Transcriptome profiling of *Eutrema salsugineum* under low phosphate and low sulfur

Transcriptome profiling of *Eutrema salsugineum* under low phosphate and low sulfur

By Si Jing Zhang, B.Sc.

A Thesis Submitted to the School of Graduate Studies in the Partial Fulfillment of the Requirements for the Degree Masters in Biology

McMaster University © Copyright by Si Jing Zhang September 16, 2020

McMaster University
Masters in Biology (2020)
Hamilton, Ontario (Department of Biology)

TITLE: Transcriptome profiling of *Eutrema salsugineum* under low phosphate and low sulfur

AUTHOR: Si Jing Zhang (McMaster University)
SUPERVISOR: Dr. Brian Golding, Dr. Elizabeth Weretilnyk

NUMBER OF PAGES: xii, 54

Abstract

Improving the efficiency by which crops use nutrients is critical for maintaining high crop productivity while reducing fertility management costs and eutrophication related to fertilizer runoff. The native crucifer and halophyte, Yukon *Eutrema salsugineum*, was used in this study. Yukon *E. salsugineum* is closely related to important Brassica crops and thrives in its native habitat on soil that is low in available phosphate (P_i) and high in sulfur (S). To determine how Yukon *E. salsugineum* copes with low P_i, leaf transcriptomes were prepared from four week-old plants grown in controlled environment chambers using soil lacking or supplemented with P_i and/or S. This thesis focused on using bioinformatic approaches to assemble, analyze and compare the transcriptome profiles produced by the Yukon *E. salsugineum* plants undergoing four nutrient combinations of high and/or low P_i and S. The objective of the study was to identify traits associated with altered S and/or P_i with the prediction based on other species that low P_i, in particular, would pose the greatest stress and hence elicit the greatest transcriptional reprogramming. Transcriptome libraries were generated from four treatment groups with three biological replicates each. Reads in each library were mapped to 23,578 genes in the *E. salsugineum* transcriptome with an average unique read mapping ratio of 99.52%. Surprisingly, pairwise comparisons of the transcriptomes showed little evidence of P_i-responsive reprogramming whereas treatments differing in soil S content showed a clear S-responsive transcriptome profile. Principal Component Analysis revealed that the low variance quaternary Principal Component distinguished the transcriptomes of plants undergoing low versus high P_i treatments with differential gene expression analysis only finding 11 P_i-responsive genes. This outcome suggests that leaf transcriptomes of Yukon *E. salsugineum* plants under low P_i are largely undifferentiated from plants provided with P_i and is consistent with Yukon *E. salsugineum* maintaining P_i homeostasis through fine-tuning the expression of protein-coding and non-coding RNA rather than large-scale transcriptomic reprogramming. Previous research has shown Yukon *E. salsugineum* to be very efficient in its use of P_i and this work suggests that the altered expression of relatively few genes may be needed to develop P_i-efficient crops to sustain the crop demand of a growing population.

Acknowledgements

Acknowledgements go to NSERC and to supervisors Dr. Golding and Dr. Weretilnyk and external supervisor Dr. Cameron.

Contents

List of Figures

List of Tables

List of Abbreviations

SO_4^{2-} sulfate

miR399 *microRNA399*

Smc1 STRUCTURAL MAINTENANCE OF CHROMOSOMES 1

A. thaliana *Arabidopsis thaliana*

E. salsugineum *Eutrema salsugineum*

APS1-3 ATP sulfurylase 1 to 3

ATP adenosine triphosphate

ATP2 ATP SYNTHASE SUBUNIT β

BHLH32 BASIC HELIX-LOOP-HELIX 32

CDC25 CELL DIVISION CYCLE 25

cDNA complementary DNA

CRE1 CYTOKININ RECEPTOR AtAHK4-LIKE PROTEIN

DAG diacylglycerol

DEG differentially expressed gene

DGD DIGALACTOSYLDIACYLGLYCEROL SYNTHASE

DGDG digalactosyldiacylglycerol

EGL3 ENHANCER OF GLABRA3

ER endoplasmic reticulum

FDR false discovery rate

FPKM fragments per kilobase per million mapped reads

GA gibberellic acid

GL HOMEOBOX-LEUCINE ZIPPER PROTEIN GLABRA

GO Gene Ontology

IPS1 induced by phosphate starvation 1

IPS2 induced by phosphate starvation 2

JA jasmonic acid

JGI Joint Genome Institute

LFC $\log_2$ fold change

lncRNA long non-coding RNA

MGD MONOGALACTOSYLDIACYLGLYCEROL SYNTHASE

MGDG monogalactosyldiacylglycerol

miRNA microRNA

mRNA messenger RNA

MYB Myb-like DNA-binding

MYB-CC Myb coiled-coil

NPC NON-SPECIFIC PLC

P phosphorus

P_i inorganic phosphate

P1BS PHR1 binding site

PA PHOSPHATASE PHOSPHATIDIC ACID PHOSPHATASE

PAP PURPLE ACID PHOSPHATASE

PC Principal Component

PC phosphatidylcholine

PCA Principal Component Analysis

PE phosphatidylethanolamine

PEP phosphoenolpyruvate

PEPC PHOSPHOENOLPYRUVATE CARBOXYLASE

PG phosphatidylglycerol

PHF1 PHOSPHATE TRANSPORTER TRAFFIC FACILIATOR 1

PHO1 PHOSPHATE 1

PHO2 PHOSPHATE 2

PHR1 PHOSPHATE STARVATION RESPONSE 1

PHT PHOSPHATE TRANSPORTER

PI-PLC PHOSPHOINOSITIDE-SPECIFIC-PHOSPHOLIPASE C

PLC PHOSPHOLIPASE C

PLD PHOSPHOLIPASE D

PPCK1 and 2 PHOSPHOENOLPYRUVATE CARBOXYKINASE 1 and 2

pre-miRNA precursor miRNA

pri-miRNA primary miRNA

PSI phosphate starvation-inducible

PSR phosphate starvation response

Rhd RHODANESE

RNA-seq RNA sequencing

RNAi RNA interference

RNase RIBONUCLEASE

RNase III RIBONUCLEASE III

RNS S-LIKE RIBONUCLEASE

RNS1 S-LIKE RIBONUCLEASE 1

Rps16 40S RIBOSOMAL PROTEIN S16

rRNA ribosomal RNA

S sulfur

SIZ1 SAP AND MIZ/SP-RING ZINC FINGER DOMAIN-CONTAINING PROTEIN 1

SLIM1 SULFUR LIMITATION 1

SPX SYG1/PHO81/XPR1 DOMAIN-CONTAINING

SQD URIDINE-PHOSPHATE-SULFOQUINOVOSE SYNTHASE

SQDG sulfoquinovosyldiacylglycerol

SRA Sequence Read Archive

SULTR SULFATE TRANSPORTER

SUMO small, ubiquitin-related modifier

TAIR The *Arabidopsis* Information Resource

TTG1 TRANSPARENT TESTA GLABRA1

ZAT6 ZINC FINGER OF *ARABIDOPSIS THALIANA*

Declaration of Authorship

I, Si Jing ZHANG, declare that this thesis titled, "Transcriptome profiling of *EUTREMA SALSUGINEUM* under low phosphate and low sulfur" and the work presented in it are my own. I confirm that:

- Dr. Elizabeth Weretilnyk and Dr. Vera Velasco conceived the experiment presented in this work
- Dr. Vera Velasco and Amanda Garvin performed the experiment, conducted physiological analysis, tissue collection, RNA extraction and RNA preparation for sequencing
- The author performed raw data processing, designed the analytical pipeline and performed data analysis, those results are presented in the work below
- Dr. Elizabeth Weretilnyk and Dr. Brian Golding supervised the project

Chapter 1

Literature Review

1.1 Phosphate Availability and Uptake

Inorganic phosphate (P_i) is the main source of phosphorus (P) in plants, a key component of crucial organic molecules including nucleic acids, phospholipids as well as adenosine triphosphate (ATP). P_i is involved in major biochemical pathways such as nucleic acid synthesis, amino acid synthesis, photosynthesis, respiration and stress responses (reviewed by Theodorou and Plaxton 1993). P is a macronutrient and plants that are deficient show severely stunted growth that leads to substantial crop loss in agriculture (Grant et al. 2001).

1.1.1 The effects of excess fertilizer use on the environment

Despite the high demand for P_i by plants, it is typically a limiting nutrient in the soil due to its presence in forms that are inaccessible such as organic P that must be mineralized for plants to use or as insoluble, hence unavailable, precipitates (Stewart and Tiessen 1987; Tisdale and Nelson 1975). As a result, P_i fertilizer (P_2O_5) is widely used to increase P_i availability to plants. There are several considerations that make a reliance on the use of fertilizers to sustain high rates of crop productivity problematic. In 2017, the total world demand for P_i fertilizer was over 43 million tonnes and this demand is predicted to increase by 2.2% annually (FAO 2017). However, even the current rate of global P_i usage is unsustainable and could deplete the world's limited supply of accessible rock P_i reserves by 2040 (Blackwell et al. 2019). Moreover, runoff of fertilizer from fields can contaminate aquatic systems leading to eutrophication (Schindler 1971). Increased nitrogen and P_i produce algal blooms in aquatic systems and accelerate the growth of aerobic bacteria that deplete oxygen from the water, while the decay of organic matter further contributes to severe hypoxia that can kill oxygen-dependent aquatic life (Weiss 1969). Other adverse changes to aquatic ecosystems are induced by the presence of high nutrients. Thick layers of algae on the water surface reduce available sunlight for bottom-dwelling animals and plants (Shaw et al. 2003), while hypoxia further alters conditions that together greatly decrease the biodiversity in the ecological community. Algal blooms that include genera such as dinoflagellates and cyanobacteria release toxins that can directly kill animals. For example, microcystins produced by some cyanobacteria species cause poisoning through hepatotoxicosis, altering cell physiology in the stomach that leads to intestinal cell damage (Carmichael and Falconer 1993). The bioaccumulation of algal toxins in fish, wild animals and livestock can render food sources toxic to other animals, including humans (Rosales-Loessener 1989; Trainer and Baden 1999). Total P_i in water bodies can be reduced by strictly regulating farming practices that can leach nutrients into aquatic systems and through remediation. However, it remains that plants have a limited pool of accessible P_i for meeting crop needs and this means more P_i must be provided. Modifying crops to make them more efficient in their use of available P_i would reduce fertilizer use and ultimately diminish its environmental impacts such as eutrophication.

1.1.2 Phosphate availability in the soil

Plants primarily take up inorganic P_i, also known as "reactive phosphate" in its orthophosphoionic form ($H_2PO_4^-$) and the availability of plant-accessible P_i in its soluble form is dependent on the pH of the soil (Tisdale and Nelson 1975). In both cultivated and uncultivated soils, P_i has a tendency to precipitate with iron, manganese and aluminum in acidic soils (pH 5.5 and below) and with calcium and magnesium in alkaline soil (pH 8 and above) (Tisdale and Nelson 1975). Generally, plants have the most access to soil P_i when the pH is between 5.5 and 7.0 (Tisdale and Nelson 1975). P_i is taken up by the roots at the root-soil interface creating a radial zone of depletion around the root through which P_i diffuses. The rate of P_i diffusion can be as low as 4×10^{-11} $cm^2 \cdot sec^{-1}$ in P_i deficient soils or 2 to 4 $\times 10^{-8}$ $cm^2 \cdot sec^{-1}$ in high-P_i soils (Barber et al. 1963; Bieleski 1973). Furthermore, the P_i concentration in roots is typically 1,000 to 10,000 times greater than the surrounding soil creating an unfavourable gradient for diffusion (Bieleski 1973; Tisdale and Nelson 1975). Additionally, the negatively charged forms of P_i available to plants must overcome a highly negative membrane potential. As such, P_i is transported across root membranes in co-transport with H^+ to prevent the cell membrane potential from becoming hyperpolarized (Ullrich-Eberius et al. 1981). Taken together, low P_i concentrations and its anionic properties pose challenges to plants for meeting their needs for this essential nutrient.

To improve P_i availability, plants actively release P_i from organic sources in soil through secretion of PURPLE ACID PHOSPHATASEs (PAPs) from roots including from phytate (inositol hexophosphate) by the action of phytase-specific PAPs (reviewed by Tran et al. 2010). The capacity to hydrolyze P_i by root secreted phosphatases is a mechanism that is up-regulated in many plants experiencing P_i starvation conditions (Misson et al. 2005; Wu et al. 2003).

1.1.3 Plant phosphate transporters

In addition to modifying root morphology and secreting P_i mobilizing enzymes, plants use P_i transporters to actively move P_i across membranes. In plant roots, a number of P_i transporter families have been characterized in the past 40 years, with the expression of a class of high-affinity transporters induced by P_i deficient soil conditions receiving the most attention (reviewed by Raghothama 2000). High-affinity transporters are mostly distributed across P_i-starved roots, but can also be expressed in other organs such as leaves and flowers (Hamburger 2002; Leggewie 1997; Liu et al. 1998). On the other hand, low affinity transporters have constitutive expression and can be expressed throughout the entire plant (reviewed by Raghothama 2000).

One of the most prominent P_i transport mechanisms includes the high-affinity family of PHOSPHATE TRANSPORTERs (PHTs) and PHOSPHATE TRANSPORTER TRAFFIC FACILIATOR 1 (PHF1). PHTs are involved in P_i uptake but they also re-mobilize P_i throughout the plant and in *Arabidopsis thaliana* (*A. thaliana*) are categorized under 5 families: PHT1, PHT2, PHT3, PHT4 and PHT5 (Bari 2006; Huang et al. 2013). Members of the PHT1 family are considered to be high-affinity transporters and several members are

up-regulated when the plant senses a low P_i environment (Raghothama 1999). The P_i-binding site of the PHT1 group is highly conserved and variations in P_i affinity between PHT1 family members are likely due to post-translational modifications (Ceasar et al. 2016). Fontenot et al. (2015) demonstrated the ability of PHT1;1 to form monomeric dimers or trimers and disruption to the oligomerization ability of PHT1;1 decreased P_i uptake in *A. thaliana*. Additionally, Ayadi et al. (2015) proposed that PHT1 might have dual affinity based on modelling and sequence alignment of the PHT1 active sites. In *A. thaliana*, 9 genes of the *PHT1* family have been identified and all, with the exception of *AtPHT1;6*, are expressed in roots, evidence of their roles in root P_i uptake (Mudge et al. 2002). *AtPHT1;1*, *AtPHT1;8* and *AtPHT1;9* encode for high-affinity transporters that are responsible for root-to-shoot P_i-remobilization (Remy et al. 2012; Shin et al. 2004). Variability has been found among plants with respect to the complement of *PHT1* members found. In the halophyte *Eutrema salsugineum* (*E. salsugineum*), there is no homolog for *AtPHT1;1* but there are seven orthologs of *EsPHT1;3* (Velasco et al. 2016; Wang et al. 2017). The association with low P_i expression may be conserved. Expression profiles for a number of *EsPHT1* members have been tested and results show that *EsPHT1;3*, *EsPHT1;4*, *EsPHT1;5* and *EsPHT1;8* are up-regulated in roots of *E. salsugineum* plants grown under conditions that produce low P_i-availability in the soil (Velasco et al. 2016). In contrast to the distribution of PHT1 family members in roots, PHT2 transporters are predominantly expressed in the shoots and are located in the plasma membranes of the chloroplast (reviewed by Liu et al. (2011)). *PHT2;1* expression in leaves was reported as being not responsive to low P_i for both *A. thaliana* and *E. salsugineum*, suggesting that regulation among *PHT* family members may be conserved and not necessarily regulated by P_i availability (Velasco et al. 2016). PHT3s are mitochondrial P_i transporters and in *A. thaliana*, there are three members in this family. All three members of the PHT3 family have been identified in *E. salsugineum* and there are two orthologs of *EsPHT3;2* (Velasco 2017; Wang et al. 2017). PHT4s transport P_i with high affinity and are located in the membranes of both the golgi apparatus and the plastids (reviewed by Młodzińska and Zboińska 2016). *E. salsugineum* plants possess all six members of the PHT4 family also identified in *A. thaliana* (Champigny et al. 2013; Guo et al. 2008a; Velasco 2017). Lastly, PHT5s are the least characterized P_i transporters in the PHT family. Currently, members of the PHT5 family consist of putative SYG1/PHO81/XPR1 DOMAIN-CONTAINING (SPX) proteins that participate in P_i transport across the vacuolar membrane (reviewed by Młodzińska and Zboińska 2016). There are three putative members of PHT5 identified in *A. thaliana* and their homologs have been identified in *E. salsugineum* (Simopoulos 2019). The functions of all PHTs identified in *E. salsugineum* are inferred from their homologs in *A. thaliana* and to date, there are only two studies that validate the biochemical roles of individual PHTs in *E. salsugineum* (Yang et al. 2020a,b).

PHF1 is a golgi-associated protein that mediates the exit of PHT1 proteins from the golgi apparatus, specifically PHT1;1 (González et al. 2005; Huang et al. 2013). The expression of *PHF1* is regulated by P_i starvation and no other nutrient deficiency triggers have been observed, including those associated with deficiencies of potassium, sulphur, iron, nitrogen or

removal of sucrose from the growing medium (González et al. 2005).

PHOSPHATE 1 (PHO1) is a family of proton-coupled P_i transporters that have a high amino acid conservation among plants but little homology to the PHT family or to solute transporters in other organisms (Hamburger 2002; Rausch and Bucher 2002). There are 11 gene members in the *PHO1* family and all encode the SPX domain in the N-terminal region. The SPX domain in yeast is well characterized and known to suppress low-affinity P_i-binding activity to modulate P_i homeostasis under low P_i conditions (Hürlimann et al. 2009). Of the characterized SPX genes in plants, each possesses a diverse function and structure other than the prominent conserved SPX domain. PHO1 is predominantly expressed in the root and in *A. thaliana*, *pho1-1* mutants exhibit deficient shoot P_i accumulation but P_i uptake into the roots is not affected while overexpression of PHO1 in *A. thaliana* leaves leads to increased P_i accumulation and P_i secretion into the extracellular medium (Poirier et al. 1991; Sfanovic et al. 2011). This suggests a role of PHO1 in regulating xylem loading of P_i rather than P_i uptake into root epidermal cells (Poirier et al. 1991).

1.2 Phosphate Starvation Response

Plants have evolved numerous strategies to closely regulate cellular P_i homeostasis under conditions of P_i deficiency. These phosphate starvation response (PSR) mechanisms are induced when the plant senses a low P_i environment and the genes that participate in PSR mechanisms are known as phosphate starvation-inducible (PSI) genes (Raghothama 1999).

1.2.1 Transcriptional regulation

PHOSPHATE STARVATION RESPONSE 1 (PHR1) is a Myb coiled-coil (MYB-CC) family transcription factor that up-regulates phosphate starvation-related genes during conditions of P_i deficiency. The MYB-CC family of transcription factors is unique to plants and is represented by 15 members in *A. thaliana* (Rubio et al. 2001). The Myb-like DNA-binding (MYB) domain is a helix-turn-helix DNA-binding motif found in eukaryotic transcription factors and the coiled-coil motif is also a common motif in transcription factors. PHR1 binds to an upstream promoter motif of GNATATNC at the PHR1 binding site (P1BS) (Rubio et al. 2001). The P1BS *cis* element is found upstream of PSI genes such as *induced by phosphate starvation 1 (IPS1)*, *Pht1;1/Pht1;2*, *microRNA399* (*miR399*), *phosphate transporter 1; homolog 1* (*PHO1;H1*) and *URIDINE-PHOSPHATE-SULFOQUINOVOSE SYNTHASE (SQD) 1* and *2* (Sobkowiak et al. 2012). Another MYB family transcription factor regulating P_i starvation response is MYB62: MYB62 is localized in the nucleus and is expressed in seedling shoots during P_i limitation. PSR is modulated by MYB62 through the control of the gibberellic acid (GA) pathway and GA biosynthesis genes (Devaiah et al. 2009).

PHR1 expression appears to be unresponsive to P_i conditions. Rather, post-translational regulation of PHR1 occurs by sumoylation, a process involving the attachment of a small, ubiquitin-related modifier (SUMO) to PHR1 by SAP AND MIZ/SP-RING ZINC FINGER

DOMAIN-CONTAINING PROTEIN 1 (SIZ1). Sumoylation of the key PSR regulator PHR1 by SIZ1 regulates low P_i-induced responses in a mechanism that stabilizes PHR1 and may enhance its binding to the P1BS motif promoter. In this manner, SIZ1 is a positive regulator of the major PHR1-regulated PSI genes *IPS1* and *S-LIKE RIBONUCLEASE 1 (RNS1)* (Miura et al. 2005). However, work with rice *siz1* plants experiencing P_i deficiency led to both the suppression and expression of genes normally responsive to low-P_i so SIZ1 likely exerts positive and negative regulation on genes related to PSR. PHR1 is up-regulated by SIZ1 through sumoylation and genes regulated by PHR1 such as *IPS1* and *RNS1* are also up-regulated as a result. However, *siz1* mutants have enhanced P_i deficiency phenotypes such as increased lateral root growth and root/shoot ratio, suggesting that SIZ1 is a negative regulator for P_i-responsive root architecture (Miura et al. 2005). The second system regulating *PHR1* expression is through the SPX proteins SPX1 and SPX2. The two SPX proteins are part of the P_i sensing system and are able to modulate the expression of *PHR1* through direct competitive binding to the promoter recognition domain under high P_i conditions and P_i concentration directly affects the binding affinity of SPX1 and 2 to PHR1 (Puga et al. 2014). The function of SPX1 is also supported by studies in rice and legumes (Yao et al. 2014; Zhang et al. 2016). There is redundancy between SPX1 and SPX2 function and they both possess an upstream P1BS motif in their promoter regions, which are recognized by PHR1 to initiate SPX1 and SPX2 expression during low P_i conditions. The expression of SPX under low P_i conditions presents a negative regulatory feedback loop that ensure the quick repression of PHR1 once P_i has been replenished in the plant (Wang et al. 2014b).

