| isotig03470 | 1,270 | Galactinol synthase family protein | XP_002265947 | *V. vinifera* | 0 | Abiotic |
| isotig20888 | 1,156 | Kda class i heat shock protein | XP_002530396 | *R. communis* | 2.69E-61 | Abiotic |
| isotig13264 | 1,149 | Plastocyanin family protein | XP_002510603 | *R. communis* | 5.77E-74 | Abiotic |
| contig09765 | 1,120 | Kda class i heat shock protein | XP_003618790 | *M. truncatula* | 1.74E-64 | Abiotic |
| isotig18771 | 1,046 | Calmodulin-like | XP_006378996 | *P. trichocarpa* | 2.84E-100 | Abiotic |
| isotig02650 | 1,014 | Thioredoxin h2-like | XP_008361428 | *Malus domestica* | 1.66E-58 | Abiotic and biotic |
| contig07917 | 1,005 | 40s ribosomal protein s4 | XP_002515668 | *R. communis* | 8.65E-166 | |

*For each sequence the information shown is: the number of reads, protein description, best blast hit including accession number of protein from RefSeq database, species and E-value, and information relative to annotation when sequences were associated to response to abiotic or biotic stimulus.*

**TABLE 2 | Changes observed in high confidence single nucleotide polymorphisms (SNPs).**

| Mutation types | Change | N° | Total N° |
|---|---|---|---|
| Transitions | | | 16,579 |
| | a–g | 4,080 | |
| | g–a | 4,007 | |
| | t–c | 4,189 | |
| | c–t | 4,303 | |
| Transversions | | | 10,765 |
| | a–c | 1,194 | |
| | a–t | 1,665 | |
| | c–a | 1,344 | |
| | c–g | 1,170 | |
| | g–c | 1,120 | |
| | g–t | 1,481 | |
| | t–a | 1,552 | |
| | t–g | 1,239 | |
| Other | * – N or N – * | 15 | 15 |

*Number of detected changes for each variation type. N corresponds to SNPs for which three different nucleotides were observed.*

## Annotation

Blast results were associated only to 56% of unique transcripts which shows the current lack of information on *Ulmus* genome. Remarkable differences in blast results were observed between isotigs and singletons. A high percentage of blast hits were found in isotigs (88.3%) in contrast to the lower percentage of blast results obtained for singletons (45.2%). This difference is probably associated with the shorter length of singletons which has been negatively correlated with annotability of sequences (Götz et al., 2008); the average sequence length for isotigs is 821 pb whereas for singletons is 303 pb. In previous works using the same sequencing methodologies a high percentage of singletons mapped with the reference genome (Wall et al., 2009) or their presence in the original sample were validated by PCR (Meyer et al., 2009). Singletons were ignored in annotation step by Büchel et al. (2012) during the analysis of *U. minor* libraries. However, our blast results highlighted a high proportion (45%) of singletons with homology with proteins from plants, many of which were not present in isotigs. Thus, in our transcriptome singletons could represent unique genes expressed at low levels taking into account that a non-normalized library was used as template. The presence of artifacts was solved as much as possible including the SeqTrimNext step. By this software 11,012 unique transcripts without blast results (10,827 singletons and 185 isotigs) were rejected using different criterion including a database for specific contaminants. Very few of these unique transcripts showed homology with *O. novo-ulmi*, *O. ulmi*, or *Daldinia* sequences (181 unique transcripts), which reflects that the transcripts of these inoculated organisms were present in low percentage in the samples. The pooling of numerous samples and the inclusion of whole plants, not restricted to the inoculation zone, may explain this scarce presence of contaminant sequences. Additionally, the spread of the fungi could be limited by plant size as the study was performed on one-year-old ramets (Solla et al., 2005b).

