THESIS

BUFFERING THE EFFECTS OF A CHANGING CLIMATE: *SALSOLA TRAGUS* AS A POTENTIAL SOURCE OF STRESS TOLERANCE GENES

Submitted by

John M. Lemas

Department of Agricultural Biology

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2024

Master's Committee:

Advisor: Todd Gaines

Cynthia Brown
James Henriksen

ABSTRACT

BUFFERING THE EFFECTS OF A CHANGING CLIMATE:

SALSOLA TRAGUS AS A POTENTIAL SOURCE OF STRESS TOLERANCE GENES

The tumbleweed *Salsola tragus* is an allotetraploid C4 weedy member of the *Salsola* polyploid complex. Commonly referred to as Russian thistle, it develops a thorny habit during inflorescence, and commonly separates at an abscission layer near the soil to form a tumbleweed. This species is economically important to all land use types and is especially impactful in the Northwestern United States where it affects spring cereal production. The International Weeds Genomics Consortium recently completed a fully annotated reference genome assembly for each of the sub genomes in the somatic cells of this allotetraploid. Polyploids, in general, are overrepresented in the most troublesome weeds globally, and *Salsola tragus* is no exception. Recurrent formation of polyploids, increased activity of transposable elements, and increased mutation rates that follow genome duplication may lead to the *de novo* formation and selection of novel highly adapted alleles over time. We utilized the reference genome assembly for this species to align a stress-response transcriptome to investigate how this species responded to two selected abiotic stressors. Many expected response pathways are represented, including response to stress phytohormones, sodium-proton antiporters, calcium exchangers, and cold-responsive binding factors. In addition, several uncharacterized proteins were differentially overexpressed in the shoot and root tissues of this species. Identified genes from this species may present novel alleles for osmotic and temperature stress tolerance. Uncharacterized genes may represent novel stress response genes and can be used to improve the provided reference annotation for this species.

These genes of interest may provide the scientific community with additional genomic resources to bolster crop production in this era of climate change.

ACKNOWLEDGEMENTS

I would first and foremost like to acknowledge my wife, Stephanie Lemas, for all the encouragement, patience, understanding, and support during this difficult time in our lives. I doubt many married couples can say they completed graduate school at the same time! How amazing it was to be in the same graduation! I would also like to acknowledge my father, Mark Lemas, and my mother, Jan Lemas, for their continued advice and support. I wouldn't be who I am today without their guidance. Thanks, of course, are owed to my two hilarious dogs, Yosef and Lucy, who have served as furry emotional support friends for the both of us.

Next, I absolutely must thank Dr. Phil Westra for seeing something in me that I did not see in myself. Phil picked me up from my day job as a Starbucks barista and pulled me into the world of Weed Science. Phil's continued encouragement and professional advice has been indispensable in the completion of this Master's thesis. Also, of course, gratitude towards Dr. Todd Gains for taking a chance on me when I had absolutely no experience in molecular plant sciences. It's been an honor to work with you, Todd. I look forward to additional future collaborations.

I'd like to thank the folks, past and present, in the Colorado State University Weed Science Laboratory for showing me the ropes and teaching me the basics of this highly impactful field of science. Thanks to Jake Montgomery and Sofia Marquez-Hill for your assistance with my many computational questions. Thanks to Luan, Carlos, and Srishti for introducing me to proper wet-lab procedures. Cheers to my new friends, André Lucas Simões Araujo and Amber Pelon, for helping me through the imposter syndrome and reminding me that my research is important. I'd also like to thank Dr. Franck Dayan for reminding me that it is still important for professionals to have fun!

DEDICATION

To Dr. Philip Westra for being a positive influence on so many researchers through the course of your career. Your contributions to this impactful field of science are immeasurable.

TABLE OF CONTENTS

CHAPTER 1: INVASION ECOLOGY OF POLYPLOID SPECIES

SUMMARY

This chapter reviews our understanding of the implications of genome duplication, and the potential for whole genome duplication to impart increased invasiveness in polyploid species. The process of genome duplication affects inheritance and increases nuclear volume, resulting in biochemical and enzymatic diversification. The increased action of novel allele generation along with the reduction in inbreeding depression may allow for the widespread prevalence of polyploids as weedy species.

INTRODUCTION

Many of the world's most troublesome weeds have one specific genomic similarity: they retained both complete parental genomes through the combination of unreduced gametes in a process termed polyploidization. In this process, meiotic mistakes during gamete formation drastically increase the number of alleles per gene locus in following generations. This process has resounding effects on these polyploid progeny, some advantageous and some deleterious (Baduel et al., 2018; Qiu et al., 2020; Te Beest et al., 2012).

Polyploids have been studied and characterized for over a century (Digby, 1912), and early work linked polyploidy to increased fitness in adverse environmental conditions (Stephens, 1951; Treier et al., 2009). The recognition of polyploids as common invaders and the understanding that these invaders may present increased allelic diversity on an organismal level may connect invasion success with the cytotype of the organism.

This review focuses on an individual in the *Salsola* polyploid complex, or the tumbleweed *Salsola tragus* (Russian thistle). This highly successful species has been present in North America for roughly 150 years and has invaded a wide range of environments (Beckie & Francis, 2009;

Mosyakin, 1996). The *Salsola* complex contains cytotypes ranging from eighteen to fifty-four chromosomes. In addition, new cytotype combinations are still being detected (Ryan & Ayres, 2000). An important question is: what makes this species so able to survive and adapt across the globe in such a wide latitudinal breadth? Does the presence of multiple genomes contribute to the invasive success of *Salsola* ssp? This review seeks to address these questions through a synthesis of understandings in the scientific community around the role of polyploidization as a mechanism of successful invasion. The hypothesis is that polyploidization allows for greater niche range and a heightened competitive advantage over native diploids when exposed to fluid ecological and environmental conditions. Alternatively, perhaps polyploidization has nothing to do with invasion success and this commonality is due to other evolutionary and ecological factors, i.e. polyploidization correlates with invasiveness but does not necessarily cause invasiveness. In either case, investigations into the role of polyploidization in invasion success will prove beneficial for understanding the *Salsola* species complex.

REVIEW

1.1 Genome duplication and inheritance

Genome duplication arises from improper chromosome segregation during mitosis or through the fusion of unreduced gametes that formed due to errors in the complex process of meiosis and cytokinesis. In meiosis, a potential disruption in anaphase I spills into subsequent telophase and cytokinesis resulting in the formation of a 2n gamete (Jackson, 1976). Improper meiosis may be connected to external stimuli such as temperature or wounding, or to mutations in the many genes that regulate meiosis. As such, environments that increase exposure to these potential meiosis-impairing stimuli may result in increased polyploidization. Finally, genome duplication through mitotic/meiotic error occurs in all forms of life in both somatic and gametic cells (Jackson, 1976; Ramsey & Schemske, 1998).

Under normal cytotype conditions, specific genes regulate synapsis and recombination between homologous chromosomes. This bivalent process regulates chiasma interactions so that chromosomes preferentially pair with their respective homolog. This bivalent interaction is referred to as disomic inheritance. However, in the presence of additional chromosomes, either bivalent or multivalent interactions may take place. In this instance, the nucleus contains both homologous chromosomes and homoeologous chromosomes. Homologous chromosomes are homologs that originated from one parent genome. Homoeologous chromosomes are those that represent homologs from the other genome. Preferential synapsis may occur normally (bivalently) if the combined genomes are dissimilar enough, i.e. if the homoeologous chromosomes are distinguishable during preferential synapsis and recombination. In this instance, disomic inheritance may occur and statistical processes that follow diploid allelic variability may be applied. However, if the combined genomes are highly similar, the process of preferential synapsis may break down, allowing for synapsis and recombination to take place between homoeologous chromosomes. Although the latter is more common in the combination of conspecific genomes (autopolyploids), either process can occur in each cytotype (Ramsey & Schemske, 1998; Stift et al., 2008).

Polyploids may arise from either interspecific genome combination (allopolyploid) or through conspecific genome doubling (autopolyploid). There are many implications that must be overcome in a newly formed autopolyploid. In an allopolyploid, although an intermediate inheritance pattern may occasionally ensue (Stift et al., 2008), the individuals will commonly follow bivalent chromosomal interactions. Thus, the mechanisms of synapsis and recombination are less affected in allo- versus autopolyploids. Autopolyploids must overcome the potentially deleterious presence of polyvalence during synapsis and recombination. Eventually,

recombination between homoeologous chromosomes may begin to establish preferential synapsis after a few generations (Baduel et al., 2018).

Polyploid meiotic processes alone provide major contributions to the variability in polyploid progeny. If the mechanistic and genetic hurdles of genome duplication can be overcome, polyploidization bolsters novel allele frequency by increasing mutation rates. Increased rates of mutation are attributed to elevated transposable element activity, which correlates with the duplication of genes inherent in polyploid genomes (Levin, 1983; Te Beest et al., 2012). Initially, these mechanisms of meiotic destabilization may break down sexual reproduction until genomic stability can be re-established. During this transitionary time, polyploids may avoid sexual reproduction for several generations through vegetative reproduction (cloning) or apomixis (the generation of propagule without meiosis and fertilization). Eventually, polyploids will return to sexual reproduction to further increase genetic diversity through synapsis and recombination (Baduel et al., 2018).

Once meiotic stability is established, sexual reproduction can take place creating a myriad of diverse genotypes and cytotypes. Sexual reproduction between polyploids may yield additional polyploids of the same cytotype, or potentially new cytotypes. The fusion of reduced gametes from tetraploid individuals with reduced gametes of diploid individuals, however, usually result in reduced fitness or complete sterility. In an ecological context, this process is termed cytotype exclusion and operates where diploids and polyploids exchange gametes. The triploid cytotype demonstrates low reproductive success due to the odd number of chromosomes during meiosis. The combination of gametes that form a zygote with an odd number of chromosomes (triploid, pentaploid, heptaploid, etcetera) will form aneuploid gametes, or cells with monovalent chromosomes during meiosis. These monovalent chromosomes pose additional meiotic

complications. Cytotype exclusion commonly limits outcrossing between polyploids and diploids through the reduced fitness of progeny; however, this process is not uncommon and may also result in the formation of new polyploids (Baduel et al., 2018; Ramsey & Schemske, 1998; Stift et al., 2008).

In addition to these meiotic processes of genotypic diversification, it has been shown that multiple polyploid events may be occur within a population, termed recurrent formation (Soltis et al., 1993; Soltis & Soltis, 1999). Recurrent polyploid formation leads to distinct polyploid populations and enables gene flow between polyploids and their progenitors, or between genetically distinct polyploid populations. Interpopulation admixture in addition to intragenomic admixture (mixture between homoeologous chromosomes within one organism) has the potential to drastically increase progeny variability (Jackson, 1976; Leitch & Bennett, 1997). As such, the various mechanisms of evolution may operate on a drastic level in polyploid species.

One final note to discuss here is the concept of inbreeding depression and genetic load. Genetic load is the accumulation of deleterious alleles in a population and is common in founding populations that have decreased allelic diversity and increased inbreeding. This accumulation may be eliminated once individuals attain homozygosity, allowing the deleterious trait to be expressed and selected against, termed purging. Gene flow from the original population may also increase genetic diversity and thus reduce inbreeding depression depending on the level of propagule introduction into the founding population.

In polyploid populations, increasing the number of alleles per gene locus effectively masks this accumulation of genetic load. Thus, merely by presenting a genome with additional trait loci, genetic load may be masked in polyploid species. This may alleviate founder affects and inbreeding depression in introduced polyploid populations. Additionally, polyploids that follow

polysomic inheritance display fixed heterozygosity because the probability of homozygosity at any trait locus is nearly zero. Fixed heterozygosity has additional potential positive implications when considering founder effects of an introduced polyploid population. Eventually, however, genetic load will reach a maximum and fixed heterozygosity will break down. It has been suggested that this process, along with other mechanisms like gene silencing and transposable element activity, may drive polyploid drop, or a transition from polyploid to diploid (See Figure 1.2; Leitch & Bennett, 1997; Soltis & Soltis, 1999; Te Beest et al., 2012).

1.2 Implications of increased nuclear volume

Doubling the number of chromosomes in a nucleus results in a doubling of cell size and a fractional increase in cell volume (Levin, 1983). As such, the size to volume ratio of cytosolic components is heavily impacted by these alterations. In addition, expression rates are also impacted as nuclear pore density decreases with increased nuclear surface area (Levin, 1983). Decreased nuclear pore density reduces rates of mRNA entering the cytosol, and thus reduces the expression of genes overall. Thus, although polyploid individuals have more genes per loci, the expression of these genes is reduced (Jackson, 1976).

Effects of this component of genome duplication arise not only in expression, but also in enzyme activity and metabolite biosynthesis. Enzyme-substrate interactions are greatly impacted by this new cell geometry. Trends in specific enzymes appear to be species-specific; however, alterations to RNA concentrations and enzyme activity are significantly altered in polyploids when compared to diploids (Baduel et al., 2018; Jackson, 1976; Levin, 1983; Te Beest et al., 2012).

On an organismal scale, this alteration to cell size may lead to larger morphologies when compared to diploid progenitors (Te Beest et al., 2012). Overall, larger cells may prove advantageous or deleterious depending on specific circumstances (See Figure 1.1). Larger organisms typically have slower metabolisms and longer development times. This may prove

disadvantageous under direct competition between polyploids and diploids. However, longer development times may result in niche differentiation between polyploids and diploids in a community (Te Beest et al., 2012).

Polyploids also present altered reproductive tissues. Such alterations may include larger flowers, novel inflorescence morphologies, and larger seeds. These traits may present additional advantages to polyploids. Larger morphologically distinct flowers may attract additional pollinators, longer flowering times may present additional fertilization opportunities, and larger seeds may provide more nutrients to offspring. Thus, it follows that progeny from these polyploids may then exhibit increased fitness (Levin, 1983; Te Beest et al., 2012).

Slower evapotranspiration rates have been observed in polyploid species as well. Like the changes to nuclear pore density, stomata density is also somewhat even across species. Thus, increased leaf area due to larger cell size may lead to decreased stomata density and reduced evapotranspiration. Reduced evapotranspiration in polyploid species may contribute to the observed shift of non-native polyploid species towards more arid climates (Figure 1.1). These factors may help to shift niche space away from diploid progenitors, thus combatting the potential reduction in direct competition between diploids and polyploids (Jackson, 1976; Te Beest et al., 2012).

1.3 Physiological diversification

The retention of additional alleles per locus results in altered physiological characteristics. Various mechanisms of allelic diversity present newly formed polyploids with a variety of physiological alterations over a relatively short timescale. Physiological alterations include increased tolerance for a range of conditions, including insolation, drought, and salinity (Te Beest et al., 2012). These alterations may arise from biochemical diversification within an individual through the assembly of potentially novel protein subunit combinations. Assemblies of multi-

subunit proteins from homoeologous alleles may lead to alterations in enzyme efficiency or activity in polyploid organisms. Consequently, increased concentrations of secondary metabolites have been observed in polyploid species (Levin, 1983).

Studies have shown that polyploids shift towards more arid climates and can withstand higher rates of salinity (Te Beest et al., 2012). This may be due to the overproduction of cytosolic protectants that offset osmotic stress in these conditions, in addition to the alteration of cytosolic volume in larger polyploid cells. As such, physiological diversification in polyploid organisms may result in the shifting of preferred niche space. Given the potential reduction in direct competition for polyploid species, such as differential growth rates or cytotype exclusion, a differentiation of niche space may be ideal for polyploid individuals.

CONCLUSION

There are many mechanisms potentially at play in the success of non-native species. These factors include pre-adaptation, enemy escape, propagule pressure, plasticity, novel allelochemical weapons, and many more (Gioria et al., 2023). For polyploids specifically, this chapter has presented an array of factors at play when considering polyploidy and invasiveness. These factors include cytotype exclusion, the generation of novel alleles, heterozygosity, and genetic load. Each of these factors play a role in the development of polyploids as invasive species. As such, though polyploids may overrepresent invasive species, polyploidization is not a prerequisite to invasiveness. In fact, this research has outlined the potentially deleterious process of establishment as a polyploid organism. In combination with findings that polyploids may not present increased fitness *per se*, what is it about polyploids that contribute to their success?

Firstly, the process of cytotype exclusion is a potential contributing factor towards polyploids as invasive species. Although this is a deleterious mechanism in outcrossing polyploids within a diploid population, in an introduced setting this mechanism may select for polyploids.

This concept may relate to the relative abundance of reduced and unreduced gametes in an outcrossing population. If diploids are overrepresented in the native range, cytotype exclusion may operate against polyploids due to a higher abundance of reduced diploid gametes. However, if there are relatively equivalent numbers of both cytotypes, and thus an equal opportunity for cytotypes to cross, this process presents one potential mechanism for the selection of polyploidy in an introduced population.

Next, the generation of novel alleles through drastic genomic recombination has many implications in introduced settings. In general, genetic variation is the stage on which natural selection acts. As outlined above, polyploidization increases the generation of new alleles to a level that purely diploid populations may never attain. The resulting increased allelic diversity leads to potentially greater adaptive advantages (Estoup et al., 2016). Gene flow through recurrent polyploid formation provides an additional avenue for the generation of novel genotypes (Baduel et al., 2018). These genomic mechanisms, along with physiological alterations such as increased secondary metabolites, may lead to the observed niche-shift of many non-native polyploids towards more arid conditions. In addition, and potentially more advantageous in non-native scenarios, extensive allelic alterations increase the potential for a highly advantageous allelic combination to arise (McIntyre & Strauss, 2017).

For polyploids exhibiting disomic inheritance patterns, fixed heterozygosity may assist in the delay of inbreeding depression. If we consider the introduction of diploid and tetraploid individuals, founder effects will be lower in polyploids due to the introduction of multiple genes per locus per introduced individual. Studies have shown increased inbreeding depression in diploid organisms, thus giving polyploids an advantage in introduced ranges (Rosche et al., 2017). This is especially true for allopolyploids that follow disomic inheritance. Autopolyploids may not benefit

as much, but masking of deleterious alleles still operates in this system due to the inherent duplication of genes.

To conclude, the recognition of polyploids as common invasive species may be due to many cytological, biochemical, and physiological aspects that are associated with whole genome duplication. Although there are a range of negative mechanistic and genomic impacts for polyploids to overcome, this process has the potential to generate highly specialized phenotypes (McIntyre & Strauss, 2017). The process of phenotypic stabilization comes along with genomic stabilization, and if this process occurs under selective pressures the resulting species may be more fit for its realized niche that its diploid predecessors. This is because the process of genomic stabilization for polyploid species has the potential to generate allele combinations not found in a purely diploid system of disomic inheritance. To address the correlation between polyploidy and invasiveness, the resulting heightened adaptability in a stabilizing polyploid genome may drive the successful establishment of these organisms under the strong selective pressures commonly associated with novel environmental conditions or anthropogenic habitat alteration. Thus, it follows that the presence of polyploids as invasive species may not be directly due to more genes per locus, but more due to how those genes have undergone selection to create a species that attains specific adaptations to its new environment. Through this process, polyploids can sidestep cytotype exclusion or lower competitive ability by shifting realized niche space into any that are available. Physiological alterations and biochemical variation lend a hand in realizing this niche space by buffering environmental stressors and increasing resource use efficiency. Thus, through the process of genome duplication, new species may be attained that have adaptive advantages unrealized by diploid progenitors. These novel allele combinations may be a factor contributing to over-representation of polyploids in lists of the globally most troublesome and invasive weeds.

FIGURES

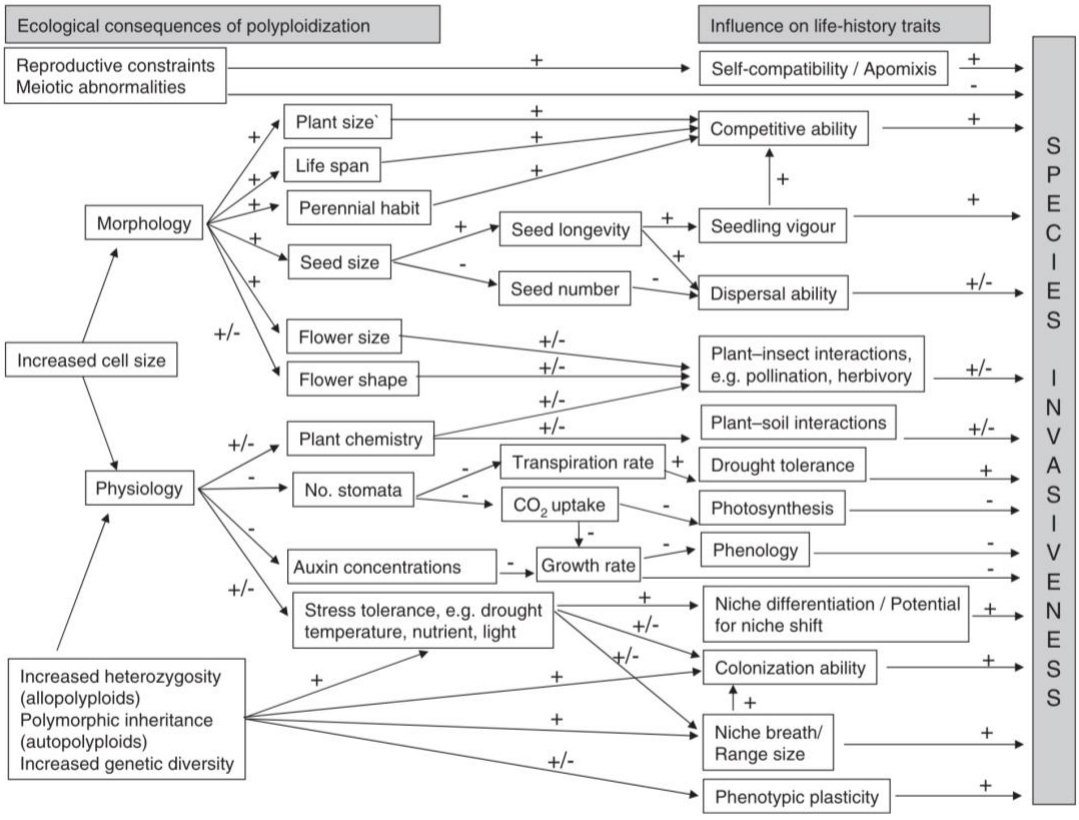


Figure 1.1: This chart outlines implications of genome duplication in invasiveness and encompasses both allopolyploids and autopolyploids. Of note here are the positive interactions between polyploidy and stress tolerance, niche differentiation, and phenotypic plasticity.
Credit: Te Beest, Le Roux, Richardson, Brysting, Suda, Kubesova, Pysek, The more the better? The role of polyploidy in facilitating plant invasions, Annals of Botany, 2012, 109, 1, 19-45, by permission of Oxford University Press.

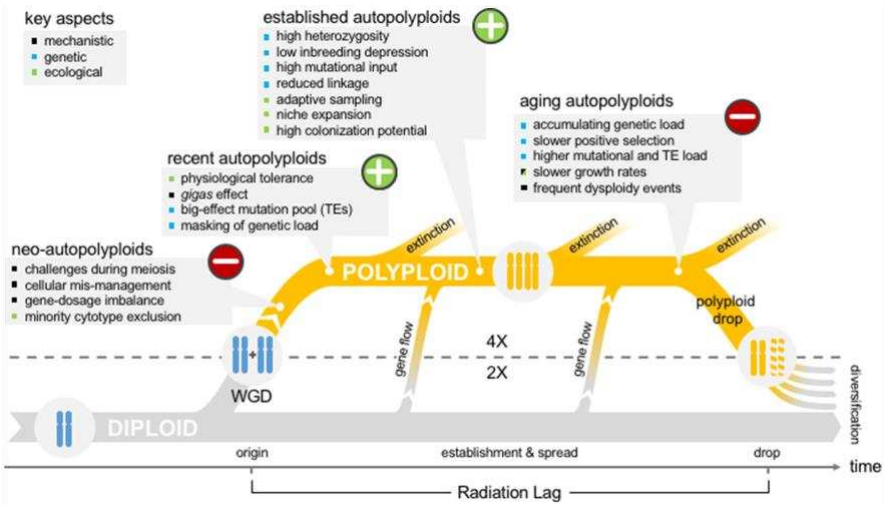


Figure 1.2: This schematic shows the proposed process of speciation and genomic and ecological implications of genome duplication. This process can lead to the formation of novel allele combinations that would not have been attained in the original diploid progenitors. The resulting diploid species following the polyploid drop diversify into those that are highly specialized for their new realized niche range. Copyright © 2018 Baduel, Bray, Vallejo-Marin, Kolář and Yant.

CHAPTER 2: ASSEMBLY AND ANNOTATION OF THE TETRAPLOID *SALSOLA TRAGUS* (RUSSIAN THISTLE) GENOME[1]

INTRODUCTION

The *Salsola* genus is a member of the Chenopodiaceae (goosefoot) family, which is a subfamily of the Amaranth family that includes beets, spinach, and quinoa, and as such is closely related to *Bassia scoparia* (Kochia) and distantly related to *Amaranthus palmeri* (Palmer amaranth) (Brignone & Denham, 2021). The *Salsola* genus contains 22 known subspecies, six or seven of which were introduced into north America in the late 1800s (Beckie & Francis, 2009; Brignone & Denham, 2021). Taxonomic understanding of this genus has been frequently revisited. Self-pollinating and outcrossing fertilization of this anemophilous species in combination with its widespread seed dispersal via a tumbleweed habit has led to intraspecific genetic variation and the formation of viable hybrid species. In addition, cryptic differences within subspecies also exist. The *Salsola* genus contains several species at differing ploidy levels (2n=18, 2n=36, 2n=56) (Ryan & Ayres, 2000) that are highly polymorphic; thus, literature subspecies designations have been frequently revised. Suffice it to say, improper taxonomic designation of this species in historical literature presents misrepresentations at the species level. *S. tragus* is synonymous with *S. australis, S. kali tenuifolia, S. kali* ssp. *ruthentica,* and *S. ruthentica. S. tragus* also has two individual identified biotypes, denoted Type A and Type B (Ryan & Ayres, 2000). Type A is akin to *S. tragus* described in Eurasia with 2n=36 chromosomes. Type B has not been found in Eurasia and is a diploid with 2n=18 chromosomes (Smith, 2005; Sobhian et al., 2003). The designation of

[1] The following authors have contributed to this section: John Lemas, Eric Patterson, Luan Cutti, Sarah Morran, Nicholas Johnson, Jacob Montgomery, Fatemeh Abdollahi, David Nelson, Victor Llaca, Kevin Fengler, Philip Westra, Todd Gaines.

Salsola tragus Type A to this assembly was identified through relevant literature and the ploidy level of this sequenced line.

Although superficially *S. tragus* resembles *B. scoparia* (Kochia), the species can be identified by several distinctive characteristics. *S. tragus* has a highly branched habit and is generally shorter than it is wide. Leaves are slender (<1mm), long (<6mm), and mainly alternate (opposite or subopposite at early development). Inflorescence are protected by opposite bracts that are swollen at the base and narrow into a sharp point. Flowers are small (4-8mm) and usually solitary though occasionally they occur in pairs or triplets. Flowers lack petals or stalks and have a pale pink to pale green color and a papery texture. *S. tragus* can be further identified as the only species of this genus with the presence of reddish-purple stripes along stems (Beckie & Francis, 2009; Mosyakin, 1996).

At higher latitudes in North America (35 – 45ºN) *S. tragus* is a summer annual, although under dry warm winter conditions or at lower latitudes this species can occur as a winter annual. *Salsola tragus* is one of the earliest emerging weed species in the United States (Beckie & Francis, 2009). At higher latitudes plants mature and flower near July and become stiff and dry through the fall. *Salsola tragus* plants at maturity either remain attached to soil or, under windy conditions, break off at an abscission layer at the soil level to become a tumbleweed. In the latter case, these tumbleweeds can distribute seeds over vast distances in relatively short periods of time. Seeds mature through August and into September, though faster maturation occurs at higher temperatures. Either on the plant or as litter, *S. tragus* seeds undergo a crucial overwintering maturation period. Optimal germination conditions are 20 °C and 9% soil moisture at a depth no greater than 6 cm (Beckie & Francis, 2009; Crompton & Bassett, 1985). Given these conditions the coiled embryos of *S. tragus* can unwind in a matter of minutes (Rhoads et. al., 1967).

Cotyledons in these coiled embryos contain photosynthetically active chloroplasts allowing seedlings to immediately begin physiological processes (Crompton & Bassett, 1985), and the mechanics of unwinding effectively drive the young taproot into the substrate. These life cycle and reproductive rate characteristics contribute to the prolific weediness of *S. tragus* across diverse ecosystems.

In general, polyploidy drives niche shifts away from diploid progenitors (Baduel et al., 2018; Te Beest et al., 2012). Non-native polyploid species such as *S. tragus* are observed to shift towards more arid climates in non-native settings (Te Beest et al., 2012). This is due to both morphological and biochemical differences that result from genome duplication. Polyploid species exhibit increased colonialization capabilities through fixed heterozygosity and a capacity to mask genetic load. Although aspects of genome duplication prove advantageous, eventually the effects of polyploidy become deleterious. It has been suggested that polyploidy is not a terminal cytotype, and that eventually polyploid species will undergo "polyploid drop". Thus, polyploidy may be a mechanism for novel allele generation and eventually acts as a speciation event. Although empirical evidence of polyploid drop has yet to be recorded, evidence suggests that it is possible (Baduel et al., 2018; Leitch & Bennett, 1997; Soltis & Soltis, 1999). The recognition of two different *S. tragus* cytotypes in North America by Ryan & Ayres (2000) may suggest that *S. tragus* has undergone such a process in North America; however, further research is needed to characterize the validity of this observation.

As an invader of native ecosystems, *Salsola tragus* impacts arid to semi-arid ruderal areas, rangeland, cropping systems, roadsides, railroads, open spaces, urban corridors, and other areas where natural ecological systems have been stressed or altered. These impacts potentially increase invasibility of native communities through decreased competitive ability. Characteristics of *S.*

tragus such as polyploidy, a C4 photosynthetic pathway with krantz-type anatomy, self and outcrossing capabilities, a tumbleweed habit, and salinity-induced succulence allows this weed to be highly resource efficient and tolerant to various types of osmotic stress (Iamonico & Mosyakin, 2018; Kadereit & Freitag, 2011). Thus, due to the tolerance of *S. tragus* to environmental stressors, biotic and abiotic pressures placed on native communities in combination with presence of *S. tragus* propagule may lead to infestations.

Impacts from *S. tragus* occur in natural and agricultural systems. Agricultural systems deal with herbicide resistance, physical and mechanical issues, yield loss, disease introduction, and fire hazard from the presence of Russian thistle. Chemical control of Russian thistle in cropping systems has been effective; however, herbicide resistance has been reported in multiple states to both ALS inhibitors and glyphosate. Spring *et. al*. (2022) discovered high genetic diversity within population clusters and concluded that this genetic variability can lead to an increase in the adaptability of the species. High connectivity and gene flow within this species is likely correlated to anemophilous fertilization, self- and out-crossing reproduction, high fecundity, and a tumbleweed habit. The rapid spread of these resistance traits is a critical management issue. Land managers dealing with larger acreages are unable to utilize chemical control for both economic and ecological reasons. Thus, control methods are limited mainly to cultural, mechanical, and biological control. Many studies have explored these methods of control with varying success. In California, a combination of grazing, reestablished native vegetation, and chemical control proved effective (Rao et al., 2022). Biological control agents from countries such as Uzbekistan (*Desertovellum stackelbergi*), Turkey (*Conorhynchus kindermanni*), and Greece (*Aceria slasolae*) have been studied for the potential application to *S. tragus* (Gültekin et al., 2021; Smith, 2005; Sobhian et al., 2003). Biological control may add a vital additional tool for land managers to reduce

the economic impacts of this tumbleweed. However, an efficient and effective method for Russian thistle control over large areas has not yet been identified.

