T.R.
BOLU ABANT İZZET BAYSAL UNIVERSITY
INSTITUTE OF GRADUATE STUDIES
Department of Biology

COMPARATIVE ANALYSIS OF TRANSCRIPTOME PROFILES DURING EARLY LEAF DEVELOPMENT IN IZA (*TRITICUM MONOCOCCUM* SSP *MONOCOCCUM*) WHEAT

DOCTOR OF PHILOSOPHY

YUNUS ŞAHİN

ACADEMIC SUPERVISOR
Prof. Dr. Nusret ZENCİRCİ
ACADEMIC CO-SUPERVISOR
Assist. Prof. Ercan Selçuk ÜNLÜ

BOLU, FEBRUARY - 2022

APPROVAL OF THE THESIS

COMPARATIVE ANALYSIS OF TRANSCRIPTOME PROFILES DURING EARLY LEAF DEVELOPMENT IN IZA (*TRITICUM MONOCOCCUM* SSP *MONOCOCCUM*) WHEAT submitted by **Yunus ŞAHİN** and defended before the Examining Committe Members listed below in partial fulfillment of the requirements for the degree of **Doctor of Philosophy** in **Department of Biology, Institute of Graduate Studies of Bolu Abant Izzet Baysal University in 3.02.2022** by

Examining Committee Members	**Signature**
Supervisor Prof. Dr. Nusret Zencirci Bolu Abant İzzet Baysal University	.....
Member Assoc. Prof. Faheem Shehzad Sivas Science and Technology University	.....
Member Assoc. Prof. Günce Şahin Bolu Abant İzzet Baysal University	.....
Member Assoc. Prof. Fatma Pehlivan Karakaş Bolu Abant İzzet Baysal University	.....
Member Assoc. Prof. İlker Büyük Ankara University	.....

Prof. Dr. İbrahim KÜRTÜL

Director of Instıtute of Graduate Studies

ETHICAL DECLARATION

In this thesis dissertation that was properly prepared according to the Thesis Writing Rules of Bolu Abant İzzet Baysal University of the Institute of Graduates Studies, I hereby declare that;

- All data, information, and documents presented in the thesis were obtained in accordance with the academic and ethical rules,
- All data, documents, assessments, and results were presented in accordance with the scientific ethical and moral rules,
- All works that were benefitted in the thesis were appropriately cited,
- No alteration was made in the data used,
- Study presented in this thesis is original,

Otherwise, I declare that I accept the loss of all my rights in case any contradiction that may arise against me.

Based on the plagiarism report that was generated on the date of 02/12/2021 by using predetermined filtrations set by Directorate of Institute of Graduate Studies of the Turnitin programme, a plagiarism detection sofware, the similarity index detected was 20%.

Yunus ŞAHİN

ABSTRACT

COMPARATIVE ANALYSIS OF TRANSCRIPTOME PROFILES DURING EARLY LEAF DEVELOPMENT IN IZA (*TRITICUM MONOCOCCUM* SSP *MONOCOCCUM*) WHEAT
PHD THESIS
YUNUS ŞAHİN
BOLU ABANT IZZET BAYSAL UNIVERSITY
INSTITUTE OF GRADUATE STUDIES
DEPARTMENT OF BIOLOGY
(SUPERVISOR: PROF. DR. NUSRET ZENCİRCİ)
(CO-SUPERVISOR: ASSIST. PROF. ERCAN SELÇUK ÜNLÜ)
BOLU, FEBRUARY 2022
xv + 134

Changing climate conditions has negative effects on the growth of plants through stress effects they create. There is an increasing need for varieties that perform better performance under adverse environmental conditions. DNA / RNA marker selection has become a promising approach to accelerate breeding application. Leaf development: in particular, it directs metabolic processes via its morphology and anatomy structure, and directly affects the yield of the plant by photosynthetic activity. Bread wheat ancestor Iza (*Triticum monococcum ssp. monococcum*), which has a higher yield compared to bread wheat in terms of cultivation in nutrient-poor soil, will provide conserved genetic mechanisms that will help us understand the developmental process about resistant to adverse conditions. Molecular information to be obtained at this stage will help us in developing molecular markers that can be used in studies of yield increase. Our general hypothesis in the dissertation is that in *Triticum monococcum ssp. monococcum* species, the active genes in the developmental mechanism of the plant from juvenile stage lead to the metabolic activities of the plant in the late stages and grain yield is directly related to these genetic arrangements in the early process. The study focused on peroxisome biogenesis and metabolism genes so far. They showed varied expression patterns in comparative developmental stages in the transcriptome analysis of some creals. The study covered the analysis of the data sets that obtained from the developmental stages over 100 RNA-seq libraries. The peroxisome metabolism genes were categorized into 21 different categories to reveal their association with growth and development. Catabolism of polyamine harbored very interesting story that has not been addressed before. The expression patterns of *BADH* and *CAT* genes were verified by qRT-PCR. In addition, peroxisome fission genes were screened under drought stress. *PEX11C* and *CAT* have shown consistent upregulation under drought stress, in terms of bioinformatics and, were confirmed by qRT-PCR.

KEYWORDS: Catalase, PEX11, GABA, Peroxisome metabolism, Juvenile and adult stages

ÖZET

IZA (*TRITICUM MONOCOCCUM* SSP *MONOCOCCUM*) BUĞDAYININ ERKEN YAPRAK GELİŞİM SÜRECİNDEKİ TRANSKRİPTOM PROFİLLERİNİN KARŞILAŞTIRILMALI ANALİZİ
DOKTORA TEZİ
YUNUS ŞAHİN
BOLU ABANT İZZET BAYSAL ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİYOLOJİ ANABİLİM DALI
(TEZ DANIŞMANI: PROF. DR. NUSRET ZENCİRCİ)
(İKİNCİ DANIŞMAN: DR. ÖĞR. ÜYESİ ERCAN SELÇUK ÜNLÜ)
BOLU, ŞUBAT - 2022
xv + 134

Değişen iklim koşulları, yarattıkları stres etkileri ile bitkilerin büyümesi üzerinde olumsuz etkiler yaratmaktadır. Olumsuz çevre koşullarında daha iyi performans gösteren çeşitlere artan bir ihtiyaç vardır. DNA / RNA markör seçimi, ıslah uygulamasını hızlandırmak için umut verici bir yaklaşım haline geldi. Yaprak gelişmi: özellikle morfolojisi ve anatomik yapısı ile metabolik süreçleri yönlendirir ve fotosentetik aktivite ile bitkinin verimini doğrudan etkiler. Besin yönünden fakir topraklarda, ekimi açısından ekmeklik buğdaya göre daha yüksek verime sahip olan ekmeklik buğday atası Iza (*Triticum monococcum ssp. monococcum*), olumsuz koşullara dayanıklılık konusundaki gelişim sürecini anlamamıza yardımcı olacak korunmuş genetik mekanizmalar sağlamaktadır. Bu aşamada elde edilecek moleküler bilgiler, verim artışı çalışmalarında kullanılabilecek moleküler belirteçlerin geliştirilmesinde bize yardımcı olacaktır. Tezdeki genel hipotezimiz, *Triticum monococcum* ssp. *monococcum* türleri, bitkinin erken gelişim döneminden itibaren gelişim mekanizmasında yer alan aktif genler, daha sonraki aşamalarda bitkinin metabolik faaliyetlerine yön verir ve tane verimi erken süreçte ki bu genetik düzenlemelerle doğrudan ilişkilidir. Çalışmamızda peroksizom biyogenezi ve metabolizma genlerine odaklandı. Bazı genlerin transkriptom analizinde karşılaştırmalı gelişim aşamalarında çeşitli ifadelenme örüntüleri gösterildi. Çalışma, 100'ün üzerinde RNA-seq kütüphanesinin gelişim aşamalarından elde edilen veri setlerinin analizini kapsamıştır. Peroksizom metabolizma genleri, büyüme ve gelişme ile olan ilişkilerini ortaya çıkarmak için 21 farklı kategoride sınıflandırıldı. Poliamin katabolizması, daha önce ele alınmamış ilginç bir hikâye ortaya koydu. *BADH* ve *CAT* genlerinin ekspresyon örüntüleri, qRT-PCR ile doğrulandı. Ek olarak, peroksizom fisyon genleri kuraklık stresi altında taranmıştır. *PEX11C* ve *CAT*, biyoinformatik açısından kuraklık stresi altında tutarlı bir artan regülasyon göstermiştir ve qRT-PCR ile doğrulanmıştır.

ANAHTAR KELİMELER: Katalaz, PEX11, GABA, Peroksizom metabolizması, Jüvenil ve yetişkin evreleri

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

LIST OF ABBREVIATIONS AND SYMBOLS

4CL **:** 4-coumarate-CoA ligase-like

4CL5 **:** 4-coumarate: CoA ligase 5

AACT1.3 **:** Acetoacetyl-CoA thiolase 1

AAE **:** Acyl-activating enzyme

ACBPs **:** Acyl-CoA-binding proteins

ACN1 **:** Arabidopsis acetate non-utilizing 1

ACX **:** Acyl-CoA oxidase

AIP **:** Peroxisomal ALS-interacting protein

ALNS **:** Allantoin synthase

ALS **:** Acetolactate synthase

APX **:** Ascorbate peroxidase

ATMS1 **:** Cobalamin-independent met synthase 1

BADH **:** Betaine aldehyde dehydrogenase

BOP **:** BLADE ON PETIOLE

CA **:** Cinnamic acid

CAT **:** Catalase

CCC : Chlormequat chloride

cDNA : Complementary DNA

cDNA : Messenger RNA

CDPK2 : Peroxisomal Ca2+-dependent protein kinase 2

CHD : Bifunctional CA-CoA hydratase/dehydrogenase

CHY1 : Peroxisomal 3-hydroxyisobutyryl-CoA hydrolase 1

CK **:** Cytokinin

CML3 : Calmodulin-like protein

CNL : Cinnamoyl-CoA ligase

CPK1 : Calcium-dependent protein kinase 1

CPK1 : Calcium-dependent protein kinase 1

CSD3 : Cu–Zn superoxide dismutase

CSY : Citrate synthase

CuAOs : copper-containing amine oxidases

CUC1 **:** CUP-SHAPED COTYLEDON1

CUC2 **:** CUP-SHAPED COTYLEDON2

CUC3 **:** CUP-SHAPED COTYLEDON3

DEG : Differentially expressed gene

DHNA-CoA : 1,4-dihydroxy-2-naphthoyl-CoA

DPA : Day post anthesis

DRN : DORNREONSCHEN

DRN : DORNREONSCHEN-LIKE

DRP : DYNAMIN RELATED PROTEIN

ECH2 : Enoyl-CoA hydratase

ER : Endoplasmic reticulum

ETT : ETTIN

FIS : FISSION

FPKM : Fragments per kilobase of transcript per million

G6PD1 : Glucose-6-phosphate dehydrogenase 1

GA : Gibberellin

GABA : 4-aminobutyrate

GGAT : Glutamate:glyoxylate aminotransferase

GLX1 : Glyoxalase I

GO : Gene ontology

GOX : Glycolate oxidase

HAOX : Hydroxyacid oxidase

HPR1 : Hydroxypyruvate reductase 1

IAA : Indole-3-acetic acid

ICL : Isocitrate lyase

IDI : Isopentenyl diphosphate isomerase

IndA : Pseudouridine-50-phosphate glycosidase

JA : Jasmonic acid

KAT : 3-ketoacyl-CoA thiolase

LBD : LOB-domain

LFS : LEAFLESS

LRR-RLKs : LEUCINE REPEAT RECEPTOR-LIKE KINASE

MFP : Multifunctional protein

MKP1 : MAP kinase phosphatases 1

MLS : Malate synthase

MP **:** MONOPTEROS

MVA : Mevalonic acid

MVD : Mevalonate 5-diphosphate decarboxylase

NADK3 : NADH kinase 3

NCBI : The national centre of biotechnology information

NS : Naphthoate synthase

OPR : OPDA reductase

OSB-CoA : o-succinylbenzoyl-CoA

PAO : flavin-containing polyamine oxidases

PAP : Purple acid phosphatase

PAT **:** PINFORMED1-dependent polar auxin transport

PBZ : Paclobutrazol

PC : Principal component

PCA : Principal component analysis

PEX : PEROXIN

PfkB : 6-phosphofructokinase isozyme 2-type carbohydrate kinase

PGD2 : 6-phosphogluconate dehydrogenase 2

PGL3 : 6-phosphogluconolactonase 3

pICDH : NADPH-generating dehydrogenase

PIN1 **:** PINFORMED1

PLL2 : POL like phosphatase 2

PMD1 : Peroxisomal and mitochondrial division 1

pMDH : Peroxisomal malate dehydrogenase

PMK : 5-phosphomevalonate kinase

PPD : PEAPOD

PS1 : Photosystem1

PUT : Putrescine

PXA1 : Peroxisomal long-chain fatty acid import protein 2

RLI : RNAse L inhibitor-like protein

RNS : Reactive nitrogen species

ROS : Reactive oxygen species

RPKM : The reads per kilobase of transcript per million

RuBisCo : Ribulose-1,5-bisphosphate carboxylase-oxygenase

SA : Salicylic acid

SCP1 : Small CDPK-interacting protein 1

SGAT : Serine:glyoxylate aminotransferase

SO : Sulfite oxidase

SOD : Superoxide dismutase

SOX : Sarcosine oxidase

SPD : Triamine spermidine

SPM : Triamine spermine

STM **:** NAC family and KNOX GENE SHOOT MERISTEMLESS

TCA : Tricarboxylic cycle

TCP : The class II TEOSINTE BRANCHED 1/ CYCLOIDEA/ PROLIFERATING CELL FACTOR

TE1 : Thioesterase 1

TF : Transcription factor

TMM : Trimmed mean of M-values

TOR : TARGET OF RAPAMYCIN

TPM : Transcript per million

WUS : WUSCHEL-RELATED HOMEOBOX

YUC : YUCCA

ACKNOWLEDGEMENTS

The author wishes to express his deepest gratitude to his supervisor Prof. Dr. Nusret Zencirci and co-supervisor Asisst. Prof. Ercan Selçuk Ünlü for their guidance, advice, criticism, encouragements and insight throughout the research.

The author would also like to thank Assoc. Prof. Andrei Semertenko for their guidance, advice, criticism, encouragements and insight throughout the research.

The technical assistance of Ms. Kathleen Frances Hickey and Dr. Taras Nazarov is gratefully acknowledged. Moreover, i would like to thank Dr. Sharol Marcec and Alyssa Parish for their support.

The author would like to offer his special thanks to the lab members of Asisst. Prof. Ercan Selçuk Ünlü.

This research was supported by TÜBİTAK with 1059B142000023 application number. I would like to express my sincere gratitude to TÜBİTAK.

A part of this study was also supported by the Scientific Research Projects, Grant No: 2020.03.01.1435.

1. INTRODUCTION

1.1 Wheat Yield and Demands

According to the USDA, Turkey -ranks 8th in world wheat exports with wheat self-sufficiency remaining between 95-100% in 2019. However, in some years the demand for wheat cannot be met due to unfavorable climatic conditions (Republic of Turkey, Minister of Agriculture and Forestry report). As of the 2019-2020 season Turkey contained 3.3 % of the total global land area cultivated for wheat. This field also represents 35% of the total cultivated agricultural land in Turkey. In the 2018/19 production season, there was a 4.8% decrease in cultivation area and a 1.1% decrease in yield compared to the previous season. Accordingly, wheat cultivation decreased by 7% in 2018/19 production season. Wheat production decreased to 19 million tons in the 2019/20 production season.

Unfavorable climate conditions are one of the main factors that cause a decrease in yield. Climate change in Turkey will cause change in precipitation patterns and a rise in temperature. Increased temperature makes drier lands causing stresses such as salt, heat, and drought in crops. Combined stress factors have a common effect on the plant. To cope with the challenging future, developmental properties and patterns would be crucial for breeding programs. For example, it is possible to improve the crops' tolerance against high temperature stress, in part with the amount of available water in the soil. It is vital that the crops be resistant to several stress factors soon. Molecular approaches will turn out to be powerful tools to achieve such purposes.

According to World Resources Institute data, Turkey will suffer from high water stress by 2040 (Figure 1.1). A growing world population demands more food, water, and energy. Besides the growing world population, climate change threatens our food resources by disrupting land areas where we cultivate crops. We need to take action to face the challenging future of Turkey. More resistant crops with high yields would be one of the solutions to this challenging future.

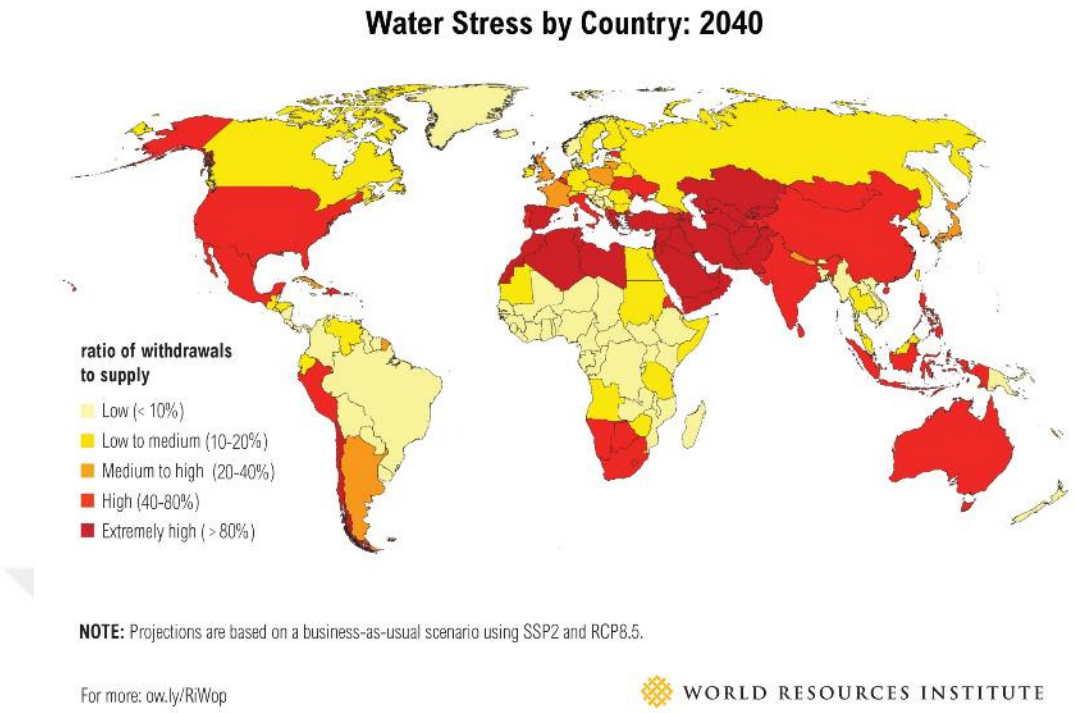


Figure 1.1. Ratio of withdrawals according to world resources institute.

Different combinations of factors might affect Turkey's land from the effects of rising temperatures or shifting precipitation patterns. Hence, drought stress has become increasingly important in plant breeding due to the declining nature and quantity of water resources around the world, leading to changes in the normal physiological functions of plants. Detailed observations and examinations should be conducted on the effects of climate change in Turkey. The crops are unarguably suffering from different kinds of or combined effects of stresses.

1.2 Einkorn

Wheat grain is composed of three layers: bran, embryo and endosperm. Like every cereal has, wheat has six developmental stages: germination, seedlings, tillering, panicle initiation, flowering, and harvest. Germination, seedling, and tillering are vegetative phases. Panicle initiation and flowering are reproductive phases, and harvest is the ripening phase.

Wheat plants are classified as their chromosome numbers: diploid (AA), tetraploid (AABB) and hexaploid (AABBDD). The diploid genome is basic to all polyploidy of wheat (Figure 1.2). *Triticum boeoticum* Boiss., *Triticum urartu* Thum., *Triticum monococcum ssp. monococcum*, *Triticum sinskajae* are diploid

wheat. *T. urartu* Thum. and *T. boeoticum* Boiss. are wild forms of the group and there is no naked seed variety in the group except *T. sinskajea*. The species *T. boeoticum* and *T. urartu* are genetically set apart due to the sterility of their hybrids. It is claimed that cultivated *T. monococcum ssp. monococum* was selected from *T. boeoticum* rather than *T. urartu* (Figure 1.2). In Turkey, *T. monococcum* is widespread in the foothills of Ağrı Mountain, the grassland of east and middle Anatolia (1).

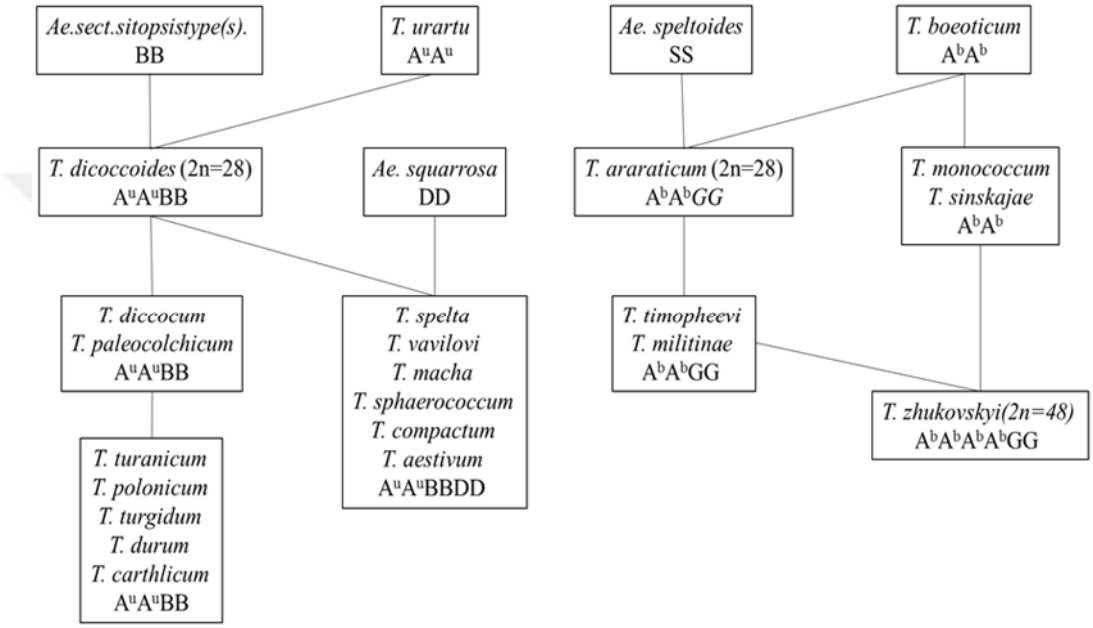


Figure 1.2. The origin of polyploid *Triticum* species. The letter representative donates the genome constitution of corresponding species.

T. monococcum ssp. *monococcum* is first called einkorn, successfully cultivated, and domesticated from its wild diploid progenitor, -*T. boeoticum*: however modern wheat genotypes are associated with *T. urartu*. Einkorn can well adapt to adverse environmental conditions. It has high nutritional value and gives acceptable yields on poor soils. Although it was widely cultivated in the beginning of agriculture, now, it is a relic crop and marginally cultivated in some provinces of Turkey (Bolu, Bilecik, Kastamonu, Çankırı, Sinop, Karabük), Italy, some Balkan countries, and Morocco (2).

According to recent studies, Einkorn has a distinct phenolic acid content. Its phenolic content is significantly higher than tetraploid and hexaploid wheat species (2,3). Most wheat breeding programs have focused on starch quality, grain size and so on, yet bioactive compound content has been underappreciated so far. Cultivated,

hulled wheat-einkorn- may supply a good genetic background to improve the bioactive compound content of modern hexaploid and tetraploid wheat species. Knowledge of transcript expression patterns in einkorn would offer valuable understanding to achieve this purpose. Detailed information on developmental mechanisms and gene expression patterns of Einkorn might provide well-timedbreeding applications. Unfortunately, knowledge of expression patterns in the vegetative phase is scarcein einkorn.

Plant performance and yield are strongly dependent on vegetative plant growth and development. Manipulations would improve developmental performance, thus eventually increasing biomass and yield through changes in developmental features, such as architectural traits, leaf, and vasculature morphology (Figure 1.3). Development-related genes, transcription factors (TFs), phytohormones and environmental factors are major external and internal factors. They can affect morphological, architectural, and vasculature traits.

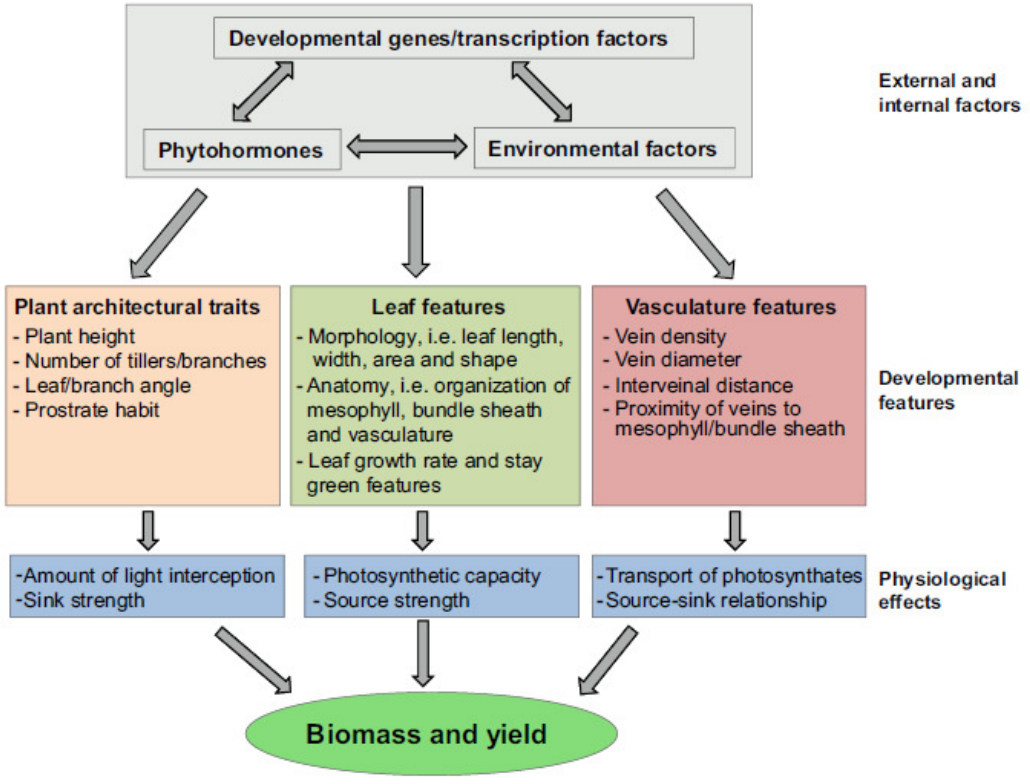


Figure 1.3. Developmental features, physiological effects, and external and internal factors is associated with biomass and yield (adapted from Mathan et al. (2016) (4)).

The optimization and/or improvement of these developmental features and traits has become essential for the efficient performance of crop plants. Important things in increasing crop yield have been achieved during the green revolution. For example, the reducing height of wheat increases tillering number in wheat, leading to more grain in wheat. Changing structural, architectural and developmental features of crops would be new research fields for breeders in the future, marking a significant step forward in increasing crop yield (5). Plant height, number of tillers, leaf/branches angles, prostrate habit affect light interception and sink strength. Morphology, anatomy, leaf growth rate and duration of staying green define photosynthetic capacity and source strength. The transport of photosynthate and the source-sink relationship are associated with vein density, diameter, interveinal distance, and vein proximity to the bundle sheath. Manipulations of those features and traits by using molecular techniques would offer a new perspective on traditional breeding programs.

Manipulation of the genetic background and network is possible using functional genomic approaches, but it takes time and requires a strategic manipulation plan for increasing or decreasing the expression of a gene of interest. Additionally, it is not always possible to find the true target in genetic manipulations. However, bioinformatics approaches have opened the door to a new plant breeding era. It reduces the cost of experiments. Understanding the molecular mechanism of the desired trait in crop plants is crucial for molecular breeding approaches. The genes of those mechanisms may be used as DNA/RNA markers which is becoming a promising approach for accelerating the breeding processes.

1.3 Molecular Mechanism of Leaf Development in Some Model Organisms

Leaf development is the most important process for vital plant activities. Leaf formation is physiologically observed as cell differentiation, division of differentiated cells and volume expansion in the shoot apical meristem. Differentiation of the cells creates adaxial/top and abaxial/bottom patterning and-forms an adaptation basis by maximizing photosynthetic capacity while preventing water loss. Three- or four-leaf primordia have already existed in the embryo. Initials of the root and first leaf are present as well. After imbibition, the coleoptile end forces its way through the soil, where it helps the emergence of the first leaf.

The leaf originates from the shoot apical meristem. The shoot apical meristem has two zones: the peripheral and the central zone. The central zone houses pluripotent cells. It contains tunica and corpus cells. The peripheral zone serves as the new organ initiation. First, primordia emergence begins in the peripheral zone. Incipient primordium is formed by the founder cells that are recruited from the peripheral zone of the shoot apical meristem. Distal growth occurs after the initiation of the leaf primordium and is typically followed by adaxial–abaxial and proximal–distal axis establishment. The lamina initiates at the site neighboring the margin. The margin separates the lamina from the petiole along the medio–lateral axis. After the occurrence of the margin, intercalary growth, which results in the overall expansion of leaf area in multiple directions, occurs throughout the entire lamina. CUP-SHAPED COTYLEDON1 (CUC1), CUC2, and CUC3, are the major transcription factors required to initiate shoot apical meristem and boundary formation (6).

Auxin is the key growth regulator in the process of leaf primordium initiation, as it specifies the cells that are distinguished from the peripheral zone. PINFORMED1 (PIN1)-dependent polar auxin transport (PAT) protein transports auxin molecules to neighboring cells in polar direction (7). Auxin maxima is established in incipient primordia by PAT. One of the main actors in the distribution of PINFORMED1 is the AUXIN1 protein. It indirectly promotes the accumulation of auxin in the cells. The rate of auxin molecules affects NAC family and KNOX GENE SHOOT MERISTEMLESS (STM) genes by which the undifferentiated state of the cells is maintained. The process has a key role in primordium initiation. MONOPTEROS (MP) activates AHP6 and inhibits ARR7/ARR15 to control meristematic fate through the regulation of cytokinin (CK) homeostasis (8) (Figure 1.4.).

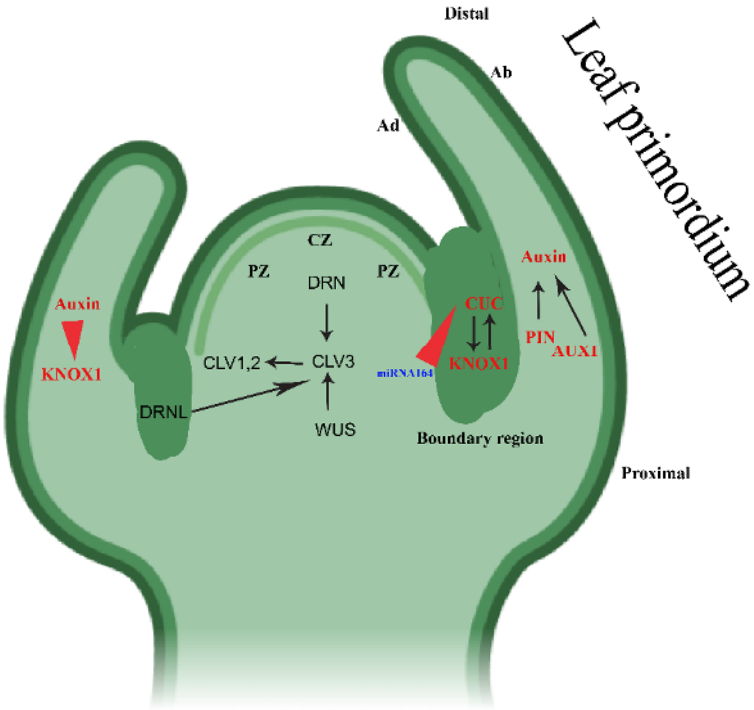


Figure 1.4. The representative demonstration of the shoot apical meristem. Peripheral zone, central zone and boundary region beside leaf primordium are marked. KNOX1 genes which are repressed by auxin molecules are indicated by red arrow. Black arrows point upregulation role of back-end genes.

Accumulation of auxin inhibits the expression of KNOX genes together with the activation of organ specific transcription factors; thus, they specify the fate of stem cells that are harbored by the shoot apical meristem (9). PIN1 and shoot apical meristem require environmental and metabolic signals. The most critical signals are light and sugar. The TARGET OF RAPAMYCIN (TOR) kinase receives the signals (10). Besides this, cells conduct light signals through a cytokinin transduction pathway (10).

Ethylene response factor-type transcription factor LEAFLESS (LFS) is induced by auxin maxima in the peripheral zone. The expression of LEAFLESS is required for leaf initiation (11). DORNREONSCHEN (DRN) and DORNREONSCHEN-LIKE (DRNL) are two factors that correlate with leaf primordia. These genes promote cell division, thus auxin maxima induces cells to develop into leaf primordia by activating LFS/DRNL expression (12).

1.4 Adaxial/Abaxial, Proximal/distal and Medio/Lateral Pattering

HD-ZIP III family (PHB, PHV, REV) specifies the adaxial cell fate of the leaf, while KAN1 domain expression cells become the abaxial domain (13). HD-

ZIP III family is regulated by miRNA164/165 (14). The regulation of miRNA164/165 is promoted by ARGONUATE10 (15). LITTLE ZIPPER also inhibits HD-ZIP III expression (16). AS2, a LOB-domain (LBD) transcription factor promotes adaxial fate (17).

KAN1 and KAN2, ARFs, ETTIN (ETT), ARF4 and ARF2 promote abaxial cell fate (18). miRNA390 promotes ETT, ARF2 and ARF4. AS1-AS2 negatively regulates ETT and ARF4. KAN1 directly represses the expression of AS2, and AS2 indirectly represses the expression of KAN1 (18). AS2 negatively regulates ETT, KAN2, and YAB5. APUM23 is a new regulator of leaf polarity, that encodes approximately 20 PUF RNA-binding proteins in *Arabidopsis thaliana* and interacts with leaf polarity that is required for the maintenance of genes (19). Moreover, AS1-AS2 inhibits the transcription of miRNA164 and ETT in the adaxial, thus indirectly affecting the cell fate of the abaxial site (20). YAB1-YAB2-YAB3 affect abaxial cell fate, being dimerized with LEUNIG family transcription corepressors (21).

How proximal-distal polarity is established remains barely known. Auxin response competency is required for the initiation of the proximal-distal axis. This region of competency is associated with the juxtaposition of the cells expressing the HD-ZIP III, KAN, miR166a and adaxial and abaxial patterning genes. The expression of these genes may reflect induced proximodistal outgrowth by auxin. The blade formation is repressed in the proximal end. BLADE ON PETIOLE (BOP1) and BOP2 activates AS2 and the boundary expressed LBD genes (22). Furthermore, BOP1 and BOP2 inhibit KNOX1.

The most important gene of medial/lateral specification is the YABBY gene family, which encodes proteins with a zinc finger and a helix-loop-helix domain. HD-ZIP III, AS, and KANADI pathways promote the YABBY genes (23). The WOX1 gene is expressed along with the adaxial-abaxial juxtaposition and overlaps with PRS at the marginal region of the leaf. WOX1 expression occurs around the meristem at the boundary between the HD-ZIP III and KAN expression domains (24). The expression of WOX1/PRS is negatively regulated by AS2 in the adaxial domain of the leaf. AS2 expression is also repressed by WOX1/PRS to restrict its expression on the adaxial side of the leaf.

1.5 Leaf Expansion and Outgrowth

As the leaf grows, the cell division at the distal end ceases and the expansion process occurs in which the cells at the base are strongly vacuolated rather than proliferating. Spatial differences in the leaf meristematic activities may explain the differences in timing and placement of lateral leaf outgrowths between monocots and eudicots. The maintenance of an undifferentiated state in various meristems is promoted by WUSCHEL-RELATED HOMEOBOX (WUS) -member of the WOX family that is typically expressed at the adaxial-abaxial junction (25). WOX activities are regulated by CLE via common activity with LEUCINE REPEAT RECEPTOR-LIKE KINASE (LRR-RLKs) (26). Leaf development begins with medio-lateral polarity development and is followed by blade initiation. Cell division is perpendicular to the medial-lateral axis, leading to a flattened leaf. The adaxial domain has lower auxin than the abaxial domain (27). This phenomenon forms a high auxin response in the middle domain. WOX1 and PRS repress adaxial and abaxial specificity of the cells in the margins of the middle domain (28) (Figure 2.). PRS is repressed by adaxial factor AS2 which requires TOPLESS family corepressors in the adaxial domain (29). In the abaxial domain, the identity of middle domain cells are repressed by KAN1 and KAN2, and ETT (ARF3), ARF2 and ARF4 (30). Members of the YAB family activate WOX1 expression in middle domain (31). In addition, WOX1 and PRS promote the expression of KLUH encoding cytochrome p450 CYP78A5. Yucca (YUC) flavin monooxygenase-like enzymes -auxin biosynthesis factor- play an important role in leaf margin development and blade outgrowth. Additionally, cell wall elasticity is crucial for the formation of asymmetry in leaf primordia. The class II TEOSINTE BRANCHED 1/ CYCLOIDEA/PROLIFERATING CELL FACTOR (TCP) inhibits marginal meristem activity (32), therefore promotes the timing of cell division to expansion. The expression of TCP is restricted to the distal region of leaves by means of miRNA319/JAW (33). GAs and brassinosteroids promote leaf growth by increasing cell proliferation and expansion. GAs involves the repression of cell-cycle inhibitors such as KIP-RELATED PROTEIN 2 and SIAMSE. BRASINOSTEROID INSENSITIVE1 or DWARF4 results in larger leaves. Brasinosteroid controls exit from mitosis. Brasinosteroid signaling represses PEAPOD1 (PPD1) and PPD2 which encode TFs that limit meristematic cell proliferation.

1.6 Transcriptomes of the Seedling Growth and Development in Crop Plants

Singh et al. (2018) observed transcriptome profiles of different tissue types in the seedlings of a wild and domesticated common bean (Phaseolus vulgaris L.) accession (34). He suggested that domesticated common bean accession has higher differentially expressed genes (DEGs) compared with wild type common bean. bHLH, HB and (R1)R2R3_Myb families TFs were present among the DEGs, suggesting the importance of these TFs in the development process. Among DEGs. Cell death, death, response to biotic stimulus, metabolic processes, oxyreductase activity and catalytic activity of gene ontology (GO) terms were overrepresented. Alternatively, because of KEGG enrichment analysis in the leaves, pathways associated with, nitrogen metabolism, photosynthesis, biosynthesis of various compounds and specific metabolic processes were significantly enriched.

Ji et al. (2019) compared the seedlings of tomato treated with two growth retardants (35). They relate chlormequat chloride (CCC), paclobutrazol (PBZ) treatments to GA metabolism genes. The GA metabolism pathway genes- GA20ox or GA3ox were decreased in CCC/PBZ-treated roots and shoots, and the degradation gene -GA2ox- was down-regulated in CCC/PBZ-treated roots and shoots. Furthermore, SlPIF1 and SlEXPs were down regulated under the treatments, suggesting the molecules- CCC and PBZ- could restrict the GA signaling pathway. Hence GA20ox or GA3ox, SlPIF1 and downstream gene SlEXPs can be good manipulation target genes to improve biomass of the crops, thereby the yield of crops.

Li et al. (2010) demonstrated transcriptome profiles of the leaf 3rd in four sampled tissues (36). They divided the leaf into four zones as a basal zone (base, 1 cm above the leaf three ligule, sink tissue), a transitional zone (−1 cm, 1 cm below the leaf two ligule, undergoing the sink-source transition), a maturing zone (+4 cm, 4 cm above the leaf two ligule) and a mature zone (tip, + 1 cm below the leaf). Alternative splicing analysis of the zones suggested that transcript processing is under developmental regulation in the maize leaf. Intron retention and alternative 5′ end processing can regulate transcript isoform abundance in the maize leaf. 16,502 genes were differentially expressed among developmental zones. They clustered these genes by using the k-means algorithm, resulting in 18 clusters. The genes that clustered in the first two groups were enriched in cell wall biosynthesis,

DNA synthesis, cell cycle regulation and chromatin structure, protein metabolism, potential signaling proteins (G proteins and leucine-rich repeat kinases), auxin and brassinosteroid biosynthesis and signaling, respiratory pathways and vesicle transport. In clusters 3rd and 4th, tetrapyrrole biosynthesis, plastid targeting, lignin and suberin production and secondary cell wall biosynthesis. In clusters 5th and 6th, photosynthesis, the Calvin cycle, redox regulation, the light reactions of photosynthesis, starch metabolism, sucrose metabolism, and minor carbohydrates were significantly enriched. While the genes that were assigned to the cluster 3rd and 4th were usually expressed in transition zone, the genes that assigned to the 5th and 6th were expressed mainly at the tip of the maize leaf. As a result, the study reveals the proximal-distal leaf gradient is produced by a localized transition in mRNA abundance.

Shi et al. (2020) documented differences in photosynthesis between rice (monocotyledonous) and *Arabidopsis* (dicotyledonous) plants (37). They sampled early growth structures called coleoptile, incomplete leaf, and true leaf in rice, while cotyledon, first true leaf and second true leaf in *Arabidopsis*. GO analysis revealed that photosynthesis, photosynthesis light reaction, photosynthetic electron transport in PSI, and photosynthesis light harvesting in PSI were enriched in incomplete leaves versus coleoptile. Photosynthesis, photosynthesis dark reaction, photosynthesis, light reaction, and photosynthesis light harvesting were enriched in true leaf versus incomplete leaves. In contrast, in Arabidopsis, no photosynthesis-related biological pathways were enriched in the early developmental stages. These data suggest that monocotyledonous and dicotyledonous have different photosynthetic development in the early growth stage.

Trevisan et al. (2020) examined the effects of exogenous Brassinosteroid in leaves, root and shoot tissues of maize (38). They identified around 2000 DEGs after the application of exogenous brassinosteroid. In leaves, responses to abiotic stimuli, biotic stimuli, endogenous stimuli, external stimuli, carbohydrate metabolic processes, lipid metabolic processes and growth terms were enriched as biological processes, while the GO terms related to the roots included cellular homeostasis, carbohydrate metabolic processes, transport, lipid metabolic processes and cell growth.

Wu et al. (2020) identified 115 members of wall-associated kinases in *Brachypodium distachyon* coleoptiles at 48 h post-germination by using

transcriptome analysis (39). As a result of de novo assembly and counting the transcription of wall-associated kinase families, these 115 members were divided into eight structural classes. Five members -BdWAK2, BdWAK10, BdWAK42, BdWAK72, and BdWAK108- were identified as the most highly expressed. Overall, they concluded that the members of wall-associated kinases play an important role in cell expansion during coleoptile growth.

Zhao et al. (2020) cloned the Seita.5G121100 into a mutant with fully droopy leaves after heading and the wild type (40). Next, they transformed the mutant by using the CRISPR/Cas9 system to knock out Seita.5G121100. Furthermore, they rescue the transformed plant with overexpression of Seita.5G121100. Results showed that Seita.5G121100 encodes droopy leaf 1. Additionally, it is a member of a member of subfamily II of LRR-RKs.

1.7 Peroxisomes and Their Associations with Development and Drought Stress

Peroxisomes are multi-facet functional, single-membrane organelles. They are involved in many metabolic pathways. Two models have been proposed for peroxisome biogenesis so far. One model suggests that peroxisomes grow and divide into two new peroxisomes while the other model supports de novo synthesis through endoplasmic reticulum. *PEX* genes are responsible for biogenesis of peroxisomes. Since all peroxisomal proteins are synthesized in the cytosol, they all need to be imported into to peroxisomal membrane through PEROXIN proteins. *PEX19, PEX3* and *PEX16* are well conserved among yeast, plant, and animal cells. *PEX16* recruits *PEX3* to the ER membrane. *PEX3* is the membrane anchor for *PEX19* while *PEX19* is an anchor for peroxisomal membrane proteins.

PEX13 and *PEX14* form a receptor-cargo docking complex in the peroxisome membrane. *PEX5* and *PEX7* are the genes that recognize peroxisomal lumen proteins. *PEX5* can be recycled from the peroxisomal matrix back to the cytosol. Recycling of *PEX5* is cooperative effort of a group of genes. *PEX4* and its membrane anchor PEX22, three RING-type ubiquitin ligases, *PEX2*, *PEX10* and *PEX12*, and two AAA ATPases, *PEX1* and *PEX6*, which are tethered to their membrane anchor *PEX26*/*APEM9*.

A pre-existing peroxisome can be divided into two new peroxisomes. First, PEX11 protein promote the process of elongation/tubulation. Five isoforms of the *PEX11* gene have been identified as *PEX11a, PEX11b, PEX11c, PEX11d, PEX11e.*

Once the peroxisome reaches a mature size, the second important process, membrane constriction and fission, is mediated by DYNAMIN RELATED protein DRP3A, DRP3B, their membrane anchors FISSION 1A (*FIS1A*) and *FIS1B* proteins. Peroxisomal and mitochondrial division 1 (*PMD1*), which is the protein that is shared with mitochondria is also involved in peroxisome proliferation.

Peroxisome can interact with other organelles through peroxule which is membrane extension. Peroxule forms physical contact and helps metabolite and protein exchange. Peroxule contacts with mitochondria and chloroplasts under abiotic stress in a PEX11a- and reactive oxygen species (ROS) dependent manner, therefore playing a key role in regulating stress perception and rapid cell responses to environmental cues.

The range of peroxisome functions has increased over the last decades. Table 1.1 shows the functions of the peroxisome in the cell.

Table 1.1. Peroxisomal metabolisms that take place in peroxisome lumen.

• ROS and RNS metabolism
• H_2O_2 and NO signaling
• Photorespiration
• Phytohormones biosynthesis
• Fatty acid β-oxidation
• Glyoxylate cycle
• Polyamine catabolism
• Amino acids metabolism
• Indole glucosinolates metabolism
• Ureide metabolism
• Purine catabolism
• Biotin biosynthesis
• Ubiquinone biosynthesis
• Phylloquinone biosynthesis
• Isoprenoids biosynthesis
• BA derivate biosynthesis
• Sulfite metabolism

Peroxisomes house many metabolic pathways. Photorespiration, fatty acid breakdown, some phytohormone biosynthesis, ROS metabolism, Glyoxylate cycle, NADPH regeneration, signaling, biosynthesis and catabolism of some substances.

Through -oxidation, fatty acids are broken down into their metabolites to generate acetyl-CoA. The enzymes acyl-CoA oxidase (ACX), multifunctional protein (MFP), and 3-ketoacyl-CoA thiolase (KAT) generate acetyl-CoA and fattyacyl-CoA. The metabolism of unsaturated fatty acids involves ECH2, which encodes an enoyl-CoA hydratase. In the absence of glucose, the glyoxylate cycle is

a part of the tricarboxylic cycle (TCA) that allows the utilization of two carbon compounds. The peroxisomal enzymes citrate synthase (CSY), isocitrate lyase (ICL), malate synthase (MLS) and Arabidopsis acetate non-utilizing 1 (ACN1) are involved in the glyoxylate cycle in peroxisomes.

The β-oxidation pathway also leads to the synthesis of a few key phytohormones such as Indole-3-acetic acid (IAA), jasmonic acid (JA), and salicylic acid (SA). Cinnamic acid (CA) is imported into the peroxisome by peroxisomal long-chain fatty acid import protein 2 (PXA1). Cinnamoyl-CoA ligase (CNL), bifunctional CA-CoA hydratase/dehydrogenase (CHD), 3-ketoacyl thiolase 1 (KAT1), thioesterase 1 (TE1), AtAIM1, OPDA reductase (OPR) participate in phytohormone biosynthesis. ACX5, MFP2, KAT2 and PEX6 are also important for JA biosynthesis.

