Pine transcriptomics — RNA-Seq data analysis of Scots pine (*Pinus sylvestris*) seedlings subjected to a wounding experiment.

Master Thesis
Department of Agricultural Sciences
University of Helsinki
Spring 2012.

Okeke Godfrey Uche
Student Number: 013707280
Supervisor: Professor Teemu Teeri

Table of Contents

List of tables
List of figures
Glossary
Abbreviations List
Abstract

List of tables:

List of figures and boxes:

Glossary

Reads: Nucleotide strings that are derived from sequencing of cDNA sample libraries.
Pair-ends: two sequencing reads that derive from both ends of the same cDNA molecule.
Contig: a stretch of DNA sequence produced by joining a collection of overlapping reads.
Scaffold: a series of ordered non-overlapping contigs that are linked by paired-end information.
Genome coverage: a parameter for sequencing amount measurement, which can be calculated by dividing the total length of all sequenced reads by the genome size.
Sequencing gaps: non-sequenced regions between contigs belonging to the same scaffold, which could be filled by walking over the corresponding contigs by tactical PCR and sequencing.
Mapping: Aligning short reads from next generation sequencers to a reference genome.
Assembly: Building contigs from sequence reads usually via an overlap consensus algorithm or de Bruijn graphs.

Abbreviations List

SNP: Single Nucleotide Polymorphism.

STS: Stilbebe Synthase

PSTS: Pine Stilbene Synthase

CHS: Chalcone Synthase

PS: Pinosylvin

PSM: Pinosylvin Monomethyl ether.

PsPAL: *Pinus sylvestris* Phenyla lanine ammonia - lyase

Ps4CL: *Pinus sylvestris* 4-coumara te-CoA ligase

PsPMT: *Pinus sylvestris* Pinosylvin-O-methyltransferase

ABSTRACT:

The stilbene biosynthetic pathway yielding products pinosylvin and pinosylvin monomethyl ether has been implicated in Scots pine's (*Pinus sylvestris*) heartwood constitutive resistance to decay. This trait, which is very valuable in forest economies has been suggested to be genetically controlled. The molecular basis for this control is not known. Wounding by drilling in stems of Scots pine seedlings have been suggested to induce production of pinosylvin and pinosylvin monomethyl ether in xylem of these seedlings which transforms to heartwood on maturity. The aim of this study was to elucidate the transcripts co-expressed with candidate genes in the stilbene biosynthesis pathway (*Ps4CL*, *PsPAL*, *PsSTS* and *PsPMT*) and their expression patterns, reconstruct a comprehensive transcriptome and decipher the difference between half sib families in the expression of transcripts reacting to the wounding of Scots pine seedlings using the next generation sequence (ABI SOLiD) profiling of RNA (RNA-Seq). Our bioinformatic pipeline involving *de novo* transcriptome assembly and scaffolding, read mapping and Digital gene expression (DGE) analysis were utilized to analyze this RNA-Seq data and its results were validated with quantitative PCR. The results show that RNA-Seq analysis generated 25,010 non redundant consensus sequences (representing expressed transcripts) and DGE revealed that the Scots pine seedling transcriptome was altered with up-regulation of 830 EST tags (genes) and down-regulation of 518 EST tags. *PsSTS* was induced in the wounding experiment while *Ps4CL*, *PsPAL* and *PsPMT* appeared to be expressed constitutively. Of the total 77,326 tags from the PGI EST list (our pseudo reference transcriptome), 29% (mean) of reads from eight RNA-Seq libraries were mapped to 13,000 EST tags. QPCR results confirmed the RNA-Seq data although the latter appeared more sensitive. Read counts from mapped RNA-Seq could be used to distinguish differentially expressed genes between half sib families comprising of high and low decay resistance lines. We conclude that wounding Scots pine seedlings induced a large number of genes including *PsSTS*. With ABI SOLiD RNA-Seq and DGE technologies, the complex transcriptome of wounded Scots pine seedlings can be reconstructed, constituent expressed transcripts elucidated and half sib family lines with divergent decay resistant trait can be distinguished.

1. INTRODUCTION

Scots pine (*Pinus sylvestris)* produces in its heartwood the most durable timber in Finland and has an established forest economy in the country. It has been estimated that during the 1990's the proportion of pine out of the total artificial reforestation area in Finland was 62 percent, spruce 29 percent and birch 8 percent although during the past few years the proportion of pine has decreased and that of spruce and birch correspondingly increased [1]. About 65 percent of forests in Finland are privately owned hence there is high demand for high timber quality and decay resistant tree lines. Scots pine is very popular for construction use owing to the decay resistance of its heartwood. Heartwood extractives of Scots pine are known to constitute of accumulated phenolic compounds namely resin acids (about 90 percent) and about 10 percent stilbenes namely pinosylvin (PS) and pinosylvin monomethyl ether (PSM) [2,3]. There has been decades of research on stilbenes to dissect their roles as constitutive antimicrobials responsible for anti-decay and durability of many tree species [4]. Stilbenes occur naturally in a number of plants and are present in wood, bark and leaves of gymnosperm trees. Being implicated in discoloration of wood and inhibition of delignification via acid pulping in wood, stilbenes are controversial as industrial compounds [5, 6, 7]. It is already known that heartwood stilbenes and other heartwood secondary metabolites are formed *in situ* in dying xylem parenchyma cells at the heartwood- sapwood boundary [8]. Evidence suggests that stilbenes are formed from carbohydrates or lipid compounds with the amount and type formed mainly dependent on the plant's genes and the physiological conditions affecting the parenchyma cells [4, 8]. The role of heartwood extractives in plant resistance against fungal and insect attacks is evident [9]. Wood that contains stilbenes show slow rates of decomposition on exposure to decay fungi and PSM has been shown to inactivate fungal enzymes containing -SH groups (cellulases, xylanases and pectinases) [10]. It has been reported that in the *Pinus* genus, PS and PSM are very vital metabolites that can kill nematodes [11]. Fungal susceptibility to stilbenes depends on synthesis and activities of their extracellular cellulolytic enzymes with increased synthesis and activities of these enzymes leading to reduced susceptibility [12]. Stilbenes have been established to be involved in resistance of heartwood to fungal decay and they may function as phytoalexins which are crucial in plant disease resistance [12, 13]. This fact (combined with the fact that the heartwood extractive constituents are under genetic control depending on the photosynthates quantity available to this region) have kept these compounds on the research spotlight with anticipated applications in tree breeding. Relatively high heritability estimates have been reported in half sib families for the concentration of PS and PSM [3] but this trait is seen only late in life at approximately 20 years of age.

1.1 Stilbene Synthase and the biosynthesis of stilbenes

Pine stilbenes are abundant in dead tissues of heartwood and bark although they have also been shown to be present in leaf tissues and sapwood in small quantities as glycoside and methyl ester complexes. Stilbenes have been known to possess inhibitory effects on photosynthesis; hence it becomes apparent why they are found in minute quantities in photosynthetic tissues. As mentioned above, they are formed *in situ* from lipids and carbohydrates by living parenchyma cells in response to stress. The genetic control of stilbene formation and the various forms of physiological conditions including stress that induce them suggest that there exist genetic variants for stilbene formation. The pine pinosylvin (PS) is the product of Pine *stilbene synthase* (*PsSTS*) encoded by a family of genes in pine (*PST1* being the most active gene of the family). *PsSTS* forms PS from malonyl-CoA and cinnamoyl-CoA and has been isolated and characterized decades ago [14, 15, 16]. There are a number of *PsSTS* variants in pines with suggestions that different isozymes of STS possess different product inhibition properties. [17]. Stilbene synthase (STS) and chalcone synthase (CHS) are members of Type III polyketide synthase (PKS) superfamily and share more than 65 percent animo acid identity. STS and CHS have common substrates (malonyl-CoA and cinnamoyl-CoA) and hence compete against each other for these precursors [18, 19]. *PsSTSs* have been classified as cinnamoyl-CoA specific STSs (EC 2.3.1.146) [17] and have been reported in *Pinus sylvestris* [20] and *Pinus strobus* [21]. A recently isolated structural variant of STS has been reported from *Pinus densiflora* (*PDSTS3*) lacking a C-terminal extension of STS due to a frame shift mutation yet it has shown a very high potential for PS formation [17]. It has been established that conserved amino acid residues in the C-terminus of STS and CHS are essential for their catalytic activities [22, 23, and 24]. There have been suggestions [25] that *PDSTS3* lacking the product inhibition normally observed for these enzymes may lead to cell death due to accumulation of PS; this type of stilbenoid formation might be operative in the heartwood [17]. Another variant *PDSTS2*, fails to accumulate stilbenoids at high levels due to the retained product inhibition and has been implicated in phytoalexin formation in response to stress [17]. Product inhibition could be critical in maintaining balance between CHS and STS biosynthesis and CHS-catalyzed reactions can be regulated by *PsSTS* biosynthesis (with suggestions that *PDSTS3* may function in down-regulating flavonoid biosynthesis) [17].The pathway for stilbene biosynthesis is shown in Figure 1.

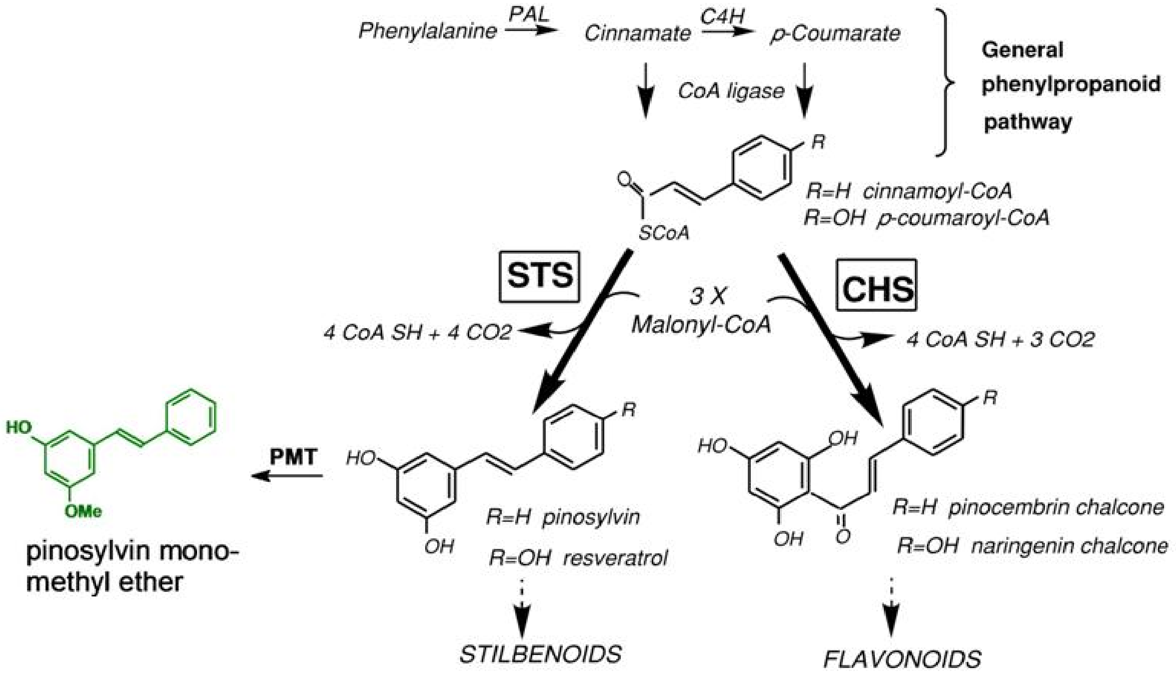


Figure 1: Stilbene and flavonoid biosynthetic pathway [17 (with modifications)].

Pine stilbenes, PS and PSM are constitutive in the heartwood as these tissues are dead but living cells of many plant species including Scots pine produce these compounds in response to stress and are especially important for plant resistance to pathogens [13]. These stress conditions include wounding, pathogens, exposure to ultraviolet radiation and ozone, desiccation and exposure to ethylene; experimental biosynthesis can thus be induced using any of these approaches. Understanding the transcriptomic profile in this biosynthetic pathway upon induction is critical for developing DNA markers, proteomic assays and metabolic fingerprinting that could be utilized for understanding the molecular mechanisms underlying decay resistance in Scots pine heartwood at an early stage in life and further applying this knowledge in stock selection.

1.2 Scots pine wounding — induced reactions and a method for genetic selection of durable Scots pine heartwood

The phloem, sapwood and cambium in trunk of the Scots pine have preformed (structural) defenses including the bark on the outside and the heartwood inside [26]. However, these preformed defenses alone cannot provide full protection against biological, mechanical and other stress conditions inflicted on the tree. It is known that pathways for secondary compounds are induced in response to stress conditions by plants as a way to support the preformed defense system. Mechanical wounding of phloem [27] and sapwood [28] induces production of phenolic compounds including PS, PSM and pinocembrin in so called decay resistant trees. Lyr [29] reported that a PS and PSM-containing heartwood was formed in the sapwood region (a somewhat abnormal occurrence) after wounding of Scots pine bark and cambium.

Harju *et al*. [30] reported that wounding (by drilling) in three year old Scots pine seedlings induced a wounding reaction in the xylem (xylem is the tissue that develops into the heartwood on maturity). However, analysis of xylem extracts showed newly synthesized products including stilbenes (PS and PSM) and lignans. Sealing of the wounds by resin flow was observed as a resultant reaction. We can recall that the Scots pine heartwood's defense mechanism in contrast to the induced reaction after wounding is constitutive due to accumulation of PS, PSM and resin acids which are known to inhibit fungal growth implicated in decay (shown above). Relatively high heritability estimates have been reported [31] for Scots pine heartwood extractives indicating that these extractives are in part genetically controlled. Resistance of Scots pine heartwood to decay is a trait that can only be manifested on maturity. Being a very desirable trait, an early testing method is needed for purposes of breeding durable and superior heartwood tree lines. Moderate to high offspring-parent heritability estimates of heartwood extractives from wounded Scots pine seedlings [30] suggest that wounding can lead to discovery of decay resistant genotypes. This is because wounding response appear to be molecularly correlated with mature Scots pine heartwood extractives responsible for decay resistance [30]. Since the stilbene pathway is turned on (via control of the *PsSTS*-1) during wound response in Scots pine seedlings at same location which forms the heartwood on maturity (xylem), it suggests that decay resistance is a trait which appear to be correlated with wounding response. Harju *et al.* [30] suggest further that wounding Scots pine seedlings by drilling holes through their stems would be an effective method for genetic testing of seedlings whose heartwood will become resistant to decay on maturity. For breeding purposes, an early testing method to identify candidate Scots pine seedlings for better heartwood properties is necessary. Wounding of Scots pine seedlings is well suited (but may be destructive in practice) as an early testing method [30] as it is able to show differential responses between individuals. This differential response to wounding due to genetic variability of Scots pine seedlings is highly desired as it will aid in elucidating superior genotypes and candidate genes involved. These variants and genes will form the basis for marker assisted selection and conventional breeding in Scots pine. At such, methods including transcriptomics will help to unravel candidate genes and SNPs which will be utilized in gene testing and marker assisted selection for Scots pine heartwood decay resistance.

1.3 The Pine Gene Index (PGI) and PGI EST resource

The DFCI (Dana-Farber Cancer Institute, Harvard USA) *Pinus* Gene Index integrates research data from international *Pinus* EST sequencing and gene research projects. The ultimate goal of the DFCI Gene Index projects, including PGI, is to represent a non-redundant view of all *Pinus* genes and data on their expression patterns, cellular roles, functions, and evolutionary relationships [68]. The PGI EST resource (PGI032811.fa) contains about 77,326 sequences alongside a PGI.GO.032811 file containing GO (Gene Ontology) annotation class of some of these EST sequences. This resource, although incomplete is very valuable and provided the pseudo reference transcriptome for our study.

1.4 RNA-Seq — Deep Sequencing of RNA reads

Understanding the complexity underlying the biological functions and processes in any organism requires a precise map of all genes, their alternative isoforms and their expression profiles across diverse cell types and tissues within the organism. However, a comprehensive study of the transcriptome of any organism given a particular condition or treatment was rare until the era of the microarrays and next generation sequencing techniques. Next generation sequencing technologies has indeed revolutionalized biology through the production of millions and billions of reads data at low cost. Their applications are enormous, spanning through all areas of biology from transcriptomics to epigenomics, genomics, quantitative genomics and so much more.

RNA sequencing (RNA-Seq) refers to the experimental procedures that generate DNA reads derived from the RNA (transcripts) of an organism. The transcriptome represents a comprehensive set of transcribed regions throughout the genome. RNA-Seq has emerged as a very powerful tool for studying transcriptomes aiming at a full reconstruction of both the entire transcripts and their expression levels of the diverse cell types and conditions of any organism. RNA-Seq can be very efficient in building a comprehensive map of any organism's transcriptome across all cell types, perturbation and states. RNA-Seq also provides a far more precise measurement of levels of transcripts and their isoforms than other methods and can reveal precisely the location of transcription boundaries to a single-base resolution [33]. RNA-Seq methods have shown to be an excellent tool for detection of alternative splicing, gene fusion events and novel classes of non-coding RNAs which are giving rise to new insights in biology.

RNA-Seq is free from many limitations of previous technologies including high background noise from microarray and limited detection capability for all arrays including the tiling array. The fact that it does not depend on prior knowledge of an organism as required by microarrays and PCR offers a huge opportunity for studying the large abundance of biodiversity worldwide which has been neglected for long as a result of concentration to model organisms and because of limiting technologies. Most RNA-Seq experiments isolate and purify RNAs from their samples, shear it, convert it into cDNA and sequence it on a high throughput platform such as SOLiD, Roche 454, Illumina GA or HiSeq [34]. This process generates millions to billions of short reads ranging from 25-300 bases taken from one end of the cDNA fragment (single end fragments) or reads from both ends of the cDNA fragment (paired-end reads). Single end fragments can easily show the connection between two consecutive exons and more information including multiple exons connections can be seen with paired-end read data.

With all these advantages, this method does not come without challenges. These challenges start from the library construction method especially for large RNA molecules which need to be fragmented into smaller fragments ranging from 150 – 500 base pairs, thus introducing bias for transcript ends especially for 3'ends of the transcript. Other challenges include the PCR artifacts which could easily be resolved with replicates. More complex are the bioinformatics challenges arising from the need to analyze store and process the huge turnover of RNA-Seq data. A further issue for RNA-Seq is the sequencing coverage, which describes the transcript percentage in focus. Better coverage entails more sequencing depth, which is required to reconstruct rare transcript variants but this is at the tradeoff of increased cost. Nagalakshmi *et al*. [35] suggested that 30 million 35-nucleotide reads from poly(A) mRNA libraries were sufficient to observe transcription from most (>90%) genes for yeast cells grown under a single condition (that is, in nutrient-rich medium). However this may be true for organisms with not so complex transcriptomic profiles like yeast but indeed more will be required for complex transcriptomes especially in the plant kingdom. The fact that the remaining genes are presumably either not expressed under the assayed condition or do not have poly(A) tails, suggests that analyzing several different conditions and using alternative methods for cDNA synthesis can further increase the coverage.

1.5 Scots pine transcriptomics— Dissecting the constituent transcribed elements underlying the Scots pine heartwood biology

For a relatively comprehensive view of the entire transcriptome playing a role in the heartwood extractives biology of the Scots pine from an early developmental stage, the RNA-Seq approach is durable. This approach is required to unveil the functional elements involved in the Scots pine heartwood and for interpreting the phenotypic variation produced by genetic and environmental interplay. Previous studies conducted in many other species have shown that not only will the transcriptome be reconstructed, also allelic variants affecting phenotypes of interest can be uncovered especially for plant breeding applications as in this experiment where constituents of the Scots pine heartwood is under genetic control [36]. However for non sequenced genomes like the Scots pine and without a closely related sequenced genome, available Expressed Sequence Tags (ESTs) have to be utilized for data analysis and processing alongside *de novo* sequence assembly methods [37], which will be discussed in detail in this work. The main focus here was a detailed bioinformatics analysis pipeline for Scots pine RNA-Seq data from transcriptome reconstruction (involving reads mapping to pine PGI ESTs [32] and *de novo* isoform compliant assembly of reads) to analysis of differential gene expression (DGE) otherwise called the digital gene expression approach.

1.6 Analyzing RNA-Seq data — Prospects and challenges

Identifying the entire transcripts including novel transcripts, gene fusion events, small RNAs and alternatively spliced variants remains the key goal of transcriptomics. However, reconstructing a comprehensive transcriptome from RNA-Seq data posses a lot of computational and bioinformatic challenges. Several assemblers have been developed including Velvet [38], ABYSS [39], ALLPATHS [40] to tackle genomic DNA assembly but these tools may not be suitable for transcriptome assembly. This is because many of these assemblers rely on sequencing depth to resolve repeats in any genome whereas in the case of a transcriptome this depth relates to the increased expression and abundance of transcripts. This abundance is critical as it is utilized in decoding gene expression using counting methods, but using genome assemblers for transcriptome reconstruction will result in marking of abundant transcripts as repeats. Secondly, strand specific sequencing often utilized in RNA-Seq, needs to be used in resolving sense and antisense transcripts but this is not often the case for genomic DNA assembly. Thirdly, the challenge of resolving alternatively spliced variants remains a big issue for genome assemblers further giving evidence that they are not well suited for transcriptome assembly.

Generally, assembly of transcriptomes requires extensive sequencing especially of the complex transcriptomes of plants and animals, and sequencing platforms including illumina, SOLiD are frequently used. The 454 Roche platform offers longer reads with lower coverage but the use of hybrid sequencing platform involving a combination of illumina or SOLiD and 454 Roche provides a very valuable strategy which gives the opportunity to extend contigs and cover gaps created by illumina or SOLiD reads using the 454 long reads. As much as long reads are valuable, paired-end sequencing protocol offers more opportunities by utilizing overlapping reads to form longer reads.

It is a very good practice to remove all artifacts from RNA-Seq datasets before any form of analysis. These artifacts include sequencing adaptors, low complexity reads [41], and PCR duplicates and can lead to mis-assemblies. Usually, quality scores of these reads are used for the removal of artifacts in a transcriptome reconstruction analysis.

1.7 Transcriptome Reconstruction strategies

The availability of a high quality reference genome or a closely related genome usually determines the strategy to be used for transcriptome reconstruction. These strategies include:

1.7.1 Reference-guided strategy

When a high quality reference genome or a closely related genome is available, the transcriptome reconstruction processes can be guided by this reference. First, RNA-Seq dataset is mapped to the reference genome using splice-compliant mappers including Tophat [42], MapSplice [43], GSNAP [44] and Blat[45]. Then overlapping reads are clustered into a graph representing transcript isoforms followed by transversing this graph to resolve individual transcript isoforms. These steps are implemented currently in Cufflinks [46] and Scripture [47]. Splice-compliant mappers can be grouped into seed-and-extend mappers (including Blat, GSNAP) and those that implement the Burrows-Wheelers Transform (BWT) algorithm (eg. Tophat). The seed-and-extend mappers start by finding a seed (a substring of a read that exactly matches the genome), then extends and accurately resolves alignment using a local alignment algorithm known as the Smith-Waterman's algorithm. These mappers also shift gaps in local alignment to match known splice-sites. The BWT mappers are optimized for mapping reads to the reference, hence they are faster than the seed-and-extend mappers. Seed-and-extend mappers create a database of all possible splice-site combinations per locus and map reads which failed to align to the reference to this database.

After aligning reads to the reference, Cufflinks creates an overlap graph from all aligned reads and transverses this graph into alternative variants by finding the minimum paths covered by this graph, that is, resolves the minimum set of transcripts that explain the intron junctions within these reads[46].
By contrast, Scripture tries to construct a splice graph containing the base of the chromosome as nodes and adds edges (connections) between the nodes if there are reads that connect these nodes. It then resolves the isoforms by finding paths through the graph which have statistically significant read coverage [47]. Scripture is more robust than cufflinks, which tends to be more stringent.

This reference-guided strategy is highly sensitive and allows for discovery of novel transcripts and reconstruction of lowly expressed transcripts. It has been used successfully on animal transcriptomes but it is not well known how these tools (Cufflinks and Scripture) may perform for polyploid plant transcriptomes. It can be used to reconstruct transcriptomes from reads with coverage as low as 10X. The reliability of this strategy on a reference makes it difficult for its use in non-model organisms whose genomes have not been sequenced. Also erroneous transcriptomes may result due to the large errors and mis-assemblies which have been incurred on many sequenced genomes. Reference-guided approaches are also unable to assemble trans-spliced genes [48].

1.7.2 *De novo* strategy

This approach does not require a reference genome but it takes advantage of overlapping reads to find a consensus and resolve transcript complexity. It does not depend on any mapping step. Some *de novo* assembly tools have been developed including Trinity [49], Oases [50], and Trans-ABySS [51]. They assemble RNA-seq data a multiple times using the De Bruijn graph algorithm [52] to reconstruct transcripts and then merge the assembled contigs into larger contigs in a post processing step. Hybrid reads are also valuable in this strategy and will improve the length and quality of the assembled transcripts.

For non-model organisms and those without reference genomes, *de novo* assembly approach provides an opportunity to study the transcriptomes of these organisms. Even when a reference is available, it can be valuable to do a *de novo* assembly as it can reconstruct trans-spliced transcripts, novel transcripts not found in the genome and exogenous sequences in the genome including parasitic, viral sequences and so on. The contigs from the *de novo* approach can further be extended by aligning reads to the assembled contigs and finding linking reads between contigs. *De novo* assembly of large plant and animal transcriptomes are computationally challenging and require a lot of RAM memory and bioinformatics expertise. It is also susceptible to sequencing errors and can therefore produce chimeric sequences which cannot be distinguished easily from trans-spliced transcripts.

1.7.3 Combined approach

The reference-guided strategy can be combined with the *de novo* approach for transcriptome reconstruction. In this combination, two methods can be used as shown below.

1.7.3.1 Assemble-then-align approach

In the cases of availability of a closely related or a low quality reference genome, a *de novo* assembly can be performed and its resulting contigs aligned to the reference so as to extend and improve the quality of the reference. Afterwards, the reference-guided approach is employed to reconstruct the transcriptome.

1.7.3.2 Align-then-assemble approach

The align-then-assemble approach starts by reconstructing the transcriptome using the reference-guided approach and then taking reads that could not map to the reference as input for the *de novo* assembly process. This can be used to recover a good number of trans-spliced transcripts and foreign sequences including pathogens from the transcriptome.

Care should be taken while trying to choose a strategy for RNA-Seq data analysis and decisions should be made considering the transcriptome complexity, availability of high quality reference genome, availability of computational resources and bioinformatics expertise.

2. OBJECTIVES

This work was aimed at presenting a comprehensive bioinformatic analysis of RNA paired-end sequence reads from a wounding experiment of Scots pine seedlings (about 5 years old). The work is part of a larger project where biosynthesis of extractives is compared between pine heartwood development and stress induction. The primary objectives were:

(a) To elucidate what genes in addition to the genes of the stilbene biosynthesis pathway are reacting to the wounding treatment.

(b) To uncover the expression patterns of these genes during wounding.

(c) To differentiate between the gene expression patterns of the High heartwood decay resistant family lines and Low resistant family lines in the wounding experiments.

(d) To reconstruct a comprehensive transcriptome from the sequenced reads.

3. METHODS

3.1 Sequencing

The sequencing data analyzed in this work was produced in the following way: Scots pine seedlings were wounded by drilling holes through their stems. Then total RNA was extracted from ground unwounded and wounded debarked stems (xylem) of the seedlings three hours, one day and four days post wounding by the method of Chang *et al.* [75]. Following extraction, DNA was removed by digesting the sample with RNAse free DNase and the RNA was purified using Qiagen RNeasy plant mini kit (Qiagenin, Finland) according to manufacturer's instructions. Finally, the libraries were prepared using SOLiD Total RNA-Seq Kit (Life Technologies, Finland) according to manufacturer's instructions starting from the RiboMinus treatment and then sequencing done using the ABI SOLiD platform (SOLiD 3).

3.2 Mapping SOLiD reads to the Pine PGI EST

Pine PGI EST set was downloaded from the site (http://compbio.dfci.harvard.edu/cgi-bin/tgi/gimain.pl?gudb=pine) containing 77326 sequences in fasta format. Paired-end sequenced reads (forward reads were 50bp long and the reverse reads 35bp long) from the ABI SOLiD were filtered to exclude low quality reads and PCR duplicates. Reads were aligned to the Pine PGI EST set using SHRiMP2 [53] which is specifically efficient for colorspace reads and gapped alignment. Mapping was done with parameters (kmer) $k = 4$, multiple spaced seeds of weight(w) = 12, and other refined parameters including allowing for better micro-indel detection using (*-a -1*) that disabled limitation of full Smith-Waterman's algorithm, enabled genome index trimming (*V*), paired-end mode of opposite in *(-p 'opp-in')* and a mapping score threshold of 80 percent(*h=80%*). Mappings were done on per library basis. Only 8 libraries, of which 4 belonged to a high decay resistant family and another 4 belonging to low resistant family line were mapped.

3.3 Normalization and counting of mapped reads

Reads that mapped to the Pine PGI EST set were counted using the *sam2count.pl* Perl script [54] producing count data for each library. Methods implemented in edgeR [55] to normalize the counts for compositional bias in libraries and for differences between libraries in sequencing depth were used for normalization. These methods accounted for under-sampling in lowly expressed genes and oversampling in highly expressed genes by building an effective library size using an offset in a generalized linear model. Normalization factors were estimated for libraries using the *calcNormFactors()* function which implements the Trimmed Mean of M values method (TMM) and the sample whose three-quarter percentile is closest to the mean of the three-quarter percentiles of the reference.

3.4 Differential Gene Expression (DGE):

The DGE was tested by utilizing a model implemented in edgeR and based on the negative binomial distribution which estimates an over-dispersion parameter from read counts. This parameter is best suited for representing biological or sample-to-sample variability. DGE was tested first by estimating the trended dispersion parameters using the Cox-Reid profile-adjusted likelihood (CR) method for multiple factors. This trended dispersion method utilized quantiles to capture expression patterns of genes especially for very lowly expressed genes. Multiple factors and contrasts were accounted for by fitting a generalized linear model (GLM) given a design matrix and the differential expression tested using a GLM likelihood ratio test [55]. The design matrix which explicitly represented our data was an additive linear model with family (High and Low) as blocking factor and sample grouping as control and treatment (in this case RNA extracted in the wounding experiment at three hour, one day and four days taken as treatment and unwounded samples as controls).

3.5 *De novo* Isofrom compliant assembly:

For assembly of short read SOLiD sequences, Trinity RNA-Seq assembler [49] was used in conjunction with the AMOS assembly package [56]. Sixteen runs of Trinity were executed in parallel (one run per library for two low decay resistant families and two high decay resistant families with each family comprising of 4 libraries) with k-mer parameter = 25 for inchworm. The Minimo pipeline [57], a light weight assembler and part of the AMOS package was run using parameters *MIN_LEN=40*, *MIN_IDENT=85* and *98*, *FASTA_EXP=1* for contig merging to yield local transcriptomes. Using paired-end joining and gap filling from uniquely mapped reads from the 5471D library on one of the assembled local transcriptomes (Transcriptome 1), the contigs were further assembled into 2,152 scaffolds using the SOPRA[58] software.

3.6 Quantitative PCR validation for the expression of selected transcripts:

Genes belonging to the stilbene biosynthesis pathway were selected from two half sib families and sixteen cDNA libraries (for the purpose of validating the result of the DGE) including *Pinus sylvestrsis* 4-coumarate-CoA ligase (4CL*)* gene (*Ps4CL*) partial cds. with Genbank accession number EU392780 [59] , *Pinus sylvestrsis isolate Lillo22 phenylalanine ammonia-lyase (pal1) gene, partial cds* (*PsPAL*) with GenBank accession number AF353986, ORF from the patent sequence *Pinus sylvestrsis* PMT *(PsPMT*)with Genebank accession number AX212322, Pinosylvin-forming stilbene synthase [*Pinus sylvestris*, seedling, mRNA, 1221 nucleotides] with Genebank accession number S50350 (*PsSTS*). Primers for quantitative PCR (QPCR) were designed using the Primer3plus software [60]. The reference gene, a partial sequence from the *Picea abies* Ubiquitin ligase (*PaUB*) [32] was selected from the result of the differential expression analysis being the most stably expressed gene and was used to design forward and reverse oligonucleotide primers. Forward and reverse primers for *PaUB, Ps4CL, PsPAL, PsPMT* and *PsSTS* are shown in Box 1. LightCycler 480 SYBR Green I Master (Roche Applied Science) was used for all QPCR experiments, following the manufacturer's protocol. All samples were measured three times for target transcript levels (*Ps4CL*, *PsPAL*, *PsPMT* and *PsSTS*) and reference transcript levels (*PaUB*). All experiments were run using the LightCycler 480 System (Roche Applied Science). Each reaction consisted of 15 *μl* containing 5 *μl* of cDNA and 10 *pmol* of each primer. The cycling conditions were 1 cycle of denaturation at 95°C/10 min, followed by 45 three-segment cycles of amplification (95°C/10 sec, 57°C/10 sec, 72°C/10 sec) where the fluorescence was automatically measured during PCR and one three-segment cycle of product melting (95°C/5 sec, 65°C/1 min, 97°C/ continuous) and cooled (40°C). The baseline adjustment method of the Light Cycler software was used to determine the Ct in each reaction. A melting curve was constructed for each primer pair to verify the presence of one gene-specific peak and the absence of primer dimmer. All samples were amplified in triplicates and the mean was used for further analysis.

Box 1 showing Forward and Reverse primers for QPCR experiment.

	Forward primers	Reverse primers
PsUB	AATCAGTGGGCCTTTGAATG	GTTCAAAGCCTGTCCAGCAT
Ps4CL	TATTAACGACCCGGAATCCA	AGCCACCTGGACAAACAAAT
PsPAL	ATGCTGAAGCTGCCAAATCT	CCGGGTAGTATCTTCGGACA
PsPMT	GACGCTGCCATTACACTTGA	CCAGGACTCCATCAACACCT
PsSTS	TTCGCTTCGATCCTTGCTAT	GCACACATCGGGATTCTTCT

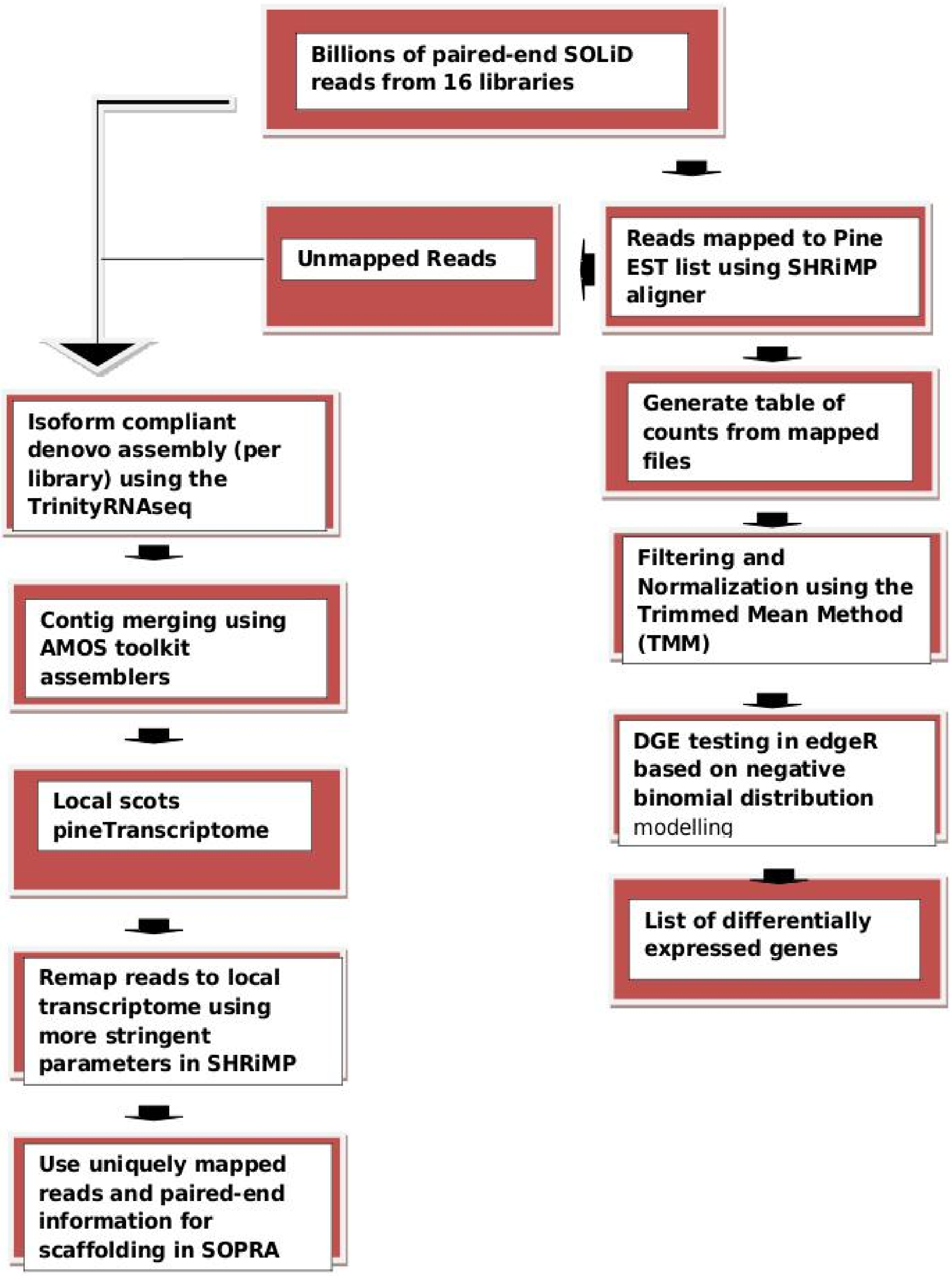


Figure 2: Analysis flowchart for Scots pine RNA-Seq data.