Resistance to chemical control increases the threat of this weed on agricultural systems throughout North America. ALS resistance first occurred in the 1980's (Saari et al., 1992; Stallings et al., 1994), and glyphosate resistance has recently been identified in North America and in Argentina (Kumar et al., 2017; Yanniccari et al., 2023). Yanniccari *et. al*. (2023) identified constitutive overexpression of the target site gene *EPSPS* as the mechanism of resistance. Therefore, parallel evolution of EPSPS tandem duplication and resulting overexpression has occurred between the closely related species *Bassia scoparia* and *Salsola tragus*. Assays for both ALS and glyphosate resistance should be performed on this reference line to characterize it as resistant or susceptible to either mode of action.

In addition to issues posed by this widespread weed, research has also been completed to understand its potential beneficial characteristics. In Russia, *Salsola* ssp. have been utilized in traditional medicine with several applications. In addition, Rosa Tundis *et. al*. (2009) characterized the potential for alkaloid extracts from *Salsola* ssp. to suppress Alzheimer's Disease (AD) progression through the inhibition of enzymes acetylcholinesterase and butyrylcholinesterase which are under study as potential AD driving enzymes. In addition to these beneficial characteristics, in North America *S. tragus* has served as a vital emergency fodder under long-term drought and increased saline conditions. Nutrition studies have found that *S. tragus* compares well to other varieties of fodder such as alfalfa (Beckie & Francis, 2009). Although it poses major issues, benefits of this species slightly offset its invasive nature.

This reference genome provides a fundamental genetic understanding of this weed. Future omics analyses could help identify mechanisms behind abiotic stress tolerance and herbicide

resistance. A fundamental genetic understanding of this weed will assist in the identification of effective control options and candidate genes to help bolster crop security during the era of global climatic flux. Our objective was to assemble to pseudo-molecule level and annotate the genome of a single *S. tragus* individual to provide genomic resources for research into herbicide resistance and weed biology (Montgomery et al., 2023).

METHODS AND MATERIALS

All materials for sequencing were gathered from one individual. Fresh tissue was used for BioNano optical mapping and flash-frozen young tissue was used for PacBio HiFi and HiC chromatin confirmation. Flash-frozen tissue of roots, stems, young leaves, and flowers were used for RNA extraction and PacBio IsoSeq. Flash-frozen tissue and RNA samples were shipped on dry ice and BioNano tissue samples were shipped at 4 ℃ to the Genome Center of Excellence at Corteva Agriscience for DNA extraction, library preparation, and sequencing as described by Raiyemo *et. al.* (2024), resulting in the assembly of two haplomes. Gene annotation using PacBio IsoSeq transcript data along with UniProt and Arabidopsis databases was conducted as described by Raiyemo *et. al.* (Raiyemo et. al., 2024). Chromosome 00 was designated for contigs with insufficient evidence for placement in pseudo-molecules. The estimated genome size was evaluated at the Flow Cytometry Facility of the Iowa State University Office of Biotechnology. Using *Zea mays* as an internal standard, the average genome size for four biological replicates was 1.319 Gbp per 1C. Chromosome counts through hybridization and imaging demonstrated the genome of this *S. tragus* individual contains 2n = 4x = 36 chromosomes per somatic cell (Montgomery et al., 2023). Assemblies for both the reference (haplome 1) and alternative (haplome 2) haplomes are available on Weedpedia (https://weedpedia.weedgenomics.org/), at NCBI (*insert BioProject number and Genome Accession Number when available*), and at CoGe

(*provide CoGe accession number*). Sequencing reads and other data used in the assembly and annotation are available at NCBI (*BioProject Number*).

Both assemblies were analyzed using the BUSCO (v4.0.2) eudicots database, which included 31 species and 2,326 BUSCOs (Manni et al., 2021). Karyotype and annotation feature files were generated for use in RStudio RIdeogram (RStudio version 4.2.1, RIdeogram version 0.2.2) for genome visualization (Hao et al., 2020). Important stress-related transcription factor families as described by Yoon *et. al.* were identified in the reference annotation file using InterPro codes and plotted as markers across the karyotype plots (Yoon et al., 2020). EDTA (v.2.2.0) (Su et al., 2021) and SubPhaser (v1.2.6) (Jia et al., 2022) were used to visualize assembly quality. Finally, chromosome 00 was indexed out of each haplome before gathering assembly statistics through Assemblathon (Pearl v5.34.1) (Earl et al., 2011).

Synteny between the A and B chromosomes of haplome 1 was investigated using RIdeogram (RIdeogram version 0.2.2). Haplome 1 was indexed using samtools (v1.16.1) (H. Li et al., 2009) and split into A and B chromosomes. Indexed B chromosomes were then used in minimap2 (v2.24) (H. Li, 2018) to align to the A chromosomes. The following .paf file was converted to a coords file through ragtag (v2.1.0) (Alonge et al., 2022) and show-coords (mummer, v3.23) (Marçais et al., 2018). The output file was used to identify syntenic regions between the A and B chromosomes (≥1 kbp, ≥95% synteny) (Figure 2.2).

RESULTS

One assembly was prepared for each haplotype present in somatic cells, termed haplome 1 and haplome 2. Each assembly contained 19 scaffolds, representing 18 pseudo-molecular chromosomes containing scaffolds greater than 10M nucleotides and Chromosome 00 containing unassigned scaffolds. Haplome 1 was 1.263 Gbp in size, with a longest and shortest scaffold length at 86.8 Mbp and 4.1 Mbp respectively. Mean and median scaffold sizes were 66.5 Mbp and 67.4

Mbp respectively. Haplome 1 N50 and L50 scaffold lengths were 70.4 Mbp and 9 bp respectively. Contigs from haplome 1 had a mean size of 1.6 Mbp with a N50 length and L50 of 11.3 Mbp and 35 bp respectively. Haplome 1 assembly contained 40.7 contigs per scaffold on average, with an average length of breaks (>24 Ns) between contigs in scaffolds of 4.4 Kbp.

Haplome 2 had a total size of 1.251 Gbp, including a longest scaffold at 86.60 Mbp. The shortest scaffold in haplome 2 was 3.42 Mbp in Chromosome 00. Mean and median scaffold sizes for the haplome 2 were 65.8 Mbp and 66.2 Mbp respectively. The N50 scaffold length for this assembly was 72.6 Mbp, and the L50 scaffold count was 8 bp. Contigs within haplome 2 had a mean length of 250 Kbp with an N50 and L50 length of 9.1 Mbp and 40 bp respectively. Finally, haplome 2 contained an average of 258 contigs per scaffold, with an average length between contigs in scaffolds of 5 Kbp.

Karyotype visualization, SubPhaser results, and the synteny analysis resulted in detailed assembly visualization and basic genomic understandings for both haplomes of *S. tragus*. Karyotype plots including gene density heatmaps indicated a higher gene density in the telomeric to sub-telomeric regions of each chromosome. SubPhaser results indicate a well phased genome with properly assigned homologous chromosomes. The synteny analysis Between the A and B chromosomes of haplome 1 showed increased translocation and inversion activity between the two subgenomes.

DISCUSSION

Analysis of the newly assembled reference genome of the allotetraploid *Salsola tragus* resulted in an increased genomic understanding of this species. The size of each haplome assembly compares well to the flow cytometry estimated sizes indicating that over 90% of the genome was successfully assembled. This pseudo-molecule telomere-to-telomere assembly will provide the

scientific community with an indispensable resource for understanding the basic biology and ecology of *Salsola tragus*.

Karyotype plots generated from haplome 1 indicate telomeric gene density and shorter B chromosomes overall. Telomeric and sub-telomeric regions of each chromosome contain a higher density of genes (Figure 2.1), which is consistent with our understanding that centromeric regions contain fewer genes, if any (Talbert & Henikoff, 2020). Although the centromeres are not indicated on these karyotype plots, one may gather their general location in the regions with decreased gene density. In addition, stress-related transcription factor families were also displayed as markers across the genome. These markers code for proteins that are integral to the expression of endogenous stress-responsive genes under both biotic and abiotic stress (Figure 2.1) (Yoon et al., 2020).

SubPhaser results indicate well phased subgenomes within each haplome. In addition, the B subgenome contains shorter chromosomes than the A subgenome (Figures 2.1 and 2.2). Although differences in chromosome lengths between subgenomes is common in allopolyploids, SubPhaser indicates another interesting pattern. Regions of k-mer characteristics that are assigned to the B subgenome are found in the telomeric regions of the A chromosomes, but not vice versa. These regions undergo synapsis and recombination during meiosis, so the integration of alleles from each subgenome is expected to occur. SubPhaser assigned k-mer characteristics for each subgenome indicates fewer incorporations of the A subgenome into the B subgenome (Figure 2.2). The increased lengths of the A chromosomes may follow this pattern of B chromosome introgression into the A subgenome. This potential phenomenon is of particular interest because these regions contain higher gene densities and may suggest the biased incorporation of one subgenome into the other.

The synteny analysis between the subgenomes of haplome 1 also indicate the combination of highly divergent species. Synteny plots in other species usually results in large spans of genetic conservation. In this case, many small fragments from either subgenome have been shuffled throughout haplome 1. This shuffling of syntenic regions between subgenomes occurred where directionality was conserved or inverted (Figure 2.3). Synteny analysis across chromosomes was not included because it did not display interpretable or understandable patterns or characteristics.

Polyploid evolution has been of particular interest in invasion ecology due to the recognition that many important non-native species are polyploids. Baduel *et. al.* proposed that polyploidy is not a terminal cytotype, but acts as a mechanism for speciation (Baduel et al., 2018). As polyploid species age, they begin to accumulate increased levels of deleterious alleles that are otherwise masked by having increased alleles per locus. Although this assists polyploids in becoming established, this accumulating genetic load can be deleterious. Baduel *et. al.* suggested that this may be one of the mechanisms driving a "polyploid drop" where the polyploid genome reduces to a diploid cytotype. At this point no empirical evidence has been found that confirms the process of polyploid drop, but mathematical evidence suggests that it is possible (Baduel et al., 2018; Leitch & Bennett, 1997; Soltis & Soltis, 1999).

In addition to these ecological considerations, we know that polyploids preset highly fluid genomes. This occurs through the increased activity of transposable elements and through the potential for multivalent chromosome interactions (Baduel et al., 2018; Levin, 1983; Stift et al., 2008; Te Beest et al., 2012). This understanding suggests that the biased incorporation of B subgenome k-mers into the A subgenome of *S. tragus* may be the result of polyploid genomic fluidity.

Finally, following the recognition of the potential for polyploid drop and the fluidity of polyploid genomes, results from these analyses may suggest the process of polyploid drop in the allotetraploid *Salsola tragus*. As previously stated, two cryptic species of differing cytotypes have been discovered in California (Ryan & Ayres, 2000). The two cytotypes reported were tetraploid and diploid. Further investigations utilizing this reference genome assembly are required to understand if the process of polyploid drop occurred in California resulting in *S. tragus* Type A and *S. tragus* Type B. Karyotype plot and SubPhaser results may suggest that the individual sequenced for this reference genome assembly represents a step in the process towards incorporating specific genetic information from the B subgenome before dropping to a diploid cytotype through the retention of the A subgenome. Thus, analysis of this reference genome assembly resulted in the hypothesis that the allotetraploid *S. tragus* individual used for this reference sequence was in the process of polyploid drop.

TABLES

Table 2.1: The reference genome of *Salsola tragus* was obtained from Weedpedia (Weedpedia, International Weeds Genomics Consortium) and run through Assemblathon. This table reports assembly statistics for haplome 1 and haplome 2 pseudo-molecule chromosomes of *Salsola tragus*, excluding Chr00. The lengths of each individual chromosome are also reported in the following table for both haplomes.

	Haplome 1	Haplome 2
Assembly		
Number of scaffolds	18	18
Total size of scaffolds	1,259,258,089	1,247,368,765
Longest scaffold	86,803,187	86,596,344
Shortest scaffold	54,248,817	54,104,439
Number of scaffolds > 10M nt	18	18
Mean scaffold size	69,958,783	69,298,265
Median scaffold size	70,363,782	70,104,450
N50 scaffold length	72,893,522	72,603,936
L50 scaffold count	8	8
scaffold %A	31.8	31.24
scaffold %C	18.04	17.73
scaffold %G	18.07	17.75
scaffold %T	31.84	31.27
scaffold %N	0	2
N50 contig length	11,492,176	9,056,315
L50 contig count	34	40
Average number of contigs per scaffold	41	271
Chromosome lengths (bp)		
Chr01A	86,803,187	86,596,344
Chr01B	66,056,523	80,735,032
Chr02A	86,165,811	80,735,032
Chr02B	67,360,422	77,475,715
Chr03A	82,575,643	76,697,057
Chr03B	62,961,061	73,798,493
Chr04A	77,496,100	72,603,936
Chr04B	59,045,003	72,972,280
Chr05A	77,695,395	70,104,450
Chr05B	62,158,187	65,849,786
Chr06A	74,663,766	66,189,443
Chr06B	54,248,817	62,069,869
Chr07A	72,893,522	58,936,479
Chr07B	63,996,572	61,584,538
Chr08A	73,186,628	54,104,439
Chr08B	62,617,531	61,135,509
Chr09A	70,363,782	61,672,209
Chr09B	58,970,139	58,934,838

FIGURES

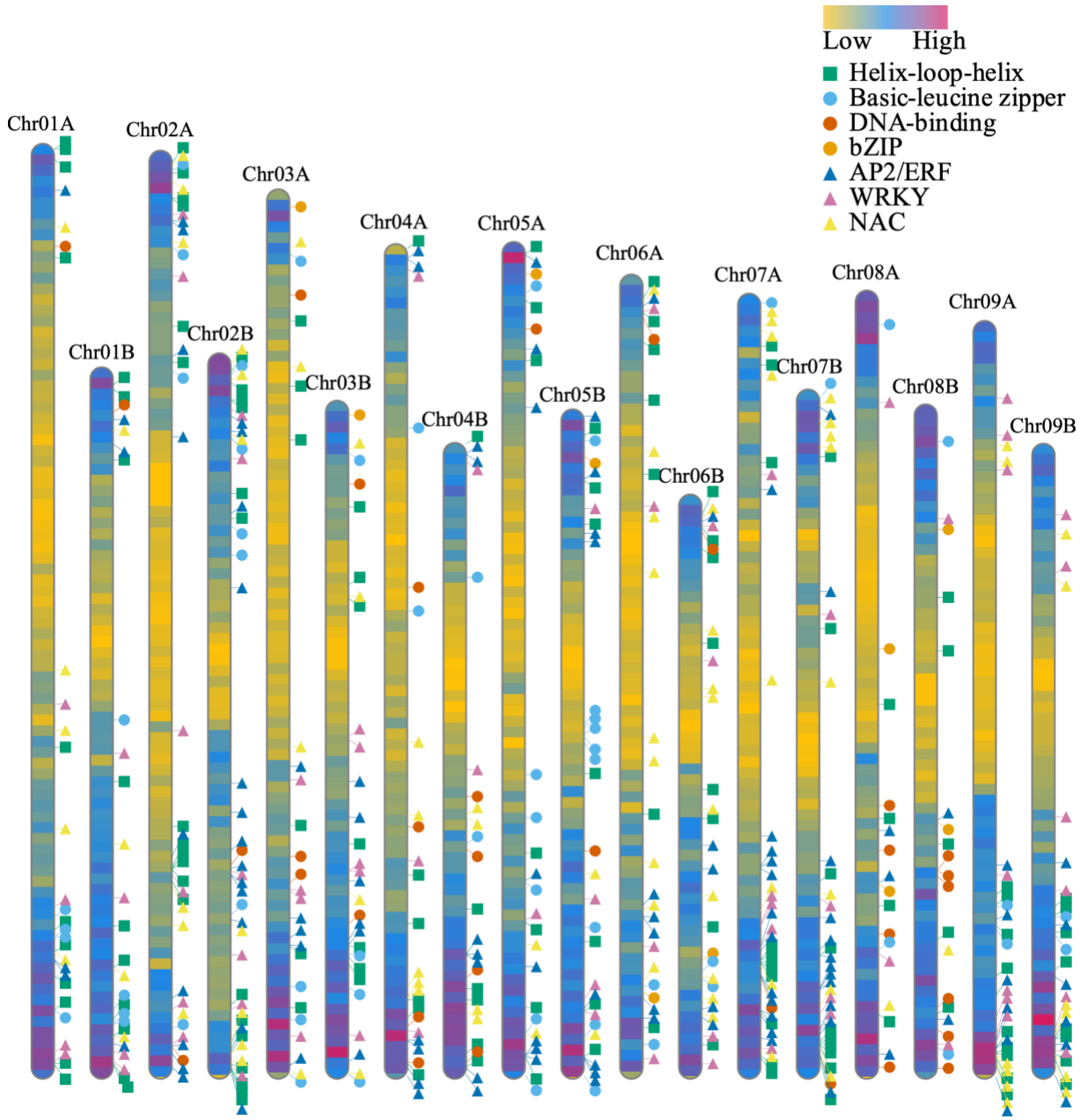


Figure 2.1: A karyotype plot with gene densities and markers for specific gene locations in the haplome 1 assembly of *Sasola tragus*. Darker coloring represents regions with increased gene density and annotation features. Markers of different shape and color indicate the locations of stress-related transcription factor families (Yoon et al., 2020). These transcription factors are integral to all types of stress response and are commonly activated by phytohormones such as abscisic acid and ethylene.

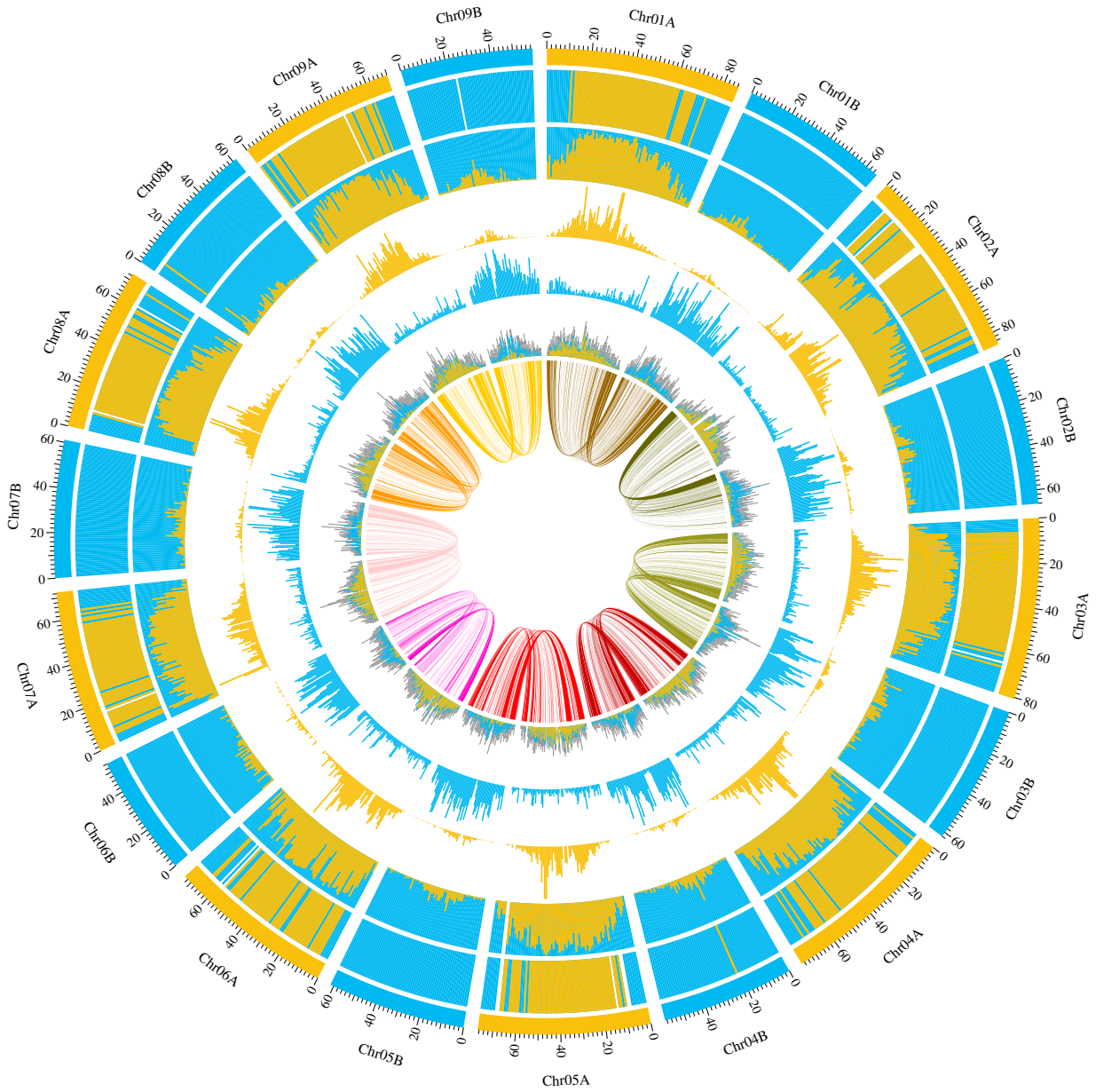


Figure 2.2: SubPhaser plot displaying quality of chromosomal designations to each subgenome in in haplome 1. Subphaser designates colors to subgenomes based on a *k*-means algorithm and is indicated in the first level of the circos plot. The second level of the plot indicates significant enrichment of subgenome-specific k-mers. The third layer indicates relative normalized proportions of these color-coded k-mers. Layers 4 and 5 represent absolute counts of each k-mer set. Finally, the 6th layer indicates long terminal repeat retrotransposon (LTR-RT) density, where each assigned color is representative of that subgenome and grey indicates non-significance. Statistical significance for each of these layers is identified in 1 Mb sliding windows. The lines in the center between proposed homologs indicate syntenic regions between the two subgenomes. The outermost level of the plot includes a scale in kilobases.

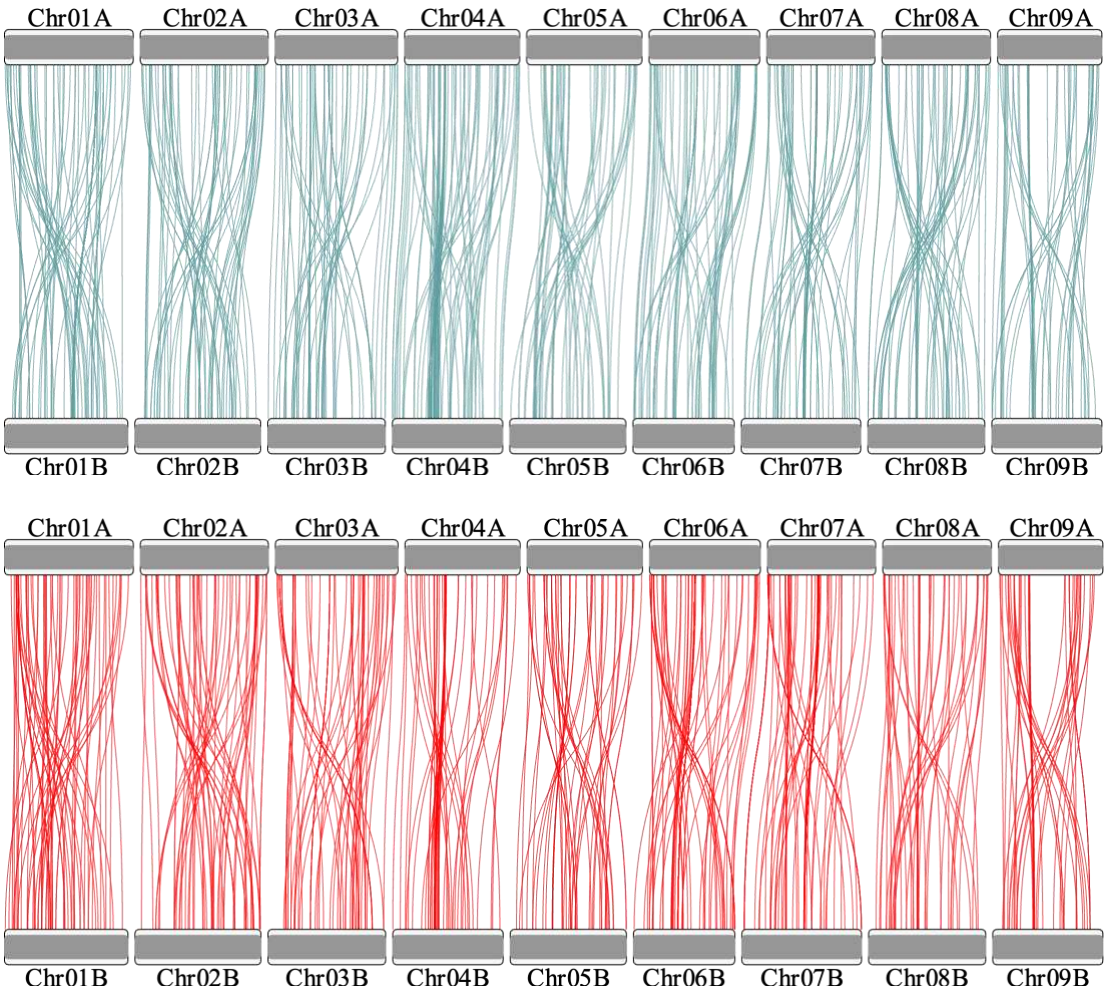


Figure 2.3: Synteny mapping between the A and B chromosomes of *S. tragus* haplome 1 syntenic regions as colored lines (a = 0.7). Each line represents > 95% synteny in regions greater than 1 kilobase. The red lines indicate an inverted region of synteny, and the blue lines indicate unidirectional regions of synteny. An alpha value was set at 0.7 to increase interpretation of overlapping lines (overplotting). These plots demonstrate the major divergence that has occurred between the subgenomes of the allotetraploid *S. tragus* as the result of an ancient polyploidization event and/or through the combined genomes of two highly divergent species.

CHAPTER 3: A TRANSCRIPTOMICS INVESTIGATION INTO ABIOTIC STRESS RESPONSE PATHWAYS IN *SALSOLA TRAGUS*

INTRODUCTION

Humans have been breeding crops and battling weeds for millennia. Through many generations, a land manager's toolkit has expanded to include various aspects of integrated pest management (IPM). Although these management practices have been quite robust, unwanted plant species persist. While domesticated cultivars are bred to feed the world, some wild plants have evolved into weeds through the various natural and anthropogenic selection pressures. The drivers of evolution, including mutation, gene flow, genetic drift, and natural selection, have enabled wild species to evolve strong mechanisms to thrive amidst various natural and anthropogenic selection pressures. The development of chemical weed control, and subsequent evolution of target site and non-target site resistance mechanisms are a primary example of the ability of wild plants to adapt to strong natural or anthropogenic selection pressures (Westwood et al., 2018).

In addition to the challenges of pest pressure, land managers are faced with two other major global implications: a growing global population and a changing climate (Calvin et al., 2023). With global population at unprecedented levels, available arable land may only increase through deforestation. Changing environmental conditions due to the increased concentrations of greenhouse gases have contributed to the loss of this arable land through desertification and salinization and to global direct yield losses (Dagar et al., 2016; Sultan, 2012). There is a general need for cultivars to withstand these pressures to bolster food production during this time of climatic flux.

Many non-native plant species have demonstrated their ability to survive and adapt to many various pressures. Unlike crop cultivars, these species have evolved entirely through the selective

pressures of their environment. Many of these species have established a global presence from increased globalization, allowing gene flow to act between continents and islands. Many non-native species have been evaluated as a nuisance to humanity, and a threat to global biodiversity (Jeschke & Strayer, 2008; Lone et al., 2019; Mack et al., 2000; Simberloff, 2005). Whatever the term applied to these species- weeds, invasives, aliens, non-natives, etc.- they will persist regardless of the IPM strategies applied to them.

Salsola tragus (Russian thistle) is an especially tenacious invasive weed hailing from the mountainous Eurasian continent. This species has established global presence and has become a thorn in the sides of all types of land managers. The *Salsola* polyploid complex contains over 30 individual species (Mosyakin, 1996) and has established weak population differentiation through widespread gene flow thanks to its tumbleweed habit and self-compatibility (Spring et al., 2022). *S. tragus* is a wind-pollinated allotetraploid with a C4 photosynthetic pathway and a tumbleweed habit (Beckie & Francis, 2009). As a halophytic species with an induced succulent habit, it presents increased temperature and osmotic stress tolerance (Rančić et al., 2019). This species has also evolved resistance to both ALS and EPSPS inhibitors (Beckie et al., 2019; Kumar et al., 2017; Morrisson & Devine, 2005; Warwick et al., 2010; Yanniccari et al., 2023).

Sequencing technology has been used to better understand many model organisms. These projects have given the world of science innumerable understandings of the innerworkings of these model organisms. Until recently, many genomics resources have been reserved for clinical studies, model organisms, or valuable cultivars mainly due to funding restrictions. Nowadays, next generation sequencing (NGS) technology has become a more affordable, accurate, and reliable tool that may be used to better understand non-model organisms. This newly available tool allows

scientists to ask the same basic questions about various biologically and economically relevant species (Ravet et al., 2018).

Given the advent and accessibility of NGS technology, we may begin to ask questions to better understand the underlying mechanisms of evolved environmental adaptation presented by these wild species. The International Weed Genomics Consortium (IWGC) has generated a set of fully annotated genomes for many of the world's most tenacious weeds (Montgomery et al., 2023). These weed genomes will be made publicly available as a reference for various genomics-related questions.

Complementary to genomics resources, RNA sequencing technology has been used as an investigatory, hypothesis driving approach to understanding various aspects of genetic responses to environmental stimuli (Conesa et al., 2016; Stark et al., 2019; Todd et al., 2016). The benefits of such a fine scale understanding of an organism's transcriptome are innumerous; however, several aspects must be considered when embarking on an RNA sequencing analysis. Firstly, library preparation may introduce bias towards specific transcripts. In addition, computational decisions may have drastic impacts on downstream results. Biologically, post-transcriptional and post-translational modifications occur frequently, especially following exposure to environmental stimuli. These alterations may have drastic effects on realized gene products. Finally, RNA extractions provide a snapshot of expression at one point in time. Thus, one must be vigilant in experimental design and computational methods, and conservative when interpreting RNA sequencing results.

Various studies have evaluated the transcriptome of various non-model species under stressful conditions in search of important signaling cascades and stress-responsive genes (Guo et al., 2020; Kreps et al., 2002; Yao et al., 2020). The IWGC has completed a genome for *Salsola*

tragus that may be used as a reference annotated genome for a similar transcriptomics analysis. This research aims to present the scientific community with a basic understanding of abiotic stress response in *S. tragus*, to demonstrate the use of newly available genomics resources to better understand plant pests and to identify genes of interest that may increase temperature and osmotic stress tolerance in crop cultivars. The overarching goal of this research is to better understand abiotic stress response in highly stress-tolerant weedy species so that we may bolster food production in this era of climate change.

