It's in the Genes! A Comparison of Drought Response Genes Between Drought Tolerant and Drought Sensitive Plants Through RNA-seq Methods

Dylan M. D'Agate
Half Hollow Hills High School, Dix Hills, NY

ACKNOWLEDGEMENTS

I would like to thank Mr. Keffy Kehrli (Stony Brook University) for providing guidance and assistance with the Gene Omnibus program, Galaxy Toolbox, and statistical analysis of the data. I would like to thank Dr. Joshua Rest (Stony Brook University), for his review of my manuscript and providing guidance.

TABLE OF CONTENTS

LIST OF TABLES/FIGURES

INTRODUCTION

Global climate change and its effect on increasing drought conditions poses a major challenge to crop production and food security in developing countries (Pachauri and Meyer, 2014). In fact it is estimated that the risk of drought affecting the global food supply could triple by 2040 (Bailey, 2015). Novel methods to help farmers combat drought conditions are urgently needed.

Genetic engineering offers great potential for combating the effects of drought on crops. Research on gene expression in a single plant species during drought conditions has provided useful information (Azevedo, 2011). However, few studies have compared plant gene expression in different plant species during drought conditions using a transcriptomic profiling technology, called RNA-Seq (Benny, 2019). RNA-Seq is a novel method used to amplify an entire transcriptome to be surveyed in a fast and efficient manner.

The National Center for Biotechnology Information (NCBI) has created an international public library of RNA-Seq data called Gene Expression Omnibus (GEO). Using GEO we were able to obtain the RNA-seq data on two important crops, Sorghum bicolor (cereal crop) and Solanum tuberosum (potato) during drought conditions. Both drought tolerant cereal crop and drought sensitive potato are important crops in developing countries at risk for food insecurity due to drought (Devaux, 2014). Cereal crop is an ideal drought tolerant model for evaluating important genes that may promote drought tolerance (Ashraf, 2010). By understanding how genes are differentially expressed in these two plant species during drought, scientists can better understand which genes to target through genetic engineering and help improve plant drought tolerance.

MATERIALS AND METHODS

We searched studies using the search criteria (Sorghum bicolor, [Organism], Solanum tuberosum [Organism]) AND "high throughput sequencing" AND "Drought") on the leaves of Sorghum bicolor and Solanum tuberosum in an international public repository, National Center for Biotechnology Information's (NCBI) Gene Expression Omnibus (GEO). GEO archives gene expression and other functional genomics data sets creating a "public library" of genomic data. Once the studies were obtained, we were able to download the raw RNA-Seq data from the leaves of 4 different conditions: Sorghum bicolor (cereal crop) normal, Sorghum bicolor (cereal crop) drought, Solanum tuberosum (potato) normal, Solanum tuberosum (potato) drought. Next we utilized an open, web-based platform for data intensive biomedical research called Galaxy Toolbox.