Photorespiration is an oxidative photosynthetic carbon cycle. It starts with carbon fixation at the stroma of the chloroplast. Ribulose-1,5-bisphosphate carboxylase-oxygenase (RuBisCo) is the first protein that initiates the carbon fixation cycle. However, oxygenation of RuBisCo catalyzes the formation of an alternative reaction in which it generates one molecule of 3PGA and one of 2PG rather than two molecules of 3PGA. 2PG is a toxic molecule for many enzymes and oxidizes organic carbon. Photorespiration is the process of converting 2PG into glycolate. The peroxisome localized photorespiratory enzymes include glycolate oxidase (GOX), glutamate:glyoxylate aminotransferase (GGAT), serine:glyoxylate aminotransferase (SGAT) and hydroxypyruvate reductase 1 (HPR1), while the NADH-producing enzymes peroxisomal malate dehydrogenase (pMDH), hydroxyacid oxidase (HAOX) and the H_2O_2-degrading enzyme catalase (CAT) are indirectly involved.

ROS and RNS metabolism can occur within peroxisomes. Ascorbate peroxidase (APX), CAT, superoxide dismutase (SOD), Cu–Zn superoxide dismutase (CSD3) are involved in ROS/RNS metabolism in peroxisome.

NADPH is a critical cofactor required by several reductive biosynthetic and detoxification pathways in peroxisomes. Regeneration of NADPH takes place in peroxisomes. Glucose-6-phosphate dehydrogenase 1 (G6PD1), 6-phosphogluconolactonase 3 (PGL3), 6-phosphogluconate dehydrogenase 2 (PGD2), NADPH-generating dehydrogenase (pICDH), NADH kinase 3 (NADK3),

and possibly betaine aldehyde dehydrogenase (BADH) are involved in NADPH regeration.

Phylloquinone is an important cofactor for photosystem I as well as a key vitamin for humans. 1,4-dihydroxy-2-naphthoyl-CoA (DHNA-CoA) thioesterases, naphthoate synthase (NS), and o-succinylbenzoyl-CoA (OSB-CoA) ligase called AAE14 are involved in phylloquinone biosynthesis. Biotin is a cofactor required in numerous carboxylation and decarboxylation reactions. Biotin F and Dephospho-CoA kinase (CoAE), the enzyme for the last step of CoA synthesis, are involved in cofactor synthesis in peroxisomes. A member of the 4-coumarate-CoA ligase-like (4CL) family of the acyl-activating enzyme (AAE) super family, 4-coumarate: CoA ligase 5 (4CL5) participiates in the phenylalanine-related biosynthetic pathway of ubiquinone.

Peroxisomes contain several enzymes involved in the mevalonic acid (MVA) pathway. Acetoacetyl-CoA thiolase 1 (AACT1.3), 5-phosphomevalonate kinase (PMK), mevalonate 5-diphosphate decarboxylase (MVD) and isopentenyl diphosphate isomerase (IDI) are involved in MVA pathway.

In peroxisome, catabolism of polyamines, urate, pseudouridine, sulfite and methylglyoxyal occurs. Polyamines include diamine putrescine (PUT), triamine spermidine (SPD) and tetramine spermine (Spm). PAs are catabolized in peroxisomes by flavin-containing polyamine oxidases (PAOs) and copper-containing amine oxidases (CuAOs). BADH converts ABAL to 4-aminobutyrate (GABA). Purine ring catabolized sequentially by urate oxidase (uricase) and the dual-functional allantoin synthase (ALNS). Peroxisomal ATP-dependent 6-phosphofructokinase isozyme 2-type carbohydrate kinase family protein (PfkB) and Pseudouridine-50-phosphate glycosidase (IndA) catalyze pseudouridine degradation in the peroxisome. Sulfite is oxidized to sulfate by the peroxisomal sulfite oxidase (SO). Methylglyoxal is removed by the action of glyoxalase I (GLX1).

Peroxisomal 3-hydroxyisobutyryl-CoA hydrolase 1 (CHY1), CHY1H1 and CHY1H2 can catabolize valine. Peroxisomal ALS-interacting protein 1 (AIP1) and AIP3 interact with acetolactate synthase (ALS), the first enzyme in BCAA synthesis. Sarcosine oxidase (SOX) oxidizes secondary or tertiary amino acids. An O-acetylserine(thiol)lyase participates in cysteine biosynthesis. Aspartate

aminotransferase isoform 3 and cobalamin-independent met synthase 1 (ATMS1) play a role in amino acid metabolism.

Several metabolic moves across the peroxisomal membrane. PXA1/CTS/PED3, LACS6 and LACS7 participitate in peroxisomal solute import. Acyl-CoA-binding proteins (ACBPs) mediate lipid transfer across membranes. PNC (PNC1 and PNC2) and PXN proteins are peroxisomal ATP and NAD+ carriers, respectively. PXN can carries many substrates such as NAD+, NADH, AMP, ADP, CoA, and acetyl-CoA. PMP22, CDC have potential pore-forming activities, however; it needs to be elucidated.

Peroxisomes harbors signaling solutes and proteins such as Ca2+, protein kinases and phosphatases. They are involved in the regulation of various metabolic pathways in the peroxisome membrane and lumen. Calcium-dependent protein kinase 1 (CPK1), the calmodulin-like protein (CML3), peroxisomal Ca2+-dependent protein kinase 2 (CDPK2) and small CDPK-interacting protein 1 (SCP1), peroxisomal CPK33, phosphatase 2A B'h subunit, PP2A-C5, PP2A-A2, PP2A-C2, POL like phosphatase 2 (PLL2), PLL3, purple acid phosphatase 7 (PAP7), PAP5, and MAP kinase phosphatases 1 (MKP1). These elements and proteins are known as participitating peroxisomal signaling events and phosphorylation, however; it needs to be further investigated via in-depth studies.

1.8 RNA-seq Technology

RNA-seq essentially serves as a gene discovery tool for identifying specific transcription factors based on their expression profiles and the profiles of potential target genes. DNA sequencing technologies have revolutionized the transcriptomics field through massively parallel sequencing of complementary DNA (cDNAs) derived from a transcript population.

Ion Torrent, PacBio, and Illumina are commercially used deep sequencing platforms. The first step is the selection of a suitable sequencing platform. For example, the de novo transcriptome assembly process is more convenient for PacBio. Illumina Hiseq can produce short reads (50–250 bp), while PacBio generates longer reads (4200–8500 bp). Longer reads will ease the detection of alternative splice isoforms compared with short reads. Additionally, sequencing from both ends of a fragment (pair-ends) is more convenient with Illumina instruments.

Total RNA isolation and selection of messenger RNA (mRNA) is the beginning of the RNA-seq experiments. High quality RNA is an important factor in successful RNA- experiments. The integrity of RNA molecules can be measured by fluorometer machines that are commercially available instruments

After the quality assessment, mRNAs are converted to a cDNA library. Each fragment of the cDNA library is read by means of sequencing adaptors that facilitate ligation into solid platforms. Reducing the variability of reads is another challenge in RNA-seq experiments. Replications facilitate interpreting optimal aspects as well as result in reproducibility. Therefore, the problem can be overcome with duplication or triplication of the same conditions. Once cDNA is cleaved into small pieces, a single 'A' base is added to the end of each cDNA fragment to ligate it with the sequencing adapters. The adapters also enable the researcher to sequence the reads from both directions (pair-end). Pair-end reads provide more reliable data for bioinformatics analysis than single-end reads.

Reads -sequenced nucleotides- are saved in a text file called fastq, which contains millions of reads. The first step in transcriptome analysis is to check the read quality. A Phred quality score (Q score) is a measure of the quality of the identification of the nucleic acids in a DNA sequence. Certain level of Phred quality score is chosen to filter the reads. Lower than Q30 is generally used to filter sequences, which means Q score > 30 are used for further analysis. Q30 corresponds to an incorrect base call of 1 in 1000 (99.9%). Sequence length is another criterion as well as Q score. Sequences of less than fifty base length are filtered by bioinformatics tools i.e., trimmomatic, Cutadapt and FASTX-Toolkit. Once the sequences are filtered, the reads are ready for assembly. There are two types of assembly processes: de novo and reference-based assembly. Bioinformatics tools that perform reference-based assembly are divided into two groups as unspliced which is the better option for prokaryotic genomes or spliced aligners. BWA, bowtie2 and MAQ are examples of unspliced aligners, whereas Tophat, Hisat2, Star and Mapsplice are the computational tools for spliced aligners.

Reference genomes can be retrieved from various platforms such as Ensembl, the national centre of biotechnology information (NCBI) and pythozome etc…The output of the aligners is SAM file that can be compressed into a BAM file.

Unlike spliced aligners, de novo assembly assembles the reads into longer sequences using certain algorithms rather than the reference genome. Several assemblers are available such as Trinity, SOAP denovo-Trans and Trans-ABySS. These tools chop off the reads into k-mer seeds to construct a unique de Bruijn graph. Subsequently, they are parsed into consensus transcripts. Several parameters are crucial in de novo assembly. For instance, the number of contigs matching the closest related genome, mis-assembly statistics, number of hybrid transcripts, assembly statistics and contig statistics are some of them.

After the assembly process, the counting of mapped reads is the next process in RNA-seq analysis. HTSeq (https://pypi.python.org/pypi/HTSeq) which is a python standalone script, featureCounts which is an application in Subread package, Cufflinks, Stringtie are commonly used tools for counting assembled or mapped reads according to coding units such as transcripts, exons, and genes. The output of the tools is a count table, which is normalized or not depending on the tool used. Normalization is another challenge to be addressed in RNA-seq analysis for accuracy. For single-end reads, the reads per kilobase of transcript per million (RPKM) mapped reads approach is used to normalize the count table according to feature (gene) length and library-size variation. RPKM is to sum all counts of the reads in a sample and divide that number by 1.000.000, and then divide the each read count by the resulting number. This normalizes for sequencing depth, giving you reads per million. Next, the result is divided the reads per million values by the length of the gene, in kilobases, resulting in RPKM value normalized by sequencing depth and length of the gene. Like RPKM, fragments per kilobase of transcript per million (FPKM) are used in the normalization process; however, they are suitable for pair-end reads. It keeps track of fragments to eliminate counting reads twice. Additionally, Transcript per million (TPM) is like RPKM and FPKM, the order of the calculations is different, tough. In addition, edgeR or DESeq2 R Bioconductor packages used their own normalization methods: trimmed mean of M-values (TMM) and median of ratio metrics, respectively.

Raw counts can be used as input in these tools for the determination of differentially expressed genes. There are some metrics that are used to cope with biases and technical/experimental variations. Other variations or biases would be sampling time, length, GC content and so on. These variations should be minimized to enhance the reliability and accuracy of differential expression analysis. There are

several software tools, which support DEG analysis, such as NOIseq, DESeq2, baySeq and edgeR. These tools can be divided into two subgroups as parametic and non-parametric. DESeq2, edgeR and baySeq adopt the negative binomial model as the main approach while NOIseq and SAMseq adopt non-parametric methods.

To sum up, this research supports the hypothesis that the development-related genes from the juvenile phase direct the biological activities in the later stages-adult stages- and the crop yield is directly related to these early genetic arrangements. Additionally, genes that are involved in peroxisome proliferation might be a good candidate to improve the crops against abiotic stresses. The genes that show significant expression patterns to be determined from this study will be evaluated if they are genetic markers that can be used in the yield improvement studies.

2. AIM AND SCOPE OF THE STUDY

Recent advances in sequencing-based technologies and functional genomics have provided the ability to quickly decipher the genetic basis of desirable plant developmental traits. Transcriptome profiling, along with the characterization of genetic variation at the level of the entire genome, and associated QTL mapping and genome-wide association studies are major strategies to identify the genes that underlie a desirable trait in crop plants.

Crop yield improvements are either the result of increased total biomass production (larger plants tend to produce greater yields), an increased harvest index, or both. Therefore, genes that control plant morphological and physiological features that lead to higher yields are important targets of crop breeding. Considerable progress has been made in deciphering the genetic and molecular basis of the developmental processes that govern the traits, particularly in the model plant species *Arabidopsis thaliana*. However, there are only limited examples of such an understanding in crop plants.

The reproducible phenotypic outcomes of leaf development may be provided by underlied genetic networks at the leaves. Multiple pathways act in regulating the same patterning processes. The regulating networks and multiple pathways have been barely known so far in plant families even though the families have conserved gene regulation networks. There is no study that examine individual leaf differences using transcriptome approaches. Most of our understanding of the molecular regulatory pathways for early inflorescence development, especially meristem development, is derived from model plants such as arabidopsis (*Arabidopsis thaliana*) and rice (*Oryza sativa*). Additionally, our knowledge of the genetic regulations that take place throughout juveline and adult phase is still rare.

Learning the role of genes in the current system is a long and arduous task. The transcriptome will be an important tool in giving us the most likely important genes to be used as biomarker(s). Although the developmental role of these factors has only just begun to be characterized in crop plants, it is likely that they will have great potential for modulating morphological and physiological features in a desirable way in crops. This understanding can be used to improve modern wheat production. In addition, einkorn may offer valuable genetic resources.

In this study, we took advantage of publicly available RNA-seq data sets to examine the expression patterns of peroxisome metabolism genes which involved in growth and development. The outcome of bioinformatic analysis was justified by qRT-PCR in *T. monococcum*.

3. MATERIALS AND METHODS

During this Ph.D. dissertation research, the following activities were undertaken: (1) cultivation of wheat plants under different development stages and drought stress (2) determination of peroxisome abundance in the samples, (3) Acquisition of sequencing data in *Arabidopsis thaliana*, *Oryza sativa*, Zea mays and *Sorghum bicolor*, *Solanum tuberosum*, *Solanum lycopersicum* (4) mapping of the sequencing data to reference genomes and principal component analysis of the samples, (6) revealing the expression patterns of peroxisome genes (7) verification and quantification of those genes by RT q-PCR.

3.1 Cultivation of Plants and Drought Stress Treatment

Triticum monococcum ssp. *monococcum* seeds were kindly provided from United State Department of Agriculture, Agricultural Research Service. *Triticum monococcum* ssp. *monococcum* seeds were grown in a growth chamber at the Institute of Biological Chemistry at Washington State University under 60% humidity, a 16/8hr light/dark cycle, 22°C during the day and 18°C at night, and light intensity at the plant level of 400 mE. Seeds were planted into 5x5x5.25'' square bins with Sungro6 peat moss potting soil. All bins were watered daily.

Drought stress was induced by withholding watering. The volumetric water content was assessed by using a ProCheck Soil Moisture Probe with a 5TC probe (Decagon now METER environment, Pullman, WA, USA). Leaf materials were collected once soil moisture reached below 0.2%.

3.2 High Throughput RNA-seq Data Acquisition

RNA-seq data sets were downloaded from the Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/). Experiments were identified using keywords for development stages and drought stress, "development" AND "species name" [organism] and "drought stress" AND "species name" [organism], respectively. Five considerations were followed for selecting the RNA-Seq data: (1) more than one replicate per treatment (2) data for the control and drought stress for each experiment; (3) RNA was extracted from the above-ground organs; (4) drought was induced by withdrawing the watering; and (5) the plants were wild type.

3.3 Data Preprocessing

The read quality of raw sequence data was attained using FastQC software ver. 0.11.9 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). FastQC software provided a modular set of analyses which we can use to give a quick impression of whether the downloaded data sets have any problems. FastQC was run using the following; code:

*fastqc *.fastq –-outdir=./fastqc_reports*

After quality check, the raw sequences that showed low quality subjected to read trimming tool. The adapter sequence, low-quality bases, and short reads were trimmed using Trimmomatic ver. 0.39 (41). Trimmomatic was run using the following; code:

Single-end;

java -jar trimmomatic.jar SE -phred33 file_1.fastq trimmed_file_1.fastq unpaired_1.fastq LEADING:30 TRAILING:30 AVGQUAL:30 MINLEN:50

Pair-end;

java -jar trimmomatic.jar PE -phred33 file_1.fastq file_2.fastq trimmed_file_1.fastq unpaired_1.fastq trimmed_file_2.fastq unpaired_2.fastq LEADING:30 TRAILING:30 AVGQUAL:30 MINLEN:50

Fastq files were screened to the level of Q30 and length > 50 bases. Trimmed files were resubjected to FastQC to warrant whether the processed sequence qualities were enough to proceed further analysis. High quality sequence files then mapped to relevant reference genome. Hisat2 ver. 2.1.0 was used as alignment software (42). The following code was run in order to generate reference sequence index for Hisat2.

hisat2-build fastq_file_name name_of_the_index

The reads were aligned to the reference genome using the following; code:

Single-end:

Hisat2 -t –new-summary name_of_the_file -x reference_genome_index_path file_name_1.fastq -S file_name.sam

Pair-end;

Hisat2 -t –new-summary name_of_the_file -x reference_genome_index_path -1 file_name_1.fastq -2 file_name_2.fastq -S file_name.sam

Arabidopsis thaliana reads were mapped to the TAIR10 genome (www.arabidopsis.org); *Zea mays* reads were mapped to the B73v4 genome (https://www.maizegdb.org); *Oryza sativa* reads was mapped to the MSU7 genome (http://rice.plantbiology.msu.edu); *Sorghum bicolor* reads were mapped to the RTx430 or Sbv3 depending on the experiment (https://phytozome.jgi.doe.gov). *Solanum tuberosum* reads were mapped to the v4.03 (https://phytozome. jgi.doe.gov). *Solanum lycopersicum* ITAG4.0 (https://phytozome.jgi.doe.gov).

SAM alignment file conversion, sorting and preparation were performed using the Samtools software ver. 1.10 (43). Implemented code is in the following; code:

samtools sort @ 4 -o ./BAM/file_name.bam file_name.sam

All bam files are used in the quantification of transcripts. Counts of the transcript were given by featureCounts software with default parameters as follow;

*featureCounts -T 4 -p -t exon -g gene_name -a path_to_gtf_file -o file_name.txt *.bam –verbose*

The output of featureCounts, which contain actual read counts per gene (with gene ID, genomic coordinates of the gene including strand and length) was used in the differentially expressed gene analysis in R programming environment.

3.4 Differential Gene Expression Analysis

3.4.1 Normalization for Sequencing Depth Differences

Deseq2 ver. 3.10 (R Bioconductor package) was used to identify differentially expressed genes between each pair of samples (44). Raw counts were normalized by library size and fit to a negative binomial model.

We generated one specific R object called DESeqDataSet. The object stores virtually all information associated with the experiment. Experimental information was stored in a data frame with sample row names. One column of the data frame contained the experimental conditions. The genes that include less than 100 counts in the *rowSums()* function were eliminated from the DESeqDataSet R object. The size factor of the samples was calculated with *estimateSizeFactor()* function and the *sizeFactors()* function added to the data set. *counts(object, normalized = TRUE)* function was used to retrieve the normalized read counts.

3.4.2 Transformation of Normalized Read Counts

Even though differential gene expression analysis uses raw read counts, transformed data works better in many downstream analyses. The Log2

transformation is a commonly used transformation technique in R programming. *log2(normalized_counts + 1)* was used to transform size-factor normalized read counts. Normalized and log2-transformed read distributions were visualized in boxplot with the *boxplot()* function.

To check whether all variables in the data show similar variance, we plot the mean versus the standard deviation. *Vsn::meanSdPlot(log.normalized.counts)* function was used to plot for the comparison To reduce the amount of heteroskedasticity VarianceStabilizedTransformation -*vst()*- transformation function applied to shrink the variance of low level read counts. Consequently, genes with low and highly variable read counts were normalized and transformed to more homogenous read count estimates so that their variance resembled the variance observed for most of the genes. Normalized and transformed read counts, including variance shrinkage, were subject to the identification of global read count patterns.

3.4.3 Principal Component Analysis

Global read count patterns were determined by principal component analysis (PCA). 2D plots using PC1 and PC2 were constructed with the built-in plotPCA function provided by the R DESeq2 package. PCAs were examined with two different approaches. First, *DESeqDataSetFromMatrix()* was created with raw counts and then all samples were normalized and logarithm transformed with *vst()* function. The genes that show greater variance were selected with the *rowVars()* function. The top 2000 genes were used to construct the 2-D PCA. Second, Each data set were visualized separately. *DESeqDataSetFromMatrix()* were created with raw counts and then all samples were normalized and logarithm transformed with the *vst()* function. PC1 and PC2 were constructed with the built-in plotPCA function provided by the R DESeq2 package.

Sample to sample distances were calculated based on the variance stabilized transformation (vst) matrix using the dist() function in the R package. The dendrograms were generated by the hclust() function.

3.4.4 Determination of DEGs Using Deseq2 Package

A DESeqDataSet object was generated with the *DESeqDataSetFromMatrix()* function. The genes that include less than 100 counts in the *rowSums()* function were eliminated from the DESeqDataSet R object. The DEG analysis was run with *DESeq()* function. The *DESeq()* function is the wrapper

of three different Deseq2 function, *estimateSizeFactors(), estimateDispersions(), nbinomWaldTest()*. The results retrieved by results (Deseq2 object, independentFiltering = TRUE, alpha = 0.05, lfcThreshold = 0). Genes with at least a |log2-foldchange| > 0 in expression and a Benjamini-Hochberg adjusted p-value (q-value) < 0.05 were considered as differentially expressed genes (DEGs).

3.4.5 GO Term Enrichment Analysis

GO term annotations of the genes were downloaded from https://phytozome.jgi.doe.gov. The topGO R Bioconductor package was used for enrichment analysis of DEGs via Fisher method (45). First, GO slim data frame was generated in R environment. A factor vector was formed with *factor ()* function to create *new ("topGOdata", ontology='BP', allGenes = factor_vector,, nodeSize = 1, annot = annFUN.gene2GO, gene2GO = GO_id_table)* object built-in topGO package. Fisher test applied to data by run Test (object, algorithm = "classic", statistic = "fisher") function. Results were screened by *GenTable (object, classicFisher = object, orderBy = "classicFisher", ranksOf = "classicFisher", topNodes = length(score(object)))*. False discovery rate of p-values were adjusted by *p.adjust (score(object), method="BH")*.

3.4.6 Determination of Peroxisome Gene Homologs in Species

Peroxisomal genes were determined according to the study carried out by Kaur and Hu (2011) in rice and Arabidopsis (46). Orthologues Homologs of peroxisomal genes encoding peroxisomal proteins were determined in other species by using orthoFinder (https://github.com/davidemms/OrthoFinder) software.

The list of homologs was represented in Appendix 5. The homologs were categorized by the information obtained from (47). The homologs were grouped according to 20 metabolic events. We also categorized the genes that have two or more roles in a cell into a new group. Heatmaps were generated with the *pheatmap()* function (NMF ver. 0.17.6) using vst values with a z-score transformation.

3.5 Quantitative Reverse Transcriptase PCR Analysis

Total RNA was extracted from three biological replicates using RNeasy plant kit (Qiagen). cDNA was synthesized using Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). Each qPCR primer was designed to target all three homeologs. The primers are listed in Table 3.1.

Table 3.1. Primers used in the growth and development study.

Name	Sequence (5'->3')	Length	Tm	GC%
BADH_035342_244_F	AACAGACGTTAGCGCACATCA	21	60.60	47.62
BADH_035342_244_R	CTCCCCACGGAGTTGAACC	19	60.00	63.16
BADH_ctrl_373_R	TACAGCTTGGATGGGCGCTG	20	62.87	60.00
BADH_ctrl_634_F	TGCAGTGCAACTTCTCGTCTAC	22	60.86	50.00
CAT1_013376_219_F	CGAGAAGATGGTGATCGAGAA	21	64.3	47.60
CAT1_013376_219_R	ACTGACCTTACATGCTCGGC	20	60.11	55.00
PEX11c_019002_152_F	GGCAGGCAGATCACTTGGAT	20	60.11	55.00
PEX11c_019002_152_R	CGAGCCTCGACATAGTCGTC	20	60.04	60.00
RLI_A_F	TTGAGCAACTCATGGACCAG	20	64.00	50.00
RLI_A_R	GCTTTCCAAGGCACAAACAT	20	63.90	45.00

qRT-PCR reactions were performed using Fast SYBR™ GreenMaster Mix (Thermo Fisher Scientific) in 96-wells plates on ViiA 7 Real-Time PCR System with default ViiA™ 7 SYBR conditions. Reactions were replicated 3 times and analyzed in QuantStudio™ Real-Time PCR Software v1.3., transcription levels were normalized to housekeeping gene RNase L inhibitor-like protein (Giménez, Pistón & Atienza 2011).

4. RESULTS

4.1 Peroxisomes in Growth and Development

In this section, I explored the expression patterns of peroxisome metabolism genes. The results of bioinformatics analysis were interpreted with the association of growth and development, and chemical reactions that take place in peroxisome lumen. The gene(s) that shows interesting expression pattern and the function of its target molecule(s) is discussed in this section.

4.1.1 Data Sets Acquisition for Growth and Development

Three considerations were followed for selecting the RNA-Seq data set: (1) more than one replicate per experiment (2) RNA was extracted from the above-ground organs (3) the plants were wild type. In GEO dataset collection, we found 33 different samples that compatible with our research aim. The information about the studies is present in Table 4.1.

Table 4.1. NCBI GEO/SRA RNA-Seq data used in the growth and development study.

SRR_number	Mbyte	Experiment	Tissue	Paper
SRR940249	613	SK064_S	Eleventh_Leaf_V9_R1	http://bit.ly/2r0L8Di
SRR940250	560	SK065_S	Eleventh_Leaf_V9_R2	http://bit.ly/2r0L8Di
SRR940252	880	SK067_S	Thirteenth_Leaf_V9_R1	http://bit.ly/2r0L8Di
SRR940253	680	SK068_S	Thirteenth_Leaf_V9_R2	http://bit.ly/2r0L8Di
SRR940254	603	SK069_S	Thirteenth_Leaf_V9_R3	http://bit.ly/2r0L8Di
SRR940260	539	SK083_S	Thirteenth_Leaf_VT_R2	http://bit.ly/2r0L8Di
SRR940274	705	SK084_S	Thirteenth_Leaf_VT_R3	http://bit.ly/2r0L8Di
SRR940276	748	SK097_S	Thirteenth_Leaf_R2_R1	http://bit.ly/2r0L8Di
SRR940278	914	SK098_S	Thirteenth_Leaf_R2_R2	http://bit.ly/2r0L8Di
SRR940237	1499	SK061_S	Eighth_Leaf_V9_R1	http://bit.ly/2r0L8Di
SRR940238	540	SK062_S	Eighth_Leaf_V9_R2	http://bit.ly/2r0L8Di
SRR1619633	1172	PP-13_S	Leaf_6DAP_R1	http://bit.ly/2r0L8Di
SRR1619661	1014	PP-30_S	Leaf_18DAP_R2	http://bit.ly/2r0L8Di
SRR1619634	1571	PP-14_S	Leaf_6DAP_R2	http://bit.ly/2r0L8Di
SRR1619636	3795	PP-21_S	Leaf_12DAP_R1	http://bit.ly/2r0L8Di
SRR1619637	797	PP-22_S	Leaf_12DAP_R2	http://bit.ly/2r0L8Di
SRR1619660	1034	PP-29_S	Leaf_18DAP_R1	http://bit.ly/2r0L8Di
SRR1619669	563	PP-5_S	Leaf_0DAP_R1	http://bit.ly/2r0L8Di
SRR1619670	784	PP-6_S	Leaf_0DAP_R2	http://bit.ly/2r0L8Di
SRR1626426	763	PP-38_S	Leaf_24DAP_R2	http://bit.ly/2r0L8Di

Table 4.1. Continued

SRR_number	*Mbyte*	*Expr.*	*Tissue*	*Paper*
SRR12416727	3500	-	mRNA_Blade_S1_rep1	https://doi.org/10.1016/j.isci.2020.101489
SRR12416728	4000	-	mRNA_Blade_S1_rep2	https://doi.org/10.1016/j.isci.2020.101489
SRR12416729	4200	-	mRNA_Blade_S2_rep1	https://doi.org/10.1016/j.isci.2020.101489
SRR12416710	3500	-	mRNA_Blade_S2_rep2	https://doi.org/10.1016/j.isci.2020.101489
SRR12416711	3900	-	mRNA_Blade_S3_rep1	https://doi.org/10.1016/j.isci.2020.101489
SRR12416712	3300	-	mRNA_Blade_S3_rep2	https://doi.org/10.1016/j.isci.2020.101489
SRR12416713	3700	-	mRNA_Blade_S4_rep1	https://doi.org/10.1016/j.isci.2020.101489
SRR12416714	3300	-	mRNA_Blade_S4_rep2	https://doi.org/10.1016/j.isci.2020.101489
SRR12416715	3400	-	mRNA_Blade_S5_rep1	https://doi.org/10.1016/j.isci.2020.101489
SRR12416716	4200	-	mRNA_Blade_S5_rep2	https://doi.org/10.1016/j.isci.2020.101489

Maize inbred B73 was used for constructing the RNA-seq libraries in the SRP014652 study. The main goal of the study that is named with SRP014652 was to construct a maize gene expression atlas based on RNA sequencing and use it to explore root development. Many tissue samples were used in SRP014652 study, yet i selected the tissues that might answer our questions to understand the debated hypothesis. Even though SRP014652 study contains multiple tissue types, only whole leaf tissue samples are used in this study Table 4.1. GSE155932 is the study that was carry out on *Oryza sativa.* The main goal of the study that named with GSE155932 was to understand the expression profiling of mRNAs and non-coding RNAs by high throughput sequencing. The GSE155932 study is presented in two superseries that are composed of laminal joint and blade tissues. Only blade tissue samples were used in this study.

4.1.2 Alignment of Reads of RNA-seq Experiments

Clean single end reads of *Zea mays* data set were mapped to the respective reference genome of the species with the alignment efficiency ranging from 70% to 89% (Table 4.2). These outcomes mean, the throughput and sequencing quality were sufficiently high to warrant further analysis in *Zea mays*.

Table 4.2. Single-end mapping results of *Zea mays* data set.

GEO/SRA CODE	SAMPLE	TOTAL PAIRS	ALIGNED 1 TIME	ALIGNED >1 TIME	OVERALL ALIGNMENT RATE
SRP014652	SRR940249	8700245	7155116 (82.24%)	321528 (3.70%)	85.94%
SRP014652	SRR940250	7939105	6468309 (81.47%)	317164 (3.99%)	85.47%
SRP014652	SRR940252	12363437	9577764 (77.47%)	454012 (3.67%)	81.14%
SRP014652	SRR940253	9592796	6967003 (72.63%)	309394 (3.23%)	75.85%
SRP014652	SRR940254	8498692	6218141 (73.17%)	260414 (3.06%)	76.23%
SRP014652	SRR940260	7610859	5560312 (73.06%)	262243 (3.45%)	76.50%
SRP014652	SRR940274	9939688	7072752 (71.16%)	605094 (6.09%)	77.24%
SRP014652	SRR940276	10640131	8221004 (77.26%)	388786 (3.65%)	80.92%
SRP014652	SRR940278	13097581	9889079 (75.50%)	447064 (3.41%)	78.92%
SRP014652	SRR940237	21370924	15732165 (73.61%)	715712 (3.35%)	76.96%
SRP014652	SRR1619668	31520063	26099314 (82.80%)	1189054 (3.77%)	86.57%
SRP014652	SRR940238	7573491	5707180 (75.36%)	387875 (5.12%)	80.48%
SRP014652	SRR1619633	16696527	13845666 (82.93%)	739502 (4.43%)	87.35%
SRP014652	SRR1619634	22282272	18355913 (82.38%)	1019526 (4.58%)	86.95%
SRP014652	SRR1619636	62489080	52362828 (83.80%)	2838099 (4.54%)	88.34%
SRP014652	SRR1619637	11236192	9431938 (83.94%)	522466 (4.65%)	88.59%
SRP014652	SRR1619660	14559869	11921114 (81.88%)	688661 (4.73%)	86.61%
SRP014652	SRR1619661	14351598	11545471 (80.45%)	601482 (4.19%)	84.64%
SRP014652	SRR1619664	17040548	12415967 (72.86%)	982135 (5.76%)	78.62%
SRP014652	SRR1619667	30686002	25326511 (82.53%)	1192049 (3.88%)	86.42%

Clean pair end reads of *Oryza sativa* data set were mapped to the respective reference genome of the species with the alignment efficiency ranging from 70% to 89% (Table 4.3). These outcomes mean, the throughput and sequencing quality were sufficiently high to warrant further analysis in *Oryza sativa*.

Table 4.3. Pair-end mapping results of *Oryza sativa* data set.

GEO/SRA CODE	SAMPLE	TOTAL PAIRS	ALIGNED CONCORDANTLY 1 TIME	OVERALL ALIGNMENT RATE
GSE155932	SRR12416727	28059822	24914368 (88.79%)	96.03%
GSE155932	SRR12416728	31257854	26462544 (84.66%)	92.89%
GSE155932	SRR12416729	33361334	30153510 (90.38%)	96.51%
GSE155932	SRR12416710	27595610	24586251 (89.09%)	95.90%
GSE155932	SRR12416711	31157819	27662246 (88.78%)	95.58%
GSE155932	SRR12416712	26045821	23160737 (88.92%)	96.20%
GSE155932	SRR12416713	29300337	24009697 (81.94%)	90.76%
GSE155932	SRR12416714	25769873	23140840 (89.80%)	96.64%
GSE155932	SRR12416715	27339070	23845292 (87.22%)	93.79%
GSE155932	SRR12416716	33694858	28877535 (85.70%)	93.30%

4.1.3 Normalization and Transformation of Raw Counts in R Environment

Our basic task was the determination of differentially expressed genes that were quantified from the read of RNA-seq experiments. The package DESeq2 offers an opportunity to test for differential expression. As input, a matrix of integer values that are obtained from the FeatureCount application is used in the analysis of differentially expressed genes. Raw counts of the matrix that belong to the i-th row and the j-th column represent mRNA counts of which can be assigned to gene i in sample j. While the package DESeq2 used for differential gene expression testing operates on the raw count integer values, most downstream analyses give better accurate results if the read counts are transformed to the log scale. Raw counts and log-scale normalized counts are represented in Appendix 2.

In theory, we assume that the variables are homoscedastic in all RNA-seq experiments. The mean vs. standard deviation data was plotted to visualize if differences among the sizes of the individual observations show homoscedastic behavior. Appendix 2 shows a comparison of log2- and vst-transformed read counts.

Several means of shrinking the variance are included in the package DESeq2. Reducing the amount of heteroskedasticity of low read counts is done by using the dispersion-mean trend. Consequently, genes with low and highly variable read counts were assigned the read count estimates that are more homogenous, and their variance was transformed into new estimates that resembled the variance observed for most of the genes. Vst transformed values were used to assess the similarity of biological replicates (Appendix 3).

4.1.4 Exploring Global Read Count Patterns

RNA-Seq data from three different studies that were acquired from two different species were evaluated to identify developmental transcriptomes. To obtain a general overview of transcription patterns of relevant species, we examined the principal components of the samples using the top 2000 genes that show the greatest expression variance (Figure 4.1).

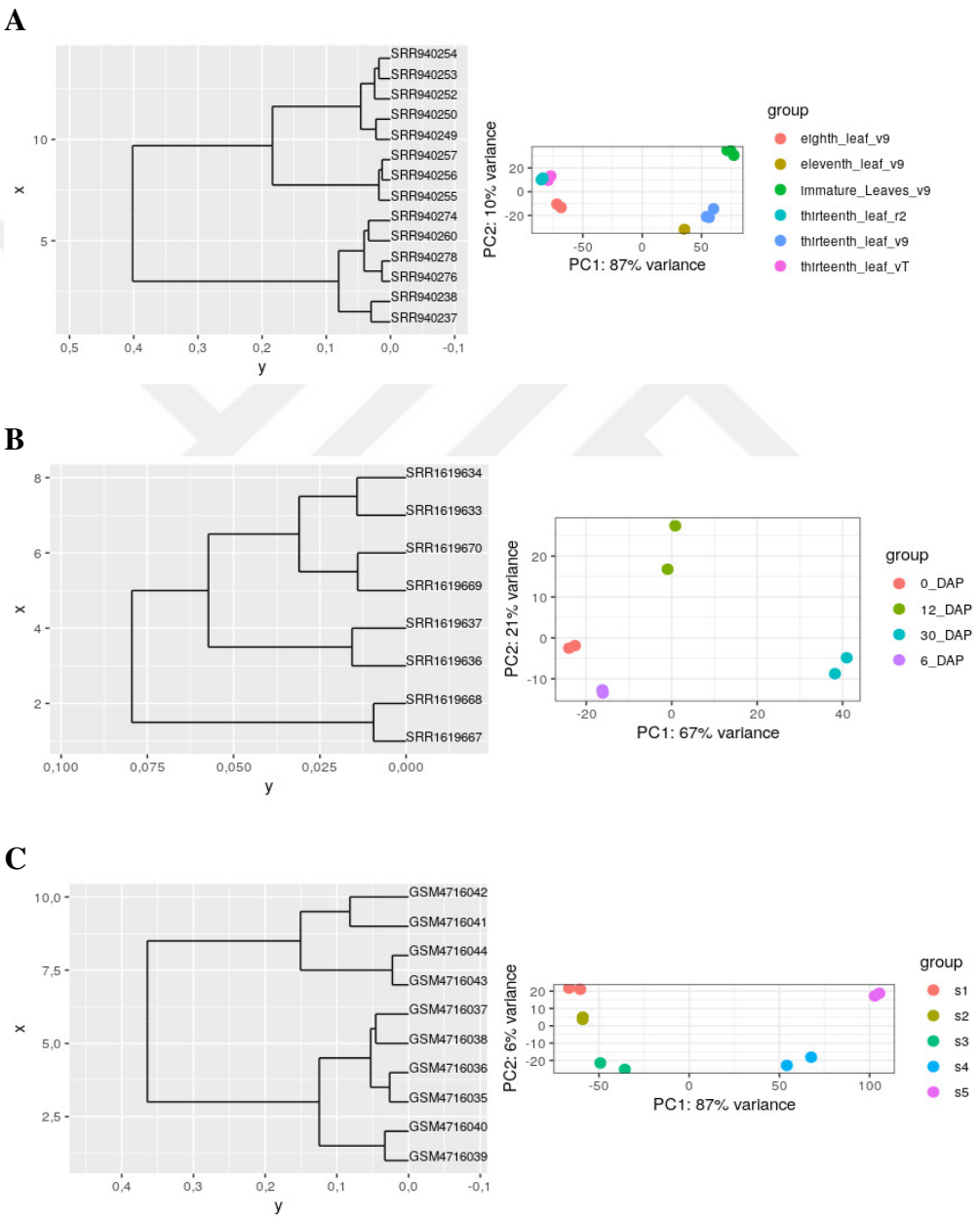


Figure 4.1. Hrierchical and principal component analysis of the data sets that are used in growth and development. A) *Zea mays;* immature and mature leaf comparison at the same and different development stages B) *Zea mays;* Time series comparison in reproductive stages C) Early seedling comparison in *Oryza sativa*

PC-1 and PC-2 represented over 85% of the total variance in maize samples; however, PC-1 and PC-2 represented over 65% of the total variance in rice samples. The Pearson correlation coefficient, r, is a measure of the strength of the linear relationship between two variables and is often used to assess the similarity of RNA-seq samples in a pairwise fashion. Pairwise correlations can be calculated with hierarchical clustering. Our hierarchical clustering verified the outcome of the PCA analysis across species (Figure 4.1). No variability was observed among replicates. Samples show greater variability among species, whereas transcription patterns of genes show little variability between samples. The two principal components explained most of the variability. The samples were separated according to the treatments in the first PC.

4.1.5 Differentially Expressed Genes

In *Zea mays*, 14 samples from the SRP014652 dataset were examined for DEG analysis. 46430 genes were counted in the SRP014652 dataset before we filtered the genes that had low counts. After filtering the genes that had less than 100 counts across samples, we kept 22024 genes according to the eighth and immature leaf comparison. 14700 genes were differentially expressed. 7176 (33%) genes were upregulated, whereas 7524 (34%) genes were downregulated. A similar approach was used to assess the expression of the genes that were compared in the eleventh and immature leaf samples. 21318 genes were kept, and 11632 genes were differentially expressed; 6024 (28%) genes were upregulated, whereas 5608 (26%) genes were downregulated in the comparison of eleventh and immature leaf. 21760 genes were kept in the comparison of thirteenth and immature leaves. 10743 genes were differentially expressed. 5580 (26%) genes were upregulated, while 5163 (24%) genes were downregulated. 20164 genes were counted in the VT vs V9. 12288 genes were differentially expressed. 5996 (30%) genes were upregulated, while 6292 (31%) genes were downregulated. 20783 genes were counted in the R2 vs V9. 12288 genes were differentially expressed. 6373 (31%) genes were upregulated, 6746 (32%) genes were downregulated. 17446 genes were counted in the R2 vs VT. 1250 genes were differentially expressed. 675 (3.9%) genes were upregulated, while 575 (3.3%) genes were downregulated. 19578 genes were counted in the eleventh vs eighth leaf. 7674 genes were differentially expressed. 3591 (18%) genes were upregulated, while 4083 (21%) genes were downregulated. 20180 genes were counted in the thirteenth vs eighth leaf. 9020 genes were

differentially expressed. 4660 (23%) genes were upregulated, 4360 (22%) genes were downregulated. 19436 genes were counted in the thirteenth vs eleventh leaf. 843 genes were differentially expressed. 220 (1.1%) genes were upregulated, 623 (3.2%) genes were downregulated. 18614 genes were counted in 6DPA vs 0DPA. 1759 genes were differentially expressed. 561 (3%) genes were upregulated, 1198 (6.4%) genes were downregulated. 20856 genes were counted in 12DPA vs 6DPA. 5563 genes were differentially expressed. 2489 (12%) genes were upregulated, and 3074 (15%) genes were downregulated. 21820 genes were counted in 30DPA vs 12DPA. 6721 genes were differentially expressed. 3140 (14%) genes were upregulated, while 3581 (16%) genes were downregulated.

In *Oryza sativa,* ten samples of the GSE155932 dataset were examined for DEG analysis. 55986 genes were counted in the GSE155932 dataset before we filtered the genes that have a low count. After filtering the genes that have less than 100 counts across samples, we kept 21120 genes according to s2 and s1 comparison. 1017 genes were differentially expressed; 618 (2.9%) genes were upregulated, whereas 399 (1.9%) genes were downregulated in the s2 and s1 comparison. A similar approach was used to assess the expression of the genes that compared in s3 vs s2 evaluation. 21262 genes were evaluated for the further analysis. 2502 genes were differentially expressed; 1449 (6.8 %) genes were upregulated, whereas 1053 (5%) genes were downregulated in the comparison of s3 vs s2. 20907 genes were counted in s4 vs s3 comparison after filtering; 5788 genes were differentially expressed. 2501 (12%) were upregulated, whereas 3287 (16%) downregulated in comparison of s4 vs s3. 20081 genes were evaluated given the filtering outcome. 1721 (8.6%) were upregulated, while 2735 genes downregulated in s5 vs s4 comparison.

4.1.6 GO Enrichment Analysis

The Gene Ontology allows users to describe a gene/gene product in detail considering three main aspects: its molecular function, the biological process in which it participates, and its cellular location. To the determine relationship of the differentially expressed genes to growth and development, we carried out GO enrichment analysis using topGO package in R so that we can overview the biological processes of differentially expressed genes. The enrichment of first three GO terms in each study is presented in Table 4.4.

Table 4.4. GO enrichment analysis.

STUDY	SPECIES	GO_ID	GENE RATIO	LOGFC	Q-VALUE
IMMATTURE VS EIGHTH LEAF	Maize	GO:0006586	9/14	-1.4/2.84	6,053860e-06
	Maize	GO:0009144	68/131	-7.13/8.29	1,428776e-04
	Maize	GO:0006090	46/77	-7.13/10.83	2,155765e-04
IMMATTURE VS ELEVEN LEAF	Maize	GO:0010557	6/6	-1.36/1.35	0,0000261746
	Maize	GO:0009144	60/131	-3.8/6.29	0,0001410707
	Maize	GO:0044743	8/14	-1.23/0.82	0,0001991505
IMMATTURE VS THIRTEENTH LEAF	Maize	GO:0019362	35/88	-3.36/4.59	4,259303e-06
	Maize	GO:0055086	110/266	-3.36/4.59	1,666720e-04
	Maize	GO:0019441	3/4	1.09/1.99	1,905828e-04
VT VS V9	Maize	GO:0034250	3/4	-2.94/0.61	0,001075112
	Maize	GO:0042402	3/4	1.56/2.1	0,001075112
	Maize	GO:0016579	17/42	-1.3/1.26	0,001075112
R2 VS V9	Maize	GO:0007219	2/2	-1/-0.86	3,044603e-05
	Maize	GO:0008593	2/2	-1/-0.86	3,044603e-05
	Maize	GO:0006370	3/3	-1/-0.86	8,412483e-05
VT VS R2	Maize	GO:1903825	4/52	0.8/2.01	0,0006706501
	Maize	GO:0019359	8/81	-0.82/0.58	0,0467806130
	Maize	GO:0019363	8/81	-0.82/0.58	0,0467806130
8^{TH} VS 11^{TH} LEAF	Maize	GO:0006535	4/4	0.78/6.01	1,675881e-08
	Maize	GO:0015780	1/2	-1.19	1,446930e-05
	Maize	GO:0005984	23/58	-3.92/5.1	2,616074e-03
8^{TH} VS 13^{TH} LEAF	Maize	GO:0006544	2/5	0.82/1.09	2,777793e-12
	Maize	GO:0006662	21/44	-6.13/2.34	8,752053e-07
	Maize	GO:1901615	13/38	-3.77/2.78	9,456633e-05
11^{TH} VS 13^{TH} LEAF	Maize	GO:0009108	7/135	-2.84/-0.7	4,867746e-29
	Maize	GO:0006099	2/17	-1.04/-1.05	3,034773e-17
	Maize	GO:0046036	1/8	-1.2	3,784271e-16
6 VS 0	Maize	GO:0007018	3/102	-0.89/0.72	0,01698319
12 VS 6	Maize	GO:0009451	11/52	-2.02/0.89	1,135620e-23
	Maize	GO:0006595	2/9	-1.38/-0.43	1,547760e-06
	Maize	GO:0046937	1/2	-0.93	3,097901e-06
30 VS 12	Maize	GO:0051704	25/73	-1.4/3.19	4,503328e-42
	Maize	GO:0006733	28/89	-1.65/2.55	2,611628e-38
	Maize	GO:0006213	5/10	1.16/1.82	1,190568e-04
S1 VS S2	Rice	GO:0046483	184/3898	-4.5/3.71	4,086014e-20
	Rice	GO:0043604	6/664	-2.14/1.2	4,086014e-20
	Rice	GO:0006950	147/3618	-4.14/4.42	1,705938e-07
S2 VS S3	Rice	GO:0019748	62/488	-4.38/3.95	1,407240e-23
	Rice	GO:0009991	12/304	3.03/3.73	1,407240e-23
	Rice	GO:0034641	408/4402	-6.79/5.05	1,961669e-19
S3 VS S4	Rice	GO:0043604	190/664	-8.35/3.99	3,980084e-08
	Rice	GO:0044271	190/664	-8.35/3.99	3,980084e-08
	Rice	GO:0016049	114/465	-9.14/6.16	3,980084e-08
S4 VS S5	Rice	GO:0050896	959/5854	-13.12/12.2	3,833150e-07
	Rice	GO:0009059	102/664	-7.84/1.91	4,402077e-05
	Rice	GO:0009987	1942/13332	-13.12/12.2	5,620946e-05

The number of enriched GO term categories was between 9 and 50 among the studies, expect one comparison of the early seedling developmental stage. Forty-six GO terms were enriched in immature vs eighth leaf comparison. Fifty GO terms were enriched in immature vs eleven leaf comparison. Forty-three GO terms

were enriched in immature vs thirteenth leaf comparison. Ten GO terms were enriched in V9 vs VT stage comparison performed in thirteenth leaf of maize. Twenty-four GO terms were enriched in V9 vs R2 comparison. Twelve GO terms were enriched in VT vs R2 comparison. Twelve GO terms were enriched in 8th and 11th leaf comparison. Twenty GO terms were enriched in 8th and 13th leaf comparison. Five GO terms were enriched in 11th and 13th leaf comparison. One GO terms were enriched in 0 vs 6 days old seedling transcriptome comparison. Eighteen GO terms were enriched in 6 vs 12 days comparison. Sixty-one GO terms were enriched in 12 vs 30 days comparison. Thirteen genes were enriched in s1 vs s2 seedling comparison in rice. Fourteen GO terms were enriched in s2 vs s3. Thirty-seven GO terms were enriched in s3 vs s4. Nine GO terms were enriched in s4 vs s5.