The WRKY gene family is a class of transcription factors almost exclusively found in plants with few exceptions in amoebozoa and fungi (Rinerson et al. 2015). The WRKY zinc finger domain interacts with the W-box (T)TGEC(C/T) *cis* element and is characterized by the WRKYGQK motif in the core sequence of the protein (Eulgem et al. 1999). Proteins with a VQ motif (FXXXVQXLTG) also interact with WRKY proteins (Cheng et al. 2012). Many of the WRKY family members have been associated with biotic and abiotic stress response pathways including drought stress, salt stress, temperature stress and nutrient deficiency stress (Chen et al. 2012). RNA interference (RNAi)-mediated silencing of *WRKY75* transcripts in *A. thaliana* resulted in accelerated anthocyanin accumulation under low P_i conditions compared to wild type, suggesting that *AtWRKY75* is differentially expressed in a P_i-dependent pattern (Devaiah et al. 2007b). AtWRKY75 silencing also reduced the expression of *AtWRKY45*, *AtIPS1* and *AtIPS2*, suggesting *AtWRKY75*'s involvement in regulating these genes (Devaiah et al. 2007b). AtWRKY6 and AtWRKY42 repress *PHO1* expression by binding to the *PHO1* W-box promoter motif when P_i is sufficient and the repression is removed under low P_i conditions (Chen et al. 2009). AtWRKY6 and AtWRKY42 may also regulate *PHT1;5* and *PHT1;8* expression because the W-box promoter element is present in the upstream promoter region of both genes that encode high-affinity P_i transporters (Chen et al. 2009). AtWRKY45 positively regulates *AtPHT1;1* expression and negatively regulates *AtWRKY75* under low P_i (Wang et al. 2014a). In rice, the WRKY gene *OsWRKY74* shares homology with the P_i starvation-related WRKY genes in *A. thaliana*. Overexpression of

OsWRKY74 in rice produces plants with heightened low P_i-responsive changes in root architecture and increased root PAP activity relative to wild-type (Dai et al. 2016).

The transcription factor BASIC HELIX-LOOP-HELIX 32 (BHLH32) negatively regulates P_i starvation response when P_i is abundant, such that *bhlh32* mutants exhibit increased total P_i content and increases of PHOSPHOENOLPYRUVATE CARBOXYKINASE 1 and 2 (PPCK1 and 2) under P_i sufficient conditions (Chen et al. 2007). PPCK1 and 2 activity is responsible for altering the kinetic properties and increasing the specific activity of PHOSPHOENOLPYRUVATE CARBOXYLASE (PEPC). PEPC catalyzes the carboxylation of phosphoenolpyruvate (PEP) producing oxaloacetate and releases P_i for re-use by the plant under P_i starvation (Gregory et al. 2009). Additionally, *bhlh32* mutant studies in *A. thaliana* show that root hair growth was not inhibited by high P_i, suggesting BHLH32 as a negative regulator for root hair formation by directly interfering with the hair cell-inducing complex HOMEOBOX-LEUCINE ZIPPER PROTEIN GLABRA (GL), ENHANCER OF GLABRA3 (EGL3) and TRANSPARENT TESTA GLABRA1 (TTG1) (Chen et al. 2007).

ZINC FINGER OF *ARABIDOPSIS THALIANA* (ZAT6) is a Cys-2/His-2 type transcription factor that is highly expressed during P_i starvation. The effect of *ZAT6* over-expression is the repression of a number of PSI genes, inhibition of primary root growth in older plants and increased P_i accumulation. Devaiah et al. (2007a) concluded that the over-expression of *ZAT6* alters root morphology with an effect on the P_i homeostasis of the plant.

1.2.2 *PHO2* and the *PHR1-miR399-PHO2* pathway

A major PSR pathway characterized in higher plants is the PHR1-*miR399*-PHO2 pathway (Bari 2006). In this pathway, PHR1 up-regulates *miR399* expression, a small RNA fragment that binds perfectly to the target messenger RNA (mRNA) sequence of *PHOSPHATE 2 (PHO2)* and directs the targeted mRNA fragment for RNA degradation (**Figure 1.1**). *PHO2* encodes a ubiquitin-conjugating E2 enzyme that interacts with target proteins by catalyzing ubiquitinylation of target proteins that are subsequently recognized by the proteasome for degradation (Bari 2006). Ubiquitinylation targets of PHO2 include PHT1 and PHF1, the removal of *PHO2* transcript by *miR399* allows P_i to be remobilized by PHT1 and PHF1 when the plant is experiencing P_i deficiency. Furthermore, *pho2* mutants of *A. thaliana* are characterized by the accumulation of toxic levels of P_i in the leaves (Delhaize and Randall 1995) and a subset of PSI genes in *A. thaliana*, including *AtIPS1* and *AtIPS2* (*At4*), are up-regulated (Bari 2006).

1.2.3 Non-coding RNAs in P_i stress response regulation

MicroRNAs (miRNAs) play important regulatory roles in nutrient assimilation pathways. MiRNAs are transcribed in the nucleus as precursor primary miRNAs (pri-miRNAs), transcripts that are typically longer than 70 nt (reviewed by O'Brien et al. 2018). The precursors are processed into precursor miRNA (pre-miRNA) by endoribonuclease Drosha, a

RIBONUCLEASE III (RNase III) and the cleaved products are exported into the cytoplasm through nuclear pores where they are processed into mature miRNA by another endoribonuclease Dicer (Lee et al. 2002). The *miR399* family is well characterized as translation regulators for *PHO2*, whose encoded protein is a critical regulator of P_i transporters and the PSR, (Bari 2006) and comprises a critical component of the *PHR1-miR399-PHO2* P_i-signaling pathway. Another miRNA that is involved in P_i stress-related regulation is *miR827*. In rice, *miR827* targets the mRNA of two vacuolar-specific P_i transporters *OsSPX-MFS1* and *OsSPX-MFS2* (Lin et al. 2010). *MiR827* is preferentially expressed in shoots of rice but is localized to roots in *A. thaliana* (Lin et al. 2010). The difference in *miR827* expression between rice and *A. thaliana* suggests that *miR827*-controlled pathways may be regulated by different mechanisms in these two species (Lin et al. 2010). In addition to *miR399* and *miR827*, two other miRNAs, miR778 and miR2111 also respond exclusively to P_i status (reviewed by Kumar et al. 2017).

IPS1 and *induced by phosphate starvation 2 (IPS2)* were described earlier in this review as key PSI genes. Both *IPS1* and *IPS2* are two long non-coding RNAs (lncRNAs) that mimic the target of *miR399* to prevent the degradation of *PHO2* mRNA transcripts and the role and significance of *PHO2* in P_i homeostasis is discussed in section 1.2.2. The defining characteristic of *IPS1* and *IPS2* is a conserved 22 bp region that is found among all plants studied to date. The 22 bp sequence does not completely match *miR399* but contains a mismatched loop at the tenth base of this region. The mismatch base does not affect the binding efficiency of *miR399* to the target transcript but it does prevent miRNA-guided cleavage of the *PHO2* transcript (Jones-Rhoades et al. 2006). Thus *IPS1* and *IPS2* are bound to *miR399* to prevent its binding to *PHO2* transcripts. The expression of *IPS1* and *IPS2* is predominantly in the root and their expression is highly up-regulated when the plant senses a P_i deficit, a response to P_i limitation shown for phylogenetically diverse plants including *A. thaliana*, rice, maize, tomato, oat, soybean, wheat and *Medicago truncatula* (Baldwin et al. 2008; Burleigh and Harrison 1998; Calderon-Vazquez et al. 2008; Franco-Zorrilla et al. 2007; Guo et al. 2008b; Oono et al. 2013a; Wang et al. 2018; Wasaki et al. 2003; Zhang et al. 2017). Mechanistically, the expression of *IPS1* and *IPS2* under P_i starvation promotes the expression of *PHO2*, which serves as a Pi-remobilization suppressor. The expression of *IPS1* and *IPS2* prevents excessive xylem loading of P_i thereby enabling a plant to fine-tune P_i homeostasis during P_i starvation. *IPS2* (formerly known as *At4*), differs from *IPS1* in the mismatch loop of the 22 bp conserved region (Franco-Zorrilla et al. 2007). However, further distinction between the two genes is still unclear, as over-expression and inactivation experiments has suggested functional redundancy between *IPS1* and *IPS2* (Franco-Zorrilla et al. 2007). *IPS2* has been characterized in *E. salsugineum* where it is constitutively expressed and transcripts are elevated under P_i-limiting conditions (Velasco et al. 2016). An orthologous *IPS1* was not identified in the published *E. salsugineum* genome annotation (Yang et al. 2013a).

1.2.4 Plant lipid membrane remodelling under Pi-limitation

P_i-starved plants increase P_i supply by replacing phospholipids with non-phsopholipids, as phospholipids represent a third of the total P_i reservoir in the plant (reviewed by Nakamura 2017). During P_i starvation, endoplasmic reticulum (ER)-derived phosphatidylcholine (PC) and phosphatidylethanolamine (PE) are converted by PHOSPHOLIPASE C (PLC) into diacylglycerol (DAG), a fundamental building block for membrane lipid molecules. There are two variants of PLC: PHOSPHOINOSITIDE-SPECIFIC-PHOSPHOLIPASE C (PI-PLC) and NON-SPECIFIC PLC (NPC). NPC family members possess a variety of dissimilar functions, the activities of two P_i-stress associated NPCs, NPC4 and NPC5, are elevated during P_i starvation (Gaude et al. 2008; Nakamura et al. 2005). Both NPCs have signature PLC activity, converting PC and PE to DAG. Although the functions of the two NPCs are not well characterized, NPC4 is localized to the plasma membrane and has a role in root development (Wimalasekera et al. 2010). NPC5 is localized to the soluble cytosolic fraction and is involved in the accumulation of galactolipids, specifically digalactosyldiacylglycerol (DGDG) (Gaude et al. 2008). In the secondary DAG conversion involving PHOSPHOLIPASE D (PLD) and PHOSPHATIDIC ACID PHOSPHATASE (PA PHOSPHATASE), two of the twelve PLD isoforms, PLDζ1 and 2, are induced by P_i starvation and are involved in low P_i-induced DGDG accumulation in plant roots (Cruz-Ramírez et al. 2006; Li et al. 2006). The two isoforms of PA PHOSPHATASEs, PA PHOSPHATASE 1 and 2 specialize in glycerolipid metabolism and dephosphorylation of lipids, respectively (Brindley and Pilquil 2009; Eastmond et al. 2010). The two PA PHOSPHATASE 1 isoforms characterized in *A. thaliana*, AtPAH1 and AtPAH2, play a role in the regulation of both DIGALACTOSYLDIACYLGLYCEROL SYNTHASE (DGD) and MONOGALACTOSYLDIACYLGLYCEROL SYNTHASE (MGD) under low P_i (Nakamura et al. 2009).

When plants experience low P_i, phospholipids are typically replaced by monogalactosyldiacylglycerol (MGDG), DGDG and sulfoquinovosyldiacylglycerol (SQDG) (reviewed by Nakamura 2013). MGDG is synthesized by Type A MGD 1 and Type B MGDG synthases (MGD2 and MGD3). Levels of MGDG do not significantly change between P_i sufficient and P_i deficient treatments, suggesting that MGDs produce the intermediate MGDGs required to produce DGDG and altering the expression of MGDs can affect DGDG accumulation under P_i starvation (reviewed by Nakamura 2013). DGDG is synthesized by two DGDG synthases, DGD1 and DGD2. DGD1 synthesizes 90% of the DGDG in plants and under normal conditions DGD2 is expressed at a low level, however, both synthases are up-regulated upon P_i starvation in *A. thaliana* and play an important role in P_i starvation-induced phospholipid replacement by glycolipids at the lipid membrane (Kelly and Dörmann 2002). The SQDG biosynthesis pathway involves two SQDG synthases: SQD1 and SQD2. SQDG production is increased in P_i deficient *A. thaliana* plants via the up-regulation of SQD1. Finally, SQDG is localized only in the thylakoid membrane and serve to replace phosphatidylglycerol (PG) under low P_i conditions (reviewed by Nakamura 2013).

1.2.5 Additional P_i stress response pathways

S-LIKE RIBONUCLEASEs (RNSs) are ribonucleases that evolved from the self-incompatibility (S) locus of family *Solanaceae* (Taylor et al. 1993). Two RIBONUCLEASE (RNase) in particular, S-like RNase 1 and 3 (RNS1 and RNS3) are highly induced in plants experiencing P_i starvation (Taylor and Green 1991). The RNS glycoproteins resupply P_i to the plant by releasing P_i molecules from mRNA or tRNA digestion (Bariola et al. 1994). RNS1 is found in the flowers while RNS3 has Pi-induced expression in roots, inflorescence, flowers not is not expressed in leaves (Bariola et al. 1994).

Cytokinins are known to negatively modulate the P_i stress response by causing the down-regulation of P_i stress-responsive genes (Martín et al. 2000). Mutations in the His protein kinase and CYTOKININ RECEPTOR AtAHK4-LIKE PROTEIN (CRE1) results in the failure of cytokinin to inhibit PSR (Franco-Zorrilla et al. 2002). Further, low P_i levels in the plant leads to the repression of CRE1, indicating the role of cytokinins in inhibiting PSR through CRE1 (Martín et al. 2000).

The jasmonic acid (JA) signaling pathway is typically associated with herbivory stress and the phytohormone JA controls a majority of the genes responsive to insect-induced wounding (reviewed by Ghasemi Pirbalouti et al. 2014). JA-responsive changes and PSR have similar characteristics such as anthocyanin accumulation and reduced growth, an observation supported by evidence that P_i starvation induces the JA biosynthetic pathway (Khan et al. 2016). The apparent role for JA in producing the reduced shoot biomass associated with the P_i starvation phenotype was shown by Khan et al. (2016), where *pho1* mutants experienced reduction in shoot growth as a result of the activation of the JA signaling pathway. Additionally, the low Pi-induced JA production provides increased insect resistance for *A. thaliana*, *Solanum lycopersicum* and *Nicotiana bethamiana* in tests with the herbivore *Spodoptera littoralis* (Khan et al. 2016).

Auxin-mediated regulation of root morphology has been studied in P_i deficient plants. As discussed earlier, root architecture is altered by P_i starvation and root sensitivity to low P_i is enhanced by exogenous auxin. However, mutants defective for auxin sensing still have wild type-like lateral root proliferation under low P_i conditions suggesting a complementary but not obligatory role of auxin in modulating root morphology under P_i limitation (López-Bucio et al. 2002; Yang and Finnegan 2010). There is also evidence supporting the role of auxin in altering the lipid membrane composition under P_i starvation by altering the expression of type B MGD (Kobayashi et al. 2006).

Plants may use sucrose as a signaling medium and sucrose plays an important role in interconnecting different signaling pathways under P_i starvation. Sucrose is required for P_i signaling in plants, as a resupply of sucrose to plants starved of P_i and sucrose show an increase in *PHT1;4* transcripts (Franco-Zorrilla et al. 2005). Sucrose and cytokinins have antagonistic effects on the regulation of PSI and promote transport and root sensitivity of auxin (Franco-Zorrilla et al. 2005). The application of exogenous sucrose exaggerates PSR

phenotypes and the inhibition of sucrose biosynthesis or defects in phloem loading result in decreased expression of PSI genes (Karthikeyan et al. 2007).

1.3 Relationship between sulfur assimilation and P_i starvation response

Uptake and movement of sulfate (SO_4^{2-}) in plants requires SULFATE TRANSPORTERs (SULTRs). However, SULTRs are implicated in nutrient acquisition pathways in addition to their primary function as SO_4^{2-} transmembrane transporters (Rouached et al. 2011). The *SULTR* gene family has four groups, *SULTR1* to *4* and all SULTR proteins contain 12 transmembrane domains (reviewed by Gigolashvili and Kopriva 2014). Group 1 SULTRs are high-affinity SO_4^{2-} transporters, group 2 are low affinity SO_4^{2-} transporters, group 3 SULTRs are plastid membrane transporters and group 4 are vacuolar membrane SO_4^{2-} transporters (Kataoka et al. 2004; Takahashi et al. 2011). In *A. thaliana*, *AtSULTR1;3*, *AtSULTR2;1* and *AtSULTR3;4* are differentially expressed under P_i starvation: *AtSULTR1;3* and *AtSULTR3;4* are up-regulated while *AtSULTR2;1* is down-regulated. AtSULTR1;3 is localized in the phloem and is responsible for SO_4^{2-} redistribution from root to shoot (Yoshimoto et al. 2003). AtSULTR2;1 is a low-affinity transporter that is up-regulated under sulfur deficiency and is localized to xylem parenchyma and root pericycle cells (Takahashi et al. 2000). AtSULTR3;4 is involved in SO_4^{2-} translocation in seedlings as well as the chloroplast (Chen et al. 2019). In *A. thaliana*, the expression of all three *AtSULTR* genes is regulated by the P_i starvation response transcription factor PHR1 under P_i starvation as seen by their decreased expression in *phr1* mutants (Rouached et al. 2011). Also, *SULTR1;3* and *AtSULTR2;1* possess a P1BS motif at the 5' promoter region allowing potential transcription initiation by PHR1 (Rouached et al. 2011). Although *AtSULTR3;4* does not possess a PHR1-binding site, it is up-regulated in Pi-deficient *phr1 A. thaliana* mutants (Rouached et al. 2011).

P_i transporters and SO_4^{2-} transporters share a number of common characteristics: both families have 12 consecutive transmembrane domains and share similar phytohormone control mechanisms. Not only does cytokinin affect the expression of PSI including P_i transporters, expression of the cytokinin receptor, CRE1, down-regulates *AtSULTR1;1* and *AtSULTR1;2* in *A. thaliana* (reviewed by Rouached 2011). P_i and SO_4^{2-} assimilation pathways also have similar miRNA control mechanisms. The *PHR1-miR399-PHO2* pathway employs *miR399* to silence *PHO2* transcripts and suppress PHO2 protein translation. *miR395* is induced by the transcription factor SULFUR LIMITATION 1 (SLIM1) that directly represses the expression of *ATP sulfurylase 1 to 3 (APS1-3)* as well as *SULTR2;1* (Ai et al. 2016; Maruyama-Nakashita et al. 2006). In rice, the disruption in the expression of the SO_4^{2-} transporter *OsSULTR3;3* through introducing a premature stop codon via a 1 bp deletion resulted in the reduction of P_i and phytate concentrations and affected the expression of P_i and SO_4^{2-} homeostasis genes (Zhao et al. 2016). However, the function of OsSULTR3;3 is not yet known, as it does not exhibit SO_4^{2-} transporter activity (Zhao et al. 2016).

Proteins of the RHODANESE (Rhd) superfamily are a group of sulfotransferases with a highly conserved catalytic domain. Rhds are closely related to the CELL DIVISION CYCLE

25 (CDC25) superfamily of phosphatases with members having a conserved catalytic loop of seven amino acid residues instead of six at Rhd catalytic sites (reviewed by Bordo and Bork 2002). The substrate specificity of Rhds appears to be defined by the length of the active-site loop. Forlani et al. (2003) transformed the substrate group specificity from SO_4^{2-} to P_i by elongating the six-amino acid loop into a seven-residue loop. This suggests the possibility of a subclass of Rhds that could have a loop conformation that binds P_i with higher affinity, but further research is required on the function of Rhd family proteins.

