

Transcriptomic response to methyl jasmonate treatment of Scots pine (*Pinus sylvestris*) seedlings

Krista Kānberga-Siliņa¹, Elza Rauda¹, Vilnis Šķipars¹, Adam Vivian-Smith², Igor Yakovlev², Baiba Krivmane¹, Ilze Šņepste¹, Dainis Ruņģis^{1*}

¹Genetic Resource Centre, Latvian State Forest Research Institute “Silava”, Rīgas 111, Salaspils LV–2169, Latvia

²Norwegian Institute of Bioeconomy Research, Section Forest Genetics and Biodiversity, Høgskoleveien 8, 1433 Ås, Norway

*Corresponding author, E-mail: dainis.rungis@silava.lv

Abstract

Scots pine (*Pinus sylvestris* L.) has high ecological and economic importance in Latvia. Induced resistance can protect trees from pathogen attack, and has potential for sustainable pest management. The effect of methyl jasmonate (MeJA) on the gene expression profile of Scots pine was determined in order to assess its potential to induce resistance. Two year old Scots pine ramets were treated with 10 mM MeJA and needle samples were collected two weeks after treatment. RNA was isolated from samples and transcription profiling was performed using the Ion Torrent PGM platform, and RT-PCR was used to confirm expression profiles for selected genes. Gene expression profiles results were summarized by a gene network analysis using gene functionality annotations. Results revealed that MeJA treatment diverts metabolic functions from growth and development to defensive pathways. Analysis of the transcriptome indicated that the MeJA treated Scots pine clone had an altered gene expression profile two weeks after the treatment, suggesting that the induction of defence mechanisms was still active at this time. The organism is in an induced state at least two weeks after the treatment, which confirms the potential of the MeJA use to induce resistance responses in Scots pine trees.

Key words: gene profiles, induced resistance, methyl jasmonate, next generation sequencing, Scots pine.

Abbreviations: DEG, differentially expressed gene; MeJA, methyl jasmonate; RT-PCR, real time polymerase chain reaction.

Introduction

Scots pine (*Pinus sylvestris* L.) has the largest natural distribution of any pine species (Richardson, Rundel 1998). In Europe, Scots pine covers at least 20% of the forested area (Masón, Alía 2000; Brus et al. 2011). While having a high economical value, conifers are susceptible to various pathogens, like root and butt rot caused by *Heterobasidion annosum* (Asiegbu et al. 2005), and in Latvia, almost half of the forested areas are affected by a variety of pathogens and pests every year (<http://www.csb.gov.lv/statistikas-temas/mezsaimnieciba-publikacijas-43247.html>). One possible approach to protect plants from pathogen attack is to prime and induce innate resistance responses, which enable plants to respond more rapidly to infections by limiting pathogen development (Hammerschmidt 1999).

It is important for a tree to protect itself from pathogens and diseases, but to also balance resource allocation between defence and growth processes, such as wood formation (Smith 2015). Phenylpropanoids and terpenes are defence compounds used by plants (Wink 1988), but the phenylpropanoid biosynthesis pathway also produces lignin precursors and other compounds related to wood formation (Whetten, Sederoff 1995; Hatfield, Vermerris 2001). *In vitro* studies have revealed a strong relationship

between the carbohydrate supply and the formation of phenolic compounds in apple tree shoot cultures, indicating competition for carbohydrates between growth-related and defence-related metabolic processes (Lux-Endrich et al. 2000).

Chitinases are enzymes that hydrolyse chitin, which is a homopolymer of N-acetylglucosamine. Chitin is a structural component in fungal and some algae cell walls, as well as in arthropod shells and cuticles (Sahai, Manocha 1993). In plants, chitinase genes are represented by large gene families. It is possible that each chitinase plays a different role in the protection of the plant against different pathogens (Schlumbaum et al. 1986), though some may also exhibit activity towards structural compounds that have similarity to chitin, which in turn, have important roles in plant development (Passarinho, De Vries 2002). Several classes of chitinase genes are distinguished based on their amino acid sequences. In plants, chitinases from classes I, II and IV are most abundant (Shinshi et al. 1990; Iseli et al. 1993; Cohen-Kupiec et al. 1998). One chitinase gene from each of these three classes was selected to validate their expression using real time polymerase chain reaction (RT-PCR). It has been previously reported that the expression of chitinase genes increases after methyl jasmonate (MeJA) treatment, and that different chitinase genes, even within

same class, can exhibit differential expression patterns (Su et al. 2015).

MeJA is a compound that participates in plant defense response mediation, including inducing expression of proteinase inhibitor genes and wound-response genes (Farmer, Ryan 1990; Creelman et al. 1997). Jasmonate has been shown to play a significant role in plant pathogen resistance (Vijayan et al. 1998). MeJA significantly increases monoterpene concentration in the xylem of *Pinus armandi* saplings (Pham et al. 2014). Many conifers possess defence responses that are induced in response to insect attack, MeJA treatment, mechanical wounding or fungal inoculation (Creelman et al. 1997; Miller et al. 2005; Moreira et al. 2009). These induced defences include the activation of existing resin ducts and the developmental formation of traumatic resin duct systems in the bark and xylem (Hudgins et al. 2003). It has been reported that application of 100 mmol L⁻¹ MeJA to intact bark of Norway spruce induces anatomical defense reactions within 30 days (Franceschi et al. 2002). In field experiments, application of MeJA to individuals prior to planting considerably reduced weevil-caused mortality in *P. sylvestris* (Zas et al. 2014). In Norway spruce, MeJA treatment resulted in up-regulation of the *PaPX3* and *PaChi4* genes, and the genes involved in lignin biosynthesis *PaPAL1*, *PaPAL2*, *PaC4H3/5* and *PaHCT1*. Likewise, the ethylene biosynthesis related genes *PaACO* and *PaACS* were also induced by MeJA treatment (Yaqoob et al. 2012).

High-throughput sequencing and bioinformatics tools have been used to investigate the influence of various factors on gene expression in plants (Hoth et al. 2002; Liu et al. 2014). Next generation sequencing allows for gene expression profiling even in non-model organisms, where genome sequences are not available (Morozova, Marra 2008). A high quality genome sequence is not available for *P. sylvestris*; however, partial transcriptomes have been published (Sun et al. 2011, Wachowiak et al. 2015). Microarrays have been successfully used to profile gene expression in *Pinus radiata* saplings exposed to ethephon, but using RNA-Seq could provide a better understanding of gene expression, since the later types of methods do not rely on sequences to designed complementary oligonucleotide probes (Dubouzet et al. 2014).

This study investigated differential gene expression profiles in one *P. sylvestris* clone two weeks after MeJA treatment using high throughput RNA sequencing methods. Differentially expressed genes (DEGs) were annotated, and gene expression networks were generated. RT-PCR was used to validate the expression of three chitinase genes.

Materials and methods

Plant material and treatment

Six two-year-old *P. sylvestris* ramets of clone Sm 9-III-2 were grown in a climate chamber for two months prior to

the experiment. Three ramets were treated with 5 mL 10 mM MeJA in 0.1% Tween 80. Three other ramets were sprayed with 5 mL 0.1% Tween 80 as a control. After treatment, all ramets were placed in clear plastic bags for two days to avoid cross reaction with volatile compounds. Growth conditions in the climate chamber were 8 h darkness and 16 h light at 17 to 22 °C temperature. Two weeks after the treatment needles were collected and stored at -80 °C.

RNA extraction

The CTAB buffer-based method (Yang et al. 2008) was used for total RNA extraction with modifications as described by Rubio-Piña and Zapata-Pérez (2011). RNA quantity was assessed fluorometrically using a Qubit fluorometer and Quant-iT reagents (Invitrogen) following the manufacturer's instructions. RNA extracts were treated using DNase I (Thermo Fisher) and purified using a standard phenol/chloroform extraction.

Next generation sequencing

Next generation sequencing was done in collaboration with the Norwegian Institute of Bioeconomy Research (NIBIO). The quality, quantity and structural integrity of total RNA was assessed using an Agilent Technologies 2100 Bioanalyzer with Agilent RNA 6000 Nano and Pico Kits according to manufacturer's instructions.

RNA samples were enriched for polyadenylated mRNAs from the total RNA samples using the Dynabeads mRNA DIRECT™ Micro Kit (Life Technologies). The quantity of total extracted from one of the treated samples (Sm8) was not sufficient for library construction; therefore, five transcriptome libraries were produced using the Ion Total RNA-Seq Kit v2 (Life Technologies) according to the manufacturer's instructions. RNA quantity and fragment size was assessed by the Qubit RNA Assay Kit on a Qubit Fluorometer and the Agilent RNA 6000 Pico Kit using an Agilent 2100 Bioanalyzer (Thermo Fisher Scientific) following the manufacturers' protocols. Sequencing was carried out using the Ion PGM™ Sequencing 200Kit v2 (catalogue No. 4482006) on Ion 318™ v2 Chips according to the manufacturer's protocol.

Real time PCR

For confirmation of expression profiles, three chitinase genes were selected, each representing a different chitinase class: I, II and VII. Sequences for primer design were obtained from the expressed sequence transcript consensus sequences. Several primer pairs for each gene were designed using the Primer3 software (<http://bioinfo.ut.ee/primer3/>). The primers with the highest efficiency (at least 90%) were used for further analysis (Table 1).

DNA was synthesized using TaqMan Reverse Transcription reagents (Applied biosystems) according to the manufacturer's protocol. Gene expression was determined using the comparative $\Delta\Delta C_t$ RT-PCR method

Table 1. Primer sequences and annealing temperatures

Primer	Sequence (5'-3')	Annealing temperature
Chit1.3F	CGGCAGCTAATAGTTTCCCA	57 °C
Chit1.3R	GAAGTTTGGCCGAAGAAAGC	
Chit2.1F	TTTAAGACGGCGGTGTGGTT	60 °C
Chit2.1R	ACCCCGCCGTTAATGATGTT	
Chit7.2F	ACTGCGATTCCACCTACACG	60 °C
Chit7.2R	AGCCCAGGTAATTTCCAGCG	
GapdhF	ACGGTTTTGGTTCGAATTGGA	60 °C
GapdhR	CCCCACGAGCTCGATATCAT	
Tub5F	CCACATTGGACAGGCCGGTATCC	60 °C
Tub5R	AATGCCGTGCTCGAGGCAGTA	
Act5F	TCATGGTTGGCATGGGACAG	60 °C
Act5R	CTCCATGTCATCCCAGTTGC	

with three endogenous control genes as references for normalization: *GAPDH* (glyceraldehyde-3-phosphate dehydrogenase; Škipars et al. 2011); *TUB* (tubulin) and *ACT* (actin; Känberga-Siliņa, unpublished). Primers and annealing temperatures are shown in Table 1. Experiments were run on a StepOnePlus Real-Time PCR System using 5x HOT FIRE Pol EvaGreen qPCR SuperMix reagents (Solis BioDyne) according to the manufacturer's protocol.

Real time PCR data was analyzed using StepOne software v2.2. Relative quantity of all candidate genes was calculated using all three endogenous control gene $\Delta\Delta C_t$ values.