4.1.7 Expression Patterns of Peroxisome Biogenesis and Metabolism Genes

We used the previous study published by Kaur and Hu (2011) to identify peroxisome-related genes (46). They compiled a list containing many of genes encoding peroxisome lumen and membrane proteins. The expression patterns of peroxisome genes were visualized with those locus IDs. The number of differentially expressed peroxisome lumen and membrane genes was shown in Figure 4.1. Most of the differentially expressed genes were observed in immature and eight leaf transcriptome comparison.

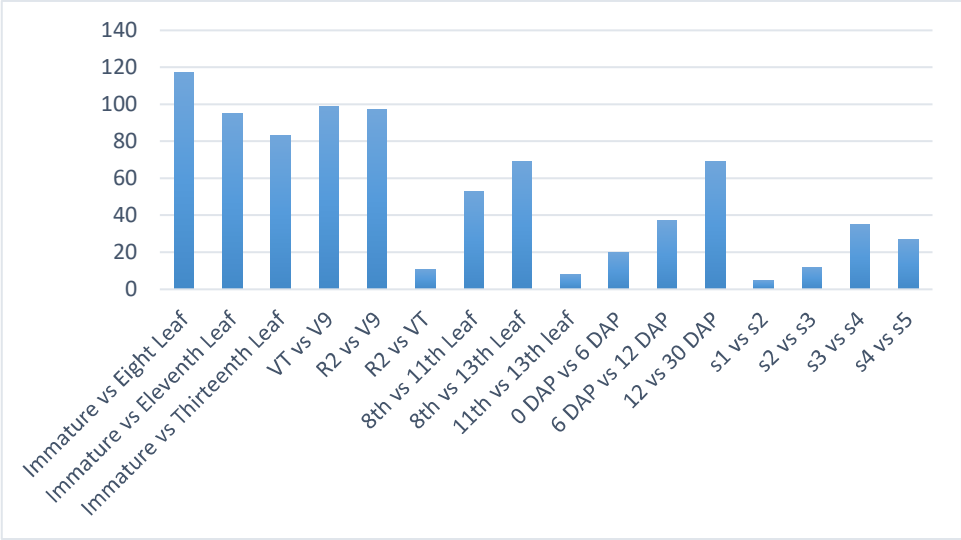
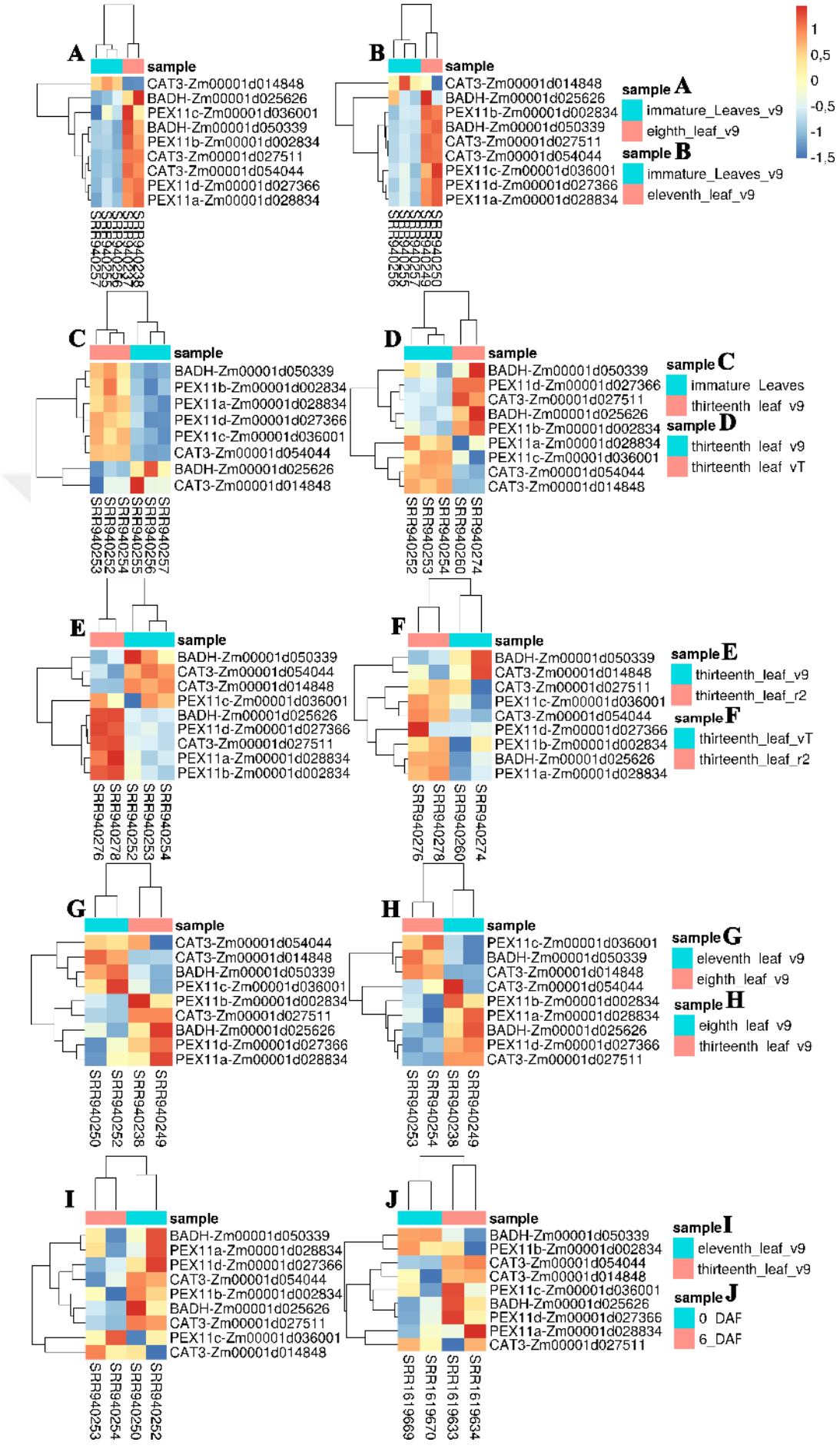


Figure 4.2. The number of differentially expressed peroxisome lumen and membrane genes.

As represented 16 different data sets in one heatmap was impractical, we prepared a representative heatmap for each species and study. Figure 4.3 shows expression patterns of some important genes that differentially expressed in each comparison. Whole gene expression heatmaps are present in Appendix 1.


A
sample
CAT3-Zm00001d014848
BADH-Zm00001d025626
PEX11c-Zm00001d036001
BADH-Zm00001d050339
PEX11b-Zm00001d002834
CAT3-Zm00001d027511
CAT3-Zm00001d054044
PEX11d-Zm00001d027366
PEX11a-Zm00001d028834
B
sample
CAT3-Zm00001d014848
BADH-Zm00001d025626
PEX11b-Zm00001d002834
BADH-Zm00001d050339
CAT3-Zm00001d027511
CAT3-Zm00001d054044
PEX11c-Zm00001d036001
PEX11d-Zm00001d027366
PEX11a-Zm00001d028834
sample A
immature_Leaves_v9
eighth_leaf_v9
sample B
immature_Leaves_v9
eleventh_leaf_v9
1
0,5
0
-0,5
-1
-1,5
C
sample
BADH-Zm00001d050339
PEX11b-Zm00001d002834
PEX11a-Zm00001d028834
PEX11d-Zm00001d027366
PEX11c-Zm00001d036001
CAT3-Zm00001d054044
BADH-Zm00001d025626
CAT3-Zm00001d014848
SRR940253
SRR940252
SRR940254
SRR940255
SRR940256
SRR940257
D
sample
BADH-Zm00001d050339
PEX11d-Zm00001d027366
CAT3-Zm00001d027511
BADH-Zm00001d025626
PEX11b-Zm00001d002834
PEX11a-Zm00001d028834
PEX11c-Zm00001d036001
CAT3-Zm00001d054044
CAT3-Zm00001d014848
SRR940252
SRR940253
SRR940254
SRR940260
SRR940274
sample C
immature Leaves
thirteenth leaf v9
sample D
thirteenth leaf v9
thirteenth leaf vT
E
sample
BADH-Zm00001d050339
CAT3-Zm00001d054044
CAT3-Zm00001d014848
PEX11c-Zm00001d036001
BADH-Zm00001d025626
PEX11d-Zm00001d027366
CAT3-Zm00001d027511
PEX11a-Zm00001d028834
PEX11b-Zm00001d002834
SRR940276
SRR940278
SRR940252
SRR940253
SRR940254
F
sample
BADH-Zm00001d050339
CAT3-Zm00001d014848
CAT3-Zm00001d027511
PEX11c-Zm00001d036001
CAT3-Zm00001d054044
PEX11d-Zm00001d027366
PEX11b-Zm00001d002834
BADH-Zm00001d025626
PEX11a-Zm00001d028834
SRR940276
SRR940278
SRR940260
SRR940274
sample E
thirteenth_leaf_v9
thirteenth_leaf_r2
sample F
thirteenth_leaf_vT
thirteenth_leaf_r2
G
sample
CAT3-Zm00001d054044
CAT3-Zm00001d014848
BADH-Zm00001d050339
PEX11c-Zm00001d036001
PEX11b-Zm00001d002834
CAT3-Zm00001d027511
BADH-Zm00001d025626
PEX11d-Zm00001d027366
PEX11a-Zm00001d028834
SRR940250
SRR940252
SRR940238
SRR940249
H
sample
PEX11c-Zm00001d036001
BADH-Zm00001d050339
CAT3-Zm00001d014848
CAT3-Zm00001d054044
PEX11b-Zm00001d002834
PEX11a-Zm00001d028834
BADH-Zm00001d025626
PEX11d-Zm00001d027366
CAT3-Zm00001d027511
SRR940253
SRR940254
SRR940238
SRR940249
sample G
eleventh leaf v9
eighth leaf v9
sample H
eighth leaf v9
thirteenth leaf v9
I
sample
BADH-Zm00001d050339
PEX11a-Zm00001d028834
PEX11d-Zm00001d027366
CAT3-Zm00001d054044
PEX11b-Zm00001d002834
BADH-Zm00001d025626
CAT3-Zm00001d027511
PEX11c-Zm00001d036001
CAT3-Zm00001d014848
SRR940253
SRR940254
SRR940250
SRR940252
J
sample
BADH-Zm00001d050339
PEX11b-Zm00001d002834
CAT3-Zm00001d054044
CAT3-Zm00001d014848
PEX11c-Zm00001d036001
BADH-Zm00001d025626
PEX11d-Zm00001d027366
PEX11a-Zm00001d028834
CAT3-Zm00001d027511
SRR1619669
SRR1619670
SRR1619633
SRR1619634
sample I
eleventh_leaf_v9
thirteenth_leaf_v9
sample J
0 DAF
6_DAF

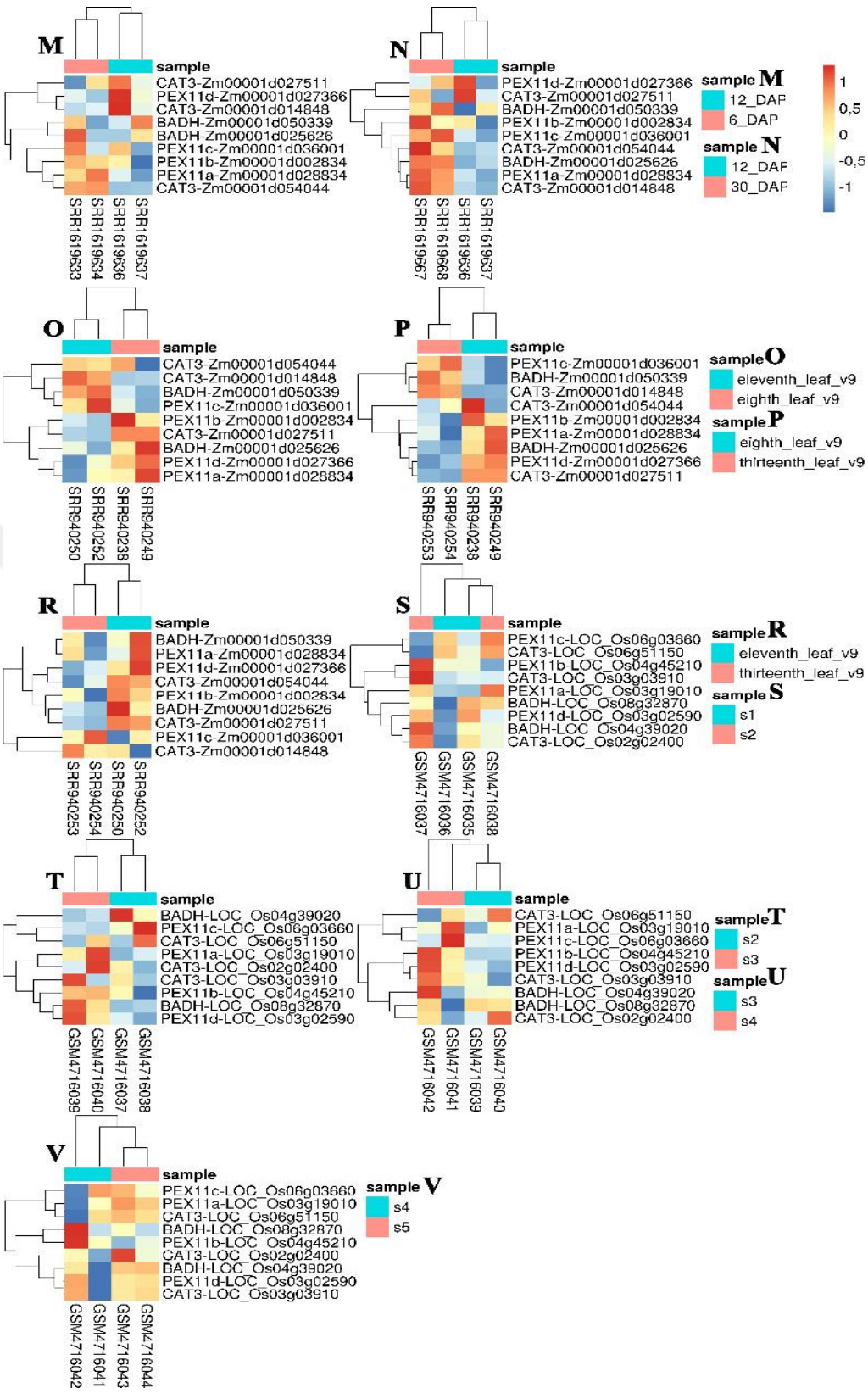


Figure 4.3. Representative figure of DEGs in growth and development study. The figure was generated with the R pheatmap package using VarianceStabilizedTransformation -vst() function- built in DESeq2 package. Vst values were represented based on the z-score transformation. Samples were clustered according to Pearson correlation analysis. Loci name of orthologs that were mapped to peroxisome genes indicated beside gene names.

Peroxisome genes are categorized in twenty-one different physiological metabolisms. Table 4.5 includes differentially expressed genes and shows how many DEGs are upregulated in the category. In the comparison of immature and adult leaves which are sampled in the same developmental stage, GGT1 and PM16 are upregulated in amino acid metabolism. AAE1, AAE12, AAE14, AAE17 and 4CL1 are upregulated in the biosynthesis of phylloquinone, biotin, CoA and ubiquinone. SO, pxPfkB and GLX1 are upregulated in catabolism of polyamines, urate, pseudouridine, sulfite and methylglyoxyal. PEX14, which is a peroxisomal import protein, is categorized in docking complex which recruit PEX5 to import peroxisomal proteins to peroxisome lumen. PEX14 was constantly upregulated in the adult leaves relative to immature leaves, except in thirteenth leaf. Additionally, Fatty acid breakdown and phytohormone biosynthesis category has consistently upregulated genes: ACX5 and KAT1. Glyoxylate cycle and acetate-malate shunt which enables an organism to use substrates that enter central carbon metabolism at the level of acetyl-CoA as the sole carbon source. CSY3 was upregulated in all three comparisons, while 6PGL is upregulated in two comparisons. NADPH regeneration and catabolism of polyamines, urate, pseudouridine, sulfite and methylglyoxyal category has only one consistently upregulated gene: BADH. PEX19a is upregulated in peroxisomal import protein. There is no consistently upregulated gene in peroxisomal solute transporters and, peroxisomal solute transporters, biosynthesis of phylloquinone, biotin, CoA, ubiquinone and phytohormone biosynthesis category. PEX11a, PEX11b and PEX11c are consistently upregulated among peroxisome fission genes. LON2 and DEG15 are upregulated in all three comparisons in terms of quality control and proteome remodelling. PEX1, PEX12 and PEX22 are upregulated in all three comparisons in the recycling of PEX5. SGAT1, MDH1, HOAX1 and HPR are upregulated in photorespiration. CAT3 is upregulated in Photorespiration and ROS/RNS metabolism and autophagy. sT1, IBR1 and sT5 are upregulated in phytohormone biosynthesis. CSD3, MDAR1, APX3, MDAR4, GR1 and DHAR was upregulated in ROS/RNS metabolism. OPR3 was upregulated in signaling Ca2+ and protein phosphorylation.

In the comparison of the genes that are upregulated in the same leaf, but at the different development stage in terms of peroxisome, SOX is upregulated in amino acid metabolism. AAE7 and AAE12 are upregulated in the biosynthesis of

phylloquinone, biotin, CoA and ubiquinone. SGAT1 and HOAX1 are upregulated in photorespiration. CAT3 is upregulated in photorespiration and ROS/RNS metabolism and autophagy. In the adult leaf comparison, AGT2 is upregulated in amino acid metabolism. Additionally, GPK1 is upregulated in glyoxylate cycle and acetate-malate shunt.

Table 4.5. 1-Amino acid metabolism, 2-Biosynthesis of phylloquinone, biotin, CoA and ubiquinone, 3-Catabolism of polyamines, urate, pseudouridine, sulfite and methylglyoxyal, 4-Docking complex, 5-Fatty acid breakdown, 6-Fatty acid breakdown and phytohormone biosynthesis, 7-Glyoxylate cycle and acetate-malate shunt, 8-NADPH regeneration, 9-NADPH regeneration and Catabolism of polyamines, urate, pseudouridine, sulfite and methylglyoxyal, 10-Peroxisomal import protein, 11-Peroxisomal solute transporters, 12-Peroxisomal solute transporters and Biosynthesis of phylloquinone, biotin, CoA and ubiquinone and Phytohormone biosynthesis, 13-Peroxisome fission gene, 14-Photorespiration, 15-Photorespiration and ROS/RNS metabolism and autophagy, 16-Phytohormone biosynthesis, 17-Quality control and proteome remodeling, 18-Recycling of PEX5, 19-Recycling of PEX5 and pexophagy and phytohormone biosynthesis, 20-ROS/RNS metabolism, 21-Signaling Ca2+ and protein phosphorylation.

	Zea mays												*Oryza sativa*			
	Juvenile vs Adult			Stage comparison in the same leaves			Adult leaf comparison			Reproductive stage comparison			Early seedling comparison			
	11_im_v9	8_im_v9	13_im_v9	vT_v9_13	R2_V9_13	R2_vT_13	11_8_v9	13_11_v9	13_8_v9	6_vs_0 DPA	12_vs_6 DPA	30_vs_12 DPA	S2_vs_S1	S3_vs_S2	S4_vs_S3	S5_vs_S4
1	9-6	9-6	8-5	10-5	8-3	2-2	6-4	3-1	8-5	2-2	4-4	5-5	-	-	4-1	4-1
2	11-8	10-8	11-8	9-4	11-7	3-3	8-5	3-0	11-4	1-0	6-2	5-4	2-2	AAE12	6-5	7-2
3	6-5	8-7	6-4	7-3	6-3	1-1	4-2	-	5-2	1-0	4-0	3-2	-	CuAO	-	3-2
4	PEX14	PEX14	-	-	PEX14	-	-	-	1-0	-	1-0	-	-	-	-	-
5	15-13	12-10	9-8	10-6	11-8	1-1	4-4	-	6-4	2-2	3-2	11-9	ACAT1.3	2-2	-	ACX4
6	3-3	4-2	2-2	3-1	3-1	-	3-2	-	3-2	-	-	4-4	-	KAT1	2-0	-
7	CSY3	4-2	3-3	4-2	4-2	-	2-2	1-1	3-2	-	1-1	4-4	MLS	-	2-0	2-2
8	5-4	8-5	5-3	8-5	10-5	-	4-2	1-0	6-4	2-1	4-1	6-4	-	-	2-1	3-2
9	BADH	BADH	BADH	BADH	BADH	1-0	BADH	-	BADH	-	-	BADH	-	-	-	-
10	2-2	3-3	-	PEX19a	3-3	-	-	-	1-0	-	APEM9	2-1	-	-	-	-
11	3-2	7-5	3-3	4-3	2-2	-	1-0	2-1	5-1	LACS6	-	4-3	-	1-0	-	2-1
12	-	PXA1	-	PXA1	1-0	-	PXA1	-	PXA1	-	-	-	-	-	PXA1	-
13	4-4	5-4	4-4	5-3	4-3	-	DRP5B	1-0	5-2	-	-	2-2	-	PEX11d	3-2	-
14	8-7	7-6	7-6	9-6	9-6	2-2	7-2	4-0	7-2	4-2	2-2	7-3	SGAT1	2-2	5-5	-
15	2-2	3-2	1-1	3-1	3-1	1-1	2-1	1-0	2-1	CAT3	CAT3	CAT3	-	-	CAT3	-
16	3-3	4-4	4-3	3-2	4-3	-	-	-	1-0	2-2	IBR1	3-3	-	st1	2-2	St5

Table 4.5. Continued.

	Zea mays												*Oryza sativa*			
	Juvenile vs Adult			Stage comparison in the same leaves			Adult leaf comparison			Reproductive stage comparison			Early seedling comparison			
	11_im_v9	8_im_v9	13_im_v9	vT_v9_13	R2_V9_13	R2_vT_13	11_8_v9	13_11_v9	13_8_v9	6_vs_0 DPA	12_vs_6 DPA	30_vs_12 DPA	S2_vs_S1	S3_vs_S2	S4_vs_S3	S5_vs_S4
17	2-2	2-2	2-2	2-2	LON2	-	-	-	1-0	-	-	-	-	LON2	-	-
18	5-4	5-5	3-3	3-3	2-2	-	1-0	2-0	3-0	-	-	-	-	-	PEX1	-
19	-	-	-	PEX6	-	-	-	-	-	-	PEX6	PEX6	-	-	-	-
20	6-5	9-8	7-6	6-6	5-5	-	4-0	1-0	6-0	1-0	1-0	6-1	-	-	2-1	1-0
21	2-2	2-2	2-2	2-1	2-1	-	1-0	-	1-0	1-0	1-0	OPR3	-	-	-	1-0

In the comparison of 0 DPA vs 6 DPA, AGT2 and ATMS1 were upregulated in amino acid metabolism. EH3 and atDCI were upregulated in fatty acid breakdown. HAOX1 and MDH1 were upregulated in photorespiration. AIM1 and IBR3 were upregulated in phythormone biosynthesis. When it comes to 6 DPA vs 12 DPA comparison. AGT2, ATMS1 and PM16 were upregulated in amino acid metabolism. 4CI3 and ACL1 were upregulated in the biosynthesis of phylloquinone, biotin, CoA and ubiquinone. atDCI and SCP2 were upregulated in Fatty acid breakdown. 6PGL was upregulated in glyoxylate cycle and acetate-malate shunt.: NDB1 was upregulated in NADPH regeneration. HAOX1 and SGAT1 were upregulated in photorespiration. IBR3 were upregulated in phythormone biosynthesis. In the comparison of 12 DPA vs 30 DPA, AGT2, ASP3, SOX and PNC1 were upregulated in amino acid metabolism. 4CI3, NS, AAE7 and AAE17 were upregulated in the biosynthesis of phylloquinone, biotin, CoA and ubiquinone. CuAO and PAO4 were upregulated in catabolism of polyamines, urate, pseudouridine, sulfite and methylglyoxyal. ECHIA, EH3, ACX6, ECH2, ACX4, ACX3, ACAT1.3, atDCI and ECI were upregulated in fatty acid breakdown. KAT1, MFP2 and ACX5 were upregulated in fatty acid breakdown and phytohormone biosynthesis. 6PGL, MLS, CSY3 and ICL were upregulated in glyoxylate cycle and acetate-malate shunt. ZnDH, 6PGDH, NDB1 and SDRc were upregulated in NADPH regeneration. APEM9 was upregulated in peroxisomal import protein. BADH was upregulated in NADPH regeneration and catabolism of polyamines, urate, pseudouridine, sulfite and methylglyoxyal. LACS6, LACS7 and PMP22 were upregulated in peroxisomal solute transporters. PEX11a and PEX11c were upregulated in peroxisomal solute transporters. HAOX1 and MDH1 were upregulated in photorespiration. AIM1, st5, IBR3 were upregulated in phytohormone biosynthesis. GR1 and MDAR1 were upregulated in ROS/NS metabolism. OPR3 was upregulated in signaling Ca2+ and protein phosphorylation.

In the comparison of s1 vs s2 in *Oryza sativa*: 4CI3 and AAE7 were upregulated in biosynthesis of phylloquinone, biotin, CoA and ubiquinone. ACAT1.3.was upregulated in biosynthesis of phylloquinone, biotin, CoA and ubiquinone. In s2 vs s3, ACX4 and ACX6 were upregulated in fatty acid breakdown. SGAT1 and HPR were upregulated in phytohormone biosynthesis. In s3 vs s4, GGT1 was upregulated in amino acid breakdown. AAE1, AAE7 and AAE12 were upregulated in biosynthesis of phylloquinone, biotin, CoA and

ubiquinone. PEX11b and PEX11d were upregulated in biosynthesis of phylloquinone, biotin, CoA and ubiquinone. NADK3 was upregulated in NADPH regeneration. SGAT1, HPR, MDH1 and HAOX1 were upregulated in photorespiration. St5 and IBR3 were upregulated in phytohormone biosynthesis. CSD3 was upregulated in ROS/NS metabolism. In s4 vs s5 SOX was upregulated in amino acid metabolism. AAE1 and AAE12 were upregulated in biosynthesis of phylloquinone, biotin, CoA and ubiquinone. CuAO and PAO4 were upregulated in catabolism of polyamines, urate, pseudouridine, sulfite and methylglyoxyal. MLS and ICL were upregulated in glyoxylate cycle and acetate-malate shunt. HBCDH and 6PGDH were upregulated in NADPH regeneration. LACS7 was upregulated in peroxisomal solute transporters.

The expression patterns of peroxisome biogenesis and metabolism genes shows varied regulation across growth and development stages. Among twenty-one peroxisomal metabolic categories, we decided to address their activation status in terms of whether differentially expressed gene and upregulated gene numbers are high relative to the other comparison. Table 4.6 presents this evaluation.

Table 4.6. Ratio of DEGs and upregulated genes in the peroxisomal metabolism categories.

	Total	11_im_v9	8_im_v9	13_im_v9	vT_v9_13	v9_R2_13	vT_R2_13	11_8_v9	13_11_v9	13_8_v9
1	14	64% | 43%	64% | 43%	58% | 36%	71% | 36%	57% | 21%	14% | 14%	43% | 31%	21% | 7%	57% | 36%
2	22	50% | 36%	46% | 36%	50% | 36%	41% | 22%	50% | 32%	14% | 14%	36% | 23%	14%-0%	50% | 18%
3	13	46% | 39%	62% | 54%	46% | 31%	54% | 23%	46% | 23%	8% | 8%	31% | 15%	-	39% | 15%
4	2	PEX14	PEX14	-	-	PEX14	-	-	-	50% | 0%
5	19	79% | 68%	63% | 53%	47% | 42%	53% | 32%	58% | 42%	5% | 5%	21% | 21%	-	32% | 21%
6	5	60% - 60%	80% | 40%	40% | 40%	60% | %20	60% | 20%	-	60% | 40%	-	60% | 40%
7	5	CSY3	80% | 40%	60% | 60%	80% | 40%	80% | 40%	-	40% | 40%	20% | 20%	60% | 40%
8	13	39% | 31%	62% | 39%	39% | 23%	62% | 39%	77% | 39%	-	31% | 15%	20% | 0%	46% | 31%
9	2	BADH	BADH	BADH	BADH	BADH	50% | 0	BADH	-	BADH
10	6	33% | 33%	50% | 50%	-	PEX19a	50% | 50%	-	-	-	17% | 0%
11	9	33% | 22%	78% | 56%	33% | 33%	44% | 33%	22% | 22%	-	11% | 0%	22% | 11%	56% | 11%
12	1	-	PXA1	-	PXA1	100% | 0%	-	PXA1	-	PXA1
13	8	50% | 50%	62.5 | 50%	50% | 50%	62.5% | 37.5%	50% | 37.5%	-	DRP5B	12.5% | 0%	62.5% | 25%
14	10	80% | 70%	70% | 60%	70% | 60%	90% | 60%	90% | 60%	20% | 20%	70% | 20%	40% | 0%	70% | 20%
15	3	67% | 67%	100% | 67%	33% | 33%	100% | 33%	100% | 33%	33% | 33%	67% | 33%	33% | 0%	67% | 33%
16	10	30% | 30%	40% | 40%	40% | 30%	30% | 20%	40% | 30%	-	-	-	10% | 0%
17	3	67% | 67%	67% | 67%	67% | 67%	67% | 67%	LON2	-	-	-	33% | 0%
18	6	83% | 67%	83% | 83%	50% | 50%	50% | 50%	33% | 33%	-	17% | 0%	33% | 0%	50% | 0%
19	2	-	-	-	PEX6	-	-	-	-	-
20	11	55% | 46%	82% | 73%	64% | 55%	55% | 55%	46% | 46%	-	36%-0%	9% | 0%	55% | 0%
21	3	67% | 67%	67% | 67%	67% | 67%	67% | 33%	67% | 33%	-	33% | 0%	-	33% | 0%

Blue represents the categories that contain less than 50% DEGs and upregulated genes.
Yellow represents the categories that contain more than 50% DEGs ratio while the ratio of upregulated genes are less than 50%.
Red represents the categories that have more than 50% DEGs and upregulated genes.

We ranked the categories according to the ratio that is calculated with the total number of genes that assigned to a category divided by the number of differentially expressed genes or the number of upregulated genes in the relevant category. This calculation offered us the opportunity of selecting the most active peroxisomal metabolism category. We assumed that the category that has more red color in the has higher rank relative to the other categories that have less red color. The same approach was used to evaluate the categories that have yellow and blue color as well. Given those ratios, Table 4.7 presents the rank of the peroxisomal metabolism categories.

Table 4.7. Ranked peroxisomal metabolic pathways.

RANK	CATEGORY
1	Photorespiration
2	Recycling of PEX5
3	Quality control and proteome remodeling
4	Signaling Ca2+ and protein phosphorylation.
5	Peroxisome fission gene
6	ROS/RNS metabolism,
7	NADPH regeneration and Catabolism of polyamines, urate, pseudouridine, sulfite and methylglyoxyal
8	Fatty acid breakdown
9	Photorespiration and ROS/RNS metabolism and autophagy
10	Docking complex
11	Fatty acid breakdown and phytohormone biosynthesis
12	Glyoxylate cycle and acetate malate shunt
13	Peroxisomal solute transporters
14	Catabolism of polyamines, urate, pseudouridine, sulfite and methylglyoxya
15	Peroxisomal import protein
16	Amino acid metabolism
17	Biosynthesis of phylloquinone, biotin, CoA and ubiquinone
18	NADPH regeneration
19	Phytohormone biosynthesis
20	Recycling of PEX5 and pexophagy and phytohormone biosynthesis

First eight categories had at least one dataset that harbors over 50% upregulated genes. Photorespiration showed the most active gene expression pattern among the categories. Photorespiration pathway was especially most active in the juvenile and adult leaf transcriptome comparison. Accordingly, we associated photorespiration with the other metabolic event that take place in peroxisome to catch interesting story.

As plants grow and develop, photosynthetic activity increases in the photosynthetic cells, therefore, it is reasonable that increment in photosynthetic activity will elevate the rate of photorespiration. While the other metabolic

pathways were associated with photorespiration, none of them contained an interesting story but NADPH regeneration and catabolism of polyamines. Moreover, the genes that are involved in photorespiration, and in the ascorbate-glutathione cycle (ROS metabolism) which is the secondary scavenging system for H_2O_2 detoxification were upregulated across the juvenile and adult leaf comparison, but GOX and GGT genes. Interestingly, Only the gene encoding BADH enzyme that catalyzes the conversion of aminobutanal to GABA was upregulated in polyamine catabolism (Figure 4.4).

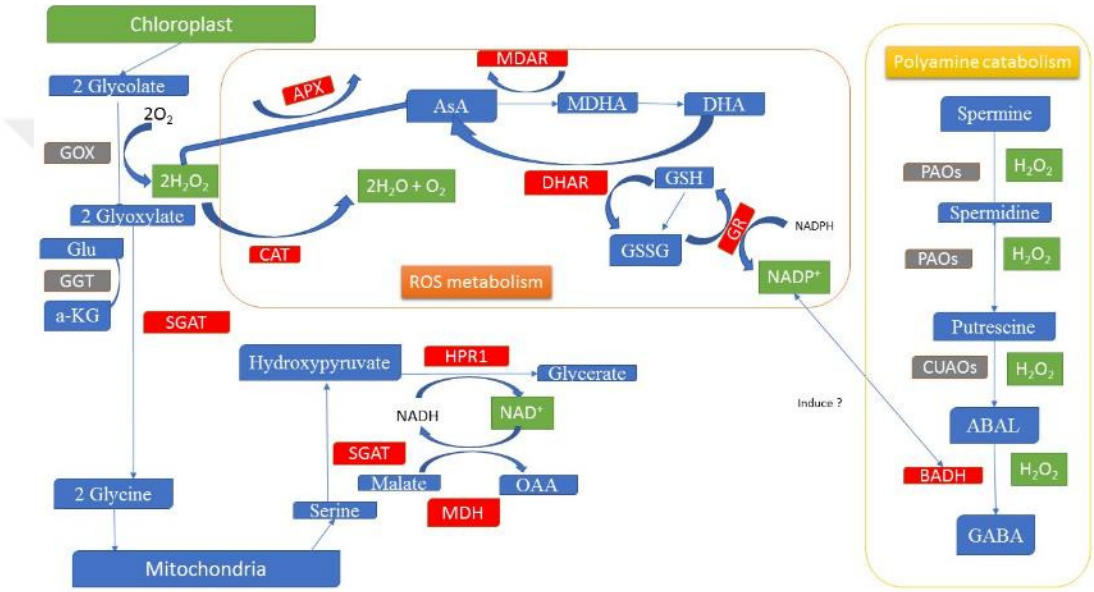


Figure 4.4. Theoretical association of photorespiration, ROS metabolism and polyamine catabolism, which take place in peroxisome lumen. The genes that are shown in red were upregulated.

4.1.8 Reverse Transcriptase Quantitative Polymerase Chain Reaction

The expression of betaine alanine dehydrogenase and catalase genes were examined via qRT-PCR. Since the reference genome of *T. monococcum* has not been completed yet, the primer of the genes was designed using the peptide seqences of *T. aestivum* and then the peptide sequence was aligned to cDNA sequences of *T. monococcum* DV92 accesion that generated by jaiswallab (http://jaiswallab.cgrb.oregonstate.edu/genomics/wheat). The isoforms (Zm00001d050339_P001 and Zm00001d025626_P002, which contain PTS1 signal [c-terminal SKL]) of maize BADH gene, were used as reference sequence to

identify bread wheat homologs. Three loci were determined in *T. aestivum* through Blast+ software; Traes_2AL_5C9497802.6, Traes_2BL_22224159C.1 and Traes_2DL_7399173C7.1. Since Traes_2BL_22224159C.1 peptide sequence is absent of peroxisomal PTS1 signal sequence [c-terminal SKL], this sequences was excluded from multiple sequence alignment. Thirty-one cDNA fragment were identified by means of three homologous bread wheat peptide sequence: TmoDV92v1_026255, TmoDV92v1_026256, TmoDV92v1_026257, TmoDV92v1_026258, TmoDV92v1_026259, TmoDV92v1_026261, TmoDV92v1_026262, TmoDV92v1_026263, TmoDV92v1_026264, TmoDV92v1_026265, TmoDV92v1_026266, TmoDV92v1_026267, TmoDV92v1_026268, TmoDV92v1_026271, TmoDV92v1_026272, TmoDV92v1_026273, TmoDV92v1_026274, TmoDV92v1_035340, TmoDV92v1_035341, TmoDV92v1_035342, TmoDV92v1_035343, TmoDV92v1_035344, TmoDV92v1_035345, TmoDV92v1_035346, but none of the peptide sequence harbors c-terminal SKL signal sequence (Figure 4.5).

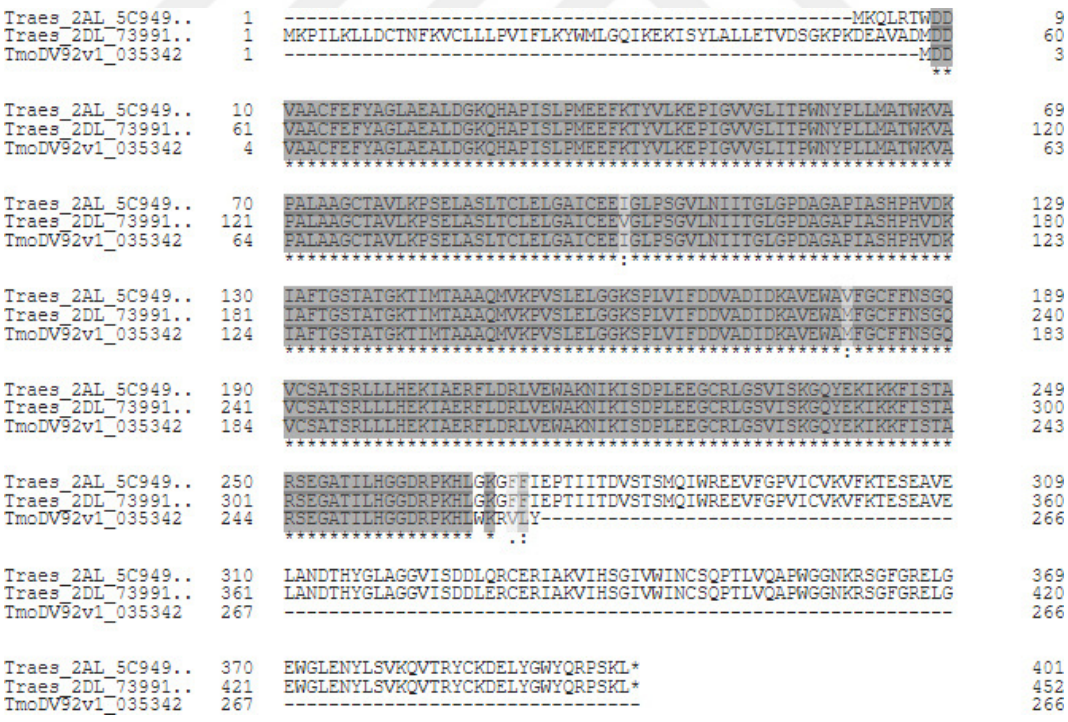


Figure 4.5. Multiple sequence alignment result of BADH peptide sequences. Identical amino acid sequences were indicated in grey color. * represents identical sequences.

As the signal sequence targets peroxisome lumen, it is important to identify the transcript that will translate into peptide sequence that contains c-terminal SKL

signal sequence. In this situation, nucleotide sequences of the same genes were taken into consideration, therefore it was realized that nucleotide sequence of einkorn has c-terminal SKL signal sequences at the 3’ end (Figure 4.6).

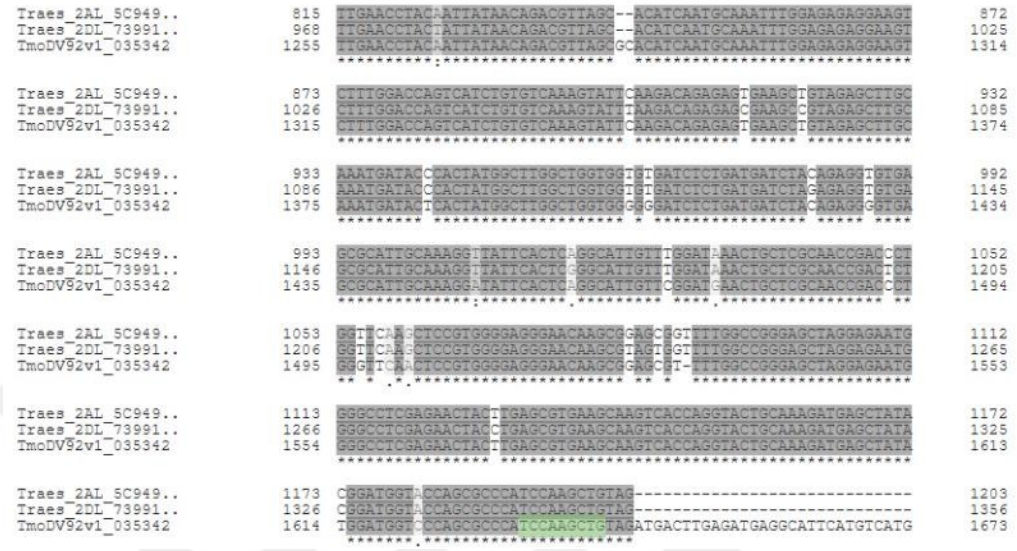


Figure 4.6. Multiple sequence alignment of nucleotide sequences. Signal sequences are indicated in green, but greys represent identical sequences.

Two different forward and reverse oligonucleotide were designed in this case. One is for qRT-PCR, but other is for validating if the amplified sequences contain c-terminal SKL signal sequence.

Catalase gene was amplified according to Tounsi et al. (2019) (48). TmoDV92v1_013359 nucleotide sequence was used to check the amplified region. RNAse L inhibitor-like protein (RLI) was endogenous control in qRT-PCR reaction.

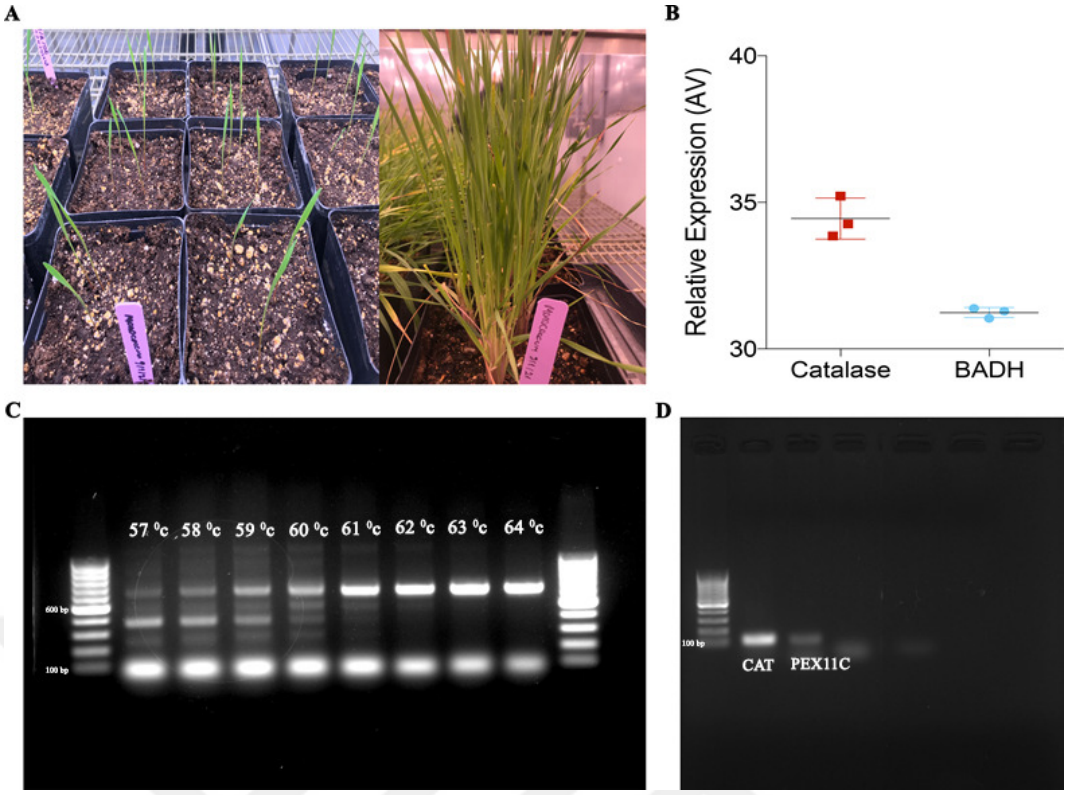


Figure 4.7. qRT-PCR results. A) the picture of cultivated plants at juvenile and adult phase. B) Relative expressions of CAT and BADH genes. C) Agarose gel picture of BADH primers. D) Agarose gel picture of CAT and PEX11C primers.

Four separate bands are exclusively visible in the 58 °C annealing temperature in gradient PCR analysis (Figure 4.7c), validating the transcript that has been working in the study contains peroxisomal targeting signal SKL. Additionally, PCR study validates catalase gene for qRT-PCR (Figure 4.7d). Given the qRT-PCR results, BADH and catalase genes are upregulated in biological replicates of adult leaf (Figure 4.7b).

4.2 Peroxisomes in Drought Stress

In this section, I explored the expression patterns of peroxisome metabolism genes. The results of bioinformatics analysis were interpreted with the association of drought stress. The gene(s) that shows interesting expression pattern and the function of its target molecule(s) is discussed in this section.

4.2.1 Data Sets Acquisition for Drought Stress

Five considerations were followed for selecting the RNA-Seq data: (1) more than one replicate per treatment (2) data for the control and drought stress for each experiment; (3) RNA was extracted from the above-ground organs; (4) drought was induced by withdrawing the watering; and (5) the plants were wild type. In GEO

dataset collection, we found 19 different studies that compatible with our research aim. The information about the studies is present in Table 4.8.

Table 4.8. NCBI data sets that are used in the drought study.