**TABLE 3 | Validation of SNPs with conventional sequencing.**

| Amplicon | Sequence | Description | Number of predicted SNPs | Validated | Non-validated | New SNPs | Abiotic or biotic stress |
|----------|----------|-------------|--------------------------|-----------|---------------|----------|--------------------------|
| Amplicon1 | contig02491 | Flavonoid 3 -monooxygenase-like | 11 | 11 | 0 | 4 | Abiotic |
| Amplicon2 | contig25532 | Protein chloroplastic-like | 10 | 10 | 0 | 0 | |
| Amplicon3 | isotig02731 | Chlorophyll a–b binding protein chloroplastic-like | 14 | 13 | 1 | 0 | Abiotic |
| Amplicon4 | isotig11158 | Snf7 family protein | 10 | 7 | 3 | 2 | |
| Amplicon5 | isotig11256 | Mannose-1-phosphate guanylyltransferase alpha-like | 9 | 9 | 0 | 3 | Abiotic and biotic |
| Amplicon6 | isotig11583 | Eukaryotic initiation factor 4a-11 | 9 | 3 | 6 | 1 | Abiotic |
| Amplicon7 | isotig13076 | Lipid-transfer protein dir1 | 10 | 10 | 0 | 1 | Abiotic |
| Amplicon8 | isotig13078 | Cationic amino acid transporter vacuolar-like | 10 | 9 | 1 | 1 | Biotic |
| Amplicon9 | isotig13259 | Transcription factor myb44-like | 14 | 13 | 1 | 0 | Abiotic and biotic |
| Amplicon10 | isotig13272 | Anthocyanidin reductase | 13 | 12 | 1 | 0 | |
| Amplicon11 | isotig18769 | Atp synthase subunit | 10 | 10 | 0 | 2 | Abiotic |
| Amplicon12 | isotig19674 | Probable aldo-keto reductase 1 | 10 | 10 | 0 | 0 | |
| Amplicon13 | isotig20847 | —NA— | 12 | 6 | 6 | 0 | |
| **Total** | | | **142** | **123 (87%)** | **19 (13%)** | **14** | |

*Number and percentage of validated, non-validated and new SNPs for each of the amplicons from the tree genotypes (J-CA2, M-DV4/5, and TO-AL1) sequenced with dideoxy-based methods. Sequences associated with response to abiotic or biotic stimulus are shown.*

The species with high proportion of best blast results were *Malus domestica* Borkh., *Vitis vinifera* L. and *Populus trichocarpa* Torr. & A. Gray, and *Morus notabilis* Scneid. The number of genes deduced from the genome of these species were 57,386, 33,514, 45,654, and 29,338, respectively (Tuskan et al., 2006; Jaillon et al., 2007; Velasco et al., 2010; He et al., 2013). A similar number of genes would be expected in *U. minor*, thus the number of unigenes identified in this transcriptome (58,429) may be overestimated, suggesting a certain degree of redundancy. Moreover, considering that several fungal endophytes naturally colonize elm trees (Martín et al., 2013), it could be possible that some unigenes without blast results correspond to genes from endophytes as their genetic information is scarce in databases.

Numerous genes were annotated with GO terms associated with environmental stress. More than 3,000 unigenes were associated with each GO term related to responses to stimulus, including abiotic and biotic ones. A set of 1,830 genes share both GO terms (abiotic and biotic stress) which could point at components involved in the crosstalk between both responses (Glombitza et al., 2004; Fujita et al., 2006). Several genes and transcription factors commonly described during molecular response to environmental stresses have been identified in the transcriptome. Expression analysis was not an initial objective of our work, therefore were used pooled samples. However, considering that the majority of samples added to the mix were inoculated, the number of reads could give an idea of transcript enrichment in stress responses. It is noteworthy that more than 50% of genes with more than 1,000 reads were annotated with GO terms "response to abiotic" or "response to biotic" stimulus. Also, 10 out of 11 isotigs identified into the transcriptome are the orthologs of *U. americana* ESTs for which a notable upregulation was observed by RT-PCR during *in vitro* interaction with *O. novo-ulmi* (Aoun et al., 2010).