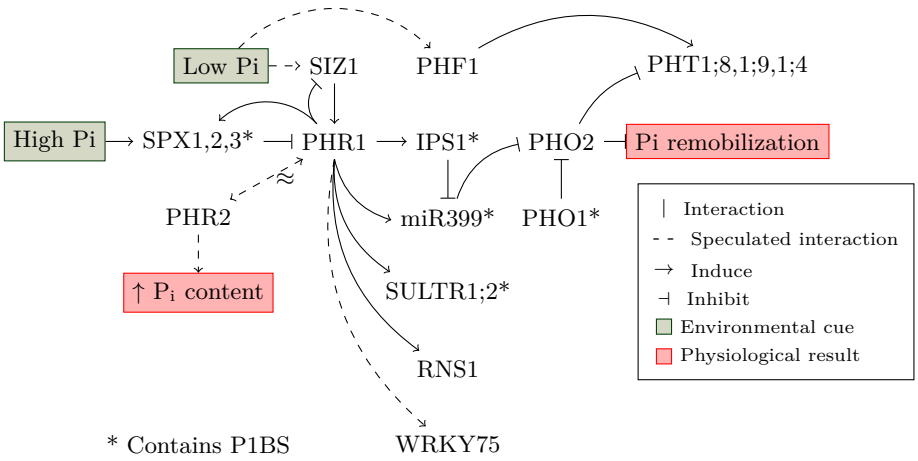

FIGURE 1.1: Model summarizing the interactions of a subset of genes implicated in a generalized plant PSR pathway

1.4 Yukon *Eutrema salsugineum* thrives in an extreme environment

The halophytic *E. salsugineum* (saltwater cress) is a highly stress-tolerant member of the *Brassicaceae* family that is able to withstand a number of abiotic stresses, including cold, high salinity, drought and nutrient deficiency (Griffith et al. 2007; Inan et al. 2004; Velasco et al. 2016; Wong et al. 2005). Many Pi-starvation studies utilize plants that typically do not thrive under stress conditions in the natural environment. Plants such as *A. thaliana* exhibit severe responses to Pi-starvation that results in stunted growth and expression of phenolic compounds such as anthocyanin (Plaxton and Carswell 1999). Difficulty in extracting RNA from tissue of stressed plants with high phenolic content in addition to decreased tissue yield presents a great challenge for consistent sampling for transcriptomic studies (Salzman et al. 1999). There is no perfect plant model organism to study every physiological response, however a case can be made for *E. salsugineum* as a candidate for studying abiotic stresses. As a close relative to *A. thaliana*, *E. salsugineum* also shares many traits with *A. thaliana* such as size, seed yield, genome ploidy, life cycles and a fully sequenced gnome (Oh et al. 2013; Yang et al. 2013a). The reference genome of *E. salsugineum*, estimated at 241 Mbp, is 2× larger than the genome of *A. thaliana* and has a gene sequence homology of 87.7% when aligned against the *A. thaliana* genome (Yang et al. 2013a). Many *A. thaliana* resources are available to study *E. salsugineum*

because of the sequence homology between the two species. *E. salsugineum* also becomes an important plant model for investigating Pi-use efficiency of agriculturally significant crops because of its higher relatedness to canola and other mustard crops compared to *A. thaliana*. Previous work by Velasco et al. (2016) showed that *E. salsugineum* ecotypes originating from the Yukon have high tolerance to Pi-deficiency. Contrary to the typical stress response manifested by *A. thaliana* to P_i deficiency (primary root growth suppression, increased lateral root hair and lateral root density), *E. salsugineum* seedling root morphology is largely unaffected by the lack of Pi, only exhibiting lower P_i levels in shoots and a minor increase in root mass (Velasco et al. 2016). Real-time quantitative PCR by Velasco et al. (2016) also showed that *E. salsugineum* plants showed increased expression of *EsIPS2* under low Pi, indicative of a response to the lack of Pi. However, the authors also reported no increased expression of *EsWRKY75* and *EsPHR1*, while their homologs in *A. thaliana* plants grown under comparable low P_i conditions showed differential expression. The disparate low Pi-responsive behaviour of *E. salsugineum* relative to *A. thaliana* offers a distinct perspective with how plants cope with low Pi, making *E. salsugineum* a favourable stress tolerance model to study for identifying low P_i tolerance traits. The current publicly-available reference genome of *E. salsugineum* was prepared using an ecotype originally collected in Shandong, China and the genome sequence was released in 2013 (Yang et al. 2013a). This project, however, focuses on the traits of the Yukon accession of *E. salsugineum* whose natural habitat would be considered "extreme" for commonly studied model plants: freezing temperatures, high salinity, low P_i content of 10-13 ppm and high sulfur content of more than 6500 ppm (Jacobson and Birks 1980; Velasco 2017). The Yukon accession of *E. salsugineum* expresses P_i stress-induced genes under well-fertilized conditions at a comparable level to a Pi-stressed *A. thaliana* plant (Velasco et al. 2016). Understanding how Yukon *E. salsugineum* copes with low P_i would identify traits needed to improve P_i-use efficiency in agriculturally significant crops, a goal that would reduce fertilizer use and help mitigate the associated environmental problems and looming rock P_i shortage.

1.5 Research objectives

This study aims to determine whether global transcriptomic reprogramming in Yukon *E. salsugineum* leaves occurs under low P_i growth conditions. Previous research reported by Velasco et al. (2016, 2020) suggests that genes responsive to low P_i in *A. thaliana* are likely constitutively expressed in leaves of the Yukon *E. salsugineum* ecotype. The outcome of finding few genes differentially expressed under low P_i conditions would be consistent with the conclusion of Velasco et al. (2020), namely that low Pi-induced traits are adaptive but fixed in this ecotype and not plastic as in the case for other plants studied to date. However, in the cited work of Velasco et al. (2016), the sample size of expressed genes reported did not include lncRNAs and novel annotated transcripts. This thesis represents a more comprehensive view of global gene expression in Yukon *E. salsugineum* shoots grown under low P_i conditions.

Plants growing in the Yukon are found on soil with low P_i availability but extremely high S

levels (Guevara et al. 2012). As such, a requirement for extra soil S was considered in the design of this study to ensure that plants were not undergoing multiple nutrient deficiencies. This consideration was particularly important because previous work in the Weretilnyk lab determined that plants grown in climate-controlled cabinets experienced deficiencies related to S even when the plants were regularly fertilized (Garvin 2016). This discovery of a S deficiency in cabinet-grown plants was unexpected as Yukon *E. salsugineum* plants do not exhibit classic S deficiency symptoms such as significantly reduced growth and biomass or chlorosis. However, to address the possible low S stress, a combination of plant nutrient treatments was performed by Dr. Vera Velasco and Ms. Amanda Garvin to generate suitable plants for the RNA-Seq work reported in this thesis. It is important to note that the soil mix used to grow the plants for this experiment contains some level of P_i and S but neither nutrient is in sufficient quantity so plants express genes related to P_i and sulfur (S) deficiencies in unmodified soil preparation (Garvin 2016; Velasco et al. 2016). By including preparations with high vs. low ${SO_4}^{2-}$ and Pi, the objectives for this thesis were the following:

- To identify genes, including patterns of expressed genes, in leaves of Yukon *E. salsugineum* that are responsive to low P_i under conditions when S is not limiting
- To identify genes, including patterns of expressed genes, in leaves of Yukon *E. salsugineum* that are responsive to low P_i under S-limiting conditions that elicit expression of S deficiency-related genes
- To determine, using CREMA (Simopoulos 2019), whether lncRNAs contribute significantly to the altered gene expression of plants experiencing low P_i and/or low S stress

Given the modest phenotypic changes reported for Yukon *E. salsugineum* experiencing low P_i, we predicted few changes in gene expression related to low P_i conditions. Moreover, we predicted that genes responsive to low P_i and low S would likely overlap given the associated gene networks regulating these deficiencies. Given our expectation of overlap between P_i- and S-responsive gene networks, we anticipated seeing the same genes differentially expressed under low P_i, low ${SO_4}^{2-}$ or combined low P_i/low ${SO_4}^{2-}$ soil conditions.

Chapter 2

Methods

2.1 Plant growth and materials

Plants of single-seed descent lines originating from a field Yukon *E. salsugineum* plant were sterilized and grown in a climate-controlled growth cabinet (Champigny et al. 2013) with a low-nutrient soil mix formulated by Velasco et al. (2016). The soil mixture was either used without modification or with 6.6 parts (w/v) $Ca_2SO_3 \cdot 2\,H_2O$ for a high S treatment. Plants were watered daily with de-ionized water. Beginning at two weeks post germination (wpg), plants were fertilized twice weekly with a fertilizer formulation (Velasco et al. 2016) modified by replacing KH_2PO_4 with KCl and/or Mg_2Cl depending on the treatment to emulate a low supply of P_i and S, respectively (**Table 2.1**). Plants were harvested at 4 wpg, three biological replicates per treatment were harvested and rosette leaf tissue from individual plants was separately flash-frozen in liquid N_2 and stored at $-80\,°C$ for RNA isolation, complementary DNA (cDNA) preparation and RNA sequencing (RNA-seq).

Table 2.1: Sample nomenclature and nutrient treatment regime

Symbol	Treatment		Description
	Pi	S	
ps	low Pi	low S	no added Pi or S
pS	low Pi	added S	no added Pi and 2.5 mM S (5000 ppm $CaSO_4$)
Ps	added Pi	low S	2.5 mM Pi and no added S
PS	added Pi	added S	2.5 mM Pi and 2.5 mM S (5000 ppm $CaSO_4$)

2.2 RNA extraction and sequencing

For this study, all steps leading to the assembly of the transcriptomes were performed by Dr. Vera Velasco and Ms. Amanda Garvin following the hot borate method (Wan and Wilkins 1994) detailed in Champigny et al. (2013). RNA quality assessment and quantification were completed using RNA Nano 6000 chips on a Bioanalyzer 2100 platform. Sequencing was carried out at the Farncombe Family Digestive Health Research Institute (McMaster University, ON, Canada). High-throughput paired-end runs were performed with an Illumina Hi-Seq 1500 platform.

2.3 Data processing, normalization and mapping

Following sequencing, the reads were assessed for read quality with FastQC v0.11.9 (Andrew 2010) and trimmed with Trimmomatic v0.34 (Bolger et al. 2014) to remove low quality reads with <30 Phred quality score. Phred scores are calculated during sequencing by determining base call peak parameters (Richterich 1998). The trimmed reads were mapped to

the reference genome assembled by Yang et al. (2013a) using the read aligner STAR v2.5.2b (Dobin et al. 2013) using the two-pass method suggested by Engström et al. (2013). The first pass-through generates an alignment file along with a splice junction file containing all identified splice junctions in each read, which was used to guide the final alignment during the second pass. The current reference genome of *E. salsugineum* was assembled from Shandong ecotype plants and there is currently no consensus genome for the Yukon accession. To increase the reliability of the mapped reads, the STAR output alignment files were examined using SAMtools v1.3.1 (Li et al. 2009) and reads with a unique mapping were separated from the multiple-mapped reads using SAMtools (Li et al. 2009). Only uniquely mapped reads as determined by the STAR read aligner were retained and used in downstream analyses. The annotation file used for alignment was downloaded from the v12 release of the Joint Genome Institute's plant portal Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html, accessed May 2019). Novel assembled transcripts that were not in the Phytozome v12 annotation were added by Champigny et al. (2013) and Simopoulos (2019).

Transcript abundances were determined using the QoRTs tool-kit v1.3.6 (Hartley and Mullikin 2015) and transcript read counts were normalized to the effective library size using the `DESeq2` Bioconductor package for R (Love et al. 2014). Genes with a total raw gene count of less than 10 across all libraries were removed during a minimal filtering step. This step was performed to improve statistical reliability of the DESeq calculations by removing reads with close to zero reads and also to reduce memory use and improve computing time for `DESeq` transformations. The preliminary removal of genes with low raw gene counts does not affect the resulting transformed data because genes with low read counts are unlikely to show evidence of significance and their removal increases differential expression detection power by `DESeq`. `DESeq` normalization applies a strict false discovery rate (FDR) filtering on the normalized mean counts and the normalization also accounts for transcript abundances of genes with high expression (Love et al. 2014).

2.4 Multivariate analysis

Transcript abundance was reported as fragments per kilobase per million mapped reads (FPKM). The raw read counts were normalized using the estimated size factors of each replicate library generated by `DESeq` as well as the median transcript length considering all RNA-seq libraries. Statistical analysis on transcript abundance among the transcriptomes was performed using R v3.6.0 (R Core Team 2013). FPKM values were log-transformed ($\log_2(FPKM + 1)$) and the mean FPKM was calculated from three biological replicates of each treatment. Principal Component Analysis (PCA) was performed on the co-variance matrix of mean FPKM values for all expressed genes between the transcriptomes of all four treatments .

2.5 Long non-coding RNA prediction

Long non-coding RNA prediction of Yukon *E. salsugineum* transcripts was performed using CREMA software (Simopoulos et al. 2018). Using gffread v0.10.8 (Pertea and Pertea 2020), transcripts that were mapped to the *E. salsugineum* reference genome including novel annotated genes (Simopoulos 2019) were extracted from the STAR sequence alignment files. Known protein-coding reads were identified using Diamond v0.9.22 (Buchfink et al. 2015) and coding potential of reads that have not been identified as protein-coding were assessed using CPAT v1.2.1 (Wang et al. 2013). Assembled transcript sequences were input into CREMA along with the output from Diamond and CPAT to generate lncRNA predictions. The default prediction cutoff of 0.5 was used as there was no difference in the number of predicted lncRNAs with cutoff scores of 0.05, 0.1 or 0.5.

2.6 Differential expression analysis

Genes showing differential expression between any two normalized transcriptomes in a pairwise comparison were identified using `DESeq2` (Love et al. 2014). In each pairwise comparison, one library in the pair was set as the basis level and deemed the "control" sample or the denominator from which the second library, referred to as the numerator, was compared against. The $\log_2$ fold change (LFC) p-values were calculated using the normalized read count of each transcript while considering the individual treatment conditions or Pi/S presence/absence as the regression model. Additionally, the LFC p-values were adjusted by an FDR correction with α of 0.1.

2.7 Enrichment of gene ontology terms

Gene Ontology (GO) annotation was used to assess the representation of different metabolic mechanisms in a group of genes of interest. GO terms were obtained from Joint Genome Institute (JGI) Phytozome v12 annotation with additional annotations added by Champigny et al. (2013) and Simopoulos (2019) based on reciprocal best BLAST hits against the *A. thaliana* transcriptome annotation. GO term enrichment of significant differentially expressed genes (DEGs) between pairwise comparisons were obtained through the `TopGO` v3.11 (Alexa et al. 2006) R package with a FDR threshold of 0.05. Semantically redundant terms were removed with `ReviGO` webtool vJan.2017 (Supek et al. 2011), resulting in a ranked list of GO terms with frequency values and GO term uniqueness: the negative similarity, or the semantic dissimilarity, of a term compared against other terms.

2.8 Data collection of non-*E. salsugineum* plants

Data of total expressed genes and Pi-responsive DEGs in *Triticum aestivum L.* (Oono et al. 2013a), *Glycine max* (Guo et al. 2008b; Zhang et al. 2017), *Oryza sativa* (Oono et al. 2013b; Park et al. 2012; Secco et al. 2013), *Plantago major* (Huang et al. 2019), *Avena sativa* (Wang

et al. 2018), *Zea mays* (Calderon-Vazquez et al. 2008), *Lupinus albus* (O'Rourke et al. 2013), *Hordeum vulgaris L.* (Ren et al. 2018), *Brachypodium distachyon* (Zhao et al. 2018) and *A. thaliana* (Liu et al. 2016) were collected based on the reported values in their respective publications. For comparison of select genes of interest in *E. salsugineum*, FastQ format raw transcriptome data from Liu et al. (2016) were downloaded from the NCBI Sequence Read Archive (SRA) (Leinonen et al. 2010) using `fasterq-dump` from the `SRA toolkit` v2.9.2. The SRA transcriptome data were mapped and processed using the STAR method described above. The *Arabidopsis* Information Resource (TAIR) v11 *A. thaliana* transcriptome annotation was used for mapping the SRA reads to the *A. thaliana* transcriptome. Orthologous genes between *E. salsugineum* and *A. thaliana* were obtained from the Phytozome v12 annotation of the *E. salsugineum* transcriptome. Novel *E. salsugineum* transcripts with no annotation for *A. thaliana* orthologs were compared against the *A. thaliana* transcriptome using the best BLAST hit approach. A heatmap of $\log_2$ FPKM values was generated for 53 genes of interest including three housekeeping genes in *E. salsugineum* and *A. thaliana* with hierarchical clustering performed on each transcriptome based on Euclidean distance.

2.9 Phylogenetic analysis

To generate the species tree, sequence data of *O. sativa*, *Z. mays*, *C. rubella*, *A. thaliana*, *B. rapa*, *B. napus*, *S. lycopersicum* and *M. truncatula* were downloaded from the NCBI RefSeq database (O'Leary et al. 2016). The reference genes chosen for generating the species tree were highly conserved housekeeping genes which are 16s *ribosomal RNA (rRNA)*, 18s *rRNA*, *40S RIBOSOMAL PROTEIN S16 (Rps16)*, *ATP SYNTHASE SUBUNIT β (ATP2)* and *STRUCTURAL MAINTENANCE OF CHROMOSOMES 1 (*Smc1*)*. MAFFT v7.205 (Katoh et al. 2002) was used with the optimized default parameters to perform alignment for both the reference sequences as well as transcript sequences for *IPS1*. The tree topologies were generated using RaxML v8.0.25 (Stamatakis 2014) using the rapid hill-climbing algorithm and the generalized time reversible substitution model with a gamma model for rate heterogeneity.

Chapter 3

Results

3.1 Characterization of the putative *IPS1* in *E. salsugineum*

To determine whether the sequence of the putative *IPS1* in Yukon *E. salsugineum* (annotated as *XLOC_008023*) corresponded to *IPS1* in other species, the transcript sequence of the putative *IPS1* was aligned with *IPS1* sequences from five other members of the *Brassicaceae* family, which are: *A. thaliana*, *Arabidopsis lyrata*, *Capsella rubella*, *Brassica rapa* and *Brassica napus* (**Figure 3.1 A**). The 22 bp conserved region of *IPS1* found in the other *Brassicaceae* members was also observed in the putative *E. salsugineum IPS1*. Bootstrap analysis of the *IPS1* alignment showed a tree topology that agrees with the species tree inferred from conserved gene alignments (**Figure 3.1 B, Figure A1.1**).

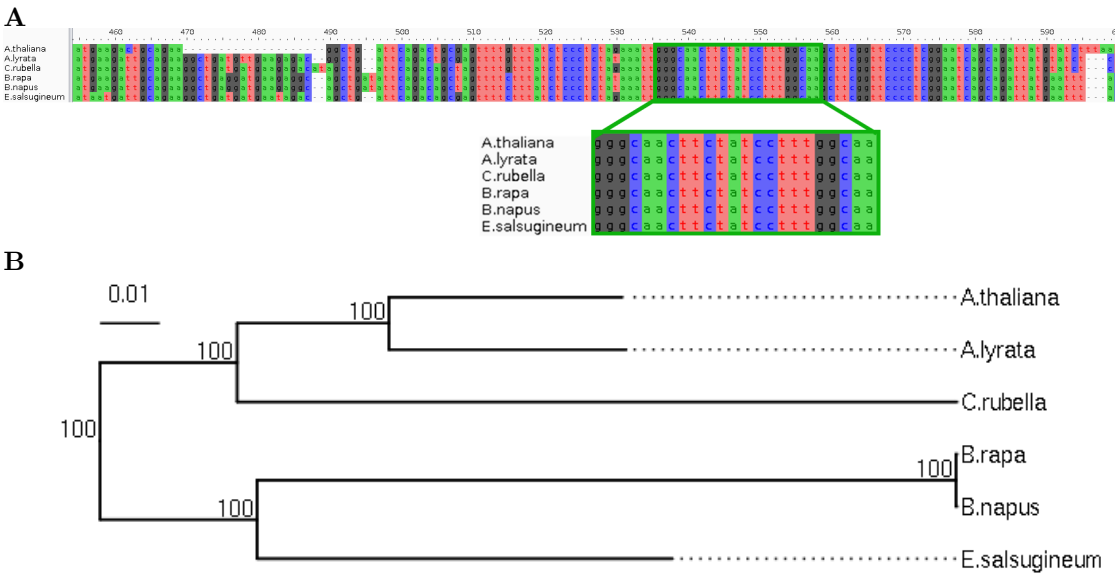

FIGURE 3.1: **Evidence identifying the unannotated locus *XLOC_008023* as the gene encoding a putative *EsIPS1*.** **A** alignment of the 22 bp conserved sequence and flanking regions of *IPS1* from *A. thaliana*, *A. lyrata*, *C. rubella*, *B. rapa*, *B. napus* with the putative *IPS1* in Yukon *E. salsugineum*. **B** *IPS1* gene tree generated from 1000 bootstraps, branch lengths were calculated with maximum likelihood estimation using a general time reversible substitution model.

In addition to the conserved 22 bp region of *IPS1*, a P1BS was observed at 470 bp upstream and a TATA-box element was observed at 26 bp upstream of the transcription start site (**Figure 3.2**). The P1BS is observed in the upstream regions of *IPS1* in other plants such as *A. thaliana* and *B. napus* (Sobkowiak et al. 2012; Yang et al. 2013b) and PHR1 is known to

```
CATATTCCATGAGATGATCCAAAATTCCAAGCATATGCACGTAATACTCCCCTATTTTCCATCATCTGCTTCCACTTATATATATGTATATATATATATA
                                ******** -470 : P1BS
TGTGTGAATTTTCAAAGTACTTTCAAACTCTCTTTATACCATATTTCATTATTTTAAATTATTCACCTTTAAGCAATATATCTCCATTTTTTCTCCTTTC

ACAAATTTTATCGTAAAGTCCTTACAAAAATACCTTTCAGAACCTTATACCAGGAGGCTTTAGCCTAAACTCGTACATTGGAGAAATTTTCAGATATTAT

TTATTGTGTAATTCAGCATGCAAAATTTTGTAAATACAGAATAAATGCGTGTATATTATTCCAAAACATAGCGTCCTGCATCGATCTATTAACAATATAT

ATATACGAGTATATCTTTTATAATGAACATTGACAAACATAAAATATATTAAACCATTGTTACAAAGAGAAAAATATATATATTGGGTGATTTGTTACAA
                                                                                        ******** -26 : TATA box
AAGTTGTCTTTTAAAAAATGAAAAAGAAAGGAGAAAAAGTCGAAAGAACACTATAGCTTATGTAAGATGTCTTCATGACCATAGACAAAGAAAAAACACA
* +1 : Transcription start site
AAAGAAAAATAAAAGCTATAAAAACCCCTAGCAAGCCTCCATTGATGAGACCTCTCTTAATTTGGCACAACACCACAAATTAAAAGAAAATGGCCATCCC

CTAAGGTAAAGATCCTCTGATTTATCTAGAGGTGTATCTTTAGGGGATGGCCTAAATGCAAGAAAAAAAAAAAGGAAAATGTGTAGTTAAGTGGTTTTGT

GTTCTCGTAAGGAAAATGTTTTAAGATATGGAGCAATAATGATTGCAGAAGGCTGATGATGAATAGACAGCTGATTCAGACAGCGAGTTTTCTTTATCTC

CCTCTAGAAATTGGGCAACTTCTATCCTTTGCAAGCTTCGGTTCCCCTCGGAATCAGCAGATTATGAATTTACTTTGTAATACTCTCTCTCTCACTCTC
            ********************** +413 : 22 bp conserved region
TCTGATCCTTTTCTCTTCTCTATGTTTTGTTTATCTTCATTTTGTGTGTGTTTTAACTGTATCAACTCCCATGGTATGTTCT
```

FIGURE 3.2: Upstream elements of the putative *IPS1* (annotated as *XLOC_008023*) in Yukon *E. salsugineum* has a PIBS and TATA box elements.

directly induce the expression of *IPS1*. The presence of the P1BS promoter suggests probable interaction between PHR1 and the putative *IPS1*.