GEO ACCESSION NO.	SPECIES	CULTIVAR	STRESS DURATION	REPLICATES	STUDY LOCATION	ARTICLE
GSE134945	*A. thaliana*	Columbia (WT)	7d,1, 21 d	3	Growth chamber	https://doi.org/10.1186/s12864-019-6350-5
GSE93979	*A. thaliana*	Columbia (WT)	21 d	3	Growth chamber	https://doi.org/10.1093/jxb/erx157
GSE78972	*Oryza Sativa*	Nipponbare (WT)	8 weeks	3	Growth chamber	https://doi.org/10.1111/pce.12760
GSE78972	*Oryza Sativa*	Nipponbare (WT)	6 weeks	3	Growth chamber	https://doi.org/10.1111/pce.12760
GSE65022	*Oryza Sativa*	Nipponbare (WT)	10 d	2	No info	No publication
GSE57950	*Oryza indica*	Huanghuazhan (P28) (Tolerant)	1-3 d	2	Green house	https://doi.org/10.1186/1471-2164-15-1026
GSE57950	*Oryza indica*	Huanghuazhan (H471) (Tolerant)	1-3 d	2	Green house	https://doi.org/10.1186/1471-2164-15-1026
GSE57950	*Oryza indica*	Huanghuazhan (HHZ) (Susceptible)	1-3 d	2	Green house	https://doi.org/10.1186/1471-2164-15-1026
GSE132113	*Zea mays*	B104	5 d	2	Field experiments	https://doi.org/10.1186/s12870-019-1941-5
SRP106756 (PRJNA385354)	*Zea mays*	M54 and 753F	1d, 2d, 3d	1	No info	No publication
SRP132035 (PRJNA432650)	*Zea mays*	No info	No info	4	No info	No publication
GSE48507	*Zea mays*	B73	6h and 66 h	2	Green house	https://doi.org/10.1186/1471-2229-14-141
GSE137780	*Zea mays*	B104	No info	2	No info	No publication
GSE137780	*Zea mays*	W22	No info	2	No info	No publication
GSE137780	*Zea mays*	DH4866	No info	2	No info	No publication
GSE128441	*Sorghum Bicolor*	RTX430	3 week	3	Field experiment	https://doi.org/10.1073/pnas.1907500116
GSE128441	*Sorghum Bicolor*	RTX430	6 week	3	Field experiment	https://doi.org/10.1073/pnas.1907500116
GSE80699	*Sorghum Bicolor*	IS20351 (Sensitive)	8 days	3	Controlled Field experiment	https://doi.org/10.1186/s12870-016-0800-x
GSE80699	*Sorghum Bicolor*	IS20330 (Tolerant)	8 days	3	Controlled Field experiment	https://doi.org/10.1186/s12870-016-0800-x

In the GSE134945 dataset, M. truncatula Jemalong A17 and A. thaliana Columbia wild-type (WT) Col-0 was used as control. Only Control samples are used to examine the GSE134945 dataset. In the GSE93979 dataset, Arabidopsis plants were grown in moist soil for 7 days. Drought stress was imposed by withholding water for 10 days. In the GSE78972 dataset, the drought stress was applied to rice plants grown under specific photoperiods and during flowering. There were two conditions in the GSE78972: long and short day. Each photoperiod treatment was considered as different study. In the GSE65022 dataset, all samples were used in our study. The GSE57950 dataset contains three different genotypes. Each genotype was considered as different study. The GSE132113 dataset contains examination of 3 different organs (ears at the V9 stage, kernels, and ear leaf at the 5 DAP) of B104 maize inbred line under normal and drought conditions. Only the ear leaf dataset was used in our study (accession codes SRP106756 and SRP132035). In the GSE48507 dataset, all samples were used in our study. In the GSE137780 dataset, the following 6 different genotypes B104, ZmNF-YB16 OE, DH4866, ZmNF-YA1 OE, W22 and zmnf-ya1 mutant were examined under normal and drought conditions. Only wild-type (B104, DH4866, W22) cultivars were selected for our study. Each genotype was considered as a different study. The GSE128441 dataset contains a large-scale field experiment, imposing pre- and post-flowering drought stress on two genotypes of sorghum. Each genotype was considered as different study. Only data for week three and six were used in our study. The GSE80699 dataset includes the drought sensitive (IS20351) and the drought tolerant (IS22330) genotypes. Each genotype was considered as different study.

4.2.2 Alignment of reads of RNA-seq experiments

Clean paired-end or single-end reads of each data set were mapped to the respective reference genome of the species whith the alignment efficiency ranging from 80% to 97% (Table 4.9).

Table 4.9. Pair-end alignment statistics in the study.

Dataset	Sample	Total pairs	Aligned concordantly 1 time	Overall alignment rate
GSE134945	GSM3980153	22689341	20885488 (92.05%)	96.71%
	GSM3980154	13933340	12806599 (91.91%)	96.58%
	GSM3980155	24828356	22485847 (90.57%)	95.54%
	GSM3980159	23588317	21917533 (92.92%)	97.17%
	GSM3980160	26477934	24567016 (92.78%)	97.06%
	GSM3980161	21191527	19774971 (93.32%)	97.37%
GSE93979	GSM2465998	6899925	6388163 (92.58%)	96.80%
	GSM2465999	6329286	5857556 (92.55%)	96.67%
	GSM2466002	6952440	6497556 (93.46%)	97.03%
	GSM2466003	6783650	6339834 (93.46%)	97.06%
GSE57950	GSM1398429	7829204	6711683 (85.73%)	91.48%
	GSM1398430	8152417	7128548 (87.44%)	93.61%
	GSM1398441	8802245	7825270 (88.90%)	94.61%
	GSM1398442	8839491	7835735 (88.64%)	94.54%
	GSM1398431	7829204	6711683 (85.73%)	91.48%
	GSM1398432	8152417	7128548 (87.44%)	93.61%
	GSM1398443	8802245	7825270 (88.90%)	94.61%
	GSM1398444	8839491	7835735 (88.64%)	94.54%
	GSM1398433	7829204	6711683 (85.73%)	91.48%
	GSM1398434	8152417	7128548 (87.44%)	93.61%
	GSM1398445	8802245	7825270 (88.90%)	94.61%
	GSM1398445	8839491	7835735 (88.64%)	94.54%
SRP132035	SRR6665368	31589335	22348551 (70.75%)	85.35%
	SRR6665369	19074135	12630462 (66.22%)	83.84%
	SRR6665370	19803856	14839818 (74.93%)	87.54%
	SRR6665371	29600397	22176443 (74.92%)	87.79%
	SRR6665372	20382076	15395653 (75.54%)	87.67%
	SRR6665373	27812511	20104134 (72.28%)	86.45%
	SRR6665374	26948651	19951947 (74.04%)	86.97%
	SRR6665375	25378192	19301683 (76.06%)	87.95%
	SRR6665376	27328553	18964501 (69.39%)	85.18%
	SRR6665377	20714648	15086340 (72.83%)	86.51%
	SRR6665378	20527936	14868098 (72.43%)	86.38%
	SRR6665379	25792919	18966597 (73.53%)	86.79%
GSE80699	GSM2133750	23400214	20355341 (86.99%)	94.19%
	GSM2133751	23403237	19627293 (83.87%)	92.34%
	GSM2133752	23831987	19929399 (83.62%)	92.05%
	GSM2133753	24041200	20247316 (84.22%)	92.19%
	GSM2133754	22800008	19156958 (84.02%)	92.22%
	GSM2133755	23458470	19862140 (84.67%)	92.69%
	GSM2133756	23964074	20250273 (84.50%)	93.14%
	GSM2133757	23391907	19664756 (84.07%)	92.78%
	GSM2133758	23901435	20160810 (84.35%)	92.89%
	GSM2133759	23409913	20078627 (85.77%)	93.55%
	GSM2133760	23335648	20059335 (85.96%)	93.59%
	GSM2133761	23514842	19687265 (83.72%)	92.21%

Table 4.10. Single-end alignment statistics in the study.

Dataset	Sample	Total pairs	Aligned 1 time	Overall alignment rate
GSE78972	GSM2082857	33012246	30429101 (92.18%)	97.94%
	GSM2082858	28693929	26341971 (91.80%)	97.79%
	GSM2082859	33738285	31011326 (91.92%)	97.23%
	GSM2082860	31951497	29561312 (92.52%)	97.87%
	GSM2082861	34717753	31981070 (92.12%)	98.34%
	GSM2082862	28803587	26466871 (91.89%)	98.21%
	GSM2082863	32798250	30092885 (91.75%)	97.19%
	GSM2082864	33838322	30839781 (91.14%)	96.46%
	GSM2082865	32797025	30360972 (92.57%)	98.44%
	GSM2082866	35848188	32745201 (91.34%)	96.55%
	GSM2082867	37177413	34085194 (91.68%)	97.89%
	GSM2082868	24165701	22277086 (92.18%)	98.01%
GSE65022	GSM1585980	32797025	30360972 (92.57%)	98.44%
	GSM1585981	35848188	32745201 (91.34%)	96.55%
	GSM1585982	37177413	34085194 (91.68%)	97.89%
	GSM1585983	24165701	22277086 (92.18%)	98.01%
GSE132113	GSM3844504	12047870	9558871 (79.34%)	86.68%
	GSM3844505	10995637	6129641 (55.75%)	67.94%
	GSM3844511	11992437	9641335 (80.40%)	87.46%
	GSM3844512	12014569	9459140 (78.73%)	85.69%
SRP106756	SRR5520482	31688970	24153506 (76.22%)	83.24%
	SRR5520483	35459053	27207803 (76.73%)	83.85%
	SRR5520484	39941093	28914034 (72.39%)	78.95%
	SRR5520485	49354932	33288636 (67.45%)	73.45%
	SRR5520486	39718688	30625128 (77.11%)	84.18%
	SRR5520487	35458504	26938820 (75.97%)	82.84%
	SRR5520488	34105269	25124208 (73.67%)	80.69%
	SRR5520489	37342033	24216921 (64.85%)	70.45%
GSE137780	GSM4087752	12047870	9558868 (79.34%)	86.68%
	GSM4087753	10995637	6129666 (55.75%)	67.94%
	GSM4087764	11992437	9641305 (80.39%)	87.46%
	GSM4087765	12014569	9459268 (78.73%)	85.69%
	GSM4087756	24013417	19501974 (81.21%)	90.85%
	GSM4087757	24021872	19606627 (81.62%)	91.35%
	GSM4087768	24061034	19828342 (82.41%)	92.02%
	GSM4087769	24023725	19708396 (82.04%)	91.93%
	GSM4087760	24012409	19593067 (81.60%)	90.95%
	GSM4087761	24005071	19558470 (81.48%)	90.99%
	GSM4087772	23921725	19370749 (80.98%)	90.70%
	GSM4087773	23971035	19374792 (80.83%)	90.67%
GSE128441	GSM3674873	20560068	19491842 (94.80%)	96.65%
	GSM3674874	22877013	21493906 (93.95%)	95.71%
	GSM3674875	20789275	19650479 (94.52%)	96.34%
	GSM3674932	16201006	15377793 (94.92%)	96.73%
	GSM3674933	20638886	19658301 (95.25%)	97.06%
	GSM3674934	21843685	20481291 (93.76%)	95.58%
	GSM3674881	19411607	18681813 (96.24%)	99.16%
	GSM3674882	20365041	19460678 (95.56%)	98.35%
	GSM3674883	22333477	21531471 (96.41%)	99.25%
	GSM3674941	18886298	17989375 (95.25%)	97.67%
	GSM3674942	20640659	19969920 (96.75%)	98.97%
	GSM3674943	15550062	14809187 (95.24%)	97.65%

These outcomes mean, the throughput and sequencing quality were sufficiently high to warrant further analysis.

4.2.3 Exploring Global Read Count Patterns

RNA-Seq data from 19 different studies that acquired from four different species were evaluated to identify drought stress transcriptomes. To obtain a general overview of transcription patterns of the relevant species, we examined the principal component of the samples using the top 2000 genes that show the greatest expression variance (Appendix 3). PC-1 and PC-2 represented over 80% of the total variance. No variability was observed among replicates. Samples show greater variability among species, whereas transcription patterns of genes show a little variety between samples. Samples were separated according to the treatments in the first PC.

4.2.4 Differentially Expressed Genes

In Arabidopsis thaliana, four samples of the GSE93979 dataset were examined for DEG analysis. 15936 genes were counted in the GSE93979 dataset. 2145 genes were differentially expressed; 1319 genes were upregulated, whereas 826 genes were downregulated. Six samples of the GSE134945 dataset were examined for DEG analysis. 19194 genes were counted in the GSE134945 dataset. 710 genes were differentially expressed. 573 genes were upregulated, whereas 137 genes were downregulated.

In Oryza sativa, 12 samples of the GSE78972 data set were examined for DEG analysis. The data set contains two treatments as long and short drought stress. The treatments were examined individually in DEG analysis. 22290 genes were counted in the GSE78972 dataset for long drought stress and 14690 genes were differentialy expressed: 7092 genes were up-regulated, whereas 7598 genes were down-regulated. 21934 genes were counted in the GSE78972 dataset for short drought stress treatment. 13938 genes were differentially expressed. 6931 genes were up-regulated, whereas 7007 genes were down-regulated. 18361 genes were counted in the GSE65022 dataset. 9346 genes were differentially expressed. 4806 genes were up-regulated, whereas 4540 genes were down-regulated. The GSE57950 dataset contains three rice genotypes as P28, H471, HHZ. 18605 genes were counted in P28 dataset. 7895 genes were differentially expressed. 3701 genes were up-regulated, whereas 4194 genes were down-regulated. 18524 genes were counted in H471 dataset. 9044 genes were differentially expressed. 4394 genes

were up-regulated, whereas 4650 genes were down-regulated. 18323 genes were counted in HHZ dataset. 9423 genes were differentially expressed. 4616 genes were up-regulated, whereas 4807 genes were down-regulated.

In Zea mays, 20290 genes were counted in the GSE132113 dataset. 9025 genes were differentially expressed. 4302 genes were up-regulated, whereas 4723 genes were down-regulated. 30194 genes were counted in the SRP106756 dataset; 2238 genes were differentially expressed. 1077 genes were up-regulated, whereas 1161 genes were down-regulated. 30194 genes were counted in SRP132035. 13795 genes were differentially expressed. 7105 genes were up-regulated, whereas 6690 genes were down-regulated. 23990 genes were counted in GSE48507. 8899 genes were differentially expressed. 4710 genes were up-regulated, whereas 4189 genes were down-regulated. Twelve samples were examined for DEG analysis in the GSE137780 dataset. This data set was separated into three sub-data set as GSE137780_b104, GSE137780_DH4866 and GSE137780_W22 according to cultivars used in the study. 20295 genes were counted in the GSE137780_b104 dataset; 9024 genes were differentially expressed. 4303 genes were up-regulated, whereas 4721 genes were down-regulated. 22049 genes were counted in the GSE137780_DH4866 dataset; 10470 genes were differentially expressed. 5381 genes were upregulated, whereas 5089 genes were downregulated. 21941 genes were counted in the GSE137780_W22 dataset. 12629 genes were differentially expressed. 6496 genes were up-regulated, whereas 6133 genes were down-regulated.

In Sorghum bicolor, twelve samples of the GSE128441 dataset were examined for DEG analysis. Dataset contains two treatments as week 3 and week 6. The treatments were examined individually in DEG analysis. 23576 genes were counted in the GSE128441 week 3 data sets; 2897 genes were differentially expressed. 1181 genes were up-regulated, whereas 1716 genes were down-regulated. 23576 genes were counted in the GSE128441 week 6 data sets. 1959 genes were differentially expressed; 931 genes were up-regulated, whereas 1028 genes were down-regulated. Twelve samples of the GSE80699 dataset were examined for DEG analysis. The GSE80699 dataset contains two genotypes-IS20351 (Sensitive) and IS20330 (Tolerant). The treatments were examined individually in DEG analysis. 24643 genes were counted in IS20351 (Sensitive); 4298 genes were differentially expressed. 1794 genes were up-regulated, whereas

2504 genes were down-regulated. 24301 genes were counted in IS20330 (Tolerant). 4576 genes were differentially expressed; 1918 genes were up-regulated, whereas 2658 genes were down-regulated.

4.2.5 GO Enrichment Analysis

The Gene Ontology allows users to describe a gene/gene product in detail considering three main aspects: its molecular function, the biological process in which it participates, and its cellular location. To determine the relationship of the differentially expressed genes to the stress-related genes, we carried out GO enrichment analysis using topGO package in R so that we can overview the biological processes of differentially expressed genes. The number of enriched GO term categories was between 40 and 241 among the studies. The enrichment of stress related GO terms in each study is presented in Table 4.11.

Table 4.11. Stress-related GO terms enrichment in each dataset. Differentially expressed and annotated genes are shown in the GENE RATIO column. LOGFC column indicates minimum and maximum log2folcahange values of the differentially expressed genes. Q-VALUE column shows the adjusted p-value of fisher exact test.

STUDY	SPECIES	TERM	GO_ID	GENE RATIO	LOGFC	Q-VALUE
GSE137780_DH4866	Maize	Defense response	GO:0006952	20/51	-6.86/4.68	0.022
		negative regulation of gene expression	GO:0045814	3/6	-2.19/-0.89	0.046
		regulation of cyclin-dependent protein s...	GO:0000079	4/18	-2.02/2.39	0.008
GSE137780_B104	Maize	Response to oxidative stress	GO:0006979	31/185	-4.68/3.74	0.032
		Response to abiotic stimulus	GO:0009628	10/27	-2.88/6.51	0.024
		Regulation of cyclin-dependent protein k...	GO:1904029	4/18	-2.09/ -0.67	0.001
GSE137780_W22	Maize	Regulation of cellular component organiz...	GO:0051128	15/56	-2.17/2.07	0.035
		Regulation of biosynthetic process	GO:0009889	413/1320	-8.30/7.35	0.035
		Regulation of cyclin-dependent protein synthesis	GO:0000079	12/18	-1.13/2.02	0.001
		Regulation of actin polymerization or de...	GO:0008064	11/26	-2.17/2.07	0.000
		Regulation of cellular component organiz...	GO:0051128	13/56	-0.70/4.25	0.000
GSE78972S	Rice	Response to external stimulus	GO:0009605	207/494	-6.65/8.27	0.000
		Regulation of gene expression, epigenetics	GO:0040029	65/128	3.28/5.08	0.000
		Regulation of anatomical structure size	GO:0090066	169/465	-12.15/5.71	0.000
		Abscission	GO:0009838	9/29	-2.93/0.56	0.000
GSE78972L	Rice	Response to abiotic stimulus	GO:0009628	1118/2194	-9.52/9.71	0.000
		Response to extracellular stimulus	GO:0009991	105/304	-9.15/9.71	0.001
		Regulation of biological process	GO:0050789	710/1588	-8.13/8.73	0.000
		Regulation of cell size	GO:0008361	179/465	-10.98/8.52	0.000
		Regulation of gene expression, epigeneti...	GO:0040029	64/128	-2.48/3.31	0.000
		Anatomical structure morphogenesis	GO:0009653	343/886	-10.98/8.52	0.000
		Abscission	GO:0009838	9/29	-2.60/3.27	0.000
GSE65022	Rice	Response to extracellular stimulus	GO:0009991	79/304	-6.71/9.40	0.000
		Regulation of cell size	GO:0008361	122/465	-6.71/7.11	0.000
		Regulation of cellular process	GO:0050794	524/1464	-8.69/8.72	0.005
		Peroxisome	GO:0005777	100/191	-5.63/8.48	0.000

Table 4.11. Continued.

GSE57950P	Rice	Response to stress	GO:0006950	1106 /3618	-10.35/9.82	0.036
		Response to abiotic stimulus	GO:0009628	736/2194	-10.35/9.82	0.00
		Regulation of anatomical structure size	GO:0090066	99/465	-7.03/4.58	0.013
		Regulation of biological process	GO:0050789	498/1588	-8.8/6.00	0.000
		Regulation of gene expression	GO:0010468	34/128	-2.68/1.70	0.000
GSE57950HZ	Rice	Regulation of anatomical structure size	GO:0090066	127/465	-9.66/4.45	0.000
		Regulation of cell size	GO:0008361	127/465	-9.66/4.45	0.001
		Regulation of gene expression	GO:0010468	46/128	-4.16/1.72	0.000
		Signal transduction	GO:0007165	533/1464	-9.75/11.15	0.001
GSE57950H	Rice	Response to stress	GO:0006950	1198/3618	-10.63/10.41	0.000
		Response to external stimulus	GO:0009605	132/494	-9.56/4.61	0.000
		Regulation of cellular process	GO:0050794	501/1464	-9.09/7.22	0.018
		Peroxisome	GO:0005777	99/191	-8.73 / 3.01	0.000
GSE128441W3	Sorghum	Regulation of metabolic process	GO:0019222	104/1170	-2.69/7.58	0.003
GSE128441W6	Sorghum	Regulation of cellular process	GO:0050794	106/1586	-2.75/3.44	0.027
		Regulation of transcription	GO:0006355	65/1077	-2.75/3.44	0.000
GSE80699_IS20351	Sorghum	Integral component of peroxisomal membra...	GO:0005779	3/7	0.86/2.01	0.000
		Regulation of transcription	GO:0006355	107/996	-8.37/7.47	0.012
GSE80699_IS22330	Sorghum	Cellular response to stress	GO:0033554	13/141	-4.67/-4.49	0.000
GSE93979	Arabidopsis	Response to stress	GO:0006950	577/3887	-9.67/7.98	0.000
		Regulation of response to osmotic stress	GO:0047484	6/37	-1.01/4.19	0.010
		Regulation of response to external stimulus	GO:0032101	20/102	-0.86/2.18	0.000
		Response to jasmonic acid	GO:0009753	42/208	-4.78/4.22	0.000
GSE134945	Arabidopsis	Auxin-activated signaling pathway	GO:0009734	13/208	-2.82/1.02	0.001

Genes involved in response to stress, to abiotic stimuli including drought, and stress-induced regulation of gene expression were enriched in all RNA-seq data sets for the stress samples relative to control.

4.2.6 Expression Patterns of Peroxisome Biogenesis

We used the previous study published by (Kaur ve Hu, 2011) to identify peroxisome-related genes. They compiled a list containing 170 genes encoding peroxisome lumen and membrane proteins. The expression patterns of peroxisome genes were screened with those gene locus IDs. Maize and sorghum gene locus IDs were identified by OrthoFinder using proteome sequences of the respective species. The number of differentially expressed peroxisome lumen and membrane genes was shown in Figure 4.8. Most of the differentially expressed genes were observed in Oryza sativa and Zea mays data sets.

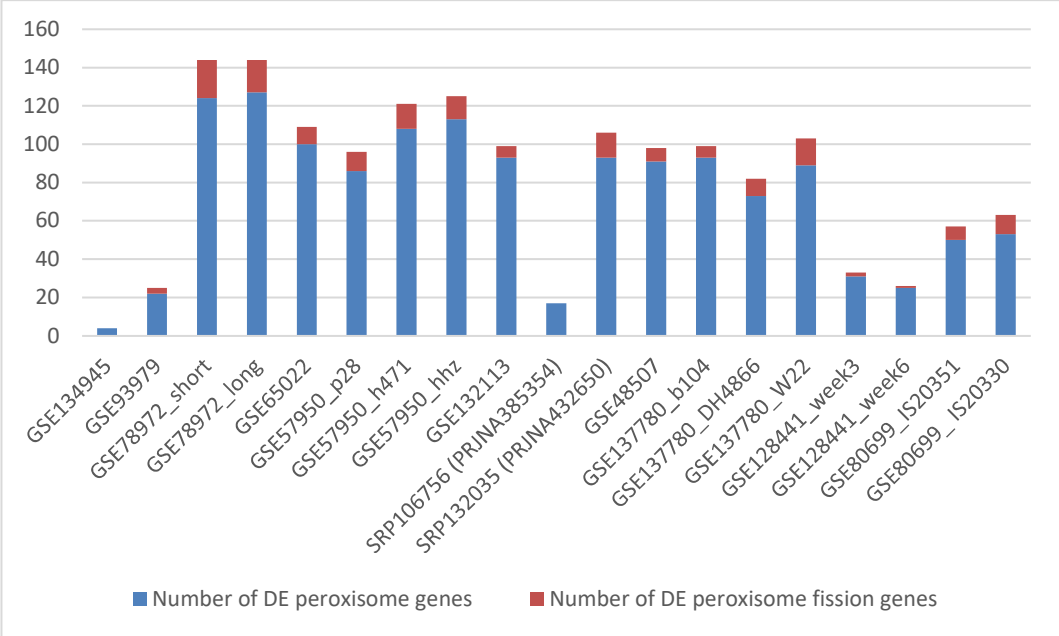


Figure 4.8. The number of differentially expressed peroxisome lumen and membrane genes. Horizontal line represents the number of differentially expressed genes. Vertical line represents the data sets.

As representing 19 different datasets in one heatmap was impractical, we prepared a representative heatmap for each species. Figure 4.9 shows expression patterns of one representative dataset for each species.

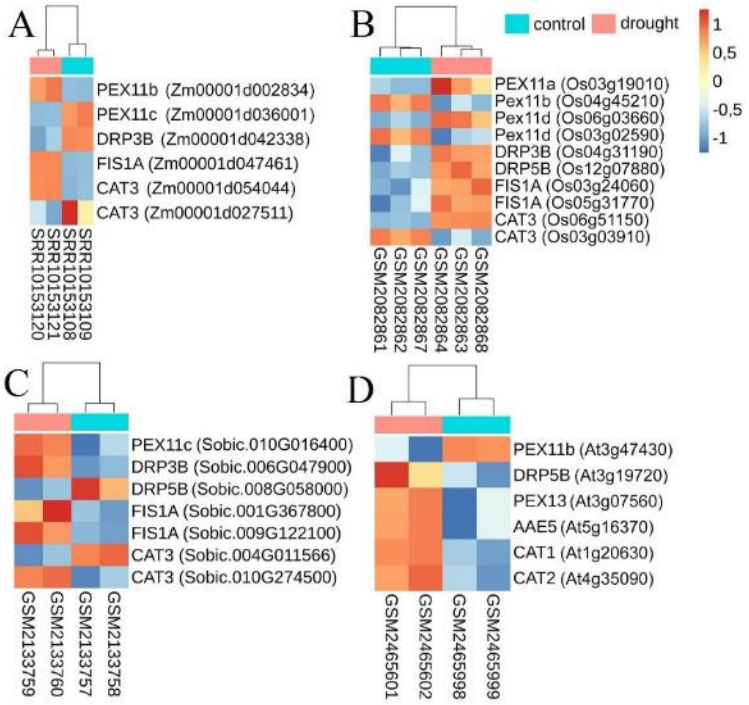


Figure 4.9. A-D, Heatmaps of peroxisome fission genes that are differentially expressed in response to drought stress in Z. mays (A), O. sativa (B), S. Bicolor (C), A. thaliana (D). The figure was generated with the R pheatmap package using VarianceStabilizedTransformation -vst() function- built in DESeq2 package. Vst values were represented based on the z-score transformation. Samples were clustered according to Pearson correlation analysis. Loci name of orthologs that were mapped to peroxisome fission genes indicated beside gene names.

We present whole expression data of peroxisome fission genes in Appendix 4. In Arabidopsis thaliana, PEX11b, DRP5B, PEX13 and PEX19B were upregulated in one data set.

In *Oryza sativa*, PEX11c gene was upregulated under drought stress in all datasets. LOC_os03g19010 transcript corresponding to PEX11a was upregulated during drought stress except two datasets. PEX11b was upregulated in all data set, however; PEX11d was not present as upregulated gene in three control groups. These results indicate that PEX11c and PEX11a are important for drought response, while PEX11b and PEX11d are crucial for homeostasis of cellular processes under normal condition. DRP3B was up-regulated only in two datasets, whereas DRP5B was up-regulated all datasets except one. It revealed that DRP5B was upregulated in both tolerant and susceptible cultivars. LOC_os06g51150 transcript of CAT3 gene was upregulated in all datasets under drought stress, however LOC_os03g03910 transcript of the CAT gene in Oryza sativa was upregulated in control groups. We found only two upregulated FISA homologs

(LOC_os03g24060) in the data set under drought stress, whereas it was upregulated in only one data set in the control. Other transcript (LOC_os05g31770) of FISA showed a contradiction in terms of expression patterns. PEX1, PEX7 and PEX13 were among genes that show a contradiction in terms of expression pattern. PEX5 and PEX3b were upregulated in two data sets. Other PEX genes generally were upregulated in the control.

In *Z. mays*, PEX11C gene was upregulated under drought stress in three datasets, which showed the same trend with the rice study. PEX11B was upregulated under drought stress in two data sets. PEX11D gene was up-regulated in the control group in one dataset; nevertheless, it was up-regulated in two datasets under drought stress. PEX11B and PEX11D were never upregulated in rice datasets. PEX11A was not a DEG in maize, while it was upregulated under drought in rice.

DRP3B was up-regulated in three datasets under normal condition; however, DRP3B was up-regulated under only drought stress in the rice datasets. DRP5B was not DEGs in maize, while it was upregulated under drought stress in rice.

FIS1A (zm00001d047461) was upregulated in one dataset under drought stress. PEX1 was upregulated in two datasets under normal condition. PEX2, PEX3B, PEX5 and PEX10 were upregulated one time under normal condition. PEX12 and PEX16 were upregulated under normal condition in three and two datasets respectively. PEX10 and PEX12 showed the same results in both datasets from C3 and C4 plants. PEX4 and PEX13 were up-regulated in three data set under drought condition, whereas they were up-regulated in one dataset under normal condition. Up-regulation of PEX4 was observed mostly under drought condition in maize, while it was up-regulated under normal condition in a rice data set. The same outcome was in case of PEX13. PEX7 and PEX14 were DEGs only in the rice datasets. PEX19A was upregulated in three datasets, two of which are the drought stress). Two transcripts of PEX22 were amongst DEGs. One of them was up-regulated under drought stress, while another one was up-regulated in the control group only in the SRP132035 dataset. zm00001d027511 (only one time up-regulated under drought) and zm00001d054044 (only one time up-regulated under drought) transcript of CAT3 gene were upregulated in control groups, except one and two datasets, respectively. CAT3 (zm00001d014848) was upregulated in three

datasets under drought condition. zm00001d014848 is the homolog of LOC_os6g51150 that was upregulated under drought. zm00001d027511 and zm00001d054044 were the homologs of LOC_os03g03910 and LOC_os002g02400, respectively. CAT3 expression pattern showed the same trend in the *Oryza sativa* datasets.

In *Sorghum bicolor*, DRP3B was upregulated only in two datasets. DRP5B was only upregulated in a control dataset. PEX11b and PEX11c were upregulated in one and two data sets, respectively. PEX11d was upregulated in only one control dataset. Two FIS1A genes were upregulated in one drought dataset. Two homologs of CAT3 were upregulated in one drought dataset. PEX22 was upregulated in three drought datasets. PEX4, PEX5, PEX10 and PEX14 were upregulated in one drought dataset.

4.2.7 Reverse Transcriptase Quantitative Polymerase Chain Reaction for Drought Stress Treatment

The expression of PEX11C and CAT genes were examined via qRT-PCR. Since the reference genome of *T. monococcum* has not been completed yet, the primer of the genes was designed using the peptide seqences of *T. aestivum* and then the peptide sequences were aligned to cDNA sequences of *T. monococcum* DV92 accesion that generated by jaiswallab (http://jaiswallab.cgrb.oregonstate.edu/genomics/wheat). Same primer was used to assess the expression of CAT gene. TmoDV92v1_019002 sequences were used to design PEX11C primer to use in qRT-PCR.

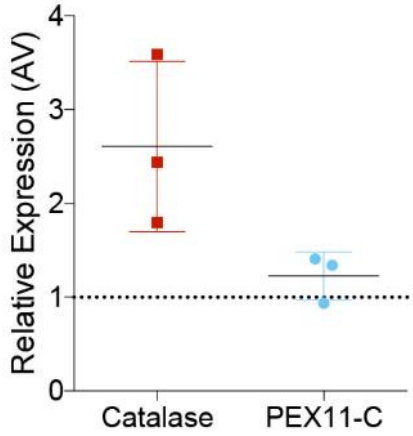


Figure 4.10. qRT-PCR outcome of CAT and PEX11C genes.

Given the qRT-PCR results for drought stress experiment, PEX11C and CAT genes are upregulated in biological replicates (Figure 4.10). CAT gene showed strong CAT upregulation response relative to PEX11C.

5. DISCUSSION

Around a hundred RNA-seq libraries that prepared at the different developmental stages were examined in this study. The outcomes of DEGs analysis revealed the varied transcription profiles in the leaf transcriptome comparisons. Additionally, the peroxisomal gene expression in terms of metabolisms that take place in peroxisome lumen were addressed, which is the first transcriptional evaluation of peroxisomal metabolic genes in the literature. The outcomes of transcriptome analysis were justified by picking up key genes that supports the hyposises of this dissertation such as BADH and CAT in growth and development, and CAT and PEX11C under drought stress.

High-rate genetic regulation was occurring in the immature and adult leaf transcriptomes that collected at the same growth stage. No significant genetic regulation was obseved in the adult leaf or early seedling transcriptome comparisons in terms of peroxisomal metabolism genes. The number of DEGs that involved in peroxisomal metabolism was the highest in juvenile and adult leaf comparison. This implies peroxisomal metabolism activities change during the developmental maturation which is controlled by independently regulated processes. The leaf maturation depends on the coordinated activity of interacting pathways. One of best example is miR156 which is necessary for the expression of the juvenile phase. miR156 operates by repressing the expression of SPL genes that act in pathways with different developmental functions. Likewise, miRNA176 operates the regulation of flowering time and floral organ identity, playing a role in vegetative phase change (49). The number of DEGs in comparison of juvenile and adult leaves points out that morphologically same tissue has altered gene expression profiles in the same developmental stages.

Rubisco, which is an enzyme that carboxylate ribulose-1,5-bisphosphate in the chloroplast, has a higher affinity either CO_2 or O_2. Once Ribusco is oxygenated in choloroplast, it forms 3-phosphoglycerate and 2-phosphoglycolate that is “toxic” compound for various enyzymes and inhibits most of the choloroplastic enzyme functions. The photorespiratory pathway converts 2-phosphoglycolate into glycolate which is processed in peroxisome and mitochondria. Glycolate are converted to glycerate by a series of enzyme-mediated reactions, which is

ultimately returned to the chloroplast to regenerate RuBP. The majority of genes that involved in the photorespiration process were upregulated in the adult phase, except *GGT* and *GOX* genes. Photosynthesis is one of growth and development responsive process in plants. As photostynthetic activity increases, photorespiration process is activated in the plant cells, therefore increament in gene transcription is expected outcome in photorespiration pathway. Polyamine catabolism and photorespiration is one of the main source of H_2O_2 which is catabolized by CAT and APX in the ROS metabolism. Hydrogen peroxide induces ROS metabolism in the peroxisome lumen. It is observed that all ROS metabolism genes were upregulated while developmental maturation occurs in leaves. Interestingly, only BADH gene is constantly upregulated among the genes that involved in polyamine catabolism. This phenomenon is interesting because the gene that encodes BADH protein is a multifunctional enzyme that is involved in various pathways e.g synthesis of glycine betaine which functions as osmotic protectant, NADP+(H) regeneration, catabolism of polamines and stress mechanisms etc… The expression of BADH gene can be strongly induced by various stresses. Given the expression pattern of BADH, we can say that BADH gene is developmentally responsive gene during maturation. Yan et al. (2012) (50) observed the overexpression of BADH gene in both transgenic T2 seeds and non-transgenic ones. The phenotypic characteristics of transgenic plants did not differ from non-transgenic plants under normal growing conditions, which indicated that overexpression of BADH had no adverse effect on plant growth, but the expression of BADH gene resulted in enhancement of the salt tolerance in alfaalfa transgenic lines via increasing betaine and proline accumulation. Similar studies have been perfomed in different plant species such as arabidopsis, tomato, potato, tabocca, einkorn etc… under stress conditions (51–54). Glycine betaine is an osmoprotectant, which electrically neutral at physiological pH and is a polar compound. The actual role is still debate, but it is thought that GB maintains the water balance between the plant cell and the environment. Moreover, it stabilizes macromolecules, enzyme activities, and membranes under stress conditions. Several studies conducted on introducing the GB pathway in plants that lack showed enhanced tolerance to drought and salinity stress. Transgenic expression of BADH gene increases the amount of GB content in plants under drought and salt stress (51). 6VS/6AS translocation caused a consistent upregulation of BADH at both transcription and translation levels in

Haynaldia villosa serving as potential genetic resources for wheat genetic improvement., playing roles in promoting GB content and the ability of drought resistance (55). Recent studies have shown that BADH gene improves plant tolerance through not only GB pathway, but also polyamine catabolism which forms GABA. BADH is broadly specific to its substrates and uses NADH as a cofactor to promote the conversion of 4-guanidinobutanal to GABA. Aditionally, BADH promote a bunch of enzymes that function in ROS metabolism.

It was reported the phosphorylation of BADH protein under salt stress in *T. monococcum* ssp. *monococcum* (53). However, we are not able to explain whether BADH protein undergoes post-translational modification in this study. The increment in Ca+ signaling can activate through MAPK and ABA signaling pathway under stress conditions. This pathway promotes the activation of SnRK2s and activated SnRK2s phosphorylates BADH, as well as superoxide dismutase (SOD), peroxidase (POD), catalase (CAT), ascorbate peroxidase (APX) and lipoxygenase (LOX), which are enzymes that are constantly upregulated in ROS metabolism in this study, implying the regulation mechanism of BADH and ROS metabolism is in common. Another report has revealed that BADH is the regulator of some ROS metabolic genes as well. Amino acid residues affect the basicity of the catalytic glutamate of the hydrolytic aldehydyde dehydrogenases, favoring NAD+ and NADP+ preference of BADH. NADPH is used in several processes, including conversion of GSSG into GSH by GR activity, NO generation by an L-arginine-dependent NOS activity and β-oxidation in peroxisomes. It has been reported several peroxisomal NADPH recycling enzymes which are isocitrate dehydrogenase, glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase. Isocitrate dehydrogenase function in nitrogen metabolism, redox control and stomatal movement (56,57). Glucose-6-phosphate dehydrogenase serves in pentose phospshate pathway, support of antioxidant system and NO generation (58). 6-phosphogluconate dehydrogenase is involved in pentose phospshate pathway and pollen-tube-ovule (59), whereas Glutathione reductase which is involved in ascorbate-glutatione pathway and allows the regeneration of the antioxidant GSH is NADPH consuming enzyme. As mentioned previously, the genes that is involved in ascorbate-glutatione pathway were upregulated during maturation and BADH protein regulates the transcription of some of those genes. Interestingly, to our knowledge, no one has addressed how BADH promote the

expression of ROS metabolism remains unclear. Given transcription pattern, BADH might promote ascorbate-glutathione pathway through the regeneration of NADPH that used as a cofactor by GR in peroxisome metabolism, but this hypothesis needs experimentally more elucidation.

GABA can be biosynthesized via glutamate decarboxylase and terminal oxidation of the diamine putrescine or the polyamine spermidine by the action of copper-containing amine oxidase (CuAO, EC 1.4.3.22) and FAD-dependent polyamine oxidase (PAO, E.C. 1.5.3.6). Peroxisomal aldehyde dehydrogenase 10A9 (ALDH10A9 which we called as BADH in this study) catalyzes the final step of Put oxidation by converting 4-aminobutanal to GABA. GABA is involved in many activities in plant growth and development. In flowering plants, GABA promotes sexual reproduction via pollen tube growth, fruit ripening, seed germination, elongation of primary and adventitious root and dark growing hypocotyls, affecting cell-wall carbon metabolism-related genes, phytohormone (IAA, ABA, and ethylene) signaling, stomatal conductance, nitrogen metabolism, anti-oxidative defense mechanism.,abiotic and biotic stresses (60).

GABA can be derived from polyamine catabolism in peroxisome, polyamine anabolism in the plastid and glutamate in cytosol. It is converted into succinic-semialdehyde by GABA-transaminase (GABA-T) for use in the tricarboxylic-acid (TCA) cycle (61). GABA permease is the influx transporter of GABA that mediate the transport of GABA to mitochondria. One study localized GABA permease in mitochondria but not choloroplast or peroxisome (62). GABA-efflux transporter was identified in wheat. Plant ALMTs (aluminum-activated malate transporters) can facilitate GABA efflux and influx to cell in response to $Al3^{+}$ at low pH and to anions at high pH (63,64). Some of the studies debate about existence of undiscovered GABA transporter(s) in the various compartments of plant cells, although the relative importance of this different GABA transporter(s) depends on stress and growth stage (65). Exogenous application of GABA significantly improved morphological growth of seedlings by regulating net photosynthesis, gas exchange capacities and chlorophyll biosynthesis (66). In this case, this evidence suggests the existence of GABA transporters that might be embedded in the cellular plasma, plastidal and peroxisomal membrane, however; the transporter(s) has not been identified yet. Notably, why the cell requires GABA in peroxisome lumen is still unclear. It raises our doubts that if no GABA

transporter exists in peroxisome membrane, it assigns a new function to GABA that have not still discovered yet. New discoveries in GABA transporters might offer us to understand the function of GABA in peroxisome lumen.

The main ROS are singlet oxygen ($^1O_2$), superoxide radical ($O^{2\cdot-}$), hydroxyl radical ($HO^{\cdot}$), and hydrogen peroxide (H_2O_2), which are the chemical products of reactions that accumulates at high rate under stress conditions. Hydrogen peroxide (H_2O_2) mainly generated by photoreduced superoxide radical via SOD activity in choloroplast and/or mitochondria. As mentioned previously, peroxisomes are involved in photorespiration, fatty acid β-oxidation and polyamine catabolism, which generate a large amount of H_2O_2 within peroxisomes. Catalases are predominately peroxisomal proteins that function as antioxidant enzymes to scavenge H_2O_2. Arabidopsis has three catalase genes each of which function in various physiological processes. CAT1 is mainly expressed in pollen and seeds, CAT2 mainly in photosynthetically active tissues, and CAT3 mainly in vascular tissues and senescent leaves. Catalase mutant showed severe growth defects with high hydrogen peroxide (H_2O_2) content and dwarft phenotype (67). Fatty acid β-oxidation in peroxisomes was impaired in catalase mutants because peroxisomal proteins were damaged by excess oxidation, indicating that peroxisome function was affected by catalase deletion (67). Beside this, the trancriptome analysis of this study suggests photorespiration and polyamine catabolism are another source of hydrogen peroxide (H_2O_2) that might affect the function peroxisome during growth and development. In addition to the rise in sensitivity to abiotic stress, around 2000 genes were differentially expressed in comparison of *cat* mutant and wild-type Arabidopsis (67), which may reveal the CAT signaling between peroxisome and nucleus. DEGs are mainly enriched in cellular processes, metabolic processes, biological regulation, developmental processes, and immune processes, as well as responses to stimuli and signaling (67), which is the same enriched biological processes that are revealed in this study.

The CAT gene expression suggests that catalases, while not being essential for survival, are important for plant development. Upregulation of CAT gene transcripts were verified by qRT-PCR in juvenile and adult stages. Futhermore, the expression of CAT gene was also verified by qRT-PCR under drought stress in the early seedling. CAT gene was upregulated at a higher rate than that of the expression of CAT gene that was examined in growth and development.

Peroxisome abundance colerates with hydrogen peroxide and ROS scavenging system, offering that ROS activity can be assessed by using peroxisome abundance (68). Moreover, our recent study reveals that peroxisome abundance varies among wheat cultivars (it is under review) depending on genetic background. It is reported that around 300 peroxisomal proteins have been annotated so far (46,69).

We analyzed transcription of all annotated peroxisomal genes in response to drought using 19 published RNA-Seq datasets from *Zea mays*, *Oryza sativa*, *Sorghum bicolor*, and *Arabidopsis thaliana*. We compared the GO term enrichment in the stressed versus control datasets to verify the impact of stress on global gene transcription. Genes involved in response to stress, to abiotic stimuli including drought, and stress-induced regulation of gene expression was enriched in the all RNA-Seq datasets for the stress samples relatively to control, which support the samples were under drought stress. The number of differentially expressed genes were arounf 75 to 120 across the species. However, bioinformatics analysis revealed that only catalase (CAT) and PEX11 were constantly upregulated in all RNA-Seq datasets. Interestingly, some members of CAT and PEX11 gene families were downregulated in response to drought. It means there is functional specialization amongst these gene families under stress. Accordingly, we verified CAT and PEX11 transcription through qRT-PCR in *T. monococcum*. They showed similar expression patterns with hexaploid wheat cultivars (publication under review). Peroxisomal proliferation is driven by a set of genes including PEX11. Of three wheat PEX11 genes, transcription of only PEX11C was constantly upregulated under drought in all species. Interestingly, some members of PEX11 gene families were down regulated in response to drought as PEX11A and PEX11B in this work.

High light, drought, salt and ABA all trigger ROS production which transcriptionally activates some PEX genes. In Arabidopsis, PEX11A is involved in the formation of peroxules under cadmium stress (70), while PEX11B is involved in light-induced peroxisome proliferation (71). Furthermore, PEX11B, C and D are upregulated under hypoxia and biotic stress whereas PEX11C is down regulated in ABA-treated samples (72). PEX11E is induced in response to salt stress (73). In this study and previous one which is under revision now, wheat species- einkorn and bread wheat, striking differences were seen between drought tolerant bread wheat and einkorn with respect to PEX gene expression. This indicates differences

in peroxisome proliferation process under drought and other stresses. Hence, CAT and PEX11C are good candidates of peroxisome proliferation biomarkers for drought tolerance, in terms of maintaining yield.

6. CONCLUSIONS AND RECOMMENDATIONS

1. Climate change in Turkey will cause changes to precipitation patterns and a rise in temperature. Increased temperature makes drier lands causing stresses such as salt, heat, and drought in crops. Drought can exacerbate the yield loss. To cope with the challenging future, developmental properties and patterns of crops would be crucial for breeding cultivars to reduce yield loss. Conventional breeding approaches are time-consuming and laborious, whereas DNA marker-assisted selection might be less prone to errors caused by environmental variability, and faster to reduce yield loss. Expanding the list of genetic markers for physiological, morphological, and molecular survival mechanisms is essential for breeding programs to cope with global warming.
2. Peroxisome harbors various metabolic pathways. Its function is important for cellular homeostasis. The most active metabolic pathway is photorespiration in peroxisome lumen. Increment in photorespiration produces ROS that triggers signal cascades, inducing the genes that involved in other metabolic pathways such as polyamine catabolism, NADPH regeneration, peroxisome fission and ascorbate-glutathione scavenging system.
3. The last step of polyamine catabolism biosynthesis GABA by BADH gene which is upregulated in the adult phase. BADH regulates ROS genes in peroxisome lumen by means of possible NADPH regeneration. However, this hypothesis needs more experimental evidence in addition to transcriptome analysis.
4. Polyamine catabolism regulates detrimental aldehyde content. Polyamine catabolism contains PAOs and CuOs and BADH genes, yet only BADH gene is upregulated in comparison with juvenile and adult leaves transcriptomes. Notably, we still cannot answer why other two genes are not differentially expressed.
5. GABA is a multifunctional molecule in plant cell. Possibly, GABA is used as signal molecule in peroxisome to trigger metabolic events.

GABA-P are mitochondrial GABA transporter that is not present in peroxisome, and we still do not know whether the GABA transporter exists in plasma or peroxisome membrane. Exogenous application of GABA improves the tolerance of plants against abiotic stress, suggesting the existence of GABA transporter in the plasma and/or peroxisome membrane. To understand this, we need to examine whether GABA is associated with peroxisome proliferation.

6. Drought reduces the soil moisture content. This is perceived by the root system, causing the stomatal closure. It affects the photosynthesis rate, leading to increase in ROS content. Peroxisome proliferation is induced to get rid of the detrimental effect of ROS compounds under drought stress, thus peroxisome might be a good tool to increase the tolerance of plants to drought stress. Notably, outcome of qRT-PCR validates the upregulation of CAT as well as BADH genes for growth and development study. Additionally, we identified that CAT gene is a stress-responsive and the biomarker of peroxisome proliferation as well as PEX11C peroxisome fission gene.
7. Both CAT and PEX11C are good candidate biomarkers for determining drought tolerant wheat species that show high peroxisome proliferation.

7. REFERENCES

The references of the study were prepared by using Mendeley software with Vancouver style.

1. Kün E. Serin iklim tahılları. Samsun: Ankara Üniv.Ziraat Fak.Yay; 1981.

2. Serpen A, Gökmen V, Karagöz A, Köksel H. Phytochemical quantification and total antioxidant capacities of Emmer (Triticum dicoccon Schrank) and Einkorn (Triticum monococcum L.) wheat landraces. J Agric Food Chem. 2008;56:7285–92.

3. Sahin Y, Yildirim A, Yucesan B, Zencirci N, Erbayram Ş, Gürel E. Phytochemical content and antioxidant activity of einkorn (Triticum monococcum ssp . monococcum), bread (Triticum aes- tivum L .), and durum (Triticum durum Desf .) wheat. Prog Nutr. 2017;19(10):450–9.

4. Mathan J, Bhattacharya J, Ranjan A. Enhancing crop yield by optimizing plant developmental features. Development. 2016;143(18):3283–94.

5. Meyer RS, Purugganan MD. Evolution of crop species: Genetics of domestication and diversification. Nat Rev Genet. 2013 Dec 18;14(12):840–52. Available from: https://www.nature.com/articles/nrg3605

6. Raman S, Greb T, Peaucelle A, Blein T, Laufs P, Theres K. Interplay of miR164, CUP-SHAPED COTYLEDON genes and LATERAL SUPPRESSOR controls axillary meristem formation in Arabidopsis thaliana. Plant J. 2008 Jul 1;55(1):65–76. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-313X.2008.03483.x

7. Reinhardt D, Pesce E-R, Stieger P, Mandel T, Baltensperger K, Bennett M, et al. Regulation of phyllotaxis by polar auxin transport. Nature. 2003;426(6964):255–60.