Within ECs and KEGGs the pathway "Phenylalanine metabolism" with 25 different ECs was remarkably represented. A central role of this pathway in the resistance to the wilt fungus *Verticillium dahliae* has been described in cotton (Xu et al., 2011). The enzyme phenylalanine ammonia lyase (PAL; EC: 4.3.1.24) has been widely used in studies focused on DED at two levels, measurement of enzymatic activity (Corchete et al., 1993) and induced transcripts (Nasmith et al., 2008). Correlation between PAL gene expression and accumulation of lignin, suberine, and phenols was observed by Aoun et al. (2009). Into this transcriptome, four ECs were identified in biosynthesis pathway of cutin, suberine, and wax and 15 ECs in flavonoids biosynthesis. The final products from these pathways have been described as resistance factors (Duchesne et al., 1985; Sticklen et al., 1991; Martín et al., 2005a). Another important pathway identified related with defense was the alpha-linolenic acid metabolism pathway that results in jasmonate biosynthesis, a plant hormone with a central role during biotic stress responses (Santino et al., 2013). It is remarkable that within T-cell receptor signaling pathway only two ECs were identified (EC:2.7.10.2 protein-tyrosine kinase and EC:3.1.3.16 phosphatase) but the number of unigenes associated to these activities (201 and 81 respectively) was higher in proportion to other pathways.

## SNP Identification

The present work represents the first massive identification of SNPs and other polymorphism in *Ulmus* species. The use of three different genotypes allowed the identification of 30,693 high fidelity polymorphic positions in 7,125 different isotigs. A validation of SNPs in a subset of isotigs by conventional PCR confirmed 87%. Similar results both in identification and validation of SNPs were observed in *Eucalyptus grandis* (Novaes et al., 2008). The majority of these new SNPs were identified by GS Reference Mapper, but they do not follow the parameters to be considered high confidence variation. Extrapolating the results

to the complete transcriptome we could expect approximately 22,500 valid SNPs, a high percentage of them localized in genes involved in abiotic or biotic response. Seven isotigs amplified for validation were associated with "response to abiotic" or "response to biotic" stimulus GO terms.

## Conclusion

The transcriptome collection represents the major genetic resource for *U. minor* focused on biotic and abiotic stress responses, including DED. This wilt disease strongly affects *Ulmus* species causing the death of numerous adult trees, which could cause genetic impoverishment of elm populations. Considering that the genotypes used as starting material showed different degree of tolerance to DED, a large number of genes and enzymes identified could be potentially involved in molecular pathways that confer this trait. In addition, the inclusion of a water stress treatment increased the number of genes involved in response to dehydration activated as an indirect effect of the pathogen which provokes vessel occlusion and cavitation. The knowledge of these genes could constitute an interesting resource for the construction of other tools such as microarrays or exome capture systems by which researchers will be able to analyze the differential expression between tolerant and susceptible genotypes as well as to study the additive effect of DED in a harsher environment caused by climate change. Identified resistance genes or transcription factors from *U. minor* could represent interesting tools for genetic manipulation of elm as a way to combat the disease; for instance, transgenic American elms overexpressing an antimicrobial peptide show reduced DED symptoms (Newhouse et al., 2007). Also a large set of SNPs markers have been discovered in this work. A lot of them are localized in genes putatively involved in response to fungus stimulus and could also be associated with genotypes showing different degrees of tolerance to DED. Thus, SNP analysis could also be designed to look for associations between one or few candidate genes and tolerance to DED in association genetics studies. Another interesting possibility would be exploring the transferability of these molecular resources involved in DED tolerance to other species generally susceptible to the disease such as *U. americana*, *U. glabra*, or *U. laevis*. Finally, the results of this study will help to better understand genome wide patterns of adaptive variation in Ulmaceae and will help to optimize a correct management of genetic resources of these important noble hardwood species.

## Author Contributions

PP performed SNPs validation, analysis of data and drafted the manuscript. MV prepared vegetal material and performed inoculations. MC conceived the research. LG provided funding. CC conceived and performed the research and drafted the paper. All authors contributed to and approved the final version of the manuscript.

