T.R.
BOLU ABANT İZZET BAYSAL UNIVERSITY
INSTITUTE OF GRADUATE STUDIES
DEPARTMENT OF BIOLOGY

META ANALYSIS OF RNA-SEQ DATA FOR SUGAR BEET (*BETA VULGARIS L.)* UNDER CERTAIN BIOTIC AND ABIOTIC STRESSES

MASTER OF SCIENCE

BURAK BULUT

ACADEMIC SUPERVISOR
ASSOC. PROF. DR. SONGÜL GÜREL
ACADEMIC CO-SUPERVIZOR
ASSIST. PROF. DR. ERCAN SELÇUK ÜNLÜ

BOLU, JANUARY - 2022

APPROVAL OF THE THESIS

META ANALYSIS OF RNA-SEQ DATA FOR SUGAR BEET (*BETA VULGARIS L.*) UNDER CERTAIN BIOTIC AND ABIOTIC STRESSES submitted by **BURAK BULUT** and defended before the Examining Committe Members listed below in partial fulfillment of the requirements for the degree of **Master of Science** in **Department of Biology, Institute of Graduate Studies of Bolu Abant Izzet Baysal University in 18.01.2022** by

Examining Committee Members	**Signature**
Supervisor Assoc. Prof. Dr. Songül GÜREL Bolu Abant İzzet Baysal University	.....
Member Prof. Dr. Nusret ZENCİRCİ Bolu Abant İzzet Baysal University	.....
Member Prof. Dr. Mehmet Cengiz BALOĞLU Kastamonu University	.....

Prof. Dr. İbrahim KÜRTÜL

Director of Institute 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 10/12/2021 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 15 %.

_____

Burak BULUT

To the *Feather* of *Peacock*,

what makes Darwin sick everytime he gazed at it...

ABSTRACT

META ANALYSIS OF RNA-SEQ DATA FOR SUGAR BEET (*BETA VULGARIS L.)* UNDER CERTAIN BIOTIC AND ABIOTIC STRESSES
MSC THESIS
BURAK BULUT
BOLU ABANT IZZET BAYSAL UNIVERSITY
INSTITUTE OF GRADUATE STUDIES
DEPARTMENT OF BIOLOGY
(SUPERVISOR: ASSOC. PROF. DR. SONGÜL GÜREL)
(CO-SUPERVISOR: ASSIST. PROF. DR. ERCAN SELÇUK ÜNLÜ)
BOLU, JANUARY 2022
XIII + 51 PAGES

Sugar beet *(Beta vulgaris)* meets the 30% of world sugar production. Soil pollution, biotic and abiotic factors in production areas greatly reduce product quantity and quality. Such stress conditions are drought, salt, heat, light, nematod bacteria and fungi infections. Sugar beet differently responds to biotic and abiotic stresses. Understanding *Beta vulgaris* response to stress at the molecular level can help both in increasing the yield and quality in crop. Transcriptome studies performed under stress conditions can shed light on the responses of plants at the molecular level. In this study four different stress condition data has been used. Two of them are biotic stresses (nematode and fungi infection), two of them are abiotic stresses (ABA treatment and salt stress). Meta-analysis can help to find common responses under different stress conditions. In this study, we performed meta-analysis of studies conducted under biotic and abiotic stress conditions and revealed ten genes that differentially expressed under both biotic and abiotic stress conditions. Biotic stresses have 460 common genes and abiotic stresses have 1031 stress related genes. Downregulation of phenophytinase and upregulation of mannitol dehydrogenase may indicate the increasing photosynthesis rate. Nevertheless, further studies are necessary for identification of genes and understand the whole mechanism of *B. vulgaris* stress response.

KEYWORDS: *Beta vulgaris*, sugarbeet, RNA-Seq, Transcriptome, Meta analysis

ÖZET

BELİRLİ BİYOTİK VE ABİYOTİK STRESLERLE İLİŞKİLİ ŞEKER PANCARI (*BETA VULGARİS L.*) RNA-SEQ VERİLERİNİN META ANALİZİ
YÜKSEK LİSANS TEZİ
BURAK BULUT
BOLU ABANT İZZET BAYSAL ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİYOLOJİ ANABİLİM DALI
(TEZ DANIŞMANI: DOÇ. DR. SONGÜL GÜREL)
(İKİNCİ DANIŞMAN: DR. ÖĞRETİM ÜYESİ ERCAN SELÇUK ÜNLÜ
BOLU, OCAK - 2022
XIII + 54 SAYFA

Şeker pancarı (*Beta vulgaris*) dünya şeker üretiminin %30'unu karşılamaktadır. Üretim alanlarındaki toprak kirliliği, biyotik ve abiyotik faktörler, ürün miktarını ve kalitesini büyük ölçüde düşürür. Bu stres koşulları kuraklık, tuz, ısı, ışık, nematod, bakteri ve mantar enfeksiyonlarıdır. Şeker pancarı, biyotik ve abiyotik streslere farklı tepkiler verir. *Beta vulgaris*'in strese tepkisini moleküler düzeyde anlamak, mahsulde hem verimi hem de kaliteyi artırmaya yardımcı olabilir. Stres koşulları altında yapılan transkriptom çalışmaları, bitkilerin moleküler düzeydeki tepkilerine ışık tutabilir. Bu çalışmada dört farklı stres durumu verisi kullanılmıştır. Bunlardan ikisi biyotik stresler (nematod ve mantar enfeksiyonu), ikisi abiyotik streslerdir (ABA tedavisi ve tuz stresi). Meta-analiz, farklı stres koşulları altında ortak tepkiler bulmaya yardımcı olabilir. Bu çalışmada, biyotik ve abiyotik stres koşulları altında yürütülen çalışmaların meta-analizini gerçekleştirdik ve hem biyotik hem de abiyotik stres koşullarında farklı şekilde ifade edilen on geni ortaya çıkardık. Biyotik streslerin 460 ortak geni ve abiyotik streslerin 1031 stresle ilgili geni vardır. Fenofitinazın aşağı regülasyonu ve mannitol dehidrojenazın yukarı regülasyonu, artan fotosentez hızının göstergesi olabilir. Bununla birlikte, genlerin tanımlanması ve *B. vulgaris* stres yanıtının tüm mekanizmasının anlaşılması için daha ileri çalışmalara ihtiyaç vardır.

ANAHTAR KELİMELER: *Beta vulgaris*, şeker pancarı, RNA-Seq, Trankriptom, Meta analiz

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

LIST OF PICTURES

Page

LIST OF ABBREVIATIONS AND SYMBOLS

ROS : Reactive oxygen species

ABA : Abscisic acid

MLPs : Major latex-like proteins

RNS : Reactive nitrogen species

OXPHOS : Oxidative phosphorylation

NADH : Nicorinamide adenine dinucleotide

TCA : Tricarboxylic acid

ADP : Adenosine diphosphate

MRR : Mitochondrial retrograde signaling

KEGG : Kyoto Encyclopedia of Genes and Genomes

NCBI : National Center for Biotechnology Information

Phytozome : Plant Comparative Genomics portal of the Department of Energy's Joint Genome Institute

GO : The Gene Ontology

RNA : ribonucleic acid

mRNA : messenger RNA

ncRNA : non-coding RNA

DNA : Deoxyribose Nucleic Acid

cDNA : complementary DNA

NGS : next-generation sequencing

TGS : thirth generation sequencing

MSA : Multiple Sequence Alignment

ACKNOWLEDGEMENTS

I would like to express my gratitude to the people who invited me to the academia supper, hope it will not be the last;

My advisor, **Assoc. Prof. Dr. Songül GÜREL**, one of the nicest people I have ever met, showed me the scientific way and her support despite her busy schedule. You gave me the courage to fly.

My thesis exam committee members: **Prof. Dr. Nusret ZENCİRCİ, Prof. Dr. Mehmet Cengiz BALOĞLU, Assoc. Prof. Dr. Fatma PEHLİVAN KARAKAŞ, Assoc. Prof. Dr. İlker BÜYÜK** for their comments and feedbacks in the dissertation.

The academician I always admired, **Prof. Dr. Ekrem GÜREL**, who helped and constantly supported me when I was at the bottom. You gave me the opportunity of flying.

My co-advisor, my friend, my big brother, **Assist. Prof. Dr. Ercan Selçuk ÜNLÜ**, who taught me bioinformatics, science, and academia. This thesis and this study would not be accomplished without you. If I am Icarus, you are the master craftsman, my father Daedalus, who gave me my wings. Thank you sir for everything you have done for me, and thank you for the most important advice you gave me, "**RTFM**".

To **Yunus Şahin** and **Uğur Çabuk**, for our meetings, personal talks, and for their guidance when I was at a dead end. You guys were my light on this road.

To the scientist, my friend, my idol, **Dr. Ömer Can ÜNÜVAR**. Just a few people enlightened me as you do. Every time I talk with you, I can feel my neurons are reshaping.

To my family and friends, for supporting all the time and believing me to keep going in the way of my academic career.

Finally, to the **Suns** of this story. I am aware of you and if one day I fly too close to you, falling will be my fault.

I thank you all from the bottom of my heart!

Burak BULUT

Bolu, January 2022

1. INTRODUCTION

1.1 Beta vulgaris

Sugar beet (*Beta vulgaris* subsp. *vulgaris*) is the main sugar production crop in the temperate climatic zone and, meets nearly 30% of the world's sugar production, the remaining 70% produced by sugar cane (1,2). *B. vulgaris* which is shown in picture 1.2, belongs to the order Caryophyllales. The subspecies of vulgaris include leafy beets (chard), garden beets (red), fodder beets (forage), and sugar beets. Among these, leafy beets are the oldest crop that has been growing by humans.

Domain: Eukaryota; Kingdom: Plantae; Phylum: Spermatophyta; Subphylum: Angiospermae; Class: Dicotyledonae; Order: Caryophyllales; Family: Chenopodiaceae; Genus: Beta; Species: *Beta vulgaris* (3)

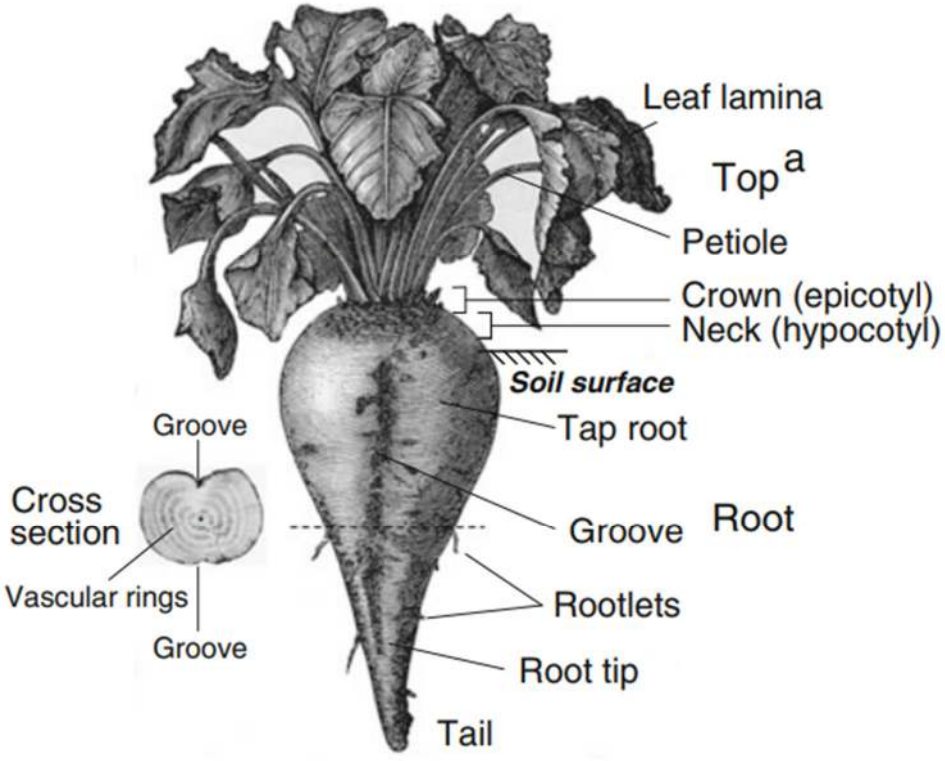

Figure 1.1 Sugar beet drawing with the common names of the main parts (4)

1.2 History of *B. vulgaris*

Sugar beet was domesticated in the 18th century. They are grown in the northern hemisphere and are usually planted in the early spring. Their vegetative

organs develop in the first and generative organs in the next year. The morphology of *Beta vulgaris* has been shown in Figure 1.1. Seeds are either monogerm or polygerm. The plant height varies between 85-180 cm depending on the soil conditions, climate, and species. Harvesting begins 5-9 months later also depending on the soil and environmental conditions (5). Picture 1.1 shows a farming area of sugar beet. After the harvesting, beets can deteriorate within a few days. Removed beet leaves are plowed into the soil but the leaves can be considered as animal feed or even feedstock for biofuels (6).

Picture 1.1 A typical *Beta vulgaris* farming area. Photo supplied by Assoc. Prof. Dr. Songül GÜREL

1.3 Properties of *B. vulgaris*

The structure of sugar beet consists of 4-5% cell tissue, 4-5% chemically bound water, and 90-95% sap. The composition of beet juice consists of 12-21% sugar (sucrose), 1-1,6% other non-sugar organic substances, 0.8% inorganic salts (6). The organism's root composition makes it valuable for sugar production. Roots have high soluble sugar content, high hemicellulose, and pectin carbohydrate, and low lignin content. Dry basis protein content varies from 2% to 10% of the roots. The surface of harvested sugar beet showed a variety of bacterial species. With the domination of *Proteobacteria*, the surface contained 530 bacterial species (5).

Table 1.1 World sugar beet production between 2016-2021, The Data obtained from Republic of Turkey Ministry of Agriculture and Forestry (7)

	2016/17	2017/18	2018/19	2019/20	2020/21	Change (%)
Cultivation Area	4.592	4.990	4.798	4.609	-	-3.9
Yield	60,72	62.93	57.05	60.42	-	5.9
Production (ton)	278.835	313.989	273.712	278.498	-	1.7
Sugar Production	174.050	194.529	179.347	165.496	121.866	9.9
Sugar Consumption	170.612	172.687	172.370	170.225	173.760	2.1
Sugar Stocks	44.482	42.550	52.116	53.232	46.243	-13.1
Sugar Imports	54.445	54.217	52.030	51.657	54.507	5.5
Sugar Export	59.039	65.097	57.090	53.273	65.333	22.6

While the medicinal and food uses are older, the very first modern sugar beet selection goes on in the middle 18th century, German Silesia. It is also known that sugar beet is used as food in ancient Egypt. In 1747, German chemist Andreas Marggraf discovered that after a crude extraction from sugar beetroot, the crystals formed as same as the sugar cane. His student Karl Achard developed the process of sugar extraction from sugar beet, and this was the first selection of high sugar type beets (8). Apart from being a food source, sugar beet is used as livestock feed and becoming a promising biofuel source alternative to fossil fuels. As an example of livestock feed, beet tops (leaves and petioles) are an excellent source of carbohydrates, vitamin A and protein. In industrial fields, they are widely used in alcohol, pharmaceuticals, and baker's yeast. The waste lime of sugar beet is also used in increasing the pH level of the soil (2,9).

Picture 1.2 *Beta vulgaris.* Photo supplied by Assoc. Prof. Dr. Songül GÜREL

The plant's genome has been sequenced in 2014 from industrial diploid sugar beet (2n=18) with an estimated size of 714-758 megabases and 27,421 predicted protein-coding genes. For molecular breeding and comparative plant

genomics study, sequencing of sugar beet's genome was very important. Yet, genome sequencing facilities need further studies for gene function characterization (10). In the same year, the plant's chloroplast genome was also sequenced (6).

1.4 Stress in Plants

Global warming is the effect of human activities on climate change where the burning of fossil fuels releases greenhouse gases into the atmosphere. The most important greenhouse gas released by human activities is carbon dioxide. Greenhouse gases absorb the infrared radiation that comes from the sun and are reflected by Earth's surface. They simply act as a blanket on the Earth and cause warming up the Earth (11). The warming up of Earth's temperature caused the sea-level rising, reduction of polar ice caps, and rising of the ocean's temperature. 130 years of recorded data shows that the global temperature has risen 0.8-1.0°C (12).