8. Zhao Z, Andersen SU, Ljung K, Dolezal K, Miotk A, Schultheiss SJ, et al. Hormonal control of the shoot stem-cell niche. Nature. 2010 Jun 24;465(7301):1089–92. Available from: https://www.nature.com/articles/nature09126

9. Tabata R, Ikezaki M, Fujibe T, Aida M, Tian CE, Ueno Y, et al. Arabidopsis AUXIN RESPONSE FACTOR6 and 8 regulate jasmonic acid biosynthesis and floral organ development via repression of class 1 KNOX genes. Plant Cell Physiol. 2010 Jan 1;51(1):164–75. Available from: www.pcp.oxfordjournals.org

10. Pfeiffer A, Janocha D, Dong Y, Medzihradszky A, Schöne S, Daum G, et al. Integration of light and metabolic signals for stem cell activation at the shoot apical meristem. Elife. 2016;5:e17023.

11. Capua Y, Eshed Y. Coordination of auxin-triggered leaf initiation by tomato LEAFLESS. Proc Natl Acad Sci U S A. 2017 Mar 21;114(12):3246–51. Available from: www.pnas.org/cgi/doi/10.1073/pnas.1617146114

12. Chandler JW, Jacobs B, Cole M, Comelli P, Werr W. DORNRÖSCHEN-LIKE expression marks Arabidopsis floral organ founder cells and precedes auxin response maxima. Plant Mol Biol. 2011 May 30;76(1–2):171–85. Available from: https://link.springer.com/article/10.1007/s11103-011-9779-8

13. Juarez MT, Twigg RW, Timmermans MCP. Specification of adaxial cell fate maize leaf development. Development. 2004 Sep 15;131(18):4533–44. Available from: https://dev.biologists.org/content/131/18/4533

14. Du Q, Wang H. The role of HD-ZIP III transcription factors and miR165/166 in vascular development and secondary cell wall formation. Plant Signal Behav. 2015;10(10):e10789551-5.

15. Ji L, Liu X, Yan J, Wang W, Yumul RE, Kim YJ, et al. ARGONAUTE10 and ARGONAUTE1 Regulate the Termination of Floral Stem Cells through Two MicroRNAs in Arabidopsis. PLoS Genet. 2011 Mar 31;7(3):e1001358. Available from: https://dx.plos.org/10.1371/journal.pgen.1001358

16. Wenkel S, Emery J, Hou BH, Evans MMS, Barton MK. A feedback regulatory module formed by Little Zipper and HD-ZIPIII genes. Plant Cell. 2007;19(11):3379–90.

17. Lin WC, Shuai B, Springer PS. The Arabidopsis LATERAL ORGAN BOUNDARIES-Domain Gene ASYMMETRIC LEAVES2 Functions in the Repression of KNOX Gene Expression and in Adaxial-Abaxial Patterning. Plant Cell. 2003 Oct;15(10):2241–52.

18. Pekker I, Alvarez JP, Eshed Y. Auxin response factors mediate Arabidopsis organ asymmetry via modulation of KANADI activity. Plant Cell. 2005;17(11):2899–910. Available from: /pmc/articles/PMC1276018/?report=abstract

19. Huang T, Kerstetter RA, Irish VF. APUM23, a PUF family protein, functions in leaf development and organ polarity in Arabidopsis. J Exp Bot. 2014 Mar;65(4):1181–91. Available from: /pmc/articles/PMC3935572/?report=abstract

20. Iwasaki M, Takahashi H, Iwakawa H, Nakagawa A, Ishikawa T, Tanaka H, et al. Dual regulation of ETTIN (ARF3) gene expression by AS1-AS2, which maintains the DNA methylation level, is involved in stabilization of leaf adaxial-abaxial partitioning in Arabidopsis. Development. 2013 May 1;140(9):1958–69. Available from: https://dev.biologists.org/content/140/9/1958

21. Stahle MI, Kuehlich J, Staron L, Von Arnim AG, Golz JF. YABBYs and the transcriptional corepressors LEUNIG and LEUNIG-HOMOLOG maintain leaf polarity and meristem activity in Arabidopsis. Plant Cell. 2009 Oct 1;21(10):3105–18. Available from: www.plantcell.org/cgi/doi/10.1105/tpc.109.070458

22. Chan MH, Ji HJ, Hong GN, Fletcher JC. BLADE-ON-PETIOLE1 and 2 control Arabidopsis lateral organ fate through regulation of LOB domain and adaxial-abaxial polarity genes. Plant Cell. 2007 Jun;19(6):1809–25. Available from: /pmc/articles/PMC1955725/?report=abstract

23. Hawker NP, Bowman JL. Roles for class III HD-zip and Kanadi genes in arabidopsis root development. Plant Physiol. 2004 Aug;135(4):2261–70. Available from: /pmc/articles/PMC520795/?report=abstract

24. Satterlee JW, Scanlon MJ. Coordination of leaf development across developmental axes [Internet]. Vol. 8, Plants. MDPI AG; 2019. p. 1–19. Available from: /pmc/articles/PMC6843618/?report=abstract

25. Ohmori Y, Tanaka W, Kojima M, Sakakibara H, Hirano HY. WUSCHEL-RELATED HOMEOBOX4 Is involved in meristem maintenance and is negatively regulated by the CLE gene FCP1 in rice. Plant Cell. 2013;25(1):229–41. Available from: /pmc/articles/PMC3584538/?report=abstract

26. Betsuyaku S, Sawa S, Yamada M. The Function of the CLE Peptides in Plant Development and Plant-Microbe Interactions. Arab B. 2011 Jan;9:e0149. Available from: /pmc/articles/PMC3268505/?report=abstract

27. Qi J, Wang Y, Yu T, Cunha A, Wu B, Vernoux T, et al. Auxin depletion from leaf primordia contributes to organ patterning. Proc Natl Acad Sci U S A. 2014 Dec 30;111(52):18769–74.

28. Nakata M, Okada K. The three-domain model: A new model for the early development of leaves in Arabidopsis thaliana. Plant Signal Behav. 2012;9:e0149. Available from: /pmc/articles/PMC3548863/?report=abstract

29. Zhang F, Tadege M. Repression of AS2 by WOX family transcription factors is required for leaf development in medicago and arabidopsis. Plant Signal Behav. 2015;10(7):e993291. Available from: /pmc/articles/PMC4623463/?report=abstract

30. Ram H, Sahadevan S, Gale N, Caggiano MP, Yu X, Ohno C, et al. An integrated analysis of cell-type specific gene expression reveals genes regulated by REVOLUTA and KANADI1 in the Arabidopsis shoot apical meristem. PLoS Genet. 2020 Apr 1;16(4):e1008661. Available from: https://doi.org/10.1371/journal.pgen.1008661

31. Nakata M, Matsumoto N, Tsugeki R, Rikirsch E, Laux T, Okada K. Roles of the middle domain-specific WUSCHEL-RELATED HOMEOBOX genes in early development of leaves in Arabidopsis. Plant Cell. 2012 Feb 1;24(2):519–35. Available from: www.plantcell.org/cgi/doi/10.1105/tpc.111.092858

32. Lan J, Qin G. The regulation of CIN-like TCP transcription factors [Internet]. Vol. 21, International Journal of Molecular Sciences. MDPI AG; 2020. p. 1–17. Available from: /pmc/articles/PMC7349945/?report=abstract

33. Challa KR, Rath M, Nath U. The CIN-TCP transcription factors promote commitment to differentiation in Arabidopsis leaf pavement cells via both auxin-dependent and independent pathways. PLoS Genet. 2019 Feb 1;15(2). Available from: /pmc/articles/PMC6386416/?report=abstract

34. Singh J, Zhao J, Vallejos CE. Differential transcriptome patterns associated with early seedling development in a wild and a domesticated common bean (Phaseolus vulgaris L.) accession. Plant Sci. 2018 Sep 1;274:153–62.

35. Ji Y, Chen G, Zheng X, Zhong Q, Zhang M, Wu Z, et al. Comprehensive Transcriptome Reveals an Opposite Regulatory Effect of Plant Growth Retardants in Controlling Seedling Overgrowth between Roots and Shoots. Int J Mol Sci. 2019 Jul 5;20(13):3307. Available from: https://www.mdpi.com/1422-0067/20/13/3307

36. Li P, Ponnala L, Gandotra N, Wang L, Si Y, Tausta SL, et al. The developmental dynamics of the maize leaf transcriptome. Nat Genet. 2010;42(12):1060–7. Available from: http://dx.doi.org/10.1038/ng.703

37. Shi Y, Chen J, Hou X. Similarities and Differences of Photosynthesis Establishment Related mRNAs and Novel lncRNAs in Early Seedlings (Coleoptile/Cotyledon vs. True Leaf) of Rice and Arabidopsis. Front Genet. 2020 Sep 8;11:1063. Available from: www.frontiersin.org

38. Trevisan S, Forestan C, Brojanigo S, Quaggiotti S, Varotto S. Brassinosteroid application affects the growth and gravitropic response of maize by regulating gene expression in the roots, shoots and leaves. Plant Growth Regul. 2020 Sep 1;92(1):117–30. Available from: https://doi.org/10.1007/s10725-020-00626-z

39. Wu X, Bacic A, Johnson KL, Humphries J. The Role of Brachypodium distachyon Wall-Associated Kinases (WAKs) in Cell Expansion and Stress Responses. Cells. 2020 Nov 14;9(11):2478. Available from: https://www.mdpi.com/2073-4409/9/11/2478

40. Zhao M, Tang S, Zhang H, He M, Liu J, Zhi H, et al. DROOPY LEAF1 controls leaf architecture by orchestrating early brassinosteroid signaling. Proc Natl Acad Sci U S A. 2020 Sep 1;117(35):21766–74. Available from: https://www.pnas.org/content/117/35/21766

41. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20.

42. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc. 2016;11(9):1650–67. Available from: http://dx.doi.org/10.1038/nprot.2016-095

43. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19505943

44. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):1–21.

45. Alexa A, Rahnenfuhrer J. topGO: Enrichment Analysis for Gene Ontology. R Packag version 2460. 2021;

46. Kaur N, Hu J. Defining the plant peroxisomal proteome: From arabidopsis to rice. Front Plant Sci. 2011 Dec 27;2(DEC):103.

47. Pan R, Liu J, Wang S, Hu J. Peroxisomes: versatile organelles with diverse roles in plants. New Phytol. 2020 Feb 1;225(4):1410–27. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/nph.16134

48. Tounsi S, Kamoun Y, Feki K, Jemli S, Saïdi MN, Ziadi H, et al. Localization and expression analysis of a novel catalase from Triticum monococcum TmCAT1 involved in response to different environmental stresses. Plant Physiol Biochem. 2019 Jun 1;139:366–78.

49. Wu G, Park MY, Conway SR, Wang JW, Weigel D, Poethig RS. The sequential action of miR156 and miR172 regulates developmental timing in Arabidopsis. Cell. 2009 Aug 21;138(4):750. Available from: /pmc/articles/PMC2732587/

50. Yan LP, Liu CL, Liang HM, Mao XH, Wang F, Pang CH, et al. Physiological responses to salt stress of T2 alfalfa progenies carrying a transgene for betaine aldehyde dehydrogenase. Plant Cell Tissue Organ Cult. 2012 Feb 22;108(2):191–9. Available from: https://link.springer.com/article/10.1007/s11240-011-0027-3

51. Zhang N, Si HJ, Wen G, Du HH, Liu BL, Wang D. Enhanced drought and salinity tolerance in transgenic potato plants with a BADH gene from spinach. Plant Biotechnol Rep. 2011 Jan 1;5(1):71–7.

52. Jia GX, Zhu ZQ, Chang FQ, Li YX. Transformation of tomato with the BADH gene from Atriplex improves salt tolerance. Plant Cell Rep. 2002;21:141–6.

53. Lv DW, Zhu GR, Zhu D, Bian YW, Liang XN, Cheng ZW, et al. Proteomic and phosphoproteomic analysis reveals the response and defense mechanism in leaves of diploid wheat T. monococcum under salt stress and recovery. J Proteomics. 2016 Jun 30;143:93–105.

54. Li QL, Gao XR, Yu XH, Wang XZ, An LJ. Molecular cloning and characterization of betaine aldehyde dehydrogenase gene from Suaeda liaotungensis and its use in improved tolerance to salinity in transgenic tobacco. Biotechnol Lett. 2003 Sep;25(17):1431–6. Available from: https://pubmed.ncbi.nlm.nih.gov/14514045/

55. Zou R, Wu J su, Wang R mei, Yan Y ming. Grain proteomic analysis reveals central stress responsive proteins caused by wheat-Haynaldia villosa 6VS/6AL translocation. J Integr Agric. 2020 Nov 1;19(11):2628–42.

56. Corpas FJ, Barroso JB, Sandalio LM, Palma JM, Lupiáñez JA, Del Río LA. Peroxisomal NADP-Dependent Isocitrate Dehydrogenase. Characterization and Activity Regulation during Natural Senescence. Plant Physiol. 1999 Nov 1;121(3):921–8. Available from: https://academic.oup.com/plphys/article/121/3/921/6098676

57. Leterrier M, Barroso JB, Valderrama R, Begara-Morales JC, Sánchez-Calvo B, Chaki M, et al. Peroxisomal NADP-isocitrate dehydrogenase is required for Arabidopsis stomatal movement. Protoplasma. 2016 Mar 1;253(2):403–15. Available from: https://link.springer.com/article/10.1007/s00709-015-0819-0

58. Corpas FJ, Barroso JB, Sandalio LM, Distefano S, Palma JM, Lupiáñez JA, et al. A dehydrogenase-mediated recycling system of NADPH in plant peroxisomes. Biochem J. 1998 Mar 1;330(Pt 2):777. Available from: /pmc/articles/PMC1219205/?report=abstract

59. Hölscher C, Lutterbey MC, Lansing H, Meyer T, Fischer K, von Schaewen A. Defects in Peroxisomal 6-Phosphogluconate Dehydrogenase Isoform PGD2 Prevent Gametophytic Interaction in Arabidopsis thaliana. Plant Physiol. 2016 May 2;171(1):192–205. Available from: https://academic.oup.com/plphys/article/171/1/192/6114929

60. Li L, Dou N, Zhang H, Wu C. The versatile GABA in plants. Plant Signal Behav. 2021;16(3):1862565. Available from: /pmc/articles/PMC7889023/

61. Clark SM, Di Leo R, Van Cauwenberghe OR, Mullen RT, Shelp BJ. Subcellular localization and expression of multiple tomato γ-aminobutyrate transaminases that utilize both pyruvate and glyoxylate. J Exp Bot. 2009 Jul 1;60(11):3255–67. Available from: https://academic.oup.com/jxb/article/60/11/3255/534622

62. Michaeli S, Fait A, Lagor K, Nunes-Nesi A, Grillich N, Yellin A, et al. A mitochondrial GABA permease connects the GABA shunt and the TCA cycle, and is essential for normal carbon metabolism. Plant J. 2011 Aug 1;67(3):485–98. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-313X.2011.04612.x

63. Ramesh SA, Tyerman SD, Xu B, Bose J, Kaur S, Conn V, et al. GABA signalling modulates plant growth by directly regulating the activity of plant-specific anion transporters. Nat Commun. 2015 Jul 29;6:7879. Available from: https://www.nature.com/ articles/ncomms8879

64. Ramesh SA, Kamran M, Sullivan W, Chirkova L, Okamoto M, Degryse F, et al. Aluminum-Activated Malate Transporters Can Facilitate GABA Transport. Plant Cell. 2018 Jun 8;30(5):1147–64. Available from: https://academic.oup.com/plcell/article/30/5/1147/ 6099104

65. Zarei A, Trobacher CP, Shelp BJ. Arabidopsis aldehyde dehydrogenase 10 family members confer salt tolerance through putrescine-derived 4-aminobutyrate (GABA) production. Sci Rep. 2016 Oct 11;6:35115. Available from: /pmc/articles/PMC5057122/

66. Li W, Liu J, Ashraf U, Li G, Li Y, Lu W, et al. Exogenous γ-aminobutyric Acid (GABA) application improved early growth, net photosynthesis, and associated physio-biochemical events in Maize. Front Plant Sci. 2016 Jun 22;7:919.

67. Su T, Wang P, Li H, Zhao Y, Lu Y, Dai P, et al. The Arabidopsis catalase triple mutant reveals important roles of catalases and peroxisome-derived signaling in plant development. J Integr Plant Biol. 2018 Jul 1;60(7):591–607. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/jipb.12649

68. Smertenko A. Can Peroxisomes Inform Cellular Response to Drought? Trends Plant Sci. 2017 Dec 1;22(12):1005–7. Available from: http://www.cell.com/article/ S1360138517302200/ fulltext

69. Reumann S, Babujee L, Changle M, Wienkoop S, Siemsen T, Antonicelli GE, et al. Proteome Analysis of Arabidopsis Leaf Peroxisomes Reveals Novel Targeting Peptides, Metabolic Pathways, and Defense Mechanisms. Plant Cell. 2007 Dec 3;19(10):3170–93. Available from: https://academic.oup.com/plcell/article/19/10/3170/6092346

70. Rodríguez-Serrano M, Romero-Puertas MC, Sanz-Fernández M, Hu J, Sandalio LM. Peroxisomes Extend Peroxules in a Fast Response to Stress via a Reactive Oxygen Species-Mediated Induction of the Peroxin PEX11a. Plant Physiol. 2016 Jul 6;171(3):1665–74. Available from: https://academic.oup.com/plphys/article/171/3/1665/6115226

71. Desai M, Hu J. Light Induces Peroxisome Proliferation in Arabidopsis Seedlings through the Photoreceptor Phytochrome A, the Transcription Factor HY5 HOMOLOG, and the Peroxisomal Protein PEROXIN11b. Plant Physiol. 2008 Mar 3;146(3):1117–27. Available from: https://academic.oup.com/plphys/article/146/3/1117/6107274

72. Ebeed HT, Stevenson SR, Cuming AC, Baker A. Conserved and differential transcriptional responses of peroxisome associated pathways to drought, dehydration and ABA. J Exp Bot. 2018;69(20):4971–85.

73. Charlton WL, Matsui K, Johnson B, Graham IA, Ohme-Takagi M, Baker A. Salt-induced expression of peroxisome-associated genes requires components of the ethylene, jasmonate and abscisic acid signalling pathways. Plant Cell Environ. 2005 Apr 1;28(4):513–24. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-3040.2004. 01293.x

8. APPENDICES

Appendix 1 Expression Patterns of Peroxisomal Genes in Growth and Development

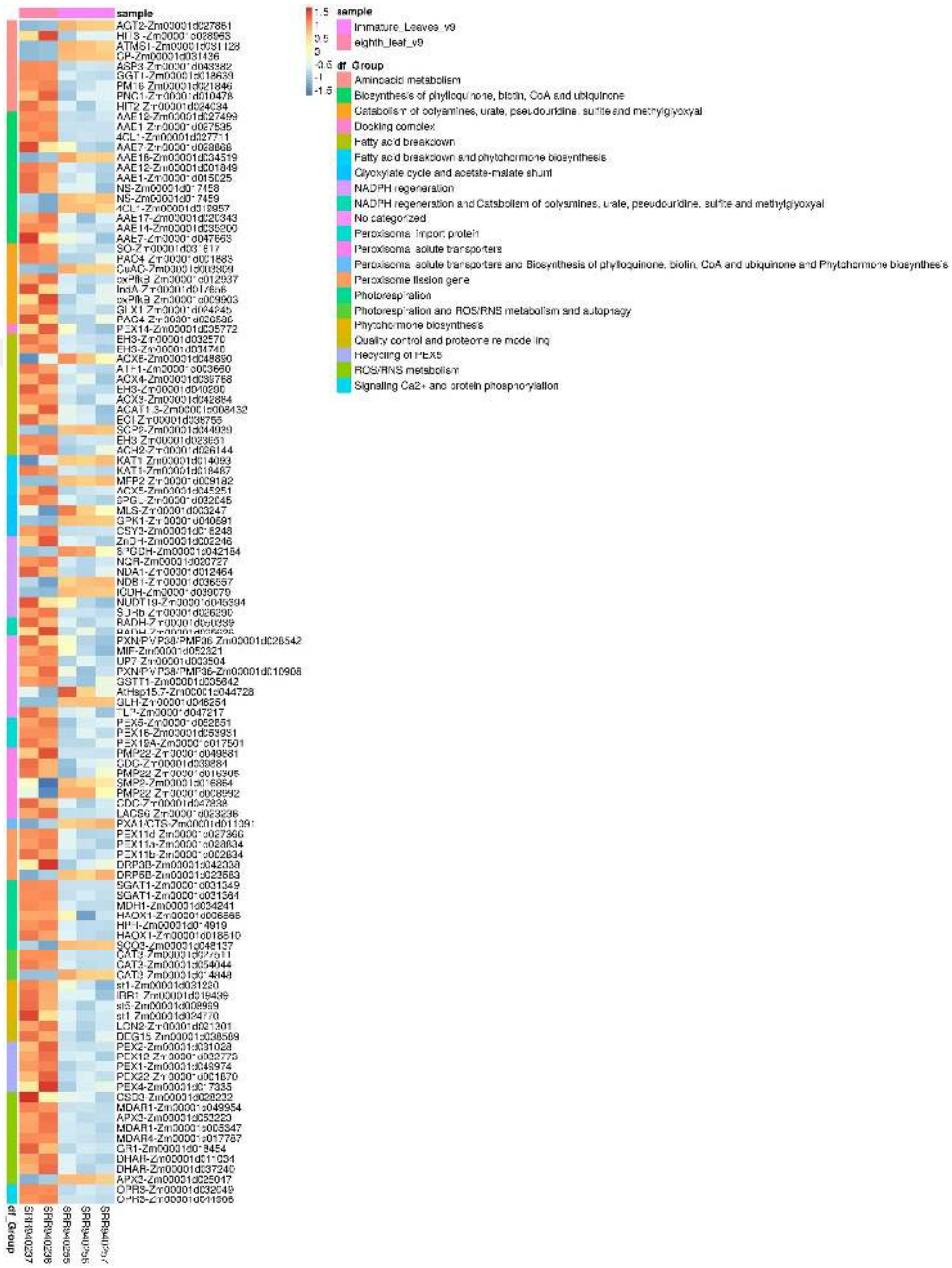


Figure 1.1. Expression patterns of peroxisome related DEGs in comparison with immature and eigth leaf at v9 stage. Heatmap was generated based on the z-score transformed vst.

Figure 1.2. Expression patterns of peroxisome related DEGs in comparison with immature and eleventh leaf at v9 stage. Heatmap was generated based on the z-score transformed vst.

Figure 1.3. Expression patterns of peroxisome related DEGs in comparison with immature and thirteenth leaf at v9 stage. Heatmap was generated based on the z-score transformed vst.

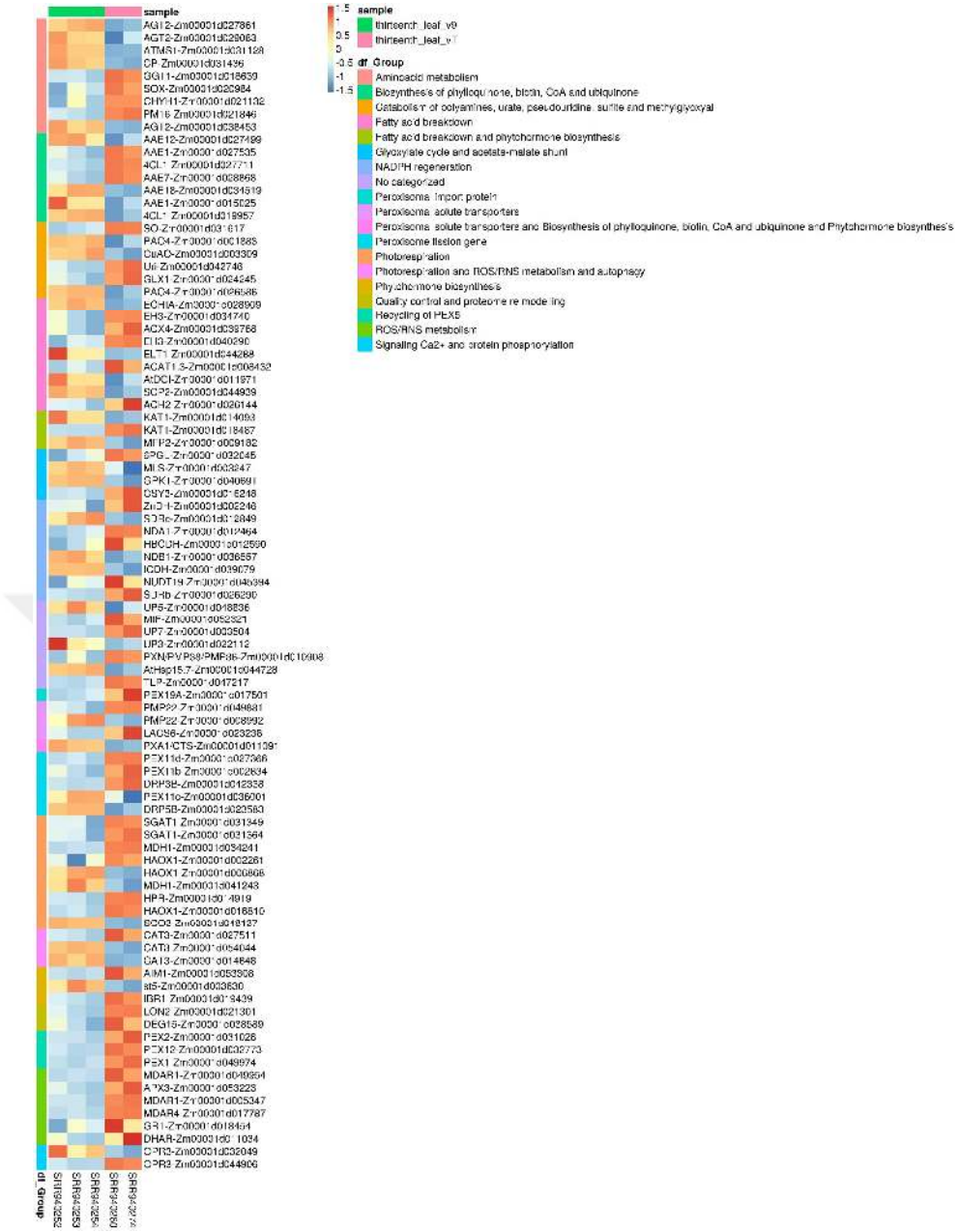


Figure 1.4. Expression patterns of peroxisome related DEGs in comparison with v9 and vT stage. Heatmap was generated based on the z-score transformed vst.

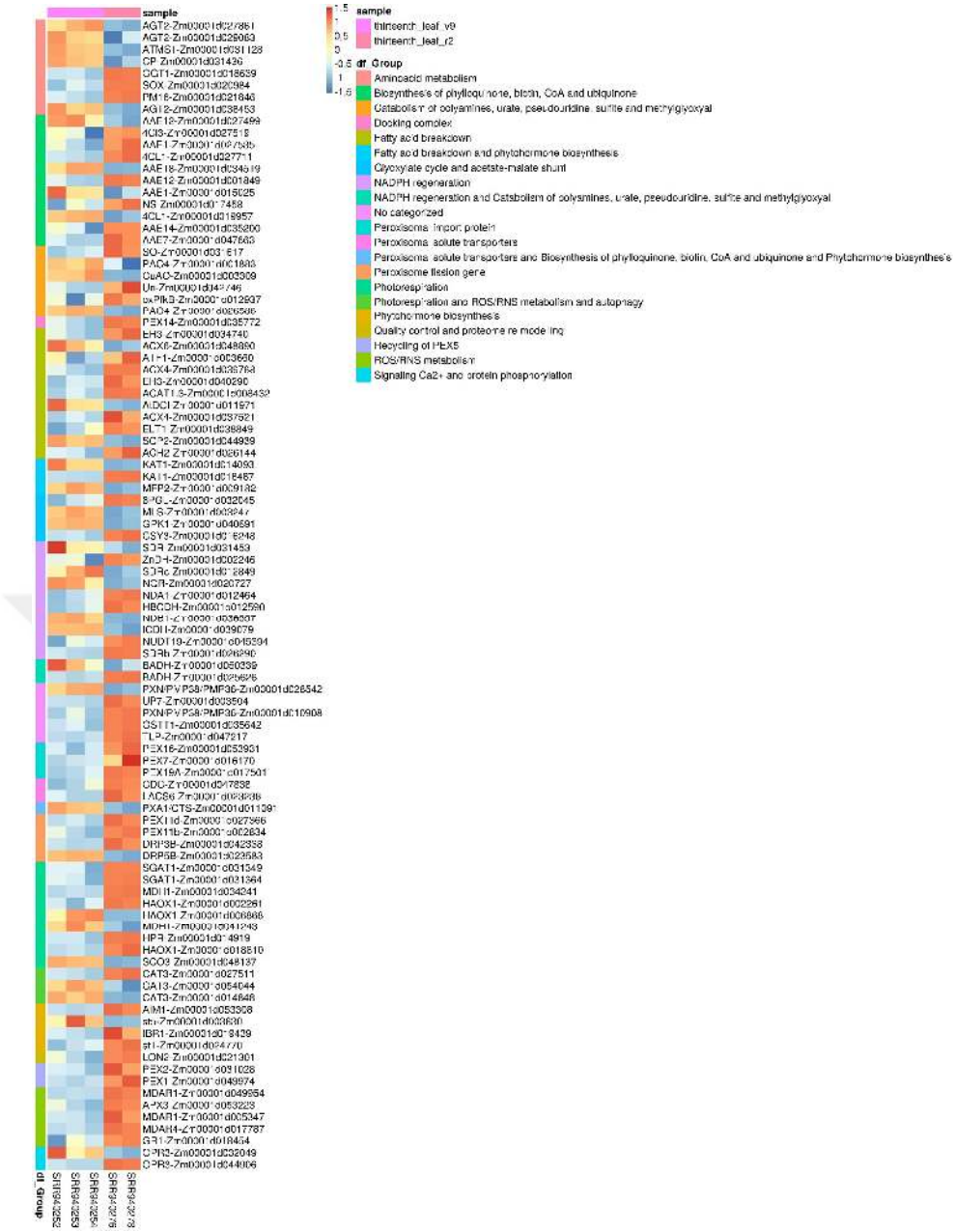


Figure 1.5. Expression patterns of peroxisome related DEGs in comparison with v9 and r2 stage. Heatmap was generated based on the z-score transformed vst.

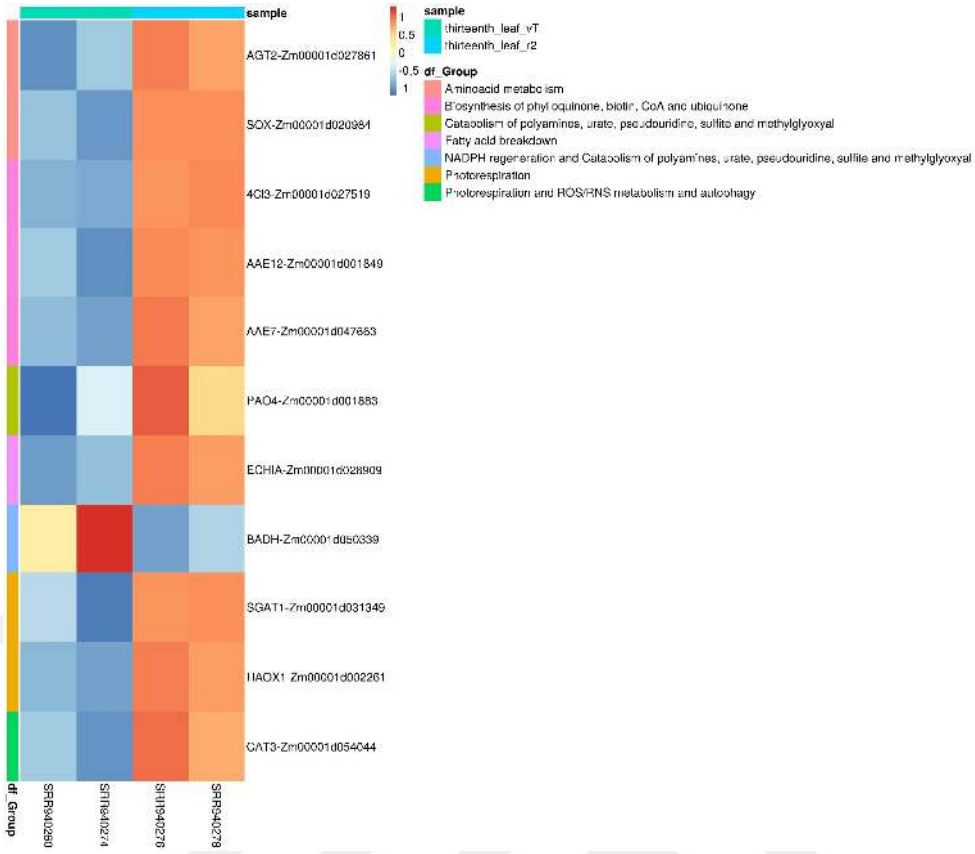


Figure 1.6. Expression patterns of peroxisome related DEGs in comparison with vT and r2 stage. Heatmap was generated based on the z-score transformed vst.

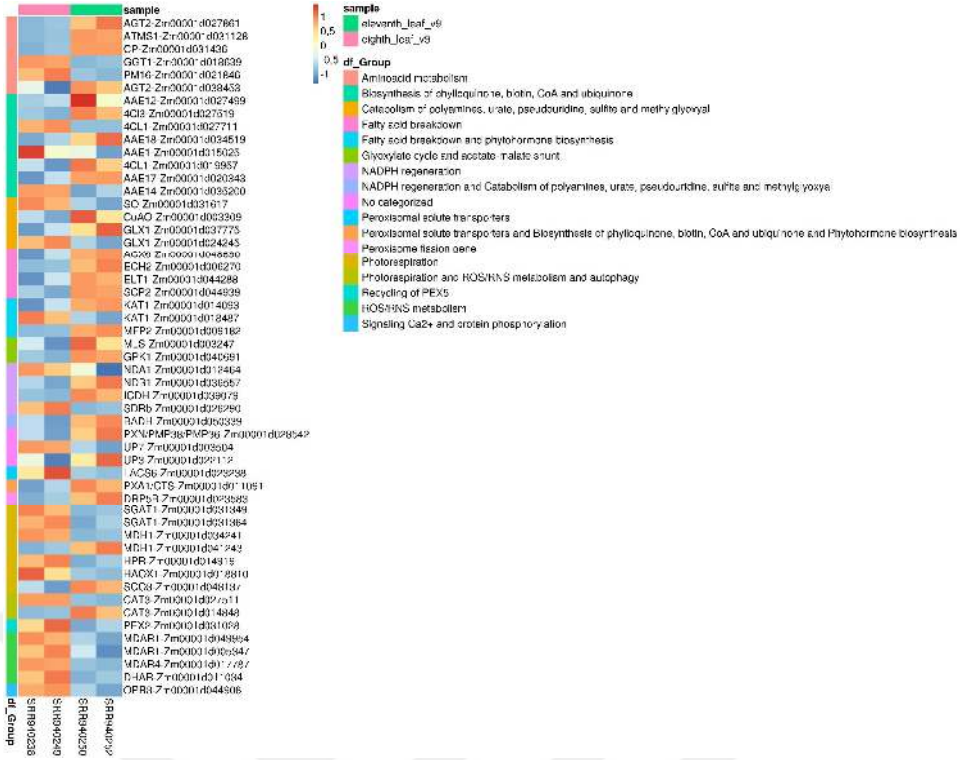


Figure 1.7. Expression patterns of peroxisome related DEGs in comparison with eigth and eleventh leaf at same development stage. Heatmap was generated based on the z-score transformed vst.

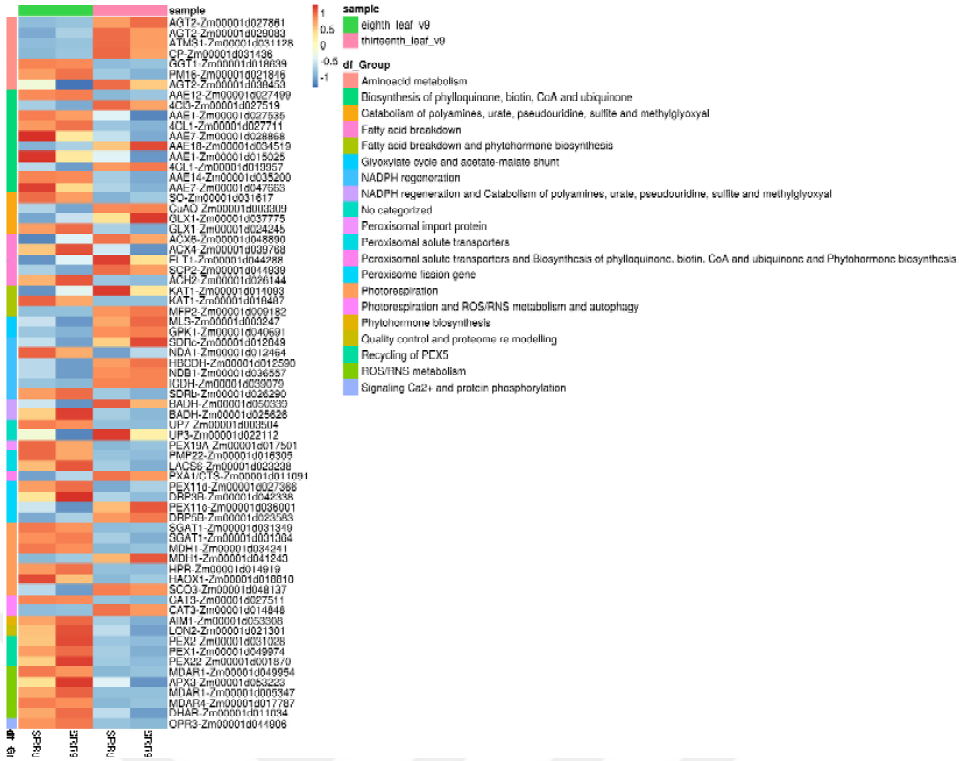


Figure 1.8. Expression patterns of peroxisome related DEGs in comparison with eighth and thirteenth leaf at same development stage. Heatmap was generated based on the z-score transformed vst.

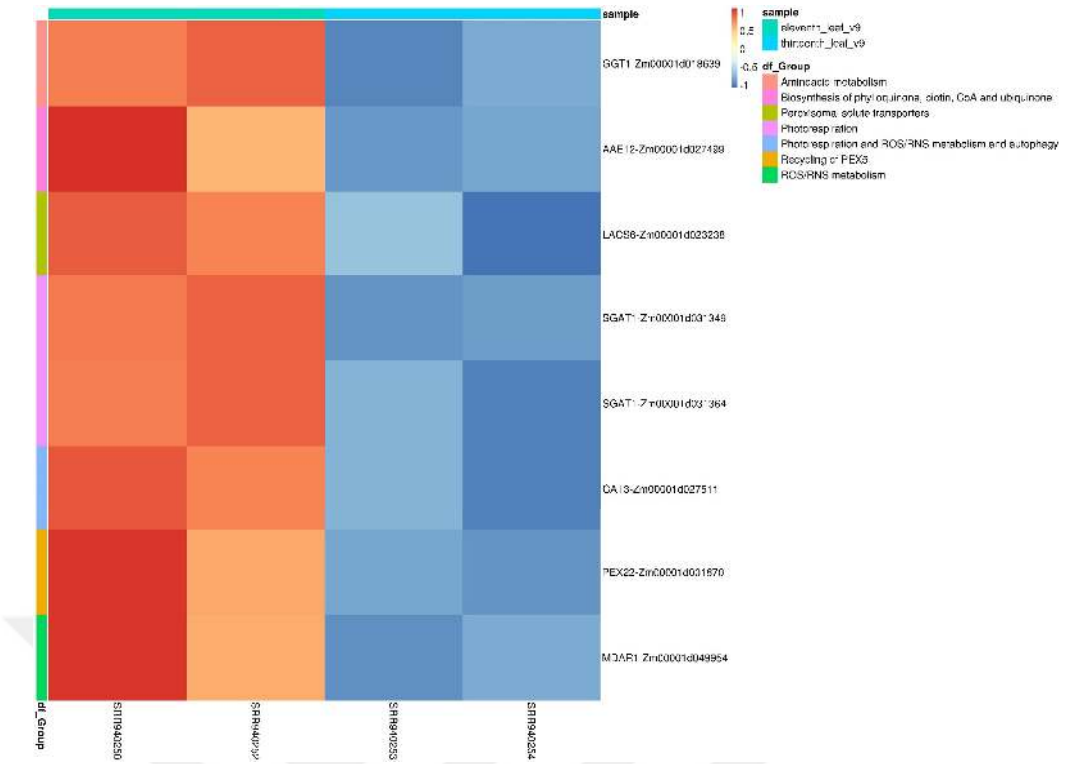


Figure 1.9. Expression patterns of peroxisome related DEGs in comparison with eleventh and thirteenth leaf at same development stage. Heatmap was generated based on the z-score transformed vst.

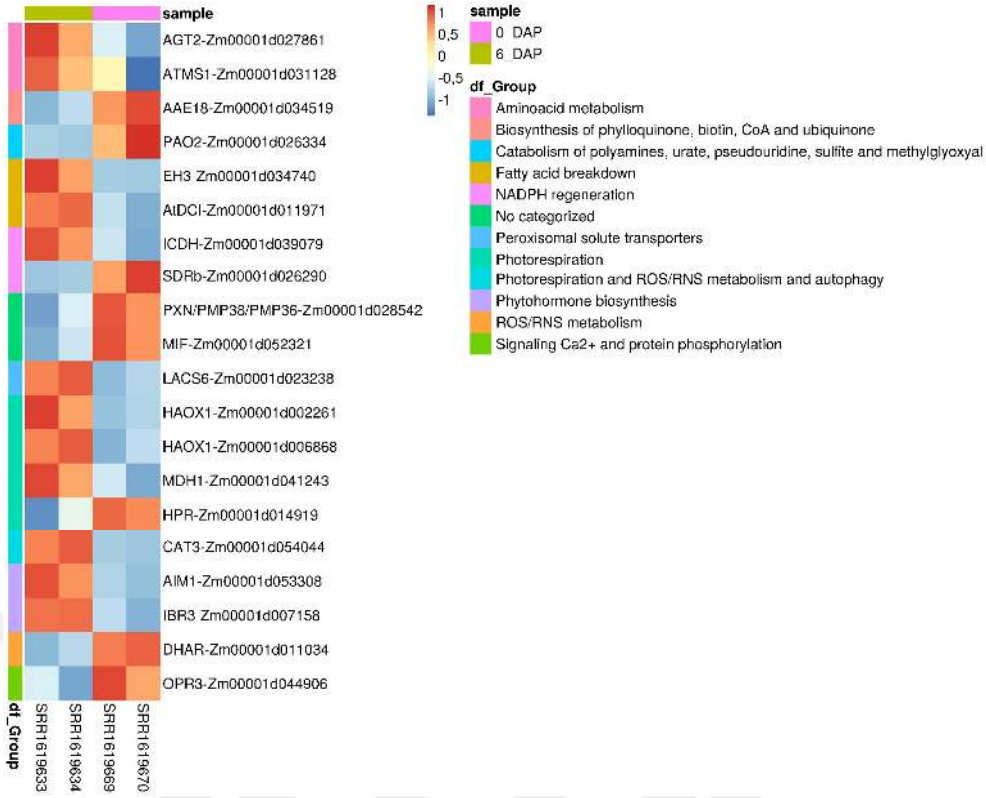


Figure 1.10. Expression patterns of peroxisome related DEGs in comparison with 0 DAP and 6 DAP at reproductive stage. Heatmap was generated based on the z-score transformed vst.

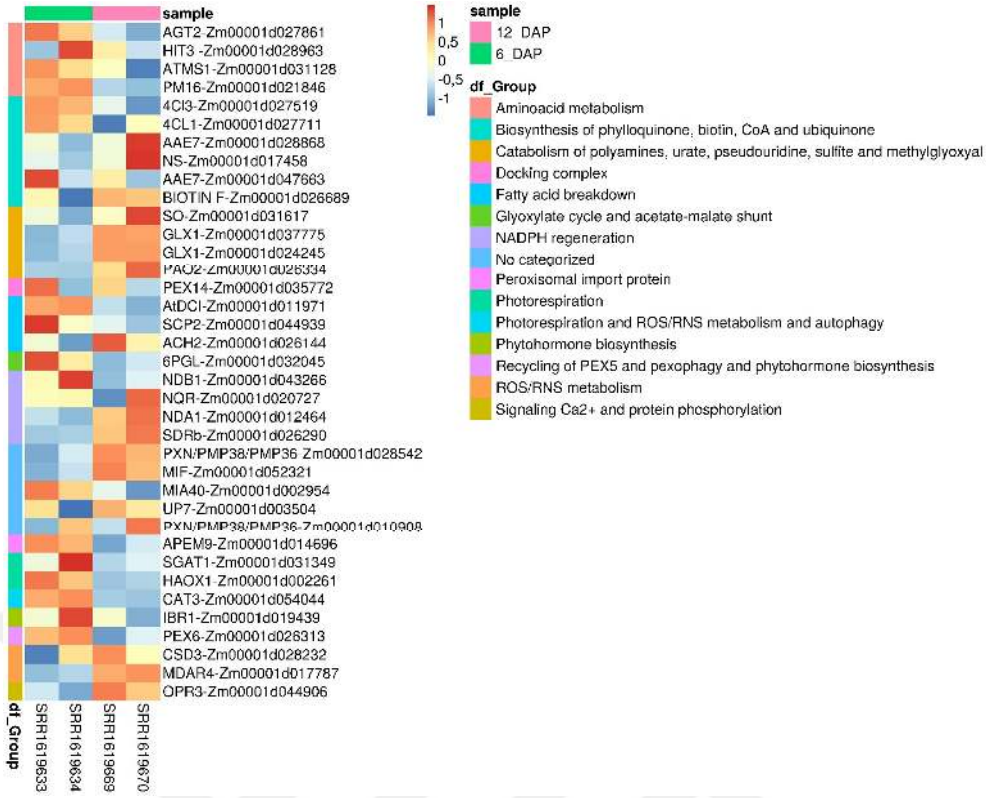


Figure 1.11. Expression patterns of peroxisome related DEGs in comparison with 6 DAP and 12 DAP at reproductive stage. Heatmap was generated based on the z-score transformed vst.

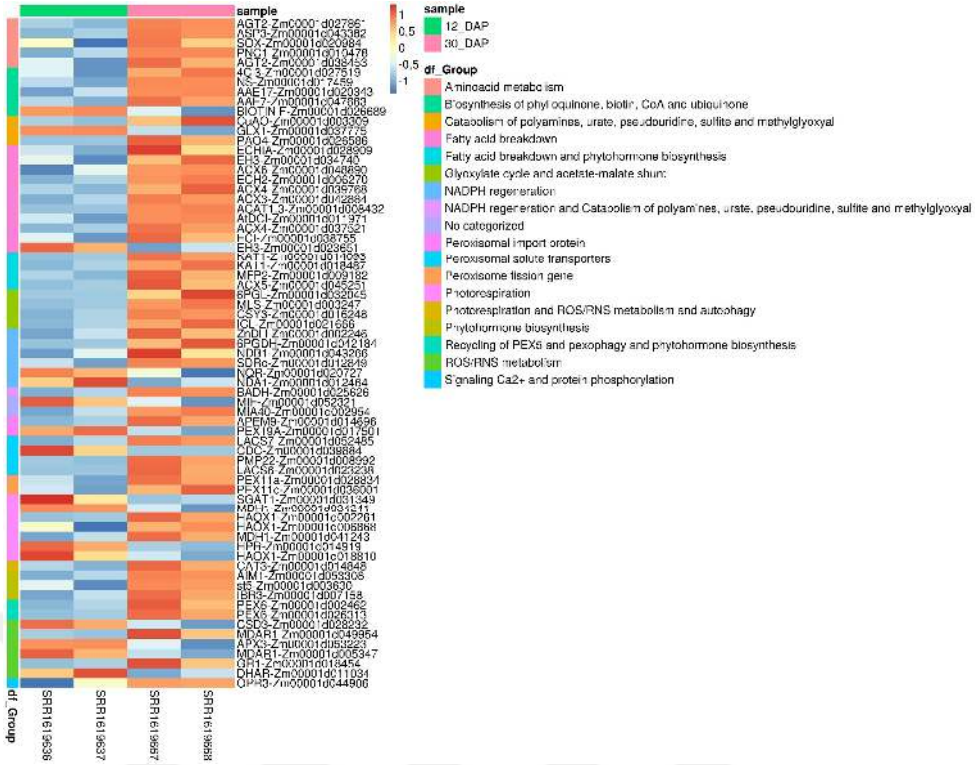


Figure 1.12. Expression patterns of peroxisome related DEGs in comparison with 12 DAP and 30 DAP at reproductive stage. Heatmap was generated based on the z-score transformed vst.