## Acknowledgments

## Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2015.00541

## References

Aoun, M., Jacobi, V., Boyle, B., and Bernier, L. (2010). Identification and monitoring of *Ulmus americana* transcripts during in vitro interactions with the Dutch elm disease pathogen *Ophiostoma novo-ulmi*. *Physiol. Mol. Plant Pathol.* 74, 254-266. doi: 10.1016/j.pmpp.2010.04.003

Aoun, M., Rioux, D., Simard, M., and Bernier, L. (2009). Fungal colonization and host defense reactions in *Ulmus americana* callus cultures inoculated with *Ophiostoma novo-ulmi*. *Phytopathology* 99, 642–650. doi: 10.1094/PHYTO-99-6-0642

Bernier, L., Aoun, M., Bouvet, G., Comeau, A., Dufour, J., Naruzawa, E., et al. (2015). Genomics of the Dutch elm disease pathosystem: are we there yet? *iForest-Biogeosci. Forestry* 8, 149–157. doi: 10.3832/ifor1211-008

Brasier, C. M. (2000). "Intercontinental spread and continuing evolution of the Dutch elm disease pathogens," in *The Elms*, ed. C. P. Dunn (Berlin: Springer), 61–72. doi: 10.1007/978-1-4615-4507-1_4

Bräutigam, A., and Gowik, U. (2010). What can next generation sequencing do for you? Next generation sequencing as a valuable tool in plant research. *Plant Biol.* 12, 831–841. doi: 10.1111/j.1438-8677.2010.00373.x

Brunet, J., Zalapa, J. E., Pecori, F., and Santini, A. (2013). Hybridization and introgression between the exotic Siberian elm, *Ulmus pumila*, and the native Field elm, *U. minor*, in Italy. *Biol. Invasions* 15, 2717–2730. doi: 10.1007/s10530-013-0486-z

Büchel, K., Mcdowell, E., Nelson, W., Descour, A., Gershenzon, J., Hilker, M., et al. (2012). An elm EST database for identifying leaf beetle egg-induced defense genes. *BMC Genomics* 13:242. doi: 10.1186/1471-2164-13-242

Chancerel, E., Lepoittevin, C., Le Provost, G., Lin, Y.-C., Jaramillo-Correa, J. P., Eckert, A. J., et al. (2011). Development and implementation of a highly-multiplexed SNP array for genetic mapping in maritime pine and comparative mapping with loblolly pine. *BMC Genomics* 12:368. doi: 10.1186/1471-2164-12-368

Chang, S., Puryear, J., and Cairney, J. (1993). A simple and efficient method for isolating RNA from pine trees. *Plant Mol. Biol. Rep.* 11, 113–116. doi: 10.1007/bf02670468

Collada, C., Fuentes -Utrilla, P., Gil, L., and Cervera, M. T. (2004). Characterization of microsatellite loci in *Ulmus minor* Miller and cross-amplification in *U. glabra* Hudson *and U. laevis* pall. *Mol. Ecol. Notes* 4, 731–732. doi: 10.1111/j.1471-8286.2004.00798.x

Comeau, A. M., Dufour, J., Bouvet, G. F., Jacobi, V., Nigg, M., Henrissat, B., et al. (2015). Functional annotation of the *Ophiostoma novo-ulmi* genome: insights into the phytopathogenicity of the fungal agent of Dutch Elm disease. *Genome Biol. Evol.* 7, 410–430. doi: 10.1093/gbe/evu281

Corchete, M., Diez, J., and Valle, T. (1993). Phenylalanine ammonia-lyase activity in suspension cultures of *Ulmus pumila* and *U. campestris* treated with spores of *Ceratocystis ulmi*. *Plant Cell Rep.* 13, 111–114. doi: 10.1007/BF002 35302

Doyle, J. J. (1990). Isolation of plant DNA from fresh tissue. *Focus* 12, 13–15.