Data analysis

DNase Fastq format sequence data were analyzed using the CLC Genomics Workbench 7.5.1. software (QIAGEN, Aarhus, Denmark). Due to the absence of a high quality *P. sylvestris* genome sequence, expressed sequences were utilized using the reference sequences (Wachowiak et al. 2015). Fastq files were imported using the SAM/BAM procedure and the low quality reads (Q value < 20) were excluded from analysis. Transcriptome libraries for each sample were separately mapped to the reference transcriptome using the RNA-Seq tool with the following parameters: gap cost 2, indel cost 3 and length and a similarity fraction of 0.8. Results were then exported for further data normalization and comparison in the R environment using DESeq library (Anders and Huber 2010). All sample data were grouped as "treated" and "control". Using the "fdr" method, logFC values were obtained, and only reads with logFC values > |2| were used for further analysis.

To annotate transcriptome sequences we used BLAST for comparing the differentially expressed sequences with the NCBI database (<http://www.ncbi.nlm.nih.gov/>). Each sequence was annotated separately at 80 % identity threshold.

Gene expression networks were generated using the Cytoscape 3.2.1 program with the GeneMANIA plug-in (Montojo et al. 2014). The program uses several databases

containing information on gene interactions to create a gene network that visually represents the relationship between genes. The NCBI BLASTx was used to identify *Arabidopsis thaliana* protein homologues for the DEGs defined in this study. Two separate networks were created depicting the interactions of the up and down regulated gene groups. The GeneMANIA plugin also provides annotations for gene functionality, which were used to classify the DEGs into functional groups.

Results

Sequencing of gene expression libraries

Five gene expression libraries were sequenced using the Ion Torrent PGM. An average of 76.38% of reads per library were successfully mapped onto the reference transcriptome (Table 2). Using RNA-Seq analysis for each read library we obtained expressed reference contigs, which were further analyzed using the FDR (False Discovery Rate) tool in R. A total of 284 sequences were found to be differentially expressed, while 206 were upregulated in MeJA treated samples and 78 sequences were more highly expressed in the control samples. DEGs were annotated with BLASTn (Appendix 1). Among the most highly expressed up regulated genes was a carbonic anhydrase (KEH16352.1), an acid phosphate (CAC84485.1), and a gene with similarity to an alpha terpineol/1,8-cineole synthase (AFU73861.1). Among the down regulated genes, three genes were highly expressed in the control libraries: an anthocyanidin reductase (AAU95082.1), a 50S ribosomal protein L3 (WP_014093814.1), and a Leucine-Rich Receptor-like protein (XP_007018364.1).

Gene networks

Protein homologues from *A. thaliana* were identified for all of the differentially expressed gene sequences, and gene networks were created for both upregulated and down regulated genes using GeneMANIA in Cytoscape. Both gene groups formed one network for each group

Table 2. RNA-Seq mapping results. Sm3 and Sm4 are treated sample libraries, Sm5, Sm6 and Sm7 are control libraries

Sample name	Sm3	Sm4	Sm5	Sm6	Sm7
Number of reads	1 975 699	1 496 742	1 807 085	4 543 929	885 484
Mapped fragments	1 619 297	1 161 871	1 352 336	3 770 254	571 301
Average read length, bp	143	185	185	194	208
Proportion of reads mapped to reference sequence	81.96%	77.63%	74.84%	82.97%	64.52%

without substructures. From the upregulated group, 12 genes (*LSU3*; *AT1G15240/F9L1_19*; *ILR2*; *T2D23.3*; *LOB*; *AT4G23895*; *NDR1*; *AT1G59312*; *AT2G1400D*; *AT3G59210*; *ATPAO3*; *PIP3*) were not included in the gene network, and likewise another 10 genes were not included from the down regulated group (*DREB2C*; *AT4G31480*; *CRRSP18*; *PLP5*; *T22P22_10*; *FIB*; *F19C17.40*; *RPS12*; *PLL3*; *BRCC36A*). The upregulated gene network was more centrally organized (Appendix 2), while the downregulated network contained an expanded and less concentrated network at specific nodes (Appendix 3).

According to gene ontology classification, the upregulated group was dominated by genes that respond to a range of stimuli and are involved in metabolic processes. Only a small fraction of these genes were directly categorized as involved in defence processes. The downregulated gene group was dominated by genes involved in metabolic, developmental processes and cell formation (Fig. 1).

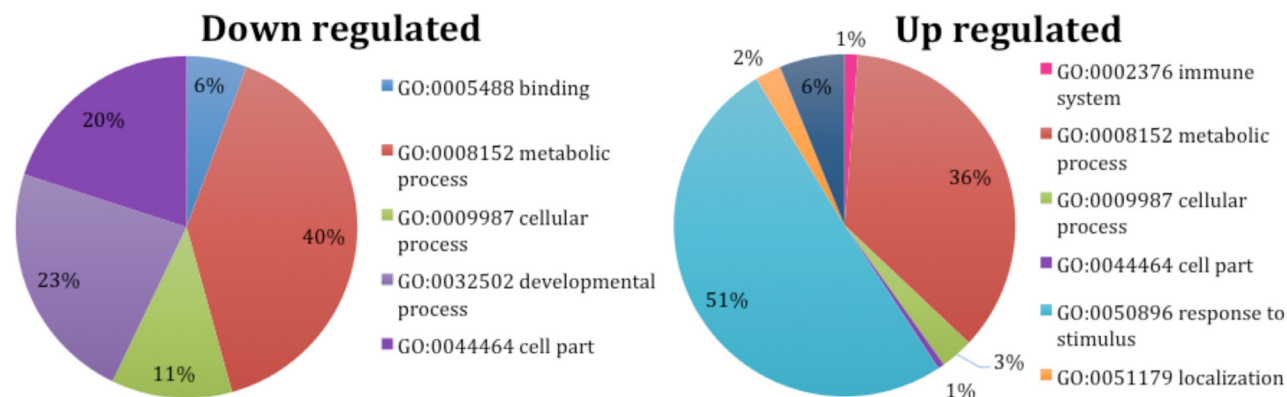
Chitinase gene expression

Gene expression of representatives of three chitinase groups was also determined using RT-PCR: group I (*Chit1*), group II (*Chit2*) and group VII (*Chit7*). The group I chitinase showed a higher expression in treated samples (Sm3, Sm4, Sm8) than in controls, but overall the expression levels were relatively low (Fig. 2). The overall relative expression of the group II chitinase gene was slightly higher and more equal between treated and control samples. Expression of the group VII chitinase gene was the lowest overall, but it was more highly expressed in the MeJA treated samples than in the controls. The highest gene expression level

for all three genes was detected in the sample Sm3 (Fig. 2). The expression profiles of these three chitinase genes were similar to the relative expressions identified from the transcriptome profiling data. However, these genes were not included in the upregulated gene group as the relative expression values between treated and control libraries did not exceed the $\log_{2}FC > |2|$ threshold.

Discussion

Efficient analysis of gene expression using high-throughput sequencing data requires a high quality reference sequence. Initially, a different Scots pine transcriptome reference sequence ERR015466 (Sun et al. 2011) was utilized for transcript mapping. This resulted in the identification of a large amount of repetitive and retrotransposon sequences in our data, which was not unexpected, as approximately 75% of the conifer genomes is composed of repetitive sequences and retrotransposons (Ahuja, Neale 2005). Analysis of the *Pinus taeda* genome sequence revealed that analysis parameters utilized for sequence assembly can influence the identification of repetitive sequences (Kovach et al. 2010), which might explain the large number of repetitive sequences identified using the initial reference sequences. Subsequent data analysis was performed with an alternative Scots pine transcriptome as a reference (Wachowiak et al. 2015), and this allowed us to obtain more informative results with a larger number of protein-coding genes being identified, since this study identified 19659 separate contigs representing a total of 40968 unigenes. This means that the reference sequences utilized for transcript

**Fig. 1.** Gene ontology annotation summary of central genes in up and down regulated gene networks.

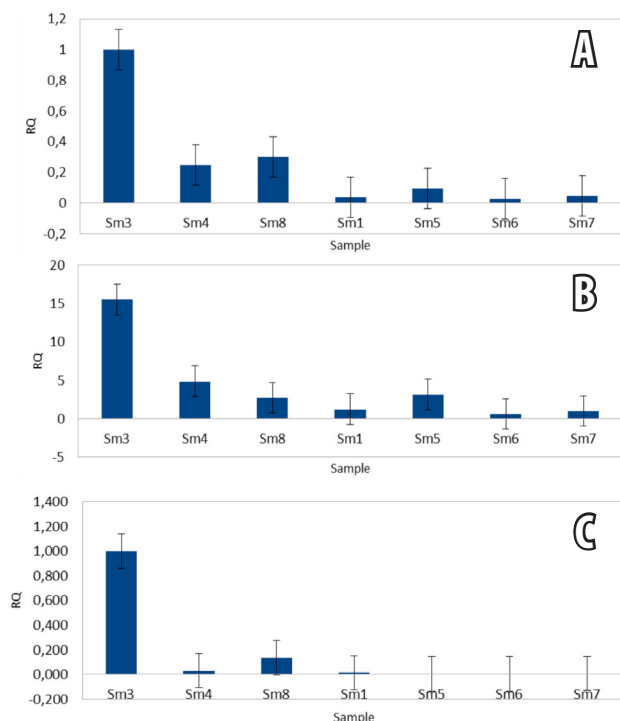


Fig. 2. Group I (A), II (B), VII (C) chitinase relative gene expression (RQ), where Sm3, Sm4, and Sm8 are treated samples and Sm1, Sm5, Sm6 and Sm7 are control samples .

mapping had a much deeper coverage than the first library (ERR015466), but even so, the later study may have not been sufficiently representative of the full genomic complement of transcribed genes. The lack of a completely annotated *P. sylvestris* genome, or exomic complement, combined with information on differential splicing, might have reduced the overall accuracy and discovery of differentially transcribed genes in this study.

Seven of the genes down regulated after MeJA treatment are photosystem related, indicating that treated samples have reduced allocation of resources to the photosystem and that biosynthetic pathways have been reorientated to different processes, which is a typical plant response to biotic attacks (Bilgin et al. 2010). Plants can maintain photosynthetic capacity despite the activation of the defence process signalling system; however, growth of the trees can be affected (Attaran et al. 2014). The induction of defence processes can result in slower growth and development, but the individuals would have increased resistance to pathogens, since metabolism, prior to the pathogen attack, would be orientated towards defence processes. Uninduced plants are not primed for the production of defence related metabolic processes, and this provides a direct advantage to the pathogen.

It has been reported that in sugarcane, the activity of Class VII chitinase gradually decreases over time after treatment with MeJA (Su et al. 2015), which could explain the relatively low expression of this gene compared to other

two chitinase genes in our study. The relatively higher expression levels of the chitinase genes detected by RT-PCR in sample Sm3 were not found in the analysis of the RNA-Seq data. The expression of the chitinase genes was elevated in all of the MeJA treated libraries, but the difference in gene expression was below the threshold set for identification of differentially expressed genes ($\log_{2}FC$ value $> |2|$). The cause of the high levels of chitinase gene expression in sample Sm3 was not immediately obvious. One possibility is that this individual was additionally traumatized at the time of sampling or shortly before, but there was no visual or other evidence to support this. Another possibility is that this individual had some undiagnosed infection or variability, which could also lead to enhanced expression of chitinases (Schlumbaum et al. 1986; Fracetto et al. 2013). Treatment with salicylic acid, MeJA, and abscisic acid has shown that different chitinase genes, even within same class, are transcribed differentially over time (Su et al. 2015). The *Chit1* gene is the most promising for induced resistance activity, as class I chitinases have the potential to directly inhibit fungal pathogens, because of their chitinolytic activity.