4. RESULTS

Our analysis pipeline is detailed in a flow chart shown in Figure 2.

4.1 Mapping reads to the reference EST list:

The use of EST abundance as an indicator for transcript abundance is a well accepted method for identification of differential gene expression. This method has been validated in numerous studies using ESTs from Sanger sequencing [70, 71]. We performed Scots pine transcriptome analysis at a single-nucleotide resolution using ABI-SOLiD RNA-Seq technology. Briefly, RNA from ground wounded (three hours, one day and four days post-wounding) and unwounded stems of Scots pine seedlings from two half sib families was sequenced. To quantify wounding response and tissue-specific transcript abundance in our Scots pine seedlings short-read dataset, millions of paired-end reads were mapped to the Pine EST list after quality filtering using SHRiMP2 [53] with a mapping score threshold of 80% and allowing for micro-indel detection. Mean percentage of reads mapped to the PGI EST list for the eight libraries were about 29 percent and mapping percentage for each library are shown in Table 1. About 70 percent of the unmapped reads were due to the incomplete reference, lack of a sequenced genome and the normal occurrence of novel transcripts from RNA-Seq experiments [69].

Table 1: Percentage of mapped reads from 8 sequence libraries from Scots pine RNA-Seq.

Library	**Mapped reads (%)**
395C	27.32
3953H	26.41
3951D	28.00
3954D	31.28
413C	30.84
4133H	30.91
4131D	31.74
4134D	31.97

395 and 413 represent two half sib families where the mother tree was characterized as low and high heartwood decay resistant lines respectively. The codes C, 3H, 1D, 4D indicate samples from Control (unwounded seedlings), 3 hours, 1day and 4days post wounding respectively.

Mapping of the RNA-Seq reads on the Pine EST revealed about 13,000 tags mapped out of the total 77,326 tags in the EST list.

4.2 Differential gene expression (DGE):

Scots pine DGE analysis was based on response to wounding stress between unwounded and wounded time series of the half sib libraries. Following mapping, reads mapped to the ESTs were counted and normalized (taking into account the compositional bias in libraries). We used edgeR [55] to fit a model using the negative binomial distribution for estimation of a parameter known as the over-dispersion parameter (which accounts for between sample variability) using the Cox-Reid (CR) profile adjusted likelihood method best suited for multiple factors. Subsequently, we fit a generalized linear model (GLM) using the sample-wise over-dispersion parameter estimates and given a design matrix representing a control versus treatment contrast with half sib family grouping as a blocking factor. DGE was tested in a GLM likelihood ratio test and multiple testing was accounted for by reporting False Discovery Rates (FDR) at 1%. Digital gene expression analysis between the libraries used generated 1,348 EST tags differentially expressed at $P<0.01$. 830 EST tags had positive log fold changes (logFC) and 518 EST tags had negative logFCs (positive logFC indicating up-regulation and negative logFC down-regulation). The top ten EST tags (ranked according to LogFC) for the up-regulated and down-regulated groups are given in Table 2 and a list of all the differentially expressed tags are shown in Appendix 1and 2. A good number of transcripts encoding dihydropinoslvin synthase (including TC154538, TC160074), pinosylvin synthase or pinosylvin-forming stilbene synthase (including TC195493, TC156839, TC168370, DR025377, DR167706) and different pathogen related defense proteins (including TC171264) were found among the up-regulated gene list (Appendix 1). A complex profile of transcripts representing genes involved in defense, metabolism, transport, transcription and other complex biological processes were found to be up-regulated. On the other hand, a good number of transcripts encoding the *Pinus taeda* (+)- alpha-pinene synthase (including TC155407, GT265035, FG616357, TC155384, TC177232), a complex profile of transcriptional regulators (including TC178048), potassium channels (TC185372), metabolism and other biological processes were found to be down-regulated (Appendix 2).

Table 2 showing top 10 each of upregulated and downregulated differentially expressed EST tags from DGE analysis of Scots pine RNA-Seq data

	logConc	**logFC**	**LR**	**P.Value**	**FDR**
DR096636	-12.25	6.1210	65.290	6.464e-16	4.157e-12
TC171087	-14.03	6.3020	64.410	1.012e-15	4.337e-12
TC171089	-12.27	5.8120	61.600	4.206e-15	1.352e-11
TC171992	-13.32	5.3170	58.830	1.721e-14	4.428e-11
TC181650	-13.75	4.7320	54.370	1.657e-13	3.552e-10
TC192333	-10.57	0.7098	51.910	5.804e-13	9.506e-10
TC186817	-13.61	4.6000	51.880	5.913e-13	9.506e-10
TC194485	-13.87	5.3020	50.760	1.042e-12	1.490e-09
TC175575	-13.32	4.6650	50.150	1.428e-12	1.713e-09
TC161878	-12.70	4.6600	49.880	1.631e-12	1.748e-09
TC199290	-10.397	-0.6218	71.720	2.474e-17	3.182e-13
TC178048	-13.135	-2.9160	50.090	1.465e-12	1.713e-09
TC190883	-12.963	-2.9330	44.550	2.484e-11	1.102e-08
TC191810	-12.962	-2.7680	43.240	4.850e-11	1.835e-08
TC174422	-12.502	-2.9600	42.610	6.670e-11	2.383e-08
TC160053	-12.140	-2.9520	38.870	4.537e-10	8.709e-08
TC156269	-12.891	-2.8780	38.050	6.892e-10	1.166e-07
TC195576	-13.066	-2.6680	38.030	6.980e-10	1.166e-07
TC193124	-12.795	-2.7680	37.840	7.698e-10	1.253e-07
BQ696511	-12.523	-2.9920	37.590	8.715e-10	1.384e-07

LogConc is a regression estimate of the abundance of each transcript in the sample (showing the base 2 logarithm relative quantity of each transcript between the test and control samples), LR is the likelihood ratio statistic, P.Value is the p-value from the LR test and FDR is the multiple testing adjusted value for false discovery rate using the Benjamini and Hochberg's method [61]. Refer to Appendices 1 and 2 for annotation of these target sequences.

4.3 Pine Transcriptome assembly:

In total, about 940 million paired-end reads of raw RNA-Seq data representing RNA from wounded stems of Scots pine seedlings from four half sib families were used for a wound response transcriptome assembly. The data comprised of sixteen libraries of unwounded and wounded time series RNA-Seq reads. Following a sequence quality filtering process, the reads were assembled using the Trinity RNA-Seq assembler [49] per library. Transcriptome assembly clipped into 25-mers resulted in a good number of contigs ranging from about 600 to 15,000 for the 16 libraries (Table 3). These contigs were further merged using the Minimo [57] pipeline with very stringent sequence identity parameter of 98% (giving rise to 25,010 contigs in Transcriptome 1 shown in Table 4) and also allowing for inclusion of more divergent sequences (85% similarity) into longer contigs (giving rise to 7,550 contigs in Transcriptome 2 also shown in Table 4) with median coverage per contig ranging from 8× to 550× (coverage of merged transcripts). The 85% sequence similarity parameter collapsed most of the variations in the sequences into consensus sequences with resultant better sequence lengths. Subsequently, we re-mapped reads from the 5471D library on Transcriptome 1 and making use of about 45% uniquely mapped paired-end reads to decipher contig connections and orientation, further scaffold these contigs into 2,152 super contigs (scaffolds) using the SOPRA[58] software.

Table 3: Number of assembled contigs per sequence library of Scots pine RNA-Seq data.

Library	**Number of Contigs**
395C	1,042
3953H	1,934
3951D	2,441
3954D	2861
413C	3,417
4133H	3,273
4131D	4,202
4134D	15,163
426C	799
4263H	919
4261D	693
4264D	964
547C	3,801
5473H	3,154
5471D	3,270
5474D	2,606

395,413, 426 and 547 represent four half sib families with 395 and 547 as low heartwood decay resistant lines and 413 and 426 the high heartwood decay resistant family lines. Codes including C, 3H, 1D, 4D represents sequenced control samples (unwounded seedlings), 3 hours, 1day and 4days post wounding samples respectively.

Table 4: Transcriptome and scaffold information of assembled Scots pine RNA-Seq data.

Transriptome 1:	
Total number of contigs	25,010
N50	354bp
Average contig size	347bp
Transcriptome 2:	
Total number of contigs	7,550
N50	509bp
Average contig size	490bp

Scaffold info:	
Scaffolds	2,152
Singletons	19,825
Largest scaffold size	17,800bp

Transcriptome 1 was generated from merged contigs from the assembly with sequence identity of 98% and Transcriptome 2 with sequence identity of 85%.

4.4 Quantitative PCR (qPCR) validation for the expression of genes in Stilbene biosynthesis pathway:

Quantitative transcriptome analysis using high-throughput sequencing platforms such as illumina and SOLiD has been shown to be accurate and highly correlated with other quantitative methods such as QPCR and microarray technologies [72, 73]. Due to limited number of samples, we utilized non sequenced half sib family samples comprising the 434 family with low heartwood decay resistance and 508 family with high heartwood decay resistance for the validation of the expression patterns of four candidate genes in the stilbene biosynthetic pathway (Figure 3, 4, 5, 6a and 6b). We observed variable expression patterns of the candidate genes *Ps4CL*, *PsPAL* and *PsPMT* (Figures 4, 5 and 6 respectively) through the post-wounding time course experiment for both the 434 and 508 families. However, the qPCR results supported the differential expression of the *PsSTS* (Figure 6a and 6b) from the DGE results which had a good number of different transcripts encoding stilbene synthase being up-regulated (Appendix 1).

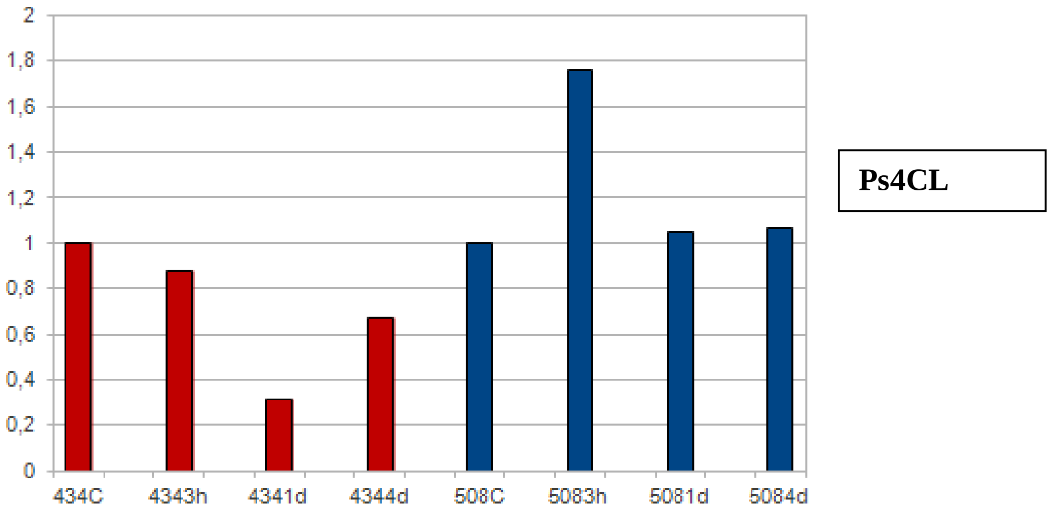


Figure 3: *Ps4CL* expression from QPCR results of two Scots pine half sib families in a post-wounding time course experiment.

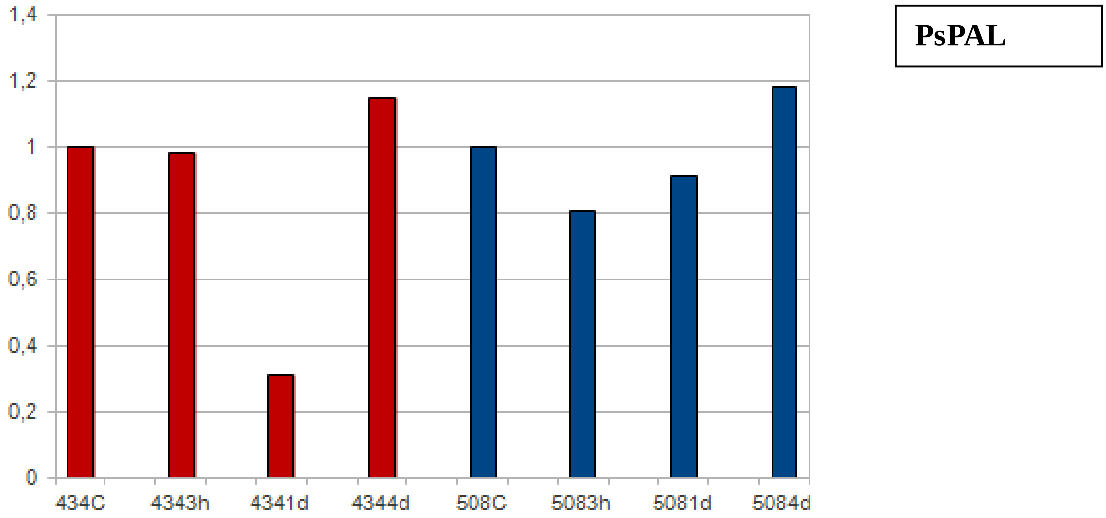


Figure 4: *PsPAL* expression from QPCR results of two Scots pine half sib families in a post-wounding time course experiment.

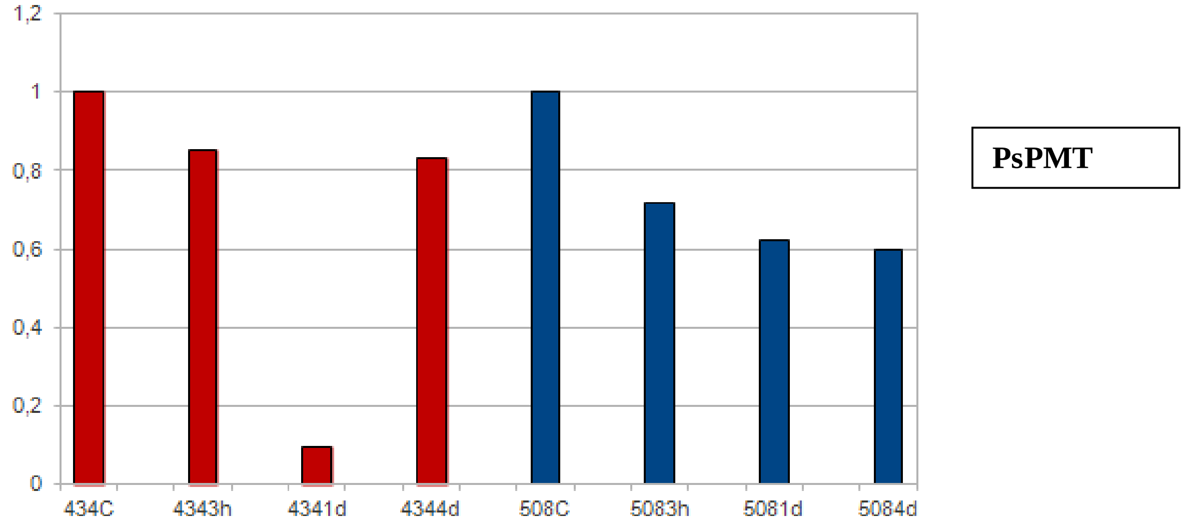
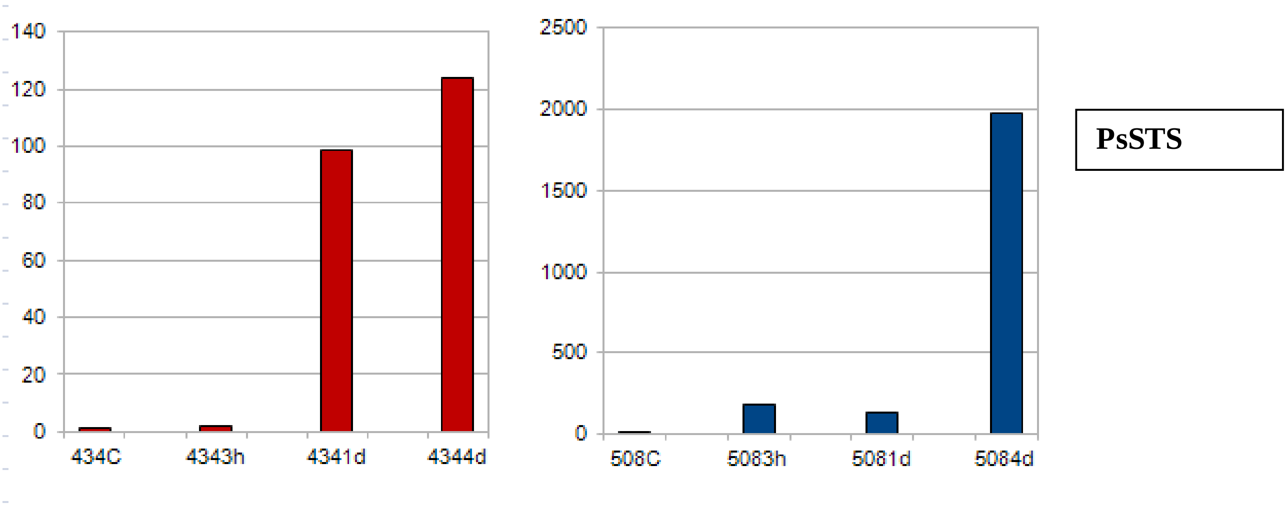


Figure 5: *PsPMT* expression from QPCR results of two Scots pine half sib families in a post-wounding time course experiment.

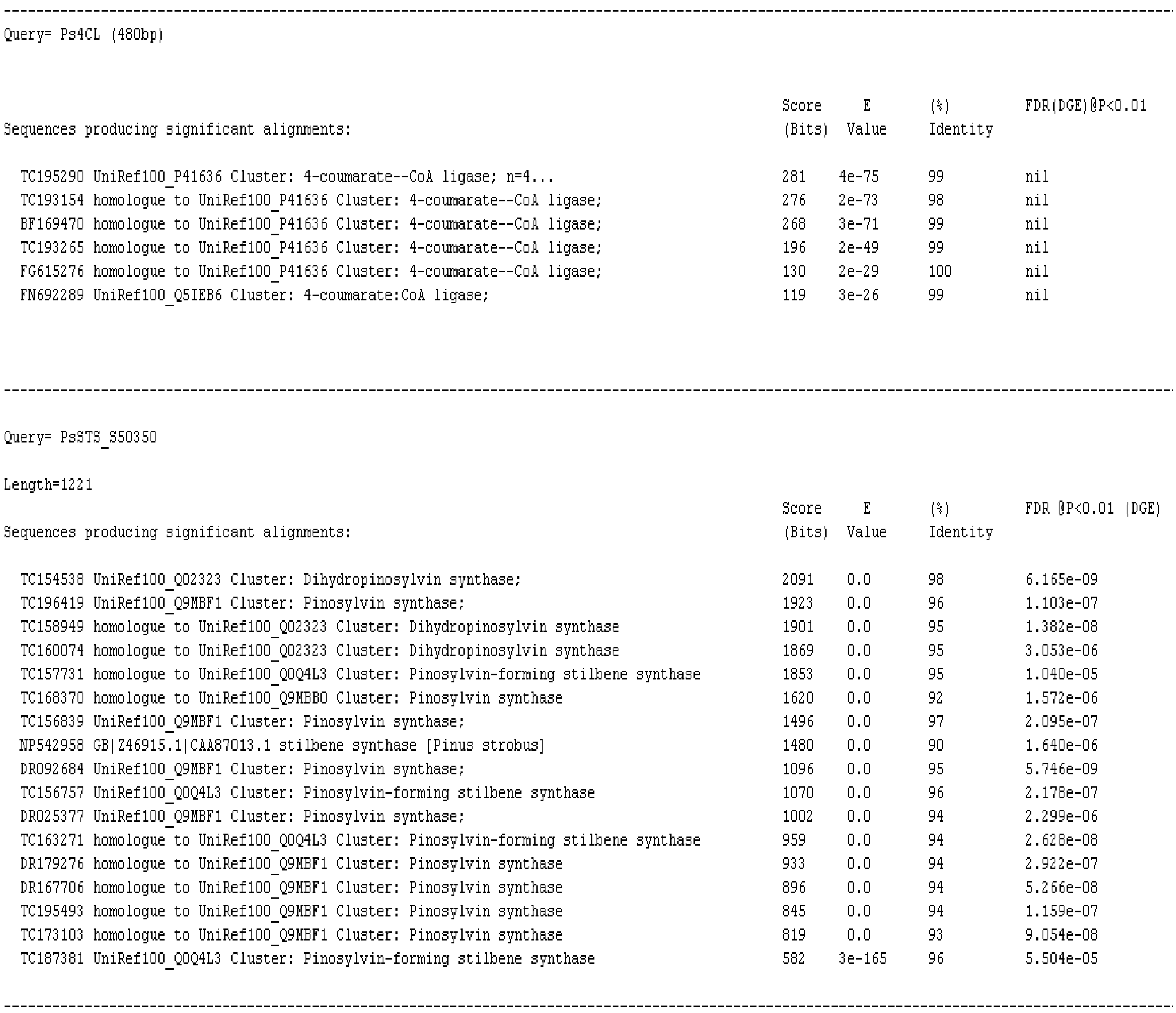


Figures 6a and 6b: *PsSTS* expression from QPCR results of two Scots pine half sib families in a post-wounding time course experiment. (Low resistant family 434 – Red and high resistant family 508 – Blue, note the different scales.)

4.5 Analysis of candidate genes from the stilbene biosynthesis pathway:

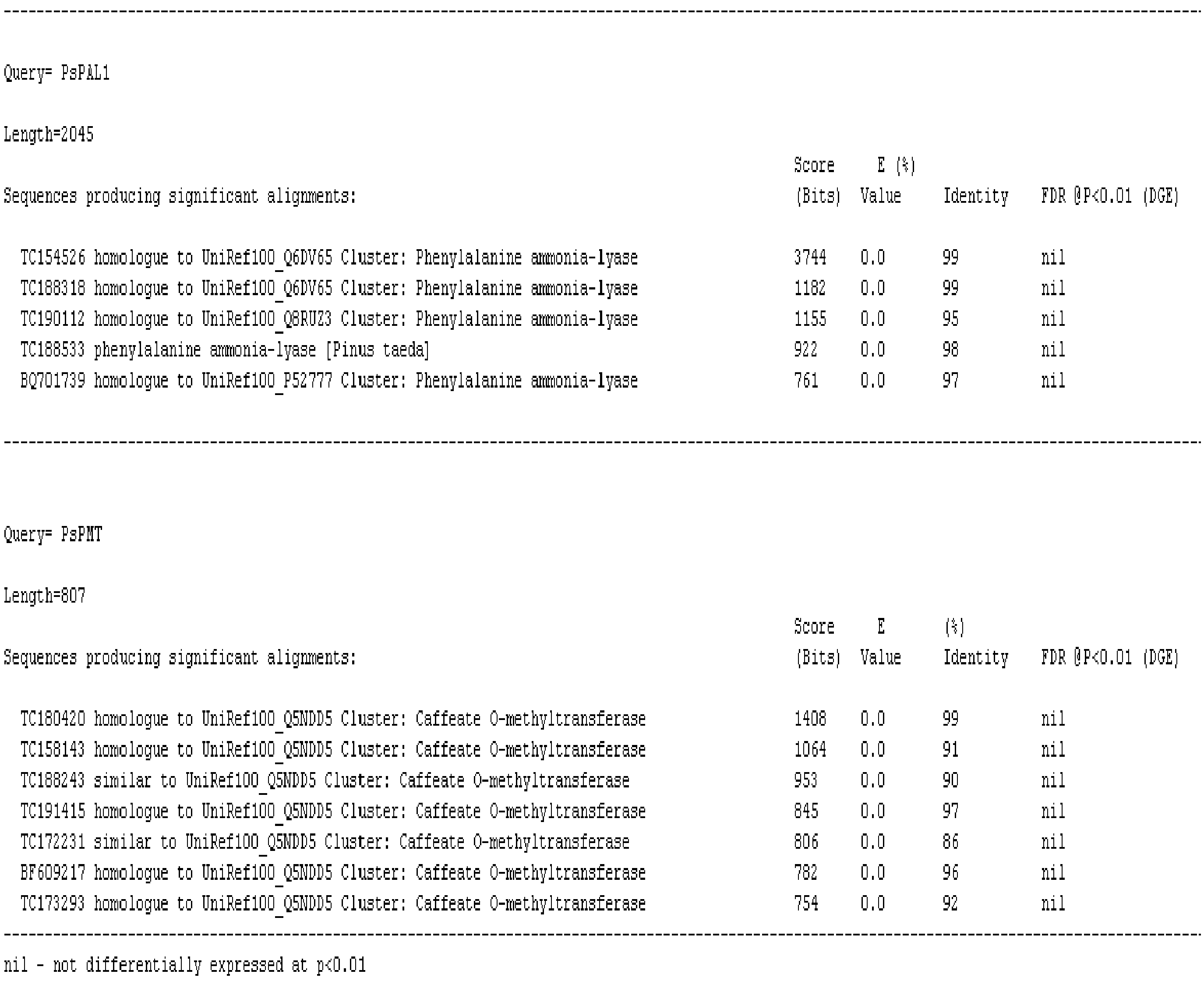
Further analysis was performed using BLASTN to extract some valuable information for our four stilbene biosynthesis pathway candidate genes (*Ps4CL*, *PsSTS*, *PsPMT* and *PsPAL*) from the pine PGI EST list. Table 5 show BLASTN hits of *Ps4CL* and *PsSTS* on the pine PGI EST list and their DGE values at P<0.01 (1% FDR) and Table 6 show BLASTN hits of *PsPMT* and *PsPAL* on the pine PGI EST list and their DGE values at P<0.01 (1% FDR). The vales represented with *nil* in Tables 5 and 6 show that the candidates *Ps4CL*, *PsPAL* and *PsPMT* were not differentially expressed (that is fluctuations in their expression patterns were not significant but rather crudely stable) given our multiple testing threshold. This was confirmed by QPCR results (Figures 3, 4 and 5). On the other hand Table 5 show clear significant differential expression of BLASTN hits of our candidate *PsSTS* which were also confirmed by our qPCR results shown in Figures 6a and 6b.

Table 5: BLASTN hits of *Ps4CL* and *PsSTS* on the pine PGI EST list and their FDR expression values at P<0.01. All nil values means values not differentially expressed at P<0.01.



```
----------------------------------------------------------------------------------------------------------------------------------------
Query= Ps4CL (480bp)


                                                                                              Score     E       (%)        FDR(DGE)@P<0.01
Sequences producing significant alignments:                                                   (Bits)  Value     Identity

  TC195290 UniRef100_P41636 Cluster: 4-coumarate--CoA ligase; n=4...                          281    4e-75      99         nil
  TC193154 homologue to UniRef100_P41636 Cluster: 4-coumarate--CoA ligase;                    276    2e-73      98         nil
  BF169470 homologue to UniRef100_P41636 Cluster: 4-coumarate--CoA ligase;                    268    3e-71      99         nil
  TC193265 homologue to UniRef100_P41636 Cluster: 4-coumarate--CoA ligase;                    196    2e-49      99         nil
  FG615276 homologue to UniRef100_P41636 Cluster: 4-coumarate--CoA ligase;                    130    2e-29      100        nil
  FN692289 UniRef100_Q5IEB6 Cluster: 4-coumarate:CoA ligase;                                  119    3e-26      99         nil



----------------------------------------------------------------------------------------------------------------------------------------

Query= PsSTS_S50350

Length=1221
                                                                                              Score     E       (%)        FDR @P<0.01 (DGE)
Sequences producing significant alignments:                                                   (Bits)  Value     Identity

  TC154538 UniRef100_Q02323 Cluster: Dihydropinosylvin synthase;                              2091     0.0      98        6.165e-09
  TC196419 UniRef100_Q9MBF1 Cluster: Pinosylvin synthase;                                     1923     0.0      96        1.103e-07
  TC158949 homologue to UniRef100_Q02323 Cluster: Dihydropinosylvin synthase                  1901     0.0      95        1.382e-08
  TC160074 homologue to UniRef100_Q02323 Cluster: Dihydropinosylvin synthase                  1869     0.0      95        3.053e-06
  TC157731 homologue to UniRef100_Q0Q4L3 Cluster: Pinosylvin-forming stilbene synthase        1853     0.0      95        1.040e-05
  TC168370 homologue to UniRef100_Q9MBB0 Cluster: Pinosylvin synthase                         1620     0.0      92        1.572e-06
  TC156839 UniRef100_Q9MBF1 Cluster: Pinosylvin synthase;                                     1496     0.0      97        2.095e-07
  NP542958 GB|Z46915.1|CAA87013.1 stilbene synthase [Pinus strobus]                           1480     0.0      90        1.640e-06
  DR092684 UniRef100_Q9MBF1 Cluster: Pinosylvin synthase;                                     1096     0.0      95        5.746e-09
  TC156757 UniRef100_Q0Q4L3 Cluster: Pinosylvin-forming stilbene synthase                     1070     0.0      96        2.178e-07
  DR025377 UniRef100_Q9MBF1 Cluster: Pinosylvin synthase;                                     1002     0.0      94        2.299e-06
  TC163271 homologue to UniRef100_Q0Q4L3 Cluster: Pinosylvin-forming stilbene synthase        959      0.0      94        2.628e-08
  DR179276 homologue to UniRef100_Q9MBF1 Cluster: Pinosylvin synthase                         933      0.0      94        2.922e-07
  DR167706 homologue to UniRef100_Q9MBF1 Cluster: Pinosylvin synthase                         896      0.0      94        5.266e-08
  TC195493 homologue to UniRef100_Q9MBF1 Cluster: Pinosylvin synthase                         845      0.0      94        1.159e-07
  TC173103 homologue to UniRef100_Q9MBF1 Cluster: Pinosylvin synthase                         819      0.0      93        9.054e-08
  TC187381 UniRef100_Q0Q4L3 Cluster: Pinosylvin-forming stilbene synthase                     582    3e-165     96        5.504e-05

----------------------------------------------------------------------------------------------------------------------------------------
```

Table 6: Blast hits of *PsPMT* and *PsPAL* on the pine PGI EST list and their FDR expression values at P<0.01.



```
-------------------------------------------------------------------------------------------------------------------------

Query= PsPAL1

Length=2045
                                                                              Score    E (%)
Sequences producing significant alignments:                                   (Bits)  Value    Identity    FDR @P<0.01 (DGE)

  TC154526 homologue to UniRef100_Q6DV65 Cluster: Phenylalanine ammonia-lyase   3744    0.0      99          nil
  TC188318 homologue to UniRef100_Q6DV65 Cluster: Phenylalanine ammonia-lyase   1182    0.0      99          nil
  TC190112 homologue to UniRef100_Q8RUZ3 Cluster: Phenylalanine ammonia-lyase   1155    0.0      95          nil
  TC188533 phenylalanine ammonia-lyase [Pinus taeda]                            922     0.0      98          nil
  BQ701739 homologue to UniRef100_P52777 Cluster: Phenylalanine ammonia-lyase   761     0.0      97          nil

-------------------------------------------------------------------------------------------------------------------------


Query= PsPMT

Length=807
                                                                              Score    E       (%)
Sequences producing significant alignments:                                   (Bits)  Value    Identity    FDR @P<0.01 (DGE)

  TC180420 homologue to UniRef100_Q5NDD5 Cluster: Caffeate O-methyltransferase  1408    0.0      99          nil
  TC158143 homologue to UniRef100_Q5NDD5 Cluster: Caffeate O-methyltransferase  1064    0.0      91          nil
  TC188243 similar to UniRef100_Q5NDD5 Cluster: Caffeate O-methyltransferase    953     0.0      90          nil
  TC191415 homologue to UniRef100_Q5NDD5 Cluster: Caffeate O-methyltransferase  845     0.0      97          nil
  TC172231 similar to UniRef100_Q5NDD5 Cluster: Caffeate O-methyltransferase    806     0.0      86          nil
  BF609217 homologue to UniRef100_Q5NDD5 Cluster: Caffeate O-methyltransferase  782     0.0      96          nil
  TC173293 homologue to UniRef100_Q5NDD5 Cluster: Caffeate O-methyltransferase  754     0.0      92          nil
-------------------------------------------------------------------------------------------------------------------------
nil - not differentially expressed at p<0.01
```

BLASTN analysis of these four stilbene biosynthesis pathway candidates (*Ps4CL*, *PsSTS*, *PsPMT* and *PsPAL*) on our local transcriptome assembly and scaffolds was continued to see if these genes were expressed although their differential expression was not tested. Observations from Table 7 show the presence of all these candidate genes in our transcriptome and scaffold sequences.

Table 7: BLASTN hits of *PsSTS*, *PsPAL*, *PsPMT*, and *Ps4CL* on our locally assembled scaffolds and Transcriptome 2.

Scaffolds BLASTN analysis of 4 candidate genes:

	(%)Identity	**E value**	**Sequence length(bp)**
Stilbene Synthase **(***PsSTS***)**	query	--	1221
Scaffold 1621	97	6e-159	931
Scaffold 1744	93	6e-144	575
PsPAL	query	--	2045
Scaffold 406	99	0.0	2416
Scaffold 1863	99	0.0	1626
Scaffold 6700(singleton)	99	0.0	1011
PsPMT	query	--	807
Scaffold 430	99	0.0	838
Scaffold 502	98	4e-139	543
Scaffold 1257	99	2e-107	795
Ps4CL	query	--	480
Scaffold 2036	99	2e-74	2498
Scaffold 8451(singleton)	100	6e-21	358

Transcriptome 2 Blast analysis of 4 candidate genes:

	(%)Identity	**E value**	**Sequence length(bp)**
Stilbene Synthase **(***PsSTS***)**	query	--	1221
NODE_1532	97	0.0	795
Contig 3150	97	3e-159	454
PsPAL	query	--	2045
NODE_1159	99	0.0	2332
Contig 7113	99	0.0	783
Contig 6621	99	0.0	1011
Contig 6740	99	0.0	816
PsPMT	query	--	807
Contig 498	99	0.0	407
Contig 5103	99	2e-139	338
Contig 607	99	8e-108	543
Ps4CL	query	--	480
NODE_948	99	2e-46	1104

4.6 Comparison of the QPCR results with RNA-Seq count data results:

Although there was no sequenced reference genome, RNA-Seq results were compared to the QPCR results using mapped read counts of the top BLASTN hits of these four candidate genes *Ps4CL*, *PsSTS*, *PsPAL* and *PsPMT* on the PGI EST list. This comparison was to test which method was more sensitive in showing the expression patterns of these genes. Figures 7, 8, 9 and 10 show the expression patterns of BLASTN top hits of *Ps4CL*, *PsSTS*, *PsPAL* and *PsPMT* using read counts. A literal view of these results show that the RNA-Seq approach appear more sensitive than the qPCR results (Figures 3-6b).

