NAIST-IS-DD1361018

Doctoral Dissertation

Curcuminoid Synthesis Pathway Analysis Based on Integration of RNA-seq and Metabolite Data

Li Donghan

September 1, 2015

Department of Information Science
Graduate School of Information Science
Nara Institute of Science and Technology

A Doctoral Dissertation
submitted to the Graduate School of Information Science,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
Doctor of Science.
Li Donghan

Thesis Committee:

Professor Shigehiko KANAYA	(Supervisor)
Professor Keiichi YASUMOTO	(Co-supervisor)
Associate Professor MD. ALTAF-UL-AMIN	(Co-supervisor)
Associate Professor Tadao SUGIURA	(Co-supervisor)
Assistant Professor Naoaki ONO	(Co-supervisor)

Curcuminoid Synthesis Pathway Analysis Based on Integration of RNA-seq and Metabolite Data *

Li Donghan

Abstract

Curcuminoids are secondary metabolites that act as the primary active constituents of turmeric. To understand the curcuminoids synthesis pathway, we compared two wild strains and two cultivars to understand the differences in the synthesis pathway. Because the fluxes of metabolic reactions depend on the amounts of their substrate and the activity of the catalysts, we analyzed the metabolite concentrations and gene expression of the enzymes. The RNA sequencing (RNA-seq) analysis based on Next-Generation Sequencing (NGS) technology was applied to detect gene expression level. By comparing the genes expression level corresponding to both leaves and rhizomes of 4 specimens (*Curcuma aromatica*, *Curcuma longa*, *Curcuma longa* cultivars Sekiyou, *Curcuma longa* cultivars Ougon), we determined the genes for which expression level significantly changes between the leaves and rhizomes. We next focused on the metabolic pathway that synthesizes curcuminoids, in order to understand the difference in curcuminoids concentrations among the four specimens. We developed a 'selection first' method for RNA-seq analysis in which short reads are mapped to selected enzymes in the biosynthetic pathways in order to reduce the effect of mapping errors. Using this method, we found that there were significant changes in expression levels of the enzymes at the branching point of the curcuminoids biosynthesis pathway, which is consistent with the observed curcuminoids contents. By the high expression level of phenylpropanoylacetyl-CoA enzymes, p-coumaroyl-CoA is expected to synthase in a fast reaction speed and make the synthesis flux to the largest accumulation of demethoxy curcumin in *Curcuma longa* cultivars Ougon and *Curcuma longa aromatica.* While in the other group (*Curcuma longa* cultivars Sekiyou and *Curcuma longa*), low expression level of phenylpropanoylacetyl-CoA results in more accumulation of feruloy-CoA instead of demethoxy curcumin. The high expression level of curcumin synthase enzymes in the following reaction steps synthesize the largest amount of curcumin.

Keywords:
metabolite analysis, RNA-seq, NGS, curcumin, gene expression analysis,

---

*Doctoral Dissertation, Department of Information Science, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-DD1361018, September 25, 2015.

Contents

List of Figures

List of Tables

List of Abbreviation

AR	*Curcuma aromatica*
4CL	4-coumarate:CoA ligase
C3H	cinnamate-3-hydroxylase
C4H	cinnamate-4-hydroxylase
CURS1	curcumin synthase1
CURS2	curcumin synthase2
CURS3	curcumin synthase3
CV	Coefficients of Variation
CUS	curcumin synthase
DCS	phenylpropanoylacetyl-CoA synthase
ESI	Electrospray Ionization.
FC	Folder Change.
FPKM	Fragments Per Kilobase of exon per Million mapped fragments
GC - MS	Gas Chromatography - Mass Spectrometry
HCT	hydroxycinnamoyl transferase
IND	Investigational New Drug Application is the way that a pharmaceutical company obtains permission to ship an experimental drug across state lines (usually to clinical investigators) before a marketing application for the drug has been approved.
LC-MS	Liquid Chromatography–Mass Spectrometry
LN	*Curcuma longa*

NDA	New Drug Application is the regulation and control of new drugs in the United States. Since 1938, every new drug has been the subject of an approved NDA before U.S. commercialization.
NGS	Next-Generation Sequencing
OMT	O-methyltransferase
OU	*Curcuma longa* cultivars Ougon
PCR	Polymerase Chain Reaction
PAL	phenylalanine ammonia-lyase
SK	*Curcuma* longa cultivars Sekiyou

Chapter I. Introduction

This doctor dissertation explained the author's research and results on the research subject: applying the next-generation sequencing (NGS) technique and gas chromatography - mass spectrometry (GC - MS) simultaneously on the analysis of the metabolomics.

The purpose of this study is to analyze the metabolite biosynthesis pathway and understanding the metabolite synthesis mechanism. The metabolites are affected by the network of synthesis reactions and expression level of enzymes. The bio-reaction network would decide synthesis process. The enzymes expression level controls the bio-reaction's rate. In this study, we choose turmeric (*Curcuma longa*) as study object and four cultivars of turmeric are selected as samples. We compare the concentration of main secondary metabolite of the samples by the GC-MS to determine the metabolite difference between the samples. On the other hand, we use NGS technique to determine the transcriptome of each sample. Not only the enzyme sequence alignment but also the enzyme expression level can be determined. The metabolite changes among the 4 samples, while the enzymes are up-regulated or down-regulated differently compared with each other sample. We found that the metabolite difference between the samples can be interpreted by the gene expression level changes between the 4 cultivars. The up-regulated or down-regulated genes affect the bio-synthesis reactions; leading to the different reaction flow; and finally cause the metabolite changes among the samples.

When NGS technique initialized to transcriptome analysis, first the gene sequences will be assembled from short reads. Then all the raw short reads will be mapped to the assembled gene sequences to calculate the short read abundance of each gene sequence which will be gene expression level. According to the purpose of the research, a specific gene set of the transcriptom will be selected by and the gene expression level will be determind. The workflow is suitable and generally applied in whole transcriptom and genome analysis. It can be called ‘ mapping first, select later’ mapping strategy. But when this ‘mapping first’ strategy applied to a limited gene set (for example the secondary metabolic pathway genes), the mapping result will be inaccurate for that in the mapping stage, compare with the targeted gene set, large numbers of whole transcriptom genes involve in and mapping gene template is not accurate. To avoid this, we develop a ‘select first’ mapping strategy in this study. For the secondary metabolic pathway gene analysis, we firstly select a targeted gene set and then map all the raw reads to the selected genes. This method lowers the coefficient of variation (CV) of estimated expression level of selected genes from 2.12 to 1.47 compared with the ‘mapping first’ strategy. We also validated this strategy by applying on four kinds of housekeeping genes. The CVs of these housekeeping genes out of smaller range.

1.1 Turmeric and curcuminoid

During the past decades, there has been increasing demand of the new medical technology and new drug to the pharmaceutical research and development. Nowadays, the new drug developing requires more developing time and cost. The Tufts Center for the Study of Drug Development (TCSDD) has conducted several studies of drug development times, which indicate that the total time from synthesis (of a compound) to approval of a New Drug Application (NDA) for self-originated New Chemical Entity (NCE) has increased significantly, from an average of 7.9 years in the 1960s to 12.8 years in the 1990s **[Fig 1.1]** (Michael et al. 2004). Much of the increase in drug development time is due to increased time for the clinical trial portion of the process. Also clinical testing is the most costly component of drug discovery, with the costs

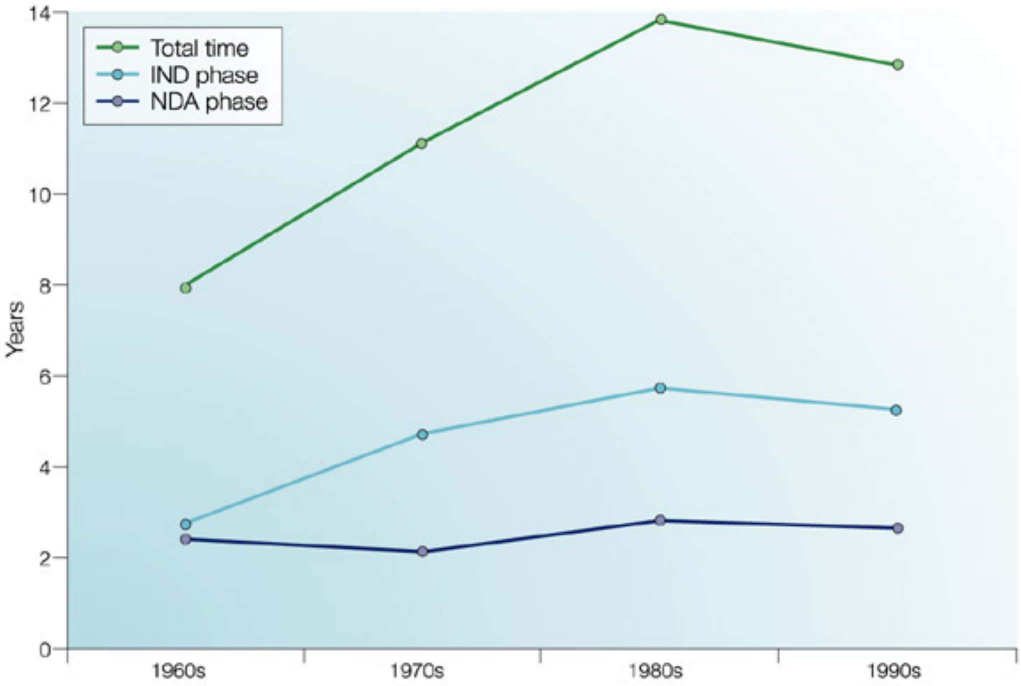


Figure 1.1 Total time cost from synthesis (of a compound) to NDA approval. (Figure adapted from Michael et al. 2004)

in each successive phase clinical testing is the most costly component of drug discovery, with the costs increasing in each successive phase **[Fig 1.2]** (Michael et al. 2004). Phase I corresponds to a 30 days of clinical trials in human. A small number of normal, healthy volunteers are conducted and safe dose is determined. Phased II, large number samples of volunteers who have the production's medical condition are intended to treat. In Phase III, large number of patients are involved. Phase III expected costs are 44% higher than Phase I. Although animal testing costs are less than those for clinical trials, the total cost of these tests is not trivial. Meanwhile the risk of the side effect became more and more uncontrollable. Because of the increasing cost of the new drug development, more and more pharmic researchers are turning into the traditional herbal medicine research.

Traditional herbs are used in many countries for centuries, for example Traditional Chinese Medicine in China, Kampo in Japan, Jamu in Indonesia (Borchers et al. 2000; Afendi et al. 2012; Tan et al. 2004). Compare with modern medicine, traditional herbs have been used for a long time and have been improved. Most herbal medicines are derived from wild plant product and they are the mixture of multiple plant secondary metabolites. Hence plant secondary metabolomics is the key to understand the bio-synthesized mechanism of medicinal components in the plants.

Turmeric (*Curcuma longa*) is a tropical plant in the family Zingiberaceae **[Fig. 1.3A]**. This genus has been considered originating in the Indo-Malayan Region (Purseglove, 1968) and has widespread occurrence in the tropics of Asia to Africa and Australia. Out of the 100 or so species were reported in the genus, and about 40 of them are of Indian origin (Velayyudhan *et al.,* 1999). The powdered rhizome of turmeric **[Fig 1.3B]** has been used widely as a spice especially in traditional Asian dishes, as a traditional medicine, and also as a bright yellow dye (Jayaprakasha et al. 2005). Traditional Indian Ayurvedic and Sidha systems of medicine have recognized the medicinal value of turmeric in its crude form since very ancient times. The last few decades have witnessed extensive research interests worldwide in the biomedical activity of turmeric and its compounds. Thus *Curcuma* is now gaining importance all over the world as a mighty

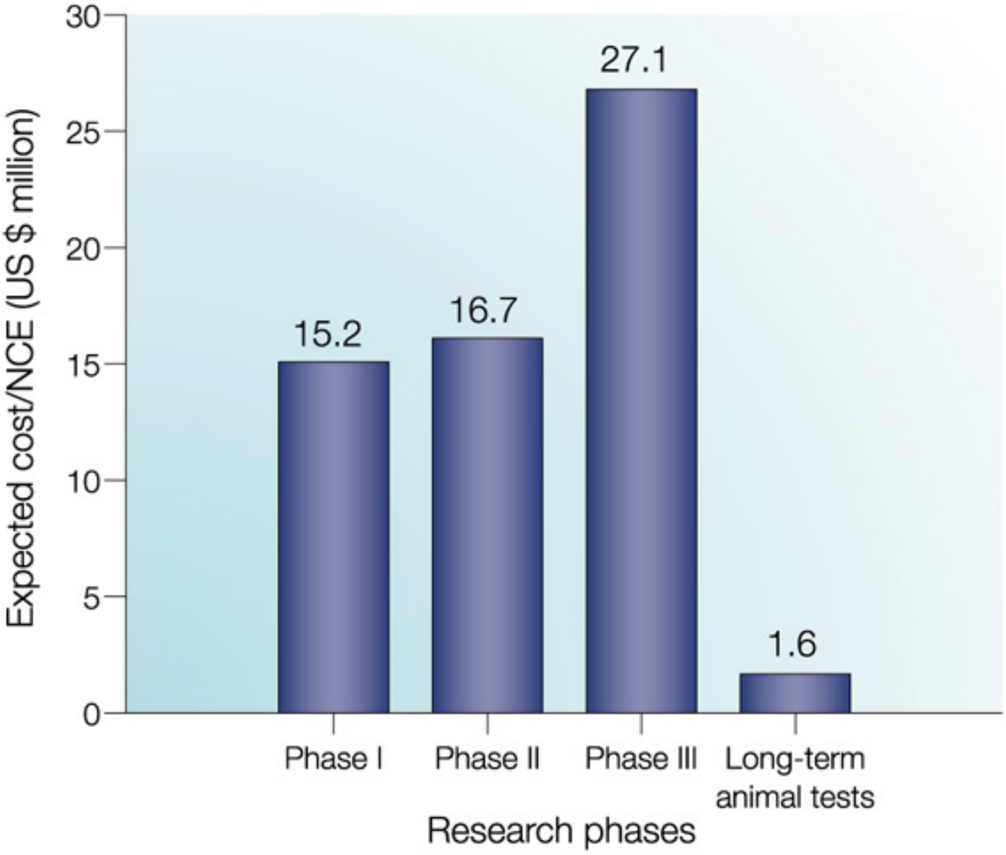


Figure 1.2 Expected cost/NCE of research phases. (Figure adapted from Michael et al. 2004)

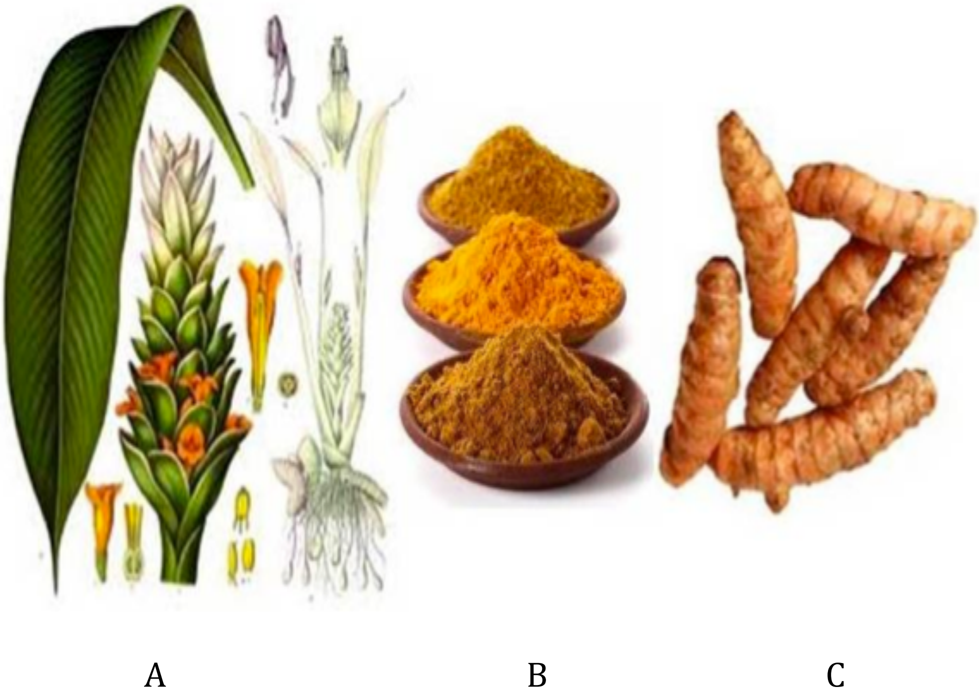


Figure 1.3 *Curcuma longa* plant, the rhizome, and the powdered rhizome.

cure to combat a variety of ailments, as the genus carries molecules credited with anti-inflammatory, hypocholestraemic, choleratic, antimicrobial, antirheumatic, antifibrotic, antivenomous, antiviral, antidiabetic, antihepatotoxic and anticancerous properties as well as insect repellent activity (Chattopadhyay et al. 2004). A US Patent (No. 20030185907) is now in place on a method of treatment of inflammation and pain in mammals including human beings using curcuminoids along with other plant-derived molecules. Turmeric oil is also now used in aromatherapy and the perfume industry. Turmeric powder has healing effect on both aseptic and septic wounds in rats and rabbits (Gujral et al. 1953). And it also shows adjuvant chemoprotection in experimental forestomach and oral cancer models of Swiss mice (Azuine et al. 1994). A research of indomethacin-induced gastric ulcer caused by reactive oxygen species shows the gastroprotective effect of curcumin thereon (Chattopadhyay et al. 2005).

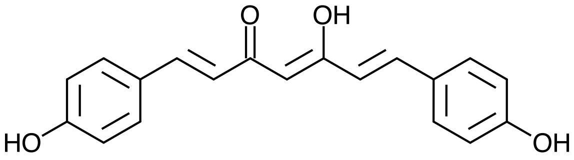
Curcumin is known as the main medicinal ingredient, and its analogs **[Fig 1.4]**—curcuminoids—also show useful bioactivity (Maheshwari et al. 2006). Various species in the genus *Curcuma* synthesize curcuminoids, and their contents and activities vary greatly (Sasikumar 2005; Lobo et al. 2009; Wang Y, et al. 1999). Jayaprakasha studied the variation of three curcuminoids in four different commercial varieties of turmeric, namely 'Salem', 'Erode', 'Balasore' and 'Mysore'. The percentage of curcumin, demethoxy curcumin and bis-demethoxy curcumin were found to be 1.06 ± 0.061 to 5.65 ± 0.040, 0.83 ± 0.047 to 3.36 ± 0.040 and 0.42 ± 0.036 to 2.16 ± 0.06, respectively in the four samples (Jayaprakasha et al. 2002). The total percentages of curcuminoids are 2.34 ± 0.171 to 9.18 ± 0.232. The curcumin contents among subspecies of *C. longa* vary from high (more than 1800mg/100gDW) to low (50~100mg/100gDW) (Toshikazu et al. 2006). Therefore, understanding the bio-production of these different analogs in the species of the genus *Curcuma* will provide valuable information for its medicinal applications. Currently, because of the increasing demand in the pharmaceutical and food industries, there is a pressing need (Kinghorn, A. et al. 2011) to understand the dynamics of the productivity of curcumin and curcuminoids in the genus *Curcuma*. To understand the differences in metabolic production between species and cultivars, both the amounts of substrates and the activities of enzymes provide important clues.

In this study, we focused on the analysis of metabolite concentrations and gene expressions of enzymes in the curcuminoids biosynthesis pathways. To date, several attempts at crop improvement for turmeric have been performed to increase yield potential and to enhance the amounts of active ingredients such as curcuminoids and the sesquiterpenoids (Ravindran et al. 2007). Turmeric, however, is a cross-pollinated triploid species, which indicates its scarce pollen fertility. They were propagated traditionally by underground rhizomes and thus crop improvement programs in turmeric were largely restricted to clonal selection or induced mutations with subsequent selection. Even though there is a long history of turmeric cultivation, only a few improved varieties of turmeric are recognized in Japan. Recently, some cultivars of turmeric with higher contents of curcuminoids than the wild-type *C. longa*, were registered in the Japanese Plant Variety Protection facility (Ministry of Agriculture, Forestry and Fisheries, Tokyo Japan). *C. longa* 'Ougon,' a commercial cultivar, was bred in Okinawa, and *C. longa* 'Sekiyo,' another commercial cultivar, was originally discovered in Southeast Asia.

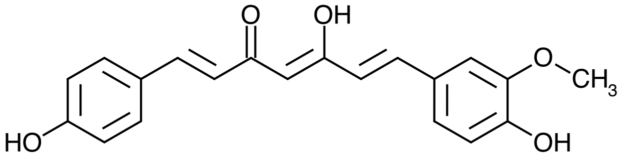
Curcumin [1,7-bis-(4-hydroxy-3-methoxyphenyl)hepta-1,6-diene-3,5-dione] **[Fig. 1.4]** is isolated as the pigment of the *C. longa* rhizome. It has wide variations in efficacy in anti-inflammatory, antiulcerogenic, and antitumor activities (Aggarwal et al. 2003). Demand for curcumin is rising because of its pharmaceutical usefulness as a dietary supplement. Curcuminoids, including de-methylated analogues of curcumin **[Fig. 1.4]**, have also received attention.

Curcuminoids are mainly accumulated in rhizomes of turmeric, resembling ginger roots. Here we focused on differences in curcuminoid contents among cultivars and analyzed gene expressions of their synthesizing enzymes to understand the metabolic reactions involved. We measured metabolite concentrations using liquid chromatography–mass spectrometry (LC-MS) and gas chromatography - mass spectrometry (GC-MS), and analyzed gene expressions using an RNA-Seq method

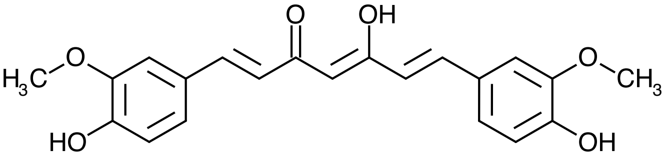
(Wang et al. 2009), for rhizome samples of the two wild-type variants (*C. aromatic* abbreviated as AR) and *C. longa* (LN)) and two cultivars (*C. longa* ‘Ougon’ (OU), and *C. longa*, ‘Sekiyo’ (SK)).



bisdemethoxy curcumin



demethoxy curcumin



curcumin

Figure 1.4: Three analogs of curcumin.

1.2 Specimens and reagents

Four samples of specimens utilized in this work are namely Okinawa *Curcuma longa*, Okinawa *Curcuma aromatica*, a new cultivar *Curcuma longa* ‘Sekiyo’ (“Sekiyo”, Japanese Plant Variety Protection No. 21486), and another cultivar *Curcuma longa* ‘Ougon’ (“Ougon”, Japanese Plant Variety Protection No. 21484) and be showed in **Table 1.1**. Three replicates of each specimen were planted in the Experimental Station for Medicinal Plant Research, University of Toyama following usual agronomical cultivation practices under the same conditions. The leaves and rhizomes were dug out after the aerial parts fully dried up. This period was considered as the initial period of dormancy. All specimens were deposited in the Museum of Materia Medica, College of Pharmaceutical Science, Ritsumeikan University (*C. longa*; Boucher No. RIN-CL-15, *C. aromatic*; Boucher No. RIN-CL-16, *C. longa*, Sekiyo; Boucher No. RIN-CL-17, *C. longa*, Okinawa Ougon; Boucher No. RIN-CL-18). All analytical grades of chemicals and LC-MS grades of chromatographic solvent reagents were purchased from Wako Chemical Co. Ltd (Tokyo, Japan). Standard compounds for HPLC quantitation of curcumin, demethoxy curcumin, and bisdemethoxy curcumin were purchased from Nagara Science Co. Ltd. (Gifu, Japan).

Table 1.1 Four sample of specimens utilized in this work

Sample name	Abbreviation	Data name from Toyama University	Sample names in Japanese	Scientific name
Curcuma aromatica_leaf	AR	Toyama-univ_1	春ウコン_葉	*Curcuma aromatica*
Curcuma aromatica_root	AR	Toyama-univ_2	春ウコン_根	*Curcuma aromatica*
Curcuma longa_leaf	LN	Toyama-univ_3	秋ウコン_葉	*Curcuma longa*
Curcuma longa_root	LN	Toyama-univ_4	秋ウコン_根	*Curcuma longa*
Curcuma ougon_leaf	OU	Toyama-univ_5	皇金_葉	*Curcuma longa*
Curcuma ougon_root	OU	Toyama-univ_6	皇金_根	*Curcuma longa*
Curcuma sekiyou_leaf	SK	Toyama-univ_7	赤陽_葉	*Curcuma longa*
Curcuma sekiyou_root	SK	Toyama-univ_8	赤陽_根	*Curcuma longa*

1.3 Next Generation Sequencing

Transcriptome Analysis

Transcriptome is the complete set of messenger RNA (mRNA) and noncoding RNA (ncRNA) transcripts produced by a particular cell, cell type, or organism. An intriguing problem in molecular biology is how the identical genetic make-up of cells can give rise to different cell types, each of which plays a defined role in the functioning of a multicellular organism. This phenotypic diversity has been linked to the fact that different cell types within the organism activate (or express) different sets of genes (transcriptomes) that lead to different cell fates and functions. The development of microarrays supplanted single-gene approaches by allowing simultaneous characterization of expression levels of thousands of known or putative transcripts (Sanger 1977). This method brought about a multitude of expression-profiling initiatives aiming to comprehensively characterize expression signatures of different cell types and disease states. Further developments in the microarray field enabled other transcriptomics applications, such as the detection of noncoding RNAs, single nucleotide polymorphisms (SNPs), and alternative splicing events (Mockler et al. 2005). Due to their cost efficiency, microarrays are a commonly used tool in transcriptomics research utilized in many laboratories around the world（Pozhitkov et al. 2007). However this method can not address several key aspects such as detect novel transcripts and ability to study the coding sequence of detected transcripts. Moreover, since microarrays are indirect methods in which transcript abundance is inferred from hybridization intensity rather than measured explicitly, the derived data are noisy, which interferes with reproducibility and cross-sample comparisons.

The development of expressed sequence tag (EST) sequencing in 1991 partially addressed the cost limitation of FLcDNA sequencing by introducing a less complete, less accurate, yet cheaper approach to the detection of expressed transcripts than was possible with sequencing full-length cDNA sequencing (FLcDNAs) (Boguski 1995). Despite the decrease in cost, however, EST sequencing with the Sanger method was still too expensive and labor intensive to be routinely used on a transcriptomewide scale. Moreover, due to the low redundancy of sequencing reads, EST data were not suitable for estimating transcript abundance.

The report of Serial Analysis of Gene Expression (SAGE) provided a key advance in transcriptome sequencing as it facilitated the use of Sanger sequencing for gene expression profiling (Velculescu VE, 1995). SAGE experiments offered many advantages over microarrays, such as the ability to detect novel transcripts, the ability to obtain direct measures of transcript abundance thus allowing easier comparisons between multiple samples, and the discovery of novel alternative splice isoforms. However, SAGE studies still involved a laborious cloning procedure, were costly, and produced short sequence tags (14 or 21 bp) that are difficult to resolve for transcripts with similar coding sequence.

As shown in **Figure 1.5**, Sanger sequencing was adopted as the primary technology in the "first generation" of laboratory and commercial sequencing applications due to its high efficiency and low radioactivity. At that time, DNA sequencing was laborious and radioactive materials were required. After years of improvement, Applied Biosystems introduced the first automatic sequencing machine (namely AB370) in 1987, adopting capillary electrophoresis which made the sequencing faster and more accurate. Since the early 1990s, DNA sequence production has almost exclusively been carried out with capillary-based, semi-automated implementations of the Sanger biochemistry (Shendure et al. 2008). This method is essentially tricking DNA polymerase into incorporating nucleotides with a slight chemical modification—the exchange of the 3′ hydroxyl group needed for chain elongation with a hydrogen atom that is functionally unable to participate in the reaction with the incoming nucleotide to extend the synthesized strand. Mixing proportions of the four native deoxynucleotides with one of four of their analogs, termed dideoxynucleotides, yields a collection of nucleotide-specific terminated fragments for each of the four bases (Sanger et al. 1977).

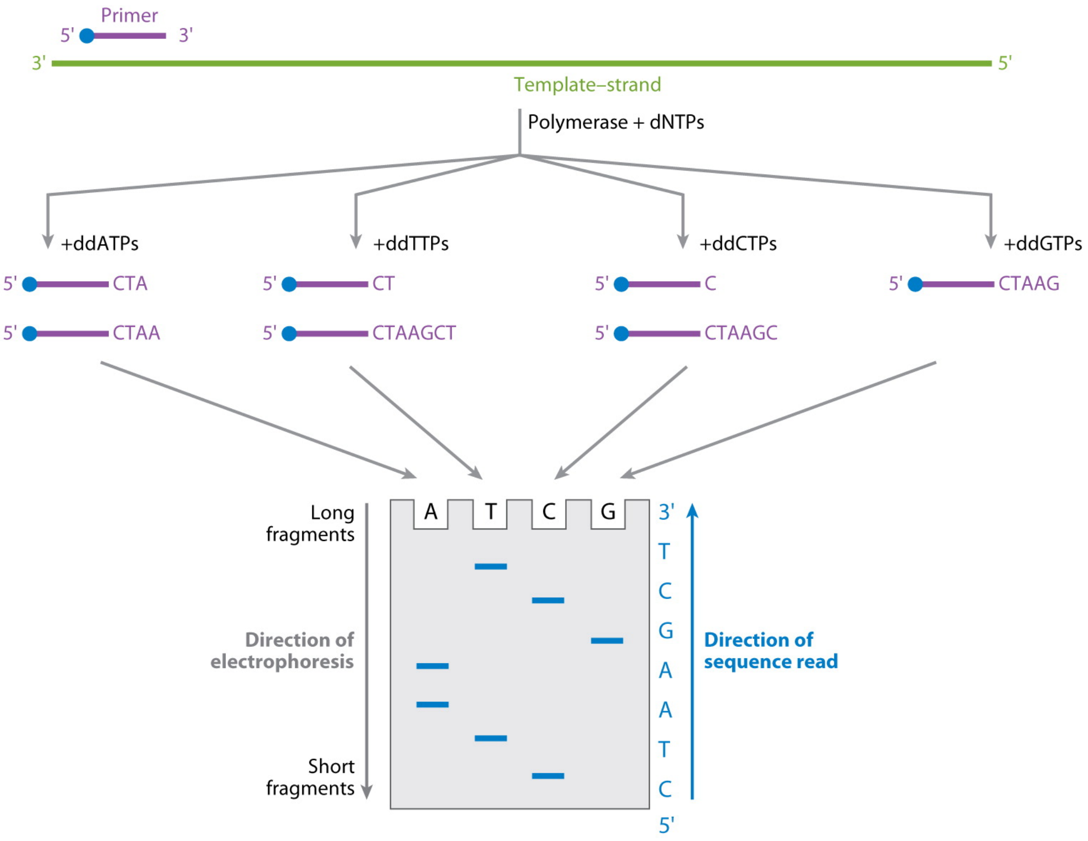


Figure 1.5 Sanger sequencing (Figure adapted from Mardis et al., 2013)

Newer methods are referred to as next-generation sequencing (NGS) (Mardis et al. 2008). And the NGS technologies are different from Sanger method in many aspects such as massively parallel analysis, high throughput, and relative lower cost.

Following the human genome project, the first NGS DNA sequencer GS20 was launched to the market by 454 Life Science in 2005, and Solexa released Genome Analyzer the next year, followed by (Sequencing by Oligo Ligation Detection) SOLiD provided from Agencourt, which are three most typical massively parallel sequencing systems in the NGS that shared good performance on throughput, accuracy, and cost compared with Sanger sequencing (**Table 1.2**). These founder companies were then purchased by other companies: in 2006 Agencourt was purchased by Applied Biosystems, and in 2007, 454 was purchased by Roche, while Solexa was purchased by Illumina. After years of evolution, these three systems exhibit better performance and their own advantages in terms of read length, accuracy, applications, consumables, man power requirement and informatics infrastructure, and so forth (Lin et al 2012).

These newer technologies constitute various strategies that rely on a combination of template preparation, sequencing and imaging, and genome alignment and assembly methods. The arrival of NGS technologies in the marketplace has changed the way we think about scientific approaches in basic, applied and clinical research. In some respects, the potential of NGS is akin to the early days of polymerase chain reaction (PCR), with one's imagination being the primary limitation to its use. The major advance offered by NGS is the ability to produce an enormous volume of data cheaply — in some cases in excess of one billion short reads per instrument run. This feature expands the realm of experimentation beyond just determining the order of bases. For example, in gene-expression studies, microarrays are now being replaced by seq-based methods, which can identify and quantify rare transcripts without prior knowledge of a particular gene and can provide information regarding alternative splicing and sequence variation in identified genes.

Table 1.2 Comparison of 3 systems of NGS and Sanger method

Sequencer	454 GS FLX	HiSeq 2000	SOLiDv4	Sanger 3730xl
Sequencing mechanism	Pyrosequencing	Sequencing by synthesis	Ligation and two-base coding	Dideoxy chain termination
Read length	700 bp	50SE, 50PE, 101PE	50 + 35 bp or 50 + 50 bp	400~900 bp
Accuracy	99.9%	98%, (100PE)	99.94% *raw data	99.999%
Reads	1 M	3 G	1200~1400 M	—
Output data/run	0.7 Gb	600 Gb	120 Gb	1.9~84 Kb
Time/run	24 Hours	3~10 Days	7 Days for SE 14 Days for PE	20 Mins~3 Hours
Advantage	Read length, fast	High throughput	Accuracy	High quality, long read length
Disadvantage	Error rate with polybase more than 6, high cost, low throughput	Short read assembly	Short read assembly	High cost low throughput

The ability of sequencing the whole genome of many related organisms has allowed large-scale comparative and evolutionary studies to be performed that were unimaginable just a few years ago. The broadest application of NGS may be the re-sequencing of human genomes to enhance our understanding of how genetic differences affect health and disease (Bras et al. 2012; Gonzaga-Jauregui et al. 2012). The variety of NGS features makes it likely that multiple platforms will coexist in the marketplace, with some having clear advantages for particular applications over others. Although NGS makes genome sequences handy, the followed data analysis and biological explanations are still the bottle-neck in understanding genomes.

1.4 Bioinformatics

In the RNA-Seq analysis, the whole transcriptome provides a huge number of short reads. In general, the raw reads are assembled to obtain expressed gene sequences and then each reads are aligned (mapped) to the assembled sequences first. Then following up reads count calculation reveals the expression level of the assembled sequences. For the short reads are from the whole mRNA of a cell, it contains all the gene information of the cell. The following up gene analysis focuses on the whole transcriptome/genome. Whole transcriptome analysis will give a all-round gene level information about the cell, which just correspond to the primary metabolic pathway analysis perfectly.

In chapter IV, we focus on a small scale of a secondary metabolic pathway -- curcuminoid synthesis pathway, which is less complex and small scale compare with the major metabolic pathways. The traditional whole transcriptome analysis is not very suitable for a specific secondary metabolic pathway analysis. Though the NGS data from transcriptome contains the all gene information including curcuminoid synthesis pathway genes. Mapping the short reads to the whole transcriptome let the multifarious gene information from primary metabolic pathway overshadowed the genes which we focused on. So we used a targeted mapping strategy--focusing on a small set of secondary metabolic pathway genes -- to analyze the gene expression.

To minimize analysis errors identified by measures such as read-quality control (Rismani-Yazdi et al. 2011, Patel and Mukesh 2012), we propose a "selection-first" RNA-Seq analysis, in which reads are mapped to particular biosynthetic pathways. Initially, we selected the gene sequences that were most homologous to the targeted gene sequences for known enzymes in the biosynthetic pathways of curcuminoids, and then we mapped the raw reads to the selected template sequences. Compared to the general "mapping-first" analysis in which all reads are mapped to the total assembled sequences by removing inaccurate templates, we expect to reduce the errors caused by misassembly and mismapping. This lowers the CVs of estimated expression levels of selected sequences.

Chapter II. Gene expression analysis

2.1 NGS data prepare and work flow

NGS analysis background

The NGS technique can be applied for different research like nucleotide mutation types detection, transcriptome analysis, and so on. One of the most common use is transcriptom/genome analysis. To get the transciptome information from billions of NGS short reads, people assemble the short reads into transcriptome in 3 kinds of strategy: a reference-based strategy, a *de novo* assemble strategy or a combined strategy that merges the two.