One of the genes with highest upregulated expression after MeJA treatment was the acidic phosphatase gene. This is a non-specific enzyme that is most likely to be needed in phosphorus hydrolysis and mobilization (Duff et al. 1994). Increased expression of this gene is also observed in senescing *A. thaliana* leaves, where it could participate in the mobilization of phosphate, which is released during nucleic acid degradation (Buchanan-Wollaston et al. 2005). In addition, three different purple acidic phosphatase genes were detected in the upregulated gene group. This enzyme catalyzes the hydrolysis of various phosphomonoesters and amide substrates, as well as the production of phosphorus in plants (Schenk et al. 2013). In *P. pinaster* it has been observed that higher availability of phosphorus reduces resin accumulation and increases its release in the needles (Blanch et al. 2012). At high concentrations, monoterpenes have pest repellent and moderate fungicidal properties (Everaerts et al. 2012). This means that MeJA could promote resin release, which has the potential to facilitate plant resistance to a variety of pests.

The upregulated gene network was highly centred, with no separate gene subgroups, indicating that the MeJA treatment triggers the activation of a related gene cascade and a unified response. The downregulated gene network suggests a more general gene expression response that is not focused on one function. In addition, a summary of GO annotations indicate that the expressed genes are more evenly distributed among various functions intended to provide basic functions of the organism. It should be noted that the results do not examine the genes that are similarly expressed in both groups (< 2 times gene expression difference).

Our results show that in samples treated with MeJA, two

weeks after treatment, the increased expression of signal transduction genes can be detected. This suggests that the organism is in an induced state, and that in the event of a pest attack, the plant may respond more rapidly to the infection process or other stress conditions. However, Scots pine as a species has high ecological plasticity and genetic variability even within a single population (Sevik et al. 2010; Salmela et al. 2013), which influences the response of individuals to a variety of factors. As only one *P. sylvestris* clone was used in our experiment, these results need to be replicated or verified using a larger number of individuals from a different population. In addition, further time points would be required to determine the duration of the response to MeJA treatment.

A similar study in *Picea sitchensis*, which investigated gene expression in injured plants affected by insect feeding, showed that several thousand genes are activated in defense responses (Ralph et al. 2006). Expression of similar genes is also evident in our results and indicates that the MeJA treated Scots pine individual had an altered gene expression profile two weeks after the treatment, potentially increasing survival and resistance to pathogen attack.

Acknowledgements

This study was performed within the framework of the Latvian Council of Science project Investigation of molecular defense mechanisms in Scots pine (*Pinus sylvestris* L.) (No. 284/2012). We would like to thank Carl Gunnar Fosdøl Norwegian Institute of Bioeconomy Research for assistance with the project.

References

Ahuja M.R., Neale D.B. 2005. Evolution of genome size in conifers. *Silvae Genet.* 54: 126–137.

Anders S., Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biol.* 11: R106.

Asiegbu F.O., Adomas A., Stenlid J. 2005. Conifer root and butt rot caused by *Heterobasidion annosum* (Fr.) Bref. sl. *Molec. Plant Pathol.* 6: 395–409.

Attaran E., Major I.T., Cruz J.A., Rosa B.A., Koo A.J., Chen J., Kramer D.M., He S. Y., Howe G.A. 2014. Temporal dynamics of growth and photosynthesis suppression in response to jasmonate signalling. *Plant Physiol.* 165: 1302–1314.

Bilgin D.D., Zavala J.A., Zhu J.I.N., Clough S.J., Ort D.R., DeLucia E.H. 2010. Biotic stress globally downregulates photosynthesis genes. *Plant Cell Environ.* 33: 1597–1613.

Blanch J.S., Sampedro L., Llusà J., Moreira X., Zas R., Peñuelas J. 2012. Effects of phosphorus availability and genetic variation of leaf terpene content and emission rate in *Pinus pinaster* seedlings susceptible and resistant to the pine weevil, *Hylobius abietis*. *Plant Biol.* 14: 66–72.

Brus D.J., Hengeveld G.M., Walvoort D.J.J., Goedhart P.W., Heidema A.H., Nabuurs G.J., Gunia K. 2011: Statistical mapping of tree species over Europe. *Eur. J. Forest Res.* 131: 145–157.

Buchanan-Wollaston V., Page T., Harrison E., Breeze E., Lim P.O., Nam H.G., Lin J.F., Wu S.H., Swidzinski J., Ishizaki K., Leaver C.J. 2005. Comparative transcriptome analysis reveals

significant differences in gene expression and signalling pathways between developmental and dark/starvation-induced senescence in Arabidopsis. *Plant J.* 42: 567–585.

Cohen-Kupiec R., Chet I., Cohen-Kupiec C. R. 1998. The molecular biology of chitin digestion. *Curr. Opin. Biotechnol.* 9: 270–277.

Creelman R.A., Mullet J.E. 1997. Biosynthesis and action of jasmonates in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 48: 355–381.

Dubouzet J.G., Donaldson L., Black M.A., McNoe L., Liu V., Lloyd-Jones G. 2014. Heterologous hybridisation to a *Pinus* microarray: profiling of gene expression in *Pinus radiata* saplings exposed to ethephon. *New Zealand J. Forestry Sci.* 44: 21.

Duff S.M., Sarath G., Plaxton W.C. 1994. The role of acid phosphatases in plant phosphorus metabolism. *Physiol. Plant.* 90: 791–800.

Everaerts C., Grégoire J.C., Merlin J. 2012. The toxicity of Norway spruce monoterpenes to two bark beetle species and their associates. In: Mattson W.J., Levieux J., Bernard-Dagan C. (eds) *Mechanisms of Woody Plant Defenses Against Insects*. Springer, New York, pp. 335–344.

Farmer E.E., Ryan C.A. 1990. Interplant communication: airborne methyl jasmonate induces synthesis of proteinase inhibitors in plant leaves. *Proc. Nat. Acad. Sci. USA* 87: 7713–7716.

Fracetto G.G., Peres L.E., Mehdy M.C., Lambais M.R. 2013. Tomato ethylene mutants exhibit differences in arbuscular mycorrhiza development and levels of plant defense-related transcripts. *Symbiosis* 60: 155–167.

Franceschi V.R., Krekling T., Christiansen E. 2002. Application of MeJA on *Picea abies* (Pinaceae) stems induces defense-related responses in phloem and xylem. *Am. J. Bot.* 89: 578–586.

Hammerschmidt R. 1999. Induced disease resistance: how do induced plants stop pathogens? *Physiol. Mol. Plant Pathol.* 55: 77–84.

Hatfield R., Vermerris W. 2001. Lignin formation in plants. The dilemma of linkage specificity. *Plant Physiol.* 126: 1351–1357.

Hoth S., Morgante M., Sanchez J. P., Hanafey M.K., Tingey S.V., Chua N.H. 2002. Genome-wide gene expression profiling in *Arabidopsis thaliana* reveals new targets of abscisic acid and largely impaired gene regulation in the abi1-1 mutant. *J. Cell Sci.* 115: 4891–4900.

Hudgins J.W., Christiansen E., Franceschi V.R. 2003. MeJA induces changes mimicking anatomical defenses in diverse members of the Pinaceae. *Tree Physiol.* 23: 361–371.

Iseli B., Boller T., Neuhaus J.M. 1993. The N-terminal cysteine-rich domain of tobacco class I chitinase is essential for chitin binding but not for catalytic or antifungal activity. *Plant Physiol.* 103: 221–226.

Kovach A., Wegrzyn J.L., Parra G., Holt C., Bruening G.E., Loopstra C.A., Hartigan J., Yandell M., Langley C.H., Korf I., Neale D.B. 2010. The *Pinus taeda* genome is characterized by diverse and highly diverged repetitive sequences. *BMC Genomics* 11: e1.

Liu Y., Liu M., Li X., Cao B., Ma X. 2014. Identification of differentially expressed genes in leaf of *Reaumuria soongorica* under PEG-induced drought stress by digital gene expression profiling. *PLoS One* 9: e94277.

Lux-Endrich A., Treutter D., Feucht W. 2000. Influence of nutrients and carbohydrate supply on the phenol composition of apple shoot cultures. *Plant Cell Tissue Organ Cult.* 60: 15–21.

Masón W.L., Alía R. 2000. Current and future status of Scots pine (*Pinus sylvestris* L.) forests in Europe. *Forest Syst.* 9: 317–335.

- Miller B., Madilao L.L., Ralph S., Bohlmann J., 2005. Insect-induced conifer defense. White pine weevil and MeJA induce traumatic resinosis, de novo formed volatile emissions, and accumulation of terpenoid synthase and putative octadecanoid pathway transcripts in Sitka spruce. *Plant Physiol.* 137: 369–382.
- Montejo J., Zuberi K., Rodriguez H., Bader G.D., Morris Q. 2014. GeneMANIA: Fast gene network construction and function prediction for Cytoscape. F1000Research 3.
- Morozova O., Marra M.A. 2008. Genomics Applications of next-generation sequencing technologies in functional genomics. *Genomics* 92: 255–264.
- Moreira X., Sampedro L. Zas R., 2009. Defensive responses of *Pinus pinaster* seedlings to exogenous application of methyl jasmonate: concentration effect and systemic response. *Environ. Exp. Bot.* 67: 94–100.
- Passarinho P.A., de Vries S.C. 2002. Arabidopsis chitinases: a genomic survey. *The Arabidopsis Book* 1: e0023.
- Pham T., Chen H., Yu J., Dai L., Zhang R., Vu T. Q.T. 2014. The Differential effects of the blue-stain fungus *Leptographium qinlingensis* on monoterpenes and sesquiterpenes in the stem of Chinese white pine (*Pinus armandi*) saplings. *Forests* 5: 2730–2749.
- Ralph S G., Yueh H., Friedmann M., Aeschliman D., Zeznik J.A., Nelson C.C., Butterfield Y.S., Kirkpatrick R., Liu J., Jones S.J., Marra M.A., Douglas C.J., Ritland K., Bohlmann J. 2006. Conifer defence against insects: microarray gene expression profiling of Sitka spruce (*Picea sitchensis*) induced by mechanical wounding or feeding by spruce budworms (*Choristoneura occidentalis*) or white pine weevils (*Pissodes strobi*) reveals large-scale changes of the host transcriptome. *Plant Cell Environ.* 29: 1545–1570.
- Richardson D. M., Rundel P. W. 1998. Ecology and biogeography of *Pinus*: an introduction. In: Richardson D.M. (ed) *Ecology and Biogeography of Pinus*. Cambridge, Cambridge University Press, UK pp. 3–46.
- Rubio-Piña J.A., Zapata-Pérez O. 2011. Isolation of total RNA from tissues rich in polyphenols and polysaccharides of mangrove plants. *Electr. J. Biotechnol.* 14: 11.
- Sahai A.S., Manoch M.S. 1993. Chitinases of fungi and plants : their involvement in morphogenesis and host-parasite interaction. *FEMS Microbiol. Rev.* 11: 317–338.
- Salmela M.J., Cavers S., Cottrell J.E., Iason G.R., Ennos R.A. 2013. Spring phenology shows genetic variation among and within populations in seedlings of Scots pine (*Pinus sylvestris* L.) in the Scottish Highlands. *Plant Ecol. Divers.* 6: 523–536.
- Schenk G., Mitić N., Hanson G.R., Comba P. 2013. Purple acid phosphatase: A journey into the function and mechanism of a colorful enzyme. *Coord. Chem. Rev.* 257: 473–482.
- Schlumberg A., Mauch E., Vögeli U., Boller, T. 1986. Plant chitinases are potent inhibitors of fungal growth. *Nature* 324: 365–367.
- Sevik H., Ayan S., Turna I., Yahyaoglu Z. 2010. Genetic diversity among populations in Scotch pine (*Pinus sylvestris* L.) seed stands of Western Black Sea Region in Turkey. *African J. Biotechnol.* 9: 7266–7272.
- Shinshi H., Neuhaus J., Ryals J., Meins F. 1990. Structure of a tobacco endochitinase gene: evidence that different chitinase genes can arise by transposition of sequences encoding a cysteine-rich domain. *Plant Mol. Biol.* 14: 35–368.
- Su Y., Xu L., Wang S., Wang Z., Yang Y., Chen Y., Que Y. 2015. Identification, phylogeny, and transcript of chitinase family genes in sugarcane. *Sci. Rep.* 5: srep10708.
- Smith K. T. 2015. Compartmentalization, resource allocation, and wood quality. *Curr. Forestry Rep.* 1: 8–15.
- Sun H., Paulin L., Alatalo E., Asiegbo F.O. 2011. Response of living tissues of *Pinus sylvestris* to the saprotrophic biocontrol fungus *Phlebiopsis gigantea*. *Tree Physiol.* 31: 438–451.
- Škipars V., Krivmane B., Ruņģis D. 2011. Thaumatin-like protein gene copy number variation in Scots pine (*Pinus sylvestris*). *Environ. Exp. Biol.* 9:75–81.
- Vijayan P., Shockey J., Lévesque C.A., Cook R.J. 1998. A role for jasmonate in pathogen defense of Arabidopsis. *Proc. Nat. Acad. Sci. USA* 95: 7209–7214.
- Wachowiak W., Trivedi U., Perry A., Cavers S. 2015. Comparative transcriptomics of a complex of four European pine species. *BMC Genomics* 16: 234.
- Whetten R., Sederoff R., 1995. Lignin biosynthesis. *Plant Cell* 7: 1001.
- Wink M. 1988. Plant breeding: importance of plant secondary metabolites for protection against pathogens and herbivores. *Theor. Appl. Genet.* 75: 225–233.
- Yang G., Zhou R., Tang T., Shi S. 2008. Preparative biochemistry and biotechnology simple and efficient isolation of high-quality total RNA from *Hibiscus tiliaceus*, a mangrove associate and its relatives. *Prep. Biochem. Biotechnol.* 38: 257–264.
- Yaqoob N., Yakovlev I. A., Krokene P., Kvaalen H., Solheim H., Fossdal C.G. 2012. Defence-related gene expression in bark and sapwood of Norway spruce in response to *Heterobasidion parviporum* and methyl jasmonate. *Physiol. Mol. Plant Pathol.* 77: 10–16.
- Zas R., Björklund N., Nordlander G., Cendán C., Hellqvist C., Sampedro L. 2014. Exploiting jasmonate-induced responses for field protection of conifer seedlings against a major forest pest, *Hylobius abietis*. *Forest Ecol. Manage.* 313: 212–223.