Duchesne, L., Jeng, R., and Hubbes, M. (1985). Accumulation of phytoalexins in *Ulmus americana* in response to infection by a nonaggressive and an aggressive strain of *Ophiostoma ulmi*. *Can. J. Bot.* 63, 678–680. doi: 10.1139/b85-086

Ďurkovič, J., Čaňová, I., Lagaňa, R., Kučerová, V., Moravčík, M., Priwitzer, T., et al. (2013). Leaf trait dissimilarities between Dutch elm hybrids with a contrasting tolerance to Dutch elm disease. *Ann. Bot.* 111, 215–227. doi: 10.1093/aob/mcs274

Ekblom, R., and Galindo, J. (2010). Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity* 107, 1–15. doi: 10.1038/hdy.2010.152

Falgueras, J., Lara, A. J., Cantón, F. R., Pérez-Trabado, G., and Claros, M. G. (2007). "SeqTrim—a validation and trimming tool for all purpose sequence reads," in *Innovations in Hybrid Intelligent Systems*, eds E. Corchado, J. M. Corchado, and A. Abraham (Berlin: Springer), 353–360.

Forgetta, V., Leveque, G., Dias, J., Grove, D., Lyons, R. Jr., Genik, S., et al. (2013). Sequencing of the Dutch Elm Disease fungus genome using the Roche/454 GS-FLX titanium system in a comparison of multiple genomics core facilities. *J. Biomol. Tech.* 24, 39–49. doi: 10.7171/jbt.12-2401-005

Fuentes-Utrilla, P., Valbuena-Carabaña, M., Ennos, R., and Gil, L. (2014a). Population clustering and clonal structure evidence the relict state of *Ulmus minor* Mill. in the Balearic Islands. *Heredity* 113, 21–31. doi: 10.1038/hdy.2014.21

Fuentes-Utrilla, P., Venturas, M., Hollingsworth, P., Squirrell, J., Collada, C., Stone, G., et al. (2014b). Extending glacial refugia for a European tree: genetic markers show that Iberian populations of white elm are native relicts and not introductions. *Heredity* 112, 105–113. doi: 10.1038/hdy.2013.81

Fujita, M., Fujita, Y., Noutoshi, Y., Takahashi, F., Narusaka, Y., Yamaguchi-Shinozaki, K., et al. (2006). Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. *Curr. Opin. Plant Biol.* 9, 436–442. doi: 10.1016/j.pbi.2006.05.014

Garg, R., and Jain, M. (2013). "RNA-Seq for transcriptome analysis in non-model plants," in *Legume Genomics*, ed. R. J. Rose (Berlin: Springer), 43–58. doi: 10.1007/978-1-62703-613-9_4

Gaur, R., Azam, S., Jeena, G., Khan, A. W., Choudhary, S., Jain, M., et al. (2012). High-throughput SNP discovery and genotyping for constructing a saturated linkage map of chickpea (*Cicer arietinum* L.). *DNA Res.* 19, 357–373. doi: 10.1093/dnares/dss018

Ge, G., Xiao, P., Zhang, Y., and Yang, L. (2011). The first insight into the tissue specific taxus transcriptome via Illumina second generation sequencing. *PLoS ONE* 6:e21220. doi: 10.1371/journal.pone.0021220

Ghelardini, L., and Santini, A. (2009). Avoidance by early flushing: a new perspective on Dutch elm disease research. *iForest-Biogeosci. Forestry* 2, 143–153. doi: 10.3832/ifor0508-002

Glombitza, S., Dubuis, P.-H., Thulke, O., Welzl, G., Bovet, L., Götz, M., et al. (2004). Crosstalk and differential response to abiotic and biotic stressors reflected at the transcriptional level of effector genes from secondary metabolism. *Plant Mol. Biol.* 54, 817–835. doi: 10.1007/s11103-004-0274-3

Götz, S., García-Gómez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., et al. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 36, 3420–3435. doi: 10.1093/nar/gkn176

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.