3.2 Transcriptome profiling of shoot response to P_i and S deficiency

Transcriptome profiles between Yukon *E. salsugineum* plants were compared to explore the effects of P_i on gene expression and transcriptome remodelling while considering the presence of S in the soil. RNA-seq was performed on cDNA prepared from RNA extracted from leaves of 4-week-old Yukon *E. salsugineum* plants grown under the double-nutrient treatment regimes with three replicates per treatment. Sequencing of each replicate produced 8.6-12.6 million paired-end reads of Phred score quality ≥ 30 in 11 out of the 12 replicates and 1.8 million paired-end reads with ≥ 30 Phred score in `ps12_S3`. Despite the low read count in 1 of 12 replicates, read mapping of the replicates showed a unique read mapping ratio of 99.52 $\pm$ 0.07 % among all 12 replicates (**Table 3.1**). The high read mapping ratio assures the reliability of `ps12_S3` reads for subsequent analyses, as more than 99% of the reads were mapped to the *E. salsugineum* reference genome (Yang et al. 2013a). Only uniquely mapped reads were considered as raw counts and are retained for downstream expression analysis. The reads in each replicate were mapped to approximately 23,578 of 28,885 genes or approximately 81.2% of genes in the *E. salsugineum* reference genome published by Yang et al. (2013a) including additional novel genes annotated by Simopoulos (2019) (**Table 3.2**). Transcript FPKM values were calculated from normalized reads of each replicate grouped with their respective treatments and total gene expression per gene per treatment was obtained from the mean FPKM value per transcriptome ($n = 3$). Genes with mean FPKM > 0 in at least one library were considered expressed. Of the unique total expressed genes among samples, 20,774 of

23,578 or 88.1% were expressed commonly between all four transcriptomes and 1,180 of 23,578 or 5.0% genes were only expressed in one library (**Figure 3.3**).

TABLE 3.1: Summary of paired-end reads and mapping counts for RNA-Seq libraries

Library ID	Treatment	# Quality PE* reads	% Uniquely mapped reads
ps12_S3	-Pi -S	1,811,886	99.31
ps48_S10	-Pi -S	12,522,875	99.56
ps49_S11	-Pi -S	9,160,586	99.49
ps11_S2	+Pi -S	8,685,209	99.53
ps40_S6	+Pi -S	8,902,553	99.61
ps46_S9	+Pi -S	12,566,038	99.56
ps10_S1	-Pi +S	11,916,816	99.49
ps38_S5	-Pi +S	9,966,796	99.61
ps44_S8	-Pi +S	11,843,493	99.55
ps37_S4	+Pi +S	9,853,915	99.61
ps41_S7	+Pi +S	12,116,544	99.52
ps50_S12	+Pi +S	8,163,181	99.58

* Paired-end

TABLE 3.2: Number of annotated and unannotated expressed genes identified in transcriptomes of P_i and S-treated plants

Name	Treatment	Expressed genes		
		Annotated	Novel	Total
ps	-Pi -S	20,299	1,660	21,959
Ps	+Pi -S	20,450	1,684	22,134
pS	-Pi +S	20,373	1,663	22,036
PS	+Pi +S	20,530	1,686	22,216

PCA was carried out to visualize the variance in transcript abundances between treatments (**Figure 3.4**). The first four Principal Components (PCs) accounted for over 99.9% of variance between the transcriptomes, with PC1 accounting for 99.31% of the variance. However, the rotations for all four transcriptome profiles on PC1 showed little difference between treatment transcriptomes as seen in the biplot visualizations of PC1 (**Figure A1.2**). Rather, PC2, representing 0.49% of variance between the four treatment transcriptomes, was able to distinguish between transcriptomes produced by plants undergoing high vs. low ${SO_4}^{2-}$ treatment. The biplot in **Figure 3.4 A**) provides arrows that position the scores associated with global transcript abundance in plants subjected to each of the four P_i and ${SO_4}^{2-}$ treatment combinations used in this study. In the biplot of **Figure 3.4 A**, PC2 separates transcriptomes of plants grown with low vs. high ${SO_4}^{2-}$ with scores positioned positive to the origin corresponding to transcriptomes of plants grown in soil where no additional ${SO_4}^{2-}$ was added (s) while scores negative to the origin correspond to those of plants given the high ${SO_4}^{2-}$ treatment (S). Conversely, PC4 accounts for approximately 0.08% of the variance but the scores associated with transcriptomes of plants given P_i (P) or not (p) were clearly separated

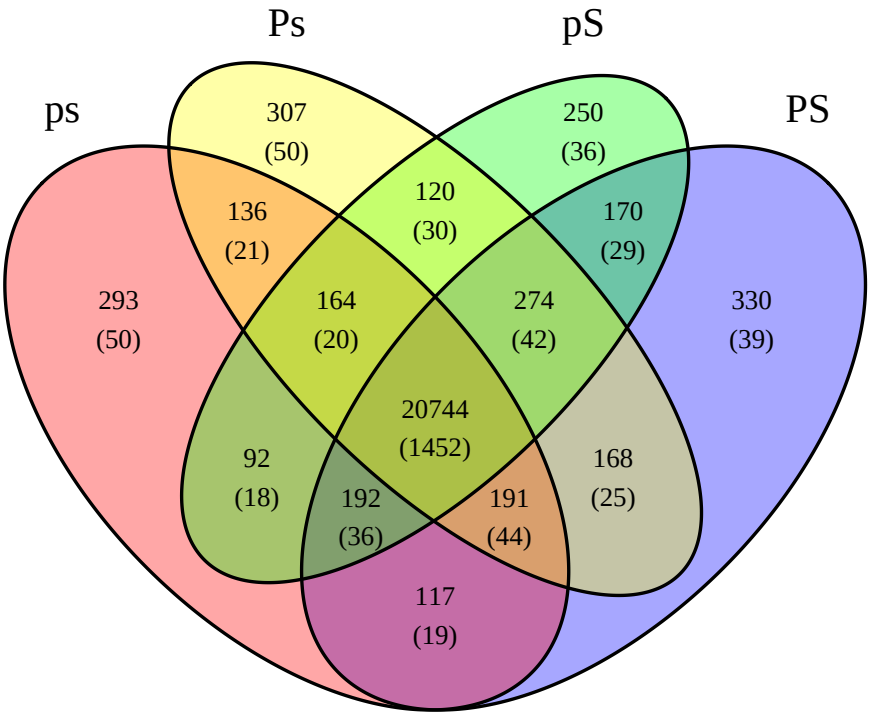

FIGURE 3.3: **Venn diagram showing the number of expressed genes, including novel genes, in each treatment combination** Normalized read counts were grouped by treatment, raw counts mapped or novel genes were converted to FPKM values and mean FPKM per gene per treatment was calculated to generate the Venn diagram ($n = 3$ biological replicates per treatment). Genes with mean FPKM > 0 in at least one library were considered expressed and the number of novel genes are shown in parentheses.

and positioned positive or negative to the origin of this axis, respectively. GO term enrichment was performed on the top 200 genes contributing to the most positive or negative factor loadings for each axis on PC2 and PC4. GO enrichment of the top-contributing genes in each PC can provide an overview of genes and mechanisms that are expressed in association with the P_i and/or S status of the plant. GO terms from the top genes contributing to PC2 reflected various mechanisms associated with S status, namely sulfur compound metabolism, S-glycoside metabolism, glycoside metabolism, glucosinolate metabolism, responses to wounding, oxidative stress and biotic stress (**Figure 3.5** and **Table A1.1**). In contrast, few of the GO terms associated with the top genes contributing to the PC4 loading have direct links to P_i metabolism with exceptions being P_i ion homeostasis and anionic inorganic ion homeostasis.

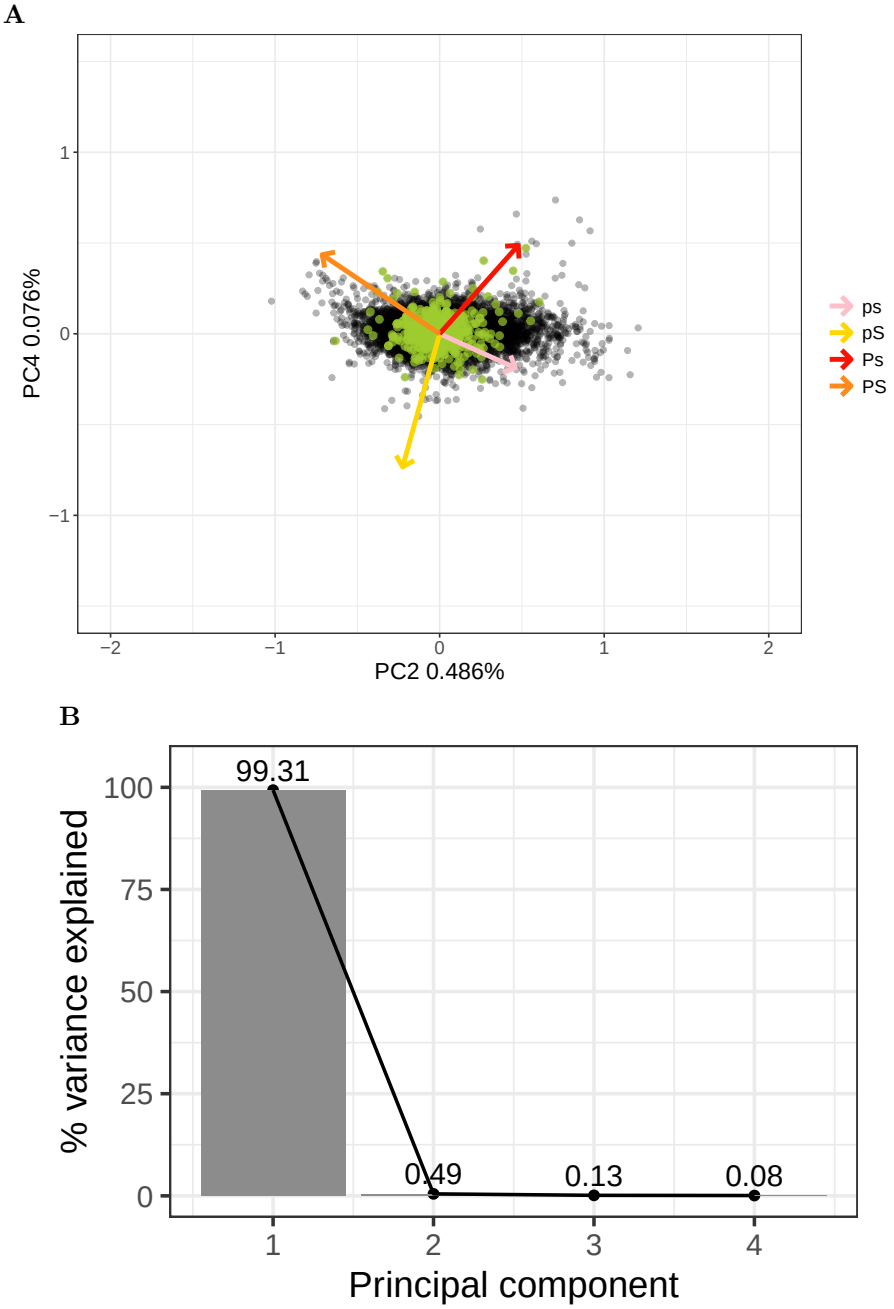

FIGURE 3.4: **PCA of PS, ps, Ps and pS transcriptomes show negligible variability beyond PC1.** PC2 and PC4 biplot (A) and scree plot for PCA (B) are shown. PC1 represents 99.31% of the variance between treatment transcriptomes, PC2 represents 0.49%, PC3 represents 0.13% and PC4 0.08%. The variations in SO_4^{2-} treatment and P_i treatment are defined by PC2 and PC4, respectively. ● = Expressed genes, ● = CREMA-predicted lncRNA

3.3 Differential expression analysis shows few low P_i-responsive genes in Yukon *E. salsugineum*

Pairwise comparisons between the expressed genes of the four treatment transcriptomes generated from the Yukon *E. salsugineum* plants in this study enable identification of genes that were differentially expressed between the four fertilizer treatments. The detailed break down of each pairwise comparison is listed in **Table 3.3**. In comparing the pair Ps vs. PS, the treatment comparison being made corresponds to a difference in ${SO_4}^{2-}$ while P_i was present and under these conditions, 265 and 49 genes were up-regulated and down-regulated, respectively. In the pS vs. PS comparison, high ${SO_4}^{2-}$ was present in the soil used for both treatments but P_i content was different and between transcriptomes associated with these treatments the expression of only two genes was significantly up-regulated and none down-regulated. The two up-regulated genes in the pS vs. PS comparison correspond to *Thhalv10018786m.g* and *Thhalv10024976m.g*, the former a S-adenosyl-L-methionine dependent carboxyl methyltransferase and the latter a cytochrome P450-like unknown protein that were up-regulated at 1.56 and 2.23 $\log_2$ fold, respectively. When ${SO_4}^{2-}$ was not added to the soil but additional P_i was present, the comparison between ps vs. Ps transcriptomes shows no genes with up-regulated expression but 11 genes were down-regulated.

In **Table 3.3**, two pair-wise comparisons between the transcriptomes consider the influence on gene expression when both nutrients were altered differentially (hence ΔPS). In the comparison between pS as the numerator vs. Ps as the denominator, transcriptomes were compared between plants not given P_i but were provided with additional ${SO_4}^{2-}$ relative to plants given P_i but no additional ${SO_4}^{2-}$. It follows from this comparison that 11 genes had up-regulated expression and 127 genes were down-regulated between transcriptomes of pS relative to Ps plants. The second condition, which is a bit easier to interpret, involves plants treated with low P_i and low ${SO_4}^{2-}$ (ps numerator) relative to plants that were provided with both nutrients (PS denominator). Interestingly, in this ps vs. PS comparison, only 20 DEGs were identified with 17 up-regulated and 3 down-regulated genes, respectively.

To summarize the above results, the greatest number of DEGs listed in **Table 3.3** were identified in comparisons involving S vs. s (hence ΔS) whether the level of P_i was varied or not as assessed by ΔP comparisons. This impression becomes more clear when data from the transcriptomes were combined in such a way that the influence of soil ${SO_4}^{2-}$ was either included or discounted from the pair-wise comparative analysis. Alternatively stated, one can remove the influence of ${SO_4}^{2-}$ addition on plants transcriptomes by combining the expression data from pS and ps transcriptomes and comparing that pooled dataset to the combined dataset of PS and Ps transcriptomes. In this pooled comparison involving p vs. P, only a single DEG was detected corresponding to the novel transcriptome *DROUGHT.22069*. Sequence alignment of *DROUGHT.22069* with the NCBI database showed a 46% DNA sequence identity to an uncharacterized lncRNA in *Brassica oleracea* and *Brassica napus*. *DROUGHT.22069* had a -21.7 $\log_2$-fold change in p compared to P treatments (p-value $= 1.86 \times 10^{-12}$). Conversely, when the abundance values for transcripts were combined to reflect the difference in soil ${SO_4}^{2-}$

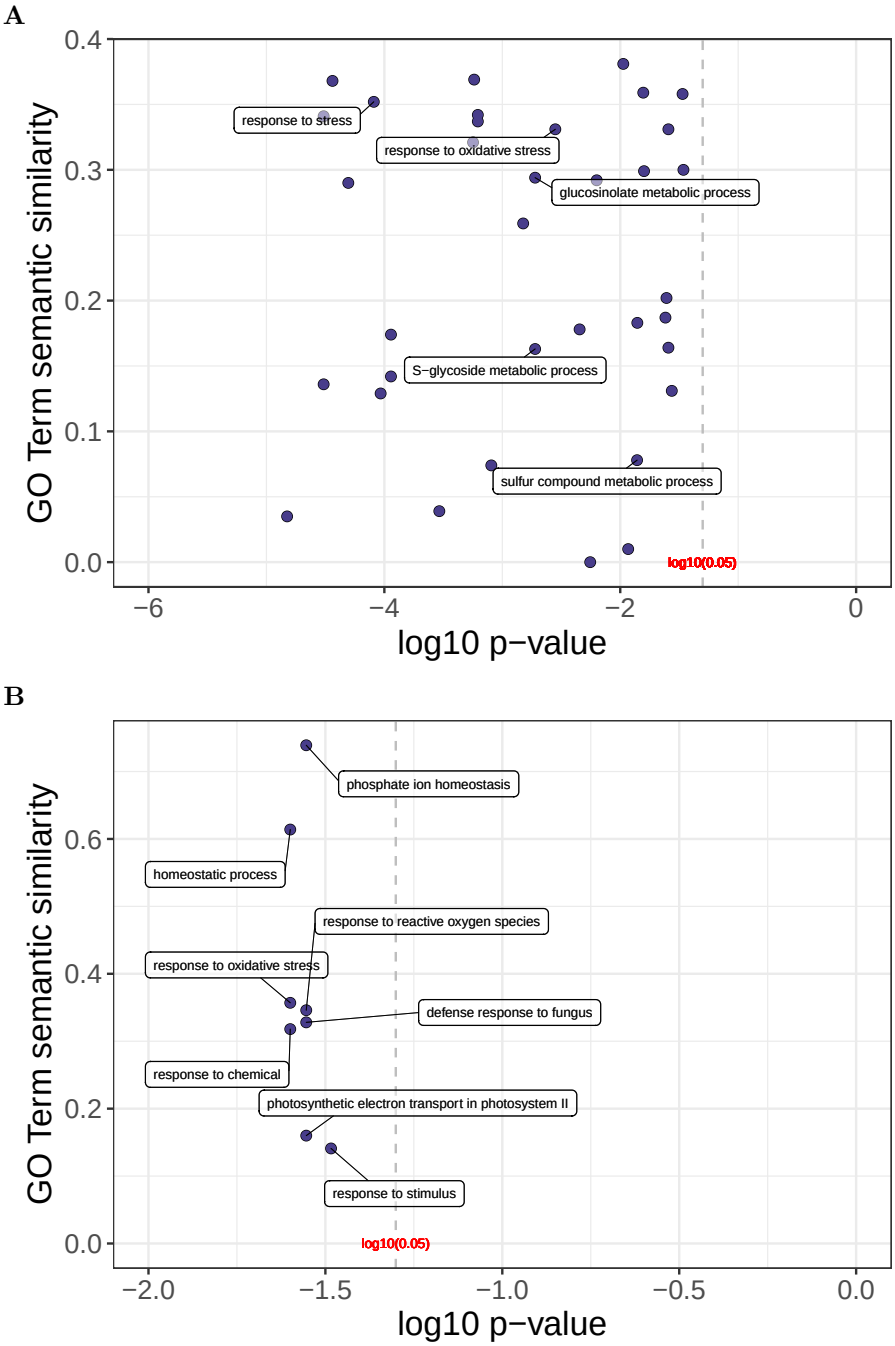

FIGURE 3.5: **GO term enrichment of top 200 genes contributing to the positive and negative axes of PC2 and PC4 biplot.** Terms associated with PC2 (A) and PC4 (B) are shown along with associated $\log_1 0$ *p*-values. PC2 terms show multiple direct connections with S metabolism while PC4 is associated with fewer terms directly related to P_i (P_i-related ion homeostasis and trivalent inorganic anion homeostasis).

content only (*i.e.* (P/p)s vs. (P/p)S), there were 288 and 83 up- and down-regulated genes, respectively, that were S-responsive. Thus whether by isolating the effects of SO_4^{2-} from *Pi* or not, it is clear that altering the SO_4^{2-} content of the soil had a greater influence on gene expression than did an altered P_i nutrient regime.

TABLE 3.3: Differentially expressed genes and predicted lncRNAs per library in parentheses

Denominator	Numerator	Comparison	Up-regulated	Down-regulated
PS	Ps	ΔS	265(5)	49(0)
	pS	ΔP	2(0)	0(0)
	ps	ΔPS	17(0)	3(0)
Ps	pS	ΔPS	11(1)	127(6)
	ps	ΔP	0(0)	11(1)
pS	ps	ΔS	73(5)	18(1)
P	p	ΔP	1(1)	0(0)
S	s	ΔS	288(10)	83(1)

GO terms associated with the pooled DEGs from all six pairwise comparisons highlight the metabolic pathways affected by the difference in both nutrients (**Figure 3.6**). These GO terms include sulfur compound metabolism, carbohydrate metabolism, glycoside metabolism and glycosinolate metabolism. Various response pathways were also affected, such as immune response to biotic stress, response to jasmonic acid, chemical stress, oxidative stress, biotic stress and wounding.