Effects of global warming can be long-lasting and even irreversible. Some of the worst scenarios are considered as rainfall and snow patterns change, drought and severe storms, melting of glaciers, increase in floods, rise in sea level, higher humidity, and changes in plant and animal behaviors but maybe the most important effect would be agricultural sustainability. Agriculture is crucial for humans (12). With the growing population number, humans need more food than ever. With these Two problems on the horizon, crops and agricultural yields need a different approach for adaptation and high demand. One of these highly demanded foods is sugar. From the evolutionary perspective, sugar increases body fat and prepares the animal for food scarcity times (13).

In nature, plants come across the stress conditions such as biotic and abiotic conditions (14). Briefly shown in Figure 1.2. These stress conditions have negative effects on the plants and decrease the crop yield by changing its physiology, phenotype, biochemical and molecular structure. Evolution helped plants to develop complex defense mechanisms to respond the environmental stresses. A better understanding of an organism's response to the stress conditions by molecular techniques with the help of bioinformatic studies and developing new molecular strategies may help us to reduce the negative effects of biotic and abiotic stresses (15,16).

Biotic Stress Factors ● Insects ● Viruses ● Bacteria ● Fungi ● Nematodes ● Herbivores ● Weeds ● Arachnids		**Abiotic Stress Factors** ● Cold ● Heat ● Salinity ● Drought ● Excess water ● Radiations ● Oxidative stress ● Chemical pollutants

Figure 1.2 Stress Factors (modified from (17))

1.4.1 Biotic Stress

Biotic stresses are the infection of plants by pathogens (viruses, nematodes, bacteria, and fungi) or the attack of herbivore pests. Biotic stress agents mainly attacks the plant's nutrients and damages the plant's vital systems. With the rising temperature, optimum living conditions arise for pathogens. Both biotic and abiotic factors cause problems in agricultural production. Hence the plants developed strong stress mechanisms to protect themselves. They generally coordinate their response to stress conditions by signaling molecules. As an example, the biotic stress response mechanism activates mainly by ethylene signaling, salicylic acid, and jasmonic acid pathways. The effects of these signaling pathways are usually reducing photosynthesis, leaf, and root growth. They also close their stomata, produce reactive oxygen species and toxic compounds, and induce cell death. Plant's cell death response for reducing the rate of spreading the pathogen through the cell by cell (18,19). According to the study of Rizhsky et al., 2002, when multiple stress conditions (heat and drought) applied to tobacco (*Nicotiana tabacum*) and Arabidopsis (*Arabidopsis thaliana*), they induce a novel gene expression that has not been transcript either of the stress conditions alone. It should be known that the stress signaling pathways might act antagonistically to each other. For example, abscisic acid (ABA) can act as a global stress response signal but also suppress the biotic stress signaling pathways. This can have positive and negative

effects on plants' defense mechanisms. ABA is known to prevent the accumulation of defense compounds such as lignins and phenylpropanoids. In Arabidopsis, ABA treatment and drought stress increase the susceptibility of the avirulent bacteria *Pseudomonas syringae* while ABA treatment in tomato (*Solanum lycopersicum*) increases the susceptibility to virulent bacteria *Erwinia chrysanthemi* and *Botrytis cinerea* (18).

Biotic stresses cause in the range of 31-42% yield loss of sugar beet(20). While there are limited studies of *B. vulgaris* under biotic stress, one study of sugar beet under fungal stress of *Rhizoctonia solani*, revealed some pathogen-related defense mechanism genes such as BvMLP1, BvMLP2, and BvMLP3(21). Major latex-like proteins (MLPs) are members of the pathogenesis-related proteins class 10 family. The two functions of MLPs are known as binding to various hydrophobic compounds and transporting those compounds to the other organs to contribute to the expression of physiologically important ligands activity in specific organs and secondly, MLPs act as the biotic and abiotic stress response. They generate resistance against pathogens by inducing pathogen-related genes(21).

1.4.2 Abiotic Stress

Plants require water, light, carbon, and nutrient sources for growth. Abiotic stress is an environmental condition that reduces plant growth under optimum levels. Abiotic stresses are drought, heat, cold, nutrient deficiency, and excess salt or toxic metals. Temperature, drought, and salt stresses are the most common ones among abiotic stresses. These stresses cause morphological, physiological, and molecular changes in plants. Plants developed responses in such conditions but the responses depend on the tissue affected by the stress. As an example, drought stress causes different responses in roots and leaves. The transcriptional response is specific to tissues or even cells (22). As an example, water deficiency reduces water uptake into expanding cell and inhibits plant growth and activate the reactive oxygen species (ROS) genes. By growth limitation, the plant starts resistance to water flow in expanding cells and develops a larger water potential gradient. Another abiotic stress molecule is reactive nitrogen species (RNS). Both ROS and RNS modify the enzyme activities and gene regulations. ROS and RNS work coordinately in a network. Many studies confirmed the negative effects of ROS in abiotic stresses but there are few studies about RNS (22).

To achieve the maximum yield, sugar beets should have enough moisture levels and atmospheric demands for transpiration. Nevertheless, these conditions are often cannot be met, and this causes the yield loss every year. The most challenging abiotic stresses of sugar beet are drought, heat, salinity, and cold. Heavy metals, nutrient deficiency, and flooding can also cause problems. Sugar beet is considered one of the most tolerant plants to hot, dry, and mildly saline conditions (23). For example, salt stress induces the accumulation the ROS in photosynthetic and respiratory electron transport systems. According to the study of Hossain et al., 2017b, H_2O_2, which is a ROS member, was found lower under the long-term salinity than control groups. This result is unusual from the responses of other plants. It is speculated that sugar beet is quite efficient in adjusting H_2O_2 levels in long term. Antioxidant-related genes, superoxide dismutases, and peroxiredoxins were observed as upregulated under salt stress. These genes are known as ROS removers (24).

1.4.3 Mitochondria and Stress Response

The major function of mitochondria is to adenosine triphosphate synthesis, which is the energy source of cells, through oxidative photorespiration. According to other phyla, mitochondria in plants is more complex and flexible with specific pathways to support photosynthesis. Apart from being the energy factory of cells, mitochondria play a key role in other biological processes associated with nitrogen, phosphorus, carbon, and sulfur metabolism in plants. Mitochondria is a key agent in how plants respond to oxidative stress. It is crucial to understand the mitochondria's role in stress response for improvement in plant growth and biomass production (25).

The ATP synthesis is placed in mitochondria by Oxidative Phosphorylation (OXPHOS). The prerequisite molecules of OXPHOS are the formation of nicotinamide adenine dinucleotide (NADH) and flavin adenine dinucleotide (FADH2) by the oxidation of organic compounds. The oxidation of organic compounds placed in the tricarboxylic acid (TCA) cycle and its side reactions. The first step of OXPHOS is, transferring the electron of reduced NADH and FADH2 to molecular oxygen. This process is placed in the electron transport chain which is located on the inner mitochondrial membrane. By this electron transfer, mitochondria created an electrochemical proton gradient and in the second part, use

this proton gradient difference to synthesize ATP by phosphorylation of adenosine diphosphate (ADP) (26).

Biotic and abiotic stresses are well known to induce oxidative stresses in plants. Under the stress conditions, the changed ROS, protein, and ROS-induced lipid peroxidations are interacting with mitochondria which makes mitochondria an environmental sensor. While the ETC system produces a big amount of ROS under normal conditions, antioxidant enzymes dealt with it. However, under extreme conditions, these defense systems are overwhelmed and that causes ROS to accumulate. Inhibition of mitochondrial enzymes by ROS is a known effect (25,26).

Under such extreme conditions, nucleus can recieve signals from mitochondria's functioning. This signaling pathway starts a transcriptional response in the nucleus and this process is called mitochondrial retrograde regulation (MRR). Many of the transcriptional regulatory factors involved in MRR was identified and they include a wide range of functional classes such as transcriptional factors and cyclin-dependent kinases. It has been shown that, between auxin and retrograde signaling, there is an antagonistic relationship. During the MRR, two molecules bind to the endoplasmic reticulum, which relocates to the nucleus to start transcriptional regulation. This makes the endoplasmic reticulum an important part of MRR.

Plant hormones play key roles in plant growth and development and also stress response. Auxin, gibberellin, cytokinin, and ethylene are the hormones involved in growth and development where, abscisic acid, salicylic acid, jasmonates, and brassinosteroids are involved in stress response. The connection between hormones (especially ABA), ROS, and MRR regulation has been shown in various studies (27–29).

1.4.4 Stress Response in *Beta vulgaris*

In sugar beet, most of the stress studies focused on salt and drought stress. Like any other plant, the stress conditions cause ROS accumulation in sugar beets. Under the salt stress which sugar beet is highly tolerant, mostly signal transduction, phosphorylation, and redox balance genes are playing important roles. Additionally, ROS scavenging system-related genes are significantly differentiated. Photosynthesis-related genes do not affect by salt and drought stress. Some studies point that some photosynthesis-related genes are even upregulated during salt and

drought stress conditions. Sugar beet's other responses to salt stress are upregulation of ribosome biogenesis, cell wall building, and cell wall expansion genes. These morphological changes are thought to be related to osmotic pressure balancing.

The biotic stress response is poorly studied in *Beta vulgaris.* One of the biotic stress condition studies in insect pest of sugar beet. The study revealed the 3 signaling pathway-related genes which are thought to be involved in biotic stress response, methyl jasmonate, calicylic acid, and ethylene. Most of the biotic stress response genes are metabolic and catalytic genes (87, 88).

1.5 Bioinformatics

Bioinformatics is a rapidly growing interdisciplinary field that combination of biology, computational sciences, mathematics, and physics. With the advances in both molecular techniques and technology, bioinformatic fields growing and evolving faster than ever. There are three main aims of bioinformatics. First, is the organization of data that is already exists and submission of new data in a form that allows scientists to reach easily. Such data storage is called databases. NCBI, ENSEMBL, and Uniprot are examples of biological databases. The second aim of bioinformatics is to develop new tools and resources that can be used in the analysis of data. For example, BLAST+ can be used for blasting a protein or DNA sequence. The third aim is to use bioinformatical tools to analyze the data and make biologically meaningful results (30). It may be considered that the greatest achievement of the bioinformatical field is the Human Genome Project. The human genome project is a scientific project that was aiming the sequence the base pairs of human DNA. It officially launched in 1990 and was completed on April 14, 2003. Nowadays the sequencing is way easier with the help of new technologies. Apart from genome sequencing techniques, bioinformatical tools are used in a vast array of other projects such as gene expression levels, prediction of protein structure and function, detection of gene regulation networks, gene-disease interactions, and prediction and analysis of molecular pathways (31).

1.5.1 Biological Databases

Biological databases are crucial for bioinformaticians. Such databases are the National Center for Biotechnology Information (NCBI), Kyoto Encyclopedia of Genes and Genomes (KEGG), Ensembl and Plant Comparative Genomics portal of the Department of Energy's Joint Genome Institute (Phytozome), etc. These

databases offer scientists a variety of biological data of a broad range of organisms(32). In 1995, KEGG was developed as a ground-in database for biological resources. The main goal was to produce map genes in the genome. When KEGG was first developed it concluded only four databases, PATHWAY, GENES, COMPOUND, and ENZYME. Also, during that time database contained only metabolic pathways. KEGG was only able to perform mapping through ENZYME. Later, the KEGG database expanded and included more data sources into their base. PATHWAY database supported with BRITE and MODULE, GENES database with GENOME, COMPOUND database with GLYCAN and REACTION, and KO replaced the ENZYME database. After that expansion, KEGG was used for not only genomics but also transcriptomics, proteomics, glycomics, metabolomics, and metagenomics (33). NCBI on the other hand was created in 1988 at the National Library of Medicine. The main mission of the NCBI database is to develop new technologies to understand the fundamental molecular and genetic processes to control health and disease. Its aim was these four major tasks, to create automated systems to create and analyze information of molecular biology, biochemistry, and genetics, computer-based biological information processing with enhanced methods, use of databases by biotechnologists and medical personnel, to gather biotechnological information worldwide (34). The Ensembl database is a genome annotation platform that is capable of integrating experimental data against a reference genome. Compared to KEGG and NCBI databases, the Ensembl database is limited. It only contains the genomic information of vertebrate subphylum and model organisms (35). Lastly, Phytozome is a database specific to plants some algae, and selected species that have been sequenced elsewhere. Phytozome provides genome, gene family data, and analysis (36). The Gene Ontology (GO) project is tool-based website and program that created evidence-supported annotations to describe the cellular, molecular and biological function of genomic products by classifying them to their ontologies. GO annotation is known as the connection between a class of annotation and a gene product. These connections are evidence-supported. The tools also provide the reference of evidence supporting the association (37).

Table 1.2 Some of the Biological Database Sources

Database	Description	Link
NCBI	Database for advances the science and medicine through providing genomic and biomedical information.	ncbi.nlm.nih.gov
KEGG	Database for understanding the biological system from molecular-level function.	genome.jp/kegg
Phytozome	Plant, algae and selected organism database.	phytozome-next.jgi.doe.gov
GO	Largest source of gene function information.	geneontology.org

1.5.2 RNA-Seq

As a field of bioinformatics and molecular biology, functional genomics is various omics techniques to describe genes and proteins functions and interactions with each other. Functional genomics techniques are known as proteomics metabolomics and transcriptomics. Transcriptomics technologies are the study of an organism's ribonucleic acid (RNA) transcripts of a tissue or cell under a certain functional or developmental stage (38). These transcripts include messenger RNA (mRNA) and non-coding RNA (ncRNA). mRNA is regulation of genomic information stored in Deoxyribose Nucleic Acid (DNA) to a protein where ncRNA regulates expression of genes, protein synthesis, and various cellular activities.

Walter Fiers and his colleagues sequenced the complete transcripts of bacteriophage MS2 in the 1970s. The MS2 bacteriophages genome length is 3569 base pairs long. Considering nowadays transcriptomic studies, the genome of the bacteriophage is very small. This is because of the RNA's single-stranded structure. RNA is easily degraded by the enzyme called RNases. Therefore, sequencing of longer transcriptomics was very hard, nevertheless, it was the very first transcriptomic study. After the discovery of the reverse transcriptase enzyme, the rigid mRNA and ncRNA molecule can be converted to a more stable, double-helix DNA molecule. The synthesized DNA molecule is called complementary DNA (cDNA). The cDNA synthesized from RNA process called Sanger method after the name of British chemist Sanger who discovered this system. Based on this new method microarray technology (also known as micro-chip) has begun. In recent years, RNA-Seq technology also has begun. This new method is based on next-generation sequencing (NGS) where thousands or even millions of short reads can sequence with a low error rate (38). Due to the limitations in NGS such as size of

length reads and the error rate, thirth generation sequencing (TGS) has invented. This new era is able to detect more errors and sequence longer reads. Pacific Bioscience and Oxford Nanopore developed methods based on TGS (39). Table 1.2 shows the sequencing technologies.

Table 1.3 Comparation of Genome Sequencing Techniques. bp means base-pair. (89)

	First Generation Sequencing	Second Generation Sequencing	Thirth Generation Sequencing
Used template DNA	Fragmented DNA	Fragmented DNA	Single strand DNA
Read Length	800-1000 bp	Short (Sanger method)	More than 1000 bp
Accuracy	High	High	High
Efficiency	Low	High	High
Cost	High	High	Low
Library	cDNA	cDNA	cDNA or RNA
Time	Hours	Days	Minutes

1.5.3 De Novo Assembly

One of the highly used genome sequencing techniques is De novo assembly. De novo assembly can be described as a process of reconstructing a genome with short reads. Unlike the other genome sequencing techniques, de novo assembly does not require a reference genome. Because whole-genome assembly cannot be accomplished at one glance reading, de novo assembly uses many fragmented short reads. It is also described as a giant jigsaw puzzle where many fragmented sequences are assembled by finding the shared regions up to chromosome level (40,41). Three basic steps of de novo assembly are contig assembly, scaffolding, and gap filling. During the contig assembly step, the short reads assembled in long consensus sequences called contigs, without any gaps in them. In the scaffolding process, the large contigs are paired up by large reads. These long inserts are several kilobases in length and are called a scaffold. When the contigs are scaffolded, if there is no overlapping between the contigs, there are some gaps that remain. These gaps are filled by independent reads which is the gap-filling step. To increase the quality of assembly, scaffolding and gap-filling steps can be repeated (40,41).