He, N., Zhang, C., Qi, X., Zhao, S., Tao, Y., Yang, G., et al. (2013). Draft genome sequence of the mulberry tree *Morus notabilis*. *Nat. Commun.* 4:2445. doi: 10.1038/ncomms3445

Heybroek, H. M. (1993). "The Dutch elm breeding program," in *Dutch Elm Disease Research*, eds M. B. Sticklen and J. L. Sherald (Berlin: Springer), 16–25. doi: 10.1007/978-1-4615-6872-8_3

Jaillon, O., Aury, J.-M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., et al. (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* 449, 463–67. doi: 10.1038/nature06148

Khoshraftar, S., Hung, S., Khan, S., Gong, Y., Tyagi, V., Parkinson, J., et al. (2013). Sequencing and annotation of the *Ophiostoma ulmi* genome. *BMC Genomics* 14:162. doi: 10.1186/1471-2164-14-162

Loureiro, J., Rodriguez, E., Gomes, A., and Santos, C. (2007). Genome size estimations on *Ulmus minor* Mill., *Ulmus glabra* Huds., and *Celtis australis* L. using flow cytometry. *Plant Biology* 9, 541–544. doi: 10.1055/s-2007-965165

Martín, J., Solla, A., Venturas, M., Collada, C., Domínguez, J., Miranda, E., et al. (2015). Seven *Ulmus minor* clones tolerant to *Ophiostoma novo-ulmi* registered as forest reproductive material in Spain. *iForest - Biogeosci. Forestry* 8, 172–180. doi: 10.3832ifor1224-008

Martín, J., Solla, A., Woodward, S., and Gil, L. (2005a). Fourier transform-infrared spectroscopy as a new method for evaluating host resistance in the Dutch elm disease complex. *Tree Physiol.* 25, 1331–1338. doi: 10.1093/treephys/25.10.1331

Martín, J. A., Solla, A., Coimbra, M. A., and Gil, L. (2005b). Metabolic distinction of *Ulmus minor* xylem tissues after inoculation with *Ophiostoma novo-ulmi*. *Phytochemistry* 66, 2458–2467. doi: 10.1016/j.phytochem.2005.08.004

Martín, J. A., Collada, C., Venturas, M., Domínguez, J., Miranda, E., Díez, J., et al. (2012). El programa español del olmo: 25 años de mejora e investigación frente a la grafiosis. *Cuad. De La Soc. Esp. De Cienc. For.* 36, 127–140.

Martín, J. A., Solla, A., Esteban, L. G., De Palacios, P., and Gil, L. (2009). Bordered pit and ray morphology involvement in elm resistance to *Ophiostoma novo-ulmi*. *Can. J. For. Res.* 39, 420–429. doi: 10.1139/X08-183

Martín, J. A., Witzell, J., Blumenstein, K., Rozpedowska, E., Helander, M., Sieber, T. N., et al. (2013). Resistance to Dutch Elm disease reduces presence of xylem endophytic fungi in Elms (*Ulmus* spp.). *PLoS ONE* 8:e56987. doi: 10.1371/journal.pone.0056987

Meyer, E., Aglyamova, G. V., Wang, S., Buchanan-Carter, J., Abrego, D., Colbourne, J. K., et al. (2009). Sequencing and de novo analysis of a coral larval transcriptome using 454 GSFlx. *BMC Genomics* 10:219. doi: 10.1186/1471-2164-10-219

Nasmith, C., Jeng, R., and Hubbes, M. (2008). Targeted gene analysis in *Ulmus americana* and *U. pumila* tissues. *For. Pathol.* 38, 90–103. doi: 10.1111/j.1439-0329.2007.00523.x

Neale, D. B., and Kremer, A. (2011). Forest tree genomics: growing resources and applications. *Nat. Rev. Genet.* 12, 111–122. doi: 10.1038/nrg2931

Neely, D., and Himelick, E. (1963). Root graft transmission of Dutch elm disease in municipalities. *Plant Dis. Rep.* 47, 83–85.