When a reference genome for the target transcriptome is available, the transcriptome assembly can be built upon it. In general, this strategy — which is known as 'reference-based' or 'ab initio' assembly — involves three steps. First, RNA-seq reads are aligned to a reference genome using a splice-aware aligner. Second, overlapping reads from each locus are clustered to build a graph representing all possible isoforms. The final step involves traversing the graph to resolve individual isoforms. With the reference-based transcriptome strategy, assembly can be solved uing parallel computing and requirement of the computer server is not too much. More importantly, the reference-based strategy is very sensitive and can assemble transcripts of low abundance (Jeffrey 2011).

The '*de novo*' transcriptome assembly strategy does not use a reference genome: it leverages the redundancy of short-read sequencing to find overlaps between the reads and assembles them into transcripts. The '*de novo*' assembly can provide an initial set of transcripts for it does not depend on any reference genome, allowing for RNA-seq expression studies. A second advantage of *de novo* assembly is that it does not depend on the correct alignment of reads to known splice sites (Burset et al. 2000) or the prediction of novel splicing sites, as required by reference-based assemblers.

To analyze the curcuminoid synthesis pathway mechanism, we select 4 cultivars (SK, OU, AR and LN as shown in **Table 1.1**) of turmeric.

We focused on the two cultivars, SK and OU by comparing them with the wild types to understand quantitative balances of curcuminoid biosynthesis in order to obtain clues for more efficient production of curcumin and curcuminoids. We analyzed the metabolite concentrations and gene expressions of these four specimens to clarify the relationships between species by curcuminoid metabolic pathway. Poly(A) + RNA was isolated from the frozen rhizomes from each specimens and cDNA libraries were constructed; the detailed protocol was described in (Suzuki 2004). The cDNA library was sequenced using the paired-end method with an Illumina HiSeq 1000 platform (Illumina Inc., San Diego, CA, USA). Each fragment was sequenced to a read length of 100 nucleotides from each end and we obtained more than 144 million pair reads. We used Trinity software (Grabherr et al. 2011) to assemble these reads to form scaffolds. The beginning of the curcuminoid synthesis is from more basic synthesis pathway such like glycolysis, and we also get 4 kinds of turmeric cultivates NGS short read data both from leaves and rhizomes. We use Trinity to assemble the short reads. **Table 2.1** shows the summary of the original data and assembling.

After getting the original NGS data of leaves and rhizomes from all 4 specimens, we used Trinity platform (http://trinityrnaseq.github.io/ Grabher, et al. 2011) to assemble the whole genome. Generally after the assembling of raw reads, all the reads would be mapped to the assembled sequences to calculate the sequence expression level. For all the reads are abstracted from all the mRNA/cDNA, this method is naturally used as genome/transcriptom analysis.

Table 2.1 8 samples original NGS data from 4 specimens of *Curcuma longa*

Sample name	Read numbers	Size
AR_leaf	9917832 × 2	5.23GB
AR_root	6706886 × 2	3.53GB
LN_leaf	6950578 × 2	3.66GB
LN_root	8829693 × 2	4.65GB
OU_leaf	8685678 × 2	4.58GB
OU_root	11669887 × 2	6.15GB
SK_leaf	12647947 × 2	6.67GB
SK_root	8987480 × 2	4.74GB
Total	74395981 × 2	39.21GB

2.2 Gene differential analysis between leaf and rhizome

Select genes with significantly change in expression level

The leaf is considered as the most active part of the whole plant. Many primary metabolic synthesis pathways like ATP, nucleotide, photosynthesis etc. are processed in the leaves. The curcuminoids are accumulated in the rhizome. The differences of gene expression level between in the leaf and rhizome will reveal the metabolic pathway diversity. We use NGS data from all 8 samples (4 are leaf data and 4 are rhizome data) to compare the gene expression changes between leaves and rhizome.

We use all 8 samples reads together to assemble the transcriptome. The Trinity assembler produce 302331 scaffolds. To eliminate redundant scaffolds (with tiny sequence difference only between the scaffolds) in the assembled scaffolds set, we clustered the scaffolds by Usearch software (http://www.drive5.com/usearch/; Edgar et al. 2010) and produce 192056 clustered scaffolds. Then the short reads of 8 samples were mapped to these 192056 scaffolds to identify the gene expression level of these 8 samples. The working flow is shown in **Figure 2.1**. Step 1, we firstly use 8 samples from both leaves and rhizomes of 4 cultivars to get sequenced by Illumina Hi-seq 1000 (Illumina Inc., San Diego, CA, USA). More than 74 million short reads are obtained by NGS. Step 2, We get 302311 assembled sequences scaffolds from 8 samples short reads together. The assembled sequence is called “scaffold”. Step 3, the Trinity assembled scafflolds are clustered by Usearch software (http://www.drive5.com/usearch/; Edgar et al. 2010), 192056 clustered scaffolds were selected. Step 4. Mapping 8 samples short reads to the clustered scaffolds separately. Step 5. The scaffold expression level was calculated by FPKM package from Trinity kit. The mapping result shows that the gene expression levels of different samples are different.

To compare the difference between the leaf and rhizome, the gene expression level ratio between the leaf and rhizome is calculate by using the following equation:

$$\mathrm{Gene}_i\ \mathrm{ratio} = \frac{\mathrm{Gene}_i\ \mathrm{expression\ level\ in\ leaf}}{\mathrm{Gene}_i\ \mathrm{expression\ level\ in\ rhizome}} \tag{1}$$

where *i* is from 1 to 192056 to represent the clustered scaffolds.

To calculate the ratio of the expression level of each gene, we firstly get rid of the genes which have 0 expression level records both in leaves and rhizome, which means the genes might be absent in the species. After removing the none-expressed genes separately from 4 cultivars, around 100000 scaffolds left for each cultivar (AR 94845 scaffolds, LN 97305 scaffolds, OU 92778 scaffolds, SK 103098 scaffolds left).

The gene ratio shows the extent of expression difference between the leaf and rhizome in a cultivar. We focused on especially high or low gene ratio which shows different biosynthesis pattern of a cell. Since the distribution of gene ratio can be approximated by Gaussian distribution [**Fig. 2.2**], we applied Z-test to evaluate significantly up/down regulated genes. To identify the significantly changed gens we set the H_0: gene expressed significantly high in the leaf compare with that in the root. Considering that around 100000 genes expressed for all 4 cultivars, we use a relatively strict threshold = 1/100000 for a doubled sided significance test. We selected the gene with gene ratio fold change (FC) bigger than 1 with the confidence level p-value smaller than 10^{-6}. The summary of significant genes shows in **Table 2.2, Figure 2.3**.

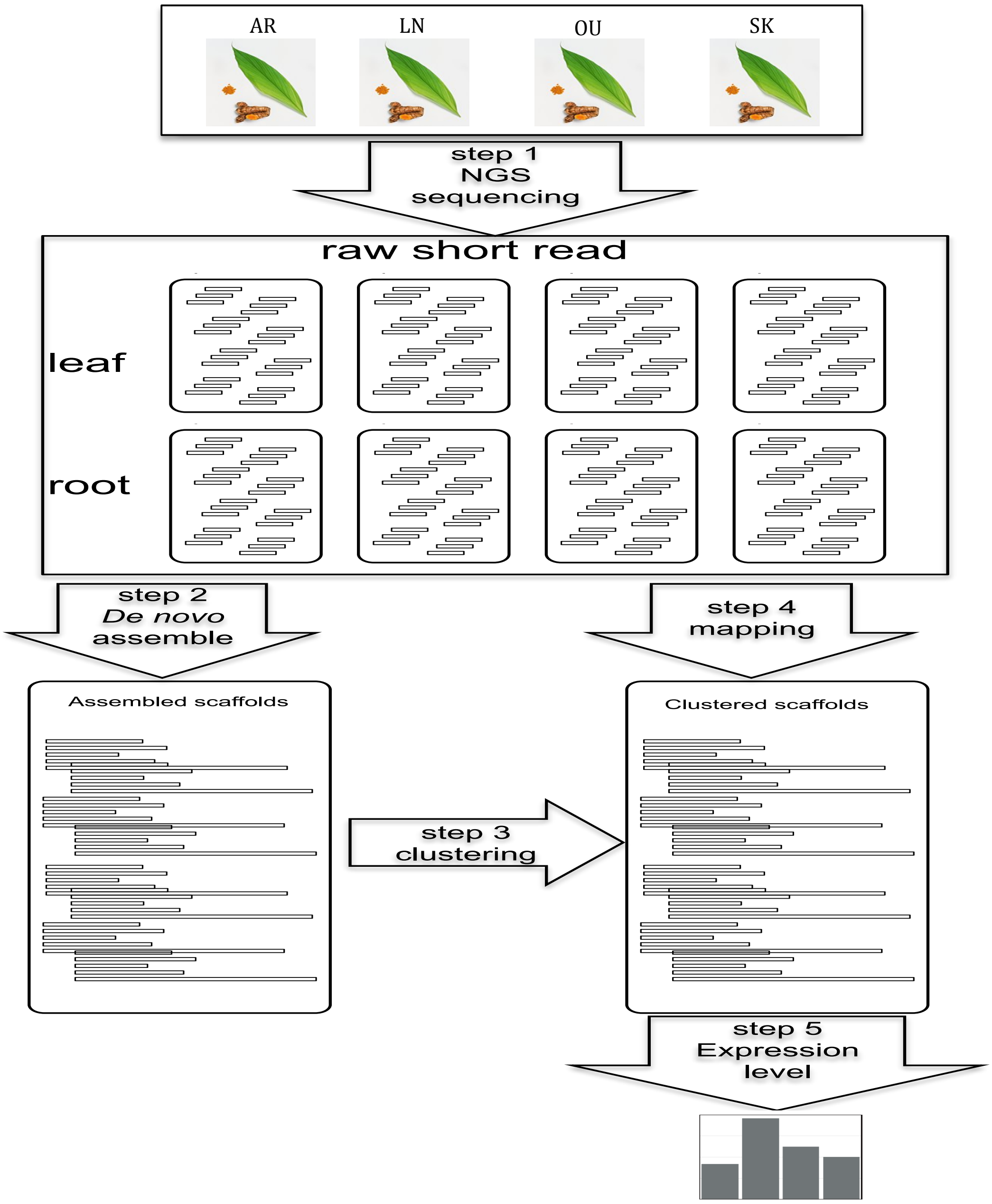


Figure 2.1 Work flow of NGS data preparation. NGS data from 4 specimens 8 sample is assembled and gene expression level is calculated.

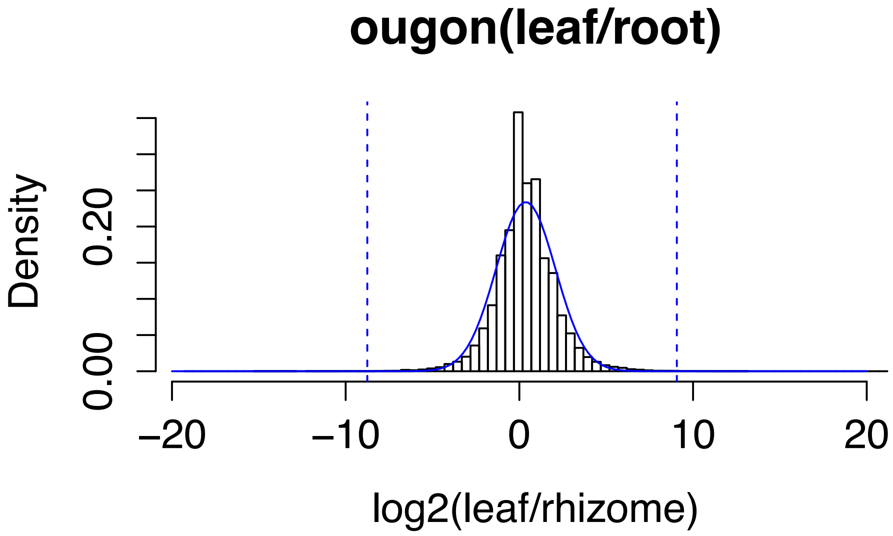


Figure 2.2 Example of *C. longa* ougon of genes expression level ratio distribution. We calculate the expression of each gene first. Then expression ratio between the leaf and root is shown as the histogram. The blue line is the FC ratio distribution. The blue dash lines are the confidence interval of significant FC with *p*-value equals 1/100000. By selecting the genes which appear out of the blue line range we can pick up the genes which expression significantly change between leaf and rhizome.

Table 2.2 Gene expression selection based on ratio = leaf/rhizome

	AR	LN	OU	SK
Genes No.*	94845	97305	92778	103098
High express in leaf	864	1880	2015	1303
High express in rhizome	706	1230	1355	1049
Mean of the ratio distribution	0.343	0.470	0.369	0.156
Sd of the ratio distribution	2.092	1.911	1.705	1.822

* We removed the genes that show 0 expression in both leaf and rhizome in every specimen.

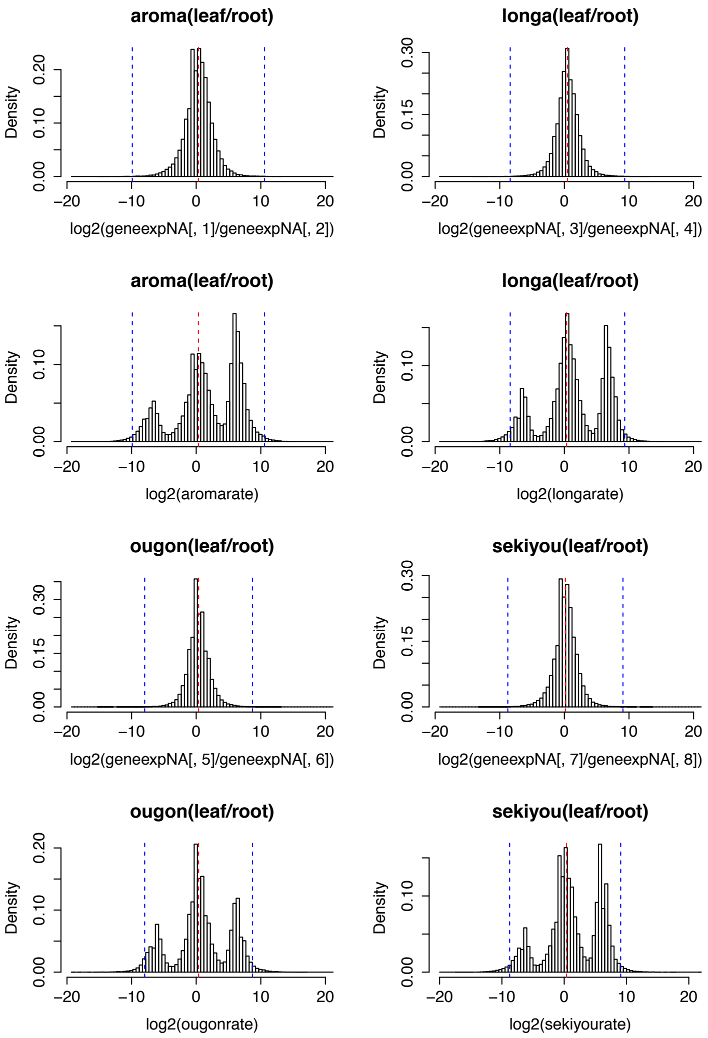


Figure 2.3 The histograms with single peak show the ratio distribution between leaf and root that all genes express both in leaf and root. The histograms with 3 peaks show the ratio distribution that the genes express level can be 0 either in leaf or root. For the expression level 0 can not be calculated by the gene ratio, we change all 0 expression to 0.01. That leads to the 2 more peaks. The red dash line indicates the mean of each sample, and the blue dash lines indicate the threshold of 1/100000. The genes exist between the blue dash lines are considered not change significantly enough. And genes outside the blue dash lines are selected as significant change genes.

2.3 Major metabolite pathway analysis

Mapping the significant genes to the major pathway

In biochemistry, a metabolite pathway is series of chemical reactions that modified the initial chemical compounds into diverse of product compounds in a cell. All these reactions are catalyzed by enzymes. The product of one reaction is catalyzed as the substrate for next step reaction. These enzymes often require dietary minerals, vitamins and other cofactors to function. Pathways are required for the maintenance of homeostasis within an organism and the flux of metabolites through a pathway is regulated depending on the needs of the cell and the availability of the substrate. The end product of a pathway may be used immediately, initiate another metabolic pathway or be stored for later use. The metabolism of a cell consists of an elaborate network of interconnected pathways that enable the synthesis and breakdown of molecules. The metabolic pathway plays import role in cell. Every metabolic pathway consists of a series of biochemical reactions that are connected by their intermediates: the products of one reaction are the substrates for subsequent reactions, and so on. Metabolic pathways are often considered to flow in one direction. Although all chemical reactions are technically reversible, conditions in the cell are often such that it is thermodynamically more favorable for flux to flow in one direction of a reaction.

To annotate the gene function we compared the all clustered scaffold to the *Arabidopsis* genome by BLAST software (http://blast.ncbi.nlm.nih.gov/Blast.cgi; Altschul et al. 1990). After annotation of each scaffold, the information of genes whose expression level significantly changes can be obtained from KEGG database (KEGG; http://www.genome.jp/kegg/; Kanehisa et al. 2000). The bio-reaction which is catalyzed by the scaffold is recorded. Then the bio-reaction can be cataloged by the pathway. Here is a 2-step mapping: from scaffold to reaction, and then to the pathway. All the genes that up-regulated in the leaf among 4 specimens are cataloged into 118 pathways **[Table 2.3]**.

Table 2.3 All up-regulated genes in leaf refer to 118 pathways.

No.	Pathway	Pathway description	Up-reg gene No.	Total gene No.
1	ath00500	Starch and sucrose metabolism	17	990
2	ath00520	Amino sugar and nucleotide sugar metabolism	16	696
3	ath01100	Metabolic pathways	185	8378
4	ath01110	Biosynthesis of secondary metabolites	82	4600
5	ath00600	Sphingolipid metabolism	7	189
6	ath04075	Plant hormone signal transduction	18	1198
7	ath04144	Endocytosis	9	779
8	ath04120	Ubiquitin mediated proteolysis	14	646
9	ath00260	Glycine, serine and threonine metabolism	11	283
10	ath00630	Glyoxylate and dicarboxylate metabolism	11	317
11	ath01130	Biosynthesis of antibiotics	46	1938
12	ath00330	Arginine and proline metabolism	9	293
13	ath01230	Biosynthesis of amino acids	31	1116
14	ath03008	Ribosome biogenesis in eukaryotes	9	387
15	ath03010	Ribosome	34	1863

16	ath00510	N-Glycan biosynthesis	7	199
17	ath04141	Protein processing in endoplasmic reticulum	31	1082
18	ath03040	Spliceosome	22	932
19	ath00561	Glycerolipid metabolism	8	361
20	ath00564	Glycerophospholipid metabolism	6	460
21	ath00020	Citrate cycle (TCA cycle)	8	274
22	ath00270	Cysteine and methionine metabolism	8	393
23	ath00620	Pyruvate metabolism	10	468
24	ath00710	Carbon fixation in photosynthetic organisms	12	412
25	ath01200	Carbon metabolism	33	1283
26	ath03015	mRNA surveillance pathway	7	656
27	ath04140	Regulation of autophagy	1	134
28	ath00400	Phenylalanine, tyrosine and tryptophan biosynthesis	5	154
29	ath03018	RNA degradation	16	549
30	ath04146	Peroxisome	8	355
31	ath00040	Pentose and glucuronate interconversions	9	252
32	ath00071	Fatty acid degradation	4	190
33	ath00592	alpha-Linolenic acid metabolism	3	189
34	ath01040	Biosynthesis of	4	167

		unsaturated fatty acids		
35	ath01212	Fatty acid metabolism	8	347
36	ath03022	Basal transcription factors	5	231
37	ath04712	Circadian rhythm - plant	8	283
38	ath00190	Oxidative phosphorylation	24	771
39	ath00195	Photosynthesis	9	213
40	ath02010	ABC transporters	2	140
41	ath00360	Phenylalanine metabolism	9	462
42	ath00940	Phenylpropanoid biosynthesis	13	677
43	ath00250	Alanine, aspartate and glutamate metabolism	7	215
44	ath00970	Aminoacyl-tRNA biosynthesis	9	235
45	ath00053	Ascorbate and aldarate metabolism	4	203
46	ath00562	Inositol phosphate metabolism	4	269
47	ath04070	Phosphatidylinositol signaling system	10	330
48	ath03410	Base excision repair	3	182
49	ath00280	Valine, leucine and isoleucine degradation	5	187
50	ath00410	beta-Alanine metabolism	4	155
51	ath00640	Propanoate metabolism	3	104
52	ath00511	Other glycan degradation	3	109
53	ath00790	Folate biosynthesis	2	95
54	ath04122	Sulfur relay system	2	47

55	ath00905	Brassinosteroid biosynthesis	1	83
56	ath00906	Carotenoid biosynthesis	4	157
57	ath00750	Vitamin B6 metabolism	2	67
58	ath00230	Purine metabolism	16	620
59	ath00240	Pyrimidine metabolism	9	544
60	ath03020	RNA polymerase	4	171
61	ath00061	Fatty acid biosynthesis	5	193
62	ath00780	Biotin metabolism	2	63
63	ath04145	Phagosome	8	510
64	ath03050	Proteasome	6	232
65	ath00290	Valine, leucine and isoleucine biosynthesis	1	89
66	ath00770	Pantothenate and CoA biosynthesis	2	98
67	ath00966	Glucosinolate biosynthesis	1	32
68	ath01210	2-Oxocarboxylic acid metabolism	5	263
69	ath00670	One carbon pool by folate	3	73
70	ath00010	Glycolysis / Gluconeogenesis	10	599
71	ath00941	Flavonoid biosynthesis	3	155
72	ath00945	Stilbenoid, diarylheptanoid and gingerol biosynthesis	2	140
73	ath00062	Fatty acid elongation	3	181
74	ath03013	RNA transport	17	801

75	ath00196	Photosynthesis - antenna proteins	3	92
76	ath00910	Nitrogen metabolism	4	168
77	ath04626	Plant-pathogen interaction	27	1082
78	ath00430	Taurine and hypotaurine metabolism	1	82
79	ath03060	Protein export	3	200
80	ath00514	Other types of O-glycan biosynthesis	2	21
81	ath00130	Ubiquinone and other terpenoid-quinone biosynthesis	6	206
82	ath00030	Pentose phosphate pathway	11	247
83	ath00051	Fructose and mannose metabolism	5	310
84	ath00900	Terpenoid backbone biosynthesis	3	273
85	ath00052	Galactose metabolism	6	261
86	ath00460	Cyanoamino acid metabolism	4	188
87	ath00450	Selenocompound metabolism	1	93
88	ath00100	Steroid biosynthesis	1	109
89	ath00860	Porphyrin and chlorophyll metabolism	4	149
90	ath00531	Glycosaminoglycan degradation	1	55

91	ath00920	Sulfur metabolismmail	2	125
92	ath00908	Zeatin biosynthesis	1	81
93	ath04130	SNARE interactions in vesicular transport	7	217
94	ath00730	Thiamine metabolism	1	84
95	ath00310	Lysine degradation	1	94
96	ath00340	Histidine metabolism	1	75
97	ath00380	Tryptophan metabolism	1	128
98	ath00903	Limonene and pinene degradation	1	121
99	ath00480	Glutathione metabolism	5	288
100	ath00590	Arachidonic acid metabolism	2	60
101	ath00350	Tyrosine metabolism	1	102
102	ath00950	Isoquinoline alkaloid biosynthesis	1	48
103	ath00960	Tropane, piperidine and pyridine alkaloid biosynthesis	1	97
104	ath00760	Nicotinate and nicotinamide metabolism	3	81
105	ath00261	Monobactam biosynthesis	1	27
106	ath00300	Lysine biosynthesis	1	49
107	ath00565	Ether lipid metabolism	2	134
108	ath00591	Linoleic acid metabolism	1	81
109	ath00072	Synthesis and degradation of ketone bodies	1	25
110	ath00650	Butanoate metabolism	2	67

111	ath03030	DNA replication	2	229
112	ath03420	Nucleotide excision repair	3	269
113	ath03430	Mismatch repair	3	176
114	ath03440	Homologous recombination	2	226
115	ath00904	Diterpenoid biosynthesis	1	91
116	ath01220	Degradation of aromatic compounds	1	44
117	ath00563	Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	1	88
118	ath00785	Lipoic acid metabolism	1	26

To understand the difference between the 4 cultivars, the significantly changed genes are considered from every cultivar **[APP. Table 1-4]**. After the BLAST selection, the scaffolds that are with function annotation based on homologous genes are kept. Then by searching the KEGG database we confirm the pathway to which the annotated scaffolds can be refereed. From the Appendix Tables, inspect the genes appearance in every cultivar. We specially focus on 4 pathways that have close relationship to the next secondary metabolic pathway of curcuminoid: carbon fixation in photosynthetic organisms, glycolysis metabolic pathway, phenylalanine synthesis pathway and pentose phosphate pathway.

AR: The genes which in Carbon fixation in photosynthetic organisms pathway, and in the primary synthesis of the phenylalanine synthesis, have a high expression level change ranking in NADP-dependent malic enzyme 2 (NADP-ME2), glyceraldehyde 3-phosphate dehydrogenase (GAPCP-2), glyceraldehyde 3-phosphate dehydrogenase (GAPC2), malate dehydrogenase 1 (mMDH1) and fructose-1,6-bisphosphatase (HCEF1) (**[APP. Table 1]** row 26, 36, 38, 93 and 116). We also find genes of glyceraldehyde 3-phosphate dehydrogenase (GAPCP-2), glyceraldehyde 3-phosphate dehydrogenase (GAPC2), fructose-1,6-bisphosphatase (HCEF1) have a significant change in glycolysis pathway (**[APP. Table 1]** row 36, 38, 116) that are also related to the phenylalanine synthesis.

LN: For Carbon fixation in photosynthetic organisms pathway, there are 6 genes: NADP-dependent malic enzyme 2 (NADP-ME2), malate dehydrogenase (MDH), phosphoenolpyruvate carboxylase 1 (PPC1), malate dehydrogenase 1 (mMDH1), ribose 5-phosphate isomerase A, fructose-1,6-bisphosphatase (HCEF1) (**[APP. Table 2]** row 21, 33,71, 87, 142, 161) have been found up-regulated in the leaf. And 4 expression level changed genes: NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (ALDH11A3), aldehyde dehydrogenase 3I1 (ALDH3I1), fructose-1,6-bisphosphatase (HCEF1), pyruvate dehydrogenase E1 beta (MAB1) (**[APP. Table 2]** row 38, 88, 161, 182) found in Glycolysis pathway.

OU: Four phenylalanine related genes are found up-regulated in the leaf: anthranilate synthase component I-2 (ASA2), tryptophan synthase beta chain (TSB2), bi-functional dehydroquinate-shikimate dehydrogenase (MEE32), tryptophan synthase alpha chain (**[APP. Table 3]** row 22, 59, 122, 184). Also other pathways like Pentose phosphate pathway related genes: NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (ALDH11A3), transketolasek, Aldolase-type TIM barrel family protein (**[APP. Table 3]** row 56, 89, 111), Carbon fixation in photosynthetic organisms genes pathway genes: NADP-dependent malic enzyme 2 (NADP-ME2), alanine aminotransferase (AlaAT1), transketolase, fructose-bisphosphate aldolase 2 (FBA2), malate dehydrogenase 1 (mMDH1) (**[APP. Table 3]** row 12, 61, 89, 178,193), Glycolysis pathway genes: NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (ALDH11A3), pyruvate dehydrogenase E1 beta (MAB1), plastidial pyruvate kinase 1 (PKP-ALPHA), 6-phosphofructokinase 3 (PFK3), FBA2; fructose-bisphosphate aldolase 2 (FBA2) (**[APP. Table 3]** row 56, 104, 131, 174, 178) are significantly changed genes in the leaf compare with in the rhizome.

SK: There are a lot of Carbon fixation in Photosynthetic organisms related genes: NADP-dependent malic enzyme 2 (NADP-ME2), ribose 5-phosphate isomerase A, phosphoenolpyruvate carboxykinase [ATP] (PCK1), phosphoenolpyruvate carboxylase 1 (PPC1), malate dehydrogenase 1 (mMDH1) (**[APP. Table 4]** row 21, 40, 45, 48, 90) and Pentose phosphate pathway related genes: ribose-phosphate pyrophosphokinase 2 (PRS2), ribose 5-phosphate isomerase A, glucose-6-phosphate dehydrogenase 1 (G6PD1), PRS2; ribose-phosphate pyrophosphokinase 2 (PRS2) (**[APP. Table 4]** row 23, 40, 88, 115) found up-regulated in the leaf.

Analysis of curcuminoid synthesis genes in major metabolite pathways

The RNA-seq based on NGS technology provides the whole transciptome of a cell. Not only the major metabolite pathway gene sequences but also secondary metabolite pathway genes also can be detected. Compare with the major metabolite pathway, secondary metabolite pathway would affect the final products more directly. We made a preliminary research on the curcuminoid synthesis genes analysis in major metabolite pathway analysis.

The curcumionid synthesis pathway from phenylalanine to curcumin and its analogs is available from KEGG (http://www.genome.jp/kegg/; Kanehisa et al. 2000) and MetaCyc (http://www.metacyc.org/; Caspi, et al 2008) database. Curcuminoid synthesis pathway refers to 11 different bio-reactions though only 4 out of these 11 reactions' enzymes/genes are identified in turmeric (CURS1, CURS2, CURS3, CUS). For other genes, we choose homologous genes from other species (ath: Arabidopsis thaliana, gmx: Glycine max, osa: Oryz sativa, zma: Zea mays), which catalyze the same reactions. The 11 genes are showed as **Table 2.4**.

Because of the curcumionds are accumulated in the rhizome, it is considered that the corresponding genes express high in the rhizome than in the leaf. We selected the FC of the express level of significantly changed genes in the rhizome **[Table 2.2]** (namely 706 up-regulated scaffolds in AR, 1230 in LN, 1355 in OU, and 1049 in SK). We compared these genes with 11 curcumin synthesis pathway genes with BLASTX and get 443 scaffolds annotated in total all through 4 cultivars **[App. Table 5]**. Some of the genes show significant FC in only one species but also there are genes shared by more than one species. **Figure 2.4** gives the detail information of the up-regulated genes in rhizome of all 4 cultivars.

Table 2.4 11 genes of curcuminoid synthesis pathway

enzyme	full name	KEGG reaction No.
PAL	phenylalanine ammonia-lyase	4.3.1.24
4CL	4-coumarate:CoA ligase	6.2.1.-
C4H	cinnamate-4-hydroxylase	1.14.13.11
HCT	hydroxycinnamoyl transferase	2.3.1.133
C3H	cinnamate-3-hydroxylase	1.14.13.36
OMT	O-methyltransferase	2.1.1.104
DCS	phenylpropanoylacetyl-CoA synthase	2.3.1.218
CURS1	Curcumin synthase	2.3.1.217
CURS2	Curcumin synthase	2.3.1.217
CURS3	Curcumin synthase demethoxycurcumin synthase	2.3.1.217 2.3.1.219
CUS	bisdemethoxycurcumin synthase	2.3.1.211

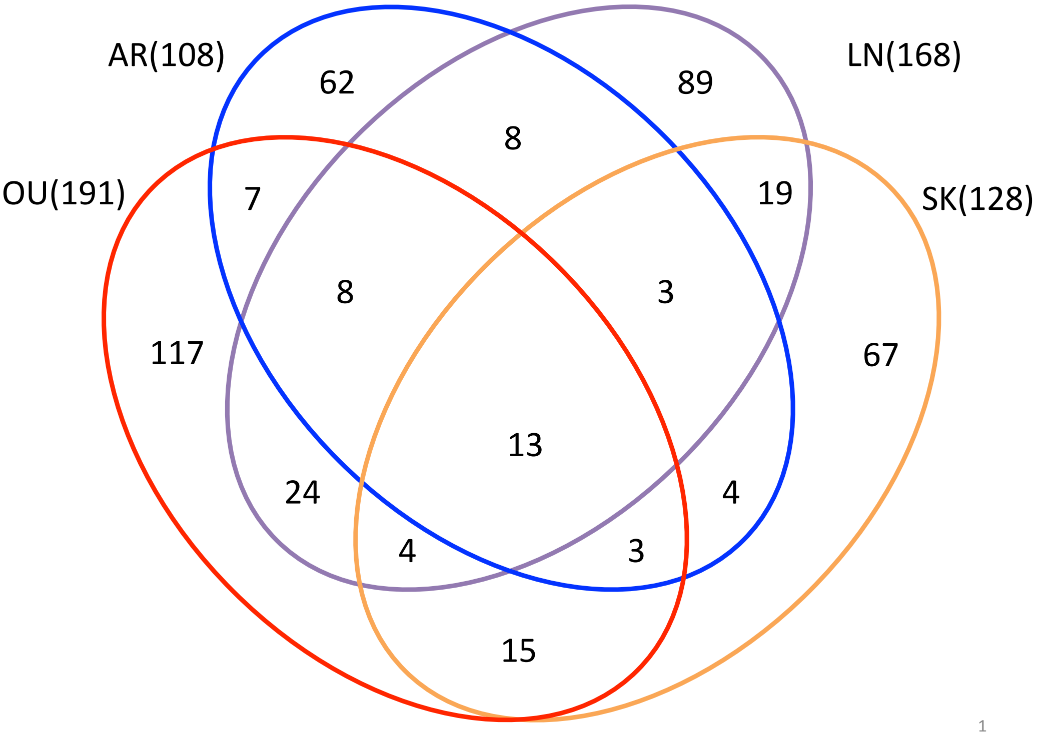


Figure 2.4 Venn diagram of 443 up-regulated genes in the rhizome among 4 cultivars are annotated by 11 curcumionoid synthesis pathway genes (namely 108 scaffolds in AR, 168 in LN, 191 in OU, and 128 in SK). There are 13 up-regulated genes shared by all 4 cultivars. And we also have a preliminary research on combination of LN and OU, which shared 24 up-regulated genes.

There are genes only up-regulated in a specific cultivar, and also some other genes have been found in 2, 3 or 4 cultivars. We summarized every possible combination and get all these genes annotated **[Table 2.5]**. Total 13 genes can be found up-regulated in the rhizome in all 4 cultivars.

We also have a preliminary research on the combination of LN and OU, which shared the most genes among all possible combinations. There are 24 genes found to be significantly up-regulated in the rhizome both in LN and OU. Among these 24 genes we found many genes appear in the same pathway: 2 genes for Ribosome; 2 genes for Oxidative phosphorylation; 2 genes for RNA transport, 2 genes for Plant circadian rhythm, 2 genes for Plant-pathogen interaction, 2 genes for Plant hormone signal transduction, 2 genes for Carbon metabolism. Considering the fact that we used a very strict threshold to distinguish significantly changed genes, LN and OU have many genes which can be found up-regulated in both of them. This fact shows that LN and OU have similar curcuminoid synthesis pattern. The statistical support will be discussed on the next step of research.