1.5.4 Transcriptome Mapping

On the other hand, if there is a reference genome of the desired organism, mapping tools can be used. However, mapping the raw reads (RNA-Seq data) to a reference genome is a bit complicated compared to de novo method. This process

includes a quality check of raw reads, alignment (mapping), and read counting (in Figure 1.4) (42).

Quality check of raw data is crucial for mapping because RNA-Seq library preparation is still developing technology and bad quality can directly affect the coverage of mapping. Quality checking includes read quality, RNA contamination, GC content, nucleotide composition bias, etc. Raw sequence quality score can be shown with Phred score. Phred quality score is developed by the Phred program to measure base-calling reliability. For example, a Phred quality score of 30 means 1 incorrectly sequenced base in 1000. Higher the Phred score, the higher the quality of reads. Nevertheless, Phred quality score is rarely used in NGS but it is widely accepted of DNA quality scores. Nowadays, Phred scores are represented with their corresponding ASCII characters which are the numbers between 33 to 126 or "!" to "V".

Nucleotide composition and GC content are another quality mark of the reads. GC content is the percentage of the guanine or cytosine bases in the DNA. Because guanine and cytosine have 3 bonds between them, it makes them more stable. Therefore, rather than AT bases, GC content carries more biological meaning. Also, exons contain higher GC than AT content.

Another quality mark is ribosomal RNA (rRNA) or transfer RNA (tRNA) contamination. Most of the RNA-Seq study's goal is to capture only mRNA, but some structural RNAs such as rRNA and tRNA are the most abundant RNAs in the cell. They contain up to 90% of the total RNA in a cell. Therefore, before the mapping, these RNA species must be cleaned from the data (43).

Sequence alignment is the arrangement is the comparing and finding the similarities of two or more DNA/RNA or protein sequences. This method is used for aligning short reads of RNA-Seq to a reference genome which can be obtained from public databases (44). This reference-based approach is a much more sensitive and computational power-efficient way than de novo (45). Sequence alignment is sensitive because of the RNA splicing which is shown in Figure 1.3. RNA splicing is the removal of intron sequences from pre-mRNA. Several tools have been developed for sequence alignment and each one has different powerful sides. For example, Bowtie is an unspliced read aligner with high computational memory requirements. TopHat, on the other hand, uses "exon first" method for alignment. It aligns the exon parts of reads first and then split unmapped reads into shorter

sequences and aligns them independently to find splice junctions among the exons (46). HISAT2 alignment tool was developed in 2019 with high efficiency and higher alignment sensitivity. Unlike Bowtie, HISAT2 differs in building indexes. Almost all alignment tools build indexes before alignment. HISAT2 builds a whole-genome index and thousands of small local indexes for spliced alignment possible (42,47).

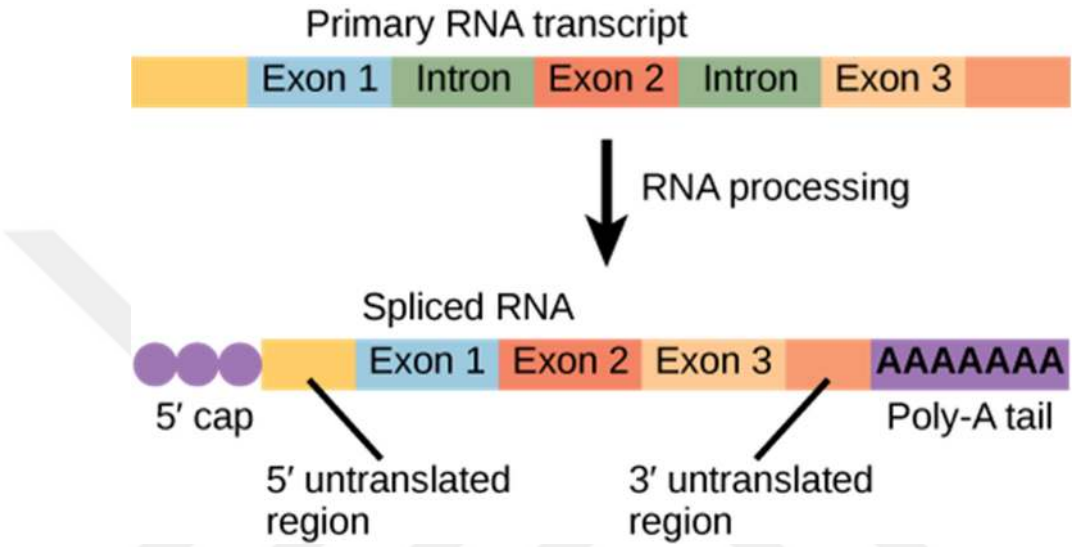

Figure 1.3 RNA splicing (48)

Read counting process starts with the genomic feature assignment before interpretation biologically. These genomic features can be genes, exons, promoters, and genomic bins. For genomic feature assignment, a reference genome and the gene annotation file are required(49). After the genomic feature assignment, quantification of gene expression is required, which can be used for differential expression of genes. Quantification is a hard step because of exon joints (50,51).

1.5.5 Differentially Expressed Genes Analysis

Differentially Expressed Genes (DEG) is a method that takes the normalized read counts data and identifies differentially expressed genes by performing statistical analysis. There is more than just one DEG analysis tool and they vary in their statistical approaches. For example, edgeR (52) and baySeq (53) uses negative binomial model where NOIseq (54) and SAMseq (55) uses non-parametric methods. Some methods have been developed to identify unknown transcripts and they are called transcript detection such as EBSeq (56) and Cuffdiff2 (57). DESeq2(58) on the other hand uses maximum likelihood estimation and takes all

gene value and move them to average value. This method is valuable for the studies where the RNA sequencing was made under different conditions and technique. Different conditions and techniques may affect the analysis and value of raw data. DESeq2 accomplish this problem by moving gene values to average gene value. This topic of bioinformation is still developing. Therefore, there is no consensus about which tool or which statistical approach is the best way to ensure the validity of data (59).

1.5.6 Meta-Analysis

Meta-analysis is a set of quantitative methodological tools that combine the outcomes of similar but independent studies to draw a general conclusion and to evaluate the similarities among studies. Meta-analysis has become one of the most important methodological developments in a variety of fields including medicine, social and biological. Because meta-analysis gives results of similar but independent studies, it is a good tool to avoid the controversial aspects of the study and make a synthesis (60).

The main goal of the meta-analysis is to understand the results of a single study despite all collected and independent studies. The collected studies (or data) should be considered if their effective size is consistent across the body of studies. If it is consistent, it means results are robust across the collected data. Effect size is a value which shows the strength of a relationship between at least two variables. If the effective size is not consistent, the possibility of implication should be considered. In conclusion, meta-analysis can address these issues where a single study cannot (61).

Most of the individual omics studies have small sample size. For seeing the general perspective, combination of multiple studies is needed. Meta-analysis improves the statistical power, accuracy, and reproducibility of combination of multiple studies. Combination of multiple transcriptomic studies also called horizontal meta-analysis. In some studies, different types of omics data are needed, this type of analysis called vertical integrative analysis. Horizontal meta-analysis is widely used meta-analysis technique for varied biological purposes. There are 3 main meta-analysis categories of transcriptomic applications. Most widely used one is two-stage method. Two-stage method firstly computes single summary statistics and then apply meta-analysis method to summary statistics. These two methods include combination of p-values, effect size or rank statistics. Second category of

method merge raw data and normalizes it, and then applies standard single-study analysis. Last method integrates the differential expressions from all studies. With the incredible amount of publicly available omics data, vertical integrative analysis method widely used. This type of method used for different levels of omics data applications. In bioinformatics field, vertical integrative analysis mostly used for clustering and prediction analysis (90,91).

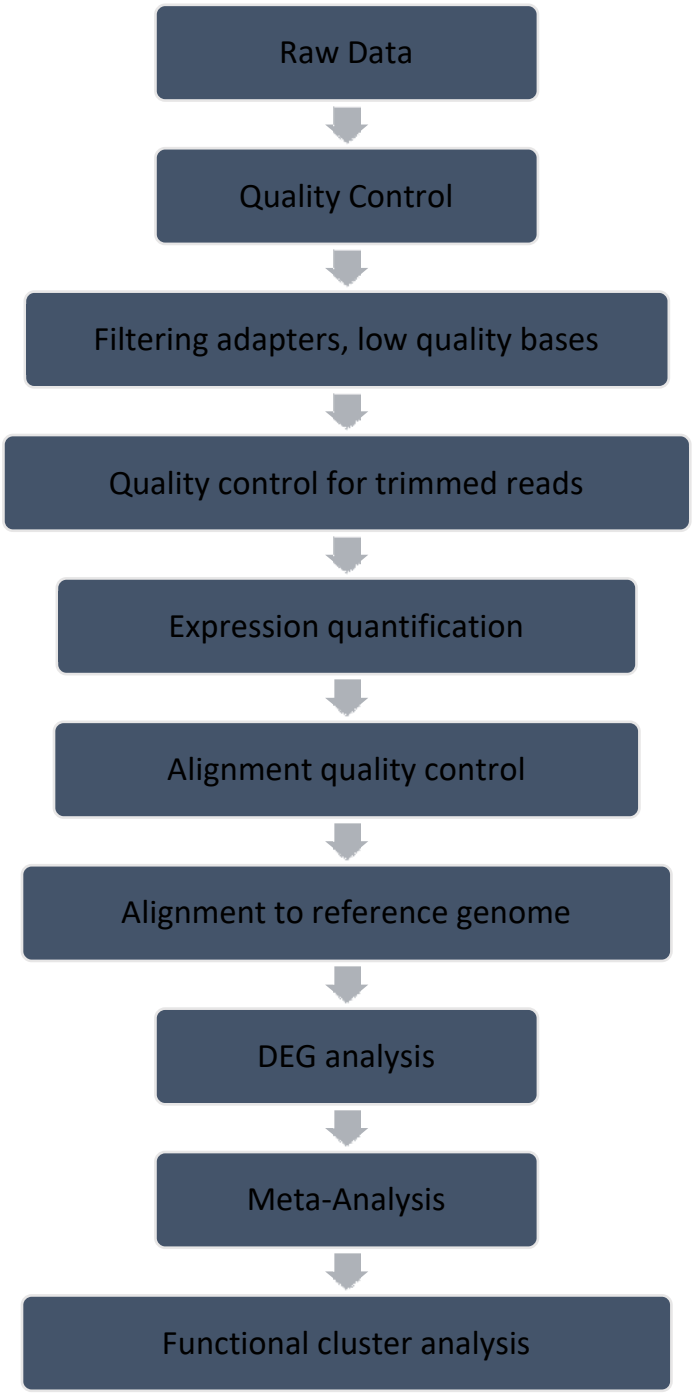

Figure 1.4 Meta analysis pipeline by mapping (42)

2. AIM AND SCOPE OF THE STUDY

Stress factors can reduce plants crop yield and product quality. These factors can be classified as biotic and abiotic factors. Biotic stress factors can be caused by bacteria, viruses, nematodes, fungi, and herbivores whereas abiotic factors are salinity, drought, heat, etc. Organisms have their unique responses to environmental conditions. To improve our cultivation systems, we should understand the stress-related responses. With the help of bioinformatics, the stress response can be understood at molecular levels. RNA-Seq technology is a good way to study the molecular response of stress conditions.

In early studies of *Beta vulgaris*, transcriptome studies are used to find stress-related genes. In this study, we aim to find the stress response genes that are differentially expressed under studied stress conditions by using the Meta analysis method.

With the premise knowledge, we hypothesized that there are common stress responsive genes that are differentially expressed independent from the stress type.

3. MATERIALS AND METHODS

11 stress-related studies data of *Beta vulgaris* selected from NCBI database on March 30th 2021, with the following filters. Source choosen as "RNA", and Platform choosen as "Illumina". Downloaded data shown in Table 3.1. Related data downloaded by SRA Toolkit version 2.11.0 (71) prefetch tool. The downloaded compressed data was converted to "fastqc" format by SRA Toolkit fasterq-dump tool version 2.11.0.

Table 3.1 Downloaded data from NCBI

Condition	Stress	SRA Acc. #	BioProject #	GEO #
Biotic	Fungal Pathogen	SRP223381	PRJNA574280	GSM4097161
Biotic	Fungal Pathogen	SRP095578	PRJNA358634	GSM2438805
Biotic	Nematod Pathogen	SRP217806	PRJNA559279	GSM4017124
Abiotic	Alkaline Stress	SRP262530	PRJNA634158	N/A
Abiotic	Alkaline Stress	SRP140444	PRJNA450324	GSM3098240
Abiotic	Alkaline Stress	SRP126020	PRJNA420895	GSM2872808
Abiotic	Salinity Stress	SRP149098	PRJNA473360	GSM3161753
Abiotic	Salinity Stress	SRP145448	PRJNA453103	N/A
Abiotic	ABA Treatment	SRP235645	PRJNA594791	N/A
Abiotic	Heat Stress	SRP044105	PRJNA254489	N/A
Abiotic	Stress Germination	SRP219737	PRJNA450098	N/A

fastqc files are checked for quality by FastQC version 0.11.9 (66) then an index file created by Hisat2 version 2.1.0 (72). To create an index file, reference genome downloaded from Ensembl plants database (RefBeet-1.2.2). Based on to index file, fastqc files are aligned to the indexed genome. Output file collected as SAM format. The data eliminated which had less than 60% of alignment.

Hisat2 codes;

For indexing;

hisat2-build beta_vulgaris_ref_genome.fastq index

For alignment;

hisat2 -x index -1 SRR"$i"_1.fastq -2 SRR"$i"_2.fastq -S SRR"$i".sam

(flag's descriptions shown in Table 3.2)

Table 3.2 Hisat2 flag explanations

Flag	Description
-x	Indicate the index file
-1	Indicate the first file of paired-end data
-2	Indicate the second file of paired-end data
-U	Used when data is single end
-S	Output file as Sam

SAM format data converted to BAM format by using samtools version 1.10 (73). Then the BAM files are sorted according to their names.

samtools codes;

samtools view -S -b SRR"$i".sam -o SRR"$i".bam

samtools sort

samtools sort SRR"$i".bam -o SRR"$i".sorted.bam

(flag's descriptions shown in Table 3.3)

Table 3.3 samtools flags explanations

Flag	Description
-S	Input file format is auto-detected
-b	Output file format is BAM
-o	Output file name

FeatureCounts version 2.0.0 (74) was used for number counts and sequences associated with their Ensembl gene IDs. The required GTF file was downloaded from Ensembl plants (75). The results are collected for differential gene expression analysis.

featureCounts code;

*featureCounts -a B_vulgaris-1.2.2.51.gtf -g gene_id -o feature_results.txt *.bam -T 4 -F GTF*

(flag's descriptions shown in Table 3.4)

Table 3.4 featureCounts flags explanations

Flag	Description
-a	Name of annotation file
-g	Specify attribute type in GTF file
-o	Name of output file
-T	Number of CPU threads
-F	Specify format of the provided annotation file

For differential expressed gene analysis R version 4.0.5 (67) package DESeq2 version 1.30.1 (58) was used. The same pipeline was used for all data. Studies are compared to their control groups and the P-adjust value is set to 0.05. First raw counts were normalized by negative binomial model. To work with better efficient, Log2 transformation technique was used in R.

Principle component analysis was made by the plotPCA function in DESeq2. 2D plots was created for each data. *vst* function has been used for normalization and log transformation of data. PC1 and PC2 created by plotPCA function.

For data visualization tidyverse version 1.1.1 (68) ggplot2 version 3.3.5. Ensembl IDs are converted to GO and KEGG ENZYME ID through R package biomaRt version 2.46.3 (69). For ID conversation, Ensembl plants database was used. *Beta vulgaris* data has been selected to convertion of Ensembl ID's to GO, KEGG ENZYME ID's.

Differentially expressed genes data was collected for meta analysis. For meta-analysis, R package metaRNASeq version 1.0.7 (76) was used. The results of meta-analysis again converted to GO and KEGG ENZYME ID. For Ensembl ID to ENTREZ ID conversion, the NCBI database was used through the E-utilities tool. ENTREZ ID's then converted to KEGG ID by using KEGG Mapper. The KEGG ID's then used for pathway analysis. Pathways of *Beta vulgaris* were created by using R package pathview version 1.30.1 (70). pathview package is using the KEGG database.