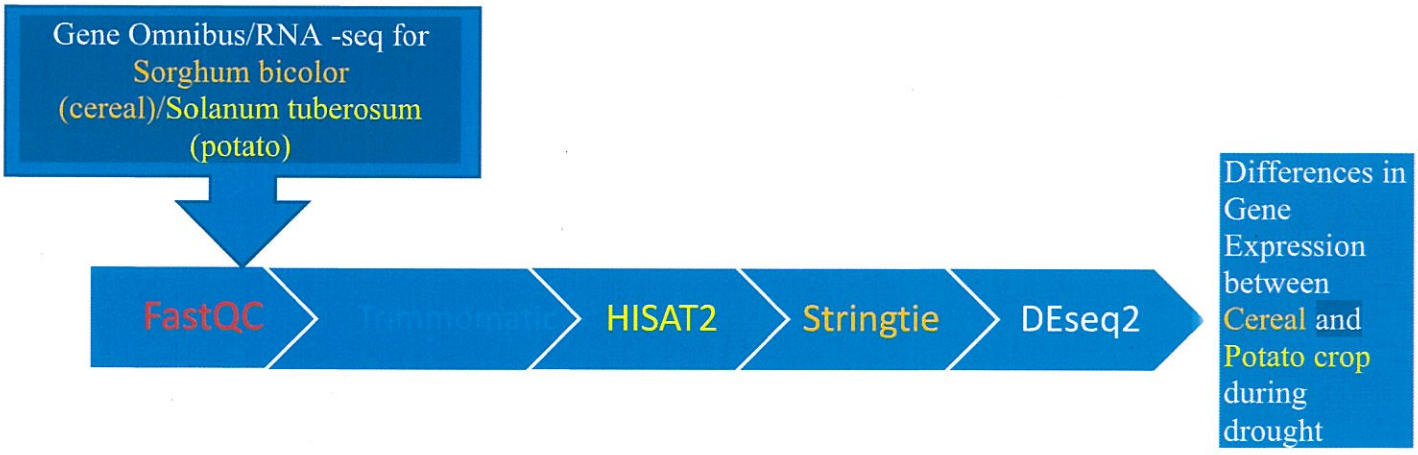

FIGURE 1. Galaxy Bioinformatics Workflow of RNA-seq Processing

Galaxy Toolbox consists of multiple programs used to help analyze RNA-seq information. The first program we used was called Fast QC, a quality control program for RNA-

seq reads. Fast QC works by importing raw RNA sequencing data and exporting a report detailing the quality of the data. Once the quality control of the data was complete we used a program called Trimmomatic to trim off bases that fall below a specified quality threshold.

Next, we performed a hierarchical indexing for spliced alignment of transcripts (HISAT2). This program efficiently aligns short sequencing reads. The program takes the trimmed sequencing reads from the RNA-seq samples and aligns them, producing a longer, more complete sequence of the original RNA transcripts. The HISAT2 output data is then used by another program called StringTie which aligns the complete sequence to a control genome and completes the tallying process for each positively identified sequence to generate a gene expression profile for the sample. Finally, we used a program called DESeq2 to determine whether or not the differences in sequence counts between two samples (Sorghum bicolor (cereal crop) drought and Solanum tuberosum (potato)) were statistically significant.

RESULTS

Several important differences were observed in our study between the Cereal species and Potato. First, the drought tolerant cereal species had a much lower number of expressed genes than the drought sensitive potato. (Table 1)

Plant Species	Total DE (%)	Upregulated (%)	Downregulated (%)
Potato	2055 (38)	1161 (57)	894 (43)
Cereal crop	653 (12)	220 (33)	433 (67)

TABLE 1. Differentially Expressed Single Copy Ortholog Genes. (Sonnhammer,2015) ($p<.05$)

Second, there were relatively few genes shared between the two plant species (Figure 2), and of the genes that were differentially expressed in both plant species, most were expressed in opposite directions (Table 2).

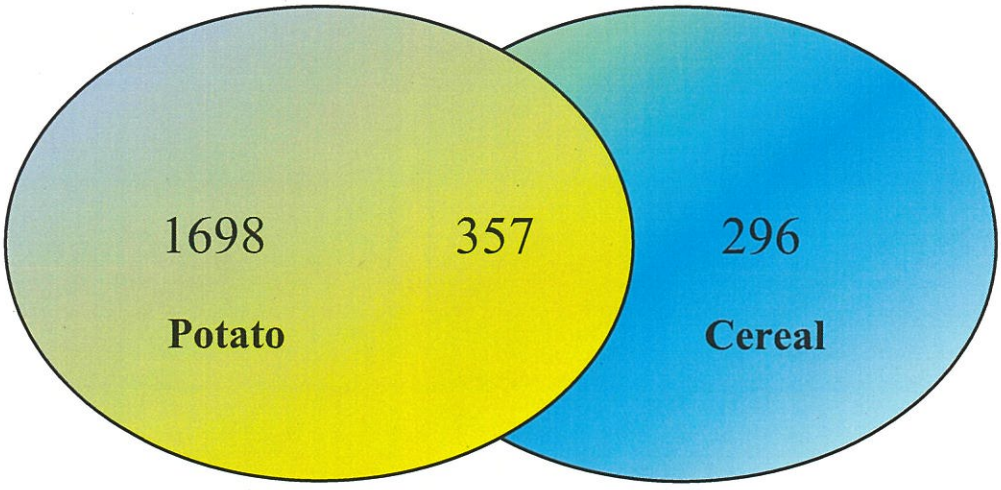

FIGURE 2. Total Differentially Expressed Ortholog Genes Individually and in Common

	Cereal crop Upregulation	*Cereal crop* Downregulation
Potato Upregulation	67	**152**
Potato Downregulation	**45**	93

TABLE 2. Differential Gene Regulation Between Cereal and Potato Species - Of the genes that were shared in both plants, 197 of 357 were differentially expressed in the opposite direction ($p<.05$)

Finally, the most significant drought responsive genetic process, Organonitrogen compound biosynthesis, was downregulated in the cereal species and upregulated in the potato (Table 3).