Newhouse, A. E., Schrodt, F., Liang, H., Maynard, C. A., and Powell, W. A. (2007). Transgenic American elm shows reduced Dutch elm disease symptoms and normal mycorrhizal colonization. *Plant Cell Rep.* 26, 977–987. doi: 10.1007/s00299-007-0313-z

Ng, K. P., Ngeow, Y. F., Yew, S. M., Hassan, H., Soo-Hoo, T. S., Na, S. L., et al. (2012). Draft genome sequence of *Daldinia eschscholzii* isolated from blood culture. *Eukaryot. Cell* 11, 703–704. doi: 10.1128/EC.00074-12

Novaes, E., Drost, D. R., Farmerie, W. G., Pappas, G. J., Grattapaglia, D., Sederoff, R. R., et al. (2008). High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome. *BMC Genomics* 9:312. doi: 10.1186/1471-2164-9-312

Parchman, T. L., Geist, K. S., Grahnen, J. A., Benkman, C. W., and Buerkle, C. A. (2010). Transcriptome sequencing in an ecologically important tree species: assembly, annotation, and marker discovery. *BMC Genomics* 11:180. doi: 10.1186/1471-2164-11-180

Pavy, N., Pelgas, B., Beauseigle, S., Blais, S., Gagnon, F., Gosselin, I., et al. (2008). Enhancing genetic mapping of complex genomes through the design of highly-multiplexed SNP arrays: application to the large and unsequenced genomes of white spruce and black spruce. *BMC Genomics* 9:21. doi: 10.1186/1471-21 64-9-21

Pinosio, S., González-Martínez, S., Bagnoli, F., Cattonaro, F., Grivet, D., Marroni, F., et al. (2014). First insights into the transcriptome and development of new genomic tools of a widespread circum-Mediterranean tree species, *Pinus halepensis* Mill. *Mol. Ecol. Resour.* 14, 846–856. doi: 10.1111/1755-0998.12232

Prasch, C. M., and Sonnewald, U. (2015). Signaling events in plants: stress factors in combination change the picture. *Environ. Exp. Bot.* 114, 4–14. doi: 10.1016/j.envexpbot.2014.06.020

Rozen, S., and Skaletsky, H. (1999). "Primer3 on the WWW for general users and for biologist programmers," in *Bioinformatics Methods and Protocols*, eds S. Misener and S. A. Krawetz (Beerlin: Springer), 365–386. doi: 10.1385/1-59259-192-2:365

Sánchez-Salguero, R., Navarro-Cerrillo, R. M., Camarero, J. J., and Fernández-Cancio, Á. (2012). Selective drought-induced decline of pine species in southeastern Spain. *Clim. Change* 113, 767–785. doi: 10.1007/s10584-011-0372-6

Santini, A., Mittempergher, M., Fagnani, A., Ferrini, F., Brunetti, M., Crivellaro, A., et al. (2004). Elm breeding for DED resistance, the Italian clones and their wood properties. *For. Syst.* 13, 179–184.

Santini, A., Pecori, F., Pepori, A., and Brookes, A. (2011). 'Morfeo' Elm: a new variety resistant to Dutch elm disease. *For. Pathol.* 42, 171–176. doi: 10.1111/j.1439-0329.2011.00737.x

Santino, A., Taurino, M., De Domenico, S., Bonsegna, S., Poltronieri, P., Pastor, V., et al. (2013). Jasmonate signaling in plant development and defense response to multiple (a)biotic stresses. *Plant Cell Rep.* 32, 1085–1098. doi: 10.1007/s00299-013-1441-2

Smalley, E., and Guries, R. (1993). Breeding elms for resistance to Dutch elm disease. *Annu. Rev. Phytopathol.* 31, 325–354. doi: 10.1146/annurev.py.31.090193.001545

Smalley, E. B., and Guries, R. P. (2000). "Asian elms: sources of disease and insect resistance," in *The Elms*, ed. C. P. Dunn (Berlin: Springer), 215–230. doi: 10.1007/978-1-4615-4507-1_13

Solla, A., Bohnens, J., Collin, E., Diamandis, S., Franke, A., Gil, L., et al. (2005a). Screening European elms for resistance to *Ophiostoma novo-ulmi*. *For. Sci.* 51, 134–141.