Table 2.5 Shared up-regulated genes in rhizome of 4 cultivars

AR + LN (8 genes)

	Gene	Pathway
1	SCL28; SC35-like splicing factor 28	Spliceosome
2	RR2; transcription factor response regulator 2	Plant hormone signal transduction
3	CNX7; molybdenum cofactor synthesis family protein	Folate biosynthesis; Metabolic pathways; Sulfur relay system
4	EP3; chitinase class IV	Amino sugar and nucleotide sugar metabolism
5	CI51; NADH dehydrogenase [ubiquinone] flavoprotein 1	Oxidative phosphorylation; Metabolic pathways
6	RNA recognition motif-containing protein	Spliceosome
7	FD3; ferredoxin 3	Photosynthesis
8	LHCA5; photosystem I light harvesting complex protein 5	Photosynthesis - antenna proteins

AR + OU (7 genes)

	Gene	Pathway
1	ESP3; DEAH RNA helicase homolog PRP2	Spliceosome
2	CYCD3;2; cyclin-D3-2	Plant hormone signal transduction
3	TWN2; Valyl-tRNA synthetase	Aminoacyl-tRNA biosynthesis
4	NCED4; nine-cis-epoxycarotenoid dioxygenase 4	Carotenoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
5	BCAT3; branched-chain-amino-	Valine, leucine and isoleucine degradation;

	acid aminotransferase 3	Valine, leucine and isoleucine biosynthesis; Pantothenate and CoA biosynthesis; Glucosinolate biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; 2-Oxocarboxylic acid metabolism; Biosynthesis of amino acids
6	dehydratase-enolase-phosphatase complex 1	Cysteine and methionine metabolism; Metabolic pathways
7	ZTL; adagio protein 1	Circadian rhythm - plant

AR + SK (4 genes)

	Gene	Pathway
1	LCB1; serine palmitoyltransferase	Sphingolipid metabolism; Metabolic pathways
2	GLX1; glyoxalase I homolog GLX1	Pyruvate metabolism
3	ATP3; ATP synthase subunit gamma	Oxidative phosphorylation; Metabolic pathways
4	NRPE5; DNA-directed RNA polymerase V subunit 5A	Purine metabolism; Pyrimidine metabolism; Metabolic pathways; RNA polymerase

LN + SK (19 genes)

	Gene	Pathway
1	PBD1; 20S proteasome beta subunit D1	Proteasome
2	HA5; H(+)-ATPase 5	Oxidative phosphorylation
3	ribose 5-phosphate isomerase A	Pentose phosphate pathway; Carbon fixation in photosynthetic organisms; Metabolic pathways;

		Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
4	40S ribosomal protein S20-1	Ribosome
5	Glycosyl hydrolase family protein	Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways
6	UGP2; UTP--glucose-1-phosphate uridylyltransferase 1	Pentose and glucuronate interconversions; Galactose metabolism; Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways; Biosynthesis of antibiotics
7	GSTL2; glutathione transferase lambda 2	Glutathione metabolism
8	PETC; cytochrome b6-f complex iron-sulfur subunit	Photosynthesis; Metabolic pathways
9	GPX1; phospholipid hydroperoxide glutathione peroxidase 1	Glutathione metabolism; Arachidonic acid metabolism
10	SKP2A; F-box protein SKP2A	Ubiquitin mediated proteolysis
11	PPC1; phosphoenolpyruvate carboxylase 1	Pyruvate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Carbon metabolism
12	PPa3; pyrophosphorylase 3	Oxidative phosphorylation
13	SDN3; small RNA degrading nuclease 3	Ribosome biogenesis in eukaryotes
14	CNX1; calnexin 1	Protein processing in endoplasmic reticulum; Phagosome

15	HSP60; heat shock protein 60	RNA degradation
16	phosphoinositide phospholipase C 6	Inositol phosphate metabolism; Metabolic pathways; Phosphatidylinositol signaling system
17	AK-LYS1; aspartokinase 1	Glycine, serine and threonine metabolism; Monobactam biosynthesis; Cysteine and methionine metabolism; Lysine biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; 2-Oxocarboxylic acid metabolism; Biosynthesis of amino acids
18	RPL24A; 60S ribosomal protein L24-1	
19	transport protein SEC13A	RNA transport; Protein processing in endoplasmic reticulum

LN + OU (24 genes)

	Gene	Pathway
1	60S ribosomal protein L22-2	Ribosome
2	60S ribosomal protein L15-1	Ribosome
3	ATP3; ATP synthase subunit gamma	Oxidative phosphorylation; Metabolic pathways
4	ubiquinol-cytochrome C reductase hinge protein	Oxidative phosphorylation; Metabolic pathways
5	CYL1; alpha-N-acetylglucosaminidase	Glycosaminoglycan degradation; Metabolic pathways
6	CKX1; cytokinin dehydrogenase 1	Zeatin biosynthesis
7	EIF4A-III; DEAD-box ATP-	RNA transport; mRNA surveillance pathway;

	dependent RNA helicase 2	Spliceosome
8	eukaryotic translation initiation factor 3 subunit 7	RNA transport
9	GI; protein GIGANTEA	Circadian rhythm - plant
10	FT; protein FLOWERING LOCUS T	Circadian rhythm - plant
11	ATHAL3B; putative phosphopantothenoylcysteine decarboxylase	Pantothenate and CoA biosynthesis; Metabolic pathways
12	ARR9; two-component response regulator ARR9	Plant hormone signal transduction
13	hypothetical protein	RNA degradation
14	SYP124; syntaxin-124	SNARE interactions in vesicular transport
15	ARR4; two-component response regulator ARR4	Plant hormone signal transduction
16	ACO1; aconitate hydratase 1	Citrate cycle (TCA cycle); Glyoxylate and dicarboxylate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; 2-Oxocarboxylic acid metabolism; Biosynthesis of amino acids
17	ALDH11A3; NADP-dependent glyceraldehyde-3-phosphate dehydrogenase	Glycolysis / Gluconeogenesis; Pentose phosphate pathway; Metabolic pathways; Carbon metabolism
18	PWWP domain-containing protein	Cysteine and methionine metabolism; Metabolic pathways
19	RPM1; disease resistance	Plant-pathogen interaction

	protein RPM1	
20	CAM7; calmodulin 7	Phosphatidylinositol signaling system; Plant-pathogen interaction
21	acyl-CoA thioesterase family protein	Fatty acid elongation; Biosynthesis of unsaturated fatty acids; Metabolic pathways; Biosynthesis of secondary metabolites
22	glutamine-dependent NAD(+) synthetase	Nicotinate and nicotinamide metabolism; Metabolic pathways
23	ELC-Like; ELC-like protein	Endocytosis
24	BGLU3; beta-glucosidase 3	Cyanoamino acid metabolism; Starch and sucrose metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites

OU + SK (15 genes)

	Genes	Pathway
1	argininosuccinate synthase	Arginine biosynthesis; Alanine, aspartate and glutamate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Biosynthesis of amino acids
2	PSAT; phosphoserine aminotransferase	Glycine, serine and threonine metabolism; Vitamin B6 metabolism; Metabolic pathways; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
3	OB-fold nucleic acid binding domain-containing protein	DNA replication; Nucleotide excision repair; Mismatch repair;

		Homologous recombination
4	pre-mRNA-splicing factor 38A	Spliceosome
5	SIP2; putative galactinol--sucrose galactosyltransferase 2	Galactose metabolism
6	SYP71; syntaxin-71	SNARE interactions in vesicular transport
7	splicing factor PWI and RNA recognition motif-containing protein	Spliceosome
8	GPAT6; bifunctional sn-glycerol-3-phosphate 2-O-acyltransferase/phosphatase	Glycerolipid metabolism; Glycerophospholipid metabolism; Metabolic pathways; Biosynthesis of secondary metabolites
9	oligosaccharyltransferase complex/magnesium transporter family protein	N-Glycan biosynthesis; Metabolic pathways; Protein processing in endoplasmic reticulum
10	CAM3; calmodulin 3	Phosphatidylinositol signaling system; Plant-pathogen interaction
11	Pre-mRNA-splicing factor 3	Spliceosome
12	ATERDJ3B; DNAJ heat shock protein ATERDJ3B	Protein processing in endoplasmic reticulum
13	SAG24; 60S ribosomal protein L10-3	Ribosome
14	GA1; Ent-copalyl diphosphate synthase	Diterpenoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
15	CHIP; E3 ubiquitin-protein ligase CHIP	Ubiquitin mediated proteolysis; Protein processing in endoplasmic reticulum

AR + LN (8 genes)

	Gene	Pathway
1	PHYB; phytochrome B	Circadian rhythm - plant
2	putative glucosamine-fructose-6-phosphate aminotransferase	Alanine, aspartate and glutamate metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways; Biosynthesis of antibiotics
3	Beta-glucosidase, GBA2 type family protein	Other glycan degradation; Sphingolipid metabolism; Metabolic pathways
4	SDH2-1; succinate dehydrogenase [ubiquinone] iron-sulfur subunit 1	Citrate cycle (TCA cycle); Oxidative phosphorylation; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
5	palmitoyl protein thioesterase family protein	Fatty acid elongation; Metabolic pathways; Fatty acid metabolism
6	RNA recognition motif-containing protein	Spliceosome
7	transport protein SEC13A	RNA transport; Protein processing in endoplasmic reticulum
8	RPL27; 50S ribosomal protein L27	Ribosome

AR + LN + SK (3 genes)

	Gene	Pathway
1	mMDH1; malate dehydrogenase 1	Citrate cycle (TCA cycle); Cysteine and methionine metabolism; Pyruvate metabolism; Glyoxylate and dicarboxylate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary

		metabolites; Biosynthesis of antibiotics; Carbon metabolism
2	EER4; transcription initiation factor TFIID subunit 12B	Basal transcription factors
3	UBC9; ubiquitin conjugating enzyme 9	Ubiquitin mediated proteolysis; Protein processing in endoplasmic reticulum

AR + OU + SK (3 genes)

	Gene	Pathway
1	SNRK2.10; SNF1-related protein kinase 2.10	Plant hormone signal transduction
2	GPAT4; bifunctional sn-glycerol-3-phosphate 2-O-acyltransferase/phosphatase	Glycerolipid metabolism; Glycerophospholipid metabolism; Metabolic pathways; Biosynthesis of secondary metabolites
3	TSB2; tryptophan synthase beta chain	Glycine, serine and threonine metabolism; Phenylalanine, tyrosine and tryptophan biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of amino acids

LN + OU + SK (4 genes)

	Gene	Pathway
1	Bet1-like protein	SNARE interactions in vesicular transport
2	RCI3; peroxidase 3	Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
3	endonuclease/exonuclease/phosp	Base excision repair

	hatase family protein	
4	UBC1; ubiquitin-conjugating enzyme E2 1	Ubiquitin mediated proteolysis

AR + LN + OU + SK (13 genes)

	Gene	Pathway
1	DNAse I-like superfamily protein	RNA degradation
2	ACX3; acyl-coenzyme A oxidase 3	Fatty acid degradation; alpha-Linolenic acid metabolism; Biosynthesis of unsaturated fatty acids; Metabolic pathways; Biosynthesis of secondary metabolites; Fatty acid metabolism; Peroxisome
3	FAB1B; phosphatidylinositol-3P 5-kinase-like	Inositol phosphate metabolism; Phosphatidylinositol signaling system; Phagosome
4	EIF2_GAMMA; translation initiation factor eIF-2 gamma subunit	RNA transport
5	NADP-ME2; NADP-dependent malic enzyme 2	Pyruvate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Carbon metabolism
6	40S ribosomal protein S16-1	Ribosome
7	VPS60.2; vacuolar protein sorting protein 60.2	Endocytosis
8	RD19; cysteine proteinase RD19a	Plant-pathogen interaction
9	hypothetical protein	Taurine and hypotaurine metabolism; Metabolic

		pathways
10	APC10; anaphase-promoting complex subunit 10	Ubiquitin mediated proteolysis
11	E3 ubiquitin-protein ligase SINAT3	Ubiquitin mediated proteolysis
12	Preprotein translocase Sec, Sec61-beta subunit protein	Protein export; Protein processing in endoplasmic reticulum; Phagosome
13	protein MIDASIN1	Ribosome biogenesis in eukaryotes

Chapter III. Rhizome gene expression analysis

3.1 Introduction

Curcuminoids mainly accumulate in rhizomes of turmeric, resembling ginger roots. Here we focused on differences in curcuminoids contents among cultivars and analyzed gene expressions of their synthesizing enzymes to understand the metabolic reactions involved. We measured metabolite concentrations using LC-MS and GC-MS, and analyzed gene expressions using an RNA-Seq method (Wang et al. 2009), for rhizome samples of the two wild-type variants.

3.2 Methodology

Analytical instruments

LC-MS analyses were performed using a Shimadzu LC-IT-TOF mass spectrometer (Shimadzu Corp., Kyoto, Japan) equipped with an electrospray ionization (ESI) interface. ESI parameters were as follows: source voltage +4.5 kV, capillary temperature 200 °C, nebulizer gas flow rate 1.5 L/min. The mass spectrometer was operated in the positive ion mode scanning from m/z 50 to 2000. LC–MS was operated using a high-performance liquid chromatography mode or an infusion injection mode. A Waters Atlantis T3 column (2.1 mm i.d. × 150 mm) was used (Waters, Milford, MA, USA), and the column temperature was maintained at 40 °C. The mobile phase was a binary eluent of (A) 0.1% HCOOH solution and (B) CH_3CN under the following gradient conditions: 0–30 min linear gradient from 20% to 100% CH_3CN, 30–40 min isocratic at 100% CH_3CN. The flow rate was 0.2 mL/min.

For GC–MS analysis, a Shimadzu QP2010 mass spectrometer equipped with Shimadzu GC2010 gas chromatograph system was used (Shimadzu Corp.). The column was a fused-silica capillary column, DB-1 MS (30 m × 0.25 mm i.d., film thickness 0.25 μm). The injector and transfer line were maintained at 280 °C. The oven temperature was programmed as follows: initial temperature, 50 °C; initial hold, 2 min; temperature program rate, 10 °C/min; final temperature, 300 °C; final hold 10 min. The flow rate of the carrier gas (helium) was 1 mL/min. The following conditions were used for mass spectrometry: ionization, EI mode; ionization current, 60 μA; ionization voltage, 70 eV. For GC–MS peak annotation, the KNApSAcK database (http://kanaya.naist.jp/KNApSAcK/; Afendi et al. 2012; Ikeda et al. 2013), NIST mass spectral database (http://www.nist.gov/srd/nist1a.cfm) and Wiley NBS mass spectral database were used.d

LC–MS sample preparation

Frozen specimens were individually pulverized to a fine powder using a multibead shocker (Model MB755U, Yasui Kikai Co., Osaka, Japan). Two grams of fine powder from each sample was weighed accurately and sonicated for 30 min with methanol (20 mL). After centrifugation, each extract was filtered through a 0.2 μm Millipore filter (Millipore Corp., Billerica, MA, USA). One microliter aliquots of the solvent were injected into the LC–MS equipment.

GC–MS sample preparation

Frozen rhizomes from individual plant specimens were homogenized to a fine powder using a multibead shocker, as above. A subsample was transferred to a 10 mL glass tube and weighed to 100 mg. Five milliliters of ethyl acetate was added to the tube, and the sample was extracted under sonication for 30 min. After being centrifuged (10,000 *g*, 2 min, 4 °C), the organic solvent was collected.

Strategy of RNA-Seq transcriptome for targeted metabolic pathways

In general whole genome information analysis is not suitable for targeted gene network analysis. For this reason, we develop a targeted genes mapping strategy. We analyzed the expression of enzymes in the curcuminoid biosynthesis pathway by an RNA-Seq method using selection-first expression analysis according to the schematic illustration [**Fig. 3.1**] to the targeted metabolic pathways. Each raw read of the four specimens were mapped to the references to count the number of assigned reads. Finally, to estimate the expression level of each gene, the number of fragments per kilobase of exon per million fragments (FPKM) was computed using the method proposed by Li and Dewey (2011). When multiple scaffolds were assigned to an identical enzyme gene, we simply added all expression values. The final gene expression results are showed in **Figure 3.2**. There are two choices for the following step: a. general 'mapping first' strategy. Firstly all the short reads will be mapped to the assembled genome to calculate every gene expression level by FPKM. Then the homologous comparison by BLAST will be applied between the genome and the reference sequences to select target genes. Here we use 11 genes from curcuminoids synthesis pathway (**Table 2.4**). We use Coefficent of Variation (CV) to evaluate the efficiency of this 'select-first' method compared with the general 'mapping first' strategy. Below we discuss the steps of mapping first and select first strategy separately.

For the mapping first strategy **[Fig. 3.1A]**:

1. All the raw reads will be assembled to genome (by Trinity software).
2. The short reads are mapped to the assembled genome to calculate the every scaffold express level (FPKM value) in the genome (by Trinity software).
3. Homologous comparison between the genome and reference gene database by BLAST to select candidate scaffold. For each Gene i ($i \in 1, \dots n$) from the reference gene database, there are always multiple scaffolds from assembled genome give good comparison hits. In this study the entries with the e-value smaller than 10^{-40} will be kept as gene candidates for each reference gene. For each Gene i from reference gene set, $\mathrm{scaffold}_{(i,1)}$, $\mathrm{scaffold}_{(i,2)}$, …, $\mathrm{scaffold}_{(i,m_i)}$

are corresponding candidate sequence with the BLAST e-value smaller than 10^{-40}. The $\text{scaffold}_{(i,1)}$, $\text{scaffold}_{(i,2)}$, …, $\text{scaffold}_{(i,m_i)}$ have corresponding FPKM value $\text{FPKM}_{(i,1)}$, $\text{FPKM}_{(i,2)}$, …, $\text{FPKM}_{(i,m_i)}$ respectively, which have already been calculated in Step 2. Notice that m_i is the total number of scaffolds corresponding to Gene *i*.

4.

$$\text{Mean}_i = \frac{1}{m_i}(\text{FPKM}_{(i,1)} + \text{FPKM}_{(i,2)} + \cdots + \text{FPKM}_{(i,m_i)}) \quad (1)$$

$$\text{Standard Deviation } i = \sqrt{\frac{1}{m_i}\sum_{k=1}^{m_i}(\text{FPKM}_{(i,k)} - \text{Mean}_i)} \quad (2)$$

$$\text{CV } i = \frac{\text{Standard Deviation } i}{\text{Mean}_i} \quad (3)$$

5. The average CV of all Gene i $(i \in 1, \ldots n)$ is:

$$\text{average CV} = \frac{\text{CV 1} + \text{CV 2} + \cdots + \text{CV n}}{\text{n}} \quad (4)$$

For the select first strategy **[Fig. 3.1B]**:

1. All the raw reads will assemble to genome (by Trinity software).
2. Homologous comparison between the genome and reference gene database by BLAST to select candidate scaffold. For each Gene i $(i \in 1, \ldots n)$ from the reference gene database, the entries with e-value smaller than 10^{-40} will be kept as candidate scaffolds.
3. All the short read will be mapped to the candidate scaffold to calculate the expression level (FPKM). For every Gene i $(i \in 1, \ldots n)$, corresponding scaffolds are $\text{scaffold}_{(i,1)}$, $\text{scaffold}_{(i,2)}$, …, $\text{scaffold}_{(i,m_i)}$ and also the corresponding FPKM value: $\text{FPKM}_{(i,1)}$, $\text{FPKM}_{(i,2)}$, …, $\text{FPKM}_{(i,m_i)}$
4. For each Gene *i*, we calculate the CV *i* in previous Step 4 Eq. (3).
5. The average CV of all Gene i $(i \in 1, \ldots n)$ is calculated by Eq. (4).

The CV measures the extent of variability in relation to the mean of the population. Lower CV shows smaller error of the sample. Comparing with standard deviation, CV is useful because the standard deviation of data must always be understood in the context of the mean of the data. The actual value of the CV is independent of the data in which the measurement has been taken.

In this study a group of 11 genes **[Table 2.4]** from curcumiod synthesis pathway are selected as reference gene database. Out of them 4 genes are already identified in turmeric (CURS1, CURS2, CURS3, CUS). But for other 7 genes, we just use the homologous genes from other species like Arabidopsis thalina, Glycine max, Oryz sativa and Zea mays. The average CV of the estimated expression level (select-first strategy) for each enzyme was 1.47 **[Fig. 3.1B]**, whereas the average over the same enzyme was 2.12 when the raw reads were mapped to the whole scaffold first (mapping first strategy) **[Fig. 3.1A]**. We also make pairwise t-test of the gene expression CV and the *p*-value shows great significance (*p*-value = 0.002). The detail is shown below:

Gene	CV(Select-first)	CV(Mapping first)	
C4H	1.930253147	1.200540752	AR(rhizome)
C3H	2.15691326	2.350557554	
OMT	0.607145613	2.7430908	
HCT	1.893796029	2.393337567	
PAL	1.701821048	2.4973892	
4CL	1.905663244	1.253397416	
CURS1			
CURS2	0.960886399		
CURS3	1.164113137	1.157909683	
DCS	0.991800027	2.3456167	
C4H	1.369285752	1.484140155	LN(rhizome)
C3H	2.33122185	2.057356	
OMT	0.801684242	3.4142136	
HCT	1.389514724	1.725945812	
PAL	1.726117766	3.6797124	
4CL	1.662068754	1.162122075	

CURS1			
CURS2	0.963588616		
CURS3	0.967049403	1.026949251	
DCS	1.722936229	1.9405002	
C4H	1.915677063	2.251602428	OU(rhizome)
C3H	1.505785729	1.763718146	
OMT	0.797309143	3.1468012	
HCT	1.502830697	1.518342096	
PAL	1.573914293	3.8253652	
4CL	2.2820564	1.330018739	
CURS1			
CURS2	0.886047917		
CURS3	1.394990222	1.416866028	
DCS	1.852040511	2.7494445	
C4H	1.377208391	1.166525446	SK(rhizome)
C3H	1.564141038	3.058198751	
OMT	1.60041939	3.1934048	
HCT	1.153789531	1.63707193	
PAL	1.717812329	3.8140662	
4CL	1.932395894	1.169040162	
CURS1			
CURS2	0.806773094		
CURS3	1.154014767	1.169554861	
DCS	1.480844936	2.3232882	
Average CV	1.465108627	2.123940245	
pairwised t-test	*p*-value = 0.002426242		

(In select-first strategy, only 2 scaffolds show similarity with CURS1 based on BLAST result. Because of the insufficient scaffolds we do not consider CURS1. So is the case of CURS1 and CURS2 in mapping first strategy. We also make pairwise t-test of the 2 strategies and calculate the *p*-value.)

We also validate the quality of expression analysis using the RNA-Seq method by evaluating expression levels of housekeeping genes. Four widely used housekeeping genes are selected as gene database as follows: glyceraldehyde 3-phosphate dehydrogenase, ubiquitin, tubulin, and elongation factor 1α (Guénin et al. 2008). We computed their expression levels to confirm that their expression levels do not vary among specimens. The CVs of these housekeeping genes among the four specimens ranged from 8.2% to 32.5%. It is worth noting that using the mapping-first analysis, the CVs of the same genes ranged from 12.8% to 46.2% and thus we could reduce the CVs by focusing on the specific pathways (select first strategy).

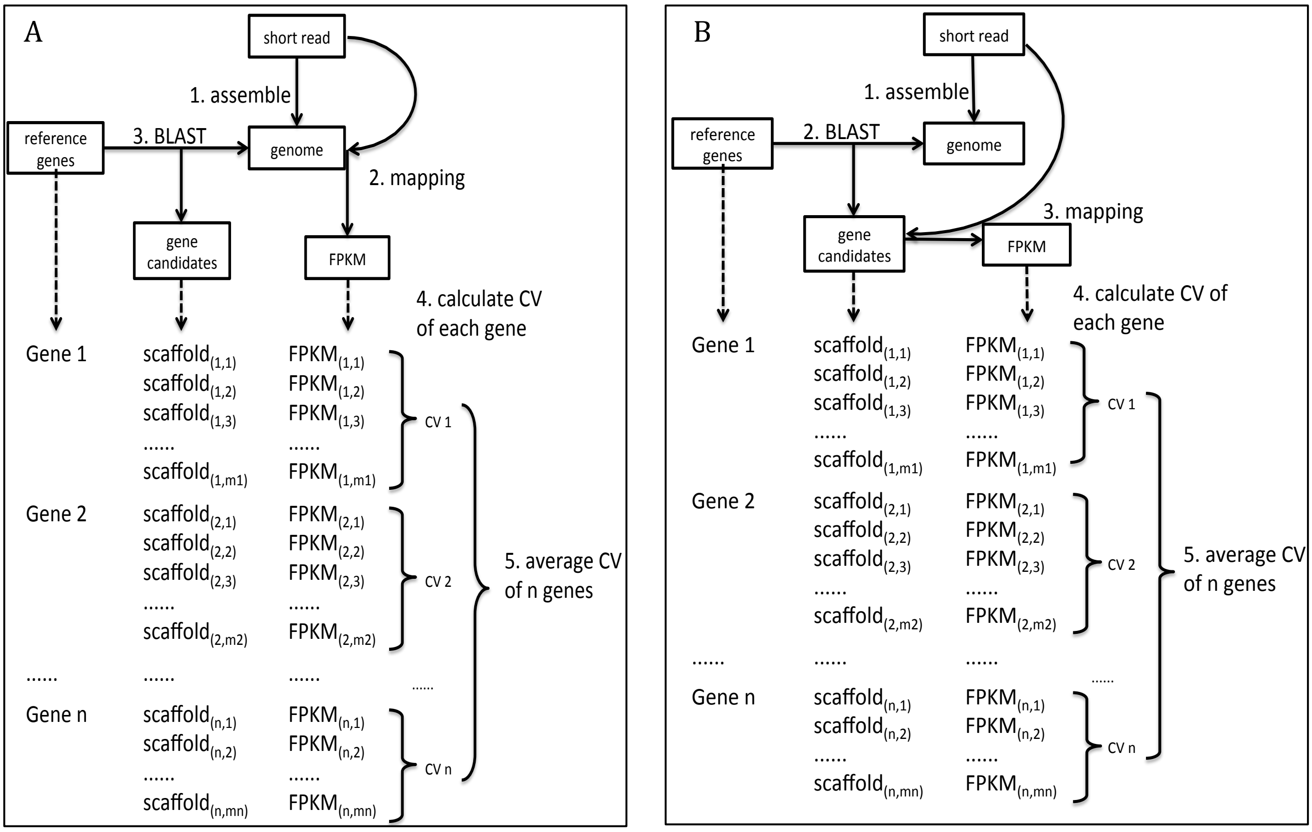


Figure 3.1 The 11 genes average CV calculation workflow. **A** is the general ‘mapping first strategy and the average CV is 2.12. **B** is the ‘select first strategy. The average CV is 1.47.

For *Curcuma* genera, only an assembled collection of expressed sequence tags is available (Koo et al. 2013), and not their genome sequences. We first assembled the raw reads sequenced by the NGS to obtain the putative gene sequences of *Curcuma* species **[Fig. 3.2**, Step 1**]**. Because the differences in gene sequences among these specimens are not large, we first assembled the total raw reads of the four specimens together to obtain a shared reference sequence set **[Fig. 3.2**, Step 2**]**. In total, we obtained 30,200 scaffolds with an average *N50* value of 1379 bp.

Next, we searched for the gene sequences involved in the curcuminoid biosynthesis, because the whole genome of *Curcuma* has not been determined. The transcription associated with curcuminoid biosynthesis is also unavailable. For this reason, we assembled short reads sequenced by RNA-Seq and compared them with enzymes of related plants to find homologous sequences. We first collected the corresponding sequences in model species (*Arabidopsis thaliana*, *Oryza sativa japonica*, *Glycine max*, and *Zea mays*) from the following databases: Kyoto Encyclopedia of Genes and Genomes (KEGG; http://www.genome.jp/kegg/; Kanehisa et al. 2000); the NCBI Reference Sequence Database (RefSeq; http://www.ncbi.nlm.nih.gov/refseq/; Pruitt et al. 2014); and KNApSAcK Motorcycle (http://kanaya.naist.jp/KNApSAcK_Family/; Afendi et al. 2012, Ikeda et al. 2013). We also obtained the gene sequences for DCS, CURS1, CURS2, and CURS3 of *C. longa* (Katsuyama et al. 2009b). By comparing these libraries with the assembled scaffolds, we selected 171 scaffolds that were highly homologous to these library genes **[Fig. 3.2**, Step 3**]**. Note that all the scaffolds that were similar to CUS library sequences were more similar to the reference sequences of CURS1–3 and DCS. We therefore did not include CUS (reported in *O. sativa* and others). as a reference. We referred to the selected scaffolds as the reference sequence set of the enzymes for curcuminoid biosynthesis. Each raw read of the four specimens was mapped to the references to count the number of assigned reads **[Fig. 3.2**, Step 4**]**. Finally, to estimate the expression level of each gene, the number of fragments per kilobase of exon per million fragments was computed using the method proposed by Li and Dewey (2011) **[Fig. 3.2**, Step 5**]**.

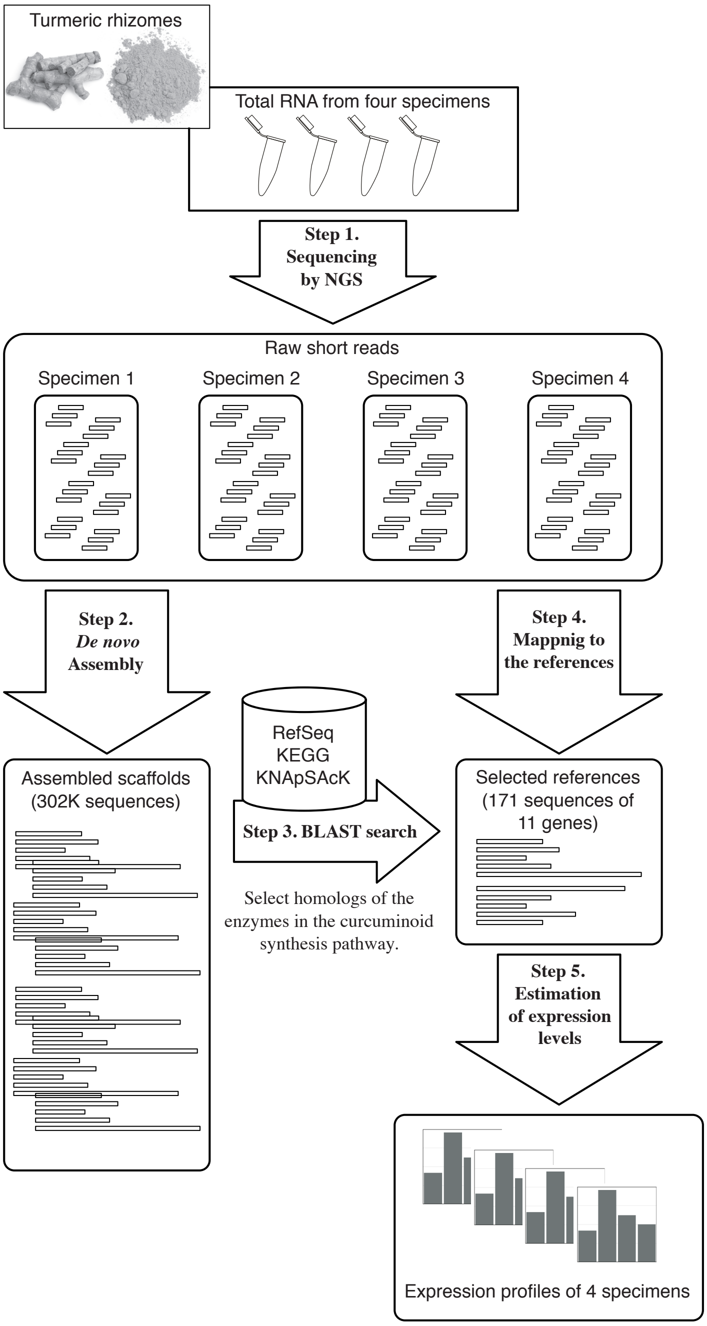


Figure 3.2 Schematic illustration of RNA-seq analysis of the curcuminoid biosynthesis pathway

3.3 Result and discussion

Gene expression level calculation

When multiple scaffolds were assigned to an identical enzyme gene, we simply added all expression values. The final genes' expression results are showed in **Fig. 3.3A**. In this method, the average CV of the estimated expression level for each enzyme was 1.47, whereas the average over the same enzyme was 2.12 when the raw reads were mapped to the whole scaffold first ($p = 0.019$ by the paired Mann–Whitney U test). We also validated the quality of expression analysis using the RNA-Seq method by evaluating expression levels of housekeeping genes. We chose scaffolds that are homologous to widely used housekeeping genes: glyceraldehyde 3-phosphate dehydrogenase, ubiquitin, tubulin, and elongation factor 1α (Guénin et al. 2008). We computed their expression levels to confirm that their expression levels do not vary among specimens. The CVs of these housekeeping genes among the four specimens ranged from 8.2% to 32.5%. It is worth noting that using the mapping-first analysis, the CVs of the same genes ranged from 12.8% to 46.2% and thus we could reduce the CVs by focusing on the specific pathways.

Species-specific biosynthesis of curcuminoids

Based on the expression analysis using the RNA-Seq method, we investigated reactions in the curcuminoid biosynthesis pathways. Curcuminoids are synthesized by Type III polyketide synthases (Katsuyama et al. 2009a) from carboxylic acid CoA ester **[Fig. 3.3A]**. First, feruloyl-CoA is produced from phenylalanine via *p*-coumaroyl-CoA. Next, DCS synthesizes feruloyldiketide-CoA from feruloyl-CoA and malonyl-CoA.

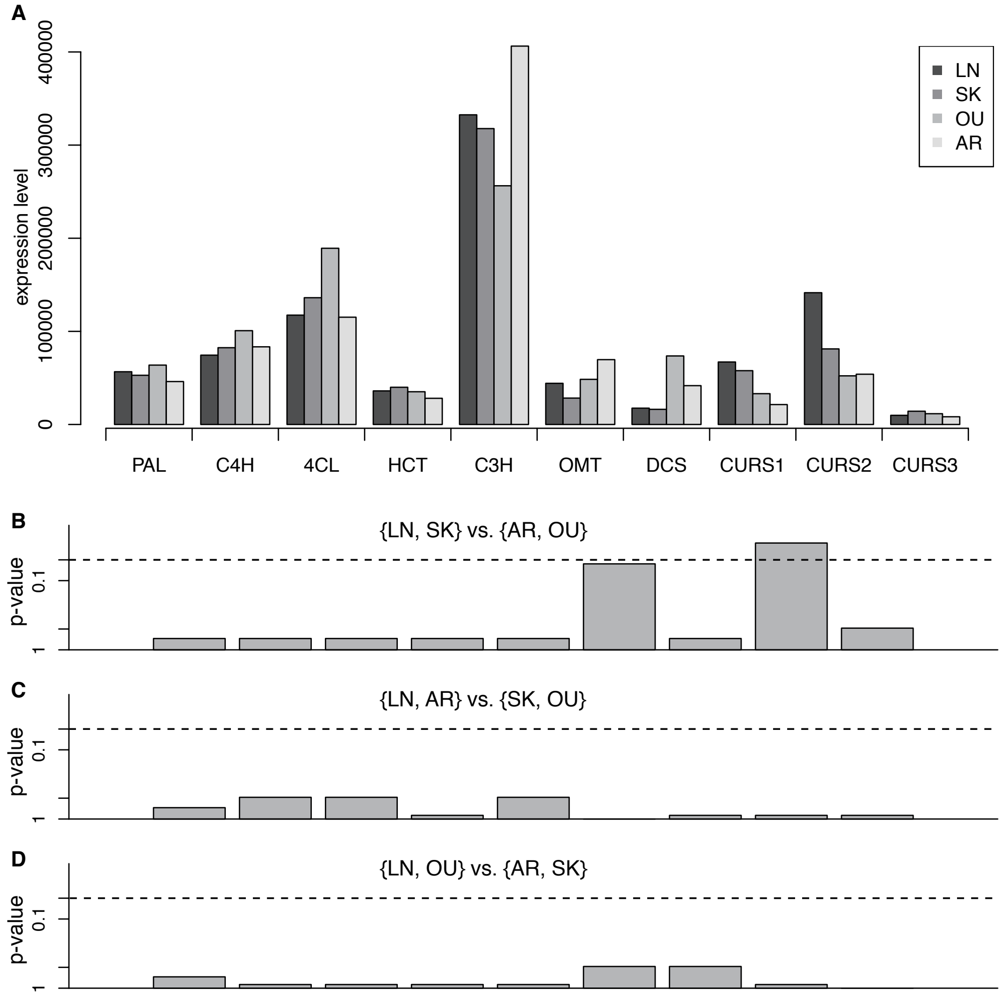


Figure 3.3 Gene expression profiles of the enzymes in the curcuminoid biosynthesis pathway. (**A**) Expression level of each enzyme. The order of enzymes corresponds to the position in the curcuminoids synthesis pathway. (**B-D**) Significance of expression difference. We compared expression levels of enzyme pairs along with the curcuminoids synthesis pathway. Each bar represents p-value of U test, weather the expression levels of adjacent enzyme pairs such as (PAL, C4H), (C4H, 4CL), ⋯ were equal between specimen sets such as {LN, SK} vs. {AR, OU}, etc. The dashed lines represent thresholds *p*=0.05.