4. RESULTS AND DISCUSSION

Eleven stress-related data study of *Beta vulgaris* downloaded from NCBI database. The data collections are aligned to the reference genome of *Beta vulgaris* and they are eliminated according to their alignment rate. The rates below %60, are eliminated so only four studies are remained, SRP095578, SRP217806, SRP149098, and SRP235645. The details of data shown in Table 4.1. Study SRP095578 and SRP217806 are fungal and nematodes stress respectfully whereas study SRP149098 and SRP235645 are salinity and ABA treatment stresses.

Table 4.1 Collected data details and Hisat2 Results

Condition	Stress	SRA Acc. #	Run #	Hisat Result (%)
Biotic	Fungal Pathogen	SRP223381	24	%12
Biotic	Fungal Pathogen	SRP095578	36	%67
Biotic	Nematode Pathogen	SRP217806	16	%65
Abiotic	Alkaline Stress	SRP262530	31	%57
Abiotic	Alkaline Stress	SRP140444	3	%29
Abiotic	Alkaline Stress	SRP126020	9	%17
Abiotic	Salinity Stress	SRP149098	12	%64
Abiotic	Salinity Stress	SRP145448	60	%16
Abiotic	ABA Treatment	SRP235645	24	%61
Abiotic	Heat Stress	SRP044105	5	%49
Abiotic	Stress Germination	SRP219737	24	%56

Genomic feature assignment and read counting process held by featureCounts. Study SRP095578 has 36 runs and 35 runs' successfully assigned alignment rate is above 80% where only one run had a 79.2% rate. Study SRP217806 has 16 runs and all the runs had more than 64% assigned alignment rate. Study 149098 has 12 runs and all of them had more than 80% assignment rate. Lastly, study SRP235645 has 24 runs, and their assignment rate is more than %79.

For differentially expressed gene analysis, R package DESeq2 was used. Every study was analyzed individually. Only 4 studies passed the determined criteria thresholds and 2 of them are biotic and 2 of them are abiotic stress studies.

4.1 DEG Analysis of Genes Under Biotic Stress

This study (Bioproject Acc. Number: PRJNA358634, SRA Acc Number: SRP095578) was conducted by (77). In this study, the effect of Rhizoctonia solani has been observed on roots. 3 conditions have been observed. Conditions are 0 days

post-infection (dpi), 2 dpi and 5 dpi. The distribution (Figure 4.1) and Venn diagram (Figure 4.2) of data is shown respectively.

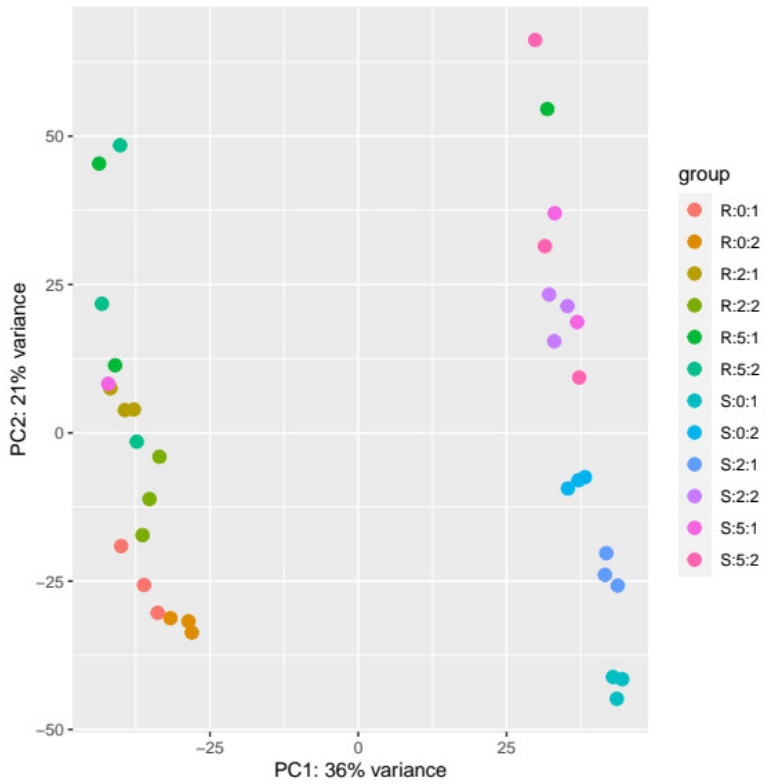

Figure 4.1 Distribution of data. R means resistant, S means susceptible. The first number indicates the dpi and the second one indicated the genotype.

In 0 dpi condition, there are 3678 DEG detected and 1787 of them are down-regulated whereas 1891 of them are upregulated. In 2 dpi conditions, 3669 DEG detected were 1773 down-regulated and 1896 up-regulated. In 5 dpi, 3347 DEGs were found, 1609 of them down-regulated and 1738 of them up-regulated. Among these 3 conditions, 758 genes are common. Between 0 dpi and 2 dpi, there are 1360 common genes, between 0 dpi and 5 dpi there are 256 common genes, between 2 dpi and 5 dpi there are 392 common genes observed.

Second biotic stress factor data (Bioproject Acc. Number: PRJNA559279, SRA Acc Number: SRP217806) is conducted by (78). In this study, a nematode species *Heterodera schachtii* was used as a stress agent. The stress conditions applied to the sugar beet's root when the beet was 15 days old. A resistant and a susceptible genotype was used. The distribution (Figure 4.3) and Venn diagram (Figure 4.4) of data is shown respectively.

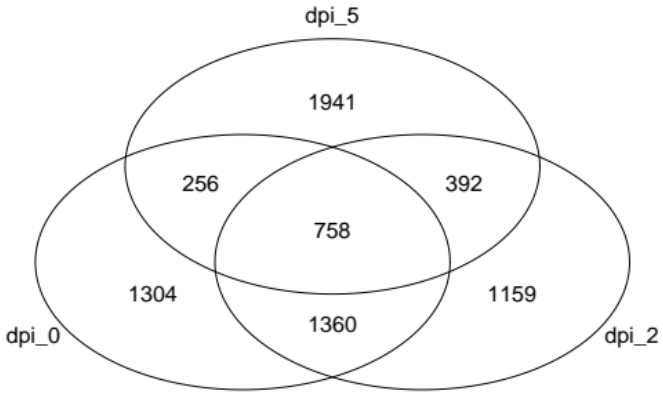

Figure 4.2 Venn Diagram of Study 095578

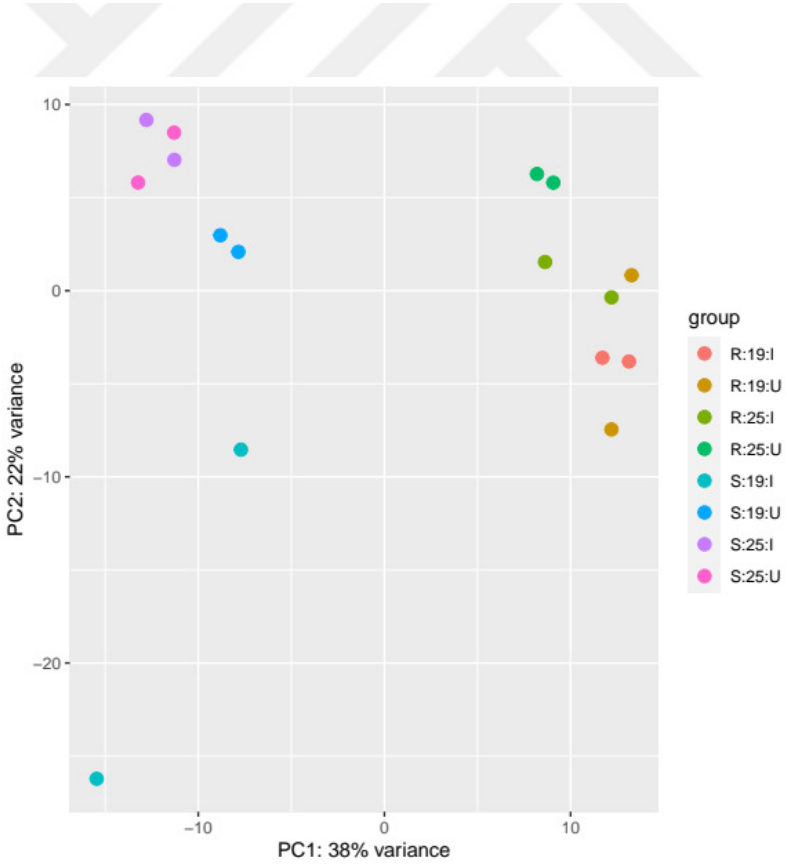

Figure 4.3 Distribution of Study 217806. R means resistant, S means susceptible, the numbers indicate day after germination of the *Beta vulgaris*, I means infected and U means uninfected.

In susceptible 19 conditions, 290 DEG was found. 122 of them are down-regulated and 148 of them are up-regulated. In the susceptible 25 conditions, there was just 61 DEG found. 37 of them are down-regulated and 24 of them are up-regulated. These conditions only 18 common genes contain.

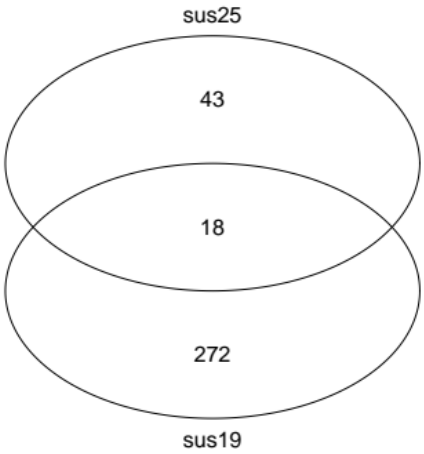

Figure 4.4 Venn diagram of Study 217806

4.2 DEG Analysis of Genes Under Abiotic Stress

One of the abiotic studies (Bioproject Acc. Number: PRJNA473360, SRA Acc. Number: SRP149098) is conducted by (79). It is a salinity stress study. The conditions are the first day and seventh day of stress. The distribution (Figure 4.5) and Venn diagram (Figure 4.6) of data is shown respectively.

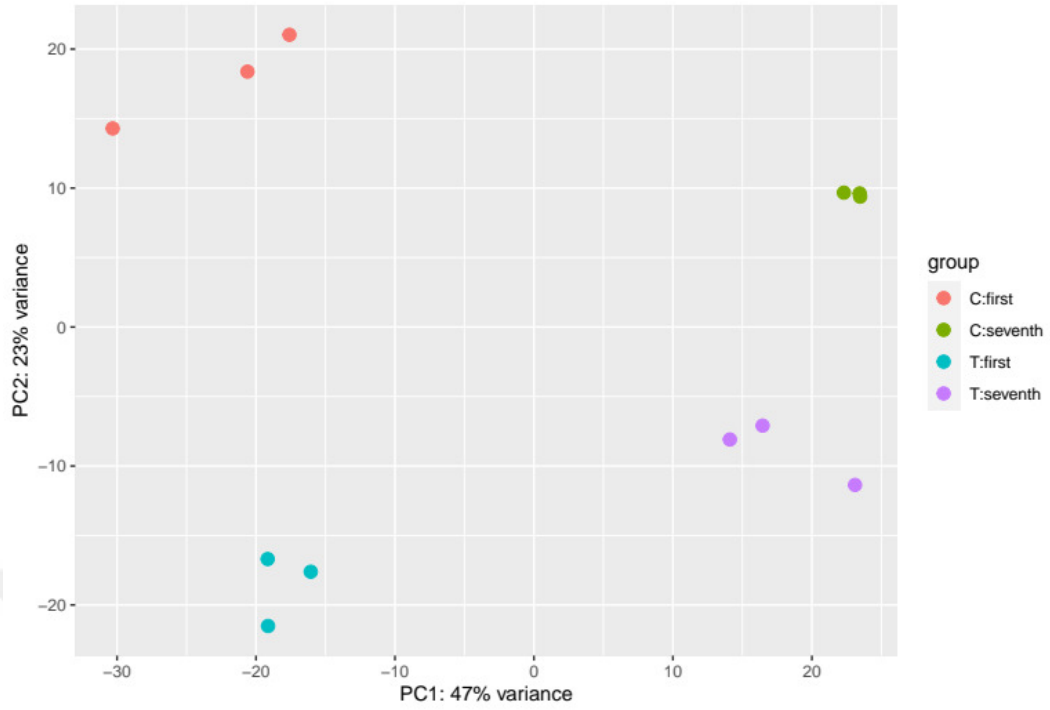

Figure 4.5 Distribution of Study 149098. C means to control, T means treatment. The first day and seventh day indicated.

In the first day condition, 6219 DEG is found which includes 2999 down-regulated and 3220 up-regulated genes. On the other hand, the seventh-day condition contains 3930 DEG where 1871 down-regulated 2059 up-regulated genes are present. There are 1688 common genes among these conditions.

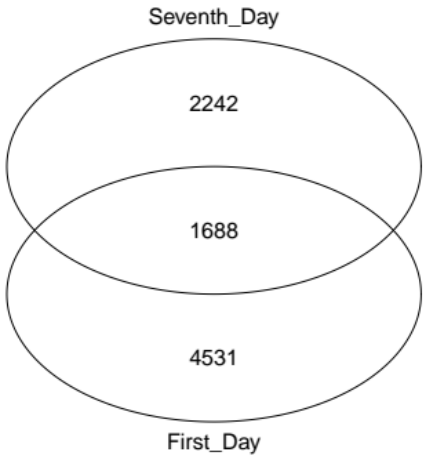

Figure 4.6 Venn Diagram of Study 149098

This study (Bioproject Acc. Number: PRJNA594791, SRA Acc Number: SRP235645) was conducted by (80). It is an ABA treatment experiment where the conditions are after 1,6,12 and 24 hours of ABA treatment. The distribution (Figure 4.7) and Venn diagram (Figure 4.8) of data is shown respectively.

In the result of hour six, there was 8706 DEG. Among them, 4458 down-regulated genes and 4248 up-regulated genes were found. The result of hour 12 contains 8894 DEG. 4546 of them are down-regulated and 4348 of them are up-regulated. Hour 24 contains 10867 DEG which are 5435 down-regulated and 5432 up-regulated genes. Among these 3, 4319 genes are common in all conditions.

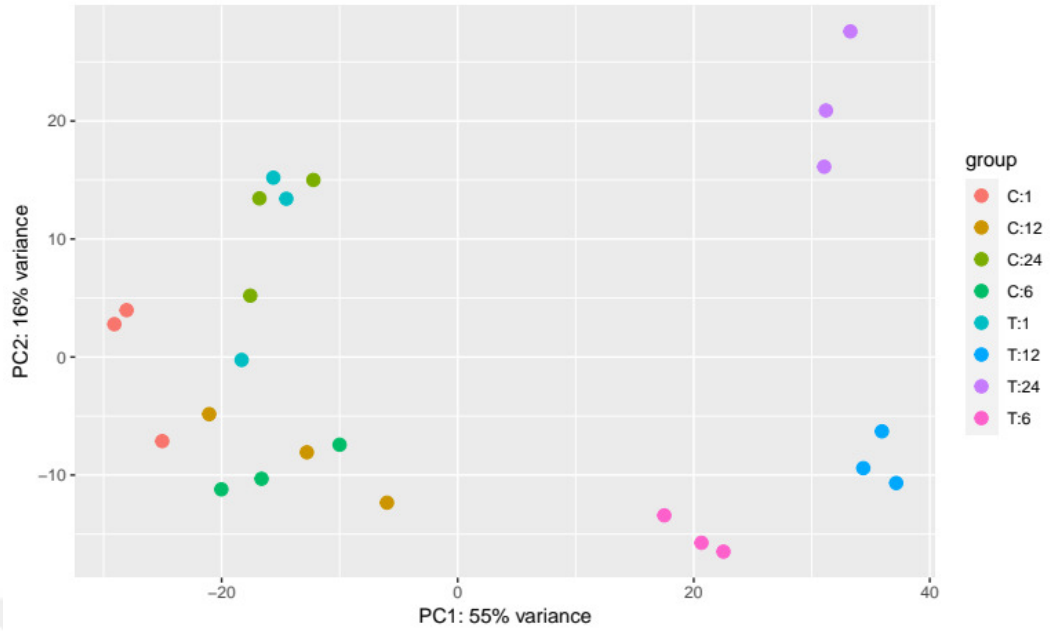

Figure 4.7 Distribution of Study 235645. T indicates treatment and C indicates control. Numbers are indicating hours.