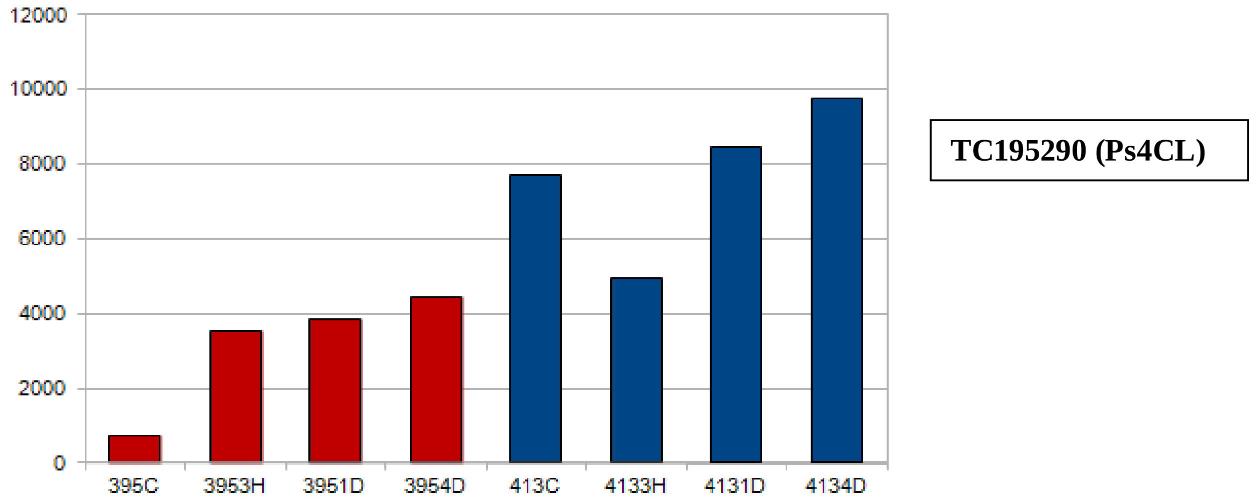


Figure 7: Expression profile of *TC195290* (a top Blast hit of *Ps4CL* on the pine EST list) using RNA-Seq count data.

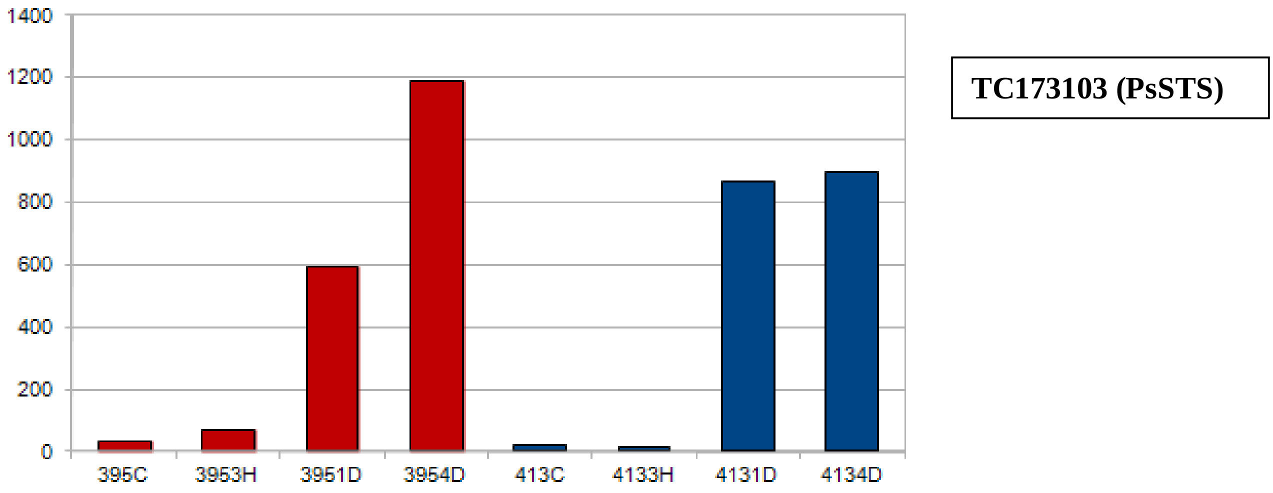


Figure 8: Expression profile of *TC173103* (a top Blast hit of *PsSTS* on the pine EST list) using RNA-Seq count data

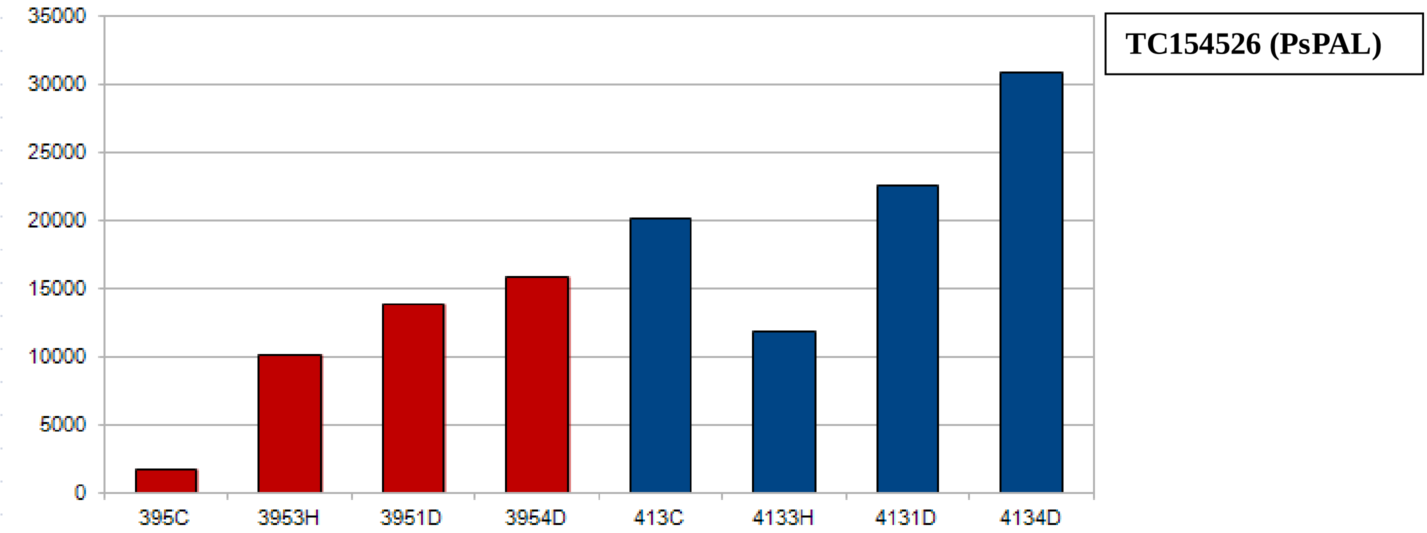


Figure 9: Expression profile of *TC154526* (a top Blast hit of *PsPAL* on the pine EST list) using RNA-Seq count data.

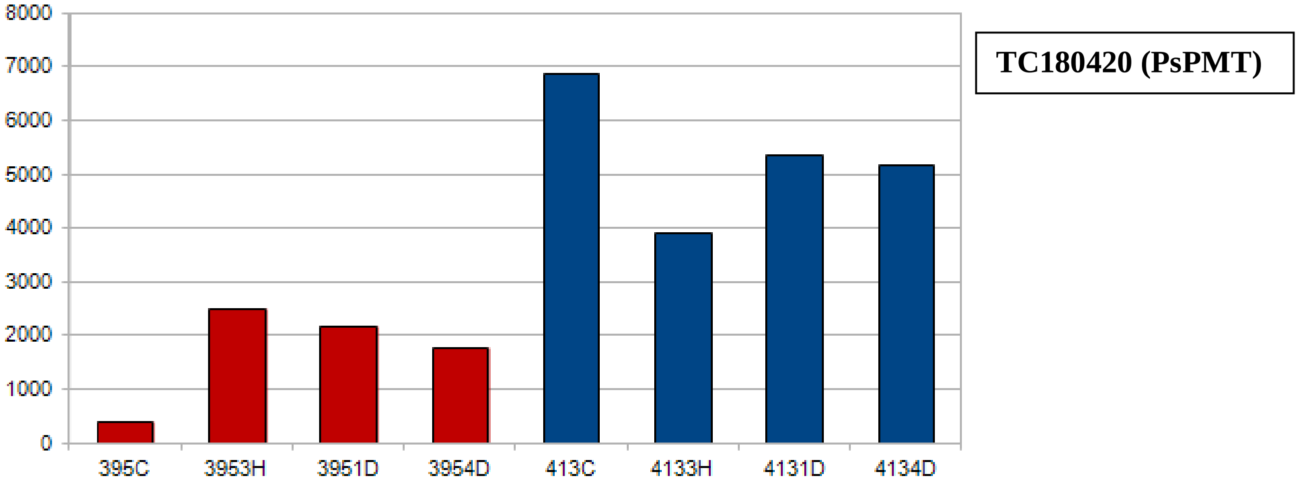


Figure 10: Expression profile of *TC180420* (a top Blast hit of *PsPMT* on the pine EST list) using RNA-Seq count data.

4.7 Differentiating the expression patterns of transcripts between half sib families:

After normalizing for compositional bias in libraries, mapped read counts were used to distinguish the expression of wound-response transcripts (for instance DR092684 - the pinosylvin synthase shown in Figure 11) between half sib family lines of low and high heartwood decay resistance. We observed from Figure 11 that for any desired transcript, we could distinguish between its expression patterns in family lines accurately using mapped read counts (more biological replicates will be needed for this).

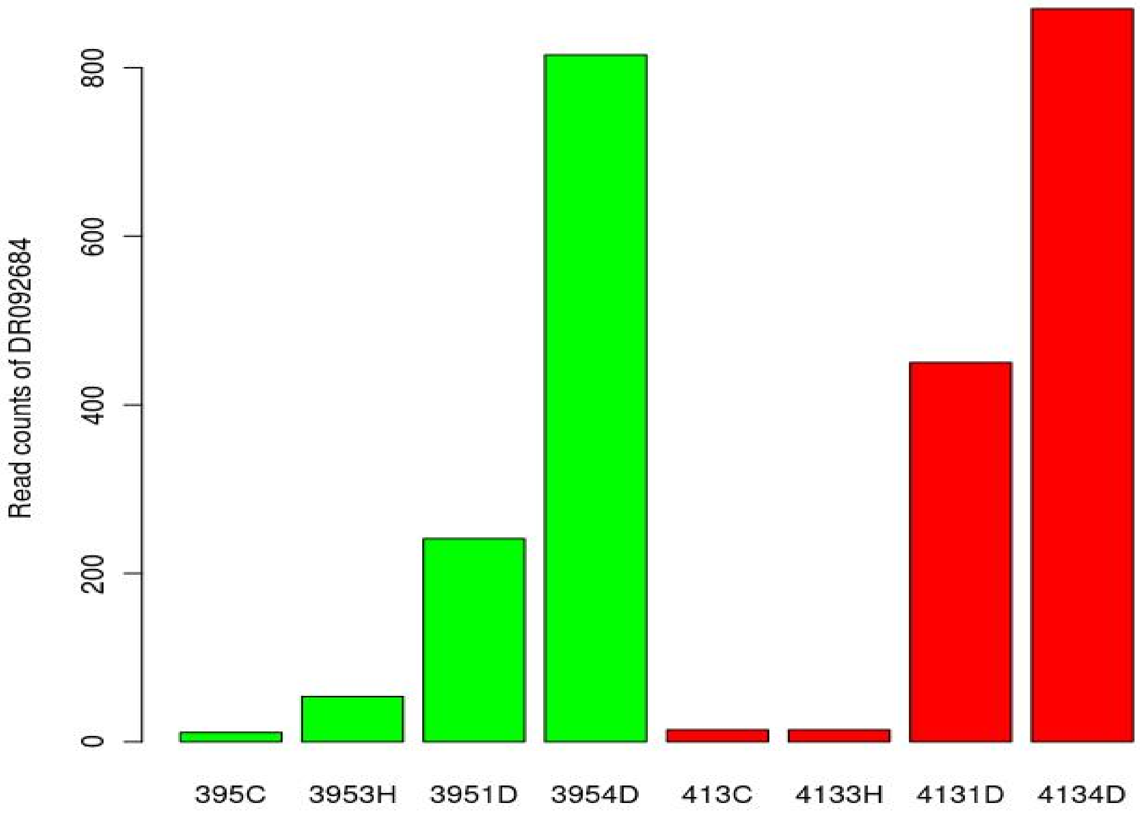


Figure 11: showing differential expression of *DR092684* EST tag between High heartwood decay resistant family line (red) and low heartwood decay resistant family line (green) and libraries.

5. DISCUSSION:

The transcriptome represents a complete collection of expressed RNAs in a cell given any condition. Characterizing the transcriptome is critical in dissecting the functional complexity of any genome and understanding cellular activities from development to disease, growth and lots more. Previously, gene expression analysis has relied on approaches including microarrays, tiling arrays; expressed sequence tags (ESTs), serial analysis of gene expression (SAGE) and other technologies. These approaches have had a number of limitations including the recognition of only specific sequences (especially for arrays), cloning bias by the ESTs and non detection of alternatively spliced variants. Transcriptome profiling and DGE using high throughput sequencing (HTS) approach is a powerful methods for dissecting complex transcriptomes exposed to desired treatments and understanding their expression patterns even with the unavailability of sequenced genomes. This approach has been applied with success to many species including human, plants, fish in different conditions [62, 63, 64, and 65].

5.1 Mapping the reads:

Percentages of mapped reads suggest that our transcriptome is biased to a limited biological process, in this case xylem development and wounding responses. This indicates that HTS transcriptomic profiling can be effectively used to track down the expression patterns of entire transcripts reacting specifically to a given condition and is in line with other studies [62, 64]. Mapping percentages also reflected, as expected, patterns of transcript expression increase and their increment trends from control libraries to 3 hours, 1 and 4 days post wounding libraries. Mapping of the Scots pine RNA-Seq reads revealed about 13,000 tags mapped out of the total 77,326 tags in the Pine EST list. This further confirms the earlier suggestion as in other studies [62, 72] that the RNA-Seq approach can accurately distinguish and extract only the transcripts expressed in an experiment from the total number of transcripts acting in the organs of the organism under survey.

5.2 The assembled transcriptome:

Although our assembled transcriptomes have not been annotated, a preliminary annotation (not shown) suggest that transcripts are mostly related to species *Pinus taeda, Picea sitchensis, Picea glauca, Populus trichocarpa,* and other plant species including *Vitis vinifera, Arabidopsis thaliana, Oryza sativa, Zea mays,* and *Glycine max*. A good number of transcripts from these transcriptomes are factors involved in known signaling pathways, cellular growth and development, plant defense mechanisms, innate immune response and other metabolic pathways. These enriched pathways especially in plant defense and innate immune responses are known to be associated with Scots pine heartwood decay resistance [12, 13, 14, 15] suggesting that valuable transcriptomes were generated and suitable for further studies.

5.3 DGE analysis - STS induction during wounding:

DGE and further analysis of its results suggest that *PsSTS* is strongly induced during the wounding experiment in both high heartwood decay resistant and low resistant families. This strong *PsSTS* induction is confirmed by the RNA-Seq count data and the QPCR expression data. On the other hand, *Ps4CL*, *PsPAL* and *PsPMT* appear to be constitutively expressed as they are not present in the top up-regulated differentially expressed tags. The above suggest that *PsSTS* is strongly induced by wounding but *Ps4CL*, *PsPAL* and *PsPMT* are constitutively expressed in the wounding experiment. This is in line with expected results since *Ps4CL* and *PsPAL* are needed for xylem differentiation but it is unclear why *PsPMT* is constitutively expressed as it is specific for the stilbene biosynthesis pathway [17].

Challenging Scots pine seedlings by wounding resulted in large alterations of its transcriptomic profile including significant ($P<0.01$) up or down-regulation of 1,348 transcripts. Several other biological processes not linked to defense including some transcripts implicated in carbohydrate metabolic processes (*TC161878*), ribosomal proteins, regulators of replication, response to light

(*TC181650*) which were significantly regulated may represent strong stress responsive genes and reflect a number of wide biological activities during Scots pine seedlings response to wounding. The occurrence of these differentially expressed transcripts represents a wider range of biological activities and is in line with other studies [62, 72]. These portray the complexity of the wounded Scots pine transcriptome.

5.4 The effect of RNA-Seq approach in elucidating the molecular mechanisms in Scots pine heartwood decay resistance and the differential gene expression patterns of genotypes:

At present transcriptome analysis involving pine species have relied mainly on the EST approach [32]. Limited information from publicly available sources suggest that very few studies have been carried out to dissect the genetic basis of Scots pine heartwood decay resistance; more so fewer were directly specialized in elucidating the molecular basis of this decay resistance. The RNA-Seq and DGE approach in this study was effective in unraveling the Scots pine seedling transcriptomic complexity when wounded and can detect large sets of transcripts including low expressing and high expressing genes. Top 40 EST tags of both the up-regulated and down regulated classes reveal transcripts which are strongly implicated in stilbene biosynthetic pathway. These transcripts including DR092684, TC171992, TC180934 have functions [32] including regulation of stilbene and flavonoid biosynthetic process, response to ultraviolet light B, response to oxidative stress, naringenin-chalcone synthase activity, regulation of chalcone biosynthetic process, regulation of anthocyanin biosynthetic process, response to wounding, negative regulation of transcription, transcription factor activity and regulation of lignan biosynthetic activity. The biological functions of the above transcripts implicate stilbenes which are known [12, 13, 14] to play vital roles in Scots pine heartwood decay resistance. Mapped reads not only generated read counts utilized for DGE, it also provided a clear basis which could be used for distinguishing between half sib families with records of varied heartwood decay resistant patterns if more biological replicates (family lines) are included in the analysis. As an example (Figure 11), read counts give a picture of the differential expression pattern between (with respect to this transcript *DR092684*) the high heartwood decay resistant family 413 and the low family 395 and provides vital information in digital form that can be applied to breeding programs (given appropriate replicates). Progressive expression patterns can be extracted from interesting transcripts in this way (using differentially expressed transcripts and comparing read counts from divergent pedigrees with sound records) and will provide input for more detailed studies which will be used to generate biomarkers that may be implemented in marker assisted schemes for breeding superior Scot pine with decay resistant heartwoods from an early stage in life.

5.5 Comparing the QPCR method with RNA-Seq and DGE:

Using the QPCR approach, the variable expression patterns of stilbene biosynthetic pathway genes *Ps4CL*, *PsPAL* and *PsPMT* from half sib families 434 and 508 were confirmed by DGE results. DGE also confirmed the expected induction of *PsSTS*. QPCR approach were less sensitive as compared to the RNA-Seq and DGE approach in reflecting gene expression patterns and distinguishing between Scots pine pedigree lines analyzed. This is in line with a study in rice [69] whose QPCR results confirmed some DGE results but was not sensitive enough to confirm the entire DGE results which were shown to be accurate when further analysed. Of great importance to a breeder, DGE approach through read counts tend to be very handy in distinguishing between heartwood decay resistance lines. These suggest that the RNA-Seq count data is more valuable than the QPCR results and could show even very slight changes in an experiment.

5.6 How good was the DGE strategy?

We confirmed the claims made in favor of DGE testing modeled by the negative binomial distribution [55] that the over-dispersion parameter was capable of extracting sample-to-sample variability through our results which showed significant differentially expressed transcripts having functions related to known information (our local unpublished knowledge) regarding our wounding experiment. The higher number of significant ($P<0.01$) up-regulated genes (830) than that of down-regulated genes also suggested the complexity of wounded Scots pine seedling transcriptome.

5.7 Assessing the assembly quality:

Accessing the quality of an assembly is not always a trivial; usually the N50 statistic which is a statistical measure of average length of a set of sequences in a draft assembly is used. A high value of this statistic is regarded as an increase in the quality of an assembly [67]. From the point of view of a biologist, an assembly is much more accurate only if it can recover the entire sequence of a gene or transcript of interest. Table 7 showed that some full length transcripts have been recovered from the transcriptomes and scaffolds recovered some other full length transcripts still fragmented in the transcriptomes. It could be seen from this that the assembly was of good quality being able to extract some full length transcripts although some are still in fragments (typical in any assembly and may be due to low sequencing coverage). Although this assembly was obtained by only a 25-mer tupule during the Trinity run and the scaffold obtained from only one library (5471D), it proved successful in capturing the transcripts expressed during our wounding experiment. This result suggest that if this assembly is extended using a multi k-mer approach (say 19, 21, 27, 31 and 45 -mers) and scaffold sequentially using the entire RNA-Seq data libraries from this experiment, it may be very possible to capture not only full length sequences of all transcripts expressed in this wounding experiment but also those that were expressed at very low levels.

5.8 RNA-Seq challenges:

Although the RNA-Seq/DGE approach seems so appealing and effective, it does not come without a huge number of challenges, especially those linked to computational resources required for bioinformatic analysis and also the bioinformatic expertise needed for such analysis. Our experience in the RNA-

Seq analysis showed that it will be more valuable to consider the use of hybrid sequencing that is using more than one sequencing technology especially the 454 Roche with good sequence length. This will enable the closing up of gaps to extract full transcripts during scaffolding and assembly steps. Due to restrictions in our computational resources, the assembly of transcriptomes was done using only a 25-mer approach but various studies [49, 51] have recommended the use of a wide range of kmer values for assembly before merging in order to capture very lowly expressed transcripts from RNA-Seq reads. It is also valuable to perform subsequent DGE analysis using local transcriptomes generated by *de novo* assembly so as to get a clear picture of the differentially expressed transcripts belonging specifically to the organism or strain under study. Although we tried to remap our 5471D library to the Transcriptome 1 with success (resulting in about 50 percent mapped reads) during the scaffolding step, subsequent DGE testing was skipped due to time restrictions, lack of computational resources and especially for lack of detailed annotation of Transcriptome 1 and 2 which are required for testing.

6. Conclusion:

This study investigated the transcriptome profile of Scots pine seedlings subjected to a wounding experiment using ABI SOLiD RNA-Seq and DGE technologies. The ample amount of transcripts generated provides a good resource for future genomic research on Scots pine stilbene biosynthesis and stress response. Preliminary annotation of assembled transcripts and some Gene Ontology annotated transcripts reveal that the stilbene biosynthesis pathway and control of decay resistance in Scots pine seedlings are complex. Read mappings and counts generated from them also suggest that they are valuable approaches for easy digital discrimination between Scots pine half sib families. This approach can be employed in a specialized breeding program design alongside the use of multiplex sequencing approach for easy discrimination of individual variants. Experience in the assembly and scaffolding of reads suggests that a hybrid sequencing approach will be valuable in transcriptomic profiling especially for organisms without sequenced genomes. The significant results obtained in this analysis also show that our bioinformatic analysis pipeline is well suited for a complex transcriptome such as the Scots pine. We conclude that *PsSTS* is strongly induced by wounding, but *Ps4CL*, *PsPAL* and *PsPMT* tend to be constitutively expressed in the wounding experiment.

DGE values based on a pseudo-reference (Pine ESTs) show that about 1,348 transcripts from about 13,000 tested were differentially expressed during this experiment with about 830 tags up-regulated and about 518 tags down-regulated during this wounding experiment. We were also able to reconstruct a nearly comprehensive local transcriptome from the RNA-Seq dataset using our bioinformatic pipeline although we skipped certain analysis largely due to lack of computational resources. It is evident from our results that a large scale transcriptomics study can be effectively carried out using the ABI SOLiD RNA-Seq and DGE technologies.

7. Acknowledgements:

I want to express my sincere thanks to Professor Teemu Teeri who provided me with supervision and the opportunity to be involved in this exciting project. I am deeply grateful to Dr. Viola Niklander-Teeri for recommending me for this job. My thanks also go to Tanja and Jin, processed the biological samples and generated the reads for this project. It’s their trusting support that ensured the smooth progress of this study. I also owe great thanks to Mr. Festus Anasonye for advice and encouragements. My grateful thanks go to all members of the Gerbera group in the Department of Agricultural Sciences, University of Helsinki.

Above all, I would like to present my special thanks to my dear family, Alfred, Joy, Ifunanya, Amarachi and Nzubechukwu for their love, prayers and support in all I do in life.

Finally, to God who makes all things possible, I give all the glory and thanks.

This work was financially supported through grants by the Department of Agricultural Sciences, University of Helsinki.

8. REFERENCES:

[1] Http://www.metla.fi/ohjelma/jal/index-en.htm. 2011.

[2] Teeri, T. 2011. [Personal communication] June 21. ID: Gerbera laboratory, Department of Agricultural Sciences, University of Helsinki.

[3] Harju, A.M., Venalainen, M., Laakso, T. and Saranpaa, P. 2008. Wounding response in xylem of Scots pine seedlings shows wide genetic variation and connection with constitutive defence of heartwood. Tree Physiology 29: 19-25.

[4] Hart, J.H. 1981. Role of phytostilbenes in decay and disease resistance. Ann. Rev. Phytopathol. 19: 437-58.

[5] Erdtman, H. 1949. Heartwood extractives of conifers. Tappi 32: 305-10.

[6] Barton, G.M. and Smith, R.S. 1971. Brown stain in kiln-dried Abies amabilis lumber. Bi-Mon. Res. Notes, Can. For. Sev. 27: 21-23.

[7] Morgan, J.W.W. and Orsler, R.J. 1968. The chemistry of color changes in wood. I. The significance of stilbenes. Holzforschung 22: 11-16.

[8] Hillis, W.E. 1977. Secondary change in wood. Recent Adv. Phytochem. 11: 247-309.

[9] Jorgensen, E. 1961. The formation of pinosylvin and its monomethyl ether in the sapwood of Pinus resinosa. Can. J. Bot. 39: 1766-72.

[10] Lyr, H. 1961. Hemmungsanalytische Untersuchungen an einigen. Ektoenzymen Holzzerstorender Pilze. Enzymologia 23:231-48.

[11] Suga, T., Ohta, S., Munesada, K., Ide, N., Kurokawa, M., Shimizu, M. and Ohta, E. (1993). Endogenous pine wood nematicidal substances in pines, *Pinus massoniana*, *P. strobus* and *P. palustris*. Phytochemistry 33: 1395-1401.

[12] Hart, J., H. and Hillis, W., E. 1974. Inhibition of wood-rotting fungi by stilbenes and other polyphenols in Eucalyptus sideroxylon. Phytopath. 64: 939-48.

[13] Hart, J.H. and Shrimpton, D.M., 1979. Role of stilbenes in resistance of wood to decay. Phytopath. 69: 1138-43.

[14] Kindl, H. 1984. Biosynthesis and biodegradation of wood components. T. Higuchi. Ed. Academic Press. New York. Pp. 349-377.

[15] Schoppner, A. and Kindl, H. 1979. Stilbene synthase (pinosylvin

forming) and its induction by ultraviolet light. FEBS Lett. 108: 349-352.

[16] Schwekendiek, A., Pfeffer, G. and Kindl, H. 1992. Pine stilbene synthase cDNA, a tool for probing environmental stress. FEBS 301, (1): 41-44.

[17] Atsushi, K., Hiroyuki, K. and Fukumi, S. 2002. A stilbene synthase from Japanese re pine (Pinus densiflora): Implications for phytoalexin accumulation and down-regulation of flavonoid biosynthesis. PNAS. 99, no.5: 3335-3339.

[18] Grisebach, H. 1985 in Biosynthesis and biodegradation of wood components, ed. Higuchi, T. (Academic,Orlando, FL), pp. 291-324.

[19] Kindl, H. 1985 in Biosynthesis and biodegradation of wood components, ed. Higuchi, T. (Academic,Orlando, FL), pp. 349-377.

[20] Fliegmann, J., Schroder, G., Schanz, S., Britsch, L. and Schroder, G. 1992. Molecular analysis of chalcone and dihydropinosylvin synthase from Scots pine (*Pinus sylvestris*) and differential regulation of these and related enzyme activities in stressed plants. Plant Mol. Biol. 18: 489-503.

[21] Raiber, S., Schroder, G. and Schroder, J. 1995. Molecular and enzymatic characterization of two stilbene synthases from Eastern white pine (*Pinus strobus*). A single Arg/His difference determine the activity and the pH dependence of the enzymes. FEBS Lett. 361: 299-302.

[22] Jez, J.M., Austin, M.B., Ferrer, J., Bowman, M.E., Schroder, J. and Noel, J.P. 2000. Structural control of polyketide formation in plant-specific polyketide synthases. Chem. Biol. 7: 919-930.

[23] Suh, D.Y., Fukuma, K., Kagami, J., Yamazaki, Y., Shibuya, M., Ebizuka, Y. and Sankawa, U. 2000. Identification of amino acid residues important in the cyclization reactions of chalcone and stilbene synthases. Biochem. J. 350: 229-235.

[24] Suh, D.Y., Kagami, J., Fukuma, K. and Sankawa, U. 2000. Evidence for Catalytic Cysteine–Histidine Dyad in Chalcone Synthase Biochem. Biophys. Res. Commun. 275: 725-730.

[25] Bonello, P., Heller, W. and Sandermann, H., Jr. 1993. Ozone effects on root-disease susceptibility and defence responses in mycorrhizal and non-mycorrhizal Scots pine (Pinus sylvestris L.) seedlings. New Phytol. 124: 653-663.

[26] Franceschi, V.R., Krokene, P., Christainsen, E. and Krekling, T. 2005. Anatomical and chemical defences of conifer bark against bark beetles and other pests. New Phytol. 167: 353-376.

[27] Bois, E. and Lieutier, F. 1997. Phenolic response of Scots pine clones to inoculation with Leptographium wingfieldii, a fungus associated with Tomicus piniperda. Plant Physiol. Biochem. 35: 819-825.

[28] Nilsson, M., Wikman, S. and Eklund, L. 2002. Induction of discoloured wood in Scots pine (Pinus sylvestris). Tree Physiol. 22: 331-338.

[29] Lyr, H. 1967. Uber den jahrezeitlichen Verlauf der Schutzkernbildung bei Pinus sylvestris nach Verwundungen. Archiv fur Forstwesen 16:51-57.

[30] Harju, A.M., Venalainen, M., Laakso, T. and Saranpaa, P. 2008. Wounding response in xylem of Scots pine seedlings shows wide genetic variation and connection with constitutive defence of heartwood. Tree Physiology 29: 19-25.

[31] Fries, A., Ericsson, T. and Gref, R. 2000. High heritability of wood extractives in Pinus sylvestris progeny tests. Can. J. For. Res. 30: 1707-1713.

[32] [Http: //compbio.dfci.harvard.edu/cgi-bin/tgi/gimain.pl?gudb=pine] May, 2011. (PGI.GO.032811).

[33] Wang, Z., Gerstein, M., Synder, M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. 10(1):57-63.

[34] Li, B., Ruotti, V., Stewart, R.M., Thomson, J.A. and Dewey, C.N. 2010. RNA-Seq gene expression estimation with read mapping uncertainty. Bioinformatics 26: 493-500.

[35] Nagalakshmi, U., Wang, Z., Shou, C., Raha, D., Gerstein, M. and Snyder, M. 2008. The transcriptional landscape of the yeast genome defined by RNA sequencing. Science 320(5881): 1344-9. [PMID: 18451266].

[36] Zhang, G., Guo, G., Hu, X., Zhang, Y., Li, Q., Li, R., Zhuang, R., Lu, Z., He, Z., Fang, X., Chen, L., Tian, W., Tao, Y., Kristainsen, K., Zhang, X., Li, S., Yang, H., Wang, J. and Wang, J. 2010. Genome Res. 20:646-654.

[37] Garber, M., Grabherr, G.M., Guttman, M. and Trapnell, C. 2011. Computational methods for transcriptome annotation and quantification using RNA-Seq. Nat. Meth. 8:469-477.

[38] Zerbino, D.R and Birney, E. 2008. Velvet: algorithms for *de novo* short

read assembly using de Bruijn graphs. Genome Res.18:821-829.

[39] Simpson J.T., Wong, K., Jackman, D.S., Schein, J.E., Jones, S.J.M. and Birol, I. 2009. ABySS: a parallel assembler for short read sequence data. Genome Res. 19, 1117–1123.

[40] Butler, J., MacCallum, I., Kleber, M., Shlyakhter, I.A., Belmonte, M.K., Lander, E.S., Nusbaum, C. and Jaffe, D.B. 2008. ALLPATHS. *De novo* assembly of whole-genome shotgun microreads. Genome Res 18: 810–820.

[41] Falgueras, J., Lara, A.J., Fernández-Pozo, N., Cantón, F.R., Pérez-Trabado, G. and Gonzalo Claros M. 2010. SeqTrim: a high-throughput pipeline for pre-processing any type of sequence read. BMC Bioinformatics 11, 38.

[42] Trapnell, C., Pachter, L. and Salzberg, S. L. 2009. TopHat: discovering splice junctions with RNA-seq. Bioinformatics 25, 1105–1111.

[43] Wang, K., Singh, D., Zeng, Z., Coleman, S. J., Huang, Y., Savich, G. L., He, X., Mieczkowski, P., Grimm, S. A., Perou, C. M., MacLeod, J. N., Chiang, D. Y., Prins, J. F. and Liu, J. (2010). MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. Nucleic Acids Res. 38, e178.

[44] Wu, T. D. and Nacu, S. 2010. Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics 26, 873–881.

[45] Kent, W. J. 2002. BLAT—the BLAST-like alignment tool. Genome Res. 12, 656–664.

[46] Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M.J., Salzberg, S.L., Wold, B.J., Pachter, L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. 28:511–515.

[47] Guttman, M., Garber, M., Levin, J.Z., Donaghey, J., Robinson, J., Adiconis, X., Fan, L., Koziol, M.J., Gnirke, A., Nusbaum, C., Rinn, J.L., Lander, E.S. and Regev, A. 2010. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nature Biotechnology 28, 503–510.

[48] Kinsella, M., Harismendy, O., Nakano, M., Frazer, K. A. and Bafna, V. 2011. Sensitive gene fusion detection using ambiguously mapping RNA-seq read pairs. Bioinformatics 27, 1068–1075.

[49] Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit., I., Adiconis, Z., Fan, L. Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, W. Nusbaum, Lindblad-Toh, K., Friedman, N. and Regev, A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology 29, 644-652.

[50] Zerbino, D.R. Oases: *De novo* transcriptome asembler for short reads. [https://www.ebi.ac.uk/~zerbino/oases/].

[51] Robertson, G., Schein, J., Chiu, R., Corbett, R., Field, M., Jackman, S.D., Mungall, K., Lee, S., Okada, H.M., Qian, J.Q., Griffith, M., Raymond, A., Thiessen, N., Cezard, T., Butterfield, S.Y., Newsome, R., Chan, S.K., She, R., Varhol, R., Kamoh, B., Prabhu, A., Tam, A., Zhao, Y., Moore, R.A., Hirst,M., Marra, A.M., Jones, S.J.M., Hoodless, P.A. and Birol, I. 2010. *De novo* assembly and analysis of RNA-Seq data. Nat. Met. 7:909-912.

[52] Pevzner, P. A., Tang, H. and Waterman, M. S. 2001. An Eulerian path approach to DNA fragment assembly. Proc. Natl Acad. Sci. USA 98, 9748–9753.

[53] Matei, D., Dzamba, M., Lister, D., Ilie, L. and Brudno, M. 2010. SHRiMP2: Sensitive yet practical short read mapping. Bioinf. 00(00):1-3

[54] Software: Sam2count.pl 2011. Https://github.com/vsbuffalo/sam2counts.

[55] Robinson, M.D., McCarthy, D.J. and Smyth, G.K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinf. 26(1):139-140.

[56] Software: AMOS package. 2011. Http://sourceforge.net/project/showfiles.php?group_id=134326.

[57] Sommer, D.D., Delcher, A.L., Salzsberg, S.L. and Pop, M. 2007. Minimus: a fast, light weight genome asembler. BMC Bioinf. 8:64.

[58] Dayarian, A., Michael, T.P. and Sengupta, A.M. 2010. SOPRA: Scaffolding algorithm for paired reads via statistical optimization. BMC Bioinf. 11:345.

[59] Database:GenBank. 2011. [http://www.ncbi.nlm.nih.gov/genbank].

[60] Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R. and

Leunissen, J.A. 2007. Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res. 35(W7):1-4. [PMID: 17485472]

[61] Benjamini, Y. and Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Statist. Soc. Ser. B 57 289-300.

[62] Hegedus, Z., Zakrzewska, A., Agoston, V., C., Ordas, A., Racz, P., Mink, M., Spaink, H., P. and Meijer A., H. 2009. Deep sequencing of the zebrafish transcriptome response to mycobacterium infection. Mol Immunol, 46(15)**:**2918-2930.