Then, CURS1–3 catalyze the formation of curcumin from feruloyl-CoA and from the feruloyldiketide-CoA (Kita et al. 2008, Katsuyama et al. 2009a). Because these DCS and CURS1–3 can catalyze *p*-coumaroyl-CoA and *p*-coumaroyldiketide-CoA, other curcuminoids i.e. demethoxycurcumin and bisdemethoxycurcumin are also produced. On the other hand, another enzyme that synthesizes bisdemethoxycurcumin directly from two *p*-coumaroyl-CoAs and one malonyl-CoA has been reported (Morita et al. 2010). Thus, the content of curcuminoids in these cultivars depends on the balance of these enzymes and the amounts in individual substrate pools.

In the GC-MS analysis, LN and SK contained large amounts of curcumin, less demethoxy curcumin, and little bisdemethoxy curcumin, while OU and AR contained large amounts of demethoxy curcumin, less curcumin, and little bisdemethoxy curcumin **[Fig. 3.4A]**. It is interesting that demethoxy curcumin is the highest curcuminoid in OU. To compare their expression profiles, we divided the four specimens into two groups according to their curcumin content: Group I comprising {LN and SK} (in which curcumin was the largest component) and Group II comprising {OU and AR} (in which demethoxy curcumin was the largest component). We selected all adjacent enzyme sets along the metabolic pathways from 2-gene to 10-gene sets (described in Statistical tests in Materials and Methods), and compared their expression levels between Group I and II using the Mann–Whitney U test. We found that the expression levels of the enzyme pair (CURS1, CURS2) were significantly higher in Group I ($p = 0.03$), and those of the pair (OMT, DCS) were higher in Group II ($p = 0.06$) **[Fig. 3.3B]**. When we compared other specimen pairs—namely, {LN, AR} vs. {SK, OU} and {LN, OU} vs. {AR, SK}, differences in any enzyme sets between them were not statistically significant **[Fig. 3.3C, 3.3D**]. It is also worth noting that the expression levels of CURS3 were low in all four specimens.

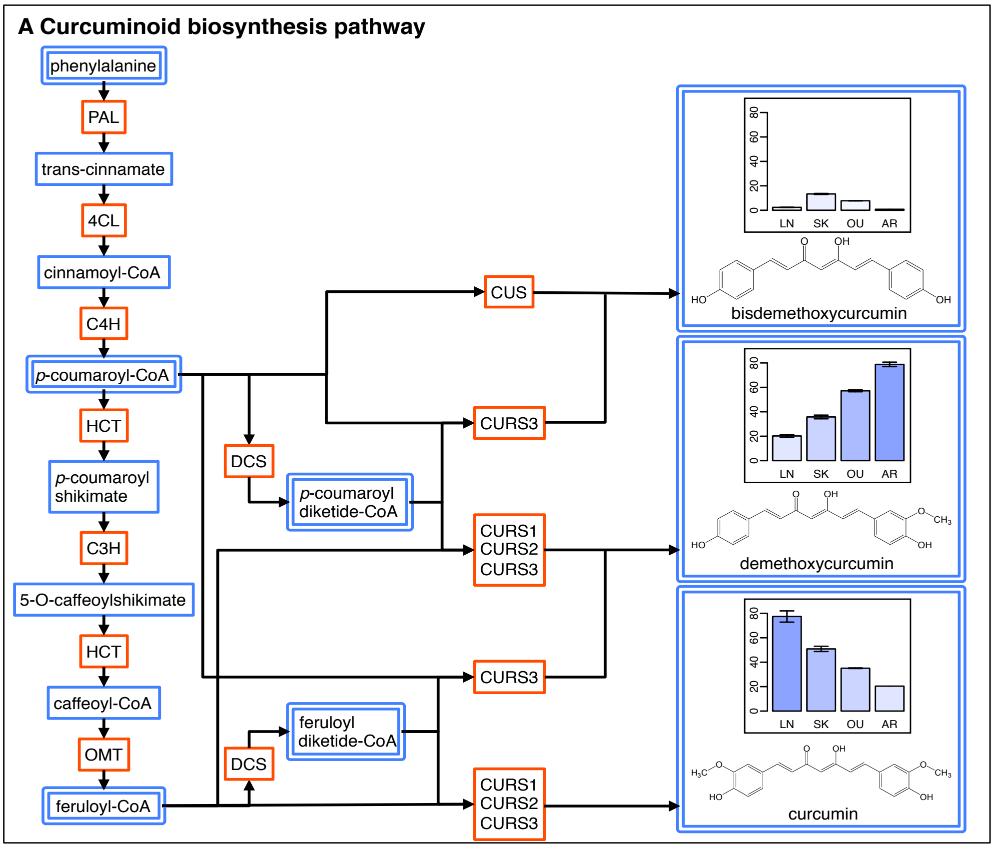


Figure 3.4 Curcuminoid biosynthesis pathway. (A) The pathway from phenylalanine to curcuminoids. The relative contents depicted in the bar graphs for curcumin, demethoxycurcumin, and bisdemethoxy curcumin, respectively, are 77.4%, 20.2%, and 2.4% for CL; 50.9%, 35.8%, and 13.3% for SK; 35.1%, 57.2%, and 7.7% for OU; and 20.4%, 78.8%, and 0.8% for AR. The error bars show the standard deviation of three.

The estimated behavior of curcuminoid pathways of Groups I and II can be summarized as shown in **Figs 3.5A** and **3.5B**, respectively. First, in Group I, the low expression of DCS results in the accumulation of feruloyl-CoA (shown in the large blue boxes in the left-side pathways), and then a high expression level of CURS1 and CURS2 (bright red arrows) explains the high curcumin content. Both CURS1 and CURS2 can also catalyze the synthesis of demethoxy curcumin, but they require the respective substrates feruloyl-CoA and *p*-coumaroyldiketide-CoA. When the concentration of *p*-coumaroyldiketide-CoA is low, the production of demethoxy curcumin is restricted. On the other hand, in Group II, because DCS is highly expressed, the flux from *p*-coumaroyl-CoA to *p*-coumaroyldiketide-CoA (thick red arrow) is larger than that to feruloyl-CoA, producing more demethoxy curcumin than curcumin. Because the syntheses of curcumin and demethoxy curcumin are catalyzed by the same enzymes, the difference in curcuminoid contents could be attributed to the difference in their substrate concentrations. To evaluate their concentrations, we measured the amount of coumarate and ferulate instead of *p*-coumaroyl-CoA and feruloyl-CoA and compared their concentration ratio between the four specimens. The ratios of the concentration of ferulate to that of coumarate in LN, SK, OU, and AR were 3.00, 2.80, 2.20, and 2.15, respectively. The differences in these ratios are consistent with the curcumin contents.

Based on the estimated expression levels, we compared the similarities of expression profiles by hierarchical clustering **[Fig. 3.6]**. The results showed that the expression patterns of LN and SK were the closest and OU was the most distant which is consistent with the fact that statistically significant expression difference was obtained in the pathways of pairs of genes OMT and DCS, and CURS1 and CURS2 in two groups {LN, SK} and {AR, OU}. Some *Curcuma* species such as *C. zedoaria* are known to produce high amounts of demethoxy curcumin (Lobo et al. 2009). Moreover, according to the phylogenetic tree estimated from the chloroplast DNA (Záveská et al. 2012), *C. zedoaria* is more distant from *C. longa* than *C. aromatica*. Our results suggest that the phenotype of OU is closer to *C. zedoaria* than to *C. longa*.

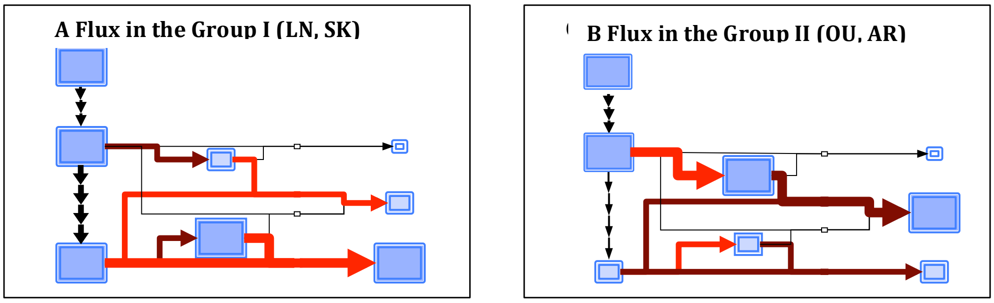


Figure 3.5 Curcuminoids synthesis flux in Group (LN, SK) and Group (OU, AR). Group I comprised LN and SK (in which curcumin was the largest component); Group II comprised OU and AR (in which demethoxycurcumin was the largest component). Bright and dark red arrows represent relative differences in expression levels for DCS vs CURS1 and CURS2. The sizes of the dashed and solid blue boxes show the expected and observed concentrations of the metabolites.

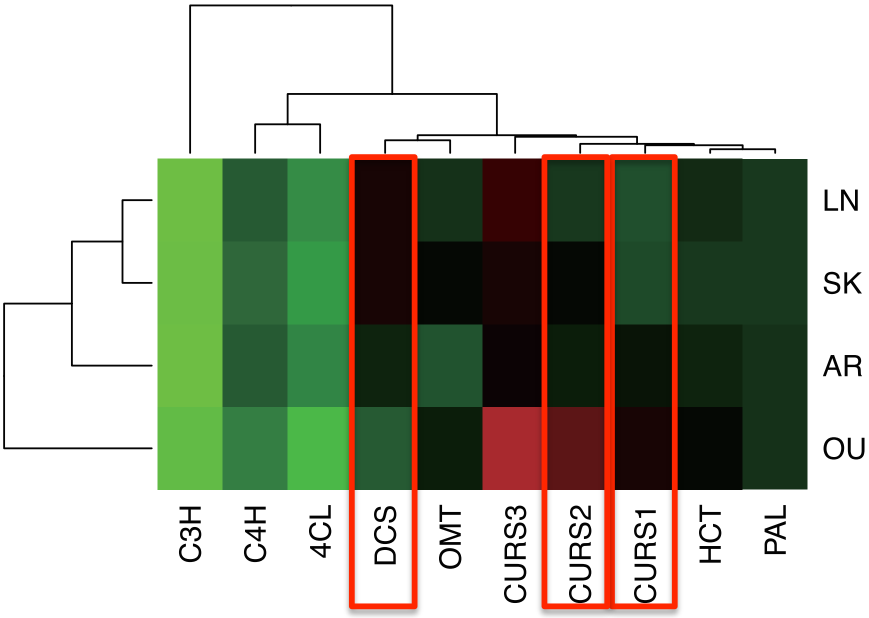


Figure 3.6 Two-dimensional hierarchical clustering of expression levels. Each row represents a specimen, and each column represents a gene. Depths of color in the green and red rectangles indicate higher and lower expression levels compared with the average, respectively. The dendrograms at the top and the left show the results of hierarchical clustering based on correlation of expression profiles among genes and specimens, respectively. The tree on the left shows that the expression of SK is close to LN while that of OU is far from them, consistently with their curcuminoid contents. The tree on the top shows that the expressions of enzymes adjacent in the pathway were correlated well. The red box of DCS gene shows low activity in LN and SK, but high in AR and OU. The CURS1 and CURS2 show high activity in AR and OU but low in LN and SK. The genes are clustered roughly according to their expression levels. It should be noted that the enzymes for synthesizing *p*-coumaroyl-CoA from cinnamate, C4H and 4CL, were also clustered into the same tree pattern, reflecting the sharing of the same regulatory.

Transcriptome analysis

For the selection-first analysis, we searched for scaffolds homologous to the amino acid sequences of the enzymes involved in curcuminoid biosynthesis pathways using the BLAST software (http://blast.ncbi.nlm.nih.gov/Blast.cgi; Altschul et al. 1990). The reference amino acid sequences for *A. thaliana*, *O. sativa japonica*, *G. max*, and *Z. mays* were obtained from RefSeq (Pruitt et al. 2014) and KNApSAcK (Afendi et al. 2012, Ikeda et al. 2013). Sequences of curcumin synthases of *C. longa* were obtained from the KEGG database (Kanehisa et al. 2000, Kanehisa et al. 2014). The scaffolds that matching the reference enzymes best were used as a library in Bowtie (Langmead et al. 2009) to map all transcriptome reads. Finally, RNA-Seq by expectation maximization value (Li and Dewey 2011) was computed using the Trinity utility (Grabherr et al. 2011) to obtain the expression level of each enzyme. For the mapping-first analysis, we mapped all transcriptome reads using Bowtie to all assembled scaffolds, and evaluated expression levels using the Trinity utility.

Statistical tests

To compare expression levels between specimens, we examined two sets of data pairs. First, we compared between the specimens that synthesize mostly curcumin or mostly demethoxy curcumin; namely, between {LN, SK} and {AR, OU}. Next, we also tested on the difference between other specimen groups, i.e. {LN, AR} vs. {SK, OU} and {LN, OU} vs. {AR, SK}. The expression levels of enzymes in a metabolic pathway are often regulated by the same transcriptional factors, so that their expression patterns are correlated (Wei et al. 2006). To detect differences in the expression levels of enzymes in curcuminoid biosynthesis pathways between specimens, we selected several series of genes along the pathway. Here, g denotes the index of 10 enzymes ordered along the curcuminoid synthesis pathway (PAL, C4H, 4CL, HCT, C3H, OMT, DCS, CURS1, CURS2, and CURS3), and X_{gh} denotes the expression level of the g-th gene of the h-th specimen. First, considering that the size of the enzyme set is two ($l = 2$), we chose pairs of adjacent enzymes (g_i, g_{i+1}), where g_i takes 1, …, 9; namely, the pairs (PAL, C4H), (C4H, 4CL), and so on. Then, we evaluated the difference in the expression levels $\{X_{gh}\}$ of genes in the i-th set ($g \in (g_i, g_{i+1})$) between specimen groups, (e.g., {LN, SK} vs. {AR, OU}), respectively, by the Mann–Whitney U test. Next, we let $l = 3$ and chose sets of three contiguous genes (g_i, g_{i+1}, g_{i+2}) from the beginning of the pathway; i.e., (PAL, C4H, 4CL), (C4H, 4CL, HCT), …, and so on. Using these enzyme sets, we compared their expressions between specimen groups. We repeated these processes, increasing the size of the gene sets, until all genes were chosen ($l = 10$).

Chapter IV. Conclusions

We utilized NGS technology on RNA-seq method to reveal the whole genome information of gene expression the cell. The whole genome contains all the secret of a cell. By comparing the gen expression in different samples, we focus on the genes for which the expression level highly changed and then on the whole metabolic pathway related to such genes.

The former part of this research mainly focuses on the major metabolic pathways, which correspond to daily bio-activities of plants. The latter part of the research focuses on the secondary metabolite pathways, which synthesize the metabolic production in a plant. As we discussed in the research, major metabolic pathways are in charge of plant daily activities and provide the substrate for the secondary metabolite synthesis pathway. We also get a preliminary research on how the secondary metabolite pathway genes express in the major pathway analysis. Two cultivars of the turmeric share many common pathways in which a number of the genes are up-regulated in the rhizome. This result indicates high possibility that cucuminoids synthesis mechanisms in these two cultivars are very similar.

We applied an error-resilient, selection-first method for a targeted expression analysis based on RNA-Seq in *Curcuma* species. Although raw reads sequenced by NGS systems may contain substantial noise, we could exploit them through focusing on specific genes and pathways by selecting target sequences of interest *a priori*. In our trial, the method reduced the variance of the estimated expression levels, and reliable differences were detected for gene expressions in curcuminoid biosynthetic pathways in the four turmeric specimens.

This method is especially effective for plant secondary metabolite biosynthesis where gene expression and metabolite concentration correlate well, and end products accumulate (or, sometimes exuded) without active catabolism. Such secondary metabolite genes are often expressed coordinately, implying that their concerted up-regulation increases end-product fluxes (Yonekura-Sakakibara et al. 2013). Yonekura-Sakakibara (2009) also showed that expression of enzymes in flavonoid biosynthesis pathways in *A. thaliana* correlated well with that of key transcription factors. We detected correlation between expression levels of transcription factors and some enzymes in the flavonoid pathways. Our results further corroborate the findings of Mewis et al. (2006) who studied defense-signaling pathways of *A. thaliana* in response to insect feeding. That work showed that contents of glucosinolate often correlate with changes in transcript levels of their biosynthetic genes. Our co-expression analyses indicate that there are similarities in co-regulation mechanisms of enzymes in the same pathway.

Genes associated with identical metabolic pathways are often coexpressed (Ihmels et al. 2003) so that they can catalyze linear chain of reactions. This is evident especially for those involved in secondary metabolic pathways (Gachon et al. 2005, Yonekura-Sakakibara et al. 2007, Aoki et al. 2007; Wada et al., 2011). In this study, we focused on the expression changes of genes adjacent along with the curcuminoid biosynthesis pathway and could classify two groups {LN, SK} and {AR, OU} based on metabolic switch obtained by co-expression of genes and accumulation of metabolites. Secondary metabolic pathways are the most diverse and rapidly evolving features of plant genomes (Wei et al. 2006; Albinsky et al. 2010; Schaefer et al. 2014). Identification of such co-expressed gene sets will provide useful information to regulate production of interesting metabolites. We could estimate changes in their metabolic flux by comparing information on enzymes and their substrates. The results consistently explained differences in curcuminoid contents between the four specimens. Whether genome sequence is available or not, our strategy, namely a targeted integration of RNA-Seq and metabolite data can be applicable to estimate expression regulation of genes in the secondary metabolic pathways of various species.

In this study we combined RNA-seq method based on NGS technique and metabolite analysis together to investigate the plant metabolic pathway. The complexity of metabolite analysis requires not only the assist from genetic aspect but also from metabolite. Though there are plenty of the factors that affect the metabolite pathways, not only the genes and the metabolite, this study is a trial to consider these basic elements.

When we analysze the curcuminoid synthesis pathway, the current whole transciptome analysis method is not very suitable for a small scale synthesis network of genes. All the raw reads map to the whole genome, the information that related the specific genes would be swallowed by the major metabolite pathway information. We developed a “selection-first” method for RNA-Seq analysis in which short reads are mapped to selected enzymes in the target biosynthetic pathways in order to reduce the effect of mapping errors. This strategy lowers coefficients of variation of gene expression level.

We made NGS data mining of both primary metabolic pathways and a specific secondary metabolic pathway. There is relationship between these two kinds of pathway. Primary metabolic pathways have significant meaning to the living creatures for which understanding the mechanism is important. The secondary metabolic pathway also have deep influence on the living creatures.

Acknowledgement:

It is a very important stage of my life to study in Nara Institute of Science as a doctoral student. In the period of my Ph.D study here, I met a lot of people, and learned a lot of knowledge.

I want to thank to my supervisor, Professor Shigehiko Kanaya, who gave me the opportunity to study in the Computational Systems Biology Laboratory. I greatly appreciate his assistance and support during my studies. Professor Kanaya not only gives me the knowledge and the skill during the doctoral study, but also shows me the example of a real scientist.

I would like to thank my thesis committee, Prefessor Shigehiko Kanaya, Professor Keichi Yasumoto, Associate Professor MD.Altafl-Ul-Amin, Associate Professor Tadao Sugihura, Assistant Professor Naoaki Ono. Thank for their valuable comments and supervision of my work. Their valuable advice, guidance, critics mde the dissertation improved greatly.

I would like to thank all the other staff in Computational Systems Biology Laboratory: Assistant Professor Tetsuo Sato, Assistant Professor Alex Ming Huang, Aki H. Morita, Minako Ohashi. It is very kind of you to assist me in daily work and life.

I want to give my great appreciation to Assistance Professor Naoaki Ono, who kindly guided my Ph.D research carefully during all my Ph.D course. And also I want give my appreciation to my friend Katsuragi Tetsuo, and Nelson Nelson Kipchirchir Kibinge, who helped me a lot in my daily research and live.

Last but not least, my appreciation also extends to my parents and my friends. With your understanding, patience, critics, companion, support, and encouragement, I could achieve my doctoral degree.

Thank you very much.

List of Publications

Journal:

1. Donghan Li, Naoki Ono, Tetsuo Sato,Tadao Sugiura, Md Altaf-Ul-Amin, Daisaku Ohta, Hideyuki Suzuki, Masanori Arita, Ken Tanaka, Zhiqiang Ma, Shigehiko Kanaya; Targeted Integration of RNA-Seq and Metabolite Data to Elucidate Curcuminoid Biosynthesis in Four *Curcuma* Species; Plant and Cell Physiology; 56(5):843-51; 2015 May

International Conference:

1. Donghan Li, Naoaki Ono, Shigehiko Kanaya, Ken Tanaka; RNA-Seq Analysis of Curcuminoid Synthesis Pathway of *Curcuma Longa*; AMT-BHI 2013 (Active Media Technology, Brain and Health Informatics); 2013 October
2. Donghan Li, Naoaki Ono, Tetsuo Sato, Tadao Sugiura, Md. Ataf-UI-Amin, Masanori Arita, Ken Tanaka, Zhiqiang Ma, Shigehiko Kanaya; Targeted Integration Between RNA-Seq and Metabolomics Data to Elucidate Curcuminoid Biosynthesis Flux in Four *Curcuma* Species; GIW ISCB-ASIA 2014 (The 25th International Conference on Genome Informatic; 2014 December
3. Donghan Li, Naoaki Ono, Tetsuo Sato, Tadao Sugiura, Md. Altaf-Ul-Amin, Masanori Arita, Ken Tanaka, Zhiqiang Ma, Shigehiko Kanaya; Comparison of curcuminoid biosynthesis of Curcuma longa and its cultivars using a pathway based RNA-Seq analysis method; Metabolomics 2015; 2015 July

Domestic Conference:

1. Donghan Li, Naoaki Ono, Ken Tanaka, Md Altaf-UI-Amin, Tadao Sugiura, Tetsuo Sato, Shigehiko Kanaya; “RNA-Seq Applied on Curcuminoid Synthesis Enzyme Finding”; NGS 現場の会第三回研究会, Kobe, Hyogo, Japan

REFERENCE:

Afendi, F. M., Okada, T., Yamazaki, M., Hirai-Morita, A., Nakamura, Y., Nakamura, K. et al. (2012). KNApSAcK family databases: integrated metabolite-plant species databases for multifaceted plant research. *Plant Cell Physiol*. 53: e1.

Aggarwal, B. B., Kumar, A. and Bharti, A.C. (2003). Anticancer potential of curcumin: preclinical and clinical studies. *Anticancer Res*. 23: 363–398.

Altaf-Ul-Amin, Md, et al. (2014). Systems biology in the context of big data and networks. *BioMed research international* 2014.

Albinsky, D., Sawada, Y., Kuwahara, A., Nagano, M., Hirai, A., Saito, K. et al. (2010). Widely targeted metabolomics and coexpression analysis as tools to identify genes involved in the side-chain elongation steps of aliphatic glucosinolate biosynthesis. *Amino Acids*, 39: 1067–1075.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215:403–410.

Aoki, K., Ogata, Y. and Shibata, D. (2007). Approaches for Extracting Practical Information from Gene Co-expression Networks in Plant Biology. *Plant Cell Physiol*. 48: 381–390.

Azuine, M. A. and Bhide, S. V., (1994). Adjuvant chemoprevention of experimental cancer: catechin and dietary turmeric in forestomach and oral cancer models. *J. Ethnopharmacol.*, 44, 211–217.

Borchers, Andrea T., et al., (2000). Shosaiko-to and other Kampo (Japanese herbal) medicines: a review of their immunomodulatory activities. *Journal of Ethnopharmacology* 73.1: 1-13.

Boguski MS. (1995). The turning point in genome research. Trends Biochem. Sci. 20:295–96

Bras, Jose, Rita Guerreiro, and John Hardy. (2012). Use of next-generation sequencing and other whole-genome strategies to dissect neurological disease. *Nature Reviews Neuroscience* 13.7: 453-464.

Burset, M., Seledtsov, I. A. & Solovyev, V. V. (2000). Analysis of canonical and non-canonical splice sites in mammalian genomes. *Nucleic Acids Res.* 28, 4364–4375.

Caspi, R., Foerster, H., Fulcher, C. A., Kaipa, P., Krummenacker, M., Latendresse, M., ... & Karp, P. D. (2008). The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. Nucleic acids research, 36(suppl 1), D623-D631.

Chattopadhyay, I., Biswas, K., Bandyopadhyay, U., & Banerjee, R. K. (2004). Turmeric and curcumin: Biological actions and medicinal applications. *Current science*, 87(1), 44-53.

Chattopadhyay, I., Bandyopadhyay, U., Biswas, K., Maity, P., & Banerjee, R. K. (2006). Indomethacin inactivates gastric peroxidase to induce reactive-oxygen-mediated gastric mucosal injury and curcumin protects it by preventing peroxidase inactivation and scavenging reactive oxygen. *Free radical biology and medicine*, *40*(8), 1397-1408.

Christopher P. A. and Van V. B. (2006). Estimating The Cost Of New Drug Development: Is It Really $802 Million? *Health Affairs* 25, 2: 420-428

DiMasi J.A. (2001). Risks in new drug development: Approval success rates for investigational drugs *Clinical Pharmacology & Therapeutics* 69, 5: 297-307

DiMasi, J. A., Hansen, R. W. & Grabowski, H. G. (2003). The price of innovation: new estimates of drug development costs. J. Health Econ. 22, 151–185.

Edgar, Robert C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26.19: 2460-2461.

E. R. Kandel, J. H. Schwartz, T. M. Jessell et al., (2000). *Principles of Neural Science*, vol. 4, McGraw-Hill, NewYork, NY, USA.

Fabricant, D. S., & Farnsworth, N. R. (2001). The value of plants used in traditional medicine for drug discovery. *Environmental health perspectives,* 109 (Suppl 1): 69.

Gachon C. M. M., Langlois-Meurinne M., Henry Y., Saindrenan P. (2005). Transcriptional co-regulation of secondary metabolism enzymes in *Arabidopsis*: functional and evolutionary implications. *Plant Mol. Biol*. 58: 229–245.

Gonzaga-Jauregui, Claudia, James R. Lupski, and Richard A. G. (2012). Human genome sequencing in health and disease. *Annual review of medicine* 63: 35.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I. et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29: 644–652.

Guénin, S., Mauriat, M., Pelloux, J., Wuytswinkel, O.V., Bellini, C. and Gutierrez, L. (2008). Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. *J. Exp. Bot.* 60: 487–493.

Gujral, M. L., Chowdhury, N. K. and Saxena, P. N., (1953).The effect of certain indigenous remedies on the healing of wounds and ulcers. *J. Indian State Med. Assoc.*, 22, 273–276.

Ihmels J., Levy, R. and Barkai N. (2003). Principles of transcriptional control in the metabolic network of *Saccharomyces cerevisiae*. *Nat. Biotechnol*. 22: 86–92.

Ikeda, S., Abe, T., Nakamura, Y., Kibinge, N., Morita, A.H., Nakatani, A. et al. (2013). Systematization of the protein sequence diversity in enzymes related to secondary metabolic pathways in plants, in the context of big data biology inspired by the KNApSAcK Motorcycle database. *Plant and Cell Physiology*, 54: 711–727.

Shita C

Jayaprakasha, G. K., Jagan, L., Rao, M. and Sakariah, K.K. (2005). Chemistry and biological activities of *C. longa*. *Trends Food Sci. Tech.* 16: 533–548.

Jeffrey A. M., Zhong W., (2011). Next-generation transcriptome assembly *Nature Reviews Genetics* 12, 671-682.

Kanehisa, M. and Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28: 27–30.

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M. and Tanabe, M. (2014). Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res.* 42: D199–D205.

Katsuyama, Y., Kita, T., Funa, N. and Horinouchi, S. (2009a). Curcuminoid biosynthesis by two type III polyketide synthases in the herb *Curcuma longa*. *J. Biol. Chem.* 284: 11160–11170.

Katsuyama, Y., Kita, T. and Horinouchi, S. (2009b). Identification and characterization of multiple curcumin synthases from the herb *Curcuma longa*. *FEBS Lett.* 583: 2799–2803.

Kita, T., Imai, S., Sawada, H., Kumagai, H. and Seto, H. (2008). The biosynthetic pathway of curcuminoid in turmeric (*Curcuma longa*) as revealed by ^{13}C-labeled precursors. *Biosci. Biotechnol. Biochem.* 72: 1789–1798.

Kinghorn, A. D., Pan, L., Fletcher, J. N., & Chai, H. (2011). The relevance of higher plants in lead compound discovery programs. Journal of natural products, 74(6), 1539-1555.

Koo, H.J., McDowell, E.T., Ma, X., Greer, K.A., Kapteyn, J., Xie, Z. et al. (2013). Ginger and turmeric expressed sequence tags identify signature genes for rhizome identity and development and the biosynthesis of curcuminoids, gingerols and terpenoids. *BMC Plant Biol.* 13: 27.

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

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

Lin Liu, Yinhu Li, Siliang Li, et al. (2012). Comparison of Next-Generation Sequencing Systems, *Journal of Biomedicine and Biotechnology*, vol. 2012, Article ID 251364, 11 pages.

Lobo, R., Prabhu, K.S., Shirwaikar, A. and Shirwaikar, A. (2009). *Curcuma zedoaria* Rosc. (white turmeric): a review of its chemical, pharmacological and ethnomedicinal properties *J. Pharm. Pharmacol*. 61: 13–21.

Maheshwari, R.K., Singh, A.K., Gaddipati, J. and Srimal, R.C. (2006). Multiple biological activities of curcumin: a short review. *Life Sci.* 78: 2081–2087.

Mardis, Elaine R. (2008). The impact of next-generation sequencing technology on genetics. *Trends in genetics* 24.3: 133-141.

Mardis, Elaine R. (2013). Next-generation sequencing platforms. *Annual review of analytical chemistry* 6: 287-303.

Metzker, M. L. (2010). Sequencing technologies—the next generation. *Nature reviews genetics*, 11(1), 31-46.

Mewis, I., Tokuhisa, J.G., Schultz, J.C., Appel, H.M., Ulrichs, C. and Gershenzon, J. (2006). Gene expression and glucosinolate accumulation in *Arabidopsis thaliana* in response to generalist and specialist herbivores of different feeding guilds and the role of defense signaling pathways. *Phytochemistry*, 67: 2450–2462.

Michael D. and Jean P. G. (2004). Key factors in the rising cost of new drug discovery and development. *Nature Reviews Drug Discovery* 3: 417-429

Morita, H., Wanibuchi, K., Nii, H., Kato, R., Sugio, S. and Abe, I. (2010). Structural basis for the one-pot formation of the diarylheptanoid scaffold by curcuminoid synthase from *Oryza sativa*. *Proc. Natl. Acad. Sci. USA*. 107: 19778–19783.

Mockler TC, Chan S, Sundaresan A, Chen H, Jacobsen SE, Ecker JR. (2005). Applications of DNA tiling arrays for whole-genome analysis. *Genomics* 85:1–15

Nguyen, Q. T., Merlo, M. E., Medema, M. H., Jankevics, A., Breitling, R., & Takano, E. (2012). Metabolomics methods for the synthetic biology of secondary metabolism. *FEBS letters*, 586(15), 2177-2183.

Patel, R.K. and Mukesh, J. (2012). NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One* 7: e30619.

Pozhitkov AE, Tautz D, Noble PA. 2007. Oligonucleotide microarrays: widely applied–poorly understood. Brief Funct. Genomic Proteomic 6:141–48

Pruitt, K.D., Brown, G.R., Hiatt, S.M., Thibaud-Nissen, F., Astashyn, A., Ermolaeva, O. et al. (2014). RefSeq: an update on mammalian reference sequences. *Nucleic Acids Res*. 42: D756–D763.

Ravindran, P.N., Nirmal Babu, K. and Sivaraman, K. (eds) (2007). Turmeric: The Genus *Curcuma*. p. 145. *CRC Press*, Boca Raton, FL, USA.

Rismani-Yazdi, H., Haznedaroglu, B.Z., Bibby, K. and Peccia, J. (2011). Transcriptome sequencing and annotation of the microalgae *Dunaliella tertiolecta*: pathway description and gene discovery for production of next-generation biofuels. *BMC Genomics* 12: 148.

Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* USA 74:5463–67

Sasikumar, B. (2005). Genetic resources of *Curcuma*: diversity, characterization and utilization. *Plant Genetic Resources* 3: 230–251.

Schaefer, R.J., Briskine, R., Springer, N.M. and Myers, C.L. (2014). Discovering Functional Modules across Diverse Maize Transcriptomes Using COB, the Co-Expression Browser. *PLoS ONE*. 9: e99193.

Shendure J., Ji H., 2008, Next-generation DNA sequencing, *Nature Biotechnology* **26**, 1135 - 1145

Suzuki, Y., Kawazu, T., Koyama, H. (2004). RNA isolation from siliques, dry seeds, and other tissues of *Arabidopsis thaliana*, *BioTechniques* 37:542–544.

Tan, Benny KH, and J. Vanitha. (2004). Immunomodulatory and antimicrobial effects of some traditional Chinese medicinal herbs: a review. *Current Medicinal Chemistry* 11.11: 1423-1430.

Toshikazu M., Hiroyasu Y., Konosuke D. (2006). Varietal difference and year-to-year variation of the contents of curcuminoids in turmeric. *The 222nd Meeting of the Crop Science Society of Japan.*P130.

Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. 1995. Serial analysis of gene expression. *Science* 270:484–87

Wada, M., Takahashi, H., Altaf-Ul-Amin, M., Nakamura, K., Hirai, M. Y., Ohta, D. Kanaya, S. (2012). Prediction of operon-like gene clusters in the *Arabidopsis thaliana* genome based on co-expression analysis of neighboring genes. *Gene*. 503: 56–64.

Wang, Z., Gerstein, M. and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10: 57–63.

Wang Y, Yan HW and Zou WM (1999). HPLC determination of three curcuminoid constituents in rhizoma Curcumae. Acta Pharmaceutica Sinica 34: 467 – 470.

Wei, H., Persson, S., Mehta, T., Srinivasasainagendra, V., Chen, L., Page, G.P., et al. (2006). Transcriptional coordination of the metabolic network in Arabidopsis. *Plant Physiol*. 142: 762–774.

Wold, B. and Myers, R. M. (2008). Sequence census methods for functional genomics. *Nature Methods* 5, 19–21.

Yonekura-Sakakibara, K., Tohge, T., Niida, R. and Saito, K. (2007). Identification of a Flavonol 7-O-Rhamnosyltransferase Gene Determining Flavonoid Pattern in Arabidopsis by Transcriptome Coexpression Analysis and Reverse Genetics. *J. Biol. Chem.* 282: 14932–14941.

Yonekura-Sakakibara, K. (2009). Functional genomics of family 1 glycosyltransferases in *Arabidopsis*. *Plant Biotechnol*. 26: 267–274.

Yonekura-Sakakibara, K., Fukushima, A., Saito, K. (2013). Transcriptome data modeling for targeted plant metabolic engineering. *Curr. Opin. Biotechnol.* 24: 285–290.