Potato ***Upregulated* Gene Processes**	***Cereal crop*** ***Downregulated* Gene Processes**
organonitrogen compound metabolic process *2.10E-57	organonitrogen compound biosynthetic process *1.42E-113
organonitrogen compound biosynthetic process *1.90E-54	amide biosynthetic process *6.86E-107
nitrogen compound metabolic process *5.00E-44	cellular biosynthetic process *4.26E-106

TABLE 3. Comparison of the Most Significant Gene Processes During Drought ($P<.001$)

DISCUSSION

In this study we observed that during drought conditions the drought tolerant Cereal crop displayed different "genetic behavior" than the drought sensitive Potato species. First, as demonstrated in Table 1, and consistent with prior studies (Fracasso, 2016), the drought tolerant Cereal crop had less expressed genes and more downregulation than the drought sensitive Potato species. Second, in Figure 2 we note there were relatively few genes shared between the two plant species. Furthermore, when examining the genes that were shared, we observed most were expressed in opposite directions with the Cereal crop displaying more downregulation and the Potato species displaying upregulation (Table 2). These differences in how genes are expressed during drought conditions may help explain the ability of the Cereal crop to tolerate drought conditions. These findings suggest that during drought conditions the Cereal crop appears to "conserve energy" and enter a "survival mode." One known effect of drought stress is "leaf senescence", which helps to decrease water loss and plays an important role in plant survival (Aoyama, 2014). Studies have demonstrated "leaf senescence" can be triggered by low Nitrogen availability (Wingler, 2006). In the present study the most significant gene processes during drought conditions between both plant species involved the regulation of organonitrogen processes. The drought sensitive Potato upregulated organonitrogen processes and the drought tolerant Cereal crop downregulated organonitrogen processes. This observation is consistent with studies that correlated low nitrogen levels (downregulation of organonitrogen species) during drought and the promotion of "leaf senescence" as a survival adaptation (Table 3, Figure 3).