METHODS AND MATERIALS

Genetic response to abiotic stressors was investigated through two separate experiments. The first experiment screened a single population of *S. tragus* for resilience to heat, drought, cold, salt, and nutrient stress. Results from this experiment informed decisions on the transcriptomics experimental design in the second experiment. *Salsola tragus* (Russian thistle; provided by Dr. Philip Westra, collected in Weld County in 2019) seeds were germinated in general all-purpose sand (Aspen Mountain Aggregates All Purpose Sand, Lowes® Companies, Inc.) at the Colorado State University Weed Research Laboratory Greenhouse in Fort Collins, Colorado. Each experiment included seedlings originating from the same maternal plant but germinated separately. Seeds were scattered on a 200 well flat filled with sand, misted (Fine Fogg-It® Mist Nozzle), and covered with a clear dome for 24 hours. Emerged seedlings were misted twice per day until the first true leaf stage. Environmental conditions were monitored in each growing location by the minute via data loggers (Model H5072, Govee Inc., Hong Kong, China) and exported in a flat file type for each experiment.

3.1 Abiotic Stress Screening experiment

A screening experiment was performed to inform decisions on the transcriptomics experimental design to follow. Seedlings were germinated as described above and grown until the

three to four node stage (the 14th through the 28th of March 2023). Seedlings were randomly selected and assigned to treatment blocks. This experiment included five treatments and two controls. Treatments were split into two blocks of 16 biological replicates resulting in 14 treatment blocks of 16 biological replicates, totaling 224 experimental units. Treatments were exposed to their respective stressors for 21 consecutive days. Osmocote® (Miracle-Gro® Osmocote® Smart-Release® Plant Food) was incorporated into the bottom inch of sand in each Cone-tainer (Stuewe & Sons, Inc. Ray Leach "Cone-tainers"TM & Trays). Each treatment block was sub-irrigated with distilled water between 8 and 10 am for 25 minutes every other day.

Treatments included an average temperature of 5ºC, an average temperature of 35ºC, 150 mM NaCl, a lack of supplied nutrients, and prolonged periods before watering. The low temperature trial was performed in a cold storage room (Colorado State University Weed Research Laboratory, Room 125). The high temperature trial was performed in an incubator (ShelLab Humidified CO_2 Incubator Series, Sheldon Manufacturing Inc., Cornelius, Oregon). A negative control for these populations was grown at ambient temperature, and all three of these treatments were provided light via LEDs (Spider Farmer, Model No SF-1000, Shenzhen Meizhiguang Technology Co., LTD) on a 14:10-hour light to dark cycle. The salt, drought, and nutrient stress trials were gown in a greenhouse along with a negative control group. A 150 mM salt solution was prepared with distilled water and sodium chloride (Cat # S9888-1KG, Sigma-Aldrich®) in a 10 L carboy and used to sub-irrigate salt stressed blocks. Drought stress was induced by watering every fourth day, and nutrient stress was induced through the omission of Osmocote® in the substrate.

After 21 consecutive days each replicate was gently removed from the substrate using a root washer. Shoot and root tissue were separated at the cotyledon. Scaled photos were taken of both tissue types and uploaded into Fiji (ImageJ2, version 2.9.0/1.53t) for surface area estimation.

Wet weight was recorded for each tissue type and replicate. Samples were then stored in a drying chamber at 63°C for 48 hours before collecting dry weights.

Surface area was calculated in Fiji by setting the scale in cm, converting photos to 8-bit, thresholding, and measuring the area of a polygon that contained each sample. The proportion of white to black pixels in each photo was used to calculate the area of the given tissue type for each of the 224 experimental units. Data were exported for statistical analysis in RStudio (version 4.2.1). Variance analysis for each treatment was performed between the two treatment blocks using ANOVA. Pairwise comparisons were performed for each tissue type and treatment to obtain statistical differences in biomass and surface area.

3.2 Tissue preparation for transcriptome analysis

This experiment aimed to understand genetic abiotic stress response in *Salsola tragus* utilizing differential gene expression analysis. *S. tragus* seeds were germinated in general all-purpose sand on 28 June of 2023 at the Colorado State University Weed Research Laboratory Greenhouse in Fort Collins, Colorado. Growing conditions included an average temperature of 20.7ºC, a light intensity of 400-600 $\mu M/m^2s$, a diel insolation cycle of 15:9 hours, and an average humidity of 33.9%. Germinated seedlings at the true leaf stage were randomly selected from this population and transplanted into Cone-tainers containing general all-purpose sand and Osmocote®. This population was sub-irrigated with distilled water in plastic containers every other day for three weeks.

At the four-node developmental stage 30 plants were randomly assigned to 6 treatment blocks each with 5 biological replicates. Treatment blocks included one control block, three salt stress blocks, and two cold stress blocks. Salt stress exposure entailed subirrigation in a 200 mM NaCl solution. Cold stress entailed prolonged exposure to an average temperature of 5 ºC, and an average humidity of 59%. Samples were collected based on timed exposure to each treatment.

Control samples were collected at 0 hours after treatment (HAT). Salt treated samples were collected at 3 HAT, 8 HAT, and 24 HAT. Cold treated samples were collected at 6 HAT and 24 HAT.

Ten L of a 200 mM salt solution was prepared with distilled water and sodium chloride (Cat # S9888-1KG, Sigma-Aldrich®) in a 10 L carboy. At 9 am on 13 July 2023, the salinity stress block was sub-irrigated in the 200 mM sodium chloride solution. The cold stress block was moved into a seed-storage cold room (Colorado State University Weed Research Laboratory, Room 125) and provided a light intensity of 550 μM/m^2s (Spider Farmer, Model No SF-1000, Shenzhen Meizhiguang Technology Co., LTD) on a 15:9-hour light to dark cycle.

3.3 Sample collection and RNA extraction

Sample collection began at 9 am on 13 July 2023. Individual experimental units in the control block were submerged horizontally in a low basin containing distilled water at ambient greenhouse temperature, the cone was gently rocked to wash the sand from the roots of the sample, then samples were laid flat on clean paper towels and gently dab dried. Root and shoot tissue were separated at the cotyledon. Roughly 100 mg of each tissue type was placed into a microcentrifuge tube and immediately placed in liquid nitrogen. The remainder of each tissue type was placed in a 15 mL centrifuge tube and placed directly on liquid nitrogen.

Samples in the salt treatment experimental block 3 HAT were gathered according to the procedure outlined above, with the substitution of the 200 mM sodium chloride solution at ambient greenhouse temperature for the distilled water bath. The sampling procedure was repeated for the cold treatment block at 6 HAT substituting a water bath pre-chilled to ambient cold room temperature. The last sample collection on 13 July 2023 included the salt treatment block at 8 HAT. Finally, both the cold and salt treatment blocks at 24 HAT were collected at 9 am on 14 July 2023

using the above procedures. The sample collection process resulted in 60 shoot and 60 root tissue samples that were flash-frozen in liquid nitrogen and stored at -70ºC.

Samples were randomly selected in increments of six from the -70ºC freezer and placed directly into liquid nitrogen for RNA extraction. Individual samples were well lysed in a pre-frozen mortar and pestle. RNA was extracted using a Spectrum™ Plant Total RNA Kit (SKU STRN50-1kt, Sigma-Aldrich, Inc.) according to manufacturer procedure. Genomic DNA was removed using an on-column DNAse treatment (SKU DNAST70, Sigma Aldrich, Inc.) following manufacturer procedure. All extractions were completed using a microcentrifuge (model # 5417C, Eppendorf, Inc.).

RNA quality and quantity was then measured using a fluorometer (Invitrogen Qubit™ 4, Thermo Fisher Scientific). A Qubit™ RNA Broad Range (BR) Assay Kit (Cat # Q10210, Invitrogen™ Thermo Fisher Scientific) was used to record individual sample RNA concentration, and a Qubit™ RNA IQ Assay Kit (Cat # Q33221, Invitrogen™ Thermo Fisher Scientific) was used to record individual sample RNA integrity and quality scores. Three out of five biological replicates from each block were selected based on required RNA yield and Integrity and Quality (IQ) scores, packaged in insulated containers with dry ice, and shipped to Novogene (Novogene Corporation Inc., 8801 Folsom Blvd #290, Sacramento, CA 95826) for RNA sequencing using Illumina.

3.4 Sequence analysis and differential expression pipeline

Illumina fastq files were downloaded from Novogene via FileZilla (v3.66.4) and uploaded onto the Alpine RC Cluster (University of Colorado Boulder Research Computing) for processing. The RNA sequencing library resulted in 165.4 GB of paired-end reads. Adapter trimming and quality control were preformed using fastp (Chen et al., 2018) (v0.23.2) and run through multiqc (Ewels et al., 2016) (1.14) for quality reporting and visualization. The reference genome (v01;

Weedpedia, IWGC) was indexed using STAR (Dobin et al., 2013) (v2.7.10b) for trimmed fastq alignment. Resulting sam files from STAR and the reference annotation file for *Salosla tragus* (v01.0; Weedpedia, International Weed Science Genomics Consortium) was used in featureCounts (Wernersson, 2006) (v2.0.3). Resulting featureCounts flat text files were downloaded from Alpine RC for local processing using DESeq2 (Love et al., 2014) (v1.38.3; https://github.com/JohnLemas/S.tragus-abioticStress-RNAseq) in RStudio (v4.2.1). DESeq2 analysis employs a negative binomial distribution to assess differential expression based on these count data. Genes were designated as significantly differentially expressed if they met thresholds of an adjusted p-value of 0.05, a $\log_2$ fold change (LFC) in expression greater than or equal to 1, and counts > 10. Contrasts were split into separate R scripts for DESeq2 analysis of all tissues 0 versus 24 HAT, shoot tissue analysis 0 versus 24 HAT, and root tissue analysis 0 versus 24 HAT. Contrasts were then split further to include one tissue type and one treatment for all timeframes. This resulted in seven differential expression contrasts.

3.5 Gene Ontology analysis

The *Salsola tragus* proteome was obtained and blasted against the Uniprot protein database to obtain Uniprot accession numbers from *S. tragus* gene IDs. The proteome was generated via the downloaded *S. tragus* CDS file (v01; Weedpedia, IWGC) and DTU Health Technology VirtualRibosome (v2.0, https://services.healthtech.dtu.dk/services/VirtualRibosome-2.0/) (Wernersson, 2006). This reference proteome was blasted against the UniProtKB swiss proteome (https://www.uniprot.org/help/downloads) (The UniProt Consortium et al., 2023). Resulting reciprocal best hits were obtained for all gene IDs in the *S. tragus* genome. Differentially expressed gene lists were then converted to UniProt accession numbers for analysis in NCBI DAVID (Sherman et al., 2022). The FDR p-value was used to identify significant upregulated biological

processes, and these results were plotted in RStudio ggplot2 (v3.5.1, https://github.com/JohnLemas/S.tragus-abioticStress-RNAseq).

RESULTS

3.6 Stress screening analysis

An overview of statistical differences between treatments can be found in Figure 3.1 and will be described here. Shoot tissue biomass for salt treated tissues was not statistically different from the greenhouse control. All other treatments were significantly different from the greenhouse control ($p<0.05$), including the headhouse control. Shoot biomass of the high and low temperature trials were not statistically different from the headhouse control ($p=0.074$ & $p=1.0$ respectively). The greenhouse and headhouse controls did not have statistically different shoot area estimations ($p=0.223$). The estimated surface area of the salt treated shoot tissues was significantly lower than the greenhouse control ($p=0.001$), but not the headhouse control ($p=0.380$). Drought, nutrient stress, high temperature, and low temperature surface area estimations were all significantly different from their respective controls ($p<0.001$). The nutrient stress trial was the only treatment that was significantly different from all other treatments for both biomass and surface area. For root tissues, all treated samples were significantly different from respective controls for both biomass and estimated surface area ($p<0.001$).

3.7 Sequencing analysis and differential expression pipeline

DESeq2 analysis resulted in 9,068 unique differentially expressed genes among all seven contrasts, representing 21% of the overall annotated genes in the reference genome. Of these, 4,594 (50.6%) originated from A chromosomes and 4,473 (49.4%) originated from the B chromosomes. Summaries of normalized count values for the eighteen samples included in the 24 HAT analysis are reported in Figure 3.2. Of these samples, the highest mean count value was 1,705, while the

first and third quartiles were 78 and 1,209 respectively. The highest median count value was 361, and the greatest maximum count value was 785,675. This highest normalized count value was for late embryogenesis related dehydrin protein in cold treated shoot tissues (*S. tragus* gene ID: SalTrChr03Ag052190). The following subsections will outline each individual DESeq contrast and report genes of greatest significance (lowest p-value) and greatest $\log_2$ fold change (LFC) in expression. These sections will begin by focusing on 0 versus 24 HAT in all tissues, in shoot tissues, and in root tissues. Finally, results from each tissue type and treatment specific contrasts for 3, 6, 8, and 24 HAT will be presented in the final subsections.

3.7.1 Differential expression analysis: 0 versus 24 HAT

The initial DESeq contrast including both tissue types at 0 versus 24 HAT resulted in the differential expression of 160 genes, 118 upregulated and 42 downregulated (Figure 3.3). Annotations from eight of the upregulated genes with the lowest p-value included “Heat-shock transcription factor family”, “PPM-type phosphatase domain superfamily”, “Tetratricopeptide-like helical domain superfamily”, “Chaperone J-domain superfamily”, “Small heat shock protein HSP20”, “Lipoxygenase”, and “Polyketide synthase”. The upregulated heat shock protein, heat shock transcription factor, J-domain heat shock factor chaperone, and tetratricopeptide cochaperone all originated from the B subgenome (SalTrChr09Bg416870, SalTrChr08Bg395090, SalTrChr03Bg288430, SalTrChr03Bg286720, SalTrChr06Bg341440).

The DESeq contrast including only shoot tissues at 0 versus 24 HAT resulted in 464 differentially expressed genes, 403 upregulated and 61 downregulated. Annotations from six of the upregulated genes with the lowest $-\log_{10}$ p-value include “Polyketide synthase”, “Senescence/spartic acid”, “Chlorophyll A-B binding protein”, “Alpha/Beta hydrolase fold”, and one uncharacterized protein. Finally, the DESeq contrast including only root tissues at 0 versus 24 HAT resulted in 146 differentially expressed genes, 60 upregulated and 86 downregulated.

Annotations from five of the upregulated genes with the lowest p-value include “Small heat shock protein HSP20”, “Heat shock transcription factor family”, “BAG family molecular chaperone regulator”, “Tetraspanin-8-like”, and one uncharacterized protein.

3.7.2 Differential expression analysis: salt treated shoot tissues

Salt treated shoot tissues resulted in 1,834 upregulated genes and 1,766 downregulated genes (Figure 3.4). Of the upregulated genes, 37 were unique to 3 HAT, 896 were unique at 8 HAT, 539 were unique at 24 HAT, and 28 were common between all timepoints. The highest normalized count for salt treated shoot tissues was for pyruvate phosphatase dikinase (PPDK) at 489,186 (*S. tragus* gene ID: SalTrChr03Ag056710).

Salt treated shoot tissues 0 versus 3 HAT resulted in 338 upregulated genes and 86 downregulated genes. Annotations from seven of the smallest p-value included “Galactose oxidase/kelch”, “Heavy metal-associated domain superfamily”, “Chaperone J-domain superfamily”, and “TAZ domain superfamily”. Genes with significant log fold change in expression included two uncharacterized genes with an upregulation of >20 $\log_2$ fold change (LFC) in expression (*S. tragus* gene IDs: SalTrChr07Bg374820 & SalTrChr07Bg372130).

Salt treated shoot tissues at 0 versus 8 HAT resulted in 1,237 upregulated and 1,416 downregulated genes. Annotations from the smallest -$\log_{10}$ p-value include “LNK family”, “CCG-binding protein 1”, “HAD-superfamily hydrolase”, “Transmembrane Fragile-X-F-associated protein”, “Galactose oxidase/kelch”, and “Ferredoxin-NADP reductase”. There were 44 genes with a < -20 LFC in expression, including “Leucine-rich repeat domain superfamily” and “GPN-loop GTPase” (*S. tragus* gene IDs: SalTrChr003Ag074460 & SalTrChr03A059800).

Salt treated shoot tissues at 0 versus 24 HAT resulted in 649 upregulated and 264 downregulated genes. Annotations from seven of the smallest p-value include “Nucleotide-binding alpha-beta plait domain superfamily”, “Lambda repressor-like”, “Small heat shock protein

HSP20", and "Heat shock HSP90 family". Two uncharacterized genes resulted in a > 20 LFC, and one uncharacterized gene resulted in a < -20 LFC in expression (*S. tragus* gene IDs: SalTrChr07Bg372130, SalTrChr07Bg374820, & SalTrChr02Bg247620).

3.7.3 Differential expression analysis: salt treated root tissues

Salt treated root tissues resulted in 375 unique upregulated genes and 271 unique downregulated genes (Figure 3.4). Of the upregulated genes, 94 were unique to 3 HAT, 101 were unique to 8 Hat, 130 were unique to 24 HAT, and six were shared between all three timepoints. The highest count for salt treated root tissues was annotated as "BURP domain" with a normalized count of 246,334 (*S. tragus* gene ID: SalTrChr7Ag149210).

Salt treated root tissues at 0 versus 3 HAT resulted in 135 upregulated and 25 downregulated genes. Annotations for genes with the smallest p-value include "Auxin-binding protein ABP19a-like", "Kiwellin-like", "Plant peroxidase", and "EF-hand domain pair". One gene annotated as "Sieve element occlusion" resulted in a > 20 LFC in expression (*S. tragus* gene ID: SalTrChr07Bg363570).

Salt treated root tissues at 0 versus 8 HAT resulted in 137 upregulated and 133 downregulated genes. Annotations from these with the smallest p-value include "JmjC domain", "Cupin-like domain 8", "Protein LHY", "LNK family", "germin", "kelch-type beta propeller", "CheY-like superfamily", "and Galactose oxidase/kelch". Forty-two genes resulted in a < -20 LFC in expression, including "FLC expressor/FLX-like", "Leucine-rich domain superfamily", and "Protein kinase-like domain superfamily" (*S. tragus* gene IDs: SalTrChr08Ag184120, SalTrChr07A146460, & SalTrChr06Bg354650).

Salt treated root tissues at 0 versus 24 HAT resulted in 159 upregulated and 144 downregulated genes. Annotations for eight genes with the smallest p-values include "Lambda repressor-like", "HSP20-like chaperone", "Small heat shock protein HSP20", "Chaperone J-

domain superfamily", "Heat shock protein HSP90 family", "Glycosyl hydrolase 36", "O-methyltransferase COMT-type" and one uncharacterized protein. One gene annotated as "Sieve element occlusion" resulted in a > 20 LFC in expression (*S. tragus* gene ID: SalTrChr07Bg363570).

3.7.4 Differential expression analysis: cold treated shoot tissues

DESeq contrasts for cold treated shoot tissues include 0 versus 6 HAT and 0 versus 24 HAT. As previously reported, the gene with the highest normalized count for cold treated shoot tissue was annotated as late embryogenesis related dehydrin protein (*S. tragus* gene ID: SalTrChr03Ag052190). Differential expression of cold treated shoot tissues resulted in the greatest number of significant differentially expressed genes, 3,492 upregulated and 2,491 downregulated genes (Figure 3.5). Of the upregulated genes, 161 were unique to the 6 HAT timepoint, 2,504 were unique to 24 HAT timepoint, and 827 were shared between both timepoints.

Cold treated shoot tissues 0 versus 6 HAT resulted in 988 upregulated and 241 downregulated genes. Annotations for genes with the lowest p-values included "CheY-like superfamily", "Chlorophyll *a/b* binding protein", "Galactose oxidase/kelch", and "Heat shock transcription factor family".

Cold treated shoot tissues 0 versus 24 HAT resulted in 3,331 upregulated and 2,406 downregulated genes. Annotations for those with the smallest p-value include "CheY-like superfamily", "Chlorophyll A-B binding protein", "Respiratory supercomplex factor 2", and "P-type ATPase". One gene annotated as "AP2/ERF" resulted in a < -20 LFC in expression (*S. tragus* gene ID: SalTrChr0Bg261570).

3.7.5 Differential expression analysis: cold treated root tissues

Cold treated root tissues resulted in 1,211 unique upregulated and 1,166 unique downregulated genes (Figure 3.6). Of the upregulated genes, 141 were unique to the 6 HAT

timepoint, 901 were unique to the 24 HAT timepoint, and 169 were shared between both timepoints. The highest read count was annotated as “HMW kininogen” with a normalized read count of 436,844 (*S. tragus* gene ID: SalTrChr06Ag130280).

Cold treated root tissues at 0 versus 6 HAT resulted in 310 upregulated and 159 downregulated genes. Four genes with the lowest p-value are annotated as “CheY-like superfamily”, “Homeobox-like domain superfamily” and “TraB family”.

Cold treated root tissues at 0 versus 24 HAT resulted in 1,070 upregulated and 1,108 downregulated genes. Five genes with the lowest p-value were annotated as “CheY-like superfamily”, “JmjC domain”, “Eukaryotic translation initiation factor SUI1”, and “AP2/ERF domain superfamily”. Thirty-four genes resulted in a < -20 LFC in expression, including those annotated as “Frigida-like” and “FLC expressor/FLX-like” (*S. tragus* gene IDs: SalTrChr04b304260 & SalTrChr08Ag184120).

3.8 Gene ontology enrichment analysis

Enrichment analysis of the upregulated genes reported for all samples 24 HAT resulted in six DAVID biological process terms, including “response to cold”, “response to water deprivation”, “response to abscisic acid”, “regulation of transcription, DNA-templated”, “cellular response to cytokinin stimulus”, and “response to salt stress” (Figure 3.7). Enrichment analysis of shoot tissues 0 versus 24 HAT resulted in 43 DAVID biological processes with FDR p-values < 0.05. The first four with the lowest p-values included “regulation of salicylic acid”, “response to cold”, “circadian rhythm”, and “phytochelatin biosynthetic process”. Others included “response to stress”, “regulation of abscisic acid biosynthetic process”, “jasmonic acid and ethylene-dependent systemic resistance”, and “response to salt stress”. Enrichment analysis of root tissues 0 versus 24 HAT resulted in only three significant DAVID biological process terms, “systemic acquired

resistance, salicylic acid mediated signaling pathway", and "cytokinin-activated signaling pathway", "regulation of transcription, DNA-templated".

Salt treated shoot tissues contrasts for 3, 8, and 24 HAT resulted in 65 unique GO terms. Salt treated root tissue contrasts for 3, 8, and 24 HAT resulted in 18 unique GO terms. Eight of these GO terms were common between tissue types, including "response to water deprivation", "extracellular transport", "response to cold", and "response to salt stress". For cold treated shoot tissues, contrasts for 6 and 24 HAT resulted in 70 unique GO terms. Cold treated root tissues for 6 and 24 HAT resulted in 45 unique GO terms. Twenty-four of these terms were common between both tissue types, including "response to water deprivation", "response to cold", "response to salt stress" and "rhythmic process". Three GO terms were shared between both treatments and tissue types, annotated as "response to water deprivation", "response to cold", and "response to salt stress".

DISCUSSION

3.9 Stress screening analysis

This research aimed to investigate transcriptional abiotic stress response in the halophyte *Salsola tragus* over a 24-hour period. A stress screening analysis was enacted to identify important stressors for further investigation in the subsequent RNA sequencing analysis. The main hypothesis for this screening was that groups exposed to stress would exhibit decreased shoot growth. An additional hypothesis was that populations would develop a more branched root system, and that the salt stress trial would not be significantly different from the control trials. These hypotheses were statistically tested to determine differences between treatment and control populations. Collected metrics included shoot wet weight, shoot dry weight, estimated shoot area, root weight, and estimated root area. Nearly all metrics produced statistically significant differences between treatment groups and control groups.

The design of this experiment was intended to control for confounding variables. Given the number of treatments and their individual requirements, all treatment blocks were designed to be identical apart from the treatment in question. Sand provided an inert, nutrient-free substrate that allowed for the control of salt-substrate interactions and ensured a lack of baseline nutrients for all groups. Also, a designated sub-irrigation interval ensured that all experimental units were exposed to the same amounts of moisture. This section of chapter three will cover experimental design issues, differences between blocks of the same trial, surface area estimations, root analyses, and the overall results of this screening considering the original hypotheses.

Difficulties in experimental design included substrate moisture retention, environmental condition differences, and cone-tainer experimental unit densities. Although the general all-purpose sand provided a reliable backdrop for the nutrient and salt stress trials, it was difficult to retain soil moisture. Evaporation and transpiration rates were high enough that the cones would dry out and plants would be inadvertently exposed to drought stress. This was mainly an issue in the headhouse control area, which displays the second highest temperature and the lowest humidity overall. Therefore, this trial did not provide an adequate control group for the temperature trails and instead represents a population that was exposed to prolonged environmental stress. In addition, cones were removed from treatment areas and individually placed into the sub-irrigation containers. This led to the inadvertent pulling of experimental units, which led to a reduction of biological replicates. This can be seen in the somewhat varied number of biological replicates per block. These issues could be avoided in the future through the employment of plastic wrap at the top of the cone-tainer to retain moisture, and deliberate and careful transferring during sub-irrigation.

In assessing variance between both blocks, the drought trial was the only treatment group that provided statistical differences between blocks. Block 2 displayed consistently higher variance than Block 1. Block 2 included the only fatality in the entire experiment (sample 141), which may be the result of these differences. These blocks were only watered every four days to induce drought stress, and Block 2 was watered half an hour after Block 1. Perhaps the extra half-hour resulted in major differences between these groups, and that Block 2 experienced increased drought conditions when compared to Block 1.

To collect metrics on root development, a root washer was used to gently remove sand from the roots of each individual sample. Although care was taken not to overly disturb the roots, differences in fine root branching were lost in the process. In addition, not all the sand was removed from every sample. Therefore, both weight and surface area comparisons were not reliable metrics using this method. An alternative, less destructive method should be used to assess differences in root development between treatment groups in the future. Unfortunately, this resulted in an inability to assess the validity of the hypothesis that different treatments would produce more or less branched root systems.

Although finer details could not be ascertained, the data do suggest some differences in root development depending on treatment groups. The greenhouse control group produced the greatest root weight and root area overall. The temperature stress and drought trials produced similar results for both root weight and root area. The nutrient stress trial produced the lowest of both metrics overall, and the headhouse control and salt stress trials produced somewhat conflicting results. Additional experimentation will need to be performed to understand the differences in root development between different stress conditions.

Pairwise comparisons for all metrics agreed well with each other; however, the dry weight metric produced the most statistically reliable results. Nearly all trials produced shoot dry weights that were significantly different from the greenhouse control group (Figure 3.1). Groups that were not significantly different include the greenhouse control versus salt stress, headhouse control versus cold stress, headhouse control versus heat stress, headhouse control versus drought, and cold stress versus drought. In addition, cold stress versus heat stress were just significantly different enough using this metric ($p = 0.047$), however this is not seen with other metrics. The effect of suspected environmental stress in the headhouse control group can be observed, as it was not significantly different from temperature and drought stress trials. The most interesting result here is that there was no significant difference in shoot dry weight between the greenhouse control and salt stress groups. This may be due to uptake and translocation of sodium ions followed by plant-wide vacuolar sequestration.

Throughout the course of the screening, the cold treated samples appeared qualitatively healthy. The stunting of the cold treated samples may be due to the process of cold acclimation and the slowing of metabolic processes. Of all the treated samples, those grown in 150 mM NaCl preformed similar to the greenhouse control in shoot biomass, but not estimated surface area. Increased shoot density due to sodium sequestration may explain the difference between metrics in these samples. Based on these results, the cold and salt stress trials were selected for transcriptomics analysis. Furthermore, the salt treatment was adjusted to 200 mM NaCl for the following experiment to increase the effect size of salt treated samples.

3.10 Differential expression analysis: An overview

The *Salsola tragus* transcriptome profile under salinity and cold stress was analyzed to understand stress response in this species through RNA sequencing. Transcriptomics in this non-model organism has not been undertaken before this study. The availability of a quality reference

genome enabled a further understanding of stress response in this species. Analysis of shoot and root tissue at 0, 3, 6, 8, and 24 HAT resulted in the differential expression of roughly 21% of the 43,354 annotated genes present in this genome (tables in SI). This regulation profile is comparable to similar transcriptome investigations (Kreps et al., 2002; Yao et al., 2020). Most of the identified upregulated genes included protein chaperones and transcription factors. It is impractical to discuss each of these differentially expressed genes individually, thus an overview of significant results common between timepoints and with previously known pathways of stress response will be presented in this section.

Although plants have an array of stress-specific response genes, abiotic stress response includes many signaling cascades, phytohormones, and kinases that are shared between different stressors. As such, many similarities in gene expression result from cold, drought, and even salinity stress (Yamaguchi-Shinozaki & Shinozaki, 2006). Calcium plays a central role in both cold and salt stress, especially when considering the Calcium-dependent protein kinases that trigger varied transcription factors in response to abiotic stress.

Regardless of the similarities in signaling and expression profiles, drought and salt stress pose different barriers to overall plant vitality. Reduced metabolic rates and dwarfism result from cold stress as it leads to membrane rigidification and the slowing of enzyme activity (Chinnusamy et al., 2010; Yamaguchi-Shinozaki & Shinozaki, 2006). Salt stress leads to the phytotoxic build-up of ions in the cytosol. These excess ions may increase ROS generation through the slowing of photosynthesis (Mansour & Hassan, 2022; Munns, 2005). Although different plants have evolved different mechanisms for dealing with salinity, cold response signaling cascades and gene expression are somewhat conserved between species.

3. 10.1 Differential expression analysis: 24 HAT

The contrast between 0 and 24 HAT resulted in the differential expression of J-domain chaperones, tetratricopeptide repeat-containing proteins, the related heat shock transcription factor family proteins, linoteate 9S-lipoxygenase, chalcone synthase, and protein phosphatase 2C. J-domain chaperones reinforce the structure of heat shock transcription factor proteins allowing them to function under stressful conditions in a variety of organisms (Walsh et al., 2004). In the family of these chaperones are the tetratricopeptide repeat-containing proteins that also function as chaperones for heat shock protein Hsp90, and play important roles in response to osmotic stress (D. Mishra et al., 2018). In other plants, these three upregulated genes are involved in a complex array of chaperones, steroids, and organizing proteins that respond to a variety of biotic and abiotic stressors. Arabidopsis mutants have demonstrated the important role of heat shock proteins, organizing proteins and the hormone brassinosteroid in salt stress response (Zhang et al., 2022). Based on these results the expression of heat shock proteins and their respective chaperones and cochaperones in *S. tragus* may be biased towards the B subgenome. Differentially expressed lipoxygenase transcripts are important for various abiotic stressors, including cold stress, and are an integral component in the biosynthesis of jasmonic acid (Nemchenko et al., 2006; J. Wang et al., 2020). Chalcone synthase is essential to flavonoid biosynthesis and thus plays a large role in a variety of plant stress response (Dao et al., 2011). Abscisic acid is a pivot point for various environmental responses, and protein phosphatase 2C is important for abscisic acid related signaling cascades (Pang et al., 2024). The identification of these genes in the first contrast including all samples 0 versus 24 hours after treatment demonstrates their plant-wide role in both cold and salt stress in *Salsola tragus.*

The 24-hour contrast was also split into two analyses including only shoot or only root tissue exposed to both cold and salt stress. Differentially expressed genes in shoot tissues included

salicylic acid-binding protein, early responsive to dehydration, early light-induced protein, chalcone synthase, and one unknown protein. Increased activity of the Alpha/Beta hydrolase fold salicylic acid-binding protein 2 in other systems has been shown to increase endogenous salicylic acid concentrations (Guan et al., 2019). Similar to jasmonic acid (JA) and abscisic acid (ABA), salicylic acid plays a crucial role in response to various biotic and abiotic stressors (Song et al., 2023). In relation to JA, senescence/spartin-associated early response to dehydration may have been induced through increased jasmonate in cold treated shoot tissues (Hu et al., 2017). Further, as the annotation suggests, the upregulated transcript annotated as chlorophyll *a/b*-binding protein are important proteins in photosynthesis and stomatal conductance. These proteins are involved in a variety of biological processes, including oxidative stress and abscisic acid signaling (Xu et al., 2012). Finally, the differential expression of one uncharacterized protein may present a novel *S. tragus* stress responsive gene in abiotic stress response.