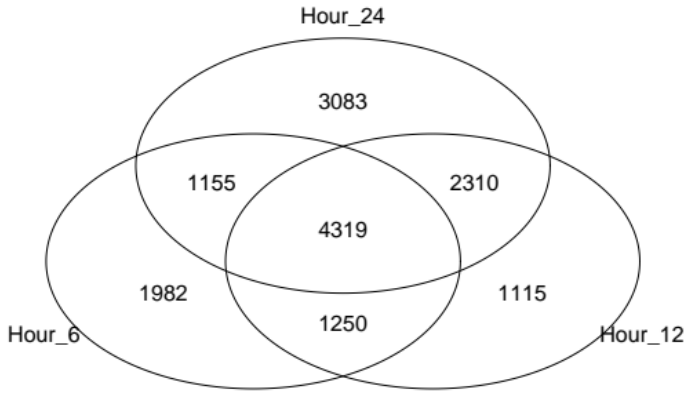

Figure 4.8 Venn diagram of study 235645.

4.3 Meta-Analysis

Meta analysis applied to the two biotic and two abiotic stress seperately to observe the differentially expressed genes of both stress conditions. Then analysis has been done to the all four stress conditions.

4.3.1 Meta-Analysis of Transcriptomes for Abiotic Stress Response

Abiotic stress data showed 1031 common stress related DEG. The venn diagram of abiotic stress meta analysis showed in Figure 4.9. GO term analysis of abiotic stress common genes are showed in Figure 4.10. In the abiotic stress data

for 161 genes, there was no GO term. For the rest of the genes there is 736 biological process, 415 cellular component and 1020 molecular function related GO terms found. The Figure 4.10 and 4.12 shows the 10 data with the most matches from the 3 major GO categories.

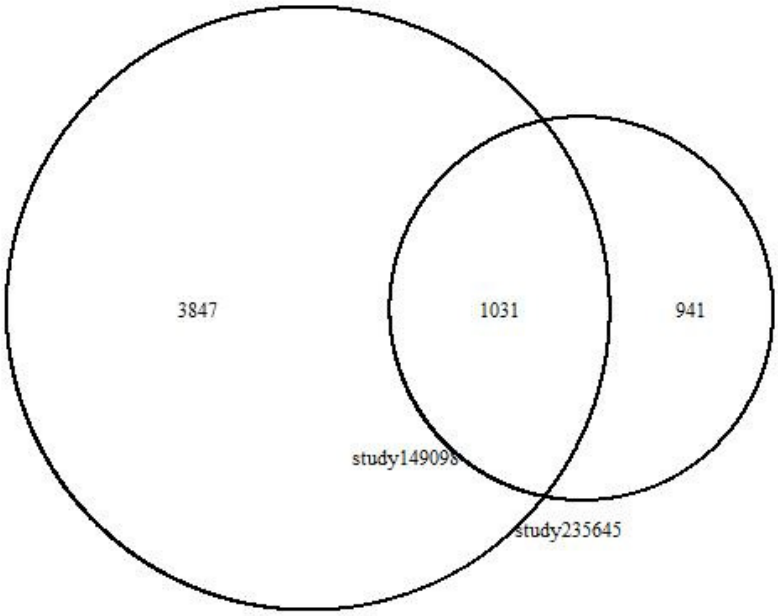

Figure 4.9 Venn diagram of Abiotic Stress Data

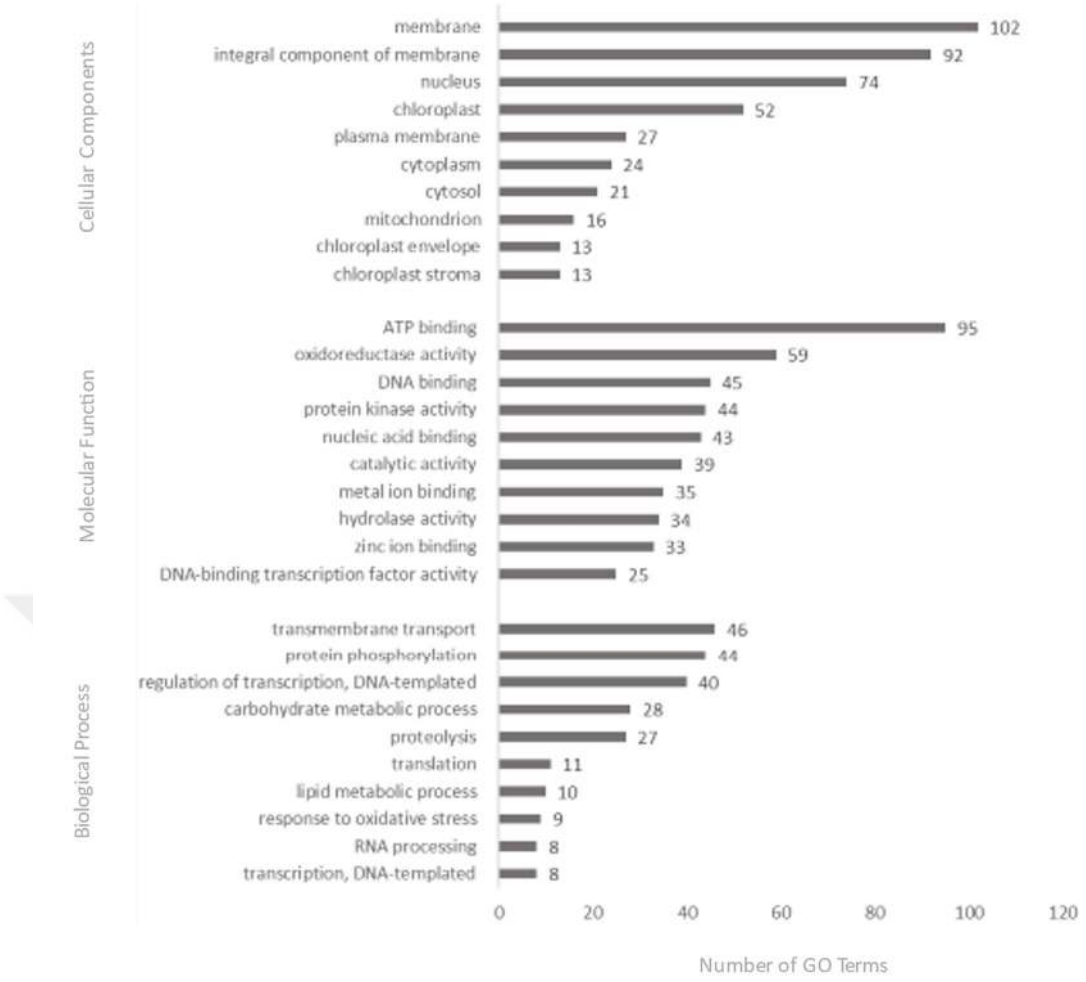

Figure 4.10 GO term distribution of Abiotic Stress Data

The differences in the GO term distribution are small between the biotic and abiotic stress data, but some key terms must be discussed. Cellular components terms of abiotic stress found abundant in mostly cell membrane and chloroplast related terms. Membrane-related terms must be connected with the balancing of the cell's osmotic pressure. As the salt treatment affects the osmotic pressure of cells, the transmembrane transport also seems to be abundant in abiotic stress data (92). Sugar beet is also known to increase the uptake of phosphorus during salt stress (93). This phosphorus uptake may lead to phosphorus poisoning, growth retardation, and necrosis.

Chloroplast-related terms are also abundant in abiotic stress data according to biotic stress data. Chlorophyll is the pigment for absorbing light energy and transforming it into chemical energy. Under salt stress, chlorophyll content is known to be decreased (94).

4.3.2 Meta Analysis of Transcriptomes for Biotic Stress Response

Biotic data showed 460 common stress related DEG. In the biotic stress data for 76 genes there was no GO term. 268 biological process, 146 cellular component and 350 molecular function related GO term was found. The venn diagram showed in Figure 4.11. GO term analysis of common genes are showed in Figure 4.12.

Figure 4.11 Venn diagram of Biotic Stress Data

Biotic stress data contains the term extracellular region where abiotic stress data does not. Extracellular region term is known to be related with the space between the cells and they express under the infection of a parasite (95).

In the biological process terms, under biotic stress data, defense response is the second most abundant term where abiotic stress data shows no sign of it. We believe that this one term is a key to an organism's response to biotic stresses. This defense response also affects molecular functions. DNA and RNA binding terms are also abundant in abiotic stress data. The molecular function terms of biotic stress data can be seen as mostly to transcript and protein synthesis related. This might be because of the biotic stress response (96).

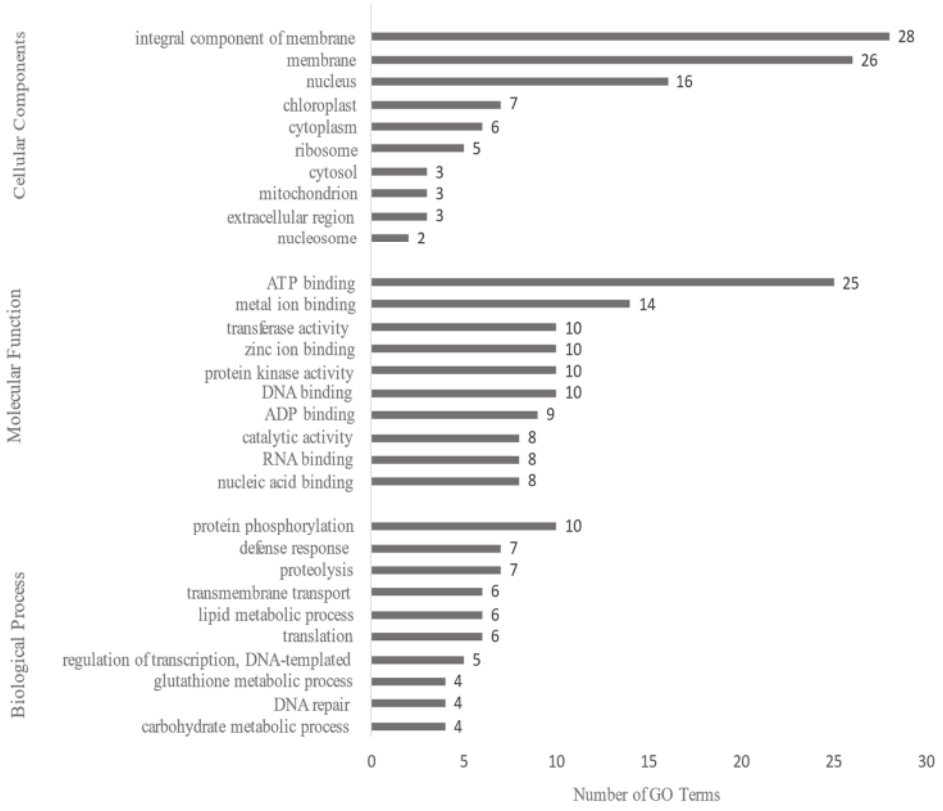

Figure 4.12 GO term distribution of Biotic Stress Data

4.3.3 Meta-Analysis of Combined Stress Conditions

The results of DEGs of the 4 different studies have been used for meta-analysis. Output files of DEG analysis were used as input files of meta-analysis. The analysis was constructed by R package metaRNASeq version 1.0.7. The results showed 10 genes that were differentially expressed in every stress condition (Figure 4.13). Six of them are up-regulated and 4 of them are down-regulated (Table 4.2).

Table 4.2 Meta-Analysis Results

	Ensembl_IDs	Log Fold Change Study 095578	Log Fold Change Study 217806	Log Fold Change Study 149098	Log Fold Change Study 235645	Regulation
1	BVRB_6g145450	-0.816	-1.498	-1.669	-0.449	Down
2	BVRB_6g133980	3.352	4.028	2.640	7.936	Up
3	BVRB_7g166460	4.651	0.967	1.054	0.919	Up
4	BVRB_2g030930	1.626	0.821	1.090	1.838	Up
5	BVRB_2g043720	0.337	0.253	0.501	0.252	Up
6	BVRB_2g030970	-1.219	-1.523	-1.160	-1.213	Down
7	BVRB_8g187810	-0.350	-0.805	-0.839	-0.439	Down
8	BVRB_4g086310	6.767	1.781	0.836	1.289	Up
9	BVRB_5g126270	-0.596	-0.788	-0.999	-0.495	Down
10	BVRB_1g021740	2.284	2.260	1.832	4.898	Up

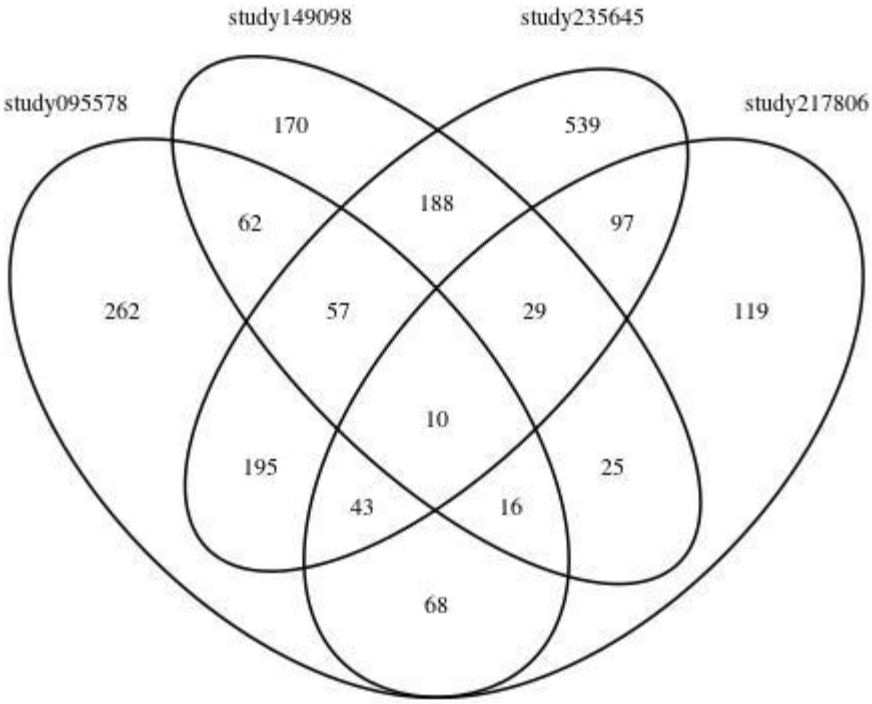

Figure 4.13 Venn diagram of Meta Analysis

The GO term of results has been compared with the biotic and abiotic stress data in Figure 4.14, 4.15, 4.16. Graphs shows the terms that common in meta, biotic and abiotic stress data.

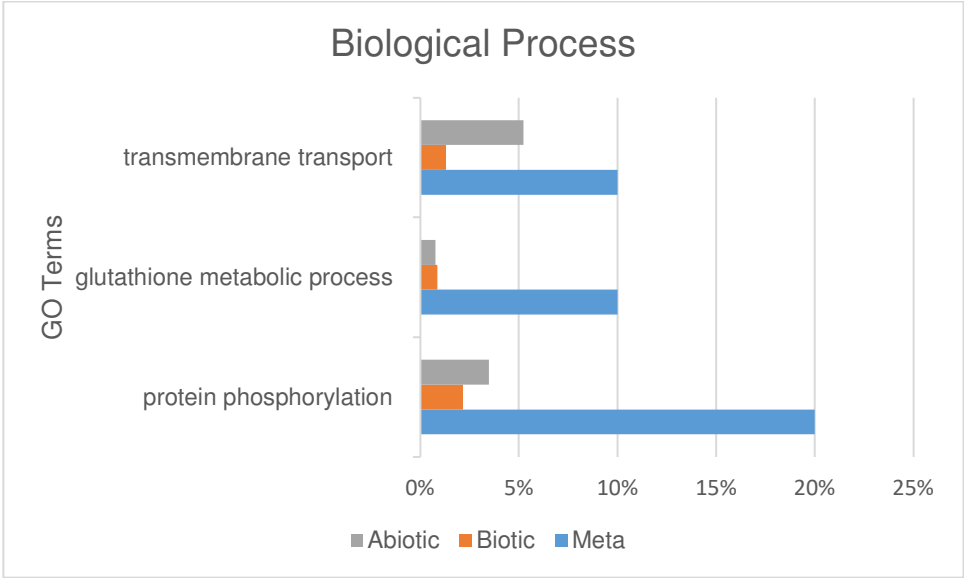

Figure 4.14 Biological Process GO term comparison. Percentages are found as term to DEG number ratio.