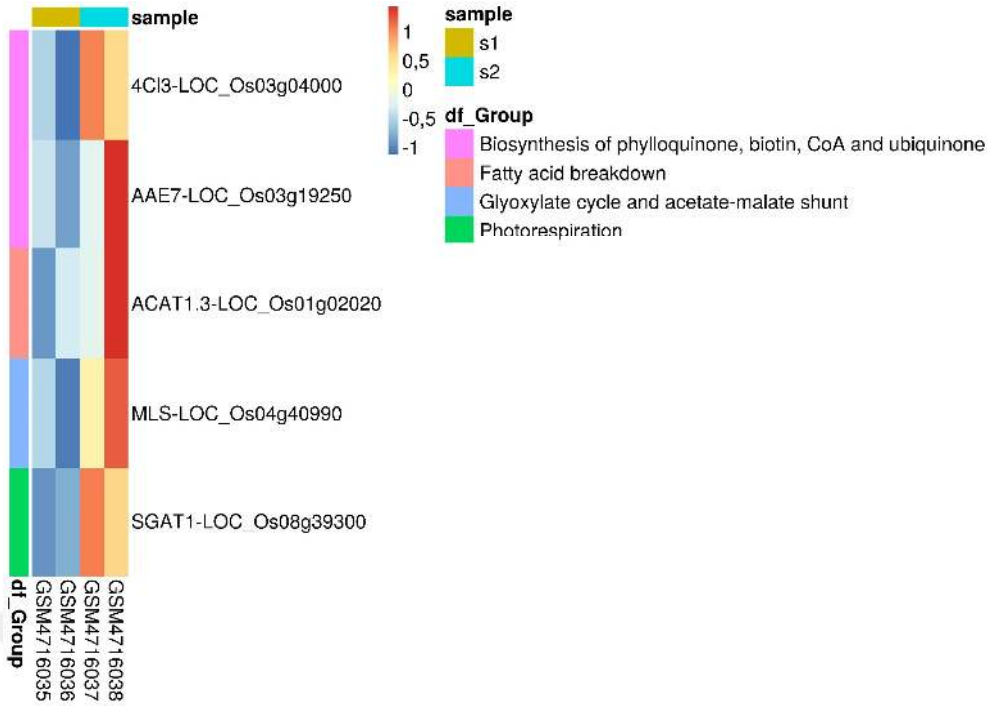


Figure 1.13. Expression patterns of peroxisome related DEGs in comparison with s1 and s2 at seedling stage. Heatmap was generated based on the z-score transformed vst.

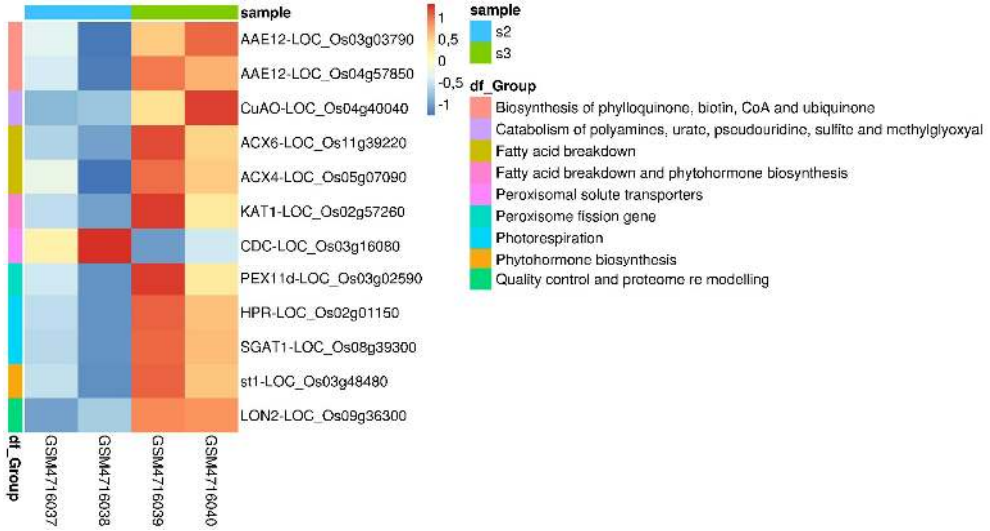


Figure 1.14. Expression patterns of peroxisome related DEGs in comparison with s3 and s2 at seedling stage. Heatmap was generated based on the z-score transformed vst.

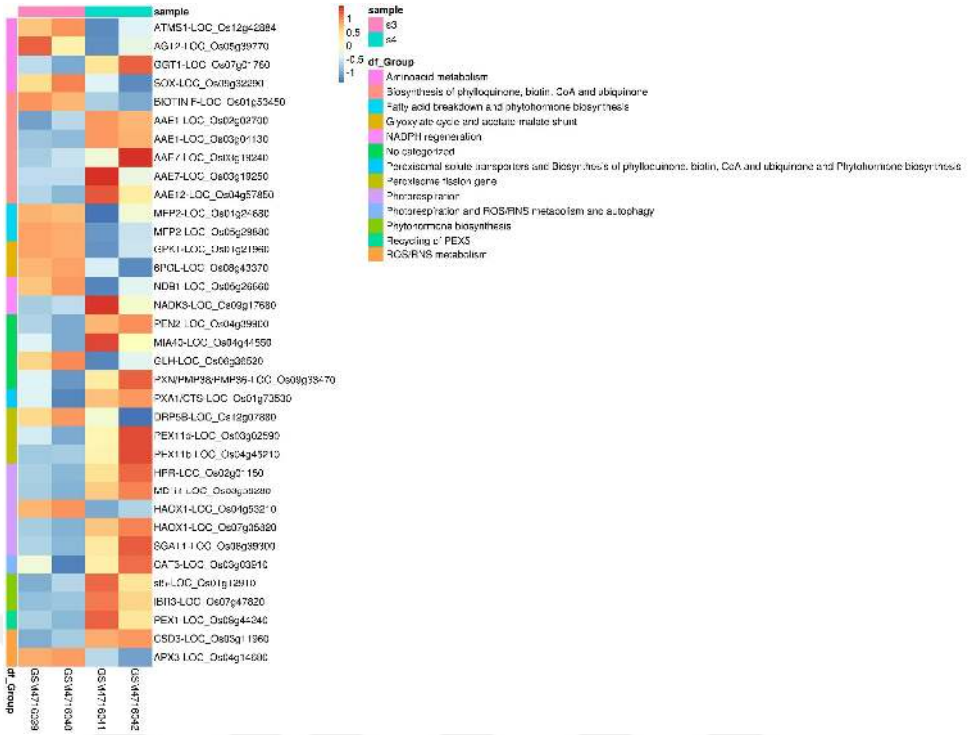


Figure 1.15. Expression patterns of peroxisome related DEGs in comparison with s3 and s4 at seedling stage. Heatmap was generated based on the z-score transformed vst.

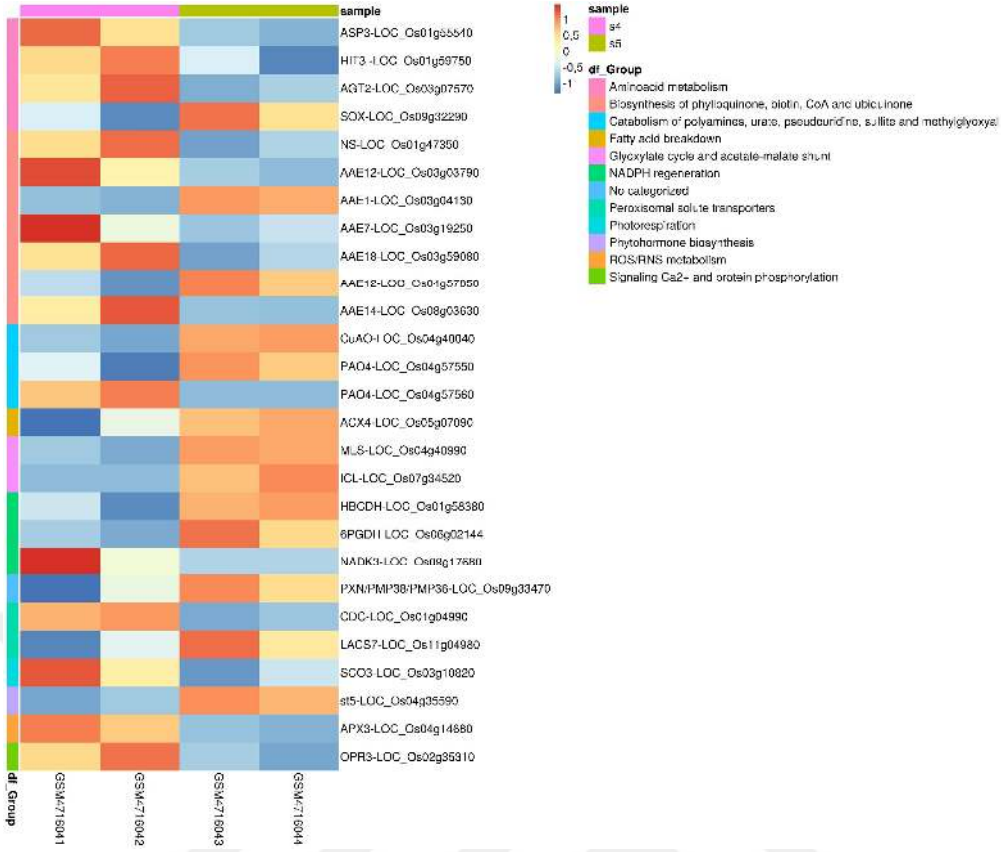
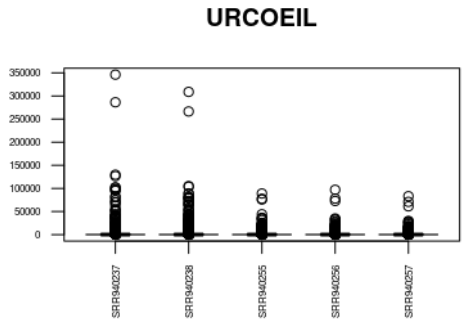
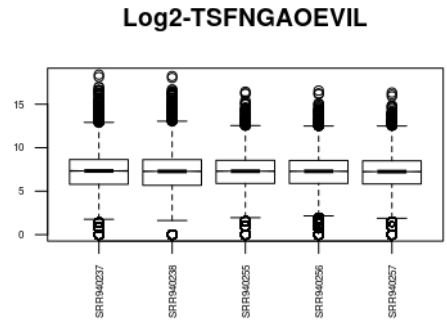
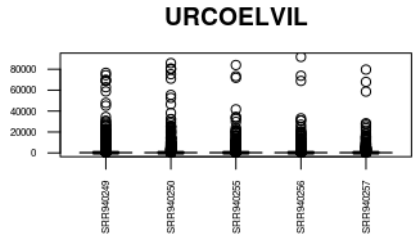


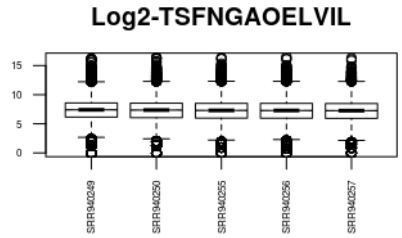
Figure 1.16. Expression patterns of peroxisome related DEGs in comparison with s4 and s5 at seedling stage. Heatmap was generated based on the z-score transformed vst.

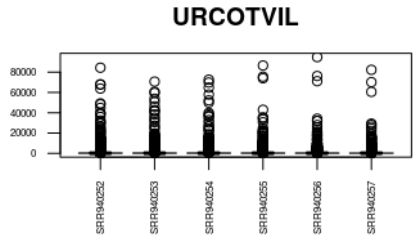
Appendix 2. Normalization and transformation of data sets.

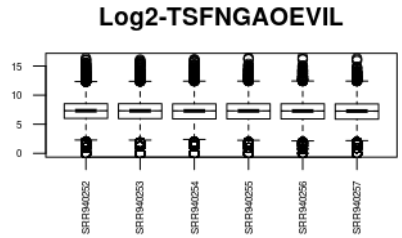


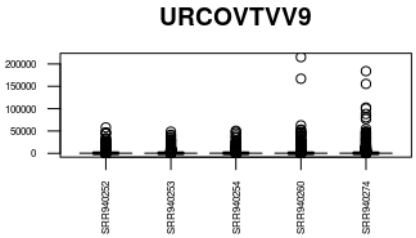


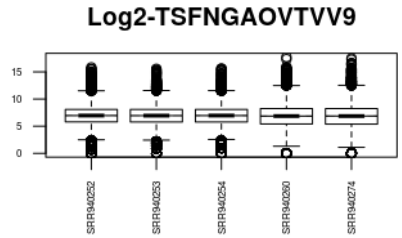




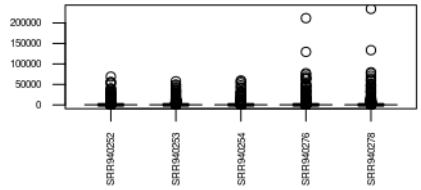




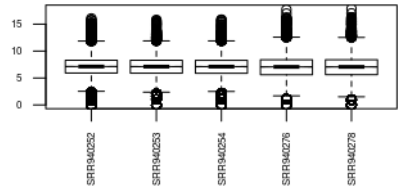




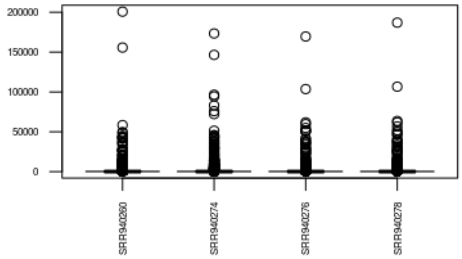
URCOR2VV9



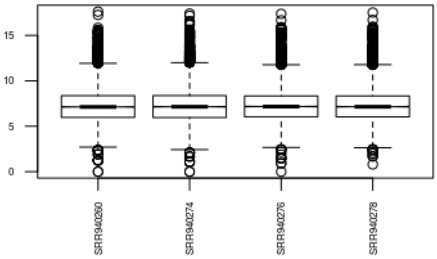
Log2-TSFNGAOR2VV9



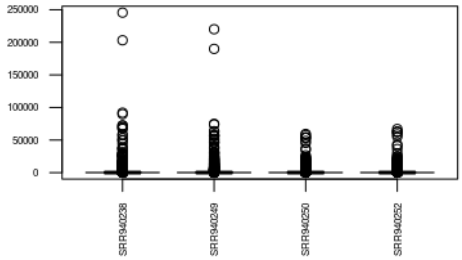
URCOVTVR2



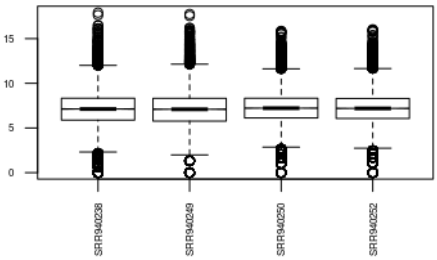
Log2-TSFNGAORVTVR2



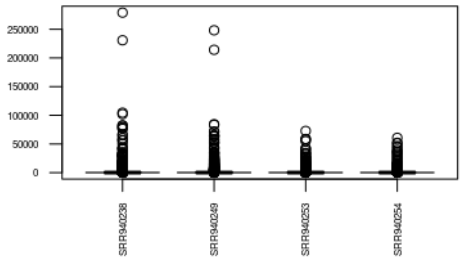
URCO8V11



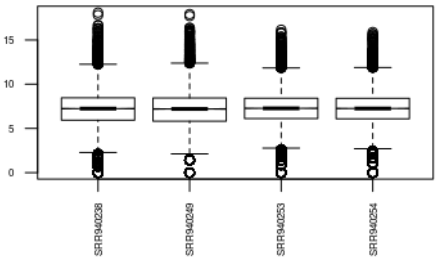
Log2-TSFNGAOR8V11



URCO8V13



Log2-TSFNGAOR8V13



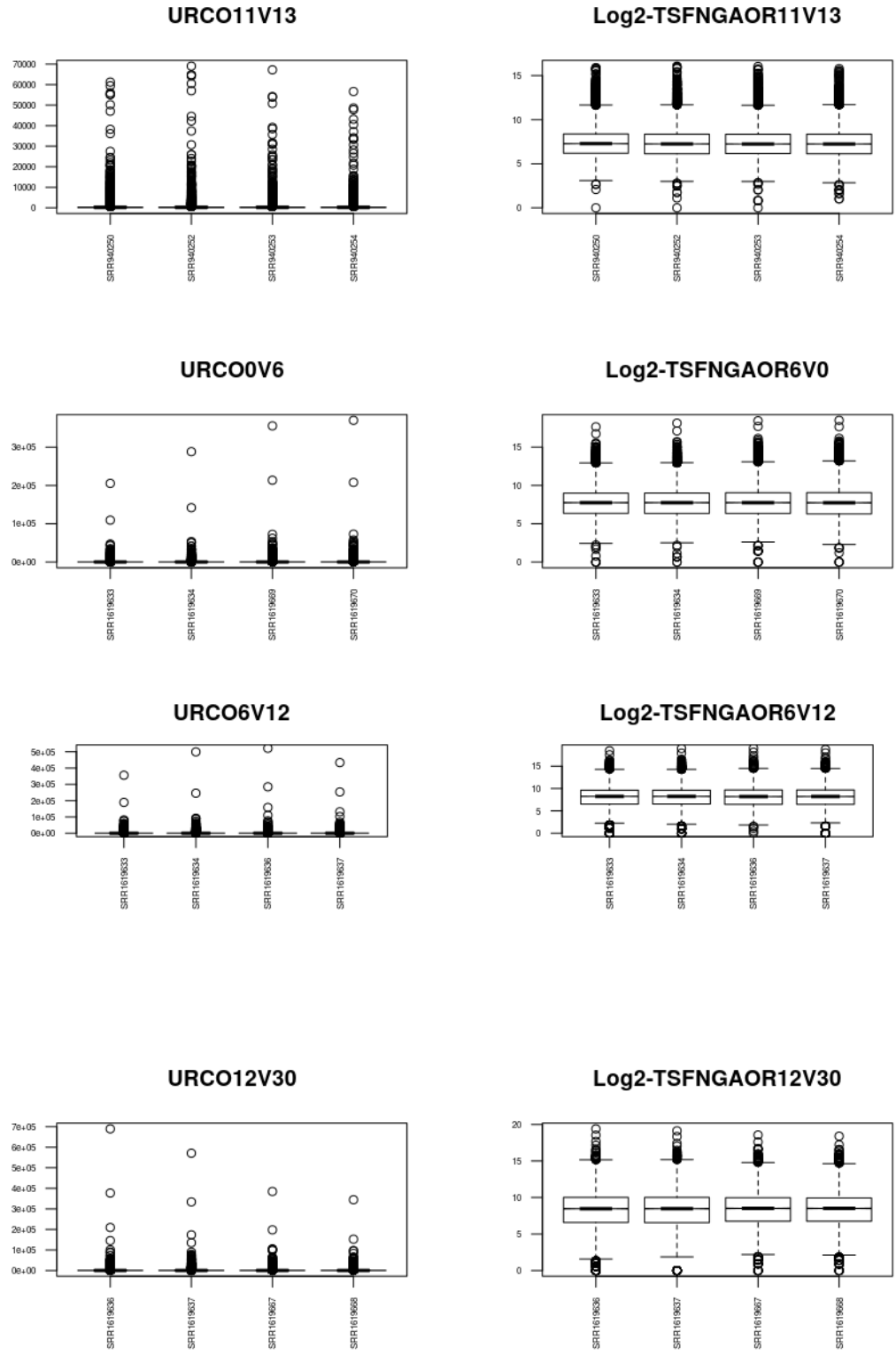
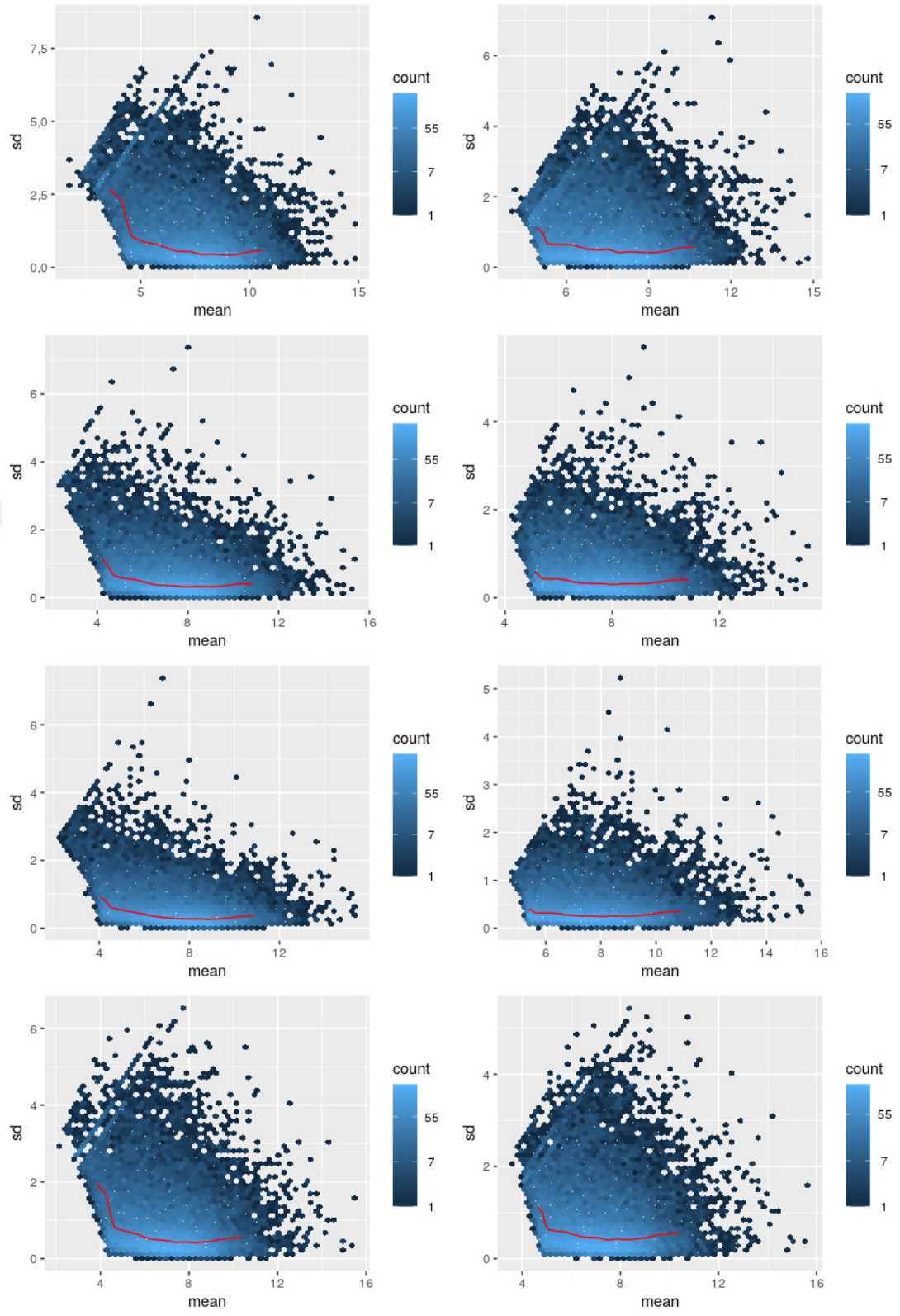
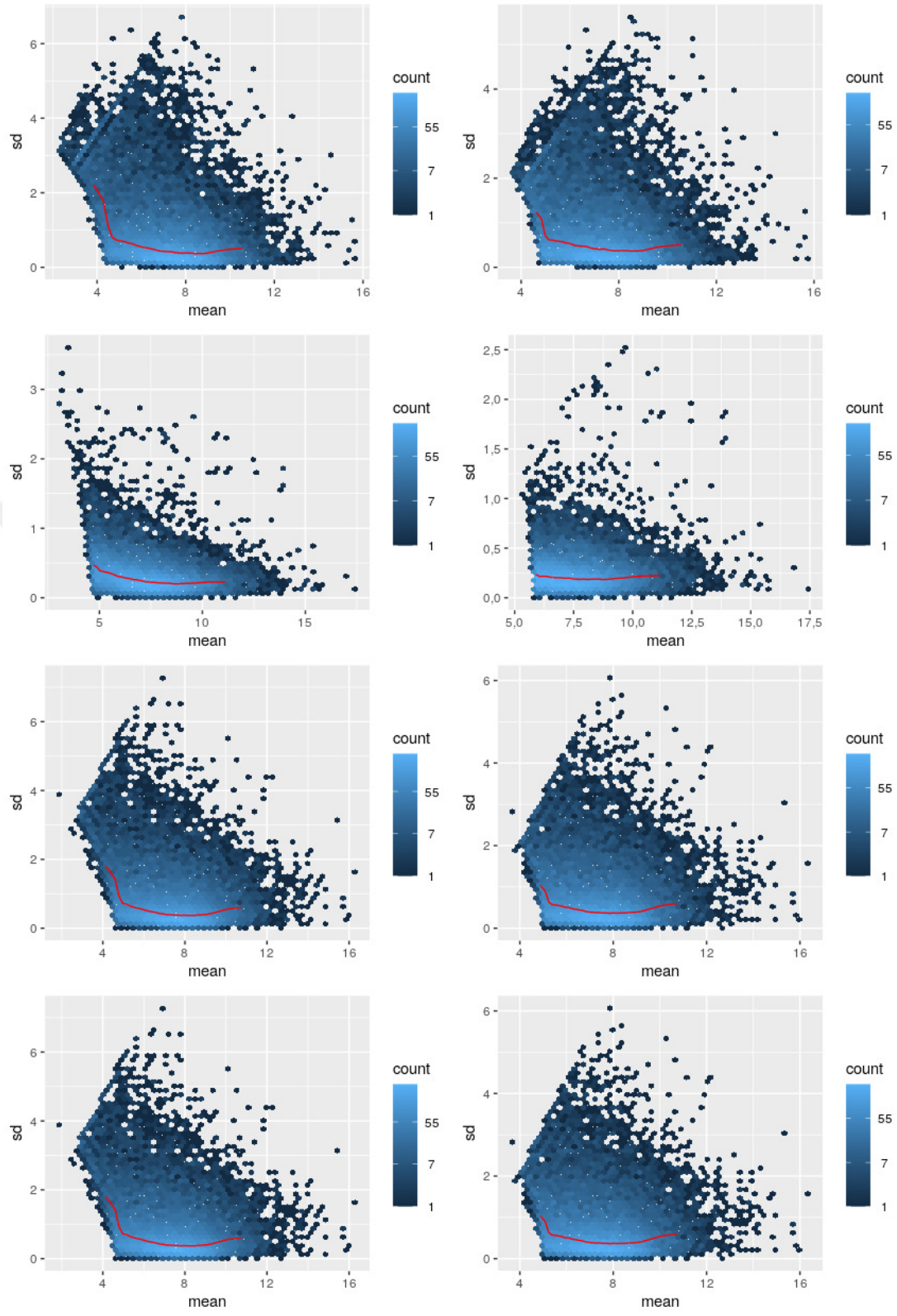
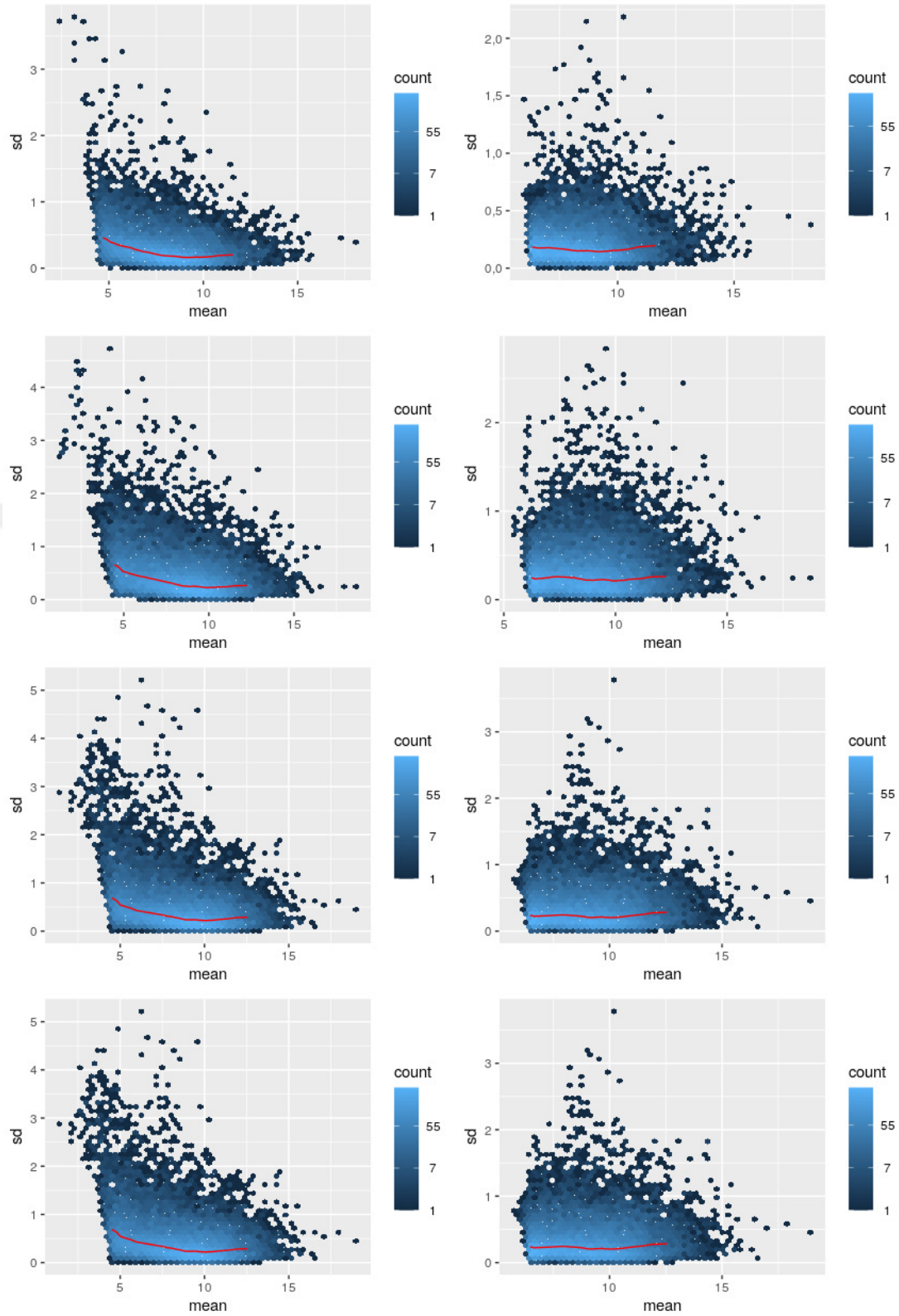


Figure 2.1. Normalization and transformation of each comparison.


sd
mean
count
55
7
1


sd
mean
count
55
7
1


sd
mean
count
55
7
1

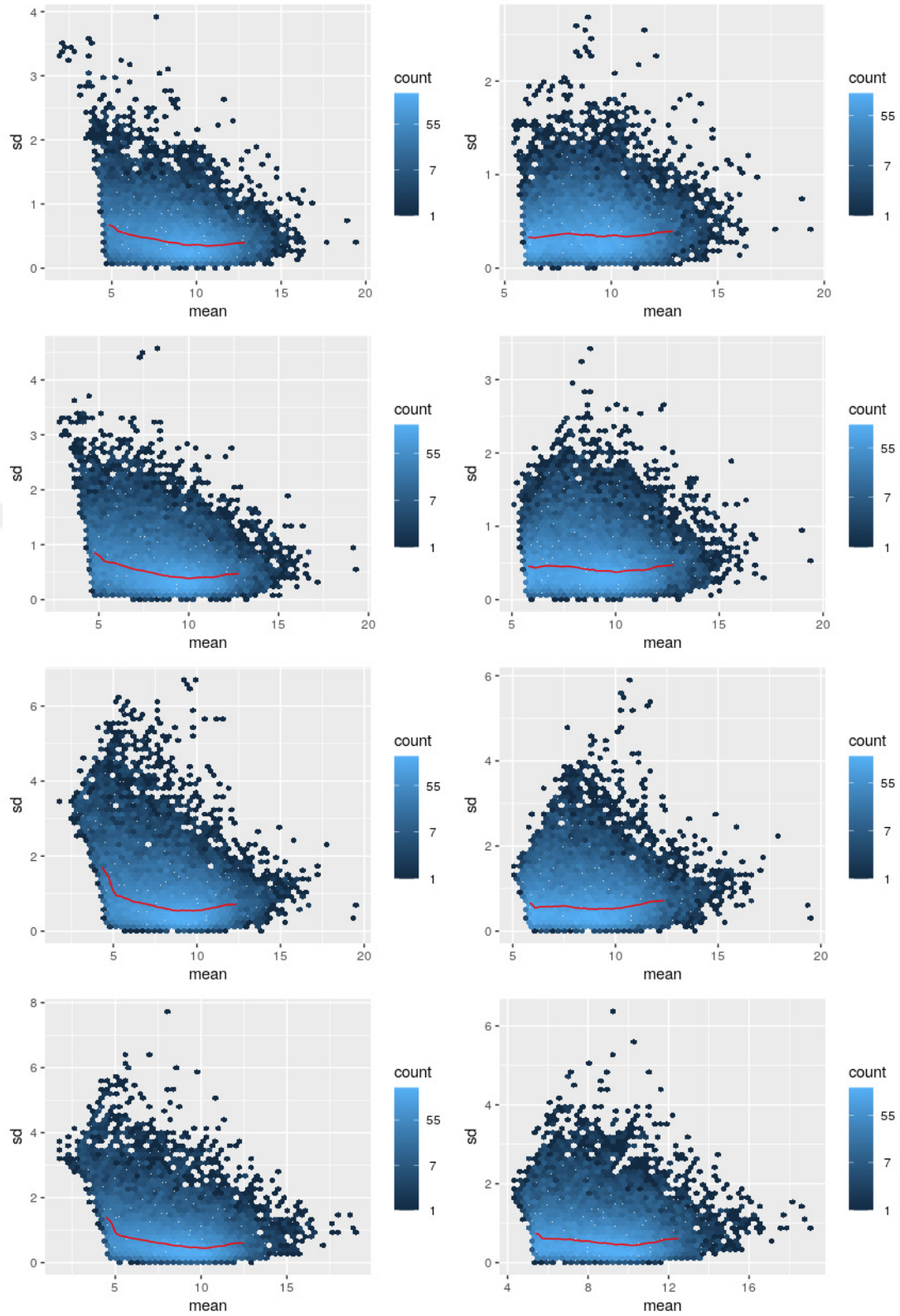


Figure 2.2. Variance stablized transformation of data sets.

Appendix 3. Hierarchical clustering and PCA of drought and control treatments

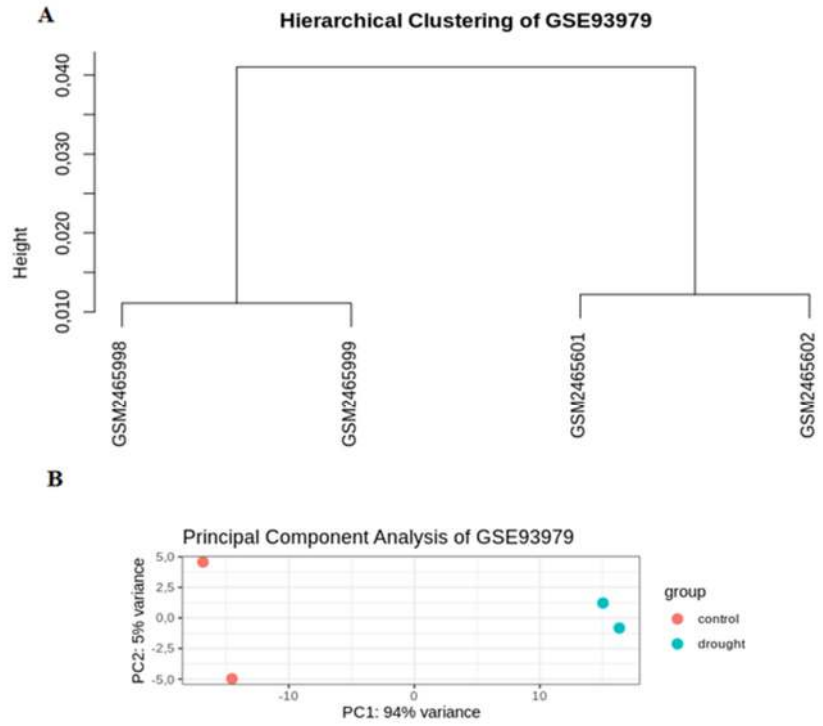


Figure 3.1. A) HC and B) PCA of GSE93979. GSM2465998 and GSM2465999 are control groups. GSM2466002 and GSM2466002 are drought-samples.

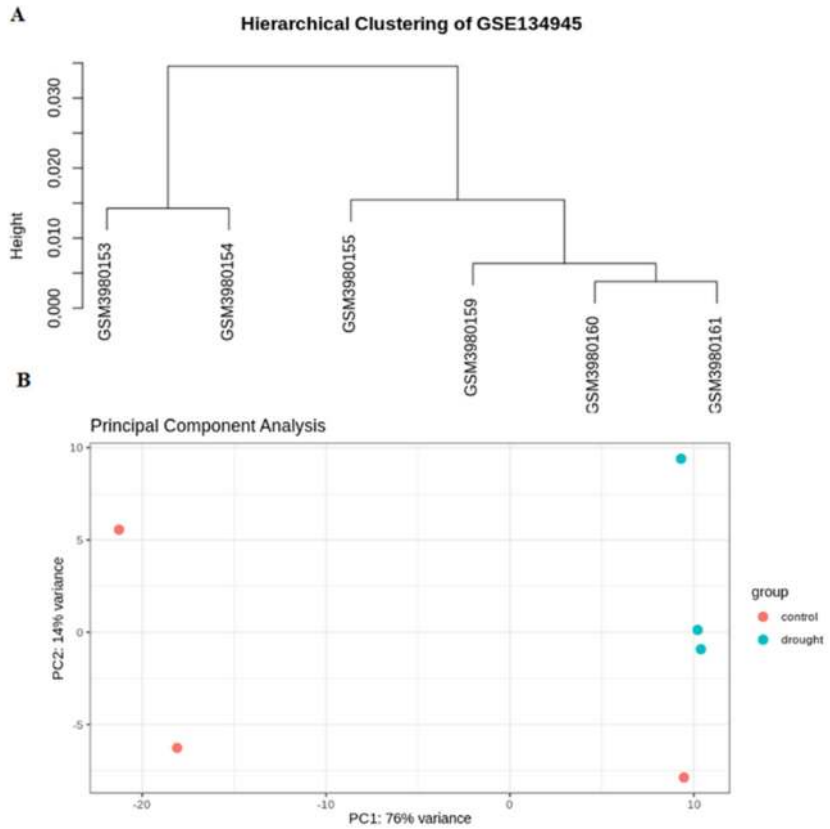


Figure 3.2. A) HC and B) PCA of GSE134945. GSM3980153, GSM3980154, GSM3980155 are control groups. GSM3980159, GSM3980160 and GSM3980161 are drought-samples.

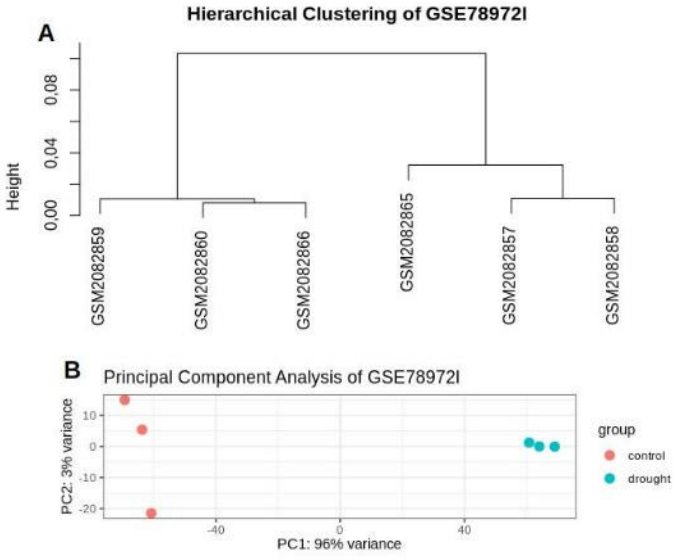


Figure 3.2. A) Hierarchical clustering analysis and B) PCA of long drought stress treatment samples of GSE78972 dataset. GSM2082857, GSM2082858 and GSM2082865 are control group. GSM2082859, GSM2082860 and GSM2082865 are drought-samples.

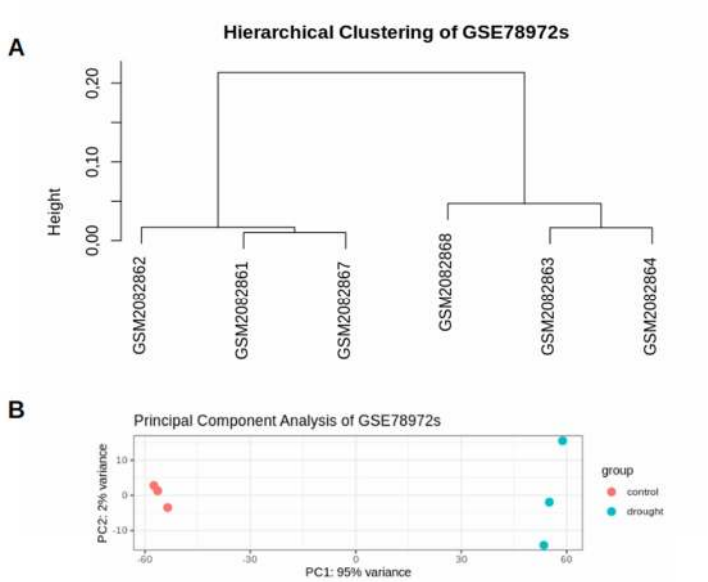


Figure 3.3. A) Hierarchical clustering analysis and B) PCA of short drought stress treatment samples of GSE78972 dataset. GSM2082861, GSM2082862 and GSM2082867 are control group. GSM2082863, GSM2082864 and GSM2082868 are drought-samples.

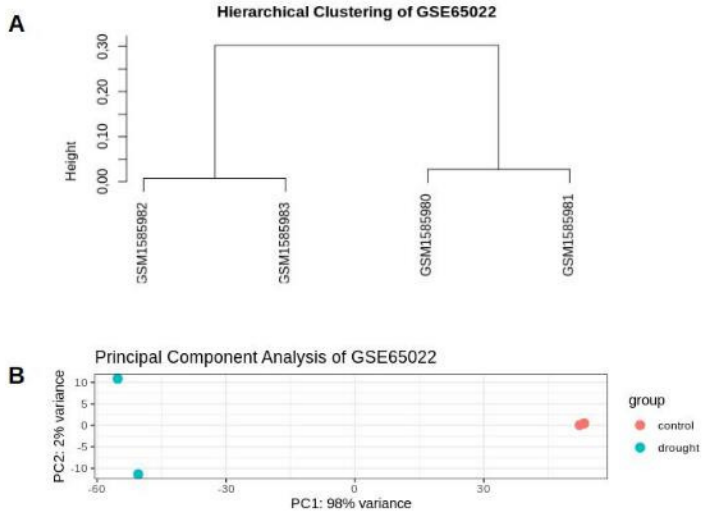


Figure 3.4. A) Hierarchical clustering analysis and B) PCA of drought stress treatment samples of GSE65022 dataset. GSM1585980 and GSM1585981 are control group. GSM1585982 and GSM1585983 are drought-samples.

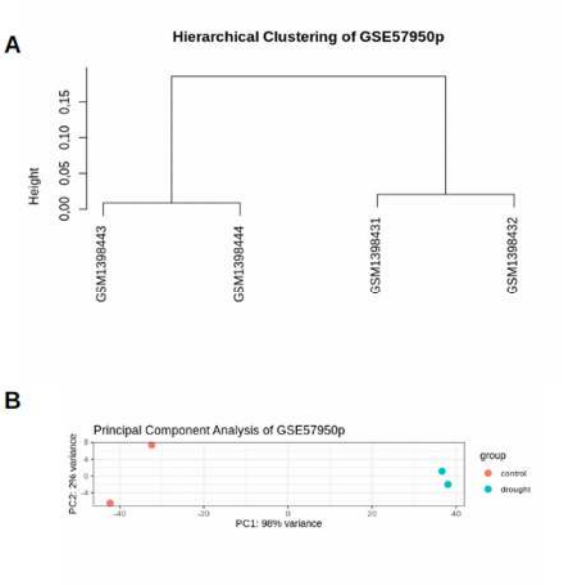


Figure 3.5. A) Hierarchical clustering analysis and B) PCA of drought stress treatment samples of P28 dataset. GSM1398431 and GSM1398432 are control group. GSM1398443 and GSM1398444 are drought-samples.

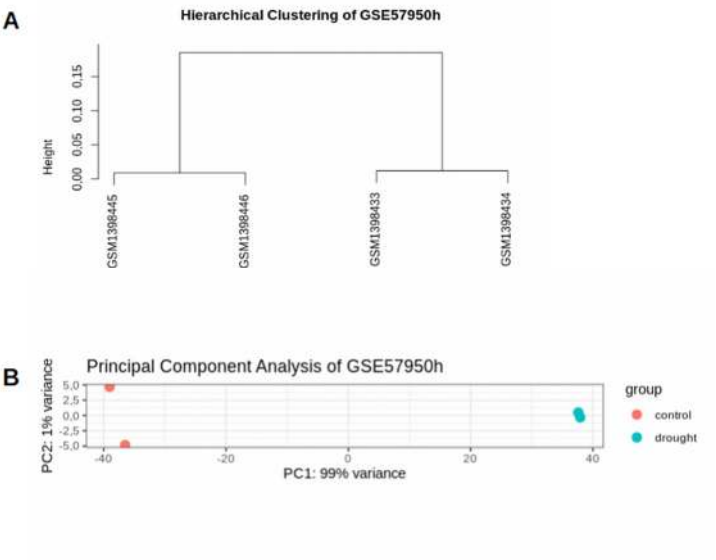


Figure 3.6. A) Hierarchical clustering analysis and B) PCA of drought stress treatment samples of H471 dataset. GSM1398433 and GSM1398434 are control group. GSM1398445 and GSM1398446 are drought-samples.