Appendix 1. Annotation of sequences differentially expressed between MeJA treated (Sm3, Sm4) and control (Sm5, Sm6, Sm7) libraries

Reference sequence	Sm3	Sm4	Sm5	Sm6	Sm7	logFC	PValue_	Gene annotation	Organism	Accession No.
comp43346_c0_seq1	250	96	0	0	0	-9.62	3.68E-07	Eukaryotic-type carbonic anhydrase	<i>Medicago truncatula</i>	KEH16352.1
comp54078_c0_seq1	91	60	0	0	0	-8.49	1.68E-07	Uninformative		
comp57289_c0_seq1	441	289	2	0	0	-8.37	0.01	Putative acid phosphatase	<i>Pinus pinaster</i>	CAC84485.1
comp20846_c0_seq1	144	190	1	0	0	-8.36	2.62E-06	Alpha terpineol /1,8-cineole synthase	<i>Pinus contorta</i>	AFU73861.1
comp53497_c0_seq2	0	89	0	0	0	-8.09	0.01	Protein HOTHEAD	<i>Gossypium arboreum</i>	KHG26487.1
comp32018_c0_seq1	85	119	0	0	1	-7.91	2.62E-06	Alpha terpineol /1,8-cineole synthase	<i>Pinus contorta</i>	AFU73861.1
comp55229_c0_seq9	47	59	0	0	0	-7.8	0.01	Chalcone synthase	<i>Pinus pinaster</i>	AAP85249.1
comp49218_c0_seq1	45	36	0	0	0	-7.64	3.68E-07	Isopentenyl diphosphate isomerase	<i>Pinus taeda</i>	ACU56978.1
comp53497_c0_seq1	0	56	0	0	0	-7.46	0.01	Protein HOTHEAD	<i>Gossypium arboreum</i>	KHG26487.1
comp53815_c0_seq5	38	29	0	0	0	-7.34	2.62E-06	Putative metallophosphatase	<i>Lupinus luteus</i>	CAD12837.1
comp35142_c0_seq1	38	22	0	0	0	-7.14	3.12E-06	Cinnamyl alcohol acyltransferase 1	<i>Larrea tridentata</i>	AHA90802.1
comp1651716_c0_seq1	16	34	0	0	0	-7.08	0	Jasmonate ZIM domain-containing protein	<i>Sonneratia caseolaris</i>	AFU90898.1
comp48680_c0_seq2	5	39	0	0	0	-7.04	0	Laccase	<i>Picea abies</i>	AFV52381.1
comp54849_c0_seq17	89	70	1	0	0	-7	0.01	Benzyl alcohol O-benzoyltransferase	<i>Morus notabilis</i>	XP_010108951.1
comp210649_c0_seq1	25	20	0	0	0	-6.78	8.21E-06	Auxin-responsive family protein	<i>Pseudotsuga macrocarpa</i>	ACH60220.1
comp53497_c0_seq3	0	80	0	0	1	-6.75	0.01	Protein HOTHEAD	<i>Gossypium arboreum</i>	KHG26487.1
comp53560_c0_seq1	0	39	0	0	0	-6.75	0.04	Glyoxal oxidase-related protein	<i>Theobroma cacao</i>	XP_007017716.1
comp50861_c0_seq1	23	72	0	1	0	-6.75	0	Serine-type endopeptidase inhibitor	<i>Arabidopsis thaliana</i>	NP_177351.1
comp47417_c0_seq1	42	4	0	0	0	-6.6	0	Isoeugenol synthase 1	<i>Glycine soja</i>	KHN26455.1
comp53815_c0_seq1	36	6	0	0	0	-6.48	0	Purple acid phosphatase	<i>Acacia mangium</i>	BAO45898.1
comp56578_c0_seq1	74	52	0	0	2	-6.42	3.68E-07	Pathogen-inducible alpha-dioxygenase	<i>Nicotiana attenuata</i>	AAG59584.1
comp47930_c0_seq1	52	34	0	1	0	-6.41	0.01	Uninformative		
comp243640_c0_seq1	26	10	0	0	0	-6.32	0	Cytochrome P450 71A1	<i>Morus notabilis</i>	XP_010102947.1
comp54314_c0_seq1	13	53	0	0	1	-6.26	0	Class I chitinase	<i>Pinus contorta</i>	AEF59005.1
comp52953_c0_seq1	52	21	0	0	1	-6.23	7.20E-06	Purple acid phosphatase	<i>Acacia mangium</i>	BAO45898.1
comp53900_c0_seq3	0	28	0	0	0	-6.21	0.05	Kunitz type trypsin inhibitor	<i>Medicago truncatula</i>	KEH38529.1
comp54934_c0_seq17	3	23	0	0	0	-6.19	0.01	Uninformative		
comp18977_c0_seq3	128	677	4	28	0	-6.13	0.02	Proline-rich protein	<i>Pinus taeda</i>	AAF75825.1
comp949065_c0_seq1	24	6	0	0	0	-6.1	0	Protein-PII uridylyltransferase	<i>Donghicola xiamenensis</i>	WP_028093236.1
comp52455_c0_seq6	28	5	0	0	0	-6.08	0.01	Amidase	<i>Pseudomonas</i> sp.	WP_008080115.1
comp54934_c0_seq15	5	20	0	0	0	-6.07	0.01	Uninformative		
comp1463951_c0_seq1	12	14	0	0	0	-6.06	0	Sulfate/thiosulfate import ATP-binding protein cysA	<i>Theobroma cacao</i>	XP_007049639.1
comp51296_c0_seq1	5	73	0	2	0	-6.05	0	Glycoside hydrolase family 71 protein	<i>Heterobasidion irregulare</i>	XP_009545841.1

Appendix 1. continued

Reference sequence	Sm3	Sm4	Sm5	Sm6	Sm7	logFC	PValue_ fdr	Gene annotation	Organism	Accession No.
comp107485_c0_seq1	14	12	0	0	0	-5.99	0	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	<i>Theobroma cacao</i>	XP_007048464.1
comp35518_c0_seq1	31	0	0	0	0	-5.98	0.05	Neutral invertase	<i>Viscum album</i> subsp. <i>album</i>	ABF50710.1
comp375741_c0_seq1	26	2	0	0	0	-5.92	0.01	1-deoxy-D-xylulose 5-phosphate synthase type II	<i>Picea abies</i>	ABS50520.1
comp55263_c1_seq19	2	18	0	0	0	-5.9	0.03	Pentatricopeptide (PPR) repeat protein	<i>Medicago truncatula</i>	KEH25607.1
comp54934_c0_seq10	2	20	0	0	0	-5.89	0.02	Uninformative		
comp494061_c0_seq1	4	18	0	0	0	-5.86	0.01	Chromatin assembly factor 1 subunit A	<i>Theobroma cacao</i>	XP_007049998.1
comp54849_c0_seq11	49	23	1	0	0	-5.81	0.04	Benzyl alcohol O-benzoyltransferase	<i>Morus notabilis</i>	XP_010108951.1
comp52965_c0_seq1	12	12	0	0	0	-5.79	0.02	Inositol-pentakisphosphate 2-kinase-like protein	<i>Medicago truncatula</i>	KEH36181.1
comp49393_c0_seq1	32	85	0	0	4	-5.75	8.36E-05	Clavamate synthase-like protein	<i>Aegilops tauschii</i>	EMT27634.1
comp1228066_c0_seq1	53	29	1	0	1	-5.68	1.25E-05	GDSL-like lipase/ Acylhydrolase superfamily protein, putative isoform 1	<i>Theobroma cacao</i>	XP_007027683.1
comp54934_c0_seq16	6	37	0	0	1	-5.65	0.01	Uninformative		
comp27055_c0_seq1	48	32	2	0	0	-5.62	7.65E-06	1-deoxy-D-xylulose 5-phosphate synthase type II	<i>Picea abies</i>	ABS50520.1
comp52828_c0_seq2	2	33	0	0	1	-5.57	0.01	Heat shock protein	<i>Ammopiptanthus mongolicus</i>	AFC01202.1
comp36780_c0_seq1	36	14	0	1	0	-5.57	0.04	Jasmonate-zim-domain protein 1	<i>Theobroma cacao</i>	XP_007012594.1
comp40675_c0_seq2	10	10	0	0	0	-5.57	0.02	ABC-2 and Plant PDR ABC-type transporter family protein isoform 5	<i>Theobroma cacao</i>	XP_007024298.1
comp45469_c0_seq4	5	12	0	0	0	-5.53	0.02	Putative LRR receptor-like serine/threonine-protein kinase	<i>Morus notabilis</i>	XP_010092221.1
comp54143_c1_seq4	10	8	0	0	0	-5.46	0	Uninformative		
comp42728_c1_seq1	6	10	0	0	0	-5.37	0.03	Sn1-specific diacylglycerol lipase alpha	<i>Gossypium arboreum</i>	KHG05530.1
comp431420_c0_seq1	26	12	0	0	1	-5.34	0	PTI1-like tyrosine-protein kinase	<i>Glycine soja</i>	KHN31557.1
comp44918_c0_seq1	16	2	0	0	0	-5.32	0.03	Uninformative		
comp68888_c0_seq1	12	40	0	0	2	-5.29	0	Thaumatococcus-like protein	<i>Pinus sylvestris</i>	AHH02776.1
comp47732_c0_seq1	0	55	0	2	1	-5.28	0.03	Asparagine synthetase	<i>Pinus pinaster</i>	ADU02856.1
comp40057_c0_seq1	4	10	0	0	0	-5.26	0.03	Uninformative		
comp19537_c0_seq1	6	9	0	0	0	-5.24	0.05	Uninformative		
comp39760_c0_seq1	42	18	2	0	0	-5.23	0	Cinnamoyl CoA reductase	<i>Arabidopsis thaliana</i>	AAM62475.1
comp46327_c0_seq1	31	6	0	0	1	-5.19	0	Uninformative		
comp50304_c0_seq1	9	49	0	2	1	-5.19	0	Response to low sulfur 3	<i>Theobroma cacao</i>	XP_007039253.1
comp23260_c0_seq1	14	2	0	0	0	-5.19	0.03	Uninformative		
comp52065_c1_seq6	9	6	0	0	0	-5.17	0.01	Tetratricopeptide repeat (TPR)-like superfamily protein isoform 1	<i>Theobroma cacao</i>	XP_007036762.1