[63] Wang B, Guo G, Wang C, Lin Y, Wang X, Zhao M, Guo Y, He M, Zhang Y, Pan L: Survey of the transcriptome of Aspergillus oryzae via massively parallel mRNA sequencing. Nucleic Acids Res 2010, in press.

[64] Morrissy, A., S., Morin, R.,D., Delaney, A., Zeng, T., McDonald, H., Jones, S., Zhao, Y., Hirst, M. and Marra M., A. 2009. Next-generation tag sequencing for cancer gene expression profiling. Genome Res 19(10):1825-1835.

[65] Zhang G, Guo G, Hu X, Zhang Y, Li Q, Li R, Zhuang R, Lu Z, He Z, Fang X, Chen L, Tian W, Tao Y, Kristiansen K, Zhang X, Li S, Yang H, Wang J and Wang J. 2010. Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome. Genome Res 20:646-654.

[66] Altschul S., F., Gish W., Miller W., Myers E., W. and Lipman D., J. 1990. Basic local alignment search tool. Journal of Molecular Biology 215 (3): 403–410. [PMID 2231712].

[67] Software: Web Forum. [http://seqanswers.com/forums/showthread.php?t=2332].

[68] http://compbio.dfci.harvard.edu/tgi/gi/pgi/GenInfo.html. 2011.

[69] Mizuno, H., Kawahara, Y., Sakai, H., Kanamori, H., Wakimoto, H., Yamagata, H., Oono, Y., Wu, J.,Ikawa, H., Itoh, T., and Matsumoto, T. 2010. Massive parallel sequencing of mRNA in identification of unannotated salinity stress-inducible transcripts in rice (*Oryza sativa* L.). BMC Genomics 11:683-696

[70] Geisler-Lee, J., Geisler, M., Coutinho, P., M., Segerman, B., Nishikubo,

N., Takahashi, J., Aspeborg, H., Djerbi, S., Master, E., Andersson-Gunneras, S., Sundberg, B., Karpinski, S., Teeri, T., T., Kleczkowski, L., A., Henrissat, B. and Mellerowicz, E., J. 2006. Poplar carbohydrate-active enzymes. Gene identification and expression analyses. Plant Physiol. 140 (3): 946-962.

[71] Pavy, N., Laroche, J., Bousquet, J. and Mackay, J. 2005. Large-scale statistical analysis of secondary xylem ESTs in pine. Plant Mol. Biol. 57 (2): 203-224.

[72] Cloonan, N., Forrest, A., R., R., Kolle, G., Gardiner, B., B., A., Faulkner, G., J., Brown, M., K., Taylor, D., F., Steptoe, A., L., Wani, S., Bethel, G., Robertson, A., J., Perkins, A., C., Bruce, J., S., Lee, C., C., Ranade, S., S., Peckham, E., H., Manning, M., J., McKernan, K., J., and Grimmond, S., M. 2008. Stem cell transcriptome profiling via massive-scale mRNA sequencing. Nat. Methods 5 (7): 613-619.

[73] Wilhelm, B., T. and Landry, J., R. 2009. RNA-Seq-quantitative measurement of expression through massively parallel RNA-sequencing. Methods 48 (3): 249-257.

[75] Chang, S., Puryear, J. and Cairney, J. 1993. A Simple and Efficient Method for Isolating RNA from Pine Trees. Plant Mol. Biol. Reporter 11: 113-116.

APPENDIX 1:

List of up-regulated EST tags in the Scots pine RNA-Seq wounding experiment

TAGS	LogConc.	LogFC	LR	P.Value	FDR	ANNOTATIONS
DR096636	-12.25	6.121	65.29	6.46E-16	4.16E-12	Cluster: Membrane protein; Magnetospirillum spp
TC171087	-14.03	6.302	64.41	1.01E-15	4.34E-12	
TC171089	-12.27	5.812	61.6	4.21E-15	1.35E-11	Os01g0208600 protein; Oryza sativa
TC171992	-13.32	5.317	58.83	1.72E-14	4.43E-11	Dirigent protein pDIR13; Picea engelmannii x
TC181650	-13.75	4.732	54.37	1.66E-13	3.55E-10	Expansin; Pinus taeda|Rep: Expansin - Pinus
TC192333	-10.57	0.709	51.91	5.80E-13	9.51E-10	S-adenosylmethionine synthetase; Vitis vinifera
TC186817	-13.61	4.6	51.88	5.91E-13	9.51E-10	Cluster: Glycine oxidase ThiO precursor; Xanthobacter
TC194485	-13.87	5.302	50.76	1.04E-12	1.49E-09	Cluster: Pectin methylesterase 4; Pyrus communis|Rep:
TC175575	-13.32	4.665	50.15	1.43E-12	1.71E-09	Pollen allergen CJP38; Cryptomeria japonica|Rep: Pollen
TC161878	-12.7	4.66	49.88	1.63E-12	1.75E-09	Pollen allergen CJP38; Cryptomeria japonica|Rep: Pollen
TC160259	-13.48	4.484	49.69	1.80E-12	1.78E-09	Cluster: Germin-like protein; Pinus taeda
TC196272	-13.94	4.675	48.65	3.05E-12	2.81E-09	Pleiotropic drug resistance protein 8; Arabidopsis
TC157262	-13.24	4.703	48.11	4.02E-12	3.23E-09	Predicted protein; Physcomitrella patens subsp.
TC165984	-13.79	4.326	48	4.25E-12	3.23E-09	Cluster: Integral membrane family protein-like; Oryza
GW772272	-13.73	4.336	48	4.27E-12	3.23E-09	Cluster: Germin-like protein; Pinus caribaea
TC159711	-13.04	4.664	47.3	6.08E-12	4.35E-09	Cluster: Predicted protein; Physcomitrella patens subsp.
DR182118	-13.33	4.5	47.14	6.60E-12	4.47E-09	Cluster: Glucanase-like protein; Thuja occidentalis
TC177871	-13.39	4.192	46.67	8.39E-12	5.39E-09	Os07g0153000 protein; Oryza sativa
TC163357	-11.57	4.122	46.54	8.97E-12	5.49E-09	Pathogenesis-related protein PsemI;Pseudotsuga menziesii
DR092684	-13.83	4.202	46.36	9.83E-12	5.75E-09	Pinus densiflora|Rep: Pinosylvin synthase - Pinus
CV032179	-13.86	4.18	46.07	1.14E-11	6.17E-09	Cluster: ACC oxidase; Trifolium repens|Rep: ACC
TC157530	-13.35	4.447	45.98	1.19E-11	6.17E-09	Thaumatin-like protein; Pinus monticola
TC154538	-11.35	3.962	45.97	1.20E-11	6.17E-09	Pinus sylvestris|Rep: Dihydropinosylvin synthase - Pinus
TC185863	-13.49	4.086	45.54	1.50E-11	7.40E-09	Class IV chitinase Chia4-Pa2; Picea abies|Rep:
TC175362	-14.03	4.925	44.91	2.07E-11	9.84E-09	Signal peptidase I; Chlorobium phaeobacteroides DSM
TC157489	-13.17	4.359	44.72	2.27E-11	1.04E-08	
TC189901	-13.6	4.085	44.4	2.68E-11	1.15E-08	Keratin, type I cytoskeletal 10; Canis
TC155065	-13.29	4.262	44.15	3.05E-11	1.27E-08	Nam-like protein; Solanum tuberosum
TC158949	-11.51	3.962	43.91	3.44E-11	1.38E-08	Dihydropinosylvin synthase; Pinus sylvestris
TC165811	-13.26	4.13	43.81	3.61E-11	1.41E-08	
TC181059	-13.48	3.868	42.64	6.57E-11	2.38E-08	
TC162570	-13.62	3.787	42.48	7.13E-11	2.48E-08	Tau class glutathione S-transferase; Pinus tabuliformis
TC163271	-13.63	3.815	42.32	7.77E-11	2.63E-08	Pinosylvin-forming stilbene synthase; Pinus massoniana
TC170735	-11.44	3.749	42.21	8.20E-11	2.71E-08	Os02g0566800 protein; Oryza sativa Japonica Group|Rep:
TC182094	-13.33	3.916	42.12	8.60E-11	2.77E-08	ASR protein; Ginkgo biloba|Rep: ASR protein
TC166781	-13.7	3.757	41.95	9.37E-11	2.88E-08	
TC172073	-13.73	3.859	41.94	9.39E-11	2.88E-08	Cluster: Basic endochitinase A precursor; Secale
TC163078	-13.75	4.373	41.72	1.06E-10	3.11E-08	Cluster: Chromosome chr15 scaffold_37, genome
TC161997	-12.6	4.135	41.7	1.06E-10	3.11E-08	Thaumatin-like protein; Cryptomeria japonica
TC171916	-12.84	4.249	41.59	1.13E-10	3.22E-08	Cluster: Oxidase-like protein; Leishmania infantum
TC168965	-13.94	4.56	41.42	1.23E-10	3.44E-08	Class IV chitinase Ab; Pinus monticola|Rep:
TC155135	-12.74	4.206	41.33	1.29E-10	3.52E-08	Cluster: LOB domain-containing protein 12; Arabidopsis
TC180455	-13.07	4.263	41.07	1.47E-10	3.93E-08	
TC155408	-13.45	3.81	40.96	1.55E-10	4.04E-08	
TC155893	-13.53	3.812	40.93	1.58E-10	4.04E-08	Cluster: Predicted protein; Physcomitrella patens subsp.
TC172038	-13.27	4.081	40.9	1.60E-10	4.04E-08	Pollen allergen CJP38; Cryptomeria japonica|Rep: Pollen
TC160864	-13.15	4.174	40.84	1.66E-10	4.09E-08	Cluster: PAR-1c protein; Nicotiana tabacum|Rep: PAR-1c
TC165228	-13.32	3.718	40.73	1.74E-10	4.23E-08	Cluster: Chromosome chr2 scaffold_105, genome
TC162898	-13.71	3.751	40.69	1.78E-10	4.25E-08	Photoassimilate-responsive Arabidopsis
TC169372	-13.8	3.964	40.2	2.29E-10	5.27E-08	Serine carboxypeptidase-like 50 precursor; Arabidopsis
TC185897	-11.56	3.971	40.18	2.31E-10	5.27E-08	ATAF1-like protein; Picea mariana protein
DR167706	-13.79	3.835	40.17	2.33E-10	5.27E-08	Pinosylvin synthase; Pinus densiflora
TC174646	-13.56	3.651	39.83	2.77E-10	6.06E-08	Cluster: Keratin 1; Sus scrofa|Rep: Keratin
TC157323	-12.81	4.075	39.82	2.78E-10	6.06E-08	Nam-like protein 10; Petunia x hybrida|Rep:
TC160843	-13.58	3.647	39.76	2.87E-10	6.16E-08	Tau class glutathione S-transferase; Pinus densata|Rep:
TC185164	-12.82	4.179	39.72	2.93E-10	6.17E-08	Potassium transporter; Physcomitrella patens subsp.
TC178842	-12.07	4.235	39.32	3.59E-10	7.46E-08	
BI416822	-13.59	3.679	39.22	3.79E-10	7.62E-08	Cluster: Integral membrane family protein-like; Oryza
TC159430	-13.15	4.111	39.22	3.79E-10	7.62E-08	Signal peptidase I; Chlorobium phaeobacteroides DSM
TC166112	-12.87	3.981	38.99	4.25E-10	8.33E-08	Histidine amino acid transporter; Oryza sativa|Rep:
TC171913	-13	4.099	38.98	4.28E-10	8.33E-08	Cluster: Chromosome chr9 scaffold_7, genome
TC173103	-13.37	3.762	38.76	4.79E-10	9.05E-08	Pinosylvin synthase; Pinus densiflora
TC156042	-13.29	4.034	38.51	5.44E-10	1.01E-07	1-aminocyclopropane-1-carboxylic acid oxidase;
TC196419	-11.64	3.795	38.28	6.14E-10	1.10E-07	Pinus densiflora|Rep: Pinosylvin synthase - Pinus
TC194996	-12.54	3.867	38.27	6.17E-10	1.10E-07	Aquaporin; Ricinus communis|Rep: Aquaporin - Ricinus
TC177372	-14.02	4.268	38.27	6.17E-10	1.10E-07	Cluster: Chromosome chr11 scaffold_13, genome
TC164934	-12.75	3.938	38.17	6.47E-10	1.14E-07	Thaumatin-like protein; Cryptomeria japonica
TC195493	-13.32	3.764	38.09	6.75E-10	1.16E-07	Pinosylvin synthase; Pinus densiflora
GW751135	-12.97	4.013	38.09	6.76E-10	1.16E-07	Pathogenesis-related protein PsemI;Pseudotsuga menziesii
GT231623	-12.52	3.932	38	7.08E-10	1.17E-07	Dirigent-like protein; Picea sitchensis
TC175346	-13.32	3.769	37.75	8.06E-10	1.30E-07	Cluster: UPF0496 protein 4; Oryza sativa
TC182293	-12.21	4.155	37.32	1.00E-09	1.57E-07	Dirigent-like protein; Picea sitchensis
TC174496	-13.3	3.728	37.26	1.03E-09	1.60E-07	Dirigent-like protein; Picea sitchensis
TC163901	-14.11	4.705	36.97	1.20E-09	1.84E-07	Peroxidase; Nicotiana tabacum

TC185039	-13.24	3.753	36.71	1.37E-09	2.08E-07	Cluster: Predicted protein; Physcomitrella patens subsp.
TC156839	-12.49	3.853	36.67	1.40E-09	2.10E-07	Pinus densiflora|Rep: Pinosylvin synthase
TC156757	-13.05	3.889	36.57	1.47E-09	2.18E-07	Pinus massoniana: Pinosylvin-forming stilbene synthase
TC158533	-13.3	3.721	36.29	1.70E-09	2.48E-07	Tau class glutathione S-transferase; Pinus densata|Rep:
TC191109	-13.99	3.86	36.2	1.79E-09	2.58E-07	
TC186301	-11.92	3.951	36.06	1.92E-09	2.74E-07	Pathogenesis-related protein PsemI;Pseudotsuga menziesii
TC176472	-13.47	3.549	35.92	2.05E-09	2.90E-07	Late embryogenesis abundant protein; Picea glauca|Rep:
GT232388	-13.57	3.415	35.87	2.11E-09	2.92E-07	Dirigent protein pDIR6; Picea engelmannii x
DR179276	-13.25	3.71	35.85	2.13E-09	2.92E-07	Pinosylvin synthase; Pinus densiflora
TC155821	-12.13	4.079	35.85	2.14E-09	2.92E-07	Dirigent-like protein; Picea sitchensis
TC192671	-12.76	3.745	35.64	2.38E-09	3.22E-07	Cluster: Extensin; Catharanthus roseus|Rep: Extensin -
TC155765	-13.34	3.432	35.48	2.58E-09	3.45E-07	ATAF1-like protein; Picea mariana
TC168517	-12.85	3.749	35.46	2.60E-09	3.45E-07	E1-E2 type:Magnesium-translocating P-type ATPase;
GW759404	-13.66	3.476	35.38	2.72E-09	3.56E-07	Predicted protein; Physcomitrella patens subsp.
TC175310	-13.49	3.687	35.32	2.80E-09	3.64E-07	
TC166498	-13.72	3.378	35.28	2.86E-09	3.67E-07	1-amino-cyclopropane-1-carboxylic acid oxidase 3:Manihot
TC178323	-13.26	3.765	35.26	2.88E-09	3.67E-07	Chromosome chr9 scaffold_7, genome
TC193559	-14.06	4.316	35.06	3.21E-09	4.00E-07	Peroxidase precursor; Picea abies
TC178656	-12.88	3.8	35.04	3.24E-09	4.00E-07	N utilization substance protein A;
TC161719	-11.95	3.9	34.99	3.32E-09	4.06E-07	Pathogenesis-related protein PsemI; Pseudotsuga menziesii
TC158614	-12.21	3.929	34.75	3.74E-09	4.50E-07	Dirigent-like protein; Picea sitchensis
TC172002	-12.14	3.877	34.67	3.90E-09	4.65E-07	
TC183848	-12.03	3.94	34.43	4.41E-09	5.20E-07	Dirigent-like protein; Picea sitchensis
TC172315	-12.16	3.974	34.24	4.88E-09	5.71E-07	Pathogenesis-related protein PsemI; Pseudotsuga menziesii
TC157432	-14.05	4.348	33.71	6.40E-09	7.35E-07	Expansin; Pinus taeda|Rep: Expansin - Pinus
TC172901	-13.97	3.953	33.67	6.53E-09	7.44E-07	Uncharacterized protein At3g14850.2; Arabidopsis
TC163616	-12.14	3.92	33.57	6.88E-09	7.76E-07	Pathogenesis-related protein PsemI; Pseudotsuga menziesii
TC171264	-13.62	3.2	33.37	7.63E-09	8.53E-07	Probable WRKY transcription factor 31; Arabidopsis
TC156363	-13.14	3.576	32.98	9.33E-09	1.03E-06	Cluster: Chromosome undetermined scaffold_230, genome
TC167708	-13.59	3.317	32.97	9.38E-09	1.03E-06	ATAF1-like protein; Picea mariana
TC168264	-13.5	3.247	32.86	9.89E-09	1.08E-06	Transglutaminase domain protein precursor;Pyrobaculum
TC157317	-13.5	3.199	32.84	9.99E-09	1.08E-06	Ethylene-responsive element-binding protein;Medicago spp
TC159060	-13.82	3.288	32.8	1.02E-08	1.09E-06	Cluster: Pinoresinol-lariciresinol reductase;Thuja pli.
TC179010	-13.64	3.347	32.75	1.05E-08	1.11E-06	Clt1; Poncirus trifoliata|Rep: Clt1 - Poncirus
CO158649	-13.59	3.215	32.67	1.09E-08	1.13E-06	Cluster: Chromosome chr10 scaffold_43, genome
TC184015	-13.83	3.399	32.49	1.20E-08	1.21E-06	Defensin; Pinus sylvestris|Rep: Defensin - Pinus
TC159421	-12.19	3.824	32.49	1.20E-08	1.21E-06	Glucose-methanol-choline oxidoreductase family protein;
TC182672	-12.76	3.541	32.1	1.47E-08	1.44E-06	
TC154758	-12.91	3.544	31.98	1.56E-08	1.52E-06	Cluster: Chromosome chr9 scaffold_7, genome
TC168370	-12.62	3.398	31.9	1.63E-08	1.57E-06	Pinosylvin synthase; Pinus densiflora
NP542958	-13.9	3.569	31.8	1.71E-08	1.64E-06	strobus
TC160061	-14.08	3.721	31.7	1.80E-08	1.71E-06	
TC154861	-13.15	3.461	31.69	1.81E-08	1.71E-06	Cluster: Predicted protein; Physcomitrella patens subsp.
TC198350	-13.09	3.506	31.67	1.82E-08	1.71E-06	Cluster: Chromosome chr7 scaffold_31, genome
TC157448	-12.15	3.679	31.36	2.14E-08	1.99E-06	Pathogenesis-related protein PsemI; Pseudotsuga
TC185670	-14.01	3.728	31.35	2.15E-08	1.99E-06	(Hypocretin) (Hcrt) [Contains: Orexin-A (Hypocretin- 1)
TC182033	-11.96	3.575	31.22	2.30E-08	2.11E-06	Cluster: Histone H2B.11; Oryza sativa|Rep: Histone
TC180004	-12.62	3.528	31.18	2.35E-08	2.15E-06	
TC169448	-12.81	3.571	31.15	2.38E-08	2.16E-06	Chromosome chr7 scaffold_20, genome
TC157135	-13.15	3.5	31.11	2.44E-08	2.19E-06	Modifier of mdg4; Drosophila melanogaster|Rep: Modifier
TC158438	-13.18	3.121	31.01	2.56E-08	2.29E-06	Late embryogenesis abundant protein; Picea glauca|Rep:
DR025377	-13.28	3.347	30.99	2.59E-08	2.30E-06	Pinus densiflora|Rep: Pinosylvin synthase - Pinus
DR097097	-13.86	3.234	30.74	2.94E-08	2.59E-06	Chromosome chr6 scaffold_28, genome
AI725288	-13.21	3.435	30.42	3.47E-08	3.04E-06	Dirigent-like protein; Picea sitchensis
TC160074	-12.04	3.636	30.4	3.52E-08	3.05E-06	Dihydropinosylvin synthase; Pinus sylvestris
TC155386	-12	3.548	30.39	3.54E-08	3.05E-06	Chromosome chr7 scaffold_20, genome
FG616308	-13.01	3.424	30.35	3.60E-08	3.09E-06	Chitinase; Zea diploperennis|Rep: Chitinase - Zea
TC159802	-12.62	3.382	30.28	3.74E-08	3.18E-06	Cluster: Chromosome chr9 scaffold_7, genome
TC196582	-13.61	3.152	30.26	3.77E-08	3.19E-06	PT1; Solanum melongena|Rep: PT1 - Solanum
NP542706	-12.41	3.348	30.19	3.91E-08	3.29E-06	
TC164481	-13.81	3.193	30.17	3.96E-08	3.29E-06	Glucanase-like protein; Thuja occidentalis|Rep:
TC158852	-13.18	3.425	30.04	4.24E-08	3.49E-06	T3H13.3 protein; Arabidopsis |Rep: T3H13.3 protein
TC158385	-13.17	3.351	29.9	4.56E-08	3.72E-06	Chromosome undetermined scaffold_377, genome
TC188417	-12.51	3.274	29.89	4.57E-08	3.72E-06	Cluster: Chromosome undetermined scaffold_469, genome
TC156188	-13.06	3.391	29.63	5.23E-08	4.23E-06	Phosphate transporter; Sesbania rostrata
TC159287	-12.13	3.535	29.61	5.28E-08	4.25E-06	Dirigent-like protein; Picea sitchensis
TC187527	-13.91	3.598	29.53	5.51E-08	4.40E-06	Expansin; Pinus taeda|Rep: Expansin - Pinus
TC156784	-13.84	3.072	29.49	5.63E-08	4.47E-06	Clavata-like receptor; Picea glauca
DR019650	-12.19	3.51	29.37	5.98E-08	4.72E-06	ACC oxidase; Picea|Rep: ACC oxidase -
TC181894	-13.7	3.116	29.32	6.15E-08	4.82E-06	Isoform mod1.8 of Q86B87 ; Drosophila
TC154749	-12.32	3.54	29.13	6.78E-08	5.29E-06	
TC162364	-13.39	2.89	28.95	7.41E-08	5.74E-06	Chromosome chr2 scaffold_105, genome
TC163125	-11.15	3.571	28.92	7.53E-08	5.80E-06	Cluster: Chromosome chr2 scaffold_112, genome
AM983083	-14.16	3.642	28.91	7.58E-08	5.80E-06	
TC186750	-12.97	3.309	28.71	8.42E-08	6.32E-06	Cluster: Chromosome chr2 scaffold_105, genome
NP542959	-14.03	3.577	28.67	8.60E-08	6.39E-06	strobus]
TC176092	-12.6	3.344	28.65	8.66E-08	6.39E-06	Pyruvate dehydrogenase complex, E2 component,
TC172794	-12.53	3.241	28.62	8.79E-08	6.39E-06	
NP542707	-13.22	3.177	28.62	8.79E-08	6.39E-06	
TC174699	-11.31	2.979	28.43	9.74E-08	7.04E-06	Chromosome chr6 scaffold_3, genome
TC159830	-13.39	3.166	28.33	1.02E-07	7.34E-06	Cluster: Chromosome chr1 scaffold_5, genome

AA556386	-12.97	3.428	28.32	1.03E-07	7.34E-06	
NP542874	-13.86	3.304	28.27	1.05E-07	7.49E-06	
TC162129	-12.59	3.169	28.05	1.18E-07	8.31E-06	Chromosome undetermined scaffold_306, genome
TC194671	-12.31	3.339	28.04	1.19E-07	8.32E-06	Cluster: Isoform 2 of Q9ZQG9 ;
TC158616	-11.87	3.293	27.9	1.28E-07	8.76E-06	Histidine amino acid transporter; Oryza sativa|Rep:
TC174998	-12.99	3.32	27.89	1.28E-07	8.76E-06	Cluster: At1g22990/F19G10_22; Arabidopsis
TC164970	-12.6	3.153	27.89	1.29E-07	8.76E-06	Chromosome undetermined scaffold_306, genome
TC168009	-13.45	3.05	27.85	1.31E-07	8.89E-06	Cluster: Pinoresinol-lariciresinol reductase; Thuja
TC161177	-12.53	3.163	27.63	1.47E-07	9.89E-06	Pollen allergen CJP38; Cryptomeria japonica|Rep: Pollen
TC196298	-12.59	3.286	27.56	1.52E-07	1.02E-05	
TC157731	-12.08	3.426	27.51	1.56E-07	1.04E-05	Pinosylvin-forming stilbene synthase; Pinus massoniana|
TC187745	-13.94	3.118	27.47	1.60E-07	1.06E-05	Cluster: 21 kD protein; Pinus radiata|Rep:
TC160527	-12.99	3.341	27.46	1.61E-07	1.06E-05	T3H13.3 protein; Arabidopsis |Rep: T3H13.3 protein
FE519583	-13.18	3.279	27.4	1.66E-07	1.08E-05	
TC170862	-13.84	3.164	27.33	1.72E-07	1.11E-05	Cold acclimation protein 1; Picea abies|Rep:
TC172356	-13.55	2.886	27.32	1.72E-07	1.11E-05	Water deficit stress inducible protein LP3-2;
TC161416	-13.91	3.11	27.26	1.78E-07	1.13E-05	Cluster: Uncharacterized protein At3g07270.2;
TC164429	-13.65	2.845	27.19	1.84E-07	1.16E-05	Chromosome chr5 scaffold_2, genome
TC179063	-13.2	2.871	27.07	1.96E-07	1.22E-05	Hydrophobic protein LTI6B; Oryza sativa|Rep: Hydrophobic
TC169078	-12.09	3.325	26.98	2.06E-07	1.27E-05	
TC155253	-14.34	3.895	26.94	2.09E-07	1.29E-05	Cluster: Chromosome chr5 scaffold_67, genome
GT242011	-12.8	3.225	26.79	2.27E-07	1.39E-05	HAK5; Solanum lycopersicum|Rep: HAK5 - Solanum
TC184166	-13.7	3.068	26.72	2.36E-07	1.44E-05	Chromosome chr11 scaffold_56, genome
TC188579	-14.24	3.674	26.63	2.46E-07	1.50E-05	Clavata-like receptor; Picea glauca
CR394078	-13.82	2.933	26.58	2.53E-07	1.53E-05	Cluster: Pollen allergen CJP38; Cryptomeria japonica|
TC190819	-14.16	3.812	26.57	2.54E-07	1.53E-05	Uncharacterized protein At3g14850.2; Arabidopsis
BQ634569	-12.75	3.198	26.5	2.64E-07	1.58E-05	
TC156148	-13.23	3.053	26.36	2.83E-07	1.69E-05	Phosphate transporter; Oryza sativa
TC184648	-13.75	3.042	26.26	2.98E-07	1.77E-05	Chromosome chr3 scaffold_199, genome
TC174383	-12	3.235	25.93	3.54E-07	2.09E-05	Cluster: Chromosome chr2 scaffold_112, genome
TC194898	-13.65	2.71	25.91	3.58E-07	2.11E-05	
NP542704	-13.15	2.963	25.69	4.01E-07	2.35E-05	
TC162349	-13.55	2.843	25.68	4.03E-07	2.35E-05	Cluster: P-glycoprotein; Arabidopsis |Rep:
TC167402	-14.15	3.379	25.66	4.08E-07	2.36E-05	Cluster: Chromosome chr6 scaffold_3, genome
TC160803	-14.01	3.282	25.39	4.69E-07	2.70E-05	
TC168704	-13.81	2.998	25.35	4.78E-07	2.75E-05	PR10 protein; Pinus monticola|Rep: PR10 protein
TC174356	-12.68	3.051	25.31	4.88E-07	2.79E-05	Cluster: Chromosome chr2 scaffold_97, genome
TC178235	-13.19	2.964	25.24	5.07E-07	2.89E-05	GDP-mannose 3,5-epimerase 1; Oryza sativa|Rep: GDP-
TC187522	-14.31	3.743	25.13	5.36E-07	3.04E-05	Expansin; Pinus taeda|Rep: Expansin - Pinus
TC173283	-13.89	2.92	25.04	5.62E-07	3.17E-05	Cluster: 21 kD protein; Pinus radiata|Rep:
TC163752	-13.89	2.922	25.02	5.66E-07	3.18E-05	Pollen allergen CJP38; Cryptomeria japonica|Rep: Pollen
TC158711	-13.03	2.981	24.86	6.17E-07	3.45E-05	Pinoresinol-lariciresinol reductase; Thuja plicata
TC193263	-13.22	2.738	24.41	7.77E-07	4.30E-05	Chromosome chr2 scaffold_97, genome
TC189279	-13.57	2.653	24.41	7.79E-07	4.30E-05	Formate dehydrogenase 2, mitochondrial precursor; Oryza
TC161130	-12.28	2.991	24.39	7.87E-07	4.33E-05	Pleiotropic drug resistance protein TUR2; Spirodela
TC198866	-12.28	3.198	24.38	7.92E-07	4.34E-05	
TC172502	-13.47	2.645	24.34	8.07E-07	4.39E-05	
TC189237	-13.89	3.171	24.34	8.09E-07	4.39E-05	Defensin; Pinus sylvestris|Rep: Defensin - Pinus
GW747487	-14.28	3.63	24.33	8.12E-07	4.39E-05	Alpha-expansin 2; Gossypium hirsutum
TC197639	-13.15	2.928	24.3	8.23E-07	4.43E-05	Cluster: O-methyltransferase, family 2; Dimerisation;
TC169441	-12.31	3.005	24.2	8.69E-07	4.64E-05	Late embryogenesis abundant protein; Picea glauca|Rep:
TC165453	-13.71	2.687	24.15	8.90E-07	4.71E-05	Cluster: Taurine catabolism dioxygenase TauD, TfdA
TC167068	-12.92	2.963	24.09	9.19E-07	4.85E-05	Chromosome chr2 scaffold_97, genome
TC162576	-13.26	2.764	24.08	9.24E-07	4.85E-05	Thaumatin-like protein; Pinus taeda
TC177086	-12.93	2.921	24.05	9.38E-07	4.91E-05	Hydrophobic protein LTI6B; Oryza sativa|Rep: Hydrophobic
TC166786	-14.17	3.533	24.03	9.47E-07	4.93E-05	
TC172481	-13.78	2.671	24.02	9.53E-07	4.94E-05	Chromosome chr3 scaffold_157, genome
TC165959	-13.16	2.889	23.89	1.02E-06	5.26E-05	Cluster: Gb|AAD32907.1; Arabidopsis |Rep: Gb|AAD32907.1
TC169001	-13.4	2.841	23.79	1.07E-06	5.50E-05	Chromosome chr18 scaffold_1, genome
TC187381	-13.05	2.929	23.79	1.08E-06	5.50E-05	Pinus massoniana Pinosylvin-forming stilbene synthase
TC171779	-12.67	2.868	23.78	1.08E-06	5.50E-05	Cluster: Chromosome undetermined scaffold_469, genome
AA556511	-12.95	3.059	23.75	1.10E-06	5.57E-05	
TC162768	-12.47	2.841	23.73	1.11E-06	5.61E-05	Cluster: O-methyltransferase, family 2; Dimerisation;
TC183794	-13.89	3.16	23.66	1.15E-06	5.74E-05	Cluster: Chromosome chr13 scaffold_48, genome
TC196696	-13.61	2.602	23.66	1.15E-06	5.74E-05	Cluster: Chromosome undetermined scaffold_302, genome
TC162373	-12.8	2.882	23.6	1.19E-06	5.91E-05	Cluster: Uncharacterized protein At5g22140.2;
TC171513	-13.26	2.878	23.55	1.22E-06	6.04E-05	Cluster: Chromosome undetermined scaffold_69, genome
TC174045	-10.98	3.586	23.54	1.22E-06	6.04E-05	ACC oxidase; Picea|Rep: ACC oxidase -
TC183475	-14.41	3.816	23.47	1.27E-06	6.26E-05	
CO165586	-14.18	3.313	23.46	1.28E-06	6.27E-05	Cluster: Flagellin-sensing 2-like protein; Brassica
TC174926	-13.35	2.642	23.38	1.33E-06	6.50E-05	Cluster: Thaumatin-like protein; Pseudotsuga menziesii
DN613379	-13.46	2.608	23.33	1.37E-06	6.63E-05	WRKY6; Nicotiana attenuata|Rep: WRKY6 - Nicotiana
TC164522	-13.46	2.592	23.32	1.37E-06	6.63E-05	Cluster: NtPRp27; Nicotiana tabacum|Rep: NtPRp27 -
TC154691	-13.83	2.915	23.22	1.45E-06	6.94E-05	AP2 domain transcription factor-like; Oryza sativa|Rep:
TC181103	-12.47	2.82	23.17	1.48E-06	7.05E-05	Formate dehydrogenase, mitochondrial precursor;
GT231151	-13.56	2.548	23.14	1.51E-06	7.15E-05	Chromosome undetermined scaffold_55, genome
TC191130	-12.43	3.044	23.02	1.60E-06	7.59E-05	
DR094801	-13.67	2.551	22.96	1.66E-06	7.81E-05	Thaumatin-like protein; Pinus monticola
TC185808	-13.33	2.73	22.94	1.67E-06	7.84E-05	
TC165234	-13.91	2.755	22.87	1.73E-06	8.07E-05	Cluster: Chromosome chr17 scaffold_85, genome
TC194955	-14.18	3.449	22.86	1.74E-06	8.07E-05	Dirigent protein pDIR2; Picea engelmannii x