3.4 Long non-coding RNA prediction

To detect the presence of lncRNAs in the Yukon *E. salsugineum* transcriptomes, CREMA (Simopoulos 2019; Simopoulos et al. 2018) was used to predict lncRNAs among total expressed genes and DEGs. Of the 23,578 expressed genes among all four transcriptomes, 615 or 2.6% were predicted as lncRNAs (**Figure 3.7**). Of the predicted lncRNAs among all of the expressed genes, 30 were present in the top 200-loading genes that contributed either positively or negatively to PC2 or PC4. Of the 615 predicted lncRNAs, 346 genes were novel and previously identified by Simopoulos (2019) and based on Blast hits of each transcript against the NCBI nucleotide database (O'Leary et al. 2016), the other 269 genes present in the JGI annotation were genes with unknown function and putative genes. **Figure 3.7** shows that only one treatment-specific lncRNA is predicted and that is from the PS transcriptome provided by plants that are not expected to be experiencing nutrient deprivation. Indeed the vast majority of predicted lncRNAs appear to be expressed in all four treatment transcriptomes suggesting that the expression of this class of gene product is not low P_i or low S-responsive. The same

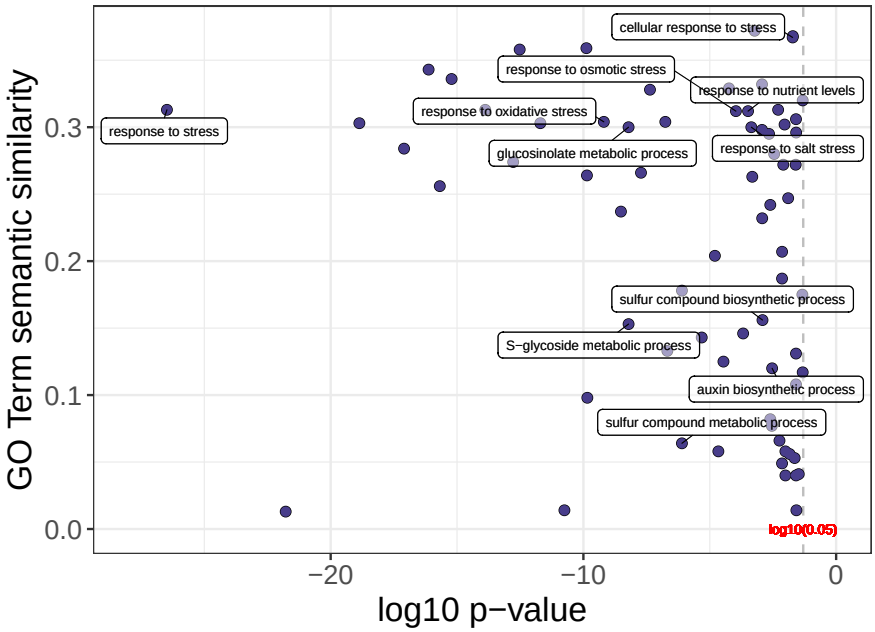

FIGURE 3.6: GO term enrichment of all differentially expressed genes pooled from all pair-wise comparisons. GO terms with associations to stress response, S metabolism and P_i metabolism are highlighted.

impression is provided by **Table 3.3**. For example, there were no predicted lncRNAs in the comparisons between ps and pS transcriptomes to that of PS and only five predicted lncRNAs when transcriptomes of Ps plants were compared to those of PS plants. The comparisons yielding the most DEG lncRNAs was seven in the paired comparisons of pS and Ps transcriptomes. If one simply considers comparisons that focus on the presence of P_i or ${SO_4}^{2-}$ in the fertilizer regime, there are still only 11 predicted lncRNAs out of 371 genes that were DEGs and that outcome was in response to a high vs. low ${SO_4}^{2-}$ treatment (s vs. S transcriptomes pooled and compared). Interestingly, the single DEG in the comparison that focused on P_i (p vs. P transcriptomes pooled and compared) was predicted as a lncRNA by CREMA

3.5 The Yukon *E. salsugineum* shows little evidence of low P_i-responsive reprogramming

Expression for a subset of PSI genes from the PS, Ps, pS and ps transcriptomes were compared to each other and to the expression of their orthologs reported in a published P_i starvation experiment for *A. thaliana* performed by Liu et al. (2016). This comparison, presented as a heatmap in **Figure 3.8**, was done to determine whether key PSI genes were differentially expressed between Yukon *E. salsugineum* transcriptomes in response to their P_i and ${SO_4}^{2-}$ treatments and to determine whether the low P_i response reported for *A. thaliana*

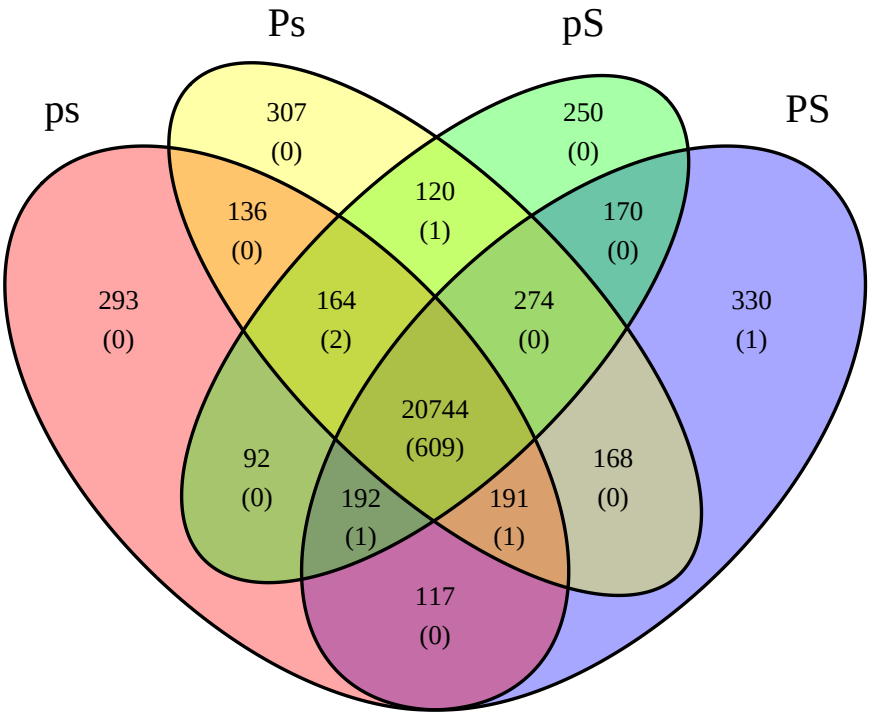

FIGURE 3.7: **Venn diagram showing the number of expressed genes, including predicted lncRNAs, in each treatment combination.** Normalized read counts were grouped by treatment, raw counts of transcripts that were mapped or predicted as lncRNA were converted to FPKM values and mean FPKM per gene per treatment was calculated to generate the Venn diagram ($n = 3$ biological replicates per treatment). Genes with mean FPKM > 0 in at least one library were considered expressed and the number of predicted lncRNAs are shown in parentheses.

was similar to the patterns found for the pS and ps transcriptomes analyzed in this study. In this study, the expression of housekeeping genes was used as internal comparisons and evidence that the treatments used did not perturb their consistent expression. However, although their expression was invariable with respect to transcript abundance between transcriptomes within each species, they were differentially expressed between the two species used in this comparison (averages of 10 and 6.3 $\log_2$ FPKM for Yukon *E. salsugineum EF1A* and *Act2*, respectively and 3.2 and 2.5 $\log_2$ FPKM for *A. thaliana EF1A* and *Act2*, respectively).

The heatmap in **Figure 3.8** showed that the expressed genes from the Yukon *E. salsugineum* transcriptomes cluster together with those from ps and pS-treated plants forming a clade with less similarity to the pattern of genes comprising the Ps-related group. A

notable feature of all the *E. salsugineum* genes used to generated the heatmap is that relatively few genes appear to be expressed differently across the four sources of the Yukon transcriptomes with, as stated above, the most deviation seen with the Ps-treated transcriptomes relative to those expressed in plants subjected to PS, pS or ps treatments. Another observation from **Figure 3.8** is that few of the expressed genes of *A. thaliana* showed similar patterns or baseline expression levels relative to their orthologs in *E. salsugineum* and, in general, expression was frequently higher in *E. salsugineum* relative to *A. thaliana* regardless of nutrient treatment. For example, PSI genes such as *PHO2* in Yukon *E. salsugineum* were often not differentially expressed between the four transcriptomes and their expression level was comparable to *PHO2* from *A. thaliana* following 3 d of low P_i stress. With respect to transcript abundance, *A. thaliana PHO2* expression went from 3.9 to 5.1 $\log_2$ FPKM between 0 d and 3 d of P_i starvation, respectively, which *PHO2* expression ranged from 7.3 to 7.9 $\log_2$ FPKM in the four Yukon *E. salsugineum* transcriptomes. However, not all *E. salsugineum* transcriptomes showed a lack of response to nutrient treatment. For example, the transcription factor *ZAT6* was P_i-responsive in *A. thaliana*, being up-regulated from 3.1 to 5.0 $\log_2$ FPKM between 0 and 3 d of P_i starvation while in Yukon *E. salsugineum*, *ZAT6* was already expressed at 5.9 $\log_2$ FPKM when P_i and ${SO_4}^{2-}$ was provided (PS treatment) but its expression was higher (about 7.0 to 7.6 $\log_2$ FPKM) for plants either lacking one or both P_i and/or S supplements (pS, ps and Ps). In addition, the expression of the classical P_i stress indicator *IPS2* in *A. thaliana* was up-regulated from 2.4 to 7.7 $\log_2$ FPKM between 0 and 3 d of P_i starvation while expression of this gene in Yukon *E. salsugineum* leaves was variable but modest compared to the *A. thaliana* data irrespective of treatment (0.03 to 0.6 $\log_2$ FPKM). Transcripts associated with *UGP1-3*, *MGD1*, *DGD1*, *PAH1* and *PHT5;1* provided a somewhat similar profile in that although they were all low P_i-responsive in *A. thaliana*, that was not the case for the expression of these genes in leaves of Yukon *E. salsugineum* where, as for the housekeeping genes, genes were consistently more highly expressed relative to their orthologs in *A. thaliana* in approximating the stress-response levels found in P_i-starved *A. thaliana*. A different outcome was found for the putative lncRNA *IPS1* which was not differentially expressed with low P_i in *A. thaliana* seedling shoots or in Yukon *E. salsugineum* leaves despite identification of a P1BS element in the upstream promoter region of the putative *IPS1* in the Shandong *E. salsugineum* genomic sequence (**Figure 3.2**). Interestingly, the genes associated with lipid remodelling under Pi stress such as *SQD1,2*, *DGD1,2*, *MGD2,3* and *UGP1-3*, were not differentially expressed in Yukon *E. salsugineum* unlike their orthologs that show differential expression in P_i-starved *A. thaliana*.

The nutrient treatments used for Yukon *E. salsugineum* in this study included manipulating S, this was not the case for the *A. thaliana* experiments performed by Liu et al. (2016). Accordingly, it is not surprising that there were few differences in the expression of S starvation-responsive genes between 0 and 3 d for P_i-starved *A. thaliana*.

Between any of the pairwise comparison of the four Yukon *E. salsugineum* transcriptomes, genes encoding the S transporters *SULTR4;2*, *SULTR3;1* and *SULTR2;1* were expressed higher in the Ps treatment than the three other transcriptomes. The low S-induced gene *SDI1*

was up-regulated in the S-deficient transcriptomes Ps and ps (6.4 and 4.3 $\log_2$ FPKM, respectively). For these genes, lower expression of S-responsive genes was found in plants grown with added ${SO_4}^{2-}$. Thus unlike the low P_i conditions, several known S starvation genes do show the expected differential expression between plants given s vs. S as a treatment.

Figure 3.8 shows the extent of transcriptome reprogramming found for Yukon *E. salsugineum* grown with low P_i or low ${SO_4}^{2-}$ is different relative to *A. thaliana* and that S availability was likely a greater factor to producing DEGs than low P_i. Indeed, this outcome is not unexpected given the data summarized in **Table 3.3**. When the contribution of P_i nutrient status was considered in altering gene expression, there was only one DEG associated with the low P_i whereas 371 DEGs were found in a comparison that focused only on the contribution of the s vs. S nutrient status (**Table 3.3**). This outcome raises an important question: is the poor response to low P_i typical of plants in general or is it a response that is distinct, perhaps even unique to Yukon *E. salsugineum*? To address this question, 28 publicly available transcriptomes were retrieved from various published P_i starvation experiments and compared with the four Yukon *E. salsugineum* transcriptomes assembled for this thesis. The data collected was intentionally broad in scope in that it corresponds to 11 phylogenetically different plants representing a variety of low P_i stress regimes and transcriptome profiling platforms (including Microarray and sequencing using Illumina or Pacific Biosciences technologies). A detailed summary of the condition, gene counts and cited source publication is given in **Table A1.3**. The outcome of the comparison is shown in **Figure 3.9** and it indicates that among all the species compared, the lowest contribution to transcriptome profiles of P_i-responsive genes was given by Yukon *E. salsugineum* with an average of 5 DEGs or 0.038% of total expressed genes being P_i-responsive in pairwise comparisons of plants grown under P_i-supplemented vs. low P_i conditions. In comparison, relative to the Yukon *E. salsugineum* transcriptomes, data from other experiments with plants exposed to low P_i had a range of 6.35% to 33.56% of total expressed genes being differentially regulated in leaves in response to P_i nutrient status. It is important to note that instead of a linear scale, the x axis was log-transformed to better display the data from Yukon *E. salsugineum* relative to other transcriptomes.

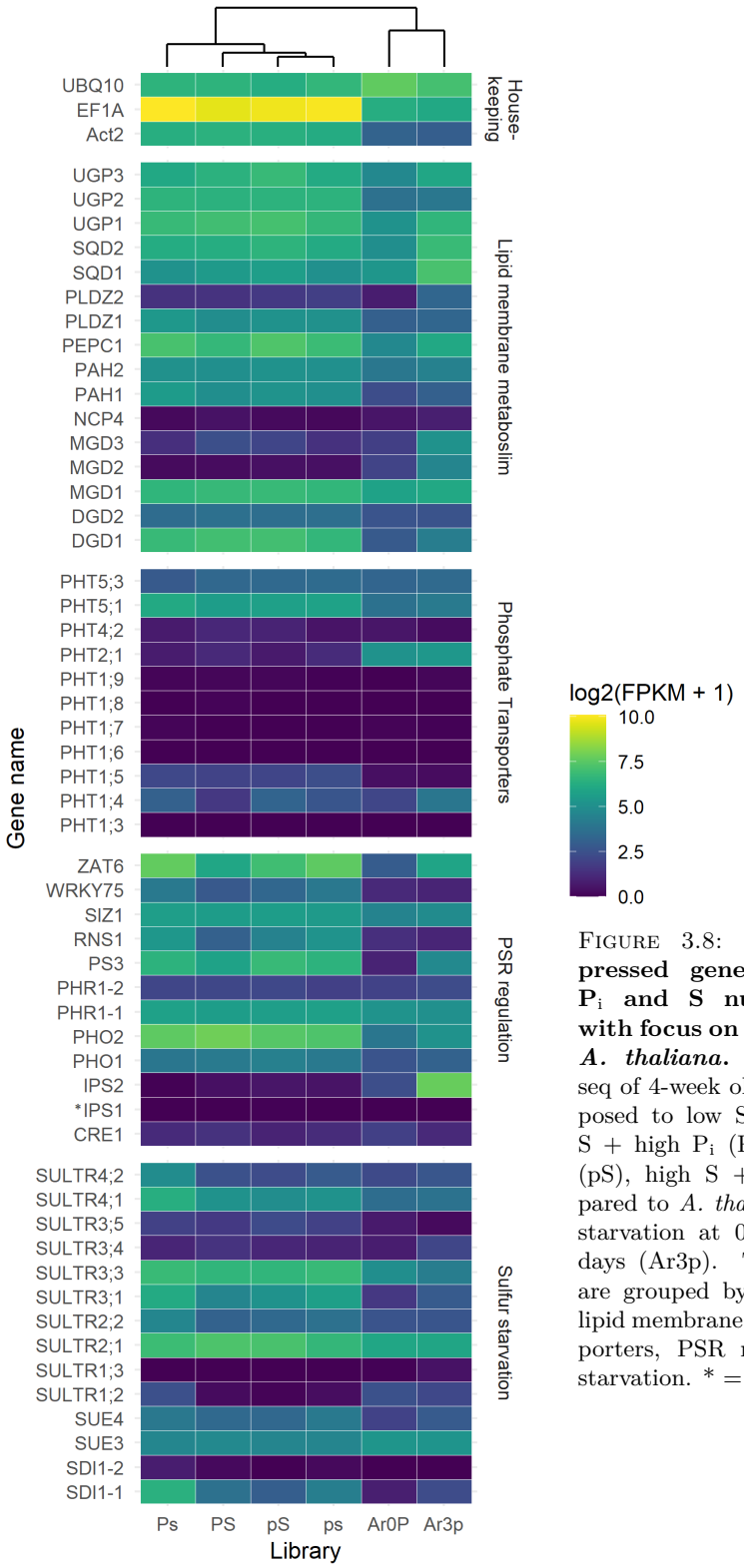

FIGURE 3.8: **Heatmap of expressed genes associated with P_i and S nutrition in plants, with focus on *E. salsugineum* and *A. thaliana*.** FPKM from RNA-seq of 4-week old Eutrema plants exposed to low S + low P_i (ps), low S + high P_i (Ps), high S + low P_i (pS), high S + high P_i (PS) compared to *A. thaliana* plants under P_i starvation at 0 days (Ar0P) and 3 days (Ar3p). The genes of interest are grouped by housekeeping genes, lipid membrane metabolism, P_i transporters, PSR regulation and sulfur starvation. * = putative *IPS1*.

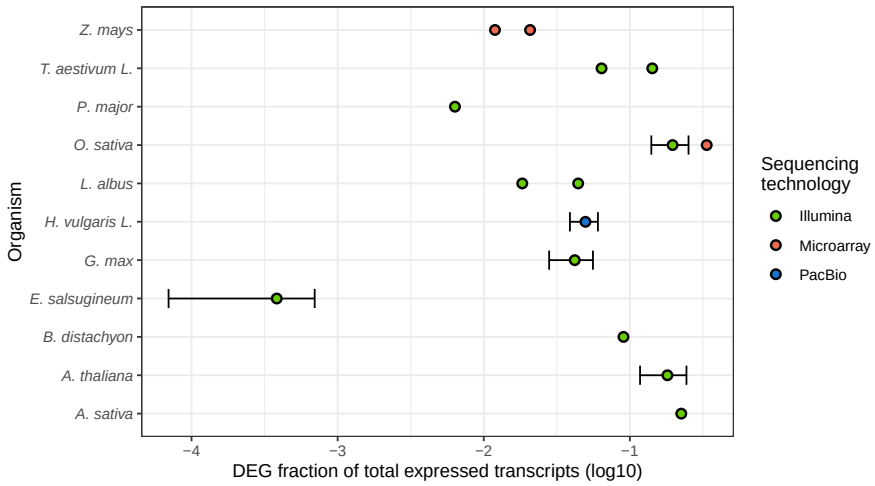

FIGURE 3.9: **Relative P_i-responsive genes between species.** The $\log_{10}$ratio of differentially expressed genes per total expressed genes in shoots of 10 species and 10 published P_i starvation experiments including results from Yukon *E. salsugineum* transcriptome. The data gathered are grouped by species and sequencing methods. The data gathered are produced with either Microarray, Illumina or Pacific Biosciences sequencing methods. Yukon *E. salsugineum* (EutremaPS) had the lowest ratio of differentially expressed genes per total expressed gene compared to the other species. Error bars represent standard deviation (data point sample sizes: *Avena sativa* = 1, *Arabidopsis thaliana* = 4, *Brachypodium distachyon* = 1, *E. salsugineum* = 3, *Glycine max* = 4, *Hordeum vulgaris* = 4, *Lupinus albus* = 2, *Oryza sativa* = 9, *Plantago major* = 1, *Triticum aestivum* = 2 and *Zea mays* = 2).

Chapter 4

Discussion

Velasco et al. (2016) first reported that Yukon *E. salsugineum* seedlings and plants grown in P_i-deficiency conditions display little to no phenotypic differences relative to plants provided with P_i. Concerns that the growing conditions were not severe enough to generate a deficiency were addressed by a direct comparison with *A. thaliana* plants grown using the same P_i-deficient media that yielded plants showing a classic P_i-deficiency response including decreased shoot biomass and altered root architecture. By contrast, Yukon *E. salsugineum* shoot and root biomass and relative growth rates were unchanged, even after four weeks of growth on soil that received no added P_i. Further, Velasco et al. (2020) reported that *E. salsugineum* does not secrete root phosphatases under low P_i conditions despite having constitutively high root PAP activity and they concluded from enzyme measurements that Yukon *E. salsugineum* likely constitutively employs glycolytic bypass mechanisms reported for low P_i-stressed plants. For example, PEPC activity can generate P_i from PEP and PEPC activity for Yukon *E. salsugineum* is typically high and not responsive to low P_i levels as is the case for *A. thaliana* (Velasco et al. 2020). Additionally, at the molecular level Velasco et al. (2016) used RT-qPCR to measure absolute transcript abundance and found that *EsIPS2*, *EsRNS1*, *EsWRKY75* and *EsPHR1* transcripts were present in plants given P_i and that only *EsIPS2* expression showed low P_i-responsive behaviour in leaves of plants not given P_i for four weeks. In all of these features, Yukon *E. salsugineum* stands out as a departure from the typical morphological, physiological and molecular low P_i responses reported in the literature, not only responses in *A. thaliana* but more broadly including crop species and the extremophyte lupins that form proteoid roots under P_i starvation (Johnson et al. 1994; Plaxton and Tran 2011).