Záveská, E., Fer, T., Sida, O., Krak, K., Marhold, K. and Leong-Skornickova, J. (2012). Phylogeny of *Curcuma* (Zingiberaceae) based on plastid and nuclear sequences: proposal of the new subgenus *Ecomata*. *Taxon* 61: 747–763

Appendix

Table 1. Up-regulated genes in leaf in cultivar AR

	Scaffold	Annotation	log2(leaf/rhizome)	Pathway
1	comp107831_c0	protein MIDASIN1	15.96	Ribosome biogenesis in eukaryotes
2	comp110162_c0	Preprotein translocase Sec, Sec61-beta subunit protein	15.84	Protein export; Protein processing in endoplasmic reticulum; Phagosome
3	comp111085_c0	RPL27; 50S ribosomal protein L27	15.37	Ribosome
4	comp114089_c0	E3 ubiquitin-protein ligase SINAT3	15.28	Ubiquitin mediated proteolysis
5	comp113126_c0	APC10; anaphase-promoting complex subunit 10	14.71	Ubiquitin mediated proteolysis
6	comp107761_c1	hypothetical protein	14.13	Taurine and hypotaurine metabolism; Metabolic pathways
7	comp102333_c0	RD19; cysteine proteinase RD19a	14.12	Plant-pathogen interaction
8	comp110408_c0	VPS60.2; vacuolar protein sorting protein 60.2	13.95	Endocytosis
9	comp106512_c0	40S ribosomal protein S16-1	13.91	Ribosome
10	comp111967_c0	ACA7; alpha carbonic anhydrase 7	13.77	Nitrogen metabolism
11	comp114024_c0	transport protein SEC13A	13.52	RNA transport; Protein processing in endoplasmic reticulum

12	comp99769_c0	U6 snRNA-associated Sm-like protein LSm1	13.45	RNA degradation
13	comp111496_c0	CRY2; cryptochrome 2	13.44	Circadian rhythm - plant
14	comp112877_c0	LHCA5; photosystem I light harvesting complex protein 5	13.31	Photosynthesis - antenna proteins
15	comp102896_c0	PUR5; phosphoribosylformylglycinamidine cyclo-ligase	13.15	Purine metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics
16	comp106357_c0	FED_A; ferredoxin-2	13.06	Photosynthesis
17	comp123248_c0	RNA recognition motif-containing protein	12.99	Spliceosome
18	comp108615_c0	2-hydroxyacyl-CoA lyase	12.87	Peroxisome
19	comp113557_c0	FD3; ferredoxin 3	12.83	Photosynthesis
20	comp118901_c0	3-oxoacyl-[acyl-carrier-protein] synthase	12.68	Fatty acid biosynthesis; Biotin metabolism; Metabolic pathways; Fatty acid metabolism
21	comp127116_c0	URE; urease	12.59	Arginine biosynthesis; Purine metabolism; Metabolic pathways
22	comp103200_c0	NDPK3; nucleoside diphosphate kinase 3	12.57	Purine metabolism; Pyrimidine metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics
23	comp119946_c0	thioredoxin family protein	12.56	Ubiquitin mediated proteolysis
24	comp121776_c0	ZTL; adagio protein 1	12.56	Circadian rhythm - plant
25	comp115622_c0	RNA recognition motif-containing protein	12.51	Spliceosome
26	comp49043_c0	NADP-ME2; NADP-dependent malic enzyme 2	12.49	Pyruvate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Carbon metabolism

27	comp115141_c0	PAC1; 20S proteasome alpha-3 subunit	12.37	Proteasome
28	comp103986_c0	EIF2_GAMMA; translation initiation factor eIF-2 gamma subunit	12.33	RNA transport
29	comp122860_c0	palmitoyl protein thioesterase family protein	12.32	Fatty acid elongation; Metabolic pathways; Fatty acid metabolism
30	comp116615_c0	SAM1; S-adenosylmethionine synthetase 1	12.31	Cysteine and methionine metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of amino acids
31	comp131243_c0	CYP98A3; cytochrome P450 98A3	12.24	Phenylpropanoid biosynthesis; Flavonoid biosynthesis; Stilbenoid, diarylheptanoid and gingerol biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
32	comp114709_c0	NRPE5; DNA-directed RNA polymerase V subunit 5A	12.24	Purine metabolism; Pyrimidine metabolism; Metabolic pathways; RNA polymerase
33	comp118847_c0	UBC9; ubiquitin conjugating enzyme 9	12.17	Ubiquitin mediated proteolysis; Protein processing in endoplasmic reticulum
34	comp117203_c0	transcription initiation factor IIB-2	12.16	Basal transcription factors
35	comp137144_c0	Beta-glucosidase, GBA2 type family protein	12.15	Other glycan degradation; Sphingolipid metabolism; Metabolic pathways
36	comp119538_c0	GAPCP-2; glyceraldehyde 3-phosphate dehydrogenase	12.14	Glycolysis / Gluconeogenesis; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
37	comp120503_c0	EMB2769; protein embryo	12.14	Spliceosome

		defective 2769		
38	comp120129_c0	GAPC2; glyceraldehyde 3-phosphate dehydrogenase GAPC2	12.10	Glycolysis / Gluconeogenesis; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
39	comp123588_c0	transcription factor IIA, alpha/beta subunit	12.10	Basal transcription factors
40	comp134068_c0	dehydratase-enolase-phosphatase complex 1	12.03	Cysteine and methionine metabolism; Metabolic pathways
41	comp125115_c0	TAG1; diacylglycerol O-acyltransferase 1	12.03	Glycerolipid metabolism; Metabolic pathways
42	comp131064_c0	phosphoribosylglycinamide formyltransferase	12.00	Purine metabolism; One carbon pool by folate; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics
43	comp124809_c1	BCAT3; branched-chain-amino-acid aminotransferase 3	11.95	Valine, leucine and isoleucine degradation; Valine, leucine and isoleucine biosynthesis; Pantothenate and CoA biosynthesis; Glucosinolate biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; 2-Oxocarboxylic acid metabolism; Biosynthesis of amino acids
44	comp111990_c0	RPT4A; 26S proteasome AAA-ATPase subunit RPT4a	11.91	Proteasome
45	comp103293_c0	FAB1B; phosphatidylinositol-3P 5-kinase-like	11.89	Inositol phosphate metabolism; Phosphatidylinositol signaling system; Phagosome

46	comp120004_c0	CI51; NADH dehydrogenase [ubiquinone] flavoprotein 1	11.86	Oxidative phosphorylation; Metabolic pathways
47	comp116938_c0	AHP1; histidine-containing phosphotransmitter 1	11.86	Plant hormone signal transduction
48	comp114748_c0	HAI2; protein phosphatase 2C 3	11.78	Plant hormone signal transduction
49	comp114727_c0	FAB1; 3-oxoacyl-[acyl-carrier-protein] synthase II	11.75	Fatty acid biosynthesis; Biotin metabolism; Metabolic pathways; Fatty acid metabolism
50	comp119518_c0	NRPC2; DNA-directed RNA polymerase RPC2	11.72	Purine metabolism; Pyrimidine metabolism; Metabolic pathways; RNA polymerase
51	comp111645_c0	protein SM-LIKE 3B	11.68	RNA degradation; Spliceosome
52	comp121549_c0	RNA helicase family protein	11.68	Spliceosome
53	comp109061_c0	SDH2-1; succinate dehydrogenase [ubiquinone] iron-sulfur subunit 1	11.64	Citrate cycle (TCA cycle); Oxidative phosphorylation; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
54	comp121527_c0	ribosomal L18p/L5e family protein	11.58	Ribosome
55	comp127845_c0	PS2; inorganic pyrophosphatase 1	11.56	Vitamin B6 metabolism; Metabolic pathways
56	comp131917_c0	NCED4; nine-cis-epoxycarotenoid dioxygenase 4	11.53	Carotenoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
57	comp103152_c0	Oxoglutarate/iron-dependent oxygenase	11.47	Arginine and proline metabolism; Metabolic pathways
58	comp100053_c0	EP3; chitinase class IV	11.47	Amino sugar and nucleotide sugar metabolism
59	comp114805_c0	40S ribosomal protein S26-1	11.46	Ribosome
60	comp118545_c0	BAS1; cytochrome P450 734A1	11.45	Brassinosteroid biosynthesis

61	comp108887_c0	CNX7; molybdenum cofactor synthesis family protein	11.43	Folate biosynthesis; Metabolic pathways; Sulfur relay system
62	comp137165_c0	Beta-glucosidase, GBA2 type family protein	11.42	Other glycan degradation; Sphingolipid metabolism; Metabolic pathways
63	comp86890_c0	RR2; transcription factor response regulator 2	11.37	Plant hormone signal transduction
64	comp118369_c0	ATP synthase subunit G protein	11.34	Oxidative phosphorylation; Metabolic pathways
65	comp110570_c0	ATP-dependent caseinolytic (Clp) protease/crotonase family protein	11.32	Valine, leucine and isoleucine degradation; beta-Alanine metabolism; Propanoate metabolism; Metabolic pathways; Carbon metabolism
66	comp99470_c0	DNA glycosylase superfamily protein	11.31	Base excision repair
67	comp115968_c0	PIP5K9; phosphatidyl inositol monophosphate 5 kinase	11.28	Inositol phosphate metabolism; Metabolic pathways; Phosphatidylinositol signaling system; Endocytosis
68	comp118078_c0	ATP3; ATP synthase subunit gamma	11.24	Oxidative phosphorylation; Metabolic pathways
69	comp125735_c0	GLCAK; glucuronokinase G	11.19	Pentose and glucuronate interconversions; Ascorbate and aldarate metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways
70	comp127499_c0	OVA4; protein ovule abortion 4	11.12	Aminoacyl-tRNA biosynthesis
71	comp112385_c0	arginine/serine-rich protein splicing factor 31A	11.11	Spliceosome
72	comp132819_c0	TWN2; Valyl-tRNA synthetase	11.11	Aminoacyl-tRNA biosynthesis
73	comp135677_c0	putative glucosamine-	11.09	Alanine, aspartate and glutamate metabolism;

		fructose-6-phosphate aminotransferase		Amino sugar and nucleotide sugar metabolism; Metabolic pathways; Biosynthesis of antibiotics
74	comp107974_c0	CUL1; cullin 1	11.08	Ubiquitin mediated proteolysis; Protein processing in endoplasmic reticulum
75	comp109295_c0	TSB2; tryptophan synthase beta chain	10.98	Glycine, serine and threonine metabolism; Phenylalanine, tyrosine and tryptophan biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of amino acids
76	comp125633_c1	peroxidase 52	10.98	Phenylalanine metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
77	comp109807_c0	GLX1; glyoxalase I homolog GLX1	10.92	Pyruvate metabolism
78	comp136955_c0	PGP6; ABC transporter B family member 6	10.91	ABC transporters
79	comp118330_c0	ATPD; F-type H+-transporting ATPase subunit delta	10.90	Oxidative phosphorylation; Photosynthesis; Metabolic pathways
80	comp125296_c0	ribosomal protein L18e/L15	10.90	Ribosome
81	comp116408_c0	PHYB; phytochrome B	10.90	Circadian rhythm - plant
82	comp129598_c0	EER4; transcription initiation factor TFIID subunit 12B	10.88	Basal transcription factors
83	comp126457_c0	choline kinase	10.88	Glycerophospholipid metabolism; Metabolic pathways
84	comp120130_c0	ACX3; acyl-coenzyme A oxidase 3	10.87	Fatty acid degradation; alpha-Linolenic acid metabolism; Biosynthesis of unsaturated fatty

				acids; Metabolic pathways; Biosynthesis of secondary metabolites; Fatty acid metabolism; Peroxisome
85	comp112893_c0	ATPMEPCRB; Probable pectinesterase/pectinesterase inhibitor 41	10.84	Pentose and glucuronate interconversions; Starch and sucrose metabolism; Metabolic pathways
86	comp108395_c0	2-hydroxyacyl-CoA lyase	10.83	Peroxisome
87	comp111876_c0	CYCD3;2; cyclin-D3-2	10.82	Plant hormone signal transduction
88	comp121185_c0	DNAse I-like superfamily protein	10.81	RNA degradation
89	comp117202_c0	TSA1; tryptophan synthase alpha chain	10.80	Glycine, serine and threonine metabolism; Phenylalanine, tyrosine and tryptophan biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of amino acids
90	comp106844_c0	KIN10; SNF1-related protein kinase catalytic subunit alpha KIN10	10.80	Regulation of autophagy
91	comp120985_c0	SCL28; SC35-like splicing factor 28	10.77	Spliceosome
92	comp124257_c0	PP2A; serine/threonine protein phosphatase 2A	10.76	mRNA surveillance pathway
93	comp119846_c0	mMDH1; malate dehydrogenase 1	10.75	Citrate cycle (TCA cycle); Cysteine and methionine metabolism; Pyruvate metabolism; Glyoxylate and dicarboxylate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism

94	comp105486_c0	40S ribosomal protein S15-4	10.74	Ribosome
95	comp133351_c0	GPAT4; bifunctional sn-glycerol-3-phosphate 2-O-acyltransferase/phosphatase	10.74	Glycerolipid metabolism; Glycerophospholipid metabolism; Metabolic pathways; Biosynthesis of secondary metabolites
96	comp82782_c0	60S ribosomal protein L15-1	10.73	Ribosome
97	comp125342_c0	ESP3; DEAH RNA helicase homolog PRP2	10.69	Spliceosome
98	comp104265_c0	STT3A; dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3A	10.68	N-Glycan biosynthesis; Metabolic pathways; Protein processing in endoplasmic reticulum
99	comp100611_c0	60S acidic ribosomal family protein	10.68	Ribosome
100	comp124879_c0	VPS24.1; vacuolar protein sorting-associated protein 24-1	10.67	Endocytosis
101	comp129824_c0	transducin/WD40 domain-containing protein	10.67	Ribosome biogenesis in eukaryotes
102	comp103745_c0	arginase	10.66	Arginine biosynthesis; Arginine and proline metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Biosynthesis of amino acids
103	comp104059_c0	glycine cleavage system H protein 2	10.64	Glycine, serine and threonine metabolism; Glyoxylate and dicarboxylate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics
104	comp104948_c0	UPL2; ubiquitin-protein	10.63	Ubiquitin mediated proteolysis

		ligase 2		
105	comp107417_c0	endosomal targeting BRO1-like domain-containing protein	10.61	Endocytosis
106	comp117156_c0	SNRK2.10; SNF1-related protein kinase 2.10	10.60	Plant hormone signal transduction
107	comp97444_c0	LCB1; serine palmitoyltransferase	10.59	Sphingolipid metabolism; Metabolic pathways
108	comp103214_c0	APL3; glucose-1-phosphate adenylyltransferase large subunit 3	10.59	Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways; Biosynthesis of secondary metabolites
109	comp109370_c0	RD19; cysteine proteinase RD19a	10.18	Plant-pathogen interaction
110	comp94346_c0	USP; UDP-sugar pyrophosphorylase	10.15	Pentose and glucuronate interconversions; Galactose metabolism; Ascorbate and aldarate metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways; Biosynthesis of antibiotics
111	comp115934_c0	RCE1; RUB1 conjugating enzyme 1	10.04	Ubiquitin mediated proteolysis
112	comp112617_c0	60S ribosomal protein L22-2	9.78	Ribosome
113	comp99451_c0	PBD1; 20S proteasome beta subunit D1	9.42	Proteasome
114	comp121314_c0	FLS1; flavonol synthase 1	9.17	Flavonoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
115	comp107312_c0	HA5; H(+)-ATPase 5	9.16	Oxidative phosphorylation
116	comp125262_c0	HCEF1; fructose-1,6-bisphosphatase	8.16	Glycolysis / Gluconeogenesis; Pentose phosphate pathway; Fructose and mannose

				metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
117	comp119277_c1	TAF13; TBP-associated factor 13	7.67	Basal transcription factors
118	comp125099_c0	EMB2719; probable 26S proteasome non-ATPase regulatory subunit 3a	7.34	Proteasome
119	comp71078_c0	SEC; putative UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase SEC	7.23	Other types of O-glycan biosynthesis
120	comp92728_c0	2-hydroxyacyl-CoA lyase	6.69	Peroxisome
121	comp112255_c0	CAM7; calmodulin 7	5.10	Phosphatidylinositol signaling system; Plant-pathogen interaction
122	comp125382_c0	HPT1; homogentisate phytyltransferase 1	3.65	Ubiquinone and other terpenoid-quinone biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
123	comp111347_c0	HDS; 4-hydroxy-3-methylbut-2-enyl diphosphate synthase	2.29	Terpenoid backbone biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics
124	comp102407_c0	DEA(D/H)-box RNA helicase family protein	1.19	RNA degradation

Table 2. Up-regulated genes in leaf in cultivar LN

	Scaffold	Annotation	log2(leaf/rhiozme)	Pathway
1	comp110162_c0	Preprotein translocase Sec, Sec61-beta subunit protein	15.60	Protein export; Protein processing in endoplasmic reticulum; Phagosome
2	comp107831_c0	protein MIDASIN1	15.25	Ribosome biogenesis in eukaryotes
3	comp114089_c0	E3 ubiquitin-protein ligase SINAT3	15.00	Ubiquitin mediated proteolysis
4	comp113126_c0	APC10; anaphase-promoting complex subunit 10	14.34	Ubiquitin mediated proteolysis
5	comp107761_c1	hypothetical protein	13.49	Taurine and hypotaurine metabolism; Metabolic pathways
6	comp110408_c0	VPS60.2; vacuolar protein sorting protein 60.2	13.40	Endocytosis
7	comp96770_c0	BGLU3; beta-glucosidase 3	13.32	Cyanoamino acid metabolism; Starch and sucrose metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites

8	comp114024_c0	transport protein SEC13A	13.32	RNA transport; Protein processing in endoplasmic reticulum
9	comp93096_c0	PLA2-ALPHA; phospholipase A2-alpha	13.16	Glycerophospholipid metabolism; Ether lipid metabolism; Arachidonic acid metabolism; Linoleic acid metabolism; alpha-Linolenic acid metabolism; Metabolic pathways; Biosynthesis of secondary metabolites
10	comp123248_c0	RNA recognition motif-containing protein	13.14	Spliceosome
11	comp103986_c0	EIF2_GAMMA; translation initiation factor eIF-2 gamma subunit	13.11	RNA transport
12	comp106512_c0	40S ribosomal protein S16-1	13.07	Ribosome
13	comp105413_c0	UBC1; ubiquitin-conjugating enzyme E2 1	13.06	Ubiquitin mediated proteolysis
14	comp129598_c0	EER4; transcription initiation factor TFIID subunit 12B	12.83	Basal transcription factors
15	comp112877_c0	LHCA5; photosystem I light harvesting complex protein 5	12.71	Photosynthesis - antenna proteins
16	comp131391_c0	COX15; cytochrome c	12.61	Oxidative phosphorylation; Porphyrin and

		oxidase assembly protein COX15		chlorophyll metabolism; Metabolic pathways; Biosynthesis of secondary metabolites
17	comp125387_c0	transport protein SEC13A	12.44	RNA transport; Protein processing in endoplasmic reticulum
18	comp120004_c0	CI51; NADH dehydrogenase [ubiquinone] flavoprotein 1	12.40	Oxidative phosphorylation; Metabolic pathways
19	comp102333_c0	RD19; cysteine proteinase RD19a	12.24	Plant-pathogen interaction
20	comp113913_c0	RPL24A; 60S ribosomal protein L24-1	12.21	Ribosome
21	comp49043_c0	NADP-ME2; NADP-dependent malic enzyme 2	12.21	Pyruvate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Carbon metabolism
22	comp99535_c0	60S ribosomal protein L15-1	12.08	Ribosome
23	comp122860_c0	palmitoyl protein thioesterase family protein	12.06	Fatty acid elongation; Metabolic pathways; Fatty acid metabolism
24	comp118847_c0	UBC9; ubiquitin conjugating enzyme 9	12.06	Ubiquitin mediated proteolysis; Protein processing in endoplasmic reticulum
25	comp113826_c0	ubiquinol-cytochrome C reductase hinge protein	12.02	Oxidative phosphorylation; Metabolic pathways

26	comp109061_c0	SDH2-1; succinate dehydrogenase [ubiquinone] iron-sulfur subunit 1	11.96	Citrate cycle (TCA cycle); Oxidative phosphorylation; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
27	comp130901_c0	ELC-Like; ELC-like protein	11.93	Endocytosis
28	comp118288_c0	CHIP; E3 ubiquitin-protein ligase CHIP	11.82	Ubiquitin mediated proteolysis; Protein processing in endoplasmic reticulum
29	comp111545_c0	OTU-like cysteine protease family protein	11.71	Protein processing in endoplasmic reticulum
30	comp115622_c0	RNA recognition motif-containing protein	11.65	Spliceosome
31	comp86890_c0	RR2; transcription factor response regulator 2	11.62	Plant hormone signal transduction
32	comp137165_c0	Beta-glucosidase, GBA2 type family protein	11.55	Other glycan degradation; Sphingolipid metabolism; Metabolic pathways
33	comp108044_c0	MDH; malate dehydrogenase	11.55	Citrate cycle (TCA cycle); Cysteine and methionine metabolism; Pyruvate metabolism; Glyoxylate and dicarboxylate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary

				metabolites; Biosynthesis of antibiotics; Carbon metabolism
34	comp103293_c0	FAB1B; phosphatidylinositol-3P 5-kinase-like	11.52	Inositol phosphate metabolism; Phosphatidylinositol signaling system; Phagosome
35	comp127244_c0	AK-LYS1; aspartokinase 1	11.39	Glycine, serine and threonine metabolism; Monobactam biosynthesis; Cysteine and methionine metabolism; Lysine biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; 2-Oxocarboxylic acid metabolism; Biosynthesis of amino acids
36	comp108887_c0	CNX7; molybdenum cofactor synthesis family protein	11.32	Folate biosynthesis; Metabolic pathways; Sulfur relay system
37	comp135677_c0	putative glucosamine-fructose-6-phosphate aminotransferase	11.27	Alanine, aspartate and glutamate metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways; Biosynthesis of antibiotics
38	comp131748_c0	ALDH11A3; NADP-dependent glyceraldehyde-3-phosphate dehydrogenase	11.19	Glycolysis / Gluconeogenesis; Pentose phosphate pathway; Metabolic pathways; Carbon metabolism

39	comp114670_c0	CML38; calcium-binding protein CML38	11.18	Plant-pathogen interaction
40	comp130472_c0	glutamine-dependent NAD(+) synthetase	11.17	Nicotinate and nicotinamide metabolism; Metabolic pathways
41	comp134956_c0	Hrd3/Sel1L-like protein	11.07	Protein processing in endoplasmic reticulum
42	comp85153_c0	FT; protein FLOWERING LOCUS T	11.06	Circadian rhythm - plant
43	comp104585_c0	phosphoinositide phospholipase C 6	11.06	Inositol phosphate metabolism; Metabolic pathways; Phosphatidylinositol signaling system
44	comp132946_c0	PAL1; phenylalanine ammonia-lyase 1	11.02	Phenylalanine metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
45	comp95101_c0	emb2742; CTP synthase family protein	11.01	Pyrimidine metabolism; Metabolic pathways
46	comp132067_c0	HSP60; heat shock protein 60	11.00	RNA degradation
47	comp108458_c0	SHM1; serine transhydroxymethyltransferase 1	10.95	Glycine, serine and threonine metabolism; Cyanoamino acid metabolism; Glyoxylate and dicarboxylate metabolism; One carbon pool by folate; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of

				antibiotics; Carbon metabolism; Biosynthesis of amino acids
48	comp113686_c0	CNX1; calnexin 1	10.91	Protein processing in endoplasmic reticulum; Phagosome
49	comp106272_c0	acyl-CoA thioesterase family protein	10.86	Fatty acid elongation; Biosynthesis of unsaturated fatty acids; Metabolic pathways; Biosynthesis of secondary metabolites
50	comp120130_c0	ACX3; acyl-coenzyme A oxidase 3	10.85	Fatty acid degradation; alpha-Linolenic acid metabolism; Biosynthesis of unsaturated fatty acids; Metabolic pathways; Biosynthesis of secondary metabolites; Fatty acid metabolism; Peroxisome
51	comp128012_c0	pfkB-like carbohydrate kinase family protein	10.71	Pentose phosphate pathway
52	comp103975_c0	CAM7; calmodulin 7	10.71	Phosphatidylinositol signaling system; Plant-pathogen interaction
53	comp111085_c0	RPL27; 50S ribosomal protein L27	10.68	Ribosome
54	comp107134_c0	SDN3; small RNA degrading nuclease 3	10.67	Ribosome biogenesis in eukaryotes

55	comp121548_c0	ribophorin I	10.67	N-Glycan biosynthesis; Metabolic pathways; Protein processing in endoplasmic reticulum
56	comp100996_c0	BGLU13; beta glucosidase 13	10.66	Cyanoamino acid metabolism; Starch and sucrose metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
57	comp121860_c0	D-glycerate 3-kinase	10.64	Glycine, serine and threonine metabolism; Glycerolipid metabolism; Glyoxylate and dicarboxylate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
58	comp80805_c0	PPa3; pyrophosphorylase 3	10.63	Oxidative phosphorylation
59	comp100053_c0	EP3; chitinase class IV	10.59	Amino sugar and nucleotide sugar metabolism
60	comp124016_c0	DFB; folylpolyglutamate synthase 1	10.57	Folate biosynthesis; Metabolic pathways
61	comp110195_c0	F-type H+-transporting ATPase subunit b	10.56	Oxidative phosphorylation; Photosynthesis; Metabolic pathways
62	comp136532_c0	NADK2; NAD kinase 2	10.54	Nicotinate and nicotinamide metabolism; Metabolic pathways
63	comp104690_c0	40S ribosomal protein S9-2	10.54	Ribosome

64	comp95027_c0	xylose isomerase	10.49	Pentose and glucuronate interconversions; Fructose and mannose metabolism; Metabolic pathways
65	comp121185_c0	DNAse I-like superfamily protein	10.49	RNA degradation
66	comp116408_c0	PHYB; phytochrome B	10.49	Circadian rhythm - plant
67	comp114387_c0	copper amine oxidase 2	10.49	Glycine, serine and threonine metabolism; Tyrosine metabolism; Phenylalanine metabolism; beta-Alanine metabolism; Isoquinoline alkaloid biosynthesis; Tropane, piperidine and pyridine alkaloid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
68	comp106000_c0	PWWP domain-containing protein	10.47	Cysteine and methionine metabolism; Metabolic pathways
69	comp41755_c0	26S proteasome regulatory complex, non-ATPase subcomplex, Rpn2/Psmd1 subunit	10.46	Proteasome
70	comp94473_c0	eukaryotic translation initiation factor 3 subunit 7	10.44	RNA transport

71	comp105717_c0	PPC1; phosphoenolpyruvate carboxylase 1	10.44	Pyruvate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Carbon metabolism
72	comp124060_c0	SKP2A; F-box protein SKP2A	10.41	Ubiquitin mediated proteolysis
73	comp113633_c0	GPX1; phospholipid hydroperoxide glutathione peroxidase 1	10.39	Glutathione metabolism; Arachidonic acid metabolism
74	comp120985_c0	SCL28; SC35-like splicing factor 28	10.38	Spliceosome
75	comp91567_c0	ACO1; aconitate hydratase 1	10.32	Citrate cycle (TCA cycle); Glyoxylate and dicarboxylate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; 2-Oxocarboxylic acid metabolism; Biosynthesis of amino acids
76	comp105625_c0	PETC; cytochrome b6-f complex iron-sulfur subunit	10.32	Photosynthesis; Metabolic pathways
77	comp108546_c0	NADH-cytochrome b5 reductase-like protein	10.31	Amino sugar and nucleotide sugar metabolism

78	comp112218_c0	AAO3; abscisic-aldehyde oxidase	10.21	Carotenoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
79	comp116160_c0	40S ribosomal protein S9-2	10.20	Ribosome
80	comp88116_c0	EIF3A; eukaryotic translation initiation factor 3A	10.18	RNA transport
81	comp136415_c0	BMS1 domain-containing protein	10.17	Ribosome biogenesis in eukaryotes
82	comp132000_c0	GSTL2; glutathione transferase lambda 2	10.16	Glutathione metabolism
83	comp103751_c0	ARR4; two-component response regulator ARR4	10.15	Plant hormone signal transduction
84	comp127808_c0	CKB4; casein kinase II beta subunit 4	10.14	Ribosome biogenesis in eukaryotes; Circadian rhythm - plant
85	comp129842_c0	Succinyl-CoA ligase [GDP-forming] subunit alpha-2	10.11	Citrate cycle (TCA cycle); Propanoate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
86	comp96613_c0	SYP124; syntaxin-124	10.11	SNARE interactions in vesicular transport
87	comp119846_c0	mMDH1; malate dehydrogenase 1	10.09	Citrate cycle (TCA cycle); Cysteine and methionine metabolism; Pyruvate metabolism;

				Glyoxylate and dicarboxylate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
88	comp97467_c0	ALDH3I1; aldehyde dehydrogenase 3I1	10.09	Glycolysis / Gluconeogenesis; Pentose and glucuronate interconversions; Ascorbate and aldarate metabolism; Fatty acid degradation; Valine, leucine and isoleucine degradation; Lysine degradation; Arginine and proline metabolism; Histidine metabolism; Tryptophan metabolism; beta-Alanine metabolism; Glycerolipid metabolism; Pyruvate metabolism; Limonene and pinene degradation; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics
89	comp107138_c0	CLA1; 1-deoxy-D-xylulose-5-phosphate synthase	10.06	Thiamine metabolism; Terpenoid backbone biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics

90	comp136154_c0	hypothetical protein	10.06	RNA degradation
91	comp116950_c0	MEE58; adenosylhomocysteinase 1	10.04	Cysteine and methionine metabolism; Metabolic pathways
92	comp112284_c1	LAX3; auxin influx carrier LAX3	10.04	Plant hormone signal transduction
93	comp113727_c0	ARR9; two-component response regulator ARR9	10.02	Plant hormone signal transduction
94	comp109260_c0	CAM2; calmodulin 2	9.99	Phosphatidylinositol signaling system; Plant-pathogen interaction
95	comp136826_c0	40S ribosomal protein S3a-1	9.96	Ribosome
96	comp82559_c0	endonuclease/exonuclease/phosphatase family protein	9.95	Base excision repair
97	comp126348_c0	AUD1; UDP-glucuronic acid decarboxylase	9.95	Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways
98	comp111549_c0	LCB1; serine palmitoyltransferase	9.95	Sphingolipid metabolism; Metabolic pathways
99	comp116279_c0	SQD2; sulfoquinovosyldiacylglycerol 2	9.95	Glycerolipid metabolism; Metabolic pathways

100	comp110433_c0	LHCA1; chlorophyll a-b binding protein 6	9.93	Photosynthesis - antenna proteins
101	comp98036_c0	CNX1; calnexin 1	9.93	Protein processing in endoplasmic reticulum; Phagosome
102	comp121778_c0	ubiquinol-cytochrome c reductase cytochrome c1 subunit	9.91	Oxidative phosphorylation; Metabolic pathways
103	comp96754_c0	VCR; varicose-related protein	9.91	RNA degradation
104	comp98731_c0	emb2394; 50S ribosomal protein L6	9.90	Ribosome
105	comp113241_c0	ATHAL3B; putative phosphopantothenoylcysteine decarboxylase	9.88	Pantothenate and CoA biosynthesis; Metabolic pathways
106	comp91637_c0	DUT1; deoxyuridine 5'-triphosphate nucleotidohydrolase	9.86	Pyrimidine metabolism; Metabolic pathways
107	comp108312_c0	DNA-directed RNA polymerase, subunit M	9.85	Purine metabolism; Pyrimidine metabolism; Metabolic pathways; RNA polymerase
108	comp95524_c0	RPM1; disease resistance protein RPM1	9.85	Plant-pathogen interaction

109	comp112489_c0	GI; protein GIGANTEA	9.85	Circadian rhythm - plant
110	comp103583_c0	Topoisomerase II-associated protein PAT1	9.84	RNA degradation
111	comp130572_c0	EIF4A-III; DEAD-box ATP-dependent RNA helicase 2	9.83	RNA transport; mRNA surveillance pathway; Spliceosome
112	comp104951_c0	HDS; 4-hydroxy-3-methylbut-2-enyl diphosphate synthase	9.82	Terpenoid backbone biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics
113	comp112994_c0	ATPase, V0 complex, subunit E	9.82	Oxidative phosphorylation; Metabolic pathways; Phagosome
114	comp111585_c0	YKT61; VAMP-like protein YKT61	9.82	SNARE interactions in vesicular transport
115	comp125922_c0	RNA helicase family protein	9.82	RNA degradation
116	comp109249_c0	EDA9; D-3-phosphoglycerate dehydrogenase	9.80	Glycine, serine and threonine metabolism; Metabolic pathways; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
117	comp123097_c0	UGP2; UTP--glucose-1-phosphate uridylyltransferase 1	9.78	Pentose and glucuronate interconversions; Galactose metabolism; Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways; Biosynthesis of

				antibiotics
118	comp123838_c0	RCI3; peroxidase 3	9.76	Phenylalanine metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
119	comp102246_c0	UBA1; ubiquitin-activating enzyme E1 1	9.75	Ubiquitin mediated proteolysis
120	comp112735_c0	CERK1; chitin elicitor receptor kinase 1	9.75	Plant-pathogen interaction
121	comp103134_c0	ATB_BETA; serine/threonine protein phosphatase 2A 55 kDa regulatory subunit B beta isoform	9.73	mRNA surveillance pathway
122	comp106412_c0	Bet1-like protein	9.73	SNARE interactions in vesicular transport
123	comp128411_c0	CKX1; cytokinin dehydrogenase 1	9.70	Zeatin biosynthesis
124	comp135921_c0	RANGAP1; RAN GTPase-activating protein 1	9.69	RNA transport
125	comp97287_c0	Glycosyl hydrolase family protein	9.69	Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways

126	comp121333_c0	CCT2; phosphorylcholine cytidylyltransferase2	9.68	Glycerophospholipid metabolism; Metabolic pathways
127	comp115034_c0	BRU6; indole-3-acetic acid-amido synthetase GH3.2	9.68	Plant hormone signal transduction
128	comp108632_c0	FNR2; ferredoxin--NADP reductase, leaf isozyme 2	9.67	Photosynthesis; Metabolic pathways
129	comp110785_c0	ACD2; red chlorophyll catabolite reductase	9.62	Porphyrin and chlorophyll metabolism; Biosynthesis of secondary metabolites
130	comp124038_c0	CYL1; alpha-N-acetylglucosaminidase	9.62	Glycosaminoglycan degradation; Metabolic pathways
131	comp132354_c0	CS26; cysteine synthase 26	9.62	Cysteine and methionine metabolism; Sulfur metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
132	comp105933_c0	APT1; adenine phosphoribosyl transferase 1	9.61	Purine metabolism; Metabolic pathways
133	comp118231_c0	ATP3; ATP synthase subunit gamma	9.61	Oxidative phosphorylation; Metabolic pathways
134	comp100237_c0	HSP81-2; heat shock protein	9.60	Protein processing in endoplasmic reticulum;

		81-2		Plant-pathogen interaction
135	comp93701_c0	SUS3; sucrose synthase 3	9.60	Starch and sucrose metabolism; Metabolic pathways
136	comp136730_c0	WRKY2; putative WRKY transcription factor 2	9.57	Plant-pathogen interaction
137	comp91392_c0	4CL3; 4-coumarate--CoA ligase 3	9.57	Ubiquinone and other terpenoid-quinone biosynthesis; Phenylalanine metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
138	comp93781_c0	polynucleotidyl transferase, ribonuclease H-like superfamily protein	9.57	RNA degradation
139	comp94547_c0	WRKY2; putative WRKY transcription factor 2	9.56	Plant-pathogen interaction
140	comp108466_c0	40S ribosomal protein S20-1	9.56	Ribosome
141	comp120952_c0	2-methoxy-6-polyprenyl-1,4-benzoquinol methylase	9.56	Ubiquinone and other terpenoid-quinone biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
142	comp121146_c0	ribose 5-phosphate isomerase A	9.55	Pentose phosphate pathway; Carbon fixation in photosynthetic organisms; Metabolic pathways;

				Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
143	comp111454_c0	HYD1; probable 3-beta-hydroxysteroid-Delta(8),Delta(7)-isomerase	9.55	Steroid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
144	comp103607_c0	probable methionine--tRNA ligase	9.55	Selenocompound metabolism; Aminoacyl-tRNA biosynthesis
145	comp106721_c0	BGLU47; beta-glucosidase 47	9.53	Cyanoamino acid metabolism; Starch and sucrose metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
146	comp118504_c0	MPK6; MAP kinase 6	9.53	Plant hormone signal transduction; Plant-pathogen interaction
147	comp103057_c0	ALATS; Alanyl-tRNA synthetase	9.52	Aminoacyl-tRNA biosynthesis
148	comp123997_c0	BCCP2; biotin carboxyl carrier protein 2	9.52	Fatty acid biosynthesis; Pyruvate metabolism; Propanoate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism;