Differential expression analysis including root tissues 0 versus 24 HAT resulted in two heat shock proteins, a tetraspanin-like protein, expansin-A12-like protein, and two uncharacterized proteins. Tetraspanin proteins were recently characterized in Arabidopsis as important extracellular transporters in the biosynthesis of vesicles. They also play a role in plasma membrane sodium signaling or response to other toxins (Liu et al., 2024). Increased expansin activity increases with ABA in many other systems (Abuqamar et al., 2013); however, in this study expansin-A12 was found to be downregulated in *S. tragus* root tissues. The complexities of expansin activity under abiotic stress conditions have yet to be clearly understood (Tenhaken, 2015). This study found downregulation of *S. tragus* expansin-12 in response to decreased temperature or increased sodium concentrations, so expansin-12 may play a larger role in membrane rigidification than those

induced by ABA. The *S. tragus* genome contains 56 genes annotated as expansins, so it's interesting that expansin-12 was the only one found differentially expressed.

3.10.2 Differential expression analysis: treatment and tissue specific contrasts

When considering all four additional DESeq contrasts, two *S. tragus* transcripts were commonly upregulated in cold treated root tissues 6 and 24 HAT, cold treated shoot tissues 6 and 24 HAT, and in salt treated root tissues 8 HAT and were annotated as "CheY-like superfamily". These two transcripts, SalTrChr05Ag121400 and SalTrChr05B336460, are annotated as a two-component response regulator like APRR5 isoform. These regulators have been identified as playing a major role in drought response through various kinases and ABA signaling cascades (Huang et al., 2018). The presence of these transcripts indicates the importance of this signaling pathway for *S. tragus* in salt and cold response.

Another prevalent differentially expressed gene was annotated as "Galactose oxidase/kelch" and was found to be differentially expressed in cold treated shoot tissue at 6 HAT, salt treated shoot tissues at 3 and 8 HAT, and in salt treated root tissues 8 HAT. The protein for this gene matched a UniProt accession ID for RING zinc finger proteins (*S. tragus* gene ID: SalTrChr06Ag129910; UniProt accession ID: Q7M3S9). Most RING zinc finger proteins fall into ubiquitin ligase activity, and enhance ABA biosynthesis (Han et al., 2022). The assumed homolog of this gene, also annotated as "galactose oxidase/kelch", matched to a different UniProt accession number relating to F-box proteins (*S. tragus* gene ID: SalTrChr06Bg344810; UniProt accession ID: Q9C8K7). Like RING zinc finger proteins, these F-box proteins are likely also E3 ligases (Kuroda et al., 2012). An additional gene annotated "F-box protein" was also differentially expressed in salt treated root tissues 8 HAT (*S. tragus* gene ID: SalTrChr07Ag146460). Temperature and osmotic stress can lead to the accumulation of misfolded or otherwise non-functional proteins; thus, the ubiquitination of these proteins is important for growth and

development. This process is carried out by a number of proteins, including RING zinc finger proteins and F-box proteins (Han et al., 2022).

Analysis of salt treated root tissues resulted in the differential expression of one lysine specific JMJ30 domain, one LHY protein, and four LNK family genes. The Jumanji C-domain containing protein JMJ30 is known to interact with several proteins to regulate circadian differences in gene expression, including one LATE ELONGATED HYPOCOTYL (LHY) protein and four NIGHT-LIGHT INDUCIBLE (LNK) proteins (Jones et al., 2019). JMJ30 was found to be upregulated and LHY and LNK genes were found to be downregulated specifically in salt treated root tissues at 8 HAT. Any indication that these genes were differentially expressed in root tissues due to salt treatment at 8 HAT cannot be made due to confounding variables. Further investigation would be required to understand how JMJ30, LNK, and LHY respond to salt stress; however, JMJ30 was also found to be upregulated in cold treated root tissues at 24 HAT. As cold and circadian rhythm expression is somewhat linked, perhaps JMJ30 regulates important gene expression in cold treated root tissues of *S. tragus*.

3.10.3 Differential expression analysis: expression of previously reported salt response genes

An additional method for analyzing salt stress response was to look at genes characterized as integral to salt response. These genes include transcription factors, ROS scavengers, osmoprotectant biosynthesis enzymes, and passive or active antiporters (A. Mishra & Tanna, 2017). The activity of several of these identified genes were screened in the expression profile of *S. tragus* samples exposed to cold and salt stress. Corresponding *Salsola tragus* gene IDs were identified in the reference assembly annotation file and screened for activity under treatment conditions. Baseline expression for this species based on the results of this analysis was 400 – 1,200 transcripts (Table 3.1). Expression of known salt responsive genes outside of this baseline included calcium antiporter CAX1, sodium antiporter NHX1, and several others (Figures 3.8

through 3.10). Calcium is a major pivot point in response to a wide range of abiotic stressors and is integral to Calcium Activated Protein Kinase (CAPK) activity. Response to salinity induces the activity of several antiporters that distribute and sequester sodium throughout the plant to sub-phytotoxic levels of Na^+ in the cytosol. These antiporters, NHX1 through NHX7, work together to translocate sodium ions that accumulate due to non-selective potassium antiporters in root tissues and are regulated through the activity of SALT OVERLY SENSITIVE 3 (SOS3). Increased cytosolic sodium concentrations also reduce rates of photosynthesis, thus this responsive spillway also works to upregulate ROS scavengers. Although the activity of these identified salt responsive genes was not found to be differentially expressed in DESeq analyses, they do provide data on how these genes are transcriptionally present under both treatment conditions in *S. tragus.*

CONCLUSION

This research sought to gain a transcriptional insight into osmotic stress response in the tetraploid tumbleweed *Salsola tragus* (Russian thistle). This species has been present in North America since the late 1800s (Beckie & Francis, 2009; Mosyakin, 1996). As a C4 weedy annual, the *Salsola* genus has demonstrated an impressive capacity to thrive in a variety of climates worldwide. Previous studies have demonstrated increased stress tolerance in desirable cultivars after the introduction of important osmotic and temperature stress genes from tolerant weedy or halophytic species (W. Li et al., 2011). Through this transcriptional insight, we hope to equip future studies with detailed information on the transcriptome changes in the widespread and highly stress tolerant species *Salsola tragus.*

Results from the stress screening analysis reinforce the characterization of *S. tragus* as a highly stress-tolerant species. This screening was performed to select abiotic stressors that *S. tragus* is particularly predisposed to tolerating. As this species displays induced succulence when exposed to elevated salinity, results from the salt treated samples in this screening assay were

expected. Along with a closely related member of the Amaranthus family, *Bassia scoparia*, *Salsola tragus* is one of the earliest emerging weedy species in North America (Beckie & Francis, 2009). As such, we also expected a high level of cold tolerance to be displayed by this species. As indicated by the near zero mortality rate under several continuous stresses, *S. tragus* is rightly characterized as a highly stress tolerant species.

The stress screening was performed to inform the selection of abiotic stressors in a transcriptomic analysis of *S. tragus.* The transcriptomes of both shoot and root tissue exposed to salinity and cold stress over a 24-hour period were analyzed through Illumina RNA sequencing. Transcription profiles of these tissue types are vastly different, thus this analysis split samples into seven different contrasts. This approach was undertaken to analyze differentially expressed transcripts in the whole plant versus that of tissue-specific expression. Genes responsible for stress response that present a slight differential expression may be masked when compared to plant-wide transcriptional changes. The use of seven different contrasts to investigate differences in expression allowed for a more thorough look at stress response in each treatment and tissue type.

Circadian differences in expression at the 3, 6, and 8 HAT timepoints may result in differentially expressed genes tied to circadian clock-related expression that do not necessarily relate to treatments (Jones et al., 2019). Our method was used to increase biological replicates in the analysis of both shoot and root tissue expression at different timepoints throughout a 24-hour period. Future analyses should utilize controls that are 24 hours apart from treatment timepoints and should consider diel-specific expression profiles. For example, sampling a control at 3 hours after sunrise and a treated sample taken 27 hours after sunrise would result in a more robust look at treatment-specific expression. This approach would require controls 24 hours before each treatment timepoint, drastically increasing sample sizes. To reduce costs and maintain a minimum

of three biological replicates in this study, the 0 versus 24 HAT comparisons were the only contrast that controlled for circadian rhythm differences in expression patterns.

Temperature and day-part response pathways share signaling pathways, transcription factors, and specific genes. This expression link is due to lower nighttime temperatures, and as such plants have evolved short-term cold expression profiles to deal with lower nighttime temperatures, and long-term cold acclimation expression patterns for extended periods of cold exposure (Jones et al., 2019). Thus, the presence of circadian rhythm-related differentially expressed genes in timepoints that do not control for day-part expression (3, 6, and 8 HAT) may result in the false identification of stress-responsive genes. Results from these timepoints should be interpreted with caution.

The identification and characterization of differentially expressed transcripts in this study was narrowed from the thousands of differentially expressed genes to those that were highly significant, displayed large magnitudes of fold-change in expression, or that were previously identified as integral to stress tolerance. The selection of differentially expressed genes based on significance led to the identification of many transcription factors and binding factors that regulate gene expression and protein function, as opposed to individual stress-responsive genes. These transcription factors may assist in cultivar stress tolerance, as they may be more efficient at their respective processes (A. Mishra & Tanna, 2017). On the contrary, many of the genes displaying large fold-changes in expression were in fact unidentified or otherwise not characterized proteins. Several of these genes were found to have a < -20 $\log_2$ fold change (LFC) in expression. The function of these highly downregulated transcripts may pose insight into the stress tolerance of this species. Those that were upregulated at levels > 20 LFC may pose important novel temperature

or salinity stress responsive genes that have yet to be characterized. As such, the functions of these unidentified transcripts require investigation in future studies.

Overall, this study concludes that *S. tragus* may not present novel stress response pathways. Besides those that were uncharacterized, all the identified differentially expressed genes in this study follow understood mechanisms of stress response. As a tetraploid species with two highly divergent genomes, one may speculate that this species contains highly efficient abiotic stress tolerance genes. These genes may have been selected for resulting in specialized genotypes at different latitudes or elevations or may have evolved *de novo* through the recombination of alleles between the two divergent genomes. The fluidity of polyploid genomes in combination with self-compatibility as an anemophilous tumbleweed drives increased gene-flow over vast distances, potentially allowing this species access to a wide range of potentially specialized stress tolerance genes. The nature of this promotes further investigation into the specialized stress-responsive genes of this species to bolster crop production under the stressful climatic conditions of the future through enhanced abiotic stress tolerance.

Table 3.1: This table reports summary statistics for normalized count values for the six control samples at 0 HAT and twelve treated samples at 24 HAT. Green rows indicate shoot samples, and tan rows indicate root samples. This table shows that shoot tissues present higher maximum counts compared to root tissues by a minimum of four-fold. Also, median values for all tissue types and treatments are comparable in magnitude reporting a standard deviation between all eighteen samples of 18.5. Finally, although shoot tissues present higher maximum values, shoot tissues also have significantly lower 1st quartile values indicating the spread of expression levels in shoot tissues.

	Min	1st Qu	Median	Mean	3rd Qu	Max
Control Samples						
Ct_sh_R1_0	0	44.8	351.1	1,331.4	1,073.3	435,372.9
Ct_sh_R2_0	0	34.7	333.5	1,383.8	1,082.5	476,971.2
Ct_sh_R3_0	0	33.8	339.2	1,465.5	1,110.6	585,570.5
Ct_rt_R1_0	0	75.8	354.9	977.3	983.4	113,968.8
Ct_rt_R2_0	0	75.5	361.4	998.3	991.5	133,116.8
Ct_rt_R3_0	0	77.9	358.9	1,003.8	1,016.3	96,130.7
Cold Treated Samples						
Cd_sh_R1_24	0	26.5	291.3	1,603.2	1,186.5	785,675.1
Cd_sh_R2_24	0	32.1	307.0	1,704.6	1,209.3	652,319.9
Cd_sh_R3_24	0	37.3	315.1	1,436.5	1,144.2	391,358.1
Cd_rt_R1_24	0	55.9	324.9	1,091.9	1,012.2	147,354.2
Cd_rt_R2_24	0	55.8	328.1	1,126.7	1,042.1	126,709.2
Cd_rt_R3_24	0	60.0	329.3	1,084.8	1,031.9	143,698.0
Salt Treated Samples						
St_sh_R1_24	0	30.3	319.3	1,465.9	1,102.3	400,756.6
St_sh_R2_24	0	39.1	346.1	1,409.6	1,101.2	399,349.4
St_sh_R3_24	0	36.8	336.1	1,404.5	1,085.7	406,795.1
St_rt_R1_24	0	59.5	340.6	1,056.2	993.7	166,624.0
St_rt_R2_24	0	73.2	344.4	1,012.9	969.8	143,497.4
St_rt_R3_24	0	68.2	345.9	1,015.5	986.1	108,949.9

Table 3.2: These tables report annotations, InterPro codes, and UniProt accession numbers for *Salsola tragus* gene IDs that were found to be differentially expressed between all seven contrasts. Tables A and B report upregulated transcripts, and downregulated transcripts are found in table C. As discussed in section 3.10, "Galactose oxidase/kelch" and "CheY-like superfamily" were common between tissue types and treatments indicating their generalized role in abiotic stress response.

A. Upregulated Transcripts Between All Contrasts

Gene ID	Uniprot. ID	InterPro ID	Description
SalTrChr04Bg308350	Q8GY91	IPR000823	Plant peroxidase
SalTrChr05Ag118930	P30946	IPR001404	Heat shock protein Hsp90 family
SalTrChr09Ag196440	Q9FMA8	IPR001929	auxin-binding protein ABP19a-like
SalTrChr05Bg326990	Q9H4I3	IPR002816	TraB family
SalTrChr04Bg313440	Q8BLR9	IPR003347	JmjC domain
SalTrChr07Bg368470	Q5UQD7	IPR005874	Eukaryotic translation initiation factor SUI1
SalTrChr08Ag183670	Q8IDA0	IPR006539	P-type ATPase
SalTrChr04Ag079170	Q94A08	IPR008811	Glycosyl hydrolases 36
SalTrChr07Ag165940	P09886	IPR008978	HSP20-like chaperone
SalTrChr04Bg300060	Q869R9	IPR009057	Homeobox-like domain superfamily
SalTrChr01Bg222620	Q6CIP4	IPR010982	Lambda repressor-like
SalTrChr01Ag006170	Q6CIP4	IPR010982	Lambda repressor-like
SalTrChr05Ag121400	Q9ZWJ9	IPR011006	CheY-like superfamily
SalTrChr05Bg336460	Q9ZWJ9	IPR011006	CheY-like superfamily
SalTrChr06Ag129910	Q7M3S9	IPR011043	Galactose oxidase/kelch
SalTrChr06Bg344810	Q9C8K7	IPR011043	Galactose oxidase/kelch

B. Upregulated Transcripts Between All Contrasts

Gene ID	Uniprot. ID	InterPro ID	Description
SalTrChr05Ag106450	Q8L3X8	IPR012677	Nucleotide-binding alpha-beta plait domain superfamily
SalTrChr01Bg228950	Q8W5R5	IPR019396	Transmembrane Fragile-X-F-associated protein
SalTrChr07Bg375940	P27516	IPR022796	Chlorophyll A-B binding protein
SalTrChr06Bg350180	No_ID	IPR026211	GIGANTEA
SalTrChr06Bg343880	O73946	IPR027417	P-loop containing nucleoside triphosphate hydrolase
SalTrChr01Bg222920	Q02953	IPR027725	Heat shock transcription factor family
SalTrChr07Bg363570	No_ID	IPR027942	Sieve element occlusion
SalTrChr03Bg286720	P02520	IPR031107	Small heat shock protein HSP20
SalTrChr03Ag067470	P27777	IPR031107	Small heat shock protein HSP20
SalTrChr08Ag187120	Q6YXY2	IPR035898	TAZ domain superfamily
SalTrChr02Ag028590	F4IC29	IPR036163	Heavy metal-associated domain superfamily
SalTrChr02Ag032720	Q5HTK3	IPR036869	Chaperone J-domain superfamily
SalTrChr09Ag201590	Q5SLW9	IPR036869	Chaperone J-domain superfamily
SalTrChr09Bg416870	Q5SLW9	IPR036869	Chaperone J-domain superfamily
SalTrChr03Ag065930	Q8H1E4	IPR036955	AP2/ERF domain superfamily
SalTrChr06Ag138210	Q7XD65	IPR039271	Kiwellin-like
SalTrChr04Ag097900	Q54FG7	IPR041667	Cupin-like domain 8

C. Downregulated Transcripts Between All Contrasts

Gene ID	Uniprot. ID	InterPro ID	Description
SalTrChr04Bg301910	Q9M9Q6	IPR001563	Peptidase S10
SalTrChr09Bg411560	Q9FMA8	IPR001929	Germin
SalTrChr03Ag059800	O42906	IPR004130	GPN-loop GTPase
SalTrChr06Bg354650	O65482	IPR011009	Protein kinase-like domain superfamily
SalTrChr05Ag103090	Q5HZL9	IPR011949	HAD-superfamily hydrolase
SalTrChr07Bg364500	P41208	IPR011992	EF-hand domain pair
SalTrChr04Bg304260	Q5XV31	IPR012474	Frigida-like
SalTrChr04Bg301870	No_ID	IPR015915	Kelch-type beta propeller
SalTrChr07Bg367600	No_ID	IPR015915	Kelch-type beta propeller
SalTrChr05Ag118100	Q53QK0	IPR016461	O-methyltransferase COMT-type
SalTrChr07Ag146460	Q8GW80	IPR032675	Leucine-rich repeat domain superfamily
SalTrChr03Ag074460	Q9FFJ3	IPR032675	Leucine-rich repeat domain superfamily
SalTrChr06Bg344510	Q9C9Y6	IPR036249	Thioredoxin-like superfamily
SalTrChr02Bg261570	O80654	IPR036955	AP2/ERF
SalTrChr03Bg288760	Q948U0	IPR039261	Ferredoxin-NADP reductase
SalTrChr06Bg349590	Q9LHH5	IPR039928	LNK family
SalTrChr06Ag141140	Q9LHH5	IPR039928	LNK family
SalTrChr08Ag184120	F4IMQ0	IPR040353	FLC EXPRESSOR/FLX-like
SalTrChr04Bg301380	A0A0S6XHF8	IPR042099	AMP-dependent synthetase-like superfamily

FIGURES

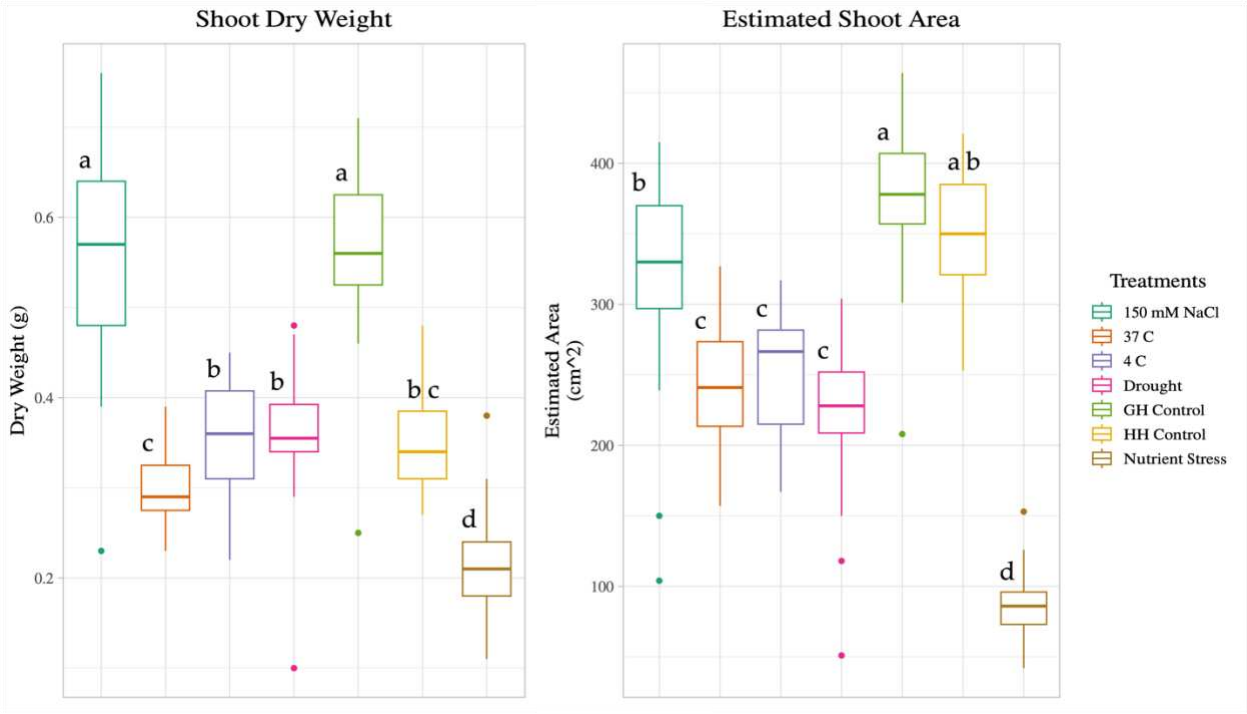


Figure 3.1: Boxplots displaying the spread of collected data on shoot biomass and estimated shoot surface area. Significant differences between treatments were identified through Tukey adjusted pairwise comparisons and are indicated as letters above each boxplot. Those with the same letter are not significantly different from each other. The greenhouse control (light green) shoot biomass was found to be significantly larger than all treatment groups except salt treated samples. The headhouse control (yellow) was significantly different from the greenhouse control in terms of biomass, but not shoot surface area. Cold treated samples were not statistically different from drought samples for either biomass or surface area.

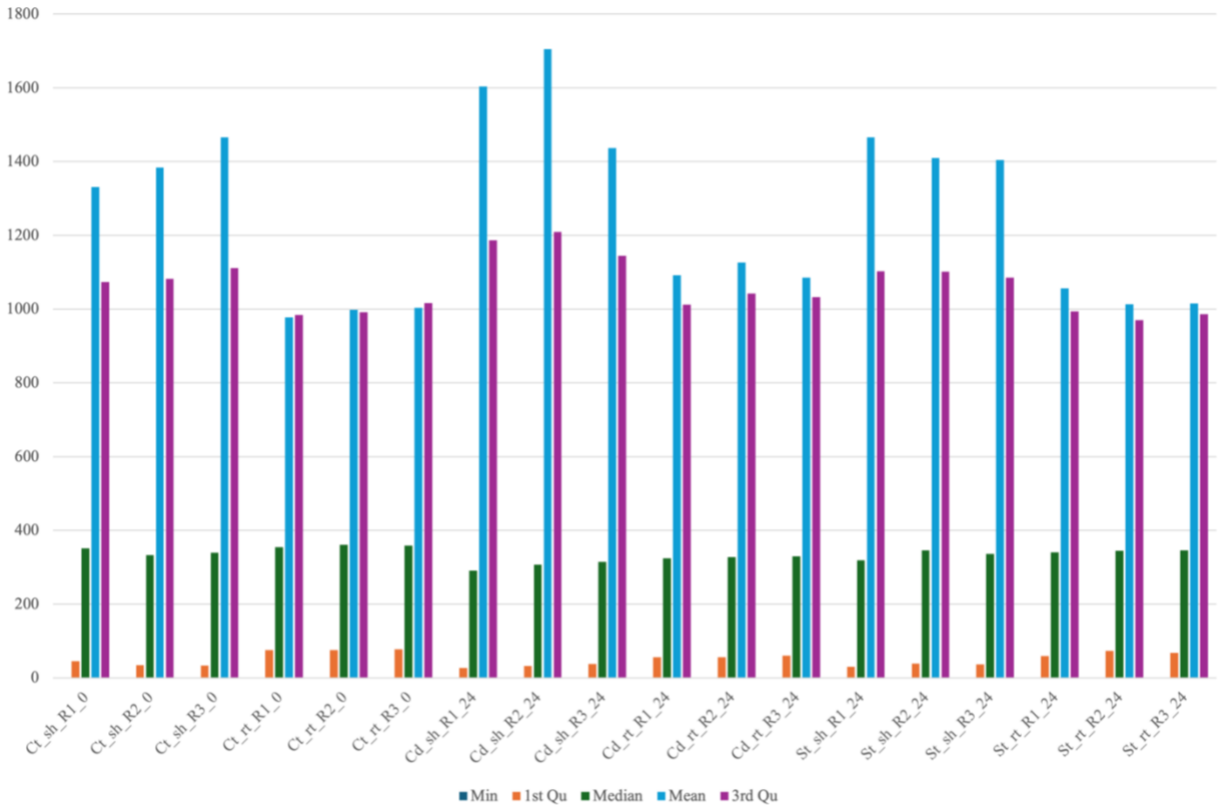
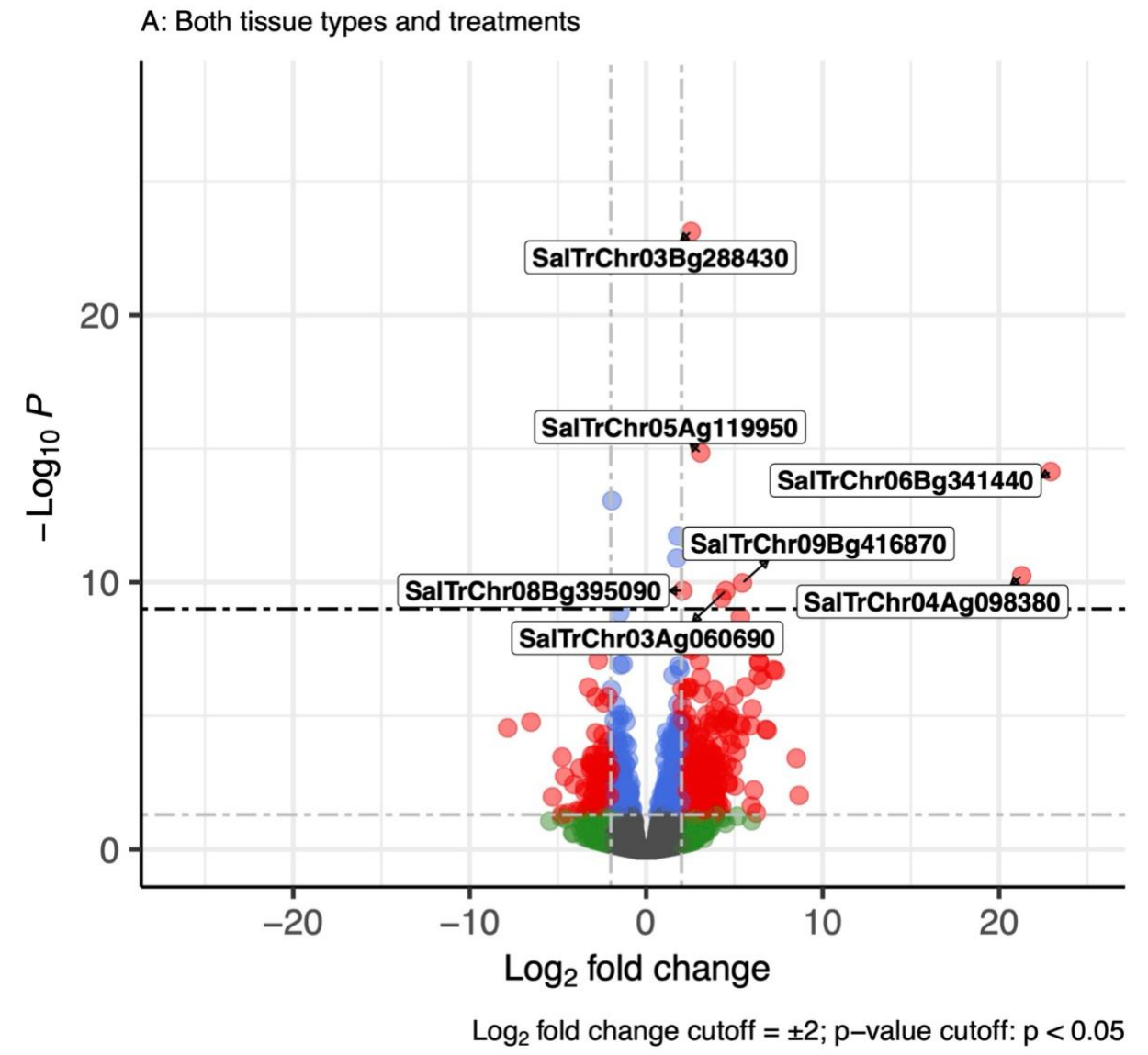
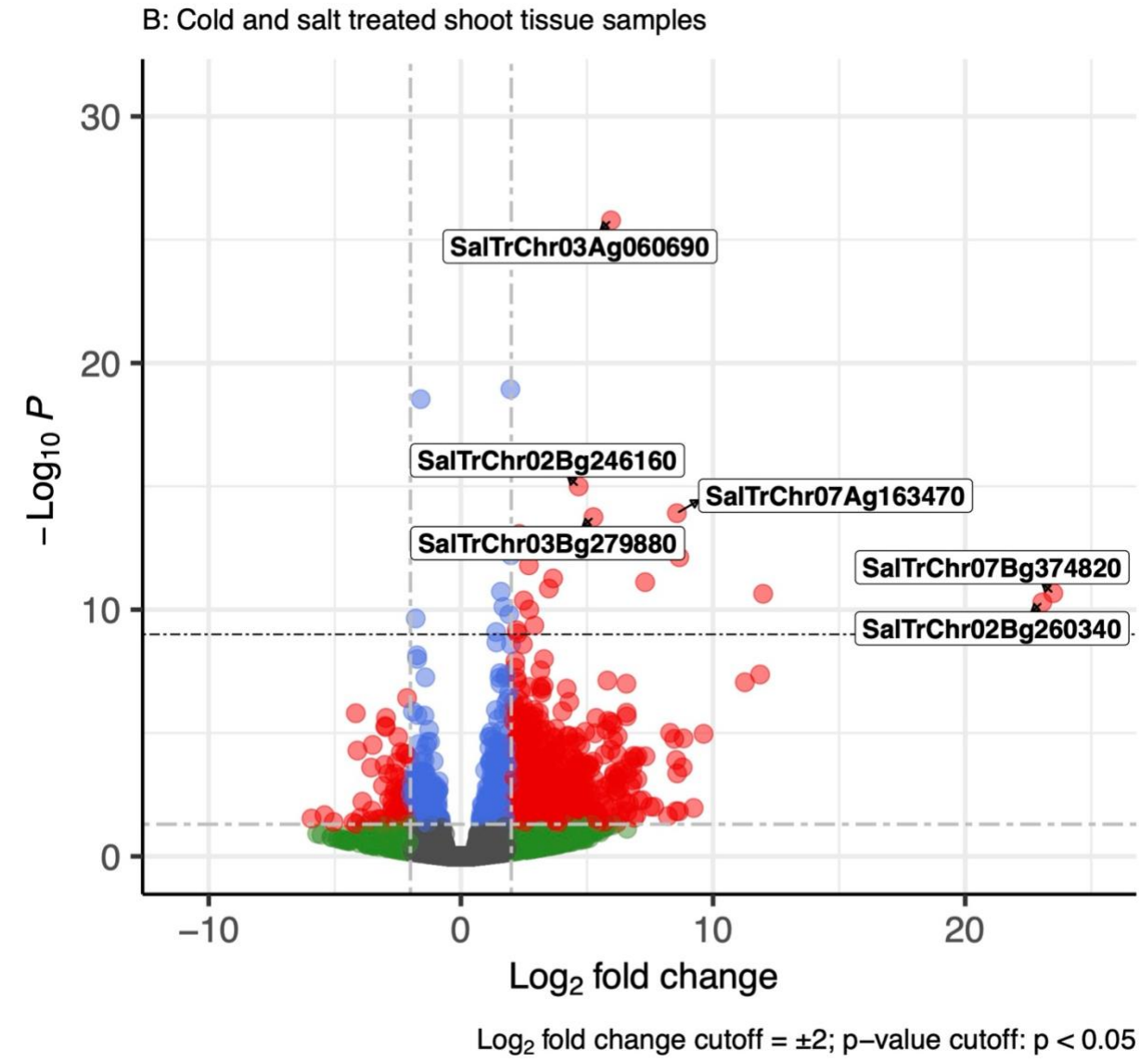


Figure 3.2: A bar chart summarizing root and shoot transcript normalized count values for the six control samples and twelve samples collected 24 hours after treatment. Mean count values (blue) indicate transcript abundance differences between shoot and root tissues in this species; however, median count values (green) are relatively consistent between all eighteen samples. Based on these data, the baseline expression across tissue types for *S. tragus* samples in this study is presumed to be a normalized count value between the median value of 400 and the mean value of 1200 transcripts. The x-axis indicates individual sample designations that include all treatment characteristics of that sample. Control (Ct), cold (Cd), and Salt (St) treated samples included both shoot (sh) and root (rt) tissue types, three biological replicates (R1, R2, R3) and different timepoints (0 and 24).