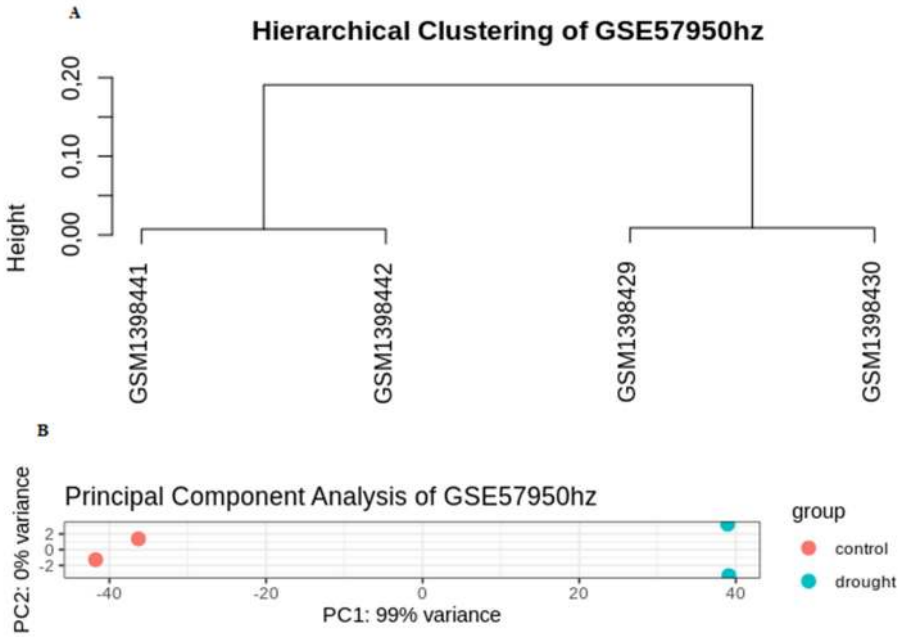


Figure 3.7. A) Hierarchical clustering analysis and B) PCA of drought stress treatment samples of HHZ dataset. GSM1398429 and GSM1398430 are control group. GSM1398441 and GSM1398442 are drought-samples.

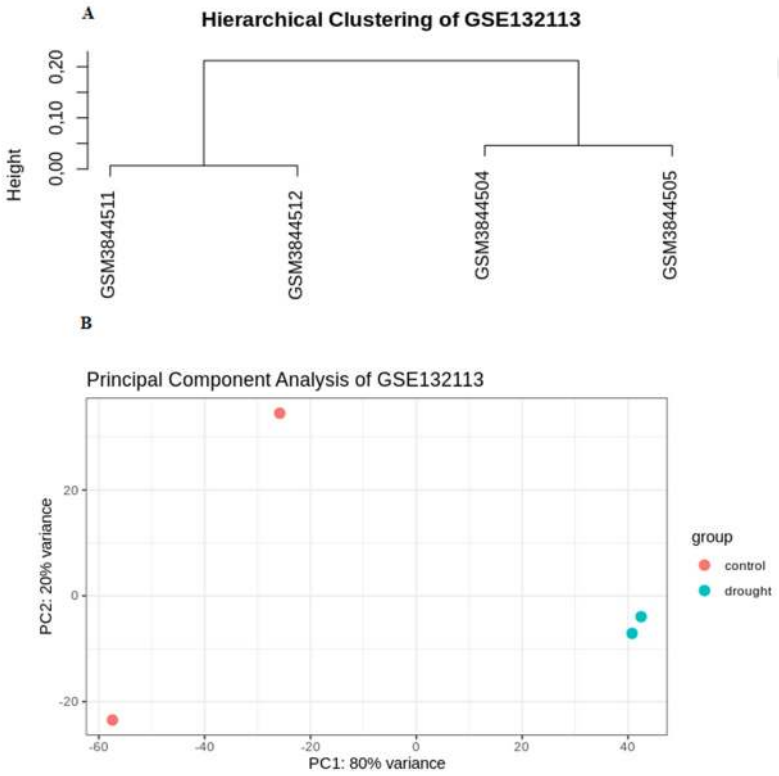


Figure 3.8. A) Hierarchical clustering analysis and B) PCA of long drought stress treatment samples of GSE78972 dataset. GSM3844504 and GSM3844505 are control group. GSM3844511 and GSM3844512 are drought-samples.

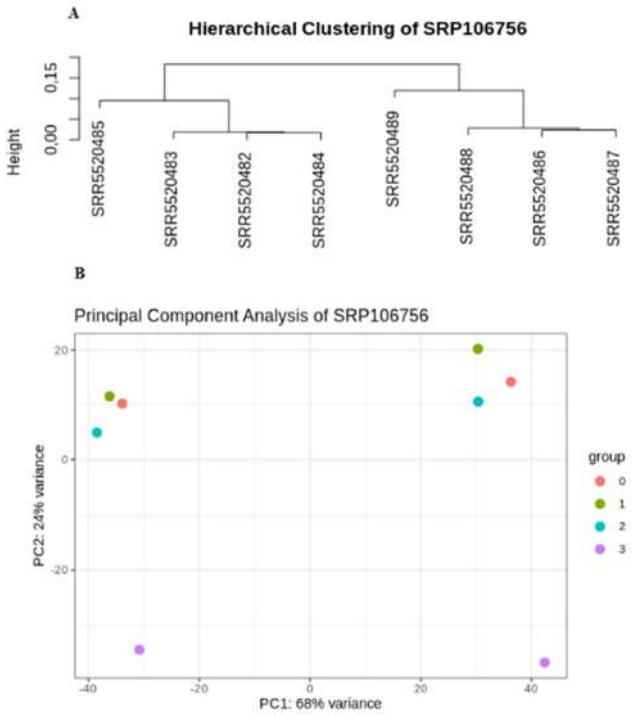


Figure 3.9. A) Hierarchical clustering analysis and B) PCA of drought stress treatment samples of SRP106756. SRR55204482, SRR55204483, SRR55204484, SRR55204485 belong to 754F inbreed line, whereas SRR55204486, SRR55204487, SRR55204488 and SRR55204489 are the samples of M45 inbreed line. 0; the sample that collected at 0 hour, 1; at 24 hours, 2; at 48 hours, 3; at 72 hours.

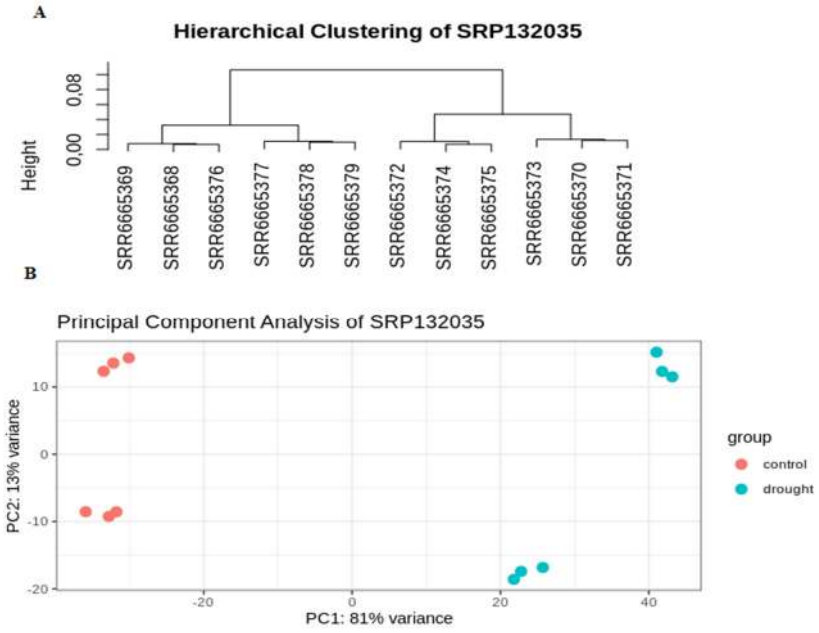


Figure 3.10. A) Hierarchical clustering analysis and B) PCA of drought stress treatment samples of SRP132035. SRR6665368, SRR6665369, SRR6665376, SRR6665377, SRR6665378 and SRR6665379 are control groups. SRR6665370, SRR6665371, SRR6665372, SRR6665373 and SRR6665374, SRR6665375 are drought-samples.

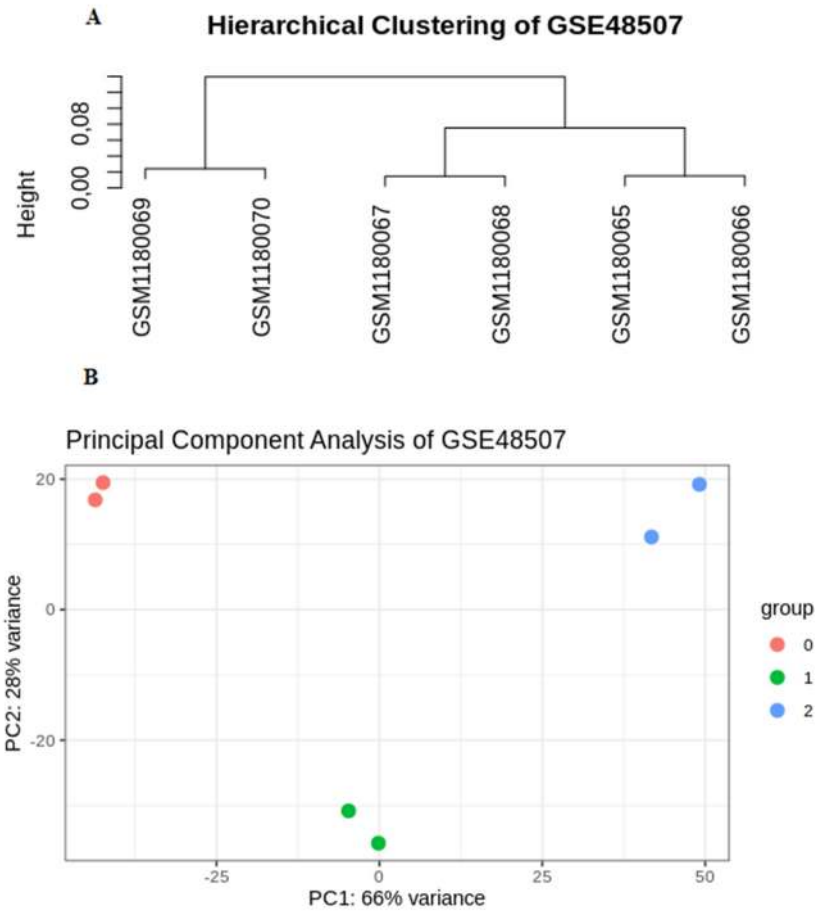


Figure 3.11. A) Hierarchical clustering analysis and B) PCA of drought stress treatment samples of GSE48507. GSM1180065 and GSM1180066 are control groups. GSM1180067, GSM1180068, GSM1180069 and GSM1180070 are drought-samples in two different time-points.

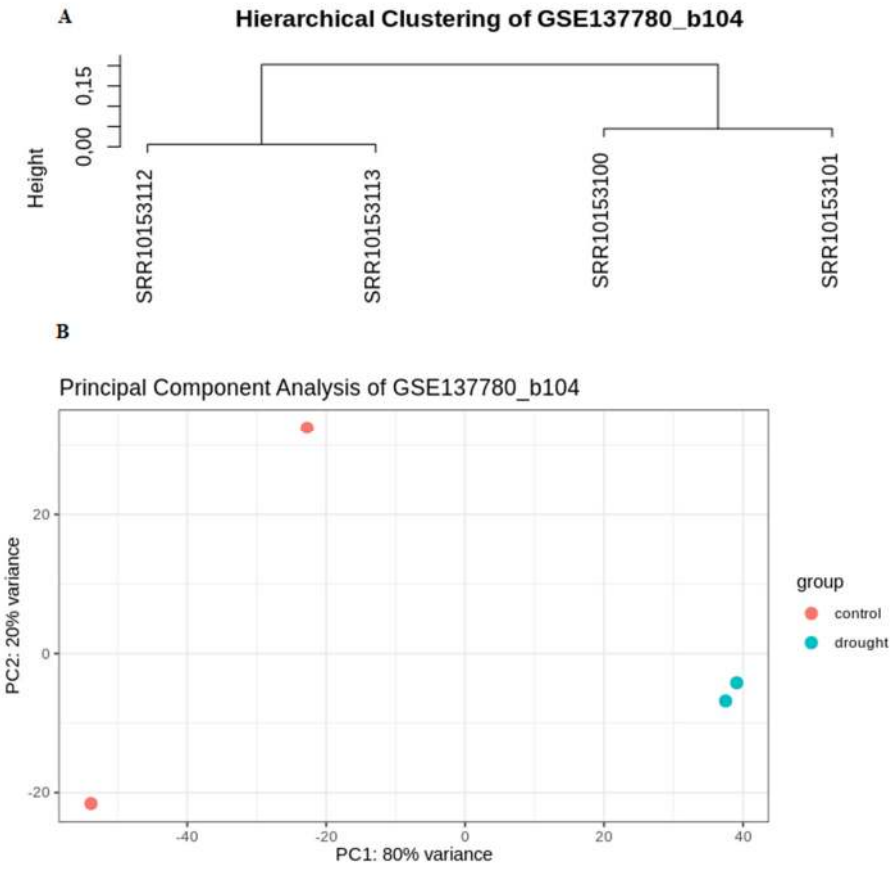


Figure 3.12. A) Hierarchical clustering analysis and B) PCA of drought stress treatment samples of GSE137780_b104. SRR10153100 and SRR10153101 are control groups. SRR10153112 and SRR10153113 are drought-samples.

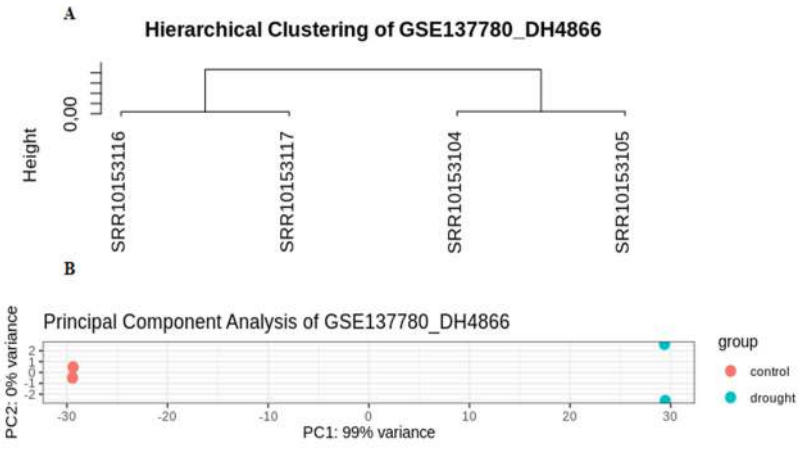


Figure 3.13. A) Hierarchical clustering analysis and B) PCA of drought stress treatment samples of GSE137780_DH4866. SRR10153104 and SRR10153105 are control groups. SRR10153116 and SRR10153117 are drought-samples.

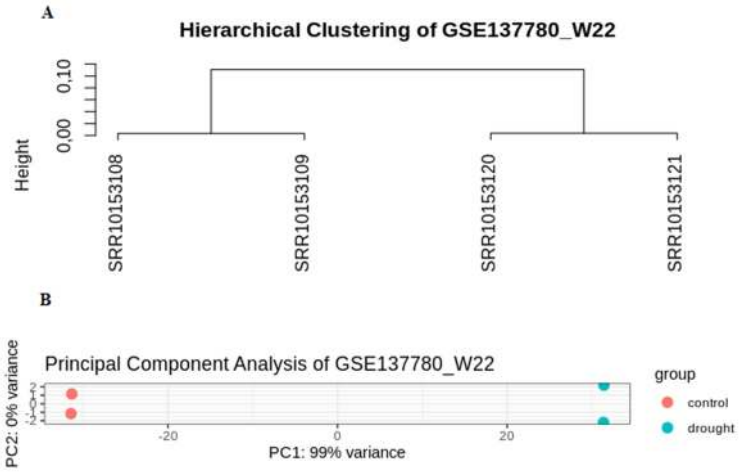


Figure 3.14. A) Hierarchical clustering analysis and B) PCA of drought stress treatment samples of GSE137780_W22. SRR10153108 and SRR10153109 are control groups. SRR10153120 and SRR10153121 are drought-samples.

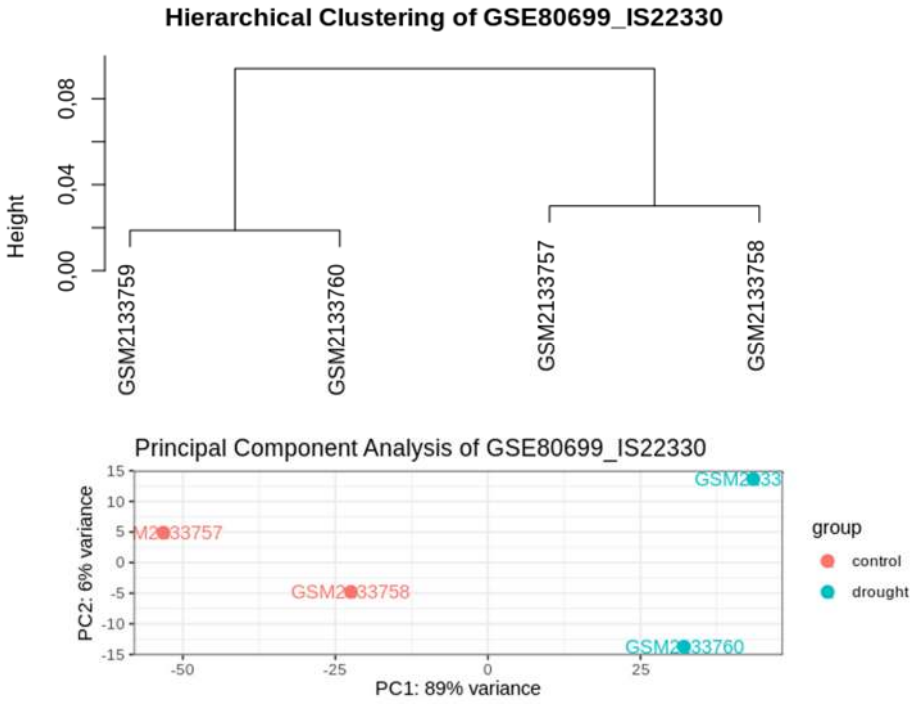


Figure 3.15. A) Hierarchical clustering analysis and B) PCA of drought stress treatment samples of GSE80699_IS2230. GSM2133759 and GSM2133760 are control groups. GSM2133757 and GSM2133758 are drought-samples.

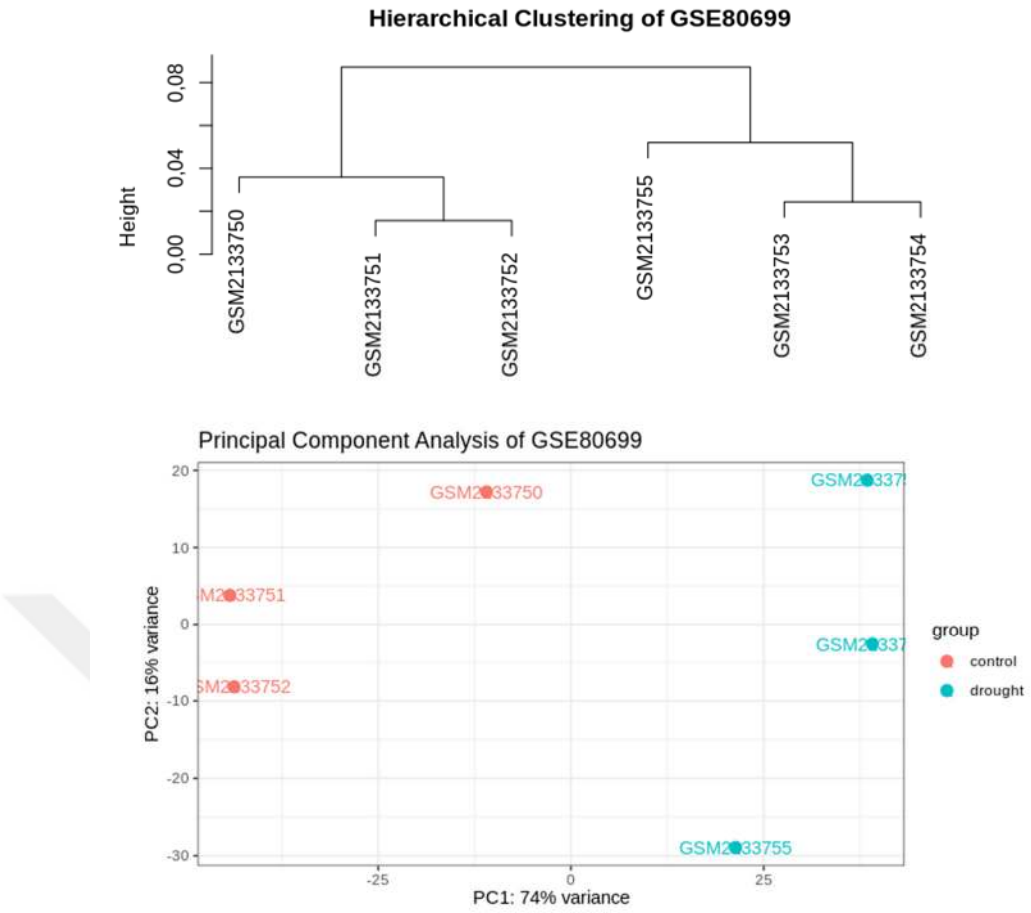


Figure 3.16. A) Hierarchical clustering analysis and B) PCA of drought stress treatment samples of GSE80699_ IS20351. GSM2133750, GSM2133751 and GSM2133752 are control groups. GSM2133753, GSM2133754 and GSM2133755 are drought-samples.

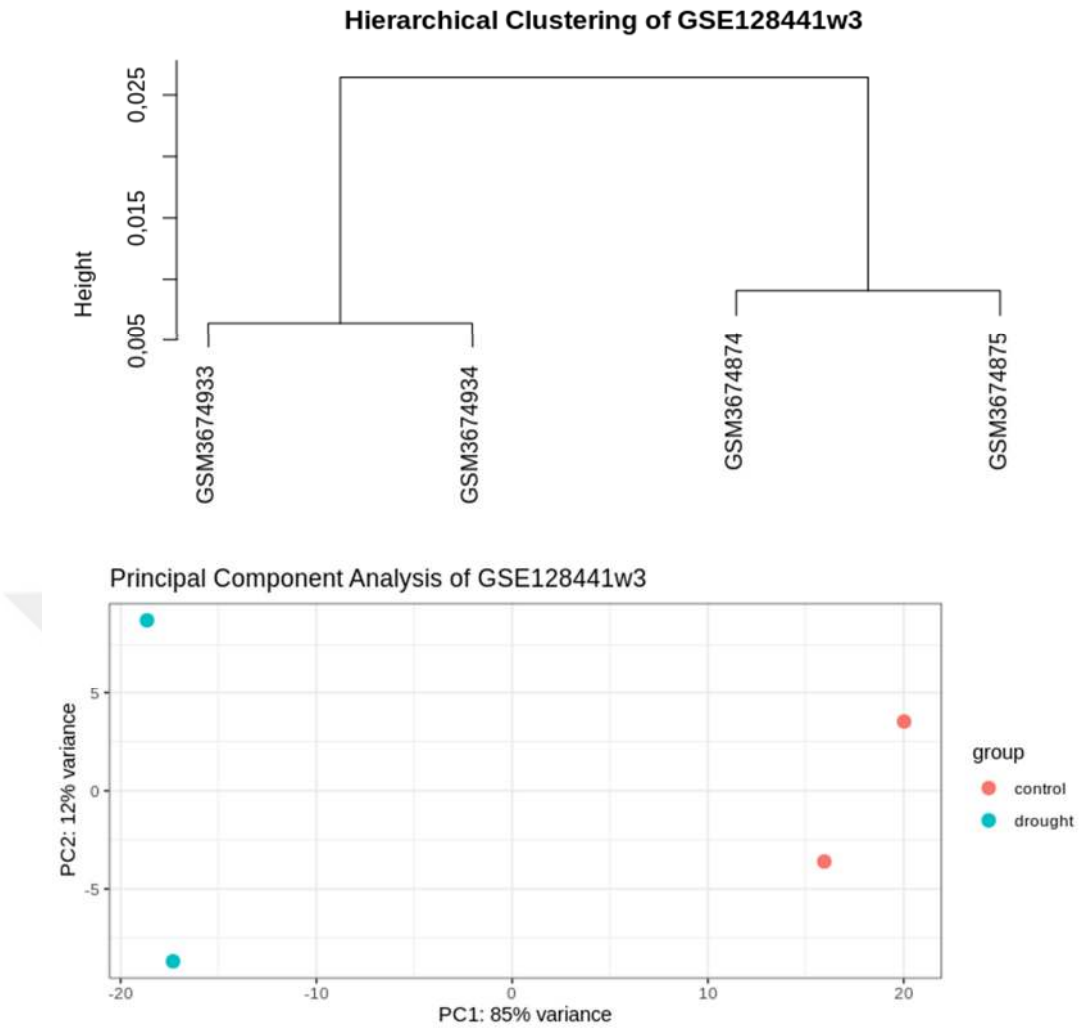


Figure 3.17. A) Hierarchical clustering analysis and B) PCA of drought stress treatment samples of GSE80699_ week 3. GSM3674933 and GSM3674934 are control groups. GSM3674974 and GSM3674975 are drought-samples.

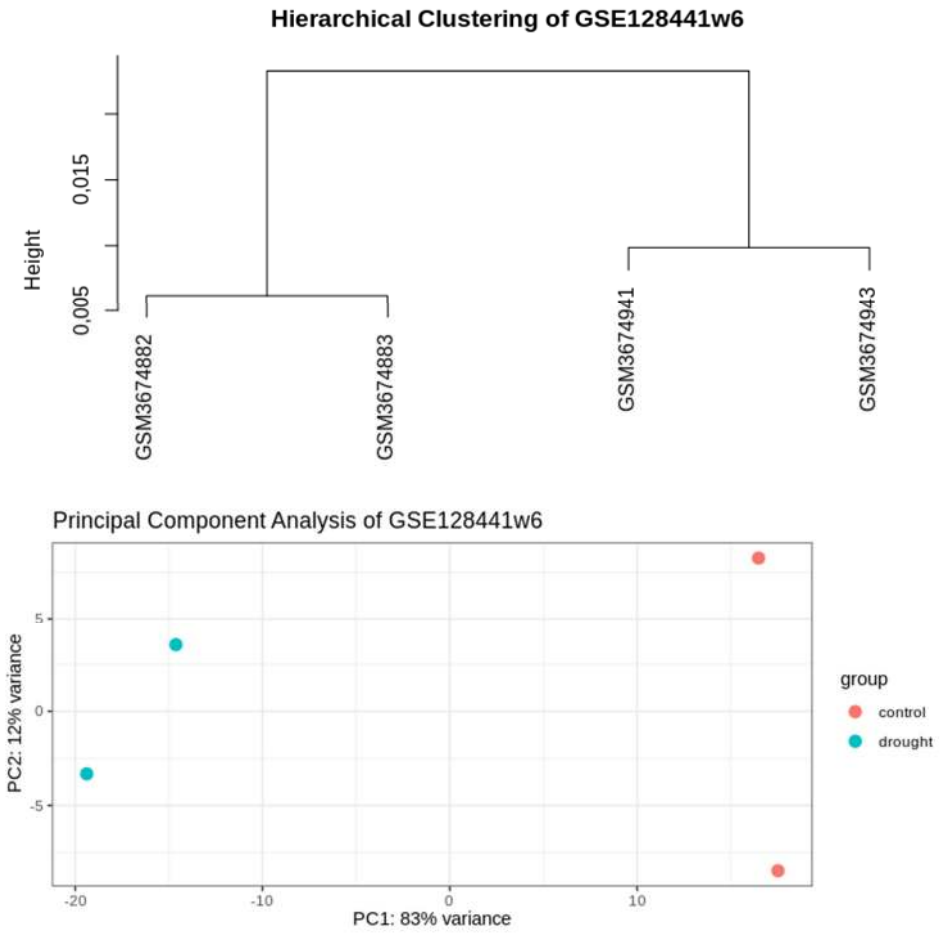


Figure 3.18. A) Hierarchical clustering analysis and B) PCA of drought stress treatment samples of GSE80699_ week 6. GSM3674941 and GSM3674943 are control groups. GSM3674982 and GSM3674983 are drought-samples.

Appendix 4. Expression patterns of peroxisomal genes under drought stress

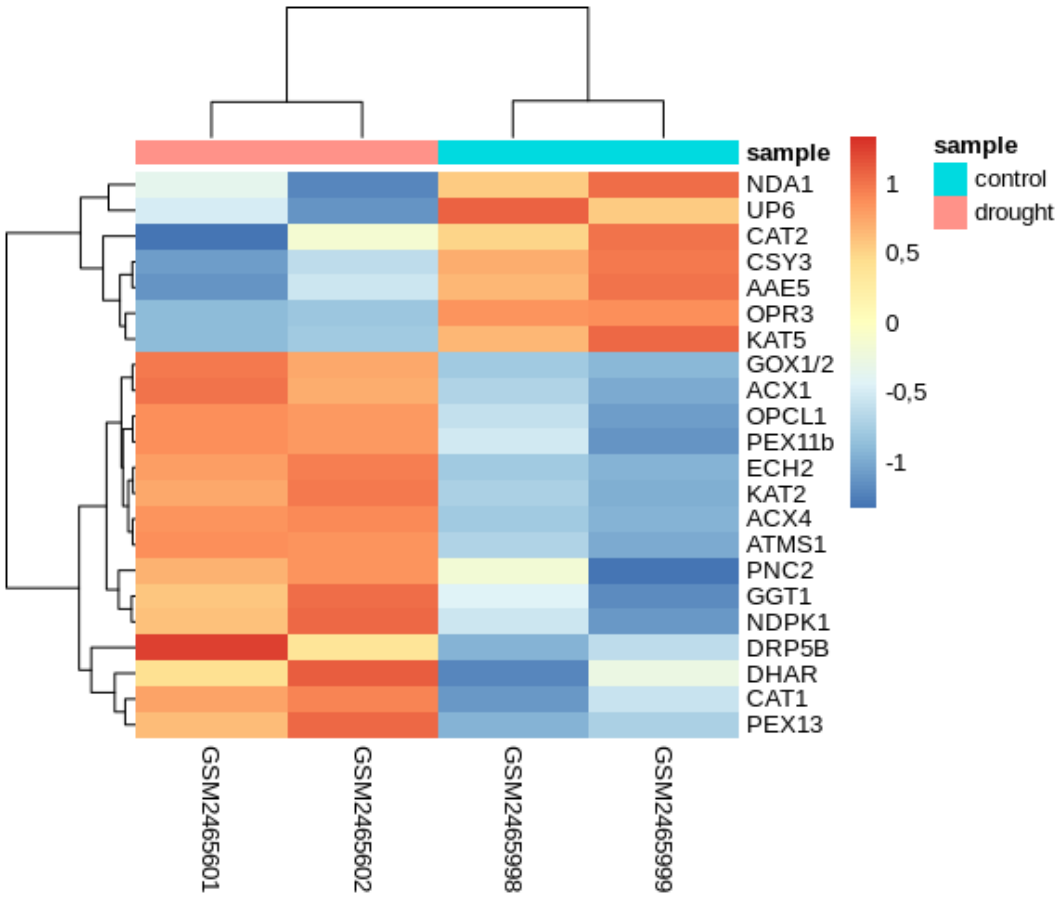


Figure 4.1. Expression patterns of twenty two peroxisome related DEGs of GSE93979. Heatmap was generated based on the z-score transformed vst value.

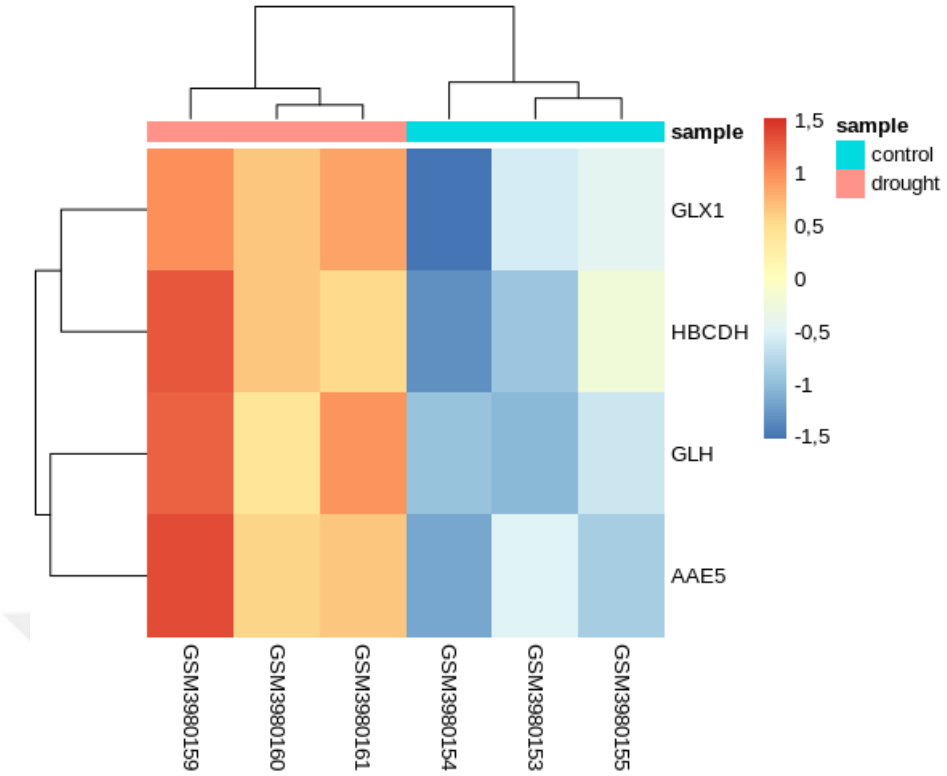


Figure 4.2. Expression patterns of four peroxisome-related DEGs of GSE134945. Heatmap was generated based on the z-score transformed vst value.

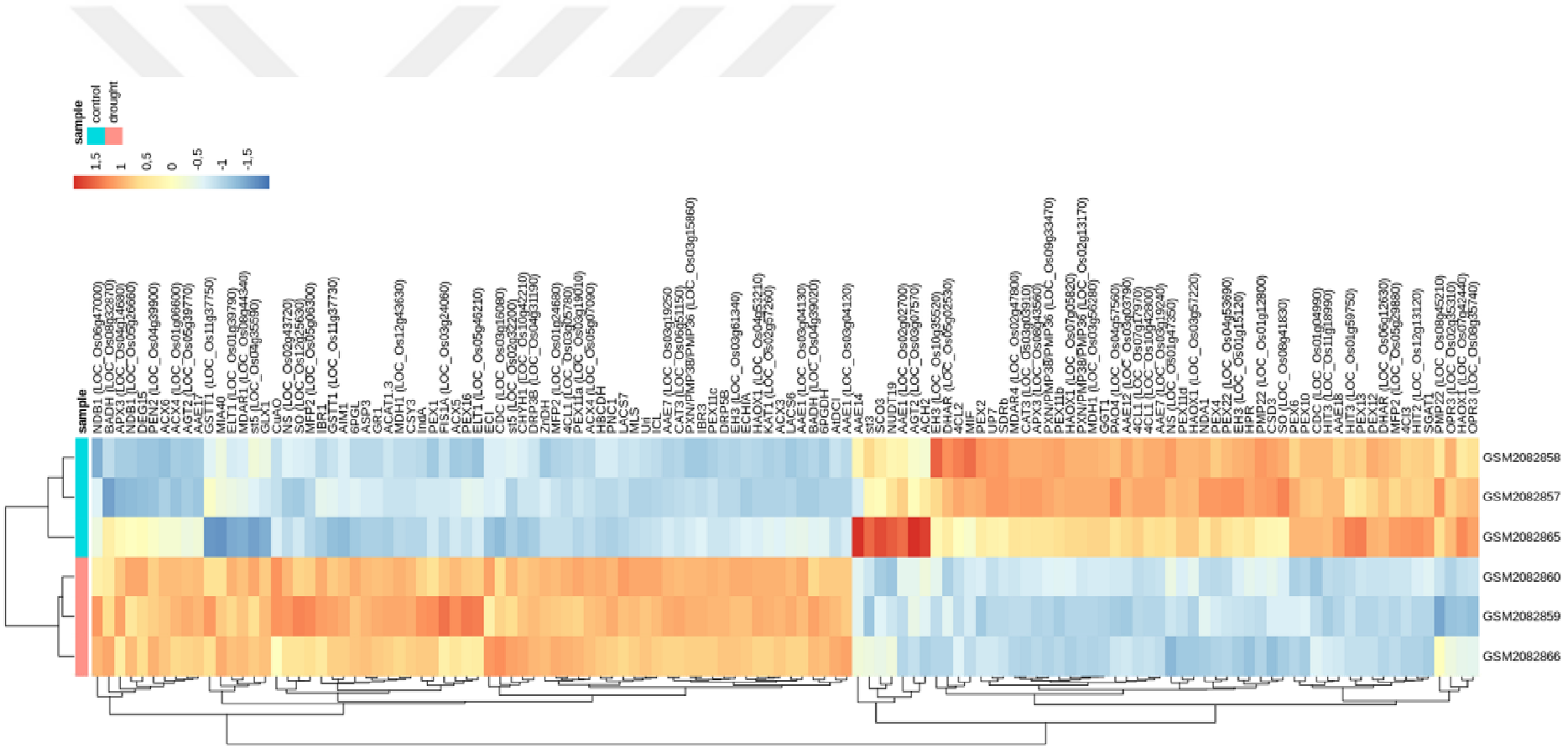


Figure 4.3. Expression patterns of 124 peroxisome related DEGs of GSE78972_long. Heatmap was generated based on the z-score transformed vst value.

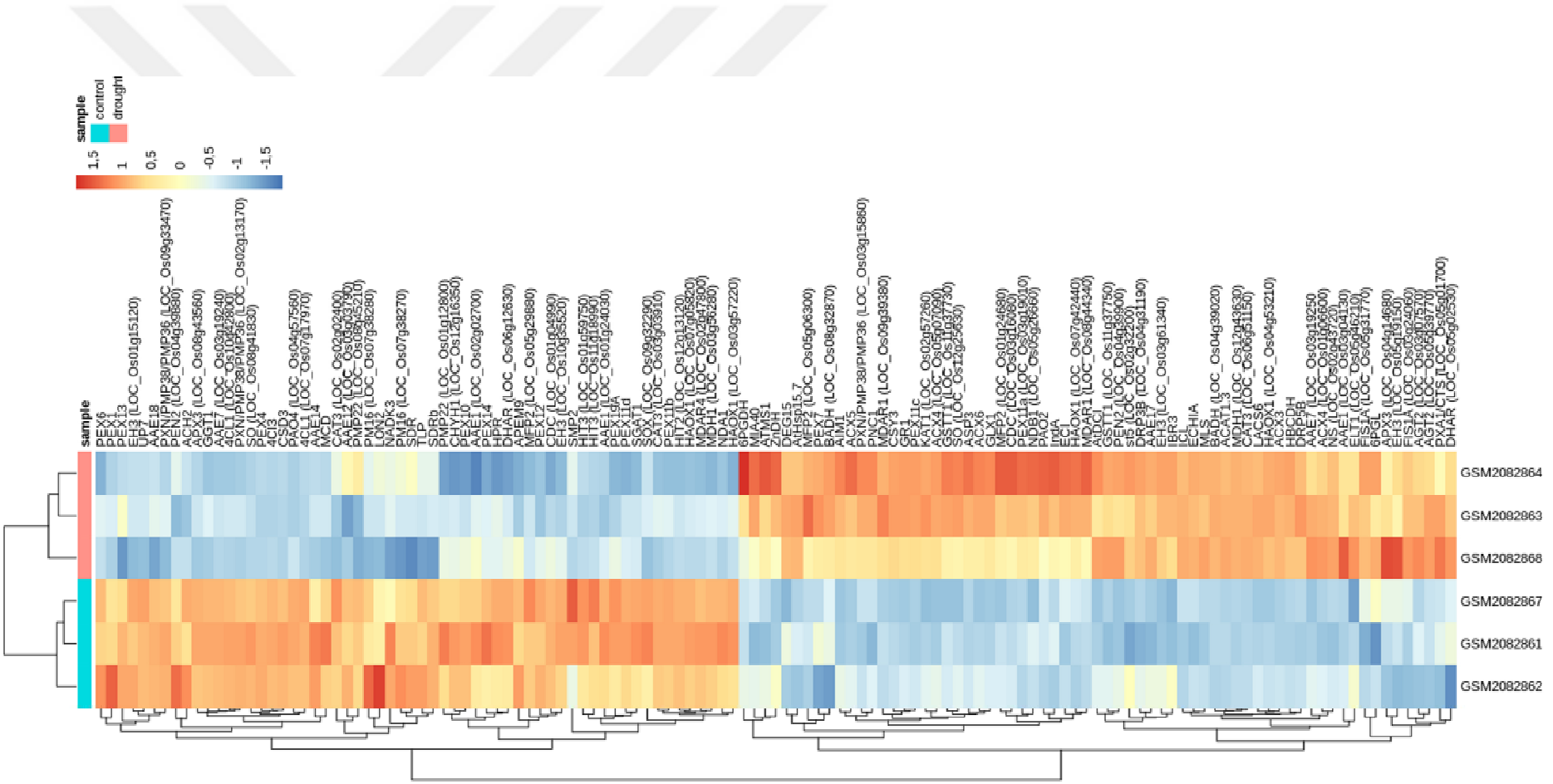


Figure 4.4. Expression patterns of 127 peroxisome related DEGs of GSE78972_short. Heatmap was generated based on the z-score transformed vst value.

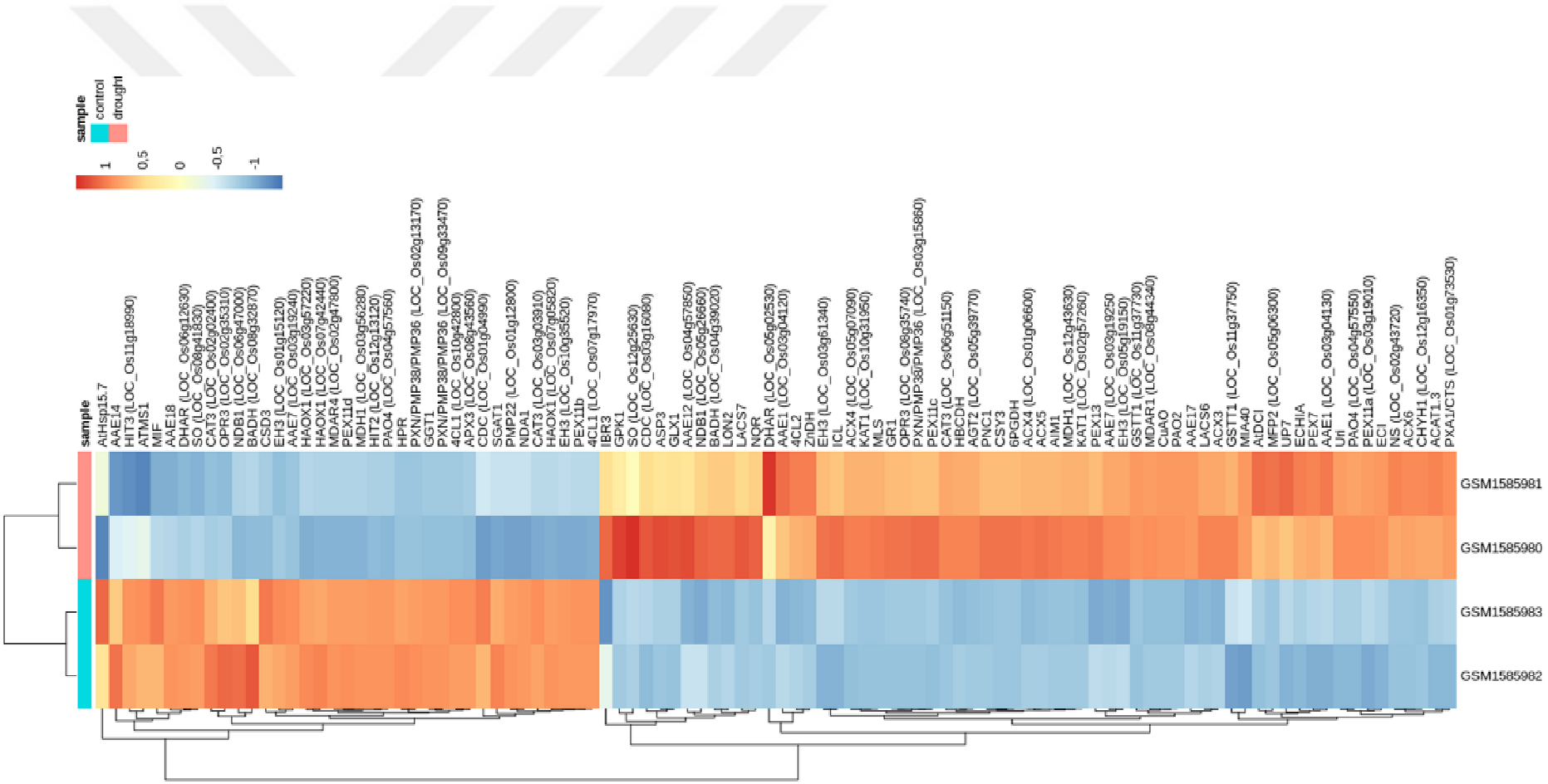


Figure 4.5. Expression patterns of 100 peroxisome related DEGs of GSE65022. Heatmap was generated based on the z-score transformed vst

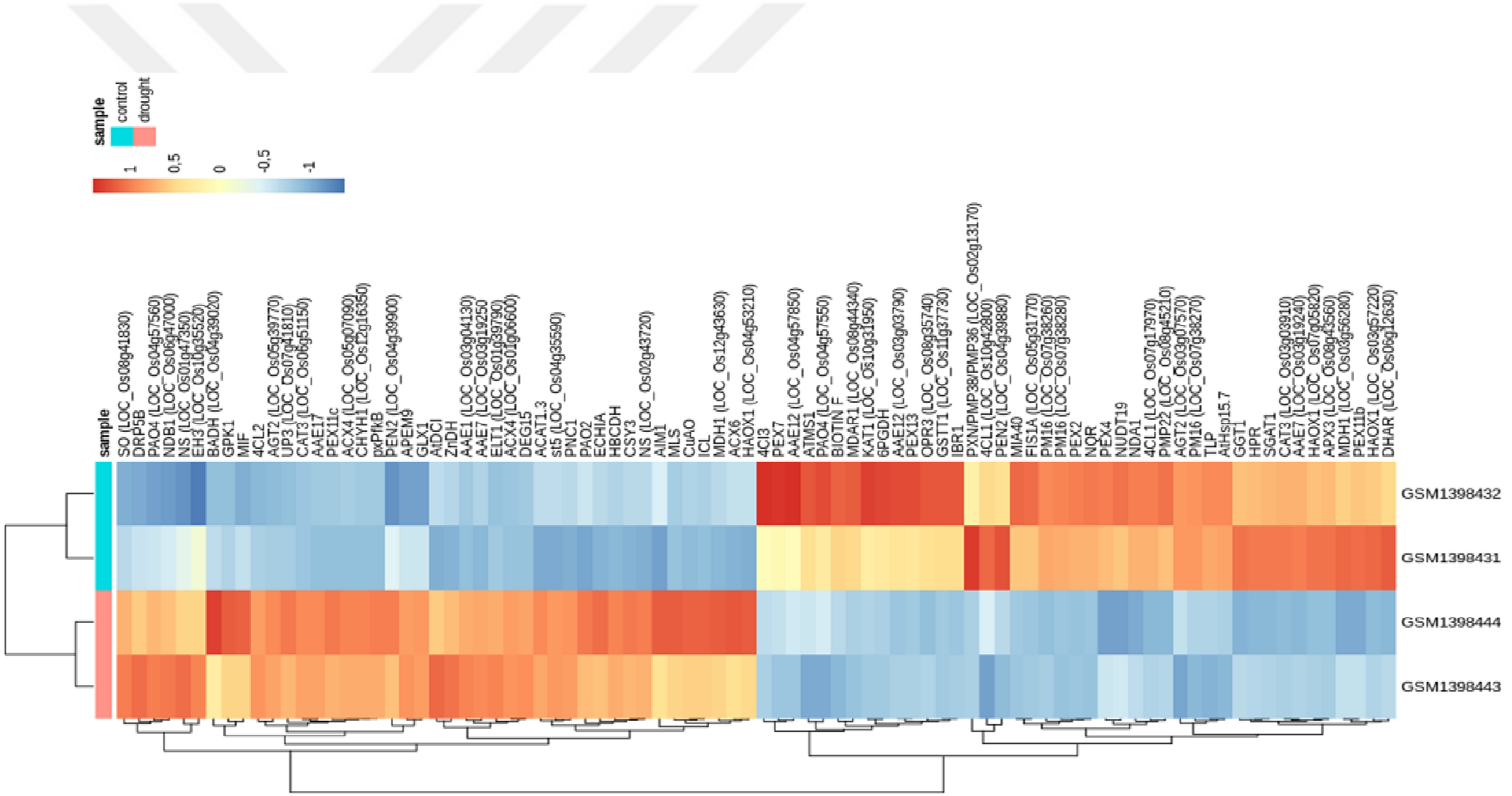


Figure 4.6. Expression patterns of 86 peroxisome related DEGs of GSE57950_p28. Heatmap was generated based on the z-score transformed vst value.