				Fatty acid metabolism
149	comp108197_c0	ubiquitin-related modifier 1	9.50	Sulfur relay system
150	comp113557_c0	FD3; ferredoxin 3	9.48	Photosynthesis
151	comp110869_c0	MTACP-1; acyl carrier protein 1	9.48	Oxidative phosphorylation; Metabolic pathways
152	comp98742_c0	UXS6; UDP-XYL synthase 6	9.48	Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways
153	comp92728_c0	2-hydroxyacyl-CoA lyase	9.45	Peroxisome
154	comp107312_c0	HA5; H(+)-ATPase 5	9.45	Oxidative phosphorylation
155	comp112255_c0	CAM7; calmodulin 7	9.43	Phosphatidylinositol signaling system; Plant-pathogen interaction
156	comp94346_c0	USP; UDP-sugar pyrophosphorylase	9.43	Pentose and glucuronate interconversions; Galactose metabolism; Ascorbate and aldarate metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways; Biosynthesis of antibiotics
157	comp111347_c0	HDS; 4-hydroxy-3-methylbut-2-enyl diphosphate synthase	9.41	Terpenoid backbone biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics

158	comp119277_c1	TAF13; TBP-associated factor 13	9.40	Basal transcription factors
159	comp121314_c0	FLS1; flavonol synthase 1	9.39	Flavonoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
160	comp109370_c0	RD19; cysteine proteinase RD19a	9.39	Plant-pathogen interaction
161	comp125262_c0	HCEF1; fructose-1,6-bisphosphatase	9.38	Glycolysis / Gluconeogenesis; Pentose phosphate pathway; Fructose and mannose metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
162	comp112617_c0	60S ribosomal protein L22-2	9.37	Ribosome
163	comp115934_c0	RCE1; RUB1 conjugating enzyme 1	9.37	Ubiquitin mediated proteolysis
164	comp99451_c0	PBD1; 20S proteasome beta subunit D1	9.37	Proteasome
165	comp102407_c0	DEA(D/H)-box RNA helicase family protein	9.36	RNA degradation
166	comp125382_c0	HPT1; homogentisate	9.35	Ubiquinone and other terpenoid-quinone

		phytyltransferase 1		biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
167	comp71078_c0	SEC; putative UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase SEC	9.34	Other types of O-glycan biosynthesis
168	comp125099_c0	EMB2719; probable 26S proteasome non-ATPase regulatory subunit 3a	9.34	Proteasome
169	comp117156_c0	SNRK2.10; SNF1-related protein kinase 2.10	9.33	Plant hormone signal transduction
170	comp123397_c0	RNR1; ribonucleoside-diphosphate reductase large subunit	9.31	Purine metabolism; Pyrimidine metabolism; Glutathione metabolism; Metabolic pathways
171	comp104081_c0	CYTC-2; cytochrome c-2	9.27	Sulfur metabolism; Metabolic pathways
172	comp54218_c0	XPO1A; exportin 1A	9.17	Ribosome biogenesis in eukaryotes; RNA transport
173	comp113416_c0	PSBP-1; photosystem II subunit P-1	9.10	Photosynthesis; Metabolic pathways

174	comp108406_c0	Hsp81.4; Heat shock protein 90-4	8.99	Protein processing in endoplasmic reticulum; Plant-pathogen interaction
175	comp99216_c0	splicing factor PWI and RNA recognition motif-containing protein	8.97	Spliceosome
176	comp108027_c0	pre-mRNA-splicing factor 38A	8.97	Spliceosome
177	comp99689_c0	40S ribosomal protein S28-1	8.97	Ribosome
178	comp109235_c0	ARF1; auxin response factor 1	8.96	Plant hormone signal transduction
179	comp93773_c0	FT; protein FLOWERING LOCUS T	8.90	Circadian rhythm - plant
180	comp89640_c0	ABCB4; auxin efflux transmembrane transporter MDR4	8.89	ABC transporters
181	comp114637_c0	DAD2; Defender against cell death 2	8.84	N-Glycan biosynthesis; Metabolic pathways; Protein processing in endoplasmic reticulum
182	comp109574_c0	MAB1; pyruvate dehydrogenase E1 beta	8.74	Glycolysis / Gluconeogenesis; Citrate cycle (TCA cycle); Pyruvate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism

183	comp113408_c0	PME2; pectin methylesterase 2	8.69	Pentose and glucuronate interconversions; Starch and sucrose metabolism; Metabolic pathways
184	comp101643_c0	40S ribosomal protein S29	8.68	Ribosome
185	comp125760_c0	RHM1; UDP-L-rhamnose synthase	8.59	Amino sugar and nucleotide sugar metabolism
186	comp111899_c0	MEKK1; mitogen-activated protein kinase kinase kinase 1	8.43	Plant-pathogen interaction
187	comp92383_c0	MGD2; monogalactosyldiacylglycerol synthase 2	8.41	Glycerolipid metabolism; Metabolic pathways
188	comp118078_c0	ATP3; ATP synthase subunit gamma	8.36	Oxidative phosphorylation; Metabolic pathways
189	comp98858_c0	ATCES1; acyl-CoA independent ceramide synthase	8.36	Sphingolipid metabolism
190	comp95616_c0	ATDAD1; dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit DAD1	8.29	N-Glycan biosynthesis; Metabolic pathways; Protein processing in endoplasmic reticulum
191	comp89148_c0	SUS5; sucrose synthase 5	8.19	Starch and sucrose metabolism; Metabolic

				pathways
192	comp114195_c0	putative NADH-ubiquinone oxidoreductase B18 subunit	7.88	Oxidative phosphorylation; Metabolic pathways
193	comp68588_c0	60S ribosomal protein L11-2	7.61	Ribosome
194	comp111097_c0	60S ribosomal protein L37-3	7.52	Ribosome
195	comp112333_c0	ABF2; abscisic acid responsive elements-binding factor 2	7.11	Plant hormone signal transduction
196	comp129724_c0	ceramide kinase	6.00	Sphingolipid metabolism

Table 3. Up-regulated genes in leaf in cultivar OU

	Scaffold	Annotation	log2(leaf/rhizome)	Pathway
1	comp107831_c0	protein MIDASIN1	16.44	Ribosome biogenesis in eukaryotes
2	comp110162_c0	Preprotein translocase Sec, Sec61-beta subunit protein	15.94	Protein export; Protein processing in endoplasmic reticulum; Phagosome
3	comp113126_c0	APC10; anaphase-promoting complex subunit 10	15.01	Ubiquitin mediated proteolysis
4	comp107761_c1	hypothetical protein	13.56	Taurine and hypotaurine metabolism; Metabolic pathways
5	comp116992_c0	Rae1-like protein	12.93	RNA transport
6	comp111085_c0	RPL27; 50S ribosomal protein L27	12.84	Ribosome
7	comp114024_c0	transport protein SEC13A	12.61	RNA transport; Protein processing in endoplasmic reticulum
8	comp120565_c0	ZAC; ADP-ribosylation factor GTPase-activating protein AGD12	12.61	Endocytosis
9	comp118352_c0	CHIP; E3 ubiquitin-protein	12.58	Ubiquitin mediated proteolysis; Protein

		ligase CHIP		processing in endoplasmic reticulum
10	comp96770_c0	BGLU3; beta-glucosidase 3	12.21	Cyanoamino acid metabolism; Starch and sucrose metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
11	comp103293_c0	FAB1B; phosphatidylinositol-3P 5-kinase-like	12.20	Inositol phosphate metabolism; Phosphatidylinositol signaling system; Phagosome
12	comp49043_c0	NADP-ME2; NADP-dependent malic enzyme 2	12.20	Pyruvate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Carbon metabolism
13	comp113826_c0	ubiquinol-cytochrome C reductase hinge protein	12.17	Oxidative phosphorylation; Metabolic pathways
14	comp99535_c0	60S ribosomal protein L15-1	12.03	Ribosome
15	comp118101_c0	SSI2; acyl-[acyl-carrier-protein] desaturase	11.96	Fatty acid biosynthesis; Biosynthesis of unsaturated fatty acids; Fatty acid metabolism
16	comp112004_c0	PLDDELTA; phospholipase D delta	11.80	Glycerophospholipid metabolism; Ether lipid metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Endocytosis
17	comp105413_c0	UBC1; ubiquitin-conjugating	11.79	Ubiquitin mediated proteolysis

		enzyme E2 1		
18	comp103847_c0	GA1; Ent-copalyl diphosphate synthase	11.74	Diterpenoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
19	comp111876_c0	CYCD3;2; cyclin-D3-2	11.65	Plant hormone signal transduction
20	comp124809_c1	BCAT3; branched-chain-amino-acid aminotransferase 3	11.58	Valine, leucine and isoleucine degradation; Valine, leucine and isoleucine biosynthesis; Pantothenate and CoA biosynthesis; Glucosinolate biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; 2-Oxocarboxylic acid metabolism; Biosynthesis of amino acids
21	comp93892_c0	SAG24; 60S ribosomal protein L10-3	11.56	Ribosome
22	comp124216_c0	ASA2; anthranilate synthase component I-2	11.52	Phenylalanine, tyrosine and tryptophan biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of amino acids
23	comp117908_c0	lipid phosphate phosphatase gamma	11.37	N-Glycan biosynthesis
24	comp121776_c0	ZTL; adagio protein 1	11.27	Circadian rhythm - plant

25	comp120130_c0	ACX3; acyl-coenzyme A oxidase 3	11.25	Fatty acid degradation; alpha-Linolenic acid metabolism; Biosynthesis of unsaturated fatty acids; Metabolic pathways; Biosynthesis of secondary metabolites; Fatty acid metabolism; Peroxisome
26	comp111389_c0	60S ribosomal protein L34-3	11.22	Ribosome
27	comp130078_c0	ATERDJ3B; DNAJ heat shock protein ATERDJ3B	11.21	Protein processing in endoplasmic reticulum
28	comp122860_c0	palmitoyl protein thioesterase family protein	11.21	Fatty acid elongation; Metabolic pathways; Fatty acid metabolism
29	comp131917_c0	NCED4; nine-cis-epoxycarotenoid dioxygenase 4	11.20	Carotenoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
30	comp94579_c0	ELC; protein ELC	11.10	Endocytosis
31	comp132819_c0	TWN2; Valyl-tRNA synthetase	11.03	Aminoacyl-tRNA biosynthesis
32	comp92492_c0	Pre-mRNA-splicing factor 3	11.03	Spliceosome
33	comp130472_c0	glutamine-dependent NAD(+) synthetase	11.03	Nicotinate and nicotinamide metabolism; Metabolic pathways
34	comp107097_c0	CAM3; calmodulin 3	10.99	Phosphatidylinositol signaling system; Plant-

				pathogen interaction
35	comp121185_c0	DNAse I-like superfamily protein	10.97	RNA degradation
36	comp124038_c0	CYL1; alpha-N-acetylglucosaminidase	10.92	Glycosaminoglycan degradation; Metabolic pathways
37	comp92408_c0	LHCB4.2; light harvesting complex photosystem II interacting protein Lhcb4.2	10.87	Photosynthesis - antenna proteins; Metabolic pathways
38	comp123485_c0	GCIP-interacting family protein	10.86	Spliceosome
39	comp96382_c0	PUX4; CDC48-interacting UBX-domain protein 4	10.81	Protein processing in endoplasmic reticulum
40	comp103672_c0	RHS2; calmodulin-like protein 7	10.78	Plant-pathogen interaction
41	comp94473_c0	eukaryotic translation initiation factor 3 subunit 7	10.72	RNA transport
42	comp95524_c0	RPM1; disease resistance protein RPM1	10.70	Plant-pathogen interaction
43	comp82559_c0	endonuclease/exonuclease/phosphatase family protein	10.67	Base excision repair

44	comp116258_c0	oligosaccharyltransferase complex/magnesium transporter family protein	10.62	N-Glycan biosynthesis; Metabolic pathways; Protein processing in endoplasmic reticulum
45	comp130901_c0	ELC-Like; ELC-like protein	10.60	Endocytosis
46	comp103975_c0	CAM7; calmodulin 7	10.60	Phosphatidylinositol signaling system; Plant-pathogen interaction
47	comp112811_c0	GPAT6; bifunctional sn-glycerol-3-phosphate 2-O-acyltransferase/phosphatase	10.54	Glycerolipid metabolism; Glycerophospholipid metabolism; Metabolic pathways; Biosynthesis of secondary metabolites
48	comp123402_c0	GAE3; UDP-D-glucuronate 4-epimerase 3	10.51	Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways
49	comp131704_c0	HSP60; heat shock protein 60	10.46	RNA degradation
50	comp135677_c0	putative glucosamine-fructose-6-phosphate aminotransferase	10.44	Alanine, aspartate and glutamate metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways; Biosynthesis of antibiotics
51	comp113408_c0	PME2; pectin methylesterase 2	10.43	Pentose and glucuronate interconversions; Starch and sucrose metabolism; Metabolic pathways
52	comp133351_c0	GPAT4; bifunctional sn-glycerol-3-phosphate 2-O-	10.39	Glycerolipid metabolism; Glycerophospholipid metabolism; Metabolic pathways; Biosynthesis of

		acyltransferase/phosphatase		secondary metabolites
53	comp113416_c0	PSBP-1; photosystem II subunit P-1	10.38	Photosynthesis; Metabolic pathways
54	comp123248_c0	RNA recognition motif-containing protein	10.37	Spliceosome
55	comp91567_c0	ACO1; aconitate hydratase 1	10.37	Citrate cycle (TCA cycle); Glyoxylate and dicarboxylate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; 2-Oxocarboxylic acid metabolism; Biosynthesis of amino acids
56	comp131748_c0	ALDH11A3; NADP-dependent glyceraldehyde-3-phosphate dehydrogenase	10.36	Glycolysis / Gluconeogenesis; Pentose phosphate pathway; Metabolic pathways; Carbon metabolism
57	comp121289_c0	PRP40A; pre-mRNA-processing protein 40A	10.35	Spliceosome
58	comp118390_c1	SAP18; histone deacetylase complex subunit SAP18	10.33	RNA transport; mRNA surveillance pathway
59	comp109295_c0	TSB2; tryptophan synthase beta chain	10.32	Glycine, serine and threonine metabolism; Phenylalanine, tyrosine and tryptophan

				biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of amino acids
60	comp116564_c0	RPL16A; 60S ribosomal protein L16A	10.29	Ribosome
61	comp115233_c0	AlaAT1; alanine aminotransferase	10.28	Arginine biosynthesis; Alanine, aspartate and glutamate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Carbon metabolism; 2-Oxocarboxylic acid metabolism; Biosynthesis of amino acids
62	comp89148_c0	SUS5; sucrose synthase 5	10.26	Starch and sucrose metabolism; Metabolic pathways
63	comp130572_c0	EIF4A-III; DEAD-box ATP-dependent RNA helicase 2	10.25	RNA transport; mRNA surveillance pathway; Spliceosome
64	comp136154_c0	hypothetical protein	10.23	RNA degradation
65	comp105675_c0	CML42; calcium-binding protein CML42	10.21	Plant-pathogen interaction
66	comp116408_c0	PHYB; phytochrome B	10.20	Circadian rhythm - plant
67	comp134068_c0	dehydratase-enolase-phosphatase complex 1	10.20	Cysteine and methionine metabolism; Metabolic pathways

68	comp104824_c0	SYNC1; asparaginyl-tRNA synthetase, cytoplasmic 1	10.20	Aminoacyl-tRNA biosynthesis
69	comp80800_c0	CAM3; calmodulin 3	10.19	Phosphatidylinositol signaling system; Plant-pathogen interaction
70	comp113241_c0	ATHAL3B; putative phosphopantothenoylcysteine decarboxylase	10.18	Pantothenate and CoA biosynthesis; Metabolic pathways
71	comp93773_c0	FT; protein FLOWERING LOCUS T	10.18	Circadian rhythm - plant
72	comp109933_c0	AVP1; Pyrophosphate-energized vacuolar membrane proton pump 1	10.16	Oxidative phosphorylation
73	comp97289_c0	U6 snRNA-associated Sm-like protein	10.15	RNA degradation; Spliceosome
74	comp54218_c0	XPO1A; exportin 1A	10.11	Ribosome biogenesis in eukaryotes; RNA transport
75	comp107164_c0	LACS9; long chain acyl-CoA synthetase 9	10.10	Fatty acid biosynthesis; Fatty acid degradation; Metabolic pathways; Fatty acid metabolism; Peroxisome
76	comp110289_c0	PAO2; polyamine oxidase 2	10.08	Arginine and proline metabolism; beta-Alanine

				metabolism
77	comp115775_c0	RPL10B; 60S ribosomal protein L10-2	10.06	Ribosome
78	comp99478_c0	large subunit ribosomal protein L24e	10.02	Ribosome
79	comp110346_c0	RNA-binding (RRM/RBD/RNP motifs) family protein	9.99	Spliceosome
80	comp111185_c0	VHA-C3; vacuolar-type H(+)-ATPase C3	9.99	Oxidative phosphorylation; Metabolic pathways; Phagosome
81	comp124776_c0	QPT; quinolinate phoshoribosyltransferase	9.96	Nicotinate and nicotinamide metabolism; Metabolic pathways
82	comp95616_c0	ATDAD1; dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit DAD1	9.95	N-Glycan biosynthesis; Metabolic pathways; Protein processing in endoplasmic reticulum
83	comp110562_c0	SEC22; vesicle transport protein SEC22	9.95	SNARE interactions in vesicular transport; Phagosome
84	comp121897_c0	CAD5; cinnamyl alcohol dehydrogenase 5	9.92	Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites

85	comp98593_c0	RGA1; DELLA protein RGA	9.86	Plant hormone signal transduction
86	comp97730_c0	splicing factor PWI and RNA recognition motif-containing protein	9.85	Spliceosome
87	comp110408_c0	VPS60.2; vacuolar protein sorting protein 60.2	9.84	Endocytosis
88	comp104123_c0	SYP71; syntaxin-71	9.84	SNARE interactions in vesicular transport
89	comp102126_c0	transketolase	9.82	Pentose phosphate pathway; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
90	comp108154_c0	PUX4; CDC48-interacting UBX-domain protein 4	9.80	Protein processing in endoplasmic reticulum
91	comp117745_c0	60S ribosomal protein L32-1	9.80	Ribosome
92	comp90645_c0	SPT1; serine palmitoyltransferase 1	9.80	Sphingolipid metabolism; Metabolic pathways
93	comp105693_c0	zinc ion binding protein	9.79	Ubiquitin mediated proteolysis
94	comp111910_c0	sec23/sec24-like transport protein	9.79	Protein processing in endoplasmic reticulum

95	comp113165_c0	PRH75; DEAD/DEAH box RNA helicase PRH75	9.75	Biosynthesis of secondary metabolites
96	comp111963_c0	AVP1; Pyrophosphate-energized vacuolar membrane proton pump 1	9.74	Oxidative phosphorylation
97	comp102780_c0	ATPase, F1 complex, alpha subunit protein	9.73	Oxidative phosphorylation; Metabolic pathways
98	comp116054_c0	SIP2; putative galactinol--sucrose galactosyltransferase 2	9.69	Galactose metabolism
99	comp96448_c0	putative calcium-binding protein CML27	9.68	Plant-pathogen interaction
100	comp112691_c0	ARF1; auxin response factor 1	9.67	Plant hormone signal transduction
101	comp85153_c0	FT; protein FLOWERING LOCUS T	9.66	Circadian rhythm - plant
102	comp119399_c0	ribosomal protein L25/Gln-tRNA synthetase	9.65	Ribosome
103	comp96613_c0	SYP124; syntaxin-124	9.62	SNARE interactions in vesicular transport
104	comp109574_c0	MAB1; pyruvate dehydrogenase E1 beta	9.62	Glycolysis / Gluconeogenesis; Citrate cycle (TCA cycle); Pyruvate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites;

				Biosynthesis of antibiotics; Carbon metabolism
105	comp110804_c0	adenosine deaminase-like protein	9.59	Purine metabolism; Metabolic pathways
106	comp137165_c0	Beta-glucosidase, GBA2 type family protein	9.58	Other glycan degradation; Sphingolipid metabolism; Metabolic pathways
107	comp116011_c0	inositol-tetrakisphosphate 1-kinase 2	9.55	Inositol phosphate metabolism; Metabolic pathways; Phosphatidylinositol signaling system
108	comp114089_c0	E3 ubiquitin-protein ligase SINAT3	9.53	Ubiquitin mediated proteolysis
109	comp108027_c0	pre-mRNA-splicing factor 38A	9.53	Spliceosome
110	comp102333_c0	RD19; cysteine proteinase RD19a	9.50	Plant-pathogen interaction
111	comp114092_c0	Aldolase-type TIM barrel family protein	9.49	Pentose phosphate pathway; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
112	comp98603_c0	CPK6; calcium dependent protein kinase 6	9.48	Plant-pathogen interaction
113	comp118621_c0	ribosomal protein L22p/L17e	9.48	Ribosome

		family protein		
114	comp123397_c0	RNR1; ribonucleoside-diphosphate reductase large subunit	9.48	Purine metabolism; Pyrimidine metabolism; Glutathione metabolism; Metabolic pathways
115	comp117156_c0	SNRK2.10; SNF1-related protein kinase 2.10	9.47	Plant hormone signal transduction
116	comp114637_c0	DAD2; Defender against cell death 2	9.46	N-Glycan biosynthesis; Metabolic pathways; Protein processing in endoplasmic reticulum
117	comp109061_c0	SDH2-1; succinate dehydrogenase [ubiquinone] iron-sulfur subunit 1	9.46	Citrate cycle (TCA cycle); Oxidative phosphorylation; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
118	comp99989_c0	4CL3; 4-coumarate--CoA ligase 3	9.45	Ubiquinone and other terpenoid-quinone biosynthesis; Phenylalanine metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
119	comp112284_c0	LAX3; auxin influx carrier LAX3	9.43	Plant hormone signal transduction
120	comp107577_c0	40S ribosomal protein S17-4	9.42	Ribosome
121	comp106000_c0	PWWP domain-containing	9.40	Cysteine and methionine metabolism; Metabolic

		protein		pathways
122	comp116445_c0	MEE32; bi-functional dehydroquinate-shikimate dehydrogenase	9.40	Phenylalanine, tyrosine and tryptophan biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Biosynthesis of amino acids
123	comp113950_c0	DGD2; digalactosyldiacylglycerol synthase 2	9.38	Glycerolipid metabolism; Metabolic pathways
124	comp102871_c0	Glycosyl hydrolase family protein	9.37	Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways
125	comp102415_c0	OB-fold nucleic acid binding domain-containing protein	9.37	DNA replication; Nucleotide excision repair; Mismatch repair; Homologous recombination
126	comp93596_c0	MSH3; DNA mismatch repair protein MSH3	9.36	Mismatch repair
127	comp123838_c0	RCI3; peroxidase 3	9.36	Phenylalanine metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
128	comp106393_c0	GLN1;4; glutamine synthetase 1;4	9.34	Arginine biosynthesis; Alanine, aspartate and glutamate metabolism; Glyoxylate and

				dicarboxylate metabolism; Nitrogen metabolism; Metabolic pathways; Biosynthesis of amino acids
129	comp106652_c0	ARFA1B; ADP-ribosylation factor A1B	9.33	Endocytosis
130	comp98858_c0	ATCES1; acyl-CoA independent ceramide synthase	9.32	Sphingolipid metabolism
131	comp109473_c0	PKP-ALPHA; plastidial pyruvate kinase 1	9.31	Glycolysis / Gluconeogenesis; Purine metabolism; Pyruvate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
132	comp125760_c0	RHM1; UDP-L-rhamnose synthase	9.28	Amino sugar and nucleotide sugar metabolism
133	comp129724_c0	ceramide kinase	9.25	Sphingolipid metabolism
134	comp88797_c0	nad1; NADH dehydrogenase subunit 1	9.23	Oxidative phosphorylation; Metabolic pathways
135	comp85108_c0	RNA recognition motif-containing protein	9.21	mRNA surveillance pathway
136	comp106272_c0	acyl-CoA thioesterase family	9.21	Fatty acid elongation; Biosynthesis of unsaturated

		protein		fatty acids; Metabolic pathways; Biosynthesis of secondary metabolites
137	comp103047_c0	CDC48C; cell division cycle 48C protein	9.20	Ribosome biogenesis in eukaryotes
138	comp114195_c0	putative NADH-ubiquinone oxidoreductase B18 subunit	9.20	Oxidative phosphorylation; Metabolic pathways
139	comp68588_c0	60S ribosomal protein L11-2	9.18	Ribosome
140	comp93017_c0	PSAT; phosphoserine aminotransferase	9.17	Glycine, serine and threonine metabolism; Vitamin B6 metabolism; Metabolic pathways; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
141	comp72237_c0	putative 3-methyladenine glycosylase I	9.16	Base excision repair
142	comp86773_c0	FUC1; alpha-L-fucosidase 1	9.15	Other glycan degradation
143	comp89296_c0	SYP121; syntaxin-121	9.14	SNARE interactions in vesicular transport
144	comp113678_c1	GSH2; glutathione synthetase 2	9.14	Glutathione metabolism; Metabolic pathways
145	comp92391_c0	Hsp81.4; Heat shock protein 90-4	9.13	Protein processing in endoplasmic reticulum; Plant-pathogen interaction
146	comp108790_c0	RPL34; 60S ribosomal protein	9.12	Ribosome

		L34-2		
147	comp97238_c0	putative UDP-glucose 6-dehydrogenase 1	9.12	Pentose and glucuronate interconversions; Ascorbate and aldarate metabolism; Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways
148	comp96338_c0	APT3; adenine phosphoribosyl transferase 3	9.07	Purine metabolism; Metabolic pathways
149	comp96159_c0	PKT3; 3-ketoacyl-CoA thiolase 2	9.06	Fatty acid degradation; Valine, leucine and isoleucine degradation; alpha-Linolenic acid metabolism; Biosynthesis of unsaturated fatty acids; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Fatty acid metabolism; Peroxisome
150	comp112489_c0	GI; protein GIGANTEA	9.06	Circadian rhythm - plant
151	comp94083_c0	40S ribosomal protein S3a-1	9.04	Ribosome
152	comp118231_c0	ATP3; ATP synthase subunit gamma	9.04	Oxidative phosphorylation; Metabolic pathways
153	comp104081_c0	CYTC-2; cytochrome c-2	9.03	Sulfur metabolism; Metabolic pathways
154	comp112333_c0	ABF2; abscisic acid responsive elements-binding	9.02	Plant hormone signal transduction

		factor 2		
155	comp109914_c0	argininosuccinate synthase	9.01	Arginine biosynthesis; Alanine, aspartate and glutamate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Biosynthesis of amino acids
156	comp106393_c1	GLN1;4; glutamine synthetase 1;4	9.01	Arginine biosynthesis; Alanine, aspartate and glutamate metabolism; Glyoxylate and dicarboxylate metabolism; Nitrogen metabolism; Metabolic pathways; Biosynthesis of amino acids
157	comp106412_c0	Bet1-like protein	9.00	SNARE interactions in vesicular transport
158	comp93778_c0	Ras-related small GTP-binding family protein	9.00	Protein processing in endoplasmic reticulum
159	comp109235_c0	ARF1; auxin response factor 1	8.99	Plant hormone signal transduction
160	comp107748_c0	DWA1; WD repeat-containing protein DWA1	8.98	RNA transport
161	comp106319_c0	OB-fold nucleic acid binding domain-containing protein	8.96	DNA replication; Nucleotide excision repair; Mismatch repair; Homologous recombination
162	comp92383_c0	MGD2; monogalactosyldiacylglycerol	8.96	Glycerolipid metabolism; Metabolic pathways

		synthase 2		
163	comp108406_c0	Hsp81.4; Heat shock protein 90-4	8.95	Protein processing in endoplasmic reticulum; Plant-pathogen interaction
164	comp89640_c0	ABCB4; auxin efflux transmembrane transporter MDR4	8.95	ABC transporters
165	comp112617_c0	60S ribosomal protein L22-2	8.94	Ribosome
166	comp100932_c0	PBD1; 20S proteasome beta subunit D1	8.94	Proteasome
167	comp100775_c0	transcription factor IIA, alpha/beta subunit	8.92	Basal transcription factors
168	comp94857_c0	EIN2; protein ETHYLENE INSENSITIVE 2	8.92	Plant hormone signal transduction
169	comp106512_c0	40S ribosomal protein S16-1	8.91	Ribosome
170	comp111542_c0	glycyl-tRNA synthetase / glycine--tRNA ligase	8.90	Aminoacyl-tRNA biosynthesis
171	comp97941_c0	eukaryotic translation initiation factor 2 (eIF-2) family protein	8.90	RNA transport
172	comp103751_c0	ARR4; two-component	8.89	Plant hormone signal transduction

		response regulator ARR4		
173	comp48176_c0	KCS17; 3-ketoacyl-CoA synthase 17	8.89	Fatty acid elongation; Biosynthesis of secondary metabolites
174	comp124287_c0	PFK3; 6-phosphofructokinase 3	8.88	Glycolysis / Gluconeogenesis; Pentose phosphate pathway; Fructose and mannose metabolism; Galactose metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids; RNA degradation
175	comp105141_c0	aminomethyltransferase	8.86	Glycine, serine and threonine metabolism; Glyoxylate and dicarboxylate metabolism; One carbon pool by folate; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
176	comp111097_c0	60S ribosomal protein L37-3	8.86	Ribosome
177	comp101643_c0	40S ribosomal protein S29	8.86	Ribosome
178	comp116598_c0	FBA2; fructose-bisphosphate aldolase 2	8.85	Glycolysis / Gluconeogenesis; Pentose phosphate pathway; Fructose and mannose metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary

				metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
179	comp116540_c0	Hydroxymethylglutaryl-CoA lyase	8.83	Synthesis and degradation of ketone bodies; Valine, leucine and isoleucine degradation; Butanoate metabolism; Metabolic pathways; Peroxisome
180	comp128411_c0	CKX1; cytokinin dehydrogenase 1	8.83	Zeatin biosynthesis
181	comp119779_c0	G4; chlorophyll synthase	8.82	Porphyrin and chlorophyll metabolism; Metabolic pathways; Biosynthesis of secondary metabolites
182	comp99689_c0	40S ribosomal protein S28-1	8.81	Ribosome
183	comp111899_c0	MEKK1; mitogen-activated protein kinase kinase kinase 1	8.80	Plant-pathogen interaction
184	comp135532_c0	tryptophan synthase alpha chain	8.77	Glycine, serine and threonine metabolism; Phenylalanine, tyrosine and tryptophan biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of amino acids
185	comp113727_c0	ARR9; two-component response regulator ARR9	8.77	Plant hormone signal transduction

186	comp115959_c0	RPM1; disease resistance protein RPM1	8.75	Plant-pathogen interaction
187	comp95573_c0	HA2; H(+)-ATPase 2	8.74	Oxidative phosphorylation
188	comp125342_c0	ESP3; DEAH RNA helicase homolog PRP2	8.73	Spliceosome
189	comp99216_c0	splicing factor PWI and RNA recognition motif-containing protein	8.73	Spliceosome
190	comp127633_c0	ATB'_ALPHA; protein phosphatase 2A B' alpha	8.71	mRNA surveillance pathway
191	comp103986_c0	EIF2_GAMMA; translation initiation factor eIF-2 gamma subunit	8.71	RNA transport
192	comp112807_c0	PRS2; ribose-phosphate pyrophosphokinase 2	8.70	Pentose phosphate pathway; Purine metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
193	comp119846_c0	mMDH1; malate dehydrogenase 1	8.68	Citrate cycle (TCA cycle); Cysteine and methionine metabolism; Pyruvate metabolism; Glyoxylate and dicarboxylate metabolism;

				Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
194	comp84903_c0	HSP20-like chaperone	8.60	Protein processing in endoplasmic reticulum
195	comp103745_c0	arginase	8.58	Arginine biosynthesis; Arginine and proline metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Biosynthesis of amino acids
196	comp98731_c0	emb2394; 50S ribosomal protein L6	8.33	Ribosome
197	comp107312_c0	HA5; H(+)-ATPase 5	8.30	Oxidative phosphorylation
198	comp93820_c0	putative splicing factor	8.22	Spliceosome
199	comp136910_c0	eukaryotic translation initiation factor 2 (eIF-2) family protein	8.15	RNA transport
200	comp114709_c0	NRPE5; DNA-directed RNA polymerase V subunit 5A	8.10	Purine metabolism; Pyrimidine metabolism; Metabolic pathways; RNA polymerase
201	comp107931_c0	NCED4; nine-cis-epoxycarotenoid dioxygenase	8.10	Carotenoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites

		4		
202	comp133130_c0	alpha-1,2-glucosyltransferase	7.93	N-Glycan biosynthesis; Metabolic pathways
203	comp129598_c0	EER4; transcription initiation factor TFIID subunit 12B	7.83	Basal transcription factors

Table 4. Up-regulated genes in leaf in cultivar SK

	Scaffold	Annotation	log2(leaf/rhizome)	Pathway
1	comp114089_c0	E3 ubiquitin-protein ligase SINAT3	14.81	Ubiquitin mediated proteolysis
2	comp103293_c0	FAB1B; phosphatidylinositol-3P 5-kinase-like	14.68	Inositol phosphate metabolism; Phosphatidylinositol signaling system; Phagosome
3	comp107761_c1	hypothetical protein	14.26	Taurine and hypotaurine metabolism; Metabolic pathways
4	comp107831_c0	protein MIDASIN1	13.76	Ribosome biogenesis in eukaryotes
5	comp108126_c0	argininosuccinate synthase	13.49	Arginine biosynthesis; Alanine, aspartate and glutamate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Biosynthesis of amino acids
6	comp135346_c0	SEC; putative UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase SEC	13.01	Other types of O-glycan biosynthesis

7	comp106512_c0	40S ribosomal protein S16-1	12.97	Ribosome
8	comp103986_c0	EIF2_GAMMA; translation initiation factor eIF-2 gamma subunit	12.95	RNA transport
9	comp110162_c0	Preprotein translocase Sec, Sec61-beta subunit protein	12.88	Protein export; Protein processing in endoplasmic reticulum; Phagosome
10	comp109807_c0	GLX1; glyoxalase I homolog GLX1	12.87	Pyruvate metabolism
11	comp113633_c0	GPX1; phospholipid hydroperoxide glutathione peroxidase 1	12.86	Glutathione metabolism; Arachidonic acid metabolism
12	comp118847_c0	UBC9; ubiquitin conjugating enzyme 9	12.82	Ubiquitin mediated proteolysis; Protein processing in endoplasmic reticulum
13	comp125387_c0	transport protein SEC13A	12.33	RNA transport; Protein processing in endoplasmic reticulum
14	comp116054_c0	SIP2; putative galactinol--sucrose galactosyltransferase 2	12.29	Galactose metabolism
15	comp93892_c0	SAG24; 60S ribosomal protein L10-3	12.28	Ribosome
16	comp71972_c0	peroxidase 52	12.16	Phenylalanine metabolism; Phenylpropanoid

				biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
17	comp118352_c0	CHIP; E3 ubiquitin-protein ligase CHIP	12.13	Ubiquitin mediated proteolysis; Protein processing in endoplasmic reticulum
18	comp127244_c0	AK-LYS1; aspartokinase 1	12.05	Glycine, serine and threonine metabolism; Monobactam biosynthesis; Cysteine and methionine metabolism; Lysine biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; 2-Oxocarboxylic acid metabolism; Biosynthesis of amino acids
19	comp113126_c0	APC10; anaphase-promoting complex subunit 10	12.03	Ubiquitin mediated proteolysis
20	comp120130_c0	ACX3; acyl-coenzyme A oxidase 3	11.88	Fatty acid degradation; alpha-Linolenic acid metabolism; Biosynthesis of unsaturated fatty acids; Metabolic pathways; Biosynthesis of secondary metabolites; Fatty acid metabolism; Peroxisome
21	comp49043_c0	NADP-ME2; NADP-dependent malic enzyme 2	11.82	Pyruvate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways;

				Carbon metabolism
22	comp104585_c0	phosphoinositide phospholipase C 6	11.80	Inositol phosphate metabolism; Metabolic pathways; Phosphatidylinositol signaling system
23	comp112807_c0	PRS2; ribose-phosphate pyrophosphokinase 2	11.77	Pentose phosphate pathway; Purine metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
24	comp127578_c0	CSY3; citrate synthase 3	11.54	Citrate cycle (TCA cycle); Glyoxylate and dicarboxylate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; 2-Oxocarboxylic acid metabolism; Biosynthesis of amino acids
25	comp108466_c0	40S ribosomal protein S20-1	11.50	Ribosome
26	comp112811_c0	GPAT6; bifunctional sn-glycerol-3-phosphate 2-O-acyltransferase/phosphatase	11.36	Glycerolipid metabolism; Glycerophospholipid metabolism; Metabolic pathways; Biosynthesis of secondary metabolites
27	comp109914_c0	argininosuccinate synthase	11.27	Arginine biosynthesis; Alanine, aspartate and glutamate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites;