Appendix 1. continued

Reference sequence	Sm3	Sm4	Sm5	Sm6	Sm7	logFC	PValue_ fdr	Gene annotation	Organism	Accession No.
comp54849_c0_seq5	47	68	1	5	0	-5.17	0.03	Benzoyl CoA benzoic acid benzoyltransferase	<i>Verbena x hybrida</i>	BAE72881.1
comp46741_c0_seq1	81	70	2	0	5	-5.14	1.08E-05	Arogenate dehydratase 3	<i>Petunia x hybrida</i>	ACY79504.1
comp54246_c0_seq2	12	3	0	0	0	-5.13	0.02	NBS/LRR	<i>Pinus taeda</i>	AAM28912.1
comp54934_c0_seq7	10	21	1	0	0	-5.12	0	Uninformative		
comp55507_c0_seq1	50	102	0	0	9	-5.11	0	Class III peroxidase PSYP1	<i>Pinus sylvestris</i>	AAG02215.1
comp49725_c0_seq6	5	8	0	0	0	-5.07	0.03	Calmodulin binding protein	<i>Ricinus communis</i>	XP_002517079.1
comp45618_c0_seq1	8	5	0	0	0	-4.98	0.01	Putative truncated TIR-NBS-LRR protein	<i>Pinus monticola</i>	ADW94528.1
comp25629_c0_seq1	26	30	0	1	2	-4.92	6.15E-05	JAZ10	<i>Gossypium barbadense</i>	AJT58397.1
comp184252_c0_seq1	6	18	0	0	1	-4.91	0.01	C2H2-like zinc finger protein	<i>Theobroma cacao</i>	XP_007021715.1
comp49872_c0_seq2	5	6	0	0	0	-4.81	0.05	F-box/LRR-repeat protein	<i>Pediculus humanus corporis</i>	XP_002429423.1
comp37005_c0_seq1	5	6	0	0	0	-4.79	0.04	Plant IF-like protein	<i>Medicago truncatula</i>	KEH43124.1
comp39224_c0_seq1	13	17	1	0	0	-4.78	0.05	Peroxidase	<i>Picea abies</i>	CAH10839.1
comp54766_c0_seq16	7	4	0	0	0	-4.75	0.04	Progesterone 5-beta-reductase	<i>Gomphocarpus fruticosus</i>	ADG56546.1
comp53815_c0_seq4	31	10	0	0	2	-4.75	0	Purple acid phosphatase	<i>Acacia mangium</i>	BAO45898.1
comp407480_c0_seq1	33	8	0	2	0	-4.74	0	ABC transporter G family member 36	<i>Gossypium arboreum</i>	KHG28037.1
comp50861_c0_seq2	21	54	1	3	1	-4.71	0.02	Serine-type endopeptidase inhibitor	<i>Arabidopsis lyrata</i> subsp. <i>lyrata</i>	XP_002887410.1
comp49971_c1_seq1	27	19	0	3	0	-4.66	3.17E-05	PAR1 protein	<i>Theobroma cacao</i>	XP_007026832.1
comp54934_c0_seq6	3	29	0	0	2	-4.66	0.05	Uninformative		
comp52407_c0_seq1	104	76	1	2	10	-4.62	0	Putative flavoprotein-containing polyamine oxidase	<i>Pinus sylvestris</i>	ADQ37306.1
comp48744_c0_seq1	22	16	1	0	1	-4.6	0	Dehydrololichyl diphosphate synthase 6	<i>Zea mays</i>	ACG35019.1
comp35804_c0_seq1	26	10	0	1	1	-4.58	0	Helix-loop-helix DNA-binding domain containing protein	<i>Oryza sativa</i> Japonica Group	ABF97395.1
comp53918_c0_seq1	6	33	1	2	0	-4.57	0.01	PREDICTED: pinin	<i>Cucumis melo</i>	XP_008453065.1
comp51065_c0_seq1	26	11	0	2	0	-4.55	0.01	Peroxidase	<i>Picea abies</i>	CAD92856.1
comp18977_c0_seq4	100	102	1	21	0	-4.54	0.03	Proline-rich protein	<i>Pinus taeda</i>	AAF75825.1
comp18977_c0_seq1	120	76	0	20	1	-4.54	0.04	Proline-rich protein	<i>Pinus taeda</i>	AAF75825.1
comp54314_c0_seq4	11	41	0	5	0	-4.49	0.01	Class I chitinase	<i>Pinus contorta</i>	AEF59005.1
comp50147_c0_seq3	67	70	4	0	6	-4.47	0	Pyrrroline-5-carboxylate synthetase isoform 3	<i>Theobroma cacao</i>	XP_007009201.1
comp52043_c1_seq2	54	99	6	5	0	-4.47	0.02	Embryo-abundant protein	<i>Picea glauca</i>	AAB01567.1
comp54813_c0_seq2	90	375	7	57	0	-4.46	0.05	PR-1	<i>Solanum torvum</i>	BAB78476.1
comp18977_c0_seq2	127	95	1	25	0	-4.44	0.04	Proline-rich protein	<i>Pinus taeda</i>	AAF75825.1
comp51867_c0_seq3	29	6	1	0	1	-4.44	0.01	Beta-lactamase	<i>Aquimarina macrocephali</i>	WP_024768563.1
comp40759_c0_seq1	42	37	5	0	0	-4.43	0	Uninformative		
comp39422_c0_seq1	18	3	0	0	1	-4.39	0.03	Pathogenesis-related homeodomain protein	<i>Medicago truncatula</i>	XP_003595073.1
comp236525_c0_seq1	33	2	0	0	2	-4.38	0.02	Zinc transporter 11-like protein	<i>Gossypium arboreum</i>	KHG22377.1
comp38640_c1_seq1	5	28	0	0	3	-4.32	0.02	Putative S-acyltransferase	<i>Glycine soja</i>	KHN33780.1

Appendix 1. continued

Reference sequence	Sm3	Sm4	Sm5	Sm6	Sm7	logFC	PValue_ fdr	Gene annotation	Organism	Accession No.
comp31109_c0_seq1	10	8	0	0	1	-4.32	0	Pentatricopeptide repeat-containing protein, putative isoform 1	<i>Theobroma cacao</i>	XP_007010632.1
comp52761_c0_seq3	148	58	0	25	1	-4.31	0.01	PDR1	<i>Picea abies</i>	ADP55080.1
comp47422_c0_seq1	10	18	2	0	0	-4.29	0.02	Primary amine oxidase isoform 2	<i>Theobroma cacao</i>	XP_007034733.1
comp53815_c0_seq2	29	5	2	0	0	-4.28	0.02	Purple acid phosphatase	<i>Acacia mangium</i>	BAO45898.1
comp52274_c3_seq6	16	3	0	0	1	-4.27	0.05	L-type lectin-domain receptor kinase S.4	<i>Medicago truncatula</i>	KEH22964.1
comp48217_c0_seq1	30	86	0	12	5	-4.27	0	Putative peptide/nitrate transporter	<i>Morus notabilis</i>	XP_010088558.1
comp49431_c0_seq1	19	9	0	0	2	-4.24	0	Double-stranded RNA-binding protein 5	<i>Morus notabilis</i>	XP_010110932.1
comp54964_c0_seq3	4	26	0	0	3	-4.23	0.03	Major facilitator superfamily protein	<i>Theobroma cacao</i>	XP_007028238.1
comp54087_c0_seq1	14	4	0	0	1	-4.22	0.01	Dynein beta chain, ciliary	<i>Gossypium arboreum</i>	KHG15580.1
comp55592_c0_seq3	23	14	1	0	2	-4.19	0	Serine/threonine protein kinase isolog	<i>Arabidopsis thaliana</i>	AAM20520.1
comp54494_c0_seq1	8	16	0	0	2	-4.19	0.01	C2H2-like zinc finger protein	<i>Theobroma cacao</i>	XP_007045432.1
comp45452_c0_seq1	16	19	0	0	3	-4.14	0.01	LOB domain protein 1	<i>Solanum demissum</i>	AAT40528.1
comp3590_c0_seq1	8	8	0	0	1	-4.13	0.02	Uninformative		
comp55572_c0_seq8	10	6	1	0	0	-4.11	0.02	Pentatricopeptide repeat-containing -like protein	<i>Gossypium arboreum</i>	KHG29369.1
comp53352_c0_seq2	10	6	1	0	0	-4.1	0.02	SWI/SNF complex subunit SMARCC2	<i>Medicago truncatula</i>	XP_003602787.1
comp14610_c0_seq1	79	14	1	2	6	-4.05	0	Triacylglycerol lipase	<i>Ricinus communis</i>	XP_002513514.1
comp29578_c0_seq1	12	10	0	0	2	-3.98	0.01	Cytochrome P450	<i>Medicago truncatula</i>	XP_003607255.1
comp40102_c0_seq1	54	51	8	0	1	-3.98	0	Dihydroflavonol-4-reductase	<i>Ginkgo biloba</i>	AGR34043
comp43206_c0_seq1	16	8	2	0	0	-3.98	0.01	Chromatin assembly factor 1 subunit A-B	<i>Gossypium arboreum</i>	KHG13021.1
comp38971_c0_seq2	10	13	2	0	0	-3.95	0.03	Plant neutral invertase family protein isoform 1	<i>Theobroma cacao</i>	XP_007041939.1
comp54755_c0_seq16	56	60	2	14	1	-3.91	0.05	2-methyl-3-buten-2-ol synthase	<i>Pinus greggii</i>	AFJ73549.1
comp42778_c0_seq1	10	16	0	0	3	-3.86	0.01	Lipase	<i>Medicago truncatula</i>	KEH44038.1
comp25525_c0_seq1	24	10	0	4	0	-3.85	0	JAZ10	<i>Gossypium barbadense</i>	AJT58397.1
comp44305_c0_seq1	11	22	1	2	1	-3.83	0	Putative quinone oxidoreductase-like protein	<i>Gardenia jasminoides</i>	AIX10953.1
comp49135_c0_seq2	17	17	3	0	0	-3.82	0.04	Antimicrobial peptide 1	<i>Pinus sylvestris</i>	AAL05052.1
comp45534_c0_seq1	159	183	17	18	8	-3.79	0.01	Phenylalanine ammonia-lyase	<i>Pinus sylvestris</i>	AAL74331.1
comp16981_c0_seq1	7	15	0	3	0	-3.78	0.02	Heavy metal transport/detoxification superfamily protein	<i>Theobroma cacao</i>	XP_007050858.1
comp569_c0_seq1	16	31	2	0	4	-3.78	0.01	WRKY family transcription factor	<i>Medicago truncatula</i>	KEH20878.1
comp26588_c0_seq1	12	6	0	2	0	-3.72	0	Cellulose-synthase-like C12	<i>Theobroma cacao</i>	XP_007018423.1