TC158770	-14.07	3.167	22.86	1.75E-06	8.07E-05	
TC162632	-12.16	2.92	22.71	1.89E-06	8.63E-05	Cluster: Chromosome chr6 scaffold_15, genome
TC187185	-12.7	2.874	22.61	1.98E-06	9.05E-05	
TC194956	-13.06	2.771	22.38	2.24E-06	1.01E-04	NBS/LRR; Pinus taeda|Rep: NBS/LRR - Pinus
TC165436	-14.18	3.477	22.36	2.26E-06	1.02E-04	PR10 protein; Pinus monticola|Rep: PR10 protein
TC180313	-13.44	2.5	22.25	2.39E-06	1.07E-04	Galactinol synthase; Brassica napus|Rep: Galactinol
TC156614	-12.37	2.826	22.23	2.42E-06	1.08E-04	Chromosome chr14 scaffold_54, genome
TC166778	-12.51	2.712	22.17	2.50E-06	1.11E-04	Cluster: O-methyltransferase, family 2; Dimerisation;
FG615126	-13.03	2.821	22.15	2.52E-06	1.12E-04	Formate dehydrogenase 2, mitochondrial precursor; Oryza
TC161468	-13.53	2.46	21.94	2.81E-06	1.23E-04	Protein phosphatase 2C; Mesembryanthemum crystallinum
TC167859	-13.86	2.615	21.94	2.82E-06	1.23E-04	Glucanase-like protein; Thuja occidentalis|Rep
TC161835	-12.37	2.804	21.92	2.84E-06	1.24E-04	Chromosome chr6 scaffold_3, genome
DR023030	-10.54	0.491	21.77	3.07E-06	1.33E-04	Cinnamate-4-hydroxylase; Gossypium arboreum|Rep:
TC156920	-13.27	2.717	21.7	3.18E-06	1.38E-04	Chromosome chr5 scaffold_67, genome
TC164540	-13.94	2.816	21.65	3.27E-06	1.41E-04	Cluster: Chromosome chr1 scaffold_46, genome
TC168057	-13.36	2.54	21.64	3.29E-06	1.41E-04	Cluster: Chromosome undetermined scaffold_100, genome
AL750771	-13.9	2.799	21.42	3.68E-06	1.57E-04	Predicted protein; Physcomitrella patens subsp.
TC159752	-12.38	2.714	21.41	3.71E-06	1.58E-04	Chromosome chr2 scaffold_105, genome
TC172415	-12.18	2.858	21.31	3.91E-06	1.65E-04	Chromosome undetermined scaffold_142, genome
TC155045	-14.18	3.021	21.26	4.02E-06	1.69E-04	Cluster: Extensin; Catharanthus roseus|Rep: Extensin -
TC155062	-13.51	2.444	21.24	4.06E-06	1.69E-04	Predicted protein; Physcomitrella patens subsp.
TC183912	-14.07	3.029	21.22	4.10E-06	1.71E-04	Chromosome chr11 scaffold_13, genome
DR017187	-13.75	2.601	21.18	4.19E-06	1.74E-04	Pleiotropic drug resistance protein 1; Nicotiana
CX648248	-13.5	2.412	21.15	4.26E-06	1.76E-04	Formate dehydrogenase, mitochondrial precursor;
TC161247	-14.21	3.145	21.11	4.34E-06	1.79E-04	Cluster: Glycoside hydrolase, family 18; Medicago
TC180073	-13.8	2.479	21.1	4.37E-06	1.80E-04	Cluster: tetra-peptide repeat homeobox-like (TPRXL)
TC169130	-14.11	2.817	21.07	4.42E-06	1.81E-04	Chromosome chr18 scaffold_24, genome
TC195171	-14.07	2.852	20.93	4.77E-06	1.94E-04	Dirigent-like protein pDIR14; Picea engelmannii x
TC155004	-13.63	2.404	20.86	4.93E-06	2.00E-04	Chromosome undetermined scaffold_143, genome
DR742536	-13.84	2.532	20.81	5.07E-06	2.05E-04	Chromosome chr6 scaffold_3, genome
NP542708	-13.31	2.482	20.74	5.27E-06	2.13E-04	
BI077276	-14.22	3.096	20.67	5.45E-06	2.18E-04	Chromosome chr18 scaffold_1, genome
TC188235	-13.76	2.446	20.65	5.52E-06	2.20E-04	Chromosome chr9 scaffold_7, genome
TC190137	-11.65	2.671	20.53	5.87E-06	2.32E-04	Protein TAR1; Kluyveromyces lactis|Rep: Protein TAR1
TC186863	-13.73	2.484	20.42	6.20E-06	2.45E-04	Alternative oxidase; Nicotiana attenuata|Rep:
TC190823	-13.18	2.546	20.29	6.64E-06	2.59E-04	Cluster: Probable mannitol dehydrogenase; Medicago spp
TC189462	-13.07	2.789	20.22	6.90E-06	2.68E-04	
TC187790	-13.58	2.318	20.09	7.38E-06	2.85E-04	Phosphoenolpyruvate carboxykinase
TC155051	-14.03	2.942	20.06	7.52E-06	2.89E-04	Chromosome chr1 scaffold_166, genome
TC156792	-12.71	2.639	19.98	7.81E-06	3.00E-04	Predicted protein; Physcomitrella patens subsp.
TC180290	-12.56	2.564	19.85	8.36E-06	3.19E-04	Cluster: glycosyl hydrolase family 17 protein;
TC182612	-13.53	2.307	19.79	8.65E-06	3.26E-04	Isoform 3 of Q8R2K1 ; Mus
TC164249	-12.29	2.755	19.78	8.67E-06	3.26E-04	Class IV chitinase Chia4-Pa1; Picea abies|Rep:
TC175906	-13.59	2.382	19.78	8.71E-06	3.27E-04	Galactinol synthase; Brassica napus|Rep:
CO198207	-13.94	2.749	19.76	8.76E-06	3.28E-04	Thaumatin-like protein; Pseudotsuga menziesii|Rep:
BQ696662	-12.56	2.563	19.74	8.86E-06	3.30E-04	Cluster: Chromosome chr9 scaffold_7, genome
TC191010	-13.69	2.499	19.71	9.00E-06	3.34E-04	
TC160710	-13.14	2.506	19.63	9.38E-06	3.47E-04	Glutathione peroxidase; Oryza sativa Indica Group|Rep:
TC187934	-12.78	2.555	19.48	1.02E-05	3.72E-04	Pleiotropic drug resistance protein 1; Nicotiana
DR111377	-13.67	2.34	19.35	1.09E-05	3.94E-04	Chromosome chr10 scaffold_43, genome
TC169717	-13.59	2.371	19.34	1.10E-05	3.96E-04	Water deficit stress inducible protein LP3-2;
TC171926	-13.22	2.414	19.28	1.13E-05	4.07E-04	
TC158960	-13.46	2.313	19.22	1.16E-05	4.17E-04	Galactinol synthase; Brassica napus|Rep:
TC160994	-13.5	2.277	19.17	1.20E-05	4.27E-04	Cluster: Predicted protein; Physcomitrella patens subsp.
TC156019	-12.13	2.633	19.07	1.26E-05	4.48E-04	Formate dehydrogenase, mitochondrial precursor;
TC171034	-13.6	2.316	18.96	1.33E-05	4.70E-04	Zinc finger, RING-type; Medicago truncatula|Rep: Zinc
TC165723	-13.2	2.397	18.96	1.34E-05	4.70E-04	Cluster: Predicted protein; Physcomitrella patens subsp.
TC178104	-13.93	2.559	18.81	1.44E-05	5.03E-04	At3g11320; Arabidopsis |Rep: At3g11320 - Arabidopsis
TC189655	-12.32	2.507	18.72	1.51E-05	5.23E-04	Protein TAR1; Kluyveromyces lactis|Rep: Protein TAR1
TC173120	-12.99	2.519	18.68	1.55E-05	5.33E-04	Cluster: Os06g0554800 protein; Oryza sativa Japonica
TC169394	-14.24	3.048	18.68	1.55E-05	5.33E-04	NADH dehydrogenase (Quinone) precursor; Salinispora
TC158642	-13.19	2.428	18.61	1.60E-05	5.49E-04	Galactinol synthase; Brassica napus|Rep: Galactinol
TC180374	-12.49	2.46	18.57	1.64E-05	5.59E-04	Pleiotropic drug resistance protein 1; Nicotiana
TC197095	-13.52	2.274	18.56	1.65E-05	5.60E-04	Cluster: Chromosome chr5 scaffold_67, genome
BG485751	-13.61	2.314	18.55	1.65E-05	5.60E-04	Chromosome undetermined scaffold_69, genome
TC165054	-13.99	2.585	18.54	1.66E-05	5.63E-04	Cluster: Predicted protein; Physcomitrella patens subsp.
TC184854	-13.44	2.324	18.52	1.68E-05	5.66E-04	Acidic ribosomal protein P1a-like; Solanum tuberosum|
TC160114	-14.17	2.924	18.52	1.68E-05	5.67E-04	Cluster: Chromosome chr1 scaffold_5, genome
TC185522	-13.83	2.324	18.49	1.71E-05	5.73E-04	
TC170763	-12.75	2.447	18.37	1.82E-05	6.07E-04	Chromosome chr1 scaffold_46, genome
TC183914	-13.62	2.234	18.37	1.82E-05	6.07E-04	Lea protein; Pseudotsuga menziesii|Rep: Lea protein
TC172483	-13.51	2.272	18.35	1.84E-05	6.10E-04	Cluster: Chromosome undetermined scaffold_503, genome
TC189641	-14.29	3.217	18.34	1.84E-05	6.11E-04	PR10 protein; Pinus monticola|Rep: PR10 protein
TC188500	-13.9	2.479	18.29	1.90E-05	6.27E-04	Clavata-like receptor; Picea glauca|Rep:
TC197678	-12.41	2.42	18.25	1.93E-05	6.34E-04	Chromosome undetermined scaffold_69, genome
TC185389	-13.47	2.224	18.24	1.95E-05	6.37E-04	Chitinase precursor; Oryza sativa|Rep:
TC178162	-12.74	2.558	18.22	1.97E-05	6.44E-04	Predicted protein; Physcomitrella patens subsp.
TC156863	-13.89	2.589	18.18	2.01E-05	6.55E-04	PaMip-2; Picea abies|Rep: PaMip-2 - Picea
TC184791	-13.58	2.199	18.12	2.07E-05	6.70E-04	Predicted protein; Physcomitrella patens subsp.
TC165858	-13.73	2.357	18.09	2.10E-05	6.78E-04	Aspartic proteinase nepenthesin II-like; Oryza sativa
TC170490	-14.04	2.707	17.99	2.22E-05	7.10E-04	Short-chain type dehydrogenase/reductase; Picea abies

TC157574	-13.71	2.204	17.97	2.24E-05	7.16E-04	Nucleoside diphosphate kinase; Pinus pinaster
TC175960	-13.54	2.267	17.88	2.35E-05	7.44E-04	Cluster: NtPRp27-like protein; Solanum tuberosum|Rep:
TC165389	-12.79	2.39	17.86	2.38E-05	7.50E-04	Cluster: Predicted protein; Physcomitrella patens subsp.
TC163333	-13.85	2.371	17.79	2.47E-05	7.76E-04	Chromosome chr1 scaffold_46, genome
TC188343	-12.72	2.505	17.77	2.49E-05	7.79E-04	Zinc finger protein YER130C; Saccharomyces cerevisiae|
TC182683	-13.69	2.158	17.76	2.51E-05	7.79E-04	Adenosine 5' phosphosulfate reductase; Populus tremula
TC186997	-14.28	3.094	17.71	2.57E-05	7.96E-04	Chromosome undetermined scaffold_3238, genome
TC169926	-13	2.435	17.67	2.62E-05	8.09E-04	Cluster: Chromosome chr4 scaffold_6, genome
TC177869	-12.59	2.392	17.63	2.68E-05	8.26E-04	Cluster: HD2 type histone deacetylase; Physcomitrella
TC171369	-13.26	2.397	17.6	2.73E-05	8.35E-04	HD2 type histone deacetylase; Physcomitrella patens
TC167029	-13.81	2.269	17.57	2.76E-05	8.44E-04	Chromosome chr18 scaffold_1, genome
TC168999	-12.69	2.374	17.54	2.82E-05	8.58E-04	Lea protein; Pseudotsuga menziesii|Rep: Lea protein
TC196902	-13.15	2.365	17.52	2.85E-05	8.67E-04	Cluster: Thaumatin-like protein; Pseudotsuga menziesii|
TC160460	-13.78	2.209	17.5	2.88E-05	8.72E-04	At1g60730/F8A5_24; Arabidopsis |Rep: At1g60730/F8A5_24
TC161914	-13.01	2.418	17.44	2.97E-05	8.95E-04	At1g60730/F8A5_24; Arabidopsis |Rep: At1g60730/F8A5_24
TC167835	-14.15	2.659	17.41	3.01E-05	9.08E-04	Chromosome undetermined scaffold_143, genome
TC177670	-13.71	2.23	17.4	3.03E-05	9.11E-04	Lipoxygenase; Vitis vinifera|Rep: Lipoxygenase - Vitis
TC165131	-13.6	2.136	17.3	3.20E-05	9.56E-04	Cluster: Chromosome chr6 scaffold_3, genome
TC192008	-14.12	2.611	17.19	3.38E-05	1.01E-03	Cluster: Chromosome chr2 scaffold_112, genome
TC163778	-12.01	2.473	17.17	3.42E-05	1.01E-03	Cluster: SAMT; Browallia americana|Rep: SAMT -
TC188216	-12.3	2.474	17.16	3.43E-05	1.02E-03	Chloroplast omega-3 fatty acid desaturase; Olea
TC166330	-13.4	2.288	17.05	3.64E-05	1.07E-03	Chromosome chr13 scaffold_17, genome
BF186114	-13.91	2.353	16.85	4.05E-05	1.18E-03	Cluster: Aldose 1-epimerase family protein, expressed;
TC163454	-13.99	2.444	16.71	4.35E-05	1.26E-03	Chromosome chr10 scaffold_138, genome
TC168815	-13.6	2.073	16.7	4.38E-05	1.27E-03	Cluster: HD2 type histone deacetylase; Physcomitrella
TC168193	-13.34	2.149	16.64	4.51E-05	1.30E-03	Predicted protein; Physcomitrella patens subsp.
TC169738	-13.27	2.235	16.61	4.59E-05	1.32E-03	Cluster: Expansin-like protein; Cunninghamia lanceolata|
TC166768	-13.47	2.127	16.57	4.69E-05	1.35E-03	Chromosome chr14 scaffold_9, genome
TC197859	-13.11	2.247	16.48	4.92E-05	1.41E-03	Chromosome undetermined scaffold_69, genome
TC183428	-14.16	2.491	16.46	4.97E-05	1.42E-03	Cluster: Ser/Thr protein phosphatase family protein,
TC192806	-14.03	2.382	16.43	5.06E-05	1.44E-03	Chromosome chr10 scaffold_179, genome
TC178402	-13.14	2.288	16.38	5.19E-05	1.47E-03	Cluster: 13-lipoxygenase; Solanum tuberosum|Rep: 13-
DR102084	-13.37	2.121	16.38	5.19E-05	1.47E-03	Cluster: Os09g0108600 protein; Oryza sativa Japonica
TC195112	-13.94	2.41	16.36	5.24E-05	1.48E-03	
TC182855	-13.97	2.597	16.29	5.44E-05	1.53E-03	Low temperature and salt responsive protein;
TC190934	-12.4	2.291	16.22	5.64E-05	1.59E-03	Cluster: Pleiotropic drug resistance protein 1;
TC168521	-13.2	2.245	16.21	5.68E-05	1.59E-03	Cluster: Chromosome chr19 scaffold_4, genome
EC428696	-13.9	2.322	16.18	5.77E-05	1.61E-03	Helix-turn-helix Fis-type protein; Nitrobacter sp.
TC173536	-13.27	2.145	16.18	5.77E-05	1.61E-03	Basic endochitinase 1 precursor; Oryza sativa
TC164478	-13.94	2.37	16.06	6.14E-05	1.69E-03	Cluster: Chromosome chr9 scaffold_7, genome
TC160983	-13.91	2.372	16.06	6.15E-05	1.69E-03	Carbohydrate kinase, YjeF related protein; Burkholderia
TC154867	-13.96	2.472	16.05	6.16E-05	1.69E-03	Cluster: S-receptor kinase -like protein; Arabidopsis
TC183805	-13.28	2.07	16.04	6.20E-05	1.70E-03	Dehydrin-like protein; Picea abies|Rep:
TC169645	-13.84	2.27	15.96	6.48E-05	1.77E-03	Lipoxygenase 3; Actinidia deliciosa|Rep: Lipoxygenase 3
TC159148	-13.9	2.199	15.95	6.52E-05	1.78E-03	Cluster: Glutathione S-transferase GST 10; Zea
TC176536	-14.22	2.65	15.93	6.58E-05	1.79E-03	Gb|AAF26953.1; Arabidopsis |Rep: Gb|AAF26953.1 -
TC162711	-13.65	2.033	15.91	6.65E-05	1.81E-03	NOX2; Striga asiatica|Rep: NOX2 - Striga
TC167723	-12.97	2.285	15.87	6.77E-05	1.83E-03	Chromosome chr4 scaffold_6, genome
TC156146	-11.78	2.379	15.8	7.03E-05	1.90E-03	Chromosome undetermined scaffold_69, genome
TC173647	-14.19	2.674	15.78	7.13E-05	1.92E-03	Galactinol synthase; Brassica napus|Rep: Galactinol
TC164027	-13.79	2.229	15.76	7.17E-05	1.92E-03	Riboflavin biosynthesis protein ribBA, chloroplast
TC170768	-13.91	2.291	15.76	7.18E-05	1.92E-03	Clavata-like receptor; Picea glauca|Rep: Clavata-like
TC193548	-14.31	3.012	15.76	7.18E-05	1.92E-03	
TC155146	-13.4	3.001	15.75	7.49E-05	1.99E-03	Cluster: glycosyl hydrolase family 17 protein;
TC161396	-12.09	2.354	15.62	7.75E-05	2.05E-03	Cluster: Chromosome chr1 scaffold_5, genome
TC166238	-12.86	2.221	15.59	7.88E-05	2.08E-03	LEA; Pinus halepensis|Rep: LEA - Pinus
TC171983	-13.81	2.183	15.53	8.10E-05	2.13E-03	LP3-1; Pinus|Rep: LP3-1 - Pinus taeda
TC185381	-13.16	2.224	15.52	8.15E-05	2.14E-03	Lea protein; Pseudotsuga menziesii|Rep: Lea protein
GW751850	-13.93	2.196	15.5	8.26E-05	2.16E-03	Cluster: Zinc finger CCHC domain-containing protein 10;
TC171549	-12.97	2.215	15.37	8.85E-05	2.30E-03	30S ribosomal protein S7; Strelitzia reginae|Rep:
TC155389	-13.67	2.06	15.21	9.63E-05	2.48E-03	Chromosome undetermined scaffold_55, genome
TC169728	-13.05	2.204	15.21	9.63E-05	2.48E-03	Arginine decarboxylase; Brassica juncea|Rep: Arginine
GW755072	-13.19	2.222	15.17	9.84E-05	2.52E-03	Cluster: HD2 type histone deacetylase; Physcomitrella
GW755937	-13.36	2.14	15.14	9.96E-05	2.55E-03	Cluster: HD2 type histone deacetylase; Physcomitrella
TC157512	-13.91	2.15	15.14	1.00E-04	2.55E-03	Nucleoside diphosphate kinase; Pinus pinaster|Rep:
TC170513	-13.23	2.165	15.11	1.01E-04	2.58E-03	Cluster: Predicted protein; Physcomitrella patens subsp.
TC182376	-13.46	2.042	15.07	1.04E-04	2.63E-03	40S ribosomal protein S15; Picea|Rep: 40S
TC183232	-13.02	2.217	15.05	1.05E-04	2.65E-03	Dirigent-like protein pDIR14; Picea engelmannii x
TC167815	-12.88	2.184	15.03	1.06E-04	2.67E-03	Epa5p; Candida glabrata|Rep: Epa5p - Candida
GT257093	-13.3	2.096	15	1.08E-04	2.70E-03	Glycosyl hydrolases family 17 protein, expressed;
TC166526	-12.7	2.182	14.94	1.11E-04	2.79E-03	Dehydrin 1; Pinus pinaster|Rep: Dehydrin 1
TC167398	-13.32	1.995	14.86	1.16E-04	2.90E-03	Cluster: Thaumatin-like protein; Pseudotsuga menziesii|
TC166735	-13.88	2.154	14.85	1.16E-04	2.91E-03	IAA-amino acid hydrolase ILR1-like 2 precursor;
TC171040	-13.75	1.971	14.85	1.17E-04	2.91E-03	Cluster: Ribonuclease, Rne/Rng family; Mycobacterium
TC156376	-13.22	2.141	14.81	1.19E-04	2.95E-03	Alternative oxidase; Nicotiana attenuata|Rep:
TC156322	-13.73	2.086	14.77	1.21E-04	3.00E-03	Pollen allergen CJP38; Cryptomeria japonica|Rep: Pollen
TC174556	-12.69	2.131	14.69	1.27E-04	3.13E-03	
TC198869	-13.36	1.965	14.65	1.29E-04	3.17E-03	Chromosome undetermined scaffold_55, genome
TC170205	-13.63	1.927	14.64	1.30E-04	3.18E-03	Chromosome chr17 scaffold_16, genome
TC169768	-12.04	2.307	14.6	1.33E-04	3.24E-03	Cluster: Os04g0412300 protein; Oryza sativa Japonica
TC154764	-11.61	2.14	14.54	1.37E-04	3.33E-03	Arginine decarboxylase; Pinus sylvestris|Rep: Arginine

TC154934	-13.29	2.098	14.52	1.39E-04	3.36E-03	Cluster: Predicted protein; Physcomitrella patens subsp.
GO096100	-13.36	1.916	14.43	1.45E-04	3.50E-03	Dehydrin 1; Pinus densata|Rep: Dehydrin 1
TC169073	-13.51	1.907	14.42	1.46E-04	3.51E-03	Chromosome chr18 scaffold_1, genome
TC166894	-13.93	2.106	14.35	1.52E-04	3.64E-03	kinase; Pinus pinaster|Rep: Nucleoside diphosphate
TC166469	-13.85	2.005	14.25	1.60E-04	3.84E-03	tetra-peptide repeat homeobox-like (TPRXL) on chromosome3
TC161665	-13.71	1.909	14.16	1.68E-04	3.97E-03	Chromosome chr6 scaffold_15, genome
TC178312	-13.66	1.936	14.14	1.70E-04	3.99E-03	Cluster: Similarity to cell wall-plasma membrane linker
TC176682	-13.52	1.913	14.14	1.70E-04	4.00E-03	Cluster: RNA recognition motif family protein, expressed;
FE520245	-13.59	1.912	14.07	1.76E-04	4.13E-03	Dehydrin 1; Pinus pinaster|Rep: Dehydrin 1
TC192263	-14.13	2.47	14.01	1.82E-04	4.26E-03	
TC165002	-13.32	1.933	14	1.83E-04	4.26E-03	Chromosome chr18 scaffold_1, genome
GT230507	-13.58	1.89	13.91	1.92E-04	4.46E-03	Predicted protein; Physcomitrella patens subsp.
TC168069	-13.61	1.95	13.84	1.99E-04	4.59E-03	Oryza sativa Japonica Group|Rep: Os04g0524800 protein
TC172763	-13.04	2.089	13.84	1.99E-04	4.59E-03	Acidic ribosomal protein P1a-like; Solanum tuberosum|
TC173952	-13.51	1.934	13.81	2.02E-04	4.64E-03	Cysteine synthase; Vitis vinifera|Rep: Cysteine synthase
TC173735	-13.29	2.064	13.81	2.03E-04	4.64E-03	Predicted protein; Physcomitrella patens subsp.
TC160565	-12.34	2.121	13.77	2.06E-04	4.71E-03	Endo-beta-1,4-glucanase; Fragaria x ananassa|Rep:
TC195769	-13.44	1.975	13.65	2.20E-04	4.98E-03	Cluster: Chromosome chr14 scaffold_21, genome
TC169669	-12.25	2.135	13.65	2.20E-04	4.98E-03	Chromosome chr4 scaffold_6, genome
TC167940	-13.88	2.098	13.62	2.24E-04	5.03E-03	Lea protein; Pseudotsuga menziesii|Rep: Lea protein
GT236181	-13.45	1.873	13.61	2.25E-04	5.05E-03	Cluster: Nodulin-like protein; Arabidopsis |Rep:
TC175318	-13.86	2.071	13.55	2.32E-04	5.19E-03	Chromosome chr18 scaffold_1, genome
GW747229	-13.04	2.046	13.49	2.40E-04	5.36E-03	
TC167855	-12.98	2.054	13.43	2.48E-04	5.52E-03	
TC189094	-12.45	2.057	13.4	2.51E-04	5.58E-03	Cluster: Chromosome chr18 scaffold_122, genome
TC163992	-12.18	2.156	13.39	2.53E-04	5.61E-03	Cluster: HD2 type histone deacetylase; Physcomitrella
TC197324	-13.07	2.069	13.37	2.55E-04	5.65E-03	Chromosome chr7 scaffold_20, genome
TC198917	-13.25	2.006	13.36	2.57E-04	5.67E-03	Chromosome chr14 scaffold_9, genome
TC173946	-13.57	1.849	13.3	2.65E-04	5.84E-03	Cluster: Chromosome chr8 scaffold_29, genome
TC185919	-13.8	2.004	13.29	2.67E-04	5.89E-03	LP3-1; Pinus|Rep: LP3-1 - Pinus taeda
TC158857	-14.06	2.163	13.28	2.68E-04	5.89E-03	Properoxidase precursor; Picea Properoxidase precursor
TC187459	-12.35	2.088	13.25	2.73E-04	5.97E-03	Cluster: Chromosome chr18 scaffold_122, genome
TC179960	-12.87	2.098	13.19	2.82E-04	6.15E-03	Chromosome chr18 scaffold_1, genome
TC173667	-13.44	1.883	13.18	2.84E-04	6.17E-03	T3H13.3 protein; Arabidopsis |Rep: T3H13.3 protein
TC187214	-12.91	2.016	13.17	2.85E-04	6.18E-03	60S ribosomal protein L17-2; Arabidopsis |Rep:
TC157989	-13.94	2.271	13.12	2.92E-04	6.33E-03	Cluster: Chromosome undetermined scaffold_155, genome
TC161868	-12.63	1.976	13.11	2.94E-04	6.35E-03	Cluster: Predicted protein; Physcomitrella patens subsp.
TC155499	-13.74	1.894	13.1	2.95E-04	6.36E-03	Cluster: PtxA protein precursor; Pisum sativum|Rep:
TC167695	-12.78	2.114	13.08	2.98E-04	6.41E-03	
TC194580	-13.44	1.8	13.05	3.03E-04	6.49E-03	UDP-glucose dehydrogenase; Bambusa oldhamii|Rep: UDP-
TC159559	-13.38	1.895	13.03	3.06E-04	6.54E-03	Chromosome chr17 scaffold_16, genome
TC179808	-13.96	2.11	13.01	3.10E-04	6.60E-03	Cluster: Qa-SNARE, Sso1/Syntaxin1-type, SYP12A-group;
TC156005	-13.83	2.009	13	3.12E-04	6.63E-03	Pinus sylvestris|Rep: Arginine decarboxylase - Pinus
TC185044	-12.72	1.961	12.98	3.14E-04	6.67E-03	Chromosome undetermined scaffold_155, genome
TC191731	-13.64	1.882	12.94	3.21E-04	6.77E-03	Basic endochitinase 1 precursor; Oryza sativa
TC190411	-12.22	2.083	12.9	3.28E-04	6.90E-03	Endo-beta-1,4-glucanase precursor; Prunus persica|Rep:
TC170786	-13.49	1.858	12.9	3.29E-04	6.91E-03	Cluster: Chromosome chr1 scaffold_75, genome
TC171936	-13.88	1.999	12.89	3.30E-04	6.92E-03	Chromosome chr12 scaffold_18, genome
TC168010	-12.36	1.967	12.88	3.32E-04	6.94E-03	Dehydrin 1; Pinus pinaster|Rep: Dehydrin 1
TC183662	-14.03	2.301	12.82	3.43E-04	7.14E-03	Class I chitinase; Triticum aestivum|Rep: Class
TC188904	-13.58	1.798	12.81	3.45E-04	7.18E-03	
TC174588	-12.3	2.03	12.8	3.46E-04	7.19E-03	Cluster: Chromosome chr3 scaffold_8, genome
TC167814	-12.56	1.958	12.74	3.58E-04	7.38E-03	Predicted protein; Physcomitrella patens subsp.
TC160832	-12.08	2.05	12.74	3.58E-04	7.38E-03	
TC179209	-12.77	1.969	12.74	3.59E-04	7.39E-03	LEA; Pinus halepensis|Rep: LEA - Pinus
TC187082	-12.39	2.003	12.7	3.66E-04	7.51E-03	Class IV chitinase Chia4-Pa1; Picea abies|Rep:
TC177638	-14.01	2.177	12.69	3.67E-04	7.51E-03	Cluster: T3H13.3 protein; Arabidopsis |Rep: T3H13.3
TC164415	-13.63	1.804	12.69	3.68E-04	7.51E-03	Cluster: At5g65400; Arabidopsis |Rep: At5g65400 -
TC165834	-13.04	2.081	12.62	3.82E-04	7.78E-03	Cluster: Chitinase A; Cycas revoluta|Rep: Chitinase
TC158601	-13.14	1.991	12.55	3.95E-04	8.02E-03	Cluster: Chromosome chr7 scaffold_42, genome
DT627566	-13.84	1.865	12.55	3.97E-04	8.03E-03	Chromosome chr4 scaffold_6, genome
TC196772	-13.29	1.836	12.48	4.11E-04	8.29E-03	
TC194792	-13.41	1.778	12.48	4.12E-04	8.30E-03	Os03g0733800 protein; Oryza sativa Japonica Group|Rep:
TC174693	-13.47	1.824	12.46	4.16E-04	8.37E-03	H/ACA ribonucleoprotein complex subunit 4; Arabidopsis
TC194944	-12.6	1.911	12.41	4.26E-04	8.52E-03	Cluster: Chromosome chr7 scaffold_42, genome
TC158128	-13.59	1.827	12.41	4.26E-04	8.52E-03	Chromosome chr12 scaffold_36, genome
TC186056	-13.54	1.741	12.41	4.28E-04	8.54E-03	Cluster: Chromosome chr9 scaffold_104, genome
TC173179	-12.94	2.005	12.35	4.42E-04	8.78E-03	Chromosome chr18 scaffold_1, genome
TC157575	-13.89	1.935	12.32	4.49E-04	8.89E-03	Proline-rich protein; Pinus taeda|Rep: Proline-rich
BQ107128	-12.32	1.915	12.3	4.53E-04	8.96E-03	Cluster: Chromosome segregation ATPases-like protein;
TC169313	-13.88	1.951	12.29	4.55E-04	8.97E-03	Pectate lyase 1; Musa acuminata|Rep: Pectate
CT582167	-14.13	2.114	12.25	4.66E-04	9.18E-03	Picea mariana|Rep: Fibrillarin - Picea mariana (Black
GW749382	-13.41	1.86	12.21	4.75E-04	9.31E-03	Oryza sativa Japonica Group|Rep: Os04g0524800 protein
TC164810	-13.82	1.829	12.2	4.78E-04	9.33E-03	Short-chain type dehydrogenase/reductase; Picea abies|
GW726377	-13.1	1.944	12.2	4.79E-04	9.33E-03	40S ribosomal protein S29; Arabidopsis |Rep:
TC192507	-13.48	1.796	12.18	4.82E-04	9.38E-03	Predicted protein; Physcomitrella patens subsp.
TC179486	-13.55	1.733	12.18	4.84E-04	9.40E-03	Dehydrin; Pinus sylvestris|Rep: Dehydrin - Pinus
GT255591	-13.33	1.763	12.16	4.88E-04	9.46E-03	
TC163737	-13.28	1.876	12.16	4.88E-04	9.46E-03	Cluster: Predicted protein; Physcomitrella patens subsp.
TC180031	-13.37	1.821	12.06	5.14E-04	9.93E-03	Ferritin, chloroplast; Physcomitrella patens subsp.
TC174826	-13.29	1.874	12.03	5.23E-04	1.01E-02	Cluster: Lipase family protein; Methylococcus

TC182150	-13.51	1.72	12.01	5.31E-04	1.02E-02	MybSt1; Solanum tuberosum|Rep: MybSt1 - Solanum
TC184325	-13.48	1.743	11.98	5.38E-04	1.03E-02	40S ribosomal protein S29; Arabidopsis |Rep:
TC154668	-13.44	1.765	11.96	5.44E-04	1.04E-02	Chromosome chr3 scaffold_199, genome
TC175882	-13.36	1.809	11.95	5.47E-04	1.04E-02	Alternative oxidase; Nicotiana attenuata|Rep:
TC174061	-13.01	1.912	11.93	5.53E-04	1.05E-02	Chromosome undetermined scaffold_2463, genome
GT231163	-13.44	1.814	11.91	5.58E-04	1.06E-02	Predicted protein; Physcomitrella patens subsp.
TC178525	-13.6	1.791	11.89	5.66E-04	1.07E-02	Chromosome chr12 scaffold_78, genome
TC172648	-12.77	1.925	11.86	5.73E-04	1.08E-02	Pinus densata|Rep: Dehydrin 1 - Pinus
TC167598	-12.69	1.896	11.84	5.78E-04	1.09E-02	Cluster: Chloroplast nucleoid DNA-binding protein-like
TC155790	-13.57	1.718	11.8	5.92E-04	1.11E-02	Wrp15a; Citrullus lanatus|Rep: Wrp15a - Citrullus
TC191743	-12.48	1.851	11.8	5.93E-04	1.11E-02	Cluster: Chromosome chr7 scaffold_42, genome
GW753016	-13.77	1.894	11.78	5.99E-04	1.12E-02	Cluster: Aspartyl protease-like; Oryza sativa Japonica
TC197070	-13.72	1.731	11.77	6.02E-04	1.13E-02	Os03g0750700 protein; n=4; Oryza sativa|Rep:
TC184813	-12.55	1.889	11.75	6.09E-04	1.14E-02	Chromosome chr18 scaffold_1, genome
TC189378	-13.01	1.918	11.73	6.16E-04	1.15E-02	Predicted protein; Physcomitrella patens subsp.
CN852188	-13	1.905	11.69	6.28E-04	1.16E-02	Protein TAR1; Kluyveromyces lactis|Rep: Protein TAR1
TC177726	-12.83	1.902	11.67	6.35E-04	1.17E-02	Cluster: Chromosome chr18 scaffold_1, genome
TC155321	-13.86	1.846	11.67	6.36E-04	1.17E-02	NAC family protein; Prunus mume|Rep: NAC
TC158019	-12.72	1.849	11.66	6.38E-04	1.18E-02	Chromosome chr17 scaffold_16, genome
TC188819	-13.15	1.884	11.66	6.40E-04	1.18E-02	Glutathione peroxidase; Oryza sativa Indica Group|Rep:
TC164013	-14.36	2.397	11.61	6.57E-04	1.20E-02	Tau class glutathione S-transferase; Pinus densata|Rep:
TC175289	-13.98	2.086	11.6	6.60E-04	1.21E-02	Alternative oxidase 2, mitochondrial precursor;
TC165939	-13.31	1.847	11.59	6.63E-04	1.21E-02	Cluster: Predicted protein; Ostreococcus lucimarinus
TC188647	-12.81	1.843	11.58	6.66E-04	1.21E-02	40S ribosomal protein S29; Arabidopsis |Rep:
TC156468	-13.89	1.912	11.58	6.68E-04	1.21E-02	Cysteine synthase; Malus x domestica|Rep: Cysteine
TC165289	-13.69	1.673	11.55	6.77E-04	1.22E-02	Wrp15a; Citrullus lanatus|Rep: Wrp15a - Citrullus
TC157341	-14.19	2.188	11.54	6.79E-04	1.22E-02	ATP sulfurylase; Camellia sinensis|Rep: ATP sulfurylase
AW754605	-12.81	1.916	11.51	6.91E-04	1.24E-02	
TC161861	-13.5	1.655	11.43	7.22E-04	1.29E-02	Chromosome chr5 scaffold_64, genome
TC157450	-12.76	1.859	11.39	7.38E-04	1.32E-02	Predicted protein; Physcomitrella patens subsp.
TC187956	-12.96	1.846	11.37	7.45E-04	1.33E-02	Dehydrin 1; Pinus pinaster|Rep: Dehydrin 1
TC197481	-12.81	1.849	11.36	7.52E-04	1.34E-02	Cluster: Predicted protein; Physcomitrella patens subsp.
TC159605	-13.86	1.813	11.35	7.56E-04	1.34E-02	3-ketoacyl-CoA thiolase 2, peroxisomal precursor;
NP542733	-12.77	1.846	11.3	7.74E-04	1.37E-02	pinaster]
TC164248	-13.79	1.679	11.29	7.78E-04	1.37E-02	Wrp15a; Citrullus lanatus|Rep: Wrp15a - Citrullus
TC166415	-12.03	1.952	11.24	8.00E-04	1.40E-02	Cluster: Os09g0108600 protein; Oryza sativa Japonica
TC166890	-13.72	1.652	11.17	8.29E-04	1.45E-02	Cluster: Prephenate dehydratase; Ipomoea trifida|Rep:
TC194378	-13.85	1.814	11.17	8.31E-04	1.45E-02	Picea mariana|Rep: Probable aquaporin - Picea
TC194064	-13.33	1.685	11.16	8.36E-04	1.46E-02	Basic endochitinase C precursor; Secale cereale|Rep:
CT578062	-13.04	1.84	11.12	8.54E-04	1.48E-02	40S ribosomal protein S29; Arabidopsis |Rep:
TC179139	-13.73	1.722	11.11	8.58E-04	1.48E-02	Cluster: PtxA protein precursor; Pisum sativum|Rep:
TC169163	-13.65	1.668	11.09	8.68E-04	1.50E-02	60S ribosomal protein L6; Vitis vinifera|Rep:
TC168687	-14.07	2.152	11.07	8.78E-04	1.51E-02	Cluster: Beta-1,3-glucanase-like protein; Arabidopsis |
TC175979	-13.44	1.643	11.05	8.88E-04	1.53E-02	Cluster: Chromosome chr7 scaffold_20, genome
GW757577	-12.46	1.795	11.03	8.96E-04	1.54E-02	Chromosome chr7 scaffold_20, genome
TC187313	-13.32	1.704	10.98	9.22E-04	1.57E-02	Chromosome chr1 scaffold_84, genome
TC171685	-12.79	1.812	10.89	9.66E-04	1.64E-02	Dehydrin-like protein; Picea abies|Rep: Dehydrin-like
TC195661	-13.35	1.705	10.89	9.68E-04	1.64E-02	Chromosome undetermined scaffold_252, genome
TC176574	-12.92	1.833	10.88	9.73E-04	1.64E-02	Dehydrin 1; Pinus pinaster|Rep: Dehydrin 1
DR023673	-13.08	1.819	10.86	9.80E-04	1.65E-02	Cluster: Chromosome chr18 scaffold_122, genome
TC198079	-12.52	1.802	10.85	9.87E-04	1.65E-02	Chromosome chr8 scaffold_29, genome
TC194745	-13.81	1.745	10.82	1.01E-03	1.68E-02	Cluster: Similarity to cell wall-plasma membrane linker
TC172986	-13.8	1.671	10.79	1.02E-03	1.70E-02	Chromosome undetermined scaffold_324, genome
TC195339	-12.93	1.807	10.79	1.02E-03	1.70E-02	Predicted protein; Physcomitrella patens subsp.
TC181728	-12.58	1.759	10.78	1.02E-03	1.70E-02	MYB transcription factor MYB52; Glycine max|Rep:
TC196254	-13.93	1.86	10.71	1.07E-03	1.76E-02	
TC168885	-13.69	1.699	10.68	1.08E-03	1.78E-02	Chromosome chr13 scaffold_149, genome
DR746310	-13.67	1.621	10.66	1.10E-03	1.80E-02	Acidic thaumatin-like protein 10; Taxus x
FE523036	-13.56	1.6	10.65	1.10E-03	1.81E-02	Uncharacterized protein At2g01250.2; Arabidopsis |Rep:
TC170084	-12.4	1.757	10.62	1.12E-03	1.83E-02	Dehydrin 4; Picea abies|Rep: Dehydrin 4
BX682909	-13.08	1.785	10.54	1.17E-03	1.91E-02	60S ribosomal protein L15-2; n=7; Picea|Rep: 60S
TC162683	-12.54	1.767	10.51	1.19E-03	1.93E-02	Chromosome chr8 scaffold_29, genome
FE524490	-12.69	1.732	10.47	1.21E-03	1.96E-02	Predicted protein; Physcomitrella patens subsp.
TC168238	-13.86	1.679	10.46	1.22E-03	1.97E-02	Chaperonin CPN60-1, mitochondrial precursor; Zea mays|
TC176635	-11.9	1.856	10.45	1.23E-03	1.98E-02	Amino acid permease; Plantago major|Rep: Amino
TC161755	-13.47	1.662	10.43	1.24E-03	2.00E-02	60S ribosomal protein L31; Picea|Rep: 60S
TC159512	-12.79	1.757	10.42	1.25E-03	2.00E-02	Chromosome chr18 scaffold_1, genome
TC160936	-13.31	1.736	10.42	1.25E-03	2.00E-02	Chromosome chr17 scaffold_16, genome
TC194347	-14.03	1.89	10.4	1.26E-03	2.02E-02	Sugar transport protein 13; Arabidopsis |Rep:
DR015600	-13.49	1.59	10.4	1.26E-03	2.02E-02	Cluster: Chromosome chr4 scaffold_333, genome
TC182899	-14.16	2.062	10.38	1.28E-03	2.03E-02	Cluster: Hexokinase 4a; Nicotiana tabacum|Rep:
FG615188	-13.42	1.634	10.35	1.29E-03	2.06E-02	Cluster: Beta-glucosidase; Cicer arietinum|Rep: Beta-
TC171842	-13	1.76	10.3	1.33E-03	2.10E-02	Basic endochitinase 1 precursor; Oryza sativa
TC160866	-13.05	1.768	10.25	1.37E-03	2.16E-02	Chromosome chr17 scaffold_16, genome
TC179461	-12.93	1.746	10.24	1.37E-03	2.17E-02	Os12g0154000 protein; Oryza sativa Japonica Group|Rep:
GW754477	-12.81	1.766	10.23	1.38E-03	2.18E-02	Pinus|Rep: Dehydrin 1 - Pinus yunnanensis
TC156152	-14.26	2.178	10.21	1.39E-03	2.19E-02	Probable pyridoxal biosynthesis protein PDX1; Ginkgo
TC161397	-13.77	1.685	10.2	1.41E-03	2.21E-02	Cluster: NtPRp27; Nicotiana tabacum|Rep: NtPRp27 -
TC157339	-12.41	1.714	10.19	1.41E-03	2.21E-02	Dehydrin 1; Pinus pinaster|Rep: Dehydrin 1
TC154805	-14.2	2.038	10.19	1.41E-03	2.21E-02	Chromosome chr2 scaffold_11, genome
TC192453	-13.91	1.68	10.17	1.43E-03	2.23E-02	