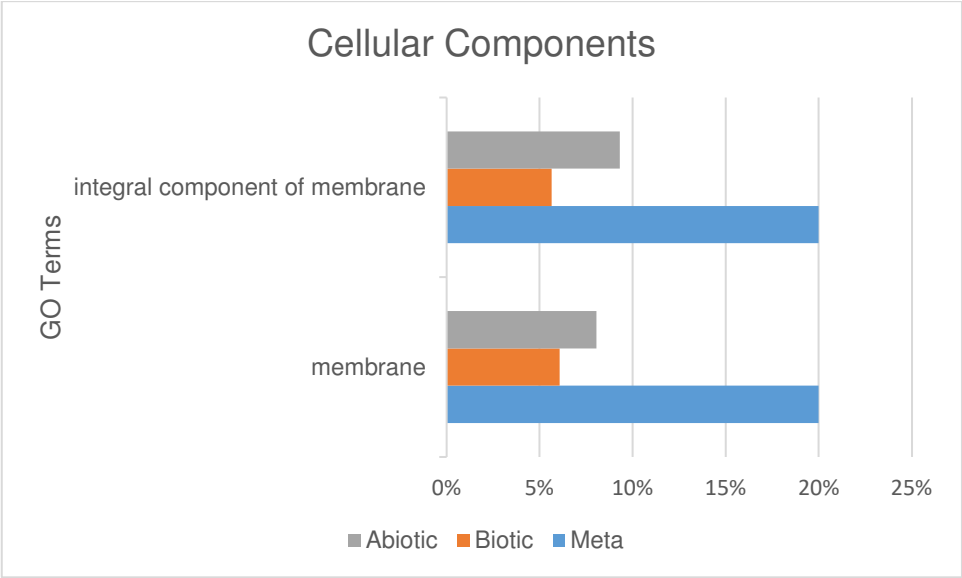

Figure 4.15 Cellular Components GO term comparison. Percentages are found as term to DEG number ratio.

Table 4.3 DEG and GO Term numbers

	GO Terms	Meta-Analysis	Biotic	Abiotic
DEG	Differentially Expressed Genes numbers	10	460	1031
Biological	transmembrane transport	1	6	54
	glutathione metabolic process	1	4	8
	protein phosphorylation	2	10	36
Cellular	integral component of membrane	2	26	96
	membrane	2	28	83
Molecular	xenobiotic transmembrane transporter activity	1	1	6
	heme binding	1	6	27
	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	1	2	21
	oxidoreductase activity	1	8	34
	antiporter activity	1	1	6
	zinc ion binding	1	10	20
	iron ion binding	1	2	27
	monooxygenase activity	1	2	21
	ATP binding	2	25	64
	protein kinase activity	2	10	36

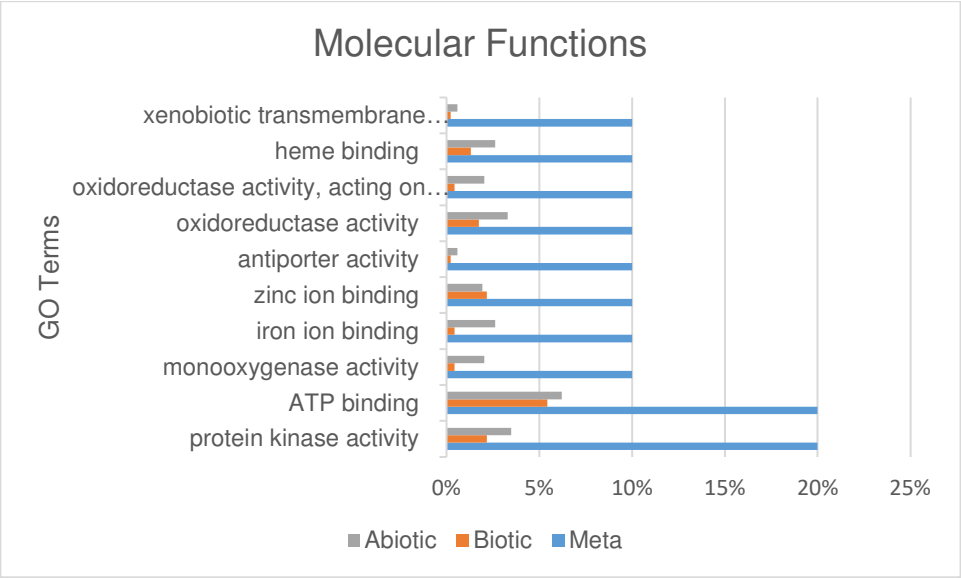

Figure 4.16 Molecular Functions GO term comparison. Percentages are found as term to DEG number ratio.

The GO term figures (Figure 4.14, Figure 4.15, and Figure 4.16) are represented as the term percentage of the total DEG numbers. In biological process GO terms, abiotic data has 54 transmembrane transport related terms where biotic data has 6. This also affect the percentage representation where abiotic stress data has 1031, and biotic has 460 DEG. All related numbers are shown in Table 4.3. Transmembrane transport terms are known to related with transportation of solutes across cells lipid bilayers. We assume that the higher amount of transmembrane transport terms related with the arrangement of cells osmotic pressure under the abiotic stresses.

In molecular functions GO terms, iron ion binding, monooxygenase activity and oxidoreductase activity. Iron is an essential element for plants. They play important roles such as respiration, chlorophyll synthesis and photosynthesis (98). They are also used as signaling molecule. The reason of abiotic stress to have higher iron ion binding term might be explained with the study of Skorupa et. al., 2016. This study revealed the up-regulated photosynthesis genes under the salt stress (97). Since the iron is important for photosynthesis, more iron related terms are understandable in abiotic stress. On the other hand, monooxygenase related GO terms are also higher in abiotic stress data. It is known that flavin and choline monooxygenase genes are increasing the drought resistance in *Arabidopsis* and increase the tolerance to abiotic stress of rice and spinach (99,100). Our results are compatible with those studies since they are higher under abiotic stress.

The Ensembl_IDs of genes are searched against to NCBI database and 6 of them are converted to Entrez gene IDs. Then those IDs are converted to KEGG gene IDs (Table 4.3). The rest of the genes stayed unknown in the KEGG database.

Table 4.4 Meta Analysis Results. N/A means not applicable

	Ensembl_ID	KEGG_ID	Gene Name
1	BVRB_1g021740	104905737	cytochrome P450 81E8-like
2	BVRB_2g030930	104882957	uncharacterized
3	BVRB_2g043720	104887745	thioredoxin domain-containing protein 2
4	BVRB_6g133980	104895532	probable mannitol dehydrogenase
5	BVRB_6g145450	104896981	probable leucine-rich repeat receptor-like protein kinase At5g49770
6	BVRB_7g166460	104899342	glutathione transferase GST 23-like
7	BVRB_4g086310	N/A	N/A
8	BVRB_5g126270	N/A	N/A
9	BVRB_8g187810	N/A	N/A
10	BVRB_2g030970	N/A	N/A

During stress conditions, ROS dramatically increases. The increase of ROS is associated with the plant's defense mechanism (11). Since ROS is harmful to cells, plants should avoid these damages through retoxification of ROS. Nevertheless, plants are adapted to use ROS as signaling molecules as a response to a variety of stress conditions. The gene products that we found are mostly related to ROS. Since ROS are used as a signal molecule in plant defense systems, our results seem reasonable.

Cytochrome P450 is one of the largest enzymatic families in bacteria, fungi, and especially plants. They are known to be a key player in multiple metabolic pathways and reactions. As a result of those reactions, various secondary metabolites are synthesized that plays important role in growth and development or protection to biotic and abiotic stress factors (82). Our study showed the related gene is upregulated under both stress conditions.

Mitochondrial thioredoxin is required for metabolic pathways to function properly. These pathways include stomatal opening, antioxidant metabolism, drought, and salinity stresses. Also detoxifies hydrogen peroxide which is a ROS. These facts are proved in arabidopsis (13). Another known ROS quencher is mannitol. Mannitols are produced by plants and fungi. Mannitol dehydrogenase is one catalyst in the catabolism of mannitol (84). Our results show the upregulation of these 2 genes under both bioitic and abiotic stresses. The results show the importance of both ROS and its detoxification under stress conditions.

Leucine-rich repeat receptor kinases are the largest subfamily of transmembrane receptor-like kinases in plants. They regulate a wide range of defense-related processes such as stem cell controlling, hormone perception, host-specific, and no-host-specific defense mechanisms, response to wounding, and symbiosis (85). Early studies shows the importance of transmembrane receptor-like kinases under both biotic and abiotic stresses but surprisingly our study showed a down-regulation of this gene. To understand this phenomenon, further study of this gene and its inducers is needed.

Glutathione S-transferases (GST) is a member of the super-family of multifunctional proteins. They are one of the most important families of detoxifying enzymes. They play important roles in plant development and both biotic, abiotic stresses. They are involved in, herbicide detoxification, plant protection against ozone damages, heavy metals, xenobiotics, and microbial infections. In the Figure 4.10, glutathione transferase-related gene's place is shown in the pathway. The different concentration of red color indicates the log-fold change of study 095578, 217806, 149098, and 235645 respectfully (86).

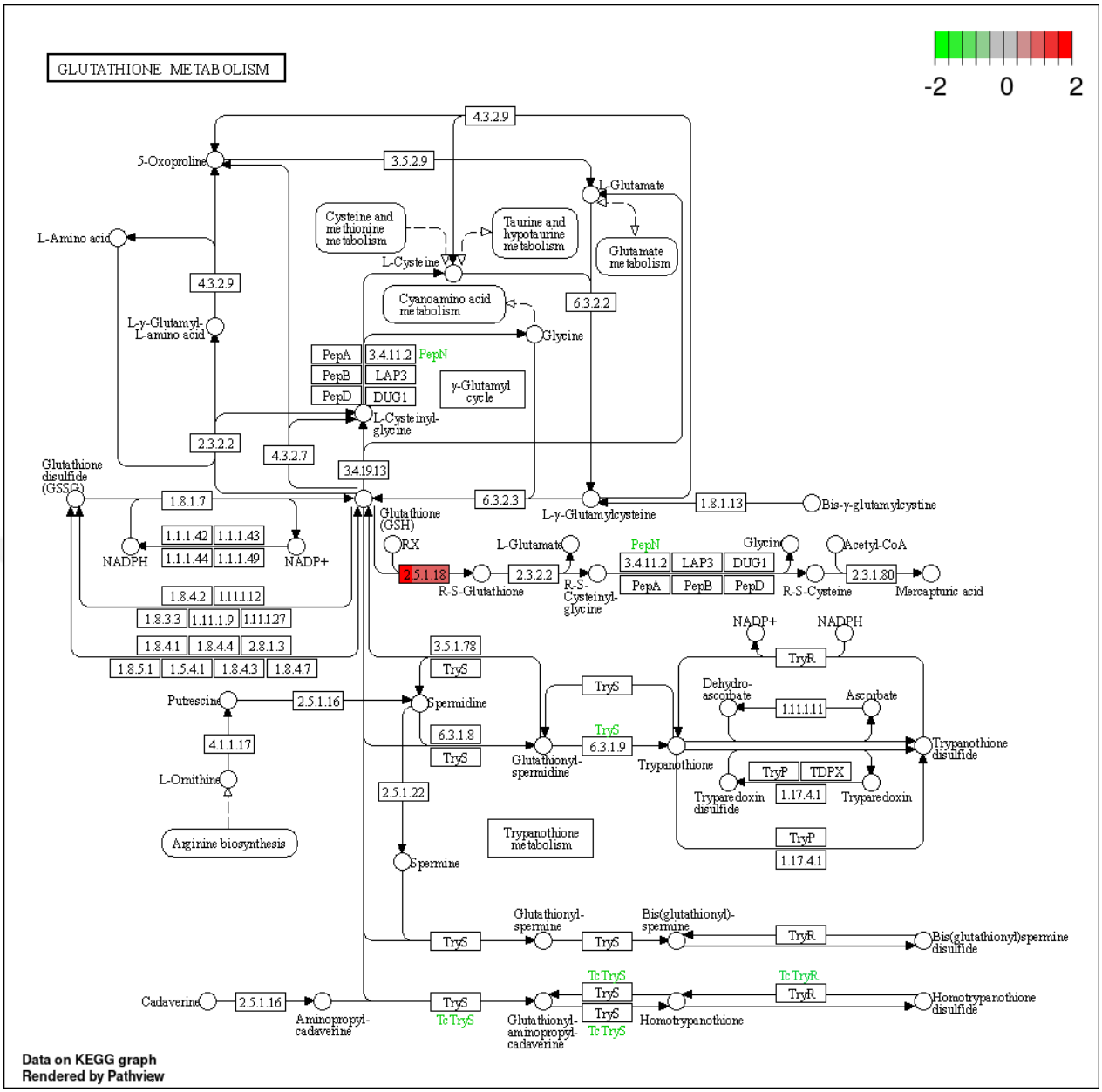

Figure 4.17 Glutathione Metabolism pathway

Our results are compatible with the literature for known genes. The rest of the genes are blasted against the NCBI protein database. The blast results are examined, hypothetical and uncharacterized proteins are avoided. The genes are selected which had the highest query coverage and percent identity (Table 4.4).

Table 4.5 NCBI Blast results of 4 uncharacterized genes

Ensembl ID	Protein	Species	Query Cover	Percentage Identity	Accession #
BVRB_2g030970	Interleukin-2 receptor subunit beta like	*Actinidia chinensis* var. *chinensis*	85%	40.46%	PSS08165.1
BVRB_4g086310	6-Phosphofructokinase 3	*Triticum urartu*	32%	78.26%	EMS55719.1
BVRB_5g126270	DETOXIFICATION 16-like isoform X2	*Spinacia oleracea*	47%	78.48%	XP_021857959.1
BVRB_8g187810	Pheophytinase	*Spatholobus suberectus*	88%	60.56%	TKY51975.1

Interleukin-2 receptor subunit beta like protein is associated human T cell-mediated immune response. Alfa, beta and omega subunits of interleukin-2 receptors are binding to interleukin 2 protein which starts a signal transduction and also plays role as T cell growth factor, proliferation, differentiation and survival (106). Our results show the down-regulation of BVRB_2g030970 gene. Blast results shows the highest similarity with the interleukin-2 receptor subunit beta like protein of *Actinidia chinensis var. chinensis.* Multiple sequence alignment (MSA) results in Figure 4.18 also shows some conserved domains of the protein among human, *Arabidopsis* and rice. Nevertheless, we strongly believe that our gene is not associated with human interleukin-2 receptor.

```
Homo_sapiens                      IPWLGHLLVGLSGAFGFIILVYLLINCRNTGPWLKKVLKCNTPDPSKFFSQLSSEHGGDV    300
Oryza_sativa_Japonica_Group       ----------------------------MAVERGERRGGGIGCL-SCCFGGGDGDGEGEE     31
Arabidopsis_thaliana              ----------------------------MTNLSLKKKLKSRFCI-AGCFRTTNHHHHDNI     36
Beta_vulgaris_subsp._vulgaris     ----------------------------HSSVSLKHKLKLILAS-CCNFSDSDEDDIKK-     37
Actinidia_chinensis_var._chinensis ----------------------------SISPSFKNKLKQTLCL-SCCFRNSRRETLGS-     35
                                                                :.        .  *    .   .

Homo_sapiens                      Q-KWLSSPFPSSSFSPGGLAPEISPLEVLERDKVTQLLLQQDKVPEPASLSSNHSLTSCF    359
Oryza_sativa_Japonica_Group       LGQRA------------ARALRTSSRWVR--DRAVELPEM-------VALAGGR------     64
Arabidopsis_thaliana              PDDLPSSPVTTEKSTQ-S--PHGGGMKNKSPRLTRTLSKSHEKCKSLIHRMGGG------     87
Beta_vulgaris_subsp._vulgaris     --NLEH-----------SDKSHTLSAWLK--SELPSIKVISAKI--FAGKLG--------     72
Actinidia_chinensis_var._chinensis --SSSSSTTSTSSDDR-PTLVRAPSMWLK--SRVHELPEIGDKCRHLINRIGRH------     84
                                    .                .             :             .

Homo_sapiens                      TNQGYFFFHLPDALEIEACQVYFTYDPYSEEDPDEGVAGAPTGSSPQPLQPLSGEDDAYC    419
Oryza_sativa_Japonica_Group       --RRKP---HLQHHQQQQLAGEFRYDPVSYALNFEED-----------------GDGEAQ    102
Arabidopsis_thaliana              --VGGV---GGHGKHIRRHTADFHYDPSSYALNFDKG---------------DEDDNID    126
Beta_vulgaris_subsp._vulgaris     -------------RNRRANSVEFGYDPLSYSLNFED------------------DDSHLD    101
Actinidia_chinensis_var._chinensis ----------HHHHRRHSSSADFRYDPLSYALNFD--------------------DSHVD    114
                                             . .     * *** *     :                       *.

Homo_sapiens                      TFPSRDDLLLFSPSLLGGPSPPSTAPGGSGAGEERMPPSLQERVPRDWDPQPLGPPTPGV    479
Oryza_sativa_Japonica_Group       PFK----YMAFSARLPASP--PPPPPPTAL------PVDRG-----S-------------    132
Arabidopsis_thaliana              RFP----LRNFSARLPHSP--PSSAKAADS------SYTVQ-----NL-LR---------    159
Beta_vulgaris_subsp._vulgaris     EYP----LRNFSSRLPPSP--VKEKNNVVV------STQIA-----GWS-----------    133
Actinidia_chinensis_var._chinensis EAP----LQNFSARVPASP--PPVDADGLA------KPAAI-----PREIACV-------    150
                                         **  :  .*
```

Figure 4.18 MSA results of Interleukin-2 receptor subunit beta like protein among different species. * means identical, : means conserved substitutions, . means semi-conserved substituons.