Solla, A., Martín, J., Ouellette, G., and Gil, L. (2005b). Influence of plant age on symptom development in *Ulmus minor* following inoculation by *Ophiostoma novo-ulmi*. *Plant Dis.* 89, 1035–1040. doi: 10.1094/PD-89-1035

Sticklen, M. B., Bolyard, M. G., Hajela, R. K., and Dufresne, L. (1991). Molecular and cellular aspects of Dutch elm disease. *Phytoprotection* 72, 1–13. doi: 10.7202/705997ar

Suzuki, N., Rivero, R. M., Shulaev, V., Blumwald, E., and Mittler, R. (2014). Abiotic and biotic stress combinations. *New Phytol.* 203, 32–43. doi: 10.1111/nph.12797

Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313, 1596–1604. doi: 10.1126/science.1128691

Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A., Cestaro, A., Kalyanaraman, A., et al. (2010). The genome of the domesticated apple (*Malus* x domestica Borkh.). *Nat. Genet.* 42, 833–839. doi: 10.1038/ng.654

Venturas, M., Fuentes-Utrilla, P., Ennos, R., Collada, C., and Gil, L. (2013a). Human-induced changes on fine-scale genetic structure in *Ulmus laevis* Pallas wetland forests at its SW distribution limit. *Plant Ecol.* 214, 317–327. doi: 10.1007/s11258-013-0170-5

Venturas, M., Iglesias, S., Nicolás, J. L., Martín, J. A., and Gil, L. (2013b). "*Ulmus minor* Mill," in *Producción y Manejo de Semillas y Plantas Forestales*, eds J. Pemán, R. M. Navarro, J. L. Nicolás, M. A. Prada, and R. Serrada (Madrid:

Organismo Autónomo de Parques Nacionales – Ministerio de Agricultura, Alimentación y Medio Ambiente), 575–597.

Venturas, M., López, R., Martín, J., Gascó, A., and Gil, L. (2014). Heritability of *Ulmus minor* resistance to Dutch elm disease and its relationship to vessel size, but not to xylem vulnerability to drought. *Plant Pathol.* 63, 500–509. doi: 10.1111/ppa.12115

Vezzulli, S., Troggio, M., Coppola, G., Jermakow, A., Cartwright, D., Zharkikh, A., et al. (2008). A reference integrated map for cultivated grapevine (*Vitis vinifera* L.) from three crosses, based on 283 SSR and 501 SNP-based markers. *Theor. Appl. Gen.* 117, 499–511. doi: 10.1007/s00122-008-07943

Wall, P. K., Leebens-Mack, J., Chanderbali, A. S., Barakat, A., Wolcott, E., Liang, H., et al. (2009). Comparison of next generation sequencing technologies for transcriptome characterization. *BMC Genom.* 10:347. doi: 10.1186/1471-2164-10-347

Whiteley, R. E., Black-Samuelsson, S., and Clapham, D. (2003). Development of microsatellite markers for the European white elm (*Ulmus laevis* Pall.) and cross-species amplification within the genus *Ulmus*. *Mol. Ecol. Notes* 3, 598–600. doi: 10.1046/j.1471-8286.2003.00525.x

Xu, L., Zhu, L., Tu, L., Liu, L., Yuan, D., Jin, L., 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

Zalapa, J., Brunet, J., and Guries, R. (2008). Isolation and characterization of microsatellite markers for red elm (*Ulmus rubra* Muhl.) and cross-species amplification with Siberian elm (*Ulmus pumila* L.). *Mol. Ecol. Resour.* 8, 109–112. doi: 10.1111/j.1471-8286.2007.01805.x

Zalapa, J. E., Brunet, J., and Guries, R. P. (2009). Patterns of hybridization and introgression between invasive *Ulmus pumila* (Ulmaceae) and native *U. rubra*. *Am. J. Bot.* 96, 1116–1128. doi: 10.3732/ajb.0800334

Zhu, Y., Machleder, E., Chenchik, A., Li, R., and Siebert, P. (2001). Reverse transcriptase template switching: a SMARTTM approach for full-length cDNA library construction. *BioTechniques* 30, 892–897.