A: Both tissue types and treatments
SalTrChr03Bg288430
SalTrChr05Ag119950
SalTrChr06Bg341440
SalTrChr09Bg416870
SalTrChr08Bg395090
SalTrChr04Ag098380
SalTrChr03Ag060690
20
10
0
−Log10 P
−20
−10
0
10
20
Log2 fold change
Log2 fold change cutoff = ±2; p−value cutoff: p < 0.05


B: Cold and salt treated shoot tissue samples
30
20
10
0
−Log10 P
SalTrChr03Ag060690
SalTrChr02Bg246160
SalTrChr07Ag163470
SalTrChr03Bg279880
SalTrChr07Bg374820
SalTrChr02Bg260340
−10
0
10
20
Log2 fold change
Log2 fold change cutoff = ±2; p−value cutoff: p < 0.05

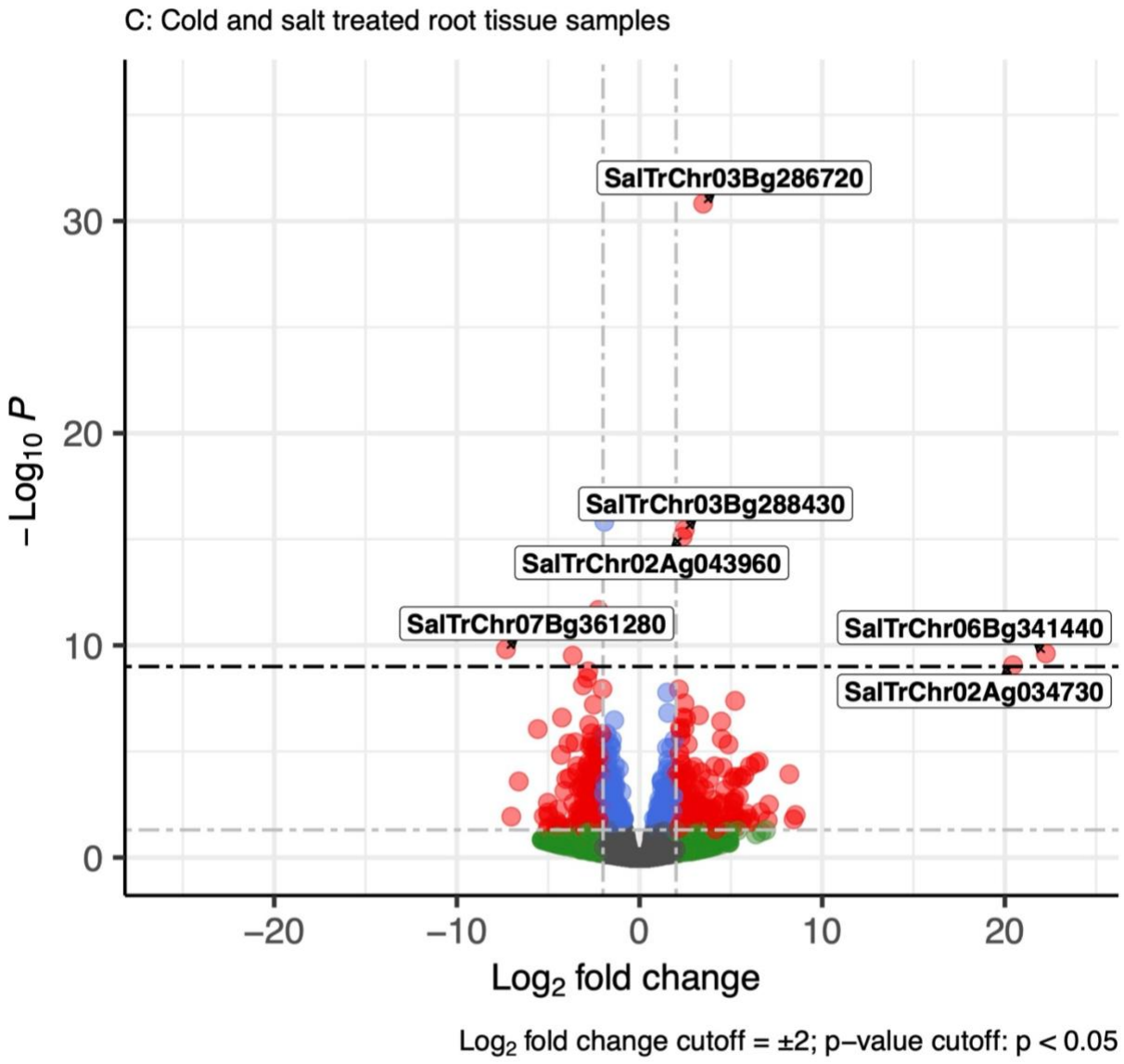
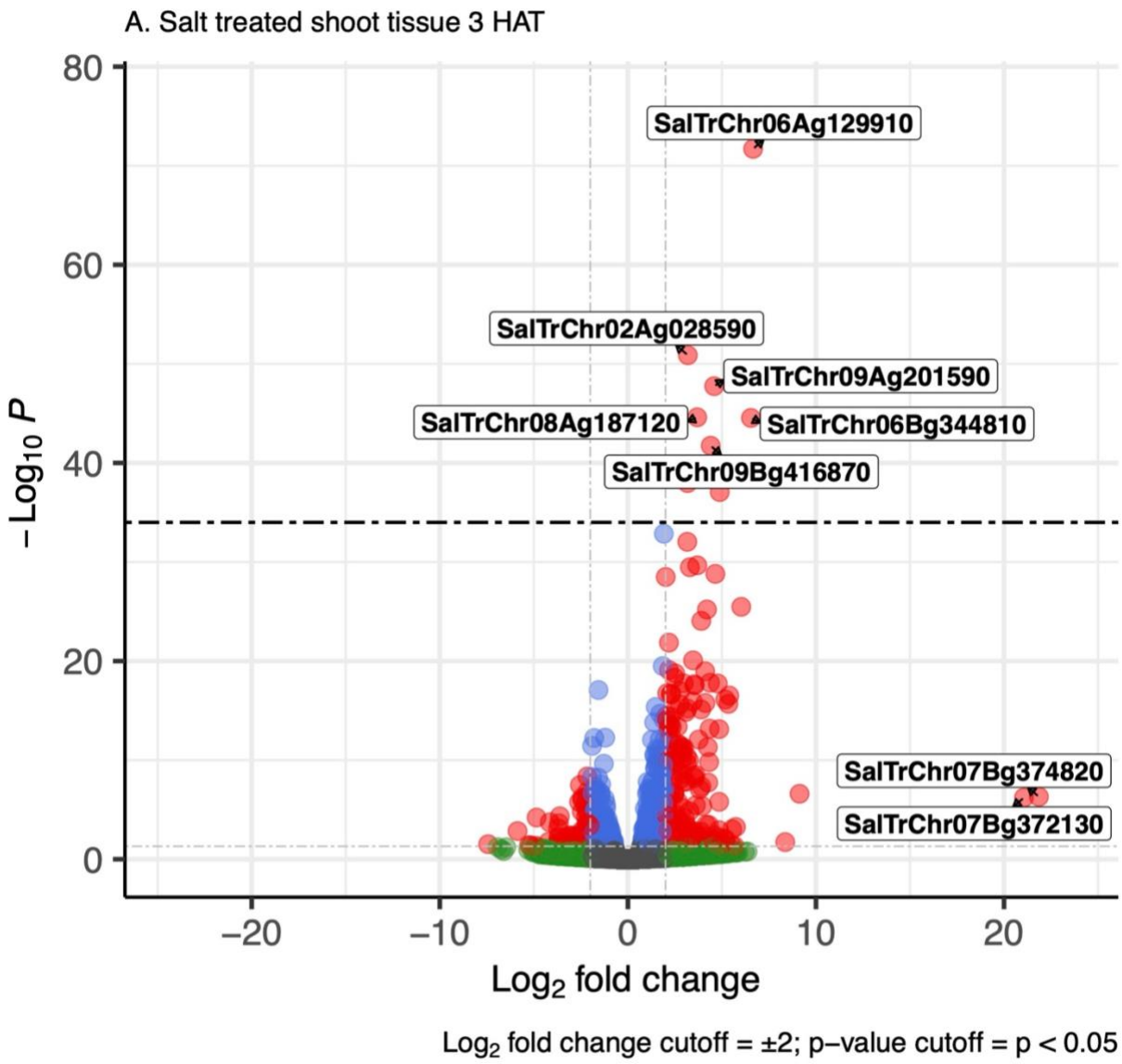
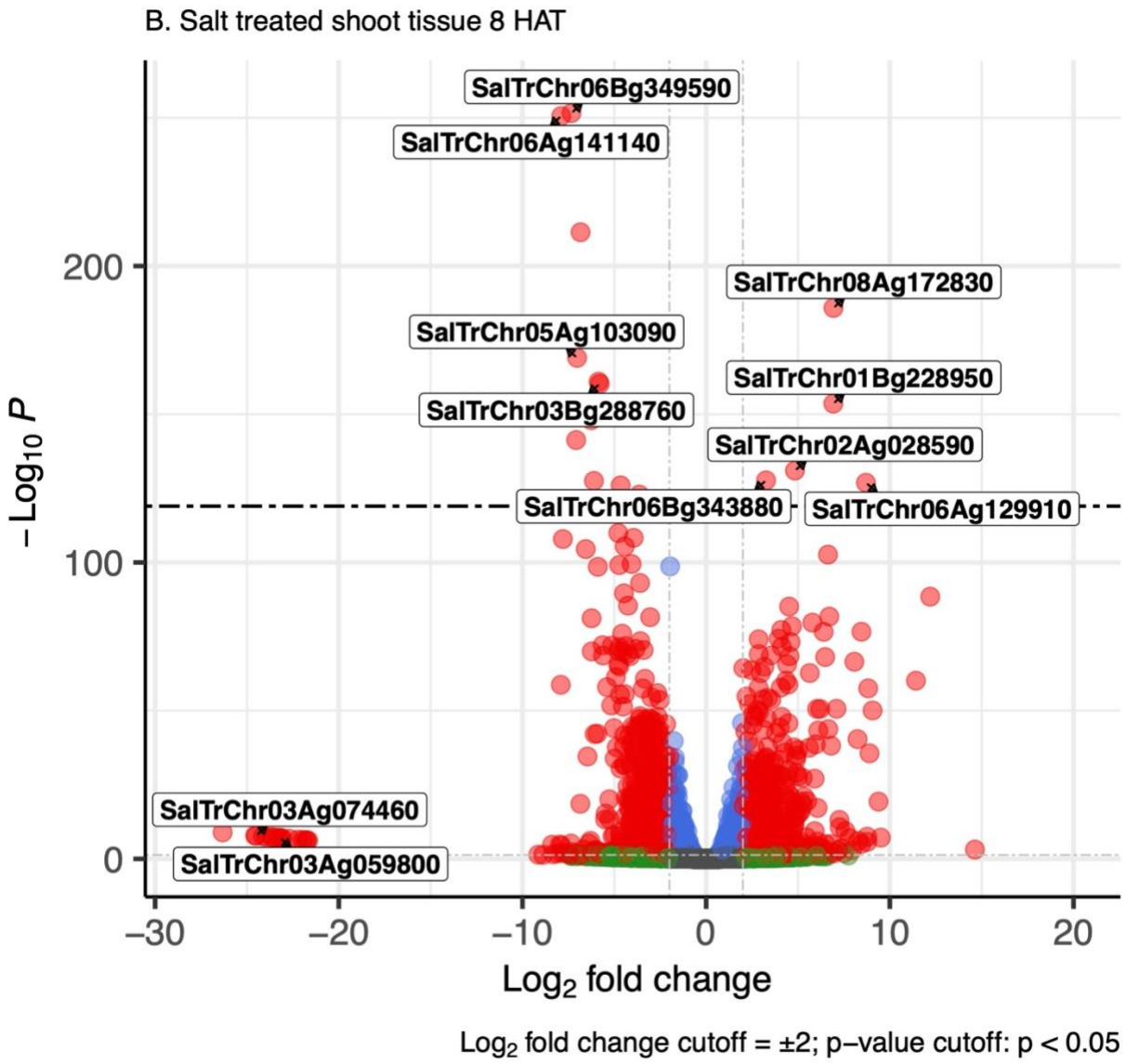
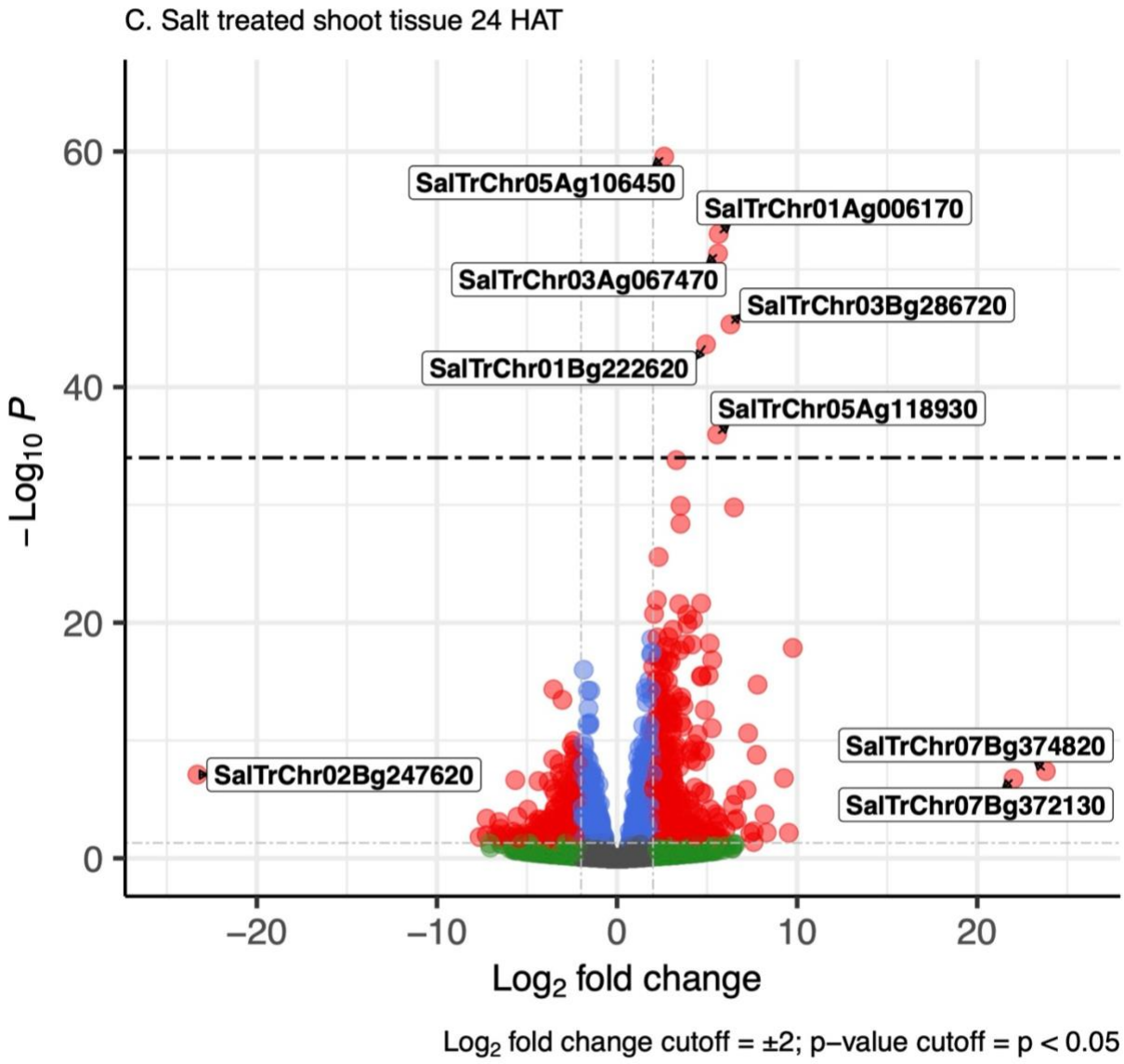
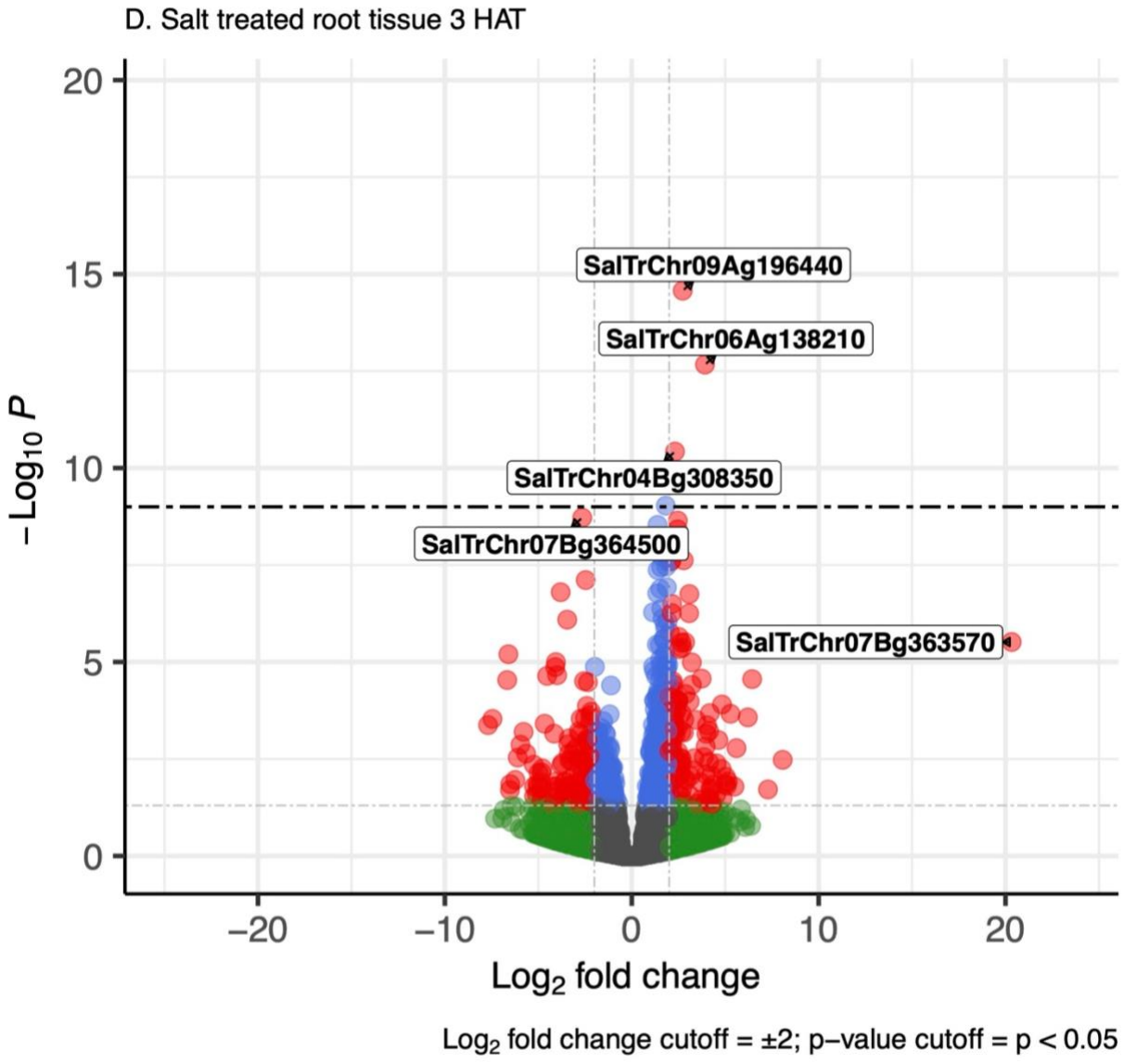


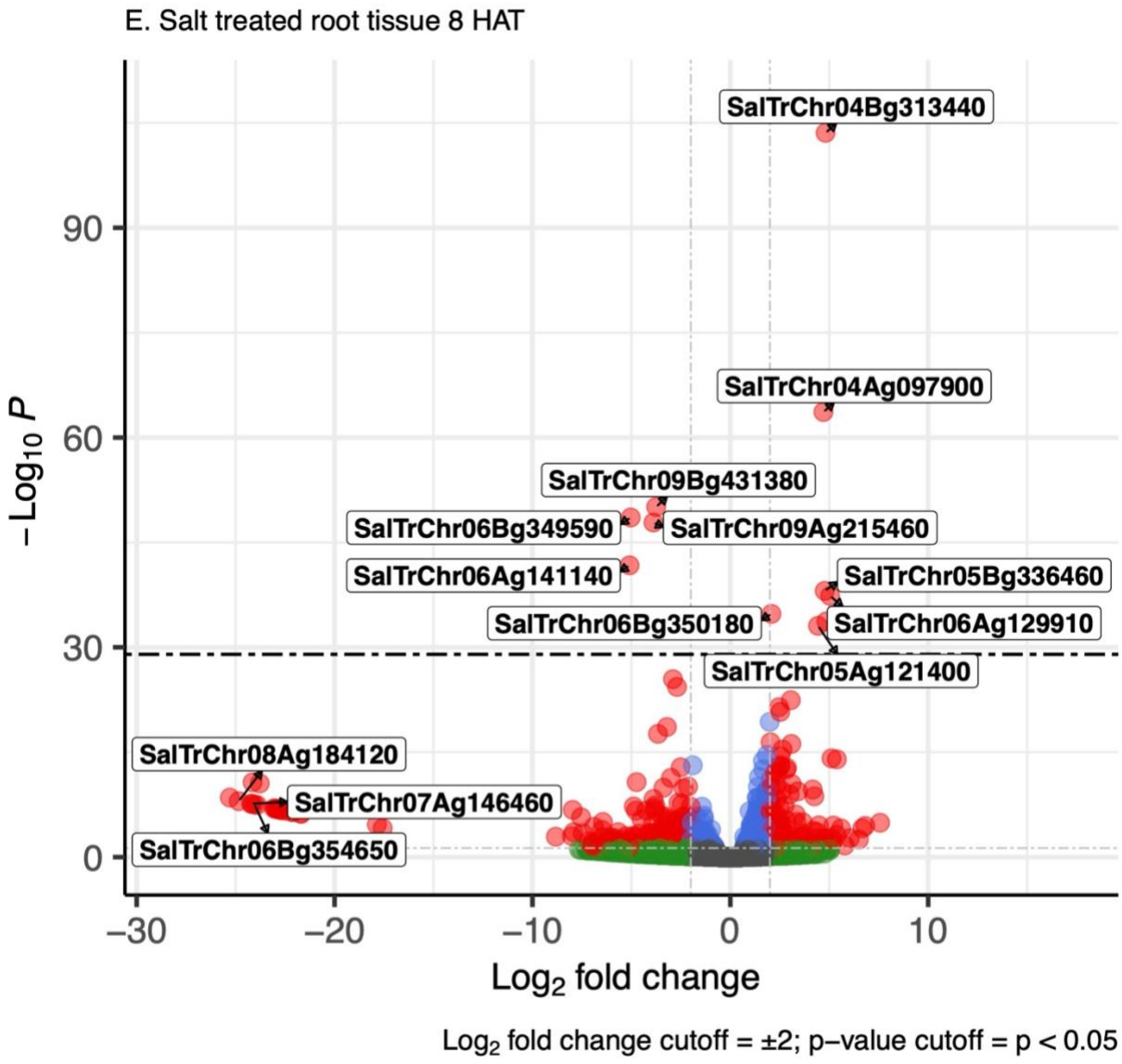
Figure 3.3: Plots A, B, and C indicate significant differentially expressed genes at 24 HAT for all tissues (A), shoot tissues (B), and root tissues (C). All significant differentially expressed genes are colored red and were defined as those with an adjusted p-value less than 0.05. This threshold is indicated by the horizontal grey line. The vertical grey lines indicate a $\log_2$ fold change (LFC) threshold of ±2. Genes that are below the p-value threshold are indicated in green if they meet the LFC threshold, or grey if they are insignificant. Blue points indicate genes that have a p-value less than 0.05 but do not exceed the LFC threshold. This plot displays the massive number of differentially expressed transcripts found in this study. As such, an additional p-value threshold is indicated by the black dashed line and was set to select a subset of highly significant differentially expressed genes and are labeled with *S. tragus* gene IDs. Several of these genes are discussed in section 3.10.1.

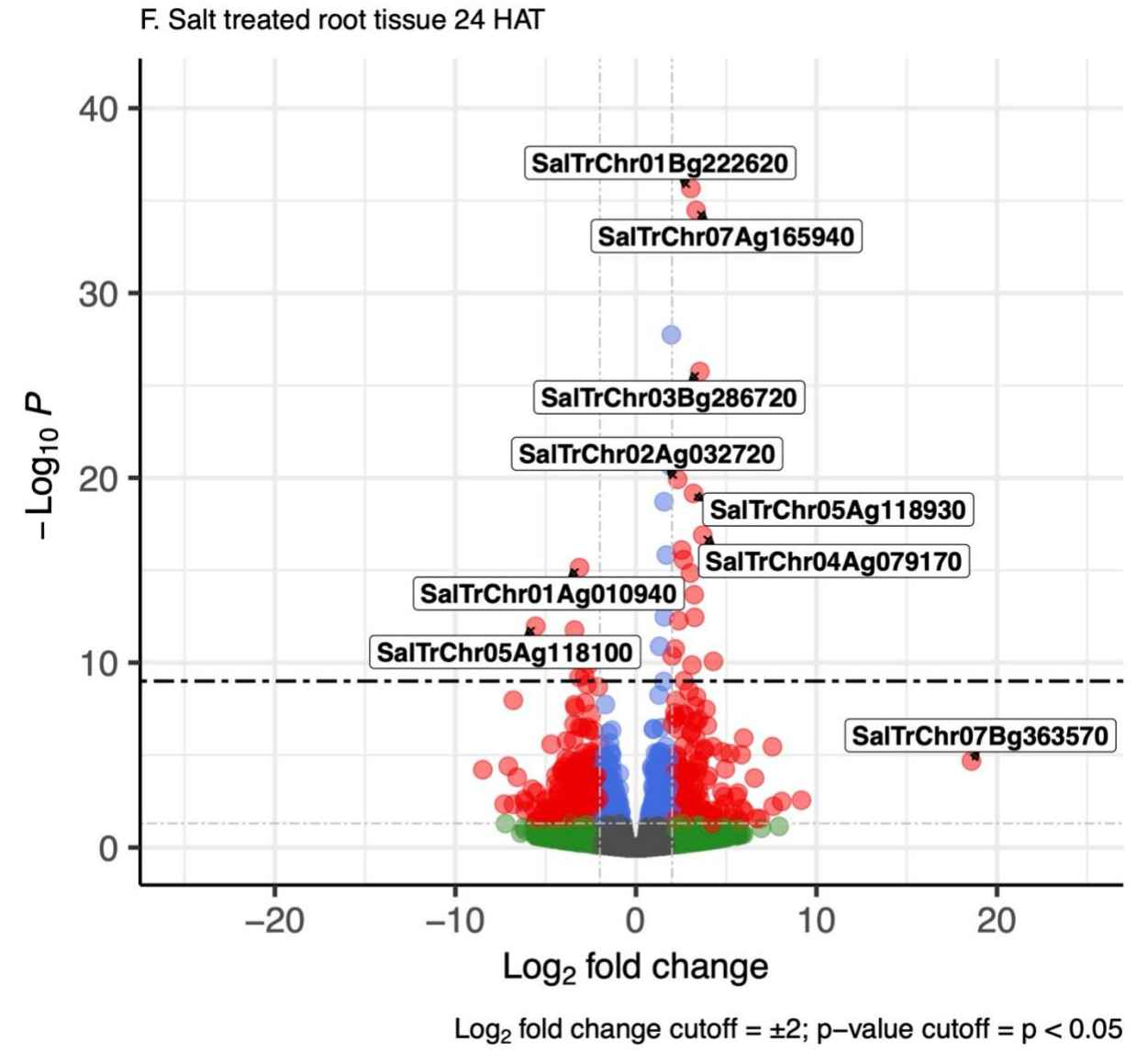

A. Salt treated shoot tissue 3 HAT
80
60
40
20
0
−Log10 P
SalTrChr06Ag129910
SalTrChr02Ag028590
SalTrChr09Ag201590
SalTrChr08Ag187120
SalTrChr06Bg344810
SalTrChr09Bg416870
SalTrChr07Bg374820
SalTrChr07Bg372130
−20
−10
0
10
20
Log2 fold change
Log2 fold change cutoff = ±2; p−value cutoff = p < 0.05