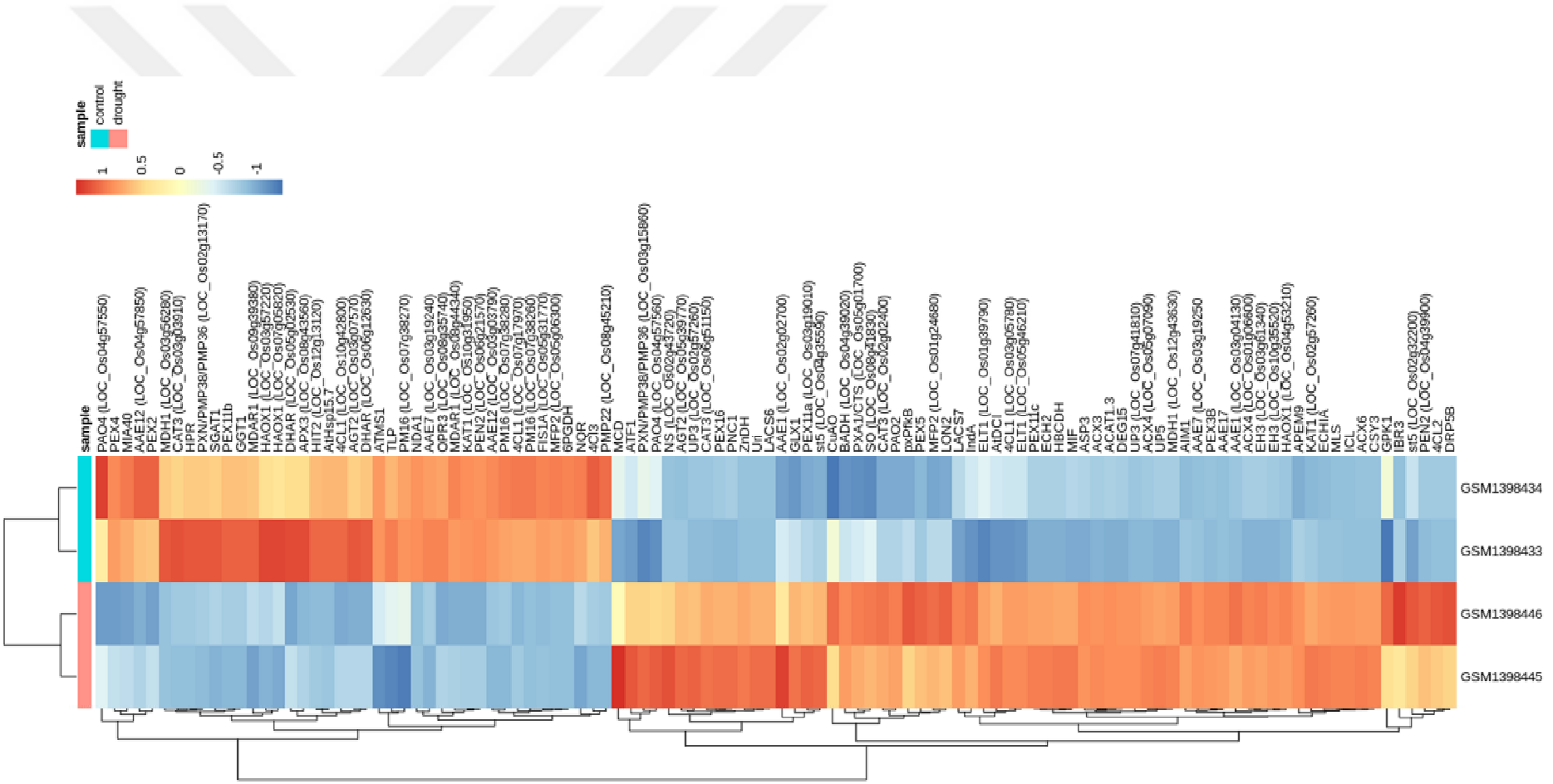


Figure 4.7. Expression patterns of 108 peroxisome related DEGs of GSE57950_h. Heatmap was generated based on the z-score transformed vst

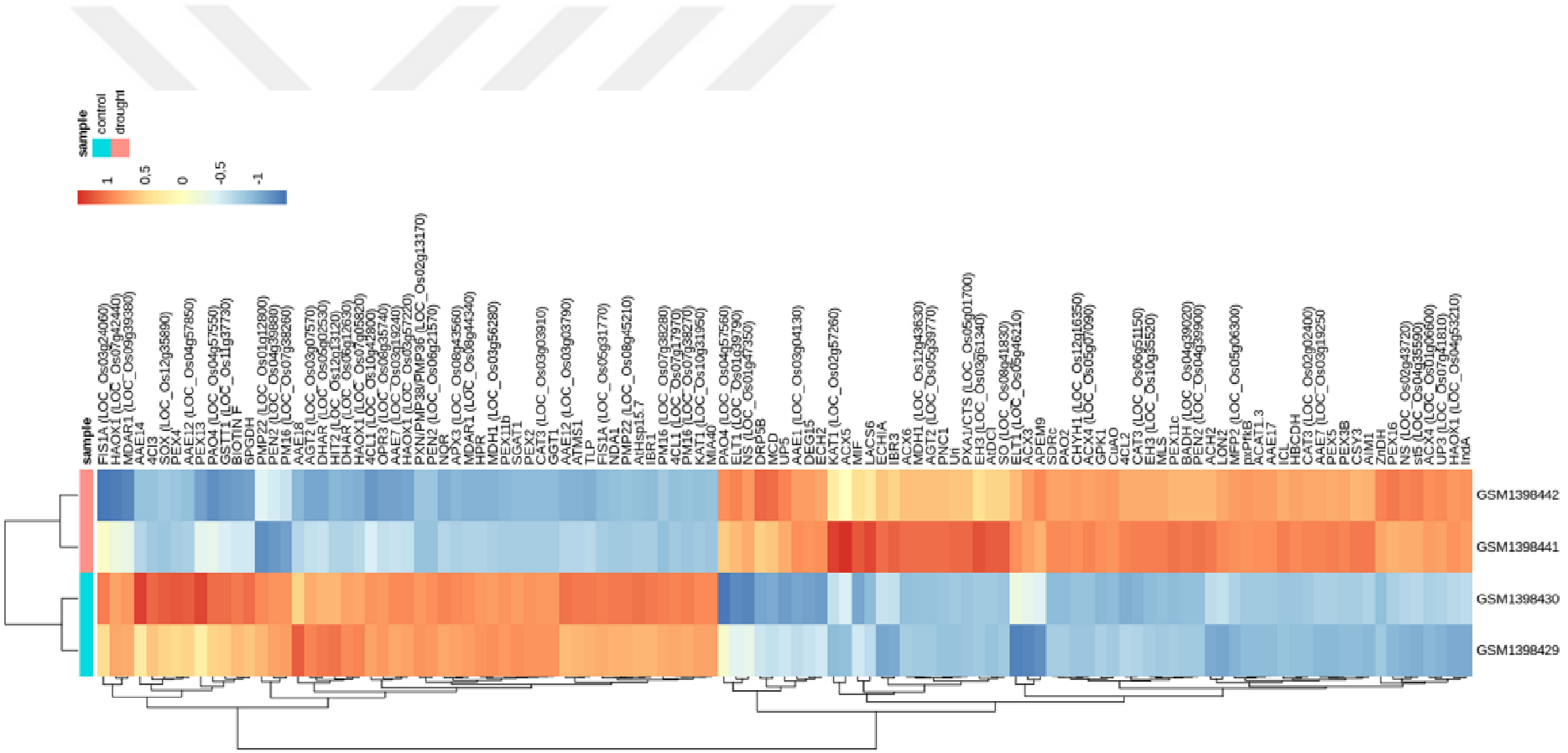


Figure 4.8. Expression patterns of 113 peroxisome related DEGs of GSE57950_hhz. Heatmap was generated based on the z-score transformed vst

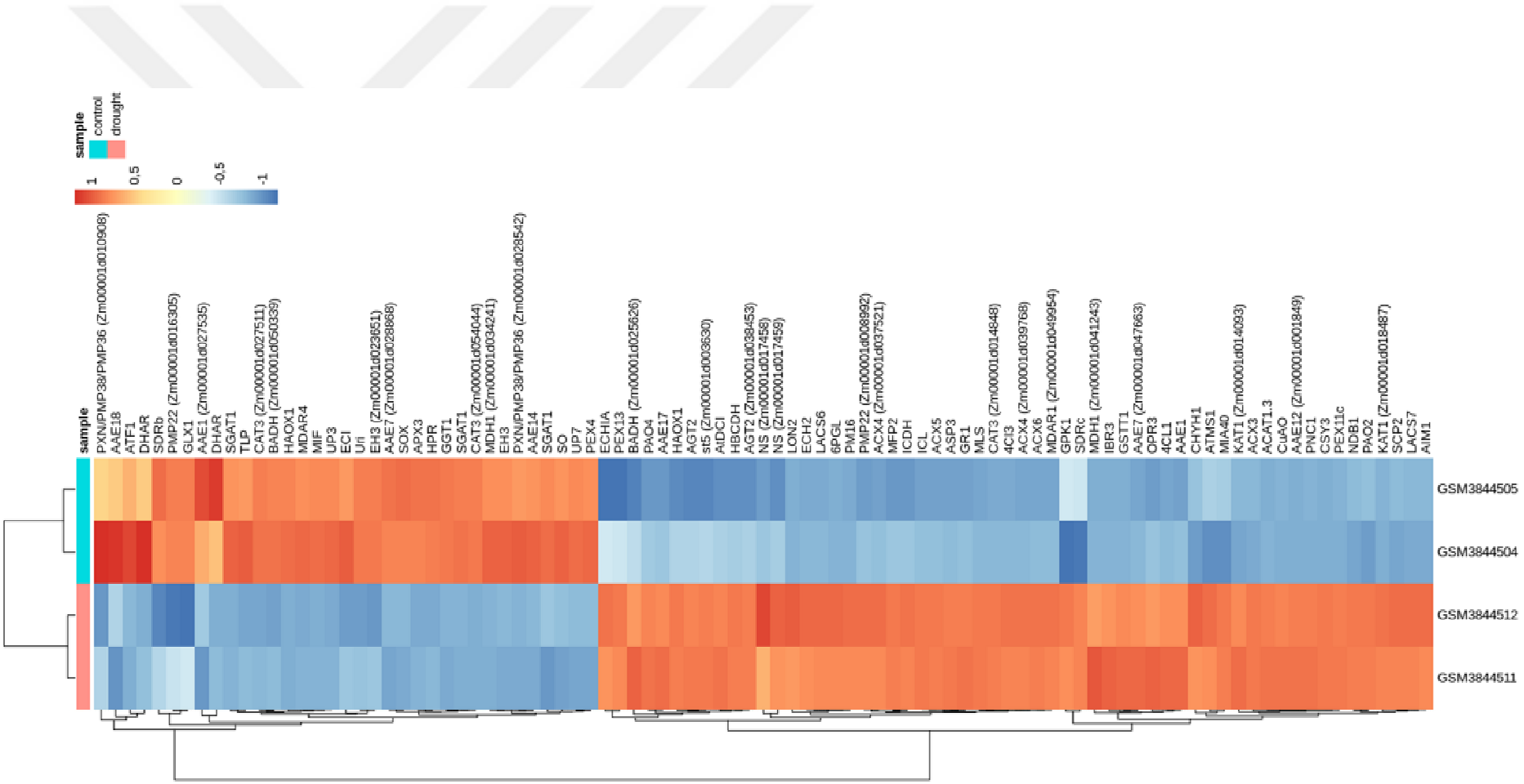


Figure 4.9. Expression patterns of 93 peroxisome related DEGs of GSE132113. Heatmap was generated based on the z-score transformed vst

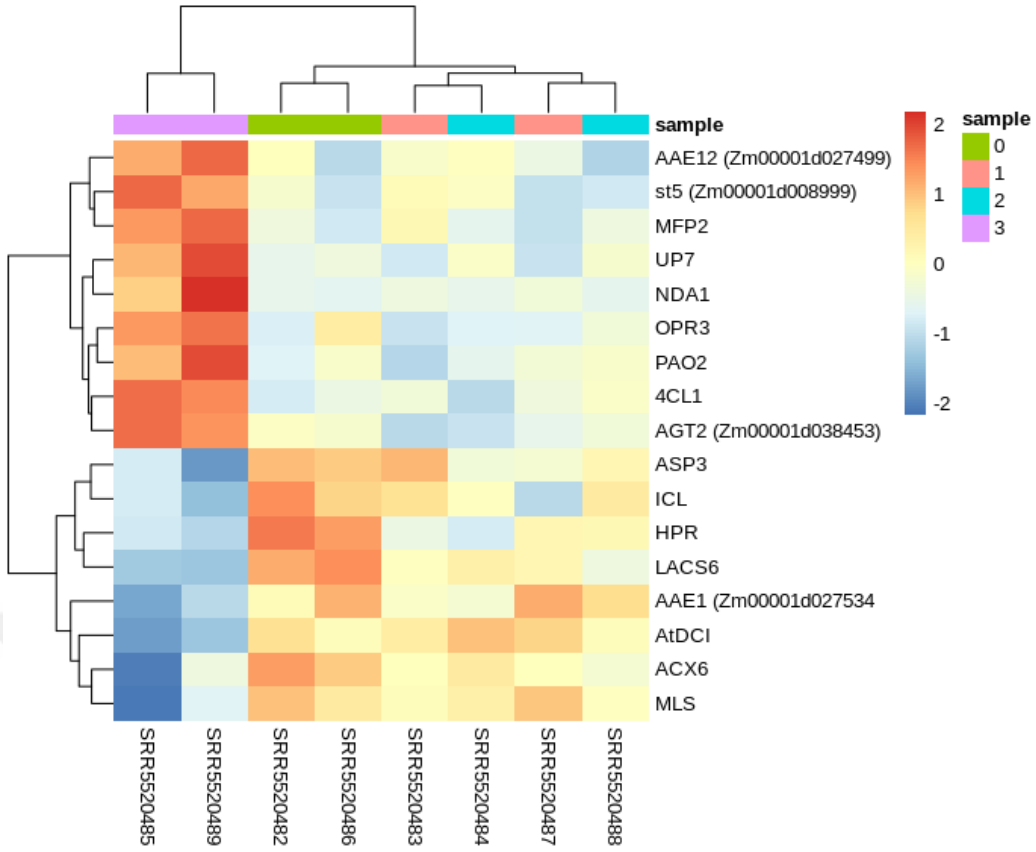


Figure 4.10. Expression patterns of 17 peroxisome related DEGs of SRP106756. Heatmap was generated based on the z-score transformed vst.

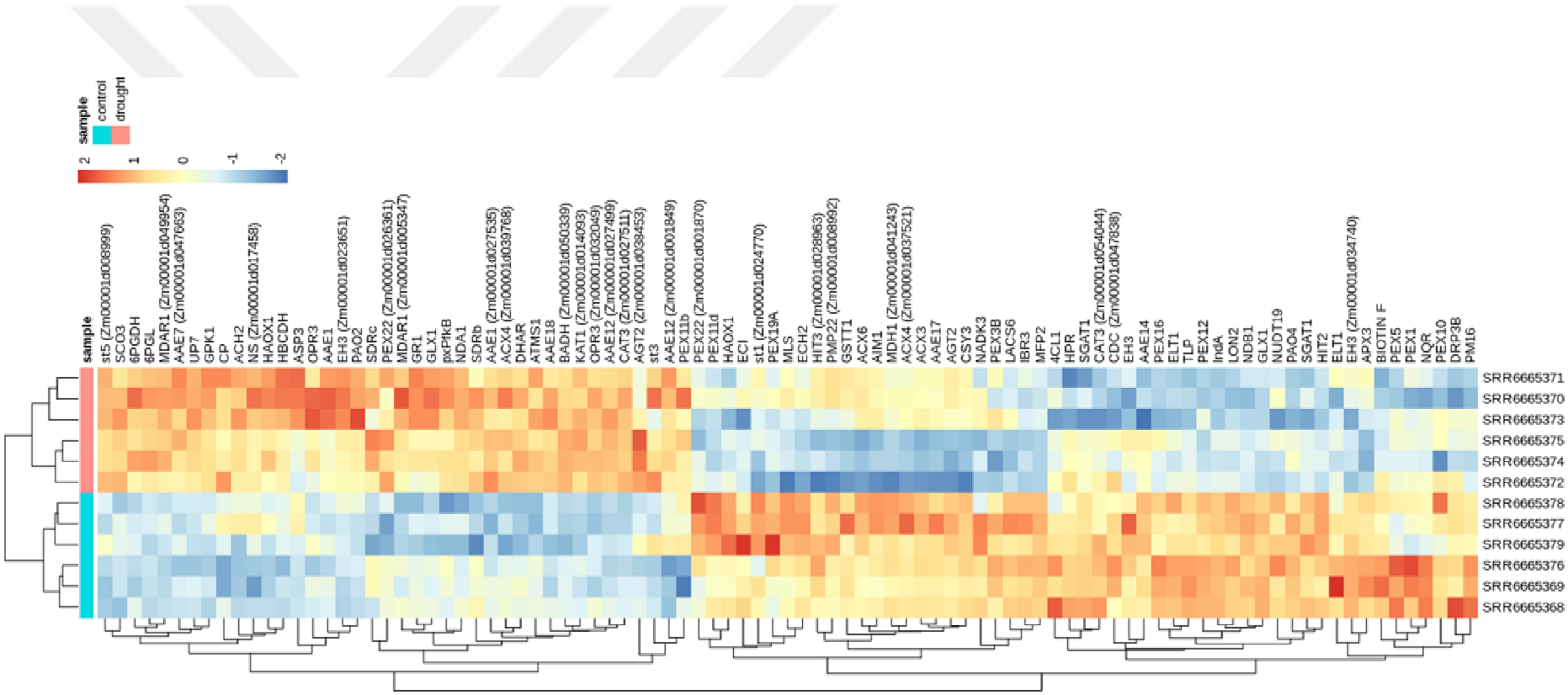


Figure 4.11. Expression patterns of 93 peroxisome related DEGs of SRP132035. Heatmap was generated based on the z-score transformed vst.

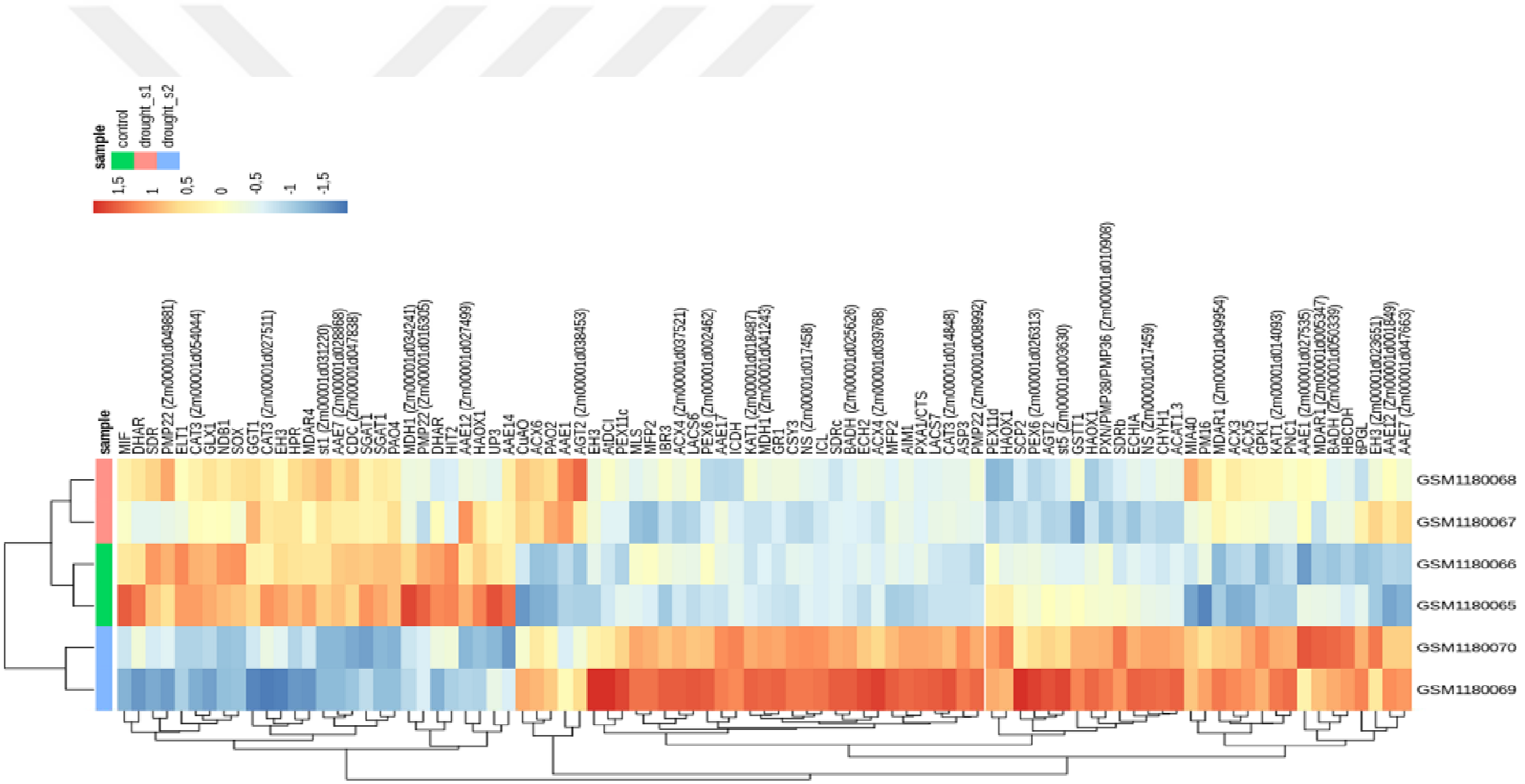


Figure 4.12. Expression patterns of 91 peroxisome related DEGs of GSE48507. Heatmap was generated based on the z-score transformed vst.

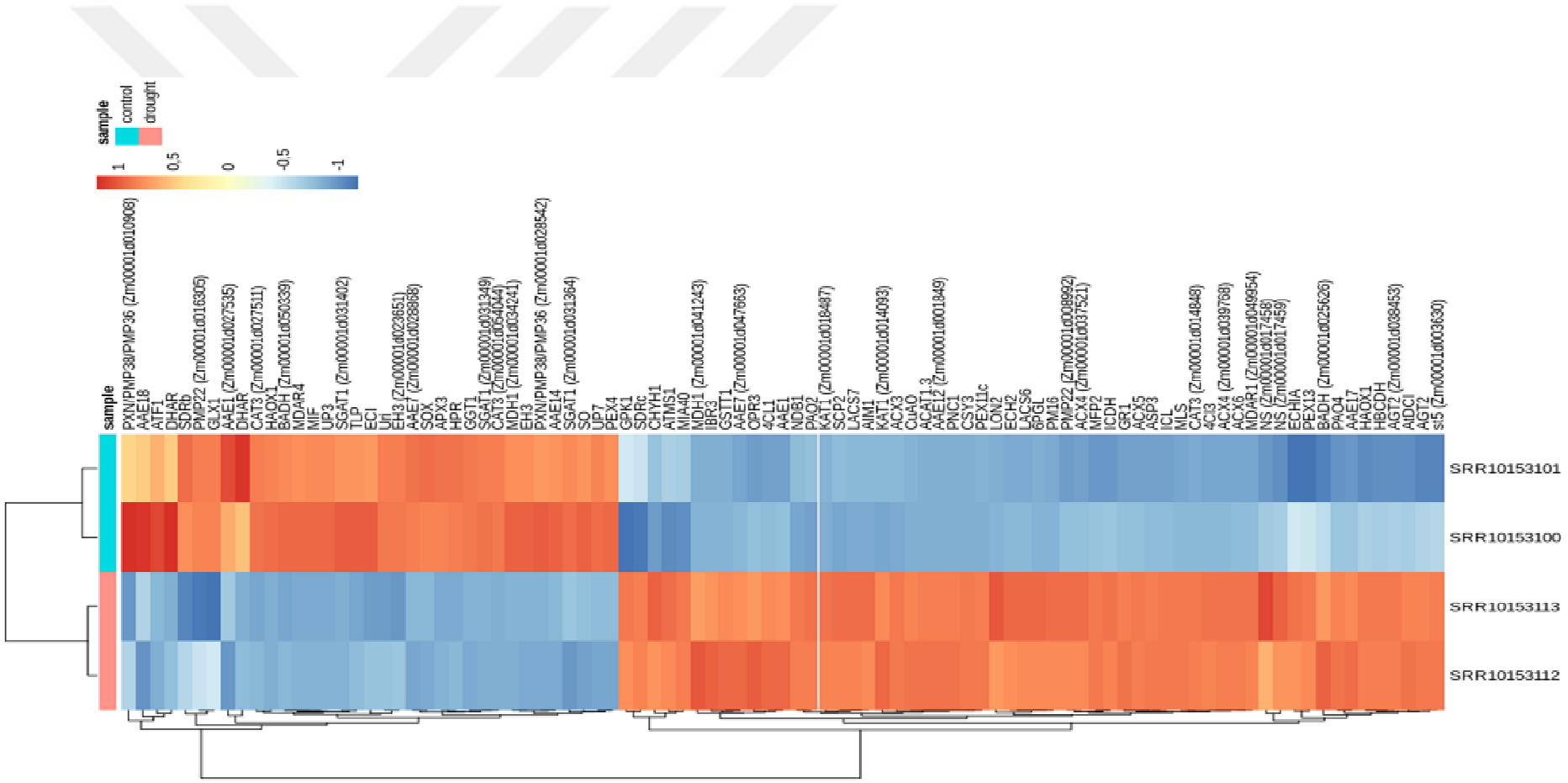


Figure 4.13. Expression patterns of 93 peroxisome related DEGs of GSE137780_b104. Heatmap was generated based on the z-score transformed vst.

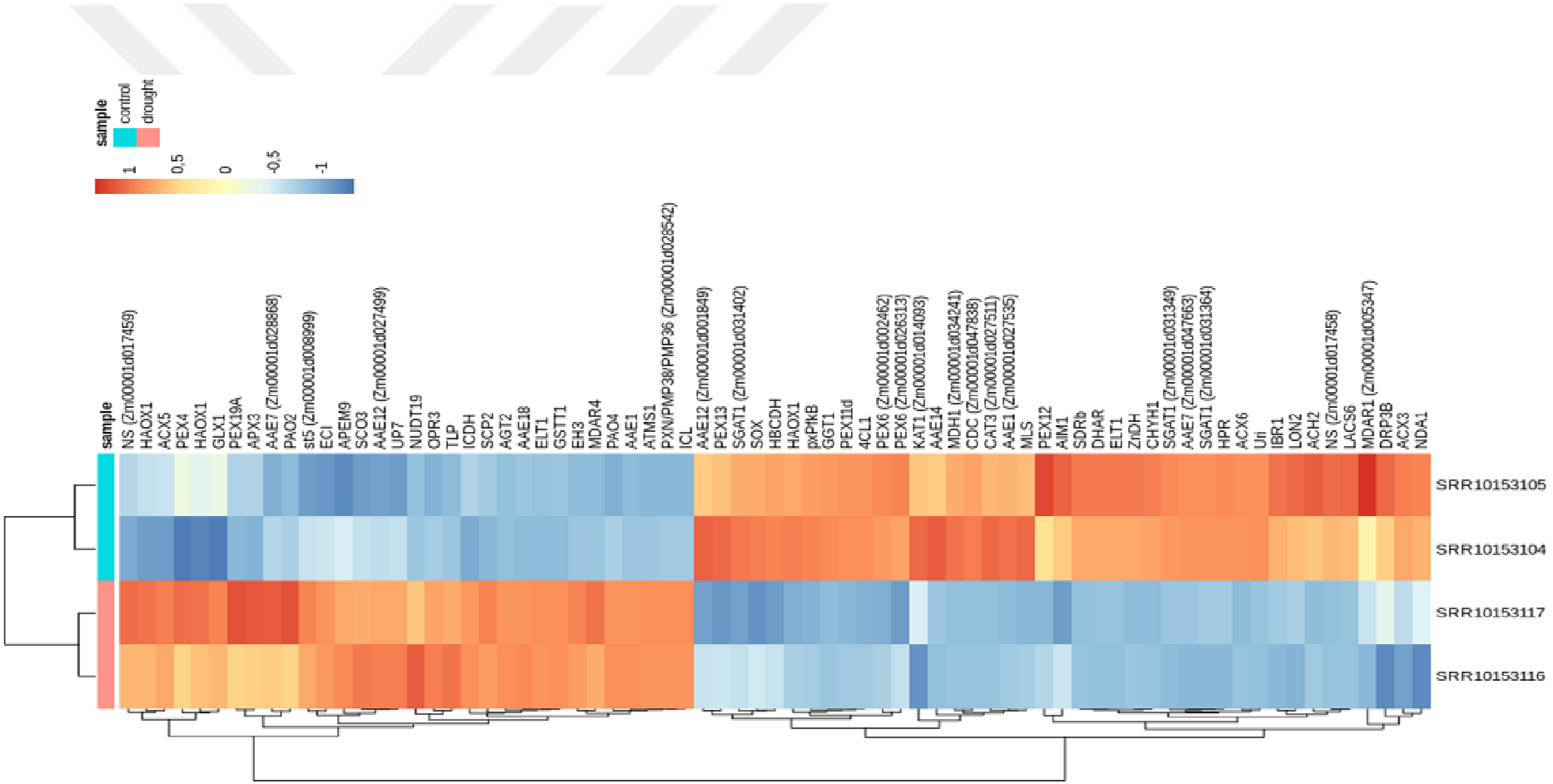


Figure 4.14. Expression patterns of 73 peroxisome related DEGs of GSE137780_DH4866. Heatmap was generated based on the z-score.

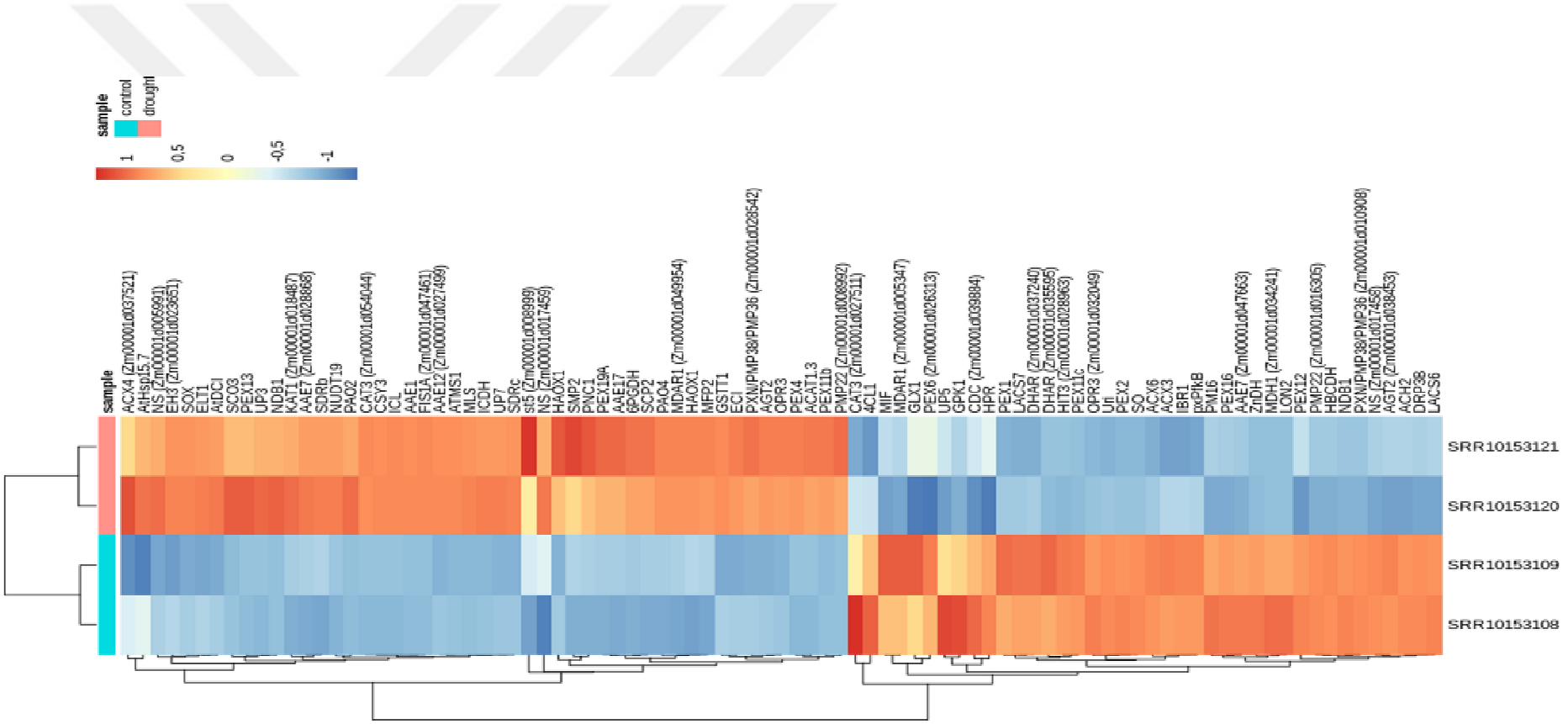


Figure 4.15. Expression patterns of 89 peroxisome related DEGs of GSE137780_W22. Heatmap was generated based on the z-score transformed vst.

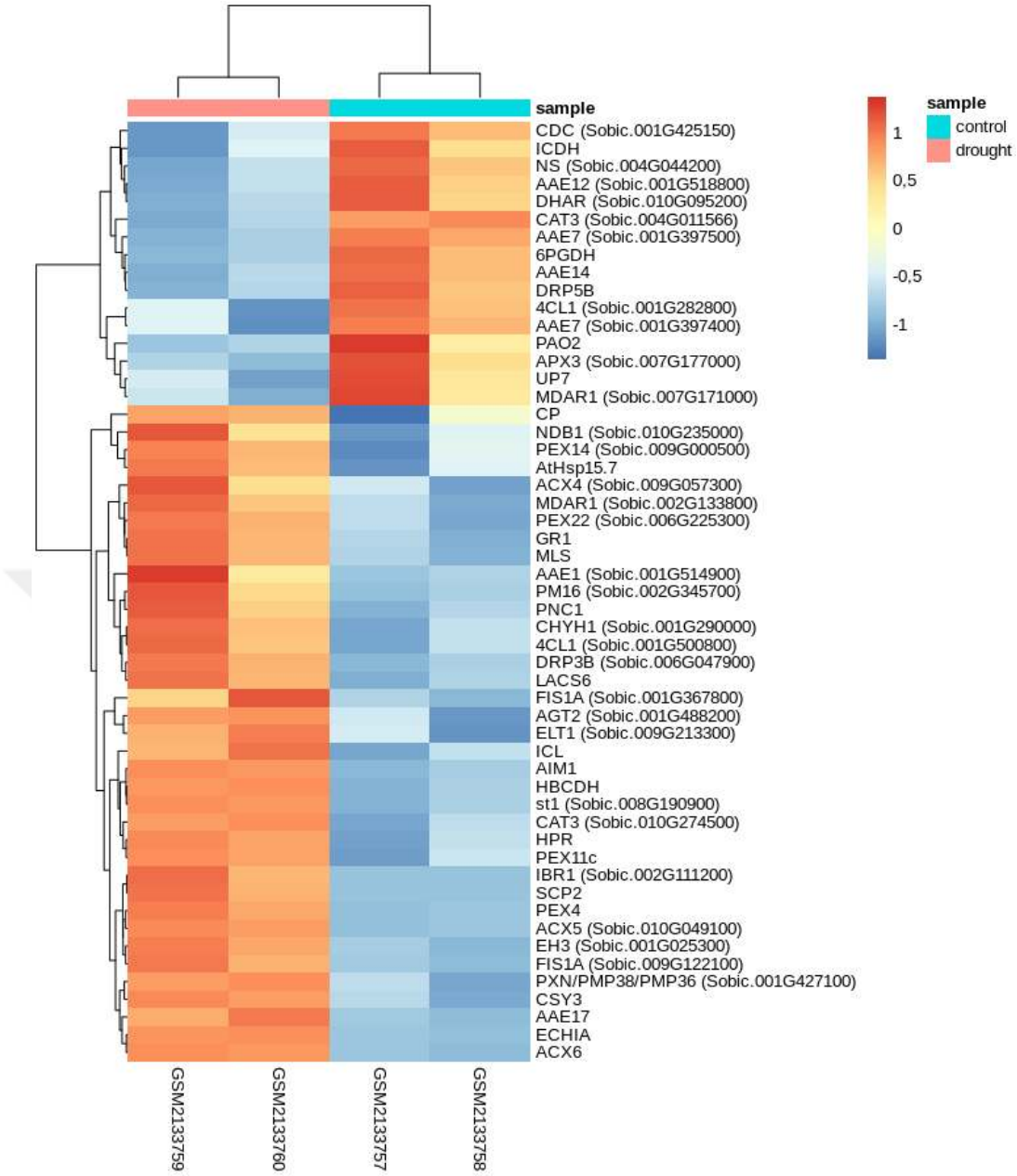


Figure 4.16. Expression patterns of 53 peroxisome related DEGs of GSE80699_IS22330. Heatmap was generated based on the z-score transformed vst.

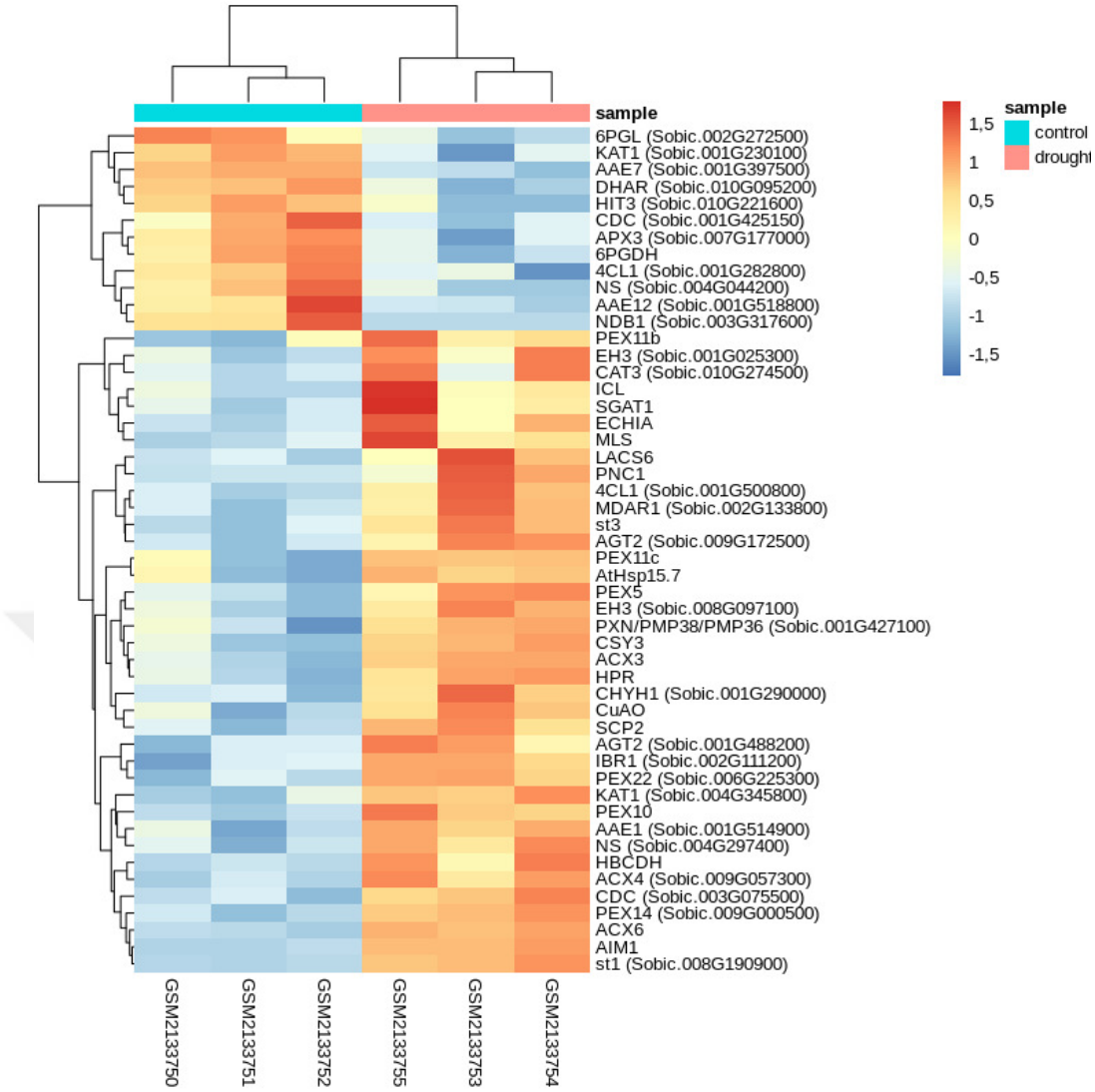


Figure 4.17. Expression patterns of 50 peroxisome related DEGs of GSE80699_IS22351. Heatmap was generated based on the z-score transformed vst.

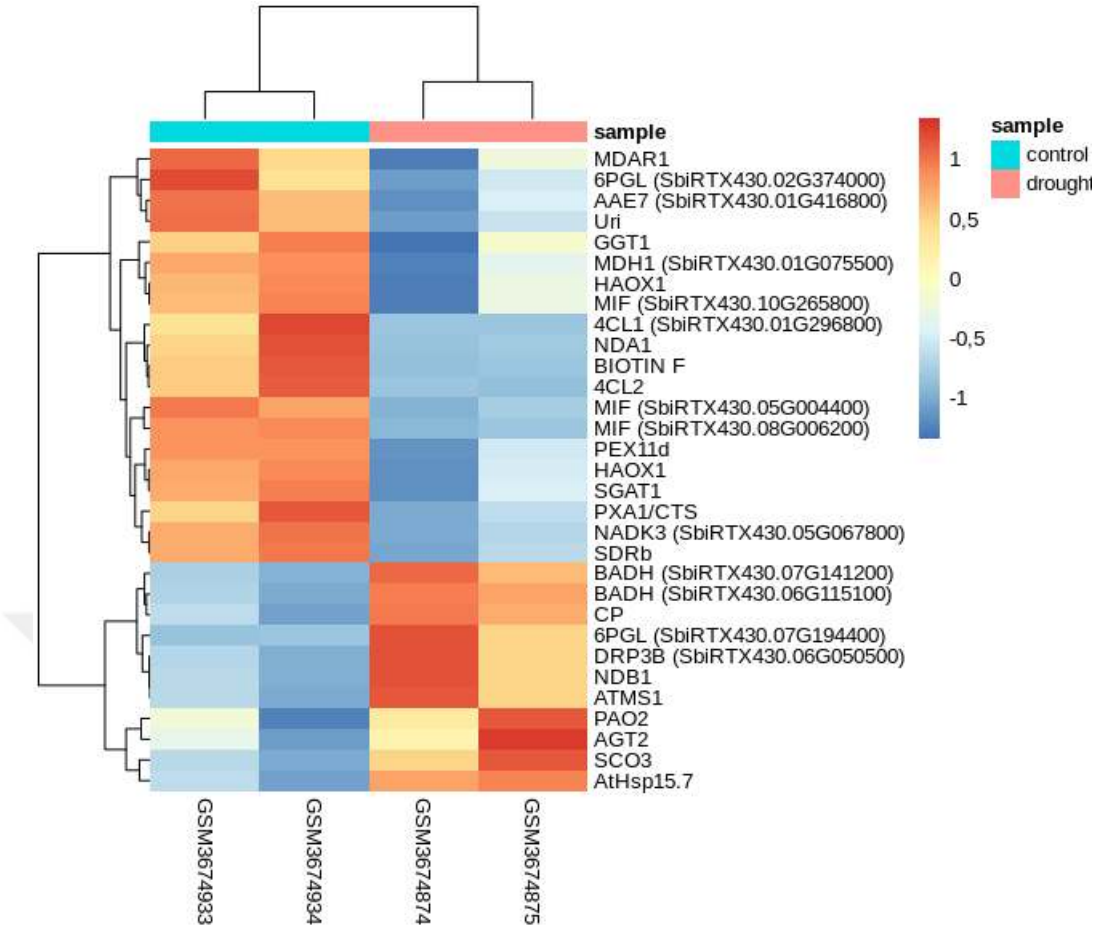


Figure 4.18. Expression patterns of 31 peroxisome related DEGs of GSE128441_week3. Heatmap was generated based on the z-score transformed vst.

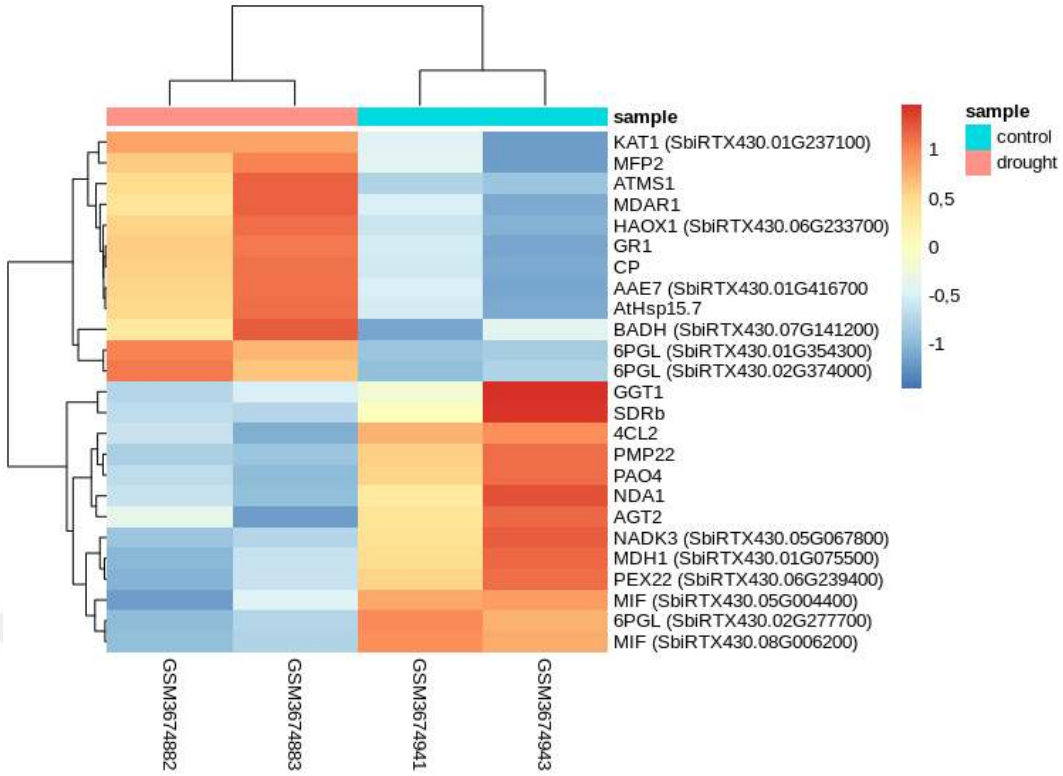


Figure 4.19. Expression patterns of 25 peroxisome related DEGs of GSE128441_week 6. Heatmap was generated based on the z-score transformed vst.