TC196803	-13.82	1.635	10.16	1.43E-03	2.24E-02	Chromosome chr19 scaffold_4, genome
TC192310	-12.95	1.756	10.16	1.44E-03	2.24E-02	Protein TAR1; Kluyveromyces lactis|Rep: Protein TAR1
TC185312	-13.57	1.592	10.13	1.46E-03	2.27E-02	Chromosome chr8 scaffold_23, genome
TC184693	-13.27	1.688	10.13	1.46E-03	2.27E-02	
TC179939	-13.88	1.684	10.12	1.47E-03	2.28E-02	T6D22.12; Arabidopsis |Rep: T6D22.12 - Arabidopsis
TC175745	-13.33	1.641	10.1	1.48E-03	2.29E-02	Cluster: Os12g0637800 protein; Oryza sativa Japonica
TC172853	-13.38	1.639	10.07	1.51E-03	2.32E-02	40S ribosomal protein S16-1; Arabidopsis |Rep:
TC159968	-13.32	1.683	10.05	1.53E-03	2.34E-02	Peptidyl-prolyl cis-trans isomerase; Helianthus annuus|
TC162177	-13.93	1.752	10.05	1.53E-03	2.34E-02	Cluster: nodulin MtN3 family protein; Arabidopsis
TC184194	-13.8	1.59	10.04	1.53E-03	2.34E-02	Cluster: Chromosome chr14 scaffold_211, genome
TC182760	-11.69	1.733	10.03	1.54E-03	2.36E-02	Monosaccharide transporter 4; Oryza sativa Japonica
TC167652	-13.95	1.827	10.01	1.55E-03	2.37E-02	Alternative oxidase; Nicotiana attenuata|Rep:
TC188814	-13.95	1.816	9.969	1.59E-03	2.42E-02	Predicted protein; Physcomitrella patens subsp.
TC158340	-13.2	1.639	9.937	1.62E-03	2.45E-02	Chromosome chr19 scaffold_4, genome
TC195940	-12.21	1.777	9.882	1.67E-03	2.50E-02	Glutathione peroxidase; Oryza sativa Indica Group|Rep:
TC157333	-12.68	1.692	9.854	1.69E-03	2.53E-02	Cluster: Chromosome chr14 scaffold_21, genome
TC156749	-13.17	1.712	9.828	1.72E-03	2.56E-02	Ribitol dehydrogenase isolog; Arabidopsis |Rep: Ribitol
GW740723	-13.57	1.533	9.828	1.72E-03	2.56E-02	Chromosome chr14 scaffold_21, genome
TC186739	-13.08	1.709	9.815	1.73E-03	2.58E-02	Glutathione peroxidase; Oryza sativa Indica Group|Rep:
TC158494	-14.01	1.917	9.776	1.77E-03	2.63E-02	Basic endochitinase C precursor; Secale cereale|Rep:
TC189658	-13.81	2.009	9.763	1.78E-03	2.64E-02	Protein TAR1; Saccharomyces cerevisiae|Rep: Protein TAR1
TC199051	-14.03	1.708	9.762	1.78E-03	2.64E-02	Type IIB calcium ATPase; Medicago truncatula|Rep:
TC185447	-13.12	1.635	9.76	1.78E-03	2.64E-02	EDS1; Nicotiana tabacum|Rep: EDS1 - Nicotiana
BI416965	-13.4	1.512	9.746	1.80E-03	2.66E-02	
BF517394	-12.44	1.666	9.73	1.81E-03	2.67E-02	Dehydrin 4; Picea abies|Rep: Dehydrin 4
TC167043	-13.56	1.619	9.72	1.82E-03	2.68E-02	Os06g0136300 protein; Oryza sativa|Rep: Os06g0136300
TC155023	-13.22	1.707	9.718	1.83E-03	2.68E-02	Predicted protein; Physcomitrella patens subsp.
TC184845	-13.47	1.531	9.714	1.83E-03	2.68E-02	Chromosome chr8 scaffold_23, genome
DR052751	-13.75	1.631	9.705	1.84E-03	2.69E-02	Homeobox protein SIX5 (DM locus-associated homeodomain
TC180233	-12.6	1.689	9.66	1.88E-03	2.75E-02	Cluster: Furcatin hydrolase; Viburnum furcatum|Rep:
BX677751	-12.82	1.69	9.659	1.88E-03	2.75E-02	Dehydrin-like protein; Picea abies|Rep: Dehydrin-like
DR019633	-13.69	1.52	9.629	1.92E-03	2.79E-02	40S ribosomal protein S16-1; Arabidopsis |Rep:
TC158442	-13.48	1.52	9.595	1.95E-03	2.83E-02	Cluster: Chromosome chr9 scaffold_104, genome
TC184326	-13.8	1.555	9.58	1.97E-03	2.85E-02	Chaperonin CPN60-2, mitochondrial precursor; Cucurbita
TC167133	-12.27	1.751	9.547	2.00E-03	2.89E-02	Predicted protein; Physcomitrella patens subsp.
TC194737	-12.4	1.693	9.545	2.00E-03	2.89E-02	Chromosome chr14 scaffold_21, genome
TC171864	-12.75	1.655	9.52	2.03E-03	2.92E-02	Chromosome chr18 scaffold_59, genome
TC160622	-11.74	1.702	9.509	2.05E-03	2.93E-02	Hexose transporter; Solanum lycopersicum|Rep: Hexose
TC159200	-11.94	1.735	9.508	2.05E-03	2.93E-02	Glutathione peroxidase; Oryza sativa Indica Group|Rep:
TC170395	-13.43	1.505	9.505	2.05E-03	2.93E-02	UDP-glucose dehydrogenase; Cinnamomum osmophloeum|Re
TC158652	-13.56	1.516	9.487	2.07E-03	2.96E-02	Chaperonin CPN60-1, mitochondrial precursor; Zea mays|
TC174512	-12.04	1.772	9.473	2.09E-03	2.97E-02	Cluster: Similarity to RNA-binding protein; Arabidopsis
TC195503	-13.04	1.674	9.451	2.11E-03	3.00E-02	Cluster: Thaumatin-like protein; Pseudotsuga menziesii|
GT235772	-13.44	1.499	9.429	2.14E-03	3.03E-02	Cluster: Chloroplast nucleoid DNA-binding protein-like
TC158858	-13.72	1.507	9.401	2.17E-03	3.07E-02	Chromosome chr13 scaffold_17, genome
BI416666	-12.93	1.684	9.385	2.19E-03	3.09E-02	Chromosome chr2 scaffold_112, genome
CT578594	-14.07	1.733	9.365	2.21E-03	3.11E-02	chain dehydrogenase/reductase family protein;
CV147414	-13.62	1.621	9.318	2.27E-03	3.18E-02	LP3-1; Pinus|Rep: LP3-1 - Pinus taeda
TC160816	-13.84	1.644	9.232	2.38E-03	3.32E-02	Flavanone 3-hydroxylase; Ginkgo biloba|Rep:
CO161945	-13.23	1.674	9.219	2.40E-03	3.34E-02	Cluster: Chromosome chr12 scaffold_78, genome
TC184174	-14.23	1.894	9.206	2.41E-03	3.35E-02	Sporulation domain protein; Polynucleobacter sp. QLW-
TC163014	-12.85	1.624	9.19	2.43E-03	3.38E-02	Predicted protein; Physcomitrella patens subsp.
TC171200	-13.12	1.613	9.183	2.44E-03	3.39E-02	Chromosome undetermined scaffold_504, genome
TC177598	-13.21	1.681	9.162	2.47E-03	3.42E-02	Chromosome chr12 scaffold_78, genome
TC158475	-14.04	1.763	9.154	2.48E-03	3.43E-02	NAC protein 1 splice variant 2;
TC193941	-11.56	1.518	9.064	2.61E-03	3.58E-02	Low molecular weight heat-shock protein; Pseudotsuga
TC166875	-13.94	1.632	9.05	2.63E-03	3.60E-02	Chromosome chr6 scaffold_25, genome
TC175107	-13.4	1.546	9.03	2.66E-03	3.64E-02	Thylakoid lumenal 17.9 kDa protein, chloroplast
CF479107	-13.87	1.587	9.008	2.69E-03	3.68E-02	Chromosome undetermined scaffold_97, genome
TC155986	-11.55	1.526	9.002	2.70E-03	3.69E-02	Predicted protein; Physcomitrella patens subsp.
TC170737	-12.42	1.626	9.001	2.70E-03	3.69E-02	Dof13; Glycine max|Rep: Dof13 - Glycine
GW726055	-12.71	1.601	8.998	2.70E-03	3.69E-02	LEA; Pinus halepensis|Rep: LEA - Pinus
TC165863	-13.53	1.613	8.983	2.73E-03	3.71E-02	Cluster: Chromosome chr18 scaffold_1, genome
CF395134	-13.87	1.682	8.97	2.74E-03	3.73E-02	
TC176110	-12.48	1.617	8.956	2.77E-03	3.75E-02	Galactinol synthase; Brassica napus|Rep: Galactinol
TC177278	-12.22	1.688	8.922	2.82E-03	3.81E-02	Uncharacterized protein At1g72370.2; Arabidopsis |Rep:
TC172034	-12.66	1.617	8.894	2.86E-03	3.85E-02	Whitefly-induced gp91-phox; Solanum lycopersicum|Rep:
TC176879	-13.84	1.508	8.887	2.87E-03	3.86E-02	Chromosome chr13 scaffold_181, genome
GT239748	-13.55	1.445	8.886	2.87E-03	3.86E-02	
TC174107	-13.05	1.637	8.871	2.90E-03	3.88E-02	Chromosome chr19 scaffold_4, genome
TC158217	-13.06	1.62	8.859	2.92E-03	3.90E-02	Chromosome chr17 scaffold_16, genome
TC163042	-13.5	1.524	8.853	2.93E-03	3.90E-02	60S ribosomal protein L6; Vitis vinifera|Rep:
TC196284	-12.29	1.652	8.852	2.93E-03	3.90E-02	Predicted protein; Physcomitrella patens subsp.
TC190872	-12.52	1.598	8.849	2.93E-03	3.90E-02	3-hydroxy-3-methylglutaryl-CoA-synthase
TC159050	-13.51	1.556	8.82	2.98E-03	3.95E-02	Tau class glutathione S-transferase; Pinus tabuliformis|
CR393088	-13.26	1.573	8.797	3.02E-03	4.00E-02	40S ribosomal protein S4; Prunus armeniaca|Rep:
CO166566	-13.68	1.438	8.743	3.11E-03	4.10E-02	Chromosome chr6 scaffold_3, genome
TC182229	-13.82	1.562	8.718	3.15E-03	4.15E-02	Spermidine synthase 2; Pisum sativum|Rep: Spermidine
TC156742	-13.62	1.569	8.717	3.15E-03	4.15E-02	Chromosome chr6 scaffold_25, genome
TC174552	-12.88	1.582	8.713	3.16E-03	4.16E-02	Predicted protein; Physcomitrella patens subsp.
TC189178	-13.26	1.544	8.705	3.17E-03	4.17E-02	Chromosome chr1 scaffold_22, genome

FG614382	-13.94	1.553	8.679	3.22E-03	4.22E-02	scaffold_86, genome Vitis
TC182510	-14.04	1.632	8.677	3.22E-03	4.23E-02	
TC164047	-13.01	1.626	8.669	3.24E-03	4.24E-02	GroEL-like chaperone, ATPase; Medicago truncatula|Rep:
TC165966	-13.12	1.602	8.661	3.25E-03	4.25E-02	2-nitropropane dioxygenase-like protein; Arabidopsis |
TC162217	-13.58	1.422	8.653	3.27E-03	4.25E-02	Chromosome chr7 scaffold_20, genome
TC160260	-13.86	1.579	8.652	3.27E-03	4.25E-02	Chromosome chr5 scaffold_2, genome
TC155371	-13.96	1.593	8.652	3.27E-03	4.25E-02	Chaperonin CPN60-2, mitochondrial precursor; Cucurbita
TC187518	-13.12	1.574	8.65	3.27E-03	4.25E-02	Whitefly-induced gp91-phox; Solanum lycopersicum|Rep:
TC193911	-12.35	1.622	8.602	3.36E-03	4.33E-02	Glutathione peroxidase; Oryza sativa Indica Group|Rep:
TC187917	-14.17	1.935	8.571	3.42E-03	4.39E-02	Chromosome undetermined scaffold_431, genome
TC190115	-13.58	1.416	8.568	3.42E-03	4.39E-02	
TC184679	-12.93	1.595	8.542	3.47E-03	4.44E-02	Cluster: NuM1 protein; Medicago sativa|Rep: NuM1
TC169265	-13.13	1.519	8.534	3.49E-03	4.45E-02	Glutamyl-tRNA reductase; Taxodium distichum var.
TC165523	-12.93	1.608	8.53	3.49E-03	4.46E-02	Chromosome chr4 scaffold_32, genome
TC182676	-13.01	1.582	8.52	3.51E-03	4.48E-02	Cluster: Chromosome chr2 scaffold_112, genome
TC187329	-13.42	1.467	8.485	3.58E-03	4.54E-02	Predicted protein; Physcomitrella patens subsp.
TC189340	-12.9	1.575	8.482	3.59E-03	4.55E-02	Predicted protein; Physcomitrella patens subsp.
TC169171	-11.89	1.663	8.464	3.62E-03	4.59E-02	Cluster: Chromosome chr11 scaffold_14, genome
TC159167	-13.43	1.484	8.456	3.64E-03	4.59E-02	Cluster: RNA helicase; Chlamydomonas reinhardtii|Rep:
TC192076	-13.39	1.493	8.454	3.64E-03	4.59E-02	
TC155537	-12.16	1.64	8.408	3.74E-03	4.70E-02	40S ribosomal protein SA; Glycine max|Rep:
TC192306	-13.61	1.402	8.399	3.75E-03	4.72E-02	Ubiquitin fusion protein; core eudicotyledons|Rep:
GO479184	-13.41	1.68	8.378	3.80E-03	4.76E-02	
TC157487	-12.11	1.635	8.373	3.81E-03	4.77E-02	Chromosome chr14 scaffold_21, genome
TC162989	-12.86	1.551	8.37	3.81E-03	4.77E-02	60S ribosomal protein L7; Solanum
TC190021	-13.64	1.395	8.363	3.83E-03	4.79E-02	Ribosomal L1 domain containing 1; Bos
BF777477	-13.96	1.583	8.353	3.85E-03	4.80E-02	Cluster: Chromosome chr8 scaffold_34, genome
TC180968	-12.1	1.635	8.35	3.86E-03	4.81E-02	40S ribosomal protein SA; Glycine max|Rep:
TC180841	-12.92	1.551	8.309	3.95E-03	4.91E-02	Cytochrome c; Fritillaria agrestis|Rep: Cytochrome c
TC183916	-13.89	1.517	8.297	3.97E-03	4.93E-02	Dehydrin-like protein; Picea abies|Rep: Dehydrin-like
AW290717	-13.15	1.525	8.289	3.99E-03	4.95E-02	Polyubiquitin; Euphorbia characias|Rep: Polyubiquitin -
TC194101	-13.53	1.42	8.279	4.01E-03	4.97E-02	Chromosome chr13 scaffold_181, genome
CR393686	-13.17	1.55	8.274	4.02E-03	4.97E-02	40S ribosomal protein S3a-1; Arabidopsis |Rep:
TC159230	-12.1	1.622	8.258	4.06E-03	5.01E-02	Chromosome chr14 scaffold_21, genome
TC199231	-13.48	1.427	8.244	4.09E-03	5.03E-02	Leucoanthocyanidin dioxygenase; Arabidopsis |Rep:
TC191151	-14.06	1.6	8.244	4.09E-03	5.03E-02	Chromosome chr13 scaffold_181, genome
TC161802	-12.46	1.549	8.184	4.23E-03	5.18E-02	Chromosome chr5 scaffold_2, genome
TC186432	-13.3	1.509	8.172	4.26E-03	5.21E-02	60S ribosomal protein L35; Euphorbia esula|Rep:
TC172335	-13.84	1.451	8.167	4.27E-03	5.21E-02	Cluster: Chromosome chr13 scaffold_17, genome
TC164709	-13.47	1.405	8.165	4.27E-03	5.21E-02	Predicted protein; Physcomitrella patens subsp.
TC161865	-12.72	1.525	8.165	4.27E-03	5.21E-02	Cluster: T30E16.31; Arabidopsis |Rep: T30E16.31 -
GT257842	-13.58	1.469	8.147	4.31E-03	5.24E-02	Chromosome chr12 scaffold_78, genome
TC180838	-13.19	1.564	8.144	4.32E-03	5.24E-02	Uncharacterized protein At3g17860.2; Arabidopsis |Rep:
TC163037	-13.87	1.605	8.144	4.32E-03	5.24E-02	
CF388187	-13.82	1.572	8.141	4.33E-03	5.24E-02	Chromosome chr12 scaffold_78, genome
CD027377	-13.15	1.51	8.136	4.34E-03	5.25E-02	Polyubiquitin; Euphorbia characias|Rep: Polyubiquitin
TC173838	-11.99	1.61	8.127	4.36E-03	5.27E-02	Dehydrin 1; Pinus pinaster|Rep: Dehydrin 1
TC191269	-14.02	1.769	8.125	4.37E-03	5.27E-02	Pentatricopeptide(PPR)repeat protein-like; Oryza
TC176041	-13.09	1.534	8.114	4.39E-03	5.30E-02	Chromosome undetermined scaffold_252, genome
TC188335	-12.22	1.582	8.084	4.47E-03	5.37E-02	Chromosome chr11 scaffold_14, genome
TC194216	-12.18	1.567	8.066	4.51E-03	5.42E-02	Dehydrin 4; Picea abies|Rep: Dehydrin 4
TC160127	-13.14	1.496	8.062	4.52E-03	5.42E-02	Chromosome chr6 scaffold_3, genome
TC195335	-13.28	1.495	8.06	4.53E-03	5.43E-02	60S acidic ribosomal protein P1-alpha 3;
BX682581	-13.19	1.49	8.049	4.55E-03	5.45E-02	Thioredoxin H-type; Picea mariana
TC186724	-10.57	0.635	8.04	4.58E-03	5.48E-02	Predicted protein; Physcomitrella patens subsp.
TC178006	-13.97	1.679	8.029	4.60E-03	5.50E-02	Mevalonate kinase; Borrelia burgdorferi
BG832530	-13.88	1.613	8.021	4.62E-03	5.52E-02	Chromosome chr14 scaffold_27, genome
TC180577	-13.03	1.535	8.016	4.64E-03	5.53E-02	Glutathione peroxidase; Oryza sativa Indica Group|Rep:
TC166524	-12.64	1.489	7.938	4.84E-03	5.75E-02	Dehydrin 1; Pinus pinaster|Rep: Dehydrin 1
TC189785	-12.26	1.564	7.935	4.85E-03	5.75E-02	Predicted protein; Physcomitrella patens subsp.
TC191236	-13.57	1.372	7.931	4.86E-03	5.76E-02	Chromosome chr2 scaffold_112, genome
TC190616	-13.58	1.373	7.927	4.87E-03	5.76E-02	Cluster: Co-chaperone-curved DNA binding protein A;
TC186035	-13.53	1.369	7.914	4.91E-03	5.79E-02	Aldehyde dehydrogenase; Oryza sativa
TC158704	-12.68	1.478	7.884	4.99E-03	5.87E-02	Heat shock protein 70A; Chlamydomonas reinhardtii|Rep:
CX646983	-13.73	1.368	7.88	5.00E-03	5.88E-02	Chromosome chr11 scaffold_118, genome
TC180742	-13.68	1.373	7.869	5.03E-03	5.90E-02	60S ribosomal protein L35a-3; Arabidopsis |Rep:
TC177508	-13.41	1.416	7.864	5.04E-03	5.91E-02	40S ribosomal protein S19-1; Arabidopsis |Rep:
TC179102	-14.06	1.552	7.854	5.07E-03	5.93E-02	Chromosome chr13 scaffold_181, genome
TC154942	-12.66	1.474	7.854	5.07E-03	5.93E-02	LEA; Pinus halepensis|Rep: LEA - Pinus
BE582385	-13.99	1.539	7.853	5.07E-03	5.93E-02	Malate dehydrogenase; Vitis vinifera
TC179863	-13.29	1.441	7.846	5.09E-03	5.95E-02	Cluster: At5g04080; Arabidopsis |Rep: At5g04080 -
TC164434	-13.53	1.355	7.822	5.16E-03	6.02E-02	Ribosomal L1 domain containing 1; Bos
GW726071	-13.75	1.362	7.815	5.18E-03	6.03E-02	Predicted protein; Physcomitrella patens subsp.
TC186475	-13.06	1.509	7.79	5.25E-03	6.08E-02	Acidic ribosomal protein; Triticum aestivum|Rep: Acidic
TC155601	-12.65	1.463	7.757	5.35E-03	6.18E-02	LEA; Pinus halepensis|Rep: LEA - Pinus
TC195787	-13.6	1.343	7.753	5.36E-03	6.19E-02	Chromosome chr11 scaffold_13, genome
TC173382	-13.37	1.427	7.74	5.40E-03	6.21E-02	Chromosome chr13 scaffold_48, genome
TC175523	-13.79	1.372	7.731	5.43E-03	6.23E-02	scaffold_2, genome Vitis
TC196113	-12.33	1.502	7.714	5.48E-03	6.27E-02	Phaseolin G-box binding protein PG1; Phaseolus
TC164466	-14.06	1.554	7.71	5.49E-03	6.28E-02	Cluster: Chromosome chr9 scaffold_33, genome
TC186095	-12.8	1.47	7.69	5.55E-03	6.33E-02	Chromosome chr14 scaffold_21, genome

TC177615	-13.11	1.494	7.686	5.57E-03	6.34E-02	Predicted protein; Physcomitrella patens subsp.
DR070405	-12.04	1.554	7.652	5.67E-03	6.45E-02	Thioredoxin H-type; Picea mariana
TC187138	-13.47	1.362	7.626	5.75E-03	6.52E-02	60S ribosomal protein L30; Lupinus luteus|Rep:
TC176383	-13.34	1.439	7.623	5.76E-03	6.53E-02	Predicted protein; Physcomitrella patens subsp.
TC197802	-13.68	1.327	7.618	5.78E-03	6.54E-02	Ribosomal protein L30e; Pisum sativum|Rep: Ribosomal
TC190945	-12.1	1.538	7.618	5.78E-03	6.54E-02	Dehydrin 1; Pinus pinaster|Rep: Dehydrin 1
TC187846	-12.71	1.45	7.594	5.86E-03	6.61E-02	LEA; Pinus halepensis|Rep: LEA - Pinus
TC191086	-13.53	1.346	7.582	5.89E-03	6.65E-02	40S ribosomal protein S19-1; Arabidopsis |Rep:
TC156149	-13.18	1.456	7.58	5.90E-03	6.65E-02	Chromosome chr11 scaffold_118, genome
TC154938	-13.17	1.459	7.572	5.93E-03	6.66E-02	Chromosome undetermined scaffold_79, genome
TC172781	-12.8	1.463	7.571	5.93E-03	6.66E-02	Ubiquitin extension protein; core eudicotyledons
TC158058	-13.95	1.54	7.568	5.94E-03	6.66E-02	Chromosome chr5 scaffold_2, genome
TC196347	-12.99	1.484	7.566	5.95E-03	6.66E-02	F26F24.20; Arabidopsis |Rep: F26F24.20 - Arabidopsis
TC195319	-13.43	1.366	7.566	5.95E-03	6.66E-02	Chromosome chr14 scaffold_21, genome
TC175506	-12.43	1.482	7.565	5.95E-03	6.66E-02	Cluster: 60S acidic ribosomal protein P2-4;
TC168424	-13.21	1.454	7.56	5.97E-03	6.67E-02	RNA recognition motif family protein, expressed;
TC197786	-13.7	1.319	7.531	6.06E-03	6.75E-02	40S ribosomal protein S17; Solanum lycopersicum|Rep:
TC165644	-13.22	1.436	7.519	6.10E-03	6.78E-02	Cluster: Chromosome chr13 scaffold_17, genome
TC171789	-12.54	1.448	7.512	6.13E-03	6.79E-02	60S ribosomal protein L15-2; n=7; Picea|Rep: 60S
TC161116	-13.18	1.538	7.508	6.14E-03	6.80E-02	LP3-1; Pinus|Rep: LP3-1 - Pinus taeda
TC175818	-13.76	1.423	7.499	6.17E-03	6.82E-02	
TC171914	-12.47	1.467	7.49	6.20E-03	6.84E-02	40S ribosomal protein SA; Daucus carota|Rep:
DR160811	-14.13	1.631	7.485	6.22E-03	6.85E-02	Chromosome chr6 scaffold_25, genome
TC156554	-13.98	1.477	7.48	6.24E-03	6.86E-02	Similarity to RNA-binding protein; Arabidopsis |Rep:
TC198415	-12.97	1.461	7.457	6.32E-03	6.94E-02	Ribosomal protein S27; Vitis vinifera|Rep: Ribosomal
BX250897	-12.67	1.433	7.451	6.34E-03	6.95E-02	Dehydrin-like protein; Picea abies Dehydrin-like protein
TC188172	-11.93	1.523	7.442	6.37E-03	6.98E-02	Thioredoxin H-type; Picea mariana
TC154592	-12.97	1.463	7.416	6.46E-03	7.05E-02	Chromosome chr1 scaffold_22, genome
DR069716	-13.81	1.398	7.416	6.46E-03	7.05E-02	Chromosome chr18 scaffold_1, genome
TC155183	-13.91	1.558	7.407	6.50E-03	7.07E-02	NADH-ubiquinone oxidoreductase chain 2;Lucasium damaeum
TC183427	-13.78	1.412	7.402	6.52E-03	7.08E-02	Cluster: At5g04080; Arabidopsis |Rep: At5g04080 -
TC155595	-12.94	1.461	7.39	6.56E-03	7.11E-02	Chaperonin CPN60, mitochondrial precursor; Brassicaceae
AW985297	-13.73	1.312	7.367	6.64E-03	7.20E-02	60S ribosomal protein L15-2; n=7; Picea|Rep: 60S
TC155435	-11.85	1.473	7.355	6.69E-03	7.23E-02	3-hydroxy-3-methylglutaryl-CoA-synthase
CT580291	-12.62	1.537	7.344	6.73E-03	7.26E-02	LP3-1; Pinus|Rep: LP3-1 - Pinus taeda
TC171827	-12.07	1.506	7.329	6.79E-03	7.29E-02	Chromosome chr12 scaffold_47, genome
TC173372	-13.04	1.451	7.3	6.90E-03	7.39E-02	Predicted protein; Physcomitrella patens subsp.
TC194171	-13.15	1.445	7.297	6.91E-03	7.39E-02	Predicted protein; Physcomitrella patens subsp.
TC161764	-13.33	1.338	7.29	6.93E-03	7.41E-02	Cell-wall invertase; Coffea canephora
TC170670	-12.65	1.411	7.272	7.00E-03	7.46E-02	Chromosome chr14 scaffold_26, genome
GW763011	-13.43	1.355	7.261	7.05E-03	7.49E-02	Chromosome chr6 scaffold_3, genome
TC172965	-13.66	1.334	7.247	7.10E-03	7.54E-02	Chromosome chr5 scaffold_2, genome
TC184186	-13.78	1.369	7.241	7.13E-03	7.56E-02	Similarity to protein-tyrosine phosphatase;Arabidopsis
TC172950	-13.61	1.287	7.229	7.18E-03	7.59E-02	LP3-1; Pinus|Rep: LP3-1 - Pinus taeda
TC175323	-12.89	1.434	7.226	7.18E-03	7.59E-02	60S acidic ribosomal protein; Hyacinthus orientalis|Rep:
CO366299	-13.89	1.412	7.22	7.21E-03	7.61E-02	Chromosome undetermined scaffold_55, genome
TC197002	-12.83	1.476	7.199	7.29E-03	7.68E-02	Chromosome chr8 scaffold_34, genome
AW043333	-13.51	1.294	7.195	7.31E-03	7.68E-02	Respiratory burst oxidase 2; Medicago truncatula|Rep:
TC157405	-13.12	1.404	7.194	7.31E-03	7.68E-02	Elicitor-responsive protein 3; Oryza sativa
TC180218	-13.25	1.418	7.191	7.33E-03	7.68E-02	RNA helicase; Vigna radiata var. radiata|Rep:
CF393854	-13.89	1.433	7.185	7.35E-03	7.70E-02	Chromosome chr5 scaffold_124, genome
TC194993	-13.59	1.277	7.172	7.40E-03	7.74E-02	Cluster: Chromosome chr9 scaffold_104, genome
TC188664	-13.15	1.438	7.165	7.43E-03	7.77E-02	60S ribosomal protein L27a-3; Arabidopsis |Rep:
TC173534	-13.12	1.413	7.164	7.44E-03	7.77E-02	Hypoxia-responsive family protein; Citrus sinensis
BI398028	-13.78	1.313	7.157	7.47E-03	7.77E-02	60S acidic ribosomal protein; Hyacinthus orientalis|Rep:
AW043385	-13.58	1.297	7.156	7.47E-03	7.77E-02	scaffold_86, genome Vitis
TC191675	-13.77	1.295	7.152	7.49E-03	7.79E-02	Predicted protein; Physcomitrella patens subsp.
TC197022	-13.67	1.28	7.132	7.57E-03	7.86E-02	Predicted protein; Physcomitrella patens subsp.
TC179978	-12.64	1.451	7.115	7.65E-03	7.91E-02	Chromosome chr6 scaffold_3, genome
TC173803	-13.06	1.429	7.106	7.69E-03	7.95E-02	Predicted protein; Physcomitrella patens subsp.
TC157972	-14.06	1.723	7.091	7.75E-03	8.00E-02	Pinus banksiana|Rep: Alcohol dehydrogenase - Pinus
TC166159	-13.3	1.392	7.085	7.77E-03	8.01E-02	Chromosome chr6 scaffold_3, genome
TC181102	-12.72	1.392	7.069	7.84E-03	8.06E-02	RPS3; Cymbidium hybrid cultivar|Rep: RPS3 -
TC184494	-12.68	1.39	7.064	7.86E-03	8.07E-02	LEA; Pinus halepensis|Rep: LEA - Pinus
TC193212	-13.06	1.42	7.053	7.91E-03	8.12E-02	Thioredoxin H-type; Picea mariana
TC155875	-12.29	1.468	7.035	7.99E-03	8.17E-02	Chromosome chr5 scaffold_2, genome
TC156777	-12.97	1.42	7.033	8.00E-03	8.17E-02	Dehydrin 4; Picea abies|Rep: Dehydrin 4
FE519797	-11.71	1.451	7.032	8.01E-03	8.17E-02	S-adenosyl methionine synthetase 2; Pinus taeda|Rep:
TC166201	-13.75	1.296	7.029	8.02E-03	8.17E-02	Predicted protein; Physcomitrella patens subsp.
TC159840	-12.94	1.437	7.018	8.07E-03	8.20E-02	Spermidine synthase; Ammopiptanthus mongolicus
TC171765	-11.72	1.413	7.01	8.11E-03	8.22E-02	Chromosome chr17 scaffold_16, genome
BE662571	-13.75	1.294	7.006	8.12E-03	8.23E-02	Ribosomal protein L19; Hyacinthus orientalis
BE643805	-12.61	1.384	7.004	8.13E-03	8.24E-02	Thioredoxin H-type; Picea mariana
BG318155	-13.16	1.361	6.997	8.16E-03	8.25E-02	ARF/SAR superfamily protein;Dictyostelium discoideum
TC191799	-9.5	1.096	6.982	8.24E-03	8.30E-02	Adenosylhomocysteinase; Nicotiana sylvestris
GT267726	-13.6	1.261	6.961	8.33E-03	8.37E-02	Predicted protein; Physcomitrella patens subsp.
CR392230	-13.27	1.381	6.96	8.34E-03	8.37E-02	Chromosome chr14 scaffold_27, genome
CX649468	-11.18	1.469	6.952	8.37E-03	8.40E-02	
CX645420	-13.97	1.403	6.94	8.43E-03	8.45E-02	Chromosome undetermined scaffold_259, genome
TC156857	-13.7	1.383	6.933	8.46E-03	8.46E-02	SJCHGC01974 protein; Schistosoma japonicum
TC176631	-13.16	1.413	6.933	8.46E-03	8.46E-02	Ubiquitin extension protein; core eudicotyledons

BF609150	-13.33	1.361	6.923	8.51E-03	8.48E-02	6-phosphogluconate dehydrogenase,decarboxylating;Vitis
TC162465	-13.87	1.331	6.921	8.52E-03	8.48E-02	Nucleosome assembly protein family; Physcomitrella patens
TC197586	-14.33	1.815	6.907	8.59E-03	8.53E-02	Cysteine synthase; Malus x domestica|Rep: Cysteine
TC186941	-13.23	1.388	6.887	8.68E-03	8.62E-02	Predicted protein; Physcomitrella patens subsp.
AW754647	-12.73	1.43	6.877	8.73E-03	8.66E-02	
TC167738	-13.6	1.256	6.87	8.77E-03	8.68E-02	Chromosome chr13 scaffold_48, genome
TC168602	-13.49	1.304	6.869	8.77E-03	8.68E-02	Small molecular heat shock protein 10;
TC158198	-12.63	1.37	6.847	8.88E-03	8.77E-02	Predicted protein; Physcomitrella patens subsp.
TC185508	-13.88	1.375	6.844	8.89E-03	8.78E-02	Chromosome chr7 scaffold_20, genome
CF473121	-14.15	1.507	6.82	9.02E-03	8.87E-02	At5g04080; Arabidopsis
TC196994	-13.07	1.407	6.817	9.03E-03	8.87E-02	Chromosome chr5 scaffold_2, genome
DR091272	-13.75	1.368	6.808	9.07E-03	8.89E-02	Subtilisin-like protein; Picea abies
TC168605	-13.09	1.398	6.796	9.14E-03	8.93E-02	60S ribosomal protein L15-2; n=7; Picea|Rep: 60S
TC190157	-12.39	1.404	6.795	9.14E-03	8.93E-02	60S ribosomal protein L15-2; n=7; Picea|Rep: 60S
TC190131	-11.71	1.384	6.786	9.19E-03	8.95E-02	Type IIIa membrane protein cp-wap13; Vigna
TC168560	-12.34	1.415	6.783	9.20E-03	8.96E-02	60S ribosomal protein L15-2; n=7; Picea|Rep: 60S
GW726206	-13.05	1.391	6.774	9.25E-03	8.99E-02	LEA; Pinus halepensis|Rep: LEA - Pinus
TC180723	-13.07	1.489	6.747	9.39E-03	9.09E-02	LP3-1; Pinus|Rep: LP3-1 - Pinus taeda
CX714488	-13.22	1.391	6.694	9.67E-03	9.33E-02	GroEL-like chaperone, ATPase; Medicago truncatula
TC196904	-12.92	1.385	6.663	9.84E-03	9.46E-02	BiP-isoform D; Glycine max|Rep: BiP-isoform D
BI416775	-12.74	1.373	6.657	9.88E-03	9.48E-02	Hydroxyproline-rich glycoprotein precursor; Volvox
TC188383	-13.38	1.276	6.653	9.90E-03	9.49E-02	Pollen allergen CJP38; Cryptomeria japonica|Rep: Pollen
TC176684	-12.95	1.42	6.651	9.91E-03	9.49E-02	Chromosome chr5 scaffold_124, genome
TC180407	-13.83	1.308	6.649	9.92E-03	9.49E-02	Chromosome chr6 scaffold_3, genome sequence
TC183197	-13.31	1.273	6.648	9.93E-03	9.49E-02	Trehalose-6-phosphate synthase; Ginkgo biloba
TC192542	-13.03	1.36	6.646	9.94E-03	9.49E-02	GDA2 protein; Pisum sativum|Rep: GDA2 protein
TC158161	-13.02	1.378	6.645	9.94E-03	9.49E-02	Predicted protein; Physcomitrella patens subsp.