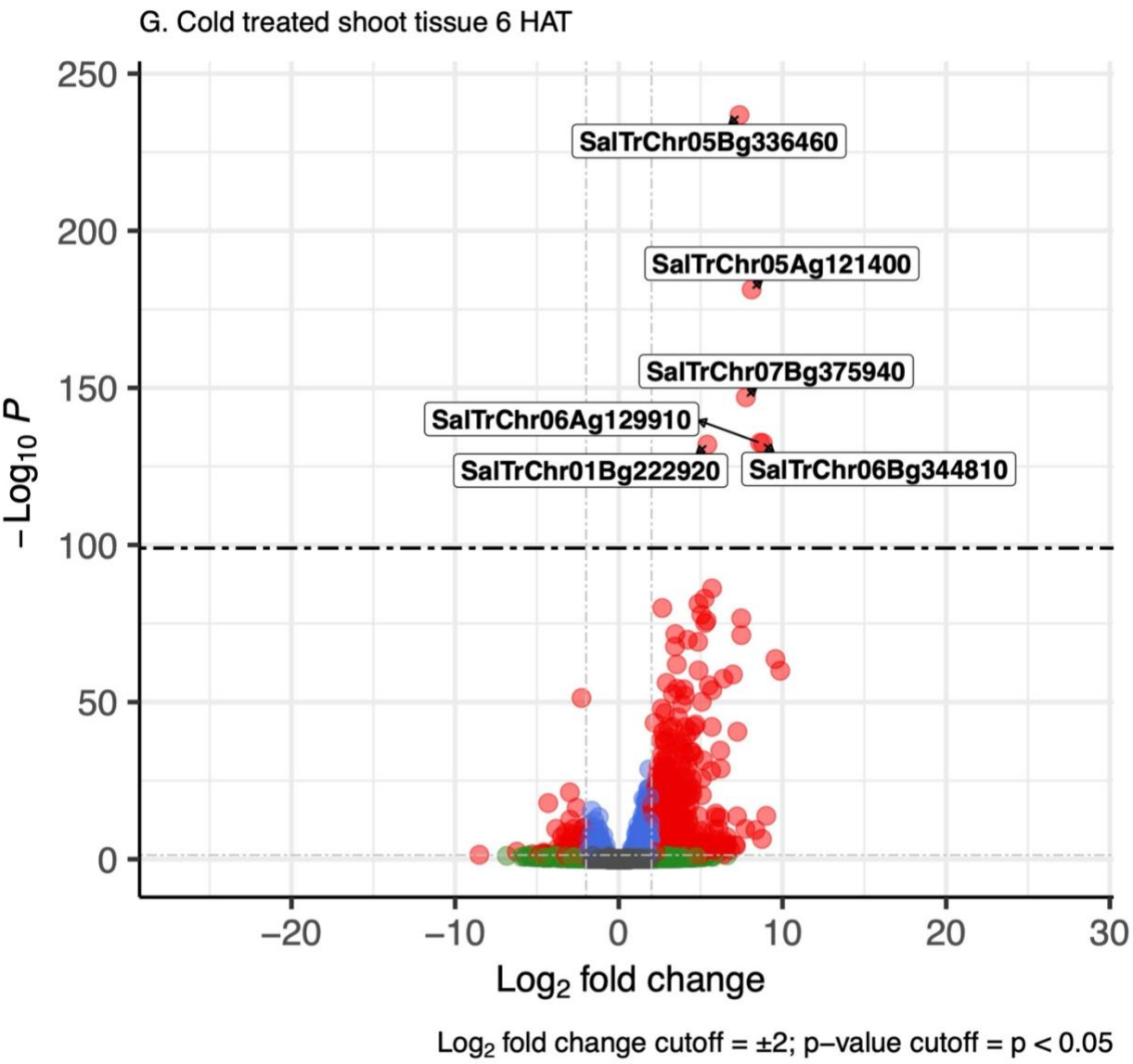

B. Salt treated shoot tissue 8 HAT
SalTrChr06Bg349590
SalTrChr06Ag141140
SalTrChr08Ag172830
SalTrChr05Ag103090
SalTrChr01Bg228950
SalTrChr03Bg288760
SalTrChr02Ag028590
SalTrChr06Bg343880
SalTrChr06Ag129910
SalTrChr03Ag074460
SalTrChr03Ag059800
200
100
0
−Log10 P
−30
−20
−10
0
10
20
Log2 fold change
Log2 fold change cutoff = ±2; p−value cutoff: p < 0.05

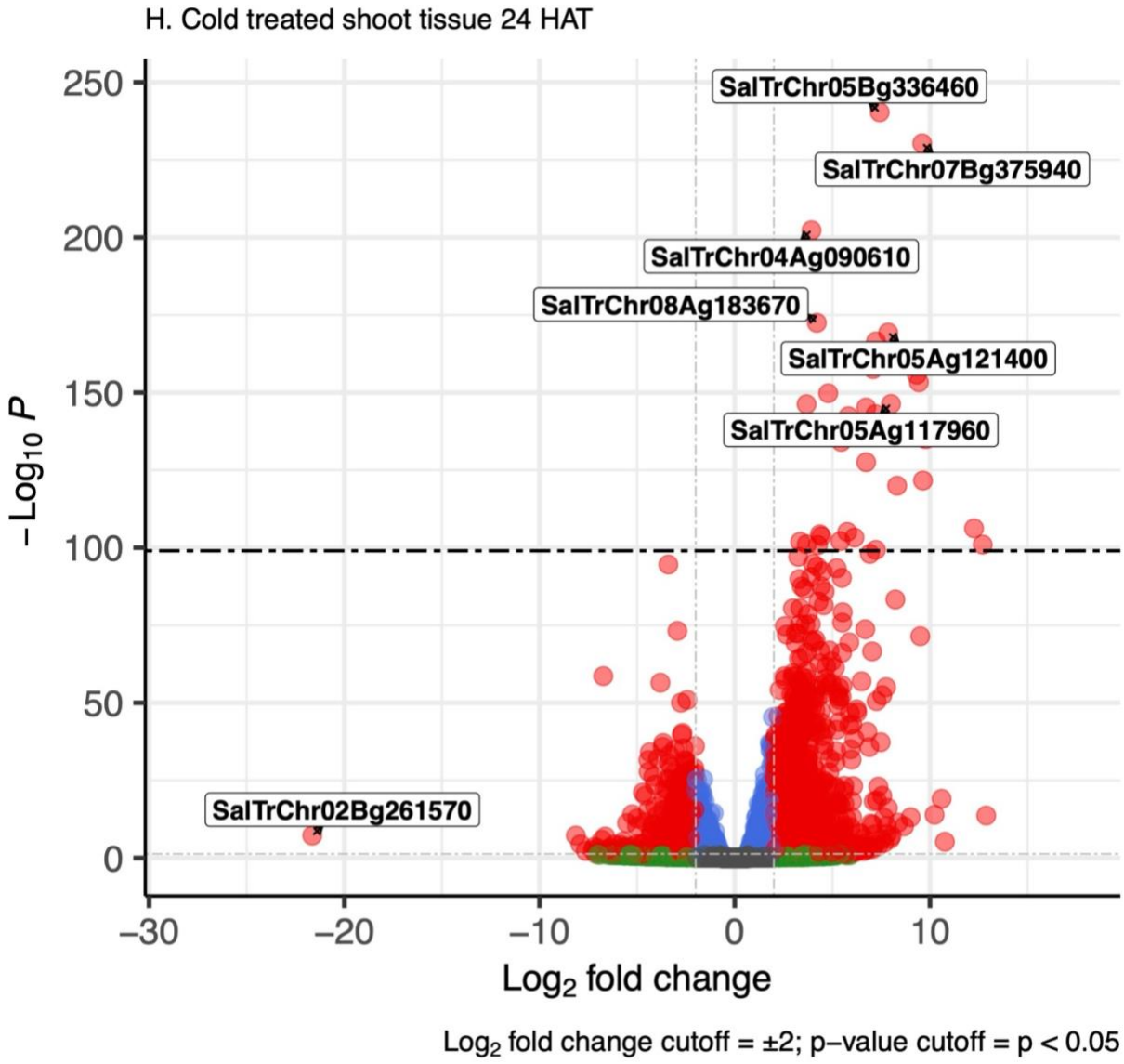

C. Salt treated shoot tissue 24 HAT
60
40
20
0
−Log10 P
SalTrChr05Ag106450
SalTrChr01Ag006170
SalTrChr03Ag067470
SalTrChr03Bg286720
SalTrChr01Bg222620
SalTrChr05Ag118930
SalTrChr07Bg374820
SalTrChr07Bg372130
SalTrChr02Bg247620
−20
−10
0
10
20
Log2 fold change
Log2 fold change cutoff = ±2; p−value cutoff = p < 0.05

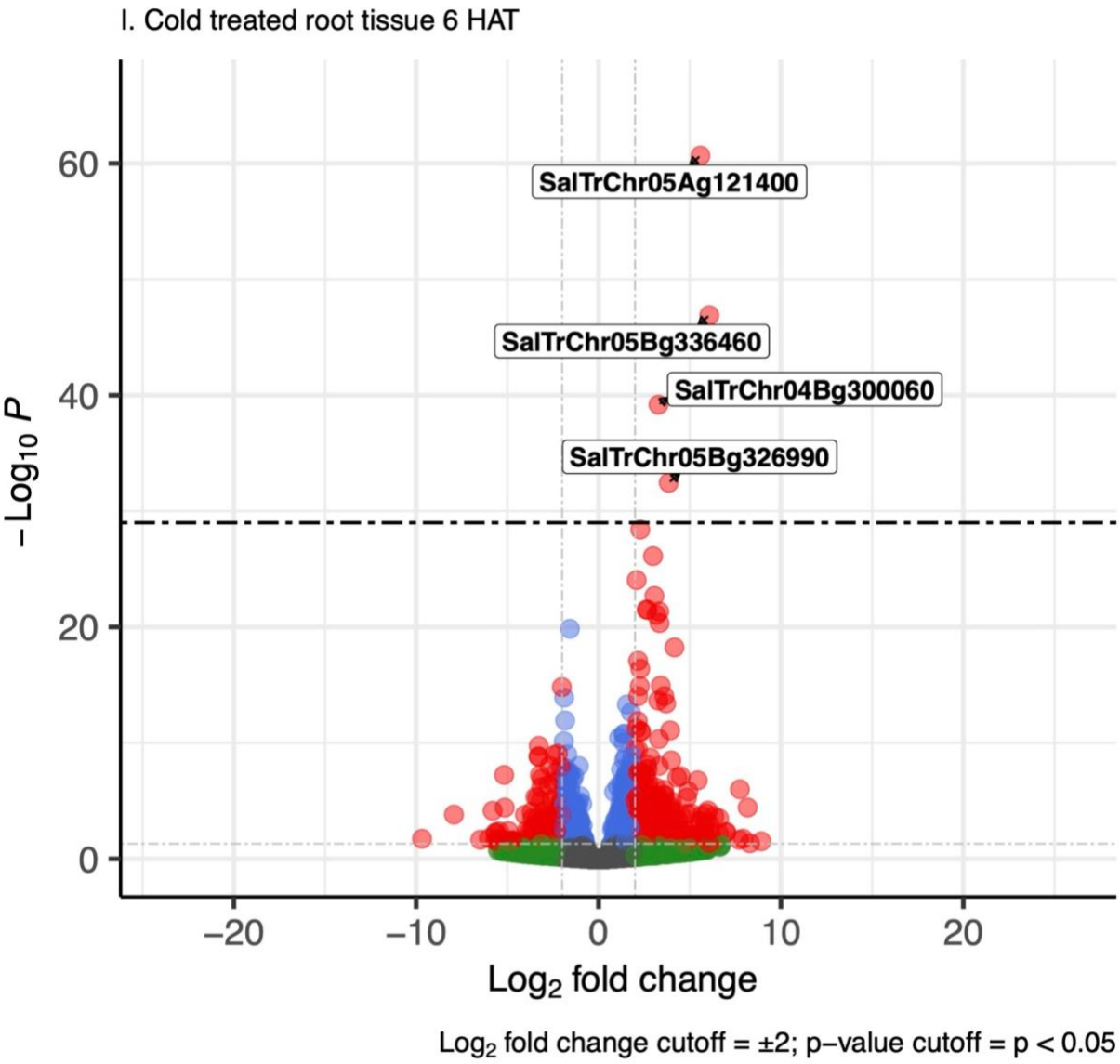

D. Salt treated root tissue 3 HAT
20
15
10
5
0
−Log10 P
SalTrChr09Ag196440
SalTrChr06Ag138210
SalTrChr04Bg308350
SalTrChr07Bg364500
SalTrChr07Bg363570
−20
−10
0
10
20
Log2 fold change
Log2 fold change cutoff = ±2; p−value cutoff = p < 0.05


E. Salt treated root tissue 8 HAT
SalTrChr04Bg313440
SalTrChr04Ag097900
SalTrChr09Bg431380
SalTrChr06Bg349590
SalTrChr09Ag215460
SalTrChr06Ag141140
SalTrChr05Bg336460
SalTrChr06Bg350180
SalTrChr06Ag129910
SalTrChr05Ag121400
SalTrChr08Ag184120
SalTrChr07Ag146460
SalTrChr06Bg354650
90
60
30
0
−Log10 P
−30
−20
−10
0
10
Log2 fold change
Log2 fold change cutoff = ±2; p−value cutoff = p < 0.05


F. Salt treated root tissue 24 HAT
SalTrChr01Bg222620
SalTrChr07Ag165940
SalTrChr03Bg286720
SalTrChr02Ag032720
SalTrChr05Ag118930
SalTrChr04Ag079170
SalTrChr01Ag010940
SalTrChr05Ag118100
SalTrChr07Bg363570
40
30
20
10
0
−Log10 P
−20
−10
0
10
20
Log2 fold change
Log2 fold change cutoff = ±2; p−value cutoff = p < 0.05


G. Cold treated shoot tissue 6 HAT
SalTrChr05Bg336460
SalTrChr05Ag121400
SalTrChr07Bg375940
SalTrChr06Ag129910
SalTrChr01Bg222920
SalTrChr06Bg344810
250
200
150
100
50
0
−Log10 P
−20
−10
0
10
20
30
Log2 fold change
Log2 fold change cutoff = ±2; p−value cutoff = p < 0.05


H. Cold treated shoot tissue 24 HAT
SalTrChr05Bg336460
SalTrChr07Bg375940
SalTrChr04Ag090610
SalTrChr08Ag183670
SalTrChr05Ag121400
SalTrChr05Ag117960
SalTrChr02Bg261570
250
200
150
100
50
0
−Log10 P
−30
−20
−10
0
10
Log2 fold change
Log2 fold change cutoff = ±2; p−value cutoff = p < 0.05


I. Cold treated root tissue 6 HAT
SalTrChr05Ag121400
SalTrChr05Bg336460
SalTrChr04Bg300060
SalTrChr05Bg326990
60
40
20
0
−Log10 P
−20
−10
0
10
20
Log2 fold change
Log2 fold change cutoff = ±2; p−value cutoff = p < 0.05

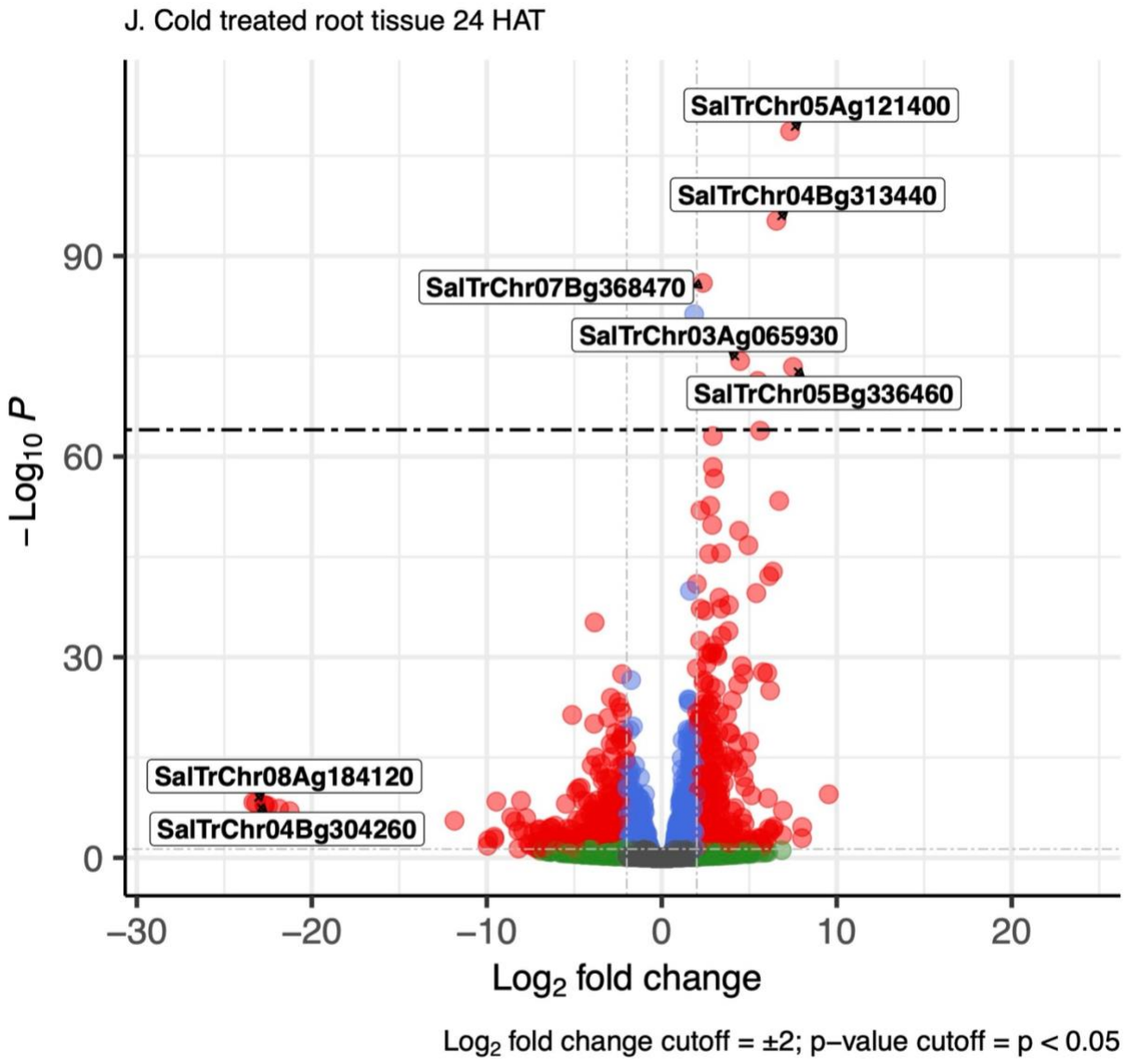


Figure 3.4: Plots A through J display differentially expressed genes for the four additional DESeq contrasts run with respective selected samples. Plots A through C included only salt treated shoot tissues at 3, 8, and 24 hours after treatment (HAT). Plots D through F included only salt treated root tissue samples at 3, 8, and 24 HAT. Plots G and H included cold treated shoot tissue samples at 6 and 24 HAT. Plots I and J included cold treated root tissue samples at 6 and 24 HAT. Grey lines indicate an adjusted p-value threshold at 0.05 (horizontal) and a $\log_2$ fold change (LFC) in expression of ±2. Green and grey points are insignificant differentially expressed genes based on these thresholds. Highly significant differentially expressed genes were selected based on the additional p-value threshold for each timepoint. *S. tragus* gene IDs for these genes are labeled in these plots and discussed in section 3.10.2. Volcano plots for 3, 6, and 8 HAT may contain significant differentially expressed genes that relate to circadian rhythm alterations to gene expression in addition to treatment-induced expression.

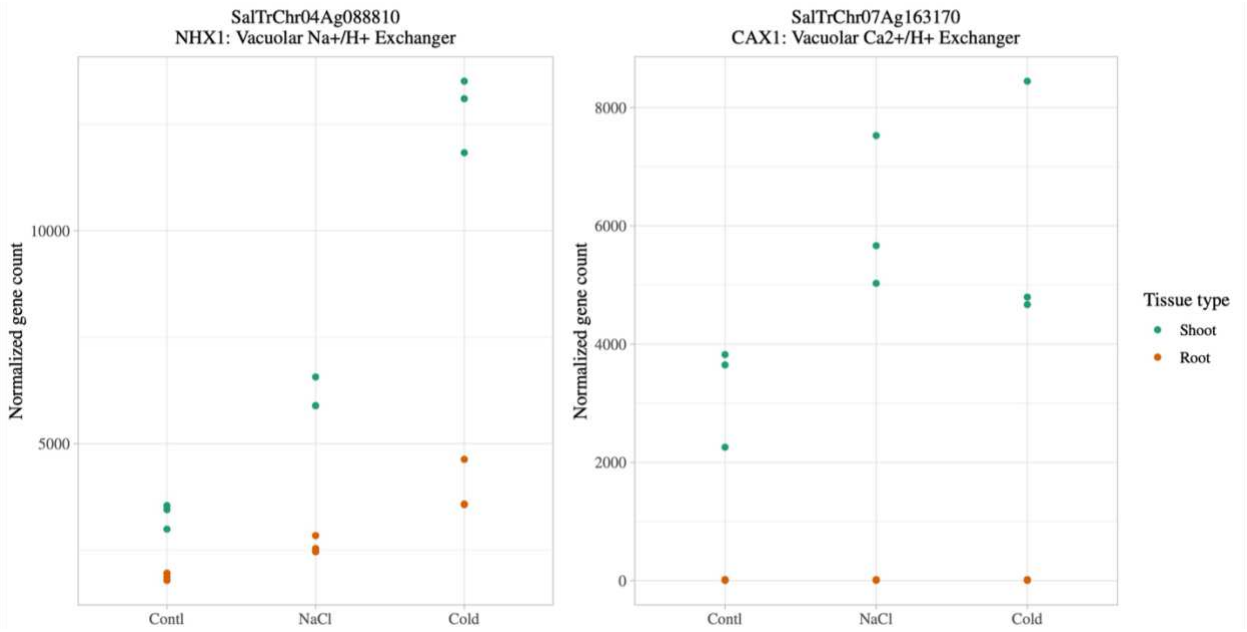


Figure 3.8: Dot plots of normalized counts for both shoot and root tissue 24 hours after treatment (HAT) for a sodium/hydrogen exchanger (presumed NHX1) and a calcium exchanger (presumed CAX1). Ion homeostasis under salinity stress is maintained through the expression of non-selective antiporters that work together to translocate sodium ions throughout the plant for vacuolar sequestration. In addition, many stress response expression pathways are reliant on cytosolic calcium concentrations. The presence of calcium activates calcium-activated protein kinase pathways that are important for both cold and salt stress response. The CAX1 plot (right) shows that the activity of this calcium antiporter in *S. tragus* is more active in shoot tissues than in root tissues. The NHX1 plot (left) indicates the importance of plant wide sodium sequestration, however the expression of this antiporter is still higher in shoot tissues.

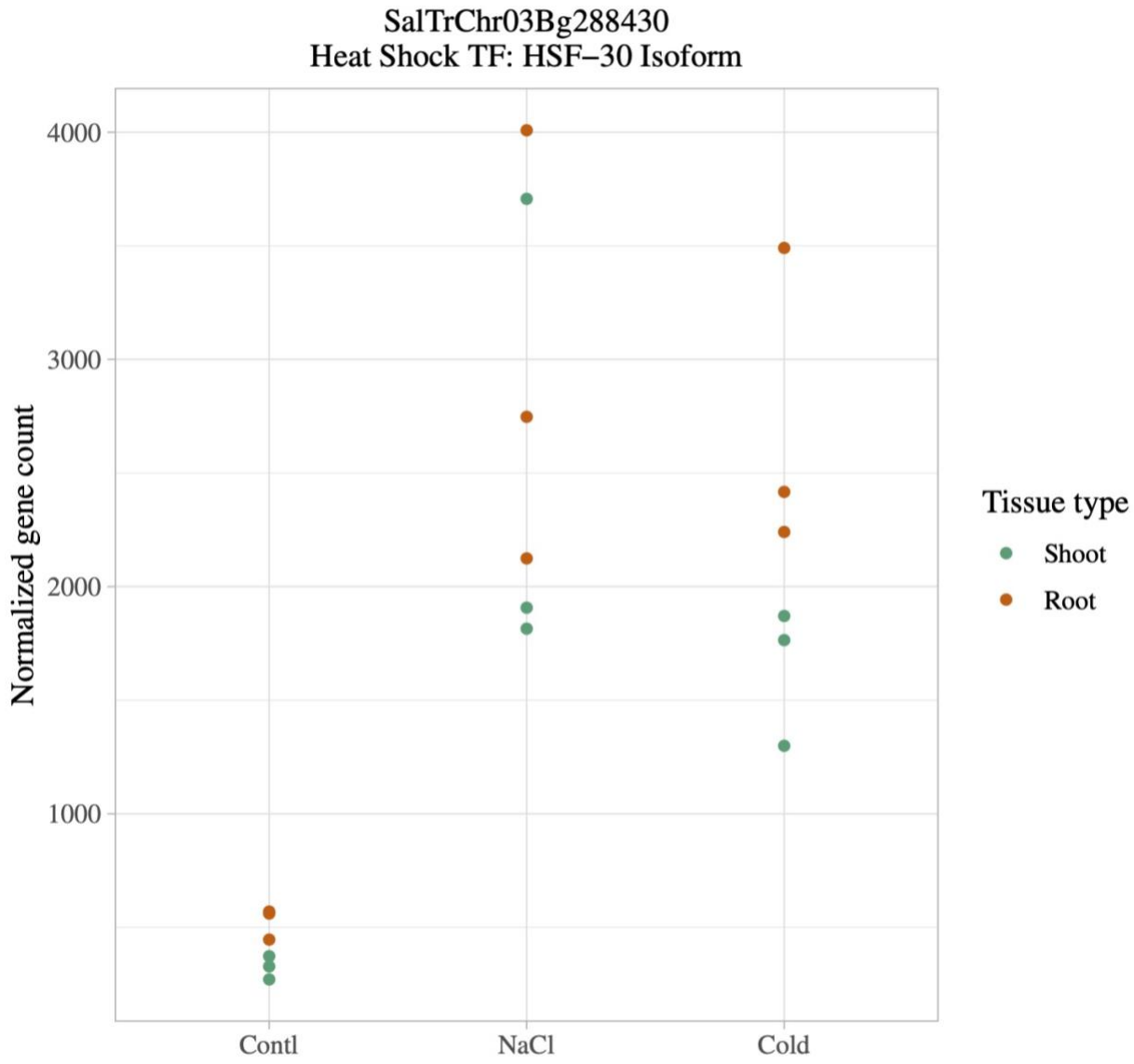


Figure 3.9: This dot plot displays the expression of the heat shock transcription factor HSF 30 in both tissue types and treatments 24 hours after treatment. As shown in figure 3.3A this gene was found to have the greatest level of significance. This plot indicates the importance of this transcription factor in both salinity and temperature stress and suggests that HSF30 may be highly active in root tissues under both cold and salinity stress.

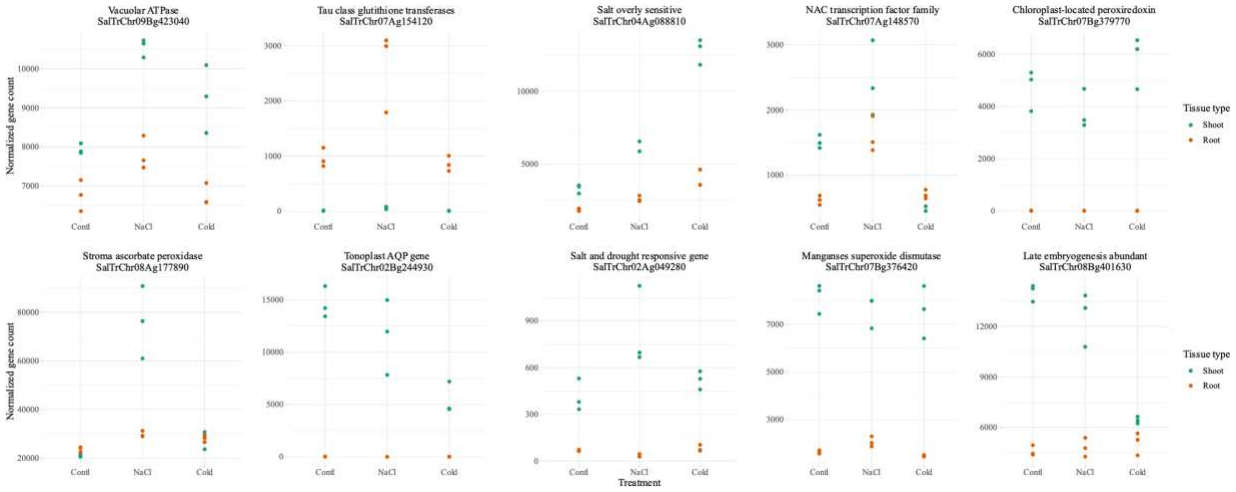
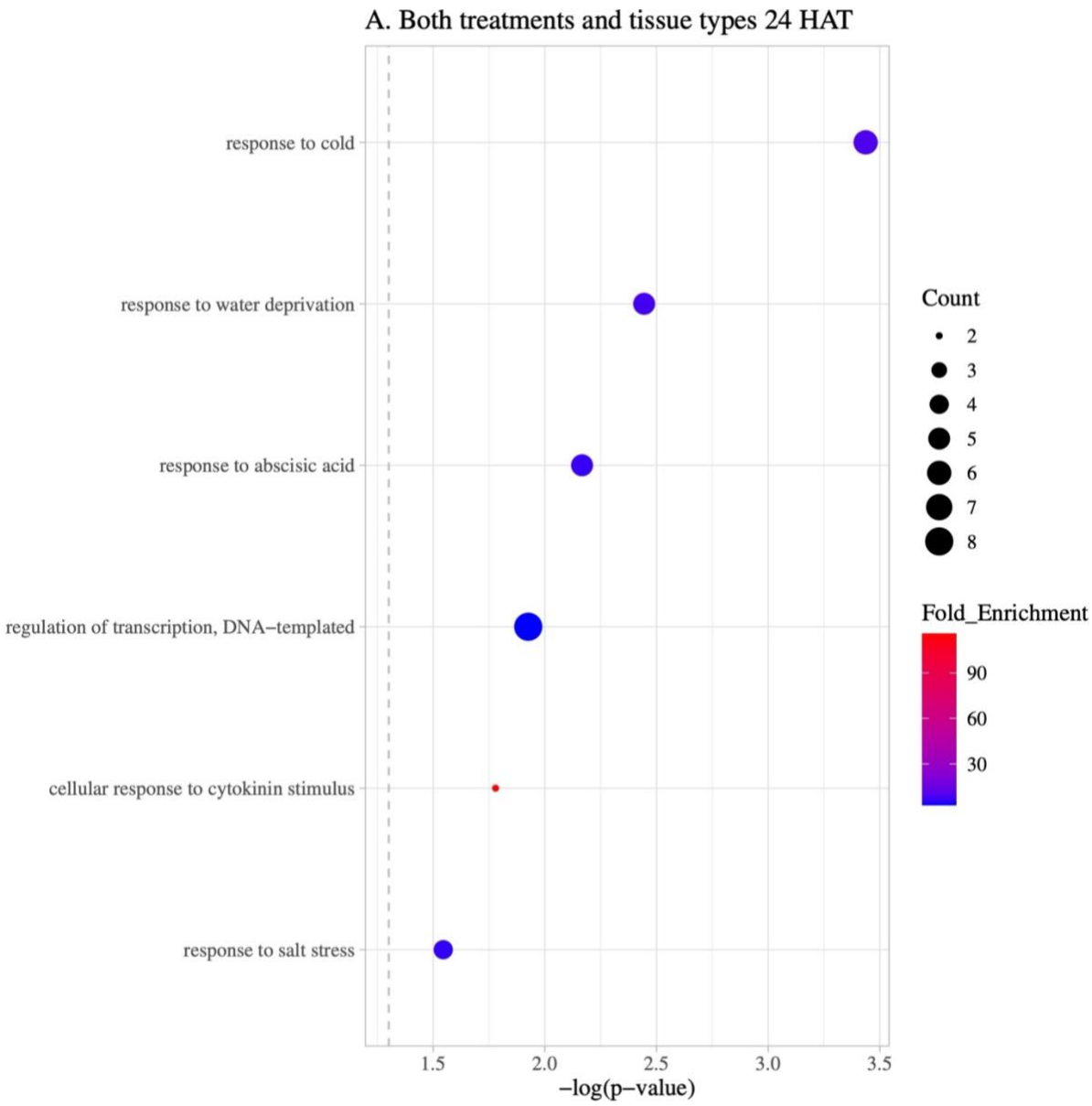


Figure 3.10: These counts plots report transcript counts for several additional important salt stress responsive genes (A. Mishra & Tanna, 2017). Each of these genes present potential utility in increasing salinity tolerance to desirable cultivars. These counts include both cold and salt treated shoot and root tissues and indicate varying levels of expression at 24 hours after treatment. Differences in expression between tissue types measured in normalized count values are displayed in these plots. Except for tau class glutathione transferases, these genes appear to be integral to salt response in shoot tissues. In addition, many of these genes do not display differential expression between control and treated tissues, potentially suggesting relatively high constitutive expression values of these important salinity responsive genes in *S. tragus*.