Appendix 1. continued

Reference sequence	Sm3	Sm4	Sm5	Sm6	Sm7	logFC	PValue_ fdr	Gene annotation	Organism	Accession No.
comp54636_c0_seq21	8	4	0	1	0	-3.72	0.04	Putative reverse transcriptase	<i>Arachis hypogaea</i>	ACN78497.1
comp52692_c0_seq1	8	4	0	0	1	-3.72	0.02	Similar to RPS-2 disease resistance protein	<i>Arabidopsis thaliana</i>	AAK96709.1
comp51409_c0_seq1	49	17	1	0	7	-3.7	0.01	Cytochrome P450 71A1	<i>Morus notabilis</i>	XP_010099912.1
comp34441_c0_seq1	14	16	0	2	2	-3.7	0	R2R3-MYB transcription factor MYB11	<i>Picea glauca</i>	ABD60286.2
comp55437_c0_seq1	565	687	31	177	35	-3.68	0	Carbonic anhydrase	<i>Ricinus communis</i>	XP_002529418.1
comp47738_c0_seq1	18	14	3	1	0	-3.66	0.01	Laccase 1	<i>Theobroma cacao</i>	XP_007036761.1
comp52474_c0_seq10	31	4	0	0	4	-3.65	0.05	Plant-specific domain TIGR01589 family protein	<i>Medicago truncatula</i>	XP_003621784.1
comp45295_c0_seq1	161	55	11	10	5	-3.64	0.01	Cinnamoyl-CoA reductase 1	<i>Betula platyphylla</i>	AIX92145.1
comp44346_c0_seq2	5	6	0	0	1	-3.63	0.04	PREDICTED: uncharacterized protein ycf45 isoform X1	<i>Nicotiana tomentosiformis</i>	XP_009631320.1
comp77850_c0_seq1	26	8	2	0	2	-3.61	0.01	myb-like HTH transcriptional regulator family protein	<i>Arabidopsis thaliana</i>	NP_973385.1
comp32428_c1_seq1	631	385	39	126	17	-3.61	0	3-deoxy-D-arabino-heptulosonate 7-phosphate synthase	<i>Picea abies</i>	AET43997.1
comp52474_c0_seq1	4	53	3	3	4	-3.6	0.03	Plant-specific domain TIGR01589 family protein	<i>Medicago truncatula</i>	XP_003621784.1
comp46505_c1_seq1	13	5	1	0	1	-3.6	0.02	DNA repair protein	<i>Aphanomyces invadans</i>	XP_008869217.1
comp54314_c0_seq5	13	43	0	6	4	-3.58	0.01	Class I chitinase	<i>Pinus contorta</i>	AEF59005.1
comp55573_c0_seq5	198	77	7	5	28	-3.58	0.01	Serine carboxypeptidase-like 20	<i>Morus notabilis</i>	XP_010095610.1
comp48058_c0_seq1	14	10	2	0	1	-3.58	0.02	TPA: calmodulin-binding heat-shock protein	<i>Zea mays</i>	DAA39600.1
comp52473_c0_seq1	17	8	1	0	2	-3.58	0.01	Dormancy/auxin associated-like protein	<i>Picea sitchensis</i>	ADP94909.1
comp42341_c0_seq1	52	32	5	0	6	-3.56	0	Acireductone dioxygenase family protein	<i>Populus trichocarpa</i>	XP_002311641.1
comp4637_c0_seq1	21	8	0	2	2	-3.55	0.01	ABC-2 and Plant PDR ABC-type transporter family protein isoform 1	<i>Theobroma cacao</i>	XP_007024294.1
comp37103_c0_seq1	13	14	0	2	2	-3.52	0.01	PREDICTED: pollen-specific leucine-rich repeat extensin-like protein 3	<i>Brassica rapa</i>	XP_009106541.1
comp32319_c0_seq1	583	320	66	29	13	-3.52	0.02	Bark storage protein A	<i>Glycine soja</i>	KHN20661.1
comp404334_c0_seq1	10	6	0	1	1	-3.49	0.02	Putative TIR-NBS-LRR protein	<i>Pinus monticola</i>	ADW94527.1
comp46803_c0_seq3	10	6	0	0	2	-3.44	0.05	Uninformative		
comp55665_c0_seq12	14	8	0	0	3	-3.41	0.05	Glutaredoxin	<i>Cucumis sativus</i>	AGX01496.1
comp52043_c1_seq1	11	30	2	3	2	-3.4	0.03	Embryo-abundant protein	<i>Picea glauca</i>	AAB01567.1
comp42931_c0_seq1	49	32	9	0	1	-3.4	0.01	Phospholipase A1-Igammal	<i>Arabidopsis thaliana</i>	NP_849603.1
comp47676_c0_seq1	32	44	4	2	7	-3.4	0	Phospholipase A1-Igammal3	<i>Morus notabilis</i>	XP_010102366.1
comp54725_c0_seq10	8	17	1	2	1	-3.39	0.03	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	<i>Theobroma cacao</i>	XP_007048464.1

Appendix 1. continued

Reference sequence	Sm3	Sm4	Sm5	Sm6	Sm7	logFC	PValue_ fdr	Gene annotation	Organism	Accession No.
comp46913_c0_seq1	15	5	0	2	1	-3.38	0.02	Plasma membrane-like protein isoform 1	<i>Theobroma cacao</i>	XP_007051655.1
comp54700_c0_seq1	20	28	0	0	9	-3.36	0.03	Ethylene response factor 14	<i>Actinidia deliciosa</i>	ADJ67443.1
comp54817_c0_seq1	347	634	33	167	48	-3.35	0.01	Class VII chitinase	<i>Pinus contorta</i>	AEF59008.1
comp41510_c0_seq1	16	13	1	2	2	-3.34	0.01	Protein-tyrosine phosphatase mitochondrial 1-like protein	<i>Morus notabilis</i>	XP_010104857.1
comp53534_c0_seq5	10	14	2	2	0	-3.34	0.03	Ramosa1 enhancer locus 2 protein	<i>Phacelurus digitatus</i>	AFJ42403.1
comp48963_c0_seq1	18	13	2	0	3	-3.33	0.01	Surfeit locus 1 cytochrome c oxidase biogenesis protein isoform 1	<i>Theobroma cacao</i>	XP_007036878.1
comp54624_c1_seq2	24	44	1	6	8	-3.29	0.01	NADH dehydrogenase subunit 7	<i>Cycas taitungensis</i>	YP_001661405.1
comp36636_c1_seq1	30	12	2	3	2	-3.28	0.01	Transmembrane protein 184A	<i>Glycine soja</i>	KHN08579.1
comp49034_c0_seq1	249	183	14	62	22	-3.27	0	Lipoxygenase	<i>Malus domestica</i>	AGI16408.1
comp28007_c0_seq1	70	38	2	6	13	-3.26	0	Nodulation-signaling pathway 1 protein	<i>Malus domestica</i>	NP_001287803.1
comp45321_c0_seq2	16	15	0	5	1	-3.25	0	Anthocyanidin reductase	<i>Ginkgo biloba</i>	AAU95082.1
comp51062_c0_seq1	9	23	2	5	0	-3.23	0.03	Blue copper protein precursor	<i>Zea mays</i>	NP_001149787.1
comp51256_c0_seq1	54	21	4	4	5	-3.23	0	MYBPA1 protein	<i>Vitis vinifera</i>	NP_001268160.1
comp51857_c0_seq1	21	8	2	2	1	-3.22	0.01	E3 ubiquitin-protein ligase RHA1B	<i>Morus notabilis</i>	XP_010112947.1
comp47291_c0_seq1	83	27	8	9	0	-3.2	0.01	ATP-binding cassette transporter	<i>Selaginella moellendorffii</i>	XP_002991414.1
comp40886_c0_seq1	20	20	0	8	1	-3.2	0	Beta-glucosidase	<i>Medicago truncatula</i>	XP_003636137.1
comp37546_c0_seq1	709	1526	208	270	40	-3.16	0.04	Nicotianamine synthase	<i>Lotus japonicus</i>	BAH22562.1
comp42663_c0_seq1	12	10	2	0	2	-3.11	0.03	Histone-lysine N-methyltransferase ASHH1	<i>Glycine soja</i>	KHN46399.1
comp35349_c0_seq1	56	18	3	4	7	-3.11	0	Zeatin O-glucosyltransferase	<i>Glycine soja</i>	KHN07349.1
comp46467_c0_seq1	6	27	0	6	4	-3.1	0.05	Protein HOTHEAD	<i>Gossypium arboreum</i>	KHG26488.1
comp39845_c0_seq1	22	41	6	2	5	-3.08	0.02	Class VII chitinase	<i>Pinus contorta</i>	AEF59007.1
comp53013_c0_seq1	10	10	0	0	4	-3.06	0.04	Sterile alpha motif domain-containing protein	<i>Theobroma cacao</i>	XP_007023327.1
comp51961_c0_seq2	16	22	4	2	2	-2.99	0.02	Histone-lysine N-methyltransferase SETD1A	<i>Theobroma cacao</i>	XP_007048538.1
comp46618_c0_seq2	123	148	7	74	6	-2.98	0.02	Thiazole synthase	<i>Gossypium arboreum</i>	KHG00545.1
comp26013_c0_seq1	13	6	1	2	1	-2.92	0.03	R2R3-MYB transcription factor MYB9	<i>Picea glauca</i>	ABQ51225.1
comp49020_c0_seq2	239	183	23	99	12	-2.91	0.02	Chitinase 2	<i>Glycine soja</i>	KHN30221.1
comp44677_c0_seq1	75	22	3	11	10	-2.89	0.02	Adenine nucleotide alpha hydrolases-like superfamily protein	<i>Theobroma cacao</i>	XP_007044293.1
comp16765_c0_seq1	15	4	1	2	1	-2.89	0.05	TPX2 protein family	<i>Theobroma cacao</i>	XP_007032694.1
comp41158_c0_seq1	16	24	3	2	5	-2.87	0.02	Uninformative		
comp53914_c0_seq1	18	34	0	14	4	-2.87	0.02	LOB domain protein	<i>Medicago truncatula</i>	XP_003597159.1