The work in this thesis was directed towards characterizing the transcriptome profiles of Yukon *E. salsugineum* plants grown under similar conditions of P_i treatment used by Velasco et al. (2016, 2020) in order to provide a more complete picture of how Yukon plants respond to a P_i-deficiency at the gene expression level. We hypothesized, based on the RT-qPCR work reported by Velasco et al. (2016), that Yukon *E. salsugineum* plants experiencing low P_i would likely have few DEGs compared to plants fertilized with 2.5 mM P_i. One difference in this study was the use of two different soil S levels leading to four treatments undergoing testing: no added P_i (p) or 2.5 mM P_i (P) in combination with no added ${SO_4}^{2-}$ (s) or 2.5 mM ${SO_4}^{2-}$ (S) (see Materials and Methods and **Table 2.1**). The addition of ${SO_4}^{2-}$ was done to prevent plants from experiencing two nutrient deficiencies given the soil favoured by Yukon *E. salsugineum* is highly enriched in S (Guevara et al. 2012). Garvin (2016) reported that genes responsive to low S were up-regulated when plants were grown in soil with a regular potting mix despite weekly fertilizer treatment with a 20-20-20 N/P/K fertilizer.

Differential gene expression analysis showed that *IPS2*, putative *IPS1*, *WRKY75*, *RNS1*, *PHR1* and *PHO2*, genes that are typically associated with PSR, were not differentially regulated in Yukon *E. salsugineum* grown under low P_i conditions (**Figure 3.8**). This finding

agrees with the RT-qPCR findings of Velasco et al. (2016), in that three of four PSI genes (*RNS1*, *WRKY75* and *PHR1*) were not differentially regulated between Yukon *E. salsugineum* grown under differing P_i treatments. Between the comparison of pS vs. PS and of P vs. p, there were three DEGs, namely S-adenosyl-L-methionine dependent carboxyl methyltransferase, cytochrome P450 unknown protein and an uncharacterized lncRNA. It is unclear whether these three genes are associated with P_i starvation and there is a lack of literature investigating the roles of these genes in relation to P_i metabolism. The original work of Velasco et al. (2016) did not include *IPS1* as evidence for the existence of this gene in the *E. salsugineum* genome annotation was not available until this study (see section 3.1). Additionally, in this study, *IPS2* was not identified as a DEG whereas it was reported as differentially expressed in the work of Velasco et al. (2016). One reason for the difference between results with respect to *IPS2* expression may reside in the choice of tissue for RNA extraction. Velasco et al. (2016) only selected fully expanded leaves and leaves are known to redistribute P_i from older to newer leaves although that capacity is poor for *E. salsugineum* (Velasco et al. 2020). Thus including tissue from immature leaves for preparing transcriptomes in the present work could have reduced *IPS2* expression or rendered it more variable and hence difficult to identify as a DEG with confidence. Whether that had a broader impact on finding P_i-responsive genes is difficult to assess, but in this study the leaf transcriptomes of Yukon *E. salsugineum* plants were substantially unchanged by the low P_i treatment in that less than 0.09% of the expressed genes were identified as DEGs when plants were grown on low vs. high/supplemented P_i soil (**Table 3.3**). However, pooled leaf tissue showed low P_i content in leaves from plants grown with low P_i (Velasco, unpublished). Moreover, the infrequency of P_i-responsive genes detected across the four Yukon *E. salsugineum* transcriptomes did not mean that the plant was completely insensitive to soil P_i levels. PCA analysis of the total expressed genes in the four transcriptomes (**Figure 3.4**) showed a distinction between the four treatments in the biplot of PC2 and PC4, albeit the variance contributing to the distinctions in this biplot is modest. Furthermore, GO enrichment analysis of the top 200 loading genes in each of the PCs (**Figure 3.5**) revealed that genes contributing to PC2, which showed distinction between s and S treatments, were associated with mechanisms relating to S-stress deficiencies. Similarly, genes contributing to PC4, which displayed distinction between P_i treatment, were associated with mechanisms such as P_i ion homeostasis and inorganic anion homeostasis (**Figure 3.5**). Together, this suggests that Yukon *E. salsugineum* plants have responded to changes in both P_i and ${SO_4}^{2-}$ content through modest adjustments in the transcriptome that were not detected with the strict shrinkage and filtering algorithms when the transcriptomes were analyzed with DESeq2.

LncRNAs have been implicated in contributing to the stress responses of plants (Xu et al. 2017; Yuan et al. 2016). The prediction of 615 lncRNAs was less than the 1,040 lncRNAs or 3% of the Yukon *E. salsugineum* genome predicted by Simopoulos (2019) with drought. The lower number of predicted lncRNAs in this work was due in part to a step in the methods taken to remove very low-expression transcripts. However, the prediction of 615 lncRNAs does agree with the overall finding of Simopoulos (2019) for *E. salsugineum* in that this species seems to

express a lower number of lncRNAs compared to other model plants such as *A. thaliana*, *Oryza sativa* and *Zea mays*, with the latter groups having over 4.8% of their respective transcriptomes predicted as lncRNAs (Simopoulos 2019). In a systematic examination of P_i-responsive lncRNAs in *A. thaliana* performed by Yuan et al. (2016), there were 309 predicted lncRNAs found whereas only one was found to be P_i-responsive in this study. The lack of stress-responsive lncRNAs was also reflected in S-limitation, as transcriptome pairs comparing differential S treatments had a higher number of DEGs but not a higher number of predicted lncRNAs. Similarly, Simopoulos et al. (2020) reported a lack of drought-responsive lncRNAs in drought-stressed Yukon *E. salsugineum*. This suggests that Yukon *E. salsugineum* does not rely on lncRNAs in mediating response to low P_i in a similar manner as *A. thaliana* where low P_i will induce up more than 300 lncRNAs (Yuan et al. 2016). Moreover, based on the lack of S-responsive lncRNAs found in this work and drought-responsive lncRNAs in the results published by Simopoulos et al. (2020), it is possible to hypothesize that Yukon *E. salsugineum* does not rely on the expression of lncRNAs as a crucial strategy for responding to stress.

Taken together, *E. salsugineum* is unique compared to other plants in the lack of a classical P_i-starvation response in the transcriptome (**Figure 3.9** and **Figure 3.8**). Yukon *E. salsugineum* had no more than 11 Pi-responsive DEGs and had the least proportion of P_i-responsive genes compared to 10 other plant species. Conceivably, Yukon *E. salsugineum* utilizes P_i efficiently in such a way that PSR-inducing conditions for *A. thaliana* that can induce many PSI genes does not induce PSR in *E. salsugineum* (**Figure 3.8**). Evidence in support of this proposal is seen in the PCA of transcriptomes of Yukon *E. salsugineum* exposed to both low P_i and low S that suggests that the transcriptome is altered in response to the change in soil nutrient levels at low variances (**Figure 3.4**). The fine-tuning strategy of the Yukon *E. salsugineum* transcriptome in response to P_i and S is a contrast to the recent work by Simopoulos et al. (2020): drought-exposed Yukon *E. salsugineum* can experience transcriptome reprogramming that involves more than 2,000 genes (Simopoulos et al. 2020). However, transcript levels do not always reflect protein abundance, minute adjustments in transcript levels can result in a significant increase or decrease in downstream protein abundance (Liu and Aebersold 2016). Schwender et al. (2014) found that transcript abundance alone could not describe changes in fluxes as well as activity in glycolysis, amino acid synthesis and fatty acid synthesis in *Brassica napus*. During glycolysis, PEP may act as an allosteric inhibitor of aldolase, a sugar converting enzyme (Plaxton 1996). Similarly, P_i itself could be an allosteric regulator in P_i signalling pathways related to P_i sensing, although research in this area is lacking. For a plant such as Yukon *E. salsugineum* that thrives under habitat that is constantly low in P_i, it is reasonable to think that P_i homeostasis is maintained through post-transcriptional regulation, translational regulation or P_i-mediated allosteric feedback control mechanisms, this would ensure that the proteins required for P_i acquisition are readily available whenever P_i is needed. Additionally, a proteome of Yukon *E. salsugineum* tissues can be created with liquid chromatography coupled with mass spectrometry to pair with the transcriptomic data. Lastly, the relationship between P_i content and gene expression of the root is not known and transcriptome profiling of Yukon *E. salsugineum* roots remain to be explored.

Appendix A

Results Supplement

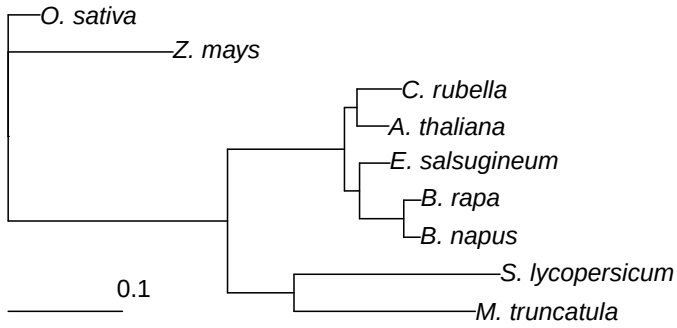

FIGURE A1.1: Species tree inferred from alignment of 16s rRNA, 18s rRNA, *Rps16*, *Atp2* and *SMC1* of *Oryza sativa*, *Zea mays*, *Capsella rubella*, *Arabidopsis thaliana*, *Eutrema salsugineum*, *Brassica rapa*, *Brassica napus*, *Solanum lycopersicum* and *Medicago truncatula*. Scale represents nucleotide substitutions per site.

TABLE A1.1: Enriched GO terms of top loading genes contributing to principal components

Term ID	Description	$\log_10$ p-value	Semantic similarity
Positive PC2			
GO:0008150	biological process	1	100.000
GO:0008152	metabolic process	0.99	75.387
GO:0050896	response to stimulus	0.965	12.210
GO:0051704	multi-organism process	0.961	0.751
GO:0005975	carbohydrate metabolic process	0.926	5.260
GO:0006790	sulfur compound metabolic process	0.922	1.822
GO:0019748	secondary metabolic process	0.871	0.138
GO:0009309	amine biosynthetic process	0.869	0.312
GO:0010260	animal organ senescence	0.864	0.000
GO:0007568	aging	0.858	0.088
GO:0016143	S-glycoside metabolic process	0.837	0.003
GO:0016137	glycoside metabolic process	0.836	0.031
GO:0090693	plant organ senescence	0.829	0.006
GO:0048366	leaf development	0.828	0.019
GO:0010150	leaf senescence	0.826	0.006
GO:0055114	oxidation-reduction process	0.822	15.060
GO:0031407	oxylipin metabolic process	0.817	0.007
GO:0009407	toxin catabolic process	0.813	0.000
GO:0042180	cellular ketone metabolic process	0.798	0.423
GO:0009404	toxin metabolic process	0.779	0.039
GO:0006082	organic acid metabolic process	0.741	9.086
GO:0042343	indole glucosinolate metabolic process	0.721	0.001
GO:0009607	response to biotic stimulus	0.71	0.342
GO:1901607	alpha-amino acid biosynthetic process	0.708	2.557
GO:0019757	glycosinolate metabolic process	0.706	0.003
GO:0019760	glucosinolate metabolic process	0.706	0.003
GO:0008652	cellular amino acid biosynthetic process	0.704	2.932
GO:1901605	alpha-amino acid metabolic process	0.704	3.625
GO:0006520	cellular amino acid metabolic process	0.703	5.591
GO:0043436	oxoacid metabolic process	0.702	9.006
GO:0009719	response to endogenous stimulus	0.701	0.526
GO:0009611	response to wounding	0.7	0.127
GO:0009415	response to water	0.68	0.026
GO:0009605	response to external stimulus	0.679	1.370
GO:0009620	response to fungus	0.675	0.035
GO:0006952	defense response	0.669	0.568
GO:0006979	response to oxidative stress	0.669	0.575
GO:0001101	response to acid chemical	0.663	0.124
GO:0050832	defense response to fungus	0.659	0.028
GO:0042221	response to chemical	0.659	3.071
GO:0009414	response to water deprivation	0.658	0.022
GO:0006950	response to stress	0.648	4.575
GO:0010035	response to inorganic substance	0.642	0.317
GO:0009725	response to hormone	0.641	0.335
GO:0043207	response to external biotic stimulus	0.634	0.300
GO:0051707	response to other organism	0.632	0.299
GO:1901700	response to oxygen-containing compound	0.631	0.503
GO:0098754	detoxification	0.619	0.804
GO:0009636	response to toxic substance	0.618	0.833
GO:0010033	response to organic substance	0.616	0.900
GO:0098542	defense response to other organism	0.616	0.220
Negative PC2			
GO:0008150	biological process	1	100.000
GO:0008152	metabolic process	0.99	75.387

Term ID	Description	$\log_10$ p-value	Semantic similarity
GO:0050896	response to stimulus	0.965	12.210
GO:0051704	multi-organism process	0.961	0.751
GO:0005975	carbohydrate metabolic process	0.926	5.260
GO:0006790	sulfur compound metabolic process	0.922	1.822
GO:0019748	secondary metabolic process	0.871	0.138
GO:0009309	amine biosynthetic process	0.869	0.312
GO:0010260	animal organ senescence	0.864	0.000
GO:0007568	aging	0.858	0.088
GO:0016143	S-glycoside metabolic process	0.837	0.003
GO:0016137	glycoside metabolic process	0.836	0.031
GO:0090693	plant organ senescence	0.829	0.006
GO:0048366	leaf development	0.828	0.019
GO:0010150	leaf senescence	0.826	0.006
GO:0055114	oxidation-reduction process	0.822	15.060
GO:0031407	oxylipin metabolic process	0.817	0.007
GO:0009407	toxin catabolic process	0.813	0.000
GO:0042180	cellular ketone metabolic process	0.798	0.423
GO:0009404	toxin metabolic process	0.779	0.039
GO:0006082	organic acid metabolic process	0.741	9.086
GO:0042343	indole glucosinolate metabolic process	0.721	0.001
GO:0009607	response to biotic stimulus	0.71	0.342
GO:1901607	alpha-amino acid biosynthetic process	0.708	2.557
GO:0019757	glycosinolate metabolic process	0.706	0.003
GO:0019760	glucosinolate metabolic process	0.706	0.003
GO:0008652	cellular amino acid biosynthetic process	0.704	2.932
GO:1901605	alpha-amino acid metabolic process	0.704	3.625
GO:0006520	cellular amino acid metabolic process	0.703	5.591
GO:0043436	oxoacid metabolic process	0.702	9.006
GO:0009719	response to endogenous stimulus	0.701	0.526
GO:0009611	response to wounding	0.7	0.127
GO:0009415	response to water	0.68	0.026
GO:0009605	response to external stimulus	0.679	1.370
GO:0009620	response to fungus	0.675	0.035
GO:0006952	defense response	0.669	0.568
GO:0006979	response to oxidative stress	0.669	0.575
GO:0001101	response to acid chemical	0.663	0.124
GO:0050832	defense response to fungus	0.659	0.028
GO:0042221	response to chemical	0.659	3.071
GO:0009414	response to water deprivation	0.658	0.022
GO:0006950	response to stress	0.648	4.575
GO:0010035	response to inorganic substance	0.642	0.317
GO:0009725	response to hormone	0.641	0.335
GO:0043207	response to external biotic stimulus	0.634	0.300
GO:0051707	response to other organism	0.632	0.299
GO:1901700	response to oxygen-containing compound	0.631	0.503
GO:0098754	detoxification	0.619	0.804
GO:0009636	response to toxic substance	0.618	0.833
GO:0010033	response to organic substance	0.616	0.900
GO:0098542	defense response to other organism	0.616	0.220
Positive PC4			
GO:0050896	response to stimulus	0.859	12.210
GO:0009772	photosynthetic electron transport in photosystem II	0.84	0.012
GO:0042221	response to chemical	0.682	3.071
GO:0050832	defense response to fungus	0.672	0.028
GO:0000302	response to reactive oxygen species	0.654	0.181
GO:0006979	response to oxidative stress	0.643	0.575
GO:0042592	homeostatic process	0.386	1.661
GO:0048878	chemical homeostasis	0.35	0.543

Term ID	Description	$\log_1 0$ p-value	Semantic similarity
GO:0098771	inorganic ion homeostasis	0.279	0.410
GO:0072506	trivalent inorganic anion homeostasis	0.261	0.035
GO:0072505	divalent inorganic anion homeostasis	0.261	0.035
GO:0055062	phosphate ion homeostasis	0.261	0.035
GO:0055083	monovalent inorganic anion homeostasis	0.259	0.038
Negative PC4			
NA			

TABLE A1.2: Enriched GO terms of DEGs in pairwise comparison

Term ID	Description	$\log_1 0$ p-value	Semantic similarity
GO:0050896	response to stimulus	-21.7788	0.013
GO:0051704	multi-organism process	-10.751	0.014
GO:0002376	immune system process	-1.572	0.014
GO:0033037	polysaccharide localization	-1.5878	0.04
GO:0009812	flavonoid metabolic process	-2.006	0.04
GO:0071554	cell wall organization or biogenesis	-1.48	0.041
GO:0010817	regulation of hormone levels	-2.1398	0.049
GO:0052545	callose localization	-1.6424	0.053
GO:0007568	aging	-1.8406	0.056
GO:0052386	cell wall thickening	-2.0066	0.0580
GO:0005975	carbohydrate metabolic process	-4.6593	0.0580
GO:0006790	sulfur compound metabolic process	-6.1013	0.0639
GO:0044262	cellular carbohydrate metabolic process	-2.2421	0.0659
GO:0009813	flavonoid biosynthetic process	-2.544	0.077
GO:0042445	hormone metabolic process	-2.6054	0.082
GO:0019748	secondary metabolic process	-9.844	0.098
GO:0019742	pentacyclic triterpenoid metabolic process	-1.5878	0.108
GO:0006722	triterpenoid metabolic process	-1.3223	0.117
GO:0009851	auxin biosynthetic process	-2.532	0.12
GO:0009308	amine metabolic process	-4.4579	0.125
GO:0019745	pentacyclic triterpenoid biosynthetic process	-1.5878	0.131
GO:0042430	indole-containing compound metabolic process	-6.6784	0.133
GO:0016137	glycoside metabolic process	-5.3178	0.143
GO:0055114	oxidation-reduction process	-3.68	0.146
GO:0016143	S-glycoside metabolic process	-8.2123	0.153
GO:0044272	sulfur compound biosynthetic process	-2.9136	0.156
GO:1901657	glycosyl compound metabolic process	-1.3366	0.175
GO:0042180	cellular ketone metabolic process	-6.1013	0.178
GO:0016138	glycoside biosynthetic process	-2.1398	0.187
GO:0044106	cellular amine metabolic process	-4.7968	0.204
GO:0046417	chorismate metabolic process	-2.1381	0.207
GO:0043648	dicarboxylic acid metabolic process	-2.925	0.232
GO:0006082	organic acid metabolic process	-8.5148	0.237
GO:0009625	response to insect	-2.6054	0.242
GO:0080167	response to karrikin	-1.8983	0.247
GO:0009607	response to biotic stimulus	-15.682	0.256
GO:0009072	aromatic amino acid family metabolic process	-3.3182	0.263
GO:0009719	response to endogenous stimulus	-9.8599	0.264
GO:0009628	response to abiotic stimulus	-7.7206	0.266
GO:0002213	defense response to insect	-2.0912	0.272
GO:0072330	monocarboxylic acid biosynthetic process	-1.6002	0.272
GO:0009611	response to wounding	-12.7788	0.274
GO:0052544	defense response by callose deposition in cell wall	-2.4514	0.28
GO:0009605	response to external stimulus	-17.0951	0.284
GO:0010200	response to chitin	-2.6541	0.295
GO:0009073	aromatic amino acid family biosynthetic process	-1.5819	0.296
GO:1901606	alpha-amino acid catabolic process	-2.9321	0.298
GO:0009651	response to salt stress	-3.3549	0.3
GO:0019760	glucosinolate metabolic process	-8.2123	0.3
GO:0009744	response to sucrose	-2.0391	0.302

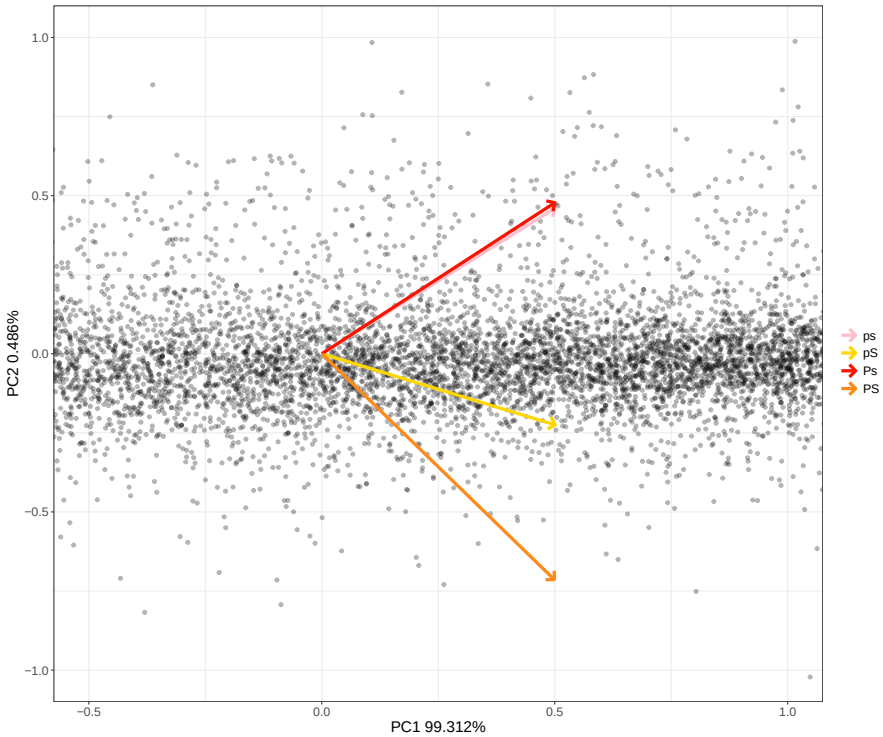

FIGURE A1.2: **Biplot graph depicting PC1 and PC2 of the PCA conducted on the PS treatment libraries.** PC1 and PC2 represent 99.31% and 0.49% of the variance between libraries, respectively. However there is high degree of overlap in PC1, resulting in uninformative distinctions between the treatment libraries.