A. Both treatments and tissue types 24 HAT
response to cold
response to water deprivation
response to abscisic acid
regulation of transcription, DNA−templated
cellular response to cytokinin stimulus
response to salt stress
1.5
2.0
2.5
3.0
3.5
−log(p−value)
Count
2
3
4
5
6
7
8
Fold_Enrichment
90
60
30

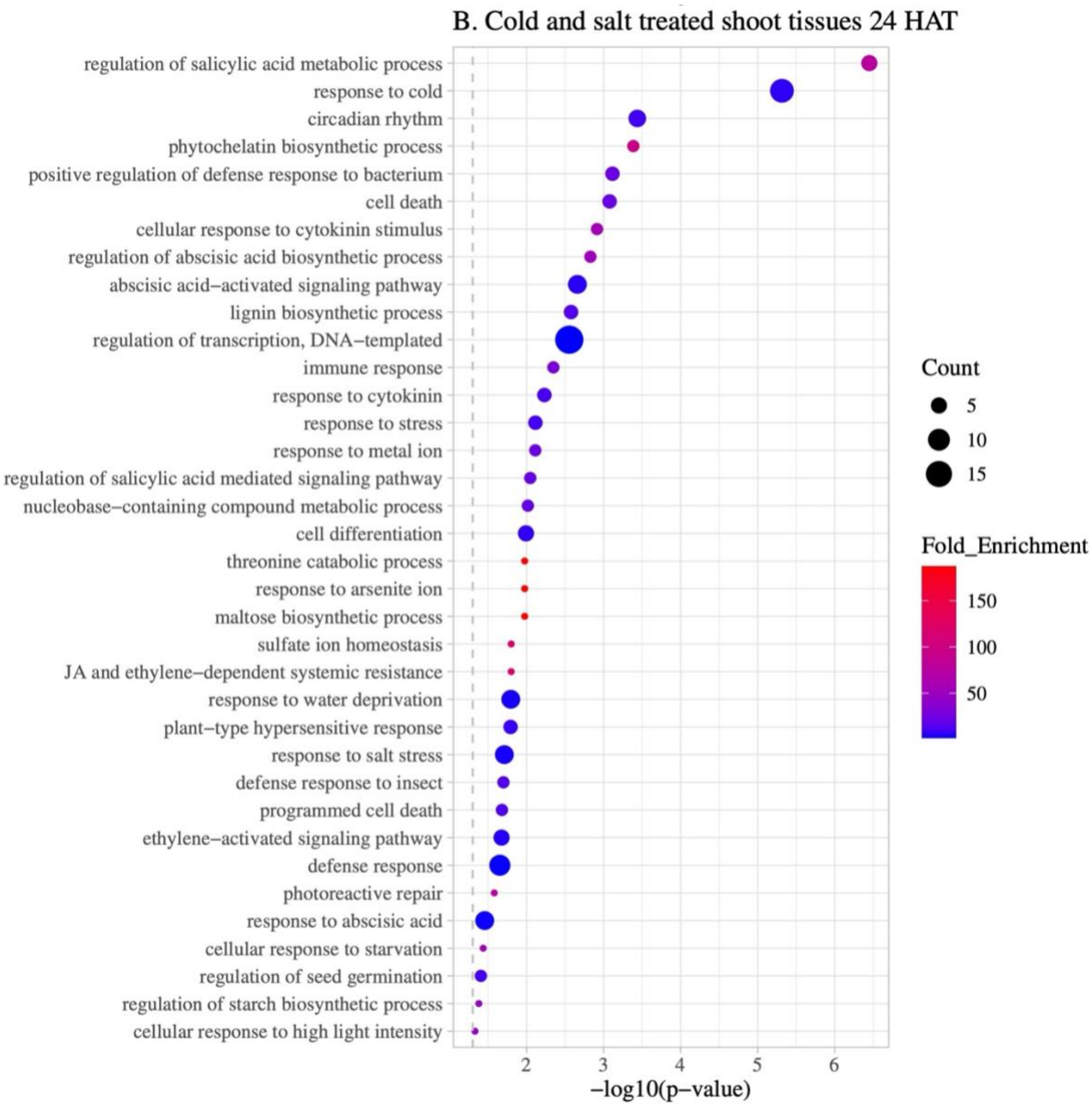


Figure 3.11: Bubble plots of GO enrichment analyses for the upregulated transcripts that were assigned UniProt accession numbers. Fold enrichment and p-values were generated with Arabidopsis as a background. Terms representing biological processes of differentially expressed transcripts are displayed along the y-axis. These terms are organized from the smallest to largest p-value, and the 'Count' metric indicates the number of UniProt accession numbers that belong to each biological process. All processes that presented an FDR p-value < 0.05 were included in these plots. Figure A displays results from the DESeq contrast that included both shoot and root tissues and both salt and cold treated samples at 24 hours after treatment. This figure represents significant upregulated biological processes for both types of stress throughout the plant. Figure B displays results from the DESeq analysis including only shoot tissues 24-hours after both cold and salt treatments. As such, Figure B is representative of the vast amount of information that is masked when including both tissue types in differential expression analyses.

REFERENCES

Abuqamar, S., Ajeb, S., Sham, A., Enan, M. R., & Iratni, R. (2013). A mutation in the *expansin-like A 2* gene enhances resistance to necrotrophic fungi and hypersensitivity to abiotic stress in *A rabidopsis thaliana* . *Molecular Plant Pathology*, *14*(8), 813–827. https://doi.org/10.1111/mpp.12049

Alonge, M., Lebeigle, L., Kirsche, M., Jenike, K., Ou, S., Aganezov, S., Wang, X., Lippman, Z. B., Schatz, M. C., & Soyk, S. (2022). Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biology*, *23*(1), 258. https://doi.org/10.1186/s13059-022-02823-7

Badillo-Vargas, I. E., Chen, Y., Martin, K. M., Rotenberg, D., & Whitfield, A. E. (2019). Discovery of Novel Thrips Vector Proteins That Bind to the Viral Attachment Protein of the Plant Bunyavirus Tomato Spotted Wilt Virus. *Journal of Virology*, *93*(21).

Baduel, P., Bray, S., Vallejo-Marin, M., Kolář, F., & Yant, L. (2018). The "Polyploid Hop": Shifting Challenges and Opportunities Over the Evolutionary Lifespan of Genome Duplications. *Frontiers in Ecology and Evolution*, *6*, 117. https://doi.org/10.3389/fevo.2018.00117

Beckie, H. J., & Francis, A. (2009). The biology of Canadian weeds. 65. *Salsola tragus* L. (updated). *Canadian Journal of Plant Science*, *89*(4), 775–789. https://doi.org/10.4141/CJPS08181

Beckie, H. J., Hall, L. M., Shirriff, S. W., Martin, E., & Leeson, J. Y. (2019). Glyphosate and acetolactate synthase inhibitor resistance in Russian thistle (*Salsola tragus* L.) in Alberta. *Canadian Journal of Plant Science*, *99*(3), 384–387. https://doi.org/10.1139/cjps-2018-0252

Brignone, N. F., & Denham, S. S. (2021). Toward an Updated Taxonomy of the South American Chenopodiaceae I: Subfamilies Betoideae, Camphorosmoideae, and Salsoloideae. *Annals of the Missouri Botanical Garden*, *106*, 10–30. https://doi.org/10.3417/2020615

Calvin, K., Dasgupta, D., Krinner, G., Mukherji, A., Thorne, P. W., Trisos, C., Romero, J., Aldunce, P., Barrett, K., Blanco, G., Cheung, W. W. L., Connors, S., Denton, F., Diongue-Niang, A., Dodman, D., Garschagen, M., Geden, O., Hayward, B., Jones, C., … Péan, C. (2023). *IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland.* (First). Intergovernmental Panel on Climate Change (IPCC). https://doi.org/10.59327/IPCC/AR6-9789291691647

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Chinnusamy, V., Zhu, J.-K., & Sunkar, R. (2010). Gene Regulation During Cold Stress Acclimation in Plants. In R. Sunkar (Ed.), *Plant Stress Tolerance* (Vol. 639, pp. 39–55). Humana Press. https://doi.org/10.1007/978-1-60761-702-0_3

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M. W., Gaffney, D. J., Elo, L. L., Zhang, X., & Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biology*, *17*(1), 13. https://doi.org/10.1186/s13059-016-0881-8

Crompton, C. W., & Bassett, I. J. (1985). THE BIOLOGY OF CANADIAN WEEDS.: 65. *Salsola pestifer* A. Nels. *Canadian Journal of Plant Science*, *65*(2), 379–388. https://doi.org/10.4141/cjps85-053

Dagar, J. C., Sharma, P. C., Chaudhari, S. K., Jat, H. S., & Ahamad, S. (2016). Climate Change vis-a-vis Saline Agriculture: Impact and Adaptation Strategies. In J. C. Dagar, P. C. Sharma, D. K. Sharma, & A. K. Singh (Eds.), *Innovative Saline Agriculture* (pp. 5–53). Springer India. https://doi.org/10.1007/978-81-322-2770-0_2

Dao, T. T. H., Linthorst, H. J. M., & Verpoorte, R. (2011). Chalcone synthase and its functions in plant resistance. *Phytochemistry Reviews*, *10*(3), 397–412. https://doi.org/10.1007/s11101-011-9211-7

Digby, L. (1912). The Cytology of Primula kewensis and of other related Primula Hybrids. *Annals of Botany*, *26*(102), 357–388.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Earl, D., Bradnam, K., St. John, J., Darling, A., Lin, D., Fass, J., Yu, H. O. K., Buffalo, V., Zerbino, D. R., Diekhans, M., Nguyen, N., Ariyaratne, P. N., Sung, W.-K., Ning, Z., Haimel, M., Simpson, J. T., Fonseca, N. A., Birol, İ., Docking, T. R., … Paten, B. (2011). Assemblathon 1: A competitive assessment of de novo short read assembly methods. *Genome Research*, *21*(12), 2224–2241. https://doi.org/10.1101/gr.126599.111

Estoup, A., Ravigné, V., Hufbauer, R., Vitalis, R., Gautier, M., & Facon, B. (2016). Is There a Genetic Paradox of Biological Invasion? *Annual Review of Ecology, Evolution, and Systematics*, *47*(1), 51–72. https://doi.org/10.1146/annurev-ecolsys-121415-032116

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047–3048. https://doi.org/10.1093/bioinformatics/btw354

Gaines, T. A., Duke, S. O., Morran, S., Rigon, C. A. G., Tranel, P. J., Küpper, A., & Dayan, F. E. (2020). Mechanisms of evolved herbicide resistance. *Journal of Biological Chemistry*, *295*(30), 10307–10330. https://doi.org/10.1074/jbc.REV120.013572

Gaines, T. A., Slavov, G. T., Hughes, D., Küpper, A., Sparks, C. D., Oliva, J., Vila-Aiub, M. M., Garcia, M. A., Merotto, A., & Neve, P. (2021). Investigating the origins and evolution of a glyphosate-resistant weed invasion in South America. *Molecular Ecology*, *30*(21), 5360–5372. https://doi.org/10.1111/mec.16221

Gaines, T. A., Zhang, W., Wang, D., Bukun, B., Chisholm, S. T., Shaner, D. L., Nissen, S. J., Patzoldt, W. L., Tranel, P. J., Culpepper, A. S., Grey, T. L., Webster, T. M., Vencill, W. K., Sammons, R. D., Jiang, J., Preston, C., Leach, J. E., & Westra, P. (2010). Gene amplification confers glyphosate resistance in *Amaranthus palmeri*. *Proceedings of the National Academy of Sciences*, *107*(3), 1029–1034. https://doi.org/10.1073/pnas.0906649107

Gioria, M., Hulme, P. E., Richardson, D. M., & Pyšek, P. (2023). Why Are Invasive Plants Successful? *Annual Review of Plant Biology*, *74*(1), 635–670. https://doi.org/10.1146/annurev-arplant-070522-071021

Guan, C., Wang, C., Li, Q., Ji, J., Wang, G., Jin, C., & Tong, Y. (2019). LcSABP2, a salicylic acid binding protein 2 gene from Lycium chinense, confers resistance to triclosan stress in Nicotiana tabacum. *Ecotoxicology and Environmental Safety*, *183*, 109516. https://doi.org/10.1016/j.ecoenv.2019.109516

Gültekin, N., Gürbüz, R., & Gözüaçik, C. (2021). Biological Observations on the Salsola tragus L. Weed Feeder Conorhynchus kindermanni Faust, 1904 (Coleoptera: Curculionidae) in Ağrı Mountain Lowland Area. *Transactions of the American Entomological Society*, *147*(4). https://doi.org/10.3157/061.147.0407

Guo, Y., Zhang, H., Yuan, Y., Cui, X., & Zhang, L. (2020). Identification and characterization of *NAC* genes in response to abiotic stress conditions in *Picea wilsonii* using transcriptome sequencing. *Biotechnology & Biotechnological Equipment*, *34*(1), 93–103. https://doi.org/10.1080/13102818.2020.1718550

Han, G., Qiao, Z., Li, Y., Yang, Z., Wang, C., Zhang, Y., Liu, L., & Wang, B. (2022). RING Zinc Finger Proteins in Plant Abiotic Stress Tolerance. *Frontiers in Plant Science*, *13*, 877011. https://doi.org/10.3389/fpls.2022.877011

Hao, Z., Lv, D., Ge, Y., Shi, J., Weijers, D., Yu, G., & Chen, J. (2020). *RIdeogram*: Drawing SVG graphics to visualize and map genome-wide data on the idiograms. *PeerJ Computer Science*, *6*, e251. https://doi.org/10.7717/peerj-cs.251

Hu, Y., Jiang, Y., Han, X., Wang, H., Pan, J., & Yu, D. (2017). Jasmonate regulates leaf senescence and tolerance to cold stress: Crosstalk with other phytohormones. *Journal of Experimental Botany*, *68*(6), 1361–1369. https://doi.org/10.1093/jxb/erx004

Huang, X., Hou, L., Meng, J., You, H., Li, Z., Gong, Z., Yang, S., & Shi, Y. (2018). The Antagonistic Action of Abscisic Acid and Cytokinin Signaling Mediates Drought Stress Response in Arabidopsis. *Molecular Plant*, *11*(7), 970–982. https://doi.org/10.1016/j.molp.2018.05.001

Iamonico, D., & Mosyakin, S. L. (2018). STUDIES ON CHENOPODIUM ALBUM S. L. (CHENOPODIACEAE / AMARANTHACEAE S. L.): CHENOPODIUM

PEDUNCULARE. *Annali Di Botanica*, *8*(Vol 8 (2018)). https://doi.org/10.4462/annbotrm-14240

Jackson, R. C. (1976). Evolution and Systematic Significance of Polyploidy. *Annual Review of Ecology and Systematics*, *7*(1), 209–234. https://doi.org/10.1146/annurev.es.07.110176.001233

Jeschke, J. M., & Strayer, D. L. (2008). Usefulness of Bioclimatic Models for Studying Climate Change and Invasive Species. *Annals of the New York Academy of Sciences*, *1134*(1), 1–24. https://doi.org/10.1196/annals.1439.002

Jia, K., Wang, Z., Wang, L., Li, G., Zhang, W., Wang, X., Xu, F., Jiao, S., Zhou, S., Liu, H., Ma, Y., Bi, G., Zhao, W., El-Kassaby, Y. A., Porth, I., Li, G., Zhang, R., & Mao, J. (2022). SUBPHASER: A robust allopolyploid subgenome phasing method based on subgenome-specific *k* -mers. *New Phytologist*, *235*(2), 801–809. https://doi.org/10.1111/nph.18173

Jones, M. A., Morohashi, K., Grotewold, E., & Harmer, S. L. (2019). Arabidopsis JMJD5/JMJ30 Acts Independently of LUX ARRHYTHMO Within the Plant Circadian Clock to Enable Temperature Compensation. *Frontiers in Plant Science*, *10*, 57. https://doi.org/10.3389/fpls.2019.00057

Kadereit, G., & Freitag, H. (2011). Molecular phylogeny of Camphorosmeae (Camphorosmoideae, Chenopodiaceae): Implications for biogeography, evolution of C_4 -photosynthesis and taxonomy. *TAXON*, *60*(1), 51–78. https://doi.org/10.1002/tax.601006

Kreps, J. A., Wu, Y., Chang, H.-S., Zhu, T., Wang, X., & Harper, J. F. (2002). Transcriptome Changes for Arabidopsis in Response to Salt, Osmotic, and Cold Stress,. *Plant Physiology*, *130*(4), 2129–2141. https://doi.org/10.1104/pp.008532

Kumar, V., Spring, J. F., Jha, P., Lyon, D. J., & Burke, I. C. (2017). Glyphosate-Resistant Russian-thistle (*Salsola tragus*) Identified in Montana and Washington. *Weed Technology*, *31*(2), 238–251. https://doi.org/10.1017/wet.2016.32

Kuroda, H., Yanagawa, Y., Takahashi, N., Horii, Y., & Matsui, M. (2012). A Comprehensive Analysis of Interaction and Localization of Arabidopsis SKP1-LIKE (ASK) and F-Box (FBX) Proteins. *PLoS ONE*, *7*(11), e50009. https://doi.org/10.1371/journal.pone.0050009

Leitch, I. J., & Bennett, M. D. (1997). Polyploidy in angiosperms. *Trends in Plant Science*, *2*(12), 470–476. https://doi.org/10.1016/S1360-1385(97)01154-0

Levin, D. A. (1983). Polyploidy and Novelty in Flowering Plants. *The American Naturalist*, *122*(1), 1–25. https://doi.org/10.1086/284115

Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*, *34*(18), 3094–3100. https://doi.org/10.1093/bioinformatics/bty191

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, W., Wang, D., Jin, T., Chang, Q., Yin, D., Xu, S., Liu, B., & Liu, L. (2011). The Vacuolar Na+/H+ Antiporter Gene SsNHX1 from the Halophyte Salsola soda Confers Salt Tolerance in Transgenic Alfalfa (Medicago sativa L.). *Plant Molecular Biology Reporter*, *29*(2), 278–290. https://doi.org/10.1007/s11105-010-0224-y

Liu, N., Hou, L., Chen, X., Bao, J., Chen, F., Cai, W., Zhu, H., Wang, L., & Chen, X. (2024). *Arabidopsis TETRASPANIN8 mediates exosome secretion and glycosyl inositol phosphoceramide sorting and trafficking*. *36*, 626–641.

Lone, P. A., Dar, J. A., Subashree, K., Raha, D., Pandey, P. K., Ray, T., Khare, P. K., & Khan, M. L. (2019). Impact of plant invasion on physical, chemical and biological aspects of ecosystems: A review. *Tropical Plant Research*, *6*(3), 528–544. https://doi.org/10.22271/tpr.2019.v6.i3.067

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Mack, R. N., Simberloff, D., Lonsdale, W. M., Evans, H., Clout, M., & Bazzaz, F. A. (2000). BIOTIC INVASIONS: CAUSES, EPIDEMIOLOGY, GLOBAL CONSEQUENCES, AND CONTROL. *Ecological Applications*, *10*(3).

Manni, M., Berkeley, M. R., Seppey, M., & Zdobnov, E. M. (2021). BUSCO: Assessing Genomic Data Quality and Beyond. *Current Protocols*, *1*(12). https://doi.org/10.1002/cpz1.323

Mansour, M. M. F., & Hassan, F. A. S. (2022). How salt stress-responsive proteins regulate plant adaptation to saline conditions. *Plant Molecular Biology*, *108*(3), 175–224. https://doi.org/10.1007/s11103-021-01232-x

Marçais, G., Delcher, A. L., Phillippy, A. M., Coston, R., Salzberg, S. L., & Zimin, A. (2018). MUMmer4: A fast and versatile genome alignment system. *PLOS Computational Biology*, *14*(1), e1005944. https://doi.org/10.1371/journal.pcbi.1005944

McIntyre, P. J., & Strauss, S. (2017). An experimental test of local adaptation among cytotypes within a polyploid complex: ADAPTATION AND NICHE BREADTH IN POLYPLOIDS. *Evolution*, *71*(8), 1960–1969. https://doi.org/10.1111/evo.13288

Medeiros, R. B., Resende, R. de O., & de Ávila, A. C. (2004). The Plant Virus *Tomato Spotted Wilt Tospovirus* Activates the Immune System of Its Main Insect Vector, *Frankliniella occidentalis*. *Journal of Virology*, *78*(10), 4976–4982. https://doi.org/10.1128/JVI.78.10.4976-4982.2004

Mishra, A., & Tanna, B. (2017). Halophytes: Potential Resources for Salt Stress Tolerance Genes and Promoters. *Frontiers in Plant Science*, *8*, 829. https://doi.org/10.3389/fpls.2017.00829

Mishra, D., Shekhar, S., Chakraborty, S., & Chakraborty, N. (2018). Carboxylate clamp tetratricopeptide repeat (TPR) domain containing Hsp90 cochaperones in Triticeace: An insight into structural and functional diversification. *Environmental and Experimental Botany*, *155*, 31–44. https://doi.org/10.1016/j.envexpbot.2018.06.020

Molin, W. T., Yaguchi, A., Blenner, M., & Saski, C. A. (2020). The EccDNA Replicon: A Heritable, Extranuclear Vehicle That Enables Gene Amplification and Glyphosate Resistance in Amaranthus palmeri[OPEN]. *The Plant Cell*, *32*(7), 2132–2140. https://doi.org/10.1105/tpc.20.00099

Montgomery, J. S., Morran, S., MacGregor, D. R., McElroy, J. S., Neve, P., Neto, C., Vila-Aiub, M. M., Sandoval, M. V., Menéndez, A. I., Kreiner, J. M., Fan, L., Caicedo, A. L., Maughan, P. J., Martins, B. A. B., Mika, J., Collavo, A., Merotto, A., Subramanian, N. K., Bagavathiannan, M. V., … Gaines, T. (2023). *The International Weed Genomics Consortium: Community Resources for Weed Genomics Research*. https://doi.org/10.1101/2023.07.19.549613

Morrisson, I. N., & Devine, M. D. (2005). Herbicide resistance in the Canadian prairie provinces: Five years after the fact. *Phytoprotection*, *75*(4), 5–16. https://doi.org/10.7202/706067ar

Mosyakin, S. L. (1996). A Taxonomic Synopsis of the Genus Salsola (Chenopodiaceae) in North America. *Annals of the Missouri Botanical Garden*, *83*(3), 387. https://doi.org/10.2307/2399867

Munns, R. (2005). Genes and salt tolerance: Bringing them together. *New Phytologist*, *167*(3), 645–663. https://doi.org/10.1111/j.1469-8137.2005.01487.x

Nemchenko, A., Kunze, S., Feussner, I., & Kolomiets, M. (2006). Duplicate maize 13-lipoxygenase genes are differentially regulated by circadian rhythm, cold stress, wounding, pathogen infection, and hormonal treatments. *Journal of Experimental Botany*, *57*(14), 3767–3779. https://doi.org/10.1093/jxb/erl137

Pang, Y., Cao, L., Ye, F., Ma, C., Liang, X., Song, Y., & Lu, X. (2024). Identification of the Maize PP2C Gene Family and Functional Studies on the Role of ZmPP2C15 in Drought Tolerance. *Plants*, *13*(3), 340. https://doi.org/10.3390/plants13030340

Qiu, T., Liu, Z., & Liu, B. (2020). The effects of hybridization and genome doubling in plant evolution via allopolyploidy. *Molecular Biology Reports*, *47*(7), 5549–5558. https://doi.org/10.1007/s11033-020-05597-y

Ramsey, J., & Schemske, D. W. (1998). PATHWAYS, MECHANISMS, AND RATES OF POLYPLOID FORMATION IN FLOWERING PLANTS. *Annual Review of Ecology and Systematics*, *29*(1), 467–501. https://doi.org/10.1146/annurev.ecolsys.29.1.467

Rančić, D., Pećinar, I., Aćić, S., & Stevanović, Z. D. (2019). Morpho-anatomical traits of halophytic species. In M. Hasanuzzaman, S. Shabala, & M. Fujita (Eds.), *Halophytes and*

climate change: Adaptive mechanisms and potential uses (1st ed., pp. 152–178). CABI. https://doi.org/10.1079/9781786394330.0152

Rao, D. R., Hovanes, K., Smith, R., Davy, J., & Gornish, E. S. (2022). Russian thistle (*Salsola* spp.) control in California rangelands over five years. *Invasive Plant Science and Management*, *15*(1), 33–40. https://doi.org/10.1017/inp.2022.9

Ravet, K., Patterson, E. L., Krähmer, H., Hamouzová, K., Fan, L., Jasieniuk, M., Lawton-Rauh, A., Malone, J. M., McElroy, J. S., Merotto, A., Westra, P., Preston, C., Vila-Aiub, M. M., Busi, R., Tranel, P. J., Reinhardt, C., Saski, C., Beffa, R., Neve, P., & Gaines, T. A. (2018). The power and potential of genomics in weed biology and management. *Pest Management Science*, *74*(10), 2216–2225. https://doi.org/10.1002/ps.5048

Reitz, S. R., Gao, Y., Kirk, W. D. J., Hoddle, M. S., Leiss, K. A., & Funderburk, J. E. (2020). Invasion Biology, Ecology, and Management of Western Flower Thrips. *Annual Review of Entomology*, *65*(1), 17–37. https://doi.org/10.1146/annurev-ento-011019-024947

Rosche, C., Hensen, I., Mráz, P., Durka, W., Hartmann, M., & Lachmuth, S. (2017). Invasion success in polyploids: The role of inbreeding in the contrasting colonization abilities of diploid versus tetraploid populations of *Centaurea stoebe* s.l. *Journal of Ecology*, *105*(2), 425–435. https://doi.org/10.1111/1365-2745.12670

Ryan, F. J., & Ayres, D. R. (2000). *Molecular markers indicate two cryptic, genetically divergent populations of Russian thistle (Salsola tragus) in California*. *78*, 9.

Saari, L. L., Cotterman, J. C., Smith, W. F., & Primiani, M. M. (1992). Sulfonylurea herbicide resistance in common chickweed, perennial ryegrass, and Russian thistle. *Pesticide Biochemistry and Physiology*, *42*(2), 110–118. https://doi.org/10.1016/0048-3575(92)90058-8

Sammons, R. D., & Gaines, T. A. (2014). Glyphosate resistance: State of knowledge. *Pest Management Science*, *70*(9), 1367–1377. https://doi.org/10.1002/ps.3743

Sherman, B. T., Hao, M., Qiu, J., Jiao, X., Baseler, M. W., Lane, H. C., Imamichi, T., & Chang, W. (2022). DAVID: A web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Research*, *50*(W1), W216–W221. https://doi.org/10.1093/nar/gkac194

Simberloff, D. (2005). Non-native Species DO Threaten the Natural Environment! *Journal of Agricultural and Environmental Ethics*, *18*(6), 595–607. https://doi.org/10.1007/s10806-005-2851-0

Smith, L. (2005). Host plant specificity and potential impact of Aceria salsolae (Acari: Eriophyidae), an agent proposed for biological control of Russian thistle (Salsola tragus). *Biological Control*, *34*(1), 83–92. https://doi.org/10.1016/j.biocontrol.2005.03.003

Sobhian, R., Fumanal, B., & Pitcairn, M. (2003). Observations on the host specificity and biology of Lixus salsolae (Col., Curculionidae), a potential biological control agent of Russian thistle, Salsola tragus (Chenopodiaceae) in North America. *Journal of Applied Entomology*, *127*(6), 322–324. https://doi.org/10.1046/j.1439-0418.2003.00761.x

Soltis, D. E., & Soltis, P. S. (1999). Polyploidy: Recurrent formation and genome evolution. *Trends in Ecology & Evolution*, *14*(9), 348–352. https://doi.org/10.1016/S0169-5347(99)01638-9

Soltis, D. E., Soltis, P. S., & Rieseberg, L. H. (1993). Molecular Data and the Dynamic Nature of Polyploidy. *Critical Reviews in Plant Sciences*, *12*(3), 243–273. https://doi.org/10.1080/07352689309701903

Song, W., Shao, H., Zheng, A., Zhao, L., & Xu, Y. (2023). *Advances in Roles of Salicylic Acid in Plant Tolerance Responses to Biotic and Abiotic Stresses*.