				Biosynthesis of antibiotics; Biosynthesis of amino acids
28	comp105413_c0	UBC1; ubiquitin-conjugating enzyme E2 1	11.11	Ubiquitin mediated proteolysis
29	comp117156_c0	SNRK2.10; SNF1-related protein kinase 2.10	11.10	Plant hormone signal transduction
30	comp135692_c0	NS1; asparaginyl-tRNA synthetase	11.10	Aminoacyl-tRNA biosynthesis
31	comp106412_c0	Bet1-like protein	10.75	SNARE interactions in vesicular transport
32	comp113913_c0	RPL24A; 60S ribosomal protein L24-1	10.73	Ribosome
33	comp92492_c0	Pre-mRNA-splicing factor 3	10.72	Spliceosome
34	comp101015_c0	GLT1; glutamate synthase 1 [NADH]	10.71	Alanine, aspartate and glutamate metabolism; Nitrogen metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Biosynthesis of amino acids
35	comp97730_c0	splicing factor PWI and RNA recognition motif-containing protein	10.68	Spliceosome

36	comp114709_c0	NRPE5; DNA-directed RNA polymerase V subunit 5A	10.63	Purine metabolism; Pyrimidine metabolism; Metabolic pathways; RNA polymerase
37	comp136174_c0	Pre-mRNA-splicing factor 3	10.62	Spliceosome
38	comp103847_c0	GA1; Ent-copalyl diphosphate synthase	10.59	Diterpenoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
39	comp109295_c0	TSB2; tryptophan synthase beta chain	10.59	Glycine, serine and threonine metabolism; Phenylalanine, tyrosine and tryptophan biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of amino acids
40	comp121146_c0	ribose 5-phosphate isomerase A	10.57	Pentose phosphate pathway; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
41	comp129598_c0	EER4; transcription initiation factor TFIID subunit 12B	10.57	Basal transcription factors
42	comp104622_c0	elongation factor 1-alpha 2	10.43	RNA transport
43	comp80805_c0	PPa3; pyrophosphorylase 3	10.37	Oxidative phosphorylation
44	comp104464_c0	CML42; calcium-binding	10.35	Plant-pathogen interaction

		protein CML42		
45	comp115865_c0	PCK1; phosphoenolpyruvate carboxykinase [ATP]	10.33	Glycolysis / Gluconeogenesis; Citrate cycle (TCA cycle); Pyruvate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
46	comp84669_c0	RANGAP2; RAN GTPase-activating protein 2	10.32	RNA transport
47	comp81670_c0	atpB; ATPase beta chain	10.28	Oxidative phosphorylation; Photosynthesis; Metabolic pathways
48	comp105717_c0	PPC1; phosphoenolpyruvate carboxylase 1	10.27	Pyruvate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Carbon metabolism
49	comp103913_c0	BS14A; BET1P/SFT1P-like protein 14A	10.26	SNARE interactions in vesicular transport
50	comp86305_c0	FKF1; flavin-binding, kelch repeat, f box 1	10.22	Circadian rhythm - plant
51	comp102333_c0	RD19; cysteine proteinase RD19a	10.21	Plant-pathogen interaction
52	comp133090_c0	GLYR1; gamma-	10.20	Glyoxylate and dicarboxylate metabolism;

		hydroxybutyrate dehydrogenase		Butanoate metabolism; Metabolic pathways; Carbon metabolism
53	comp111785_c0	60S ribosomal protein L38	10.17	Ribosome
54	comp136910_c0	eukaryotic translation initiation factor 2 (eIF-2) family protein	10.15	RNA transport
55	comp133351_c0	GPAT4; bifunctional sn-glycerol-3-phosphate 2-O-acyltransferase/phosphatase	10.14	Glycerolipid metabolism; Glycerophospholipid metabolism; Metabolic pathways; Biosynthesis of secondary metabolites
56	comp95804_c0	AAO3; abscisic-aldehyde oxidase	10.10	Carotenoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
57	comp113686_c0	CNX1; calnexin 1	10.09	Protein processing in endoplasmic reticulum; Phagosome
58	comp132067_c0	HSP60; heat shock protein 60	10.08	RNA degradation
59	comp92632_c0	cwINV4; beta-fructofuranosidase, insoluble isoenzyme CWINV4	9.99	Galactose metabolism; Starch and sucrose metabolism; Metabolic pathways
60	comp105591_c0	Asx tRNA synthetase (AspRS/AsnRS) class II core domain-contating protein	9.98	Aminoacyl-tRNA biosynthesis

61	comp94366_c0	lipoyl synthase	9.92	Lipoic acid metabolism; Metabolic pathways
62	comp112112_c0	SHD; HSP90-like protein GRP94	9.90	Protein processing in endoplasmic reticulum; Plant-pathogen interaction
63	comp132000_c0	GSTL2; glutathione transferase lambda 2	9.90	Glutathione metabolism
64	comp107493_c0	CML37; calcium-binding protein CML37	9.89	Plant-pathogen interaction
65	comp123669_c0	STS; stachyose synthase	9.88	Galactose metabolism
66	comp107134_c0	SDN3; small RNA degrading nuclease 3	9.86	Ribosome biogenesis in eukaryotes
67	comp120381_c0	small nuclear ribonucleoprotein D2	9.86	Spliceosome
68	comp115543_c0	60S ribosomal protein L32-1	9.85	Ribosome
69	comp97285_c0	SNRK2.3; serine/threonine-protein kinase SRK2I	9.85	Plant hormone signal transduction
70	comp97949_c0	NAP57; putative pseudouridine synthase NAP57	9.84	Ribosome biogenesis in eukaryotes
71	comp108187_c0	40S ribosomal protein S18	9.83	Ribosome
72	comp131650_c0	GAE6; UDP-D-glucuronate 4-	9.81	Starch and sucrose metabolism; Amino sugar and

		epimerase 6		nucleotide sugar metabolism; Metabolic pathways
73	comp123838_c0	RCI3; peroxidase 3	9.80	Phenylalanine metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
74	comp118078_c0	ATP3; ATP synthase subunit gamma	9.76	Oxidative phosphorylation; Metabolic pathways
75	comp132448_c0	CAM3; calmodulin 3	9.76	Phosphatidylinositol signaling system; Plant-pathogen interaction
76	comp116258_c0	oligosaccharyltransferase complex/magnesium transporter family protein	9.75	N-Glycan biosynthesis; Metabolic pathways; Protein processing in endoplasmic reticulum
77	comp107931_c0	NCED4; nine-cis-epoxycarotenoid dioxygenase 4	9.73	Carotenoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
78	comp99451_c0	PBD1; 20S proteasome beta subunit D1	9.71	Proteasome
79	comp93820_c0	putative splicing factor	9.69	Spliceosome
80	comp80882_c0	HSP60; heat shock protein 60	9.69	RNA degradation
81	comp123414_c0	SNRK2.5; serine/threonine-	9.65	Plant hormone signal transduction

		protein kinase SRK2H		
82	comp121185_c0	DNAse I-like superfamily protein	9.65	RNA degradation
83	comp94302_c0	RNA recognition motif-containing protein	9.64	mRNA surveillance pathway
84	comp104123_c0	SYP71; syntaxin-71	9.64	SNARE interactions in vesicular transport
85	comp130078_c0	ATERDJ3B; DNAJ heat shock protein ATERDJ3B	9.62	Protein processing in endoplasmic reticulum
86	comp73304_c0	nad2; NADH dehydrogenase subunit 2 (EC:1.6.99.3)	9.60	Oxidative phosphorylation; Metabolic pathways
87	comp93017_c0	PSAT; phosphoserine aminotransferase	9.58	Glycine, serine and threonine metabolism; Vitamin B6 metabolism; Metabolic pathways; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
88	comp97601_c0	G6PD1; glucose-6-phosphate dehydrogenase 1	9.58	Pentose phosphate pathway; Glutathione metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
89	comp100334_c0	EMB2770; pre-mRNA-processing factor 6-like	9.53	Spliceosome

		protein STA1		
90	comp119846_c0	mMDH1; malate dehydrogenase 1	9.51	Citrate cycle (TCA cycle); Cysteine and methionine metabolism; Pyruvate metabolism; Glyoxylate and dicarboxylate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
91	comp121552_c0	ADK2; adenosine kinase 2	9.50	Purine metabolism; Metabolic pathways
92	comp119736_c0	uridylate kinase-like protein	9.50	Pyrimidine metabolism; Metabolic pathways
93	comp112601_c0	polynucleotidyl transferase, ribonuclease H-like superfamily protein	9.49	RNA degradation
94	comp123097_c0	UGP2; UTP--glucose-1-phosphate uridylyltransferase 1	9.48	Pentose and glucuronate interconversions; Galactose metabolism; Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways; Biosynthesis of antibiotics
95	comp122600_c0	RAD23C; UV excision repair protein RAD23C	9.48	Nucleotide excision repair; Protein processing in endoplasmic reticulum

96	comp102415_c0	OB-fold nucleic acid binding domain-containing protein	9.48	DNA replication; Nucleotide excision repair; Mismatch repair; Homologous recombination
97	comp126789_c0	SYCO_ARATH; cysteinyl-tRNA synthetase	9.48	Aminoacyl-tRNA biosynthesis
98	comp126756_c0	AXS2; UDP-D-apiose/UDP-D-xylose synthase 2	9.47	Amino sugar and nucleotide sugar metabolism; Metabolic pathways
99	comp124060_c0	SKP2A; F-box protein SKP2A	9.46	Ubiquitin mediated proteolysis
100	comp107097_c0	CAM3; calmodulin 3	9.46	Phosphatidylinositol signaling system; Plant-pathogen interaction
101	comp129041_c0	putative glycosylphosphatidylinositol-anchor transamidase	9.45	Glycosylphosphatidylinositol(GPI)-anchor biosynthesis; Metabolic pathways
102	comp109155_c0	RNA polymerase Rpb7 N-terminal domain-containing protein	9.43	Purine metabolism; Pyrimidine metabolism; Metabolic pathways; RNA polymerase
103	comp104414_c0	HEMA1; glutamyl-tRNA reductase 1	9.42	Porphyrin and chlorophyll metabolism; Metabolic pathways; Biosynthesis of secondary metabolites
104	comp133130_c0	alpha-1,2-glucosyltransferase	9.42	N-Glycan biosynthesis; Metabolic pathways
105	comp97444_c0	LCB1; serine palmitoyltransferase	9.40	Sphingolipid metabolism; Metabolic pathways

106	comp105625_c0	PETC; cytochrome b6-f complex iron-sulfur subunit	9.38	Photosynthesis; Metabolic pathways
107	comp84608_c0	protein translation factor SUI1-2	9.37	RNA transport
108	comp84903_c0	HSP20-like chaperone	9.37	Protein processing in endoplasmic reticulum
109	comp108016_c0	RHD2; Respiratory burst oxidase homolog protein C	9.35	Plant-pathogen interaction
110	comp124714_c0	IBR1; short-chain dehydrogenase/reductase A	9.34	Metabolic pathways; Peroxisome
111	comp111429_c0	4CL1; 4-coumarate--CoA ligase 1	9.31	Ubiquinone and other terpenoid-quinone biosynthesis; Phenylalanine metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
112	comp101440_c0	UBC11; ubiquitin-conjugating enzyme 11	9.27	Ubiquitin mediated proteolysis; Protein processing in endoplasmic reticulum
113	comp107312_c0	HA5; H(+)-ATPase 5	9.23	Oxidative phosphorylation
114	comp103206_c0	protein transport protein SEC61 subunit beta	9.23	Protein export; Protein processing in endoplasmic reticulum; Phagosome
115	comp115039_c0	PRS2; ribose-phosphate pyrophosphokinase 2	9.23	Pentose phosphate pathway; Purine metabolism; Metabolic pathways; Biosynthesis of secondary

				metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
116	comp110408_c0	VPS60.2; vacuolar protein sorting protein 60.2	9.21	Endocytosis
117	comp96610_c0	C4H; trans-cinnamate 4-monooxygenase	9.20	Ubiquinone and other terpenoid-quinone biosynthesis; Phenylalanine metabolism; Phenylpropanoid biosynthesis; Flavonoid biosynthesis; Stilbenoid, diarylheptanoid and gingerol biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Degradation of aromatic compounds
118	comp108027_c0	pre-mRNA-splicing factor 38A	9.19	Spliceosome
119	comp82559_c0	endonuclease/exonuclease/phosphatase family protein	9.18	Base excision repair
120	comp98194_c0	KDTA; 3-deoxy-D-manno-octulosonic-acid transferase	9.17	Metabolic pathways
121	comp121161_c0	UBA2A; UBP1-associated protein 2A	9.16	Spliceosome
122	comp91912_c0	XRN3; 5'-3' exoribonuclease 3	9.13	Ribosome biogenesis in eukaryotes; RNA

				degradation
123	comp109503_c0	F-type H+-transporting ATPase subunit b	9.13	Oxidative phosphorylation; Photosynthesis; Metabolic pathways
124	comp97287_c0	Glycosyl hydrolase family protein	9.12	Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways
125	comp133435_c0	xylose isomerase	9.10	Pentose and glucuronate interconversions; Fructose and mannose metabolism; Metabolic pathways
126	comp102927_c0	EIF3C; eukaryotic translation initiation factor 3 subunit C	9.09	RNA transport
127	comp124630_c0	ARM repeat superfamily protein	9.08	Protein processing in endoplasmic reticulum
128	comp108991_c0	BIP1; Luminal-binding protein 1	9.07	Protein export; Protein processing in endoplasmic reticulum

Table 5. Up-regulated genes in rhizome of 4 cultivars

Seq_ID	AR	LN	OU	SK	Protein	Pathway
comp121185_c0	10.81	10.49	10.97	9.65	DNase I-like superfamily protein	RNA degradation
comp120130_c0	10.87	10.85	11.25	11.88	ACX3; acyl-coenzyme A oxidase 3	Fatty acid degradation; alpha-Linolenic acid metabolism; Biosynthesis of unsaturated fatty acids; Metabolic pathways; Biosynthesis of secondary metabolites; Fatty acid metabolism; Peroxisome
comp103293_c0	11.89	11.52	12.2	14.68	FAB1B; phosphatidylinositol-3P 5-kinase-like	Inositol phosphate metabolism; Phosphatidylinositol signaling system; Phagosome
comp103986_c0	12.33	13.11	8.71	12.95	EIF2_GAMMA; translation initiation factor eIF-2 gamma subunit	RNA transport
comp49043_c0	12.49	12.21	12.2	11.82	NADP-ME2; NADP-dependent malic enzyme 2	Pyruvate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways;

						Carbon metabolism
comp106512_c0	13.91	13.07	8.91	12.97	40S ribosomal protein S16-1	Ribosome
comp110408_c0	13.95	13.4	9.84	9.21	VPS60.2; vacuolar protein sorting protein 60.2	Endocytosis
comp102333_c0	14.12	12.24	9.5	10.21	RD19; cysteine proteinase RD19a	Plant-pathogen interaction
comp107761_c1	14.13	13.49	13.56	14.26	hypothetical protein	Taurine and hypotaurine metabolism; Metabolic pathways
comp113126_c0	14.71	14.34	15.01	12.03	APC10; a phase-promoting complex subunit 10	Ubiquitin mediated proteolysis
comp114089_c0	15.28	15	9.53	14.81	E3 ubiquitin-protein ligase SINAT3	Ubiquitin mediated proteolysis
comp110162_c0	15.84	15.6	15.94	12.88	Preprotein translocase Sec, Sec61-beta subunit protein	Protein export; Protein processing in endoplasmic reticulum; Phagosome
comp107831_c0	15.96	15.25	16.44	13.76	protein MIDASIN1	Ribosome biogenesis in eukaryotes
comp106412_c0		9.73	9	10.75	Bet1-like protein	SNARE interactions in vesicular transport
comp123838		9.76	9.36	9.8	RCI3; peroxidase 3	Phenylpropanoid biosynthesis;

_c0						Metabolic pathways; Biosynthesis of secondary metabolites
comp82559_c0		9.95	10.67	9.18	endonuclease/exonuclease/phosphatase family protein	Base excision repair
comp105413_c0		13.06	11.79	11.11	UBC1; ubiquitin-conjugating enzyme E2 1	Ubiquitin mediated proteolysis
comp117156_c0	10.6		9.47	11.1	SNRK2.10; SNF1-related protein kinase 2.10	Plant hormone signal transduction
comp133351_c0	10.74		10.39	10.14	GPAT4; bifunctional sn-glycerol-3-phosphate 2-O-acyltransferase/phosphatase	Glycerolipid metabolism; Glycerophospholipid metabolism; Metabolic pathways; Biosynthesis of secondary metabolites
comp109295_c0	10.98		10.32	10.59	TSB2; tryptophan synthase beta chain	Glycine, serine and threonine metabolism; Phenylalanine, tyrosine and tryptophan biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of amino acids
comp109914_c0			9.01	11.27	argininosuccinate synthase	Arginine biosynthesis; Alanine, aspartate and glutamate metabolism;

				Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Biosynthesis of amino acids
comp93017_c0	9.17	9.58	PSAT; phosphoserine aminotransferase	Glycine, serine and threonine metabolism; Vitamin B6 metabolism; Metabolic pathways; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
comp102415_c0	9.37	9.48	OB-fold nucleic acid binding domain-containing protein	DNA replication; Nucleotide excision repair; Mismatch repair; Homologous recombination
comp108027_c0	9.53	9.19	pre-mRNA-splicing factor 38A	Spliceosome
comp116054_c0	9.69	12.29	SIP2; putative galactinol--sucrose galactosyltransferase 2	Galactose metabolism
comp104123_c0	9.84	9.64	SYP71; syntaxin-71	SNARE interactions in vesicular transport
comp97730_	9.85	10.68	splicing factor PWI and RNA recognition	Spliceosome

c0			motif-containing protein	
comp112811_c0	10.54	11.36	GPAT6; bifunctional sn-glycerol-3-phosphate 2-O-acyltransferase/phosphatase	Glycerolipid metabolism; Glycerophospholipid metabolism; Metabolic pathways; Biosynthesis of secondary metabolites
comp116258_c0	10.62	9.75	oligosaccharyltransferase complex/magnesium transporter family protein	N-Glycan biosynthesis; Metabolic pathways; Protein processing in endoplasmic reticulum
comp107097_c0	10.99	9.46	CAM3; calmodulin 3	Phosphatidylinositol signaling system; Plant-pathogen interaction
comp92492_c0	11.03	10.72	Pre-mRNA-splicing factor 3	Spliceosome
comp130078_c0	11.21	9.62	ATERDJ3B; DnaJ heat shock protein ATERDJ3B	Protein processing in endoplasmic reticulum
comp93892_c0	11.56	12.28	SAG24; 60S ribosomal protein L10-3	Ribosome
comp103847_c0	11.74	10.59	GA1; Ent-copalyl diphosphate synthase	Diterpenoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp118352	12.58	12.13	CHIP; E3 ubiquitin-protein ligase	Ubiquitin mediated proteolysis;

_c0					CHIP	Protein processing in endoplasmic reticulum
comp119846_c0	10.75	10.09		9.51	mMDH1; malate dehydroge se 1	Citrate cycle (TCA cycle); Cysteine and methionine metabolism; Pyruvate metabolism; Glyoxylate and dicarboxylate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
comp129598_c0	10.88	12.83		10.57	EER4; transcription initiation factor TFIID subunit 12B	Basal transcription factors
comp118847_c0	12.17	12.06		12.82	UBC9; ubiquitin conjugating enzyme 9	Ubiquitin mediated proteolysis; Protein processing in endoplasmic reticulum
comp99451_c0		9.37		9.71	PBD1; 20S proteasome beta subunit D1	Proteasome
comp107312_c0		9.45		9.23	HA5; H(+)-ATPase 5	Oxidative phosphorylation

comp121146_c0	9.55	10.57	ribose 5-phosphate isomerase A	Pentose phosphate pathway; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
comp108466_c0	9.56	11.5	40S ribosomal protein S20-1	Ribosome
comp97287_c0	9.69	9.12	Glycosyl hydrolase family protein	Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways
comp123097_c0	9.78	9.48	UGP2; UTP--glucose-1-phosphate uridylyltransferase 1	Pentose and glucuro te interconversions; Galactose metabolism; Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways; Biosynthesis of antibiotics
comp132000	10.16	9.9	GSTL2; glutathione transferase	Glutathione metabolism

_c0			lambda 2	
comp105625_c0	10.32	9.38	PETC; cytochrome b6-f complex iron-sulfur subunit	Photosynthesis; Metabolic pathways
comp113633_c0	10.39	12.86	GPX1; phospholipid hydroperoxide glutathione peroxidase 1	Glutathione metabolism; Arachidonic acid metabolism
comp124060_c0	10.41	9.46	SKP2A; F-box protein SKP2A	Ubiquitin mediated proteolysis
comp105717_c0	10.44	10.27	PPC1; phosphoenolpyruvate carboxylase 1	Pyruvate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Carbon metabolism
comp80805_c0	10.63	10.37	PPa3; pyrophosphorylase 3	Oxidative phosphorylation
comp107134_c0	10.67	9.86	SDN3; small R degrading nuclease 3	Ribosome biogenesis in eukaryotes
comp113686_c0	10.91	10.09	CNX1; calnexin 1	Protein processing in endoplasmic reticulum; Phagosome
comp132067_c0	11	10.08	HSP60; heat shock protein 60	R degradation
comp104585	11.06	11.8	phosphoinositide phospholipase C 6	Inositol phosphate metabolism;

_c0					Metabolic pathways; Phosphatidylinositol sig ling system
comp127244_c0		11.39	12.05	AK-LYS1; aspartoki se 1	Glycine, serine and threonine metabolism; Monobactam biosynthesis; Cysteine and methionine metabolism; Lysine biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; 2-Oxocarboxylic acid metabolism; Biosynthesis of amino acids
comp113913_c0		12.21	10.73	RPL24A; 60S ribosomal protein L24-1	Ribosome
comp125387_c0		12.44	12.33	transport protein SEC13A	R transport; Protein processing in endoplasmic reticulum
comp97444_c0	10.59		9.4	LCB1; serine palmitoyltransferase	Sphingolipid metabolism; Metabolic pathways
comp109807_c0	10.92		12.87	GLX1; glyoxalase I homolog GLX1	Pyruvate metabolism

comp118078_c0	11.24	9.76	ATP3; ATP synthase subunit gamma	Oxidative phosphorylation; Metabolic pathways
comp114709_c0	12.24	10.63	NRPE5; DNA-directed RNA polymerase V subunit 5A	Purine metabolism; Pyrimidine metabolism; Metabolic pathways; RNA polymerase
comp108991_c0		9.07	BIP1; Luminal-binding protein 1	Protein export; Protein processing in endoplasmic reticulum
comp124630_c0		9.08	ARM repeat superfamily protein	Protein processing in endoplasmic reticulum
comp102927_c0		9.09	EIF3C; eukaryotic translation initiation factor 3 subunit C	RNA transport
comp133435_c0		9.1	xylose isomerase	Pentose and glucuronate interconversions; Fructose and mannose metabolism; Metabolic pathways
comp109503_c0		9.13	F-type H+-transporting ATPase subunit b	Oxidative phosphorylation; Photosynthesis; Metabolic pathways
comp91912_c0		9.13	XRN3; 5'-3' exoribonuclease 3	Ribosome biogenesis in eukaryotes; RNA degradation
comp121161		9.16	UBA2A; UBP1-associated protein 2A	Spliceosome

_c0			
comp98194_c0	9.17	KDTA; 3-deoxy-D-manno-octulosonic-acid transferase	Metabolic pathways
comp96610_c0	9.2	C4H; trans-cinnamate 4-monooxygenase	Ubiquinone and other terpenoid-quinone biosynthesis; Phenylalanine metabolism; Phenylpropanoid biosynthesis; Flavonoid biosynthesis; Stilbenoid, diarylheptanoid and gingerol biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Degradation of aromatic compounds
comp103206_c0	9.23	protein transport protein SEC61 subunit beta	Protein export; Protein processing in endoplasmic reticulum; Phagosome
comp115039_c0	9.23	PRS2; ribose-phosphate pyrophosphokinase 2	Pentose phosphate pathway; Purine metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism;

			Biosynthesis of amino acids
comp101440_c0	9.27	UBC11; ubiquitin-conjugating enzyme 11	Ubiquitin mediated proteolysis; Protein processing in endoplasmic reticulum
comp111429_c0	9.31	4CL1; 4-coumarate--CoA ligase 1	Ubiquinone and other terpenoid-quinone biosynthesis; Phenylalanine metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp124714_c0	9.34	IBR1; short-chain dehydrogese/reductase A	Metabolic pathways; Peroxisome
comp108016_c0	9.35	RHD2; Respiratory burst oxidase homolog protein C	Plant-pathogen interaction
comp84903_c0	9.37	HSP20-like chaperone	Protein processing in endoplasmic reticulum
comp84608_c0	9.37	protein translation factor SUI1-2	R transport
comp104414_c0	9.42	HEMA1; glutamyl-tR reductase 1	Porphyrin and chlorophyll metabolism; Metabolic pathways;

			Biosynthesis of secondary metabolites
comp133130_c0	9.42	alpha-1,2-glucosyltransferase	N-Glycan biosynthesis; Metabolic pathways
comp109155_c0	9.43	RNA polymerase Rpb7 N-terminal domain-containing protein	Purine metabolism; Pyrimidine metabolism; Metabolic pathways; RNA polymerase
comp129041_c0	9.45	putative glycosylphosphatidylinositol-anchor transamidase	Glycosylphosphatidylinositol(GPI)-anchor biosynthesis; Metabolic pathways
comp126756_c0	9.47	AXS2; UDP-D-apiose/UDP-D-xylose synthase 2	Amino sugar and nucleotide sugar metabolism; Metabolic pathways
comp126789_c0	9.48	SYCO_ARATH; cysteinyl-tRNA synthetase	Aminoacyl-tRNA biosynthesis
comp122600_c0	9.48	RAD23C; UV excision repair protein RAD23C	Nucleotide excision repair; Protein processing in endoplasmic reticulum
comp112601_c0	9.49	polynucleotidyl transferase, ribonuclease H-like superfamily protein	RNA degradation

comp119736_c0	9.5	uridylate kinase-like protein	Pyrimidine metabolism; Metabolic pathways
comp121552_c0	9.5	ADK2; adenosine kinase 2	Purine metabolism; Metabolic pathways
comp100334_c0	9.53	EMB2770; pre-mRNA-processing factor 6-like protein STA1	Spliceosome
comp97601_c0	9.58	G6PD1; glucose-6-phosphate dehydrogenase 1	Pentose phosphate pathway; Glutathione metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
comp73304_c0	9.6	nad2; NADH dehydrogenase subunit 2 (EC:1.6.99.3)	Oxidative phosphorylation; Metabolic pathways
comp94302_c0	9.64	RNA recognition motif-containing protein	mRNA surveillance pathway
comp123414_c0	9.65	SNRK2.5; serine/threonine-protein kinase SRK2H	Plant hormone signal transduction
comp80882_c0	9.69	HSP60; heat shock protein 60	RNA degradation
comp93820_	9.69	putative splicing factor	Spliceosome

c0			
comp107931_c0	9.73	NCED4; nine-cis-epoxycarotenoid dioxygenase 4	Carotenoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp132448_c0	9.76	CAM3; calmodulin 3	Phosphatidylinositol signaling system; Plant-pathogen interaction
comp131650_c0	9.81	GAE6; UDP-D-glucuronate 4-epimerase 6	Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways
comp108187_c0	9.83	40S ribosomal protein S18	Ribosome
comp97949_c0	9.84	P57; putative pseudouridine synthase P57	Ribosome biogenesis in eukaryotes
comp97285_c0	9.85	SNRK2.3; serine/threonine-protein kinase SRK2I	Plant hormone signal transduction
comp115543_c0	9.85	60S ribosomal protein L32-1	Ribosome
comp120381_c0	9.86	small nuclear ribonucleoprotein D2	Spliceosome
comp123669	9.88	STS; stachyose synthase	Galactose metabolism

_c0			
comp107493_c0	9.89	CML37; calcium-binding protein CML37	Plant-pathogen interaction
comp112112_c0	9.9	SHD; HSP90-like protein GRP94	Protein processing in endoplasmic reticulum; Plant-pathogen interaction
comp94366_c0	9.92	lipoyl synthase	Lipoic acid metabolism; Metabolic pathways
comp105591_c0	9.98	Asx tR synthetase (AspRS/AsnRS) class II core domain-contating protein	Aminoacyl-tR biosynthesis
comp92632_c0	9.99	cwINV4; beta-fructofuranosidase, insoluble isoenzyme CWINV4	Galactose metabolism; Starch and sucrose metabolism; Metabolic pathways
comp95804_c0	10.1	AAO3; abscisic-aldehyde oxidase	Carotenoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp136910_c0	10.15	eukaryotic translation initiation factor 2 (eIF-2) family protein	R transport
comp111785_c0	10.17	60S ribosomal protein L38	Ribosome

comp133090_c0	10.2	GLYR1; gamma-hydroxybutyrate dehydroge se	Glyoxylate and dicarboxylate metabolism; Butanoate metabolism; Metabolic pathways; Carbon metabolism
comp86305_c0	10.22	FKF1; flavin-binding, kelch repeat, f box 1	Circadian rhythm - plant
comp103913_c0	10.26	BS14A; BET1P/SFT1P-like protein 14A	S RE interactions in vesicular transport
comp81670_c0	10.28	atpB; ATPase beta chain	Oxidative phosphorylation; Photosynthesis; Metabolic pathways
comp84669_c0	10.32	RANGAP2; RAN GTPase-activating protein 2	R transport
comp115865_c0	10.33	PCK1; phosphoenolpyruvate carboxyki se [ATP]	Glycolysis / Gluconeogenesis; Citrate cycle (TCA cycle); Pyruvate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism

comp104464_c0	10.35	CML42; calcium-binding protein CML42	Plant-pathogen interaction
comp104622_c0	10.43	elongation factor 1-alpha 2	R transport
comp136174_c0	10.62	Pre-mR -splicing factor 3	Spliceosome
comp101015_c0	10.71	GLT1; glutamate synthase 1 [DH]	Alanine, aspartate and glutamate metabolism; Nitrogen metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Biosynthesis of amino acids
comp135692_c0	11.1	NS1; asparaginyl-tR synthetase	Aminoacyl-tR biosynthesis
comp127578_c0	11.54	CSY3; citrate synthase 3	Citrate cycle (TCA cycle); Glyoxylate and dicarboxylate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; 2-

			Oxocarboxylic acid metabolism; Biosynthesis of amino acids
comp112807_c0	11.77	PRS2; ribose-phosphate pyrophosphoki se 2	Pentose phosphate pathway; Purine metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
comp71972_c0	12.16	peroxidase 52	Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp135346_c0	13.01	SEC; putative UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase SEC	Other types of O-glycan biosynthesis
comp108126_c0	13.49	argininosucci te synthase	Arginine biosynthesis; Alanine, aspartate and glutamate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Biosynthesis of amino acids

comp116408_c0	10.9	10.49	10.2	PHYB; phytochrome B	Circadian rhythm - plant
comp135677_c0	11.09	11.27	10.44	putative glucosamine-fructose-6-phosphate aminotransferase	Alanine, aspartate and glutamate metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways; Biosynthesis of antibiotics
comp137165_c0	11.42	11.55	9.58	Beta-glucosidase, GBA2 type family protein	Other glycan degradation; Sphingolipid metabolism; Metabolic pathways
comp109061_c0	11.64	11.96	9.46	SDH2-1; succi te dehydroge se [ubiquinone] iron-sulfur subunit 1	Citrate cycle (TCA cycle); Oxidative phosphorylation; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
comp122860_c0	12.32	12.06	11.21	palmitoyl protein thioesterase family protein	Fatty acid elongation; Metabolic pathways; Fatty acid metabolism
comp123248_c0	12.99	13.14	10.37	R recognition motif-containing protein	Spliceosome

comp114024_c0	13.52	13.32	12.61	transport protein SEC13A	RNA transport; Protein processing in endoplasmic reticulum
comp111085_c0	15.37	10.68	12.84	RPL27; 50S ribosomal protein L27	Ribosome
comp112617_c0		9.37	8.94	60S ribosomal protein L22-2	Ribosome
comp118231_c0		9.61	9.04	ATP3; ATP synthase subunit gamma	Oxidative phosphorylation; Metabolic pathways
comp124038_c0		9.62	10.92	CYL1; alpha-N-acetylglucosaminidase	Glycosaminoglycan degradation; Metabolic pathways
comp128411_c0		9.7	8.83	CKX1; cytokinin dehydrogenase 1	Zeatin biosynthesis
comp130572_c0		9.83	10.25	EIF4A-III; DEAD-box ATP-dependent RNA helicase 2	RNA transport; mRNA surveillance pathway; Spliceosome
comp112489_c0		9.85	9.06	GI; protein GIGANTEA	Circadian rhythm - plant
comp95524_c0		9.85	10.7	RPM1; disease resistance protein RPM1	Plant-pathogen interaction
comp113241_c0		9.88	10.18	ATHAL3B; putative phosphopantothenoylcysteine	Pantothenate and CoA biosynthesis; Metabolic pathways

			decarboxylase	
comp113727_c0	10.02	8.77	ARR9; two-component response regulator ARR9	Plant hormone signal transduction
comp136154_c0	10.06	10.23	hypothetical protein	RNA degradation
comp96613_c0	10.11	9.62	SYP124; syntaxin-124	SNARE interactions in vesicular transport
comp103751_c0	10.15	8.89	ARR4; two-component response regulator ARR4	Plant hormone signal transduction
comp91567_c0	10.32	10.37	ACO1; aconitate hydratase 1	Citrate cycle (TCA cycle); Glyoxylate and dicarboxylate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; 2-Oxocarboxylic acid metabolism; Biosynthesis of amino acids
comp94473_c0	10.44	10.72	eukaryotic translation initiation factor 3 subunit 7	RNA transport
comp106000	10.47	9.4	PWWP domain-containing protein	Cysteine and methionine

_c0				metabolism; Metabolic pathways
comp103975_c0	10.71	10.6	CAM7; calmodulin 7	Phosphatidylinositol signaling system; Plant-pathogen interaction
comp106272_c0	10.86	9.21	acyl-CoA thioesterase family protein	Fatty acid elongation; Biosynthesis of unsaturated fatty acids; Metabolic pathways; Biosynthesis of secondary metabolites
comp85153_c0	11.06	9.66	FT; protein FLOWERING LOCUS T	Circadian rhythm - plant
comp130472_c0	11.17	11.03	glutamine-dependent D(+) synthetase	Nicotinate and nicotinamide metabolism; Metabolic pathways
comp131748_c0	11.19	10.36	ALDH11A3; NADP-dependent glyceraldehyde-3-phosphate dehydrogenase	Glycolysis / Gluconeogenesis; Pentose phosphate pathway; Metabolic pathways; Carbon metabolism
comp130901_c0	11.93	10.6	ELC-Like; ELC-like protein	Endocytosis
comp113826_c0	12.02	12.17	ubiquinol-cytochrome C reductase hinge protein	Oxidative phosphorylation; Metabolic pathways
comp99535_	12.08	12.03	60S ribosomal protein L15-1	Ribosome

c0					
comp96770_c0		13.32	12.21	BGLU3; beta-glucosidase 3	Cyanoamino acid metabolism; Starch and sucrose metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp125342_c0	10.69		8.73	ESP3; DEAH RNA helicase homolog PRP2	Spliceosome
comp111876_c0	10.82		11.65	CYCD3;2; cyclin-D3-2	Plant hormone signal transduction
comp132819_c0	11.11		11.03	TWN2; Valyl-tRNA synthetase	Aminoacyl-tRNA biosynthesis
comp131917_c0	11.53		11.2	NCED4; nine-cis-epoxycarotenoid dioxygenase 4	Carotenoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp124809_c1	11.95		11.58	BCAT3; branched-chain-amino-acid aminotransferase 3	Valine, leucine and isoleucine degradation; Valine, leucine and isoleucine biosynthesis; Pantothenate and CoA biosynthesis; Glucosinolate biosynthesis;

				Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; 2-Oxocarboxylic acid metabolism; Biosynthesis of amino acids
comp134068_c0	12.03	10.2	dehydratase-enolase-phosphatase complex 1	Cysteine and methionine metabolism; Metabolic pathways
comp121776_c0	12.56	11.27	ZTL; adagio protein 1	Circadian rhythm - plant
comp127633_c0		8.71	ATB'_ALPHA; protein phosphatase 2A B' alpha	mR surveillance pathway
comp99216_c0		8.73	splicing factor PWI and R recognition motif-containing protein	Spliceosome
comp95573_c0		8.74	HA2; H(+)-ATPase 2	Oxidative phosphorylation
comp115959_c0		8.75	RPM1; disease resistance protein RPM1	Plant-pathogen interaction
comp135532_c0		8.77	tryptophan synthase alpha chain	Glycine, serine and threonine metabolism; Phenylalanine, tyrosine and tryptophan biosynthesis;

			Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of amino acids
comp111899_c0	8.8	MEKK1; mitogen-activated protein ki se ki se ki se 1	Plant-pathogen interaction
comp99689_c0	8.81	40S ribosomal protein S28-1	Ribosome
comp119779_c0	8.82	G4; chlorophyll synthase	Porphyrin and chlorophyll metabolism; Metabolic pathways; Biosynthesis of secondary metabolites
comp116540_c0	8.83	Hydroxymethylglutaryl-CoA lyase	Synthesis and degradation of ketone bodies; Valine, leucine and isoleucine degradation; Butanoate metabolism; Metabolic pathways; Peroxisome
comp116598_c0	8.85	FBA2; fructose-bisphosphate aldolase 2	Glycolysis / Gluconeogenesis; Pentose phosphate pathway; Fructose and mannose metabolism; Carbon fixation in photosynthetic

			organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
comp101643 _c0	8.86	40S ribosomal protein S29	Ribosome
comp111097 _c0	8.86	60S ribosomal protein L37-3	Ribosome
comp105141 _c0	8.86	aminomethyltransferase	Glycine, serine and threonine metabolism; Glyoxylate and dicarboxylate metabolism; One carbon pool by folate; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
comp124287 _c0	8.88	PFK3; 6-phosphofructoki se 3	Glycolysis / Gluconeogenesis; Pentose phosphate pathway; Fructose and mannose metabolism; Galactose metabolism; Metabolic

			pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids; R degradation
comp48176_c0	8.89	KCS17; 3-ketoacyl-CoA synthase 17	Fatty acid elongation; Biosynthesis of secondary metabolites
comp97941_c0	8.9	eukaryotic translation initiation factor 2 (eIF-2) family protein	R transport
comp111542_c0	8.9	glycyl-tR synthetase / glycine--tR ligase	Aminoacyl-tR biosynthesis
comp94857_c0	8.92	EIN2; protein ETHYLENE INSENSITIVE 2	Plant hormone sig l transduction
comp100775_c0	8.92	transcription factor IIA, alpha/beta subunit	Basal transcription factors
comp100932_c0	8.94	PBD1; 20S proteasome beta subunit D1	Proteasome
comp89640_c0	8.95	ABCB4; auxin efflux transmembrane transporter MDR4	ABC transporters
comp108406	8.95	Hsp81.4; Heat shock protein 90-4	Protein processing in endoplasmic

_c0			reticulum; Plant-pathogen interaction
comp106319_c0	8.96	OB-fold nucleic acid binding domain-containing protein	D replication; Nucleotide excision repair; Mismatch repair; Homologous recombination
comp92383_c0	8.96	MGD2; monogalactosyldiacylglycerol synthase 2	Glycerolipid metabolism; Metabolic pathways
comp107748_c0	8.98	DWA1; WD repeat-containing protein DWA1	R transport
comp109235_c0	8.99	ARF1; auxin response factor 1	Plant hormone signal transduction
comp93778_c0	9	Ras-related small GTP-binding family protein	Protein processing in endoplasmic reticulum
comp106393_c1	9.01	GLN1;4; glutamine synthetase 1;4	Arginine biosynthesis; Alanine, aspartate and glutamate metabolism; Glyoxylate and dicarboxylate metabolism; Nitrogen metabolism; Metabolic pathways; Biosynthesis of amino acids
comp112333	9.02	ABF2; abscisic acid responsive	Plant hormone signal transduction

_c0		elements-binding factor 2	
comp104081 _c0	9.03	CYTC-2; cytochrome c-2	Sulfur metabolism; Metabolic pathways
comp94083_ c0	9.04	40S ribosomal protein S3a-1	Ribosome
comp96159_ c0	9.06	PKT3; 3-ketoacyl-CoA thiolase 2	Fatty acid degradation; Valine, leucine and isoleucine degradation; alpha-Linolenic acid metabolism; Biosynthesis of unsaturated fatty acids; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Fatty acid metabolism; Peroxisome
comp96338_ c0	9.07	APT3; adenine phosphoribosyl transferase 3	Purine metabolism; Metabolic pathways
comp108790 _c0	9.12	RPL34; 60S ribosomal protein L34-2	Ribosome
comp97238_ c0	9.12	putative UDP-glucose 6-dehydroge se 1	Pentose and glucuro te interconversions; Ascorbate and

			aldarate metabolism; Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways
comp92391_c0	9.13	Hsp81.4; Heat shock protein 90-4	Protein processing in endoplasmic reticulum; Plant-pathogen interaction
comp113678_c1	9.14	GSH2; glutathione synthetase 2	Glutathione metabolism; Metabolic pathways
comp89296_c0	9.14	SYP121; syntaxin-121	S RE interactions in vesicular transport
comp86773_c0	9.15	FUC1; alpha-L-fucosidase 1	Other glycan degradation
comp72237_c0	9.16	putative 3-methyladenine glycosylase I	Base excision repair
comp68588_c0	9.18	60S ribosomal protein L11-2	Ribosome
comp114195_c0	9.2	putative DH-ubiquinone oxidoreductase B18 subunit	Oxidative phosphorylation; Metabolic pathways
comp103047	9.2	CDC48C; cell division cycle 48C	Ribosome biogenesis in eukaryotes

_c0		protein	
comp85108_c0	9.21	RNA recognition motif-containing protein	mRNA surveillance pathway
comp88797_c0	9.23	nad1; NADH dehydrogenase subunit 1	Oxidative phosphorylation; Metabolic pathways
comp129724_c0	9.25	ceramide kinase	Sphingolipid metabolism
comp125760_c0	9.28	RHM1; UDP-L-rhamnose synthase	Amino sugar and nucleotide sugar metabolism
comp109473_c0	9.31	PKP-ALPHA; plastidial pyruvate kinase 1	Glycolysis / Gluconeogenesis; Purine metabolism; Pyruvate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
comp98858_c0	9.32	ATCES1; acyl-CoA independent ceramide synthase	Sphingolipid metabolism
comp106652_c0	9.33	ARFA1B; ADP-ribosylation factor A1B	Endocytosis

comp106393_c0	9.34	GLN1;4; glutamine synthetase 1;4	Arginine biosynthesis; Alanine, aspartate and glutamate metabolism; Glyoxylate and dicarboxylate metabolism; Nitrogen metabolism; Metabolic pathways; Biosynthesis of amino acids
comp93596_c0	9.36	MSH3; D mismatch repair protein MSH3	Mismatch repair
comp102871_c0	9.37	Glycosyl hydrolase family protein	Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways
comp113950_c0	9.38	DGD2; digalactosyldiacylglycerol synthase 2	Glycerolipid metabolism; Metabolic pathways
comp116445_c0	9.4	MEE32; bi-functional dehydroquinate-shikimate dehydrogenase	Phenylalanine, tyrosine and tryptophan biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Biosynthesis of amino acids
comp107577	9.42	40S ribosomal protein S17-4	Ribosome

_c0			
comp112284_c0	9.43	LAX3; auxin influx carrier LAX3	Plant hormone signal transduction
comp99989_c0	9.45	4CL3; 4-coumarate--CoA ligase 3	Ubiquinone and other terpenoid-quinone biosynthesis; Phenylalanine metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp114637_c0	9.46	DAD2; Defender against cell death 2	N-Glycan biosynthesis; Metabolic pathways; Protein processing in endoplasmic reticulum
comp123397_c0	9.48	RNR1; ribonucleoside-diphosphate reductase large subunit	Purine metabolism; Pyrimidine metabolism; Glutathione metabolism; Metabolic pathways
comp118621_c0	9.48	ribosomal protein L22p/L17e family protein	Ribosome
comp98603_c0	9.48	CPK6; calcium dependent protein kinase 6	Plant-pathogen interaction
comp114092	9.49	Aldolase-type TIM barrel family	Pentose phosphate pathway;

_c0		protein	Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
comp116011_c0	9.55	inositol-tetrakisphosphate 1-kinase 2	Inositol phosphate metabolism; Metabolic pathways; Phosphatidylinositol signaling system
comp110804_c0	9.59	adenosine deaminase-like protein	Purine metabolism; Metabolic pathways
comp109574_c0	9.62	MAB1; pyruvate dehydrogenase E1 beta	Glycolysis / Gluconeogenesis; Citrate cycle (TCA cycle); Pyruvate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
comp119399_c0	9.65	ribosomal protein L25/Gln-tR synthetase	Ribosome
comp112691_c0	9.67	ARF1; auxin response factor 1	Plant hormone signal transduction

comp96448_c0	9.68	putative calcium-binding protein CML27	Plant-pathogen interaction
comp102780_c0	9.73	ATPase, F1 complex, alpha subunit protein	Oxidative phosphorylation; Metabolic pathways
comp111963_c0	9.74	AVP1; Pyrophosphate-energized vacuolar membrane proton pump 1	Oxidative phosphorylation
comp113165_c0	9.75	PRH75; DEAD/DEAH box R helicase PRH75	Biosynthesis of secondary metabolites
comp111910_c0	9.79	sec23/sec24-like transport protein	Protein processing in endoplasmic reticulum
comp105693_c0	9.79	zinc ion binding protein	Ubiquitin mediated proteolysis
comp90645_c0	9.8	SPT1; serine palmitoyltransferase 1	Sphingolipid metabolism; Metabolic pathways
comp117745_c0	9.8	60S ribosomal protein L32-1	Ribosome
comp108154_c0	9.8	PUX4; CDC48-interacting UBX-domain protein 4	Protein processing in endoplasmic reticulum
comp102126_c0	9.82	transketolase	Pentose phosphate pathway; Carbon fixation in photosynthetic

			organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
comp98593_c0	9.86	RGA1; DELLA protein RGA	Plant hormone signal transduction
comp121897_c0	9.92	CAD5; cinnamyl alcohol dehydrogenase 5	Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp110562_c0	9.95	SEC22; vesicle transport protein SEC22	SNARE interactions in vesicular transport; Phagosome
comp95616_c0	9.95	ATDAD1; dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit DAD1	N-Glycan biosynthesis; Metabolic pathways; Protein processing in endoplasmic reticulum
comp124776_c0	9.96	QPT; quinolinate phoshoribosyltransferase	Nicotinate and nicotinamide metabolism; Metabolic pathways
comp110346_c0	9.99	RNA-binding (RRM/RBD/RNP motifs) family protein	Spliceosome
comp111185	9.99	VHA-C3; vacuolar-type H(+)-ATPase	Oxidative phosphorylation;

_c0		C3	Metabolic pathways; Phagosome
comp99478_c0	10.02	large subunit ribosomal protein L24e	Ribosome
comp115775_c0	10.06	RPL10B; 60S ribosomal protein L10-2	Ribosome
comp110289_c0	10.08	PAO2; polyamine oxidase 2	Arginine and proline metabolism; beta-Alanine metabolism
comp107164_c0	10.1	LACS9; long chain acyl-CoA synthetase 9	Fatty acid biosynthesis; Fatty acid degradation; Metabolic pathways; Fatty acid metabolism; Peroxisome
comp54218_c0	10.11	XPO1A; exportin 1A	Ribosome biogenesis in eukaryotes; RNA transport
comp97289_c0	10.15	U6 snRNA-associated Sm-like protein	RNA degradation; Spliceosome
comp109933_c0	10.16	AVP1; Pyrophosphate-energized vacuolar membrane proton pump 1	Oxidative phosphorylation
comp93773_c0	10.18	FT; protein FLOWERING LOCUS T	Circadian rhythm - plant
comp80800_c0	10.19	CAM3; calmodulin 3	Phosphatidylinositol signaling system; Plant-pathogen interaction

comp104824_c0	10.2	SYNC1; asparaginyl-tR synthetase, cytoplasmic 1	Aminoacyl-tR biosynthesis
comp105675_c0	10.21	CML42; calcium-binding protein CML42	Plant-pathogen interaction
comp89148_c0	10.26	SUS5; sucrose synthase 5	Starch and sucrose metabolism; Metabolic pathways
comp115233_c0	10.28	AlaAT1; alanine aminotransferase	Arginine biosynthesis; Alanine, aspartate and glutamate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Carbon metabolism; 2-Oxocarboxylic acid metabolism; Biosynthesis of amino acids
comp116564_c0	10.29	RPL16A; 60S ribosomal protein L16A	Ribosome
comp118390_c1	10.33	SAP18; histone deacetylase complex subunit SAP18	R transport; mR surveillance pathway
comp121289_c0	10.35	PRP40A; pre-mR -processing protein 40A	Spliceosome
comp113416	10.38	PSBP-1; photosystem II subunit P-1	Photosynthesis; Metabolic pathways

_c0			
comp113408_c0	10.43	PME2; pectin methylesterase 2	Pentose and glucuro te interconversions; Starch and sucrose metabolism; Metabolic pathways
comp131704_c0	10.46	HSP60; heat shock protein 60	R degradation
comp123402_c0	10.51	GAE3; UDP-D-glucuro te 4-epimerase 3	Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways
comp103672_c0	10.78	RHS2; calmodulin-like protein 7	Plant-pathogen interaction
comp96382_c0	10.81	PUX4; CDC48-interacting UBX-domain protein 4	Protein processing in endoplasmic reticulum
comp123485_c0	10.86	GCIP-interacting family protein	Spliceosome
comp92408_c0	10.87	LHCB4.2; light harvesting complex photosystem II interacting protein Lhcb4.2	Photosynthesis - anten proteins; Metabolic pathways
comp94579_c0	11.1	ELC; protein ELC	Endocytosis

comp111389_c0	11.22	60S ribosomal protein L34-3	Ribosome
comp117908_c0	11.37	lipid phosphate phosphatase gamma	N-Glycan biosynthesis
comp124216_c0	11.52	ASA2; anthranilate synthase component I-2	Phenylalanine, tyrosine and tryptophan biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of amino acids
comp112004_c0	11.8	PLDDELTA; phospholipase D delta	Glycerophospholipid metabolism; Ether lipid metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Endocytosis
comp118101_c0	11.96	SSI2; acyl-[acyl-carrier-protein] desaturase	Fatty acid biosynthesis; Biosynthesis of unsaturated fatty acids; Fatty acid metabolism
comp120565_c0	12.61	ZAC; ADP-ribosylation factor GTPase-activating protein AGD12	Endocytosis
comp116992_c0	12.93	Rae1-like protein	R transport

comp120985_c0	10.77	10.38	SCL28; SC35-like splicing factor 28	Spliceosome
comp86890_c0	11.37	11.62	RR2; transcription factor response regulator 2	Plant hormone signal transduction
comp108887_c0	11.43	11.32	CNX7; molybdenum cofactor synthesis family protein	Folate biosynthesis; Metabolic pathways; Sulfur relay system
comp100053_c0	11.47	10.59	EP3; chitinase class IV	Amino sugar and nucleotide sugar metabolism
comp120004_c0	11.86	12.4	CI51; NADH dehydrogenase [ubiquinone] flavoprotein 1	Oxidative phosphorylation; Metabolic pathways
comp115622_c0	12.51	11.65	RNA recognition motif-containing protein	Spliceosome
comp113557_c0	12.83	9.48	FD3; ferredoxin 3	Photosynthesis
comp112877_c0	13.31	12.71	LHCA5; photosystem I light harvesting complex protein 5	Photosynthesis - antenna proteins
comp125099_c0		9.34	EMB2719; probable 26S proteasome non-ATPase regulatory subunit 3a	Proteasome
comp71078_c0		9.34	SEC; putative UDP-N-acetylglucosamine--peptide N-	Other types of O-glycan biosynthesis

		acetylglucosaminyltransferase SEC	
comp125382_c0	9.35	HPT1; homogentisate phytyltransferase 1	Ubiquinone and other terpenoid-quinone biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp102407_c0	9.36	DEA(D/H)-box R helicase family protein	R degradation
comp115934_c0	9.37	RCE1; RUB1 conjugating enzyme 1	Ubiquitin mediated proteolysis
comp125262_c0	9.38	HCEF1; fructose-1,6-bisphosphatase	Glycolysis / Gluconeogenesis; Pentose phosphate pathway; Fructose and mannose metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
comp109370_c0	9.39	RD19; cysteine protei se RD19a	Plant-pathogen interaction
comp121314	9.39	FLS1; flavonol synthase 1	Flavonoid biosynthesis; Metabolic

_c0			pathways; Biosynthesis of secondary metabolites
comp119277_c1	9.4	TAF13; TBP-associated factor 13	Basal transcription factors
comp111347_c0	9.41	HDS; 4-hydroxy-3-methylbut-2-enyl diphosphate synthase	Terpenoid backbone biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics
comp112255_c0	9.43	CAM7; calmodulin 7	Phosphatidylinositol signaling system; Plant-pathogen interaction
comp94346_c0	9.43	USP; UDP-sugar pyrophosphorylase	Pentose and glucuronate interconversions; Galactose metabolism; Ascorbate and aldarate metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways; Biosynthesis of antibiotics
comp92728_c0	9.45	2-hydroxyacyl-CoA lyase	Peroxisome
comp110869	9.48	MTACP-1; acyl carrier protein 1	Oxidative phosphorylation;

_c0			Metabolic pathways
comp98742_c0	9.48	UXS6; UDP-XYL synthase 6	Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways
comp108197_c0	9.5	ubiquitin-related modifier 1	Sulfur relay system
comp123997_c0	9.52	BCCP2; biotin carboxyl carrier protein 2	Fatty acid biosynthesis; Pyruvate metabolism; Propanoate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Fatty acid metabolism
comp103057_c0	9.52	ALATS; Alanyl-tR synthetase	Aminoacyl-tR biosynthesis
comp106721_c0	9.53	BGLU47; beta-glucosidase 47	Cyanoamino acid metabolism; Starch and sucrose metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites

comp118504_c0	9.53	MPK6; MAP kinase 6	Plant hormone signal transduction; Plant-pathogen interaction
comp103607_c0	9.55	probable methionine--tRNA ligase	Selenocompound metabolism; Aminoacyl-tRNA biosynthesis
comp111454_c0	9.55	HYD1; probable 3-beta-hydroxysteroid-Delta(8),Delta(7)-isomerase	Steroid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp120952_c0	9.56	2-methoxy-6-polyprenyl-1,4-benzoquinol methylase	Ubiquinone and other terpenoid-quinone biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp94547_c0	9.56	WRKY2; putative WRKY transcription factor 2	Plant-pathogen interaction
comp93781_c0	9.57	polynucleotidyl transferase, ribonuclease H-like superfamily protein	RNA degradation
comp136730_c0	9.57	WRKY2; putative WRKY transcription factor 2	Plant-pathogen interaction
comp91392_c0	9.57	4CL3; 4-coumarate--CoA ligase 3	Ubiquinone and other terpenoid-quinone biosynthesis; Phenylalanine

			metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp93701_c0	9.6	SUS3; sucrose synthase 3	Starch and sucrose metabolism; Metabolic pathways
comp100237_c0	9.6	HSP81-2; heat shock protein 81-2	Protein processing in endoplasmic reticulum; Plant-pathogen interaction
comp105933_c0	9.61	APT1; adenine phosphoribosyl transferase 1	Purine metabolism; Metabolic pathways
comp110785_c0	9.62	ACD2; red chlorophyll catabolite reductase	Porphyrin and chlorophyll metabolism; Biosynthesis of secondary metabolites
comp132354_c0	9.62	CS26; cysteine synthase 26	Cysteine and methionine metabolism; Sulfur metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino

			acids
comp108632_c0	9.67	FNR2; ferredoxin--NADP reductase, leaf isozyme 2	Photosynthesis; Metabolic pathways
comp115034_c0	9.68	BRU6; indole-3-acetic acid-amido synthetase GH3.2	Plant hormone signal transduction
comp121333_c0	9.68	CCT2; phosphorylcholine cytidylyltransferase2	Glycerophospholipid metabolism; Metabolic pathways
comp135921_c0	9.69	RANGAP1; RAN GTPase-activating protein 1	RNA transport
comp103134_c0	9.73	ATB_BETA; serine/threonine protein phosphatase 2A 55 kDa regulatory subunit B beta isoform	mRNA surveillance pathway
comp112735_c0	9.75	CERK1; chitin elicitor receptor kinase 1	Plant-pathogen interaction
comp102246_c0	9.75	UBA1; ubiquitin-activating enzyme E1 1	Ubiquitin mediated proteolysis
comp109249_c0	9.8	EDA9; D-3-phosphoglycerate dehydrogenase	Glycine, serine and threonine metabolism; Metabolic pathways; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino

			acids
comp111585_c0	9.82	YKT61; VAMP-like protein YKT61	S RE interactions in vesicular transport
comp125922_c0	9.82	R helicase family protein	R degradation
comp112994_c0	9.82	ATPase, V0 complex, subunit E	Oxidative phosphorylation; Metabolic pathways; Phagosome
comp104951_c0	9.82	HDS; 4-hydroxy-3-methylbut-2-enyl diphosphate synthase	Terpenoid backbone biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics
comp103583_c0	9.84	Topoisomerase II-associated protein PAT1	R degradation
comp108312_c0	9.85	D -directed R polymerase, subunit M	Purine metabolism; Pyrimidine metabolism; Metabolic pathways; R polymerase
comp91637_c0	9.86	DUT1; deoxyuridine 5'-triphosphate nucleotidohydrolase	Pyrimidine metabolism; Metabolic pathways
comp98731_c0	9.9	emb2394; 50S ribosomal protein L6	Ribosome

comp96754_c0	9.91	VCR; varicose-related protein	RNA degradation
comp121778_c0	9.91	ubiquinol-cytochrome c reductase cytochrome c1 subunit	Oxidative phosphorylation; Metabolic pathways
comp98036_c0	9.93	CNX1; calnexin 1	Protein processing in endoplasmic reticulum; Phagosome
comp110433_c0	9.93	LHCA1; chlorophyll a-b binding protein 6	Photosynthesis - antenna proteins
comp116279_c0	9.95	SQD2; sulfoquinovosyldiacylglycerol 2	Glycerolipid metabolism; Metabolic pathways
comp111549_c0	9.95	LCB1; serine palmitoyltransferase	Sphingolipid metabolism; Metabolic pathways
comp126348_c0	9.95	AUD1; UDP-glucuronic acid decarboxylase	Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways
comp136826_c0	9.96	40S ribosomal protein S3a-1	Ribosome
comp109260_c0	9.99	CAM2; calmodulin 2	Phosphatidylinositol signaling system; Plant-pathogen interaction
comp112284	10.04	LAX3; auxin influx carrier LAX3	Plant hormone signal transduction

_c1			
comp116950_c0	10.04	MEE58; adenosylhomocystei se 1	Cysteine and methionine metabolism; Metabolic pathways
comp107138_c0	10.06	CLA1; 1-deoxy-D-xylulose-5-phosphate synthase	Thiamine metabolism; Terpenoid backbone biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics
comp97467_c0	10.09	ALDH3I1; aldehyde dehydroge se 3I1	Glycolysis / Gluconeogenesis; Pentose and glucuro te interconversions; Ascorbate and aldarate metabolism; Fatty acid degradation; Valine, leucine and isoleucine degradation; Lysine degradation; Arginine and proline metabolism; Histidine metabolism; Tryptophan metabolism; beta-Alanine metabolism; Glycerolipid metabolism; Pyruvate metabolism; Limonene and pinene degradation;

			Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics
comp129842_c0	10.11	Succinyl-CoA ligase [GDP-forming] subunit alpha-2	Citrate cycle (TCA cycle); Propanoate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
comp127808_c0	10.14	CKB4; casein ki se II beta subunit 4	Ribosome biogenesis in eukaryotes; Circadian rhythm - plant
comp136415_c0	10.17	BMS1 domain-containing protein	Ribosome biogenesis in eukaryotes
comp88116_c0	10.18	EIF3A; eukaryotic translation initiation factor 3A	R transport
comp116160_c0	10.2	40S ribosomal protein S9-2	Ribosome
comp112218_c0	10.21	AAO3; abscisic-aldehyde oxidase	Carotenoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp108546	10.31	DH-cytochrome b5 reductase-like	Amino sugar and nucleotide sugar

_c0		protein	metabolism
comp41755_c0	10.46	26S proteasome regulatory complex, non-ATPase subcomplex, Rpn2/Psmd1 subunit	Proteasome
comp114387_c0	10.49	copper amine oxidase 2	Glycine, serine and threonine metabolism; Tyrosine metabolism; Phenylalanine metabolism; beta-Alanine metabolism; Isoquinoline alkaloid biosynthesis; Tropane, piperidine and pyridine alkaloid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp95027_c0	10.49	xylose isomerase	Pentose and glucuronate interconversions; Fructose and mannose metabolism; Metabolic pathways
comp104690_c0	10.54	40S ribosomal protein S9-2	Ribosome
comp136532	10.54	DK2; D kinase 2	Nicotinate and nicotinamide

_c0			metabolism; Metabolic pathways
comp110195_c0	10.56	F-type H+-transporting ATPase subunit b	Oxidative phosphorylation; Photosynthesis; Metabolic pathways
comp124016_c0	10.57	DFB; folylpolyglutamate synthase 1	Folate biosynthesis; Metabolic pathways
comp121860_c0	10.64	D-glycerate 3-ki se	Glycine, serine and threonine metabolism; Glycerolipid metabolism; Glyoxylate and dicarboxylate metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
comp100996_c0	10.66	BGLU13; beta glucosidase 13	Cyanoamino acid metabolism; Starch and sucrose metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp121548_c0	10.67	ribophorin I	N-Glycan biosynthesis; Metabolic pathways; Protein processing in

			endoplasmic reticulum
comp128012_c0	10.71	pfkB-like carbohydrate ki se family protein	Pentose phosphate pathway
comp108458_c0	10.95	SHM1; serine transhydroxymethyltransferase 1	Glycine, serine and threonine metabolism; Cyanoamino acid metabolism; Glyoxylate and dicarboxylate metabolism; One carbon pool by folate; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
comp95101_c0	11.01	emb2742; CTP synthase family protein	Pyrimidine metabolism; Metabolic pathways
comp132946_c0	11.02	PAL1; phenylalanine ammonia-lyase 1	Phenylalanine metabolism; Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp134956_c0	11.07	Hrd3/Sel1L-like protein	Protein processing in endoplasmic reticulum

comp114670_c0	11.18	CML38; calcium-binding protein CML38	Plant-pathogen interaction
comp108044_c0	11.55	MDH; malate dehydroge se	Citrate cycle (TCA cycle); Cysteine and methionine metabolism; Pyruvate metabolism; Glyoxylate and dicarboxylate metabolism; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism
comp111545_c0	11.71	OTU-like cysteine protease family protein	Protein processing in endoplasmic reticulum
comp118288_c0	11.82	CHIP; E3 ubiquitin-protein ligase CHIP	Ubiquitin mediated proteolysis; Protein processing in endoplasmic reticulum
comp131391_c0	12.61	COX15; cytochrome c oxidase assembly protein COX15	Oxidative phosphorylation; Porphyrin and chlorophyll metabolism; Metabolic pathways; Biosynthesis of secondary

				metabolites
comp93096_c0		13.16	PLA2-ALPHA; phospholipase A2-alpha	Glycerophospholipid metabolism; Ether lipid metabolism; Arachidonic acid metabolism; Linoleic acid metabolism; alpha-Linolenic acid metabolism; Metabolic pathways; Biosynthesis of secondary metabolites
comp103214_c0	10.59		APL3; glucose-1-phosphate adenylyltransferase large subunit 3	Starch and sucrose metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways; Biosynthesis of secondary metabolites
comp107417_c0	10.61		endosomal targeting BRO1-like domain-containing protein	Endocytosis
comp104948_c0	10.63		UPL2; ubiquitin-protein ligase 2	Ubiquitin mediated proteolysis
comp104059_c0	10.64		glycine cleavage system H protein 2	Glycine, serine and threonine metabolism; Glyoxylate and dicarboxylate metabolism;

			Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics
comp103745_c0	10.66	argi se	Arginine biosynthesis; Arginine and proline metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Biosynthesis of amino acids
comp129824_c0	10.67	transducin/WD40 domain-containing protein	Ribosome biogenesis in eukaryotes
comp124879_c0	10.67	VPS24.1; vacuolar protein sorting-associated protein 24-1	Endocytosis
comp100611_c0	10.68	60S acidic ribosomal family protein	Ribosome
comp104265_c0	10.68	STT3A; dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3A	N-Glycan biosynthesis; Metabolic pathways; Protein processing in endoplasmic reticulum
comp82782_c0	10.73	60S ribosomal protein L15-1	Ribosome

comp105486_c0	10.74	40S ribosomal protein S15-4	Ribosome
comp124257_c0	10.76	PP2A; serine/threonine protein phosphatase 2A	mRNA surveillance pathway
comp106844_c0	10.8	KIN10; SNF1-related protein kinase catalytic subunit alpha KIN10	Regulation of autophagy
comp117202_c0	10.8	TSA1; tryptophan synthase alpha chain	Glycine, serine and threonine metabolism; Phenylalanine, tyrosine and tryptophan biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of amino acids
comp108395_c0	10.83	2-hydroxyacyl-CoA lyase	Peroxisome
comp112893_c0	10.84	ATPMEPCRB; Probable pectinesterase/pectinesterase inhibitor 41	Pentose and glucuronate interconversions; Starch and sucrose metabolism; Metabolic pathways
comp126457_c0	10.88	choline kinase	Glycerophospholipid metabolism; Metabolic pathways
comp125296	10.9	ribosomal protein L18e/L15	Ribosome

_c0			
comp118330_c0	10.9	ATPD; F-type H+-transporting ATPase subunit delta	Oxidative phosphorylation; Photosynthesis; Metabolic pathways
comp136955_c0	10.91	PGP6; ABC transporter B family member 6	ABC transporters
comp125633_c1	10.98	peroxidase 52	Phenylpropanoid biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp107974_c0	11.08	CUL1; cullin 1	Ubiquitin mediated proteolysis; Protein processing in endoplasmic reticulum
comp112385_c0	11.11	arginine/serine-rich protein splicing factor 31A	Spliceosome
comp127499_c0	11.12	OVA4; protein ovule abortion 4	Aminoacyl-tRNA biosynthesis
comp125735_c0	11.19	GLCAK; glucuronokinase G	Pentose and glucuronate interconversions; Ascorbate and aldarate metabolism; Amino sugar and nucleotide sugar metabolism; Metabolic pathways

comp115968_c0	11.28	PIP5K9; phosphatidyl inositol monophosphate 5 kinase	Inositol phosphate metabolism; Metabolic pathways; Phosphatidylinositol signaling system; Endocytosis
comp99470_c0	11.31	DNA glycosylase superfamily protein	Base excision repair
comp110570_c0	11.32	ATP-dependent caseinolytic (Clp) protease/crotonase family protein	Valine, leucine and isoleucine degradation; beta-Alanine metabolism; Propanoate metabolism; Metabolic pathways; Carbon metabolism
comp118369_c0	11.34	ATP synthase subunit G protein	Oxidative phosphorylation; Metabolic pathways
comp118545_c0	11.45	BAS1; cytochrome P450 734A1	Brassinosteroid biosynthesis
comp114805_c0	11.46	40S ribosomal protein S26-1	Ribosome
comp103152_c0	11.47	Oxoglutarate/iron-dependent oxygenase	Arginine and proline metabolism; Metabolic pathways
comp127845	11.56	PS2; inorganic pyrophosphatase 1	Vitamin B6 metabolism; Metabolic

_c0			pathways
comp121527_c0	11.58	ribosomal L18p/L5e family protein	Ribosome
comp121549_c0	11.68	R helicase family protein	Spliceosome
comp111645_c0	11.68	protein SM-LIKE 3B	R degradation; Spliceosome
comp119518_c0	11.72	NRPC2; D -directed R polymerase RPC2	Purine metabolism; Pyrimidine metabolism; Metabolic pathways; R polymerase
comp114727_c0	11.75	FAB1; 3-oxoacyl-[acyl-carrier-protein] synthase II	Fatty acid biosynthesis; Biotin metabolism; Metabolic pathways; Fatty acid metabolism
comp114748_c0	11.78	HAI2; protein phosphatase 2C 3	Plant hormone sig l transduction
comp116938_c0	11.86	AHP1; histidine-containing phosphotransmitter 1	Plant hormone sig l transduction
comp111990_c0	11.91	RPT4A; 26S proteasome AAA-ATPase subunit RPT4a	Proteasome
comp131064	12	phosphoribosylglyci mide	Purine metabolism; One carbon

_c0		formyltransferase	pool by folate; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics
comp125115_c0	12.03	TAG1; diacylglycerol O-acyltransferase 1	Glycerolipid metabolism; Metabolic pathways
comp123588_c0	12.1	transcription factor IIA, alpha/beta subunit	Basal transcription factors
comp120129_c0	12.1	GAPC2; glyceraldehyde 3-phosphate dehydroge se GAPC2	Glycolysis / Gluconeogenesis; Carbon fixation in photosynthetic organisms; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
comp120503_c0	12.14	EMB2769; protein embryo defective 2769	Spliceosome
comp119538_c0	12.14	GAPCP-2; glyceraldehyde 3-phosphate dehydroge se	Glycolysis / Gluconeogenesis; Carbon fixation in photosynthetic organisms; Metabolic pathways;

			Biosynthesis of secondary metabolites; Biosynthesis of antibiotics; Carbon metabolism; Biosynthesis of amino acids
comp137144_c0	12.15	Beta-glucosidase, GBA2 type family protein	Other glycan degradation; Sphingolipid metabolism; Metabolic pathways
comp117203_c0	12.16	transcription initiation factor IIB-2	Basal transcription factors
comp131243_c0	12.24	CYP98A3; cytochrome P450 98A3	Phenylpropanoid biosynthesis; Flavonoid biosynthesis; Stilbenoid, diarylheptanoid and gingerol biosynthesis; Metabolic pathways; Biosynthesis of secondary metabolites
comp116615_c0	12.31	SAM1; S-adenosylmethionine synthetase 1	Cysteine and methionine metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of amino acids

comp115141_c0	12.37	PAC1; 20S proteasome alpha-3 subunit	Proteasome
comp119946_c0	12.56	thioredoxin family protein	Ubiquitin mediated proteolysis
comp103200_c0	12.57	NDPK3; nucleoside diphosphate ki se 3	Purine metabolism; Pyrimidine metabolism; Metabolic pathways; Biosynthesis of secondary metabolites; Biosynthesis of antibiotics
comp127116_c0	12.59	URE; urease	Arginine biosynthesis; Purine metabolism; Metabolic pathways
comp118901_c0	12.68	3-oxoacyl-[acyl-carrier-protein] synthase	Fatty acid biosynthesis; Biotin metabolism; Metabolic pathways; Fatty acid metabolism
comp108615_c0	12.87	2-hydroxyacyl-CoA lyase	Peroxisome
comp106357_c0	13.06	FED_A; ferredoxin-2	Photosynthesis
comp102896_c0	13.15	PUR5; phosphoribosylformylglycimidine cyclo-ligase	Purine metabolism; Metabolic pathways; Biosynthesis of

				secondary metabolites; Biosynthesis of antibiotics
comp111496_c0	13.44		CRY2; cryptochrome 2	Circadian rhythm - plant
comp99769_c0	13.45		U6 snR -associated Sm-like protein LSm1	R degradation
comp111967_c0	13.77		ACA7; alpha carbonic anhydrase 7	Nitrogen metabolism