TABLE A1.3: Gene expression data across multiple published Pi experiments of different plants.

Species	Pi Treatment	Tissue	Up-regulated	Down-regulated	# expressed genes	Technology	Citation
Eutrema salsugineum	Ps vs. pS	shoot	17	3	23,578	Illumina	This work
	ps vs. Ps	shoot	0	11			
	pS vs. PS	shoot	2	0			
Arabidopsis thaliana	10 µM vs. 250 µM	root	1,532	516	21,979	Illumina	Liu et al. (2016)
		shoot	2,425	1,458			
		root	3,624	1,545			
		shoot	2,869	1,916			
Avena sativa	1 µM vs. 100 µM	root	7,817	1,554	41,679	Illumina	Wang et al. (2018)
Brachypodium distachyon	10 µM vs. 500 µM	root	1,175	565	19,242	Illumina	Zhao et al. (2018)
Glycine max	0.2 µM vs. 250 µM	root	554	596	39,652	Illumina	Guo et al. (2008b)
		shoot	1,113	1,232			
		root	874	412			
		shoot	1,284	590			
Hordeum vulgaris L.	0.039 mM vs. 0.39 mM	root	1,132	1,057	54,400	PacBio	Ren et al. (2018)
		shoot	1,001	1,207			
		root	2,390	716			
		shoot	1,945	1,358			
Lupinus albus	0 µM vs. 1 mM	root	535	369	50,734	Illumina	O'Rourke et al. (2013)
		shoot	355	987			
Oryza sativa	0 mM vs. 0.32 mM	shoot	8,043	12,900	62,400	Microarray	Park et al. (2012)
	0.003 23 mM vs. 0.323 mM	root	4,909	4,703	52,640	Illumina	Oono et al. (2013b)
		shoot	6,294	5,112			
		root	5,617	5,653			
		shoot	8,410	5,215			
		root	5,442	6,561			
		shoot	5,769	5,544			
		root	4,713	5,766			
		shoot	7,963	5,125			
	0 mM vs. 0.323 mM	root	1,703	2,058	34,296	Illumina	Secco et al. (2013)
		shoot	1,374	1,655			
Plantago major	0 µM vs. 500 µM	shoot	109	128	37,309	Illumina	Huang et al. (2019)
Triticum aestivum L.	0 mM vs. 0.323 mM	root	1,004	892	29,617	Illumina	Oono et al. (2013a)

Species	Pi Treatment	Tissue	Up-regulated	Down-regulated	# expressed genes	Technology	Citation
		shoot	2,833	1,382			
Zea mays	5 µM vs. 1 mM	root	270	279	46,000	Microarray	**li2012phosphate**
	2 µM vs. 200 µM	root	820	363	57,000	Microarray	Calderon-Vazquez et al. (2008)

Bibliography

Ai, Q, Liang, G, Zhang, H, and Yu, D (2016). Control of sulfate concentration by miR395-targeted APS genes in *Arabidopsis thaliana*. *Plant Diversity* 38(2), pp. 92–100.

Alexa, A, Rahnenführer, J, and Lengauer, T (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* 22(13), pp. 1600–1607.

Andrew, S (2010). *A Quality Control Tool for High Throughput Sequence Data*. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Ayadi, A, David, P, Arrighi, J, Chiarenza, S, Thibaud, M, Nussaume, L, and Marin, E (2015). Reducing the genetic redundancy of *Arabidopsis* PHOSPHATE TRANSPORTER1 transporters to study phosphate uptake and signaling. *Plant Physiology* 167(4), pp. 1511–1526.

Baldwin, JC, Karthikeyan, AS, Cao, A, and Raghothama, KG (2008). Biochemical and molecular analysis of LePS2;1: a phosphate starvation induced protein phosphatase gene from tomato. *Planta* 228(2), p. 273.

Barber, SA, Walker, JM, and Vasey, EH (1963). Mechanisms for movement of plant nutrients from soil and fertilizer to plant root. *Journal of Agricultural and Food Chemistry* 11(3), pp. 204–207.

Bari, R (2006). PHO2, MicroRNA399, and PHR1 Define a Phosphate-Signaling Pathway in Plants. *Plant Physiology* 141(3), pp. 988–999.

Bariola, PA, Howard, CJ, Taylor, CB, Verburg, MT, Jaglan, VD, and Green, PJ (1994). The *Arabidopsis* ribonuclease gene RNS1 is tightly controlled in response to phosphate limitation. *The Plant Journal* 6(5), pp. 673–685.

Bieleski, RL (1973). Phosphate pools, phosphate transport, and phosphate availability. *Annual Review of Plant Physiology* 24(1), pp. 225–252.

Blackwell, MSA, Darch, T, and Haslam, RP (2019). Phosphorus use efficiency and fertilizers: future opportunities for improvements. *Frontiers of Agricultural Science and Engineering* 6(4), pp. 332–340.

Bolger, AM, Lohse, M, and Usadel, B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15), pp. 2114–2120.

Bordo, D and Bork, P (2002). The RHODANESE/CDC25 phosphatase superfamily. *EMBO Reports* 3(8), pp. 741–746.

Brindley, DN and Pilquil, C (2009). Lipid phosphate phosphatases and signaling. *Journal of Lipid Research* 50(Supplement), S225–S230.

Buchfink, B, Xie, C, and Huson, DH (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Methods* 12(1), pp. 59–60.

Burleigh, SM and Harrison, MJ (1998). Characterization of the Mt4 gene from *Medicago truncatula*. *Gene* 216(1), pp. 47–53.

Calderon-Vazquez, C, Ibarra-Laclette, E, Caballero-Perez, J, and Herrera-Estrella, L (2008). Transcript profiling of *Zea mays* roots reveals gene responses to phosphate deficiency at the plant-and species-specific levels. *Journal of Experimental Botany* 59(9), pp. 2479–2497.

Carmichael, WW and Falconer, IR (1993). *Diseases related to freshwater blue-green algal toxins, and control measures.* Academic Press London, pp. 187–209.

Ceasar, SA, Baker, A, Muench, SP, Ignacimuthu, S, and Baldwin, SA (2016). The conservation of phosphate-binding residues among PHT1 transporters suggests that distinct transport affinities are unlikely to result from differences in the phosphate-binding site. *Biochemical Society Transactions* 44(5), pp. 1541–1548.

Champigny, MJ, Sung, WWL, Catana, V, Salwan, R, Summers, PS, Dudley, SA, Provart, NJ, Cameron, RK, Golding, GB, and Weretilnyk, EA (2013). RNA-Seq effectively monitors gene expression in *Eutrema salsugineum* plants growing in an extreme natural habitat and in controlled growth cabinet conditions. *BMC Genomics* 14(1), p. 578.

Chen, L, Song, Y, Li, S, Zhang, L, Zou, C, and Yu, D (2012). The role of WRKY transcription factors in plant abiotic stresses. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* 1819(2), pp. 120–128.

Chen, Y, Li, L, Xu, Q, Kong, Y, Wang, H, and Wu, W (2009). The WRKY6 transcription factor modulates PHOSPHATE1 expression in response to low P_i stress in *Arabidopsis. The Plant Cell* 21(11), pp. 3554–3566.

Chen, Z, Nimmo, GA, Jenkins, GI, and Nimmo, HG (2007). BHLH32 modulates several biochemical and morphological processes that respond to P_i starvation in *Arabidopsis. Biochemical Journal* 405(1), pp. 191–198.

Chen, Z, Zhao, P, Miao, Z, Qi, G, Wang, Z, Yuan, Y, Ahmad, N, Cao, M, Hell, R, Wirtz, M, et al. (2019). SULTR3s function in chloroplast sulfate uptake and affect ABA biosynthesis and the stress response. *Plant Physiology* 180(1), pp. 593–604.

Cheng, Y, Zhou, Y, Yang, Y, Chi, Y, Zhou, J, Chen Jand Wang, F, Fan, B, Shi, K, Zhou, Y, et al. (2012). Structural and functional analysis of VQ motif-containing proteins in *Arabidopsis* as interacting proteins of WRKY transcription factors. *Plant Physiology* 159(2), pp. 810–825.

Cruz-Ramírez, A, Oropeza-Aburto, A, Razo-Hernández, F, Ramírez-Chávez, E, and Herrera-Estrella, L (2006). PHOSPHOLIPASE DZ2 plays an important role in extraplastidic galactolipid biosynthesis and phosphate recycling in *Arabidopsis* roots. *Proceedings of the National Academy of Sciences* 103(17), pp. 6765–6770.

Dai, X, Wang, Y, and Zhang, W (2016). OsWRKY74, a WRKY transcription factor, modulates tolerance to phosphate starvation in rice. *Journal of Experimental Botany* 67(3), pp. 947–960.

Delhaize, E and Randall, PJ (1995). Characterization of a phosphate-accumulator mutant of *Arabidopsis thaliana. Plant Physiology* 107(1), pp. 207–213.

Devaiah, BN, Nagarajan, VK, and Raghothama, KG (2007a). Phosphate homeostasis and root development in *Arabidopsis* are synchronized by the zinc finger transcription factor ZAT6. *Plant Physiology* 145(1), pp. 147–159.

Devaiah, BN, Karthikeyan, AS, and Raghothama, KG (2007b). WRKY75 transcription factor is a modulator of phosphate acquisition and root development in *Arabidopsis. Plant Physiology* 143(4), pp. 1789–1801.

Devaiah, BN, Madhuvanthi, R, Karthikeyan, AS, and Raghothama, KG (2009). Phosphate starvation responses and gibberellic acid biosynthesis are regulated by the MYB62 transcription factor in *Arabidopsis*. *Molecular Plant* 2(1), pp. 43–58.

Dobin, A, Davis, CA, Schlesinger, F, Drenkow, J, Zaleski, C, Jha, S, Batut, P, Chaisson, M, and Gingeras, TR (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1), pp. 15–21.

Eastmond, PJ, Quettier, A, Kroon, JTM, Craddock, C, Adams, N, and Slabas, AR (2010). PHOSPHATIDIC ACID PHOSPHOHYDROLASE1 and 2 regulate phospholipid synthesis at the endoplasmic reticulum in *Arabidopsis*. *The Plant Cell* 22(8), pp. 2796–2811.

Engström, PG, Steijger, T, Sipos, B, Grant, GR, Kahles, A, Alioto, T, Behr, J, Bertone, P, Bohnert, R, Campagna, D, et al. (2013). Systematic evaluation of spliced alignment programs for RNA-seq data. *Nature Methods* 10(12), pp. 1185–1191.

Eulgem, T, Rushton, PJ, Schmelzer, E, Hahlbrock, K, and Somssich, IE (1999). Early nuclear events in plant defence signalling: rapid gene activation by WRKY transcription factors. *The EMBO Journal* 18(17), pp. 4689–4699.

Fontenot, EB, Ditusa, SF, Kato, N, Olivier, DM, Dale, R, Lin, W, Chiou, T, Macnaughtan, MA, and Smith, AP (2015). Increased phosphate transport of *Arabidopsis thaliana* Pht1;1 by site-directed mutagenesis of tyrosine 312 may be attributed to the disruption of homomeric interactions. *Plant, Cell & Environment* 38(10), pp. 2012–2022.

Food and Agricultural Organization of the United Nations (2017). World fertilizer trends and outlook to 2020, pp. 3–5.

Forlani, F, Carpen, A, and Pagani, S (2003). Evidence that elongation of the catalytic loop of the *Azotobacter vinelandii* rhodanese changed selectivity from sulfur-to phosphate-containing substrates. *Protein Engineering* 16(7), pp. 515–519.

Franco-Zorrilla, JM, Martin, AC, Solano, R, Rubio, V, Leyva, A, and Paz-Ares, J (2002). Mutations at CRE1 impair cytokinin-induced repression of phosphate starvation responses in *Arabidopsis*. *The Plant Journal* 32(3), pp. 353–360.

Franco-Zorrilla, JM, Martín, AC, Leyva, A, and Paz-Ares, J (2005). Interaction between phosphate-starvation, sugar, and cytokinin signaling in *Arabidopsis* and the roles of cytokinin receptors CRE1/AHK4 and AHK3. *Plant Physiology* 138(2), pp. 847–857.

Franco-Zorrilla, JM, Valli, A, Todesco, M, Mateos, I, Puga, MI, Rubio-Somoza, I, Leyva, A, Weigel, D, García, JA, and Paz-Ares, J (2007). Target mimicry provides a new mechanism for regulation of microRNA activity. *Nature Genetics* 39(8), pp. 1033–1037.

Garvin, A (2016). Sulfur nutrition studies using the halophyte *Eutrema salsugineum*. MA thesis. McMaster University.

Gaude, N, Nakamura, Y, Scheible, W, Ohta, H, and Dörmann, P (2008). PHOSPHOLIPASE C5 (NPC5) is involved in galactolipid accumulation during phosphate limitation in leaves of *Arabidopsis*. *The Plant Journal* 56(1), pp. 28–39.

Ghasemi Pirbalouti, A, Sajjadi, SE, and Parang, K (2014). A review (research and patents) on jasmonic acid and its derivatives. *Archiv der pharmazie* 347(4), pp. 229–239.

Gigolashvili, T and Kopriva, S (2014). Transporters in plant sulfur metabolism. *Frontiers in Plant Science* 5, p. 442.

González, E, Solano, R, Rubio, V, Leyva, A, and Paz-Ares, J (2005). PHOSPHATE TRANSPORTER TRAFFIC FACILITATOR1 is a plant-specific SEC12-related protein that enables the endoplasmic reticulum exit of a high-affinity phosphate transporter in *Arabidopsis*. *The Plant Cell* 17(12), pp. 3500–3512.

Grant, CA, Flaten, DN, Tomasiewicz, DJ, and Sheppard, SC (2001). The importance of early season phosphorus nutrition. *Canadian Journal of Plant Science* 81(2), pp. 211–224.

Gregory, AL, Hurley, BA, Tran, HT, Valentine, AJ, She, Y, Knowles, VL, and Plaxton, WC (2009). *In vivo* regulatory phosphorylation of the phosphoenolpyruvate carboxylase AtPPC1 in phosphate-starved *Arabidopsis thaliana*. *Biochemical Journal* 420(1), pp. 57–65.

Griffith, M, Timonin, M, Wong, ACE, Gray, GR, Akhter, SR, Saldanha, M, Rogers, MiA, Weretilnyk, EA, and Moffatt, B (2007). *Thellungiella*: an *Arabidopsis*-related model plant adapted to cold temperatures. *Plant, Cell & Environment* 30(5), pp. 529–538.

Guevara, DR, Champigny, MJ, Tattersall, A, Dedrick, J, Wong, CE, Li, Y, Labbe, A, Ping, C, Wang, Y, Nuin, P, et al. (2012). Transcriptomic and metabolomic analysis of Yukon *Thellungiella* plants grown in cabinets and their natural habitat show phenotypic plasticity. *BMC Plant Biology* 12(1), pp. 1–17.

Guo, B, Jin, Y, Wussler, C, Blancaflor, EB, Motes, CM, and Versaw, WK (2008a). Functional analysis of the *Arabidopsis* PHT4 family of intracellular phosphate transporters. *New Phytologist* 177(4), pp. 889–898.

Guo, W, Zhang, L, Zhao, J, Liao, H, Zhuang, C, and Yan, X (2008b). Identification of temporally and spatially phosphate-starvation responsive genes in *Glycine max*. *Plant Science* 175(4), pp. 574–584.

Hamburger, D (2002). Identification and characterization of the *Arabidopsis* PHO1 gene involved in phosphate loading to the xylem. *The Plant Cell* 14(4), pp. 889–902.

Hartley, SW and Mullikin, JC (2015). QoRTs: a comprehensive toolset for quality control and data processing of RNA-Seq experiments. *BMC Bioinformatics* 16(1), p. 224.

Huang, J, Huang, Z, Zhou, X, Xia, C, Imran, M, Wang, S, Xu, C, Zha, M, Liu, Y, and Zhang, C (2019). Tissue-specific transcriptomic profiling of *Plantago major* provides insights for the involvement of vasculature in phosphate deficiency responses. *Molecular Genetics and Genomics* 294(1), pp. 159–175.

Huang, T, Han, C, Lin, S, Chen, Y, Tsai, Y, Chen, Y, Chen, J, Lin, W, Chen, P, Liu, T, et al. (2013). Identification of downstream components of ubiquitin-conjugating enzyme PHOSPHATE2 by quantitative membrane proteomics in *Arabidopsis* roots. *The Plant Cell* 25(10), pp. 4044–4060.

Hürlimann, HC, Pinson, B, Stadler-Waibel, M, Zeeman, SC, and Freimoser, FM (2009). The SPX domain of the yeast low-affinity phosphate transporter Pho90 regulates transport activity. *EMBO Reports* 10(9), pp. 1003–1008.

Inan, G, Zhang, Q, Li, P, Wang, Z, Cao, Z, Zhang, H, Zhang, C, Quist, TM, Goodwin, SM, Zhu, J, et al. (2004). Salt cress. A halophyte and cryophyte *Arabidopsis* relative model system and its applicability to molecular genetic analyses of growth and development of extremophiles. *Plant Physiology* 135(3), pp. 1718–1737.

Jacobson, GL and Birks, HJB (1980). Soil development on recent end Moraines of the Klutlan Glacier, Yukon Territory, Canada? *Quaternary Research* 14(1), pp. 87–100.

Johnson, JF, Allan, DL, and Vance, CP (1994). Phosphorus stress-induced proteoid roots show altered metabolism in Lupinus albus. *Plant Physiology* 104(2), pp. 657–665.

Jones-Rhoades, MW, Bartel, DP, and Bartel, B (2006). MicroRNAs and their regulatory roles in plants. *Annual Review of Plant Biology* 57, pp. 19–53.

Karthikeyan, AS, Varadarajan, DK, Jain, A, Held, MA, Carpita, NC, and Raghothama, KG (2007). Phosphate starvation responses are mediated by sugar signaling in *Arabidopsis*. *Planta* 225(4), pp. 907–918.

Kataoka, T, Watanabe-Takahashi, A, Hayashi, N, Ohnishi, M, Mimura, T, Buchner, P, Hawkesford, MJ, Yamaya, T, and Takahashi, H (2004). Vacuolar sulfate transporters are essential determinants controlling internal distribution of sulfate in *Arabidopsis*. *The Plant Cell* 16(10), pp. 2693–2704.

Katoh, K, Misawa, K, Kuma, K, and Miyata, T (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* 30(14), pp. 3059–3066.

Kelly, AA and Dörmann, P (2002). DGD2, an *Arabidopsis* gene encoding a UDP-galactose-dependent digalactosyldiacylglycerol synthase is expressed during growth under phosphate-limiting conditions. *Journal of Biological Chemistry* 277(2), pp. 1166–1173.

Khan, GA, Vogiatzaki, E, Glauser, G, and Poirier, Y (2016). Phosphate deficiency induces the jasmonate pathway and enhances resistance to insect herbivory. *Plant Physiology* 171(1), pp. 632–644.

Kobayashi, K, Masuda, T, Takamiya, K, and Ohta, H (2006). Membrane lipid alteration during phosphate starvation is regulated by phosphate signaling and auxin/cytokinin cross-talk. *The Plant Journal* 47(2), pp. 238–248.

Kumar, S, Verma, S, and Trivedi, PK (2017). Involvement of small RNAs in phosphorus and sulfur sensing, signaling and stress: current update. *Frontiers in Plant Science* 8, p. 285.

Lee, Y, Jeon, K, Lee, J, Kim, S, and Kim, VN (2002). MicroRNA maturation: stepwise processing and subcellular localization. *The EMBO Journal* 21(17), pp. 4663–4670.

Leggewie, G (1997). Two cDNAs from potato are able to complement a phosphate uptake-deficient yeast mutant: identification of phosphate transporters from higher plants. *The Plant Cell* 9(3), pp. 381–392.

Leinonen, R, Sugawara, H, Shumway, M, and Collaboration, International Nucleotide Sequence Database (2010). The sequence read archive. *Nucleic Acids Research* 39(suppl_1), pp. D19–D21.

Li, H, Handsaker, B, Wysoker, A, Fennell, T, Ruan, J, Homer, N, Marth, G, Abecasis, G, and Durbin, R (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25(16), pp. 2078–2079.

Li, M, Welti, R, and Wang, X (2006). Quantitative profiling of *Arabidopsis* polar glycerolipids in response to phosphorus starvation. Roles of phospholipases Dζ1 and Dζ2 in phosphatidylcholine hydrolysis and digalactosyldiacylglycerol accumulation in phosphorus-starved plants. *Plant Physiology* 142(2), pp. 750–761.

Lin, S, Santi, C, Jobet, E, Lacut, E, El Kholti, N, Karlowski, WM, Verdeil, J, Breitler, JC, Périn, C, Ko, S, et al. (2010). Complex regulation of two target genes encoding SPX-MFS proteins by rice miR827 in response to phosphate starvation. *Plant and Cell Physiology* 51(12), pp. 2119–2131.

Liu, C, Muchhal, US, Uthappa, M, Kononowicz, AK, and Raghothama, KG (1998). Tomato Phosphate Transporter Genes Are Differentially Regulated in Plant Tissues by Phosphorus. *Plant Physiology* 116(1), pp. 91–99.

Liu, F, Chang, X, Ye, Y, Xie, W, Wu, P, and Lian, X (2011). Comprehensive sequence and whole-life-cycle expression profile analysis of the phosphate transporter gene family in rice. *Molecular Plant* 4(6), pp. 1105–1122.

Liu, T, Huang, T, Yang, S, Hong, Y, Huang Sand Wang, F, Chiang, S, Tsai, Sh, Lu, W, and Chiou, T (2016). Identification of plant vacuolar transporters mediating phosphate storage. *Nature Communications* 7(1), pp. 1–11.