Among these results, phosphofructokinase enzymes are studied and identified in mulberry. They are play role in drought and bicarbonate stress response. Specifically they are responsible from glycolysis pathways which provides ATP, NADH and other precursor metabolites for to meet the energy demanding of cell (101). Our meta-analysis results showed this related gene to be upregulated which might interfere the energy metabolism in positive way. The MSA results are in Figure 4.19.

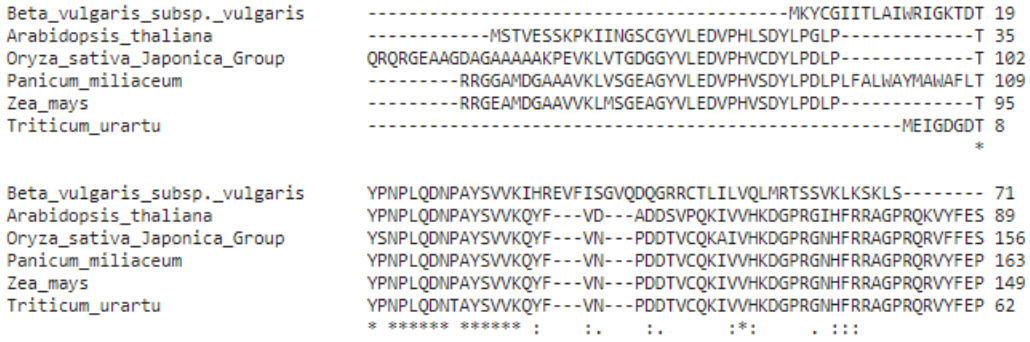

Figure 4.19 Phosphofructokinase MSA results. * means identical, : means conserved substitutions, . means semi-conserved substituons.

On the other hand pheophytinase is identified as chloroplast-located and senescence-induced enzyme which is common in algae and land plants. This enzyme degrades the chlorophyll pigment during senescence (102). Our pheophytinase similar gene might have the same duty. Our results show down-regulation of this gene and this supports the study of Skorupa et. al., 2016. Their study showed the up-regulation of photosynthesis genes under salt stress and our study suggest that *Beta vulgaris* might increase the photosynthesis under stress conditions. MSA results of phenophytinase i shown in Figure 4.20.

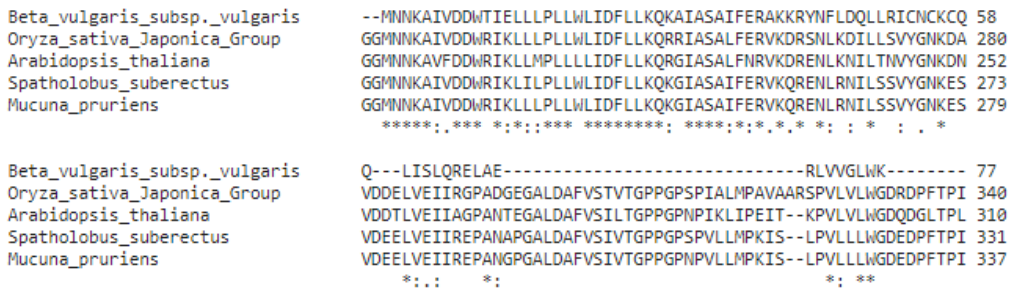

Figure 4.20 MSA results of protein phenophytinase. * means identical, : means conserved substitutions, . means semi-conserved substituons.

DETOXIFICATION 16-like isoform X2 enzyme belongs to the multidrug and toxic compound extrusion or multi-antimicrobial extrusion (MATE, also known as DETOXIFICATION proteins in *Arabidopsis*) family. These transporter genes are present in all kingdoms of life. It is very common and diversed in plant kingdom. This gene family are usually 400-550 amino acid long and responsible for transport of primary and secondary metobolites such as alkoloids, flavonoids, and hormones. Another duty of these gene products is detoxification of heavy metals, aluminum toxicity tolerance and disease resistance. This gene has not been identified in *B. vulgaris* yet but the MSA results in Figure 4.21 and diversification of the gene among plants gives the idea of another MATE family member protein (103).

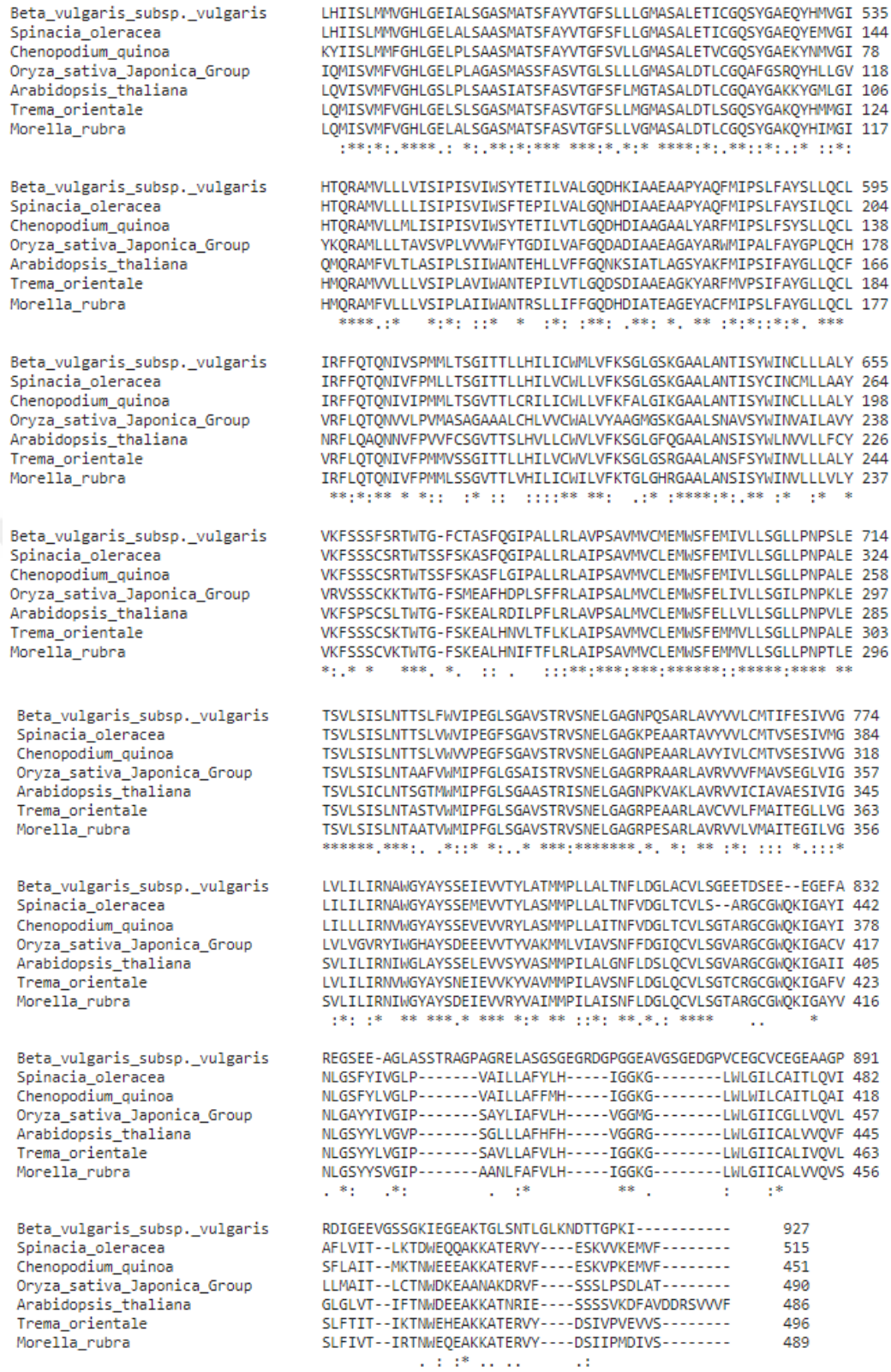

Figure 4.21 DETOXIFICATION 16-like isoform X2 MSA results. * means identical, : means conserved substitutions, . means semi-conserved substituons.

Overall results are showing the increasing of photosynthesis during both biotic and abiotic stress factors. Plants are usually decreases the photosynthesis rate under stress conditions, especially under the salt stress (104). Plants are closing the stomata under stress condition to stop evaporation, thus they become unable to intake CO_2. *B. vulgaris* on the other hand, known to highly salt tolerant crop by maintain the osmotic pressure and they take more water than other plants (105) . So we think that the evaporation is not importantly affecting the sugar beet under stress conditions. Down regulation of pheophytinase gene is supporting this idea. Another supporters of this result is thioredoxin and mannitol genes up regulation. Both are ROS detoxifiers where thioredoxin also responsible of stomatal opening. Another evidence of this idea may be the upregulation of phosphofructokinase gene. It is known to play role in glycolysis and ATP synthesize, because during the glycolysis glucose is used which is the product of photosynthesis. *B. vulgaris* not only seems to continue to photosynthesize, but also seems to increase it.

5. CONCLUSIONS AND RECOMMENDATIONS

In this study, we collected data from NCBI database that is related with stress conditions of *Beta vulgaris.* The data processed and aligned to the reference gonome and then genetic features assigned to the genes and counted. The counted genes analysed for differentially expressed genes and those results are used for Meta-analysis. These results contains the genes that differentially expressed under both biotic and abiotic conditions.

With this study, we revealed the genes that express differentially no matter which stress condition affects *Beta vulgaris*. These results can be usefull for sugarbeet production and crop yield.

Nevertheless, the unknown genes must be examined and characterized. The characterization of these genes needs further studies.

This study will contribute to future studies and help to understand stress response of both *Beta vulgaris* and other plants. Furthermore, this study can be used as a reference in the future RNA-Seq, Meta-analysis of transcriptome data studies.

6. REFERENCES

Vancouver citation system was used in this thesis.

1. Rozman Č, Kljajić M, Pažek K. Sugar Beet Production: A System Dynamics Model and Economic Analysis. Organizacija. 2015 Aug 1;48(3):145–54.

2. Gurel E, Gurel S, Lemaux PG. Biotechnology Applications for Sugar Beet. Crit Rev Plant Sci. 2008 Jun 25;27(2):108–40.

3. Lange W, Brandenburg WA, Bock TSMD. Taxonomy and cultonomy of beet (Beta vulgaris L.). Bot J Linn Soc. 1999 May;130(1):81–96.

4. Biancardi E, McGrath J, Panella L, Lewellen RT, Stevanato P. Sugar Beet. In 2010. p. 173–219.

5. Zicari S, Zhang R, Kaffka S. Sugar Beet. In: Integrated Processing Technologies for Food and Agricultural By-Products [Internet]. Elsevier; 2019 [cited 2021 Nov 24]. p. 331–51. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780128141380000137

6. Li H, Cao H, Cai Y-F, Wang J-H, Qu S-P, Huang X-Q. The complete chloroplast genome sequence of sugar beet (Beta *vulgaris* ssp. *vulgaris*). Mitochondrial DNA. 2014 Jun;25(3):209–11.

7. Şeker Pancarı, Ocak-2021, Tarım Ürünleri Piyasa Raporu Hazır.pdf [Internet]. [cited 2021 Nov 24]. Available from: https://arastirma.tarimorman.gov.tr/tepge/Belgeler/PDF%20Tar%C4%B1m%20%C3%9Cr%C3%BCnleri%20Piyasalar%C4%B1/2021-Ocak%20Tar%C4%B1m%20%C3%9Cr%C3%BCnleri%20Raporu/%C5%9Eeker%20Pancar%C4%B1,%20Ocak-2021,%20Tar%C4%B1m%20%C3%9Cr%C3%BCnleri%20Piyasa%20Raporu%20Haz%C4%B1r.pdf

8. Sugarbeet History – American Sugar Beet Association [Internet]. [cited 2021 Nov 24]. Available from: https://americansugarbeet.org/who-we-are/sugarbeet-history/

9. Sugarbeets [Internet]. [cited 2021 Nov 24]. Available from: https://hort.purdue.edu/newcrop/afcm/sugarbeet.html

10. Dohm JC, Minoche AE, Holtgräwe D, Capella-Gutiérrez S, Zakrzewski F, Tafer H, et al. The genome of the recently domesticated crop plant sugar beet (Beta vulgaris). Nature. 2014 Jan;505(7484):546–9.

11. Houghton J. Global warming. Rep Prog Phys. 2005 Jun 1;68(6):1343–403.

12. Jan MQ, Muazim K, Ashraf A. Climate change and its threats to sustainable agriculture. Reg Probl. 2018;21(3 (1)):37–40.

13. Wiss DA, Avena N, Rada P. Sugar Addiction: From Evolution to Revolution. Front Psychiatry. 2018 Nov 7;9:545.

14. Ku Y-S, Sintaha M, Cheung M-Y, Lam H-M. Plant Hormone Signaling Crosstalks between Biotic and Abiotic Stress Responses. Int J Mol Sci. 2018 Oct 17;19(10):3206.

15. Zharkov DO, editor. Mechanisms of Genome Protection and Repair [Internet]. Cham: Springer International Publishing; 2020 [cited 2021 Nov 24]. (Advances in Experimental Medicine and Biology; vol. 1241). Available from: http://link.springer.com/10.1007/978-3-030-41283-8

16. Zhang Y, Nan J, Yu B. OMICS Technologies and Applications in Sugar Beet. Front Plant Sci [Internet]. 2016 Jun 22 [cited 2021 Nov 24];7. Available from: http://journal.frontiersin.org/Article/10.3389/fpls.2016.00900/abstract

17. Hand drawn sugar beet stock vector. Illustration of agricultural - 149269293 [Internet]. [cited 2021 Nov 25]. Available from: https://www.dreamstime.com/colorful-hand-drawn-sugar-beet-isolated-white-background-vector-illustration-hand-drawn-sugar-beet-image149269293

18. Atkinson NJ, Lilley CJ, Urwin PE. Identification of Genes Involved in the Response of Arabidopsis to Simultaneous Biotic and Abiotic Stresses. Plant Physiol. 2013 Jul 31;162(4):2028–41.

19. Chalker-Scott L. Environmental Significance of Anthocyanins in Plant Stress Responses. Photochem Photobiol. 1999 Jul;70(1):1–9.

20. Shameer K, Naika MBN, Shafi KM, Sowdhamini R. Decoding systems biology of plant stress for sustainable agriculture development and optimized food production. Prog Biophys Mol Biol. 2019 Aug;145:19–39.

21. Fujita K, Inui H. Review: Biological functions of major latex-like proteins in plants. Plant Sci. 2021 May;306:110856.

22. Cramer GR, Urano K, Delrot S, Pezzotti M, Shinozaki K. Effects of abiotic stress on plants: a systems biology perspective. BMC Plant Biol. 2011;11(1):163.