Appendix 1. continued

Reference sequence	Sm3	Sm4	Sm5	Sm6	Sm7	logFC	PValue_ fdr	Gene annotation	Organism	Accession No.
comp53386_c0_seq1	19	23	4	0	6	-2.79	0.05	Caffeic acid ortho-methyltransferase	<i>Pinus radiata</i>	AAD24001.1
comp242031_c0_seq1	12	4	0	4	0	-2.78	0.03	Retrotransposon protein	<i>Oryza sativa</i> <i>Japonica Group</i>	ABA99600.2
comp49900_c0_seq1	10	10	0	2	3	-2.78	0.05	Nucleolar protein gar2-related isoform 2	<i>Theobroma cacao</i>	XP_007028895.1
comp52440_c0_seq1	58	23	5	5	10	-2.77	0.01	Cytochrome P450	<i>Gossypium arboreum</i>	KHG08008.1
comp54305_c0_seq10	25	11	2	3	4	-2.75	0.02	Putative potassium transporter/channel	<i>Arabidopsis thaliana</i>	BAC42897
comp54748_c0_seq12	14	18	3	6	1	-2.66	0.04	RING/U-box superfamily protein	<i>Theobroma cacao</i>	XP_007018124.1
comp45543_c0_seq2	59	17	6	10	4	-2.62	0.01	Lanatoside 15'-O-acetyltransferase	<i>Digitalis lanata</i>	AAAY42522.1
comp38572_c0_seq1	134	66	11	42	11	-2.6	0.02	Clavamate synthase-like protein	<i>Arabidopsis thaliana</i>	NP_188773.1
comp55534_c0_seq33	170	99	16	63	14	-2.59	0.03	GCR2-like 1	<i>Theobroma cacao</i>	XP_007047870.1
comp51584_c0_seq1	129	91	6	55	20	-2.58	0.04	ABC-2 and Plant PDR ABC-type transporter family protein isoform 1	<i>Theobroma cacao</i>	XP_007024294.1
comp52761_c0_seq1	39	40	10	13	3	-2.52	0.04	PDR1	<i>Picea abies</i>	ADP55080.1
comp55358_c0_seq23	19	26	2	9	5	-2.5	0.02	Glutamate decarboxylase	<i>Pinus pinaster</i>	ABA18653.1
comp54745_c1_seq1	34	22	4	5	9	-2.47	0.02	PPPDE peptidase domain-containing 2	<i>Gossypium arboreum</i>	KHG00551.1
comp38121_c0_seq1	20	11	0	7	4	-2.46	0.03	AT-hook motif nuclear localized protein 20	<i>Theobroma cacao</i>	XP_007035040.1
comp46420_c0_seq1	36	28	2	14	9	-2.44	0.01	Pleckstrin homology domain-containing family M member 3	<i>Morus notabilis</i>	XP_010087674.1
comp51011_c0_seq2	20	9	2	6	2	-2.4	0.03	Uninformative		
comp49687_c0_seq1	59	33	8	8	16	-2.26	0.04	Uninformative		
comp52634_c0_seq1	49	28	6	26	5	-2.25	0.03	3-dehydroquinate dehydratase/shikimate 5-dehydrogenase	<i>Juglans regia</i>	AAW65140.1
comp43649_c0_seq1	56	26	10	8	10	-2.25	0.05	IAA-leucine resistant-like gene 6 isoform 1	<i>Theobroma cacao</i>	XP_007019095.1
comp31662_c0_seq1	43	41	7	17	12	-2.21	0.03	Ribulose-5-phosphate-3-epimerase	<i>Ricinus communis</i>	XP_002509539.1
comp50798_c0_seq1	48	46	11	20	12	-2.13	0.04	Putative serine/threonine-protein kinase WNK11	<i>Glycine soja</i>	KHN43812.1
comp47976_c0_seq1	62	33	6	21	16	-2.08	0.04	Protein ETHYLENE INSENSITIVE 3	<i>Morus notabilis</i>	XP_010106128.1
comp54938_c0_seq1	7	2	14	75	29	2.65	0.03	Cobalamin binding protein	<i>Desulfosporosinus meridiei</i>	WP_014903849.1
comp19487_c0_seq1	10	2	28	88	48	2.86	0.01	Zinc-binding protein	<i>Cryptomeria japonica</i>	BAE92295.1
comp54809_c0_seq5	7	2	19	118	38	3.15	0.04	Alpha-mannosidase	<i>Pyrus x bretschneideri</i>	AGR44468.1
comp50587_c0_seq1	4	0	7	80	9	3.44	0.05	30S ribosomal protein S1	<i>Glycine soja</i>	KHN23663.1
comp45265_c0_seq2	0	2	9	27	13	3.47	0.04	F-box/kelch-repeat protein	<i>Morus notabilis</i>	XP_010090191.1
comp62062_c0_seq1	4	0	15	74	12	3.52	0.02	Rhodanese/Cell cycle control phosphatase superfamily protein	<i>Theobroma cacao</i>	XP_007037783.1
comp556453_c0_seq1	3	0	16	42	17	3.61	0.02	Pleiotropic drug resistance 11	<i>Theobroma cacao</i>	XP_007023609.1

Appendix 1. continued

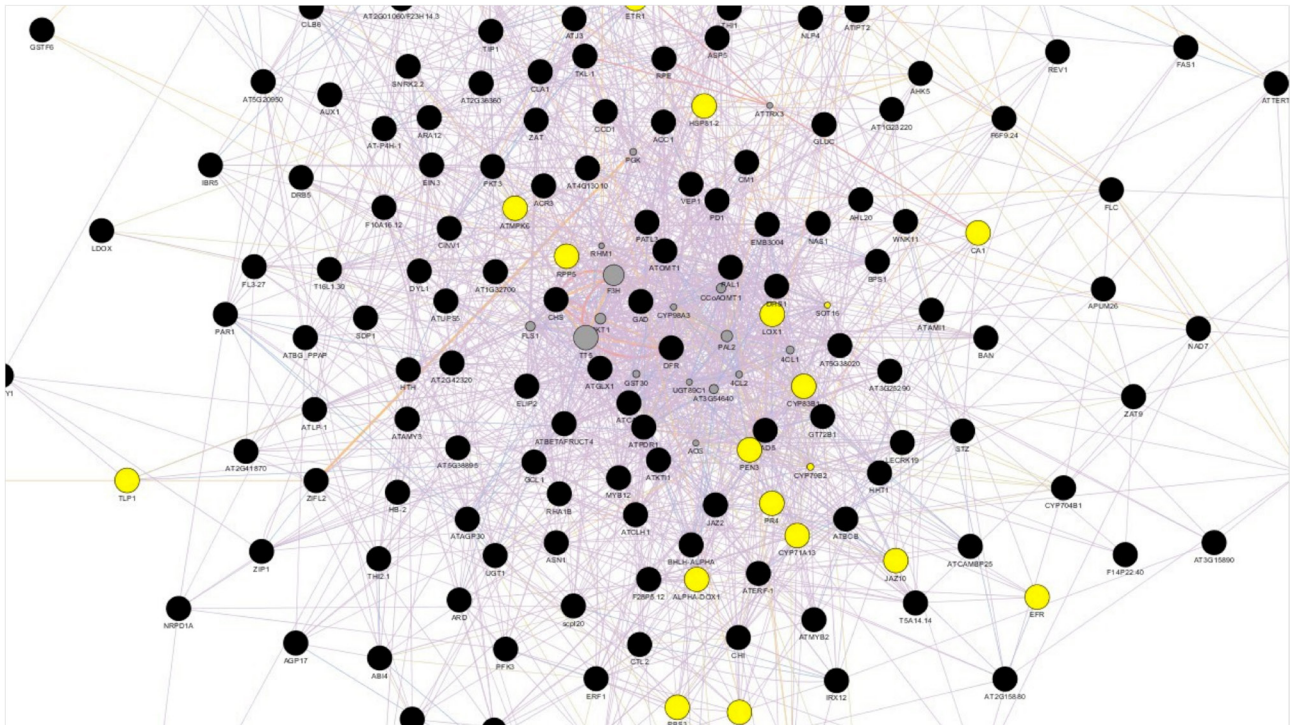
Reference sequence	Sm3	Sm4	Sm5	Sm6	Sm7	logFC	PValue_ fdr	Gene annotation	Organism	Accession No.
comp32378_c0_seq1	2	2	10	85	19	3.64	0.01	ATP-dependent Clp protease proteolytic subunit	<i>Pinus thunbergii</i>	NP_042379.1
comp54841_c0_seq18	2	1	9	77	10	3.64	0.04	NBS/LRR	<i>Pinus taeda</i>	AAM28912.1
comp47096_c0_seq1	2	0	13	43	6	3.71	0.03	Pentatricopeptide repeat-containing protein	<i>Ricinus communis</i>	XP_002511505.1
comp42175_c0_seq1	0	3	13	76	19	3.81	0.03	Cytochrome P450 78A3	<i>Morus notabilis</i>	XP_010100198.1
comp52935_c0_seq2	2	0	12	40	12	3.81	0.02	Nitrate transporter	<i>Cucumis sativus</i>	AFR11354.1
comp70039_c0_seq1	2	2	4	134	8	3.83	0.05	Regulator of G-protein signaling	<i>Chara braunii</i>	AHB52760.1
comp55154_c0_seq1	2	0	10	58	5	3.87	0.03	Myosin heavy chain-related	<i>Theobroma cacao</i>	XP_007041107.1
comp52063_c0_seq8	2	0	11	44	14	3.89	0.02	Phosphate transporter 1,4	<i>Theobroma cacao</i>	XP_007050310.1
comp43157_c0_seq1	2	0	13	64	7	3.9	0.05	Clavata-like receptor	<i>Picea glauca</i>	ABF73316.1
comp49546_c1_seq1	2	0	8	64	6	3.92	0.04	Putative AT-hook DNA-binding protein	<i>Oryza sativa</i> <i>Japonica Group</i>	BAD10063.1
comp55194_c0_seq1	3	2	13	145	40	3.93	0.01	Too large		
comp49025_c0_seq3	1	0	8	18	9	3.93	0.02	Heme oxygenase-like, multi-helical isoform 1	<i>Theobroma cacao</i>	XP_007030327.1
comp48011_c0_seq1	4	0	17	96	17	3.94	0	WRKY transcription factor 14-1	<i>Dimocarpus longan</i>	AEO31477.1
comp42365_c0_seq1	3	0	12	83	22	3.96	0.01	Properoxidase	<i>Picea abies</i>	CAL25300.1
comp41730_c0_seq1	3	0	13	81	18	3.99	0.01	Adenylate kinase isoenzyme	<i>Medicago truncatula</i>	XP_003617100.1
comp55326_c0_seq4	1	0	5	25	14	4	0.03	F-box/LRR-repeat 3-like protein	<i>Gossypium arboreum</i>	KHG22457.1
comp35564_c0_seq1	2	0	4	82	12	4.01	0.05	Pre-mRNA cleavage complex II protein family isoform 1	<i>Theobroma cacao</i>	XP_007010520.1
comp53805_c0_seq1	1	0	6	31	10	4.09	0.02	Biotin/lipoate A/B protein ligase family protein	<i>Arabidopsis thaliana</i>	NP_189543.2
comp55489_c0_seq4	1	0	11	21	9	4.09	0.02	Pentatricopeptide repeat (PPR) superfamily protein	<i>Theobroma cacao</i>	XP_007043514.1
comp55660_c0_seq1	2	0	6	80	7	4.1	0.04	SWIM zinc finger family protein	<i>Theobroma cacao</i>	XP_007040728.1
comp51780_c0_seq2	1	0	5	33	14	4.18	0.02	Boron transporter	<i>Ricinus communis</i>	XP_002519293.1
comp48504_c0_seq1	0	2	9	93	8	4.21	0.04	Subtilisin-like protease	<i>Glycine soja</i>	KHN12975.1
comp50903_c0_seq3	1	0	5	46	8	4.24	0.02	Splicing factor, arginine/serine-rich 12	<i>Zea mays</i>	ACG33547.1
comp51704_c0_seq1	2	0	6	84	13	4.32	0.01	NAC domain-containing protein	<i>Glycine soja</i>	KHN03216.1
comp55152_c0_seq2	1	0	12	34	13	4.51	0.01	Isoamylase 2	<i>Morus notabilis</i>	XP_010090367.1
comp47615_c0_seq1	1	0	3	56	15	4.59	0.03	Cornifelin	<i>Medicago truncatula</i>	XP_003629262.1
comp43387_c0_seq3	1	0	2	66	10	4.61	0.04	Fibrillin 8	<i>Coffea canephora</i>	ABD39695.1
comp55616_c0_seq1	1	0	10	45	17	4.68	0.01	LRR receptor-like serine/threonine-protein kinase	<i>Theobroma cacao</i>	XP_007029832.1
comp43106_c0_seq1	0	2	46	33	27	4.72	0.01	Leucine-rich repeat family protein	<i>Populus trichocarpa</i>	XP_002314568.2
comp42365_c0_seq2	3	2	42	271	33	4.76	0.01	Properoxidase	<i>Picea abies</i>	CAL25300.1
comp18457_c0_seq1	1	0	6	86	3	4.89	0.02	Uninformative		
comp55645_c0_seq1	2	0	19	113	22	4.89	0	Alpha 1,3 fucosyltransferase	<i>Populus tremula</i> × <i>Populus alba</i>	CAI70373.1
comp30617_c0_seq1	0	2	7	156	40	5.02	0.02	PLATZ transcription factor family protein isoform 2	<i>Theobroma cacao</i>	XP_007018634.1