APPENDIX 2:

List of down-regulated EST tags in Scots pine RNA-Seq wounding experiment:

TAGS	LogConc.	LogFC	LR	P.Value	FDR	ANNOTATIONS
TC199290	-10.397	-0.621	71.72	2.47E-17	3.18E-13	NAD(P)H dehydrogenase;Ralstonia pickettii dehydrogenase;
TC178048	-13.135	-2.916	50.09	1.47E-12	1.71E-09	Transcriptional regulator;Stigmatella aurantiaca
TC190883	-12.963	-2.933	44.55	2.48E-11	1.10E-08	Cluster: Metallothionein-like protein;Arabidopsis
TC191810	-12.962	-2.768	43.24	4.85E-11	1.84E-08	Low molecular weight heat-shock protein;
TC174422	-12.502	-2.96	42.61	6.67E-11	2.38E-08	
TC160053	-12.14	-2.952	38.87	4.54E-10	8.71E-08	CYPA;Pinus taeda; CytochromeP450 CYPA
TC156269	-12.891	-2.878	38.05	6.89E-10	1.17E-07	Cluster: Chromosome chr1 scaffold_5, whole Chromosome
TC195576	-13.066	-2.668	38.03	6.98E-10	1.17E-07	Cytochrome P450 CYPA;Pinus P450 taeda; Cytochrome
TC193124	-12.795	-2.768	37.84	7.70E-10	1.25E-07	
BQ696511	-12.523	-2.992	37.59	8.72E-10	1.38E-07	Cytochrome P450 CYPA;Pinus P450 taeda; Cytochrome
TC185698	-13.527	-2.595	35.22	2.94E-09	3.71E-07	Cluster: Metallothionein-like protein;Arabidopsis
TC195238	-12.179	-2.607	34.85	3.56E-09	4.32E-07	Uncharacterized protein At5g19140.2;Arabidopsis protein
TC166978	-12.11	-2.607	33.95	5.67E-09	6.57E-07	Cluster: Uncharacterized protein
TC167447	-11.289	-2.512	32.76	1.04E-08	1.11E-06	Cluster: Uncharacterized protein
DR089624	-12.06	-2.565	32.7	1.07E-08	1.12E-06	Cluster: Uncharacterized protein
TC187546	-11.491	-2.717	32.54	1.17E-08	1.20E-06	Cluster: Integrase;uncultured bacterium; Integrase;
TC172136	-11.769	-2.821	32.51	1.19E-08	1.21E-06	Cytochrome P450 CYPA;Pinus P450 taeda; Cytochrome
FG615045	-12.053	-2.547	32.4	1.25E-08	1.25E-06	Dormancy-associated protein;Codonopsis lanceolata;
DR746324	-12.21	-2.498	32.29	1.33E-08	1.32E-06	Cluster: Genomic DNA, chromosome 3, Genomic P1 clone:
DR078367	-13.072	-2.298	30.18	3.94E-08	3.29E-06	Uncharacterized protein At5g19140.2;Arabidopsis protein
TC172311	-9.156	-1.891	28.77	8.13E-08	6.16E-06	during drought stress, clone Pp.ap18 drought
DR057561	-10.291	-1.892	28.77	8.14E-08	6.16E-06	Probable multidrug resistance transmembrane protein;
CV032775	-9.656	-1.889	28.7	8.45E-08	6.32E-06	
TC156705	-13.757	-2.873	28.62	8.79E-08	6.39E-06	Cytochrome c oxidase subunit I; cMoneilema
TC159763	-9.282	-1.874	28.23	1.08E-07	7.60E-06	Isoform 3 of Q8R2K1 ; 3Mus
TC195725	-13.488	-2.246	27.96	1.24E-07	8.58E-06	70 kDa heat shock protein; kDaArabidopsis
TC165847	-13.733	-2.825	27.96	1.24E-07	8.58E-06	
TC177293	-13.745	-2.602	27.42	1.64E-07	1.08E-05	Cluster: Metallothionein-like protein;Arabidopsis
TC182917	-12.511	-2.352	27.32	1.73E-07	1.11E-05	Low molecular weight heat-shock protein;
TC180248	-10.123	-1.844	27.29	1.75E-07	1.12E-05	Cluster: Ribosomal protein L35;Ribosomal Hydra vulgaris;
TC197374	-9.801	-1.842	27.22	1.82E-07	1.15E-05	3-isopropylmalate dehydratase small subunit;dehydratase
TC176999	-12.701	-2.34	27.19	1.85E-07	1.16E-05	Dormancy-associated protein;Codonopsis lanceolata;
DR078762	-10.001	-1.841	27.16	1.88E-07	1.17E-05	Glutathione synthase;Sodalis glossinidius synthase; str.
TC186072	-13.142	-2.04	24.42	7.73E-07	4.30E-05	Chromosome chr9 scaffold_7, whole genome chr9 sequence;
BX784377	-11.423	-2.507	24.29	8.29E-07	4.44E-05	Villin 2;Nicotiana tabacum; 2; Villin 2
TC168951	-13.504	-2.037	24.18	8.76E-07	4.66E-05	
TC158215	-13.527	-1.984	23.73	1.11E-06	5.61E-05	GDA2 protein;Pisum sativum; protein; GDA2 protein
TC156038	-13.023	-2.149	23.35	1.35E-06	6.57E-05	Glucose-1-phosphate adenylyltransferase;Citrullus
TC185372	-13.738	-2.221	23.28	1.40E-06	6.75E-05	Cluster: Potassium channel KAT1;Potassium Arabidopsis
CK909821	-12.425	-2.186	23.2	1.46E-06	6.98E-05	
TC157709	-13.17	-2.009	22.88	1.72E-06	8.05E-05	GDA2 protein;Pisum sativum; protein; GDA2 protein
CO165393	-13.266	-1.961	22.83	1.77E-06	8.17E-05	Cluster: Plasma membrane associated protein-like;
TC162294	-13.914	-2.32	22.8	1.80E-06	8.26E-05	Cluster: Chromosome chr6 scaffold_25, whole Chromosome
TC172461	-13.589	-2.023	22.49	2.11E-06	9.58E-05	Chromosome chr9 scaffold_7, whole genome chr9 sequence;
TC173020	-12.61	-2.234	22.36	2.26E-06	1.02E-04	Laccase;Pinus taeda; Laccase- Pinus
TC179274	-12.528	-2.118	22.17	2.49E-06	1.11E-04	Cluster: Os02g0211200 protein;Oryza Os02g0211200 sativa
TC157156	-12.272	-2.109	21.97	2.78E-06	1.22E-04	HSP70;Triticum aestivum; HSP70- Triticum
TC196878	-13.241	-1.926	21.66	3.26E-06	1.41E-04	Cluster: Plasma membrane associated protein-like;
TC158923	-13.198	-1.944	21.55	3.44E-06	1.47E-04	Cluster: Plasma membrane associated protein-like;
TC182239	-13.443	-1.976	21.32	3.88E-06	1.64E-04	Late embryogenesis abundant protein;embryogenesis Picea;
TC187998	-13.117	-1.968	21.25	4.04E-06	1.69E-04	Trehalose-6-phosphate synthase;Ginkgo biloba; synthase;
TC181016	-13.858	-2.121	21.06	4.45E-06	1.82E-04	Nodulin 21;Glycine max; 21; Nodulin 21
TC154733	-13.117	-2.149	20.72	5.31E-06	2.13E-04	Type 2 1-hydroxy-2-methyl-2-(E)-butenyl-4-phosphate
TC187638	-12.567	-2.064	20.67	5.45E-06	2.18E-04	Low molecular weight heat-shock protein;
TC176417	-13.136	-1.954	20.64	5.55E-06	2.20E-04	Chlorophyll a/b-binding protein precursor;a/b-binding
TC175441	-13.22	-2.036	20.41	6.25E-06	2.46E-04	Geranyl diphosphate synthase;Abies diphosphate grandis;
TC170301	-12.961	-1.999	20.36	6.43E-06	2.52E-04	GDA2 protein;Pisum sativum; protein; GDA2 protein
TC190346	-12.664	-2.049	20.31	6.59E-06	2.58E-04	Low molecular weight heat-shock protein;
TC184935	-13.072	-2.008	20.16	7.13E-06	2.76E-04	101 kDa heat shock protein; kDaNicotiana
TC185310	-12.902	-2.037	19.94	7.98E-06	3.06E-04	Chromosome chr19 scaffold_4, whole genome chr19
TC154831	-13.412	-1.857	19.82	8.53E-06	3.24E-04	Gb|AAF02142.1;Arabidopsis thaliana; Gb|AAF02142.1-
TC164987	-13.875	-2.203	19.81	8.54E-06	3.24E-04	Dormancy-associated protein;Arabidopsis thaliana;
TC162804	-13.624	-1.968	19.8	8.58E-06	3.25E-04	Cluster: Receptor protein kinase-like protein;
TC179722	-13.495	-1.898	19.75	8.81E-06	3.28E-04	Late embryogenesis abundant protein;embryogenesis Picea;
GW771587	-13.908	-2.206	19.56	9.75E-06	3.59E-04	Cluster: Inward potassium channel alpha Inward subunit;
CO200270	-13.401	-1.841	19.49	1.01E-05	3.71E-04	Late embryogenesis abundant protein;embryogenesis Picea;
TC194759	-11.563	-2.064	19.38	1.07E-05	3.92E-04	Chromosome chr5 scaffold_67, whole genome chr5 sequence;
TC199135	-13.547	-1.858	19.37	1.08E-05	3.92E-04	
TC158360	-12.959	-1.961	19.36	1.08E-05	3.93E-04	Snakin-like cysteine rich protein;cysteine Phaseolus
DR078366	-13.723	-1.938	19.22	1.16E-05	4.17E-04	Chromosome chr18 scaffold_61, whole genome chr18
TC181911	-13.618	-1.925	19.21	1.17E-05	4.19E-04	70 kDa heat shock cognate kDa protein 3;
TC191505	-12.851	-1.956	19	1.31E-05	4.63E-04	Cluster: Plasma membrane associated protein-like;
TC191475	-13.589	-1.933	18.97	1.33E-05	4.69E-04	Chromosome undetermined scaffold_174, whole genome
TC196964	-13.338	-1.767	18.91	1.37E-05	4.80E-04	
TC193042	-13.668	-2.008	18.85	1.42E-05	4.95E-04	HSP70;Triticum aestivum; HSP70- Triticum
TC169331	-13.902	-2.209	18.77	1.48E-05	5.14E-04	Heat shock 70 kDa protein; shockGlycine
TC155419	-12.77	-2.055	18.74	1.50E-05	5.20E-04	101 kDa heat shock protein; kDaNicotiana
TC180353	-13.368	-1.852	18.61	1.61E-05	5.49E-04	Thiazole biosynthetic enzyme;Pseudotsuga biosynthetic
TC180716	-11.453	-2.051	18.42	1.78E-05	5.95E-04	Heat shock protein HSP101;shock Zea mays;
TC196477	-13.576	-1.901	18.33	1.86E-05	6.13E-04	Chloroplast glyceraldehyde-3-phosphate dehydrogenase
TC155029	-12.751	-1.922	18.27	1.92E-05	6.32E-04	Phenylpropenal double-bond reductase;Pinus double-bond
DR093944	-13.205	-1.99	18.17	2.02E-05	6.56E-04	Chromosome chr4 scaffold_32, whole genome chr4 sequence;
TC199342	-13.324	-1.762	18.14	2.05E-05	6.65E-04	Cluster: Protein kinase-like protein;Protein Arabidopsis
TC161088	-12.546	-1.908	18.08	2.12E-05	6.82E-04	Cluster: Chromosome undetermined scaffold_186, whole
TC155407	-11.424	-2.173	18.03	2.18E-05	6.98E-04	Pinus taeda; (+)-alpha-pinene synthase Pinus - Pinus
TC156438	-12.918	-1.906	17.96	2.26E-05	7.19E-04	Phenylpropenal double-bond reductase;Pinus double-bond

TC197407	-12.737	-2.381	17.92	2.31E-05	7.33E-04	(-)-alpha-pinene synthase;Pinus taeda; synthase; (-)-
TC156489	-11.187	-1.945	17.88	2.35E-05	7.44E-04	Pinus sylvestris; L-asparaginase - Pinus sylvestris;
TC175164	-13.463	-1.731	17.78	2.48E-05	7.79E-04	Pinus banksiana; Alcohol dehydrogenase Pinus - Pinus
TC156698	-12.294	-1.954	17.77	2.49E-05	7.79E-04	Heat shock protein;Glycine shock max; Heat
TC163792	-13.963	-2.121	17.76	2.50E-05	7.79E-04	COG0845: Membrane-fusion protein;Yersinia Membrane-fusion
GW725865	-13.975	-2.298	17.72	2.57E-05	7.95E-04	Late embryogenesis abundant protein;embryogenesis Picea;
GT265035	-12.543	-2.11	17.6	2.72E-05	8.35E-04	(+)-alpha-pinene synthase;Pinus taeda; synthase; (+)
TC172390	-13.48	-1.927	17.5	2.87E-05	8.69E-04	Os10g0459400 protein;Oryza sativa protein; Japonica
TC194293	-13.529	-1.716	17.37	3.08E-05	9.22E-04	Probable histone H2A.2;Oryza histone sativa; Probable
FR730409	-12.299	-1.839	17.18	3.40E-05	1.01E-03	Short-chain dehydrogenase/reductase SDR precursor;
TC170338	-12.45	-2.043	17.15	3.45E-05	1.02E-03	Geranyl diphosphate synthase;Abies diphosphate grandis;
TC174423	-13.344	-1.707	17.12	3.52E-05	1.04E-03	Cluster: Chromosome chr11 scaffold_56, whole Chromosome
AI725242	-13.3	-1.736	17.11	3.53E-05	1.04E-03	Late embryogenesis abundant protein;embryogenesis Picea;
FG616357	-13.17	-2.342	17.1	3.54E-05	1.04E-03	(-)-alpha-pinene synthase;Pinus taeda; synthase;
TC163983	-13.328	-1.706	16.95	3.83E-05	1.12E-03	Heat shock protein 81-1;shock Arabidopsis thaliana;
TC195386	-13.383	-1.7	16.94	3.85E-05	1.12E-03	Chromosome undetermined scaffold_91, whole genome
TC166245	-12.623	-2.027	16.77	4.23E-05	1.23E-03	1-deoxy-D-xylulose 5-phosphate synthase type II; 5-ph
DR018625	-12.866	-2.01	16.52	4.81E-05	1.38E-03	1-deoxy-D-xylulose 5-phosphate synthase type II; 5-ph
TC157654	-12.41	-1.985	16.21	5.66E-05	1.59E-03	1-deoxy-D-xylulose 5-phosphate synthase type II; 5-ph
TC171021	-13.038	-1.774	16.15	5.85E-05	1.63E-03	Heat shock protein;Hevea shock brasiliensis; Heat
TC159454	-13.045	-1.894	16.13	5.92E-05	1.64E-03	Cluster: Inward-rectifying K+ channel;Inward-rectifying
TC171759	-12.03	-1.851	16.13	5.93E-05	1.64E-03	Small heat-shock protein;Pseudotsuga heat-shock
GT245002	-13.781	-1.832	16.08	6.08E-05	1.68E-03	Cluster: Chromosome chr11 scaffold_56, whole Chromosome
TC169724	-13.577	-1.907	15.96	6.46E-05	1.77E-03	Os10g0459400 protein;Oryza sativa protein; Japonica
TC168739	-11.646	-2.109	15.79	7.09E-05	1.91E-03	1-hydroxy-2-methyl-2-(E)-butenyl-4-phosphate reductase;
TC166290	-13.865	-1.877	15.77	7.15E-05	1.92E-03	Chromosome chr7 scaffold_20, whole genome chr7 sequence;
TC176960	-13.625	-1.724	15.73	7.32E-05	1.95E-03	Octicosapeptide/Phox/Bem1p;Medicago truncatula;
TC172451	-12.947	-1.856	15.62	7.74E-05	2.05E-03	Thiazole biosynthetic enzyme;Pseudotsuga biosynthetic
TC155384	-12.263	-2.296	15.61	7.77E-05	2.05E-03	Pinus taeda; (-)-alpha-pinene synthase Pinus - Pinus
TC178580	-13.654	-1.769	15.52	8.18E-05	2.14E-03	At2g47485;Arabidopsis thaliana; At2g47485- Arabidopsis
TC181231	-13.32	-1.626	15.45	8.46E-05	2.21E-03	Zeta-carotene desaturase,chloroplast/chromoplast precurso
TC162341	-13.554	-1.604	15.37	8.82E-05	2.30E-03	Cluster: GABA-A epsilon subunit splice GABA-A variant;
TC180513	-13.292	-1.623	15.34	8.99E-05	2.33E-03	Nuclear transcription factor Y subunit transcription
TC165456	-14.11	-2.263	15.28	9.28E-05	2.40E-03	Predicted protein;Physcomitrella patens protein; subsp.
DR182086	-11.487	-1.756	15.26	9.39E-05	2.43E-03	
TC155699	-13.922	-1.951	15.17	9.81E-05	2.52E-03	
TC157755	-13.053	-1.861	15.14	9.97E-05	2.55E-03	CYPB;Pinus taeda; CytochromeP450 CYPB
TC177970	-13.964	-1.913	15.07	1.04E-04	2.63E-03	COG0845: Membrane-fusion protein;Yersinia Membrane-fusion
TC162493	-13.17	-1.781	15.03	1.06E-04	2.67E-03	Chromosome undetermined scaffold_52, whole genome
TC190178	-13.061	-1.68	14.81	1.19E-04	2.95E-03	T6J4.14 protein;Arabidopsis thaliana; protein; T6J4.14
TC194727	-12.792	-1.755	14.73	1.24E-04	3.07E-03	Chromosome undetermined scaffold_30, whole genome
TC193991	-12.706	-1.842	14.72	1.25E-04	3.08E-03	Heat shock protein;Glycine shock max; Heat
TC196327	-13.578	-1.574	14.67	1.28E-04	3.14E-03	Major ampullate spidroin 2-2;ampullate Kukulcania
TC155115	-13.228	-1.64	14.67	1.28E-04	3.14E-03	Pinus radiata; Endo-beta-1,4-glucanase - Pinus radiata;
BM492497	-12.866	-1.761	14.65	1.30E-04	3.17E-03	
TC185003	-13.284	-1.611	14.63	1.31E-04	3.19E-03	F28K19.24;Arabidopsis thaliana; F28K19.24- Arabidopsis
GW772842	-13.5	-1.552	14.5	1.40E-04	3.39E-03	Chromosome chr1 scaffold_5, whole genome chr1 sequence;
TC182802	-13.257	-1.603	14.45	1.44E-04	3.48E-03	Late embryogenesis abundant protein;embryogenesis Picea;
TC163025	-12.387	-1.76	14.41	1.47E-04	3.53E-03	Chromosome chr1 scaffold_5, whole genome chr1 sequence;
TC159501	-13.484	-1.746	14.23	1.62E-04	3.87E-03	Predicted protein;Physcomitrella patens protein; subsp.
TC158390	-13.122	-1.652	14.22	1.63E-04	3.88E-03	Chromosome chr13 scaffold_286, whole genome chr13
TC187205	-13.456	-1.548	14.21	1.64E-04	3.89E-03	Cluster: Chromosome undetermined scaffold_52, whole
TC177232	-12.518	-1.81	14.16	1.68E-04	3.97E-03	Pinus taeda; (+)-alpha-pinene synthase Pinus - Pinus
TC187060	-13.168	-1.602	14.16	1.68E-04	3.97E-03	Uncharacterized protein At1g56220.2;Arabidopsis protein
TC184239	-13.717	-1.69	14.15	1.69E-04	3.99E-03	9 protein homolog;Mus protein musculus; Retinitis
TC165388	-13.398	-1.556	14.11	1.72E-04	4.05E-03	Cluster: Chromosome undetermined scaffold_52, whole
TC161272	-13.616	-1.611	13.99	1.84E-04	4.29E-03	Early-responsive to dehydration stress protein; toPicea
TC161612	-13.422	-1.531	13.9	1.93E-04	4.48E-03	MYB transcription factor MYB51;transcription Glycine max;
TC187415	-13.168	-1.835	13.86	1.97E-04	4.57E-03	Non-specific lipid-transfer protein;
TC162237	-12.238	-1.674	13.85	1.98E-04	4.59E-03	Isoform 3 of Q8R2K1 ; 3Mus
CT583541	-13.656	-1.679	13.81	2.03E-04	4.64E-03	Chromosome 5 SCAF15026, whole genome 5 sequence;
TC187268	-11.712	-1.737	13.8	2.03E-04	4.64E-03	Lysine exporter protein (LYSE/YGGA) Lysine precursor
TC173437	-12.542	-1.651	13.71	2.13E-04	4.86E-03	TrPRP2;Trifolium repens; TrPRP2- Trifolium
TC154541	-9.857	-1.404	13.7	2.14E-04	4.88E-03	44 kDa reaction center protein; kDaPinus
TC170082	-13.25	-1.538	13.7	2.15E-04	4.88E-03	Uncharacterized protein At1g56220.2;Arabidopsis protein
TC185393	-13.575	-1.607	13.65	2.20E-04	4.98E-03	Beta-galactosidase;Raphanus sativus; Beta-galactosidase-
TC162922	-13.289	-1.656	13.65	2.21E-04	4.98E-03	Alginate biosynthesis protein Alg8;biosynthesis
TC160219	-12.402	-1.897	13.64	2.21E-04	4.98E-03	Chromosome chr4 scaffold_32, whole genome chr4 sequence;
TC172319	-13.538	-1.552	13.63	2.23E-04	5.02E-03	Probable WRKY transcription factor 21; WRKYArabidopsis
BF609124	-13.044	-1.625	13.51	2.37E-04	5.29E-03	AT3g01590/F4P13_13;Arabidopsis thaliana; AT3g01590/
TC180504	-14.05	-2.081	13.28	2.69E-04	5.89E-03	70 kDa heat shock cognate kDa protein 3;
CX649641	-13.568	-1.612	13.24	2.75E-04	6.01E-03	Chromosome chr14 scaffold_21, whole genome chr14
TC178296	-12.5	-1.634	13.18	2.82E-04	6.15E-03	Predicted protein;Physcomitrella patens protein; subsp.
TC164593	-13.415	-1.594	13.13	2.90E-04	6.29E-03	Cluster: Chromosome chr18 scaffold_1, whole Chromosome
TC167660	-12.416	-1.617	13.1	2.95E-04	6.36E-03	S-adenosylmethionine synthetase;Dendrobium cv.
TC174017	-13.08	-1.571	13.07	3.00E-04	6.44E-03	Uncharacterized protein At1g56220.2;Arabidopsis protein
TC176589	-12.675	-1.619	13.03	3.06E-04	6.54E-03	Expressed protein;Arabidopsis thaliana; protein;
TC183582	-13.181	-1.543	13.02	3.08E-04	6.56E-03	Late embryogenesis abundant protein;embryogenesis Picea;
TC189789	-10.728	-1.877	12.98	3.15E-04	6.67E-03	Cluster: Secreted protein;Saccharophagus Secreted
TC170679	-13.659	-1.717	12.95	3.19E-04	6.75E-03	Snakin-like cysteine rich protein;cysteine Phaseolus
TC163557	-12.441	-1.602	12.95	3.20E-04	6.76E-03	
TC174637	-12.952	-1.62	12.89	3.31E-04	6.93E-03	
GT244470	-13.207	-1.543	12.82	3.42E-04	7.14E-03	Chromosome chr1 scaffold_135, whole genome chr1
TC178301	-13.231	-1.502	12.8	3.47E-04	7.21E-03	Cluster: Subtilisin-like proteinase-like;Oryza
TC158569	-13.647	-1.602	12.76	3.54E-04	7.34E-03	Probable aquaporin;Picea mariana; aquaporin; Probable
CD021029	-13.799	-1.615	12.7	3.65E-04	7.51E-03	
TC184902	-13.851	-1.648	12.69	3.67E-04	7.51E-03	Chromosome chr5 scaffold_2, whole genome chr5 sequence;
TC163206	-12.905	-1.615	12.68	3.70E-04	7.55E-03	Chromosome chr18 scaffold_24, whole genome chr18
TC177567	-13.327	-1.458	12.57	3.92E-04	7.98E-03	Cluster: Uncharacterized protein
CF667185	-13.336	-1.496	12.55	3.95E-04	8.02E-03	Pinus sylvestris; Polyubiquitin
TC196543	-13.745	-1.8	12.55	3.96E-04	8.03E-03	Os10g0459400 protein;Oryza sativa protein; Japonica
TC187271	-12.793	-1.631	12.45	4.17E-04	8.38E-03	Lodderomyces elongisporus protein; NRRL YB-4239;

TC163354	-12.601	-1.617	12.43	4.23E-04	8.49E-03	At5g14500;Arabidopsis thaliana; At5g14500- Arabidopsis
TC166571	-13.511	-1.49	12.4	4.30E-04	8.57E-03	Chromosome chr2 scaffold_196, whole genome chr2
TC156472	-12.185	-1.614	12.39	4.32E-04	8.60E-03	Isoform 3 of Q8R2K1 ; 3Mus
TC164396	-11.63	-1.68	12.36	4.39E-04	8.73E-03	Cluster: Predicted protein;Physcomitrella Predicted
TC178924	-11.649	-1.681	12.33	4.47E-04	8.85E-03	Cluster: Predicted protein;Physcomitrella Predicted
TC184267	-12.796	-1.61	12.24	4.69E-04	9.22E-03	Cluster: Peroxidase 10 precursor;Peroxidase Arabidopsis
TC161893	-12.43	-1.561	12.22	4.72E-04	9.27E-03	
TC178067	-13.765	-1.753	12.2	4.78E-04	9.33E-03	Chromosome undetermined scaffold_52, whole genome
FE524387	-13.017	-1.554	12.2	4.78E-04	9.33E-03	Chromosome chr1 scaffold_5, whole genome chr1 sequence;
TC156867	-13.732	-1.509	12.09	5.08E-04	9.82E-03	Cluster: Os05g0226800 protein;Oryza Os05g0226800 sativa
TC171025	-13.683	-1.558	12.05	5.18E-04	9.98E-03	
TC171642	-13.597	-1.505	12.04	5.21E-04	1.00E-02	Cluster: Structural protein-like protein;Structural
TC184093	-12.893	-1.609	12.01	5.31E-04	1.02E-02	
TC179762	-12.188	-1.569	12	5.31E-04	1.02E-02	Uncharacterized protein At1g56220.2;Arabidopsis protein
TC194936	-12.649	-1.559	11.9	5.61E-04	1.06E-02	Auxin influx transport protein;influx Casuarina glauca;
TC178647	-14.218	-1.895	11.89	5.66E-04	1.07E-02	
TC185302	-13.813	-1.592	11.86	5.75E-04	1.08E-02	Cluster: Chromosome chr4 scaffold_6, whole Chromosome
TC154533	-9.285	-1.268	11.81	5.89E-04	1.11E-02	carboxylase large chain precursor;large Pinus; Ribulose
TC184132	-11.426	-1.622	11.74	6.13E-04	1.14E-02	Pinus thunbergii; ORF77 - Pinus thunbergii; thunbergii
TC160446	-13.196	-1.483	11.73	6.16E-04	1.15E-02	Receptor-like kinase;Marchantia polymorpha; kinase;
DR055494	-11.838	-1.658	11.71	6.20E-04	1.15E-02	T6J4.14 protein;Arabidopsis thaliana; protein; T6J4.14
TC161179	-13.416	-1.431	11.64	6.46E-04	1.19E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC180204	-11.327	-1.489	11.64	6.47E-04	1.19E-02	Catalase;Brassica napus; Catalase- Brassica
TC197777	-12.781	-1.544	11.63	6.50E-04	1.19E-02	Lysine exporter protein;Chloroflexus exporter aurantiacus
TC188369	-13.571	-1.455	11.6	6.59E-04	1.21E-02	Pinus banksiana; Alcohol dehydrogenase Pinus - Pinus
TC157648	-13.259	-1.449	11.59	6.61E-04	1.21E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC175089	-11.666	-1.647	11.59	6.65E-04	1.21E-02	Cluster: Predicted protein;Physcomitrella Predicted
DR118781	-12.693	-1.578	11.57	6.70E-04	1.21E-02	Chromosome chr14 scaffold_21, whole genome chr14
TC165129	-13.919	-1.807	11.55	6.76E-04	1.22E-02	Snakin-like cysteine rich protein;cysteine Phaseolus
TC163669	-13.655	-1.623	11.55	6.76E-04	1.22E-02	Unknow protein;Oryza sativa protein; Japonica Group;
TC177895	-11.747	-1.59	11.53	6.85E-04	1.23E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC187780	-10.711	-1.801	11.44	7.18E-04	1.29E-02	Cluster: Secreted protein;Saccharophagus Secreted
TC161801	-12.751	-1.54	11.39	7.37E-04	1.32E-02	Pinus banksiana; Alcohol dehydrogenase Pinus - Pinus
TC176757	-13.243	-1.435	11.35	7.53E-04	1.34E-02	Chromosome chr6 scaffold_3, whole genome chr6 sequence;
TC181690	-12.383	-1.532	11.34	7.60E-04	1.35E-02	Cluster: Chromosome undetermined scaffold_349, whole
TC188391	-12.996	-1.625	11.33	7.63E-04	1.35E-02	Alteromonadales bacterium TW-7; Glutathione
TC157560	-12.908	-1.558	11.31	7.71E-04	1.36E-02	
TC192183	-13.378	-1.459	11.26	7.94E-04	1.40E-02	Alteromonadales bacterium TW-7; Glutathione
TC193097	-13.554	-1.541	11.25	7.97E-04	1.40E-02	Receptor-like protein kinase 1;protein Glycine max;
TC178562	-13.008	-1.508	11.25	7.98E-04	1.40E-02	Quinonprotein;Sorangium cellulosum 'Soce 56';
TC155628	-13.052	-1.506	11.18	8.26E-04	1.45E-02	Chromosome chr1 scaffold_135, whole genome chr1
DR071651	-13.032	-1.541	11.18	8.28E-04	1.45E-02	Late embryogenesis abundant protein;embryogenesis Picea
TC157507	-12.289	-1.693	11.17	8.32E-04	1.45E-02	(E)-4-hydroxy-3-methyl-but-2-enyl diphosphate synthase
TC174932	-13.66	-1.414	11.15	8.40E-04	1.46E-02	Cluster: Predicted protein;Physcomitrella Predicted
AM983212	-13.247	-1.394	11.14	8.45E-04	1.47E-02	Chromosome chr2 scaffold_196, whole genome chr2
TC170112	-12.245	-1.703	11.11	8.58E-04	1.48E-02	(E)-4-hydroxy-3-methyl-but-2-enyl diphosphate synthase
TC189520	-13.319	-1.566	11.09	8.68E-04	1.50E-02	Pinus taeda; Diterpene synthase Pinus - Pinus
TC171454	-13.641	-1.442	11.06	8.81E-04	1.52E-02	SRR383;Striga asiatica; SRR383- Striga
TC168182	-13.402	-1.411	11.02	9.03E-04	1.55E-02	
TC155346	-12.61	-1.539	11	9.12E-04	1.56E-02	Hyothetical protein;Toxoplasma Hyothetical gondii RH;
TC169058	-12.59	-1.48	10.99	9.14E-04	1.56E-02	Chromosome chr5 scaffold_2, whole genome chr5 sequence;
TC191922	-13.658	-1.484	10.98	9.20E-04	1.57E-02	Alcohol dehydrogenase;Pinus banksiana; dehydrogenase;
TC177757	-11.649	-1.588	10.96	9.31E-04	1.58E-02	Chromosome chr14 scaffold_21, whole genome chr14
TC156550	-13.736	-1.401	10.96	9.33E-04	1.59E-02	Chromosome chr10 scaffold_282, whole genome chr10
TC193427	-14.106	-1.839	10.94	9.41E-04	1.60E-02	Thiazole biosynthetic enzyme;Pseudotsuga biosynthetic
TC178400	-7.041	-0.3785	10.88	9.74E-04	1.64E-02	Chromosome chr18 scaffold_1, whole genome chr18
GW761469	-12.837	-1.528	10.88	9.75E-04	1.64E-02	Cluster: Predicted protein;Physcomitrella Predicted
TC165364	-12.819	-1.522	10.87	9.80E-04	1.65E-02	Cluster: ERD4 protein;Brassica ERD4 juncea; ERD4
TC198262	-12.74	-1.512	10.86	9.83E-04	1.65E-02	
DR388555	-13.488	-1.4	10.85	9.87E-04	1.65E-02	Cluster: Acid phosphatase type 5 Acid precursor;
TC182400	-13.473	-1.361	10.85	9.91E-04	1.66E-02	Cluster: Peroxidase;Oryza sativa; Peroxidase; Peroxidase
TC165664	-12.118	-1.511	10.82	1.00E-03	1.68E-02	Uncharacterized protein At1g56220.2;Arabidopsis protein
CF471271	-13.241	-1.392	10.8	1.02E-03	1.69E-02	Cluster: Predicted protein;Physcomitrella Predicted
AI920180	-13.404	-1.348	10.77	1.03E-03	1.71E-02	Odorant receptor;Danio rerio; receptor; Odorant receptor
DR744760	-13.471	-1.382	10.77	1.03E-03	1.71E-02	Uncharacterized protein At2g01890.2;Arabidopsis protein
TC198060	-13.519	-1.36	10.76	1.04E-03	1.71E-02	
TC165083	-13.874	-1.541	10.75	1.04E-03	1.72E-02	Chromosome chr16 scaffold_10, whole genome chr16
CF670789	-13.34	-1.497	10.72	1.06E-03	1.75E-02	(E)-4-hydroxy-3-methyl-but-2-enyl diphosphate synthase;
TC191832	-11.425	-1.548	10.69	1.08E-03	1.77E-02	ORF77;Pinus thunbergii; ORF77- Pinus
TC157678	-13.433	-1.465	10.64	1.11E-03	1.81E-02	Caffeic acid O-methyltransferase;Pinus acid pinaster;
TC191477	-13.471	-1.346	10.62	1.12E-03	1.83E-02	Predicted protein;Physcomitrella patens protein; subsp.
CK909868	-13.412	-1.536	10.51	1.19E-03	1.93E-02	Diterpene synthase;Pinus taeda; synthase; Diterpene
TC175122	-12.475	-1.446	10.49	1.20E-03	1.95E-02	Cluster: Subtilisin-like proteinase-like;Oryza
TC165161	-13.747	-1.452	10.49	1.20E-03	1.95E-02	52O08_7;Brassica rapa subsp.pekinensis; 52O08_7
TC162251	-13.777	-1.488	10.48	1.21E-03	1.96E-02	Chromosome undetermined scaffold_550, whole genome
TC176608	-12.915	-1.478	10.45	1.23E-03	1.98E-02	Exocellular acid invertase 1;acid Beta vulgaris;
TC155445	-12.405	-1.652	10.45	1.23E-03	1.98E-02	Pinus taeda; Diterpene synthase Pinus - Pinus
TC163303	-13.976	-1.645	10.42	1.25E-03	2.00E-02	Cluster: Os06g0131500 protein;Oryza Os06g0131500 sativa
TC158216	-13.719	-1.442	10.37	1.28E-03	2.04E-02	Chromosome chr12 scaffold_18, whole genome chr12
TC181175	-9.869	-1.253	10.35	1.29E-03	2.06E-02	Pinus koraiensis; Cytochrome b6 Pinus - Pinus
TC164548	-13.05	-1.413	10.32	1.32E-03	2.10E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC158705	-13.957	-1.647	10.31	1.33E-03	2.10E-02	Cluster: Chromosome chr11 scaffold_14, whole Chromosome
TC179786	-13.856	-1.643	10.27	1.35E-03	2.14E-02	Snakin-like cysteine rich protein;cysteine Phaseolus
TC189585	-13.388	-1.74	10.22	1.39E-03	2.19E-02	Cluster: Chromosome chr8 scaffold_23, whole Chromosome
TC171694	-13.33	-1.339	10.21	1.40E-03	2.19E-02	Chromosome chr18 scaffold_1, whole genome chr18
TC185086	-13.414	-1.328	10.18	1.42E-03	2.21E-02	Cluster: Chromosome chr8 scaffold_88, whole Chromosome
TC195849	-13.192	-1.4	10.12	1.47E-03	2.28E-02	Chromosome chr6 scaffold_3, whole genome chr6 sequence;
TC195583	-12.063	-1.488	10.11	1.48E-03	2.28E-02	Catalase;Oryza sativa; Catalase- Oryza
TC158746	-13.449	-1.547	10.11	1.48E-03	2.28E-02	Chromosome chr2 scaffold_187, whole genome chr2
TC164435	-12.947	-1.439	10.1	1.48E-03	2.29E-02	Beta-galactosidase;Vitis vinifera; Beta-galactosidase-
TC173448	-12.802	-1.491	10.08	1.50E-03	2.30E-02	Cluster: Chromosome chr6 scaffold_3, whole Chromosome
CO167776	-12.709	-1.443	10.08	1.50E-03	2.30E-02	