23. Ober ES, Rajabi A. Abiotic Stress in Sugar Beet. Sugar Tech. 2010 Dec;12(3–4):294–8.

24. Lv X, Chen S, Wang Y. Advances in Understanding the Physiological and Molecular Responses of Sugar Beet to Salt Stress. Front Plant Sci. 2019 Nov 6;10:1431.

25. Jacoby RP, Li L, Huang S, Lee CP, Millar AH, Taylor NL. Mitochondrial Composition, Function and Stress Response in Plants. J Integr Plant Biol. 2012 Oct;no-no.

26. Braun H-P. The Oxidative Phosphorylation system of the mitochondria in plants. Mitochondrion. 2020 Jul;53:66–75.

27. Laluk K, AbuQamar S, Mengiste T. The Arabidopsis Mitochondria-Localized Pentatricopeptide Repeat Protein PGN Functions in Defense against Necrotrophic Fungi and Abiotic Stress Tolerance. Plant Physiol. 2011 Aug 5;156(4):2053–68.

28. Sechet J, Roux C, Plessis A, Effroy D, Frey A, Perreau F, et al. The ABA-Deficiency Suppressor Locus HAS2 Encodes the PPR Protein LOI1/MEF11 Involved in Mitochondrial RNA Editing. Mol Plant. 2015 Apr;8(4):644–56.

29. Murayama M, Hayashi S, Nishimura N, Ishide M, Kobayashi K, Yagi Y, et al. Isolation of Arabidopsis ahg11, a weak ABA hypersensitive mutant defective in nad4 RNA editing. J Exp Bot. 2012;63(14):5301–10.

30. Luscombe NM, Greenbaum D, Gerstein M. What is bioinformatics? An introduction and overview. Yearb Med Inform. 2001 Aug;10(01):83–100.

31. Bayat A. Science, medicine, and the future: Bioinformatics. BMJ. 2002 Apr 27;324(7344):1018–22.

32. Baxevanis AD, Bateman A. The Importance of Biological Databases in Biological Discovery. Curr Protoc Bioinforma [Internet]. 2015 Jun [cited 2021 Nov 24];50(1). Available from: https://onlinelibrary.wiley.com/doi/10.1002/0471250953.bi0101s50

33. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 2017 Jan 4;45(D1):D353–61.

34. Jenuth JP. Publicly Available Tools and Resources on the Web. :12.

35. Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, et al. Ensembl 2020. Nucleic Acids Res. 2019 Nov 6;gkz966.

36. Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, et al. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. 2012 Jan;40(D1):D1178–86.

37. The Gene Ontology Consortium. Gene Ontology Consortium: going forward. Nucleic Acids Res. 2015 Jan 28;43(D1):D1049–56.

38. Zhang H. The review of transcriptome sequencing: principles, history and advances. IOP Conf Ser Earth Environ Sci. 2019 Oct 1;332(4):042003.

39. Ambardar S, Gupta R, Trakroo D, Lal R, Vakhlu J. High Throughput Sequencing: An Overview of Sequencing Chemistry. Indian J Microbiol. 2016 Dec 1;56(4):394–404.

40. Sohn J, Nam J-W. The present and future of *de novo* whole-genome assembly. Brief Bioinform. 2016 Oct 14;bbw096.

41. Liao X, Li M, Zou Y, Wu F-X, Yi-Pan, Wang J. Current challenges and solutions of de novo assembly. Quant Biol. 2019 Jun;7(2):90–109.

42. Gupta A, Gupta S, Jatawa SK, Kumar A, Suravajhala P. A Simplest Bioinformatics Pipeline for Whole Transcriptome Sequencing: Overview of the Processing and Steps from Raw Data to Downstream Analysis. bioRxiv. 2019;836973.

43. Li X, Nair A, Wang S, Wang L. Quality Control of RNA-Seq Experiments. In: Picardi E, editor. RNA Bioinformatics [Internet]. New York, NY: Springer New York; 2015 [cited 2021 Nov 24]. p. 137–46. (Methods in Molecular Biology; vol. 1269). Available from: http://link.springer.com/10.1007/978-1-4939-2291-8_8

44. Rockmann C, Endrullat C, Frohme M, Pospisil H. Next Generation Sequence Analysis. In: Encyclopedia of Bioinformatics and Computational Biology [Internet]. Elsevier; 2019 [cited 2021 Nov 24]. p. 352–63. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780128096338202069

45. Hu J, Ge H, Newman M, Liu K. OSA: a fast and accurate alignment tool for RNA-Seq. Bioinformatics. 2012 Jul 15;28(14):1933–4.

46. Ghosh S, Chan C-KK. Analysis of RNA-Seq Data Using TopHat and Cufflinks. In: Edwards D, editor. Plant Bioinformatics [Internet]. New York, NY: Springer New York; 2016 [cited 2021 Nov 24]. p. 339–61. (Methods in Molecular Biology; vol. 1374). Available from: http://link.springer.com/10.1007/978-1-4939-3167-5_18

47. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019 Aug;37(8):907–15.

48. RNA Processing in Eukaryotes | OpenStax Biology 2e [Internet]. [cited 2021 Nov 25]. Available from: https://courses.lumenlearning.com/suny-osbiology2e/chapter/rna-processing-in-eukaryotes/

49. Annotating Genomes with GFF3 or GTF files [Internet]. [cited 2021 Nov 24]. Available from: https://www.ncbi.nlm.nih.gov/genbank/genomes_gff/

50. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014 Apr 1;30(7):923–30.

51. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res. 2019 May 7;47(8):e47.

52. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinforma Oxf Engl. 2010 Jan 1;26(1):139–40.

53. Hardcastle TJ, Kelly KA. baySeq: empirical Bayesian methods for identifying differential expression in sequence count data. BMC Bioinformatics. 2010 Aug 10;11:422.

54. Tarazona S, Furió-Tarí P, Turrà D, Pietro AD, Nueda MJ, Ferrer A, et al. Data quality aware analysis of differential expression in RNA-seq with NOISeq R/Bioc package. Nucleic Acids Res. 2015;43(21):e140–e140.

55. Li J, Tibshirani R. Finding consistent patterns: a nonparametric approach for identifying differential expression in RNA-Seq data. Stat Methods Med Res. 2013 Oct;22(5):519–36.

56. EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments - PubMed [Internet]. [cited 2021 Nov 25]. Available from: https://pubmed.ncbi.nlm.nih.gov/23428641/

57. Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol. 2013 Jan;31(1):46–53.

58. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

59. Costa-Silva J, Domingues D, Lopes FM. RNA-Seq differential expression analysis: An extended review and a software tool. PLOS ONE. 2017 Ara;12(12):e0190152.

60. Nakagawa S, Santos ESA. Methodological issues and advances in biological meta-analysis. Evol Ecol. 2012 Sep;26(5):1253–74.

61. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. Introduction to meta-analysis. John Wiley & Sons; 2021.

62. Villa V, Negrini S, Capodaglio P, Zaina F, Cimolin V, Vismara L, et al. Osteopathic manipulative treatment in obese patients with chronic low back pain: A pilot study. Man Ther. 2012;17(5):451–5.

63. Salman MC. Pelvik organ desteğinin epidemiyolojisi, patofizyolojisi ve değerlendirilmesi. In: Ayhan A, Taşkıran Ç, Dursun P, editors. Danforth's obstetrik ve jinekoloji. 10th ed. Güneş Tıp Kitabevleri; 2010. p. 819–99.

64. Güner H. Ürojinekoloji ve Pelvik Rekonstrüktif Cerrahi. 2. Ankara: Güneş Tıp Kitabevleri; 2008.

65. Kaya S. İdiopatik detrüsör aşırı aktivitesi olan hastalarda farklı tedavi modalitelerinin karşılaştırılması: fizik tedavi ve rehabilitasyon, medikal tedavi. Hacettepe Üniversitesi; 2008.

66. Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

67. R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

68. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019). "Welcome to the tidyverse." Journal of Open Source Software, 4(43), 1686. doi: 10.21105/joss.01686.

69. Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, Huber W (2005). "BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis." Bioinformatics, 21, 3439–3440.

70. Luo, Weijun, Brouwer, Cory (2013). “Pathview: an R/Bioconductor package for pathway-based data integration and visualization.” Bioinformatics, 29(14), 1830-1831. doi: 10.1093/bioinformatics/btt285.

71. Sherry S, Xiao C, Durbrow K, Kimelman M, Rodarmer K, Shumway M, et al. Ncbi sra toolkit technology for next generation sequence data. In: Plant and Animal Genome XX Conference (January 14-18, 2012) Plant and Animal Genome. 2012.

72. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019 Aug;37(8):907–15.

73. Sequence Alignment/Map format and SAMtools | Bioinformatics | Oxford Academic [Internet]. [cited 2021 Dec 6]. Available from: https://academic.oup.com/bioinformatics/article/25/16/2078/204688

74. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014 Apr 1;30(7):923–30.

75. Rodríguez del Río Á, Minoche AE, Zwickl NF, Friedrich A, Liedtke S, Schmidt T, et al. Genomes of the wild beets *Beta patula* and *Beta vulgaris* ssp. *maritima*. Plant J. 2019 Sep;99(6):1242–53.

76. Marot G, Jaffrezic F, Rau A. metaRNASeq: Differential meta-analysis of RNA-seq data. :13.

77. Holmquist L, Dölfors F, Fogelqvist J, Cohn J, Kraft T, Dixelius C. Major latex protein-like encoding genes contribute to Rhizoctonia solani defense responses in sugar beet. Mol Genet Genomics. 2021 Jan;296(1):155–64.

78. Ghaemi R, Pourjam E, Safaie N, Verstraeten B, Mahmoudi SB, Mehrabi R, et al. Molecular insights into the compatible and incompatible interactions between sugar beet and the beet cyst nematode. BMC Plant Biol. 2020 Oct 22;20:483.

79. Liu L, Wang B, Liu D, Zou C, Wu P, Wang Z, et al. Transcriptomic and metabolomic analyses reveal mechanisms of adaptation to salinity in which carbon and nitrogen metabolism is altered in sugar beet roots. BMC Plant Biol. 2020 Apr 3;20:138.

80. Xing W, Pi Z, Liu J, Li X, Zou Y, Wang M, et al. Comparative transcriptome analysis reveals an ABA-responsive regulation network associated with cell wall organization and oxidation reduction in sugar beet. Plant Growth Regul. 2020 May;91(1):127–41.

81. Thioredoxin-mediated redox signalling in plant immunity - ScienceDirect [Internet]. [cited 2021 Dec 4]. Available from: https://www.sciencedirect.com/science/article/pii/S0168945217312517

82. The cytochrome P450 superfamily: Key players in plant development and defense | Elsevier Enhanced Reader [Internet]. [cited 2021 Dec 4]. Available from: https://reader.elsevier.com/reader/sd/pii/S2095311914609801?token=FA874A354A2DC0E1ADD2306C92A7FB6E7C8ECFD94FC321E8CCB79A69CACB331656B01E4E27E2E4B2ED760BA19D8AFB23&originRegion=eu-west-1&originCreation=20211204173450

83. Da Fonseca-Pereira P, Daloso DM, Gago J, Nunes-Nesi A, Araújo WL. On the role of the plant mitochondrial thioredoxin system during abiotic stress. Plant Signal Behav. 2019;14(6):1592536.

84. Upadhyay RS, Meena M, Prasad V, Zehra A, Gupta VK. Mannitol metabolism during pathogenic fungal–host interactions under stressed conditions. Front Microbiol. 2015;6:1019.

85. Torii KU. Leucine-rich repeat receptor kinases in plants: structure, function, and signal transduction pathways. Int Rev Cytol. 2004;234(243):1–46.

86. Estévez IH, Hernández MR. Plant glutathione S-transferases: an overview. Plant Gene. 2020;23:100233.

87. Lv, X., Chen, S., & Wang, Y. (2019). Advances in understanding the physiological and molecular responses of sugar beet to salt stress. *Frontiers in plant science*, *10*, 1431.

88. Yu, B., Chen, M., Grin, I., & Ma, C. (2020). Mechanisms of sugar beet response to biotic and abiotic stresses. *Mechanisms of Genome Protection and Repair*, 167-194.

89. Bolukbasi, E., & Aras, E. S. (2015). Third generation DNA sequencing technologies. *DNA*, *1*(3).

90. Tseng, G. C., Ghosh, D., & Feingold, E. (2012). Comprehensive literature review and statistical considerations for microarray meta-analysis. *Nucleic acids research*, *40*(9), 3785-3799.

91. Ma, T. (2018). *Differential Expression and Feature Selection in the Analysis of Multiple Omics Studies* (Doctoral dissertation, University of Pittsburgh).

92. Liang, W., Ma, X., Wan, P., & Liu, L. (2018). Plant salt-tolerance mechanism: A review. Biochemical and biophysical research communications, 495(1), 286-291.

93. Zhu, J. K. (2003). Regulation of ion homeostasis under salt stress. Current opinion in plant biology, 6(5), 441-445.

94. Hubbard, M. J., & Cohen, P. (1993). On target with a new mechanism for the regulation of protein phosphorylation. Trends in biochemical sciences, 18(5), 172-177.

95. Bult CJ, Blake JA, Smith CL, Kadin JA, Richardson JE, the Mouse Genome Database Group. 2019. Mouse Genome Database (MGD) 2019. Nucleic Acids Res. 2019 Jan. 8;47 (D1): D801–D806.

96. Cohen, S. P., & Leach, J. E. (2019). Abiotic and biotic stresses induce a core transcriptome response in rice. Scientific reports, 9(1), 1-11.

97. Skorupa, M., Gołębiewski, M., Domagalski, K., Kurnik, K., Nahia, K. A., Złoch, M., ... & Tyburski, J. (2016). Transcriptomic profiling of the salt stress response in excised leaves of the halophyte Beta vulgaris ssp. maritima. *Plant Science*, *243*, 56-70.

98. Kobayashi, T., Nozoye, T., & Nishizawa, N. K. (2019). Iron transport and its regulation in plants. *Free Radical Biology and Medicine*, *133*, 11-20.

99. Shirasawa, K., Takabe, T., Takabe, T., & Kishitani, S. (2006). Accumulation of glycinebetaine in rice plants that overexpress choline monooxygenase from spinach and evaluation of their tolerance to abiotic stress. *Annals of botany*, *98*(3), 565-571.

100. Lee, M., Jung, J. H., Han, D. Y., Seo, P. J., Park, W. J., & Park, C. M. (2012). Activation of a flavin monooxygenase gene YUCCA7 enhances drought resistance in Arabidopsis. *Planta*, *235*(5), 923-938.

101. Yao, K., & Wu, Y. Y. (2016). Phosphofructokinase and glucose-6-phosphate dehydrogenase in response to drought and bicarbonate stress at transcriptional and functional levels in mulberry. *Russian Journal of Plant Physiology*, *63*(2), 235-242.

102. Schelbert, S., Aubry, S., Burla, B., Agne, B., Kessler, F., Krupinska, K., & Hörtensteiner, S. (2009). Pheophytin pheophorbide hydrolase (pheophytinase) is involved in chlorophyll breakdown during leaf senescence in Arabidopsis. *The Plant Cell*, *21*(3), 767-785.

103. Dos Santos, A. L., Chaves-Silva, S., Yang, L., Maia, L. G. S., Chalfun-Júnior, A., Sinharoy, S., ... & Benedito, V. A. (2017). Global analysis of the MATE gene family of metabolite transporters in tomato. *BMC plant biology*, *17*(1),

104. Ashraf, M. H. P. J. C., & Harris, P. J. (2013). Photosynthesis under stressful environments: an overview. *Photosynthetica*, *51*(2), 163-190.

105. da Silva, P. F., de Matos, R. M., Borges, V. E., Sobrinho, T. G., Neto, J. D., de Lima, V. L. A., ... & de Oliveira Pereira, M. (2018). Water consumption of'Beta vulgaris' L. cultivated in greenhouse under fertigation and types of foundation fertilization. *Australian Journal of Crop Science*, *12*(8), 1335-1341.

106. Cavanillas, M. L., Alcina, A., Núñez, C., De Las Heras, V., Fernández-Arquero, M., Bartolomé, M., ... & Urcelay, E. (2010). Polymorphisms in the IL2, IL2RA and IL2RB genes in multiple sclerosis risk. *European Journal of Human Genetics*, *18*(7), 794-79