Liu, Y and Aebersold, R (2016). The interdependence of transcript and protein abundance: new data–new complexities. *Molecular Systems Biology* 12(1), p. 856.

López-Bucio, J, Hernández-Abreu, E, Sánchez-Calderón, L, Nieto-Jacobo, MF, Simpson, J, and Herrera-Estrella, L (2002). Phosphate availability alters architecture and causes changes in hormone sensitivity in the *Arabidopsis* root system. *Plant Physiology* 129(1), pp. 244–256.

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

Martín, AC, Del Pozo, JC, Iglesias, J, Rubio, V, Solano, R, De La Peña, A, Leyva, A, and Paz-Ares, J (2000). Influence of cytokinins on the expression of phosphate starvation responsive genes in *Arabidopsis*. *The Plant Journal* 24(5), pp. 559–567.

Maruyama-Nakashita, A, Nakamura, Y, Tohge, T, Saito, K, and Takahashi, H (2006). *Arabidopsis* SLIM1 is a central transcriptional regulator of plant sulfur response and metabolism. *The Plant Cell* 18(11), pp. 3235–3251.

Misson, J, Raghothama, KG, Jain, A, Jouhet, J, Block, MA, Bligny, R, Ortet, P, Creff, A, Somerville, S, Rolland, N, et al. (2005). A genome-wide transcriptional analysis using *Arabidopsis thaliana* Affymetrix gene chips determined plant responses to phosphate deprivation. *Proceedings of the National Academy of Sciences* 102(33), pp. 11934–11939.

Miura, K, Rus, A, Sharkhuu, A, Yokoi, S, Karthikeyan, AS, Raghothama, KG, Baek, D, Koo, YD, Jin JB Bressan, RA, et al. (2005). The *Arabidopsis* SUMO E3 ligase SIZ1 controls phosphate deficiency responses. *Proceedings of the National Academy of Sciences* 102(21), pp. 7760–7765.

Młodzińska, E and Zboińska, M (2016). Phosphate uptake and allocation–a closer look at *Arabidopsis thaliana* L. and *Oryza sativa* L. *Frontiers in Plant Science* 7, p. 1198.

Mudge, SR, Rae, AL, Diatloff, E, and Smith, FW (2002). Expression analysis suggests novel roles for members of the Pht1 family of phosphate transporters in *Arabidopsis*. *The Plant Journal* 31(3), pp. 341–353.

Nakamura, Y (2013). Phosphate starvation and membrane lipid remodeling in seed plants. *Progress in Lipid Research* 52(1), pp. 43–50.

Nakamura, Y (2017). Plant phospholipid diversity: emerging functions in metabolism and protein–lipid interactions. *Trends in Plant Science* 22(12), pp. 1027–1040.

Nakamura, Y, Awai, K, Masuda, T, Yoshioka, Y, Takamiya, K, and Ohta, H (2005). A novel phosphatidylcholine-hydrolyzing phospholipase C induced by phosphate starvation in *Arabidopsis*. *Journal of Biological Chemistry* 280(9), pp. 7469–7476.

Nakamura, Y, Koizumi, R, Shui, G, Shimojima, M, Wenk, MR, Ito, T, and Ohta, H (2009). *Arabidopsis* lipins mediate eukaryotic pathway of lipid metabolism and cope critically with phosphate starvation. *Proceedings of the National Academy of Sciences* 106(49), pp. 20978–20983.

O'Brien, J, Hayder, H, Zayed, Y, and Peng, C (2018). Overview of microRNA biogenesis, mechanisms of actions, and circulation. *Frontiers in Endocrinology* 9, p. 402.

O'Leary, NA, Wright, MW, Brister, JR, Ciufo, S, Haddad, D, McVeigh, R, Rajput, B, Robbertse, B, Smith-White, B, Ako-Adjei, D, et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research* 44(D1), pp. D733–D745.

O'Rourke, JA, Yang, SS, Miller, SS, Bucciarelli, B, Liu, J, Rydeen, A, Bozsoki, Z, Uhde-Stone, C, Tu, ZJ, Allan, D, et al. (2013). An RNA-Seq transcriptome analysis of orthophosphate-deficient white lupin reveals novel insights into phosphorus acclimation in plants. *Plant Physiology* 161(2), pp. 705–724.

Oh, D, Dassanayake, M, Bohnert, HJ, and Cheeseman, JM (2013). Life at the extreme: lessons from the genome. *Genome Biology* 13(3), pp. 1–9.

Oono, Y, Kobayashi, F, Kawahara, Y, Yazawa, T, Handa, H, Itoh, T, and Matsumoto, T (2013a). Characterisation of the wheat (*Triticum aestivum L.*) transcriptome by de novo assembly for the discovery of phosphate starvation-responsive genes: gene expression in Pi-stressed wheat. *BMC Genomics* 14(1), p. 77.

Oono, Y, Kawahara, Y, Yazawa, T, Kanamori, H, Kuramata, M, Yamagata, H, Hosokawa, S, Minami, H, Ishikawa, S, Wu, J, et al. (2013b). Diversity in the complexity of phosphate starvation transcriptomes among rice cultivars based on RNA-Seq profiles. *Plant Molecular Biology* 83(6), pp. 523–537.

Park, MR, Baek, S, Benildo, G, Yun, SJ, and Hasenstein, KH (2012). Transcriptome profiling characterizes phosphate deficiency effects on carbohydrate metabolism in rice leaves. *Journal of Plant Physiology* 169(2), pp. 193–205.

Pertea, G and Pertea, M (2020). GFF Utilities: GffRead and GffCompare. *F1000Research* 9.

Plaxton, WC (1996). The organization and regulation of plant glycolysis. *Annual Review of Plant Biology* 47(1), pp. 185–214.

Plaxton, WC and Carswell, MC (1999). Metabolic aspects of the phosphate starvation response in plants. *Plant responses to environmental stresses: from phytohormones to genome reorganization*, pp. 349–372.

Plaxton, WC and Tran, HT (2011). Metabolic adaptations of phosphate-starved plants. *Plant Physiology* 156(3), pp. 1006–1015.

Poirier, Y, Thoma, S, Somerville, C, and Schiefelbein, J (1991). Mutant of *Arabidopsis* deficient in xylem loading of phosphate. *Plant Physiology* 97(3), pp. 1087–1093.

Puga, MI, Mateos, I, Charukesi, R, Wang, Z, Franco-Zorrilla, JM, Lorenzo, L de, Irigoyen, ML, Masiero, S, Bustos, R, Rodríguez, J, et al. (2014). SPX1 is a phosphate-dependent inhibitor of Phosphate Starvation Response 1 in *Arabidopsis*. *Proceedings of the National Academy of Sciences* 111(41), pp. 14947–14952.

R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria.

Raghothama, KG (1999). Phosphate Acquisition. *Annual Review of Plant Physiology and Plant Molecular Biology* 50(1), pp. 665–693.

Raghothama, KG (2000). Phosphate transport and signaling. *Current Opinion in Plant Biology* 3(3), pp. 182–187.

Rausch, C and Bucher, M (2002). Molecular mechanisms of phosphate transport in plants. *Planta* 216(1), pp. 23–37.

Remy, E, Cabrito, TR, Batista, RA, Teixeira, MC, Sá-Correia, I, and Duque, P (2012). The PHT1;9 and PHT1;8 transporters mediate inorganic phosphate acquisition by the *Arabidopsis thaliana* root during phosphorus starvation. *New Phytologist* 195(2), pp. 356–371.

Ren, P, Meng, Y, Li, B, Ma, X, Si, E, Lai, Y, Wang, J, Yao, L, Yang, K, Shang, X, et al. (2018). Molecular mechanisms of acclimatization to phosphorus starvation and recovery underlying full-length transcriptome profiling in barley (*Hordeum vulgare L.*) *Frontiers in Plant Science* 9, p. 500.

Richterich, P (1998). Estimation of errors in raw DNA sequences: a validation study. *Genome Research* 8(3), pp. 251–259.

Rinerson, CI, Rabara, RC, Tripathi, P, Shen, QJ, and Rushton, PJ (2015). The evolution of WRKY transcription factors. *BMC Plant Biology* 15(1), pp. 1–18.

Rosales-Loessener, F (1989). Management of red tides and paralytic shellfish poisoning in Guatemala. *Biology, Epidemiology and Management of Pyrodinium Red Tides*, p. 153.

Rouached, H (2011). Multilevel coordination of phosphate and sulfate homeostasis in plants. *Plant Signaling & Behavior* 6(7), pp. 952–955.

Rouached, H, Secco, D, Arpat, B, and Poirier, Y (2011). The transcription factor PHR1 plays a key role in the regulation of sulfate shoot-to-root flux upon phosphate starvation in *Arabidopsis*. *BMC Plant Biology* 11(1), p. 19.

Rubio, V, Linhares, F, Solano, R, Martín, AC, Iglesias, J, Leyva, A, and Paz-Ares, J (2001). A conserved MYB transcription factor involved in phosphate starvation signaling both in vascular plants and in unicellular algae. *Genes & Development* 15(16), pp. 2122–2133.

Salzman, RA, Fujita, T, Zhu-Salzman, K, Hasegawa, PM, and Bressan, RA (1999). An improved RNA isolation method for plant tissues containing high levels of phenolic compounds or carbohydrates. *Plant Molecular Biology Reporter* 17(1), pp. 11–17.

Schindler, DW (1971). Carbon, nitrogen, and phosphorus and the eutrophication of freshwater lakes 1. *Journal of Phycology* 7(4), pp. 321–329.

Schwender, J, König, C, Klapperstück, M, Heinzel, N, Munz, E, Hebbelmann, I, Hay, JO, Denolf, P, De Bodt, S, Redestig, H, et al. (2014). Transcript abundance on its own cannot be used to infer fluxes in central metabolism. *Frontiers in Plant Science* 5, p. 668.

Secco, D, Jabnoune, M, Walker, H, Shou, H, Wu, P, Poirier, Y, and Whelan, J (2013). Spatio-temporal transcript profiling of rice roots and shoots in response to phosphate starvation and recovery. *The Plant Cell* 25(11), pp. 4285–4304.

Sfanovic, A, Arpat, AB, Bligny, R, Gout, E, Vidoudez, C, Bensimon, M, and Poirier, Y (2011). Over-expression of PHO1 in *Arabidopsis* leaves reveals its role in mediating phosphate efflux. *The Plant Journal* 66(4), pp. 689–699.

Shaw, GR, Moore, DP, and Garnett, C (2003). Eutrophication and algal blooms. *Environmental and Ecological Chemistry* 2.

Shin, H, Shin, H, Dewbre, GR, and Harrison, MJ (2004). Phosphate transport in *Arabidopsis*: PHT1;1 and PHT1;4 play a major role in phosphate acquisition from both low-and high-phosphate environments. *The Plant Journal* 39(4), pp. 629–642.

Simopoulos, CMA (2019). Using machine learning to predict long non-coding RNAs and exploring their evolutionary patterns and prevalence in plant transcriptomes. PhD thesis.

Simopoulos, CMA, Weretilnyk, EA, and Golding, GB (2018). Prediction of plant lncRNA by ensemble machine learning classifiers. *BMC Genomics* 19(1), p. 316.

Simopoulos, CMA, MacLeod, MJR, Irani, S, Sung, WWL, Champigny, MJ, Summers, PS, Golding, GB, and Weretilnyk, EA (2020). Coding and long non-coding RNAs provide evidence of distinct transcriptional reprogramming for two ecotypes of the extremophile plant Eutrema salsugineum undergoing water deficit stress. *BMC Genomics* 21(1), pp. 1–16.

Sobkowiak, L, Bielewicz, D, Małecka, E, Jakobsen, I, Albrechtsen, M, Szweykowska-Kulinska, Z, and Pacak, AM (2012). The role of the P1BS element containing promoter-driven genes in P_i transport and homeostasis in plants. *Frontiers in Plant Science* 3, p. 58.

Stamatakis, A (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9), pp. 1312–1313.

Stewart, JWB and Tiessen, H (1987). Dynamics of soil organic phosphorus. *Biogeochemistry* 4(1), pp. 41–60.

Supek, F, Bošnjak, M, Škunca, N, and Šmuc, T (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *PloS One* 6(7).

Takahashi, H, Watanabe-Takahashi, A, Smith, FW, Blake-Kalff, M, Hawkesford, MJ, and Saito, K (2000). The roles of three functional sulphate transporters involved in uptake and translocation of sulphate in *Arabidopsis thaliana*. *The Plant Journal* 23(2), pp. 171–182.

Takahashi, H, Kopriva, S, Giordano, M, Saito, K, and Hell, R (2011). Sulfur assimilation in photosynthetic organisms: molecular functions and regulations of transporters and assimilatory enzymes. *Annual Review of Plant Biology* 62, pp. 157–184.

Taylor, CB and Green, PJ (1991). Genes with homology to fungal and S-gene RNases are expressed in *Arabidopsis thaliana*. *Plant Physiology* 96(3), pp. 980–984.

Taylor, CB, Bariola, PA, Raines, RT, Green, PJ, et al. (1993). RNS2: a senescence-associated RNase of *Arabidopsis* that diverged from the S-RNases before speciation. *Proceedings of the National Academy of Sciences* 90(11), pp. 5118–5122.

Theodorou, ME and Plaxton, WC (1993). Metabolic adaptations of plant respiration to nutritional phosphate deprivation. *Plant Physiology* 101(2), pp. 339–344.

Tisdale, SL and Nelson, WL (1975). *Soil Fertility and Fertilizers Third Edition*. 3rd ed. Macmillan Publishing Co., Inc. 196–202.

Trainer, VL and Baden, DG (1999). High affinity binding of red tide neurotoxins to marine mammal brain. *Aquatic Toxicology* 46(2), pp. 139–148.

Tran, HT, Hurley, BA, and Plaxton, WC (2010). Feeding hungry plants: the role of purple acid phosphatases in phosphate nutrition. *Plant Science* 179(1-2), pp. 14–27.

Ullrich-Eberius, CI, Novacky, A, Fischer, E, and Lüttge, U (1981). Relationship between energy-dependent phosphate uptake and the electrical membrane potential in *Lemna gibba* G1. *Plant Physiology* 67(4), pp. 797–801.

Velasco, VME (2017). Traits underlying phosphorus use by the extremophyte *Eutrema salsugineum*. PhD thesis.

Velasco, VME, Mansbridge, J, Bremner, S, Carruthers, K, Summers, PS, Sung, WWL, Champigny, MJ, and Weretilnyk, EA (2016). Acclimation of the crucifer *Eutrema salsugineum* to phosphate limitation is associated with constitutively high expression of phosphate-starvation genes. *Plant, Cell & Environment* 39(8), pp. 1818–1834.

Velasco, VME, Irani, S, Axakova, A, Silva, R da, Summers, PS, and Weretilnyk, EA (2020). Evidence that tolerance of *Eutrema salsugineum* to low phosphate conditions is hard-wired by constitutive metabolic and root-associated adaptations. *Planta* 251(1), p. 18.

Wan, C and Wilkins, TA (1994). A modified hot borate method significantly enhances the yield of high-quality RNA from cotton (*Gossypium hirsutum L.*) *Analytical Biochemistry* 223(1), pp. 7–12.

Wang, D, Lv, S, Jiang, P, and Li, Y (2017). Roles, regulation, and agricultural application of plant phosphate transporters. *Frontiers in Plant Science* 8, p. 817.

Wang, H, Xu, Q, Kong Yand Chen, Y, Duan, J, Wu, W, and Chen, Yi (2014a). *Arabidopsis* WRKY45 transcription factor activates PHOSPHATE TRANSPORTER1;1 expression in response to phosphate starvation. *Plant Physiology* 164(4), pp. 2020–2029.

Wang, L, Park, HJ, Dasari, S, Wang, S, Kocher, J, and Li, W (2013). CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. *Nucleic Acids Research* 41(6), e74–e74.

Wang, Y, Lysøe, E, Armarego-Marriott, T, Erban, A, Paruch, L, Van Eerde, A, Bock, R, and Liu-Clarke, J (2018). Transcriptome and metabolome analyses provide insights into root and root-released organic anion responses to phosphorus deficiency in oat. *Journal of Experimental Botany* 69(15), pp. 3759–3771.

Wang, Z, Ruan, W, Shi, J, Zhang, L, Xiang, D, Yang, C, Li, C, Wu, Z, Liu, Y, Yu, Y, et al. (2014b). Rice SPX1 and SPX2 inhibit phosphate starvation responses through interacting with PHR2 in a phosphate-dependent manner. *Proceedings of the National Academy of Sciences* 111(41), pp. 14953–14958.

Wasaki, J, Yonetani, R, Kuroda, S, Shinano, T, Yazaki, J, Fujii, F, Shimbo, K, Yamamoto, K, Sakata, K, Sasaki, T, et al. (2003). Transcriptomic analysis of metabolic changes by phosphorus stress in rice plant roots. *Plant, Cell & Environment* 26(9), pp. 1515–1523.

Weiss, CM (1969). Relation of phosphates to eutrophication. *Journal-American Water Works Association* 61(8), pp. 387–391.

Wimalasekera, R, Pejchar, P, Holk, A, Martinec, J, and Scherer, GFE (2010). Plant phosphatidylcholine-hydrolyzing PHOSPHOLIPASES C NPC3 and NPC4 with roles in root development and brassinolide signaling in *Arabidopsis thaliana*. *Molecular Plant* 3(3), pp. 610–625.

Wong, CE, Li, Y, Whitty, BR, Diaz-Camino, C, Akhter, SR, Brandle, JE, Golding, GB, Weretilnyk, EA, Moffatt, BA, and Griffith, M (2005). Expressed sequence tags from the Yukon ecotype of *Thellungiella* reveal that gene expression in response to cold, drought and salinity shows little overlap. *Plant Molecular Biology* 58(4), pp. 561–574.

Wu, P, Ma, L, Hou, X, Wang, M, Wu, Y, Liu, F, and Deng, XW (2003). Phosphate starvation triggers distinct alterations of genome expression in *Arabidopsis* roots and leaves. *Plant Physiology* 132(3), pp. 1260–1271.

Xu, J, Wang, Q, Freeling, M, Zhang, X, Xu, Y, Mao, Y, Tang, X, Wu, F, Lan, H, Cao, M, et al. (2017). Natural antisense transcripts are significantly involved in regulation of drought stress in maize. *Nucleic Acids Research* 45(9), pp. 5126–5141.

Yang Rand Jarvis, DJ, Chen, H, Beilstein, M, Grimwood, J, Jenkins, J, Shu, S, Prochnik, S, Xin, M, Ma, C, et al. (2013a). The reference genome of the halophytic plant *Eutrema salsugineum*. *Frontiers in Plant Science* 4, p. 46.

Yang, G, Du, H, Xu, H, and Ding Gand Xu, F (2013b). Identification of phosphate-starvation-inducible gene *BnIPS1* in *Brassica napus*. *Acta Physiologiae Plantarum* 35(7), pp. 2085–2094.

Yang, S, Zhao, Y, and Wang, J (2020a). Function and application of the Eutrema salsugineum PHT1;1 gene in phosphate deficiency stress. *Plant Biology*. To appear.

Yang, S, Feng, Y, Zhao, Y, Bai, J, and Wang, J (2020b). Overexpression of a Eutrema salsugineum phosphate transporter gene EsPHT1;4 enhances tolerance to low phosphorus stress in soybean. *Biotechnology Letters*, pp. 1–15.

Yang, XJ and Finnegan, PM (2010). Regulation of phosphate starvation responses in higher plants. *Annals of Botany* 105(4), pp. 513–526.

Yao, Z, Liang, C, Zhang, Q, Chen, Z, Xiao, B, Tian, J, and Liao, H (2014). SPX1 is an important component in the phosphorus signalling network of common bean regulating root growth and phosphorus homeostasis. *Journal of Experimental Botany* 65(12), pp. 3299–3310.

Yoshimoto, N, Inoue, E, Saito, K, Yamaya, T, and Takahashi, H (2003). Phloem-localizing sulfate transporter, SULTR1;3, mediates re-distribution of sulfur from source to sink organs in *Arabidopsis*. *Plant Physiology* 131(4), pp. 1511–1517.

Yuan, J, Zhang, Y, Dong, J, Sun, Y, Lim, BL, Liu, D, and Lu, ZJ (2016). Systematic characterization of novel lncRNAs responding to phosphate starvation in *Arabidopsis thaliana*. *BMC Genomics* 17(1), p. 655.

Zhang, D, Zhang, H, Chu, S, Li, H, Chi, Y, Triebwasser-Freese, D, Lv, H, and Yu, D (2017). Integrating QTL mapping and transcriptomics identifies candidate genes underlying QTLs associated with soybean tolerance to low-phosphorus stress. *Plant Molecular Biology* 93(1-2), pp. 137–150.

Zhang, J, Zhou, X, Xu, Y, Yao, M, Xie, F, Gai, J, Li, Y, and Yang, S (2016). Soybean SPX1 is an important component of the response to phosphate deficiency for phosphorus homeostasis. *Plant Science* 248, pp. 82–91.

Zhao, H, Frank, T, Tan, Y, Zhou, C, Jabnoune, M, Arpat, AB, Cui, H, Huang, J, He, Z, Poirier, Y, et al. (2016). Disruption of Os SULTR 3;3 reduces phytate and phosphorus concentrations and alters the metabolite profile in rice grains. *New Phytologist* 211(3), pp. 926–939.

Zhao, P, Wang, L, and Yin, H (2018). Transcriptional responses to phosphate starvation in *Brachypodium distachyon* roots. *Plant Physiology and Biochemistry* 122, pp. 113–120.