Appendix 1. continued

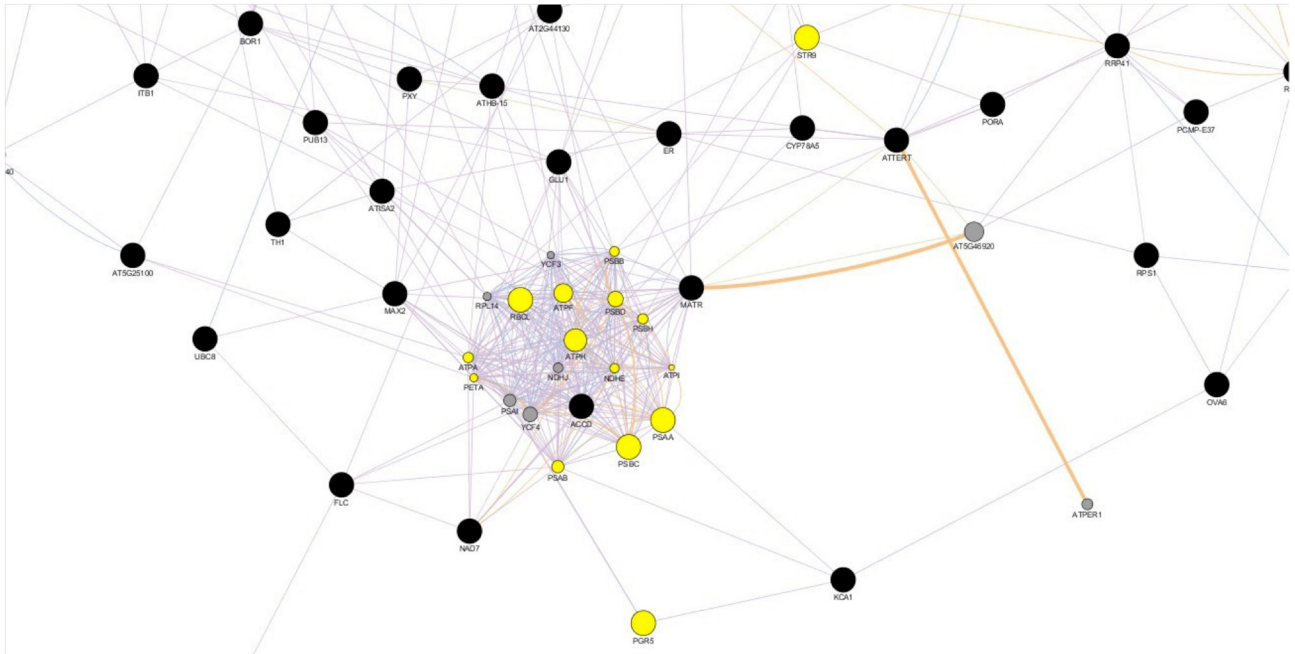
Reference sequence	Sm3	Sm4	Sm5	Sm6	Sm7	logFC	PValue_ fdr	Gene annotation	Organism	Accession No.
comp53594_c0_seq1	1	0	4	96	9	5.05	0.02	Nucleotide-diphospho- sugar transferases superfamily protein	<i>Theobroma cacao</i>	XP_007029705.1
comp54467_c0_seq3	1	0	7	118	9	5.23	0.02	Exosome complex exonuclease RRP46 like	<i>Glycine soja</i>	KHM99392.1
comp20002_c0_seq1	0	0	5	22	6	5.3	0.05	Protein lifeguard 4	<i>Arabidopsis thaliana</i>	NP_171806.1
comp52704_c0_seq1	0	0	7	18	7	5.36	0.04	Pentatricopeptide repeat- containing protein	<i>Ricinus communis</i>	XP_002532091.1
comp51350_c0_seq1	0	0	6	19	10	5.46	0.04	Subtilisin-like protein	<i>Picea abies</i>	BAA13135.1
comp55244_c0_seq5	0	0	4	27	8	5.5	0.04	LRR receptor-like serine/ threonine-protein kinase FLS2	<i>Triticum urartu</i>	EMS54167.1
comp50151_c0_seq1	0	0	6	26	7	5.51	0.03	Protein FAM91A1	<i>Gossypium arboreum</i>	KHG03579.1
comp53367_c0_seq2	0	0	11	19	4	5.56	0.05	TPA: putative WAK receptor-like protein kinase family protein	<i>Zea mays</i>	DAA38848.1
comp18781_c0_seq1	0	0	12	20	6	5.66	0.02	F-box family protein	<i>Populus trichocarpa</i>	XP_002315636.1
comp41403_c0_seq1	0	0	8	37	4	5.77	0.03	Condensin-2 complex subunit D3 isoform 3	<i>Theobroma cacao</i>	XP_007035610.1
comp50394_c0_seq2	0	0	10	24	9	5.79	0.02	Circadian clock coupling factor family protein	<i>Populus trichocarpa</i>	XP_002306041.1
comp38917_c0_seq1	0	0	6	30	11	5.81	0.02	Coatomer subunit beta	<i>Medicago truncatula</i>	KEH28616.1
comp42062_c0_seq1	0	0	6	37	7	5.84	0.03	Serine-threonine protein kinase, plant-type	<i>Ricinus communis</i>	XP_002526889.1
comp55102_c0_seq1	0	0	12	31	2	5.85	0.03	Putative TIR-NBS-LRR protein	<i>Pinus monticola</i>	ADW94527.1
comp48926_c0_seq1	0	0	8	35	7	5.9	0.02	Zinc finger protein 4	<i>Theobroma cacao</i>	XP_007032204.1
comp42562_c0_seq1	0	0	13	18	16	5.9	0.05	Plant-specific domain TIGR01615 family protein	<i>Zea mays</i>	NP_001147221.1
comp49784_c0_seq1	0	0	13	32	3	5.94	0.02	Uninformative		
comp20321_c0_seq2	0	0	9	46	2	6	0.03	BRCA1-A complex subunit Abraxas	<i>Gossypium arboreum</i>	KHG22147.1
comp50442_c0_seq1	0	0	10	34	4	6	0.01	U-box domain-containing protein 4	<i>Morus notabilis</i>	XP_010110625.1
comp41706_c0_seq1	0	0	5	59	2	6.06	0.05	Plastid RNA polymerase sigma factor X	<i>Marchantia polymorpha</i>	BAM95333.1
comp470497_c0_seq1	0	1	9	217	18	6.1	0	Patatin-T5	<i>Aegilops tauschii</i>	EMT08201.1
comp54342_c0_seq1	0	0	9	47	6	6.11	0.01	long cell-linked locus protein	<i>Zea mays</i>	NP_001147138.1
comp53191_c0_seq1	0	0	6	50	8	6.15	0.02	Putative leucine-rich repeat receptor-like protein kinase family protein	<i>Zea mays</i>	AFW83643.1
comp41173_c0_seq1	0	0	2	52	17	6.19	0.04	Subtilase family protein	<i>Theobroma cacao</i>	XP_007024651.1
comp55364_c0_seq1	0	0	2	71	6	6.22	0.05	Proline--tRNA ligase	<i>Gossypium arboreum</i>	KHG24304.1
comp52104_c0_seq1	0	0	5	70	7	6.39	0.02	Aspartic proteinase Asp1 precursor	<i>Ricinus communis</i>	XP_002514831.1
comp35394_c0_seq1	0	0	8	74	4	6.48	0.02	DIS3-like exonuclease 1	<i>Bos taurus</i>	NP_001071624.1
comp51856_c0_seq1	0	0	1	72	16	6.53	0.02	Transmembrane protein 8B	<i>Glycine soja</i>	KHN26950.1
comp54808_c0_seq4	0	0	5	83	11	6.58	0.01	Putative TIR-NBS-LRR protein	<i>Pinus monticola</i>	ADW94527.1

Appendix 1. continued

Reference sequence	Sm3	Sm4	Sm5	Sm6	Sm7	logFC	PValue_	Gene annotation	Organism	Accession No.
								fd		
comp41304_c1_seq1	0	0	11	78	2	6.6	0.01	MFS transporter	<i>Medicago truncatula</i>	AES74091.2
comp41406_c0_seq1	0	0	13	71	12	6.63	0	Class III HD-Zip protein HDZ31	<i>Pinus taeda</i>	ABG73245.1
comp54020_c0_seq9	0	0	6	81	11	6.73	0	Aspartic proteinase nepenthesin-1	<i>Morus notabilis</i>	XP_010095993.1
comp56560_c0_seq1	0	0	11	102	9	6.8	0.01	Anthocyanidin reductase	<i>Ginkgo biloba</i>	AAU95082.1
comp45074_c0_seq1	0	0	4	118	7	6.84	0.02	Catechol-O-methyltransferase	<i>Xenopus laevis</i>	NP_001090344.1
comp45070_c0_seq1	0	0	2	112	11	6.85	0.02	Wound-induced protein 1-like protein	<i>Phaseolus vulgaris</i>	AGV54254.1
comp50246_c0_seq1	0	0	30	36	22	6.9	0	Leucine-rich receptor-like protein kinase family protein	<i>Theobroma cacao</i>	XP_007018364.1
comp39355_c0_seq1	0	0	31	84	16	7.32	0	50S ribosomal protein L3	<i>Listeria ivanovii</i>	WP_014093814.1
comp57238_c0_seq1	0	0	19	169	12	7.52	0	Anthocyanidin reductase	<i>Ginkgo biloba</i>	AAU95082.1



Appendix 2. The center of the up regulated gene network. Black nodes are genes from imported list, grey nodes are genes added by the Cytoscape 3.2.1 program. Genes highlighted in yellow are involved in immune responses.



Appendix 3. The center of the down regulated gene network. Black nodes are genes from imported list, grey nodes are genes added by the Cytoscape 3.2.1 program. Genes highlighted in yellow are coding proteins associated with the chloroplast thylakoid membrane.