Spring, J. F., Revolinski, S. R., Young, F. L., Lyon, D. J., & Burke, I. C. (2022). Weak population differentiation and high diversity in *Salsola tragus* in the inland Pacific Northwest, USA. *Pest Management Science*, *78*(11), 4728–4740. https://doi.org/10.1002/ps.7093

Stafford, C. A., Walker, G. P., & Ullman, D. E. (2012). Hitching a ride: Vector feeding and virus transmission. *Communicative & Integrative Biology*, *5*(1), 43–49. https://doi.org/10.4161/cib.18640

Stallings, G. P., Thill, D. C., & Mallory-Smith, C. A. (1994). Sulfonylurea-Resistant Russian Thistle (*Salsola iberica*) Survey in Washington State. *Weed Technology*, *8*(2), 258–264. https://doi.org/10.1017/S0890037X00038744

Stark, R., Grzelak, M., & Hadfield, J. (2019). RNA sequencing: The teenage years. *Nature Reviews Genetics*, *20*(11), 631–656. https://doi.org/10.1038/s41576-019-0150-2

Stephens, S. G. (1951). *Possible Significance of Duplication in Evolution* (M. Demerec, Ed.; Vol. 4, pp. 247–265). Academic Press. https://doi.org/10.1016/S0065-2660(08)60237-0

Stift, M., Berenos, C., Kuperus, P., & Van Tienderen, P. H. (2008). Segregation Models for Disomic, Tetrasomic and Intermediate Inheritance in Tetraploids: A General Procedure Applied to Rorippa (Yellow Cress) Microsatellite Data. *Genetics*, *179*(4), 2113–2123. https://doi.org/10.1534/genetics.107.085027

Su, W., Ou, S., Hufford, M. B., & Peterson, T. (2021). A Tutorial of EDTA: Extensive De Novo TE Annotator. In J. Cho (Ed.), *Plant Transposable Elements* (Vol. 2250, pp. 55–67). Springer US. https://doi.org/10.1007/978-1-0716-1134-0_4

Sultan, B. (2012). Global warming threatens agricultural productivity in Africa and South Asia. *Environmental Research Letters*, *7*(4), 041001. https://doi.org/10.1088/1748-9326/7/4/041001

Talbert, P. B., & Henikoff, S. (2020). What makes a centromere? *Experimental Cell Research*, *389*(2), 111895. https://doi.org/10.1016/j.yexcr.2020.111895

Te Beest, M., Le Roux, J. J., Richardson, D. M., Brysting, A. K., Suda, J., Kubesova, M., & Pysek, P. (2012). The more the better? The role of polyploidy in facilitating plant invasions. *Annals of Botany*, *109*(1), 19–45. https://doi.org/10.1093/aob/mcr277

Tenhaken, R. (2015). Cell wall remodeling under abiotic stress. *Frontiers in Plant Science*, *5*. https://doi.org/10.3389/fpls.2014.00771

The UniProt Consortium, Bateman, A., Martin, M.-J., Orchard, S., Magrane, M., Ahmad, S., Alpi, E., Bowler-Barnett, E. H., Britto, R., Bye-A-Jee, H., Cukura, A., Denny, P., Dogan, T., Ebenezer, T., Fan, J., Garmiri, P., Da Costa Gonzales, L. J., Hatton-Ellis, E., Hussein, A., … Zhang, J. (2023). UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Research*, *51*(D1), D523–D531. https://doi.org/10.1093/nar/gkac1052

Todd, E. V., Black, M. A., & Gemmell, N. J. (2016). The power and promise of RNA-seq in ecology and evolution. *Molecular Ecology*, *25*(6), 1224–1241. https://doi.org/10.1111/mec.13526

Treier, U. A., Broennimann, O., Normand, S., Guisan, A., Schaffner, U., Steinger, T., & Müller-Schärer, H. (2009). Shift in cytotype frequency and niche space in the invasive plant *Centaurea maculosa*. *Ecology*, *90*(5), 1366–1377. https://doi.org/10.1890/08-0420.1

Tzin, V., & Galili, G. (2010). New Insights into the Shikimate and Aromatic Amino Acids Biosynthesis Pathways in Plants. *Molecular Plant*, *3*(6), 956–972. https://doi.org/10.1093/mp/ssq048

Walsh, P., Bursać, D., Law, Y. C., Cyr, D., & Lithgow, T. (2004). The J-protein family: Modulating protein assembly, disassembly and translocation. *EMBO Reports*, *5*(6), 567–571. https://doi.org/10.1038/sj.embor.7400172

Wang, J., Song, L., Gong, X., Xu, J., & Li, M. (2020). Functions of Jasmonic Acid in Plant Regulation and Response to Abiotic Stress. *International Journal of Molecular Sciences*, *21*(4), 1446. https://doi.org/10.3390/ijms21041446

Wang, Y., Zhu, X., Jin, Y., Duan, R., Gu, Y., Liu, X., Qian, L., & Chen, F. (2023). Selection Behavior and OBP-Transcription Response of Western Flower Thrips, Frankliniella occidentalis, to Six Plant VOCs from Kidney Beans. *International Journal of Molecular Sciences*, *24*(16), 12789. https://doi.org/10.3390/ijms241612789

Warwick, S. I., Sauder, C. A., & Beckie, H. J. (2010). Acetolactate Synthase (*ALS*) Target-Site Mutations in ALS Inhibitor-Resistant Russian Thistle (*Salsola tragus*). *Weed Science*, *58*(3), 244–251. https://doi.org/10.1614/WS-D-09-00083.1

Wernersson, R. (2006). Virtual Ribosome—A comprehensive DNA translation tool with support for integration of sequence feature annotation. *Nucleic Acids Research*, *34*(Web Server), W385–W388. https://doi.org/10.1093/nar/gkl252

Westwood, J. H., Charudattan, R., Duke, S. O., Fennimore, S. A., Marrone, P., Slaughter, D. C., Swanton, C., & Zollinger, R. (2018). Weed Management in 2050: Perspectives on the Future of Weed Science. *Weed Science*, *66*(3), 275–285. https://doi.org/10.1017/wsc.2017.78

Xu, Y.-H., Liu, R., Yan, L., Liu, Z.-Q., Jiang, S.-C., Shen, Y.-Y., Wang, X.-F., & Zhang, D.-P. (2012). Light-harvesting chlorophyll a/b-binding proteins are required for stomatal response to abscisic acid in Arabidopsis. *Journal of Experimental Botany*, *63*(3), 1095–1106. https://doi.org/10.1093/jxb/err315

Yamaguchi-Shinozaki, K., & Shinozaki, K. (2006). TRANSCRIPTIONAL REGULATORY NETWORKS IN CELLULAR RESPONSES AND TOLERANCE TO DEHYDRATION AND COLD STRESSES. *Annual Review of Plant Biology*, *57*(1), 781–803. https://doi.org/10.1146/annurev.arplant.57.032905.105444

Yanniccari, M., Palma-Bautista, C., Vázquez-García, J. G., Gigón, R., Mallory-Smith, C. A., & De Prado, R. (2023). Constitutive overexpression of *EPSPS* by gene duplication is involved in glyphosate resistance in *SALSOLA TRAGUS* . *Pest Management Science*, *79*(3), 1062–1068. https://doi.org/10.1002/ps.7272

Yao, Y., Zhang, X., Wang, N., Cui, Y., Zhang, L., & Fan, S. (2020). Transcriptome analysis of salt stress response in halophyte Atriplex centralasiatica leaves. *Acta Physiologiae Plantarum*, *42*(1), 3. https://doi.org/10.1007/s11738-019-2989-4

Yoon, Y., Seo, D. H., Shin, H., Kim, H. J., Kim, C. M., & Jang, G. (2020). The Role of Stress-Responsive Transcription Factors in Modulating Abiotic Stress Tolerance in Plants. *Agronomy*, *10*(6), 788. https://doi.org/10.3390/agronomy10060788

Zhang, K., Duan, M., Zhang, L., Li, J., Shan, L., Zheng, L., & Liu, J. (2022). HOP1 and HOP2 are involved in salt tolerance by facilitating the brassinosteroid-related nucleo-cytoplasmic partitioning of the HSP90-BIN2 complex. *Plant, Cell & Environment*, *45*(12), 3551–3565. https://doi.org/10.1111/pce.14441

APPENDIX A: INVESTIGATING *FRANKLINIELLA OCCIDENTALIS* AS A POTENTIAL VECTOR OF GLYPHOSATE RESISTANCE GENES IN *AMARANTHUS PALMERI*[2]

INTRODUCTION

The spread of glyphosate resistant (GR) *Amaranthus palmeri* throughout the Midwest severely impacts agricultural production systems that rely on GR crops, such as Roundup Ready© *Beta vulgaris*. Producers are forced to either preform multiple applications of glyphosate throughout the growing season in the hopes that sequential applications will result in weed control or adopt more labor-intensive integrated pest management (IPM) strategies. *B. vulgaris* production throughout Colorado has been drastically impacted by *A. palmeri* demonstrating extremely high levels of glyphosate resistance. As a diecious species, the heritable spread of resistance alone may not explain the current spread of high-level glyphosate resistance in *A. palmeri* populations throughout the Midwest. This research attempts to determine an uncharacterized alternative mechanism for the spread of glyphosate resistance in *A. palmeri*.

The target site of glyphosate lies in the shikimate pathway and catalyzes an essential step in aromatic amino acid biosynthesis. The enzyme targeted by glyphosate in this pathway is *5-enolpyruvvylshikimate-3-phosphate synthase (EPSPS).* In contrast to tandem *EPSPS* duplication within the genome of *Bassia scoparia*, the mechanism for glyphosate resistance that has evolved in *A. palmeri* is found in extra-chromosomal circular DNA. Termed eccDNA, these large structures contain genes of various function in just under 400k base pairs (Molin et al., 2020). The overall biological understanding of how eccDNA structures form, how genes are incorporated into them,

[2]The following authors have contributed to this section: John Lemas, Jacob Pitt, Olivia Carter, Crystal Sparks, Punya Nachappa, Todd Gaines.

and how they are passed to progeny remains unclear. As opposed to the somewhat restricted gene copy number through tandem duplication within the genome, the presence of eccDNA in *A. palmeri* allows the species to attain a much higher level of glyphosate resistance. The identification and elucidation of these DNA structures has largely been a result of GR *A. palmeri* research, by which eccDNA of varying gene content is understood to be found throughout eukaryotic organisms (Gaines et al., 2020; Sammons & Gaines, 2014).

Thrips are common agricultural pests that can be difficult to identify at the species level because hundreds of species fall under this common name. Specific thrips species, such as *Frankliniella occidentalis* (Western Flower Thrips), serve as important vectors for most plant diseases. Their anatomy allows them to harbor, replicate, and transfer viruses between hosts. Specific tissues can carry out pathogen replication, and inoculation of an individual vector only occurs in the fist-instar larval stage of *F. occidentalis* development (Badillo-Vargas et al., 2019; Medeiros et al., 2004; Stafford et al., 2012). As vectoring viruses is common in thrips species, could *F. occidentalis* also be vectoring eccDNA?

During the fist-instar larval stage of *F. occidentalis*, host inflorescence may be targeted due to the increased concentrations of volatile organic compounds in these tissues. *F. occidentalis* has been shown to select hosts based on olfactory sensations sensing volatile compounds released from inflorescent tissues (Y. Wang et al., 2023). Aromatic amino acids are frequently the precursors of these secondary metabolites. Herein lies a potential a correlation between aromatic amino acid biosynthesis and the target site of glyphosate. Glyphosate targets the shikimate pathway, essentially starving the plant of aromatic amino acids. As such, overexpression of *EPSPS* via eccDNA in tissues actively concentrating aromatic amino acids may be preferred by this proposed herbicide-resistance vector (Tzin & Galili, 2010).

In this study, *F. occidentalis* (order *Thysanoptera,* family *Thripidae*) were studied as a possible vector to spread eccDNA containing the gene *EPSPS* in populations of *A. palmeri* in Northeastern Colorado. The hypothesis for this research was that *F. occidentalis* collected in the field from glyphosate resistant *A. palmeri* would contain detectible levels of glyphosate resistance genes. The first alternative hypothesis is that *F. occidentalis* collected from glyphosate resistant *A. palmeri* and allowed to feed will contain plant material in addition to measurable amounts glyphosate resistance genes. The second alternative hypothesis is that *F. occidentalis* collected from glyphosate resistant *A. palmeri* and allowed to starve will have digested plant material and will contain measurable levels of glyphosate resistance genes. The null hypothesis for this research is that *F. occidentalis* will not contain any measurable levels of glyphosate resistance genes.

METHODS

A.1: A. palmeri and F. occidentalis sample collection

Seven sugar beet fields in northeastern Colorado, sites A through G, were visited in July of 2022. The identification of potential GR *A. palmeri* plants in the field was assessed given their stature and overall robust vitality in fields of Roundup Ready© *B. vulgaris*. Five *A. palmeri* tissue samples were collected throughout each field and stored on ice in 15mL conical tubes. A total of thirty-five *A. palmeri* tissue samples were collected from seven collection sites and stored at -80°C. In addition to these samples, one presumed GR *A. palmeri* plant per field was transplanted into a 1-gal pot and transported to the Colorado State University's Entomology department greenhouse as a source for additional thrips colonies.

F. occidentalis samples were swept from each of the five *A. palmeri* biological replicates using a sweep net. *F. occidentalis* were then collected from the net using a respirator and were either placed immediately into 70% EtOH or contained in a clean sample jar. Those placed directly

into EtOH represent the 'not-starved' samples (NS), and those contained in a clean sample jar represent the 'starved' (S) samples. In total, sixty-seven *F. occidentalis* samples were collected from seven locations.

A.2: A. palmeri DNA extraction and PCR analysis

Sample analysis began with DNA extraction. *A. palmeri* tissue samples were frozen on liquid nitrogen and lysed using a Tissuelyzer II (Manufacture date 2012, Qiagen Inc., Hilden, Germany). DNA was then extracted using a DNeasy Plant Mini Kit (Cat. No. 69204, Qiagen Inc., Hilden, Germany) following manufacturer's instructions. All extractions were completed using a microcentrifuge (model # 5417C, Eppendorf, Inc.). Nucleic acid concentrations per sample were measured on a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific) after dilution with HPLC grade dH_2O to 5 ng/µL.

A thermocycler (T100™ Thermal Cycler, Bio Rad Laboratories Inc.) was used to amplify Acetolactate synthase (ALS) and eccDNA genes from *A. palmeri* DNA samples. A PCR gradient was performed to identify proper annealing temperature for each primer. A PCR mastermix was prepared including 12.5 µL Dreamtaq (Cat # FERK1082, Thermo Fisher Scientific), 1 µL forward primer (10µM), 1 µL reverse primer (10 µM) and 1 µL of *A. palmeri* template DNA per well. The thermocycler was programmed to initiate at 94℃ for 4 minutes, run through 30 cycles of 94℃ for 30 seconds, 55℃ for 30 seconds, and 72℃ for 1.5 minutes, and terminate with a final extension period of 72℃ for 5 minutes. In addition to the 35 *A. palmeri* samples, a non-template control (NTC), known glyphosate resistant controls from Arizona and Georgia as a positive control, and known glyphosate susceptible controls also from Arizona and Georgia were also included (Gaines et al., 2021). The NTC contained HPLC grade dH_2O.

Electrophoresis was preformed to identify the presence of both ALS and eccDNA amplicons in all 35 collected samples. PCR Primers were designed to amplify a 300 bp section of

both genes. A 1% agarose gel (Cat # GR1140-500, Hoefer Inc.) was prepared and run at 95 V for 1 hour (Owl™ A2 Large Gel system, Thermo Scientific; PowerPac™ HC, Bio Rad). Images were taken with an Azure 300 gel imager (Azure Biosystems Inc., Dublin, California) under default settings with an automatic exposure time.

Quantitative real-time PCR (RT q-PCR) conditions for *A. palmeri* EPSPS copy number variation followed those published by Gaines *et. al.* (Gaines et al., 2010, 2021). Primers for RT q-PCR included ALS as a single-gene reference and an EPSPS primer, and EPSPS copy number variations were computed for all samples including the Arizona susceptible and Arizona resistant controls (Table A1.1) (Gaines et al., 2010, 2021). EPSPS copy number variation (△Ct) was calculated based on the ALS amplification cycle number (ALS_{Cq}) and the EPSPS amplification cycle number ($EPSPS_{Cq}$) as described in Gaines *et. al.* (2010). Two technical replicates were included per biological replicate. △Ct values for each technical replicate were calculated and averaged into the reported △Ct value for each sample.

A.2: F. occidentalis DNA extraction and PCR analysis

F. occidentalis DNA was extracted using CTAB and nucleic acid concentration was measured on nanodrop. PCR was then preformed with the ALS, EPSPS, and eccDNA *A. palmeri* primers as well as a positive control primer, EF-1. Mastermix preparation included 12.5 µL Dreamtaq (Cat # FERK1082, Thermo Fisher Scientific), 1 µL forward primer (10 µM), 1 µL reverse primer (10 µM), and 1 µL template DNA. PCR conditions varied per primer based on gradient assay results. Thermocycler conditions for the EF-1 primer included an initial 95℃ for 2 minutes, 40 cycles of 95℃ for 30 seconds, 58℃ for 30 seconds, and 72℃ for 60 seconds ending in a final extension period of 72℃ for 7 minutes. Thermocycler conditions for the ALS and EPSPS

primers included an initial 95℃ for 4 minutes followed by 35 cycles of 95℃ for 30 seconds, 52℃ for 30 seconds, and 72℃ for 60 seconds ending in a final extension period of 72℃ for 7 minutes. Thermocycler conditions for the eccDNA primer included an initial 95℃ for 4 minutes followed by 35 cycles of 95℃ for 30 seconds, 62℃ for 30 seconds, and 72℃ for 60 seconds ending in a final extension period of 72℃ for 7 minutes. A 1% agarose gel (Cat # GR1140-500, Hoefer Inc.) was prepared for electrophoresis, and each *F. occidentalis* PCR product was run at 100 V for 45 minutes. Gel images were again taken with an Azure© Biosystems 300 gel imager under default settings with an automatic exposure time.

RESULTS

This experiment was designed to search for the presence of glyphosate resistance genes in *F. occidentalis* populations sampled on *A. palmeri* plants in Northeastern Colorado. Results include amplification for ALS and eccDNA primers for *A. palmeri* tissue samples, and amplification for EF-1, ALS, EPSPS, and eccDNA primers for *F. occidentalis* samples. Additional RT q-PCR results were obtained to confirm the glyphosate resistance mechanism.

From sites A through F, all 5 *A. palmeri* samples showed PCR amplification for both ALS and eccDNA primers. All five *A. palmeri* samples from site G showed PCR amplification for the ALS primer, however samples G2 and G4 did not show amplification for the eccDNA primer. RT q-PCR results show increased EPSPS copy number for sites A through F samples 1 through 5, and for samples G1, G3, and G5 (Table A1.1). Mean ALS primer cycle amplification was 11.6±0.5 cycles and mean EPSPS primer amplification was 6.7±2.4 cycles. Mean calculated gene duplication from qPCR results was 63.5±51.8 *EPSPS* gene copies between all 35 samples.

Overall *F. occidentalis* PCR product amplification resulted in 98.5% positive amplification for the EF-1 primer (Figure 2). Only one sample, A1NS, showed amplification for the ALS primer.

For the EPSPS primer, 15 samples showed amplification representing 22.7% of the total samples. Lastly, the eccDNA primer had 15.2% amplification with only 10 amplicons.

Considering the 33 not-starved samples, 30.3% amplified for EPSPS and 24.2% amplified for the eccDNA primers. In the 34 starved samples, 17.6% amplified for the EPSPS primer and 5.9% amplified for the eccDNA primer.

DISCUSSION

The objective of this research was to investigate *F. occidentalis* as a potential vector for the spread of extra-chromosomal circular DNA (eccDNA) containing duplications of the gene *EPSPS*. This eccDNA imparts high levels of glyphosate resistance to populations of *A. palmeri* in *Beta vulgaris* (sugar beet) fields throughout northern Colorado (Gaines et al., 2021; Molin et al., 2020). Multiple assays were employed to interrogate this proposed mechanism, including PCR, real-time q-PCR, and gel electrophoresis.

Glyphosate resistant *A. palmeri* were identified based on overall vitality. Samples were obtained from 15 to 40 yards into the sugar beet fields from the access point. All samples were collected in a 13-hour period and stored on ice. Over the course of sample collection, the quality of *A. palmeri* tissue samples from the earlier fields may have been impacted. *F. occidentalis* samples were either stored in ethanol or were contained in a cool environment, and therefore it is unlikely that the quality of these samples was impacted through the course of sample collection.

Polymerase chain-reaction (PCR) amplification followed by gel electrophoresis for the 35 collected *A. palmeri* tissue samples were successful in determining the presence of eccDNA. Samples 32 and 34 only amplified the ALS control primer, so it is likely that the negative amplification results for the eccDNA primer confirm no presence of eccDNA in these two samples. *A. palmeri* present in site G had recently been mechanically removed before sampling occurred.

Due to this we were restricted to locate GR *A. palmeri* samples either on the outskirts of the field or as remnants within the field. This introduced border effects or subsequent potential bias in the collection of samples from site G.

Negative results for samples G2 and G4 correlate with additional results from q-PCR and with *F. occidentalis* PCR results. The absence of *EPSPS* gene duplication from q-PCR confirms the absence of eccDNA in samples G2 and G4. When considering *F. occidentalis, EPSPS* amplification from sample G4, neither G4S nor G4NS showed amplification. *EPSPS* amplification for G2NS was also negative, however G2S did show slight *EPSPS* amplification compared to the NTC. This may be a false positive due to the nature of qualitative amplification, or it may have been due to the presence of *F. occidentalis* individuals that relocated to the outskirts of the field after the destruction of GR *A. palmeri* within the field.

Real-time q-PCR (RT q-PCR) nresults show increased *EPSPS* gene copy numbers for 94.3% of *A. palmeri* samples. Although RT q-PCR is a reliable tool to measure gene amplification, a few samples showed irregular cycle counts that led to incorrect measurements. An ALS cycle number that differed from the mean of 11 and an EPSPS cycle number that differed from the mean of 6.7 impacted the calculation for gene amplification. The most drastic examples include D5, E4, and G5 where gene amplification calculations resulted in more than 190 copies. This is due to the decreased EPSPS cycle number and increased ALS cycle number resulting in a drastically increased copy number value. Also, technical replicates from G2 and G4 display higher than average cycle numbers for EPSPS, which may have reduced their calculated copy numbers. The variance in these results severely impacted the copy number statistics resulting in a large standard deviation. Additional technical replication would have effectively reduced this variance producing more reliable EPSPS copy number results.

PCR results from all 67 *F. occidentalis* samples were difficult to attain given the low concentration of plant DNA present in each sample. Positive EF-1 amplification from 66 out of 67 *F. occidentalis* samples reflect viable results for all samples except C4S. The lack of an ALS primer amplicon in 66 out of 67 samples shows that the ingested plant material was digested in both starved and not-starved samples, or that this portion of plant DNA was quickly degraded through digestion. Amplification of the EPSPS gene was obtained regardless of ALS amplification.

The eccDNA amplification region was very specific, thus the lack of eccDNA amplification could have been due to the breaking down of the extrachromosomal circular DNAs. Although the presence of eccDNA in these samples is not absolute based on these results, the presence of the EPSPS amplicon suggests further investigation. Future studies should include the life stage of both *A. palmeri* and *F. occidentalis* and utilize a controlled environment. As reviewed by Reitz *et. al.*, the first larval instage of *F. occidentalis* prefers to feed on inflorescence and is particularly attracted to certain aromatic amino acids (Reitz et al., 2020). Aromatic amino acids are synthesized via the shikimate pathway, which is the target site of glyphosate. One may speculate that *F. occidentalis* at the correct life stage may be exposed to increased levels of eccDNA in the inflorescent tissue of GR *A. palmeri*, especially when exposed to glyphosate. Thus, future work should involve a colony of *F. occidentalis* that can reproduce on GR female *A. palmeri* flowers to increase the chance of that colony acquiring eccDNA. These thrips may be transferred to the flowers of a GS female *A. palmeri* plant, which may then produce GR offspring containing the eccDNA after crossing with a GS male.

Table A.1.1: This table reports calculated *EPSPS* copy numbers for *Amaranthus palmeri* samples collected at seven sites throughout Northeastern Colorado. The last four samples in this table report *EPSPS* gene copy numbers for known resistant and known susceptible samples from both Georgia and Arizona. Samples from sites A through G were identified as glyphosate resistant due to their visible vitality in Roundup Ready © sugar beet fields. The high level of gene duplication seen in these samples is characteristic of extra chromosomal circular DNA as opposed to the tandem gene duplication identified in the related species *Bassia scoparia*.

Sample Number	EPSPS Ct	ALS Ct	EPSPS Gene Copy Number
A1	5.22	11.69	90.70
A2	6.13	9.16	71.80
A3	9.15	11.84	55.20
A5	5.81	12.29	89.35
A5	6.89	12.05	35.90
B1	5.97	11.87	59.80
B2	6.35	11.45	34.20
B3	15.96	11.70	25.25
B4	5.06	11.49	86.55
B5	5.74	11.42	51.50
C1	5.88	11.76	59.30
C2	5.85	11.46	48.95
C3	7.21	11.88	25.55
C4	5.32	11.65	80.85
C5	7.07	11.53	22.10
D1	5.24	11.47	75.50
D2	7.30	12.06	27.20
D3	6.31	11.84	46.30
D4	6.03	11.64	48.95
D5	3.71	11.26	193.50
E1	6.09	11.75	50.20
E2	5.46	11.66	91.57
E3	4.91	12.08	147.94
E4	4.81	12.14	223.75
E5	6.11	11.40	39.29
F1	6.21	11.77	48.43
F2	7.44	11.55	17.32
F3	6.92	11.56	24.90
F4	5.99	11.44	44.06
F5	7.51	11.65	18.04
G1	5.95	11.53	48.00
G2	11.49	11.19	0.84
G3	6.46	11.32	29.16
G4	13.28	11.28	0.25
G5	4.00	11.47	184.03
Arizona Resistant	5.18	11.50	96.08
Arizona Susceptible	12.66	12.00	0.63
Georgia Resistance	5.71	11.56	58.31
Georgia Susceptible	12.30	11.66	0.64

FIGURES

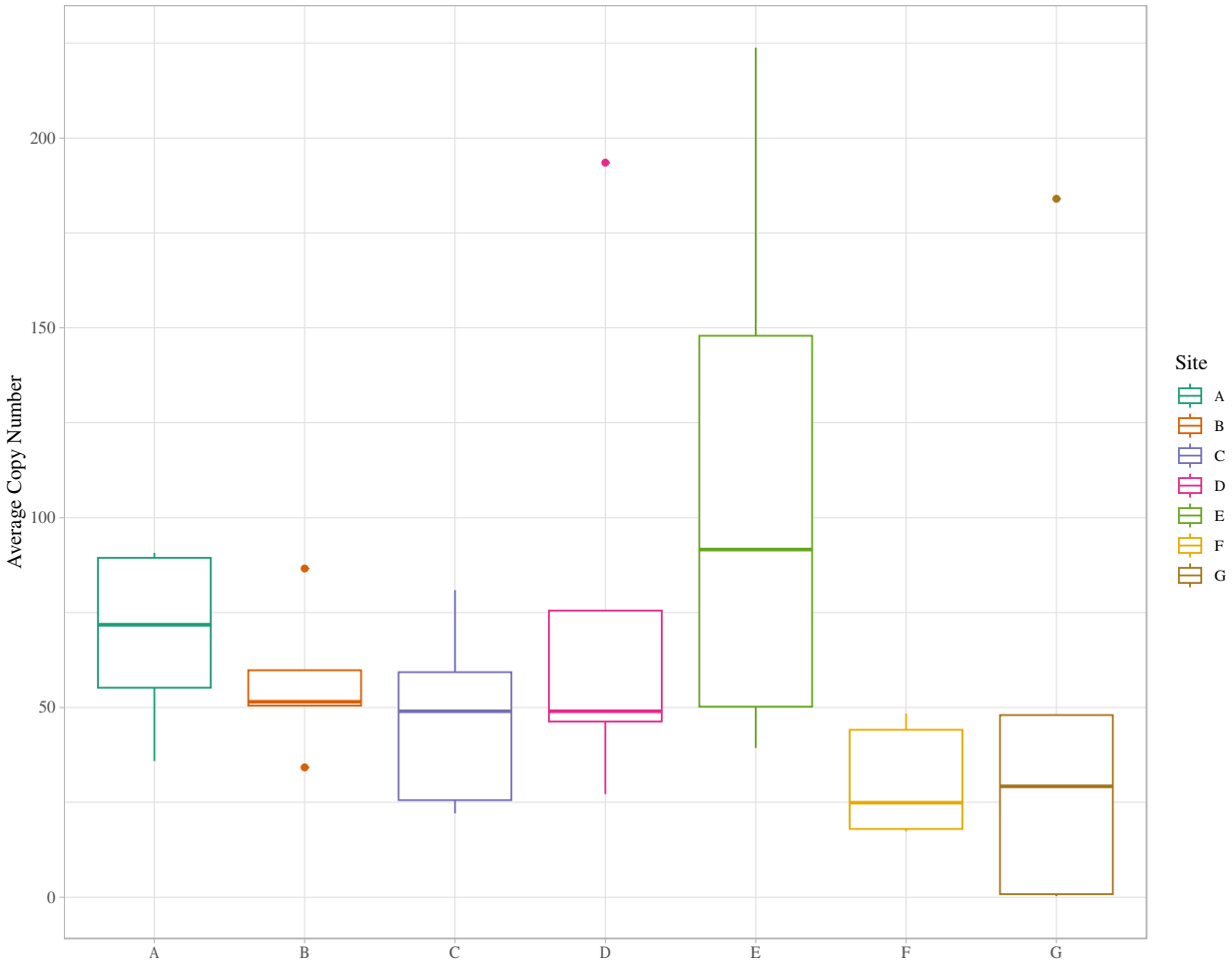


Figure A.1.1: A survey of *A. palmeri* glyphosate resistance in seven *Beta vulgaris* plots throughout Northeastern Colorado. These boxplots display the spread of calculated gene copy numbers between each collection site. Each of these sites included five biological replicates. DNA from each replicate was extracted and run on quantitative real-time PCR with ALS and EPSPS primers.

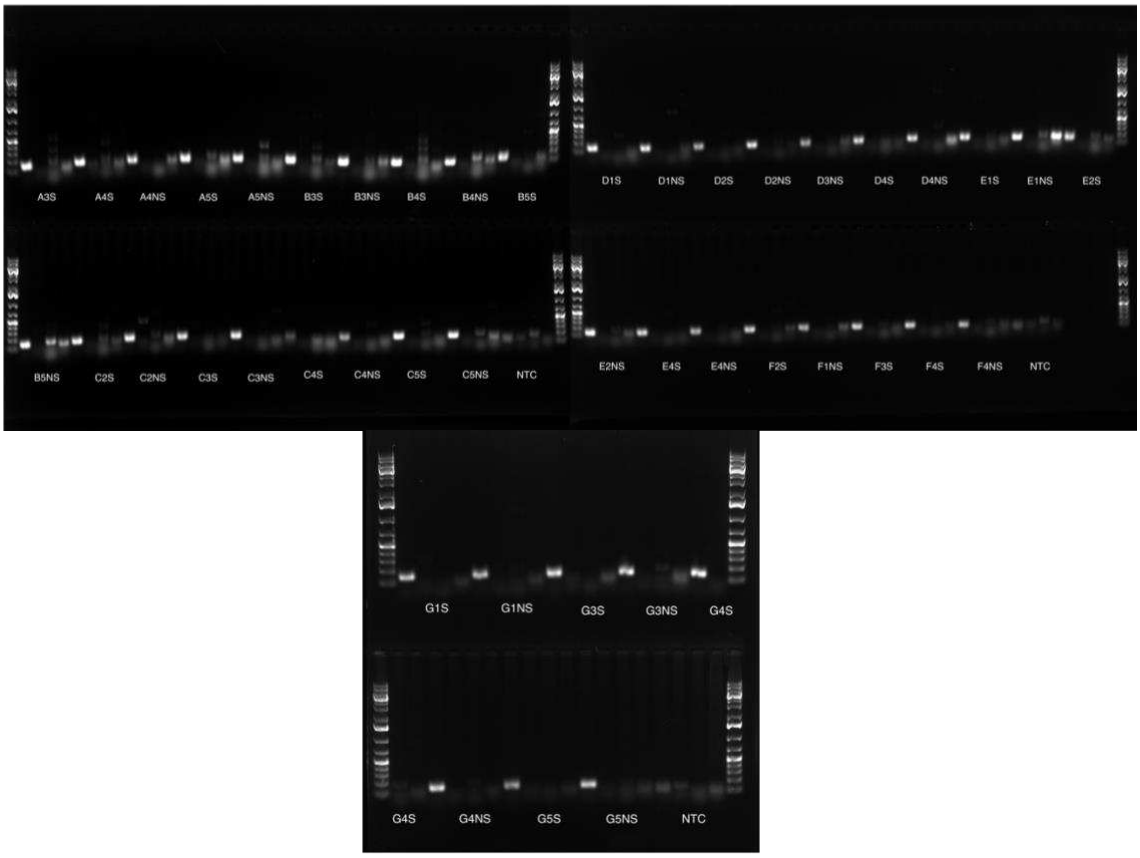


Figure A.1.2: Polymerase chain reaction results for all collected *F. occidentalis* samples collected on glyphosate resistant *A. palmeri* from Northeastern Colorado. Each sample is labeled with the collection site letter, the abbreviation 'NS' for not-starved, and 'S' for starved samples. Not starved samples were allowed to feed on glyphosate resistant *A. palmeri* until DNA extraction. Starved samples were not allowed to feed before DNA extraction. Each sample was run with four primers in the following respective order: EF-1, ALS, EPSPS, eccDNA. The EF-1 primer amplifies a common housekeeping gene in *F. occidentalis* and was used as a positive control. The ALS primer is a common housekeeping gene for plants and was used to assess the presence of plant DNA in the sample. The EPSPS primer amplifies the target site of glyphosate, and the eccDNA primer amplifies a site-specific region of the large extrachromosomal genetic construct. Non-template controls (NTC) were run in each PCR for a negative control.