CV146979	-13.475	-1.336	10.05	1.52E-03	2.34E-02	Predicted protein;Physcomitrella patens protein; subsp.
GW766216	-13.231	-1.427	10.02	1.55E-03	2.37E-02	Os05g0409500 protein;Oryza sativa protein; Japonica
TC171846	-12.99	-1.41	9.966	1.60E-03	2.42E-02	Subtilisin-like protease;Arabidopsis Subtilisin-like
TC162136	-14.096	-1.731	9.966	1.60E-03	2.42E-02	Thiazole biosynthetic enzyme;Pseudotsuga biosynthetic
TC168643	-13.328	-1.326	9.956	1.60E-03	2.43E-02	Cluster: Chromosome chr8 scaffold_23, whole Chromosome
TC194930	-11.303	-1.368	9.941	1.62E-03	2.45E-02	Catalase;Nicotiana glutinosa; Catalase- Nicotiana
AA739630	-11.256	-1.377	9.936	1.62E-03	2.45E-02	Ubiquitin extension protein;core extension
TC170208	-13.077	-1.379	9.932	1.62E-03	2.45E-02	Ser/Thr protein phosphatase family protein, protein
TC163907	-11.345	-1.379	9.923	1.63E-03	2.46E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC197400	-13.341	-1.321	9.92	1.64E-03	2.46E-02	Retinoblastoma protein;Nicotiana benthamiana; protein;
TC174429	-12.961	-1.429	9.909	1.64E-03	2.47E-02	Beta-galactosidase;Vitis vinifera; Beta-galactosidase-
TC159902	-12.65	-1.543	9.888	1.66E-03	2.50E-02	Gb|AAD25141.1;Arabidopsis thaliana; Gb|AAD25141.1-
TC156603	-11.742	-1.587	9.882	1.67E-03	2.50E-02	Cluster: Pectinesterase precursor;Ficus Pectinesterase
DR746251	-13.422	-1.294	9.862	1.69E-03	2.53E-02	Retinoblastoma-like protein;Cocos nucifera; protein;
TC193702	-11.144	-1.306	9.83	1.72E-03	2.56E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC157659	-11.293	-1.344	9.749	1.79E-03	2.66E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC157252	-13.329	-1.293	9.741	1.80E-03	2.66E-02	Beta-galactosidase;Fragaria x Beta-galactosidase;
TC154564	-13.566	-1.302	9.73	1.81E-03	2.67E-02	Cluster: Histone H1 WH1B.1;Histone Triticum aestivum;
TC176304	-13.382	-1.316	9.723	1.82E-03	2.68E-02	Cluster: Chromosome chr8 scaffold_23, whole Chromosome
TC183681	-14.009	-1.554	9.7	1.84E-03	2.70E-02	Probable histone H2A.2;Oryza histone sativa; Probable
TC174709	-13.376	-1.411	9.647	1.90E-03	2.77E-02	2-C-methyl-D-erythritol 2,4-cyclodiphosphate
TC187417	-13.699	-1.345	9.638	1.91E-03	2.78E-02	Cluster: Histone H1;Triticum Histone aestivum; Histone
TC154593	-12.421	-1.438	9.618	1.93E-03	2.80E-02	Zinc finger A20 and AN1 finger domain-containing stress-
TC174654	-13.043	-1.392	9.604	1.94E-03	2.82E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC196667	-12.233	-1.579	9.569	1.98E-03	2.86E-02	Diterpene synthase;Pinus taeda; synthase;
TC168339	-13.616	-1.312	9.541	2.01E-03	2.90E-02	Cryptochrome 1;Orobanche minor; 1; Cryptochrome 1
TC183245	-13.355	-1.336	9.53	2.02E-03	2.91E-02	Chromosome chr2 scaffold_113, whole genome chr2
TC158325	-12.221	-1.497	9.522	2.03E-03	2.92E-02	Cluster: Chromosome chr8 scaffold_29, whole Chromosome
TC186757	-12.398	-1.387	9.513	2.04E-03	2.93E-02	18.2 kDa class I heat kDa shock protein;
TC190102	-11.019	-1.323	9.506	2.05E-03	2.93E-02	
TC190720	-8.472	-0.124	9.453	2.11E-03	3.00E-02	Protein TAR1;Kluyveromyces lactis; TAR1; Protein TAR1
TC195178	-13.635	-1.394	9.451	2.11E-03	3.00E-02	Chloroplast 30S ribosomal protein S15; 30S n=5; Strobus;
TC154873	-12.772	-1.435	9.426	2.14E-03	3.03E-02	Chromosome chr1 scaffold_135, whole genome chr1
CR392867	-14.055	-1.564	9.406	2.16E-03	3.06E-02	
TC184490	-13.522	-1.337	9.394	2.18E-03	3.08E-02	Cluster: Sperm protamine-P1;Mus Sperm musculus; Sperm
TC158139	-12.204	-1.581	9.37	2.21E-03	3.11E-02	Pinus taeda; Diterpene synthase Pinus - Pinus
TC166991	-13.83	-1.444	9.366	2.21E-03	3.11E-02	Glutamate decarboxylase;Pinus pinaster; decarboxylase;
TC169966	-11.657	-1.472	9.345	2.24E-03	3.14E-02	Chromosome chr11 scaffold_13, whole genome chr11
TC164700	-13.498	-1.276	9.335	2.25E-03	3.16E-02	Cluster: Uncharacterized protein At4g32270.2
TC155302	-12.254	-1.393	9.331	2.25E-03	3.16E-02	Chromosome undetermined scaffold_113, whole genome
TC155600	-12.725	-1.432	9.293	2.30E-03	3.22E-02	Chromosome chr1 scaffold_135, whole genome chr1
TC162039	-11.511	-1.494	9.276	2.32E-03	3.25E-02	Inositol-3-phosphate synthase;Sesamum indicum; synthase;
TC159668	-13.892	-1.506	9.234	2.38E-03	3.32E-02	Endo-beta-1,4-glucanase;Pinus radiata;
TC170506	-13.602	-1.334	9.215	2.40E-03	3.34E-02	
TC170159	-13.283	-1.305	9.213	2.40E-03	3.34E-02	Predicted protein;Chlamydomonas reinhardtii; protein;
TC155520	-12.291	-1.372	9.147	2.49E-03	3.44E-02	Chromosome undetermined scaffold_113, whole genome
TC154832	-11.898	-1.479	9.136	2.51E-03	3.46E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC160455	-11.988	-1.459	9.124	2.52E-03	3.48E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC175140	-13.842	-1.479	9.084	2.58E-03	3.55E-02	Cluster: Predicted protein;Physcomitrella Predicted
TC160016	-11.391	-1.407	9.065	2.61E-03	3.58E-02	ORF83a;Pinus koraiensis; ORF83a- Pinus
AM983425	-13.421	-1.271	9.052	2.62E-03	3.60E-02	Monosiga brevicollis MX1; Predicted Monosiga protein -
FE522852	-13.557	-1.255	8.995	2.71E-03	3.69E-02	
TC173716	-13.676	-1.331	8.968	2.75E-03	3.73E-02	Abscisic acid-induced protein-like;Oryza acid-induced
TC158666	-12.979	-1.367	8.93	2.81E-03	3.80E-02	Os04g0591100 protein;Oryza sativa protein; Japonica
TC183116	-12.412	-1.373	8.92	2.82E-03	3.81E-02	CC-NBS-LRR resistance-like protein;CC-NBS-LRR Pinus
TC164060	-13.62	-1.575	8.906	2.84E-03	3.84E-02	Chromosome chr2 scaffold_187, whole genome chr2
TC165044	-13.305	-1.355	8.903	2.85E-03	3.84E-02	Cluster: Chromosome chr4 scaffold_6, whole Chromosome
TC175271	-13.594	-1.261	8.903	2.85E-03	3.84E-02	Cryptochrome 1a;Triticum aestivum; 1a; Cryptochrome 1a
TC177291	-13.325	-1.267	8.893	2.86E-03	3.85E-02	Chromosome chr8 scaffold_29, whole genome chr8 sequence;
CV035531	-13.806	-1.344	8.886	2.87E-03	3.86E-02	Cluster: Chromosome chr8 scaffold_23, whole Chromosome
CV138312	-13.155	-1.308	8.871	2.90E-03	3.88E-02	Zinc finger an1-like family protein; fingerBrassica
TC177657	-13.55	-1.286	8.864	2.91E-03	3.89E-02	Expressed protein;Oryza sativa; protein; Expressed
TC155189	-13.752	-1.321	8.848	2.93E-03	3.90E-02	Cluster: Chromosome chr13 scaffold_48, whole Chromosome
TC156222	-13.817	-1.341	8.848	2.93E-03	3.90E-02	Cluster: Disease resistance associated protein;
TC188430	-12.729	-1.353	8.773	3.06E-03	4.05E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC162032	-13.257	-1.254	8.76	3.08E-03	4.07E-02	
TC185915	-13.662	-1.3	8.752	3.09E-03	4.09E-02	Structural protein-like protein;Structural Arabidopsis
GW725610	-13.351	-1.27	8.743	3.11E-03	4.10E-02	Cluster: Predicted protein;Nematostella Predicted
TC174322	-13.739	-1.461	8.668	3.24E-03	4.24E-02	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase;
TC182421	-12.055	-1.451	8.657	3.26E-03	4.25E-02	Interleukin-10 receptor 2;Gallus receptor gallus;
GT270115	-13.246	-1.232	8.655	3.26E-03	4.25E-02	Ser/Thr protein phosphatase family protein, protein
GW773648	-13.271	-1.264	8.649	3.27E-03	4.25E-02	Secretogranin-2 precursor (Secretogranin II) (SgII)
TC157914	-13.426	-1.222	8.645	3.28E-03	4.25E-02	Cluster: Predicted protein;Physcomitrella Predicted
TC193006	-12.605	-1.407	8.64	3.29E-03	4.26E-02	Metallothionein-like protein;Arabidopsis Metallothionein-
TC164498	-13.21	-1.293	8.621	3.32E-03	4.30E-02	
TC175560	-13.716	-1.32	8.613	3.34E-03	4.31E-02	Chromosome undetermined scaffold_465, whole genome
TC189736	-12.733	-1.398	8.607	3.35E-03	4.32E-02	Cluster: Predicted protein;Physcomitrella Predicted
TC183596	-13.246	-1.253	8.602	3.36E-03	4.33E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC166191	-11.71	-1.365	8.599	3.36E-03	4.33E-02	Chromosome chr7 scaffold_20, whole genome chr7 sequence;
TC170915	-13.721	-1.324	8.594	3.37E-03	4.34E-02	Sucrose-phosphate synthase isoform B;synthase Nicotiana
TC196149	-7.091	-0.484	8.544	3.47E-03	4.44E-02	Chromosome chr10 scaffold_76, whole genome chr10
CO200583	-13.361	-1.255	8.542	3.47E-03	4.44E-02	Cluster: Predicted protein;Monosiga Predicted
TC170295	-11.977	-1.436	8.538	3.48E-03	4.45E-02	Cluster: Pectinesterase precursor;Ficus Pectinesterase
TC168167	-12.945	-1.349	8.507	3.54E-03	4.51E-02	Cluster: Chromosome chr6 scaffold_3, whole Chromosome
TC171002	-13.432	-1.221	8.503	3.55E-03	4.51E-02	Clavata-like receptor;Picea glauca; receptor;
TC193472	-13.613	-1.238	8.5	3.55E-03	4.51E-02	
TC186900	-13.753	-1.45	8.485	3.58E-03	4.54E-02	Cluster: Chromosome chr13 scaffold_17, whole Chromosome
TC199351	-13.441	-1.217	8.459	3.63E-03	4.59E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC171149	-13.585	-1.263	8.459	3.63E-03	4.59E-02	Cluster: Chromosome chr9 scaffold_7, whole Chromosome
TC180308	-13.732	-1.336	8.413	3.73E-03	4.69E-02	D7-related protein;Culicoides sonorensis; protein;
TC181556	-12.19	-1.328	8.401	3.75E-03	4.72E-02	Cluster: Catalase;Capsicum annuum; Catalase; Catalase -
TC186136	-12.33	-1.313	8.391	3.77E-03	4.73E-02	Chromosome chr8 scaffold_88, whole genome chr8 sequence;

GO096284	-13.294	-1.251	8.36	3.84E-03	4.79E-02	Photosystem II D2 protein;II Liliales; Photosystem
TC178325	-12.576	-1.358	8.324	3.91E-03	4.87E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC166564	-12.159	-1.427	8.302	3.96E-03	4.92E-02	Cluster: Chromosome chr12 scaffold_18, whole Chromosome
TC189761	-11.501	-1.335	8.279	4.01E-03	4.97E-02	60S ribosomal protein L34;ribosomal Nicotiana tabacum;
TC187616	-11.305	-1.285	8.266	4.04E-03	4.99E-02	D2 protein (Photosystem Q(A) protein); proteinMarchantia
TC155598	-13.743	-1.325	8.258	4.06E-03	5.01E-02	Cluster: Chromosome chr3 scaffold_8, whole Chromosome
TC186972	-13.181	-1.25	8.245	4.09E-03	5.03E-02	Cluster: Gag;Zea mays; Gag; Gag -
TC191015	-12.725	-1.318	8.223	4.14E-03	5.08E-02	Glyceraldehyde-3-phosphate dehydrogenase, cytosolic;
BX250219	-13.684	-1.374	8.213	4.16E-03	5.11E-02	Outer membrane lipoprotein Slp;membrane Geobacter
DR111093	-13.656	-1.251	8.189	4.21E-03	5.17E-02	
TC158669	-13.244	-1.221	8.163	4.28E-03	5.21E-02	Heat shock 70 kDa protein; shockGlycine
TC182106	-11.74	-1.426	8.154	4.30E-03	5.23E-02	Pinus thunbergii; ORF77 - Pinus thunbergii; thunbergii
BQ697477	-13.629	-1.215	8.148	4.31E-03	5.24E-02	Cluster: Chromosome chr13 scaffold_17, whole Chromosome
TC164455	-12.269	-1.297	8.143	4.32E-03	5.24E-02	Ser/Thr protein phosphatase family protein, protein
TC186618	-10.997	-1.287	8.137	4.34E-03	5.25E-02	P700 chlorophyll a apoprotein A1; chlorophyllPinus
TC180849	-12.504	-1.313	8.105	4.42E-03	5.32E-02	core eudicotyledons; AT4g05320/C17L7_240 - Arabidopsis
FE523044	-13.523	-1.199	8.074	4.49E-03	5.40E-02	LIM domain protein WLIM-1;domain Helianthus annuus;
TC178841	-10.742	-1.485	8.024	4.62E-03	5.51E-02	Catalase-1/2;Glycine max; Catalase-1/2- Glycine
TC185766	-14.032	-1.489	8.009	4.66E-03	5.54E-02	Thiazole biosynthetic enzyme;Pseudotsuga biosynthetic
TC156251	-12.992	-1.301	8.008	4.66E-03	5.54E-02	Cluster: Chromosome chr8 scaffold_23, whole Chromosome
TC191487	-13.623	-1.194	7.979	4.73E-03	5.63E-02	Chromosome chr6 scaffold_3, whole genome chr6 sequence;
TC173335	-13.438	-1.175	7.928	4.87E-03	5.76E-02	Chromosome chr5 scaffold_2, whole genome chr5 sequence;
TC168158	-13.306	-1.199	7.917	4.90E-03	5.79E-02	Retinoblastoma-like protein;Cocos nucifera; protein;
GW726284	-13.432	-1.209	7.914	4.91E-03	5.79E-02	
TC170750	-12.867	-1.303	7.912	4.91E-03	5.79E-02	Selenocysteine methyltransferase;Brassica oleracea
CO165540	-13.18	-1.219	7.871	5.02E-03	5.90E-02	Cluster: Polysaccharide biosynthesis related protein;
FG618203	-13.288	-1.205	7.871	5.02E-03	5.90E-02	
TC162932	-13.392	-1.169	7.84	5.11E-03	5.96E-02	Chromosome chr18 scaffold_1, whole genome chr18
TC163811	-12.471	-1.259	7.82	5.17E-03	6.02E-02	Chromosome chr10 scaffold_138, whole genome chr10
TC189850	-12.344	-1.369	7.813	5.19E-03	6.03E-02	Cluster: Aspartic proteinase nepenthesin-1 precursor;
TC183560	-13.369	-1.213	7.798	5.23E-03	6.08E-02	
NP1369776	-12.847	-1.319	7.795	5.24E-03	6.08E-02	
TC184285	-13.065	-1.398	7.794	5.24E-03	6.08E-02	Pinus taeda; Diterpene synthase Pinus - Pinus
TC183987	-13.398	-1.171	7.79	5.25E-03	6.08E-02	
TC173978	-13.49	-1.162	7.788	5.26E-03	6.08E-02	Cluster: Retinoblastoma-related protein 1;Retinoblastoma-
TC165977	-13.519	-1.161	7.766	5.32E-03	6.15E-02	At1g04920/F13M7_7;Arabidopsis thaliana; At1g04920/
TC194337	-13.573	-1.19	7.746	5.38E-03	6.21E-02	Chromosome chr1 scaffold_84, whole genome chr1 sequence;
TC164132	-13.977	-1.477	7.742	5.40E-03	6.21E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC170388	-12.13	-1.311	7.734	5.42E-03	6.23E-02	Chromosome chr7 scaffold_20, whole genome chr7 sequence;
TC172437	-13.641	-1.231	7.72	5.46E-03	6.27E-02	Uncharacterized protein At3g27210;Arabidopsis protein
BQ698297	-13.169	-1.375	7.718	5.47E-03	6.27E-02	Diterpene synthase;Pinus taeda; synthase; Diterpene
TC157495	-13.377	-1.175	7.715	5.48E-03	6.27E-02	Chromosome chr1 scaffold_84, whole genome chr1 sequence;
TC156511	-13.715	-1.332	7.705	5.51E-03	6.29E-02	Predicted protein;Chlamydomonas reinhardtii; protein;
TC197467	-13.625	-1.206	7.704	5.51E-03	6.29E-02	ribosomal protein S18;Pinus protein koraiensis;
TC154971	-12.766	-1.301	7.663	5.64E-03	6.41E-02	Chromosome chr1 scaffold_135, whole genome chr1
TC170711	-13.05	-1.25	7.647	5.69E-03	6.46E-02	Ascorbate peroxidase;Pinaceae; Ascorbate peroxidase;
TC183424	-13.663	-1.204	7.634	5.73E-03	6.50E-02	Cluster: Chromosome chr13 scaffold_17, whole Chromosome
TC191431	-12.291	-1.248	7.601	5.83E-03	6.59E-02	Catalase isozyme 3;Cucurbita isozyme pepo; Catalase
TC174690	-13.589	-1.234	7.577	5.91E-03	6.66E-02	Chromosome chr8 scaffold_23, whole genome chr8 sequence;
TC196273	-13.552	-1.149	7.569	5.94E-03	6.66E-02	
TC174802	-13.558	-1.208	7.568	5.94E-03	6.66E-02	
TC183105	-11.902	-1.301	7.558	5.97E-03	6.67E-02	Predicted protein;Monosiga brevicollis protein; MX1;
TC160766	-12.554	-1.238	7.542	6.03E-03	6.72E-02	Cluster: Catalase isozyme 2;Catalase Cucurbita pepo;
GO096289	-13.158	-1.26	7.536	6.05E-03	6.74E-02	Photosystem II CP43 protein;II Cedrus deodara;
GW750811	-13.331	-1.161	7.535	6.05E-03	6.74E-02	scaffold_17 whole sequence; wholeParamecium,
TC160267	-13.431	-1.156	7.527	6.08E-03	6.76E-02	Chromosome chr5 scaffold_2, whole genome chr5 sequence;
AI725133	-12.323	-1.238	7.517	6.11E-03	6.78E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC170220	-12.295	-1.244	7.512	6.13E-03	6.79E-02	Ser/Thr protein phosphatase family protein, protein
TC164307	-12.849	-1.305	7.503	6.16E-03	6.81E-02	Gb|AAC63835.1;Arabidopsis thaliana; Gb|AAC63835.1-
TC183023	-13.9	-1.27	7.494	6.19E-03	6.83E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC176501	-13.706	-1.204	7.493	6.19E-03	6.83E-02	Cluster: Chromosome chr13 scaffold_48, whole Chromosome
TC189345	-13.922	-1.318	7.459	6.31E-03	6.94E-02	Cluster: Chromosome chr13 scaffold_17, whole Chromosome
TC194333	-13.091	-1.238	7.451	6.34E-03	6.95E-02	Chromosome undetermined scaffold_38, whole genome
TC155963	-13.424	-1.131	7.432	6.41E-03	7.01E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC178002	-11.199	-1.188	7.423	6.44E-03	7.04E-02	
TC194228	-13.414	-1.131	7.418	6.46E-03	7.05E-02	Histone H2A;Hyacinthus orientalis; H2A; Histone H2A
TC191500	-12.587	-1.281	7.409	6.49E-03	7.07E-02	Ferredoxin-like protein;Mariprofundus ferrooxydans
GW740250	-13.541	-1.194	7.402	6.52E-03	7.08E-02	scaffold_35 whole sequence; wholeParamecium,
TC168690	-13.815	-1.267	7.394	6.54E-03	7.11E-02	Chromosome chr9 scaffold_90, whole genome chr9 sequence;
TC164688	-13.338	-1.153	7.392	6.55E-03	7.11E-02	Abscisic acid-induced protein-like;Oryza acid-induced
BX681215	-10.357	-0.306	7.357	6.68E-03	7.23E-02	
TC170228	-13.58	-1.149	7.356	6.68E-03	7.23E-02	Chromosome chr1 scaffold_46, whole genome chr1 sequence;
TC192709	-12.224	-1.25	7.351	6.70E-03	7.24E-02	Aspartic proteinase;Theobroma cacao; proteinase;
CO364594	-13.301	-1.177	7.348	6.72E-03	7.25E-02	Cluster: Chromosome chr14 scaffold_54, whole Chromosome
TC181337	-13.721	-1.199	7.341	6.74E-03	7.26E-02	Chromosome chr1 scaffold_46, whole genome chr1
TC164736	-12.644	-1.232	7.332	6.77E-03	7.29E-02	Chromosome chr4 scaffold_6, whole genome chr4
GO096290	-13.739	-1.245	7.329	6.78E-03	7.29E-02	Photosystem II protein D2;II Ipomoea purpurea;
TC166361	-13.815	-1.261	7.322	6.81E-03	7.31E-02	Cluster: RL11;Human herpesvirus RL11; 5; RL11
TC196692	-13.378	-1.161	7.301	6.89E-03	7.39E-02	Cluster: Ribokinase;Haemophilus influenzae Ribokinase;
TC184917	-13.345	-1.156	7.296	6.91E-03	7.39E-02	DNA mismatch repair protein MutS; mismatchMicroscilla
TC165532	-13.436	-1.17	7.295	6.92E-03	7.39E-02	Chromosome chr12 scaffold_93, whole genome chr12
TC175169	-13.286	-1.168	7.285	6.95E-03	7.42E-02	Chromosome chr1 scaffold_84, whole genome chr1
TC179104	-12.817	-1.24	7.282	6.97E-03	7.43E-02	Cell wall protein;Saccharomyces wall cerevisiae YJM789;
TC179329	-13.596	-1.143	7.281	6.97E-03	7.43E-02	Histone H2A;Hyacinthus orientalis; H2A; Histone H2A
TC156059	-13.703	-1.157	7.269	7.02E-03	7.47E-02	Cluster: Chromosome chr12 scaffold_36, whole Chromosome
TC172004	-12.9	-1.255	7.257	7.06E-03	7.50E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC193433	-12.564	-1.261	7.235	7.15E-03	7.57E-02	Photosystem II protein D2;II Ipomoea purpurea;
TC170726	-13.467	-1.115	7.229	7.17E-03	7.59E-02	Cluster: Chromosome chr19 scaffold_4, whole Chromosome
TC157493	-13.472	-1.115	7.225	7.19E-03	7.59E-02	Cluster: Chromosome chr19 scaffold_4, whole Chromosome
TC174890	-12.869	-1.346	7.21	7.25E-03	7.64E-02	1-deoxy-D-xylulose-5-phosphate reductoisomerase;
TC176847	-14.087	-1.423	7.21	7.25E-03	7.64E-02	Cytochrome P450 CYPB;Pinus P450 taeda; Cytochrome
TC158547	-13.496	-1.138	7.197	7.30E-03	7.68E-02	Chromosome chr1 scaffold_84, whole genome chr1 sequence;

TC191284	-13.422	-1.118	7.191	7.33E-03	7.68E-02	
TC156654	-13.607	-1.125	7.176	7.39E-03	7.73E-02	Peptide transporter-like;Oryza sativa transporter-like;
TC187811	-13.497	-1.116	7.161	7.45E-03	7.77E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC175588	-13.208	-1.176	7.157	7.47E-03	7.77E-02	
GO096291	-11.96	-1.344	7.135	7.56E-03	7.85E-02	CP43 chlorophyll apoprotein of photosystem chlorophyll 2
TC177931	-11.534	-1.244	7.12	7.62E-03	7.90E-02	Granule-bound starch synthase 1,
TC182783	-12.867	-1.273	7.119	7.63E-03	7.90E-02	Cluster: Gag-pol polyprotein;Glycine Gag-pol max; Gag-pol
TC165693	-13.195	-1.171	7.117	7.64E-03	7.91E-02	Maltose ABC transporter, permease protein; ABCVibrio
TC173857	-13.099	-1.199	7.098	7.72E-03	7.97E-02	Heat shock protein DnaJ;shock Medicago truncatula;
TC180154	-13.698	-1.189	7.089	7.76E-03	8.00E-02	Clavata-like receptor;Picea glauca; receptor;
FG617140	-13.766	-1.221	7.078	7.81E-03	8.04E-02	Chromosome chr9 scaffold_90, whole genome sequence;
TC196223	-9.665	-1.024	7.074	7.82E-03	8.05E-02	
TC182930	-12.517	-1.252	7.064	7.86E-03	8.07E-02	polymerase subunit alpha;Pinus subunit thunbergii;
TC176896	-13.628	-1.124	7.047	7.94E-03	8.14E-02	Predicted protein;Physcomitrella patens protein; subsp.
NP543241	-12.881	-1.28	7.044	7.95E-03	8.15E-02	
FN693640	-12.967	-1.255	7.039	7.98E-03	8.16E-02	Os05g0374700 protein;Oryza Os05g0374700 sativa Japonica
TC171074	-13.189	-1.18	7.032	8.01E-03	8.17E-02	
TC180668	-10.42	-0.771	7.031	8.01E-03	8.17E-02	C chain;Pinus; ATP chain; synthase C
TC164128	-13.442	-1.133	7.023	8.05E-03	8.19E-02	
BF609787	-13.382	-1.104	7.019	8.07E-03	8.20E-02	BH04450;Bartonella henselae; UPF0301protein BH04450
CF474284	-13.546	-1.174	7.017	8.07E-03	8.20E-02	At3g56260;Arabidopsis thaliana; At3g56260- Arabidopsis
TC197169	-13.948	-1.359	7.001	8.15E-03	8.24E-02	
TC158530	-12.101	-1.262	6.999	8.15E-03	8.25E-02	At1g11080;Arabidopsis thaliana; At1g11080- Arabidopsis
TC161674	-13.827	-1.232	6.997	8.17E-03	8.25E-02	Ribosomal protein L7/L12;Stenotrophomonas protein
CF666457	-12.598	-1.261	6.989	8.20E-03	8.28E-02	Os06g0536800 protein;Oryza Os06g0536800 sativa Japonica
TC170022	-13.516	-1.122	6.977	8.26E-03	8.32E-02	Chromosome chr1 scaffold_84, whole genome sequence;
TC192853	-13.208	-1.122	6.975	8.26E-03	8.32E-02	Cluster: Chromosome chr18 scaffold_1, whole Chromosome
TC165978	-12.927	-1.222	6.97	8.29E-03	8.33E-02	Probable serine/threonine-specific protein kinase;
TC157248	-13.458	-1.106	6.949	8.39E-03	8.41E-02	Cluster: Chromosome chr6 scaffold_3, whole Chromosome
TC189730	-13.51	-1.115	6.932	8.47E-03	8.46E-02	Clavata-like receptor;Picea glauca; receptor;
FG616103	-13.235	-1.17	6.931	8.47E-03	8.46E-02	
TC154996	-13.576	-1.124	6.93	8.48E-03	8.46E-02	Cluster: Chromosome chr17 scaffold_12, whole Chromosome
FG615694	-12.105	-1.248	6.927	8.49E-03	8.46E-02	
TC154639	-13.31	-1.149	6.907	8.59E-03	8.53E-02	Signal transduction histidine kinase regulating citrate/
TC173492	-12.778	-1.218	6.87	8.76E-03	8.68E-02	Chromosome chr8 scaffold_23, whole genome sequence;
TC157737	-12.227	-1.224	6.87	8.77E-03	8.68E-02	Chromosome chr7 scaffold_31, whole genome sequence;
TC167803	-13.728	-1.12	6.863	8.80E-03	8.70E-02	Chromosome chr14 scaffold_164, whole genome chr14
TC183801	-13.365	-1.109	6.837	8.93E-03	8.80E-02	Pb-fam-2 protein;Plasmodium berghei; protein;
EL342390	-12.768	-1.208	6.837	8.93E-03	8.80E-02	Cluster: Osr40g3 protein;Oryza Osr40g3 sativa; Osr40g3
TC159233	-13.467	-1.089	6.831	8.96E-03	8.82E-02	Chromosome chr18 scaffold_137, whole genome chr18
TC189513	-13.471	-1.121	6.816	9.03E-03	8.87E-02	
TC175664	-13.373	-1.109	6.816	9.04E-03	8.87E-02	Predicted protein;Monosiga Predicted brevicollis MX1;
TC196265	-12.765	-1.219	6.812	9.05E-03	8.88E-02	reaction center subunit IX;center Pinus thunbergii;
TC169649	-12.909	-1.215	6.81	9.07E-03	8.89E-02	Chromosome chr15 scaffold_19 whole genome sequence;
TC182193	-12.869	-1.213	6.801	9.11E-03	8.92E-02	Auxin-regulated protein;Zinnia elegans; protein;
TC193802	-13.259	-1.131	6.795	9.14E-03	8.93E-02	Chromosome chr12 scaffold_47, whole genome chr12
TC158207	-11.558	-1.199	6.789	9.17E-03	8.95E-02	Aspartic proteinase;Theobroma cacao; proteinase;
TC194497	-13.564	-1.098	6.787	9.18E-03	8.95E-02	Histone H2A;Hyacinthus orientalis; H2A; Histone H2A
TC189298	-11.145	-1.184	6.782	9.21E-03	8.96E-02	Cytochrome b6-f complex subunit 4; b6-fPinus
TC157452	-13.165	-1.138	6.768	9.28E-03	9.01E-02	Expressed protein;Oryza sativa; protein
TC159037	-13.696	-1.149	6.758	9.33E-03	9.06E-02	Predicted protein;Physcomitrella patens protein; subsp.
TC166741	-13.424	-1.088	6.753	9.36E-03	9.08E-02	Ser/Thr protein phosphatase family protein, protein
TC162429	-11.855	-1.345	6.751	9.37E-03	9.08E-02	Cytochrome b6-f complex subunit 4; b6-fPicea
TC172072	-13.639	-1.123	6.746	9.39E-03	9.09E-02	meprin and TRAF homology domain-containing and protein /
TC184833	-13.474	-1.116	6.736	9.45E-03	9.14E-02	Chromosome undetermined scaffold_139, whole
TC165813	-13.433	-1.081	6.72	9.53E-03	9.21E-02	Formamidase-like protein;Arabidopsis thaliana; protein
TC184955	-13.885	-1.289	6.719	9.54E-03	9.21E-02	Auxin response factor 23;response Oryza sativa
TC161302	-11.34	-1.108	6.696	9.66E-03	9.32E-02	
TC179673	-13.807	-1.204	6.691	9.69E-03	9.33E-02	Clavata-like receptor;Picea glauca; receptor;
TC199268	-12.407	-1.194	6.687	9.71E-03	9.35E-02	CROC-1-like protein;Picea mariana; protein; CROC-1-like
CO166849	-13.258	-1.116	6.665	9.83E-03	9.45E-02	Cytosolic ascorbate peroxidase;Vitis ascorbate vinifera;
TC187404	-13.275	-1.117	6.665	9.83E-03	9.45E-02	
AI919903	-12.75	-1.195	6.654	9.89E-03	9.49E-02	
TC184452	-13.475	-1.076	6.645	9.94E-03	9.49E-02	Cluster: Chromosome chr8 scaffold_23, whole Chromosome