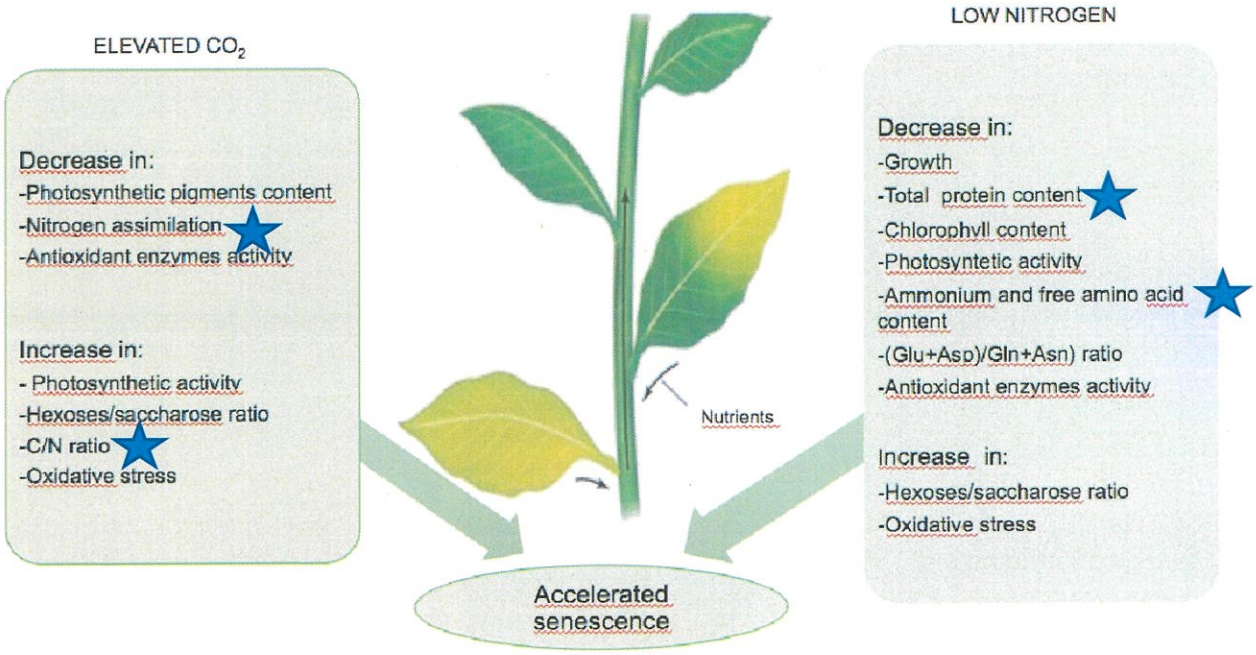

Figure 3 Factors that Determine the Induction of Leaf Senescence ("Survival Mode")
★-Processes observed in Cereal crop during drought (Aoyama, 2014).

CONCLUSION

In this study, the drought tolerant Cereal crop displayed different "genetic behavior" during drought conditions compared to the drought sensitive Potato. During drought conditions the Cereal crop appears to "conserve energy" and enter a "survival mode" with less expressed genes and greater downregulation. The Potato species demonstrated an opposite response. The most significant drought responsive gene processes, Organonitrogen compounds, had different expression between the two plant species. The downregulation of Organonitrogen compounds in the Cereal crop may improve plant drought tolerance by promoting "leaf senescence". The above differences in gene expression may help elucidate how drought tolerant plants "behave" genetically during drought conditions. Finally, by understanding how genes are differentially expressed in these two plant species during drought, scientists can better understand which genes to target through genetic engineering and help improve plant drought tolerance.

Future studies can improve on the above study by including more plant species to check for patterns of gene expression during drought conditions. Future research may build on these findings in improving our understanding of drought-tolerant genes and help decipher gene networks involved in drought. Future genetic engineering may focus on these differentially expressed gene processes as molecular targets and improve drought tolerance among various drought sensitive crops. Finally, we hope the above findings may be used to help improve crop yield and food security in developing countries during drought conditions.

REFERENCES

Aoyama S, Huarancca Reyes T, Guglielminetti L, Lu Y, Morita Y, Sato T, Yamaguchi J. 2014. Ubiquitin ligase ATL31 functions in leaf senescence in response to the balance between atmospheric CO2 and nitrogen availability in Arabidopsis. Plant Cell Physiol. Feb; 55(2):293-305.

Ashraf, M. 2010. Inducing drought tolerance in plants: recent advances. Biotechnology advances, 28(1), 169–83.

Azevedo H, Silva-Correia J, Oliveira J, et al. 2015. A strategy for the identification of new abiotic stress determinants in Arabidopsis using web-based data mining and reverse genetics. OMICS.2011;15(12):935–47.

Bailey R. 2015. Extreme weather and resilience of the global food system, Global Food Security Programme.

Benny J, Pisciotta A, Tiziano,C et al. 2019. Identification of key genes and its chromosome regions linked to drought responses in leaves across different crops through meta-analysis of RNA-Seq data. BMC Plant Biology;194(19):1-18.

Devaux, André, Kromann, Peter & Ortiz, Oscar. 2014. Potatoes for Sustainable Global Food Security. Potato Research. Vol. 57. 10.1007/s11540-014-9265-1.

Fracasso A, Trindade LM, Amaducci S. 2016. Drought stress tolerance strategies revealed by RNA-Seq in two sorghum genotypes with contrasting WUE. BMC Plant Biol. BMC Plant Biology; 16: 115.

Lobell, D. B. 2012. The influence of climate change on global crop productivity. *Plant Physiol.* 160, 1686–1697.

Sonnhammer E and Östlund G. 2015. Nucleic Acids Res. 43:D234-D239.

Wang, Z., Gerstein, M., & Snyder, M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet, 10(1), 57-63.

Wingler A, Purdy S, MacLean JA, Pourtau N. 2006. The role of sugars in integrating environmental signals during the regulation of leaf senescence. J Exp Bot.:57(2):391-9.